

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO**  
**CENTRO TECNOLÓGICO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**

**LEONARDO DELARMELINA SECCHIN**

**PROGRAMAÇÃO EM DOIS NÍVEIS: TEORIA E  
ALGORITMOS**

**VITÓRIA**

**2010**

LEONARDO DELARMELINA SECCHIN

**PROGRAMAÇÃO EM DOIS NÍVEIS: TEORIA E  
ALGORITMOS**

Dissertação submetida ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do grau de Mestre em Informática na área de Otimização.

Orientadores:

Prof. Dr. Hannu Tapio Ahonen

Prof. Dr. Arlindo Gomes de Alvarenga

**VITÓRIA**

**2010**

Dados Internacionais de Catalogação-na-publicação (CIP)  
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

---

S444p Secchin, Leonardo Delarmelina, 1982-  
Programação em dois níveis : teoria e algoritmos / Leonardo  
Delarmelina Secchin. – 2010.  
139 f. : il.

Orientador: Hannu Tapio Ahonen.  
Co-Orientador: Arlindo Gomes de Alvarenga.  
Dissertação (mestrado) – Universidade Federal do Espírito  
Santo, Centro Tecnológico.

1. Programação (Matemática). 2. Otimização matemática. I.  
Ahonen, Hannu Tapio. II. Alvarenga, Arlindo Gomes de. III.  
Universidade Federal do Espírito Santo. Centro Tecnológico. IV.  
Título.

CDU: 004

---

# Programação em Dois Níveis: Teoria e Algoritmos

**Leonardo Delarmelina Secchin**

Dissertação submetida ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Mestre em Informática.

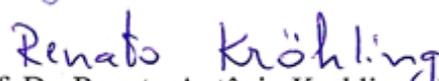
Aprovada em 18/03/10 por:



Prof. Dr. Hannu Tapio Ahonen - DI/UFES



Prof. Dr. Arlindo Gomes de Alvarenga - DI/UFES



Prof. Dr. Renato Antônio Krohling - DI/UFES



Prof. Dr. Henrique Pacca Loureiro Luna - IC/UFAL

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Vitória-ES, março de 2010

# Agradecimentos

Agradeço a minha família que me aguenta sem quase reclamar.

Agradeço aos meus orientadores e mestres Arlindo Gomes de Alvarenga e Hannu Tapio Ahonen pelos ensinamentos.

Agradeço a minha amada e queridinha Shirley pela paciência e por me envolver em seus estudos sobre geolinguística, mesmo que não saiba em que isso possa me ser útil.

Agradeço aos meus amigos da turma de graduação Andréia, Karlo, Kenia, Luciana, Matheus e Rick. Agradeço também aos meus amigos do curso de matemática André (Pacato), Andréia (Foguinho), Átila, Daniel Caxias, Diogo, Fábio Lima (Toddy), Fábio Simas, Gabriel, João Paulo (Maluco Beleza), Marcia (Marcinha), Michel (666!!!), Paulo Cesar, Robson (Barba) e todos os outros que de uma forma ou outra contribuíram para minha formação acadêmica, em conversas motivadoras ou simplesmente pela presença que tanto me alegra.

Agradeço aos meus amigos do mestrado Diego e Kamila pelos momentos agradáveis que passamos no Laboratório de Otimização.

Agradeço aos professores da graduação Ricardo Soares Leite e Moacir Rosado Filho por acreditarem em meu potencial e me incentivarem sempre.

Agradeço ao professor do mestrado Renato Krohling por seus valiosos conselhos e ensinamentos.

Agradeço aos pesquisadores que gentilmente enviaram-me seus artigos e teses quando solicitados.

Agradeço a Raul Seixas por suas palavras que tanto ensinam.

Agradeço a CAPES pelo apoio financeiro, sem o qual a realização deste trabalho seria muito mais difícil.

*Ah, mas que sujeito chato sou eu que não acha nada engraçado...  
Macaco, praia, carro, jornal, tobogã... eu acho tudo isso um saco!*

Raul Seixas

# Resumo

Este trabalho aborda de forma rigorosa o problema de dois níveis, sobretudo o caso linear. Resultados conhecidos da literatura tiveram suas demonstrações reproduzidas, ou refeitas. Como motivação para o leitor, formulações de problemas clássicos como problemas de dois níveis foram expostas. No aspecto teórico, destacam-se como contribuições a formalização das relações entre os modelos usualmente encontrados na literatura; suas extensões para problemas multinível; o resultado que complementa a equivalência entre soluções ótimas dos modelos para o caso linear otimista; e a generalização do método de Calamai e Vicente para geração de problemas-teste lineares. No aspecto prático, destaca-se o novo método para soluções ótimas locais de problemas lineares, cujo diferencial diante de outros métodos é a generalidade: engloba ilimitabilidade, e exige apenas que o poliedro do problema não tenha faces degeneradas.

**Palavras-chave:** Programação em dois níveis. Programação multinível. Problemas-teste.

# Abstract

This work gives a rigorous approach of bilevel problems, especially the linear case. Proofs of known results in the literature are reproduced or remade. As motivation for the reader, classic problems are reformulated as bilevel problems. In theoretical point of view, some contributions are the formalization of relations between models of literature; their extensions to multilevel problems; the result that complements the equivalence between optimal solutions of the models in linear optimistic case; and the generalization of the method of Calamai and Vicente for generation of linear test problems. In practical point of view, the contribution is a new algorithm for local optimal solutions of linear problems, which differs from other methods in generality: treat unlimited problems, and only requires that the problem's polyhedron does not have degenerate faces.

**Key words:** Bilevel programming. Multilevel programming. Test problems.

# Lista de Figuras

1	Ilustração para o Exemplo 2.0.5 mostrando a diferença em considerar ou não restrições de alto nível em $\mathfrak{S}$ . . . . .	24
2	Relações entre os modelos (2.2) otimista, (2.4)-(2.5) e BLP. . . . .	25
3	Projeção $\bar{\Omega}_x = \pi(\Omega_l)$ de $\Omega_l$ sobre $\mathbb{R}^{n_x}$ . . . . .	32
4	Ilustração para o Exemplo 3.1.11 mostrando o movimento de restrições. . .	39
5	À esquerda, comportamento do problema (*); à direita, comportamento do problema (**). . . . .	40
6	Região admissível $\Upsilon$ desconexa (linhas escuras). A linha pontilhada representa uma restrição de alto nível. . . . .	40
7	Situação hipotética em que $r = 0$ e todas faces de dimensão 2 adjacentes a $(x^r, y^r)$ são racionais. As linhas em destaque representam faces analisadas no estado corrente. A face de dimensão 2 relativa à restrição 1 não foi analisada. Aqui foram representados as execuções de dois laços do passo 3. . . . .	68
8	Resolução do problema $p(x_k)$ . À esquerda, geometria para $x_k < \frac{\rho_k+5}{2}$ . Ao centro, $x_k = \frac{\rho_k+5}{2}$ . À direita, $x_k > \frac{\rho_k+5}{2}$ . As linhas em destaque são as soluções ótimas. . . . .	88

# Lista de Tabelas

1	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 25$ e $n_y = 40$ . . . . .	95
2	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 75$ e $n_y = 35$ . . . . .	95
3	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 75$ e $n_y = 90$ . . . . .	96
4	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 100$ e $n_y = 70$ . . . . .	96
5	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 100$ e $n_y = 150$ . . . . .	97
6	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 200$ e $n_y = 150$ . . . . .	97
7	Testes computacionais - Problemas gerados segundo Calamai e Vicente com $n_x = 200$ e $n_y = 350$ . . . . .	98
8	Testes computacionais - Problemas do tipo GRUPO 1 com $n_x = 15$ . . . . .	98
9	Testes computacionais - Problemas do tipo GRUPO 1 com $n_x = 25$ . . . . .	99
10	Testes computacionais - Problemas do tipo GRUPO 1 com $n_x = 50$ . . . . .	99
11	Testes computacionais - Problemas do tipo GRUPO 1 com $n_x = 70$ . . . . .	100
12	Testes computacionais - Problemas do tipo GRUPO 2 com $n_x = 15$ . . . . .	100
13	Testes computacionais - Problemas do tipo GRUPO 2 com $n_x = 25$ . . . . .	101
14	Testes computacionais - Problemas do tipo GRUPO 2 com $n_x = 50$ . . . . .	101
15	Testes computacionais - Problemas do tipo GRUPO 2 com $n_x = 70$ . . . . .	102
16	Relações entre problemas primal e dual. . . . .	124

17	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x = 25$ e $n_y = 40$ . . . . .	132
18	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x = 75$ e $n_y = 35$ . . . . .	133
19	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x = 75$ e $n_y = 90$ . . . . .	133
20	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x =$ $100$ e $n_y = 70$ . . . . .	134
21	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x =$ $100$ e $n_y = 150$ . . . . .	134
22	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x =$ $200$ e $n_y = 150$ . . . . .	135
23	Parâmetros dos problemas gerados segundo Calamai e Vicente com $n_x =$ $200$ e $n_y = 350$ . . . . .	135
24	Parâmetros dos problemas-teste do tipo GRUPO 1 com $n_x = 15$ . . . . .	136
25	Parâmetros dos problemas-teste do tipo GRUPO 1 com $n_x = 25$ . . . . .	136
26	Parâmetros dos problemas-teste do tipo GRUPO 1 com $n_x = 50$ . . . . .	137
27	Parâmetros dos problemas-teste do tipo GRUPO 1 com $n_x = 70$ . . . . .	137
28	Parâmetros dos problemas-teste do tipo GRUPO 2 com $n_x = 15$ . . . . .	138
29	Parâmetros dos problemas-teste do tipo GRUPO 2 com $n_x = 25$ . . . . .	138
30	Parâmetros dos problemas-teste do tipo GRUPO 2 com $n_x = 50$ . . . . .	139
31	Parâmetros dos problemas-teste do tipo GRUPO 2 com $n_x = 70$ . . . . .	139

# Lista de Símbolos

$\arg \min$	conjunto dos elementos que atingem o mínimo
$\arg \max$	conjunto dos elementos que atingem o máximo
s.a.	“sujeito a”
$M^t$	transposta da matriz $M$
$M_j$	$j$ -ésima linha da matriz $M$
$[M]_{j \in J}$	matriz cujas linhas são as linhas $M_j$ , com $j \in J$ , onde $J$ é conjunto ordenado de índices
$ E $	cardinalidade do conjunto finito $E$
$\complement E$	complementar do conjunto $E$ (o conjunto universo fica claro no contexto)
$\text{conv } E$	fecho convexo do conjunto $E$
$\text{cone } F$	fecho cônico do conjunto $F$
$\mathbb{R}_+^n$	conjunto $\{x \in \mathbb{R}^n; x \geq 0\}$
$\mathbb{Z}_+^n$	conjunto $\{x \in \mathbb{Z}^n; x \geq 0\}$
$\mathbb{R}^{m \times n}$	conjunto das matrizes reais de $m$ linhas e $n$ colunas
$\ \cdot\ $	norma euclidiana
$\ x\ _\infty$	norma do infinito do vetor $x \in \mathbb{R}^n$ , ou seja, $\ x\ _\infty = \max_{1 \leq i \leq n}  e_i x $ , onde $e_i$ é o $i$ -ésimo vetor canônico de $\mathbb{R}^n$
$B_\epsilon(\bar{x})$	bola aberta definida pela norma euclidiana, de centro em $\bar{x}$ e raio $\epsilon > 0$ , ou seja, $B_\epsilon(\bar{x}) = \{y; \ \bar{x} - y\  < \epsilon\}$
$B_\epsilon^\infty(\bar{x})$	bola aberta definida pela norma do infinito, de centro em $\bar{x}$ e raio $\epsilon > 0$ , ou seja, $B_\epsilon^\infty(\bar{x}) = \{y; \ \bar{x} - y\ _\infty < \epsilon\}$
$\ M\ _\infty$	norma da matrix $M = [m_{ij}]$ definida por $\ M\ _\infty = \max_i \left( \sum_j  m_{ij}  \right)$
$\text{dist}(E, F)$	distância entre os conjuntos $E$ e $F$
$\text{int } E$	interior do conjunto $E$

# Sumário

<b>1</b>	<b>Introdução</b>	<b>14</b>
1.1	Organização deste trabalho . . . . .	15
<b>2</b>	<b>O problema de dois níveis</b>	<b>16</b>
<b>3</b>	<b>O caso linear</b>	<b>27</b>
3.1	Delimitação teórica . . . . .	27
3.2	Existência de soluções ótimas . . . . .	41
<b>4</b>	<b>Formulações dois níveis de problemas clássicos</b>	<b>42</b>
4.1	Problema de programação linear 0-1 misto . . . . .	42
4.2	Problema de programação bilinear . . . . .	43
4.3	Problema max-min linear . . . . .	43
4.4	Um problema específico: o caixeiro viajante (TSP) . . . . .	44
4.4.1	O problema de caminho mínimo . . . . .	44
4.4.2	O problema de tarifação (TOP) . . . . .	45
4.4.3	Um modelo TOP para o TSP . . . . .	48
<b>5</b>	<b>Algoritmos existentes para o caso linear</b>	<b>52</b>
5.1	O algoritmo do K-ésimo melhor vértice . . . . .	52
5.2	Abordagem via complementariedade . . . . .	53
5.3	Eliminação de variáveis . . . . .	54
5.4	Outras abordagens . . . . .	59

<b>6</b>	<b>Um novo algoritmo para soluções ótimas locais no caso linear</b>	<b>60</b>
6.1	Formulação . . . . .	60
6.2	Obtendo $(x^0, y^0) \in \Upsilon$ quando $B_1 = 0$ . . . . .	75
6.3	Testes computacionais . . . . .	76
6.3.1	Geração de problemas-teste . . . . .	76
6.3.2	Generalização do método de Calamai e Vicente para geração de problemas-teste . . . . .	82
6.3.3	Resultados . . . . .	92
<b>7</b>	<b>Outras questões em problemas de dois níveis</b>	<b>105</b>
7.1	Regularização de problemas com variáveis de baixo nível inteiras . . . . .	105
<b>8</b>	<b>Conclusões e trabalhos futuros</b>	<b>108</b>
	<b>Referências</b>	<b>111</b>
	<b>Apêndice A – Revisão de pré-requisitos</b>	<b>114</b>
A.1	Álgebra Linear . . . . .	114
A.2	Topologia no $\mathbb{R}^n$ . . . . .	116
A.3	Poliedros, programação linear e dualidade . . . . .	120
A.4	Funções ponto a conjunto . . . . .	125
	<b>Apêndice B – Problemas multinível</b>	<b>127</b>
	<b>Apêndice C – Parâmetros dos problemas-teste usados</b>	<b>132</b>

# 1 Introdução

Problemas de dois níveis foram considerados primeiramente por Bracken and McGill (*apud* [12]), onde eram chamados de *problemas com problemas de otimização nas restrições*<sup>1</sup>. Os termos *dois níveis* e *multinível*<sup>2</sup> foram introduzidos mais tarde por Candler e Norton (*apud* [12]). Todos esses autores trabalhavam com modelos onde as restrições do alto nível não continham variáveis de baixo nível. A formulação geral foi introduzida por Aiyoshi e Shimizu (*apud* [12]).

Problemas de dois níveis têm vasta aplicação. Dentre elas, aplicações na economia; determinação de preços ótimos em estradas tarifadas, redes de eletricidade, redes de distribuição de gás e petróleo; questões militares; engenharia; gerência e administração; planejamentos para agricultura; monitoramento da emissão de gases poluentes; problema de projetos de redes; teoria dos jogos.

Atualmente, problemas de dois níveis são estudados por muitos pesquisadores, dos quais destacamos Stephan Dempe (Friburgo, Alemanha), Gilles Savard, Charles Audet e Patrice Marcotte (os três últimos de Montreal, Canadá). Esses autores destacam-se por seus trabalhos teóricos. Em relação a algoritmos, a maior parte dos trabalhos são adaptações de metaheurísticas existentes (veja a subseção 6.3.3). Pelo menos de nosso conhecimento, poucos algoritmos exatos foram desenvolvidos nos últimos anos, o que torna os problemas de dois níveis ainda mais interessantes de serem investigados.

Como referência, convidamos o leitor a consultar as revisões de literatura [21, 12, 15, 31, 22]. Outras referências interessantes são as teses de S. Dempe [14] e M. B. Campêlo Neto [11], e os livros de J. F. Bard [3] e de Shimizu *et al.* [40].

Nosso objetivo com o presente trabalho é analisar o problema de dois níveis, em especial o caso linear (funções objetivo e restrições lineares). Sempre que possível, demonstrações dos resultados da literatura serão dadas quando contribuir para o entendimento dos conceitos e algoritmos. Cabe destacarmos que este trabalho propõe-se ao

---

<sup>1</sup>Originalmente, *mathematical programs with optimization problems in the constraints*.

<sup>2</sup>Em inglês, *bilevel* e *multilevel*.

estudo teórico e prático do problema em questão. Portanto, vários resultados são dados, mesmo que não usados nos algoritmos.

## 1.1 Organização deste trabalho

No Capítulo 2 apresentamos o problema de dois níveis e seu modelo geral. No Capítulo 3 discutimos o caso linear. Resultados importantes para confecção de algoritmos são dados. No Capítulo 4 apresentamos exemplos de como o problema de dois níveis pode ser usado para modelar problemas clássicos da literatura, tais como o problema do caixeiro viajante (TSP). Detalhes e demonstrações de resultados são omitidos. No Capítulo 5 mostramos alguns algoritmos existentes na literatura para o caso linear. No Capítulo 6 introduzimos um novo algoritmo para soluções ótimas locais neste caso. No Capítulo 7 abordamos questões específicas que não são necessárias aos outros capítulos. Finalmente, no Capítulo 8 fazemos nossas conclusões e apontamos algumas possíveis direções a seguir.

Encontram-se ainda no Apêndice A uma breve revisão de alguns dos pré-requisitos para o entendimento deste trabalho, e no Apêndice B discussões sobre problemas de vários níveis.

## 2 O problema de dois níveis

Problemas de dois níveis são compostos por dois agentes que tomam decisões em níveis de hierarquia distintos. Enquanto o primeiro agente toma sua decisão  $x$ , o segundo agente obtém sua(s) solução(ões) ótima(s)  $y$  de acordo com a escolha  $x$ . Neste sentido, o agente no topo da hierarquia, o líder, tem controle parcial sobre as decisões do agente inferior, o escravo, e acaba por escolher, dentre as soluções ótimas que o escravo lhe oferece, aquela que melhor lhe convém.

Inicialmente, consideremos

$$S(x) = \arg \min_y \{f(x, y); y \in \Omega_y(x)\} \quad (2.1)$$

o conjunto das soluções ótimas do agente escravo para a decisão  $x$  do líder, onde  $\Omega_y(x) \subset \mathbb{R}^{n_y}$  para todo  $x$ . Estamos supondo que o problema

$$\inf_y \{f(x, y); y \in \Omega_y(x)\}$$

atinge seus ínfimos para todo  $x$ , quando houverem (por isso usamos “min” em (2.1)). O problema em (2.1) é chamado *problema de baixo nível* (LLP( $x$ )<sup>1</sup>) ou *problema escravo*. Para discussões futuras, definimos o *domínio de S*

$$\Omega_x := \{x \in \mathbb{R}^{n_x}; S(x) \neq \emptyset\}$$

(usualmente na literatura denomina-se  $\text{dom } S$  - veja [2]).

Há questões interessantes em problemas de dois níveis não presentes em problemas de um único nível. Por exemplo, pode-se considerar que o escravo, tendo em mãos a decisão  $x$ , retorne ao líder suas decisões que são as melhores do ponto de vista do líder, ou então retorne aquelas que são as piores para o líder. Esses princípios, de cooperação ou não entre os agentes, são tratados na literatura como formulações *otimista* e *pessimista* do problema de dois níveis. Abordaremos-as logo adiante. Antes disso, observamos que se  $S(x)$  reduz-se

---

<sup>1</sup>Do inglês *Lower Level Problem*.

a um único elemento para cada  $x$ , digamos  $S(x) = \{y_x\}$ ,  $\forall x \in \Omega_x$ , o problema de dois níveis fica bem definido, pois o retorno do escravo é único, e não há distinção entre as abordagens otimista e pessimista (ou qualquer outra que se possa imaginar). Neste caso a aplicação  $S$  induz a função  $y : \Omega_x \rightarrow \mathbb{R}^{n_y}$  dada por  $y(x) = y_x$ . Daí o problema de dois níveis pode ser descrito pelo modelo

$$\min_x F(x, y(x)) \quad \text{s.a.} \quad (x, y(x)) \in \Omega_u, \quad x \in \Omega_x,$$

onde  $\Omega_u \subset \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ , isto é, um problema de um único nível. Aqui a noção de solução ótima local/global é a usual.

Agora suponha que  $S(x)$  tenha mais de um elemento para algum  $x \in \Omega_x$ . Considerando que o líder tenha controle total apenas sobre  $x$ , não sabemos a política que o agente escravo adota ao entregar suas decisões, um subconjunto  $\mathfrak{S}(x)$  de  $S(x)$ , ao líder. Assim, podemos formular o *problema de dois níveis* como

$$\min_{(x,y)} F(x, y) \quad \text{s.a.} \quad (x, y) \in \Omega_u, \quad y \in \mathfrak{S}(x) \subset S(x). \quad (2.2)$$

O problema (2.2) sem a restrição  $y \in \mathfrak{S}(x)$  é chamado *problema de alto nível* (ULP<sup>2</sup>) ou *problema líder*. Neste sentido,  $(x, y) \in \Omega_u$  ( $y \in \Omega_y(x)$ ) são restrições de alto (baixo) nível,  $x$  ( $y$ ) é variável de alto (baixo) nível e  $F$  ( $f$ ) é função objetivo de alto (baixo) nível.

Trabalhemos agora os dois casos citados na literatura. De acordo com o que foi dito anteriormente, na abordagem otimista faz-se

$$\mathfrak{S}(x) = \arg \min_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\},$$

e na pessimista

$$\mathfrak{S}(x) = \arg \max_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\}.$$

Neste capítulo estamos supondo que o problema

$$\inf_{(x,y)} \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\} \quad (2.3)$$

atinge seus ínfimos, quando houverem. Isto permite utilizar “min” na definição de  $\mathfrak{S}(x)$  para o caso otimista. No caso da abordagem pessimista, estamos supondo que o mesmo problema com “sup” atinge seus supremos, quando houverem.

Essas suposições permitirão utilizar “min” ou “max” nos modelos que se seguirão. Isso ficará claro ao leitor, e por isso não faremos mais comentários.

---

<sup>2</sup>Do inglês *Upper Level Problem*.

Mostraremos adiante (Teorema 2.0.2) que podemos escrever a abordagem otimista como

$$\min_x \varphi_o(x) \quad \text{s.a.} \quad x \in \text{dom } \varphi_o \quad (2.4)$$

onde

$$\begin{aligned} \varphi_o(x) &= \min_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\} \text{ e} \\ \text{dom } \varphi_o &= \{x \in \mathbb{R}^{n_x}; \varphi_o(x) > -\infty\}. \end{aligned} \quad (2.5)$$

Veremos também (Teorema 2.0.3) que as soluções ótimas de (2.4)-(2.5) têm relação íntima com o modelo

$$\text{BLP}^3 : \min_{(x,y)} F(x, y) \quad \text{s.a.} \quad (x, y) \in \Omega_u \quad y \in S(x).$$

Notemos que BLP é o modelo (2.2) quando  $\mathfrak{S} \equiv S$ . Escreveremos  $\min_{x,y} F(x, y)$  no lugar de  $\min_{(x,y)} F(x, y)$ . Se quisermos minimizar a função  $\min_y F(x, y)$  em  $x$ , escreveremos  $\min_x \min_y F(x, y)$ , ou como em (2.4)-(2.5).

Da mesma forma, a abordagem pessimista pode ser descrita como

$$\min_x \varphi_p(x) \quad \text{s.a.} \quad x \in \text{dom } \varphi_p \quad (2.6)$$

onde

$$\begin{aligned} \varphi_p(x) &= \max_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\} \text{ e} \\ \text{dom } \varphi_p &= \{x \in \mathbb{R}^{n_x}; \varphi_p(x) < \infty\}. \end{aligned} \quad (2.7)$$

Esta abordagem é mais complicada de ser tratada. Neste trabalho daremos ênfase à abordagem otimista. Em [14] encontra-se vasto conteúdo sobre a abordagem pessimista.

Usando a noção usual de solução ótima local/global dos modelos (2.4)-(2.5) e (2.6)-(2.7) fazemos a seguinte definição:

**Definição 2.0.1.** *i) Um ponto  $(\bar{x}, \bar{y})$  é dito solução local otimista de (2.2) se*

$$\bar{y} \in S(\bar{x}), \quad (\bar{x}, \bar{y}) \in \Omega_u, \quad F(\bar{x}, \bar{y}) = \varphi_o(\bar{x})$$

*e existir uma vizinhança aberta  $V(\bar{x})$  de  $\bar{x}$  tal que*

$$\varphi_o(\bar{x}) \leq \varphi_o(x), \quad \forall x \in V(\bar{x}) \cap \text{dom } \varphi_o.$$

*Dizemos que  $(\bar{x}, \bar{y})$  é solução global otimista de (2.2) se puder ser  $V(\bar{x}) = \mathbb{R}^{n_x}$ .*

---

<sup>3</sup>Do inglês *BiLevel Problem*.

ii) Um ponto  $(\bar{x}, \bar{y})$  é dito solução local pessimista de (2.2) se

$$\bar{y} \in S(\bar{x}), \quad (\bar{x}, \bar{y}) \in \Omega_u, \quad F(\bar{x}, \bar{y}) = \varphi_p(\bar{x})$$

e existir uma vizinhança aberta  $V(\bar{x})$  de  $\bar{x}$  tal que

$$\varphi_p(\bar{x}) \leq \varphi_p(x), \quad \forall x \in V(\bar{x}) \cap \text{dom } \varphi_p.$$

Dizemos que  $(\bar{x}, \bar{y})$  é solução global pessimista de (2.2) se puder ser  $V(\bar{x}) = \mathbb{R}^{n_x}$ .

O próximo Teorema relaciona as soluções ótimas dos modelos (2.4)-(2.5) e (2.6)-(2.7) com as soluções ótimas do modelo (2.2) nas abordagens otimista e pessimista, respectivamente. Como consequência, mostra que a Definição 2.0.1 é coerente na descrição das soluções ótimas de (2.2). Ele ainda relaciona tais modelos nas situações de inviabilidade e ilimitabilidade.

**Teorema 2.0.2.** *i) (2.4)-(2.5) (respectivamente (2.6)-(2.7)) é inviável se, e somente se, (2.2) otimista (respectivamente pessimista) é inviável;*

*ii) (2.4)-(2.5) (respectivamente (2.6)-(2.7)) é ilimitado se, e somente se, (2.2) otimista (respectivamente pessimista) é ilimitado;*

*iii) Uma solução ótima local (respectivamente global)  $\bar{x}$  de (2.4)-(2.5) induz uma solução ótima local (respectivamente global)  $(\bar{x}, \bar{y})$  do modelo (2.2) otimista ao tomarmos  $\bar{y}$  proveniente de  $\varphi_o(\bar{x})$ , isto é,  $\bar{y} \in \{y; (\bar{x}, y) \in \Omega_u, y \in S(\bar{x}), F(\bar{x}, y) = \varphi_o(\bar{x})\}$ . Reciprocamente, uma solução ótima local (respectivamente global)  $(\bar{x}, \bar{y})$  do modelo (2.2) otimista é tal que  $\bar{x}$  é solução ótima local (respectivamente global) de (2.4)-(2.5).*

*O mesmo ocorre na abordagem pessimista, levando em consideração o modelo (2.6)-(2.7).*

*Demonstração.* *i)* Considere o caso otimista. Mostremos a implicação  $\Leftarrow$ ). Para cada  $x$ , definimos  $\Omega_u(x) = \{y; (x, y) \in \Omega_u\}$ . Suponhamos que (2.4)-(2.5) é viável, e fixemos  $x \in \text{dom } \varphi_o$ . Então existe  $y$  tal que  $y \in \Omega_u(x) \cap S(x)$  e  $\varphi_o(x) = F(x, y)$ . Isso implica que  $y \in \Omega_u(x) \cap \mathfrak{S}(x)$ , e portanto (2.2) otimista é viável.

Mostremos agora a implicação  $\Rightarrow$ ). Se (2.4)-(2.5) é inviável,  $\text{dom } \varphi_o = \emptyset$ , isto é,  $\varphi_o(x)$  não está definida para todo  $x$ . Então para todo  $x$ ,  $\Omega_u(x) \cap S(x) = \emptyset$  ou existe uma sequência  $\{y_n\} \subset \Omega_u(x) \cap S(x)$  tal que  $F(x, y_n) \rightarrow -\infty$ . Isso implica que  $\mathfrak{S}(x) = \emptyset$  para todo  $x$ , e o problema (2.2) otimista é inviável.

A argumentação para o caso pessimista é análoga.

ii) Considere o caso otimista. Mostremos a implicação  $\Leftarrow$ ). Se (2.2) otimista é ilimitado, existe uma sequência  $\{(x_n, y_n)\} \subset \Omega_u$  tal que  $y_n \in \mathfrak{S}(x_n), \forall n$ , e  $F(x_n, y_n) \rightarrow -\infty$ . Como  $y_n \in \mathfrak{S}(x_n), \forall n$ , vem  $\varphi_o(x_n) = F(x_n, y_n), \forall n$ . Daí  $\varphi_o(x_n) \rightarrow -\infty$ , isto é, (2.4)-(2.5) é ilimitado.

Mostremos agora a implicação  $\Rightarrow$ ). Se (2.4)-(2.5) é ilimitado, existe uma sequência  $\{x_n\} \subset \text{dom } \varphi_o$  tal que  $\varphi_o(x_n) \rightarrow -\infty$ . Para cada  $x_n$  tomemos um  $y_n \in \Omega_u(x_n) \cap S(x_n)$  com  $F(x_n, y_n) = \varphi_o(x_n)$ . Temos  $(x_n, y_n) \in \Omega_u$  e  $y_n \in \mathfrak{S}(x_n)$  para todo  $n$ , e de  $F(x_n, y_n) = \varphi_o(x_n) \rightarrow -\infty$ , segue que o problema (2.2) otimista é ilimitado.

A argumentação para o caso pessimista é análoga.

iii) Considere o modelo (2.2) otimista. Como  $\mathfrak{S}(x) = \arg \min_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\}$ , temos que

$$(x, y) \in (\Omega_x \times \mathfrak{S}(x)) \cap \Omega_u \Leftrightarrow F(x, y) = \varphi_o(x), (x, y) \in \Omega_u \text{ e } y \in S(x). \quad (2.8)$$

Ou seja,  $(x, y)$  é viável para (2.2) otimista se, e somente se,  $y \in S(x)$ ,  $(x, y) \in \Omega_u$  e  $F(x, y) = \varphi_o(x)$ . Seja  $\bar{x}$  solução ótima local de (2.4)-(2.5). Tomando um  $\bar{y}$  proveniente de  $\varphi_o(\bar{x})$ , temos  $F(\bar{x}, \bar{y}) = \varphi_o(\bar{x}) \leq \varphi_o(x), \forall x \in V(\bar{x}) \cap \text{dom } \varphi_o$ , e daí

$$F(\bar{x}, \bar{y}) \leq F(x, y), \forall (x, y) \in (\Omega_x \times \mathfrak{S}(x)) \cap \Omega_u \cap (V(\bar{x}) \times \mathbb{R}^{n_y}).$$

Notemos que para cada par  $(x, y)$  na expressão anterior, vem de (2.8) que  $x \in \text{dom } \varphi_o$ . Isso mostra que toda solução ótima local (ou global quando  $V(\bar{x}) = \mathbb{R}^{n_x}$ )  $\bar{x}$  de (2.4)-(2.5) induz uma solução ótima local (ou global)  $(\bar{x}, \bar{y})$  de (2.2) tomando um  $\bar{y}$  proveniente de  $\varphi_o(\bar{x})$ .

Reciprocamente, se  $(\bar{x}, \bar{y})$  é solução ótima local de (2.2) otimista, isto é, se

$$F(\bar{x}, \bar{y}) \leq F(x, y), \forall (x, y) \in (\Omega_x \times \mathfrak{S}(x)) \cap \Omega_u \cap (V(\bar{x}) \times V(\bar{y})),$$

então, como  $F(x, y_1) = F(x, y_2), \forall y_1, y_2 \in \mathfrak{S}(x)$ , temos

$$F(\bar{x}, \bar{y}) \leq F(x, y), \forall (x, y) \in (\Omega_x \times \mathfrak{S}(x)) \cap \Omega_u \cap (V(\bar{x}) \times \mathbb{R}^{n_y}).$$

Daí, para cada  $x \in V(\bar{x}) \cap \text{dom } \varphi_o$ , segue de (2.8) que  $\varphi_o(\bar{x}) \leq \varphi_o(x)$ , e logo  $\bar{x}$  é solução ótima local (ou global quando  $V(\bar{x}) = \mathbb{R}^{n_x}$ ) de (2.4)-(2.5).

A argumentação para a abordagem pessimista é análoga.  $\square$

Como já mencionado, o Teorema 2.0.2 mostra que a Definição 2.0.1 é coerente. Por

isso nos reportaremos às soluções ótimas de (2.2) otimista (respectivamente pessimista) como soluções otimistas (respectivamente pessimistas) de (2.2) no sentido da Definição 2.0.1.

O resultado a seguir relaciona soluções otimistas de (2.2) e soluções ótimas de BLP:

**Teorema 2.0.3.** *i) Se BLP é inviável então (2.2) otimista é inviável;*

*ii) Se (2.2) otimista é ilimitado então BLP é ilimitado;*

*iii) Suponhamos que BLP tenha solução ótima. Temos que se  $(\bar{x}, \bar{y})$  é solução local otimista (respectivamente global otimista) de (2.2) então  $(\bar{x}, \bar{y})$  é solução ótima local (respectivamente global) de BLP;*

*iv) Se  $(\bar{x}, \bar{y})$  é solução ótima global de BLP então  $(\bar{x}, \bar{y})$  é solução global otimista de (2.2).*

*Demonstração.* *i)* Para cada  $x$ , definimos  $\Omega_u(x) = \{y; (x, y) \in \Omega_u\}$ . Se BLP é inviável então  $\Omega_u(x) \cap \mathfrak{S}(x) \subset \Omega_u(x) \cap S(x) = \emptyset$  para todo  $x$  e logo (2.2) otimista é inviável.

*ii)* Se (2.2) otimista é ilimitado, existe uma sequência  $\{(x_n, y_n)\} \subset \Omega_u$  com  $y_n \in \mathfrak{S}(x_n), \forall n$ , tal que  $F(x_n, y_n) \rightarrow -\infty$ . Como  $\mathfrak{S}(x_n) \subset S(x_n), \forall n$ , BLP é ilimitado.

*iii)* Por definição existe uma vizinhança aberta  $V(\bar{x})$  de  $\bar{x}$  com  $\varphi_o(\bar{x}) \leq \varphi_o(x), \forall x \in V(\bar{x}) \cap \text{dom } \varphi_o$ . Para cada  $x \in V(\bar{x}) \cap \text{dom } \varphi_o$ , qualquer  $y \in S(x)$  com  $(x, y) \in \Omega_u$  é tal que  $F(\bar{x}, \bar{y}) = \varphi_o(\bar{x}) \leq \varphi_o(x) \leq F(x, y)$ . Daí

$$F(\bar{x}, \bar{y}) \leq F(x, y), \forall (x, y) \in (V(\bar{x}) \times S(x)) \cap \Omega_u, \quad x \in \text{dom } \varphi_o. \quad (2.9)$$

Agora, afirmamos que se para algum  $x \in V(\bar{x})$  com  $(\{x\} \times S(x)) \cap \Omega_u \neq \emptyset$ , tivermos  $x \notin \text{dom } \varphi_o$ , temos que BLP é ilimitado. De fato, suponha que exista um tal  $x$ . Como  $(\{x\} \times S(x)) \cap \Omega_u \neq \emptyset$ , a única possibilidade de se ter  $x \notin \text{dom } \varphi_o$  é que exista uma sequência  $\{y_n\} \subset S(x)$  tal que  $(x, y_n) \in \Omega_u, \forall n$ , e  $F(x, y_n) \rightarrow -\infty$ . Mas  $(x, y_n)$  é viável para BLP para todo  $n$ , e logo BLP é ilimitado.

Segue então da hipótese de BLP ter solução ótima que todo  $(x, y) \in (V(\bar{x}) \times S(x)) \cap \Omega_u$  é tal que  $x \in \text{dom } \varphi_o$ , e de (2.9) vem que  $(\bar{x}, \bar{y})$  é solução ótima local de BLP. A argumentação para solução ótima global é a mesma, onde temos  $V(\bar{x}) = \mathbb{R}^{n_x}$ .

*iv)* Temos  $F(\bar{x}, \bar{y}) \leq F(x, y), \forall (x, y) \in \Omega_u$  com  $y \in S(x)$ . Em particular,  $F(\bar{x}, \bar{y}) \leq F(\bar{x}, y), \forall y \in S(\bar{x})$  com  $(\bar{x}, y) \in \Omega_u$ . Logo  $\varphi_o(\bar{x}) = F(\bar{x}, \bar{y})$ . Também temos  $\varphi_o(\bar{x}) = F(\bar{x}, \bar{y}) \leq \min_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\} = \varphi_o(x), \forall x \in \text{dom } \varphi_o$ , isto é,  $(\bar{x}, \bar{y})$  é solução global otimista de (2.2).  $\square$

Em geral a recíproca do item (iii) do Teorema 2.0.3 para solução ótima local é falsa, como se vê no próximo exemplo.

**Exemplo 2.0.4.** [17] Considere o problema otimista de dois níveis<sup>4</sup>

$$\min_{x,y} x \quad \text{s.a.} \quad -1 \leq x \leq 1, \quad y \in S(x)$$

onde

$$S(x) = \arg \min_y \{xy; 0 \leq y \leq 1\}.$$

O modelo acima corresponde ao modelo BLP, onde  $\mathfrak{S} \equiv S$ . O ponto  $(\bar{x}, \bar{y}) = (0, 0)$  é solução ótima local de BLP, mas  $\bar{x} = 0$  não é solução ótima local de (2.4)-(2.5). De fato, qualquer vizinhança  $W_\epsilon(0, 0) = (-\epsilon, \epsilon) \times (-\epsilon, \epsilon)$  de  $(\bar{x}, \bar{y})$ , com  $\epsilon \in (0, 1)$ , é tal que  $F(x, y) = x \geq 0$  para todo  $(x, y) \in W_\epsilon(0, 0)$  com  $y \in S(x)$  e  $-1 \leq x \leq 1$ . Isto porque não há  $y \in (-\epsilon, \epsilon) \cap S(x)$  para  $x < 0$ , visto que  $S(x) = \{1\}$  para  $x < 0$ . Isto mostra que  $(\bar{x}, \bar{y})$  é solução ótima local de BLP. Por outro lado, tome uma vizinhança  $V(0) = (-\epsilon, \epsilon)$  de  $\bar{x} = 0$ . Para qualquer  $x \in V(0) \setminus \{0\}$ , existe  $y \in S(x)$  com  $(x, y) \in \Omega_u = \{(x, y); -1 \leq x \leq 1\}$ . De fato, para  $x < 0$  tome  $y = 1 \in \{1\} = S(x)$ , e para  $x > 0$  tome  $y = 0 \in \{0\} = S(x)$ . Isto mostra que  $x \in \text{dom } \varphi_o$  e que  $\bar{x}$  não é minimizador local de  $\varphi_o(x) = x$ , isto é,  $\bar{x}$  não é solução ótima local de (2.4)-(2.5).  $\square$

O Exemplo 2.0.4 denuncia a diferença entre tomar, para cada  $x \in \Omega_x$ ,  $\mathfrak{S}(x) = \arg \min_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\}$  e  $\mathfrak{S}(x) = S(x)$ . É uma questão de interpretação: a primeira igualdade diz que o escravo é quem decide que soluções ótimas entrega para o líder, no caso, as melhores. Só daí o líder escolhe, dentre estas soluções que lhe foram oferecidas, aquelas que melhor lhe convém; já a segunda igualdade dá liberdade para o líder escolher, dentre todas as soluções ótimas do escravo, aquela que melhor lhe convém. Vários autores, no entanto, tratam a segunda interpretação, o BLP, como o modelo do problema de dois níveis (veja por exemplo [12]). Stephen Dempe em sua tese [14] é um dos poucos que consideram essa diferença. Mostraremos no Capítulo 3 que a recíproca do item (iii) do Teorema 2.0.3 para solução ótima local é verdadeira quando as restrições e funções do problema são lineares. Neste caso, é indiferente qual igualdade se toma. Isso evidentemente facilita o estudo da abordagem otimista dessa classe de problemas, e será esse nosso principal objetivo.

Notemos que ao fazer  $\mathfrak{S}(x) = \arg \min_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\}$  ou  $\mathfrak{S}(x) = \arg \max_y \{F(x, y); (x, y) \in \Omega_u, y \in S(x)\}$ , é indiferente minimizar  $F$  em  $x$  e  $y$  ou somente

<sup>4</sup>A fim de simplificar a escrita, usaremos o termo “problema otimista de dois níveis” para nos referirmos tanto ao modelo (2.2) quanto à BLP. Isso será empregado durante todo o trabalho quando conveniente. Quando não houver confusão, poderemos ainda omitir o termo “otimista”.

em  $x$  no modelo (2.2).

Também, há uma diferença sutil em considerar ou não a restrição  $(x, y) \in \Omega_u$  em  $\mathfrak{S}$ , como mostra o próximo exemplo:

**Exemplo 2.0.5.** Considere o problema otimista de dois níveis

$$\min_{x, y_1, y_2} -y_2 \quad \text{s.a.} \quad -2y_1 + y_2 \leq 0, \quad 0 \leq x \leq 3, \quad (y_1, y_2) \in \mathfrak{S}(x) \subset S(x) \quad (2.10)$$

onde

$$S(x) = \arg \min_{y_1, y_2} \{-y_1 - y_2; y_1 + y_2 \leq x, y_1, y_2 \geq 0\}.$$

É fácil ver que para  $x \geq 0$  temos  $S(x) = \{(y_1, y_2); y_1 + y_2 = x, y_1, y_2 \geq 0\}$ , e para  $x < 0$ ,  $S(x) = \emptyset$ .

**CASO 1:** Se  $\mathfrak{S}(x) = \arg \min_{y_1, y_2} \{-y_2; (y_1, y_2) \in S(x)\}$ , temos

$$\mathfrak{S}(x) = \arg \min_{y_1, y_2} \{-y_2; y_1 + y_2 = x, y_1, y_2 \geq 0\} = \{(y_1, y_2) = (0, x)\}.$$

Com isso, a restrição de alto nível  $-2y_1 + y_2 \leq 0$  implica  $x \leq 0$ , e como  $x \geq 0$ , necessariamente  $(0, 0, 0)$  é o único ponto viável para (2.10). Portanto, neste caso,  $(0, 0, 0)$  é a solução ótima global do problema de dois níveis.

**CASO 2:** Se  $\mathfrak{S}(x) = \arg \min_{y_1, y_2} \{-y_2; -2y_1 + y_2 \leq 0, 0 \leq x \leq 3, (y_1, y_2) \in S(x)\}$ , temos

$$\mathfrak{S}(x) = \arg \min_{y_1, y_2} \{-y_2; -2y_1 + y_2 \leq 0, 0 \leq x \leq 3, y_1 + y_2 = x, y_1, y_2 \geq 0\}.$$

Resolvendo o problema em  $\mathfrak{S}(x)$  (nas variáveis  $y_1$  e  $y_2$ ), encontramos  $\mathfrak{S}(x) = \{(y_1, y_2) = (\frac{x}{3}, \frac{2x}{3})\}$  para  $0 \leq x \leq 3$  e  $\mathfrak{S}(x) = \emptyset$  para o resto. Daí  $\{(x, y_1, y_2); 0 \leq x \leq 3, (y_1, y_2) \in \mathfrak{S}(x)\}$  é o conjunto viável de (2.10), e neste caso  $(3, 1, 2)$  é a solução ótima global do problema (2.10).

