

# **Programação em dois níveis: reformulação utilizando as condições KKT**

Francisco Nogueira Calmon Sobral

DISSERTAÇÃO APRESENTADA  
AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA  
OBTENÇÃO DO TÍTULO  
DE  
MESTRE EM CIÊNCIAS

Área de Concentração: Ciência da Computação  
Orientador: Prof. Dr. Ernesto G. Birgin

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq e da FAPESP proc. no. 2005/57824-9

São Paulo, abril de 2008



## **Programação em dois níveis: reformulação utilizando as condições KKT**

Este exemplar corresponde à redação  
final da dissertação devidamente corrigida  
e defendida por Francisco Nogueira Calmon Sobral  
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Ernesto G. Birgin (orientador) – IME-USP
- Prof. Dr. Paulo José da Silva e Silva – IME-USP
- Profa. Dra. Ana Friedlander de Martínez Peres – IMECC - UNICAMP



# Agradecimentos

Agradeço ao Ernesto por estar sempre presente e aberto a discussões quando necessário. Em especial, quero agradecer por me ensinar como pesquisar e ser o modelo de pesquisador que pretendo me tornar.

Agradeço principalmente à minha querida Ane, que compreendeu (e aceitou) meu trabalho, ajudou-me nas correções e deu um imenso apoio, sem o qual talvez não fosse capaz de terminar este trabalho.

Também agradeço muito aos meus pais por sempre me incentivarem a fazer o que gosto, à minha irmã Cecília por me acompanhar nesses últimos anos e ao meu irmão Fernando pela força à distância.

Finalmente, agradeço a todos os meus familiares e amigos, que me deram força e incentivo ao longo destes 2 anos, em especial aos amigos do laboratório conjunto, meus grandes amigos da graduação e muitas outras pessoas que com certeza estou esquecendo, mas não por isso são menos importantes.



*People optimize...*  
*Nature optimizes...*  
Nocedal e Wright





## Resumo

Em um problema de natureza hierárquica, o nível mais influente toma certas decisões que afetam o comportamento dos níveis inferiores. Cada decisão do nível mais influente é considerada como fixa pelos níveis inferiores, que, com tais informações, tomam decisões que maximizam seus objetivos. Essas decisões podem influenciar os resultados obtidos pelo nível superior, que, por sua vez, também anseia pela decisão ótima. Em programação matemática, este problema é modelado como um problema de programação em níveis. Neste trabalho, consideramos uma classe particular de problemas de programação em níveis: os problemas de programação matemática em dois níveis. Estudamos uma técnica de resolução que consiste em substituir o problema do nível inferior por suas condições necessárias de primeira ordem, que podem ser formuladas de diversas maneiras, conforme as restrições de complementaridade são modificadas. O novo problema torna-se um problema de programação não linear e pode ser resolvido com algoritmos clássicos de otimização. Com o auxílio de condições de otimalidade de primeira e segunda ordem mostramos as relações entre o problema original e o problema reformulado. Aplicamos a técnica a problemas encontrados na literatura, analisamos o seu comportamento e apresentamos estratégias para eliminar certos inconvenientes encontrados.

**Palavras-chave:** complementaridade, condições KKT, funções NCP, programação em dois níveis, programação não linear.



## Abstract

In problems of hierarchical nature, the choices made by the most influential level - the so-called leader - affect the behavior of the lower levels. For each one of the leader's decisions there is a response from the lower levels, which maximizes the value of their respective objectives. These optimal choices, in return, may have influence in the results achieved by the leader, which also wants to make the optimal choices. In mathematical programming, this kind of problem is described as a multilevel programming problem. The present work considers a specific kind of multilevel problem: the bilevel mathematical problem. We study a resolution technique which consists in replacing the lower level problem by its necessary first order conditions, which can be formulated in various ways, as complementarity constraints occur and are modified. The new reformulated problem is a nonlinear programming problem which can be solved by classical optimization methods. Using first and second order optimality conditions, we show the relations between the original bilevel problem and the reformulated problem. We apply the described technique to solve a set of bilevel problems taken from the literature, analyse their behavior and discuss strategies to prevent undesirable difficulties that may arise.

**Keywords:** bilevel programming, complementarity, KKT conditions, NCP functions, nonlinear programming.



# Sumário

<b>Introdução</b>	<b>1</b>
<b>1 Conceitos de programação não linear</b>	<b>5</b>
1.1 Condições de qualificação das restrições . . . . .	5
1.2 Condições de otimalidade . . . . .	10
1.3 Problemas de complementaridade não linear . . . . .	14
<b>2 Programação em dois níveis</b>	<b>19</b>
2.1 Programação matemática em dois níveis . . . . .	20
2.2 Métodos . . . . .	31
2.2.1 Algoritmo <i>K-th Best</i> . . . . .	31
2.2.2 Reformulação utilizando condições KKT . . . . .	31
2.2.3 Branch and bound . . . . .	32
2.2.4 Penalidades . . . . .	32
2.2.5 Direções de descida . . . . .	33
2.2.6 Regiões de confiança . . . . .	34
2.2.7 Restauração inexata . . . . .	34
<b>3 Reformulação através das condições KKT</b>	<b>37</b>
3.1 O problema reformulado . . . . .	38
3.2 Restrições de complementaridade . . . . .	45
3.3 Técnicas de resolução da reformulação . . . . .	48
<b>4 Resultados numéricos</b>	<b>61</b>
4.1 Problemas considerados . . . . .	62
4.2 Resolução das reformulações . . . . .	88
<b>5 Considerações finais</b>	<b>95</b>
5.1 Linhas para futuros estudos . . . . .	96
<b>Referências Bibliográficas</b>	<b>99</b>



# Introdução

Em 1952 Heinrich von Stackelberg escreveu o livro “*The Theory of the Market Economy*” (do original *Grundlagen der Theoretischen Volkswirtschaftslehre*) [Sta52], no qual estudou principalmente a formação de preços no mercado.

Stackelberg afirmou que se existe apenas um produtor de um determinado bem no mercado, e este produtor tem o poder de alterar o preço deste bem como melhor lhe convir, então seu objetivo é tomar decisões que maximizem seu lucro líquido com a venda do produto. O produtor sabe que se aumentar a disponibilidade do bem produzido, o preço abaixará. Por outro lado, se aumentar o preço deste bem, terá que diminuir sua disponibilidade, pois as vendas diminuirão.

Quando o mercado, por sua vez, possui dois produtores do mesmo bem, a quantidade disponível deste bem será dada pela soma das quantias que cada produtor deseja, ou pode, disponibilizar. Ambos anseiam pela decisão ótima (aumentar preço ou aumentar produção) que lhes proporcione o maior lucro líquido possível.

Cada produtor pode se considerar um líder no mercado e tomar suas decisões independentemente, escolhendo, neste caso, a produção que maximize seus lucros sem levar em consideração seu concorrente. Haverá uma disputa de preços/produção até que um dos dois desista e aceite as imposições do outro.

Alternativamente, um produtor pode esperar pelas decisões do outro, para depois definir sua estratégia ótima, considerando as informações do concorrente como fatos (dados fixos). Da mesma forma, o outro produtor observa as decisões tomadas por seu concorrente e define novas estratégias que maximizem seu lucro líquido. Este processo se repete até que atinge-se o equilíbrio. Quando um produtor é reconhecidamente mais forte no mercado que o outro, ao mais fraco nada mais resta a não ser esperar as decisões do primeiro e depois escolher suas decisões ótimas.

Vinte anos depois, e aparentemente sem influências de Stackelberg, Bracken e McGill [BM73] apresentam à comunidade matemática problemas de programação matemática que possuem problemas de otimização como restrições. Os autores se depararam com o problema de alocar armas de destruição para atacar depósitos de armamentos e recursos pré-determinados, de maneira a minimizar o gasto com o ataque. Estes alvos, porém, podem deslocar os armamentos e recursos entre depósitos para minimizar os danos causados pelo ataque.

Os problemas estudados tanto por Stackelberg quanto por Bracken e McGill, embora tenham aplicações diferentes, possuem um ponto em comum: a existência de uma entidade dominante que deseja maximizar seus lucros, mas sua função objetivo está relacionada a decisões tomadas por uma classe dominada. Esta classe dominada também deve fazer

escolhas que maximizem seu lucro, mas suas opções só estão definidas após as decisões da entidade dominante serem feitas, ou seja, há uma ordem nas escolhas.

A estrutura hierárquica é a principal característica destes problemas, que foram denominados, na área de programação matemática, por problemas em níveis [CN77]. Os problemas de programação matemática em dois níveis são os mais conhecidos e sua definição foi fortemente influenciada pelo trabalho de Bracken e McGill. Em problemas de programação em dois níveis, desejamos minimizar a função denominada do primeiro nível ou do líder e queremos que um subconjunto das variáveis do problema seja minimizador global de um problema de programação não linear denominado problema do segundo nível ou do seguidor. Logo, o problema pode ser formulado como:

$$\begin{array}{ll} \text{Minimizar}_{x,y} & F(x, y) \\ \text{s. a} & (x, y) \in \Omega_1 \\ & y \in \arg \min_{y'} \{f(x, y') \mid y' \in \Omega_2(x)\}, \end{array}$$

onde  $\Omega_1$  e  $\Omega_2(x)$  são os conjuntos viáveis do primeiro e do segundo nível, respectivamente.

Na década de 80 os problemas de programação em níveis receberam uma atenção maior do meio acadêmico, motivo pelo qual vários algoritmos foram desenvolvidos. Como exemplos, podemos citar métodos que usam *branch and bound* [Bar88, BF82], de máxima descida [SG94], penalidades [AS84, AW90] e, mais recentemente, regiões de confiança [CMS05] e restauração inexata [ACC<sup>+</sup>07], entre outros.

Muitos problemas, devido à sua estrutura fortemente hierárquica, puderam ser melhor modelados utilizando a programação em níveis. Em economia, utiliza-se programação em dois níveis para investigar a estrutura hierárquica de empresas com muitas divisões descentralizadas e, principalmente, problemas de alocação de recursos [ODJ95]. Também estudam-se meios de determinação do preço de energia elétrica [LLS00], gás [DK02] e biocombustível [BPS00]. Em ecologia, estudam-se maneiras de determinar políticas para a emissão de poluentes na atmosfera [AM99]. Uma infinidade de aplicações, nas mais variadas áreas, pode ser encontrada nos trabalhos de Dempe [Dem03] e Vicente e Calamai [VC94].

Em [BS08] os autores estudaram um problema de empacotamento que consiste em, dado um número fixo  $N$  de itens circulares ou esféricos, encontrar a dimensão mínima que objetos pré-determinados devem ter para que contenham todos os itens, sem que haja sobreposição entre eles. No caso bidimensional e com o objeto circular, este problema pode ser formulado como segue:

$$\begin{array}{ll} \text{Minimizar}_{R,c} & R \\ \text{s. a} & (c_i^x - c_j^x)^2 + (c_i^y - c_j^y)^2 \geq (r_i + r_j)^2, \quad \forall i < j \\ & (c_i^x)^2 + (c_i^y)^2 \leq (R - r_i)^2, \quad i = 1, \dots, N \\ & R \geq r_{\max} \equiv \max_{i=1, \dots, N} \{r_i\}, \end{array}$$

onde  $(c_i^x, c_i^y)$  representa o centro do  $i$ -ésimo item circular de raio  $r_i$  fixo, para  $i = 1, \dots, N$ . Observemos que encontrar a circunferência de menor área, ou menor perímetro, é equivalente a encontrar a circunferência de menor raio. Para  $N = 8$ , a solução deste problema está representada na Figura 1(a).



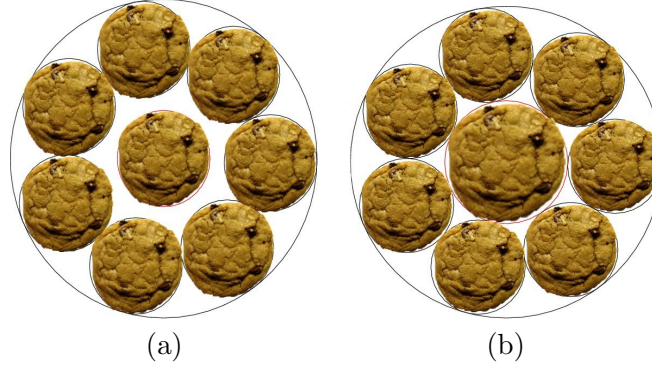


Figura 1: (a) um problema de empacotamento de círculos em círculos com a abordagem usual e  $N = 8$ . (b) o problema de empacotamento resolvido com uma formulação em dois níveis e  $N = 8$ . Note que o biscoito que estava “solto” no centro de (a) teve seu raio aumentado em (b).