A figura 1 ilustra o problema. Nela são representados o conjunto  $S(\bar{x})$ , a restrição  $-2y_1 + y_2 \leq 0$  e o subconjunto do  $\mathbb{R}^3$  formado pela restrição  $y_1 + y_2 \leq x$  quando variamos  $x$ , os dois últimos pelas linhas com pontilhado forte. O ponto em destaque sobre o eixo  $y_2$  representa o conjunto  $\mathfrak{S}(\bar{x})$  no CASO 1, e o outro ponto sobre a reta  $y_1 + y_2 = \bar{x}$ , o conjunto  $\mathfrak{S}(\bar{x})$  no CASO 2. Observamos que o ponto do CASO 1 não satisfaz  $-2y_1 + y_2 \leq 0$  quando  $\bar{x} > 0$ .  $\square$

Campêlo [11], apesar de considerar o problema de dois níveis com restrições no alto nível, as retira da escrita de  $\mathfrak{S}$ . O principal motivo de considerarmos essas restrições em

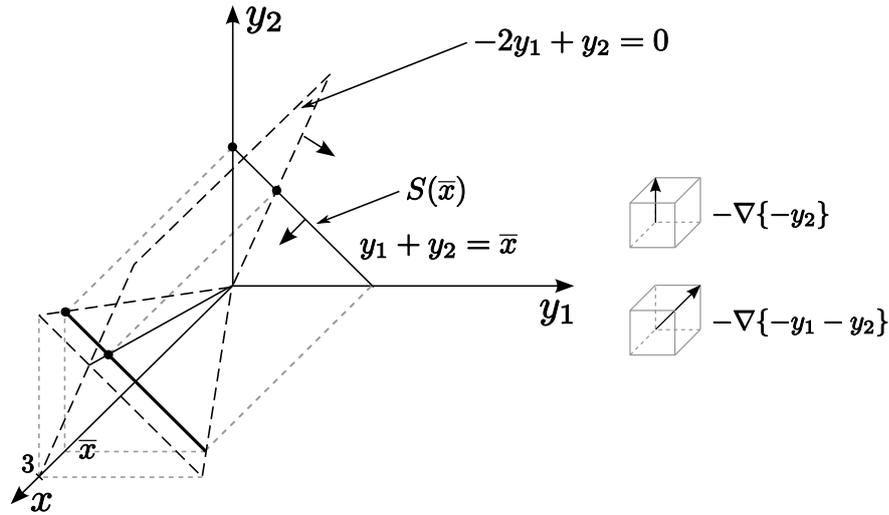


Figura 1: Ilustração para o Exemplo 2.0.5 mostrando a diferença em considerar ou não restrições de alto nível em  $\mathfrak{S}$ .

$\mathfrak{S}(x)$  é que, sem elas, a equivalência entre soluções ótimas globais de (2.2) otimista e de BLP, dada pelo Teorema 2.0.3, se perde quando as restrições de alto nível dependem da escolha do baixo nível. De fato, observemos o Exemplo 2.0.5. Se  $\mathfrak{S}(x)$  for como no CASO 1, onde (2.10) corresponde ao modelo BLP pois  $\mathfrak{S} \equiv S$ , a solução ótima global de BLP seria  $(3, 1, 2)$ , e não  $(0, 0, 0)$ .

Seguiremos agora com o conceito de semicontinuidade de funções ponto a conjunto (veja o Apêndice A.4), que será útil no relacionamento entre soluções locais otimistas de (2.2) e soluções ótimas locais de BLP.

**Definição 2.0.6.** *Uma função ponto a conjunto  $\Gamma : Z \subset \mathbb{R}^p \rightarrow 2^{\mathbb{R}^q}$  é dita semicontínua superior em  $\bar{z} \in Z$  se para cada aberto  $A \supset \Gamma(\bar{z})$  existir um aberto  $V \ni \bar{z}$  tal que  $\Gamma(z) \subset A$  para todo  $z \in V \cap Z$ .*

$\Gamma$  é dita semicontínua inferior em  $\bar{z} \in Z$  se para cada aberto  $A \subset \mathbb{R}^q$  com  $\Gamma(\bar{z}) \cap A \neq \emptyset$  existir um aberto  $V \ni \bar{z}$  tal que  $\Gamma(z) \cap A \neq \emptyset$  para todo  $z \in V \cap Z$ .

$\Gamma$  é dita contínua<sup>5</sup> em  $\bar{z} \in Z$  se é semicontínua inferior e superior em  $\bar{z}$ .

**Teorema 2.0.7.** *Se  $(\bar{x}, \bar{y})$  é solução ótima local de BLP onde  $S$  é semicontínua superior em  $\bar{x}$ , e  $S(\bar{x})$  contém um único elemento, então  $(\bar{x}, \bar{y})$  é solução local otimista de (2.2).*

*Demonstração.* Primeiramente vemos que  $|S(\bar{x})| = 1 \Rightarrow S(\bar{x}) = \{\bar{y}\} \Rightarrow F(\bar{x}, \bar{y}) = \varphi_o(\bar{x})$ . Daí  $\bar{x}$  é viável para (2.4)-(2.5). Por hipótese, existe uma vizinhança aberta  $V(\bar{x}, \bar{y})$  de

<sup>5</sup>O leitor não deve confundir continuidade de funções ponto a conjunto com continuidade de funções de imagens com único valor.

$(\bar{x}, \bar{y})$  tal que  $F(\bar{x}, \bar{y}) \leq F(x, y), \forall (x, y) \in \Omega_u \cap V(\bar{x}, \bar{y})$  com  $y \in S(x)$ . Seja  $A \supset S(\bar{x})$  aberto. Como  $S$  é semicontínua superior em  $\bar{x}$ , existe uma vizinhança aberta  $V(\bar{x})$  de  $\bar{x}$  tal que  $S(x) \subset A, \forall x \in V(\bar{x})$  (observamos que  $x \notin \Omega_x \Rightarrow S(x) = \emptyset$ ). Do fato de  $S(\bar{x})$  conter um único elemento, conseguimos que  $A$  e  $V(\bar{x})$  sejam tais que  $V(\bar{x}) \times A \subset V(\bar{x}, \bar{y})$ . Então para cada  $x \in V(\bar{x}) \cap \text{dom } \varphi_o$  temos  $\varphi_o(\bar{x}) = F(\bar{x}, \bar{y}) \leq F(x, y), \forall y \in S(x) \subset A$  com  $(x, y) \in \Omega_u$ . Daí  $\varphi_o(\bar{x}) \leq \varphi_o(x), \forall x \in V(\bar{x}) \cap \text{dom } \varphi_o$ , isto é,  $(\bar{x}, \bar{y})$  é solução local otimista de (2.2).  $\square$

A condição  $|S(\bar{x})| = 1$  imposta no Teorema 2.0.7 é importante, como mostra o Exemplo 2.0.4: temos  $\varphi_o(0) = F(0, 0) = 0, S$  semicontínua superior em 0 (de fato, qualquer  $A \supset S(0) = [0, 1]$  contém  $S(x) = \{0\}, x > 0$ , e  $S(x) = \{1\}, x < 0$ ), mas  $S(0)$  contém mais de um elemento.

O quadro 2 a seguir resume os resultados deste capítulo para o caso otimista.

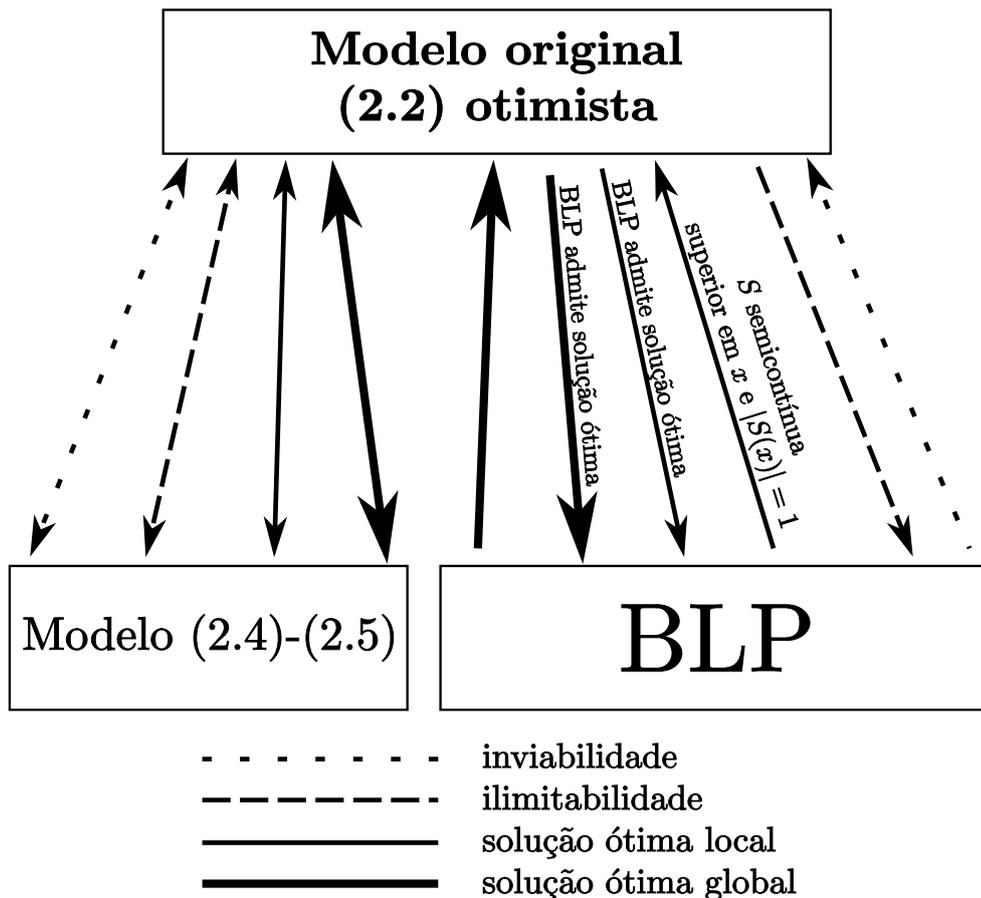


Figura 2: Relações entre os modelos (2.2) otimista, (2.4)-(2.5) e BLP.

A seguir, uma observação que nos ajudará a entender posteriormente a geometria de um problema de dois níveis com funções objetivo e restrições lineares. Considere a função

$S : \Omega_x \rightarrow 2^{\mathbb{R}^{n_y}}$ , definida por (2.1) (página 16). Definimos o *gráfico de  $S$*  como sendo

$$\text{graf } S := \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; y \in S(x)\}.$$

Assim podemos escrever BLP como

$$\min_{x,y} F(x, y) \quad \text{s.a.} \quad (x, y) \in \Omega_u \cap \text{graf } S.$$

Fizemos essa observação aqui simplesmente porque  $\text{graf } S$  é definido de maneira geral, sem que necessariamente  $S(x)$  seja um problema de programação linear.

## 3 O caso linear

Neste capítulo, estudaremos os *problemas lineares de dois níveis* onde as funções objetivo e restrições de alto e baixo níveis são lineares. Daremos descrições geométricas do comportamento desses problemas, bem como alguns outros resultados específicos. Usaremos adiante esses resultados para desenvolver técnicas de resolução dessa classe de problemas. Vale ressaltar que o modelo aqui estudado corresponde ao modelo BLP, onde  $\mathfrak{S} \equiv S$ .

### 3.1 Delimitação teórica

A menos que mencionado explicitamente, trabalharemos com o problema linear otimista dado pelo modelo LBLP<sup>1</sup>

$$\begin{aligned} \text{LBLP: } \min_{x,y} \{c_1x + d_1y\} \\ \text{s.a. } A_1x + B_1y \leq b_1, \quad x \geq 0 \\ y \in \arg \min_y d_2y \\ \text{s.a. } A_2x + B_2y \leq b_2, \quad y \geq 0 \end{aligned}$$

onde  $c_1 \in \mathbb{R}^{n_x}$ ,  $d_1, d_2 \in \mathbb{R}^{n_y}$ ,  $A_1 \in \mathbb{R}^{m_u \times n_x}$ ,  $B_1 \in \mathbb{R}^{m_u \times n_y}$ ,  $A_2 \in \mathbb{R}^{m_l \times n_x}$ ,  $B_2 \in \mathbb{R}^{m_l \times n_y}$ ,  $b_1 \in \mathbb{R}^{m_u}$  e  $b_2 \in \mathbb{R}^{m_l}$ .

Relativo ao modelo acima, temos que

- o *poliedro do ULP* é  $\Omega_u = \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; A_1x + B_1y \leq b_1, x \geq 0\}$ ;
- $\Omega_y(x) = \{y \in \mathbb{R}^{n_y}; B_2y \leq b_2 - A_2x, y \geq 0\}$ ;
- $S(x) = \arg \min_y \{d_2y; y \in \Omega_y(x)\}$ .

Além disso, faremos a seguinte definição:

---

<sup>1</sup>Do inglês *Linear BiLevel Problem*.

**Definição 3.1.1.** 1) O poliedro do problema de baixo nível é

$$\Omega_l = \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; A_2x + B_2y \leq b_2, y \geq 0\}.$$

2) O conjunto viável é

$$\Omega = \Omega_u \cap \Omega_l = \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; A_kx + B_ky \leq b_k, k = 1, 2, x, y \geq 0\}.$$

$(x, y)$  é dito racional se  $x \in \Omega_x$  e  $y \in S(x)$  (ou equivalentemente, se  $(x, y) \in \text{graf } S$ ).

3) A função valor ótimo é a função  $v : \Omega_x \rightarrow \mathbb{R}$  definida por

$$v(x) = d_2y, \quad y \in S(x) \quad \text{ou ainda} \quad v(x) = \min_y \{d_2y; y \in \Omega_y(x)\}.$$

4) O conjunto admissível (ou região induzida) é

$$\Upsilon = \text{graf } S \cap \Omega_u = \{(x, y) \in \Omega_u; y \in S(x)\}.$$

$(x, y)$  é dito admissível se  $(x, y) \in \Upsilon$ .

Um conceito que usaremos é o de face de conjuntos poliedrais.

**Definição 3.1.2.** Dado um conjunto poliedral  $D = \{z \in \mathbb{R}^n; Az \leq b\}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ , um subconjunto  $Q \subset D$  é dito ser uma face de  $D$  se existir um conjunto  $J \subset \{1, \dots, m\}$  tal que

$$Q = Q(J) = \{z \in D; A_jz = b_j, j \in J\}.$$

Dados  $J$  e sua face associada  $Q(J)$ , dizemos que  $Q$  é face  $d$ -dimensional de  $D$  (ou face de dimensão  $d$ ),  $0 \leq d \leq n$ , se existir  $z_0 \in Q(J)$  tal que  $A_jz_0 < b_j, \forall j \notin J$ , e a matriz  $[A_i], i \in J$ , tem posto  $n - d$ . Mais ainda, dizemos que  $Q(J)$  é face não degenerada se  $|J| = n - d$ .

De acordo com o Teorema A.3.3 (página 122), um vértice é uma face 0-dimensional. Comumente diz-se que uma face 1-dimensional é uma *aresta*. Se  $\text{int } D \neq \emptyset$  então  $D$  é uma face  $n$ -dimensional de  $D$ . Notemos ainda que cada face é um conjunto poliedral.

A grande dificuldade dos problemas lineares de dois níveis está na determinação de  $S(x)$ . É importante então estudarmos características da função  $S$  que mapeia  $\Omega_x$  às soluções ótimas do problema de baixo nível. Começaremos com observações importantes para os próximos resultados deste capítulo. Usando dualidade,  $y \in S(x)$  se, e somente se,

existir  $\lambda \in \mathbb{R}^{m_l}$  tal que

$$\begin{aligned} A_2x + B_2y \leq b_2, \quad y \geq 0, \quad \lambda \geq 0, \quad \lambda^t(A_2x + B_2y - b_2) = 0, \\ B_2^t\lambda + d_2^t \geq 0, \quad y^t(B_2^t\lambda + d_2^t) = 0. \end{aligned} \quad (3.1)$$

Para conjuntos  $I \subset \{1, \dots, n_y\}$  e  $J \subset \{1, \dots, m_l\}$ , consideremos o conjunto  $M(I, J)$  das soluções  $(x, y, \lambda) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{m_l}$  do sistema

$$(A_2x + B_2y - b_2)_i = 0, i \in J \quad (3.2)$$

$$(A_2x + B_2y - b_2)_i \leq 0, i \notin J \quad (3.3)$$

$$y_j = 0, j \in I, \quad y_j \geq 0, j \notin I \quad (3.4)$$

$$\lambda_i = 0, i \notin J, \quad \lambda_i \geq 0, i \in J \quad (3.5)$$

$$(B_2^t\lambda + d_2^t)_j = 0, j \notin I, \quad (B_2^t\lambda + d_2^t)_j \geq 0, j \in I. \quad (3.6)$$

Existem apenas finitos  $M(I, J)$  pois existem apenas finitos pares  $(I, J)$ . Agora, se  $W$  é o conjunto das soluções de (3.1), temos  $W = \bigcup M(I, J)$ . De fato, a inclusão  $\supset$  é clara, e para a inclusão  $\subset$  basta definirmos os conjuntos  $I$  e  $J$  de acordo com as coordenadas nulas de  $y$  e  $\lambda$ .

Notemos ainda que pelo sistema (3.2)-(3.6) é suficiente considerar apenas finitos vetores duais  $\lambda^1, \dots, \lambda^r$  na descrição de graf  $S$ , um para cada par  $(I, J)$ . Isso será usado na demonstração de alguns resultados adiante.

O resultado a seguir revela uma importante característica da região dos pontos racionais do problema linear de dois níveis. É a característica central para outros resultados que se seguirão.

**Teorema 3.1.3.** *Consideremos a função ponto a conjunto  $S : \mathbb{R}^{n_x} \rightarrow 2^{\mathbb{R}^{n_y}}$  dada por*

$$S(x) = \arg \min_y \{d_2y; B_2y \leq b_2 - A_2x, y \geq 0\}.$$

*Então  $\text{graf } S = \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; y \in S(x)\}$  é igual à uma união finita de faces de  $\Omega_l$ . Em particular,  $\text{graf } S$  é fechado.*

*Demonstração.* Consideremos o sistema (3.2)-(3.6). Se  $M(I, J) \neq \emptyset$ , definimos o conjunto

$$PM(I, J) = \{(x, y); \text{ existe } \lambda \text{ tal que } (x, y, \lambda) \in M(I, J)\}.$$

$PM(I, J)$  é o conjunto das soluções do sistema (3.2)-(3.4), visto que essas restrições são independentes de (3.5) e (3.6) e um  $\lambda$  correspondente existe. Imediatamente segue que

cada  $PM(I, J)$  é face de  $\Omega_l$ , e é um conjunto fechado.

Vamos mostrar que  $\text{graf } S = \bigcup PM(I, J)$ , donde, pelo fato de existirem finitos  $PM(I, J)$ , segue diretamente o resultado. Dado  $(x, y) \in \text{graf } S$ , existe  $\lambda$  tal que  $(x, y, \lambda) \in W = \bigcup M(I, J)$ . Daí  $(x, y, \lambda) \in M(I_0, J_0)$  para algum par  $(I_0, J_0)$  e logo  $(x, y) \in PM(I_0, J_0) \subset \bigcup PM(I, J)$ . Por outro lado, dado  $(x, y) \in \bigcup PM(I, J)$ , consideremos um par  $(I_0, J_0)$  tal que  $(x, y) \in PM(I_0, J_0)$ . A esse conjunto temos associado um  $\lambda$ , de forma que  $(x, y, \lambda) \in M(I_0, J_0) \subset W$ . Portanto  $(x, y) \in \text{graf } S$ .  $\square$

A fim de estendermos nosso estudo sobre as características do problema de baixo nível, definimos

$$\bar{\Omega}_x := \{x \in \mathbb{R}^{n_x}; \Omega_y(x) \neq \emptyset\} = \{x \in \mathbb{R}^{n_x}; \exists y \text{ tal que } B_2 y \leq b_2 - A_2 x, y \geq 0\}.$$

O próximo resultado revela uma questão de ordem prática: se quisermos saber se  $S(x)$  é não vazio, basta vermos se  $\Omega_y(x)$  o é, o que é mais razoável do ponto de vista computacional. Além disso será usado na demonstração de alguns resultados posteriores.

**Lema 3.1.4.** *i) O conjunto  $\bar{\Omega}_x$  é poliedral (portanto convexo e fechado);*

*ii) Se  $\Omega_x \neq \emptyset$  ou  $\Omega_l$  for compacto então  $\Omega_x = \bar{\Omega}_x$ . Em particular,  $\Omega_x$  é poliedral.*

*Demonstração.* *i)* Consideremos a aplicação  $\pi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_x}$  dada por  $\pi(x, y) = x$ . Observamos que  $\bar{\Omega}_x = \pi(\Omega_l)$ . Pelo Teorema da representação de Minkowski-Weyl (Teorema A.3.2, página 121), existem um conjunto finito  $E = \{(x^k, y^k) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; k = 1, \dots, K\}$  e um cone  $\mathcal{R} = \text{cone}\{d^p \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; p = 1, \dots, P\}$  finitamente gerado tais que

$$\begin{aligned} \Omega_l &= \text{conv } E + \mathcal{R} \\ &= \left\{ (x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; (x, y) = \sum_{k=1}^K a^k (x^k, y^k) + (d_x, d_y), \right. \\ &\quad \left. \sum_{k=1}^K a^k = 1, a^k \in \mathbb{R}_+, (d_x, d_y) \in \mathcal{R} \right\}. \end{aligned} \tag{3.7}$$

Mostremos que  $\pi(\Omega_l) = \text{conv } \pi(E) + \pi(\mathcal{R})$ . Seja  $x \in \pi(\Omega_l)$ . De  $\bar{\Omega}_x = \pi(\Omega_l)$ , segue

que existe  $y \in \Omega_y(x)$  (e portanto  $(x, y) \in \Omega_l$ ). Escrevamos  $(x, y)$  como em (3.7). Temos

$$\begin{aligned} x &= \pi(x, y) = \pi \left( \sum_{k=1}^K a^k(x^k, y^k) + (d_x, d_y) \right) \\ &= \sum_{k=1}^K a^k x^k + d_x \in \text{conv } \pi(E) + \pi(\mathcal{R}). \end{aligned}$$

Seja agora  $x \in \text{conv } \pi(E) + \pi(\mathcal{R})$ . Como  $\pi(E) = \{x^k; k = 1, \dots, K\}$ , existem finitos  $a^k$  tais como em (3.7) e  $d_x \in \pi(\mathcal{R})$  de forma que  $x = \sum_{k=1}^K a^k x^k + d_x$ . Pela definição de  $\pi$ , existem  $d_y$  e  $y_k \in \Omega_y(x_k)$ ,  $k = 1, \dots, K$ , tais que  $(d_x, d_y) \in \mathcal{R}$  e  $(x, y) \in \Omega_l$ , onde  $y = \sum_{k=1}^K a^k y^k + d_y$ . Segue que  $x \in \pi(\Omega_l)$  e logo  $\pi(\Omega_l) = \text{conv } \pi(E) + \pi(\mathcal{R})$ .

Temos que  $\pi(\mathcal{R}) = \text{cone } \{\pi(d^p); p = 1, \dots, P\}$ , isto é,  $\pi(\mathcal{R})$  é um cone finitamente gerado. Então, novamente pelo Teorema da representação de Minkowski-Weyl,  $\bar{\Omega}_x = \pi(\Omega_l)$  é poliedral, como queríamos demonstrar. A Figura 3 adiante ilustra a situação.

ii) A inclusão  $\subset$  é óbvia, visto que  $S(x) \subset \Omega_y(x)$  para todo  $x$ . Suponhamos que  $\Omega_x \neq \emptyset$  e mostremos a inclusão contrária. Considerando a função valor ótimo  $v(x) = \min_y \{d_2 y; y \in \Omega_y(x)\}$ , temos que

$$v(x) > -\infty \text{ para algum } x \in \bar{\Omega}_x \quad \Rightarrow \quad v(x) > -\infty \text{ para todo } x \in \bar{\Omega}_x$$

[43, Lema A2]. Tomando então um  $x \in \Omega_x \subset \bar{\Omega}_x$ , vem que  $S(x) \neq \emptyset$  para todo  $x \in \bar{\Omega}_x$ , isto é,  $\Omega_x \supset \bar{\Omega}_x$ .

Agora, supomos que  $\Omega_l$  seja compacto e mostremos que  $\Omega_x \supset \bar{\Omega}_x$ . Se  $\Omega_l = \emptyset$  nada há a fazer pois neste caso  $\Omega_x \subset \bar{\Omega}_x = \emptyset$ . Caso contrário, em que  $\Omega_l \neq \emptyset$ , existe  $x \in \bar{\Omega}_x$ , e temos que  $\Omega_y(x)$  é compacto. Como  $v$  é contínua (item (i) do Teorema 3.1.6 adiante, página 33), segue do Teorema de Weierstrass (Teorema A.2.4, página 120) que  $v(x) > -\infty$ , isto é,  $S(x) \neq \emptyset$ . Daí  $\Omega_x \neq \emptyset$  e caímos na argumentação anterior.  $\square$

A seguir, estudaremos mais algumas características geométricas do problema de baixo nível.

**Lema 3.1.5.** *Consideremos a função ponto a conjunto  $S : \mathbb{R}^{n_x} \rightarrow 2^{\mathbb{R}^{n_y}}$  dada por*

$$S(x) = \arg \min_y \{d_2 y; B_2 y \leq b_2 - A_2 x, y \geq 0\}.$$

*Então  $S$  é semicontínua inferior em todo  $x \in \Omega_x = \{x \in \mathbb{R}^{n_x}; S(x) \neq \emptyset\}$ . Se  $\Omega_l$  for compacto então  $S$  é contínua em todo  $x \in \Omega_x$ .*

*Demonstração.* Mostremos inicialmente a semicontinuidade superior no caso  $\Omega_l$  compacto.

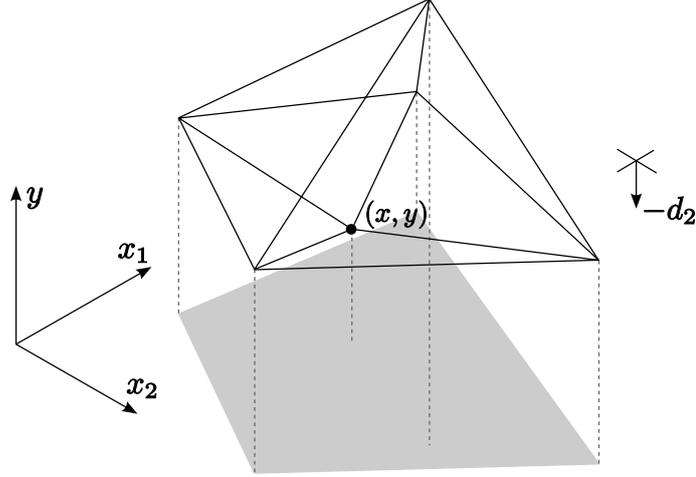


Figura 3: Projeção  $\bar{\Omega}_x = \pi(\Omega_l)$  de  $\Omega_l$  sobre  $\mathbb{R}^{n_x}$ .

Como  $\Omega_x$  é fechado (item (ii) do Lema 3.1.4), pelo Teorema A.4.3 (página 126) é suficiente mostrar que, para todo fechado  $M \subset \mathbb{R}^{n_y}$ , a imagem inversa de  $M$  por  $S$ ,

$$S^{-1}(M) = \{x \in \mathbb{R}^{n_x}; S(x) \cap M \neq \emptyset\},$$

é um conjunto fechado. Tomemos  $M \subset \mathbb{R}^{n_y}$  fechado. Se  $S^{-1}(M) = \emptyset$ , nada há a fazer. Suponhamos o contrário e tomemos uma sequência  $\{x_n\} \subset S^{-1}(M) \subset \Omega_x$  com  $\lim x_n = \bar{x}$ . Afirmamos que  $\text{dist}(\{\bar{x}\} \times M, \text{graf } S) = 0$ . De fato, para cada  $x_n$  existe  $y_n \in S(x_n) \cap M$ . Temos  $(x_n, y_n) \in \text{graf } S$ ,  $(\bar{x}, y_n) \in \{\bar{x}\} \times M$  e  $\|(x_n, y_n) - (\bar{x}, y_n)\|_\infty = \|x_n - \bar{x}\|_\infty \rightarrow 0$  quando  $n \rightarrow \infty$ .

Agora, como  $\{\bar{x}\} \times M$  é fechado,  $\text{graf } S$  compacto (pois  $\Omega_l$  é compacto) e  $\text{dist}(\{\bar{x}\} \times M, \text{graf } S) = 0$ , existe  $\bar{y} \in M$  com  $(\bar{x}, \bar{y}) \in \text{graf } S$  (Teorema A.2.5, página 120). Daí  $\bar{y} \in S(\bar{x}) \cap M \Rightarrow \lim x_n = \bar{x} \in S^{-1}(M)$ , e logo  $S$  é semicontínua superior em  $\bar{x}$ .

Mostremos agora a semicontinuidade inferior. Fixemos  $\bar{x} \in \Omega_x$  e tomemos um aberto  $A \in \mathbb{R}^{n_y}$  tal que  $S(\bar{x}) \cap A \neq \emptyset$ . Fixado  $\bar{y} \in S(\bar{x}) \cap A$ , existe uma bola aberta  $B_\delta^\infty(\bar{y})$  tal que  $B_\delta^\infty(\bar{y}) \subset A$ . Se  $A_2 = 0$  então  $S(x) = S(\bar{x})$  para todo  $x \in \Omega_x$ , e nada há a fazer. Supomos então  $A_2 \neq 0$  ( $\Rightarrow \|A_2\|_\infty > 0$ ) e fazemos

$$\epsilon = \frac{\delta}{n_y \Delta \|A_2\|_\infty},$$

onde  $\Delta > 0$  é tal que, para cada submatriz inversível  $B$  de  $A_2$ , os módulos das entradas de  $B^{-1}$  não ultrapassem  $\Delta$ . Pelo Teorema A.3.5 (página 122), para cada  $x \in \Omega_x \cap B_\epsilon^\infty(\bar{x})$

existe  $y \in S(x)$  com

$$\begin{aligned} \|\bar{y} - y\|_\infty &\leq n_y \Delta \|(b_2 - A_2 \bar{x}) - (b_2 - A_2 x)\|_\infty \\ &= n_y \Delta \|A_2(x - \bar{x})\|_\infty \leq n_y \Delta \|A_2\|_\infty \|\bar{x} - x\|_\infty, \end{aligned} \quad (3.8)$$

onde a última desigualdade segue das propriedades das normas. Daí  $\|\bar{y} - y\|_\infty \leq n_y \Delta \|A_2\|_\infty \epsilon = \delta$ , e logo  $y \in S(x) \cap B_\delta^\infty(\bar{y}) \subset A$ . Segue que  $S(x) \cap A \neq \emptyset$ , donde concluímos que  $S$  é semicontínua inferior em  $\bar{x}$ .  $\square$

Sobre o Lema anterior, Böhm [6] demonstra a semicontinuidade superior de  $S$  supondo apenas que  $S(x)$  seja compacto para algum  $x \in \Omega_x$ .

**Teorema 3.1.6.** *i) A função valor ótimo  $v(x) = \min_y \{d_2 y; B_2 y \leq b_2 - A_2 x, y \geq 0\}$  é linear por partes, contínua e convexa;*

*ii) graf  $S$  é conexo por caminhos<sup>2</sup>.*

*Demonstração.* *i)* Tomemos vetores duais  $\lambda^1, \dots, \lambda^r$  que descrevem graf  $S$  pelo sistema (3.2)-(3.6) (página 29). É claro que  $B_2^t \lambda^i + d_2^t \geq 0, \lambda^i \geq 0, \forall i$ . Usando o dual do problema  $v(x) = \min_y \{d_2 y; y \in \Omega_y(x)\}$  ( $x \in \Omega_x$ ) temos

$$\max_{1 \leq i \leq r} \{\lambda^{it}(A_2 x - b_2); B_2^t \lambda^i + d_2^t \geq 0, \lambda^i \geq 0\} = \max_{1 \leq i \leq r} \{(\lambda^{it} A_2)x - \lambda^{it} b_2\} = v(x),$$

donde segue que  $v$  é linear por partes.

Tomemos  $\bar{x} \in \Omega_x$  e  $\bar{y} \in S(\bar{x})$ , e mostremos que  $v$  é contínua em  $\bar{x}$ . Pela Desigualdade de Cauchy-Schwarz (Teorema A.2.2, página 118), temos

$$|v(\bar{x}) - v(x)| = |d_2 \bar{y} - d_2 y| \leq \|d_2\| \|\bar{y} - y\|,$$

onde  $y \in S(x)$  é escolhido de modo a satisfazer (3.8). Tomemos  $a > 0$  tal que  $\|\bar{y} - y\| \leq a \|\bar{y} - y\|_\infty, \forall y \in \mathbb{R}^{n_y}$ , cuja existência é assegurada pelo Teorema A.2.1 (página 117). Assim, pela expressão (3.8) vem

$$|v(\bar{x}) - v(x)| \leq \|d_2\| \|\bar{y} - y\| \leq a \|d_2\| \|\bar{y} - y\|_\infty \leq a n_y \Delta \|d_2\| \|A_2\|_\infty \|\bar{x} - x\|_\infty,$$

onde  $\Delta$  é como na demonstração do Lema 3.1.5. Isto mostra que  $v$  é contínua em  $\bar{x}$ .<sup>3</sup>

<sup>2</sup>Um conjunto conexo por caminhos é sempre conexo, mas a recíproca nem sempre é verdadeira. Veja [28, página 61].

<sup>3</sup>Note que a demonstração da continuidade de  $v$  não depende do Lema 3.1.4 e portanto podemos usá-la no referido lema como fizemos (página 30).

Mostremos agora que  $v$  é convexa. Tomemos  $x_1, x_2 \in \Omega_x$ ,  $\alpha \in [0, 1]$  e consideremos  $x_\alpha = \alpha x_2 + (1 - \alpha)x_1$ . Temos  $\Omega_x \neq \emptyset$  e logo  $\Omega_x = \bar{\Omega}_x$  é poliedral (Lema 3.1.4). Daí  $x_\alpha \in \Omega_x$ . Tomemos  $y_1 \in S(x_1)$  e  $y_2 \in S(x_2)$  e consideremos  $y_\alpha = \alpha y_2 + (1 - \alpha)y_1$ . Afirmamos que  $y_\alpha \in \Omega_y(x_\alpha)$ . De fato,  $y_\alpha \geq 0$  e

$$\begin{aligned} B_2 y_\alpha &= B_2(\alpha y_2 + (1 - \alpha)y_1) = \alpha B_2 y_2 + (1 - \alpha)B_2 y_1 \\ &\leq \alpha(b_2 - A_2 x_2) + (1 - \alpha)(b_2 - A_2 x_1) = b_2 - A_2 x_\alpha. \end{aligned}$$

Daí

$$\begin{aligned} v(x_\alpha) &\leq d_2 y_\alpha = d_2(\alpha y_2 + (1 - \alpha)y_1) \\ &= \alpha(d_2 y_2) + (1 - \alpha)(d_2 y_1) = \alpha v(x_2) + (1 - \alpha)v(x_1), \end{aligned}$$

e logo  $v$  é convexa.

*ii)*  $\Omega_l$  é uma união finita de faces  $Q_1, \dots, Q_R$ , sendo algumas dessas, digamos  $Q_1, \dots, Q_r$  ( $r \leq R$ ), faces de  $\text{graf } S$  (Teorema 3.1.3). Dados  $(\tilde{x}, \tilde{y}), (\bar{x}, \bar{y}) \in \text{graf } S$ , temos que mostrar que existe um caminho em  $\text{graf } S$  ligando esses pontos. Neste caso,  $\Omega_x \neq \emptyset$  e do item *(ii)* do Lema 3.1.4 segue que  $\Omega_x = \bar{\Omega}_x$  e  $x_\alpha = \alpha \bar{x} + (1 - \alpha)\tilde{x} \in \Omega_x$ ,  $\forall \alpha \in [0, 1]$ . O segmento  $U = \{x_\alpha; \alpha \in [0, 1]\}$  está contido numa união de projeções de finitas faces de  $\text{graf } S$ , digamos  $U \subset \bigcup_{i=1}^k \pi(Q_i)$  ( $k \leq r$ ). Reindexando se necessário, podemos supor que  $\pi(Q_1), \dots, \pi(Q_k)$  aparecem em ordem crescente de índice  $i$  à medida que  $\alpha$  cresce, ou seja,  $\alpha < \beta \Leftrightarrow Q_i \ni x_\alpha, Q_j \ni x_\beta$  sejam tais que  $i \leq j$ . Então ao longo de  $U$  existem finitos  $x_{\alpha_1} = \tilde{x}, x_{\alpha_2}, \dots, x_{\alpha_k}, x_{\alpha_{k+1}} = \bar{x}$  tais que  $x_{\alpha_i}, x_{\alpha_{i+1}} \in \pi(Q_i)$ ,  $1 \leq i \leq k$ , e  $x_{\alpha_i} \in \pi(Q_{i-1}) \cap \pi(Q_i)$ ,  $2 \leq i \leq k$ . Agora tomemos para cada  $2 \leq i \leq k$  um  $y_i \in S(x_{\alpha_i})$ . Construimos assim o caminho

$$(\tilde{x}, \tilde{y}) = (x_{\alpha_1}, \tilde{y}) \rightarrow (x_{\alpha_2}, y_2) \rightarrow \dots \rightarrow (x_{\alpha_k}, y_k) \rightarrow (x_{\alpha_{k+1}}, \bar{y}) = (\bar{x}, \bar{y})$$

que pelo fato das faces  $Q_1, \dots, Q_k \subset \text{graf } S$  serem convexas, está contido em  $\text{graf } S$ , como queríamos.  $\square$

Um fato interessante para o caso linear é que a recíproca do item *(iii)* do Teorema 2.0.3 (página 21) vale para soluções ótimas locais. Esse fato é mais forte que o Teorema 2.0.7 (página 24), e deixa-nos mais confortáveis para confecção de algoritmos para soluções locais otimistas do problema linear de dois níveis. Antes de enunciá-lo e demonstrá-lo, precisamos do seguinte resultado:

**Lema 3.1.7.** *Dados vetores duais  $\lambda^1, \dots, \lambda^r$  que descrevem  $\text{graf } S$  pelo sistema (3.2)-(3.6)*

(página 29), definimos para cada  $x \in \Omega_x$

$$\lambda(x) = \{\lambda^i; \text{existem } y, I \text{ e } J \text{ tais que } (x, y, \lambda^i) \in M(I, J)\}.$$

Então

i) para cada  $\bar{x} \in \Omega_x$ , existe uma vizinhança aberta  $V(\bar{x})$  de  $\bar{x}$  tal que  $\lambda(x) \cap \lambda(\bar{x}) \neq \emptyset$ ,  
 $\forall x \in V(\bar{x}) \cap \Omega_x$ ;

ii) para cada  $\bar{x} \in \Omega_x$  tem-se  $\lambda^{it}(A_2\bar{x} - b_2) = \lambda^{jt}(A_2\bar{x} - b_2)$ ,  $\forall \lambda^i, \lambda^j \in \lambda(\bar{x})$ .

*Demonstração.* i) Seja  $\bar{x} \in \Omega_x$  e tome  $\bar{y} \in S(\bar{x})$ . Tomemos conjuntos de índices  $\bar{I}, \bar{J}$  e  $\bar{\lambda} \in \mathbb{R}^{m_i}$  tais que

$$\begin{aligned} (A_2\bar{x} + B_2\bar{y} - b_2)_i &= 0, i \in \bar{J} \\ (A_2\bar{x} + B_2\bar{y} - b_2)_i &< 0, i \notin \bar{J} \\ \bar{y}_j &= 0, j \in \bar{I}, \quad \bar{y}_j > 0, j \notin \bar{I} \\ \bar{\lambda}_i &= 0, i \notin \bar{J}, \quad \bar{\lambda}_i \geq 0, i \in \bar{J} \\ (B_2^t\bar{\lambda} + d_2^t)_j &= 0, j \notin \bar{I}, \quad (B_2^t\bar{\lambda} + d_2^t)_j \geq 0, j \in \bar{I}. \end{aligned} \tag{3.9}$$

Temos  $\bar{\lambda} \in \lambda(\bar{x})$ . Existem bolas abertas  $B_\epsilon^\infty(\bar{x})$  e  $B_\epsilon^\infty(\bar{y})$  tais que

$$(A_2\tilde{x} + B_2\tilde{y} - b_2)_i < 0, i \notin \bar{J} \quad \text{e} \quad \tilde{y}_j > 0, j \notin \bar{I}$$

para todo  $(\tilde{x}, \tilde{y}) \in B_\epsilon^\infty(\bar{x}) \times B_\epsilon^\infty(\bar{y}) = B_\epsilon^\infty(\bar{x}, \bar{y})$  (no caso extremo em que todas as restrições são ativas em  $(\bar{x}, \bar{y})$ , podemos tomar qualquer bola aberta contendo  $(\bar{x}, \bar{y})$ ). Temos  $B_\epsilon^\infty(\bar{y}) \cap S(\bar{x}) \supset \{\bar{y}\} \neq \emptyset$  e, como  $S$  é semicontínua inferior em  $\bar{x}$  (Lema 3.1.5), existe  $\delta \leq \epsilon$  tal que  $S(x) \cap B_\delta^\infty(\bar{y}) \neq \emptyset, \forall x \in B_\delta^\infty(\bar{x}) \cap \Omega_x$ . Isto implica que para cada  $x \in B_\delta^\infty(\bar{x}) \cap \Omega_x$ , existe  $y \in S(x)$  tal que  $\|\bar{y} - y\|_\infty < \epsilon$ . Portanto, para estes pontos  $(x, y)$ ,

$$\|(\bar{x}, \bar{y}) - (x, y)\|_\infty = \max\{\|\bar{x} - x\|_\infty, \|\bar{y} - y\|_\infty\} < \max\{\delta, \epsilon\} = \epsilon,$$

isto é,  $(x, y) \in B_\epsilon(\bar{x}, \bar{y})$ . Observamos que as restrições de igualdade em (3.9) podem tornar-se desigualdades estritas em  $(x, y)$ . No entanto, um  $\lambda$  solução de (3.9) associado a  $(x, y)$  é tal que todas as igualdades em (3.9) relativas a  $\bar{\lambda}$  verificam-se também para  $\lambda$ , isto porque as desigualdades estritas associadas a  $(\bar{x}, \bar{y})$  continuam para  $(x, y)$ . Segue que  $\lambda \in \lambda(\bar{x}) \cap \lambda(x)$  e que  $B_\delta^\infty(\bar{x})$  é uma tal vizinhança de  $\bar{x}$  procurada.

ii) Seja  $\bar{x} \in \Omega_x$  e tomemos  $\lambda^1, \lambda^2 \in \lambda(\bar{x})$ . Existem  $y_1, y_2 \in S(\bar{x})$  satisfazendo (3.1) (página 29) para  $\lambda^1$  e  $\lambda^2$ , respectivamente. Agora, temos  $\lambda^{1t}(A_2\bar{x} - b_2) = d_2y_1$  pois

$\lambda^{1t}(A_2\bar{x} - b_2) - d_2y_1 = -\lambda^{1t}B_2y_1 - d_2y_1 = 0$  (o mesmo vale para  $\lambda^2$ ). Daí

$$\lambda^{1t}(A_2\bar{x} - b_2) = d_2y_1 = v(\bar{x}) = d_2y_2 = \lambda^{2t}(A_2\bar{x} - b_2).$$

□

A geometria por detrás da definição de  $\lambda(x)$  no Lema 3.1.7 é interessante. Podem haver, na descrição de graf  $S$ , várias faces ativas num dado ponto  $(x, y)$  (de dimensões possivelmente distintas). Qualquer variável dual que se tenha pré-fixado para essas faces é solução do sistema (3.2)-(3.6) para esse dado  $(x, y)$ . Observamos na figura 3 a situação. Temos um vértice, arestas e pedaços de planos ativos em  $(x, y)$ . Qualquer uma dessas faces poderia entrar na descrição de graf  $S$ . Mais ainda: o vértice em questão e as arestas adjacentes não são necessárias na descrição de graf  $S$ . Bastam as faces de dimensões maiores. Isso revela que graf  $S$  não tem uma escrita única, pelo menos não minimal, do ponto de vista da quantidade de faces envolvidas. No Capítulo 6 exploraremos essa idéia para confeccionar um algoritmo para soluções ótimas locais.

Finalmente enunciamos a recíproca do item (iii) do Teorema 2.0.3 no caso linear:

**Teorema 3.1.8.** *Se  $(\bar{x}, \bar{y})$  é solução ótima local de LBLP então  $(\bar{x}, \bar{y})$  é solução local otimista do modelo (2.2) correspondente (página 17).*

*Demonstração.* Suponhamos por contradição que  $(\bar{x}, \bar{y})$  seja solução ótima local de LBLP e não solução local otimista de (2.2). Então existem uma bola aberta  $B_\epsilon^\infty(\bar{x}, \bar{y})$  tal que  $c_1\bar{x} + d_1\bar{y} \leq c_1x + d_1y, \forall (x, y) \in B_\epsilon(\bar{x}, \bar{y}) \cap \Upsilon$ , e também  $(\tilde{x}, \tilde{y}) \in \Upsilon$  com  $\|\bar{x} - \tilde{x}\|_\infty < \epsilon$  e  $c_1\tilde{x} + d_1\tilde{y} < c_1\bar{x} + d_1\bar{y}$  (podemos ter até  $\tilde{x} = \bar{x}$ ).

Definimos  $(x_\alpha, y_\alpha) = \alpha(\tilde{x}, \tilde{y}) + (1 - \alpha)(\bar{x}, \bar{y})$ . Como  $\Omega$  é convexo, temos  $(x_\alpha, y_\alpha) \in \Omega, \forall \alpha \in [0, 1]$ . Tomemos  $\epsilon_0 > 0$  tal que  $B_{\epsilon_0}(\bar{x})$  tenha a propriedade do item (i) do Lema 3.1.7, em relação à  $\bar{x}$ . Podemos supor sem perda de generalidade que  $\epsilon \leq \epsilon_0$ . Nosso objetivo é mostrar que  $y_\alpha \in S(x_\alpha), \forall \alpha \leq \epsilon$ . Para  $\alpha \leq \epsilon$  fixado, tome  $\lambda \in \lambda(\bar{x}) \cap \lambda(\tilde{x})$ , cuja existência é assegurada pelo item (i) do Lema 3.1.7. Mostremos primeiro que  $(\bar{x}, \bar{y}, \lambda)$  e  $(\tilde{x}, \tilde{y}, \lambda)$  são soluções do sistema (3.1) (página 29). Para isto, basta verificar as condições de complementariedade (as outras já valem). Façamos a argumentação para  $(\bar{x}, \bar{y}, \lambda)$ . Como  $\bar{y} \in S(\bar{x})$ , existe  $\bar{\lambda} \in \lambda(\bar{x})$  com  $\bar{\lambda}^t(A_2\bar{x} - b_2) = d_2\bar{y}$ . Pelo item (ii) do Lema 3.1.7, temos  $\lambda^t(A_2\bar{x} - b_2) = d_2\bar{y}$ . Agora se fosse  $(\lambda^tB_2 + d_2)\bar{y} > 0$  (notemos que  $\lambda^tB_2 + d_2 \geq 0$  e  $\bar{y} \geq 0$ ), teríamos  $\lambda^t(A_2\bar{x} - b_2) = d_2\bar{y} > -\lambda^tB_2\bar{y} \Rightarrow \lambda^t(A_2\bar{x} + B_2\bar{y} - b_2) > 0$ , um absurdo pois  $\lambda \geq 0$  e  $A_2\bar{x} + B_2\bar{y} - b_2 \leq 0$ . Por outro lado, se fosse  $\lambda^t(A_2\bar{x} + B_2\bar{y} - b_2) < 0$  (notemos que  $\lambda \geq 0$  e  $A_2\bar{x} + B_2\bar{y} - b_2 \leq 0$ ), teríamos  $d_2\bar{y} = \lambda^t(A_2\bar{x} - b_2) < -\lambda^tB_2\bar{y} \Rightarrow (\lambda^tB_2 + d_2)\bar{y} < 0$ , um

absurdo pois  $\lambda^t B_2 + d_2 \geq 0$  e  $\bar{y} \geq 0$ . Isto mostra que  $\lambda^t(A_2\bar{x} + B_2\bar{y} - b_2) = \bar{y}^t(B_2^t\lambda + d_2^t) = 0$ . Assim,  $(\bar{x}, \bar{y}, \lambda)$  é solução de (3.1). Pela mesma argumentação, mostramos que  $(\tilde{x}, \tilde{y}, \lambda)$  é também solução de (3.1).

Agora, tendo em vista que  $(x_\alpha, y_\alpha) \in \Omega$ , que  $\lambda$  é dual-viável em relação ao problema de baixo nível parametrizado em  $x_\alpha$  (isto é, que  $B_2^t\lambda + d_2^t \geq 0$  e  $\lambda \geq 0$ ), que

$$y_\alpha^t(B_2^t\lambda + d_2^t) = \alpha\tilde{y}^t(B_2^t\lambda + d_2^t) + (1 - \alpha)\bar{y}^t(B_2^t\lambda + d_2^t) = 0,$$

e que

$$\begin{aligned} \lambda^t(A_2x_\alpha + B_2y_\alpha - b_2) &= \lambda^t[\alpha A_2\tilde{x} + (1 - \alpha)A_2\bar{x} + \alpha B_2\tilde{y} + \\ &\quad (1 - \alpha)B_2\bar{y} - \alpha b_2 - (1 - \alpha)b_2] \\ &= \alpha\lambda^t(A_2\tilde{x} + B_2\tilde{y} - b_2) + (1 - \alpha)\lambda^t(A_2\bar{x} + B_2\bar{y} - b_2) = 0, \end{aligned}$$

vem que  $(x_\alpha, y_\alpha, \lambda)$  é solução de (3.1), e daí,  $y_\alpha \in S(x_\alpha)$ .

Com isso em mãos, tomemos  $\alpha \in (0, \epsilon]$  suficientemente pequeno para que  $\|\bar{y} - y_\alpha\| < \epsilon$ . Temos

$$\begin{aligned} c_1x_\alpha + d_1y_\alpha &= \alpha(c_1\tilde{x} + d_1\tilde{y}) + (1 - \alpha)(c_1\bar{x} + d_1\bar{y}) \\ &= \alpha[(c_1\tilde{x} + d_1\tilde{y}) - (c_1\bar{x} + d_1\bar{y})] + c_1\bar{x} + d_1\bar{y} \\ &< c_1\bar{x} + d_1\bar{y}. \end{aligned}$$

Ora,  $(x_\alpha, y_\alpha) \in B_\epsilon(\bar{x}, \bar{y}) \cap \Upsilon$  e da última desigualdade obtemos uma contradição. Portanto  $(\bar{x}, \bar{y})$  é solução local otimista de (2.2).  $\square$

Usando as características geométricas do problema linear de dois níveis expostas no Teorema 3.1.3 (página 29) e no Teorema 3.1.6 (página 33), veremos agora características geométricas do conjunto admissível de problemas lineares de dois níveis.

**Teorema 3.1.9.** *i)  $\Upsilon$  é uma união finita de faces de  $\Omega$ ;*

*ii) Se  $B_1 = 0$  então  $\Upsilon$  é conexo por caminhos.*

*Demonstração.* *i)* Por definição  $\Upsilon = \text{graf } S \cap \Omega_u$ . Então pelo Teorema 3.1.3,  $\Upsilon$  é a interseção de  $\Omega_u$  com uma união finita de faces de  $\Omega_l$ . Analisando as relações (3.2)-(3.4) (página 29) e as restrições  $A_1x + B_1y \leq b_1, x \geq 0$  que definem  $\Omega_u$  é fácil concluir o resultado.

*ii)* Sendo  $B_1 = 0$ , podemos passar as restrições de alto nível para o baixo nível

sem modificar  $\Upsilon$  (veja o Teorema 3.1.12 adiante, página 41). Obtemos assim um novo problema linear de dois níveis, onde  $\Omega_l = \Omega$  e  $\text{graf } S = \Upsilon$ . O resultado segue então do fato de  $\text{graf } S$  ser conexo por caminhos (item (ii) do Teorema 3.1.6, página 33).  $\square$

**Corolário 3.1.10.** *Considere o problema otimista*

$$\min_{x,y} F(x,y) \quad \text{s.a. } A_1x + B_1y \leq b_1, y \in S(x), x \geq 0$$

onde  $F$  é côncava sobre  $\Omega$ . Supondo que  $\Omega$  possua vértice (satisfeito se  $\Omega$  é compacto), se este problema admite solução ótima, uma ocorre num vértice de  $\Omega$ .

*Demonstração.* Consideremos o modelo

$$-\max_{x,y} -F(x,y) \quad \text{s.a. } A_1x + B_1y \leq b_1, (x,y) \in \text{graf } S, x \geq 0, \quad (3.10)$$

equivalente ao problema do enunciado. Pelo item (i) do Teorema 3.1.9, a região de busca  $\Upsilon = \text{graf } S \cap \Omega_u$  do problema (3.10) é uma união finita de conjuntos poliedrais, digamos  $\Upsilon_1, \dots, \Upsilon_r$ . O problema (3.10) é resolvido resolvendo-o sobre cada  $\Upsilon_i$ . Digamos que uma solução ótima esteja em  $\Upsilon_k$  que, pela prova do Teorema 3.1.3 (página 29), podemos escrever como  $\Upsilon_k = \Omega_u \cap PM(I_k, J_k)$ .  $PM(I_k, J_k)$  é o conjunto das soluções do sistema (3.2)-(3.4) (com  $I = I_k$  e  $J = J_k$ , é claro). Como  $-F$  é convexa, segue do Teorema A.3.4 (página 122) que um vértice de  $\Upsilon_k$  é ótimo, digamos  $(\bar{x}, \bar{y})$ . O Teorema A.3.3 (página 122) diz que exatamente  $n_x + n_y$  restrições de  $A_1x + B_1y \leq b_1, x \geq 0, (3.2) - (3.4)$  são ativas em  $(\bar{x}, \bar{y})$ , e são linearmente independentes, isto é,  $(\bar{x}, \bar{y})$  é vértice de  $\Omega$ , como queríamos demonstrar.  $\square$

O Corolário 3.1.10 é exposto em [30, 41], sem prova, para  $F$  linear. Em [14] encontra-se uma versão particular, também sem prova. Observamos que se  $d_2$  dependesse de  $x$ , a prova do Teorema 3.1.3 (página 29) não valeria (e portanto a prova do Corolário 3.1.10) pois (3.2)-(3.4) dependeria de (3.5) e (3.6) (página 29).

O Corolário 3.1.10 evidencia a característica combinatória dos problemas lineares de dois níveis: afim de resolvê-los, basta que analisemos os vértices de  $\Omega$ . Essa idéia é usada por algoritmos, como veremos no Capítulo 5, e também possibilita a adaptação de ferramentas de programação linear. Esses problemas podem ser resolvidos em tempo polinomial quando as variáveis de baixo nível têm dimensão finita [18, Teorema 6.3].

Fazemos agora uma última observação. Em geral, ao mover restrições do alto para o baixo nível, ou vice-versa, modificamos o conjunto  $\Upsilon$  dos pontos admissíveis. O próximo exemplo ilustra uma típica situação.

**Exemplo 3.1.11.** [30] Considere o problema bidimensional otimista de dois níveis (\*) e o problema (\*\*) obtido pela troca da restrição (4) do baixo para o alto nível:

$$(*) \max_{x,y} \{-x - 4y\}$$

$$\text{s.a. } x \geq 0$$

$$y \in \arg \max_y y$$

$$\text{s.a. } -2x - y \leq 8 \quad (1)$$

$$-3x + 2y \leq 6 \quad (2)$$

$$-5x + 6y \leq 60 \quad (3)$$

$$2x + y \leq 16 \quad (4)$$

$$2x - 5y \leq 0 \quad (5)$$

$$y \geq 0$$

$$(**) \max_{x,y} \{-x - 4y\}$$

$$\text{s.a. } 2x + y \leq 16 \quad (4)$$

$$x \geq 0$$

$$y \in \arg \max_y y$$

$$\text{s.a. } -2x - y \leq 8 \quad (1)$$

$$-3x + 2y \leq 6 \quad (2)$$

$$-5x + 6y \leq 60 \quad (3)$$

$$2x - 5y \leq 0 \quad (5)$$

$$y \geq 0.$$

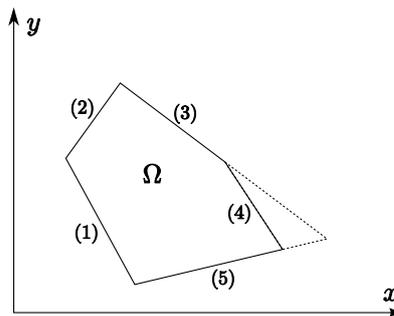


Figura 4: Ilustração para o Exemplo 3.1.11 mostrando o movimento de restrições.

Analisemos geometricamente os problemas (\*) (figuras à esquerda) e (\*\*) (figuras à direita):

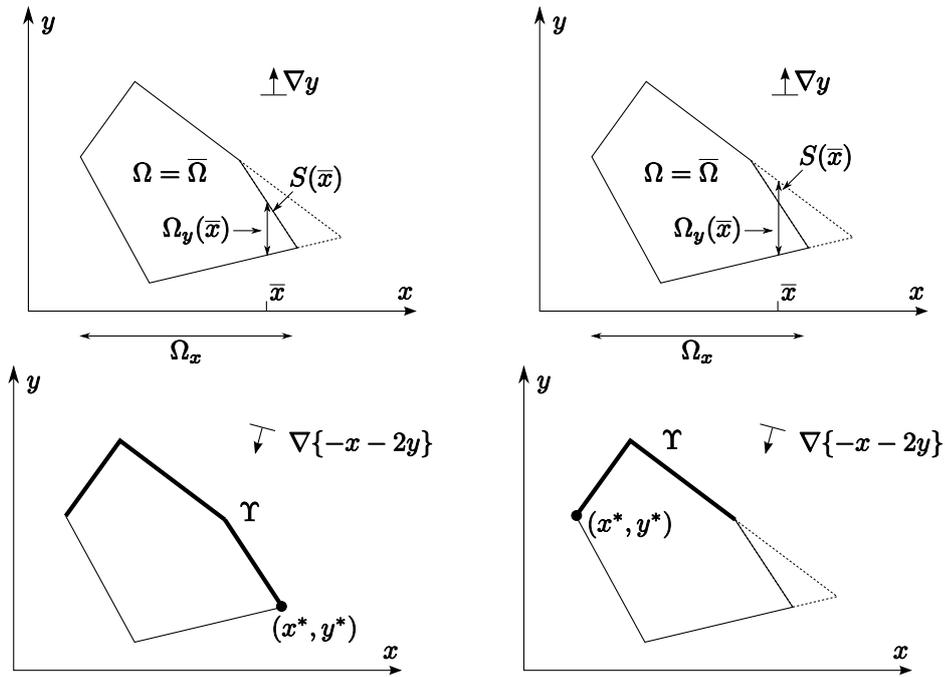


Figura 5: À esquerda, comportamento do problema (\*); à direita, comportamento do problema (\*\*).

Observamos que  $\Upsilon$  é distinto nos dois casos. Note que podemos ter até  $\Upsilon$  desconexo, ou mesmo vazio, mediante a introdução de restrições de alto nível.

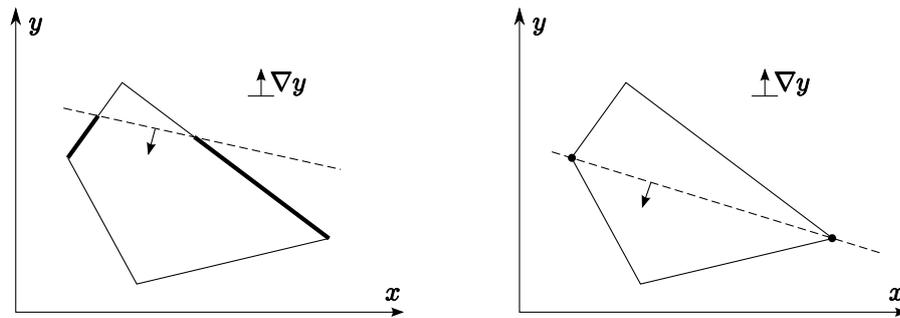


Figura 6: Região admissível  $\Upsilon$  desconexa (linhas escuras). A linha pontilhada representa uma restrição de alto nível.

□

Vale ressaltar que a argumentação acima vale também quando as restrições não são lineares. O resultado a seguir, cuja demonstração é trivial, relaciona os conjuntos admissíveis de um problema e do obtido ao mover restrições para o baixo nível. Apesar do problema de seu enunciado ser linear, também é válido quando as restrições e funções são

não lineares, donde entende-se que “restrições que não contém variáveis de baixo nível” seja, no caso geral, “restrições que não dependem das escolhas do baixo nível”.

**Teorema 3.1.12.** *Consideremos o problema de dois níveis*

$$\begin{aligned}
 P1: \quad & \min_{x,y} \{c_1x + d_1y\} \\
 \text{s.a.} \quad & A_1x + B_1y \leq b_1, \quad x \geq 0 \\
 & C_1x + D_1y \leq e_1 \\
 & y \in \arg \min_y d_2y \\
 & \text{s.a. } A_2x + B_2y \leq b_2, \quad y \geq 0
 \end{aligned}$$

e o problema  $P2$  obtido de  $P1$  ao mover as restrições  $C_1x + D_1y \leq e_1$  para o baixo nível. Então  $\Upsilon^1 \subset \Upsilon^2$ , onde o super-índice  $i$  é relativo ao problema  $Pi$ . Em particular, se as restrições  $C_1x + D_1y \leq e_1$  não contiverem variáveis de baixo nível, ou seja, se  $D_1 = 0$ , então  $\Upsilon^1 = \Upsilon^2$ .

## 3.2 Existência de soluções ótimas

O resultado a seguir reúne condições suficientes para assegurar a existência de soluções globais otimizadas para o problema linear de dois níveis.

**Teorema 3.2.1.** *Suponhamos que  $\Omega$  seja compacto e não vazio. Temos que se  $\Upsilon = \text{graf } S \cap \Omega_u \subset \Omega$  for não vazio ou  $B_1 = 0$  então o problema linear de dois níveis admite uma solução global otimista.*

*Demonstração.* Suponhamos que  $\Upsilon \neq \emptyset$ . Pelo item (i) do Teorema 3.1.9 (página 37), o problema linear de dois níveis consiste em minimizar a função contínua  $c_1x + d_1y$  em finitas faces de  $\Omega$ . Pelo fato de  $\Omega$  ser compacto, essas faces também são, e o resultado segue então do Teorema de Weierstrass (Teorema A.2.4, página 120).

Agora suponhamos que  $B_1 = 0$  e consideremos o problema

$$\min_{x,y} d_2y \quad \text{s.a.} \quad (x, y) \in \Omega.$$

Pelo Teorema de Weierstrass, o problema acima admite solução ótima, digamos  $(\bar{x}, \bar{y})$ . Temos que  $d_2\bar{y} \leq d_2y, \forall (x, y) \in \Omega$ . Em particular, segue do fato de  $B_1 = 0$  que  $d_2\bar{y} \leq d_2y, \forall y \in \Omega_y(\bar{x})$  e logo  $\bar{y} \in S(\bar{x})$ . Daí  $(\bar{x}, \bar{y}) \in \Upsilon$  e o resultado segue da discussão anterior.  $\square$

## 4 Formulações dois níveis de problemas clássicos

Neste capítulo apresentaremos exemplos de formulações dois níveis de problemas clássicos da literatura. Nosso intuito não é fazer aqui um estudo aprofundado, mas simplesmente mostrar o quão rico é o mundo dos problemas multinível. Por isso, detalhes e demonstrações dos resultados serão omitidos. Vale lembrar que esta é uma pequena amostra do grande número de trabalhos nessa linha.

### 4.1 Problema de programação linear 0-1 misto

Considere o problema de programação linear 0-1 misto [30]

$$\text{MIP}_{0-1}: \max_{x,u} \{cx + eu\} \quad \text{s.a.} \quad Ax + Eu \leq b, \quad x \geq 0, \quad u \in \{0, 1\}^n.$$

A condição de integralidade  $u \in \{0, 1\}^n$  é equivalente à

$$0 \leq u \leq 1, \quad 0 = \min_{1 \leq i \leq n} \{u_i, 1 - u_i\}.$$

Introduzindo uma variável  $y$ , podemos reescrever  $\text{MIP}_{0-1}$  como

$$\begin{aligned} & \max_{x,y,u} \{cx + eu\} \\ & \text{s.a.} \quad Ax + Eu \leq b \\ & \quad 0 \leq u \leq 1_n, \quad x \geq 0, \quad y = 0 \\ & \quad y \in \arg \max_w 1_n w \\ & \quad \text{s.a.} \quad w \leq u, \quad w \leq 1_n - u \end{aligned}$$

onde  $1_n$  representa o vetor formado por 1's. A restrição de alto nível  $y = 0$  força a integralidade de  $u$ . De fato, se  $y = 0 \in \arg \max 1_n w$  e  $u$  não for inteiro, a restrição  $0 \leq u \leq 1_n$  diz que  $u_i, 1 - u_i > 0$  para algum  $i$  e logo  $w_i > 0$  na otimalidade, isto é,

$\max 1_n w > 0$ . Daí  $y = 0 \notin \arg \max 1_n w$ , uma contradição.

## 4.2 Problema de programação bilinear

Considere o problema de programação bilinear [11]

$$\text{BP: } \max_{x,y} \{c^t x + x^t Q y + d^t y\} \quad \text{s.a.} \quad Ax \leq a, \quad x \geq 0, \quad By \leq b, \quad y \geq 0.$$

Definimos

$$f(x) = \sup \{(Q^t x + d)^t y; By \leq b, y \geq 0\}.$$

Fazendo o dual do problema acima, temos que

$$f(x) = \inf \{b^t \lambda; B^t \lambda \geq Q^t x + d, \lambda \geq 0\}.$$

Finalmente, é possível mostrar que BP é globalmente equivalente ao problema de dois níveis

$$\begin{aligned} & \max_{x,\lambda} \{c^t x + b^t \lambda\} \\ & \text{s.a. } Ax \leq a, \quad x \geq 0 \\ & \quad y \in \arg \max_y -b^t \lambda \\ & \quad \text{s.a. } B^t \lambda \geq Q^t x + d, \quad \lambda \geq 0. \end{aligned}$$

## 4.3 Problema max-min linear

A formulação do problema max-min linear [11] é dada por

$$\text{MM: } \max_{x \geq 0} \min_{y \geq 0} \{cx + dy; Ax + By \leq b\}.$$

Um par  $(x, y)$  é viável para MM se  $Ax + By \leq b$  e  $y \in \arg \min_{y \geq 0} \{cx + dy; Ax + By \leq b\}$ .

Fica definido naturalmente o problema de baixo nível

$$\max_y -dy \quad \text{s.a.} \quad Ax + By \leq b, \quad y \geq 0,$$

relativo à função  $\min_{y \geq 0} \{cx + dy; Ax + By \leq b\}$  parametrizada em  $x$  (o termo  $-cx$  pode ser descartado pois é constante). Imediatamente obtemos o problema de dois níveis

$$\begin{aligned} & \max_{x,y} \{cx + dy\} \\ & \text{s.a. } x \geq 0 \\ & \quad y \in \arg \max_y -dy \\ & \quad \text{s.a. } Ax + By \leq b, \quad y \geq 0. \end{aligned}$$

Observamos que não há diferença em maximizar a função objetivo do alto nível em  $x$  e  $y$  ou apenas em  $x$ , pois a parcela  $dy$  já é otimizada pelo baixo nível.

## 4.4 Um problema específico: o caixeiro viajante (TSP<sup>1</sup>)

Mostraremos agora como o TSP pode ser formulado como um problema de dois níveis. Para isso, lançaremos mão do modelo do problema de tarifação (TOP<sup>2</sup>), usado no estudo de preços ótimos para estradas tarifadas [23, 20, 19]. Descreveremo-lo na subseção 4.4.2. Logo depois, formularemos o TSP como TOP. Antes, relembremos o clássico problema de caminho mínimo.

### 4.4.1 O problema de caminho mínimo

O problema de caminho mínimo é formulado como segue. Dada uma rede  $G(N, A)$ , destacamos dois nós em  $N$ ,  $o$  e  $d$ , que serão denominados *origem* e *destino*, respectivamente. A cada  $(i, j) \in A$  associamos um custo  $c_{ij}$ , de forma que  $G(N, A)$  não possua ciclos de peso negativo. Deseja-se então encontrar um caminho de menor custo partindo da origem e chegando ao destino.

Introduzimos as variáveis

$$z_{ij} = \begin{cases} 1, & \text{se } (i, j) \text{ faz parte do caminho} \\ 0, & \text{caso contrário} \end{cases}.$$

Como  $G$  não possui ciclos de peso negativo, um caminho  $P$  pode ser determinado pelas

<sup>1</sup>Do inglês *Travelling Salesman Problem*.

<sup>2</sup>Do inglês *Toll Setting Problem*.

variáveis  $z_{ij}$  sujeitas às restrições

$$\sum_{i:(i,j) \in A} z_{ij} - \sum_{i:(j,i) \in A} z_{ji} = \begin{cases} -1, & \text{se } j = o \\ 1, & \text{se } j = d \\ 0, & \text{caso contrário} \end{cases} \quad \forall j \in N, \quad (4.1)$$

onde descartamos possíveis ciclos de peso nulo. Assim, o problema de caminho mínimo é formulado como

$$\begin{aligned} \min_z cz &= \sum_{(i,j) \in A} c_{ij} z_{ij} \\ \text{s.a. (4.1), } & z_{ij} \in \{0, 1\}, \forall (i, j) \in A. \end{aligned} \quad (4.2)$$

Agora, seja  $M$  a matriz relativa às restrições (4.1) e  $b$  a coluna formada pelos termos à direita da igualdade dessas restrições. A matriz  $M$  é unimodular [4] (isto é, o determinante de toda submatriz quadrada é -1, 0 ou 1), e vemos facilmente que a matriz  $\begin{bmatrix} M & b \end{bmatrix}$  também é. Daí segue que uma solução ótima do problema

$$\min_z cz \quad \text{s.a. (4.1), } z_{ij} \geq 0, \forall (i, j) \in A \quad (4.3)$$

é composta somente de 0's e 1's, isto é, os problemas (4.2) e (4.3) são globalmente equivalentes. O último é um problema de fluxo mínimo, e as restrições (4.1) são as equações de balanceamento. Neste problema, enviamos uma unidade de fluxo de  $o$  para  $d$ .

A equivalência entre (4.2) e (4.3) será usada no modelo do problema de tarifação descrito na subseção seguinte.

#### 4.4.2 O problema de tarifação (TOP)

Como referência para esta subseção, o leitor pode consultar a tese de S. Dewez [19] e o artigo de S. Dewez *et al.* [20].

Imaginemos uma rede de tráfego, composta de cidades (nós) e estradas ligando estas cidades (arcos). Custos fixos de viagem são associados a cada arco. Suponhamos que uma parte da rede seja controlada por uma empresa que tem a liberdade de cobrar tarifas pelas viagens em suas estradas, de acordo com a quantidade de carregamento transportado. O problema de tarifação consiste no seguinte: a cada política de tarifação, os usuários que trafegam na rede percorrem um caminho de menor custo, desde sua origem ao seu destino, ao mesmo tempo que o dono da rede tarifada procura maximizar seu ganho na cobrança

de tarifas.

Afim de modelar o problema, seja  $G(K, N, \bar{A})$  uma rede onde  $K$  é o conjunto cujos elementos representam usuários de mesmo par origem-destino (*commodities*). Seja  $\{(o^k, d^k) \in N^2; k \in K\}$  o conjunto desses pares origem-destino, onde demandas  $\eta^k > 0$  são associadas a cada destino  $d^k$ . Seja também  $A \subset \bar{A}$  o subconjunto dos arcos tarifados (*toll arcs*) e  $B = \bar{A} \setminus A$  o subconjunto dos arcos não tarifados (*toll-free arcs*). A cada arco  $(i, j) \in A$  associamos um custo fixo  $c_{ij}$  e uma tarifação adicional  $t_{ij} \geq 0$ . A cada arco  $(i, j) \in B$  associamos um custo fixo  $d_{ij}$ . Sejam ainda

$$x_{ij}^k = \begin{cases} 1, & \text{se } k \in K \text{ usa } (i, j) \in A \\ 0, & \text{caso contrário} \end{cases} \quad \text{e} \quad y_{ij}^k = \begin{cases} 1, & \text{se } k \in K \text{ usa } (i, j) \in B \\ 0, & \text{caso contrário} \end{cases}.$$

O modelo do problema de tarifação consiste em um problema otimista de dois níveis, onde o baixo nível representa os caminhos mínimos adotados pelos usuários da rede, e o alto nível o ganho em tarifas cobradas pela empresa gestora. Formulamos-o portanto como

$$\begin{aligned} \text{TOP: } \max_{x, y, t} \quad & \sum_{k \in K} \sum_{(i, j) \in A} \eta^k t_{ij} x_{ij}^k \\ \text{s.a. } \quad & (0 \leq) t_{ij}^{\min} \leq t_{ij} \leq t_{ij}^{\max}, \quad \forall (i, j) \in A \\ & (x, y) \in \arg \min_{x, y} \sum_{k \in K} \left( \sum_{(i, j) \in A} (c_{ij} + t_{ij}) x_{ij}^k + \sum_{(i, j) \in B} d_{ij} y_{ij}^k \right) \\ & \text{s.a. } \sum_{i; (i, j) \in A} x_{ij}^k + \sum_{i; (i, j) \in B} y_{ij}^k - \sum_{i; (j, i) \in A} x_{ji}^k - \sum_{i; (j, i) \in B} y_{ji}^k \\ & \quad = \begin{cases} -1, & \text{se } j = o^k \\ 1, & \text{se } j = d^k \\ 0, & \text{caso contrário} \end{cases} \quad \forall k \in K, \forall j \in N \\ & x_{ij}^k \geq 0, \quad \forall k \in K, \forall (i, j) \in A \\ & y_{ij}^k \geq 0, \quad \forall k \in K, \forall (i, j) \in B. \end{aligned}$$

Na otimalidade,  $x^k$  e  $y^k$  são binários  $\forall k \in K$ , visto que cada problema em  $k \in K$  é de caminho mínimo (subseção 4.4.1). Também, se para algum  $k \in K$  só existem caminhos de  $o^k$  a  $d^k$  que contenham algum arco  $(i, j)$  tarifado, a condição  $t_{ij} \leq t_{ij}^{\max}$  é importante para prevenir que o problema não possua solução ótima. Se o limite superior não for finito neste caso, poderíamos tomar  $t_{ij}$  arbitrariamente grande, já que a *commodity*  $k$  é obrigada a passar pelo arco  $(i, j)$ .

Em vista do fato de que na otimalidade  $x^k$  e  $y^k$  são binários, podemos, para cada  $k \in K$ , substituir o problema de baixo nível por suas condições de otimalidade, obtendo

$$\max_{x,y,t,\lambda} \sum_{k \in K} \sum_{(i,j) \in A} \eta^k t_{ij} x_{ij}^k \quad (4.4)$$

$$\begin{aligned} \text{s.a. } (0 \leq) t_{ij}^{\min} \leq t_{ij} \leq t_{ij}^{\max}, \quad \forall (i,j) \in A \\ \sum_{i:(i,j) \in A} x_{ij}^k + \sum_{i:(i,j) \in B} y_{ij}^k - \sum_{i:(j,i) \in A} x_{ji}^k - \sum_{i:(j,i) \in B} y_{ji}^k \\ = \begin{cases} -1, & \text{se } j = o^k \\ 1, & \text{se } j = d^k \\ 0, & \text{caso contrário} \end{cases} \quad \forall k \in K, \forall j \in N \end{aligned}$$

$$\lambda_j^k - \lambda_i^k \leq c_{ij} + t_{ij}, \quad \forall k \in K, \forall (i,j) \in A$$

$$\lambda_j^k - \lambda_i^k \leq d_{ij}, \quad \forall k \in K, \forall (i,j) \in B$$

$$\sum_{(i,j) \in A} (c_{ij} + t_{ij}) x_{ij}^k + \sum_{(i,j) \in B} d_{ij} y_{ij}^k = \lambda_{d^k}^k - \lambda_{o^k}^k, \quad \forall k \in K \quad (4.5)$$

$$x_{ij}^k \geq 0, \quad \forall k \in K, \forall (i,j) \in A$$

$$y_{ij}^k \geq 0, \quad \forall k \in K, \forall (i,j) \in B.$$

Os termos bilineares  $t_{ij} x_{ij}^k$  em (4.4) e (4.5) podem ser substituídos por expressões lineares como se segue. Façamos  $\tau_{ij}^k = t_{ij} x_{ij}^k$ . Como na otimalidade  $x_{ij}^k$  é binário, temos que  $x_{ij}^k = 0 \Rightarrow \tau_{ij}^k = 0$  e  $x_{ij}^k = 1 \Rightarrow \tau_{ij}^k = t_{ij}$ . Essas restrições são capturadas por

$$\begin{aligned} \tau_{ij}^k &\leq M_{ij}^k x_{ij}^k && \forall k \in K, \forall (i,j) \in A \\ t_{ij} - \tau_{ij}^k &\leq N_{ij} (1 - x_{ij}^k) && \forall k \in K, \forall (i,j) \in A \\ \tau_{ij}^k &\leq t_{ij} && \forall k \in K, \forall (i,j) \in A \\ \tau_{ij}^k &\geq 0 && \forall k \in K, \forall (i,j) \in A \end{aligned}$$

onde  $M_{ij}^k \geq \tau_{ij}^k$  e  $N_{ij} \geq t_{ij}$ , para alguns  $\tau_{ij}^k$  e  $t_{ij}$  ótimos. O problema se escreve então como

o MIP

$$\begin{aligned}
\text{TOP-ARCS: } \max_{x,y,t,\lambda} & \sum_{k \in K} \sum_{(i,j) \in A} \eta^k \tau_{ij}^k \\
\text{s.a. } & (0 \leq) t_{ij}^{\min} \leq t_{ij} \leq t_{ij}^{\max}, \quad \forall (i,j) \in A \\
& \sum_{i:(i,j) \in A} x_{ij}^k + \sum_{i:(i,j) \in B} y_{ij}^k - \sum_{i:(j,i) \in A} x_{ji}^k - \sum_{i:(j,i) \in B} y_{ji}^k \\
& \qquad \qquad \qquad = \begin{cases} -1, & \text{se } j = o^k \\ 1, & \text{se } j = d^k \\ 0, & \text{caso contrário} \end{cases} \quad \forall k \in K, \forall j \in N \\
& \lambda_j^k - \lambda_i^k \leq c_{ij} + t_{ij}, \quad \forall k \in K, \forall (i,j) \in A \\
& \lambda_j^k - \lambda_i^k \leq d_{ij}, \quad \forall k \in K, \forall (i,j) \in B \\
& \sum_{(i,j) \in A} (c_{ij} x_{ij}^k + \tau_{ij}^k) + \sum_{(i,j) \in B} d_{ij} y_{ij}^k = \lambda_{d^k}^k - \lambda_{o^k}^k, \quad \forall k \in K \\
& \tau_{ij}^k \leq M_{ij}^k x_{ij}^k, \quad \forall k \in K, \forall (i,j) \in A \\
& t_{ij} - \tau_{ij}^k \leq N_{ij} (1 - x_{ij}^k), \quad \forall k \in K, \forall (i,j) \in A \\
& \tau_{ij}^k \leq t_{ij}, \quad \forall k \in K, \forall (i,j) \in A \\
& x_{ij}^k \in \{0, 1\}, \quad \forall (i,j) \in A \\
& y_{ij}^k \geq 0, \tau_{ij}^k \geq 0, \quad \forall k \in K, \forall (i,j) \in B.
\end{aligned}$$

### 4.4.3 Um modelo TOP para o TSP

Os resultados apresentados nesta subseção encontram-se no artigo de P. Marcotte *et al.* [32].

Seja  $G(N_0, A_0)$  um grafo direcionado. Denotaremos os nós por  $N_0 = \{1, 2, \dots, n\}$ . A cada  $(i, j) \in A_0$  associamos um custo  $c_{ij}$ . O problema do caixeiro viajante consiste em encontrar um circuito hamiltoniano em  $G(N_0, A_0)$  de menor custo. A fim de formular o TSP como TOP, adicionamos o nó  $n+1$  e substituímos todos  $(i, 1) \in A_0$  por  $(i, n+1)$ , onde  $i \in N_0 \setminus \{1\}$ . Seja  $A$  o conjunto dos arcos após a modificação. Adicionamos ainda o arco  $(1, n+1)$ .  $A$  será o conjunto dos arcos tarifados e  $B = \{(1, n+1)\}$  o conjunto dos arcos não tarifados. Temos então um novo grafo  $G(N, \bar{A})$ , onde  $N = N_0 \cup \{n+1\}$  e  $\bar{A} = A \cup B$ ,  $A \cap B = \emptyset$ . A cada  $(i, j) \in A$  associamos o custo fixo  $-1$  e uma tarifa  $t_{ij} \geq 2$  e ao arco não tarifado  $(1, n+1)$  associamos o custo fixo  $n$ . Consideremos o problema de tarifação onde  $(1, n+1)$  é o único par origem-destino (*single commodity*) demandando

uma unidade:

$$\begin{aligned}
 \text{TOP1: } & \max_{x,y,t} \sum_{(i,j) \in A} t_{ij} x_{ij} \\
 \text{s.a. } & t_{ij} \geq 2, \quad \forall (i,j) \in A \\
 & (x,y) \in \arg \min_{x,y} \sum_{(i,j) \in A} (-1 + t_{ij}) x_{ij} + ny \\
 \text{s.a. } & \sum_{i:(i,j) \in A} x_{ij} + y - \sum_{i:(j,i) \in A} x_{ji} = \begin{cases} -1, & \text{se } j = 1 \\ 1, & \text{se } j = n + 1 \\ 0, & \text{caso contrário} \end{cases} \quad \forall j \in N \\
 & x_{ij} \geq 0, \quad \forall (i,j) \in A \\
 & y \geq 0.
 \end{aligned}$$

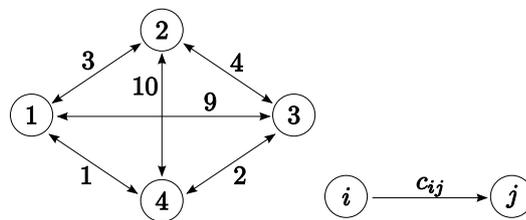
Notemos que limites superiores para os  $t_{ij}$  não são necessários pois existe um caminho não tarifado de 1 à  $n + 1$ , a saber,  $1 \rightarrow n + 1$ .

A argumentação segue agora no sentido de relacionar soluções de TOP1 a circuitos hamiltonianos em  $G(N_0, A_0)$ . Dado um caminho  $P$ , diremos que o circuito  $C$  obtido de  $P$  ligando seus nós extremos é induzido por  $P$ , ou que  $P$  induz  $C$ .

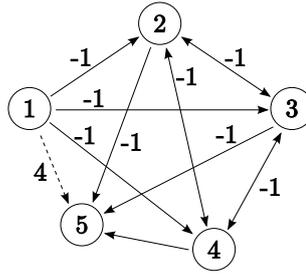
**Teorema 4.4.1.** *Existe um circuito hamiltoniano em  $G(N_0, A_0)$  induzido por um caminho tarifado em  $A$  se, e somente se, o valor ótimo de TOP1 for igual a  $2n$ .*

O próximo exemplo mostra que os arcos tarifados de uma solução ótima de TOP1 (que formam um caminho) não necessariamente induzem uma solução ótima do TSP, isto é, um circuito hamiltoniano em  $G(N_0, A_0)$  de custo mínimo.

**Exemplo 4.4.2.** Consideremos a rede  $G(N_0, A_0)$



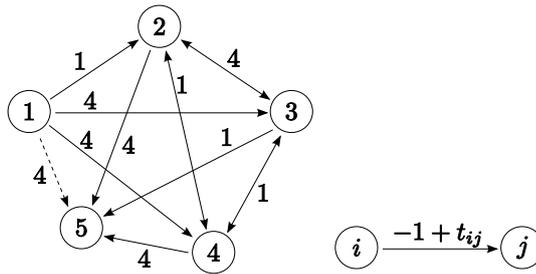
Vemos facilmente que o circuito  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ , de custo 10, é ótimo para o TSP. A rede  $G(N, \bar{A})$  correspondente é representada a seguir:



Tomemos  $t_0 = (t_{ij})_{4 \times 5}$  como sendo

$$t_0 = \begin{bmatrix} \cdot & 2 & 5 & 5 & \cdot \\ \cdot & \cdot & 5 & 2 & 5 \\ \cdot & 5 & \cdot & 2 & 2 \\ \cdot & 2 & 2 & \cdot & 5 \end{bmatrix},$$

onde  $\cdot$  denota que não atribuímos valor (ausência da variável correspondente). A rede parametrizada em  $t_0$  é representada como



O caminho tarifado  $P : 1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 5$  em  $G(N, \overline{A})$  representa uma solução ótima para o baixo nível de TOP1 parametrizado em  $t_0$  (existe mais uma somente, a representada pelo caminho tarifado  $1 \rightarrow n + 1$ ). Também,  $P$  e  $t_0$  são tais que o valor da função objetivo de alto nível de TOP1 é  $8 = 2n$ . Pela Proposição 4.4.1,  $P$  e  $t_0$  representam portanto uma solução ótima de TOP1 (observamos que se considerarmos o caminho  $1 \rightarrow n + 1$ , o valor da função objetivo seria 0). No entanto,  $P$  induz o circuito hamiltoniano  $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 1$  em  $G(N_0, A_0)$ , que não é ótimo para o TSP pois seu custo é 24.  $\square$

Consideramos agora uma versão perturbada de TOP1. Fazemos os custos fixos dos arcos de  $A$  iguais a  $-1 + \frac{c_{ij}}{L}$  e os limites inferiores das tarifas iguais a  $2 - \frac{c_{ij}}{L}$ , onde  $c_{ij}$  são os custos dos arcos em  $A_0$ , relativos ao TSP, e  $L$  é um limite superior estrito para o custo

do circuito ótimo do TSP (por exemplo,  $L = n \cdot \max_{(i,j) \in A_0} c_{ij} + 1$ ). Obtemos o problema

$$\begin{aligned}
 \text{TOP-TSP: } & \max_{x,y,t} \sum_{(i,j) \in A} t_{ij} x_{ij} \\
 \text{s.a. } & t_{ij} \geq 2 - \frac{c_{ij}}{L}, \quad \forall (i,j) \in A \\
 & (x,y) \in \arg \min_{x,y} \sum_{(i,j) \in A} \left( \left( -1 + \frac{c_{ij}}{L} \right) + t_{ij} \right) x_{ij} + ny \\
 \text{s.a. } & \sum_{i:(i,j) \in A} x_{ij} + y - \sum_{i:(j,i) \in A} x_{ji} \\
 & = \begin{cases} -1, & \text{se } j = 1 \\ 1, & \text{se } j = n + 1 \\ 0, & \text{caso contrário} \end{cases} \quad \forall j \in N \\
 & x_{ij} \geq 0, \quad \forall (i,j) \in A \\
 & y \geq 0.
 \end{aligned}$$

O resultado a seguir estabelece a relação entre as soluções ótimas de TOP-TSP e as soluções ótimas do TSP, e conclui nossa discussão.

**Teorema 4.4.3.** *Se TSP admite solução ótima, o caminho tarifado relativo à uma solução ótima de TOP-TSP induz um circuito ótimo para o TSP.*

## 5 Algoritmos existentes para o caso linear

Neste capítulo, apresentaremos resumidamente algumas técnicas existentes na literatura para o caso linear. Para uma referência mais completa, veja [3, 40, 11].

Durante todo o capítulo, estamos considerando o problema

$$\begin{aligned} \text{LBLP: } \min_{x,y} \{c_1x + d_1y\} \\ \text{s.a. } A_1x + B_1y \leq b_1, \quad x \geq 0 \\ y \in \arg \min_y d_2y \\ \text{s.a. } A_2x + B_2y \leq b_2, \quad y \geq 0 \end{aligned}$$

investigado no Capítulo 3.

### 5.1 O algoritmo do K-ésimo melhor vértice

Do Corolário 3.1.10 (página 38), podemos estabelecer uma maneira de resolver LBLP globalmente. Consideremos a relaxação de LBLP

$$\text{LR}^1 : \min_{x,y} \{c_1x + d_1y\} \quad \text{s.a. } A_1x + B_1y \leq b_1, A_2x + B_2y \leq b_2, x, y \geq 0$$

cujo espaço de busca é  $\Omega$ . Enumeram-se as soluções básicas viáveis (vértices), dispostas em ordem crescente de valor da função objetivo. Testa-se, nessa ordem, se cada uma é ótima para o baixo nível, isto é, se  $y \in S(x)$  (para efeito prático essa busca se faz analisando os vértices adjacentes ao corrente). A busca termina ao encontrar a primeira solução básica com esta propriedade (uma solução ótima do problema de dois níveis), ou ao terminar a lista de soluções básicas viáveis. Este método é conhecido na literatura por método do *K-ésimo melhor vértice*<sup>2</sup> e foi introduzido por Bialas e Karwan (*apud* [40]).

---

<sup>1</sup>Do inglês *Leader Relaxation*.

<sup>2</sup>Em inglês, *Kth-best*.

Apesar desses autores suporem  $\Omega$  compacto e que, para cada  $x$ , o problema de baixo nível retorne uma única solução ótima (*apud* [11]), isso não é necessário como mostra o próprio Corolário 3.1.10.