Podemos formular uma versão em dois níveis deste problema, que denominaremos “Problema do padeiro benévolo”. A formulação do problema é dada por:

$$\begin{aligned}
 & \underset{R, c, r}{\text{Minimizar}} && R \\
 \text{s. a} &&& R \geq r_{\min} \\
 &&& r \in \arg \max_{r'} \pi \sum_{i=1}^N r_i'^2 \\
 \text{s. a} &&& (c_i^x - c_j^x)^2 + (c_i^y - c_j^y)^2 \geq (r_i + r_j)^2, \quad \forall i < j \\
 &&& (c_i^x)^2 + (c_i^y)^2 \leq (R - r_i')^2, \quad i = 1, \dots, N \\
 &&& r_{\min} \leq r_i' \leq R, \quad i = 1, \dots, N.
 \end{aligned}$$

Esse modelo representa o problema de um padeiro que deseja encontrar o tamanho mínimo de uma massa circular da qual seja possível cortar  $N$  biscoitos circulares com raio pelo menos  $r_{\min}$ . Uma vez minimizado o tamanho da massa, deseja maximizar a soma dos tamanhos dos  $N$  biscoitos. A solução para esse novo problema está representada na Figura 1(b).

Uma técnica de resolução conhecida consiste em substituir o problema do segundo nível por suas condições KKT [BF82]. As condições de Karush-Kuhn-Tucker, ou KKT como são conhecidas, são condições de otimalidade que, sob certas hipóteses, devem ser satisfeitas pelos minimizadores do problema. Uma vez feita a substituição, o novo problema de otimização, que chamaremos de problema reformulado, é um problema de programação não linear e pode ser resolvido por algoritmos já existentes.

O objetivo principal deste trabalho é estudar essa técnica de resolução. Sabe-se que o problema formado pelas condições KKT não possui todas as características desejadas para ser resolvido por algoritmos de programação não linear, porém estamos interessados em estudar os tipos de problemas que podem ser resolvidos e qual é a dificuldade encontrada pelos métodos. Apresentamos um grande número de exemplos que ilustram diversos casos encontrados em problemas com dois níveis.

Discutimos outras formas que possibilitam um problema em dois níveis ser escrito como um problema não linear, baseadas em problemas de complementaridade não linear. Para

estudar o comportamento dos algoritmos de programação não linear, todas as formulações estudadas são resolvidas com o método baseado em Lagrangianos Aumentados ALGEN-CAN [ABMS07]. Os testes realizados têm o propósito de verificar o desempenho da técnica estudada principalmente nos problemas encontrados na literatura.

No Capítulo 1, fazemos uma revisão dos principais elementos de programação não linear que serão usados em todo o texto. Além dos conceitos básicos, também deixamos o leitor a par dos termos e siglas usados. O final desse capítulo está reservado para uma introdução aos problemas de complementaridade não linear, dos quais retiramos alguns conceitos que serão posteriormente aplicados às reformulações dos problemas em dois níveis.

Uma introdução à programação em dois níveis é feita no Capítulo 2, na qual definimos o problema considerado, outros tipos de abordagens e fazemos uma breve recapitulação dos métodos mais conhecidos existentes na literatura, além de alguns métodos recentes.

No Capítulo 3 descrevemos com detalhes a técnica de resolução do problema em dois níveis através de sua reformulação como um problema de programação não linear. Apresentamos uma classe de problemas em dois níveis para a qual o método gera um problema equivalente e mostramos, com muitos exemplos, as conseqüências de aplicar o método para um problema qualquer.

Usamos as funções NCP (funções de complementaridade não linear) da teoria de problemas de complementaridade não linear para discutir vantagens e desvantagens de reformulações alternativas do problema em dois níveis. Por fim, apresentamos duas estratégias que tentam contornar um inconveniente freqüentemente encontrado neste método: a existência de pontos que são viáveis do problema reformulado, mas não são viáveis no problema original.

O Capítulo 4 dedica-se a apresentar os resultados numéricos e comparações feitas entre as implementações das reformulações consideradas neste trabalho. Para todos os problemas de teste considerados, apresentamos uma representação gráfica de seu conjunto viável (quando possível) e encontramos analiticamente ou por meio de algoritmos de otimização global os seus minimizadores globais.

Por fim, no Capítulo 5 apresentamos as conclusões tiradas do estudo de problemas de programação em dois níveis e suas técnicas de resolução, bem como algumas considerações finais e indicações para futuros estudos.

## Notação

Seja uma matriz  $A \in \mathbb{R}^{m \times n}$ , denotaremos por  $(A)_i$  sua  $i$ -ésima linha e por  $(A)^j$  sua  $j$ -ésima coluna. Uma função  $f$  é de classe  $\mathcal{C}^k$  se possuir derivadas parciais de ordem  $k$  e estas derivadas forem contínuas. Por fim,  $B_\varepsilon(x)$  representará o conjunto  $\{y \in \mathbb{R}^n \mid \|y - x\|_2 < \varepsilon\}$ .

# Capítulo 1

## Conceitos de programação não linear

O objetivo deste capítulo é introduzir conceitos básicos de programação não linear que serão usados ao longo de todo o trabalho. O foco desta introdução é apresentar definições, nomenclaturas e, principalmente, condições necessárias e condições suficientes de otimalidade para problemas de programação não linear.

Na Seção 1.1 são apresentadas muitas das definições básicas, com o objetivo principal de introduzir as condições de qualificação. Na Seção 1.2, utilizamos as condições de qualificação para apresentar as condições de otimalidade existentes na teoria de programação não linear. Por fim, na Seção 1.3 fazemos uma rápida introdução aos problemas de complementaridade não linear e sua aplicação nas condições de otimalidade.

### 1.1 Condições de qualificação das restrições

Da mesma forma que um médico, ao realizar um procedimento cirúrgico em um paciente, parte do princípio que ele/ela é um ser humano, em programação não linear com restrições, determinadas condições de otimalidade partem do pressuposto de que o ponto considerado já possui certas características. Estas características, quando estudamos condições necessárias de otimalidade, são chamadas de *condições de qualificação* (do inglês *Constraint Qualifications*) ou qualificações das restrições. Quanto mais fraca uma condição de qualificação, mais abrangente (forte) torna-se a condição de otimalidade que a utiliza. A seguir definimos alguns conceitos primordiais de programação não linear (PNL) e em seguida apresentamos as condições de qualificação mais conhecidas e que serão utilizadas ao longo do trabalho.

O problema de programação não linear com restrições de igualdade e desigualdade é definido por:

$$\begin{aligned} & \underset{x}{\text{Minimizar}} && f(x) \\ & \text{s. a} && g(x) \leq 0 \\ & && h(x) = 0 \\ & && x \in \mathbb{R}^n, \end{aligned} \tag{1.1}$$

onde  $f : \mathbb{R}^n \mapsto \mathbb{R}$ ,  $g : \mathbb{R}^n \mapsto \mathbb{R}^p$  e  $h : \mathbb{R}^n \mapsto \mathbb{R}^m$ . Por simplicidade, consideramos que

as restrições de caixa ( $l \leq x \leq u$ ,  $l, u \in \mathbb{R}^n$ ) são um caso particular das restrições de desigualdade  $g(x) \leq 0$ . Denominaremos por  $\Omega = \{x \in \mathbb{R}^n \mid g(x) \leq 0 \text{ e } h(x) = 0\}$  o **conjunto viável** do problema. Um ponto  $x$  é um **ponto viável** de (1.1) se  $x \in \Omega$ .

Dizemos que  $\bar{x}$  é **minimizador global** de (1.1) se  $\forall x \in \Omega \quad f(x) \geq f(\bar{x})$  (caso valha a desigualdade estrita,  $\bar{x}$  é o minimizador global estrito). Quando  $\bar{x}$  é um minimizador global de (1.1), chamamos o valor da função objetivo em  $\bar{x}$  de **mínimo global** (mínimo global estrito, respectivamente). Se existe  $\varepsilon > 0$  tal que  $\forall x \in \Omega \cap B_\varepsilon(\bar{x}), f(x) \geq f(\bar{x})$ , então  $\bar{x}$  é denominado **minimizador local** de  $f$  (caso valha a desigualdade estrita,  $\bar{x}$  é minimizador local estrito de  $f$ ).

#### 1.1.1 DEFINIÇÃO

Dado  $x \in \mathbb{R}^n$  e o problema definido em (1.1), denominamos **conjunto das restrições** (de desigualdade) **ativas** em  $x$  o conjunto

$$I(x) = \{i \in \{1, \dots, p\} \mid g_i(x) = 0\}.$$

#### 1.1.2 DEFINIÇÃO

Dado  $x \in \Omega$ , se as funções  $g_i$ ,  $i \in I(x)$  e  $h_i$ ,  $i = 1, \dots, m$  forem diferenciáveis em  $x$ ,  $x$  é denominado **regular** se os elementos do conjunto

$$\{\nabla g_i(x), i \in I(x) \quad \text{e} \quad \nabla h_i(x), i = 1, \dots, m\}$$

forem linearmente independentes.

Dado  $x \in \Omega$ , se as funções de desigualdade  $g_i$ ,  $i \in I(x)$ , forem diferenciáveis, o conjunto

$$G_0(x) = \{d \in \mathbb{R}^n \mid \nabla g_i(x)^T d \leq 0, i \in I(x)\}$$

é denominado conjunto das direções viáveis de primeira ordem associadas às restrições de desigualdade. Similarmente, se as funções  $h_i$ , para  $i = 1, \dots, m$ , forem diferenciáveis, o conjunto

$$H_0(x) = \{d \in \mathbb{R}^n \mid \nabla h_i(x)^T d = 0, i = 1, \dots, m\}$$

é denominado conjunto das direções viáveis de primeira ordem associadas às restrições de igualdade.

#### 1.1.3 DEFINIÇÃO

Dado  $x \in \Omega$ , o conjunto formado por  $G_0(x) \cap H_0(x)$  é denominado **conjunto das direções viáveis de primeira ordem** em  $x$ .

Uma condição de qualificação é uma propriedade sobre os pontos viáveis do problema de otimização que, quando satisfeita por um minimizador local do problema, garante a existência dos multiplicadores de Lagrange [Sch06]. Os multiplicadores de Lagrange são variáveis auxiliares que foram inicialmente utilizadas para caracterizar soluções de problemas com restrições de igualdade. Posteriormente foram generalizados nas condições de Karush-Kuhn-Tucker, mais conhecidas como condições KKT, que consideram simultaneamente ambos os tipos de restrições.

Dados  $\bar{x}$  e o problema (1.1), a existência dos multiplicadores de Lagrange está atrelada ao fato de existirem escalares  $\bar{\lambda}_i$  para as restrições de igualdade e  $\bar{\mu}_i$  para as restrições de desigualdade que sejam solução do seguinte sistema:

$$\begin{aligned} \nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) &= 0 \\ \bar{\mu}_i g_i(\bar{x}) &= 0, \quad i = 1, \dots, p \\ g(\bar{x}) &\leq 0 \\ h(\bar{x}) &= 0 \\ \bar{\mu} &\geq 0 \\ \bar{\lambda} &\in \mathbb{R}^n, \end{aligned} \tag{1.2}$$

onde  $\nabla L(x, \lambda, \mu)$  é o gradiente da **função Lagrangiana** definida como:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{i=1}^p \mu_i g_i(x). \tag{1.3}$$

Este sistema também é chamado de sistema KKT ou de **condições KKT**. As restrições da forma

$$\bar{\mu}_i g_i(\bar{x}) = 0, \quad i = 1, \dots, p,$$

juntamente com  $g_i(x) \leq 0$  e  $\mu_i \geq 0$  são chamadas de *restrições de complementaridade*. Como veremos na Seção 1.3, existem diversas maneiras de representar as restrições de complementaridade. A forma usada no sistema (1.2) é a forma mais usual, conhecida como suave, pois é diferenciável desde que  $g_i$  também o seja.

Existem várias condições de qualificação conhecidas na literatura: regularidade (também conhecida como independência linear), Slater [Sla50], Mangasarian-Fromovitz [MF67], CPLD [QW00, AMS05], quase-regularidade, entre outras. Neste trabalho usaremos principalmente as condições de qualificação de Mangasarian-Fromovitz, restrições lineares e de quase-regularidade.

A quase-regularidade também é conhecida como condição de qualificação de Abadie. Tal condição é considerada muito fraca, no sentido de que as outras condições de qualificação conhecidas implicam na quase-regularidade. Logo, uma condição de otimalidade que utilize esta condição de qualificação é mais geral do que uma condição de otimalidade que use a independência linear, por exemplo. Antes de apresentarmos a condição de qualificação de quase-regularidade precisamos introduzir os cones tangentes.

#### 1.1.4 DEFINIÇÃO

Seja  $S$  um conjunto não vazio e  $\bar{x} \in \bar{S}$  (onde  $\bar{S}$  é o fecho de  $S$ ). Denomina-se **cone tangente de  $S$  em  $\bar{x}$**  o conjunto:

$$T(\bar{x}) = \left\{ d \mid d = \lim_{k \rightarrow \infty} \lambda_k (x_k - \bar{x}), \text{ onde } \lambda_k > 0, x_k \in S \text{ para cada } k \text{ e } x_k \rightarrow \bar{x} \right\}.$$

Isto quer dizer que um vetor  $d$  é tangente a  $S$  em  $\bar{x}$  se é possível construir uma seqüência de pontos viáveis  $\{x_k\}$  que se aproxima de  $\bar{x}$  e uma seqüência  $\{\lambda_k\}$  com  $\lambda_k > 0$  de modo que  $\lambda_k(x_k - \bar{x})$  converge para  $d$ .

## 1.1.5 DEFINIÇÃO (QUASE-REGULARIDADE)

Seja  $x \in \Omega$ . Diz-se que  $x$  satisfaz a condição de qualificação de quase-regularidade, ou, equivalentemente, que  $\bar{x}$  é um ponto quase-regular se

$$T(\bar{x}) = G_0(\bar{x}) \cap H_0(\bar{x}).$$

Para verificarmos que a quase-regularidade é uma condição de qualificação, devemos mostrar que se um minimizador do problema (1.1) é um ponto quase-regular, existem multiplicadores de Lagrange que satisfazem (1.2). Esta afirmação será enunciada como uma condição necessária de otimalidade na próxima seção. Sua prova pode ser encontrada em [Ber99].

Entre todas as condições de qualificação que implicam na quase-regularidade, podemos citar a independência linear (regularidade) e as restrições lineares. Embora estas condições sejam mais restritivas, permitem uma fácil verificação algébrica, o que as torna muito utilizadas em provas de convergência para algoritmos de otimização. Abaixo apresentamos a condição de qualificação relativa às restrições lineares.

## 1.1.6 DEFINIÇÃO (RESTRICÇÕES LINEARES)

Seja  $x \in \Omega$ . Diz-se que  $x$  satisfaz a condição de qualificação de restrições lineares se todas as restrições de (1.1) forem lineares.

Podemos encontrar no livro [Ber99] uma demonstração de que restrições lineares implicam em outra condição de qualificação denominada quase-normalidade e que esta, por sua vez, implica na quase-regularidade. Optamos por provar que restrições lineares implicam diretamente na quase-regularidade.

## 1.1.7 TEOREMA

Suponha que as funções  $g_i$ ,  $i = 1, \dots, p$ , e  $h_i$ ,  $i = 1, \dots, m$ , do problema (1.1) sejam lineares. Se  $\bar{x} \in \Omega$  então  $\bar{x}$  é quase-regular.

PROVA:

A prova consiste em mostrar que se  $\bar{x}$  é um ponto viável do problema, então  $\bar{x}$  é quase-regular, ou seja,  $T(\bar{x}) = G_0(\bar{x}) \cap H_0(\bar{x})$ .

- $T(\bar{x}) \subseteq G_0(\bar{x}) \cap H_0(\bar{x})$

Seja  $y \in T(\bar{x})$ . Então existe uma seqüência de pontos viáveis  $\{x^k\}$  que converge para  $\bar{x}$  tal que  $y = \lim_{k \rightarrow \infty} \lambda_k(x^k - \bar{x})$ , onde  $\lambda_k > 0$ . Isso significa que

$$\frac{y}{\|y\|} = \lim_{k \rightarrow \infty} \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|},$$

pela definição de  $y$  e pela continuidade da função  $\|\cdot\|$ .

Considere uma restrição de desigualdade  $g_i$ ,  $i \in I(\bar{x})$ . Embora as funções  $g_i$ ,  $i = 1, \dots, p$ , sejam todas lineares, tal hipótese não é necessária na prova desta parte.

Desta forma

$$\begin{aligned} g_i(x^k) &= g_i(\bar{x}) + \nabla g_i(\bar{x})^T(x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \\ \frac{g_i(x^k) - g_i(\bar{x})}{\|x^k - \bar{x}\|} &= \nabla g_i(\bar{x})^T \left[ \frac{1}{\|x^k - \bar{x}\|} (x^k - \bar{x}) \right] + \frac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \\ \underbrace{\lim_{k \rightarrow \infty} \frac{g_i(x^k) - g_i(\bar{x})}{\|x^k - \bar{x}\|}}_{\leq 0} &= \lim_{k \rightarrow \infty} \left\{ \nabla g_i(\bar{x})^T \underbrace{\left[ \frac{1}{\|x^k - \bar{x}\|} (x^k - \bar{x}) \right]}_{\rightarrow y/\|y\|} \right\} + \underbrace{\lim_{k \rightarrow \infty} \left\{ \frac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} \right\}}_{\rightarrow 0}. \end{aligned}$$

Logo,  $\nabla g_i(\bar{x})^T(y/\|y\|) \leq 0$ . Como  $\|y\| \geq 0$ , a desigualdade não é afetada, o que implica que  $\nabla g_i(\bar{x})^T y \leq 0$ . De maneira análoga, para as restrições de igualdade  $h_i$ ,  $i = 1, \dots, m$ , obtém-se

$$\nabla h_i(\bar{x})^T y = 0.$$

Logo, pode-se concluir que  $y \in G_0(\bar{x}) \cap H_0(\bar{x})$ .

- $T(\bar{x}) \supseteq G_0(\bar{x}) \cap H_0(\bar{x})$

Seja  $y \in G_0(\bar{x}) \cap H_0(\bar{x})$  e considere o ponto  $\bar{x} + \lambda y$ ,  $\lambda > 0$ . Suponha que as restrições lineares são da forma

$$\begin{aligned} Ax &\leq b, & A &\in \mathbb{R}^{p \times n} \\ Cx &= d, & C &\in \mathbb{R}^{m \times n}. \end{aligned}$$

As restrições ativas em  $\bar{x}$  satisfazem as seguintes restrições:

$$\begin{aligned} (A)_i(\bar{x} + \lambda y) &= \underbrace{(A)_i \bar{x}}_{b_i} + \underbrace{\lambda (A)_i y}_{\leq 0} \leq b_i, & \forall i \in I(\bar{x}), \\ (C)_i(\bar{x} + \lambda y) &= \underbrace{(C)_i \bar{x}}_{d_i} + \underbrace{\lambda (C)_i y}_0 = d_i, & \forall i = 1, \dots, m. \end{aligned}$$

Para as restrições de desigualdade inativas em  $\bar{x}$  vale que

$$(A)_i(\bar{x} + \lambda y) = \underbrace{(A)_i \bar{x}}_{< b_i} + \lambda (A)_i y \leq b_i, \quad \forall i \notin I(\bar{x}),$$

desde que seja escolhido  $\lambda$  suficientemente pequeno, ou seja,  $\lambda \in (0, \delta)$ , para algum  $\delta > 0$ . Isso significa que o ponto  $\bar{x} + \lambda y$  é um ponto viável, para  $\lambda \in (0, \delta)$ .

Considere então a sequência  $\{x^k\}$  onde  $x^k = \bar{x} + \lambda_k y$ , para  $\lambda_k \in (0, \delta)$ ,  $\lambda_k \rightarrow 0^+$  quando  $k \rightarrow \infty$ . Logo,

$$\lim_{k \rightarrow \infty} \frac{1}{\lambda_k} (x^k - \bar{x}) = \lim_{k \rightarrow \infty} \frac{1}{\lambda_k} (\bar{x} + \lambda_k y - \bar{x}) = y$$

o que implica  $y \in T(\bar{x})$ , pela definição de cone tangente.

Estes argumentos finalizam a prova e implicam que todo ponto viável do problema é um ponto quase-regular. ■

## 1.2 Condições de otimalidade

Na abordagem estudada neste trabalho, para resolver problemas em dois níveis, o problema do nível inferior é substituído por suas condições KKT. Para verificarmos se o problema agora reformulado é equivalente ao original, temos que verificar para quais casos de problemas de programação não linear as condições KKT são necessárias e suficientes.

Se um problema de otimização for tal que as suas soluções globais (e apenas elas) satisfizerem as condições KKT, encontrar sua solução é equivalente a resolver o sistema KKT. Esta propriedade pode ser utilizada na programação em dois níveis. Como veremos no Capítulo 3, se a solução global do problema do segundo nível satisfizer as condições necessárias e suficientes de otimalidade, então o problema em dois níveis poderá ser reformulado equivalentemente como um problema de programação não linear.

As primeiras condições a serem apresentadas são as condições KKT que são condições necessárias de otimalidade, o que significa que se um ponto for minimizador local e satisfizer alguma condição de qualificação, então deverá satisfazê-las. Originalmente pede-se que o minimizador satisfaça a condição de qualificação de regularidade (ou independência linear, veja a Definição 1.1.2). Em [Ber99] prova-se que a quase-regularidade é uma condição de qualificação e que todas as outras condições de qualificação implicam na quase-regularidade. Deste modo, podemos enunciar as condições KKT utilizando a condição mais fraca de quase-regularidade como condição de qualificação.

### 1.2.1 TEOREMA (CONDIÇÃO NECESSÁRIA DE PRIMEIRA ORDEM)

Seja  $\bar{x}$  um ponto viável do problema (1.1). Suponha que  $f$  e  $g_i$ ,  $i = 1, \dots, p$ , são diferenciáveis em  $\bar{x}$  e que  $h_i$  é de classe  $\mathcal{C}^1$ ,  $i = 1, \dots, m$ . Além disso, suponha que  $\bar{x}$  é quase-regular. Se  $\bar{x}$  é um minimizador local de (1.1), então existem escalares  $\bar{\lambda}_i$ ,  $i = 1, \dots, m$ , e  $\bar{\mu}_i$ ,  $i = 1, \dots, p$ , que satisfazem as condições KKT (1.2).

Para que os teoremas que se seguem possam ser melhor compreendidos, é necessário introduzirmos os conceitos de convexidade, pseudo-convexidade e quase-convexidade de funções, ilustrados na Figura 1.1. Em programação não linear, embora seja uma hipótese forte, a convexidade garante que, se um problema de otimização possui minimizadores, então estes minimizadores são globais. Na tentativa de englobar uma variedade maior de funções, tentamos garantir a existência de minimizadores a partir de hipóteses mais fracas do que convexidade, como a pseudo-convexidade e a quase-convexidade.

### 1.2.2 DEFINIÇÃO

Sejam  $S$  um conjunto convexo e  $f : S \mapsto \mathbb{R}$ .

- (a)  $f$  é denominada **convexa** se  $\forall x, y \in S$  e  $\lambda \in [0, 1]$

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

Uma função  $f$  é **côncava** se  $-f$  for convexa. Dado  $\bar{x} \in S$ ,  $f$  é **convexa em**  $\bar{x}$  se,  $\forall y \in S$  e  $\lambda \in [0, 1]$

$$f(\lambda \bar{x} + (1 - \lambda)y) \leq \lambda f(\bar{x}) + (1 - \lambda)f(y).$$



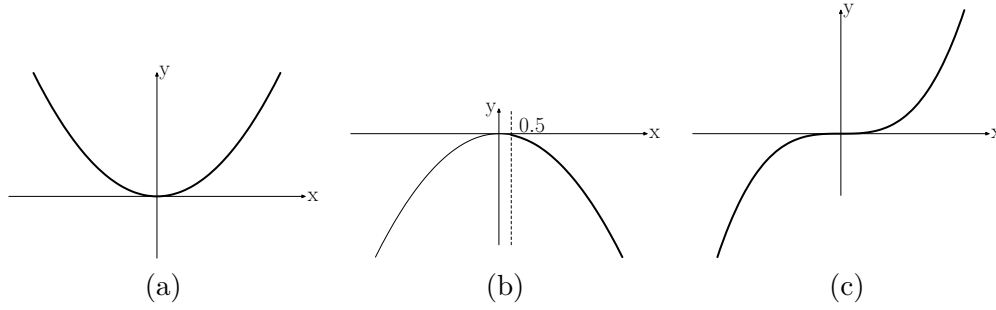


Figura 1.1: (a) A função  $x^2$  é convexa. (b) A função  $-x^2$  é pseudo-convexa para o intervalo  $x \geq 0.5$ , por exemplo. (c) A função  $x^3$  é quase-convexa.