## 5.2 Abordagem via complementariedade

No Capítulo 3 usamos dualidade para escrever as condições de otimalidade do problema do baixo nível no sistema (3.1) (página 29). Então, fazendo  $w = b_2 - A_2x - B_2y$  e  $\mu = d_2^t + B_2^t\lambda$ , o modelo

$$\begin{aligned} \min_{x,y} \{c_1x + d_1y\} \\ \text{s.a. } w &= b_2 - A_2x - B_2y \\ \mu &= d_2^t + B_2^t\lambda \\ \lambda^t w &= \mu^t y = 0 \\ x, y, \lambda, \mu, w &\geq 0 \end{aligned}$$

é equivalente à LBLP no caso em que restrições de alto nível  $A_1x + B_1y \leq b_1$  não são presentes. A fim de minimizar a função  $c_1x + d_1y$ , Júdice e Faustino (*apud* [40]) propuseram substituí-la pela restrição  $c_1x + d_1y \leq t$ , obtendo assim o *problema de complementariedade linear paramétrica* (LCP( $t$ ))<sup>3</sup> dado por

$$\begin{aligned} \text{LCP}(t): \quad \begin{pmatrix} w \\ \mu \\ \alpha \end{pmatrix} &= \begin{pmatrix} b_2 \\ d_2^t \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} t + \begin{pmatrix} -A_2 & -B_2 & 0 \\ 0 & 0 & B_2^t \\ -c_1 & -d_1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} \\ \lambda^t w &= \mu^t y = 0, \quad x, y, \lambda, \mu, w, \alpha \geq 0. \end{aligned}$$

Uma solução ótima global de LBLP é aproximada então por uma solução de LCP( $t^*$ ), onde  $t^*$  é suficientemente pequeno. Para encontrar uma tal solução ótima, Júdice e Faustino propõem o método descrito a seguir. Inicia-se resolvendo o sistema LCP obtido de LCP( $t$ ) ao omitir a restrição  $c_1x + d_1y + \alpha = t$  (se LBLP tem solução ótima, o sistema LCP tem solução). Sendo  $(x_0, y_0)$  um ponto encontrado, tomemos  $t_0 = c_1x_0 + d_1y_0$ . O algoritmo então resolve uma sequência LCP( $t_k$ ) de sistemas, onde  $\{t_k\}$  é uma sequência decrescente definida por

$$t_k = c_1x_{k-1} + d_1y_{k-1} - \gamma|c_1x_{k-1} + d_1y_{k-1}|$$

<sup>3</sup>Do inglês *Linear Complementarity Problem*.

com  $(x_{k-1}, y_{k-1})$  solução de  $\text{LCP}(t_{k-1})$  e  $\gamma > 0$  pequeno. O processo pára, digamos na iteração  $r$ , quando  $\text{LCP}(t_r)$  não é viável. Quando isso ocorre, a solução  $(x_{r-1}, y_{r-1})$  de  $\text{LCP}(t_{r-1})$  satisfaz

$$0 \leq c_1 x_{r-1} + d_1 y_{r-1} - F^* \leq \gamma |c_1 x_{r-1} + d_1 y_{r-1}| = \epsilon$$

onde  $F^*$  é o valor ótimo de LBLP. Portanto se LBLP admite solução ótima, esse método sequencial (SLCP<sup>4</sup>) encontra uma solução  $\epsilon$ -ótima. Resumindo, temos o algoritmo seguinte.

---

**Algoritmo 1** SLCP

---

**Entrada:** Ausência de restrições de alto nível, isto é,  $A_1 = 0$ ,  $B_1 = 0$  e  $b_1 = 0$ .

**Saída:** Uma solução  $\epsilon$ -ótima de LBLP.

---

1. Tome  $\gamma > 0$  pequeno. Resolva o sistema LCP obtido de  $\text{LCP}(t)$  ao omitir a restrição  $c_1 x + d_1 y + \alpha = t$ . Seja  $(x_0, y_0)$  uma solução de LCP. Faça  $t_0 = c_1 x_0 + d_1 y_0$ ,  $t_1 = t_0 - \gamma |t_0|$ ,  $k = 1$  e vá para o passo 2.
  2. Resolva o sistema  $\text{LCP}(t_k)$ . Se for inviável, pare e retorne  $(x_{k-1}, y_{k-1})$  como solução  $\epsilon$ -ótima de LBLP. Caso contrário, seja  $(x_k, y_k)$  solução de  $\text{LCP}(t_k)$ . Faça  $t_{k+1} = c_1 x_k + d_1 y_k - \gamma |c_1 x_k + d_1 y_k|$ ,  $k \leftarrow k + 1$  e repita o passo 2.
- 

## 5.3 Eliminação de variáveis

O método aqui exposto foi desenvolvido por Hansen, Jaumard e Savard (*apud* [3, 40, 13]). É um método do tipo *branch and bound* que, na atualidade, mostra-se eficiente dentre os algoritmos exatos [13].

Tendo em vista a prova do Teorema 3.1.3 (página 29), podemos explorar soluções ótimas de LBLP usando conjuntos de índices  $I, J$  para o baixo nível, onde supomos sem perda de generalidade que pelo menos uma das restrições de baixo nível é ativa numa solução ótima de LBLP. Aqui, consideramos que  $\Omega$  seja compacto e não vazio. Para cada  $i \in \{1, \dots, m_l + n_y\}$ , definimos

$$\alpha_i = \begin{cases} 1, & \text{se a } i\text{-ésima restrição do baixo nível realiza-se como igualdade} \\ 0, & \text{se tal restrição realiza-se como desigualdade estrita} \end{cases}$$

---

<sup>4</sup>Do inglês *Sequential Linear Complementarity Problem*.

(os índices  $i > m_l$  capturam as restrições  $y \geq 0$ ). Os autores demonstram que as relações

$$\begin{aligned} \sum_{i \in \{i; B_{2ij} > 0\}} \alpha_i &\geq 1, & \text{se } d_{2j} > 0 \\ \alpha_{m_l+j} + \sum_{i \in \{i; B_{2ij} < 0\}} \alpha_i &\geq 1, & \text{se } d_{2j} < 0, \end{aligned} \quad (5.1)$$

para  $1 \leq j \leq n_y$ , são válidas em qualquer ponto admissível de LBLP, onde  $B_{2ij}$  é a componente da posição  $(i, j)$  de  $B_2$ . O processo de ramificação é feito da seguinte forma: primeiramente, fixa-se  $\alpha_i$  em 1 e 0; na sequência, uma das relações (5.1) é escolhida, digamos  $\alpha_{i_1} + \alpha_{i_2} + \alpha_{i_3} \cdots \geq 1$ , e a ramificação segue a regra  $\alpha_{i_1} = 1$  ou ( $\alpha_{i_1} = 0$  e  $\alpha_{i_2} = 1$ ) ou ( $\alpha_{i_1} = \alpha_{i_2} = 0$  e  $\alpha_{i_3} = 1$ ) e assim sucessivamente.

Quando  $\alpha_i$  é fixado em 1, é exigido que a  $i$ -ésima restrição de baixo nível seja de igualdade. Se  $i > m_l$ , podemos então eliminar a variável  $y_{m_l-i}$ , que é fixada em 0. Agora, se  $i \leq m_l$  podemos eliminar uma variável de baixo nível introduzindo uma restrição extra, como segue. Supondo que seja exigido que a restrição  $\sum_{i=1}^{n_y} a_i y_i \leq b_i$  seja de igualdade, podemos, por exemplo, escrever  $y_1 = b_i - \sum_{i=2}^{n_y} \frac{a_i}{a_1} y_i$  caso  $a_1 \neq 0$  (caso não fosse escolheríamos outro  $y_j$ ). Podemos substituir então  $y_1$  por essa expressão em todas as outras restrições e na função objetivo, e a restrição  $y_1 \geq 0$  substituímos por  $\sum_{i=2}^{n_y} \frac{a_i}{a_1} y_i - b_i \leq 0$ . Aqui podemos escolher qualquer  $y_j$  cujo  $a_j \neq 0$ . Hansen, Jaumard e Savard fazem a escolha da variável  $y_j$  que, ao ser substituída nas outras restrições, produza o menor número de coeficientes não nulos na matriz das restrições resultante (menor preenchimento<sup>5</sup>).

Quando  $\alpha_i$  é fixado em 0, é exigido que a  $i$ -ésima restrição de baixo nível seja de desigualdade estrita. Devido à dificuldade computacional de tratar restrições deste tipo, recorreremos ao dual do problema de baixo nível. Neste caso, a variável ou restrição dual relativa à  $i$ -ésima restrição ou variável é eliminada.

Fixados  $\alpha_{i_1}, \dots, \alpha_{i_r}$  (não necessariamente todo um vetor  $\alpha$ ), resolvemos a relaxação LR correspondente (página 52) para obter limitantes inferiores para LBLP, que denotaremos por  $\underline{F}$ . Limitantes superiores são dados pela melhor solução encontrada. Hansen, Jaumard e Savard usam penalizadores, a fim de obter limitantes inferiores mais fortes,

<sup>5</sup>Em inglês, *fill-in*.

descritos a seguir. Considere as expressões

$$F = \underline{F} + \sum_{j \in R} \bar{w}_j z_j$$

$$z_i = \bar{b}_i - \sum_{j \in R} \bar{C}_{ij} z_j, \quad i \in P$$

correspondentes a um quadro simplex ótimo da relaxação LR do problema corrente, onde  $F$  é a função objetivo (de alto nível),  $R$  o conjunto de índices das variáveis não básicas,  $P$  o conjunto de índices das variáveis básicas,  $\bar{w}$  o vetor de coeficientes reduzidos (relativo à  $[c_1 \quad d_1]$ ),

$$\bar{b} = [b_1^t \quad b_2^t]^t, \quad \bar{C} = \begin{bmatrix} A_1 & I_{m_u} & 0 & B_1 \\ A_2 & 0 & I_{m_l} & B_2 \\ I_{n_x} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{n_y} \end{bmatrix} \quad \text{e} \quad z = [x^t \quad s_u^t \quad s_l^t \quad y^t]^t,$$

com  $s_u \in \mathbb{R}_+^{m_u}$  e  $s_l \in \mathbb{R}_+^{m_l}$  variáveis de folga. Para cada  $i$ , seja  $R(i) = \{j \in R; \bar{C}_{ij} > 0\}$ .

Definimos

$$p_i = \begin{cases} \bar{b}_i \min_{j \in R(i)} \left\{ \frac{\bar{w}_j}{\bar{C}_{ij}} \right\}, & \text{se } R(i) \neq \emptyset \\ 0, & \text{caso contrário} \end{cases}.$$

O termo  $p_i$  corresponde ao incremento na função objetivo de alto nível ao retirarmos  $z_i$  da base (caso  $i \in P$ ).

Agora, em cada nó da árvore de enumeração temos relações tipo (5.1) correspondentes. Seja  $r_k : \sum_{i \in I_k} \alpha_i \geq 1$  uma delas ( $I_k \subset \{1, \dots, m_l + n_y\}$ ). Como já foi dito,  $r_k$  vale para qualquer ponto viável de LBLP (considerando o nó atual da árvore de ramificação), e então pelo menos uma variável  $z_{n_x+m_u+i}$  ( $i \in I_k$ ) deve ser fixada em 0 (pois isso corresponde a fazer  $\alpha_i = 1$ ). Logo

$$\underline{F}' = \underline{F} + \min_{i \in I_k} p_{n_x+m_u+i}$$

é um limitante inferior para o subproblema corrente. Finalmente, considerando o conjunto  $H$  de todas as relações do tipo (5.1) para o nó corrente, temos que

$$\underline{F}'' = \underline{F} + \max_{k \in \{k; r_k \in H\}} \min_{i \in I_k} p_{n_x+m_u+i}$$

é limitante inferior para o subproblema corrente.

Com isto, enunciamos o algoritmo de Hansen, Jaumard e Savard (HJS).

**Algoritmo 2** HJS**Entrada:**  $\Omega$  compacto e não vazio.**Saída:** Uma solução ótima global de LBLP.

1. (Inicialização) Obtenha uma solução inicial  $(x^0, y^0) \in \Upsilon$  usando uma heurística. Faça  $(\bar{x}, \bar{y}) \leftarrow (x^0, y^0)$  e  $\bar{F} \leftarrow c_1 x^0 + d_1 y^0$ . Se nenhuma solução inicial pode ser encontrada, inicialize  $(x^0, y^0)$  arbitrariamente e faça  $\bar{F} \leftarrow \infty$ . Todas as variáveis  $\alpha_i$ ,  $1 \leq i \leq m_l + n_y$  são livres. Faça  $R = \emptyset$ . Vá para o passo 2.
2. (Primeiro teste de otimalidade) Resolva a relaxação LR e tome  $(x, y)$  uma solução ótima sua. Se  $F = c_1 x + d_1 y \geq \bar{F}$ , vá para o passo 13. Caso contrário, vá para o passo 3.
3. (Primeiro teste de viabilidade) Resolva o dual de LR. Se tal problema for inviável, vá para o passo 13. Caso contrário, vá para o passo 4.
4. (Teste de solução, parte 1) Considere novamente a solução ótima  $(x, y)$  de LR. Verifique se este ponto é racional para o subproblema corrente, isto é, resolva o problema de baixo nível do subproblema corrente, tome  $\tilde{y}$  uma solução ótima sua e verifique se  $d_2 y = d_2 \tilde{y}$ . Se  $\bar{F} > c_1 x + d_1 y$  faça  $\bar{F} \leftarrow c_1 x + d_1 y$  e  $(\bar{x}, \bar{y}) \leftarrow (x, y)$  e vá para o passo 13. Caso contrário, vá para o passo 5.
5. (Teste de solução, parte 2) Considere novamente a solução ótima  $(x, y)$  de LR. Verifique se este ponto é racional ( $y \in S(x)$ ), isto é, resolva o problema de baixo nível LLP( $x$ ), tome  $\tilde{y}$  uma solução ótima sua e verifique se  $d_2 y = d_2 \tilde{y}$ . Se  $\bar{F} > c_1 x + d_1 y$  faça  $\bar{F} \leftarrow c_1 x + d_1 y$  e  $(\bar{x}, \bar{y}) \leftarrow (x, y)$  e vá para o passo 13. Caso contrário, vá para o passo 6.
6. (Segundo teste de otimalidade) Calcule todos  $p_i$  associados à variáveis estritamente positivas no quadro simplex ótimo de LR. Faça os outros  $p_i$  iguais a 0. Para cada relação  $r_k \in H$ , calcule

$$\pi_k = \min_{i \in I_k} p_i$$

e faça

$$\Pi = \max_{k \in \{k; r_k \in H\}} \pi_k.$$

Se  $\bar{F} \leq (c_1 x + d_1 y) + \Pi$ , vá para o passo 13. Caso contrário, vá para o passo 7.

7. (Segundo teste de viabilidade) Se LR é inviável, vá para o passo 13. Caso contrário, vá para o passo 8.

8. (Primeiro teste de otimalidade condicional) Considere novamente a solução ótima  $(x, y)$  de LR e os termos  $p_i$ . Para todo  $i$  tal que  $\bar{F} \leq (c_1x + d_1y) + \pi_i$ , fixe  $\alpha_i = 0$  e atualize  $H$ . Vá para o passo 9.
9. (Terceiro teste de otimalidade) Se  $H$  contém uma relação  $r_k$  tal que  $\alpha_j = 0, \forall j \in I_k$ , vá para o passo 13. Caso contrário, vá para o passo 10.
10. (Teste de otimalidade relacional) Para todo  $j$  onde  $y_j$  ainda é presente no subproblema corrente (não foi eliminada), adicione a  $H$  todas relações do tipo (5.1) que não são redundantes. Remova de  $H$  aquelas relações que se tornam redundantes. Vá para o passo 11.
11. (Segundo teste de otimalidade condicional) Se  $H$  contém uma relação  $r_k$  tal que  $\alpha_j = 0$ , para todo  $j \in I_k$  exceto para um índice  $i \in I_k$ , fixe  $\alpha_i = 1$ , elimine do subproblema a variável  $y_j$  presente na  $i$ -ésima restrição que produz o menor preenchimento e vá para o passo 2. Caso contrário, vá para o passo 12.
12. (Ramificação) Aplique uma regra de ramificação para fixação de alguma variável livre  $\alpha_i$  ou uma relação  $r_k \in H$  tal que todas variáveis  $\alpha_i$  com  $i \in I_k$  são livres. No primeiro caso, fixe  $\alpha_i = 1$ , e no segundo caso fixe em 1 a primeira variável em  $r_k$ . Elimine do subproblema a variável  $y_j$  presente na  $i$ -ésima restrição que produz o menor preenchimento. Vá para o passo 2.
13. (*Backtracking*) Se o nó corrente foi obtido pela fixação de uma variável  $\alpha_k$  (e não por uma relação do tipo (5.1)), encontre a última variável  $\alpha_i$  fixada em 1, fixe esta variável em 0 e libere todas as variáveis  $\alpha_j$  fixadas em 0 depois que  $\alpha_i$  foi fixada em 1. Caso contrário, encontre a última relação  $r_k$  tal que menos que  $|I_k|$  ramificações foram exploradas (isto é, existem possibilidades não exploradas de vetores  $\alpha$  que satisfazem  $r_k$ ) e considere a ramificação seguinte. Se não houver tais variáveis ou relações, pare e retorne  $(\bar{x}, \bar{y})$  como solução ótima ( $\bar{F}$  é valor ótimo). Caso contrário, vá para o passo 2.

---

Quando  $B_1 = 0$ , um ponto inicial  $(x^0, y^0) \in \Upsilon$  pode ser obtido de acordo com a seção 6.2 (página 75). Várias regras de ramificação são apresentadas por Hansen, Jaumard e Savard. Uma delas, que apresenta melhores resultados computacionais [13] é a seguinte: (BR6) selecione uma relação  $r_k \in H$  com pelo menos duas variáveis  $\alpha_{i_1}$  e  $\alpha_{i_2}$  tais que suas variáveis de folga correspondentes  $z_{n_x+m_u+i_1}$  e  $z_{n_x+m_u+i_2}$  são básicas e

ainda  $\sum_{i \in I_k} \lambda_i z_{n_x+m_u+i}$  é máximo, onde  $\lambda_i$  é variável dual correspondente (para o baixo nível). Se não existe tal relação, selecione uma variável  $\alpha_i$  com máximo  $p_{n_x+m_u+i}$  e tal que  $\lambda_i z_{n_x+m_u+i} > 0$ .

## 5.4 Outras abordagens

Abordagens via penalização são expostas na tese de M. Campêlo [11]. Em [33] encontra-se uma boa discussão sobre penalizações para uma generalização do problema de dois níveis.

Outras abordagens via heurísticas são utilizadas na literatura. Em [3] encontram-se métodos baseados em inteligência artificial e busca tabu; Em [10, 9], uma abordagem com algoritmo genético; Em [36], simulated annealing; Em [29, 42], busca tabu; Em [25], *particle swarm*; Em [37], uma abordagem baseada em técnicas de decomposição. Outras abordagens interessantes para modelos lineares ou não podem ser vistas em [16, 8].

## 6 Um novo algoritmo para soluções ótimas locais no caso linear

Neste Capítulo proporemos uma técnica para encontrar soluções ótimas locais do modelo

$$\begin{aligned} \text{LBLP: } \min_{x,y} \{c_1x + d_1y\} \\ \text{s.a. } A_1x + B_1y \leq b_1, \quad x \geq 0 \\ y \in \arg \min_y d_2y \\ \text{s.a. } A_2x + B_2y \leq b_2, \quad y \geq 0 \end{aligned}$$

estudado no Capítulo 3.

A idéia é a seguinte: Dado um ponto admissível inicial, procuramos em suas faces (de  $\Omega_l$ ) adjacentes (de maior dimensão quanto possível) aquelas que sejam racionais, isto é, que são subfaces de faces de uma escrita de graf  $S$ , e que melhoram a função objetivo de alto nível, quando minimizamos a mesma sobre tal face. Caso encontremos uma face que tenha um ponto admissível que diminua a função objetivo de alto nível, atualizamos tal ponto como estado corrente, e repetimos o processo para este ponto, de modo a não verificar faces já analisadas. O procedimento pára quando encontramos um  $(x^r, y^r)$  tal que todas suas faces adjacentes são não racionais ou não dão soluções melhores. Neste caso, o ponto  $(x^r, y^r)$  é uma solução ótima local de LBLP.

### 6.1 Formulação

Sob a hipótese de que todas as faces de  $\Omega$  são não degeneradas (veja a Definição 3.1.2, página 28), procurar faces de  $\Omega_l$  adjacentes a um ponto e que estão em graf  $S$  pode ser feito zerando as variáveis duais correspondentes às restrições ativas do baixo

nível (veja o sistema (3.2)-(3.6), página 29). Ao fazer isso, “liberamos” uma restrição para que seja não ativa. No algoritmo seguinte, utilizamos três conjuntos para controle dos índices das restrições:  $L$  para índices de restrições que foram liberadas em todas as faces adjacentes estudadas num ponto corrente;  $Q$  para índices de restrições que foram liberadas no estudo de uma face adjacente a um ponto corrente; e  $R$  para índices de restrições que foram testadas (quanto à racionalidade da face que induzem) no estudo de uma face adjacente a um ponto corrente.

---

**Algoritmo 3** Solução ótima local
 

---

**Entrada:** LBLP viável e não ilimitado. Supõe-se todas as faces de  $\Omega$  não degeneradas.

**Saída:** Uma solução ótima local de LBLP.

---

1. (Inicialização) Calcule  $(x^0, y^0) \in \Upsilon$  usando uma heurística. Faça  $L \leftarrow \emptyset$ ,  $k \leftarrow 0$  e defina

$$\Lambda^k = \{i; (A_2 x^k + B_2 y^k - b_2)_i = 0\} \cup \{m_l + i; y_i^k = 0\}.$$

Se  $(x^0, y^0)$  for vértice de  $\Omega_l$ , vá para o passo 2. Caso contrário, execute o passo 5 com  $Q = \emptyset$ .

2. Faça  $Q \leftarrow \emptyset$  e  $R \leftarrow \emptyset$ . Vá para o passo 3.

3. (Detecção de faces racionais adjacentes) Se  $Q \neq \emptyset$ , escolha  $j \in \Lambda^k \setminus R$ .

Se  $Q = \emptyset$ , escolha se possível  $j \in \Lambda^k \setminus (L \cup R)$ . Se tal escolha não for possível, vá para o passo 4.

Verifique se o sistema

$$\begin{aligned} (B_2^t \lambda + d_2^t)_i &= 0, & m_l + i &\in \{j\} \cup \mathcal{C}\Lambda \\ (B_2^t \lambda + d_2^t)_i &\geq 0, & m_l + i &\in \Lambda \setminus \{j\} \\ \lambda_i &= 0, & i &\in \{j\} \cup \mathcal{C}\Lambda \\ \lambda_i &\geq 0, & i &\in \Lambda \setminus \{j\} \end{aligned}$$

admite solução, onde  $\Lambda = \Lambda^k \setminus R$  e  $\mathcal{C}\Lambda = \{1, \dots, m_l + n_y\} \setminus \Lambda$ . Se admite solução, digamos  $\lambda^k$ , faça

$$R \leftarrow R \cup A, \quad Q \leftarrow Q \cup A \quad \text{e} \quad L \leftarrow L \cup A$$

onde  $A = \{i \in \Lambda^k; \lambda_i^k = 0\} \cup \{m_l + i \in \Lambda^k; (B_2^t \lambda^k + d_2^t)_i = 0\} \subset \Lambda^k$ . Se o sistema anterior não admite solução, faça  $R \leftarrow R \cup \{j\}$ . Se  $\Lambda^k \setminus R = \emptyset$  vá para o passo 4. Se  $\Lambda^k \setminus R \neq \emptyset$  repita o passo 3.

4. (Teste de parada) Se  $Q \neq \emptyset$ , vá para o passo 5. Se  $Q = \emptyset$  pare e retorne  $(x^k, y^k)$  como solução ótima local de LBLP.
5. (Passo de minimização) Resolva

$$\begin{aligned}
& \min_{x,y} \{c_1x + d_1y\} \\
& \text{s.a. } A_1x + B_1y \leq b_1, \quad x \geq 0 \\
& \quad (A_2x + B_2y - b_2)_i = 0, \quad i \in \Lambda \\
& \quad (A_2x + B_2y - b_2)_i \leq 0, \quad i \notin \Lambda \\
& \quad y_i = 0, \quad m_l + i \in \Lambda \\
& \quad y_i \geq 0, \quad m_l + i \notin \Lambda
\end{aligned}$$

onde  $\Lambda = \Lambda^k \setminus Q$ , e tome uma solução ótima  $(x^{k+1}, y^{k+1})$  sua (este problema admite solução pois  $(x^k, y^k)$  é um ponto viável seu). Se  $c_1x^{k+1} + d_1y^{k+1} < c_1x^k + d_1y^k$ , faça  $k \leftarrow k + 1$  e  $L \leftarrow \emptyset$ . Vá para o passo 2.

**Teorema 6.1.1.** *O Algoritmo 3 encontra uma solução ótima local de LBLP em finitos passos.*

*Demonstração.* Tomemos  $(x^0, y^0) \in \Upsilon$ . O algoritmo gera um número finito  $(x^0, y^0), \dots, (x^r, y^r)$  de pontos em  $\Upsilon$  visto que gerado  $(x^{k+1}, y^{k+1})$  com  $c_1x^{k+1} + d_1y^{k+1} < c_1x^k + d_1y^k$ , o algoritmo não volta a analisar  $(x^0, y^0), \dots, (x^k, y^k)$ ; para todo  $(x^k, y^k)$ , o laço 3 termina pois a cada iteração,  $R$  é acrescido até que  $\Lambda^k \setminus R = \emptyset$ ; para todo  $(x^k, y^k)$  no passo 5, ou passamos a um novo  $(x^{k+1}, y^{k+1}) \neq (x^k, y^k)$  ou repetimos esse processo com sucessivos fracassos de diminuição da função  $c_1x + d_1y$ , e neste caso  $L$  é acrescido até que  $\Lambda^k \setminus (L \cup R) = \emptyset$ . No último caso, de acordo com o passo 3,  $Q = \emptyset$  e o algoritmo termina com o passo 4. Vale observarmos que esta situação sempre ocorre pois existem finitas faces em  $\Omega$ .

Seja  $(x^r, y^r)$  um ponto obtido no fim do algoritmo e  $U^r$  o conjunto dos índices das restrições de alto nível ativas em  $(x^r, y^r)$ . Tomemos  $\delta > 0$  tal que para cada  $(x, y) \in B_\delta(x^r, y^r)$  tenhamos

$$\begin{aligned}
(A_1x + B_1y - b_1)_i &< 0, \quad x_i > 0, \quad \forall i \notin U^r \\
(A_2x + B_2y - b_2)_i &< 0, \quad y_i > 0, \quad \forall i \notin \Lambda^r.
\end{aligned} \tag{6.1}$$

Para cada  $j \in \Lambda^r$ , consideremos a face  $u_j$  de  $\Omega_u$ , adjacente à  $(x^r, y^r)$ , definida por

$$\begin{aligned} u &= \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; (A_1x + B_1y - b_1)_i = 0, \quad i \in U^r \setminus \{j\}, \\ &\quad (A_1x + B_1y - b_1)_i \leq 0, \quad i \notin U^r \setminus \{j\}, \\ &\quad x_i = 0, \quad m_u + i \in U^r \setminus \{j\}, \\ &\quad x_i \geq 0, \quad m_u + i \notin U^r \setminus \{j\}\}. \end{aligned}$$

Seja  $u^r$  a face de  $\Omega_u$  cujas restrições/variáveis de igualdade têm índices em  $U^r$ . Da mesma forma, para cada  $j \in \Lambda^r$ , consideremos a face  $q_j$  de  $\Omega_l$ , adjacente a  $(x^r, y^r)$ , caracterizada por

$$\begin{aligned} q_j &= \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; (A_2x + B_2y - b_2)_i = 0, \quad i \in \Lambda^r \setminus \{j\}, \\ &\quad (A_2x + B_2y - b_2)_i \leq 0, \quad i \notin \Lambda^r \setminus \{j\}, \\ &\quad y_i = 0, \quad m_l + i \in \Lambda^r \setminus \{j\}, \\ &\quad y_i \geq 0, \quad m_l + i \notin \Lambda^r \setminus \{j\}\}. \end{aligned}$$

Seja  $q^r$  a face de  $\Omega_l$  cujas restrições/variáveis de igualdade têm índices em  $\Lambda^r$ .

Agora, suponhamos por contradição que  $(x^r, y^r)$  não seja solução ótima local de LBLP. Existe então um  $(\bar{x}, \bar{y}) \in B_\delta(x^r, y^r) \cap \Upsilon$  com

$$c_1\bar{x} + d_1\bar{y} < c_1x^r + d_1y^r.$$

Afirmamos que existem  $n_x + n_y - |U^r| - |\Lambda^r|$  vetores linearmente independentes (LI), que denotaremos por  $a_j$ ,  $j \in R$ , tais que  $(x^r, y^r) + a_j \in q^r \cap u^r$ . De fato, consideremos a matriz

$$C = \begin{bmatrix} \left[ \begin{array}{cc} A_1 & B_1 \end{array} \right]_{j \in U^r} \\ \left[ \begin{array}{cc} A_2 & B_2 \end{array} \right]_{j \in \Lambda^r} \end{bmatrix}$$

e a aplicação linear  $T$  associada a  $C$ . Pela não degeneração das faces de  $\Omega$ ,  $C$  tem posto igual ao seu número de linhas, a saber,  $|U^r| + |\Lambda^r|$ . Assim,  $\dim \text{Im}(T) = |U^r| + |\Lambda^r|$  e o Teorema do Núcleo e da Imagem (Teorema A.1.1, página 115) nos diz que  $\dim \text{Ker}(T) = n_x + n_y - |U^r| - |\Lambda^r|$ . Logo, existem  $n_x + n_y - |U^r| - |\Lambda^r|$  vetores  $d_j$  LI entre si satisfazendo

$Cz = 0$ . Como

$$C \begin{bmatrix} x^r \\ y^r \end{bmatrix} = \begin{bmatrix} \left[ \begin{array}{c} b_1 \\ b_2 \end{array} \right]_{j \in U^r} \\ \left[ \begin{array}{c} b_1 \\ b_2 \end{array} \right]_{j \in \Lambda^r} \end{bmatrix} = b',$$

$b' \in \text{Im}(T)$  e do Teorema A.1.2 (página 116) segue que o conjunto das soluções de  $Cz = b'$  é uma variedade afim paralela à  $\text{Ker}(T)$ . Levando em consideração (6.1), existem  $\mu_j > 0$ ,  $j \in R$ , suficientemente pequenos tais que  $(x^r, y^r) + \mu_j d_j \in q^r \cap u^r$  (observamos que  $(x^r, y^r) + \mu_j d_j$  é solução de  $Cz = b'$ ). Tomemos então  $a_j = \mu_j d_j$ .

Podemos supor sem perda de generalidade que os vetores  $a_j$  são ortogonais entre si (isto é,  $a_j a_i^t = 0$  sempre que  $j \neq i$ ) [27, Processo de ortogonalização de Gram-Schmidt, página 128]. Observamos ainda que todos  $a_j$  são ortogonais às linhas de  $C$ , visto que são soluções do sistema homogêneo  $Cz = 0$ . Então os vetores  $C_1, \dots, C_{|U^r|+|\Lambda^r|}$  e  $a_j$ ,  $j \in R$ , são LI. De fato, se fosse  $a_j = \sum_{i \neq j} \eta_i a_i + \sum_i \beta_i C_i$  teríamos

$$0 < a_j a_j^t = \sum_{i \neq j} \eta_i (a_i a_j^t) + \sum_i \beta_i (C_i a_j^t) = 0,$$

um absurdo. Da mesma forma, se  $C_j = \sum_i \eta_i a_i + \sum_{i \neq j} \beta_i C_i$ , pelo menos um  $\eta_k$  é não nulo pois as linhas de  $C$  são LI entre si. Daí

$$0 = C_j a_k^t = \sum_i \eta_i (a_i a_k^t) + \sum_{i \neq j} \beta_i (C_i a_k^t) = \eta_k (a_k a_k^t) > 0,$$

absurdo.

Com isto, para cada  $j \in R$ , construímos uma restrição de desigualdade  $\sigma_j$  ativa em  $(x^r, y^r)$  e que satisfaz  $(\bar{x}, \bar{y})$ , como segue. Se  $a_j \begin{bmatrix} \bar{x}^t & \bar{y}^t \end{bmatrix} \leq a_j \begin{bmatrix} x^{rt} & y^{rt} \end{bmatrix}^t$  tomamos  $\sigma_j : a_j \begin{bmatrix} x^t & y^t \end{bmatrix}^t \leq a_j \begin{bmatrix} x^{rt} & y^{rt} \end{bmatrix}^t$  e caso contrário, tomamos  $\sigma_j : (-a_j) \begin{bmatrix} x^t & y^t \end{bmatrix}^t \leq (-a_j) \begin{bmatrix} x^{rt} & y^{rt} \end{bmatrix}^t$ . Consideremos então, para cada  $j \in R$ , o conjunto

$$w_j = \{(x, y) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}; (x, y) \text{ satisfaz } \sigma_j\}.$$

Tomemos  $W = \bigcap_{j \in R} w_j$ , e consideremos o problema linear de dois níveis

$$\begin{aligned} \text{LBLP}^+ : \quad & \min_{x, y} \{c_1 x + d_1 y\} \\ \text{s.a.} \quad & (x, y) \in \Omega_u \cap W \\ & y \in \arg \min_y d_2 y \\ & \text{s.a. } (x, y) \in \Omega_l. \end{aligned}$$

Os pontos  $(x^r, y^r)$  e  $(\bar{x}, \bar{y})$  são admissíveis para LBLP<sup>+</sup>. Mais ainda,  $(x^r, y^r)$  é vértice não degenerado de  $\Omega^+ = \Omega_u \cap \Omega_l \cap W = \Omega \cap W$ . Definindo  $W_j = \bigcup_{j \in R \setminus \{j\}} w_j$ , temos que as  $n_x + n_y$  arestas de  $\Omega^+$  emanando de  $(x^r, y^r)$  são  $q_j \cap u^r \cap W$ ,  $j \in \Lambda^r$ ,  $q^r \cap u_j \cap W$ ,  $j \in U^r$ , e  $q^r \cap u^r \cap W_j$ ,  $j \in R$ . Também,

$$\begin{aligned} \emptyset \neq q_j \cap u^r \cap W &\subset q_j \cap \Omega_u, \quad \forall j \in \Lambda^r, \\ \emptyset \neq q^r \cap u_j \cap W &\subset q^r \cap \Omega_u, \quad \forall j \in U^r, \quad \text{e} \\ \emptyset \neq q^r \cap u^r \cap W_j &\subset q^r \cap \Omega_u, \quad \forall j \in R. \end{aligned} \tag{6.2}$$

Para simplificar a escrita, vamos supor que os conjuntos de índices  $U^r$ ,  $\Lambda^r$  e  $R$  sejam disjuntos, e que  $U^r \cap \Lambda^r \cap R = \{1, \dots, n_x + n_y\}$ . As arestas de  $\Omega^+$  emanando de  $(x^r, y^r)$  definem direções  $\gamma_1, \dots, \gamma_{n_x+n_y}$ . Sendo  $\gamma = (\bar{x}, \bar{y}) - (x^r, y^r)$ , existem  $\alpha_j \geq 0$  tais que

$$\gamma = \sum_i \alpha_i \gamma_i. \tag{6.3}$$

Como  $(\bar{x}, \bar{y}) \in B_\delta(x^r, y^r)$ , segue de (6.1) que todas as restrições não ativas em  $(x^r, y^r)$  também são não ativas em  $(\bar{x}, \bar{y})$ . Por outro lado, pode ocorrer de restrições ativas em  $(x^r, y^r)$  serem não ativas em  $(\bar{x}, \bar{y})$ . Sejam portanto  $\bar{U} \subset U^r$  e  $\bar{\Lambda} \subset \Lambda^r$  os conjuntos dos índices das restrições do alto e baixo níveis, respectivamente, que são ativas em  $(x^r, y^r)$  mas não em  $(\bar{x}, \bar{y})$ . Para cada  $j$ , escrevendo  $\gamma_j = (x_j, y_j) - (x^r, y^r)$ , de (6.1) e do fato de  $(x^r, y^r)$  ser vértice não degenerado de  $\Omega^+$ , podemos supor que  $(x_j, y_j)$  realiza como desigualdade estrita todas as restrições de desigualdade em sua respectiva aresta. Temos então que, para  $j \in \Lambda^r \setminus \bar{\Lambda}$ , se  $j \leq m_l$ ,

$$\begin{aligned} 0 &= (A_2 \bar{x} + B_2 \bar{y} - b_2)_j \\ &= A_2 \left( x^r + \sum_i \alpha_i (x_i - x^r) \right)_j + B_2 \left( y^r + \sum_i \alpha_i (y_i - y^r) \right)_j - (b_2)_j \\ &= \underbrace{(A_2 x^r + B_2 y^r - b_2)_j}_0 + \sum_i \alpha_i [A_2 (x_i - x^r) + B_2 (y_i - y^r)]_j \\ &= \sum_i \alpha_i (A_2 x_i + B_2 y_i - b_2)_j = \alpha_j \underbrace{(A_2 x_j + B_2 y_j - b_2)_j}_{<0} \Rightarrow \alpha_j = 0, \end{aligned}$$

e se  $j > m_l$ ,

$$0 = \bar{y}_{j-m_l} = \underbrace{(y^r)_{j-m_l}}_0 + \sum_i \alpha_i (y_i - y^r)_{j-m_l} = \sum_i \alpha_i (y_i)_{j-m_l} = \alpha_j \underbrace{(y_j)_{j-m_l}}_{>0} \Rightarrow \alpha_j = 0.$$

Ou seja,  $\alpha_j = 0$  sempre que  $j \in \Lambda^r \setminus \bar{\Lambda}$ . Logo, de (6.3) obtemos

$$\gamma = \sum_{i \in \bar{\Lambda}} \alpha_i \gamma_i + \sum_{i \in U^r \cup R} \alpha_i \gamma_i. \quad (6.4)$$

Agora, como  $(\bar{x}, \bar{y})$  e  $(x^r, y^r)$  são admissíveis para LBLP, as faces  $q^r$  e  $q_j$ ,  $j \in \bar{\Lambda}$ , são racionais (estão em graf  $S$ ). Segue portanto de (6.2) que as arestas de  $\Omega^+$  emanando de  $(x^r, y^r)$  são formadas de pontos admissíveis de LBLP<sup>+</sup>, exceto possivelmente as relativas aos índices em  $\Lambda^r \setminus \bar{\Lambda}$ . Logo as direções  $\gamma_j$ ,  $j \in U^r \cup \bar{\Lambda} \cup R$ , são viáveis para LBLP<sup>+</sup>.

Ao fim do algoritmo, o passo 5 garante que  $(x^r, y^r)$  é solução ótima do problema

$$\min_{x,y} \{c_1 x + d_1 y\} \quad \text{s.a.} \quad (x, y) \in q^r \cap \Omega_u. \quad (6.5)$$

É importante observarmos que se  $r = 0$ , a execução do passo 5 logo após o cálculo de  $(x^0, y^0)$  garantiria tal afirmação. Não é necessário executar o passo 5 se  $(x^0, y^0)$  for vértice de  $\Omega_l$  pois teríamos  $q^0 = \{(x^0, y^0)\}$ . Com isto, as duas últimas inclusões em (6.2) dizem que as direções  $\gamma_j$ ,  $j \in U^r \cup R$ , não são de descida. Também, ao fim do algoritmo temos  $\Lambda^r \setminus (L \cup R) = Q = \emptyset$ . Isto significa que cada índice em  $\Lambda^r$  foi analisado pelo algoritmo, mesmo que vários índices foram escolhidos de uma só vez no passo 3 (neste caso todas as faces  $q_j$  com esses índices são analisadas de uma só vez). Neste caso o passo 5 garante que  $(x^r, y^r)$  é solução ótima dos problemas

$$\min_{x,y} \{c_1 x + d_1 y\} \quad \text{s.a.} \quad (x, y) \in q_j \cap \Omega_u,$$

para todos  $j \in \bar{\Lambda}$  ( $j \in \bar{\Lambda}$  é escolhido com sucesso no passo 3 pois  $q_j$  é racional). Daí segue da primeira inclusão em (6.2) que as direções  $\gamma_j$ ,  $j \in \bar{\Lambda}$ , não são de descida.

Finalmente mostremos que  $(x^r, y^r)$  é solução ótima local de LBLP. Pela nossa suposição inicial, a direção  $\gamma = (\bar{x}, \bar{y}) - (x^r, y^r)$  é de descida para a função  $c_1 x + d_1 y$ . Segue de (6.4) que

$$0 > \begin{bmatrix} c_1 & d_1 \end{bmatrix} \gamma = \sum_{i \in \bar{\Lambda}} \alpha_i \begin{bmatrix} c_1 & d_1 \end{bmatrix} \gamma_i + \sum_{i \in U^r \cup R} \alpha_i \begin{bmatrix} c_1 & d_1 \end{bmatrix} \gamma_i \geq 0,$$

um absurdo. Concluimos portanto que  $(x^r, y^r)$  é solução ótima local de LBLP.  $\square$

O próximo exemplo ilustra a necessidade da execução do passo 5 logo após o cálculo de  $(x^0, y^0)$ , caso este não seja vértice de  $\Omega_l$ .

**Exemplo 6.1.2.** Considere o problema bidimensional de dois níveis

$$\min_{x,y} x - y$$

$$\text{s.a. } 0 \leq x \leq 3$$

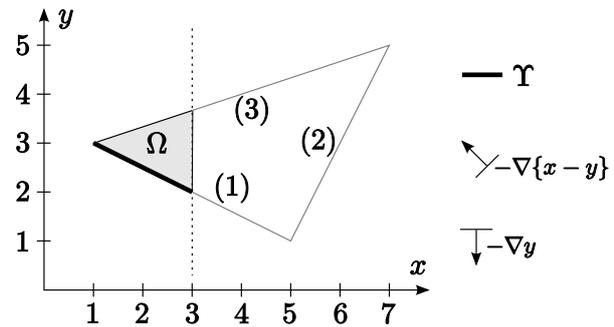
$$y \in \arg \min_y y$$

$$\text{s.a. } -x - 2y \leq -7 \quad (1)$$

$$2x - y \leq 9 \quad (2)$$

$$-x + 3y \leq 8 \quad (3)$$

$$y \geq 0$$



Suponhamos que o passo 1 do Algoritmo 3 retorne  $(x^0, y^0) = (3, 2)$  (esse é o retorno utilizando a técnica da seção 6.2). Se não executássemos o passo 5, o sistema resultante do passo 3 seria

$$\begin{bmatrix} 2 & -1 & 3 \end{bmatrix} \lambda + \begin{bmatrix} 1 \end{bmatrix} = 0, \\ \lambda = 0,$$

onde  $\lambda_1$  foi fixada em 0. Esse sistema não possui solução. Não há mais restrições ativas em  $(3, 2)$  a testar, e então o algoritmo terminaria ( $Q = \emptyset$  no passo 4) com  $(x^0, y^0)$ , que não é solução ótima local. Geometricamente vemos que a única solução ótima local (e global) é  $(1, 3)$ , exatamente o ponto que seria obtido no passo 5 com  $\Lambda = \Lambda^0$ .  $\square$

Apesar de o algoritmo analisar todas as arestas num dado ponto final  $(x^r, y^r)$ , não é garantido que todas suas faces adjacentes são analisadas, como ilustra a figura 7.

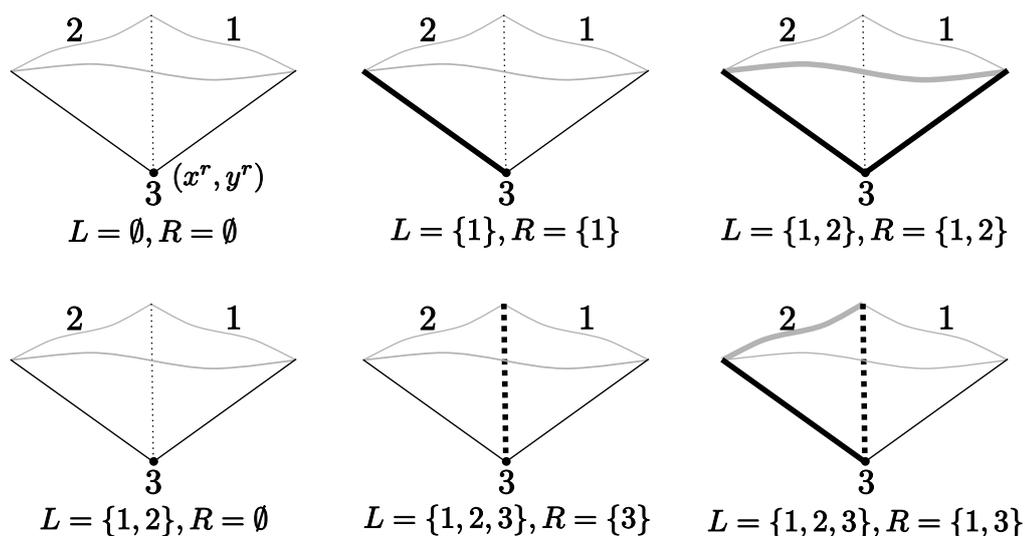


Figura 7: Situação hipotética em que  $r = 0$  e todas faces de dimensão 2 adjacentes a  $(x^r, y^r)$  são racionais. As linhas em destaque representam faces analisadas no estado corrente. A face de dimensão 2 relativa à restrição 1 não foi analisada. Aqui foram representados as execuções de dois laços do passo 3.

A hipótese de que todas as faces de  $\Omega$  sejam não degeneradas não é forte na prática. O próximo resultado justifica esse fato.

**Teorema 6.1.3.** *O conjunto dos problemas lineares de dois níveis cujas faces de  $\Omega$  são todas não degeneradas é aberto e denso no conjunto dos problemas lineares de dois níveis.*

*Demonstração.* Ver [41, Teorema 3]. □

Em outras palavras, o Teorema 6.1.3 diz que a possibilidade de gerarmos aleatoriamente um problema linear de dois níveis com alguma face degenerada é praticamente nula, e que com pequenas perturbações nos dados de um problema desses podemos gerar um outro com todas as faces não degeneradas.

Vejamos agora um exemplo da aplicação do Algoritmo 3 para o caso em que  $B_1 = 0$  (veja a seção 6.2 para um método de obtenção de  $(x^0, y^0)$  no passo 1).

**Exemplo 6.1.4.** Considere o problema bidimensional de dois níveis

$$\min_{x,y} F(x,y) = -x - 2y$$

$$\text{s.a. } 0 \leq x \leq 7$$

$$y \in \arg \min_y$$

$$\text{s.a. } -x - 2y \leq -6 \quad (1)$$

$$-x - y \leq -4 \quad (2)$$

$$-3x - y \leq -6 \quad (3)$$

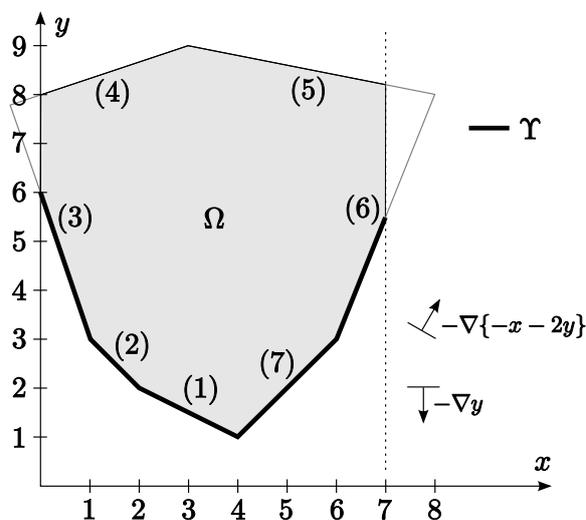
$$-x + 3y \leq 24 \quad (4)$$

$$x + 5y \leq 48 \quad (5)$$

$$5x - 2y \leq 24 \quad (6)$$

$$x - y \leq 3 \quad (7)$$

$$y \geq 0 \quad (8)$$



Temos  $B_2^t = \begin{bmatrix} -2 & -1 & -1 & 3 & 5 & -2 & -1 \end{bmatrix}$  e  $d_2 = \begin{bmatrix} 1 \end{bmatrix}$ . Apliquemos o algoritmo 3.

**PASSO 1:** O problema

$$\min_{x,y} y \quad \text{s.a.} \quad 0 \leq x \leq 7, \quad (1) - (8)$$

tem como única solução ótima  $(x^0, y^0) = (4, 1)$ , que é um ponto admissível e vértice de  $\Omega_t$ . Fazemos  $L = \emptyset$ ,  $k = 0$  e  $\Lambda^0 = \{1, 7\}$ .

**PASSO 2:** Fazemos  $Q = R = \emptyset$ .

**PASSO 3:** Como  $Q = \emptyset$ , escolhemos  $j = 1 \in \Lambda^0 \setminus (L \cup R) = \{1, 7\}$ . Temos  $\Lambda = \Lambda^0 \setminus R = \{1, 7\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, 2, 3, 4, 5, 6, 8\}$  e  $\Lambda \setminus \{j\} = \{7\}$ .

O sistema

$$-2\lambda_1 - \lambda_2 - \lambda_3 + 3\lambda_4 + 5\lambda_5 - 2\lambda_6 - \lambda_7 + 1 = 0$$

$$\lambda_1 = \dots = \lambda_6 = 0, \quad \lambda_7 \geq 0$$

tem solução, a saber,  $\lambda = (0, 0, 0, 0, 0, 0, 1)$ . Temos  $A = \{1\}$ . Fazemos  $L = Q = R = \emptyset \cup A = \{1\}$ . Temos  $\Lambda^0 \setminus R = \{7\} \neq \emptyset$ .

**PASSO 3:** Como  $Q \neq \emptyset$ , escolhemos  $j = 7 \in \Lambda^0 \setminus R = \{7\}$ . Temos  $\Lambda = \Lambda^0 \setminus R = \{7\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, \dots, 8\}$  e  $\Lambda \setminus \{j\} = \{1\}$ .

O sistema

$$\begin{aligned} -2\lambda_1 - \lambda_2 - \lambda_3 + 3\lambda_4 + 5\lambda_5 - 2\lambda_6 - \lambda_7 + 1 &= 0 \\ \lambda_1 = \dots = \lambda_7 &= 0 \end{aligned}$$

não tem solução. Fazemos  $R = \{1\} \cup \{7\} = \Lambda^0 \Rightarrow \Lambda^0 \setminus R = \emptyset$ .

**PASSO 4:** Temos  $Q \neq \emptyset$ .

**PASSO 5:** Temos  $\Lambda = \Lambda^0 \setminus Q = \{1, 7\} \setminus \{1\} = \{7\}$ . O problema

$$\begin{aligned} \min_{x,y} F(x, y) &= -x - 2y \\ \text{s.a. } 0 &\leq x \leq 7 \\ x - y &= 3 \quad ((7) \text{ está ativa}) \\ (1) - (6), (8) \end{aligned}$$

tem como única solução ótima  $(x^1, y^1) = (6, 3)$ . Como  $F(6, 3) = -12 < -6 = F(4, 1)$ , fazemos  $k = 1$  e  $L = \emptyset$ .

**PASSO 2:** Fazemos  $Q = R = \emptyset$ .

**PASSO 3:** Temos  $\Lambda^1 = \{6, 7\}$ .

Como  $Q = \emptyset$ , escolhemos  $j = 6 \in \Lambda^1 \setminus (L \cup R) = \{6, 7\}$ . Temos  $\Lambda = \Lambda^1 \setminus R = \{6, 7\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, 2, 3, 4, 5, 6, 8\}$  e  $\Lambda \setminus \{j\} = \{7\}$ . O sistema correspondente é o mesmo que o da primeira execução do passo 3 cuja solução é  $\lambda = (0, 0, 0, 0, 0, 0, 1)$ . Temos  $A = \{6\}$ . Fazemos  $L = Q = R = \{6\}$ . Temos  $\Lambda^1 \setminus R = \{7\} \neq \emptyset$ .

**PASSO 3:** Como  $Q \neq \emptyset$ , escolhemos  $j = 7 \in \Lambda^1 \setminus R = \{7\}$ . Temos  $\Lambda = \Lambda^1 \setminus R = \{7\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, \dots, 8\}$  e  $\Lambda \setminus \{j\} = \{6\}$ . É fácil verificar que neste caso o sistema do passo 3 não possui solução. Fazemos  $R = \{6\} \cup \{7\} = \Lambda^1 \Rightarrow \Lambda^1 \setminus R = \emptyset$ .

**PASSO 4:** Temos  $Q \neq \emptyset$ .

**PASSO 5:** Temos  $\Lambda = \Lambda^1 \setminus Q = \{6, 7\} \setminus \{6\} = \{7\}$ . Neste caso o problema do passo 5 é o mesmo que o do passo 5 anterior, e a solução corrente não é melhorada.

**PASSO 2:** Fazemos  $Q = R = \emptyset$ .

**PASSO 3:** Como  $Q = \emptyset$ , escolhemos  $j = 7 \in \Lambda^1 \setminus (L \cup R) = \{6, 7\} \setminus \{6\} = \{7\}$ . Temos  $\Lambda = \Lambda^1 \setminus R = \{6, 7\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, 2, 3, 4, 5, 7, 8\}$  e  $\Lambda \setminus \{j\} = \{6, 7\} \setminus \{7\} = \{6\}$ .

O sistema

$$\begin{aligned} -2\lambda_1 - \lambda_2 - \lambda_3 + 3\lambda_4 + 5\lambda_5 - 2\lambda_6 - \lambda_7 + 1 &= 0 \\ \lambda_1 = \dots = \lambda_5 = \lambda_7 &= 0, \quad \lambda_6 \geq 0 \end{aligned}$$

tem solução, a saber,  $\lambda = (0, 0, 0, 0, 0, \frac{1}{2}, 0)$ . Temos  $A = \{7\}$ . Fazemos  $L = \{6, 7\}$ ,  $Q = R = \{7\}$  e  $j = 1$ . Temos  $\Lambda^1 \setminus R = \{7\} \neq \emptyset$ .

**PASSO 3:** Como  $Q \neq \emptyset$ , escolhemos  $j = 6 \in \Lambda^1 \setminus R = \{6\}$ . Temos  $\Lambda = \Lambda^1 \setminus R = \{6\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, \dots, 8\}$  e  $\Lambda \setminus \{j\} = \emptyset$ . Neste caso o sistema do passo 3 não possui solução. Fazemos  $R = \{7\} \cup \{6\} = \Lambda^1 \Rightarrow \Lambda^1 \setminus R = \emptyset$ .

**PASSO 4:** Temos  $Q \neq \emptyset$ .

**PASSO 5:** Temos  $\Lambda = \Lambda^1 \setminus Q = \{6, 7\} \setminus \{7\} = \{6\}$ . O problema

$$\begin{aligned} \min_{x,y} F(x, y) &= -x - 2y \\ \text{s.a. } 0 &\leq x \leq 7 \\ 5x - 2y &= 24 \quad ((6) \text{ está ativa}) \\ (1) - (5), (7), (8) \end{aligned}$$

tem como única solução ótima  $(x^2, y^2) = (7, \frac{11}{2})$ . Como  $F(7, \frac{11}{2}) = -18 < -12 = F(6, 3)$ , fazemos  $k = 2$  e  $L = \emptyset$ .

**PASSO 2:** Fazemos  $Q = R = \emptyset$ .

**PASSO 3:** Temos  $\Lambda^2 = \{6\}$ .

Como  $Q = \emptyset$ , escolhemos  $j = 6 \in \Lambda^2 \setminus (L \cup R) = \{6\}$ . Temos  $\Lambda = \Lambda^2 \setminus R = \{6\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, \dots, 8\}$  e  $\Lambda \setminus \{j\} = \emptyset$ . Neste caso o sistema do passo 3 não possui solução. Fazemos  $R = \{6\} = \Lambda^2 \Rightarrow \Lambda^2 \setminus R = \emptyset$ .

**PASSO 4:** Como  $Q = \emptyset$ , terminamos o algoritmo e retornamos  $(x^2, y^2) = (7, \frac{11}{2})$  como solução ótima local de LBLP (veja a figura deste exemplo).  $\square$

**Exemplo 6.1.5.** Considere o mesmo problema do Exemplo 6.1.4, sem a restrição de alto nível  $x \leq 7$ , e apliquemos o Algoritmo 3. O processo é o mesmo até a penúltima execução do passo 4. Continuemos a partir daí então.

**PASSO 5:** Temos  $\Lambda = \Lambda^1 \setminus Q = \{6, 7\} \setminus \{7\} = \{6\}$ . O problema

$$\begin{aligned} \min_{x,y} F(x,y) &= -x - 2y \\ \text{s.a. } x &\geq 0 \\ 5x - 2y &= 24 \quad ((6) \text{ está ativa}) \\ (1) - (5), (7), (8) \end{aligned}$$

tem como única solução ótima  $(x^2, y^2) = (8, 8)$ . Como  $F(8, 8) = -24 < -12 = F(6, 3)$ , fazemos  $k = 2$  e  $L = \emptyset$ .

**PASSO 2:** Fazemos  $Q = R = \emptyset$ .

**PASSO 3:** Temos  $\Lambda^2 = \{5, 6\}$ .

Como  $Q = \emptyset$ , escolhemos  $j = 5 \in \Lambda^2 \setminus (L \cup R) = \{5, 6\}$ . Temos  $\Lambda = \Lambda^2 \setminus R = \{5, 6\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, 2, 3, 4, 5, 7, 8\}$  e  $\Lambda \setminus \{j\} = \{6\}$ . O sistema correspondente do passo 3 já foi resolvido anteriormente, com solução  $\lambda = (0, 0, 0, 0, 0, \frac{1}{2}, 0)$ . Temos  $A = \{5\}$ . Fazemos  $L = Q = R = \{5\}$ . Temos  $\Lambda^2 \setminus R = \{6\} \neq \emptyset$ .

**PASSO 3:** Como  $Q \neq \emptyset$ , escolhemos  $j = 6 \in \Lambda^2 \setminus R = \{6\}$ . Temos  $\Lambda = \Lambda^2 \setminus R = \{6\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, \dots, 8\}$  e  $\Lambda \setminus \{j\} = \emptyset$ . Neste caso o sistema do passo 3 não possui solução. Fazemos  $R = \{5\} \cup \{6\} = \Lambda^2 \Rightarrow \Lambda^2 \setminus R = \emptyset$ .

**PASSO 4:** Temos  $Q \neq \emptyset$ .

**PASSO 5:** Temos  $\Lambda = \Lambda^2 \setminus Q = \{5, 6\} \setminus \{5\} = \{6\}$ . Neste caso o problema do passo 5 é o mesmo que o do passo 5 anterior, e a solução corrente não é melhorada.

**PASSO 2:** Fazemos  $Q = R = \emptyset$ .

**PASSO 3:** Temos  $\Lambda^2 = \{5, 6\}$ .

Como  $Q = \emptyset$ , escolhemos  $j = 6 \in \Lambda^2 \setminus (L \cup R) = \{5, 6\} \setminus \{5\} = \{6\}$ . Temos  $\Lambda = \Lambda^2 \setminus R = \{5, 6\}$ ,  $\{j\} \cup \mathbb{C}\Lambda = \{1, 2, 3, 4, 6, 7, 8\}$  e  $\Lambda \setminus \{j\} = \{5\}$ .

O sistema

$$\begin{aligned} -2\lambda_1 - \lambda_2 - \lambda_3 + 3\lambda_4 + 5\lambda_5 - 2\lambda_6 - \lambda_7 + 1 &= 0 \\ \lambda_1 = \dots = \lambda_6 = \lambda_7 &= 0, \quad \lambda_5 \geq 0 \end{aligned}$$

não tem solução (da primeira equação e das entradas nulas de  $\lambda$ , vem  $\lambda_5 = -\frac{1}{5} < 0$ ). Fazemos  $R = \{5\}$ . Temos  $\Lambda^2 \setminus R = \{6\} \neq \emptyset$ .

**PASSO 3:** Temos  $Q = \emptyset$  e  $\Lambda^2 \setminus (L \cup R) = \emptyset$ .

**PASSO 4:** Como  $Q = \emptyset$ , terminamos o algoritmo e retornamos  $(x^2, y^2) = (8, 8)$  como solução ótima local de LBLP.  $\square$

Se o problema não contiver as restrições de não negatividade  $y \geq 0$ , podemos adaptar o Algoritmo 3. Neste caso, as condições de otimalidade para o problema de baixo nível, obtidas via dualidade, exigem que as restrições duais  $B_2^t \lambda + d_2^t \geq 0$  sejam de igualdade. Podemos desconsiderar a parcela  $\{m_l + i; y_i = 0\}$  na definição de  $\Lambda^k$  e reescrever os passos 3 e 5 como

3'. (Detecção de faces racionais adjacentes) Se  $Q \neq \emptyset$ , escolha  $j \in \Lambda^k \setminus R$ .

Se  $Q = \emptyset$ , escolha se possível  $j \in \Lambda^k \setminus (L \cup R)$ . Se tal escolha não for possível, vá para o passo 4.

Verifique se o sistema

$$\begin{aligned} B_2^t \lambda + d_2^t &= 0 \\ \lambda_i &= 0, \quad i \in \{j\} \cup \mathbf{C}\Lambda \\ \lambda_i &\geq 0, \quad i \in \Lambda \setminus \{j\} \end{aligned}$$

admite solução, onde  $\Lambda = \Lambda^k \setminus R$  e  $\mathbf{C}\Lambda = \{1, \dots, m_l\} \setminus \Lambda$ . Se admite solução, digamos  $\lambda^k$ , faça

$$R \leftarrow R \cup A, \quad Q \leftarrow Q \cup A \quad \text{e} \quad L \leftarrow L \cup A$$

onde  $A = \{i \in \Lambda^k; \lambda_i^k = 0\} \subset \Lambda^k$ . Se o sistema anterior não admite solução, faça  $R \leftarrow R \cup \{j\}$ . Se  $\Lambda^k \setminus R = \emptyset$  vá para o passo 4. Se  $\Lambda^k \setminus R \neq \emptyset$  repita o passo 3'.

5'. (Passo de minimização) Resolva

$$\begin{aligned} \min_{x,y} \{c_1 x + d_1 y\} \\ \text{s.a. } A_1 x + B_1 y &\leq b_1, \quad x \geq 0 \\ (A_2 x + B_2 y - b_2)_i &= 0, \quad i \in \Lambda \\ (A_2 x + B_2 y - b_2)_i &\leq 0, \quad i \notin \Lambda \end{aligned}$$

onde  $\Lambda = \Lambda^k \setminus Q$ , e tome uma solução ótima  $(x^{k+1}, y^{k+1})$  sua (este problema admite solução pois  $(x^k, y^k)$  é um ponto viável seu). Se  $c_1 x^{k+1} + d_1 y^{k+1} < c_1 x^k + d_1 y^k$ , faça  $k \leftarrow k + 1$ ,  $L \leftarrow \emptyset$  e vá para o passo 2. Caso contrário, vá para o passo 2.

Uma outra análise interessante é quando LBLP é ilimitado. Neste caso também podemos adaptar o Algoritmo 3 modificando o passo 5 da seguinte forma:

5" (Passo de minimização) Resolva

$$\begin{aligned} & \min_{x,y} \{c_1x + d_1y\} \\ \text{s.a. } & A_1x + B_1y \leq b_1, \quad x \geq 0 \\ & (A_2x + B_2y - b_2)_i = 0, \quad i \in \Lambda \\ & (A_2x + B_2y - b_2)_i \leq 0, \quad i \notin \Lambda \\ & y_i = 0, \quad i + m_i \in \Lambda \quad (\text{se } y \geq 0 \text{ estiverem presentes em LBLP}) \\ & y_i \geq 0, \quad i + m_i \notin \Lambda \quad (\text{se } y \geq 0 \text{ estiverem presentes em LBLP}) \end{aligned}$$

onde  $\Lambda = \Lambda^k \setminus Q$ . Se o problema é ilimitado (ele é viável pois  $(x^k, y^k)$  é um ponto viável seu), pare e reporte que LBLP é ilimitado. Se o problema admite solução ótima, tome uma solução ótima  $(x^{k+1}, y^{k+1})$  sua. Se  $c_1x^{k+1} + d_1y^{k+1} < c_1x^k + d_1y^k$ , faça  $k \leftarrow k + 1$ ,  $L \leftarrow \emptyset$  e vá para o passo 2. Caso contrário, vá para o passo 2.

No entanto, lembremos que se LBLP for ilimitado, não sabemos se o problema otimista de dois níveis original dado por (2.2) (página 17) é ilimitado.

Por último, observamos que sob a hipótese de que  $\Omega$  possua vértice (garantida se  $\Omega$  for compacto), a não degeneração de todas as faces de  $\Omega$  é assegurada pela não degeneração de seus vértices somente, como mostra o resultado a seguir.

**Lema 6.1.6.** *Seja  $D = \{z \in \mathbb{R}^n; Az \leq b\}$  um poliedro que possui vértice. Então todas as faces de  $D$  são não degeneradas se, e somente se, todos os vértices de  $D$  são não degenerados.*

*Demonstração.* A implicação  $\Rightarrow$  é óbvia. Mostremos a implicação  $\Leftarrow$ ). Suponhamos que exista uma face degenerada  $Q = Q(J) = \{z \in D; A_jz = b_j, \forall j \in J\} \subset D$  ( $J$  é tal que existe  $z_0 \in Q(J)$  com  $A_jz_0 < b_j, \forall j \notin J$ ), de dimensão  $d$ . Como  $D$  possui vértice, não contém qualquer reta (Teorema A.3.1, página 121). Logo  $Q(J) \subset D$  não possui qualquer reta, isto é, possui um vértice de  $D$  (Teorema A.3.1), digamos,  $\bar{z}$ . Pela Definição 3.1.2 (página 28) tem-se  $|J| > n - d$  e daí, dentre as restrições de índices em  $J$ , há pelo menos uma que é linearmente dependente das outras. Dentre as restrições ativas em  $\bar{z}$ , estão as de índices em  $J$  pois  $\bar{z} \in Q(J)$ . Daí as restrições ativas em  $\bar{z}$  são linearmente dependentes, e logo  $\bar{z}$  é vértice degenerado de  $D$ .  $\square$

## 6.2 Obtendo $(x^0, y^0) \in \Upsilon$ quando $B_1 = 0$

Supomos nesta seção que  $B_1 = 0$  e que LBLP seja viável. Apresentaremos maneiras de obter um ponto admissível  $(x^0, y^0)$ , necessário ao passo 1 do Algoritmo 3.

No caso em que  $\Omega$  é compacto, a prova do Teorema 3.2.1 (página 41) garante que  $(x^0, y^0) \in \Upsilon$  pode ser tomado como uma solução ótima do problema

$$\min_{x,y} d_2 y \quad \text{s.a.} \quad A_1 x \leq b_1, \quad A_2 x + B_2 y \leq b_2, \quad x, y \geq 0. \quad (6.6)$$

Agora, no caso em que  $\Omega$  não é compacto, o problema (6.6) pode ser ilimitado. No entanto, isso não implica que LBLP seja ilimitado, como mostra o próximo exemplo:

**Exemplo 6.2.1.** Considere o problema bidimensional de dois níveis

$$\min_{x,y} x - y$$

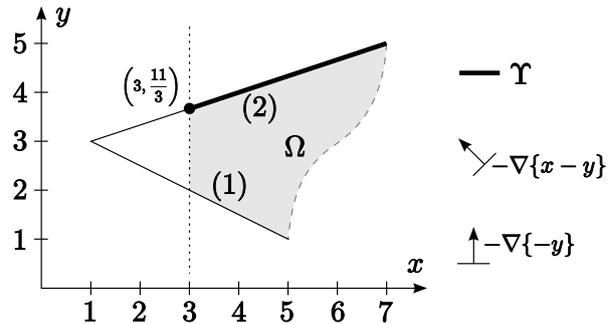
$$\text{s.a. } x \geq 3$$

$$y \in \arg \min_y -y$$

$$\text{s.a. } -x - 2y \leq -7 \quad (1)$$

$$-x + 3y \leq 8 \quad (2)$$

$$y \geq 0$$



Geometricamente, vê-se que  $S(x) \neq \emptyset$  para todo  $x \geq 3$ , que  $v(x) = -\frac{1}{3}(x+8) \rightarrow -\infty$  quando  $x \rightarrow \infty$  (e portanto o problema (6.6) correspondente é ilimitado), mas o problema de dois níveis admite solução ótima, a saber,  $(3, \frac{11}{3})$ .  $\square$

Se LBLP admite solução ótima, existe  $x$  com  $S(x) \neq \emptyset \Rightarrow \Omega_x \neq \emptyset$ . Pelo item (ii) do Lema 3.1.4 (página 30) temos que  $\Omega_x = \bar{\Omega}_x = \{x \in \mathbb{R}^{n_x}; \Omega_y(x) \neq \emptyset\}$ . Fixando  $(x^0, \bar{y}) \in \Omega$ , temos imediatamente  $x^0 \in \bar{\Omega}_x$ , isto é,  $x^0$  tal que

$$A_1 x^0 \leq b_1, \quad x^0 \geq 0 \quad \text{e existe } y \text{ tal que } B_2 y \leq b_2 - A_2 x^0, \quad y \geq 0, \quad \text{a saber, } \bar{y}.$$

Portanto, o problema

$$\min_y d_2 y \quad \text{s.a.} \quad B_2 y \leq b_2 - A_2 x^0, \quad y \geq 0 \quad (6.7)$$

tem solução ótima, digamos  $y^0$ . Imediatamente segue que  $(x^0, y^0) \in \Upsilon$ . Vale ressaltar que o método anterior é válido também no caso em que  $\Omega$  é compacto.

Agora, por [43, Lema A2], temos

$$v(x) > -\infty \text{ para algum } x \in \bar{\Omega}_x \Leftrightarrow v(x) > -\infty \text{ para todo } x \in \bar{\Omega}_x$$

(a implicação  $\Rightarrow$  foi usada na prova do item (ii) do Lema 3.1.4) e portanto se tomarmos  $(x^0, \bar{y}) \in \Omega$ , e o problema (6.7) for ilimitado, podemos concluir que LBLP é inviável, pois neste caso o problema de baixo nível é ilimitado para todo  $x \in \bar{\Omega}_x \supset \Omega_x = \emptyset$ . Formulamos então o algoritmo seguinte para o caso em que  $B_1 = 0$ .

---

**Algoritmo 4** Solução ótima local caso  $B_1 = 0$

---

**Entrada:** LBLP não ilimitado. Supõe-se todas as faces de  $\Omega$  não degeneradas e  $\Omega \neq \emptyset$ .

**Saída:** Uma solução ótima local de LBLP ou a constatação de sua inviabilidade.

Tome  $(x^0, \bar{y}) \in \Omega$  e resolva o problema viável

$$\min_y d_2 y \quad \text{s.a.} \quad B_2 y \leq b_2 - A_2 x^0, \quad y \geq 0.$$

Se o problema anterior for ilimitado, LBLP é inviável. Caso contrário, tome uma solução ótima  $y^0$  sua e execute o Algoritmo 3 usando  $(x^0, y^0)$  como ponto admissível inicial.

---

Podemos usar o Algoritmo 3 com o passo 5” no Algoritmo anterior. Neste caso, a ilimitabilidade de LBLP pode ser constatada.

## 6.3 Testes computacionais

### 6.3.1 Geração de problemas-teste

Infelizmente não existem, pelo menos de nosso conhecimento, um catálogo de problemas lineares de dois níveis disponível publicamente. A página de Leyffer [26] reúne, de vários trabalhos na literatura, *problemas com restrições de equilíbrio*, que são globalmente equivalentes à problemas de dois níveis. Dentre esses problemas, alguns são lineares, como queremos. Os problemas com restrições de equilíbrio descritos na página de Leyffer trazem, no lugar da função objetivo de baixo nível, as condições de otimalidade de Kuhn-Tucker para o problema de baixo nível, que são equivalentes ao sistema (3.1) (página 29). Apesar de ser possível derivar problemas lineares de dois níveis a partir desses problemas, não conheceríamos suas soluções ótimas locais de antemão (não estão tabeladas), e logo decidimos não utilizá-los. Ao contrário, geraremos nossos próprios problemas.

Até o momento somente os trabalhos de Calamai e Vicente [7] e Khosrow Moshirvaziri *et al.* [34] propõem métodos para geração de problemas-teste. Ao contrário do primeiro

trabalho, o segundo gera problemas com variáveis de baixo nível no alto nível, isto é,  $B_1 \neq 0$ . Se exigirmos que  $B_1 = 0$ , o método de Khosrow Moshirvaziri *et al.* perde uma característica fundamental: a não trivialidade do problema gerado. Por isso optamos por usar o trabalho de Calamai e Vicente. Este trabalho também é usado por outros autores (veja por exemplo [11]).

A seguir descreveremos como o método de Calamai e Vicente gera um problema linear de dois níveis. Primeiro, consideremos o problema bidimensional

$$\begin{aligned}
 & \min_{x_k, y_k} \{-x_k + y_k\} \\
 & \text{s.a. } y_k \in \arg \min_{y_k} -y_k \\
 & \text{s.a. } x_k + y_k \leq \rho_k \\
 & \quad -2x_k + y_k \leq 0 \\
 & \quad 1 \leq x_k \leq 3.
 \end{aligned} \tag{6.8}$$

Os autores consideram a função objetivo de alto nível como sendo  $3 - x_k + y_k$ . Retiramos o termo constante, escrevendo o problema no modelo geral (do capítulo 3). Isto não provoca modificações nas soluções.

As soluções ótimas do problema (6.8) quando  $3 \leq \rho_k \leq 9$  são facilmente calculadas. Observamos que restrições de não negatividade não estão presentes, apesar de poderem ser incluídas sem prejuízo ao problema. Escolhidos  $n_x$  e  $n_y$  o número de variáveis de alto e baixo níveis, respectivamente, para o problema que queremos gerar, consideremos também o problema

$$\begin{aligned}
 & \min_{x_k, y_k} \left\{ \sum_{m < k \leq n_x} (-x_k) + \sum_{m < k \leq n_y} y_k \right\} \\
 & \text{s.a. } y_k \in \arg \min_{y_k} 0 \\
 & \text{s.a. } y_k \geq 0, \quad m < k \leq n_y \\
 & \quad x_k \leq 3, \quad m < k \leq n_x,
 \end{aligned} \tag{6.9}$$

onde  $m = \min\{n_x, n_y\}$ . Também aqui, os autores escrevem  $3 - x_k$  no lugar de  $-x_k$ . Escolhendo valores para  $\rho_k$ , gera-se a partir dos problemas (6.8) e (6.9) um problema de

várias variáveis, separável, dado por

$$\begin{aligned} \min_{x,y} & \begin{bmatrix} -1_{n_x} & 1_{n_y} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ \text{s.a. } & y \in \arg \min_y \begin{bmatrix} 0_{n_x} & -1_m & 0_{n_y-m} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ & \text{s.a. } \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq b, \end{aligned} \quad (6.10)$$

onde  $1_r$  ( $0_r$ ) é o vetor de  $r$  1's (0's),

$$A = \begin{bmatrix} I_m & 0_{m \times (n_x - m)} \\ -2I_m & 0_{m \times (n_x - m)} \\ \frac{1}{3}I_m & 0_{m \times (n_x - m)} \\ -I_m & 0_{m \times (n_x - m)} \\ Q_x \end{bmatrix}, \quad B = \begin{bmatrix} I_m & 0_{m \times (n_y - m)} \\ I_m & 0_{m \times (n_y - m)} \\ 0_{m \times n_y} \\ 0_{m \times n_y} \\ Q_y \end{bmatrix}, \quad b = \begin{bmatrix} \rho \\ 0_m \\ 1_m \\ -1_m \\ e \end{bmatrix}$$

e  $0_{r \times t}$  é a matriz nula de dimensão  $r \times t$ .

Seja  $\alpha = \max\{n_x, n_y\} - m$ . Se  $\alpha > 0$  (isto é, se  $n_x \neq n_y$ ), as matrizes  $Q_x$ ,  $Q_y$  e  $e$  são dadas por

$$Q_x = \begin{cases} 0_{\alpha \times n_x} & \text{se } n_x < n_y \\ \begin{bmatrix} 0_{\alpha \times m} & \frac{1}{3}I_\alpha \end{bmatrix} & \text{se } n_x > n_y \end{cases}, \quad Q_y = \begin{cases} \begin{bmatrix} 0_{\alpha \times m} & -I_\alpha \end{bmatrix} & \text{se } n_x < n_y \\ 0_{\alpha \times n_y} & \text{se } n_x > n_y \end{cases} e \\ e = \begin{cases} 1_\alpha & \text{se } n_x > n_y \\ 0_\alpha & \text{se } n_x < n_y \end{cases}.$$

Se  $\alpha = 0$ , essas matrizes não estão definidas. O vetor  $\rho \in \mathbb{R}^m$  é construído da seguinte forma: escolhemos inteiros  $m_i \geq 0$ ,  $i \in \{1, \dots, 5\}$ , tais que

$$\sum_{i=1}^5 m_i = m.$$

Daí escolhemos

$$\begin{aligned} \rho_i &= 3, & 1 \leq i \leq m_1, \\ \rho_i &= 7, & m_1 < i \leq m_1 + m_2, \\ \rho_i &= 9, & m_1 + m_2 < i \leq m_1 + m_2 + m_3, \\ 3 < \rho_i < 7, & m_1 + m_2 + m_3 < i \leq m_1 + \dots + m_4, \\ 7 < \rho_i < 9, & m_1 + \dots + m_4 < i \leq m. \end{aligned}$$

Calamai e Vicente reportam que para  $1 \leq i \leq m_1$ , as restrições  $x_i \geq 1$  e  $-2x_i + y_i \leq 0$  são redundantes. Também para  $m_1 + m_2 < i \leq m_1 + m_2 + m_3$ , as restrições  $x_i \leq 3$  e  $x_i + y_i \leq 9$  são redundantes. Portanto são  $2(m_1 + m_3)$  restrições redundantes, que devem ser eliminadas nas matrizes  $\begin{bmatrix} A & B \end{bmatrix}$  e  $b$  para evitar degeneração.

O problema (6.10) possui as seguintes propriedades:

- possui  $2^{m_2}$  soluções ótimas globais;
- possui  $2^{m_2+m_4+m_5} - 2^{m_2}$  soluções ótimas locais/não globais. Logo, possui soluções ótimas locais/não globais se, e somente se,  $m_4 + m_5 > 0$ ;
- o valor ótimo é dado por

$$-3m_1 + m_2 + m_3 + m_5 + \sum_{i=m_1+m_2+m_3+1}^{m_1+\dots+m_4} (\rho_i - 6) - 3(n_x - m)$$

(este valor difere ligeiramente do trabalho de Calamai e Vicente dado o descarte dos termos constantes nas funções objetivo de alto nível);

- o problema de baixo nível retorna uma única solução ótima para cada  $x$  fixado no alto nível.

A fim de disfarçar a separabilidade do problema (6.10), fazemos uma mudança de variáveis

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = M^{-1} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (6.11)$$

Seguindo a proposta de Calamai e Vicente, tomaremos  $M$  como sendo  $M = HDH$ , onde

$$H = \begin{bmatrix} I_{n_x} - 2v_x v_x^t & 0 \\ 0 & I_{n_y} - 2v_y v_y^t \end{bmatrix},$$

$v_x \in \mathbb{R}^{n_x}$  e  $v_y \in \mathbb{R}^{n_y}$  são esparsos,  $v_x^t v_x = v_y^t v_y = 1$ ,  $D = \text{diag}(d_1, \dots, d_{n_x+n_y})$  com  $d_1, \dots, d_{n_x+n_y} > 0$  e o número de condicionamento da norma euclidiana da matriz  $D$  igual a  $10^\delta$  para  $\delta > 0$  dado, isto é,

$$\frac{\max_i d_i}{\min_i d_i} = 10^\delta$$

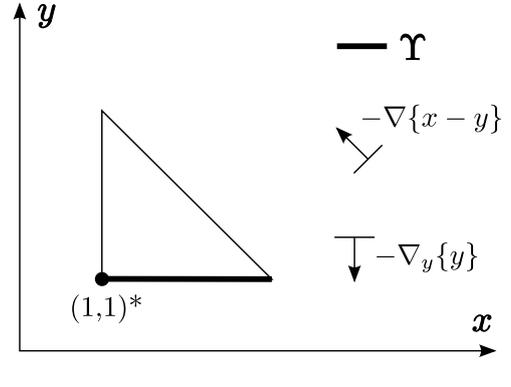
(a norma euclidiana  $\|A\|_2$  de uma matriz real  $A$  é a raiz do maior autovalor da matriz semidefinida positiva  $A^t A$  [28, Seção 15 do capítulo I], e o número de condicionamento de  $A$  em relação a essa norma é dado por  $\|A\|_2 \|A^{-1}\|_2$ ).

Assim, obtemos o problema

$$\begin{aligned} \min_{x', y'} & \left( \begin{bmatrix} -1_{n_x} & 1_{n_y} \end{bmatrix} M \right) \begin{bmatrix} x' \\ y' \end{bmatrix} \\ \text{s.a. } & y' \in \arg \min_{y'} \left( \begin{bmatrix} 0_{n_x} & -1_m & 0_{n_y-m} \end{bmatrix} M \right) \begin{bmatrix} x' \\ y' \end{bmatrix} \\ & \text{s.a. } \left( \begin{bmatrix} A & B \end{bmatrix} M \right) \begin{bmatrix} x' \\ y' \end{bmatrix} \leq b. \end{aligned} \quad (6.12)$$

Calamai e Vicente mostram que, para esta matriz  $M$ , as soluções ótimas locais/globais dos problemas (6.10) e (6.12) estão em correspondência biunívoca mediante (6.11) (é fácil mostrar que  $M^{-1} = HD^{-1}H$ ). De fato, a matriz  $M$  em questão satisfaz as hipóteses do Teorema 6.3.3 (página 85). Vale notarmos que, diferentemente dos problemas de programação linear, esta equivalência não é preservada para quaisquer matrizes inversíveis, como mostra o exemplo a seguir.

**Exemplo 6.3.1.** Consideremos o problema de dois níveis

$$\begin{aligned} \min_{x,y} & \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ \text{s.a. } & y \in \arg \min_y \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ & \text{s.a. } \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}. \end{aligned}$$


Geometricamente, vemos que a região admissível é  $\Upsilon = \{(x, y) \in \mathbb{R}^2; y = 1, 1 \leq x \leq 2\}$  e o valor ótimo  $F^* = 1 - 1 = 0$ . Considere agora matriz inversível

$$R = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

que corresponde à rotação em torno da origem por um ângulo de 45 graus no sentido anti-horário. O problema obtido pela mudança

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

é dado por

$$\begin{aligned}
& \min_{x,y} \begin{bmatrix} \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
& \text{s.a. } y \in \arg \min_y \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
& \text{s.a. } \begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix},
\end{aligned}$$

donde vemos geometricamente que a região admissível não corresponde à original, e cujo valor ótimo é  $-1 \neq 0$ .  $\square$

Para efeito prático,  $v$  será calculado aleatoriamente usando um parâmetro  $\tau \in [0; 0, 6]$  para a percentagem de zeros. Escolheremos  $\delta > 0$  inteiro e  $d_i > 0$ ,  $i \in \{1, \dots, n_x + n_y\}$ , da seguinte forma: escolhemos um índice  $\alpha \in \{1, \dots, n_x + n_y\}$  e  $d_\alpha$  aleatoriamente entre  $10^\delta/2$  e  $10^\delta$ . Daí escolhemos  $\beta \in \{1, \dots, n_x + n_y\} \setminus \{\alpha\}$  e fazemos  $d_\beta = d_\alpha/10^\delta$ . Então, como  $\delta \geq 1$ , temos

$$\frac{10^\delta}{2} \leq d_\alpha \leq 10^\delta \Rightarrow \frac{1}{2} \leq d_\beta \leq 1 < \frac{10^\delta}{2} \leq d_\alpha \Rightarrow d_\beta < d_\alpha.$$

Finalmente, escolhemos os outros  $d_i$ 's aleatoriamente em  $[d_\beta, d_\alpha]$ . Assim

$$\frac{\max_i d_i}{\min_i d_i} = \frac{d_\alpha}{d_\beta} = \frac{d_\alpha}{d_\alpha/10^\delta} = 10^\delta$$

e logo vale a relação de condicionamento da norma euclidiana.

Sendo  $d_x = (d_1, \dots, d_{n_x})$  e  $d_y = (d_{n_x+1}, \dots, d_{n_x+n_y})$ , é possível mostrar que

$$M = \begin{bmatrix} (m_{ij}^x)_{n_x \times n_x} & 0 \\ 0 & (m_{ij}^y)_{n_y \times n_y} \end{bmatrix}$$

com

$$m_{ij}^z = \frac{\delta_{ij}}{d_{zi}} + 2v_{zi}v_{zj} \left( \bar{v}_z - \frac{1}{d_{zi}} - \frac{1}{d_{zj}} \right),$$

onde  $\bar{v}_z = \sum_{k=1}^{n_z} (2v_{zk}^2/d_{zk})$  e  $\delta_{ij}$  é o delta de Kronecker, para  $z = x, y$ . Com isto fica fácil calcular o produto das matrizes do problema por  $M$ .

Observamos que em geral, após a mudança de variáveis, as restrições de não negatividade são não satisfeitas pelas soluções, isto é, se inseridas podem modificar o problema. Por isso consideramos a versão do nosso algoritmo em que essas restrições não são pre-

sentes (veja seção 6.1).

O gerador foi implementado em linguagem C. Utilizamos o método *minimal standard* (versão 3) desenvolvido por Park e Miller [35] para o cálculo de números pseudo-aleatórios. O código para este método foi adaptado do relatório técnico de Saucier [38].

### 6.3.2 Generalização do método de Calamai e Vicente para geração de problemas-teste

Apresentaremos agora uma generalização natural do método de Calamai e Vicente para obtenção de problemas-teste em dois níveis. Vale ressaltar que os modelos aqui obtidos são do tipo BLP, e nem sempre correspondem à descrição do problema otimista de dois níveis dada no Capítulo 2 (veja o Exemplo 2.0.4, página 22).

Consideramos o problema parametrizado em  $k$  dado por

$$\begin{aligned} P(k): \quad & \min_{x_k, y_k} F_k(x_k, y_k) \\ & \text{s.a. } g_k(x_k, y_k) \leq 0 \\ & y_k \in \arg \min_{y_k} f_k(x_k, y_k) \\ & \text{s.a. } h_k(x_k, y_k) \leq 0, \end{aligned}$$

onde  $x_k \in \mathbb{R}^{n_{xk}}$ ,  $y_k \in \mathbb{R}^{n_{yk}}$ ,  $F_k, f_k : \mathbb{R}^{n_{xk}+n_{yk}} \rightarrow \mathbb{R}$ ,  $g_k : \mathbb{R}^{n_{xk}+n_{yk}} \rightarrow \mathbb{R}^{m_{uk}}$  e  $h_k : \mathbb{R}^{n_{xk}+n_{yk}} \rightarrow \mathbb{R}^{m_{lk}}$ . Fixando  $k_1, \dots, k_K$  e renumerando os índices, construímos o problema separável

$$\begin{aligned} \min_{x, y} F(x, y) &= \sum_{k=1}^K F_k(x_k, y_k) \\ \text{s.a. } g(x, y) &\leq 0 \\ y_k &\in \arg \min_{y_k} f_k(x_k, y_k) \\ \text{s.a. } h_k(x_k, y_k) &\leq 0, \quad k = 1, \dots, K, \end{aligned}$$

onde  $(x, y) = (x_1, \dots, x_K, y_1, \dots, y_K)$  e  $g(x, y) = (g_1(x_1, y_1), \dots, g_K(x_K, y_K))$ . Os problemas de baixo nível são independentes entre si, e portanto podemos escrever este problema

como

$$\begin{aligned}
 P: \quad & \min_{x,y} F(x,y) \\
 \text{s.a.} \quad & g(x,y) \leq 0 \\
 & y \in \arg \min_y f(x,y) = \sum_{k=1}^K f_k(x_k, y_k) \\
 & \text{s.a. } h(x,y) \leq 0,
 \end{aligned}$$

onde  $h(x,y) = (h_1(x_1, y_1), \dots, h_K(x_K, y_K))$ . Temos o seguinte resultado:

**Teorema 6.3.2.** *Fixados  $k_1, \dots, k_K$  temos que*

- i)  $(x,y)$  é solução ótima global de  $P$  se, e somente se,  $(x_{k_i}, y_{k_i})$  é solução ótima global de  $P(k_i)$  para todo  $i$ ;*
- ii)  $(x,y)$  é solução ótima local/não global de  $P$  se, e somente se,  $(x_{k_i}, y_{k_i})$  é solução ótima local ou global de  $P(k_i)$  para todo  $i$ , onde pelo menos um  $(x_{k_j}, y_{k_j})$  é solução ótima local/não global de  $P(k_j)$ ;*
- iii)  $P$  é inviável (respectivamente ilimitado) se, e somente se,  $P(k_i)$  é inviável (respectivamente ilimitado) para algum  $i$ ;*
- iv) Se  $n_i < \infty$  é o número de soluções ótimas globais de  $P(k_i)$  então  $P$  tem  $n_1 n_2 \cdots n_K$  soluções ótimas globais;*
- v) Se  $n_i < \infty$  e  $m_i < \infty$  são o número de soluções ótimas globais e de soluções ótimas locais/não globais de  $P(k_i)$ , respectivamente, então  $P$  tem*

$$\prod_{k=1}^K (m_k + n_k) - \prod_{k=1}^K n_k$$

*soluções ótimas locais/não globais;*

- vi) O problema do baixo nível de  $P$  retorna uma única solução ótima para cada  $x$  se, e somente se, para todo  $i \in \{1, \dots, K\}$ , o problema do baixo nível de  $P(k_i)$  retorna uma única solução ótima para cada  $x_{k_i}$ .*

*Demonstração.* Para simplificar, renumeramos os índices  $k_1, \dots, k_K$  como  $1, \dots, K$ .