(b) Dado  $\bar{x} \in S$ ,  $f$  é **pseudo-convexa em  $\bar{x}$**  se  $\forall y \in S$

$$\nabla f(\bar{x})^T(y - \bar{x}) \geq 0 \Rightarrow f(y) \geq f(\bar{x}).$$

$f$  é denominada **pseudo-convexa** se  $\forall x, y \in S$

$$\nabla f(x)^T(y - x) \geq 0 \Rightarrow f(y) \geq f(x).$$

(c) Dado  $\bar{x} \in S$ ,  $f$  é **quase-convexa em  $\bar{x}$**  se  $\forall y \in S$  e  $\lambda \in [0, 1]$

$$f(\lambda \bar{x} + (1 - \lambda)y) \leq \max\{f(\bar{x}), f(y)\}.$$

$f$  é denominada **quase-convexa** se,  $\forall x, y \in S$  e  $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}.$$

$f$  é denominada **quase-côncava** se e somente se  $-f$  for quase-convexa.

Em [BSS93] prova-se que convexidade com diferenciabilidade implica na pseudo-convexidade e que a pseudo-convexidade implica na quase-convexidade. Se não houver diferenciabilidade a convexidade implica na quase-convexidade. No teorema a seguir utilizamos os conceitos de convexidade e as informações de primeira ordem para determinar a otimalidade de um ponto viável, quando este satisfaz as condições KKT do problema (1.1). A prova deste teorema pode ser encontrada em [BSS93].

### 1.2.3 TEOREMA (CONDIÇÃO SUFICIENTE DE PRIMEIRA ORDEM)

Seja  $\bar{x}$  uma solução viável de (1.1) que satisfaz as condições KKT (1.2). Seja  $J^+ = \{i \mid \bar{\lambda}_i > 0\}$  e  $J^- = \{i \mid \bar{\lambda}_i < 0\}$ . Se  $f$  é pseudo-convexa em  $\bar{x}$ ,  $g_i$  é quase-convexa em  $\bar{x}$  para  $i \in I(\bar{x})$ ,  $h_i$  é quase-convexa em  $\bar{x}$  para  $i \in J^+$  e  $h_i$  é quase-côncava em  $\bar{x}$  para  $i \in J^-$ , então  $\bar{x}$  é minimizador global de (1.1).

Utilizando os Teoremas 1.2.1 e 1.2.3 podemos afirmar que, se todos os minimizadores do problema de PNL forem quase-regulares e, além disso, cada minimizador satisfizer as condições do Teorema 1.2.3, então as condições KKT serão necessárias e suficientes. Porém, encontrar todos os minimizadores de (1.1) é uma tarefa difícil, além de inviável na prática.

O que podemos encontrar são tipos de problemas para os quais todos os pontos satisfazem as condições dos teoremas acima. Consideremos o seguinte problema:

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{s. a} & Ax \leq b \\ & Cx = d, \end{array} \quad (1.4)$$

onde  $f : \mathbb{R}^n \mapsto \mathbb{R}$  é uma função pseudo-convexa,  $A \in \mathbb{R}^{p \times n}$ ,  $b \in \mathbb{R}^p$ ,  $C \in \mathbb{R}^{m \times n}$  e  $d \in \mathbb{R}^m$ .

#### 1.2.4 PROPOSIÇÃO

*Suponha que o conjunto solução de (1.4) é não vazio. Um ponto viável  $\bar{x}$  é um minimizador global de (1.4) se e somente se  $\bar{x}$  satisfaz as condições KKT.*

PROVA:

Como, pelo Teorema 1.1.7, as restrições lineares são uma condição de qualificação, sabe-se que para todo minimizador de (1.4) existem multiplicadores de Lagrange associados tais que as condições (1.2) são satisfeitas, em particular para os minimizadores globais. Logo,  $\bar{x}$  satisfaz as condições KKT.

Suponha que  $\bar{x}$  satisfaz as condições KKT. Observe primeiramente que, como todas as restrições do problema são lineares, elas são quase-convexas e quase-côncavas. Como a função objetivo é pseudo-convexa, pelo Teorema 1.2.3,  $\bar{x}$  é um minimizador global de (1.4). ■

Quando alguma hipótese da Proposição 1.2.4 é retirada, podemos correr o risco de não existirem multiplicadores de Lagrange associados ao minimizador local, pois se o ponto não satisfaz nenhuma condição de qualificação, não há garantia de sua existência. Além disso, quando a função objetivo não é pseudo-convexa, muitos pontos que não são minimizadores globais podem satisfazer as condições KKT, como os minimizadores locais, os pontos de sela e até alguns maximizadores. O ponto de máximo da função  $-x^2$ , por exemplo, é  $\hat{x} = 0$  e satisfaz as condições KKT.

As condições de segunda ordem podem evitar que os minimizadores locais, os pontos de sela e os maximizadores sejam considerados, impondo propriedades sobre a Hessiana da função Lagrangiana (1.3) associada ao problema de PNL. No Capítulo 3 discutiremos o caso no qual estas condições, aplicadas a problemas de programação em dois níveis, eliminam pontos que não são minimizadores.

Apresentamos abaixo uma condição necessária e uma condição suficiente de segunda ordem e discutimos suas implementações como restrições de problemas de programação não linear. As provas podem ser encontradas em [BSS93].

#### 1.2.5 TEOREMA (CONDIÇÃO NECESSÁRIA DE SEGUNDA ORDEM)

*Seja  $\bar{x}$  um minimizador local do problema (1.1). Suponha que  $\bar{x}$  é um ponto regular e que  $\bar{x}$  satisfaz as condições KKT, com multiplicadores  $\bar{\lambda} \in \mathbb{R}^m$  e  $\bar{\mu} \in \mathbb{R}^p$ . Suponha ainda que as funções  $f$ ,  $g_i$ ,  $i = 1, \dots, p$ , e  $h_i$ ,  $i = 1, \dots, m$ , são duas vezes diferenciáveis. Considere os conjuntos*

$$I^+(\bar{x}) = \{i \in I(\bar{x}) \mid \bar{\mu}_i > 0\} \quad \text{e} \quad I^0(\bar{x}) = \{i \in I(\bar{x}) \mid \bar{\mu}_i = 0\}.$$

Então para todo  $d$  pertencente a

$$C = \left\{ d \neq 0 \left| \begin{array}{l} \nabla g_i(\bar{x})^T d = 0, \quad i \in I^+(\bar{x}) \\ \nabla g_i(\bar{x})^T d \leq 0, \quad i \in I^0(\bar{x}) \\ \nabla h_i(\bar{x})^T d = 0, \quad i = 1, \dots, m \end{array} \right. \right\}$$

deve-se ter

$$d^T \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}) d \geq 0.$$

#### 1.2.6 TEOREMA (CONDIÇÃO SUFICIENTE DE SEGUNDA ORDEM)

Considere o problema (1.1), onde  $f$ ,  $g_i$ ,  $i = 1, \dots, p$ , e  $h_i$ ,  $i = 1, \dots, m$ , são funções duas vezes diferenciáveis. Seja  $\bar{x}$  um ponto que satisfaz as condições KKT, com multiplicadores  $\bar{\lambda} \in \mathbb{R}^m$  e  $\bar{\mu} \in \mathbb{R}^p$  associados. Se

$$d^T \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}) d > 0, \quad \forall d \in C,$$

então  $\bar{x}$  é minimizador local estrito de (1.1), onde  $C$  é o conjunto apresentado no Teorema 1.2.5.

O que os teoremas acima pedem, em outras palavras, é que a Hessiana da função Lagrangiana seja semi-definida positiva (Teorema 1.2.5) ou definida positiva (Teorema 1.2.6) no espaço tangente aos gradientes das restrições ativas (aqui denominado conjunto  $C$ ). Dada uma matriz simétrica  $A \in \mathbb{R}^{n \times n}$ , denominamos  $A$  **definida positiva** se  $\forall x \in \mathbb{R}^n$ ,  $x \neq 0$ ,

$$x^T A x > 0.$$

Da mesma maneira,  $A$  é denominada **semi-definida positiva** se

$$x^T A x \geq 0, \quad \forall x \in \mathbb{R}^n.$$

Dado um subconjunto  $S$  de  $\mathbb{R}^n$ ,  $A$  é **semi-definida positiva em  $S$**  se considerarmos os vetores  $x \in S$  ao invés de  $x \in \mathbb{R}^n$ .

Nosso objetivo é usar a condição suficiente de segunda ordem para eliminar os pontos que não são minimizadores globais, mas satisfazem as condições KKT. No exemplo da função  $-x^2$  no ponto  $\hat{x} = 0$ , temos que a condição suficiente pede que  $\nabla_{xx}^2 f$  seja definida positiva, condição que obviamente não é satisfeita em  $\hat{x}$ . Infelizmente, por ser apenas uma condição suficiente, para a função  $x^4$ , por exemplo, o minimizador  $\bar{x} = 0$  também não a satisfaz.

Na prática, condições de segunda ordem não podem ser utilizadas como restrições em algoritmos de programação não linear exatamente como foram formuladas nos Teoremas 1.2.5 e 1.2.6. Há dois inconvenientes:

- (i) encontrar todos os vetores pertencentes ao conjunto  $C$ ,
- (ii) representar restrições de desigualdade estrita.

Os inconvenientes podem ser contornados através de um resultado de álgebra linear computacional conhecido como Decomposição de Cholesky. Basicamente, uma matriz  $A$  é definida positiva se e somente se existe uma única matriz triangular inferior  $G$  com os

elementos da diagonal positivos, tal que  $A = GG^T$  (veja [Wat91]). Exigimos, então, que a Hessiana da função Lagrangiana  $\nabla_{xx}^2 L$  satisfaça:

$$\begin{aligned} \nabla_{xx}^2 L(x, \lambda, \mu) &= GG^T \\ g_{ii} &\geq \varepsilon \quad i = 1, \dots, n, \end{aligned} \quad (1.5)$$

onde

$$G = \begin{bmatrix} g_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ g_{n1} & \dots & g_{nn} \end{bmatrix}$$

é uma matriz triangular inferior cujos elementos não nulos são considerados variáveis e  $\varepsilon$  é um valor positivo pequeno, necessário para a implementação prática da restrição de desigualdade estrita.

As restrições (1.5) implicam que  $\nabla_{xx}^2 L$  seja definida positiva, mas não são equivalentes à nenhuma condição de otimalidade de segunda ordem apresentada anteriormente. De fato, exigir que  $\nabla_{xx}^2 L$  seja definida positiva é uma restrição mais forte que a condição suficiente do Teorema 1.2.6, basta notar que se  $d^T \nabla_{xx}^2 L d > 0$  para todo  $d \in \mathbb{R}^n$ , então  $d^T \nabla_{xx}^2 L d > 0$  para  $d \in C \subseteq \mathbb{R}^n$ . A seguir mostramos um caso simples no qual o minimizador não satisfaz as condições suficientes de segunda ordem que, neste caso, são equivalentes à condição (1.5):

$$\text{Minimizar}_x \quad x^4.$$

Nesse caso, o minimizador é  $\bar{x} = 0$ , mas  $\nabla_{xx}^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})$  não é definida positiva, pois  $\nabla_{xx}^2 L(x, \lambda, \mu) = 12x^2 \Rightarrow \nabla_{xx}^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$ .

É interessante mencionarmos que Nocedal e Wright [NW99] apresentam um método de verificação das condições necessárias e suficientes de segunda ordem que pode ser implementado na prática. O método utiliza matrizes denominadas *Hessianas Projetadas* para projetar a Hessiana da função Lagrangiana no espaço tangente aos gradientes das restrições ativas. Depois podemos verificar se a matriz projetada é definida (ou semi-definida) positiva utilizando métodos numéricos para determinar os autovalores de uma matriz. O inconveniente desta técnica é o custo computacional para calcular a Hessiana projetada e para determinar os autovalores da matriz.

### 1.3 Problemas de complementaridade não linear

Um problema de complementaridade não linear (conhecido como NCP) consiste em encontrarmos  $x \in \mathbb{R}^n$  que satisfaça [Isa92]:

$$x^T G(x) = 0, \quad x \geq 0, \quad G(x) \geq 0, \quad (1.6)$$

onde  $G(x) : \mathbb{R}^n \mapsto \mathbb{R}^p$ . Problemas desse tipo possuem aplicações em desigualdades variacionais, problemas de engenharia, economia, etc. Em otimização, as condições (1.6) são equivalentes às condições de otimalidade de primeira ordem do seguinte problema [FP97]:

$$\begin{aligned} &\text{Minimizar}_x \quad f(x) \\ \text{s. a} \quad &-g_i(x) \geq 0, \quad i = 1, \dots, p \\ &x_i \geq 0, \quad i = 1, \dots, n. \end{aligned}$$

Podemos verificar tal afirmação escrevendo suas condições KKT:

$$\begin{aligned}\nabla f(x) + \sum_{i=1}^p \mu_i \nabla g_i(x) &= \mu^x \quad (\doteq \mu^x(x, \mu)) \\ -g(x) &\geq 0 \\ -\mu_i g_i(x) &= 0 \\ \mu_i^x x_i &= 0 \\ \mu, \mu^x, x &\geq 0.\end{aligned}$$

Logo, se definirmos  $z = \begin{bmatrix} x \\ \mu \end{bmatrix}$  e  $G(z) = \begin{bmatrix} \mu^x(x, \mu) \\ -g(x) \end{bmatrix}$ , as condições KKT são equivalentes a pedirmos que  $z$  e  $G(z)$  satisfaçam

$$z \geq 0, G(z) \geq 0 \text{ e } z^T G(z) = 0,$$

exatamente como queríamos. Notemos que  $z^T G(z) = 0 \iff z_i G_i(z) = 0$ , pois  $z_i \geq 0$  e  $G_i(z) \geq 0$ .

Na definição das condições KKT (1.2) para o problema de programação não linear mais geral (1.1), usamos um conjunto de restrições denominadas restrições de complementaridade. O objetivo de uma restrição de complementaridade associada à variável  $\mu_i$  e à restrição  $g_i(x)$  é fazer  $\mu_i = 0$  se  $g_i(x) < 0$  ou  $g_i(x) = 0$  se  $\mu_i > 0$ , o que pode ser representado pelas seguintes condições:

$$g_i(x) \leq 0 \tag{1.7}$$

$$\mu_i \geq 0 \tag{1.8}$$

$$\mu_i g_i(x) = 0. \tag{1.9}$$

É fácil perceber que existe uma grande similaridade entre as restrições de complementaridade e um problema de complementaridade do tipo (1.6). O principal objetivo desta seção é mostrar que esta similaridade pode originar novas maneiras de escrever as restrições de complementaridade.

A formulação dada por (1.7–1.9) é a mais conhecida para restrições de complementaridade, sendo utilizada na maioria das definições de condições KKT. Nós a denominaremos formulação **suave**, visto que a sua diferenciabilidade depende apenas da diferenciabilidade da função  $g_i(x)$ , que geralmente é suave.

A forte semelhança entre o NCP (1.6) e as restrições de complementaridade sugere utilizarmos conceitos existentes na literatura de problemas de complementaridade não linear para representar de forma alternativa e equivalente as restrições (1.7–1.9). Sabe-se que é possível substituir o NCP clássico por uma função equivalente, de forma que o novo problema consista em satisfazer uma única equação. Funções com esta propriedade são denominadas funções NCP.

#### 1.3.1 DEFINIÇÃO (FUNÇÃO NCP)

Seja  $\phi(a, b) : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ .  $\phi$  é denominada **função NCP** se

$$\phi(a, b) = 0 \iff a^T b = 0, \quad a, b \geq 0.$$

Com esta definição, o problema (1.6) transforma-se no problema de encontrar  $x \in \mathbb{R}^n$  tal que  $\phi(x, G(x)) = 0$ .

Podemos, então, usar as funções NCP para substituir equivalentemente as restrições de complementaridade por uma única restrição de igualdade. Além de reduzirmos o número de restrições, cada função possui características próprias, tornando-se mais ou menos adequada a determinada situação. Na literatura existem muitas funções NCP, cada qual com suas vantagens e desvantagens, mas apresentaremos no trabalho apenas duas das funções mais conhecidas.

Vamos considerar  $\phi : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$  e definir a seguinte função:

$$\phi(a, b) = \min(a, b),$$

onde

$$\min(a, b) = \begin{cases} a, & \text{se } a < b \\ a \text{ ou } b, & \text{se } a = b \\ b, & \text{se } a > b. \end{cases}$$

É fácil verificar que  $\phi(a, b) = 0$  se e somente se  $ab = 0$  e  $a, b \geq 0$ , logo,  $\phi(a, b)$  é uma função NCP. Esta função é chamada **função mínimo**. Usando a função mínimo, as restrições de complementaridade suaves são substituídas por:

$$\min(\mu_i, -g_i(x)) = 0. \quad (1.10)$$

Esta função claramente não é diferenciável para todo ponto tal que  $\mu_i = -g_i(x)$ , mesmo que  $g_i(x)$  o seja, como verificamos a seguir:

$$\begin{aligned} \frac{\partial}{\partial x} \min(\mu_i, -g_i(x)) &= \begin{cases} 0, & \text{se } \mu_i < -g_i(x) \\ 0 \text{ ou } -\nabla g_i(x), & \text{se } \mu_i = -g_i(x) \\ -\nabla g_i(x), & \text{se } \mu_i > -g_i(x) \end{cases} \\ \frac{\partial}{\partial \mu_i} \min(\mu_i, -g_i(x)) &= \begin{cases} 1, & \text{se } \mu_i < -g_i(x) \\ 1 \text{ ou } 0, & \text{se } \mu_i = -g_i(x) \\ 0, & \text{se } \mu_i > -g_i(x) \end{cases} \end{aligned}$$

Se utilizarmos (1.10) como restrição em um problema de PNL, a falta de diferenciabilidade é um fator que pode fazer muitos métodos de programação não linear funcionarem incorretamente. O caso crítico ocorre quando  $-g_i(x) = \mu_i$ , pois o ponto possui duas derivadas. É possível que a escolha de uma derivada, em detrimento de outra, influencie na solução final encontrada pelo método, ou na existência dos multiplicadores de Lagrange para a solução. Por esta razão, decidimos definir a função mínimo com três possibilidades ao invés de duas, pois o desempenho da formulação torna-se dependente da escolha feita na possibilidade  $a = b$ .

Vamos definir agora outra função NCP, baseada na função de Fischer-Burmeister [Fis92]. Seja  $\phi : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ . A função de Fischer-Burmeister é definida por:

$$\phi(a, b) = a + b - \sqrt{a^2 + b^2}.$$

## 1.3.2 PROPOSIÇÃO

A função de Fischer-Burmeister é uma função NCP.

PROVA:

Suponha que  $a, b \geq 0$  e  $ab = 0$ . Então

$$\begin{aligned} a + b - \sqrt{a^2 + b^2} &= a + b - \sqrt{a^2 + 2ab + b^2} \\ &= a + b - \sqrt{(a + b)^2} \\ &= a + b - |a + b| \\ &= a + b - (a + b) \\ &= 0 \end{aligned}$$

Suponha agora que  $a + b - \sqrt{a^2 + b^2} = 0$ . Então

$$a + b = \sqrt{a^2 + b^2} \geq 0.$$

Como  $a + b \geq 0$ , é possível elevar ao quadrado os dois lados da igualdade acima e obter que  $ab = 0$ . Elevando ao quadrado a desigualdade  $a + b \geq 0$ , obtém-se que  $ab \geq 0$ . Então ou  $a \geq 0$  e  $b \geq 0$  ou  $a \leq 0$  e  $b \leq 0$ . Mas sabe-se que  $a + b \geq 0$ , o que implica que  $a, b \geq 0$ . Logo,  $a, b \geq 0$  e  $ab = 0$ .

■

Dentre suas muitas características, a função de Fischer-Burmeister é suave e diferenciável em todos os pontos, exceto em  $(a, b) = (0, 0)$ . A falta de diferenciabilidade na origem também é observada na prática, pois ao implementarmos a função de Fischer-Burmeister em um problema de programação não linear, o algoritmo realmente foi incapaz de calcular o gradiente em  $(0, 0)$ . Para contornar este inconveniente, Kanzow [Kan96] sugeriu uma função posteriormente denominada de **Fischer-Burmeister perturbada** definida por:

$$\phi_\varepsilon(a, b) = a + b - \sqrt{a^2 + b^2 + \varepsilon}, \quad \varepsilon > 0.$$

Esta função é diferenciável em todos os pontos, com gradiente:

$$\nabla \phi_\varepsilon(a, b) = \begin{bmatrix} 1 - \frac{a}{\sqrt{a^2 + b^2 + \varepsilon}} \\ 1 - \frac{b}{\sqrt{a^2 + b^2 + \varepsilon}} \end{bmatrix}$$

e possui a seguinte característica:

$$\phi_\varepsilon(a, b) = 0 \iff a, b > 0 \text{ e } ab = \varepsilon/2.$$

Embora, de acordo com a definição, a versão perturbada da função de Fischer-Burmeister não seja uma função NCP, ela pode ser utilizada na prática. Basta tomarmos um valor de  $\varepsilon$  grande o suficiente para que a função  $\phi_\varepsilon$  seja diferenciável. Claramente, quanto maior o valor de  $\varepsilon$  menos semelhante a uma função NCP será  $\phi_\varepsilon$ , de forma que verificamos na prática

o funcionamento de valores suficientemente pequenos para  $\varepsilon$ . Neste caso, as restrições de complementaridade suaves (1.7–1.9) são substituídas pela igualdade:

$$\mu_i - g_i(x) - \sqrt{\mu_i^2 + g_i(x)^2 + \varepsilon} = 0. \quad (1.11)$$

Para calcular o gradiente de  $\phi_\varepsilon(\mu_i, -g_i(x))$ , basta aplicar a regra da cadeia:

$$\begin{aligned} \frac{\partial}{\partial x_j} \phi_\varepsilon(\mu_i, -g_i(x)) &= \left[ \frac{\partial \mu_i}{\partial x_j} \quad -\frac{\partial g_i(x)}{\partial x_j} \right] \cdot \nabla \phi_\varepsilon(\mu_i, -g_i(x)) = -\frac{\partial g_i(x)}{\partial x_j} \left( 1 + \frac{g_i(x)}{\sqrt{\mu_i^2 + g_i(x)^2 + \varepsilon}} \right) \\ \frac{\partial}{\partial \mu_i} \phi_\varepsilon(\mu_i, -g_i(x)) &= \left[ \frac{\partial \mu_i}{\partial \mu_i} \quad -\frac{\partial g_i(x)}{\partial \mu_i} \right] \cdot \nabla \phi_\varepsilon(\mu_i, -g_i(x)) = 1 - \frac{\mu_i}{\sqrt{\mu_i^2 + g_i(x)^2 + \varepsilon}}. \end{aligned}$$

Convém lembrarmos que a diferenciabilidade de  $\phi_\varepsilon(\mu_i, -g_i(x))$  depende da diferenciabilidade da função  $g_i(x)$ .



## Capítulo 2

# Programação em dois níveis

Um problema de programação em dois níveis (também conhecido como BLPP) é um problema de programação não linear que possui um subconjunto de variáveis restrito a ser solução de um outro problema de otimização. Os problemas de programação em níveis foram originalmente introduzidos na programação matemática em meados da década de 70, através do trabalho de Bracken e McGill [BM73]. A formulação e nomenclatura dos problemas como conhecemos atualmente foi introduzida no trabalho de Candler e Norton [CN77]. Anteriormente ao trabalho de Bracken e McGill, na década de 50 Stackelberg apresentou em [Sta52] um estudo de problemas hierárquicos de Economia muito similares aos problemas em níveis. Seu trabalho é considerado um dos primordiais na teoria de programação em dois (ou mais) níveis.