*i) Mostremos a implicação  $\Rightarrow$ .* Seja  $(\bar{x}, \bar{y})$  solução ótima global de  $P$ . Então  $F(\bar{x}, \bar{y}) \leq F(x,y), \forall (x,y) \in \Upsilon$ , onde  $\Upsilon$  é a região admissível do problema  $P$ . Suponhamos por

contradição que  $(\bar{x}_i, \bar{y}_i)$  não seja solução ótima de  $P(i)$ . Então existem  $\delta > 0$  e  $(\tilde{x}_i, \tilde{y}_i) \in B_\delta^\infty(\bar{x}_i, \bar{y}_i) \cap \Upsilon_i$  com  $F_i(\tilde{x}_i, \tilde{y}_i) < F_i(\bar{x}_i, \bar{y}_i)$ , onde  $\Upsilon_i$  é a região admissível de  $P(i)$ . Daí

$$\begin{aligned} F(\bar{x}_1, \dots, \tilde{x}_i, \dots, \bar{x}_K, \bar{y}_1, \dots, \tilde{y}_i, \dots, \bar{y}_K) &= F_i(\tilde{x}_i, \tilde{y}_i) + \sum_{k=1, k \neq i}^K F_k(\bar{x}_k, \bar{y}_k) \\ &< \sum_{k=1}^K F_k(x_k, y_k) = F(\bar{x}, \bar{y}) \end{aligned}$$

com

$$\begin{aligned} (\bar{x}_1, \dots, \tilde{x}_i, \dots, \bar{x}_K, \bar{y}_1, \dots, \tilde{y}_i, \dots, \bar{y}_K) &\in \\ (B_\delta^\infty(\bar{x}_1) \times \dots \times B_\delta^\infty(\bar{x}_K) \times B_\delta^\infty(\bar{y}_1) \times \dots \times B_\delta^\infty(\bar{y}_K)) &\cap \\ \{(x, y); (x_i, y_i) \in \Upsilon_i, \forall i\} &= B_\delta^\infty(\bar{x}, \bar{y}) \cap \Upsilon, \end{aligned}$$

uma contradição. Portanto vale a implicação  $\Rightarrow$ ). A implicação  $\Leftarrow$ ) é feita de forma análoga.

*ii)* A prova é análoga à do item *(i)*.

*iii)*  $P$  inviável  $\Leftrightarrow \Upsilon = \{(x, y); (x_i, y_i) \in \Upsilon_i, \forall i\} = \emptyset \Leftrightarrow$  para algum  $i$ ,  $\Upsilon_i = \emptyset$ , isto é,  $P(i)$  inviável.

$P$  ilimitado  $\Leftrightarrow$  existe uma sequência  $\{(x^n, y^n)\} \subset \Upsilon$  com  $\lim F(x^n, y^n) \rightarrow -\infty \Leftrightarrow$  para algum  $i$ , a sequência  $\{(x_i^n, y_i^n)\} \subset \Upsilon_i$  é tal que  $\lim F_i(x_i^n, y_i^n) \rightarrow -\infty$ . Assim, vale a afirmação sobre a ilimitabilidade.

*iv)* Pelo item *(i)*, cada solução ótima global de  $P$  é formada pela justaposição de soluções ótimas globais dos  $P(k)$ 's, como na construção do vetor  $(x, y)$ . Com isto, basta notarmos que existem  $n_1 n_2 \dots n_K$  possibilidades de justaposição das soluções ótimas globais dos  $P(k)$ 's.

*v)* O número de soluções ótimas locais/não globais de  $P$  é o seu número de soluções ótimas (globais ou locais) menos o seu número de soluções ótimas globais. O resultado segue então dos itens *(ii)* e *(iv)*.

*vi)* Mostremos a implicação  $\Rightarrow$ ). Se o problema do baixo nível de  $P$  retorna uma única solução ótima para cada  $x$ , fica bem definida uma função  $y = y(x)$ . Agora, o problema do baixo nível de um  $P(i)$  qualquer é parametrizado somente por  $x_i$ . Logo podemos escrever

$$y = y(x) = (y_1(x_1), \dots, y_K(x_K)). \quad (6.13)$$

Portanto ficam definidas funções  $y_i = y_i(x_i)$ ,  $i = 1, \dots, K$ , ou seja, o problema do baixo

nível de cada  $P(i)$  retorna uma única solução ótima para cada  $x_i$ , e provamos  $\Rightarrow$ ).

Mostremos agora a implicação  $\Leftarrow$ ). Se para cada  $i$  o problema do baixo nível de  $P(i)$  retorna uma única solução ótima para todo  $x_i$ , ficam bem definidas funções  $y_i(x_i)$ ,  $i = 1, \dots, K$ . Então naturalmente temos (6.13), e o problema do baixo nível de  $P$  retorna uma única solução ótima para cada  $x$ .  $\square$

Obtido o problema separável  $P$ , prosseguimos de forma análoga à proposta de Calamai e Vicente. A fim de disfarçar a separabilidade de  $P$ , fazemos uma mudança de variáveis

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = M^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \quad (6.14)$$

onde  $M \in \mathbb{R}^{r \times r}$ ,  $r = \sum_{k=1}^K (n_{xk} + n_{yk})$ . Assim obtemos o problema

$$\begin{aligned} & \min_{x', y'} F^M(x', y') \\ & \text{s.a. } g^M(x', y') \leq 0 \\ & \quad y' \in \arg \min_{y'} f^M(x', y') \\ & \quad \text{s.a. } h^M(x', y') \leq 0, \end{aligned}$$

onde  $F^M(x', y') = F\left(M^{-1} \begin{bmatrix} x \\ y \end{bmatrix}\right)$  e  $f^M, g^M, h^M$  análogos. Como visto no Exemplo 6.3.1 (página 80), nem toda matriz inversível  $M$  deriva problemas equivalentes aos seus originais. No entanto, o resultado a seguir é válido.

**Teorema 6.3.3.** *Consideremos o problema de dois níveis*

$$\begin{aligned} & \min_{x, y} F(x, y) \\ & \text{s.a. } g(x, y) \leq 0 \\ & \quad y \in \arg \min_y dy \\ & \quad \text{s.a. } By \leq b - Ax, \end{aligned} \quad (6.15)$$

onde  $x \in \mathbb{R}^{n_x}$  e  $y \in \mathbb{R}^{n_y}$ . Para matrizes inversíveis  $M \in \mathbb{R}^{(n_x+n_y) \times (n_x+n_y)}$  do tipo

$$\begin{bmatrix} M_x & 0 \\ 0 & M_y \end{bmatrix}$$

com  $M_x \in \mathbb{R}^{n_x \times n_x}$  e  $M_y \in \mathbb{R}^{n_y \times n_y}$ , o problema

$$\begin{aligned} & \min_{x', y'} F^M(x', y') \\ & \text{s.a. } g^M(x', y') \leq 0 \\ & \quad y' \in \arg \min_{y'} (dM_y)y' \\ & \quad \text{s.a. } (BM_y)y' \leq b - (AM_x)x', \end{aligned}$$

onde  $F^M(x', y') = F\left(M^{-1} \begin{bmatrix} x \\ y \end{bmatrix}\right)$  e  $g^M$  análogo, é equivalente (global e localmente) ao problema (6.15).

*Demonstração.* Sendo

$$\begin{bmatrix} x \\ y \end{bmatrix} = M \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} M_x x' \\ M_y y' \end{bmatrix},$$

basta mostrarmos que  $y \in S(x) \Leftrightarrow y' \in S^M(x')$ , onde  $S^M(x') = \arg \min_{\bar{y}'} \{(dM_y)\bar{y}'; (BM_y)\bar{y}' \leq b - (AM_x)x'\}$ . De fato,

$$\begin{aligned} y \in S(x) & \Leftrightarrow dy \leq d\bar{y}, \quad \forall \bar{y} \text{ tal que } B\bar{y} \leq b - Ax \\ & \Leftrightarrow d(M_y y') \leq d(M_y \bar{y}'), \quad \forall \bar{y}' \text{ tal que } B(M_y \bar{y}') \leq b - A(M_x x') \\ & \Leftrightarrow (dM_y)y' \leq (dM_y)\bar{y}', \quad \forall \bar{y}' \text{ tal que } (BM_y)\bar{y}' \leq b - (AM_x)x' \\ & \Leftrightarrow y' \in S^M(x'). \end{aligned}$$

□

Suponha agora que as restrições dos  $P(k)$ 's sejam lineares, isto é,

$$g_k(x_k, y_k) = \begin{bmatrix} A_1^k & B_1^k \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} - b_1^k \quad \text{e} \quad h_k(x_k, y_k) = \begin{bmatrix} A_2^k & B_2^k \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} - b_2^k.$$

Então as restrições do problema separável  $P$  são

$$g(x, y) = \begin{bmatrix} A_1^1 & & B_1^1 & & \\ & \ddots & & \ddots & \\ & & A_1^K & & B_1^K \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} b_1^1 \\ \vdots \\ b_1^K \end{bmatrix}$$

e

$$h(x, y) = \begin{bmatrix} A_2^1 & & B_2^1 & & \\ & \ddots & & \ddots & \\ & & A_2^K & & B_2^K \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} b_2^1 \\ \vdots \\ b_2^K \end{bmatrix}$$

onde os espaços vazios representam zeros. Aplicando a mudança de variáveis (6.14) obtemos

$$g^M(x, y) = \left( \begin{bmatrix} A_1^1 & & B_1^1 & & \\ & \ddots & & \ddots & \\ & & A_1^K & & B_1^K \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} b_1^1 \\ \vdots \\ b_1^K \end{bmatrix}$$

e

$$h^M(x, y) = \left( \begin{bmatrix} A_2^1 & & B_2^1 & & \\ & \ddots & & \ddots & \\ & & A_2^K & & B_2^K \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} b_2^1 \\ \vdots \\ b_2^K \end{bmatrix}.$$

Com a generalização proposta, derivamos problemas-teste em dois níveis fixando  $k_1, \dots, k_K$ , definindo os problemas iniciais  $P(k_i)$ ,  $i = 1, \dots, K$ , (do qual devemos conhecer suas soluções ótimas analíticas) e a matriz  $M$  (como no Teorema 6.3.3). As propriedades do problema obtido dependem fortemente dos problemas iniciais envolvidos, cujas relações são dadas pelo Teorema 6.3.2.

Calamai e Vicente, sabendo o número de variáveis no alto e baixo níveis ( $n_x$  e  $n_y$  respectivamente), tomam  $m = \min\{n_x, n_y\}$  problemas da forma (6.8) e um outro da forma (6.9) (subseção 6.3.1, página 77). Além dessa escolha, trabalharemos com outras duas, descritas a seguir, que por serem lineares equivalem ao modelo do problema otimista de dois níveis dado no Capítulo 2 (Teoremas 2.0.3, página 21, e 3.1.8, página 36).

**GRUPO 1 DE PROBLEMAS:** Considere o problema parametrizado em  $x_k$  dado por

$$\begin{aligned} p(x_k): \quad & \min_{y_{1k}, y_{2k}} \{y_{1k} + y_{2k}\} \\ & \text{s.a.} \quad -x_k + y_{1k} - 2y_{2k} \leq -7 \\ & \quad \quad -3x_k - 2y_{1k} - y_{2k} \leq -6 \\ & \quad \quad -y_{1k} - y_{2k} \leq \rho_k, \end{aligned}$$

onde  $\rho_k \in [-11, 5]$ . Resolvendo-o geometricamente (veja figura 8), encontramos que

- para  $x_k \leq \frac{\rho_k + 5}{2}$ ,  $p(x_k)$  tem solução ótima única, a saber,  $(1 - x_k, 4 - x_k)$ ;
- para  $x_k > \frac{\rho_k + 5}{2}$ ,  $p(x_k)$  tem infinitas soluções ótimas, a saber, o conjunto limitado

$$U_k(x_k) = \left\{ \begin{array}{l} (y_{1k}, y_{2k}); \quad y_{1k} - 2y_{2k} \leq -7 + x_k, \\ \quad \quad \quad -2y_{1k} - y_{2k} \leq -6 + 3x_k \\ \quad \quad \quad -y_{1k} - y_{2k} = \rho_k \end{array} \right\}.$$

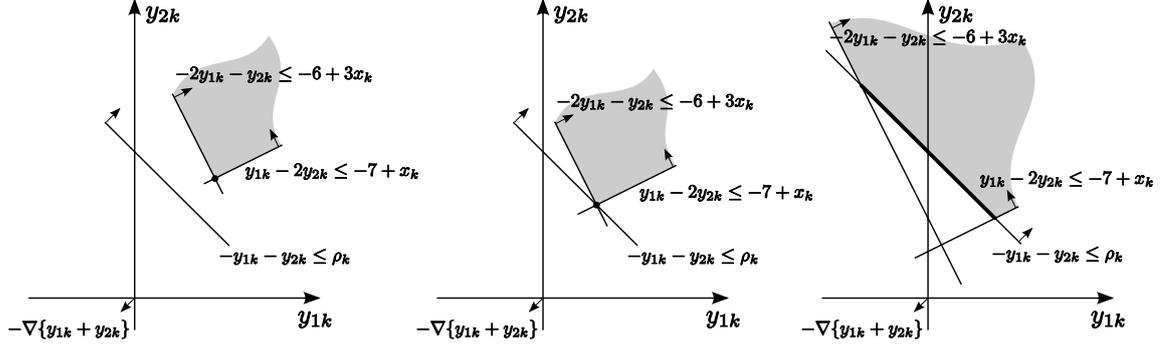


Figura 8: Resolução do problema  $p(x_k)$ . À esquerda, geometria para  $x_k < \frac{\rho_k + 5}{2}$ . Ao centro,  $x_k = \frac{\rho_k + 5}{2}$ . À direita,  $x_k > \frac{\rho_k + 5}{2}$ . As linhas em destaque são as soluções ótimas.

Considere agora o problema de dois níveis dado por

$$\begin{aligned}
 P(k): \quad & \min_{x_k, y_{1k}, y_{2k}} F(x_k, y_{1k}, y_{2k}) = -x_k + y_{1k} - 2y_{2k} \\
 & \text{s.a. } (y_{1k}, y_{2k}) \in \arg \min_{y_{1k}, y_{2k}} \{y_{1k} + y_{2k}\} \\
 & \text{s.a. } -x_k + y_{1k} - 2y_{2k} \leq -7 \\
 & \quad -3x_k - 2y_{1k} - y_{2k} \leq -6 \\
 & \quad -y_{1k} - y_{2k} \leq \rho_k \\
 & \quad -5 \leq x_k \leq 5.
 \end{aligned}$$

Para cada  $k$ , temos  $g_k \equiv 0$ ,

$$A_2^k = A = \begin{bmatrix} -1 \\ -3 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \quad B_2^k = B = \begin{bmatrix} 1 & -2 \\ -2 & -1 \\ -1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad b_2^k = b^k = \begin{bmatrix} -7 \\ -6 \\ \rho_k \\ 5 \\ 5 \end{bmatrix}. \quad (6.16)$$

O problema escravo de  $P(k)$  é  $p(x_k)$  com as restrições adicionais  $-5 \leq x_k \leq 5$ . Essas restrições não influenciam nas soluções ótimas do baixo nível, que são as mesmas de  $p(x_k)$ , pois só possuem variáveis de alto nível (Teorema 3.1.12, página 41). Portanto, como

$$-11 \leq \rho_k \leq 5 \Leftrightarrow -3 \leq \frac{\rho_k + 5}{2} \leq 5,$$

$P(k)$  pode ser resolvido separadamente pelos problemas

$$P_1(k): \min_{x_k, y_{1k}, y_{2k}} F(x_k, y_{1k}, y_{2k}) = -x_k + y_{1k} - 2y_{2k}$$

$$\text{s.a. } (y_{1k}, y_{2k}) = (1 - x_k, 4 - x_k)$$

$$-5 \leq x_k \leq \frac{\rho_k + 5}{2}$$

e

$$P_2(k): \min_{x_k, y_{1k}, y_{2k}} F(x_k, y_{1k}, y_{2k}) = -x_k + y_{1k} - 2y_{2k}$$

$$\text{s.a. } y_{1k} - 2y_{2k} \leq -7 + x_k,$$

$$-2y_{1k} - y_{2k} \leq -6 + 3x_k$$

$$-y_{1k} - y_{2k} = \rho_k$$

$$\frac{\rho_k + 5}{2} < x_k \leq 5,$$

ou seja,

$$P(k): \min \{F(x_k, y_{1k}, y_{2k}); (x_k, y_{1k}, y_{2k}) \in \arg \min P_1(k) \cup \arg \min P_2(k)\}.$$

O problema  $P_1(k)$  tem função objetivo constante igual a  $-7$ , e suas soluções ótimas são  $(x_k, 1 - x_k, 4 - x_k)$  onde  $-5 \leq x_k \leq \frac{\rho_k + 5}{2}$ . O problema  $P_2(k)$  tem solução ótima única  $(5, \rho_k - 9, -2\rho_k + 9)$  com valor ótimo  $5\rho_k - 32 < -7$ . Portanto, se  $\rho_k$  for tal que  $P_2(k)$  é viável, isto é, se  $\frac{\rho_k + 5}{2} < 5$  ( $\Leftrightarrow \rho_k < 5$ ), a solução ótima de  $P(k)$  é a solução ótima de  $P_2(k)$ . Caso  $\rho_k = 5$ , as soluções ótimas de  $P(k)$  são as soluções ótimas de  $P_1(k)$ . Mais ainda, neste caso a restrição  $-y_{1k} - y_{2k} \leq 5$  em  $P(k)$  torna-se redundante, e deve ser eliminada para evitar degeneração.

Finalmente, o problema gerado é dado por

$$P: \min_{x, y} \left( \begin{bmatrix} -1_K & d \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\text{s.a. } y \in \arg \min_y \left( \begin{bmatrix} 0_K & 1_{2K} \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\text{s.a. } \left( \begin{bmatrix} A & & B & & \\ & \ddots & & \ddots & \\ & & A & & B \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} b^1 \\ \vdots \\ b^K \end{bmatrix},$$

onde  $d = \begin{bmatrix} 1 & -2 & \dots & 1 & -2 \end{bmatrix}$  tem dimensão  $1 \times 2K$ , e onde devemos eliminar as

restrições  $-y_{1k} - y_{2k} \leq 5$ .

Sendo  $I_{<} = \{k; \rho_k < 5\}$ , o problema  $P$  possui as seguintes propriedades, asseguradas pelo Teorema 6.3.2:

- Se  $|I_{<}| = 0$ ,  $P$  possui infinitas soluções ótimas globais, a saber, as soluções formadas pela justaposição das soluções ótimas de  $P_1(k)$ ,  $k = 1, \dots, K$ . Neste caso,  $P$  não tem soluções ótimas locais/não globais;
- Se  $0 < |I_{<}| < K$ ,  $P$  tem infinitas soluções ótimas globais, a saber, as soluções formadas pela justaposição de soluções ótimas de  $P_2(k)$ ,  $k \notin I_{<}$ , com soluções ótimas de  $P_1(k)$ ,  $k \in I_{<}$ . Neste caso,  $P$  possui infinitas soluções ótimas locais/não globais, a saber, as soluções cuja justaposição tenha pelo menos uma solução ótima de algum  $P_1(k)$ ,  $k \notin I_{<}$ , tal que  $x_k \in [-5, \frac{\rho_k+5}{2}]$ ;
- Se  $|I_{<}| = K$ ,  $P$  possui solução ótima global única, a saber, a solução formada pela justaposição das soluções ótimas de  $P_2(k)$ ,  $k = 1, \dots, K$ . Neste caso,  $P$  possui infinitas soluções ótimas locais/não globais, a saber, as soluções cuja justaposição tenha pelo menos uma solução ótima de algum  $P_1(k)$  tal que  $x_k \in [-5, \frac{\rho_k+5}{2}]$ ;
- O valor ótimo é  $F^* = 5 \sum_{k=1}^K \rho_k - 32K$ ;
- O problema do baixo nível não retorna uma única solução ótima para todo  $x$  se, e somente se,  $|I_{<}| > 0$ .

**GRUPO 2 DE PROBLEMAS:** Geraremos este grupo de problemas-teste modificando ligeiramente o GRUPO 1. Consideremos o mesmo problema  $p(x_k)$  anterior, com  $\rho_k \in [-11, 5]$ , e o problema de dois níveis dado por

$$\begin{aligned}
 Q(k): \quad & \min_{x_k, y_{1k}, y_{2k}} F(x_k, y_{1k}, y_{2k}) = -2x_k - y_{1k} - 2y_{2k} \\
 & \text{s.a. } (y_{1k}, y_{2k}) \in \arg \min_{y_{1k}, y_{2k}} \{y_{1k} + y_{2k}\} \\
 & \text{s.a. } -x_k + y_{1k} - 2y_{2k} \leq -7 \\
 & \quad -3x_k - 2y_{1k} - y_{2k} \leq -6 \\
 & \quad -y_{1k} - y_{2k} \leq \rho_k \\
 & \quad -7 \leq x_k \leq 5.
 \end{aligned}$$

Da mesma forma,  $Q(k)$  pode ser resolvido separadamente pelos problemas

$$Q_1(k): \min_{x_k, y_{1k}, y_{2k}} F(x_k, y_{1k}, y_{2k}) = -2x_k - y_{1k} - 2y_{2k}$$

$$\text{s.a. } (y_{1k}, y_{2k}) = (1 - x_k, 4 - x_k)$$

$$-7 \leq x_k \leq \frac{\rho_k + 5}{2}$$

e

$$Q_2(k): \min_{x_k, y_{1k}, y_{2k}} F(x_k, y_{1k}, y_{2k}) = -2x_k - y_{1k} - 2y_{2k}$$

$$\text{s.a. } y_{1k} - 2y_{2k} \leq -7 + x_k,$$

$$-2y_{1k} - y_{2k} \leq -6 + 3x_k$$

$$-y_{1k} - y_{2k} = \rho_k$$

$$\frac{\rho_k + 5}{2} < x_k \leq 5.$$

O problema  $Q_1(k)$  tem função objetivo  $F(x_k, y_{1k}, y_{2k}) = x_k - 9$ . Portanto possui solução ótima única  $(-7, 8, 11)$  e valor ótimo  $-16$ . Notemos ainda que  $Q_1(k)$  não possui soluções ótimas locais/não globais. O problema  $Q_2(k)$  tem solução ótima única  $(5, \rho_k - 9, -2\rho_k + 9)$  e valor ótimo  $3\rho_k - 19$ . Também,  $Q_2(k)$  não possui soluções ótimas locais/não globais. Como  $3\rho_k - 19 < -16 \Leftrightarrow \rho_k < 1$  e  $3\rho_k - 19 = -16 \Leftrightarrow \rho_k = 1$ , temos que

- se  $\rho_k < 1$  então  $Q(k)$  possui solução ótima única, a saber, a solução ótima de  $Q_2(k)$ , e solução ótima local/não global única, a saber, a solução ótima de  $Q_1(k)$ ;
- se  $\rho_k = 1$  então  $Q(k)$  possui duas soluções ótimas, a saber, as soluções ótimas de  $Q_1(k)$  e  $Q_2(k)$ , e nenhuma solução ótima local/não global;
- se  $\rho_k > 1$  então  $Q(k)$  possui solução ótima única, a saber, a solução ótima de  $Q_1(k)$ , e solução ótima local/não global única, a saber, a solução ótima de  $Q_2(k)$ .

No caso em que  $\rho_k = 5$ , a restrição  $-y_{1k} - y_{2k} \leq 5$  deve ser eliminada de  $Q(k)$  para evitar degeneração.

O problema gerado é dado por

$$\begin{aligned}
 Q: \min_{x,y} & \left( \begin{bmatrix} (-2)1_K & e \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix} \\
 \text{s.a. } & y \in \arg \min_y \left( \begin{bmatrix} 0_K & 1_{2K} \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix} \\
 \text{s.a. } & \left( \begin{bmatrix} A & & B & & \\ & \ddots & & \ddots & \\ & & A & & B \end{bmatrix} M \right) \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} c^1 \\ \vdots \\ c^K \end{bmatrix},
 \end{aligned}$$

onde  $e = \begin{bmatrix} -1 & -2 & \dots & -1 & -2 \end{bmatrix}$  tem dimensão  $1 \times 2K$ ,  $A, B$  são como em (6.16),

$$c^k = \begin{bmatrix} -7 & -6 & \rho_k & 5 & 7 \end{bmatrix}^t,$$

e onde devemos eliminar as restrições  $-y_{1k} - y_{2k} \leq 5$ .

Sendo

$$I_{<} = \{k; \rho_k < 1\},$$

$$I_1 = \{k; \rho_k = 1\} \text{ e}$$

$$I_{>} = \{k; \rho_k > 1\},$$

o problema  $Q$  possui as seguintes propriedades, provenientes do Teorema 6.3.2:

- Possui  $2^{|I_1|}$  soluções ótimas globais e  $2^K - 2^{|I_1|}$  soluções ótimas locais/não globais. Logo possui solução ótima local/não global se, e somente se,  $|I_1| < K$ . Neste caso, são as soluções cuja justaposição tenha pelo menos uma solução ótima local/não global de algum  $Q(k)$ ,  $k \notin I_1$ ;
- O valor ótimo é  $F^* = 3 \sum_{k \in I_{<}} \rho_k - 19|I_{<}| - 16(|I_1| + |I_{>}|)$ ;
- O problema do baixo nível não retorna uma única solução ótima para todo  $x$  se, e somente se, algum  $\rho_k < 5$  for escolhido.

### 6.3.3 Resultados

Para os testes computacionais, agrupamos os problemas-teste de acordo com os seguintes critérios:

- O gerador usado - problemas-teste gerados segundo Calamai e Vicente (subseção 6.3.1, página 76) ou gerados como na subseção 6.3.2 (GRUPO 1 e 2);
- A quantidade de variáveis de alto e baixo níveis, no caso do gerador de Calamai e Vicente, e a quantidade de problemas iniciais usados, nos outros casos;
- Os parâmetros usados.

Cada grupo é composto por 20 problemas e tem parâmetros escolhidos aleatoriamente respeitando as especificidades de cada gerador (veja as subseções 6.3.1 e 6.3.2). Para os problemas gerados segundo Calamai e Vicente, 7 dimensões distintas foram consideradas, e para os problemas dos tipos GRUPO 1 e 2, 4 dimensões distintas cada. Para cada dimensão, 10 grupos, cada um com parâmetros diferentes, foram considerados. Foram analisados portanto  $7 \times 10 \times 20 = 1400$  problemas segundo Calamai e Vicente, e  $4 \times 10 \times 20 = 800$  problemas dos tipos GRUPO 1 e 2 cada. A matriz  $M$  foi calculada da mesma forma da subseção 6.3.1, para todos os geradores, com parâmetros  $\tau$  e  $\delta$  escolhidos aleatoriamente. Todos os números pseudo-randômicos foram gerados pelo método *minimal standard* (versão 3) comentado no fim da subseção 6.3.1.

O algoritmo implementado foi o Algoritmo 4 (página 76) com o passo 5”, que trata o caso de ilimitabilidade. Todos os códigos foram escritos em linguagem C, compilados no gcc versão 4.3 e executados em um computador com processador AMD X2 5600+, 2Gb de memória, sobre o sistema GNU/Linux Debian 5.0 64 bits versão estável. Para a resolução de problemas lineares, bem como a resolução do sistema do passo 3 do algoritmo, utilizamos o pacote livre Soplex [44] sobre o *framework* SCIP [1]. O Soplex permite o armazenamento das matrizes do problema como matrizes esparsas, e tal otimização foi utilizada.

Foi implementado também um pré-condicionador que realiza os seguintes procedimentos antes de resolver qualquer problema:

- Busca restrições de baixo nível que contenham somente variáveis de alto nível e as move para o alto nível. Isto não modifica o problema (Teorema 3.1.12, página 41);
- Busca restrições do tipo  $0x + 0y \leq \beta < 0$  e, caso encontre, retorna imediatamente que o problema é inviável;
- Busca restrições do tipo  $0x + 0y \leq \beta, \beta \geq 0$ , e, caso encontre, as elimina do problema.

Em especial, a primeira verificação é interessante visto que ao movermos restrições para o alto nível, reduzimos a dimensão do vetor de variáveis duais (o sistema do passo 3 fica menor).

A seguir, apresentamos as tabelas referentes aos resultados computacionais, separadas por tipo de problema e dimensão. A coluna “Gr.” indexa os grupos de 20 problemas. Os dados válidos para todos os grupos são apresentados no cabeçalho da tabela. As colunas em “Média dos tempos” representam a média dos tempos de execução do Passo 1, do Passo 3, do Passo 5 e a média dos tempos totais de execução (colunas “P1”, “P3”, “P5” e “Total” respectivamente). A coluna “P3/Total” indica a percentagem do tempo de execução do Passo 3 em relação ao tempo total, quando o cálculo for possível.

As colunas em “Média das iterações” relatam a média das iterações do Passo 3, a quantidade média de variáveis liberadas no Passo 3 ao longo de toda a execução e a média de iterações globais (colunas “P3”, “Lib.” e “Gl.” respectivamente). A coluna “Efic. P3” representa a percentagem de variáveis liberadas no Passo 3 em relação ao seu número de iterações, e será usada como medida de eficiência do Passo 3.

Nos problemas tabelados foi aplicado o pré-condicionador citado anteriormente, e por isso  $m_u > 0$ , apesar de originalmente os problemas gerados não terem restrições de alto nível. Também, a soma  $m_u + m_l$  varia pois restrições redundantes foram descartadas.

Optamos por primeiro apresentar todas as tabelas e somente depois fazermos a análise dos dados.

$n_x = 25, n_y = 40$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	38	53	< 0,01	< 0,01	< 0,01	< 0,01	–	78	33	42,38%	3
2	39	54	< 0,01	< 0,01	< 0,01	< 0,01	–	77	32	41,92%	3
3	36	51	< 0,01	0,01	< 0,01	0,01	80,77%	78	32	41,09%	3
4	49	64	< 0,01	< 0,01	< 0,01	0,01	–	78	35	45,05%	3
5	48	63	< 0,01	0,01	< 0,01	0,01	92,31%	78	37	46,77%	3
6	40	55	< 0,01	< 0,01	< 0,01	0,01	–	78	34	42,99%	3
7	36	51	< 0,01	< 0,01	< 0,01	< 0,01	–	78	33	42,12%	3
8	34	49	< 0,01	< 0,01	< 0,01	0,01	–	78	32	41,58%	3
9	50	65	< 0,01	< 0,01	< 0,01	0,01	–	78	34	44,34%	3
10	38	53	< 0,01	< 0,01	< 0,01	0,01	–	78	33	41,64%	3

Tabela 1: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 25$  e  $n_y = 40$ .

$n_x = 75, n_y = 35$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	96	56	< 0,01	< 0,01	< 0,01	< 0,01	–	71	9	12,81%	2
2	93	53	< 0,01	< 0,01	< 0,01	< 0,01	–	71	8	10,99%	2
3	90	50	< 0,01	< 0,01	< 0,01	< 0,01	–	71	6	8,52%	2
4	98	58	< 0,01	< 0,01	< 0,01	0,01	–	71	9	12,82%	2
5	89	49	< 0,01	< 0,01	< 0,01	< 0,01	–	71	6	8,18%	2
6	100	60	< 0,01	< 0,01	< 0,01	0,01	–	71	9	12,25%	2
7	88	48	< 0,01	< 0,01	< 0,01	0,01	–	71	6	8,25%	2
8	78	38	< 0,01	< 0,01	< 0,01	< 0,01	–	71	2	2,18%	2
9	95	55	< 0,01	< 0,01	< 0,01	< 0,01	–	71	8	11,90%	2
10	93	53	< 0,01	< 0,01	< 0,01	0,01	–	71	7	10,46%	2

Tabela 2: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 75$  e  $n_y = 35$ .

$n_x = 75, n_y = 90$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	111	126	0,01	0,14	0,02	0,17	84,43%	228	45	19,85%	3
2	110	125	< 0,01	0,11	0,01	0,13	83,27%	228	44	19,21%	3
3	95	110	< 0,01	0,13	< 0,01	0,14	88,85%	228	38	16,75%	3
4	87	102	< 0,01	0,11	< 0,01	0,13	87,78%	228	36	15,66%	3
5	112	127	0,01	0,15	0,02	0,17	85,06%	228	44	19,50%	3
6	147	162	0,01	0,16	0,02	0,19	84,42%	228	59	25,83%	3
7	120	135	< 0,01	0,16	< 0,01	0,17	90,72%	228	49	21,35%	3
8	120	135	0,01	0,11	0,01	0,14	81,79%	228	50	22,00%	3
9	141	156	0,01	0,17	0,02	0,20	85,86%	228	56	24,71%	3
10	101	116	< 0,01	0,12	< 0,01	0,13	90,97%	228	41	18,15%	3

Tabela 3: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 75$  e  $n_y = 90$ .

$n_x = 100, n_y = 70$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	154	124	0,01	0,06	0,02	0,09	66,49%	141	21	15,21%	2
2	147	117	< 0,01	0,06	0,01	0,07	75,51%	141	20	14,01%	2
3	129	99	0,01	0,05	0,02	0,08	63,47%	141	11	8,00%	2
4	104	74	< 0,01	0,05	< 0,01	0,06	87,80%	141	3	1,93%	2
5	143	113	< 0,01	0,05	0,01	0,07	69,40%	141	18	12,55%	2
6	144	114	< 0,01	0,08	0,01	0,09	80,42%	201	32	15,76%	3
7	137	107	0,01	0,05	0,01	0,08	68,48%	141	15	10,74%	2
8	139	109	< 0,01	0,06	0,01	0,08	70,97%	141	15	10,96%	2
9	144	114	< 0,01	0,04	0,01	0,06	67,50%	141	19	13,44%	2
10	147	117	< 0,01	0,06	0,01	0,08	75,15%	141	19	13,72%	2

Tabela 4: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 100$  e  $n_y = 70$ .

$n_x = 100, n_y = 150$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	176	226	0,02	0,68	0,03	0,74	92,36%	303	132	43,44%	3
2	170	220	0,03	0,87	0,04	0,94	92,39%	303	129	42,62%	3
3	114	164	0,01	0,73	0,02	0,77	95,47%	303	108	35,53%	3
4	168	218	0,02	0,92	0,03	0,97	94,45%	398	145	36,42%	4
5	149	199	0,02	0,76	0,03	0,81	94,16%	303	121	39,94%	3
6	200	250	0,04	0,96	0,05	1,05	91,59%	303	143	47,16%	3
7	172	222	0,03	0,73	0,04	0,79	91,86%	303	132	43,43%	3
8	182	232	0,02	0,64	0,03	0,69	92,36%	303	134	44,13%	3
9	200	250	0,04	0,95	0,05	1,04	91,37%	303	140	46,27%	3
10	175	225	0,03	0,82	0,04	0,89	92,29%	303	132	43,68%	3

Tabela 5: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 100$  e  $n_y = 150$ .

$n_x = 200, n_y = 150$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	284	234	0,07	0,90	0,11	1,08	83,01%	301	42	13,85%	2
2	235	185	0,06	0,80	0,09	0,95	84,11%	301	15	4,98%	2
3	299	249	0,11	0,78	0,15	1,04	75,12%	301	44	14,58%	2
4	308	258	0,10	0,82	0,14	1,06	77,23%	301	47	15,66%	2
5	256	206	0,09	0,90	0,14	1,13	79,73%	301	25	8,36%	2
6	299	249	0,07	0,78	0,10	0,95	82,08%	301	47	15,58%	2
7	350	300	0,11	0,78	0,16	1,05	74,37%	301	65	21,48%	2
8	311	261	0,09	0,77	0,13	0,99	77,78%	301	50	16,53%	2
9	258	208	0,05	0,69	0,08	0,82	84,01%	301	27	9,00%	2
10	255	205	0,05	0,66	0,08	0,79	83,91%	301	25	8,34%	2

Tabela 6: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 200$  e  $n_y = 150$ .

$n_x = 200, n_y = 350$ , problemas gerados segundo Calamai e Vicente											
			Média dos tempos (segundos)					Média das iterações			
Gr.	$m_u$	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	291	441	0,26	15,41	0,36	16,03	96,15%	603	343	56,92%	3
2	229	379	0,16	16,51	0,26	16,93	97,52%	603	317	52,59%	3
3	387	537	0,28	21,67	0,42	22,36	96,90%	603	389	64,55%	3
4	255	405	0,15	15,96	0,25	16,37	97,52%	603	332	55,03%	3
5	400	550	0,29	20,08	0,45	20,83	96,44%	603	397	65,85%	3
6	296	446	0,19	15,43	0,28	15,90	97,06%	593	336	56,64%	3
7	208	358	0,10	13,86	0,18	14,14	98,03%	593	296	49,95%	3
8	319	469	0,22	23,10	0,36	23,67	97,57%	603	359	59,47%	3
9	339	489	0,23	21,19	0,37	21,80	97,23%	603	369	61,21%	3
10	368	518	0,25	17,93	0,38	18,57	96,59%	603	375	62,16%	3

Tabela 7: Testes computacionais - Problemas gerados segundo Calamai e Vicente com  $n_x = 200$  e  $n_y = 350$ .

$n_x = 15, n_y = 30, m_u = 30$ , problemas do tipo GRUPO 1											
		Média dos tempos (segundos)					Média das iterações				
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.	
1	34	< 0,01	0,02	< 0,01	0,02	97,96%	210	25	11,97%	7	
2	41	< 0,01	0,04	< 0,01	0,04	93,83%	393	141	35,98%	16	
3	36	< 0,01	0,06	< 0,01	0,07	89,29%	232	29	12,51%	8	
4	36	< 0,01	0,06	< 0,01	0,06	90,24%	239	37	15,42%	9	
5	38	< 0,01	0,04	< 0,01	0,04	98,88%	357	92	25,71%	13	
6	30	< 0,01	< 0,01	< 0,01	< 0,01	—	30	0	0,00%	1	
7	40	< 0,01	0,05	< 0,01	0,06	94,21%	386	108	27,97%	15	
8	31	< 0,01	0,01	< 0,01	0,01	100,00%	123	6	4,92%	4	
9	35	< 0,01	0,03	< 0,01	0,03	86,36%	230	33	14,27%	8	
10	32	< 0,01	0,02	< 0,01	0,02	95,65%	152	8	5,28%	5	

Tabela 8: Testes computacionais - Problemas do tipo GRUPO 1 com  $n_x = 15$ .

$n_x = 25, n_y = 50, m_u = 50$ , problemas do tipo GRUPO 1										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	63	< 0,01	0,39	0,01	0,41	94,95%	665	127	19,12%	15
2	63	< 0,01	0,22	< 0,01	0,23	96,06%	778	202	26,00%	18
3	64	< 0,01	0,35	0,01	0,36	96,29%	852	209	24,50%	19
4	53	< 0,01	0,16	< 0,01	0,17	94,79%	331	27	8,29%	7
5	70	< 0,01	0,52	0,03	0,56	94,25%	1011	382	37,75%	25
6	62	< 0,01	0,65	0,03	0,68	95,47%	677	135	19,91%	15
7	52	< 0,01	0,13	< 0,01	0,14	95,68%	239	8	3,37%	5
8	61	< 0,01	0,74	0,03	0,78	95,43%	652	109	16,68%	14
9	59	< 0,01	0,44	0,01	0,45	96,04%	635	99	15,63%	14
10	65	< 0,01	1,16	0,05	1,22	95,11%	837	217	25,90%	19

Tabela 9: Testes computacionais - Problemas do tipo GRUPO 1 com  $n_x = 25$ .

$n_x = 50, n_y = 100, m_u = 100$ , problemas do tipo GRUPO 1										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	101	0,04	3,51	0,07	3,62	96,92%	401	4	0,94%	4
2	141	0,01	12,88	0,19	13,08	98,47%	5479	2744	50,08%	70
3	117	0,02	8,08	0,14	8,24	98,10%	2542	472	18,55%	28
4	141	0,03	18,75	0,40	19,17	97,78%	4103	1894	46,15%	53
5	117	< 0,01	4,44	0,08	4,53	98,08%	2632	482	18,32%	29
6	147	0,02	10,71	0,23	10,96	97,74%	4706	2590	55,02%	63
7	144	0,07	26,51	0,65	27,24	97,33%	4216	1974	46,82%	54
8	124	0,02	8,39	0,18	8,59	97,63%	3001	806	26,87%	35
9	149	0,07	23,55	0,56	24,17	97,43%	4107	2094	50,97%	54
10	142	0,06	20,77	0,49	21,32	97,42%	3859	1724	44,67%	49

Tabela 10: Testes computacionais - Problemas do tipo GRUPO 1 com  $n_x = 50$ .

$n_x = 70, n_y = 140, m_u = 140$ , problemas do tipo GRUPO 1										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	180	0,07	62,73	0,88	63,68	98,51%	6861	2228	32,47%	59
2	210	0,22	111,11	2,07	113,41	97,98%	7810	4061	52,00%	74
3	150	0,03	16,54	0,19	16,77	98,67%	2341	173	7,41%	17
4	160	0,09	71,45	0,93	72,47	98,59%	4654	707	15,19%	36
5	193	0,16	88,88	1,52	90,56	98,15%	6549	2603	39,75%	58
6	146	0,07	24,46	0,34	24,86	98,37%	1766	78	4,43%	13
7	204	0,18	160,84	2,10	163,12	98,61%	10164	5219	51,35%	94
8	183	0,05	84,34	0,95	85,34	98,83%	8389	2931	34,94%	72
9	207	0,09	76,21	1,04	77,35	98,54%	9081	4891	53,86%	86
10	159	0,07	62,21	0,76	63,04	98,68%	4135	582	14,06%	32

Tabela 11: Testes computacionais - Problemas do tipo GRUPO 1 com  $n_x = 70$ .

$n_x = 15, n_y = 30, m_u = 30$ , problemas do tipo GRUPO 2										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	43	< 0,01	0,01	< 0,01	0,02	78,57%	70	3	4,72%	2
2	41	< 0,01	0,01	< 0,01	0,02	93,75%	74	2	3,13%	3
3	39	< 0,01	0,01	< 0,01	0,01	93,10%	65	2	3,27%	2
4	40	< 0,01	< 0,01	< 0,01	< 0,01	—	67	3	3,74%	2
5	41	< 0,01	0,02	< 0,01	0,02	88,89%	72	2	3,04%	2
6	44	< 0,01	0,02	< 0,01	0,03	64,18%	63	3	3,99%	2
7	39	< 0,01	0,02	< 0,01	0,02	77,08%	78	4	5,21%	3
8	41	< 0,01	0,02	< 0,01	0,02	83,33%	70	3	4,56%	2
9	45	< 0,01	0,02	< 0,01	0,03	61,90%	71	5	7,45%	3
10	45	< 0,01	0,01	< 0,01	0,02	83,33%	72	4	5,41%	3

Tabela 12: Testes computacionais - Problemas do tipo GRUPO 2 com  $n_x = 15$ .

$n_x = 25, n_y = 50, m_u = 50$ , problemas do tipo GRUPO 2										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	68	< 0,01	0,04	< 0,01	0,05	80,18%	117	5	4,12%	2
2	66	< 0,01	0,05	< 0,01	0,06	83,67%	129	4	2,96%	3
3	69	< 0,01	0,04	< 0,01	0,06	78,13%	115	4	3,67%	2
4	72	< 0,01	0,06	< 0,01	0,08	83,71%	126	7	5,45%	3
5	67	0,02	0,14	0,02	0,18	78,18%	121	4	3,28%	3
6	62	< 0,01	0,04	< 0,01	0,05	76,92%	120	3	2,78%	2
7	75	< 0,01	0,03	< 0,01	0,04	69,39%	127	13	10,33%	3
8	70	< 0,01	0,04	< 0,01	0,05	76,86%	126	7	5,53%	3
9	75	< 0,01	0,10	< 0,01	0,12	90,64%	135	14	10,07%	3
10	64	< 0,01	0,06	< 0,01	0,08	78,31%	128	5	3,77%	3

Tabela 13: Testes computacionais - Problemas do tipo GRUPO 2 com  $n_x = 25$ .

$n_x = 50, n_y = 100, m_u = 100$ , problemas do tipo GRUPO 2										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	133	0,06	1,33	0,07	1,46	91,20%	218	6	2,93%	2
2	139	0,02	0,39	0,02	0,42	92,53%	203	6	2,96%	2
3	150	0,02	0,47	0,02	0,51	92,57%	204	7	3,67%	2
4	133	0,06	1,31	0,07	1,43	91,39%	265	11	4,28%	3
5	132	0,02	0,73	0,03	0,78	93,55%	258	14	5,38%	3
6	133	< 0,01	0,30	0,01	0,32	93,59%	251	10	3,84%	3
7	118	0,03	1,00	0,04	1,08	92,71%	267	7	2,80%	3
8	150	0,09	1,71	0,11	1,92	89,13%	264	21	7,99%	3
9	141	0,06	1,46	0,06	1,57	92,60%	200	6	3,10%	2
10	141	0,01	0,34	0,01	0,36	93,17%	259	18	6,91%	3

Tabela 14: Testes computacionais - Problemas do tipo GRUPO 2 com  $n_x = 50$ .

$n_x = 70, n_y = 140, m_u = 140$ , problemas do tipo GRUPO 2										
		Média dos tempos (segundos)					Média das iterações			
Gr.	$m_l$	P1	P3	P5	Total	P3/Total	P3	Lib.	Efic. P3	Gl.
1	190	0,06	2,16	0,06	2,27	94,99%	271	9	3,18%	2
2	180	0,07	2,75	0,07	2,90	95,09%	330	12	3,56%	2
3	201	0,09	3,39	0,12	3,61	94,10%	367	23	6,15%	3
4	163	0,13	5,15	0,16	5,44	94,68%	369	9	2,51%	3
5	196	0,22	7,89	0,22	8,32	94,77%	354	21	5,86%	3
6	197	0,05	1,98	0,06	2,09	94,69%	373	23	6,04%	3
7	178	0,22	7,13	0,26	7,62	93,65%	358	17	4,81%	3
8	165	0,13	5,28	0,16	5,57	94,68%	345	12	3,44%	3
9	205	0,11	3,80	0,15	4,05	93,70%	371	26	7,04%	3
10	190	0,05	1,55	0,06	1,65	93,71%	359	21	5,88%	3

Tabela 15: Testes computacionais - Problemas do tipo GRUPO 2 com  $n_x = 70$ .

Identificamos alguns fatores que ocorrem com regularidade nos testes, e que merecem ser analisados:

- Em todos os grupos, a maior parcela do tempo de execução foi do passo 3. Os tempos do passo 3 foram mais significativos nos grupos com  $n_x < n_y$ , em sua maioria, acima dos 90%. Os grupos 1, 6, 7, 9 da Tabela 12 (tipo GRUPO 2 com  $n_x = 15$ ) e os grupos 3, 5, 6, 7, 8 e 10 da Tabela 13 (tipo GRUPO 2 com  $n_x = 25$ ) tiveram uma participação do passo 3 no tempo de execução total abaixo dos 80%. Para estes grupos, não notamos regularidade em seus parâmetros (Tabelas 28 e 29, página 138). Apesar dessa pequena irregularidade, indiscutivelmente o passo 3 é o gargalo do processo. Torna-se necessário portanto o estudo de técnicas de escolha das variáveis/restrições duais a serem liberadas nesse passo (nossa implementação tenta liberar sempre as variáveis/restrições de maior índice). Vale lembrar que os problemas-teste usados não possuem restrições de não negatividade e portanto as restrições duais  $B_2^t \lambda + d_2^t \geq 0$  sempre são de igualdade. Se as restrições de não negatividade estivessem presentes, provavelmente o passo 3 consumiria mais tempo na execução.
- Nos grupos de problemas gerados segundo Calamai e Vicente, a eficiência do passo 3 (coluna “Efic. P3”) foi maior quando  $n_x < n_y$ . Mesmo nos grupos com valores

$m_l$  próximos (Tabelas 1 e 2, 3 e 4, 5 e 6), cujo número de variáveis duais é parecido, a diferença permanece. Nestes grupos, os que possuem maior  $n_y$  tiveram maior eficiência do passo 3, apesar de terem mais restrições duais de igualdade que os problemas com menor  $n_y$  (o número de restrições duais é  $n_y$ ). Isto ocorreu possivelmente pelo fato de mais restrições duais implicar que mais variáveis duais sejam zeradas numa mesma iteração do passo 3 (o que aumenta sua eficiência).

Destacamos neste contexto os grupos da Tabela 7 (problemas gerados segundo Calamai e Vicente com  $n_x = 200$  e  $n_y = 350$ ). Neles, a eficiência do passo 3 ultrapassou 50%, com exceção do grupo 7, que atingiu 49,95%. Isto mostra que nestes grupos, a cada iteração do passo 3, foi liberada em média uma ou mais variáveis duais. Logo, para estes grupos, a estratégia adotada no algoritmo é melhor do que liberar uma única variável de cada vez (como no algoritmo desenvolvido por Still [41]). Vale ressaltar que no fim do algoritmo, possivelmente o passo 3 é executado sem que alguma variável ou restrição dual seja liberada. Isto é evidente no grupo 6 da Tabela 8 (tipo GRUPO 1 com  $n_x = 15$ ): em 30 iterações do passo 3, nenhuma variável foi liberada. Isto ocorreu pois nos problemas deste grupo, não há soluções ótimas locais/não globais (veja Tabela 24) e o passo 1 encontrou uma solução ótima global.

- Outro fator é a relação entre as iterações do passo 3 e globais. Observamos em todos os grupos que o número de iterações do passo 3 é baixo quando o número de iterações globais o é, e vice-versa. Nos grupos do tipo GRUPO 1, esta relação é mais evidente pois o número de iterações globais é significativamente maior que nos outros grupos (e logo foram mais trabalhosos para nosso algoritmo) e varia muito em grupos de mesma dimensão. Isto indica que o número de iterações do passo 3 varia pouco a cada iteração global (e claramente depende da dimensão do problema). Observamos ainda que a eficiência do passo 3 acompanha esta relação: quanto mais iterações, maior a eficiência do passo 3. Com isto, em grandes problemas com poliedro  $\Omega$  muito recortado (com muitas restrições) e cujas soluções ótimas locais sejam “de difícil acesso”, é de se esperar que a eficiência do passo 3 seja grande. Por soluções ótimas “de difícil acesso” queremos dizer daquelas distantes do ponto admissível inicial obtido no passo 1 (que geralmente é calculado da mesma forma em outros métodos).

Concluimos nossa análise dizendo que, a despeito de comparações com outros métodos, o nosso algoritmo é viável do ponto de vista prático, e que possivelmente pode ser melhorado. Uma de suas boas características é a generalidade. Um dos grandes desafios para

---

resolver problemas com variáveis de baixo nível no alto nível, tanto para o nosso algoritmo como para outros, é encontrar um primeiro ponto admissível (George Still mostra que essa tarefa é NP-completa [41]).

Por fim, relatamos que, dentre os problemas que possuem soluções ótimas locais/não globais (veja o Apêndice C para saber quais grupos), 19,03% dos problemas gerados segundo Calamai e Vicente, 25,56% dos problemas do tipo GRUPO 1 e 3,56% dos problemas do tipo GRUPO 2 foram resolvidos globalmente. Apesar dessas quantidades serem expressivas para os dois primeiros tipos, não acreditamos que seja um dado relevante, visto que a estrutura específica de cada problema é crucial para tal fato.

## 7 Outras questões em problemas de dois níveis

Neste capítulo abordaremos questões que não são necessárias aos outros capítulos.

### 7.1 Regularização de problemas com variáveis de baixo nível inteiras

Como dito no Capítulo 2, quando um problema de dois níveis tem como retorno do baixo nível uma única solução ótima para cada parâmetro  $x$  do alto nível, sua representação (e possivelmente sua resolução) se torna mais fácil. Nosso objetivo aqui é estudar métodos para, dado  $x$ , ser suficiente a escolha de uma única solução ótima do baixo nível, garantindo sua otimalidade.

Consideremos o problema

$$\begin{aligned} \text{LBLP1: } \min_{x,y} \{c_1x + d_1(x)y\} \\ \text{s.a. } A_1x \leq b_1, \quad x \geq 0 \\ y \in \arg \min_y d_2(x)y \\ \text{s.a. } A_2x + B_2y \leq b_2, \quad y \in \mathbb{Z}_+^{n_y}, \end{aligned}$$

onde  $d_2(x) \in \mathbb{Z}^{n_y}$  para todo  $x \in \Omega_x$ .

Para cada  $x \in \Omega_x$ , para todo  $y \in S(x)$ ,  $(x, y)$  é admissível pois não há variáveis de baixo nível nas restrições do alto nível. Supomos  $\Omega_l$  compacto. Seja  $L > 0$  tal que

$c_1x + d_1(x)y \in \left(-\frac{L}{2}, \frac{L}{2}\right)$ ,  $\forall (x, y) \in \Omega_l \supset \Upsilon$ , e consideremos o modelo

$$\begin{aligned} \text{LBLP2: } \min_{x,y} \{c_1x + d_1(x)y\} \\ \text{s.a. } A_1x \leq b_1, \quad x \geq 0 \\ y \in \arg \min_y \left\{ d_2(x)y + \frac{1}{L}[c_1x + d_1(x)y] \right\} \\ \text{s.a. } B_2y \leq b_2 - A_2x, \quad y \in \mathbb{Z}_+^{n_y}. \end{aligned}$$

Temos o seguinte resultado:

**Teorema 7.1.1.** *Toda solução ótima de LBLP2 é solução ótima de LBLP1.*

*Demonstração.* No que se segue, usaremos superíndices para indicar estruturas de LBLP1 ou LBLP2. Mostremos que  $S^2(x) \subset S^1(x)$ . Seja  $y \in S^2(x)$  e suponha que  $y \notin S^1(x)$ . Então existe  $y^* \in \Omega_y^1(x)$  com  $d_2(x)y^* < d_2(x)y$ . Ora,  $d_2(x)$ ,  $y$  e  $y^*$  são vetores de inteiros e logo

$$\begin{aligned} d_2(x)y^* + \frac{1}{L}[c_1x + d_1(x)y^*] &< d_2(x)y^* + \frac{1}{2} \leq d_2(x)y - \frac{1}{2} \\ &< d_2(x)y + \frac{1}{L}[c_1x + d_1(x)y] \end{aligned}$$

pois  $\left| \frac{1}{L}[c_1x + d_1(x)y^*] \right| < \frac{1}{2}$ . Isto quer dizer que  $y \notin S^2(x)$ , absurdo.

Agora, do fato de  $\Omega_u^1 = \Omega_u^2$ , segue que  $\Upsilon^2 \subset \Upsilon^1$ . A fim de concluirmos o resultado, suponhamos que  $(x, y) \in \Upsilon^1 \setminus \Upsilon^2$  seja solução ótima de LBLP1. Então necessariamente  $y \in S^1(x) \setminus S^2(x)$ . Temos o seguinte:

- $\exists y^* \in \Omega_y(x) = \Omega_y^1(x) = \Omega_y^2(x)$  tal que  $d_2(x)y^* + \frac{1}{L}[c_1x + d_1(x)y^*] < d_2(x)y + \frac{1}{L}[c_1x + d_1(x)y]$ . Em particular, podemos considerar  $y^* \in S^2(x)$ ;
- $d_2(x)y \leq d_2(x)y^*$  pois  $y \in S^1(x)$ ;
- $c_1x + d_1(x)y \leq c_1x + d_1(x)y^*$ , visto que LBLP1 é uma relaxação de LBLP2 ( $\Upsilon^2 \subset \Upsilon^1$ ).

Daí

$$0 \leq d_2(x)y^* - d_2(x)y < \frac{1}{L}[(c_1x + d_1(x)y) - (c_1x + d_1(x)y^*)] \leq 0,$$

absurdo. Isto mostra que toda solução ótima de LBLP1 está em  $\Upsilon^2$ , e como LBLP1 é relaxação de LBLP2, concluímos o resultado.  $\square$

**Teorema 7.1.2.** *A fim de resolver LBLP2, é suficiente considerar uma única solução ótima  $y$  do problema de baixo nível  $LLP^2(x)$ , para cada  $x$ .*

*Demonstração.* Devemos mostrar que, dado  $x \in \Omega_x^2$ , uma solução ótima  $y$  de  $\text{LLP}^2(x)$  é uma melhor escolha para o alto nível, relativo a este  $x$ . Seja dado  $x \in \Omega_x^2$  e tome  $y^* \in S^2(x)$ , isto é,  $y^* \in \Omega_y^2(x)$  tal que

$$d_2(x)y^* + \frac{1}{L}[c_1x + d_1(x)y^*] \leq d_2(x)y + \frac{1}{L}[c_1x + d_1(x)y], \quad \forall y \in \Omega_y^2(x).$$

Para cada  $y \in S^2(x)$ , a desigualdade acima se realiza como igualdade e logo

$$\begin{aligned} \frac{1}{L}[(c_1x + d_1(x)y^*) - (c_1x + d_1(x)y)] &= d_2(x)y - d_2(x)y^* = 0 \\ &\Rightarrow c_1x + d_1(x)y = c_1x + d_1(x)y^*, \end{aligned}$$

onde  $d_2(x)y - d_2(x)y^* = 0$  decorre do fato de  $d_2(x)y - d_2(x)y^* \in \mathbb{Z}$  e  $|\frac{1}{L}[(c_1x + d_1(x)y^*) - (c_1x + d_1(x)y)]| < 1$ .  $\square$

Neste sentido, para cada  $x \in \Omega_x^2$ , escolhemos um  $y \in S^2(x)$  e definimos a função

$$\begin{aligned} \varphi : \Omega_x^2 &\rightarrow \mathbb{R}^{n_y} \\ x &\rightarrow y \text{ escolhido} \end{aligned}$$

(esta escolha pode ser feita por algum método determinístico de resolução de  $\text{LLP}^2(x)$  que não faça escolhas aleatórias). Obtemos então o problema

$$\begin{aligned} \min_{x,y} \{c_1x + d_1(x)\varphi(x)\} \\ \text{s.a. } A_1x \leq b_1, \quad x \in \Omega_x^2, \quad x \geq 0 \end{aligned}$$

que resolve LBLP pelos resultados anteriores.

Toda a discussão feita nessa seção vale também quando  $x \in \mathbb{Z}_+^{n_x}$ . Uma outra observação, de ordem prática, é a preocupação com a estabilidade numérica. De fato, o termo  $\frac{1}{L}[c_1x + d_1(x)y]$  pode ser numericamente instável se  $\frac{1}{L}$  for muito pequeno. Isto pode ser amenizado da seguinte forma: se multiplicarmos a função objetivo do baixo nível por qualquer  $M > 0$ , os pontos admissíveis do problema de dois níveis (e suas soluções ótimas) não são afetados. Passamos então à nova função de baixo nível

$$Md_2(x)y + \frac{M}{L}[c_1x + d_1(x)y]$$

onde o termo  $\frac{M}{L}[c_1x + d_1(x)y]$  é mais numericamente estável que o original (quando  $M = 1$ ) se  $M > 1$ . Por outro lado,  $M$  deve ser pequeno o suficiente para que essa função não extrapole o limite do tipo numérico utilizado na programação.

## 8 Conclusões e trabalhos futuros

Traçamos como objetivo deste trabalho o estudo teórico e prático dos problemas de dois níveis. Os Capítulos 2 e 3 abordaram de forma rigorosa o problema em questão em sua versão geral e linear, respectivamente. Neles encontra-se a primeira contribuição, do ponto de vista teórico, deste trabalho. Trata-se da formalização entre os modelos apresentados para a abordagem otimista: os modelos (2.2) e (2.4)-(2.5), em concordância com Dempe (que concebe tais modelos quando restrições de alto nível não dependem da escolha do baixo nível), e o modelo BLP, considerado por muitos autores como o modelo do problema de dois níveis. Também, a preocupação com a teoria traz à tona demonstrações dos principais resultados da literatura, sejam elas novas ou reproduzidas de trabalhos já publicados, reunindo-as e minimizando a dispersão presente na literatura. Nesse contexto, estabelecemos a equivalência entre soluções ótimas locais dos modelos (2.2) linear otimista e BLP, resultado que de nosso conhecimento não havia sido posto (nem mesmo para problemas com restrições de alto nível independentes da escolha do baixo nível).

Ainda no campo teórico, estendemos resultados dos problemas de dois níveis a problemas multinível, pouco estudados na literatura. Também, dada a escassez de métodos para geração de problemas-teste, propomos uma generalização do método de Calamai e Vicente. Isso abre possibilidade para geração de um número muito variado de problemas-teste (na verdade, de geradores de problemas-teste). Esse pode ser inclusive um importante passo para minimizar as dificuldades encontradas na literatura quanto à comparação de algoritmos: atualmente, códigos dos métodos existentes não estão disponíveis à comunidade, e geralmente os trabalhos não usam a mesma metodologia na geração de problemas.

Quanto ao aspecto prático, a maior parte dos trabalhos dos últimos anos baseia-se na adaptação de metaheurísticas. Pouco se fez em relação a métodos exatos, ou mesmo em relação a algoritmos para soluções ótimas locais. O último trabalho de que temos notícia nessa linha é o de Still, de 2002. O algoritmo proposto traz novo fôlego à discussão. No entanto, não inova muito, na medida em que se baseia em reescrever o problema de baixo

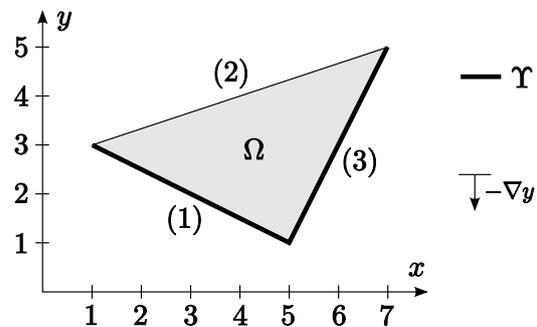
nível em suas condições de otimalidade de Kuhn-Tucker.

Como já comentamos, comparações verdadeiras de nosso algoritmo com outros métodos ainda são difíceis de serem feitas. Os testes computacionais indicam que o algoritmo proposto encontra uma solução ótima local em tempo razoável, a exceção dos problemas do GRUPO 1. Vale ressaltarmos que os problemas desse grupo não são usados em outros trabalhos, e possuem em geral retorno não único do baixo nível. É sabido na literatura que problemas com essa característica são mais difíceis para os métodos existentes. Portanto, nada pode ser afirmado a respeito da eficiência do algoritmo. Acreditamos, no entanto, que o algoritmo proposto é viável e passível de melhora, sobretudo no passo 3, o gargalo do processo.

Os problemas de dois níveis abrem várias possibilidades de investigação. A seguir, apontamos algumas delas, que julgamos pertinentes:

1. Estudo de problemas com variáveis inteiras;
2. Estudo de problemas não lineares de dois níveis;
3. Estudo das relações entre problemas de dois níveis com problemas com restrições de equilíbrio e problemas multiobjetivo;
4. Estudo de condições suficientes mais fortes que as do Teorema 3.1.12 (página 41) para mover restrições do baixo para o alto nível, sem prejuízo ao conjunto admissível  $\Upsilon$  do problema. Por exemplo, nos problemas lineares bidimensionais podemos passar a restrição  $\beta y \leq b_i - \alpha x$  ( $x, y \in \mathbb{R}$ ) do baixo para o alto nível sempre que  $\beta d_2 \geq 0$ , onde  $d_2 \in \mathbb{R}$  é o coeficiente de  $y$  na função objetivo de baixo nível. Isso ocorre devido ao fato de que, fixado  $x$ , na otimalidade do problema de baixo nível, sempre há um cone tangente que contém  $-d_2$ . Para ilustrar, considere o problema

$$\begin{aligned} \min_{x,y} \{x + y\} \\ \text{s.a. } y \in \arg \min_y y \\ \text{s.a. } x + 2y \leq 7 \quad (1) \\ \quad -x + 3y \leq 8 \quad (2) \\ \quad -2x - y \leq 9 \quad (3). \end{aligned}$$



Note que apesar da restrição (2) ter a variável do baixo nível, pode passar para o alto nível, e temos  $3 \cdot 1 \geq 0$ ;

5. Desenvolvimento de novos métodos de resolução suficientemente eficientes para problemas com variáveis de baixo nível no alto nível. Neste caso faz-se necessária a construção de métodos para obtenção de um ponto admissível inicial.

Acreditamos que os itens (1), (2), (3) e (5) sejam caminhos naturais a serem seguidos, e já existem vários trabalhos nesse sentido. Para o item (4), pelo menos de nosso conhecimento, não há trabalhos que procuram atacar a questão. Nos métodos de resolução que testam a otimalidade do baixo nível para cada  $x$  escolhido (como é o caso do nosso algoritmo e do algoritmo de Hansen, Jaumard e Savard), a questão torna-se interessante pois aí o grande número de restrições do baixo nível, em geral, eleva o tempo computacional.

Por fim, julgamos ser relevante para o entendimento de problemas de dois níveis o estudo de funções ponto a conjunto, cuja referência é o livro de Aubin e Frankowska [2].

# Referências

- [1] T. Achterberg. *Constraint Integer Programming*. Tese (Doutorado), Technische Universität Berlin, 2007. Disponível em <<http://opus.kobv.de/tuberlin/volltexte/2007/1611/>>.
- [2] J. P. Aubin e H. Frankowska. *Set-valued analysis*. Birkhäuser Boston, 1990.
- [3] J. F. Bard. *Practical bilevel optimization: algorithms and applications*, volume 30 de *Nonconvex Optimization and Its Applications*. Kluwer Academic Publishers, 1998.
- [4] M. S. Bazaraa e J. J. Jarvis. *Linear programming and network flows*. John Wiley & Sons, 1977.
- [5] D. P. Bertsekas, A. Nedic, e A. E. Ozdaglar. *Convex analysis and optimization*. Athena Scientific, Belmont, Massachusetts, 2003.
- [6] V. Böhm. On the continuity of the optimal policy set for linear programs. *SIAM Journal of Applied Mathematics*, 28(2):303–306, 1975.
- [7] P. Calamai e L. Vicente. Generating linear and linear-quadratic bilevel programming problems. *SIAM Journal on Scientific and Statistical Computing*, 14(4):770–782, 1993.
- [8] H. I. Calvete e C. Galé. Bilevel multiplicative problems: A penalty approach to optimality and a cutting plane based algorithm. *Journal of Computational and Applied Mathematics*, 218(2):259–269, 2008.
- [9] H. I. Calvete, C. Galé, e P. M. Mateo. A new approach for solving linear bilevel problems using genetic algorithms. *European Journal of Operational Research*, 188(1):14–28, 2008.
- [10] H. I. Calvete, C. Galé, e P. M. Mateo. A genetic algorithm for solving linear fractional bilevel problems. *Annals of Operations Research*, 166(1):39–56, 2009.
- [11] M. B. C. Campêlo. *Programação linear em dois níveis: uma abordagem teórica e computacional*. Tese (Doutorado), COPPE/UFRJ, Rio de Janeiro, Dezembro 1999.
- [12] B. Colson, P. Marcotte, e G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153(1):235–256, 2007.
- [13] C. H. M. de Sabóia, M. B. C. Campêlo, e S. Scheimberg. A computational study of global algorithms for bilevel linear programming. *Numerical Algorithms*, 35(2-4):155–173, 2004.
- [14] S. Dempe. *Foundations of bilevel programming*, volume 61 de *Nonconvex Optimization and Its Applications*. Kluwer Academic Publishers, 2002.

- [15] S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52:333–359, 2003.
- [16] S. Dempe, V. Kalashnikov, e R. Z. Ríos-Mercado. Discrete bilevel programming: Application to a natural gas cash-out problem. *European Journal of Operational Research*, 166(2):469–488, 2005.
- [17] S. Dempe, V. V. Kalashnikov, e N. Kalashnykova. Optimality conditions for bilevel programming problems. Em S. Dempe e V. V. Kalashnikov, editores, *Optimization with Multivalued Mappings*, volume 2 de *Springer Series in Optimization and Its Applications*, páginas 3–28. Springer Science+Business Media, LLC, 2006.
- [18] X. Deng. Complexity issues in bilevel linear programming. Em A. Migdalas, P. M. Pardalos, e P. Värbrand, editores, *Multilevel optimization: algorithms and applications*, Nonconvex Optimization and Its Applications, capítulo 6, páginas 149–164. Kluwer Academic Publishers, 1998.
- [19] S. Dewez. *On the toll setting problem*. Tese (Doutorado), Université Libre de Bruxelles, 2004.
- [20] S. Dewez, M. Labbè, P. Marcotte, e G. Savard. New formulations and valid inequalities for a bilevel pricing problem. *Operations Research Letters*, 36:141–149, 2008.
- [21] C. A. Floudas e C. E. Gounaris. A review of recent advances in global optimization. *Journal of Global Optimization*, 45(1):3–38, 2009.
- [22] Z. H. Günius, K. Sahin, e A. Ciric. Bilevel programming: applications in engineering. Em C. A. Floudas e P. M. Pardalos, editores, *Encyclopedia of Optimization 2nd Edition*. Springer Science+Business Media, LLC, 2008.
- [23] G. Heilporn, M. Labbè, P. Marcotte, e G. Savard. The highway problem: models, complexity and valid inequalities. Em *Proceedings of TRISTAN VI*. Sixth Triennial Symposium on Transportation Analysis, Phuket Island, Tailândia, Junho 2007.
- [24] A. Izmailov e M. Solodov. *Otimização vol. 1 - Condições de otimalidade, elementos de análise convexa e de dualidade*. IMPA, 2005.
- [25] R. J. Kuo e C. C. Huang. Application of particle swarm optimization algorithm for solving bi-level linear programming problem. *Computers and Mathematics with Applications*, 58(4):678–685, 2009.
- [26] S. Leyffer. Página de Sven Leyffer. Disponível em <<http://wiki.mcs.anl.gov/leyffer/index.php/MacMPEC>>. Acesso em 27/09/09.
- [27] E. L. Lima. *Álgebra Linear*. Coleção Matemática Universitária. IMPA, Rio de Janeiro, 5ª edição, 2001.
- [28] E. L. Lima. *Curso de análise*, volume 2 de *Projeto Euclides*. IMPA, Rio de Janeiro, 8ª edição, 2005.
- [29] Z. Lukac, K. Soric, e V. V. Rosenzweig. Production planning problem with sequence dependent setups as a bilevel programming problem. *European Journal of Operational Research*, 187(3):1504–1512, 2008.

- [30] P. Marcotte e G. Savard. Bilevel programming: A combinatorial perspective. Em D. Avis, A. Hertz, e O. Marcotte, editores, *Graph Theory and Combinatorial Optimization*, capítulo 7, páginas 191–218. Springer US, 2005.
- [31] P. Marcotte e G. Savard. Bilevel programming: applications. Em C. A. Floudas e P. M. Pardalos, editores, *Encyclopedia of Optimization 2nd Edition*. Springer Science+Business Media, LLC, 2008.
- [32] P. Marcotte, G. Savard, e F. Semet. A bilevel approach to the travelling salesman problem. *Operations Research Letters*, 32(3):240–248, 2004.
- [33] P. Marcotte e D. L. Zhu. Exact and inexact penalty methods for the generalized bilevel programming. *Mathematical Programming*, 74(2):141–157, 1996.
- [34] K. Moshirvaziri, M. A. Amouzegar, e S. E. Jacobsen. Test problem construction for linear bilevel programming problems. *Journal of Global Optimization*, 8(3):235–243, 1996.
- [35] S. K. Park e K. W. Miller. Random number generators: good ones are hard to find. *Communications of the ACM*, 31(10):1192–1201, 1988.
- [36] K. H. Sabin e A. R. Ciric. A dual temperature simulated annealing approach for solving bilevel programming problems. *Computers and Chemical Engineering*, 23(1):11–25, 1998.
- [37] G. K. Saharidis e M. G. Ierapetritou. Resolution method for mixed integer bi-level linear problems based on decomposition technique. *Journal of Global Optimization*, 44(1):29–51, 2009.
- [38] R. Saucier. Computer generation of statistical distributions. Relatório técnico, Army Research Laboratory, Março 2000. Disponível em <<http://ftp.arl.mil/random/random.pdf>>.
- [39] A. Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1998.
- [40] K. Shimizu, Y. Ishizuka, e J. F. Bard. *Nondifferentiable and two-level mathematical programming*. Kluwer Academic Publishers, 1997.
- [41] G. Still. Linear bilevel problems: genericity results and an efficient method for computing local minima. *Mathematical Methods of Operations Research*, 55(3):383–400, 2002.
- [42] U. P. Wen e A. D. Huang. A simple tabu search method to solve the mixed-integer linear bilevel programming problem. *European Journal of Operational Research*, 88(3):563–571, 1996.
- [43] R. J. B. Wets. Programming under uncertainty: the equivalent convex program. *SIAM Journal of Applied Mathematics*, 14(1):89–105, 1966.
- [44] R. Wunderling. *Paralleler und objektorientierter Simplex-Algorithmus*. Tese (Doutorado), Technische Universität Berlin, 1996. Disponível em <<http://www.zib.de/Publications/abstracts/TR-96-09/>>.

# APÊNDICE A – Revisão de pré-requisitos

Neste apêndice apresentamos conceitos sobre alguns dos pré-requisitos para o entendimento deste trabalho. Abordamos o mínimo necessário de cada assunto, omitindo as demonstrações dos resultados. Obviamente, a leitura não é necessária aos leitores familiarizados com os assuntos.

## A.1 Álgebra Linear

Discutiremos alguns conceitos e resultados em Álgebra Linear. Como referência, indicamos o livro do Elon Lages Lima [27].

Um conjunto  $E$  munido de operações de *soma* ( $E \times E \ni (u, v) \rightarrow u + v \in E$ ) e *multiplicação por escalar* ( $\mathbb{R} \times E \ni (\alpha, u) \rightarrow \alpha u \in E$ ) é dito *espaço vetorial* se, para todos  $u, v, w \in E$  e  $\alpha, \beta \in \mathbb{R}$ , as condições abaixo são satisfeitas:

- $u + v = v + u$  (comutatividade)
- $(u + v) + w = u + (v + w)$  e  $(\alpha\beta)v = \alpha(\beta v)$  (associatividade)
- existe um  $0 \in E$ , chamado *vetor nulo*, tal que  $v + 0 = 0 + v = v$  para todo  $v \in E$
- para cada  $v \in E$  existe  $-v \in E$ , chamado *simétrico de  $v$* , tal que  $-v + v = v + (-v) = 0$
- $(\alpha + \beta)v = \alpha v + \beta v$  e  $\alpha(u + v) = \alpha u + \alpha v$  (distributividade)
- $1v = v$ .

Os elementos de um espaço vetorial são chamados *vetores*. Daqui para frente,  $E$  sempre denotará um espaço vetorial.

Um conjunto  $F \subset E$  é dito ser um *subespaço vetorial de  $E$*  se

$$\begin{aligned} 0 &\in F, \\ u, v \in F &\Rightarrow u + v \in F \text{ e} \\ v \in F &\Rightarrow \alpha v \in F, \forall \alpha \in \mathbb{R}. \end{aligned}$$

Em particular, todo subespaço vetorial é um espaço vetorial.

Dizemos que um conjunto finito  $X \subset E$  é linearmente independente (LI) quando nenhum vetor  $v \in X$  é *combinação linear* dos outros vetores de  $X$ , isto é, não se pode ter

$$v = \sum_{u \in X \setminus \{v\}} \alpha_u u,$$

$\alpha_u \in \mathbb{R}$ . Observamos que se  $X$  é LI então  $0 \notin X$ . Equivalentemente, é possível mostrar que  $X$  é LI se, e somente se,  $\sum_{u \in X} \alpha_u u = 0 \Rightarrow \alpha_u = 0, \forall u \in X$ . Quando  $X$  não é LI, dizemos que  $X$  é *linearmente dependente* (LD), ou que seus vetores são LD.

Dizemos que  $B = \{v_1, \dots, v_n\} \subset E$  é uma *base* de  $E$  se  $B$  é LI e  $B$  gera  $E$ , isto é, se qualquer  $v \in E$  é combinação linear dos vetores de  $B$ . É possível mostrar que se  $E$  admite uma base  $B$  (finita), qualquer outra base de  $E$  tem o mesmo número de elementos de  $B$ . A este número chamamos de *dimensão de  $E$* , e denotamos por  $\dim E$ .

Sejam  $E, F$  espaços vetoriais. Uma *transformação linear* é uma função  $T : E \rightarrow F$ ,  $v \rightarrow Tv$ , que satisfaz  $T(u + v) = Tu + Tv$  e  $T(\alpha v) = \alpha Tv$  para quaisquer  $u, v \in E$  e  $\alpha \in \mathbb{R}$ . Os conjuntos  $\text{Im}(T) = \{Tv; v \in E\}$  e  $\text{Ker}(T) = \{v; Tv = 0\}$  são chamados *imagem* e *núcleo* de  $T$ , respectivamente. É possível mostrar que  $\text{Im}(T)$  e  $\text{Ker}(T)$  são subespaços de  $F$  e  $E$ , respectivamente. Também não é difícil ver que  $T$  fica bem definida dizendo apenas quem são suas imagens sobre vetores de uma base de  $E$ .

O próximo resultado relaciona as dimensões de  $\text{Im}(T)$  e  $\text{Ker}(T)$  para qualquer transformação linear  $T$ .

**Teorema A.1.1** (do Núcleo e da Imagem). *Sejam  $E, F$  espaços vetoriais de dimensão finita. Para toda transformação linear  $T : E \rightarrow F$  tem-se  $\dim E = \dim \text{Ker}(T) + \dim \text{Im}(T)$ .*

*Demonstração.* Ver [27, Teorema 6.6]. □

Um importante espaço vetorial é o conjunto

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ fatores}}$$

munido das operações

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \quad (\text{soma}) \text{ e}$$

$$\alpha x = (\alpha x_1, \dots, \alpha x_n) \quad (\text{multiplicação por escalar}),$$

onde  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ ,  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$  e  $\alpha \in \mathbb{R}$ . Uma base de  $\mathbb{R}^n$ , chamada *base canônica*, é o conjunto  $\{e_1, \dots, e_n\}$ , onde

$$e_i = (0, \dots, 0, \underbrace{1}_{\text{posição } i}, 0, \dots, 0)$$

é o *i*-ésimo vetor canônico de  $\mathbb{R}^n$ . Segue imediatamente que  $\dim \mathbb{R}^n = n$ .

Dada uma matriz  $A \in \mathbb{R}^{m \times n}$ , associamos a transformação linear  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  definida por  $T_A(e_i) = Ae_i$ ,  $\forall i = 1, \dots, n$ , onde em  $Ae_i$  interpretamos os vetores  $e_i$  como matrizes de dimensão  $1 \times n$ .

Um conjunto  $V \subset \mathbb{R}^n$  é uma *variedade afim* em  $\mathbb{R}^n$  quando  $\{(1-t)x + ty; t \in \mathbb{R}\} \subset V$  para quaisquer  $x, y \in V$ , isto é, a reta que une dois pontos quaisquer de  $V$  está contida em  $V$ . Toda variedade afim em  $\mathbb{R}^n$  não vazia pode ser obtida trasladando-se um subespaço vetorial de  $\mathbb{R}^n$ . Mais precisamente, dada uma variedade afim não vazia  $V$  em  $\mathbb{R}^n$ , é possível mostrar que existe um único subespaço vetorial  $F$  de  $\mathbb{R}^n$  tal que, para todo  $x \in V$ , tem-se  $V = x + F = \{x + v; v \in F\}$ . Neste caso dizemos que  $V$  é *paralela* a  $F$ . No presente contexto, temos o

**Teorema A.1.2.** *Seja  $A \in \mathbb{R}^{m \times n}$  e  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  sua transformação linear associada. Para todo  $b \in \text{Im}(T_A)$ , o conjunto  $V = \{z \in \mathbb{R}^n; Az = b\}$  é uma variedade afim em  $\mathbb{R}^n$ , paralela a  $\text{Ker}(T_A)$ .*

*Demonstração.* Ver [27, Teorema 6.4]. □

## A.2 Topologia no $\mathbb{R}^n$

Serão apresentados conceitos de topologia no espaço euclidiano  $\mathbb{R}^n$ . Uma boa referência no assunto é o livro do Elon Lages Lima [28].

Um *produto interno* num espaço vetorial  $E$  de dimensão finita é uma função  $\langle \cdot, \cdot \rangle_E :$

$E \times E \rightarrow \mathbb{R}$  tal que para todos  $x, y, z \in E$  e  $\alpha \in \mathbb{R}$  temos

$$\begin{aligned}\langle x, y \rangle_E &= \langle y, x \rangle_E, \\ \langle x + z, y \rangle_E &= \langle x, y \rangle_E + \langle z, y \rangle_E, \\ \langle \alpha x, y \rangle_E &= \alpha \langle x, y \rangle_E \text{ e} \\ x \neq 0 &\Rightarrow \langle x, x \rangle_E > 0.\end{aligned}$$

Em particular,  $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  definido por

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

é um produto interno em  $\mathbb{R}^n$ , chamado *produto interno canônico* de  $\mathbb{R}^n$ . Às vezes é conveniente olhar este produto interno como uma multiplicação de matrizes, da seguinte forma: a cada  $x \in \mathbb{R}^n$ , associamos a matriz  $\mathbf{x} = [x_i] \in \mathbb{R}^{n \times 1}$  de ordem  $n \times 1$ . Temos então

$$\langle x, y \rangle = \mathbf{x}^t \mathbf{y}.$$

Quando não houver perigo de confusão, escrevemos simplesmente  $x^t y$ .

Uma *norma* num espaço vetorial  $E$  de dimensão finita é uma função  $\|\cdot\|_E : E \rightarrow \mathbb{R}_+$  tal que para todos  $x, y \in E$  e  $\alpha \in \mathbb{R}$  temos

$$\begin{aligned}\|x + y\|_E &\leq \|x\|_E + \|y\|_E, \\ \|\alpha x\|_E &= |\alpha| \|x\|_E \text{ e} \\ x \neq 0 &\Rightarrow \|x\|_E > 0.\end{aligned}$$

Duas normas  $\|\cdot\|_1$  e  $\|\cdot\|_2$  em  $\mathbb{R}^n$  são equivalentes se existirem  $a, b > 0$  tais que

$$\|x\|_1 \leq a \|x\|_2 \quad \text{e} \quad \|x\|_2 \leq b \|x\|_1, \quad \forall x \in \mathbb{R}^n.$$

O resultado a seguir nos permite trabalhar em  $\mathbb{R}^n$  com a norma mais conveniente para cada situação.

**Teorema A.2.1.** *Duas normas quaisquer em  $\mathbb{R}^n$  são equivalentes.*

*Demonstração.* Ver [28, Teorema 8 do capítulo I]. □

Em  $\mathbb{R}^n$ , é comum trabalharmos com as normas

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2} \quad (\text{norma euclidiana}) \text{ e}$$

$$\|x\|_\infty = \max \{|x_1|, \dots, |x_n|\} \quad (\text{norma do infinito}).$$

Um importante resultado é o

**Teorema A.2.2** (Desigualdade de Cauchy-Schwarz). *Para quaisquer  $x, y \in \mathbb{R}^n$  temos  $|x^t y| \leq \|x\| \|y\|$ .*

*Demonstração.* Ver [28, Teorema 1 do capítulo I]. □

A expressão

$$\|M\| = \sup \{\|Mx\|_{\mathbb{R}^m}; x \in \mathbb{R}^n, \|x\|_{\mathbb{R}^n} = 1\}$$

define uma norma no espaço vetorial das matrizes  $M \in \mathbb{R}^{m \times n}$ , onde  $\|\cdot\|_{\mathbb{R}^m}$  e  $\|\cdot\|_{\mathbb{R}^n}$  são normas quaisquer de  $\mathbb{R}^m$  e  $\mathbb{R}^n$ , respectivamente. Temos que a desigualdade

$$\|Mx\|_{\mathbb{R}^m} \leq \|M\| \|x\|_{\mathbb{R}^n}$$

vale para todo  $x \in \mathbb{R}^n$ . Denotaremos por  $\|M\|_\infty$  a norma obtida quando escolhermos as normas do infinito de  $\mathbb{R}^m$  e  $\mathbb{R}^n$ . É possível mostrar que

$$\|M\|_\infty = \max_i \left( \sum_j |m_{ij}| \right),$$

onde  $M = [m_{ij}]$ .

Dada uma norma  $\|\cdot\|_{\mathbb{R}^n}$  em  $\mathbb{R}^n$  e  $x, y \in \mathbb{R}^n$ , a *distância* entre  $x$  e  $y$  com respeito à norma  $\|\cdot\|_{\mathbb{R}^n}$  é dada por

$$d(x, y) = \|x - y\|_{\mathbb{R}^n}.$$

Agora, fixemos uma norma  $\|\cdot\|_{\mathbb{R}^n}$  em  $\mathbb{R}^n$  e tomemos um conjunto  $X \subset \mathbb{R}^n$ . Dizemos que  $a \in X$  é *ponto interior* a  $X$  quando existe  $\delta > 0$  tal que  $\|x - a\|_{\mathbb{R}^n} < \delta \Rightarrow x \in X$ . O conjunto dos pontos interiores a  $X$  é denotado por  $\text{int } X$ . O conjunto  $X$  é dito *aberto* quando  $X = \text{int } X$ , isto é, quando todos seus pontos são interiores. Se  $\bar{x} \in \text{int } V$ , dizemos que  $V$  é uma *vizinhança* de  $\bar{x}$ . Em particular, o conjunto

$$B_\epsilon(\bar{x}) = \{y; \|\bar{x} - y\|_{\mathbb{R}^n} < \epsilon\} \subset \mathbb{R}^n,$$

com  $\epsilon > 0$ , é uma vizinhança aberta de  $\bar{x}$ , chamada *bola aberta* definida pela norma  $\|\cdot\|_{\mathbb{R}^n}$ ,

de centro em  $\bar{x}$  e raio  $\epsilon$ . A bola aberta

$$B_\epsilon^\infty(\bar{x}) = \{y; \|\bar{x} - y\|_\infty < \epsilon\}$$

definida pela norma do infinito é o produto cartesiano dos intervalos  $(x_i - \epsilon, x_i + \epsilon)$ , isto é,

$$B_\epsilon^\infty(\bar{x}) = (x_1 - \epsilon, x_1 + \epsilon) \times \cdots \times (x_n - \epsilon, x_n + \epsilon).$$

Em geral, temos

$$B_\epsilon^\infty(\bar{x}, \bar{y}) = B_\epsilon^\infty(\bar{x}) \times B_\epsilon^\infty(\bar{y}).$$

Isto torna a norma do infinito mais conveniente em relação ao produto cartesiano.

A *fronteira*  $\text{fr } X$  do conjunto  $X \subset \mathbb{R}^n$  é o conjunto dos pontos  $x \in \mathbb{R}^n$  com a seguinte propriedade: dada qualquer vizinhança aberta  $V(x)$  de  $x$ , existem  $y_1, y_2 \in V(x)$  tais que  $y_1 \in X$  e  $y_2 \in \complement X = \mathbb{R}^n \setminus X$ . O conjunto  $X$  é dito *fechado* se  $\text{fr } X \subset X$ , isto é, se contém sua fronteira. Vale uma outra caracterização:  $X$  é fechado se, e somente se, seu complementar  $\complement X$  é aberto. Dizemos que o conjunto  $\text{fr } X \cup X$  é o *fecho* de  $X$ . Claramente o fecho de qualquer conjunto é fechado.

Algumas características de conjuntos abertos e fechados:

**Teorema A.2.3.** *i)  $\emptyset$  e  $\mathbb{R}^n$  são os únicos subconjuntos de  $\mathbb{R}^n$  simultaneamente abertos e fechados;*

*ii) A interseção  $A = A_1 \cap \cdots \cap A_k$  de um número finito de conjuntos abertos  $A_i$  é um conjunto aberto;*

*iii) A união  $A = \bigcup_{\beta \in L} A_\beta$  de uma família qualquer  $(A_\beta)_{\beta \in L}$  de conjuntos abertos  $A_\beta$  é um conjunto aberto;*

*iv) A união  $F = F_1 \cup \cdots \cup F_k$  de um número finito de conjuntos fechados  $F_i$  é um conjunto fechado;*

*v) A interseção  $F = \bigcap_{\beta \in L} F_\beta$  de uma família qualquer  $(F_\beta)_{\beta \in L}$  de conjuntos fechados  $F_\beta$  é um conjunto fechado.*

*Demonstração.* Ver [28, Teoremas 15 e 18 do capítulo I]. □

Dizemos que  $X \subset \mathbb{R}^n$  é *limitado* quando existe  $M > 0$  tal que  $\|x\|_{\mathbb{R}^n} < M, \forall x \in X$ , onde  $\|\cdot\|_{\mathbb{R}^n}$  é uma norma em  $\mathbb{R}^n$ . Se, além de ser limitado,  $X$  for também fechado, dizemos que  $X$  é *compacto*.