Nas décadas de 80 e 90, a programação em níveis recebeu uma maior atenção do meio acadêmico, período no qual surgiram os primeiros algoritmos capazes de resolver determinadas classes de problemas em dois níveis. Entre os algoritmos mais conhecidos encontram-se métodos de máxima descida [SG94], métodos de *branch and bound* [Bar88], métodos de pontos extremos [BK78, BK82], métodos de penalidades [AS84, SA81], entre outros. Recentemente, em [ACC<sup>+</sup>07] técnicas de restauração inexata foram utilizadas para resolver problemas em dois níveis e um método que utiliza regiões de confiança foi apresentado em [CMS05].

O desenvolvimento de novos métodos fez o número de aplicações crescer rapidamente, com destaque para os problemas contendo estruturas hierárquicas. Os problemas em dois níveis, sob certas hipóteses, puderam ser classificados como um caso particular de uma classe denominada *problemas de programação matemática com restrições de equilíbrio* (MPECs), problemas de programação matemática cujas restrições principais são desigualdades variacionais paramétricas ou sistemas de complementaridade [LPR96].

Neste capítulo apresentamos as definições de problemas de programação em dois níveis. Discutimos suas características e as técnicas mais conhecidas para resolver determinados tipos de problemas. Na Seção 2.1 discorreremos sobre as definições encontradas na literatura, suas possíveis interpretações e definiremos o problema considerado neste trabalho. Na Seção 2.2 apresentamos as técnicas de resolução mais conhecidas e os tipos de problemas em dois níveis resolvidos por cada uma.

## 2.1 Programação matemática em dois níveis

O problema de programação matemática em dois níveis (BLPP) geral utilizado neste trabalho pode ser formulado como segue:

$$\begin{aligned}
 & \underset{x,y}{\text{Minimizar}} && F(x,y) \\
 & \text{s. a} && H(x,y) = 0 \\
 & && G(x,y) \leq 0 \\
 & && y \in \underset{y'}{\arg \min} f(x,y') \\
 & && \text{s. a} && h(x,y') = 0 \\
 & && && g(x,y') \leq 0.
 \end{aligned} \tag{2.1}$$

O problema:

$$\begin{aligned}
 & \underset{y'}{\text{Minimizar}} && f(x,y') \\
 & \text{s. a} && h(x,y') = 0 \\
 & && g(x,y') \leq 0,
 \end{aligned} \tag{2.2}$$

onde  $x \in \mathbb{R}^{n_1}$ ,  $y' \in \mathbb{R}^{n_2}$ ,  $f(x,y') : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}$ ,  $g(x,y') : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{p_2}$  e  $h(x,y') : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{m_2}$ , é chamado de **problema do segundo nível** ou do seguidor. Este problema é um problema de otimização parametrizado em  $x$  no qual o seguidor possui apenas o controle das variáveis  $y'$ . Denominemos por  $\Psi(x)$  o **conjunto dos minimizadores globais** de (2.2) para um dado  $x$ . Desta forma, definimos o **problema do primeiro nível** ou do líder:

$$\begin{aligned}
 & \underset{x,y}{\text{Minimizar}} && F(x,y) \\
 & \text{s. a} && H(x,y) = 0 \\
 & && G(x,y) \leq 0 \\
 & && y \in \Psi(x),
 \end{aligned} \tag{2.3}$$

onde  $x \in \mathbb{R}^{n_1}$ ,  $y \in \mathbb{R}^{n_2}$ ,  $F(x,y) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}$ ,  $G(x,y) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{p_1}$  e  $H(x,y) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{m_1}$ .

A seguir definimos alguns conjuntos encontrados na literatura e que serão usados ao longo do trabalho. O conjunto

$$\Omega \doteq \{(x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mid H(x,y) = 0, G(x,y) \leq 0, h(x,y) = 0 \text{ e } g(x,y) \leq 0\}$$

é denominado **conjunto viável** do BLPP. Quando apenas as variáveis do primeiro nível precisam ser analisadas, definimos como a **projeção de  $\Omega$  no espaço de decisão do primeiro nível** o seguinte conjunto

$$\Omega_x \doteq \{x \mid (x,y) \in \Omega \text{ para algum } y\}.$$

Finalmente, o conjunto

$$\text{IR} \doteq \{(x,y) \mid H(x,y) = 0, G(x,y) \leq 0 \text{ e } y \in \Psi(x)\}$$

é denominado **região induzida** do BLPP e representa o conjunto de pontos viáveis nos quais a função objetivo do primeiro nível será minimizada. Utilizando o conjunto IR, o BLPP (2.1) pode ser reescrito simplificadamente da seguinte forma:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) \\ \text{s. a} & (x,y) \in \text{IR}. \end{array}$$

Definimos ainda um outro conjunto que não é encontrado na literatura:

$$\text{IR}(x) \doteq \{y \mid H(x,y) = 0, G(x,y) \leq 0 \text{ e } y \in \Psi(x)\}.$$

O Exemplo 2.1.1 que segue representa um problema típico de programação em dois níveis. Através dele ilustramos os conjuntos que foram definidos anteriormente.

### 2.1.1.1 EXEMPLO ([SC98])

Considere o problema:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & F(x,y) = (x-3)^2 + (y-2)^2 \\ \text{s. a} & -2x + y - 1 \leq 0 \\ & x - 2y + 2 \leq 0 \\ & x + 2y - 14 \leq 0 \\ & 0 \leq x \leq 8 \\ & y \in \arg \min_{y'} f(x,y') = (y' - 5)^2. \end{array}$$

Observe a Figura 2.1. O **conjunto dos minimizadores globais** do segundo nível é formado por  $\Psi(x) = \{5\}$ , independentemente do valor de  $x$ , pois o minimizador do problema do segundo nível é  $\bar{y} = 5$ . Como  $\Psi(x)$  é unitário, pode ser substituído pela função  $y(x) = 5$ . O **conjunto viável do BLPP** (conjunto  $\Omega$ ) é formado pela intersecção de todas as restrições de desigualdade do primeiro nível. O segmento de reta entre os pontos  $(0,0)$  e  $(6,0)$  representa  $\Omega_x$ , a **projeção do conjunto  $\Omega$  no espaço de decisão do primeiro nível**. A **região induzida**, conjunto no qual a função  $F(x,y)$  será minimizada, é a intersecção da reta  $y(x) = 5$  com o conjunto  $\Omega$ . O minimizador deste BLPP é o ponto  $(3,5)$ , que pode ser encontrado eliminando o problema do segundo nível e substituindo  $y = 5$  no primeiro nível.

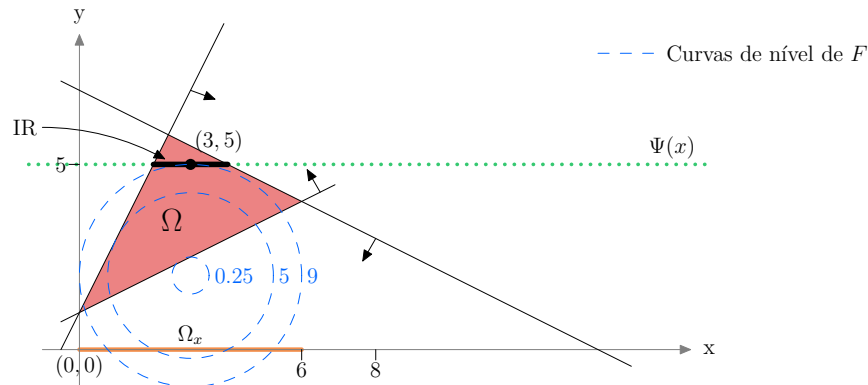


Figura 2.1: Ilustração de um problema em dois níveis e seus conjuntos mais importantes.

Em uma revisão bibliográfica feita por Vicente e Calamai [VC94] o problema (2.1) é denominado BLPP. Já para Dempe [Dem02] e Bard [Bar98] o BLPP é definido como:

$$\begin{aligned}
 & \underset{x}{\text{Minimizar}} && F(x, y) \\
 & \text{s. a} && H(x, y) = 0 \\
 & && G(x, y) \leq 0 \\
 & && y \in \arg \min_{y'} f(x, y') \\
 & && \text{s. a} && h(x, y') = 0 \\
 & && && g(x, y') \leq 0,
 \end{aligned} \tag{2.4}$$

onde a única diferença com relação a (2.1) é que a função objetivo do primeiro nível é minimizada apenas com relação às variáveis do líder ( $x$ ). Vicente e Calamai [VC94] denominam (2.4) como Problema de Stackelberg Estático (SSP).

Ao contrário do BLPP (2.1), os problemas do tipo (2.4) permitem duas abordagens: a abordagem otimista e a pessimista. Utilizando o conjunto  $\Psi(x)$ , temos que a abordagem **otimista** pode ser representada pelo problema:

$$\begin{aligned}
 & \underset{x}{\text{Min}} && \underset{y}{\text{Min}} && F(x, y) \\
 & && \text{s. a} && H(x, y) = 0 \\
 & && && G(x, y) \leq 0 \\
 & && && y \in \Psi(x).
 \end{aligned} \tag{2.5}$$

Podemos interpretar esta abordagem como um caso no qual o líder possui alguma influência sobre as escolhas do seguidor. Devido à influência do líder (ou à afinidade com ele), dentre os elementos  $y \in \Psi(x)$ , o seguidor escolherá aquele que minimiza o valor de  $F(x, y)$ .

A abordagem **pessimista** é representada pelo seguinte problema:

$$\begin{aligned}
 & \underset{x}{\text{Min}} && \underset{y}{\text{Max}} && F(x, y) \\
 & && \text{s. a} && H(x, y) = 0 \\
 & && && G(x, y) \leq 0 \\
 & && && y \in \Psi(x).
 \end{aligned} \tag{2.6}$$

Neste caso supomos que não há uma cooperação entre os dois níveis. A solução encontrada pelo líder será aquela que minimizará a perda causada pela pior escolha do seguidor dentre as soluções em  $\Psi(x)$ .

Quando, para todo  $x$ , o conjunto  $\Psi(x)$  é unitário, as influências e/ou cooperações entre os níveis não têm mais efeito. Neste caso, para qualquer escolha  $x$  do primeiro nível, existe apenas uma escolha no segundo nível, que denominamos  $y(x)$ , e o valor de  $F(x, y(x))$  torna-se dependente somente das variáveis  $x$ , não importando a abordagem usada.

O problema (2.3) é uma forma equivalente de descrevermos o BLPP (2.1), mas, da maneira como está apresentado, permite que possamos interpretá-lo sem dificuldades como sendo um problema de programação não linear. Logo, para estudar suas soluções, usamos as definições de minimizadores locais e globais conhecidas em programação não linear.

As abordagens otimista e pessimista do problema (2.4) não são problemas de programação não linear da forma (1.1), o que nos impede de aplicar as definições de minimizadores locais e globais conhecidas para estudar suas soluções. Precisamos definir os minimizadores do problema otimista e do pessimista para depois traçar um paralelo com os minimizadores dos problemas de programação não linear.

Colson, Marcotte e Savard [CMS07] consideram em seu trabalho problemas de programação em dois níveis gerais definidos da seguinte maneira:

$$\begin{aligned}
 & \underset{x,y}{\text{Minimizar}} && F(x,y) \\
 & \text{s. a} && G(x,y) \leq 0 \\
 & && x \in X \\
 & && y \in \underset{y'}{\arg \min} f(x,y') \\
 & && \text{s. a} && g(x,y') \leq 0.
 \end{aligned} \tag{2.7}$$

É fácil verificar que esta definição é uma outra forma de representação do BLPP (2.1), pois toda restrição de igualdade pode ser escrita como duas restrições de desigualdade. Os autores de [CMS07] definem minimizador local otimista do problema (2.7) como segue. O ponto  $(\bar{x}, \bar{y}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$  é um *minimizador local otimista* se

$$\begin{aligned}
 & \bar{x} \in X, \bar{y} \in \Psi(\bar{x}), \\
 & G(\bar{x}, \bar{y}) \leq 0, \\
 & F(\bar{x}, \bar{y}) \leq F(\bar{x}, y) \quad \forall y \in \Psi(\bar{x})
 \end{aligned}$$

e, além disso, existe uma vizinhança aberta  $U_\delta(\bar{x})$ ,  $\delta > 0$ , de centro em  $\bar{x}$  tal que

$$\min_{y'} \{F(\bar{x}, y') \mid y' \in \Psi(\bar{x})\} \leq \min_{y'} \{F(x, y') \mid y' \in \Psi(x)\} \quad \forall x \in U_\delta(\bar{x}) \cap X.$$

O ponto  $(\bar{x}, \bar{y})$  é um *minimizador global otimista* se uma vizinhança de tamanho  $\delta = \infty$  puder ser escolhida.