Um importante resultado é o

**Teorema A.2.4** (de Weierstrass). *Toda função real contínua  $f : K \rightarrow \mathbb{R}$ , definida num compacto  $K \subset \mathbb{R}^n$ , atinge seu máximo e seu mínimo em  $K$ , isto é, existem pontos  $x_0, x_1 \in K$  tais que  $f(x_0) \leq f(x) \leq f(x_1), \forall x \in K$ .*

*Demonstração.* Ver [28, Corolário 1 do Teorema 20 do capítulo I]. □

A distância entre dois conjuntos não vazios  $X, Y \subset \mathbb{R}^n$  com respeito à uma norma  $\|\cdot\|_{\mathbb{R}^n}$  de  $\mathbb{R}^n$  é definida como

$$d(X, Y) = \inf \{\|x - y\|_{\mathbb{R}^n}; x \in X, y \in Y\}.$$

Quando um conjunto se reduz a um ponto, digamos  $X = \{x\}$ , temos a distância entre  $x$  e  $Y$  sendo

$$d(x, Y) = \inf \{\|x - y\|_{\mathbb{R}^n}; y \in Y\}.$$

O resultado a seguir diz que o ínfimo da distância entre um conjunto compacto e outro fechado é atingido.

**Teorema A.2.5.** *Seja  $K \subset \mathbb{R}^n$  compacto e não vazio,  $F \subset \mathbb{R}^n$  fechado e não vazio, e  $\|\cdot\|_{\mathbb{R}^n}$  uma norma em  $\mathbb{R}^n$ . Então existem  $x_0 \in K$  e  $y_0 \in F$  tais que  $d(K, F) = \|x_0 - y_0\|_{\mathbb{R}^n}$ . Em particular, se  $K \cap F = \emptyset$  então  $d(K, F) > 0$ .*

*Demonstração.* Ver [28, Teorema 26 do capítulo I]. □

## A.3 Poliedros, programação linear e dualidade

Abordaremos alguns conceitos sobre análise convexa, programação linear e dualidade. Boas referências nesses assuntos são os livros de Dimitri P. Bertsekas [5], de Mokhtar S. Bazaraa e John J. Jarvis [4], de Alexey Izmailov e Mikhail Solodov [24], e de Alexander Schrijver [39].

Um conjunto  $D \subset \mathbb{R}^n$  é dito *convexo* se para cada par  $z_1, z_2 \in D$ , tem-se que  $\alpha z_1 + (1 - \alpha)z_2 \in D$  para todo  $\alpha \in [0, 1]$ . Dizemos que  $z = \alpha z_1 + (1 - \alpha)z_2$  com  $\alpha \in [0, 1]$  é uma *combinação convexa* de  $z_1$  e  $z_2$ . Um conjunto da forma  $\{z; z = \alpha z_0, \alpha \in \mathbb{R}\}$  com  $z_0 \neq 0$  é um conjunto convexo chamado *reta*. Dado qualquer conjunto  $E \subset \mathbb{R}^n$ , o *fecho convexo* de  $E$ , denotado por  $\text{conv } E$ , é o menor conjunto convexo que contém  $E$ . Equivalentemente,  $\text{conv } E$  é a interseção de todos os conjuntos convexos em  $\mathbb{R}^n$  que

contêm  $E$  (a interseção de uma família qualquer de conjuntos convexos é convexa - veja [5, Proposição 1.2.1(a)]).

Dado um conjunto convexo  $D$ ,  $z \in D$  é dito *ponto extremo* de  $D$  se não pode ser representado por uma combinação convexa de dois pontos distintos de  $z$  em  $D$ , isto é, não existem  $z_1, z_2 \in D \setminus \{z\}$  e  $\alpha \in (0, 1)$  tais que  $z = \alpha z_1 + (1 - \alpha)z_2$ . O próximo resultado relaciona pontos extremos com retas num conjunto convexo, fechado e não vazio.

**Teorema A.3.1.** *Seja  $D \subset \mathbb{R}^n$  um conjunto convexo, fechado e não vazio. Então  $D$  tem pontos extremos se, e somente se,  $D$  não contém qualquer reta.*

*Demonstração.* Ver [24, Proposição 3.2.10] ou [5, Proposição 3.3.1(b)].  $\square$

Um conjunto convexo  $C$  tal que  $\alpha z \in C$  para todos  $z \in C$  e  $\alpha \geq 0$  é dito *cone convexo*. Notemos que um cone convexo contém a origem (basta fazer  $\alpha = 0$ ). Dado qualquer conjunto  $F \subset \mathbb{R}^n$ , o *fecho cônico* de  $F$ , denotado por  $\text{cone } F$ , é o menor cone convexo que contém  $F$ . Equivalentemente,  $\text{cone } F$  é a interseção de todos os cones convexos em  $\mathbb{R}^n$  que contêm  $F$ .

Dados  $z_1, \dots, z_K \in \mathbb{R}^n$ , ao cone convexo

$$\mathcal{R} = \text{cone} \{z_i; i = 1, \dots, K\} = \left\{ \sum_{k=1}^K \alpha^k z_k, \alpha^k \geq 0 \right\}$$

formado por todas as combinações não negativas de  $z_1, \dots, z_K$  chamamos de *cone finitamente gerado* por  $z_1, \dots, z_K$ .

Uma importante classe de conjuntos convexos são os *poliedros*, cuja forma geral é

$$P = \{z \in \mathbb{R}^n; Az \leq b\}$$

onde  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ . Dizemos também que  $P$  é um conjunto *poliedral*. Por vezes é comum considerarmos restrições de não negatividade  $z \geq 0$ . Observamos também que restrições de igualdade  $az = c$  podem ser postas na forma anterior considerando as desigualdades  $az \leq c$  e  $az \geq c$  em  $Az \leq b$ .

**Teorema A.3.2** (da representação de Minkowski-Weyl). *Um conjunto  $P$  é poliedral se, e somente se, existem um conjunto finito não vazio  $E = \{z_1, \dots, z_K\} \subset \mathbb{R}^n$  e um cone finitamente gerado  $\mathcal{R}$  tais que  $P = \text{conv } E + \mathcal{R}$ , isto é,*

$$P = \left\{ z \in \mathbb{R}^n; z = \sum_{k=1}^K a^k z_k + d, \sum_{k=1}^K a^k = 1, a^k \in \mathbb{R}_+, d \in \mathcal{R} \right\}.$$

*Demonstração.* Ver [5, Proposição 3.2.2] ou [4, Teorema 2 da seção 2.6].  $\square$

Um ponto extremo de um conjunto poliedral é dito *vértice*. Tal vértice é caracterizado pelo clássico resultado:

**Teorema A.3.3.** *Seja  $\bar{z} \in D = \{z \in \mathbb{R}^n; Az \leq b\}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ . Então  $\bar{z}$  é vértice de  $D$  se, e somente se, a matriz  $[A]_{i \in I(\bar{z})}$ ,  $I(\bar{z}) = \{i; A_i \bar{z} = b_i\}$ , das restrições ativas em  $\bar{z}$  tem posto  $n$ .*

*Demonstração.* Ver [24, Proposição 3.2.11].  $\square$

Seja  $D \subset \mathbb{R}^n$  um conjunto convexo. Uma função  $f : D \rightarrow \mathbb{R}$  é dita *convexa* se para quaisquer  $z_1, z_2 \in D$  e  $\alpha \in [0, 1]$  tem-se  $f(\alpha z_1 + (1 - \alpha)z_2) \leq \alpha f(z_1) + (1 - \alpha)f(z_2)$ . A função  $f$  é dita *côncava* se  $-f$  é convexa. Um ponto  $\bar{z}$  é dito *viável* para o problema

$$\min f(z) \quad \text{s.a.} \quad z \in D \quad (\text{A.1})$$

se  $\bar{z} \in D$ . É dito *solução ótima* de (A.1) se, além de  $\bar{z} \in D$ , tivermos  $f(\bar{z}) \leq f(z)$ ,  $\forall z \in D$  (no caso de maximização,  $f(\bar{z}) \geq f(z)$ ,  $\forall z \in D$ ). Se  $\bar{z}$  é solução ótima de (A.1), dizemos que  $f(\bar{z})$  é o *valor ótimo* de (A.1). A seguir um importante resultado para problemas de minimização de funções convexas.

**Teorema A.3.4.** *Sejam  $D \subset \mathbb{R}^n$  poliedral que possua vértice e  $f : D \rightarrow \mathbb{R}$  convexa. Suponha que*

$$\max f(z) \quad \text{s.a.} \quad z \in D$$

*tenha solução ótima. Então existe uma solução ótima deste problema que é um vértice de  $D$ .*

*Demonstração.* Ver [4, Teorema 3 da seção 3.2] ou [24, Teorema 3.4.6], tendo em vista o Teorema A.3.1.  $\square$

A afirmação do Teorema anterior não é verdadeira em geral se trocarmos maximização por minimização. Mas quando  $f$  é linear, a afirmação vale para minimização. Isto porque  $-f$  é convexa.

Quando  $f$  é linear e  $D$  poliedral, chamamos o problema (A.1) de *problema de programação linear*. Esses problemas possuem a seguinte propriedade de estabilidade:

**Teorema A.3.5.** *Seja*

$$\xi(b) = \min_z \{cz; Az \leq b\}$$

com  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ , um problema de programação linear, e seja  $\Delta$  tal que, para cada submatriz inversível  $B$  de  $A$ , os módulos das entradas de  $B^{-1}$  não ultrapassem  $\Delta$ . Em outras palavras,  $\Delta$  é tal que

$$\Delta \geq \{\max_{i,j} |b_{ij}|; B^{-1} = [b_{ij}], B \text{ submatriz inversível de } A\}.$$

Sejam ainda  $b_1, b_2 \in \mathbb{R}^m$  tais que  $\xi(b_1), \xi(b_2) > -\infty$ . Então para cada solução ótima  $z_1 \in \mathbb{R}^n$  de  $\xi(b_1)$ , existe uma solução ótima  $z_2 \in \mathbb{R}^n$  de  $\xi(b_2)$  tal que

$$\|z_1 - z_2\|_\infty \leq n\Delta \|b_1 - b_2\|_\infty.$$

*Demonstração.* Ver [39, Teorema 10.5]. □

A demonstração do Teorema anterior não requer que  $\xi$  seja contínua, nem que a função ponto a conjunto que mapeia as soluções ótimas do problema seja semicontínua inferior (veja seção A.4 para definições). Por isso, o resultado é usado na demonstração da semicontinuidade inferior de  $S$  no Lema 3.1.5 (página 31).

Agora, dado o problema de programação linear

$$\min cz \quad \text{s.a.} \quad Az \geq b, \quad z \geq 0, \tag{A.2}$$

definimos o seu problema *dual* como sendo

$$\max b^t \lambda \quad \text{s.a.} \quad A^t \lambda \leq c^t, \quad \lambda \geq 0. \tag{A.3}$$

Chamamos o problema (A.2) de problema *primal*. O dual do dual é seu problema primal (veja [4, Lema 1 da seção 6.1]). Existe exatamente uma variável dual para cada restrição primal e exatamente uma restrição dual para cada variável primal. Todo problema de programação linear pode ser posto na forma (A.2). É possível no entanto derivar de (A.2) e seu dual, os duais de problemas de programação linear em outras formas. A Tabela 16 resume o comportamento do dual para outras formas de problemas primais.

	Minimização		Maximização	
	$\geq 0$	$\leftrightarrow$	$\leq$	
variáveis	$\leq 0$	$\leftrightarrow$	$\geq$	restrições
	irrestrito	$\leftrightarrow$	=	
	$\geq$	$\leftrightarrow$	$\geq 0$	
restrições	$\leq$	$\leftrightarrow$	$\leq 0$	variáveis
	=	$\leftrightarrow$	irrestrito	

Tabela 16: Relações entre problemas primal e dual.

Assim o problema parametrizado em  $x$

$$\min d_2 y \quad \text{s.a.} \quad B_2 y \leq b_2 - A_2 x, \quad y \geq 0$$

do Capítulo 3 tem como dual

$$\max (b_2 - A_2 x)^t \lambda \quad \text{s.a.} \quad B_2^t \lambda \leq d_2^t, \quad \lambda \leq 0,$$

que é equivalente ao problema

$$\max (A_2 x - b_2)^t \lambda \quad \text{s.a.} \quad B_2^t \lambda + d_2^t \geq 0, \quad \lambda \geq 0.$$

A importância do dual fica evidente no próximo resultado.

**Teorema A.3.6** (Dualidade em programação linear). *Dado um problema de programação linear primal, consideramos seu dual. Uma, e somente uma das alternativas abaixo ocorre:*

- i) Ambos os problemas têm conjunto viável vazio (são inviáveis);*
- ii) Um problema é ilimitado. Neste caso o outro é inviável;*
- iii) Um problema tem solução. Neste caso o outro também tem, e seus valores ótimos coincidem.*

*Demonstração.* Ver [24, Teorema 5.1.2] ou [4, Teorema 1 da seção 6.2]. □

Dado um problema de programação linear e seu dual, podemos derivar condições de otimalidade através do Teorema anterior. Especificamente,  $z$  é solução ótima do problema (A.2) se, e somente se, seu problema dual (A.3) tem solução ótima  $\lambda$  (neste caso,  $b^t \lambda = cz$ ).

Fazendo alguns cálculos (veja [4, Seção 6.2]), temos que  $z$  é solução ótima do problema primal (A.2) se, e somente se, existir  $\lambda$  tal que

$$\begin{aligned} Az &\geq b, \quad z \geq 0, \quad \lambda \geq 0, \quad \lambda^t(Az - b) = 0, \\ A^t\lambda &\leq c^t, \quad z^t(A^t\lambda - c^t) = 0. \end{aligned} \tag{A.4}$$

As restrições não lineares em (A.4) são chamadas restrições de *folga complementar*. É desta forma que chegamos ao sistema (3.1) do Capítulo 3 (página 29). O sistema (A.4) é também chamado de *condições de Kuhn-Tucker*.

## A.4 Funções ponto a conjunto

Discutiremos o conceito de continuidade de funções ponto a conjunto. Uma análise completa sobre o assunto é feita no livro de Jean-Pierre Aubin e Hélène Frankowska [2].

Uma *função ponto a conjunto*<sup>1</sup> é uma relação  $\Gamma$  que leva pontos de um conjunto  $Z$  a subconjuntos de um conjunto  $Y$ . Denotaremos por  $2^Y$  o conjunto dos subconjuntos de  $Y$ . Neste trabalho usamos em geral a função ponto a conjunto

$$\Gamma : Z \subset \mathbb{R}^p \rightarrow 2^{\mathbb{R}^q}.$$

Ao conjunto  $\text{dom } \Gamma = \{z \in Z; \Gamma(z) \neq \emptyset\}$  damos o nome de *domínio* de  $\Gamma$ . A *imagem inversa* de um conjunto  $M$  por  $\Gamma$  é o conjunto

$$\Gamma^{-1}(M) = \{z \in Z; \Gamma(z) \cap M \neq \emptyset\}.$$

Há ainda outra maneira de analisar a imagem inversa de uma função ponto a conjunto, que não usamos neste trabalho: pelo *núcleo* de  $M$  por  $\Gamma$  definido por  $\Gamma^{+1}(M) = \{z \in Z; \Gamma(z) \subset M\}$ .

A noção de continuidade de funções ponto a conjunto claramente não pode ser a mesma que a de funções com imagens de um único elemento. No entanto, deve englobar o último caso, que pode ser visto como o de funções ponto a conjunto onde as imagens são conjuntos de cardinalidade 1.

**Definição A.4.1.** *Uma função ponto a conjunto  $\Gamma : Z \subset \mathbb{R}^p \rightarrow 2^{\mathbb{R}^q}$  é dita semicontínua superior em  $\bar{z} \in Z$  se para cada aberto  $A \supset \Gamma(\bar{z})$  existir um aberto  $V \ni \bar{z}$  tal que  $\Gamma(z) \subset A$  para todo  $z \in V \cap Z$ .*

<sup>1</sup>Em inglês, *set-valued map* ou *multifunction*.

$\Gamma$  é dita semicontínua inferior em  $\bar{z} \in Z$  se para cada aberto  $A \subset \mathbb{R}^q$  com  $\Gamma(\bar{z}) \cap A \neq \emptyset$  existir um aberto  $V \ni \bar{z}$  tal que  $\Gamma(z) \cap A \neq \emptyset$  para todo  $z \in V \cap Z$ .

$\Gamma$  é dita contínua em  $\bar{z} \in Z$  se é semicontínua inferior e superior em  $\bar{z}$ .

Dentre os inúmeros resultados da vasta teoria de análise de funções ponto a conjunto, destacamos a seguir os usados neste trabalho.

**Teorema A.4.2.** *Uma função ponto a conjunto  $\Gamma : Z \subset \mathbb{R}^p \rightarrow 2^{\mathbb{R}^q}$  é semicontínua inferior em  $z \in Z$  se, e somente se, para cada  $w \in \Gamma(z)$  e para cada sequência  $\{z_n\}$  de elementos em  $\text{dom } \Gamma$  convergindo a  $z$ , existe uma sequência  $\{w_n\}$  convergindo a  $w$  tal que  $w_n \in \Gamma(z_n), \forall n$ .*

*Demonstração.* Ver [2, Observação da Definição 1.4.2]. □

**Teorema A.4.3.** *Seja  $\Gamma : Z \subset \mathbb{R}^p \rightarrow 2^{\mathbb{R}^q}$  uma função ponto a conjunto com  $\text{dom } \Gamma$  fechado. Temos que  $\Gamma$  é semicontínua superior em cada  $z \in \text{dom } \Gamma$  se, e somente se, a imagem inversa de qualquer conjunto fechado é um conjunto fechado.*

*Demonstração.* Ver [2, Proposição 1.4.4]. □

## APÊNDICE B – Problemas multinível

Neste apêndice, estenderemos alguns resultados de problemas de dois níveis feitos no Capítulo 2 a problemas com vários níveis.

Suponha que tenhamos  $k \geq 2$  níveis, cujas respectivas variáveis são  $x_1, \dots, x_k$  ( $x_i \in \mathbb{R}^{n_i}$ ). Para o nível  $i$ , seja  $F^i(x_1, \dots, x_k)$  sua função objetivo e  $\Omega^i(x_1, \dots, x_{i-1})$  suas restrições (entendendo que  $\Omega^1$  do primeiro nível não é parametrizado por qualquer variável  $x_j$ ). O penúltimo nível  $k - 1$  controla a variável  $x_{k-1}$ , tem como resposta do nível abaixo a variável  $x_k$  e é parametrizado pelas variáveis  $x_1, \dots, x_{k-2}$  dos níveis acima. Seu modelo é

$$\begin{aligned} \min_{x_{k-1}, x_k} F^{k-1}[1, k-2](x_{k-1}, x_k) \\ \text{s.a. } (x_{k-1}, x_k) \in \Omega^{k-1}[1, k-2] \\ x_k \in \mathfrak{S}^{k-1}[1, k-2](x_{k-1}) \subset S^{k-1}[1, k-2](x_{k-1}) \end{aligned}$$

onde

$$S^{k-1}[1, k-2](x_{k-1}) = \arg \min_{x_k} \{F^k[1, k-1](x_k); x_k \in \Omega^k[1, k-1]\}$$

e  $\mathfrak{S}^{k-1}[1, k-2]$  é a política adotada pelo nível  $k$  ao entregar suas soluções ótimas ao nível  $k-1$ . A notação  $\cdot [1, i]$  explicita as variáveis  $x_1, \dots, x_i$  de níveis acima que parametrizam todo o problema do nível  $i+1$ . Quando  $i=1$ ,  $[1, i-1]$  significará que a estrutura em questão não é parametrizada. Vale notarmos que o problema do último nível  $k$  é o descrito em  $S^{k-1}[1, k-2](x_{k-1})$ .

Da mesma forma, o nível  $k-2$  é dado por

$$\begin{aligned} \min_{x_{k-2}, x_{k-1}, x_k} F^{k-2}[1, k-3](x_{k-2}, x_{k-1}, x_k) \\ \text{s.a. } (x_{k-2}, x_{k-1}, x_k) \in \Omega^{k-2}[1, k-3] \\ (x_{k-1}, x_k) \in \mathfrak{S}^{k-2}[1, k-3](x_{k-2}) \subset S^{k-2}[1, k-3](x_{k-2}) \end{aligned}$$

onde

$$S^{k-2}[1, k-3](x_{k-2}) = \arg \min_{(x_{k-1}, x_k)} \{F^{k-1}[1, k-2](x_{k-1}, x_k);$$

$$(x_{k-1}, x_k) \in \Omega^{k-1}[1, k-2], x_k \in \mathfrak{S}^{k-1}[1, k-2](x_{k-1})\}$$

e  $\mathfrak{S}^{k-2}[1, k-3]$  é política adotada pelo nível  $k-1$  ao entregar suas soluções ótimas ao nível  $k-2$ . A restrição  $x_k \in \mathfrak{S}^{k-1}[1, k-2](x_{k-1})$  na expressão de  $S^{k-2}[1, k-3](x_{k-2})$  é necessária pois é restrição do problema do nível  $k-1$ . Indutivamente, o problema do nível  $i \in \{k-2, \dots, 1\}$  é dado por

$$\min_{x_i, \dots, x_k} F^i[1, i-1](x_i, \dots, x_k)$$

$$\text{s.a. } (x_i, \dots, x_k) \in \Omega^i[1, i-1]$$

$$(x_{i+1}, \dots, x_k) \in \mathfrak{S}^i[1, i-1](x_i) \subset S^i[1, i-1](x_i)$$

onde

$$S^i[1, i-1](x_i) = \arg \min_{(x_{i+1}, \dots, x_k)} \{F^{i+1}[1, i](x_{i+1}, \dots, x_k);$$

$$(x_{i+1}, \dots, x_k) \in \Omega^{i+1}[1, i], (x_{i+2}, \dots, x_k) \in \mathfrak{S}^{i+1}[1, i](x_{i+1})\}.$$

Com o que foi visto, formulamos o *problema de  $k$  níveis* como

$$\min_{x_1, \dots, x_k} F^1(x_1, \dots, x_k)$$

$$\text{s.a. } (x_1, \dots, x_k) \in \Omega^1$$

$$(x_2, \dots, x_k) \in \mathfrak{S}^1(x_1) \subset S^1(x_1)$$

$$\vdots$$

$$(x_{k-1}, x_k) \in \mathfrak{S}^{k-2}[1, k-3](x_{k-2}) \subset S^{k-2}[1, k-3](x_{k-2})$$

$$x_k \in \mathfrak{S}^{k-1}[1, k-2](x_{k-1}) \subset S^{k-1}[1, k-2](x_{k-1}).$$

O modelo acima pode ser simplificado usando o próximo resultado:

**Lema B.0.4.**  $(x_2, \dots, x_k) \in S^1(x_1) \Rightarrow (x_{i+1}, \dots, x_k) \in \mathfrak{S}^i[1, i-1](x_i), \forall i \in \{2, \dots, k-1\}$ .

*Demonstração.*  $(x_2, \dots, x_k) \in S^1(x_1) \Rightarrow (x_3, \dots, x_k) \in \mathfrak{S}^2[1, 1](x_2) \subset S^2[1, 1](x_2) \Rightarrow \dots \Rightarrow x_k \in \mathfrak{S}^{k-1}[1, k-2](x_{k-1}) \subset S^{k-1}[1, k-2](x_{k-1})$ .  $\square$

Podemos então escrever o problema de  $k$  níveis como

$$\begin{aligned} \min_{x_1, \dots, x_k} F^1(x_1, \dots, x_k) \\ \text{s.a. } (x_1, \dots, x_k) \in \Omega^1 \\ (x_2, \dots, x_k) \in \mathfrak{S}^1(x_1) \subset S^1(x_1). \end{aligned} \quad (\text{B.1})$$

Este modelo é um problema de dois níveis, onde o primeiro nível se mantém, e o segundo nível é o problema de  $k - 1$  níveis formado pelos  $k - 1$  últimos níveis. No caso,  $S^1(x_1)$  é o conjunto das soluções ótimas globais deste problema de  $k - 1$  níveis, parametrizado em  $x_1$ .

Consideremos agora o caso em que cada nível coopera com seu nível acima, isto é, para cada  $i = k - 1, \dots, 1$ , o nível  $i + 1$  retorna as melhores soluções ótimas do ponto de vista do nível  $i$ . Em outras palavras, fazemos indutivamente

$$\begin{aligned} \mathfrak{S}^i[1, i - 1](x_i) = \arg \min_{(x_{i+1}, \dots, x_k)} \{F^i[1, i](x_i, \dots, x_k); \\ (x_i, \dots, x_k) \in \Omega^i[1, i - 1], (x_{i+1}, \dots, x_k) \in S^i[1, i - 1](x_i)\} \end{aligned}$$

para  $i = k - 1, \dots, 1$ . Do mesmo modo do Capítulo 2, consideramos o modelo

$$\min_{x_1} \varphi_o(x_1) \quad \text{s.a.} \quad x_1 \in \text{dom } \varphi_o \quad (\text{B.2})$$

onde

$$\begin{aligned} \varphi_o(x_1) = \min_{x_2, \dots, x_k} \{F^1(x_1, \dots, x_k); (x_1, \dots, x_k) \in \Omega^1, (x_2, \dots, x_k) \in S^1(x_1)\}, \quad (\text{B.3}) \\ \text{dom } \varphi_o = \{x_1 \in \mathbb{R}^{n_1}; \varphi_o(x_1) > -\infty\} \text{ e} \\ \Omega_x = \{x_1; S^1(x_1) \neq \emptyset\}. \end{aligned}$$

Como (B.1) é um problema de dois níveis, é equivalente ao modelo (B.2)-(B.3) (Teorema 2.0.2, página 19). Temos então a

**Definição B.0.5.** *Um ponto  $(\bar{x}_1, \dots, \bar{x}_k)$  é dito solução local  $k$ -otimista de (B.1) se*

$$(\bar{x}_2, \dots, \bar{x}_k) \in S^1(x_1), \quad (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k) \in \Omega^1, \quad F(\bar{x}_1, \dots, \bar{x}_k) = \varphi_o(\bar{x}_1)$$

*e existir uma vizinhança aberta  $V(\bar{x}_1)$  de  $\bar{x}_1$  tal que*

$$\varphi_o(\bar{x}_1) \leq \varphi_o(x_1), \quad \forall x_1 \in V(\bar{x}_1) \cap \text{dom } \varphi_o$$

*Dizemos que  $(\bar{x}_1, \dots, \bar{x}_k)$  é solução global  $k$ -otimista de (B.1) se puder ser  $V(\bar{x}_1) = \mathbb{R}^{n_1}$ .*

Considere agora o modelo MLP<sup>1</sup>, análogo do modelo BLP:

$$\begin{aligned}
 \text{MLP: } \min_{x_1, \dots, x_k} F^1(x_1, \dots, x_k) \\
 \text{s.a. } (x_1, \dots, x_k) \in \Omega^1 \\
 (x_2, \dots, x_k) \in \arg \min F^2(x_2, \dots, x_k) \\
 \text{s.a. } (x_2, \dots, x_k) \in \Omega^2 \\
 \vdots \\
 x_k \in \arg \min F^k(x_k) \\
 \text{s.a. } x_k \in \Omega^k.
 \end{aligned}$$

MLP corresponde a fazer indutivamente, para todo  $(x_1, \dots, x_{i-1})$ ,  $\mathfrak{S}^i[1, i-1] \equiv S^i[1, i-1]$ ,  $i = k-1, \dots, 1$ . A fim de relacionar MLP com o modelo (B.1), notemos que para qualquer  $(x_1, \dots, x_{k-2})$  fixado,  $S^{k-2}[1, k-3](x_{k-2})$  é o conjunto das soluções globais otimizadas  $(x_{k-1}, x_k)$  do problema de dois níveis formado pelos níveis  $k-1$  e  $k$ . Portanto, pelos itens (iii) e (iv) do Teorema 2.0.3 (página 21), podemos fazer  $\mathfrak{S}^{k-1}[1, k-2] \equiv S^{k-1}[1, k-2]$  para todo  $(x_1, \dots, x_{k-2})$  (observamos que o problema de dois níveis formado pelos níveis  $k-1$  e  $k$  deve ter solução ótima para uma boa definição de  $\mathfrak{S}^{k-1}[1, k-2]$ ). Da mesma forma, para qualquer  $(x_1, \dots, x_{k-3})$  fixado,  $S^{k-3}[1, k-4](x_{k-3})$  é o conjunto das soluções globais otimizadas  $(x_{k-2}, x_{k-1}, x_k)$  do problema de dois níveis

$$\begin{aligned}
 \min_{x_{k-2}, \dots, x_k} F^{k-2}[1, k-3](x_{k-2}, \dots, x_k) \\
 \text{s.a. } (x_{k-2}, \dots, x_k) \in \Omega^{k-2}[1, k-3] \\
 (x_{k-1}, x_k) \in \mathfrak{S}^{k-2}[1, k-3](x_{k-2}) \subset S^{k-2}[1, k-3](x_{k-2})
 \end{aligned}$$

donde segue novamente do Teorema 2.0.3 que podemos fazer  $\mathfrak{S}^{k-2}[1, k-3] \equiv S^{k-2}[1, k-3]$  (novamente o problema acima tem solução ótima). Usando o mesmo raciocínio, obtemos indutivamente que ao fazer  $\mathfrak{S}^i[1, i-1] \equiv S^i[1, i-1]$  para todo  $(x_1, \dots, x_{i-1})$  e para  $i = k-1, \dots, 2$ , não modificamos soluções (locais e globais)  $k$ -otimizadas de (B.1). A diferença entre os modelos MLP e (B.1) fica portanto somente na descrição de  $\mathfrak{S}^1$ . Levando em consideração que (B.1) é um problema de dois níveis, aplicamos o Teorema 2.0.3 para obter o

**Teorema B.0.6.** *i) Se MLP é inviável então (B.1)  $k$ -otimista é inviável;*

*ii) Se (B.1)  $k$ -otimista é ilimitado então MLP é ilimitado;*

---

<sup>1</sup>Do inglês *Multilevel Problem*.

- iii) Suponhamos que MLP tenha solução ótima. Temos que se  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_k)$  é solução local  $k$ -otimista (respectivamente global  $k$ -otimista) de (B.1) então  $\bar{x}$  é solução ótima local (respectivamente global) de MLP;
- iv) Se  $\bar{x}$  é solução ótima global de MLP então  $\bar{x}$  é solução global  $k$ -otimista de (B.1).

## APÊNDICE C – Parâmetros dos problemas-teste usados

Neste apêndice relatamos os parâmetros de cada grupo de problemas-teste utilizados na subseção 6.3.3 (página 92). É importante observarmos que todos os problemas foram gerados apenas com restrições de baixo nível, e que nas tabelas que se seguem são apresentados os dados dos problemas após a aplicação do pré-condicionador discutido na subseção 6.3.3. A quantidade total de restrições ( $m_u + m_l$ ) varia entre os problemas pois restrições redundantes foram descartadas na geração dos problemas ou na aplicação do pré-condicionador. Relatamos também o número aproximado de soluções ótimas locais/não globais<sup>1</sup> (coluna “Locais”).

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	25	40	38	53	0	3	12	6	4	8,18E+03	0,593764	1
2	25	40	39	54	11	0	0	5	9	1,64E+04	0,583646	1
3	25	40	36	51	6	0	8	8	3	2,05E+03	0,567926	1
4	25	40	49	64	1	16	0	0	8	1,67E+07	0,570833	1
5	25	40	48	63	0	4	2	18	1	8,39E+06	0,572732	1
6	25	40	40	55	0	4	10	11	0	3,28E+04	0,575014	1
7	25	40	36	51	14	8	0	3	0	1,79E+03	0,597629	1
8	25	40	34	49	4	0	12	9	0	5,11E+02	0,562758	1
9	25	40	50	65	0	9	0	8	8	3,36E+07	0,580064	1
10	25	40	38	53	0	0	12	7	6	8,19E+03	0,596106	1

Tabela 17: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 25$  e  $n_y = 40$ .

<sup>1</sup>Quando tal número for zero, é porque os problemas em questão realmente não têm soluções ótimas locais/não globais.

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	75	35	96	56	1	0	13	21	0	2,10E+06	0,578889	1
2	75	35	93	53	14	0	3	16	2	2,62E+05	0,588755	1
3	75	35	90	50	16	0	4	0	15	3,28E+04	0,555153	1
4	75	35	98	58	7	0	5	11	12	8,39E+06	0,565294	1
5	75	35	89	49	5	7	16	3	4	1,63E+04	0,593232	1
6	75	35	100	60	8	8	2	0	17	3,36E+07	0,575497	1
7	75	35	88	48	22	0	0	1	12	8,19E+03	0,551136	1
8	75	35	78	38	22	0	10	3	0	7,00E+00	0,593174	1
9	75	35	95	55	4	0	11	13	7	1,05E+06	0,571894	1
10	75	35	93	53	3	2	14	2	14	2,62E+05	0,595753	1

Tabela 18: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 75$  e  $n_y = 35$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	75	90	111	126	36	0	3	0	36	6,87E+10	0,586451	1
2	75	90	110	125	0	32	40	0	3	3,01E+10	0,582652	1
3	75	90	95	110	32	0	23	0	20	1,05E+06	0,579506	1
4	75	90	87	102	25	0	38	0	12	4,10E+03	0,580922	1
5	75	90	112	127	34	32	4	5	0	1,33E+11	0,558337	1
6	75	90	147	162	3	30	0	35	7	4,72E+21	0,579024	1
7	75	90	120	135	0	5	30	40	0	3,52E+13	0,552424	1
8	75	90	120	135	30	16	0	24	5	3,52E+13	0,591463	1
9	75	90	141	156	9	20	0	31	15	7,38E+19	0,553942	1
10	75	90	101	116	3	2	46	11	13	6,71E+07	0,584279	1

Tabela 19: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 75$  e  $n_y = 90$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	100	70	154	124	3	14	13	0	40	1,80E+16	0,563601	1
2	100	70	147	117	0	17	23	30	0	1,41E+14	0,558198	1
3	100	70	129	99	38	0	3	3	26	5,37E+08	0,550411	1
4	100	70	104	74	21	0	45	4	0	1,50E+01	0,551086	1
5	100	70	143	113	25	31	2	12	0	8,79E+12	0,584329	1
6	100	70	144	114	0	23	26	3	18	1,76E+13	0,578829	1
7	100	70	137	107	31	37	2	0	0	0	0,584265	1
8	100	70	139	109	31	0	0	17	22	5,50E+11	0,570848	1
9	100	70	144	114	20	0	6	11	33	1,76E+13	0,570667	1
10	100	70	147	117	16	15	7	24	8	1,41E+14	0,556122	1

Tabela 20: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 100$  e  $n_y = 70$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	100	150	176	226	0	0	24	20	56	7,56E+22	0,599863	1
2	100	150	170	220	30	19	0	0	51	1,18E+21	0,564429	1
3	100	150	114	164	47	14	39	0	0	0	0,577627	1
4	100	150	168	218	0	8	32	0	60	2,95E+20	0,566721	1
5	100	150	149	199	37	30	14	0	19	5,63E+14	0,593526	1
6	100	150	200	250	0	66	0	34	0	1,27E+30	0,554216	1
7	100	150	172	222	0	18	28	2	52	4,72E+21	0,570501	1
8	100	150	182	232	0	0	18	22	60	4,84E+24	0,596029	1
9	100	150	200	250	0	34	0	19	47	1,27E+30	0,553660	1
10	100	150	175	225	11	5	14	70	0	3,78E+22	0,562022	1

Tabela 21: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 100$  e  $n_y = 150$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	200	150	284	234	19	0	47	80	4	1,93E+25	0,560259	1
2	200	150	235	185	66	0	49	35	0	3,44E+10	0,563476	1
3	200	150	299	249	51	72	0	0	27	6,34E+29	0,568268	1
4	200	150	308	258	42	9	0	0	99	3,25E+32	0,583656	1
5	200	150	256	206	42	0	52	22	34	7,21E+16	0,557245	1
6	200	150	299	249	0	0	51	44	55	6,34E+29	0,587970	1
7	200	150	350	300	0	77	0	73	0	1,43E+45	0,571908	1
8	200	150	311	261	34	58	5	53	0	2,60E+33	0,579293	1
9	200	150	258	208	20	0	72	0	58	2,88E+17	0,583206	1
10	200	150	255	205	45	54	50	1	0	1,80E+16	0,594724	1

Tabela 22: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 200$  e  $n_y = 150$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	Locais	$\tau$	$\delta$
1	200	350	291	441	109	85	0	1	5	2,44E+27	0,594187	1
2	200	350	229	379	93	0	78	29	0	5,37E+08	0,571745	1
3	200	350	387	537	0	94	13	12	81	1,96E+56	0,571488	1
4	200	350	255	405	58	29	87	26	0	3,60E+16	0,585331	1
5	200	350	400	550	0	62	0	75	63	1,61E+60	0,561633	1
6	200	350	296	446	19	0	85	0	96	7,92E+28	0,591598	1
7	200	350	208	358	95	8	97	0	0	0	0,595965	1
8	200	350	319	469	9	10	72	75	34	6,65E+35	0,552652	1
9	200	350	339	489	3	0	58	61	78	6,97E+41	0,553572	1
10	200	350	368	518	21	15	11	99	54	3,74E+50	0,588509	1

Tabela 23: Parâmetros dos problemas gerados segundo Calamai e Vicente com  $n_x = 200$  e  $n_y = 350$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_< $	Locais	$\tau$	$\delta$
1	15	30	30	34	4	infinitos	0,446998	2
2	15	30	30	41	11	infinitos	0,527606	2
3	15	30	30	36	6	infinitos	0,124071	1
4	15	30	30	36	6	infinitos	0,086715	2
5	15	30	30	38	8	infinitos	0,535746	1
6	15	30	30	30	0	0	0,141284	2
7	15	30	30	40	10	infinitos	0,450813	1
8	15	30	30	31	1	infinitos	0,510201	1
9	15	30	30	35	5	infinitos	0,389486	2
10	15	30	30	32	2	infinitos	0,371541	1

Tabela 24: Parâmetros dos problemas-teste do tipo GRUPO 1 com  $n_x = 15$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_< $	Locais	$\tau$	$\delta$
1	25	50	50	63	13	infinitos	0,336906	2
2	25	50	50	63	13	infinitos	0,510409	2
3	25	50	50	64	14	infinitos	0,486739	1
4	25	50	50	53	3	infinitos	0,396593	1
5	25	50	50	70	20	infinitos	0,366431	2
6	25	50	50	62	12	infinitos	0,101509	2
7	25	50	50	52	2	infinitos	0,252448	2
8	25	50	50	61	11	infinitos	0,082880	1
9	25	50	50	59	9	infinitos	0,315126	1
10	25	50	50	65	15	infinitos	0,033932	1

Tabela 25: Parâmetros dos problemas-teste do tipo GRUPO 1 com  $n_x = 25$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_{<} $	Locais	$\tau$	$\delta$
1	50	100	100	101	1	infinitos	0,013252	1
2	50	100	100	141	41	infinitos	0,536309	1
3	50	100	100	117	17	infinitos	0,365370	2
4	50	100	100	141	41	infinitos	0,296934	2
5	50	100	100	117	17	infinitos	0,501598	2
6	50	100	100	147	47	infinitos	0,476498	2
7	50	100	100	144	44	infinitos	0,183121	2
8	50	100	100	124	24	infinitos	0,405734	2
9	50	100	100	149	49	infinitos	0,220597	2
10	50	100	100	142	42	infinitos	0,237395	2

Tabela 26: Parâmetros dos problemas-teste do tipo GRUPO 1 com  $n_x = 50$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_{<} $	Locais	$\tau$	$\delta$
1	70	140	140	180	40	infinitos	0,369768	2
2	70	140	140	210	70	infinitos	0,226187	2
3	70	140	140	150	10	infinitos	0,404650	2
4	70	140	140	160	20	infinitos	0,234800	1
5	70	140	140	193	53	infinitos	0,246026	2
6	70	140	140	146	6	infinitos	0,191423	2
7	70	140	140	204	64	infinitos	0,296097	1
8	70	140	140	183	43	infinitos	0,398053	1
9	70	140	140	207	67	infinitos	0,417759	2
10	70	140	140	159	19	infinitos	0,245242	1

Tabela 27: Parâmetros dos problemas-teste do tipo GRUPO 1 com  $n_x = 70$ .

Para os problemas do tipo GRUPO 2 (subseção 6.3.2), definimos  $I_{<<} = \{k; 1 < \rho_k < 5\}$  e  $I_5 = \{k; \rho_k = 5\}$ . Com isto representamos melhor as escolhas dos parâmetros  $\rho_k$ 's.

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_{<} $	$ I_1 $	$ I_{<<} $	$ I_5 $	Locais	$\tau$	$\delta$
1	15	30	30	43	9	3	1	2	3,28E+04	0,304775	1
2	15	30	30	41	5	0	6	4	3,28E+04	0,375543	1
3	15	30	30	39	4	0	5	6	3,28E+04	0,335530	1
4	15	30	30	40	7	3	0	5	3,28E+04	0,496627	1
5	15	30	30	41	4	5	2	4	3,27E+04	0,314370	1
6	15	30	30	44	6	1	7	1	3,28E+04	0,004183	1
7	15	30	30	39	0	7	2	6	3,26E+04	0,147399	2
8	15	30	30	41	1	6	4	4	3,27E+04	0,281821	1
9	15	30	30	45	6	0	9	0	3,28E+04	0,164559	2
10	15	30	30	45	0	11	4	0	3,07E+04	0,331106	1

Tabela 28: Parâmetros dos problemas-teste do tipo GRUPO 2 com  $n_x = 15$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_{<} $	$ I_1 $	$ I_{<<} $	$ I_5 $	Locais	$\tau$	$\delta$
1	25	50	50	68	3	10	5	7	3,36E+07	0,540166	1
2	25	50	50	66	7	7	2	9	3,36E+07	0,458804	1
3	25	50	50	69	10	7	2	6	3,36E+07	0,499451	1
4	25	50	50	72	5	15	2	3	3,35E+07	0,442099	1
5	25	50	50	67	1	16	0	8	3,35E+07	0,040153	1
6	25	50	50	62	0	6	6	13	3,36E+07	0,513084	1
7	25	50	50	75	11	0	14	0	3,36E+07	0,580696	2
8	25	50	50	70	13	0	7	5	3,36E+07	0,500528	2
9	25	50	50	75	0	13	12	0	3,35E+07	0,573317	2
10	25	50	50	64	14	0	0	11	3,36E+07	0,355690	2

Tabela 29: Parâmetros dos problemas-teste do tipo GRUPO 2 com  $n_x = 25$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_{<} $	$ I_1 $	$ I_{<<} $	$ I_5 $	Locais	$\tau$	$\delta$
1	50	100	100	133	12	13	8	17	1,13E+15	0,155848	1
2	50	100	100	139	10	13	16	11	1,13E+15	0,515984	1
3	50	100	100	150	11	17	22	0	1,13E+15	0,483179	1
4	50	100	100	133	0	17	16	17	1,13E+15	0,172352	2
5	50	100	100	132	5	21	6	18	1,13E+15	0,375956	2
6	50	100	100	133	29	4	0	17	1,13E+15	0,598083	2
7	50	100	100	118	0	4	14	32	1,13E+15	0,280812	2
8	50	100	100	150	19	20	11	0	1,13E+15	0,069238	2
9	50	100	100	141	21	12	8	9	1,13E+15	0,106864	1
10	50	100	100	141	23	15	3	9	1,13E+15	0,565995	2

Tabela 30: Parâmetros dos problemas-teste do tipo GRUPO 2 com  $n_x = 50$ .

Grupo	$n_x$	$n_y$	$m_u$	$m_l$	$ I_{<} $	$ I_1 $	$ I_{<<} $	$ I_5 $	Locais	$\tau$	$\delta$
1	70	140	140	190	2	0	48	20	1,18E+21	0,396921	1
2	70	140	140	180	19	21	0	30	1,18E+21	0,396983	1
3	70	140	140	201	36	9	16	9	1,18E+21	0,319101	2
4	70	140	140	163	14	0	9	47	1,18E+21	0,206135	2
5	70	140	140	196	25	29	2	14	1,18E+21	0,014032	1
6	70	140	140	197	19	10	28	13	1,18E+21	0,522264	2
7	70	140	140	178	24	12	2	32	1,18E+21	0,070004	2
8	70	140	140	165	17	0	8	45	1,18E+21	0,177991	2
9	70	140	140	205	28	37	0	5	1,18E+21	0,283506	2
10	70	140	140	190	22	0	28	20	1,18E+21	0,522219	2

Tabela 31: Parâmetros dos problemas-teste do tipo GRUPO 2 com  $n_x = 70$ .