De acordo com o que apresentamos até o presente momento, parece existir uma incompatibilidade entre o problema e a definição apresentados em [CMS07]. O problema (2.7), assim como o problema (2.1), não permite diferentes abordagens. No primeiro nível, a função  $F(x, y)$  é minimizada com relação a  $x$  e  $y$ , de forma que está implícito no problema uma colaboração do segundo nível, visto que este fará suas escolhas visando minimizar a função objetivo do líder. Como veremos posteriormente, este problema é uma versão *relaxada* da abordagem otimista.

Concluimos que há duas possibilidades: (i) esta definição recebeu uma denominação infeliz e coincide na verdade com a definição de minimizador local encontrada em PNL; (ii) a definição não está associada ao problema (2.7) e sim ao problema (2.4), pois este último realmente possui as abordagens otimista e pessimista. A seguir mostramos que tanto (i) quanto (ii) não ocorrem e, portanto, a definição está simplesmente errada.

Para mostrar que a possibilidade (i) não ocorre, apresentamos no Exemplo 2.1.2 um problema da forma (2.7) que possui um minimizador local (de acordo com a teoria de programação não linear) para o qual a definição de minimizador local otimista apresentada em [CMS07] não se aplica.

## 2.1.2 EXEMPLO ([DEM02])

Considere o seguinte BLPP:

$$\begin{array}{ll} \text{Minimizar} & x \\ \text{s. a} & \begin{array}{l} -1 \leq x \leq 1 \\ y \in \arg \min_{y'} xy' \\ \text{s. a} \quad 0 \leq y' \leq 1. \end{array} \end{array}$$

Este problema foi adaptado de [Dem02], pois na formulação original a função do primeiro nível é minimizada apenas com relação a  $x$ . No problema do segundo nível pode-se verificar que, dado  $x$ ,

$$\Psi(x) = \begin{cases} 1, & \text{se } x < 0 \\ [0, 1], & \text{se } x = 0 \\ 0, & \text{se } x > 0. \end{cases}$$

Usando o conjunto  $\Psi(x)$ , o problema considerado pode ser reescrito como se segue:

$$\begin{array}{ll} \text{Minimizar} & x \\ \text{s. a} & \begin{array}{l} -1 \leq x \leq 1 \\ y \in \Psi(x). \end{array} \end{array}$$

O conjunto viável está representado na Figura 2.2. O minimizador global do problema é o ponto  $(\bar{x}, \bar{y}) = (-1, 1)$ . O problema também possui infinitos minimizadores locais da forma  $(0, \hat{y})$  para  $\hat{y} \in [0, 1[$ , de acordo com a definição de minimizador local para problemas de PNL.

Observe que nenhum ponto da forma  $(0, \hat{y})$ , com  $\hat{y} \in [0, 1[$ , é um minimizador local, de acordo com a definição de minimizador local otimista apresentada em [CMS07]. Dado  $0 < \delta < 1$  escolha  $\xi \in ]0, \delta[$ . Desta forma  $-\xi$  pertence à vizinhança  $U_\delta(0) = ]-\delta, +\delta[$  em torno de 0 e também pertence ao conjunto  $X = \{x \mid -1 \leq x \leq 1\}$ . Então, para toda vizinhança  $U_\delta(0)$ , existe  $x = -\xi$ , com  $\xi$  definido anteriormente, tal que

$$\min_{y'} \{-\xi \mid y' \in \Psi(-\xi)\} = -\xi < 0 = \min_{y'} \{0 \mid y' \in \Psi(0)\},$$

o que implica que  $(0, \hat{y})$ ,  $\hat{y} \in [0, 1[$ , não é um minimizador local do problema, de acordo com [CMS07].

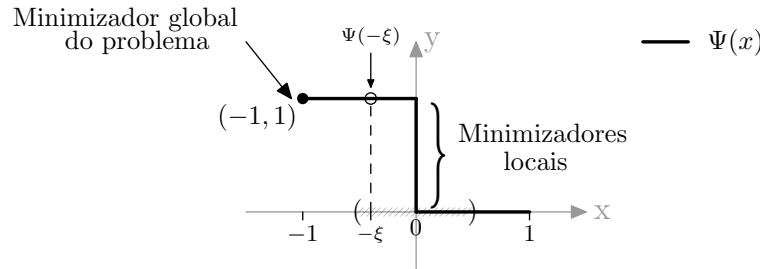


Figura 2.2: Exemplo de um problema da forma (2.7). Os pontos da forma  $(0, \hat{y})$ ,  $\hat{y} \in [0, 1[$ , não são minimizadores locais, de acordo com [CMS07].

Para mostrar que a possibilidade (ii) não ocorre, apresentamos um BLPP da forma (2.4) para o qual a definição de minimizador local otimista de [CMS07] não se aplica. O Exemplo 2.1.3 mostra que esta definição não identifica corretamente o minimizador global otimista do problema.

## 2.1.3 EXEMPLO

Considere o seguinte problema de programação em dois níveis:

$$\begin{array}{ll} \text{Minimizar} & x \\ \text{s. a} & x + y \geq 0 \\ & y \in \arg \min_{y'} -y'^2 \\ & \text{s. a} \quad x^2 + y'^2 \leq 4. \end{array}$$

A restrição do segundo nível pede que, para  $-2 \leq x \leq 2$ ,  $y'$  pertença ao intervalo  $[-\sqrt{4-x^2}, \sqrt{4-x^2}]$ . Como  $-y'^2$  é estritamente côncava, sabe-se que seus minimizadores se encontram nos extremos do intervalo. O valor mínimo de  $-y'^2$  restrita a este intervalo é obtido em dois minimizadores globais:  $\bar{y}' = -\sqrt{4-x^2}$  e  $\bar{y}' = \sqrt{4-x^2}$ , para  $-2 \leq x \leq 2$ . Quando  $x < -2$  ou  $x > 2$  o problema do segundo nível torna-se inviável. Desta forma, para todo  $x$ :

$$\Psi(x) = \begin{cases} \{\pm\sqrt{4-x^2}\}, & \text{se } -2 \leq x \leq 2 \\ \emptyset, & \text{c.c.} \end{cases}$$

A união dos conjuntos  $\Psi(x)$ , para todo  $x$ , forma uma circunferência de centro na origem e raio 2. A intersecção desta circunferência com a restrição do primeiro nível  $x + y \geq 0$  forma o conjunto  $\mathbb{R}$ , representado em negrito na Figura 2.3 e no qual a função do primeiro nível  $F(x, y) = x$  será minimizada. Considere a formulação otimista do problema acima, apresentada a seguir:

$$\begin{array}{ll} \text{Min} & \text{Min}_y x \\ \text{s. a} & x + y \geq 0 \\ & y \in \Psi(x). \end{array}$$

Como a função objetivo é independente de  $y$ , minimizar  $F(x, y) = x$  no conjunto  $\mathbb{R}$  leva ao ponto  $(\bar{x}, \bar{y}) = (-\sqrt{2}, \sqrt{2})$ , único minimizador global deste problema. O ponto  $(-\sqrt{2}, \sqrt{2})$  é, portanto, o minimizador global otimista do problema.

Seja  $U_\delta(-\sqrt{2}) = ]-\sqrt{2}-\delta, -\sqrt{2}+\delta[$  a vizinhança aberta de raio  $\delta > 0$  em torno de  $-\sqrt{2}$ . Neste problema  $X = \mathbb{R}^{n_1}$ , o que implica que  $U_\delta(-\sqrt{2}) \cap X = U_\delta(-\sqrt{2})$ . Logo, para todo  $\delta > 0$  é possível encontrar  $x = -\sqrt{2} - \xi$  com  $0 < \xi < \delta$ , tal que  $x \in U_\delta(-\sqrt{2})$ ,  $x \geq -2$  (caso contrário o conjunto  $\Psi(x)$  seria vazio) e ainda

$$\min_{y'} \{F(x, y') \mid y' \in \Psi(x)\} = -\sqrt{2} - \xi < -\sqrt{2} = \min_{y'} \{F(-\sqrt{2}, y') \mid y' \in \Psi(-\sqrt{2})\},$$

como pode ser visto na Figura 2.3. Logo, este ponto não é um minimizador local otimista, de acordo com a definição apresentada em [CMS07].

Outra definição de minimizador local otimista pode ser encontrada em [Dem02], onde os problemas em dois níveis considerados não possuem restrições no primeiro nível que dependem de variáveis do segundo nível. O BLPP é definido como se segue:

$$\begin{array}{ll} \text{Minimizar} & F(x, y) \\ \text{s. a} & x \in X \\ & y \in \arg \min_{y'} f(x, y') \\ & \text{s. a} \quad h(x, y') = 0 \\ & \quad g(x, y') \leq 0. \end{array} \tag{2.8}$$

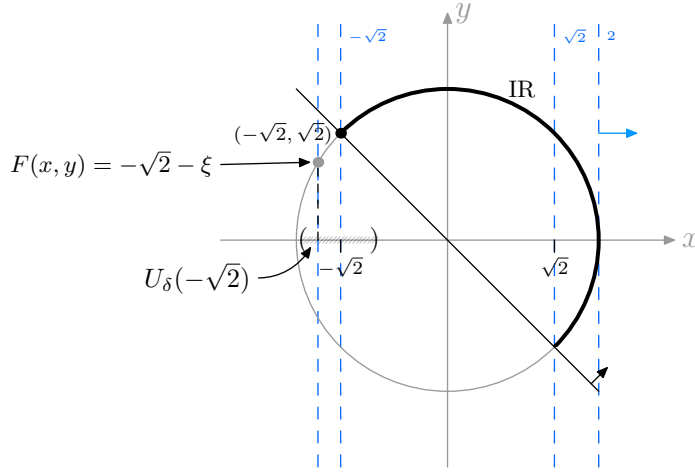


Figura 2.3: No intervalo aberto  $U_\delta(-\sqrt{2})$  em torno de  $-\sqrt{2}$  existe  $-\sqrt{2} - \xi$ ,  $\xi > 0$ , tal que, para  $\tilde{y} \in \Psi(-\sqrt{2} - \xi)$ ,  $F(-\sqrt{2} - \xi, \tilde{y}) < F(-\sqrt{2}, \sqrt{2})$ . Note que  $(-\sqrt{2} - \xi, \tilde{y})$  não é um ponto viável do BLPP. As **retas tracejadas** representam as curvas de nível de  $F(x, y)$ .

De acordo com o autor,  $(\bar{x}, \bar{y}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$  é um minimizador local otimista se

$$\begin{aligned} \bar{x} &\in X, \bar{y} \in \Psi(\bar{x}), \\ F(\bar{x}, \bar{y}) &\leq F(\bar{x}, y) \quad \forall y \in \Psi(\bar{x}) \end{aligned}$$

e, além disso, existe uma vizinhança aberta  $U_\delta(\bar{x})$ ,  $\delta > 0$ , com centro em  $\bar{x}$  tal que

$$\min_{y'} \{F(\bar{x}, y') \mid y' \in \Psi(\bar{x})\} \leq \min_{y'} \{F(x, y') \mid y' \in \Psi(x)\} \quad \forall x \in U_\delta(\bar{x}) \cap X.$$

O ponto  $(\bar{x}, \bar{y})$  é um minimizador global otimista se uma vizinhança de tamanho  $\delta = \infty$  puder ser escolhida.

Intuitivamente, qualquer definição de minimizador local necessita que a função objetivo tenha valor mínimo quando comparada com uma vizinhança de pontos em torno do minimizador. Porém, os pontos devem ser viáveis, motivo pelo qual fazemos a intersecção da vizinhança com a região viável do problema. A definição de minimizador local otimista apresentada em [CMS07] considera pontos que pertencem a uma vizinhança do minimizador, mas não são necessariamente viáveis do problema em dois níveis, pois não leva em conta todas as restrições do primeiro nível. Desta forma, incorretamente o minimizador global otimista do problema não é classificado como tal.

Vale ressaltar que a definição encontrada em [CMS07] é correta se considerarmos o problema (2.4) sem as restrições do primeiro nível dependentes de variáveis do segundo nível. Neste caso, a definição de [CMS07] coincide com a definição de [Dem02]. Isso nos leva a crer que a definição [CMS07] resulta de uma tentativa frustrada de estender a definição de Dempe para problemas da forma (2.4), como mostramos nos Exemplos 2.1.2 e 2.1.3. Este fato foi a motivação para que construíssemos uma definição baseada em [Dem02] que pudesse definir minimizadores locais para os problemas gerais.



## 2.1.4 DEFINIÇÃO

Um ponto  $(\bar{x}, \bar{y})$  é denominado **minimizador local otimista** do problema (2.4) se

$$\begin{aligned} (\bar{x}, \bar{y}) &\in \text{IR}, \\ F(\bar{x}, \bar{y}) &\leq F(\bar{x}, y) \quad \forall y \in \text{IR}(\bar{x}) \end{aligned}$$

e, além disso, existe uma vizinhança aberta  $U_\delta(\bar{x})$ ,  $\delta > 0$ , em torno de  $\bar{x}$ , tal que

$$\min_{y'} \{F(\bar{x}, y') \mid y' \in \text{IR}(\bar{x})\} \leq \min_{y'} \{F(x, y') \mid y' \in \text{IR}(x)\} \quad \forall x \in \Omega_x \cap U_\delta(\bar{x}).$$

O problema considerado neste trabalho é o BLPP (2.1). Apesar de sua formulação também ser considerada otimista, ela se apresenta em uma forma relaxada de (2.5), de modo que a denominaremos formulação **otimista relaxada**. Temos que os minimizadores globais da abordagem otimista (problema (2.5)) e da otimista relaxada (problema (2.1)) coincidem. Esta afirmação já não é verdadeira para os minimizadores locais, como mostra o Exemplo 2.1.5.

## 2.1.5 EXEMPLO (ABORDAGENS OTIMISTAS)

Considere o BLPP apresentado no Exemplo 2.1.3. Primeiramente, pode-se verificar que o minimizador global otimista  $(\bar{x}, \bar{y}) = (-\sqrt{2}, \sqrt{2})$  é considerado como tal, de acordo com a Definição 2.1.4.

Com a abordagem otimista relaxada, resolve-se o seguinte problema:

$$\begin{aligned} &\text{Minimizar}_{x,y} \quad x \\ \text{s. a} \quad &x + y \geq 0 \\ &y \in \Psi(x), \end{aligned}$$

que está representado na Figura 2.4 e cujo minimizador global é o ponto  $(\bar{x}, \bar{y}) = (-\sqrt{2}, \sqrt{2})$ , o mesmo encontrado na abordagem otimista. O problema também possui um minimizador local em  $(\hat{x}, \hat{y}) = (\sqrt{2}, -\sqrt{2})$ , de acordo com a definição de minimizador local em programação não linear, mas este ponto não é um minimizador local otimista, de acordo com a Definição 2.1.4, pois para toda vizinhança aberta  $U_\delta(\sqrt{2})$ ,  $\delta > 0$ , centrada em  $\sqrt{2}$ , existe  $x = \sqrt{2} - \xi$ , com  $0 < \xi < \min\{\delta, 2\sqrt{2}\}$ , tal que

$$\min_{y'} \{\sqrt{2} \mid y' \in \text{IR}(\sqrt{2})\} = \sqrt{2} > \sqrt{2} - \xi = \min_{y'} \{\sqrt{2} - \xi \mid y' \in \text{IR}(\sqrt{2} - \xi)\}.$$

O termo  $\min\{\delta, 2\sqrt{2}\}$  garante que  $\sqrt{2} - \xi$  pertença à vizinhança  $U_\delta(\sqrt{2})$  e também pertença ao conjunto  $\Omega_x$  que neste caso é dado por  $\{x \mid -\sqrt{2} \leq x \leq 2\}$ . Veja a Figura 2.4.

Assim como fizemos com a abordagem otimista, também estamos interessados em discutir uma definição de minimizadores para a abordagem pessimista do problema geral (2.4). Devida à grande similaridade entre as duas abordagens, faremos sempre que possível analogia com as definições e conceitos usados na abordagem otimista.

Em [CMS07], no qual o problema (2.7) é considerado,  $(\bar{x}, \bar{y}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$  é denominado **minimizador local pessimista** se

$$\begin{aligned} \bar{x} &\in X, \bar{y} \in \Psi(\bar{x}), \\ G(\bar{x}, \bar{y}) &\leq 0, \\ F(\bar{x}, \bar{y}) &\geq F(\bar{x}, y) \quad \forall y \in \Psi(\bar{x}) \end{aligned}$$



O ponto é um **minimizador global pessimista** se uma vizinhança arbitrariamente grande puder ser escolhida.

No exemplo a seguir, mostramos que a definição encontrada em [CMS07] de maneira incorreta não classifica um minimizador global pessimista como tal, enquanto a Definição 2.1.6 o faz adequadamente. Notemos que não é possível fazer uma relação entre a abordagem pessimista e o problema (2.1) pois sua interpretação é implicitamente otimista.

### 2.1.7 EXEMPLO

Considere o problema em dois níveis apresentado no Exemplo 2.1.3. Sua formulação como um problema pessimista é como se segue:

$$\begin{array}{ll} \text{Min}_{x} & \text{Max}_{y} \quad x \\ \text{s. a} & x + y \geq 0 \\ & y \in \Psi(x). \end{array}$$

Este problema está ilustrado na Figura 2.3. Como a função objetivo  $F(x, y) = x$  não depende de  $y$ , é simples ver que o minimizador global deste problema pessimista é  $(\bar{x}, \bar{y}) = (-\sqrt{2}, \sqrt{2})$ . Através de argumentos semelhantes aos usados no Exemplo 2.1.3, pode-se verificar que, de acordo com a definição encontrada em [CMS07], o ponto  $(\bar{x}, \bar{y})$  não é um minimizador local pessimista. Se a Definição 2.1.6 for usada, o ponto é um minimizador global pessimista.

Terminamos o assunto sobre as diferentes abordagens de um problema em dois níveis com o Exemplo 2.1.8, que mostra um problema no qual o minimizador global otimista não coincide com o minimizador global pessimista.

### 2.1.8 EXEMPLO

Seja o seguinte problema de programação em dois níveis:

$$\begin{array}{ll} \text{Minimizar}_{x,y} & x + y \\ \text{s. a} & y \in \arg \min_{y'} -y'^2 \\ & \text{s. a} \quad y'^2 \leq 4 - x^2. \end{array}$$

Já se sabe que o conjunto  $\Psi(x)$  do problema do seguidor é dado por (veja o Exemplo 2.1.3):

$$\Psi(x) = \begin{cases} \{\pm\sqrt{4-x^2}\}, & \text{se } -2 \leq x \leq 2 \\ \emptyset, & \text{c.c.} \end{cases}$$

Neste caso, no qual o problema em dois níveis não possui restrições no primeiro nível, o conjunto IR é composto pela união de  $\Psi(x)$ , para todo  $x$ , e está ilustrado na Figura 2.5. Observe agora a abordagem otimista e a pessimista do problema:

Otimista	Pessimista
$\begin{array}{ll} \text{Min}_{x} & \text{Min}_{y} \quad x + y \\ \text{s. a} & y \in \Psi(x) \end{array}$	$\begin{array}{ll} \text{Min}_{x} & \text{Max}_{y} \quad x + y \\ \text{s. a} & y \in \Psi(x). \end{array}$

O problema otimista parte do pressuposto de que o segundo nível sempre escolherá a solução que minimize a função objetivo do primeiro nível, quando há a possibilidade de escolha. No presente

exemplo, dado  $x$ , dentre as duas possíveis escolhas existentes no conjunto  $\Psi(x) = \{-\sqrt{4-x^2}, \sqrt{4-x^2}\}$ , o segundo nível escolherá  $-\sqrt{4-x^2}$ . Deste modo, o problema otimista resume-se a

$$\begin{array}{ll} \text{Minimizar} & x + y \\ \text{s. a} & y = -\sqrt{4-x^2} \end{array}$$

e o minimizador global otimista encontrado é  $(\bar{x}_o, \bar{y}_o) = (-\sqrt{2}, -\sqrt{2})$ . Note que este ponto é um minimizador global otimista de acordo com a definição encontrada em [Dem02] e com a Definição 2.1.4.

Por outro lado,  $(\bar{x}_o, \bar{y}_o) = (-\sqrt{2}, -\sqrt{2})$  não é minimizador global pessimista de acordo com a Definição 2.1.6, pois  $F(\bar{x}_o, \bar{y}_o) < F(\bar{x}_o, \sqrt{4-x_o^2})$ , ou seja, para o mesmo  $\bar{x}_o$  existe outro  $y \in \Psi(\bar{x}_o)$  tal que a função objetivo do primeiro nível tem valor maior. Na abordagem pessimista, o líder supõe que o segundo nível sempre escolherá a solução com o maior valor da função objetivo, o que implica neste exemplo, que, dado  $x$ , o segundo nível escolherá  $\sqrt{4-x^2}$  dentre as soluções em  $\Psi(x)$ . Logo, o problema pessimista resume-se a

$$\begin{array}{ll} \text{Minimizar} & x + y \\ \text{s. a} & y = \sqrt{4-x^2}, \end{array}$$

cujo minimizador global é  $(\bar{x}_p, \bar{y}_p) = (-2, 0)$ . Note que este ponto é um minimizador global pessimista de acordo com a Definição 2.1.6.

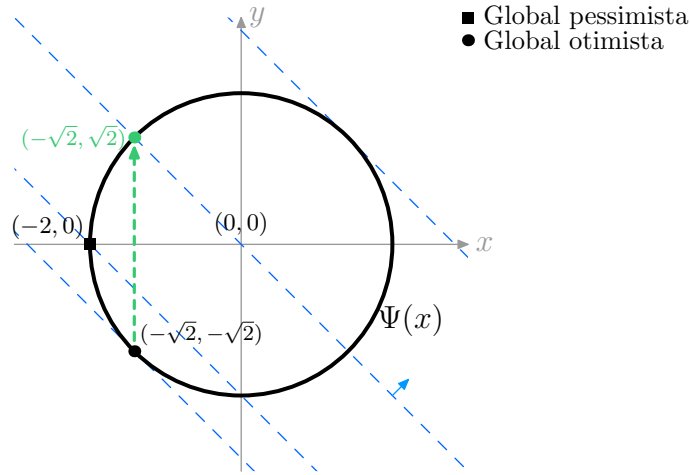


Figura 2.5: A abordagem otimista e a abordagem pessimista podem resultar em diferentes minimizadores globais. Neste exemplo, o minimizador global otimista  $(-\sqrt{2}, -\sqrt{2})$  não é minimizador pessimista pois  $F(-\sqrt{2}, \sqrt{2})$  possui maior valor funcional que  $F(-\sqrt{2}, -\sqrt{2})$ .

Os BLPPs são classificados de acordo com as funções que os compõem. Se todas as funções, tanto do líder quanto do seguidor, forem lineares, o problema é denominado problema de programação em dois níveis linear, BLPP linear, BLPP linear-linear ou apenas BLP. Se o problema do primeiro nível for linear e o problema do segundo nível for de programação quadrática, denomina-se BLPP linear-quadrático. Se ambos os problemas forem de programação quadrática, denomina-se BLPP quadrático. Caso as funções não possuam classificação, ou haja mistura com outros tipos de função, o problema é simplesmente BLPP (geral).

Os problemas do tipo BLPP linear são considerados os mais simples mas, mesmo assim, Ben-Ayed e Blair [BAB90] provaram que são NP-difíceis, de forma que não se conhecem algoritmos que os resolvam em tempo polinomial. A prova dessa afirmação consiste em mostrar que se existisse um algoritmo que resolvesse o BLPP linear em tempo polinomial, então seria possível resolver o problema da mochila em tempo polinomial. Como o problema da mochila é um problema comprovadamente NP-difícil, tal algoritmo não deve existir.

## 2.2 Métodos

Da maneira como os BLPPs são definidos não é possível utilizar algoritmos de programação não linear para resolvê-los. Um grande esforço foi realizado por parte do meio científico para analisar a estrutura de cada tipo de BLPP e desenvolver um método eficiente de resolvê-lo. Atualmente existe uma grande variedade de métodos que resolve diversos tipos de problemas em dois níveis, mas devido à estrutura complexa destes problemas nem sempre é possível mostrar convergência para minimizadores locais ou pontos estacionários.

Nesta seção fazemos um pequeno resumo de alguns métodos desenvolvidos para programação em dois níveis. O objetivo não é rever com detalhes todos os métodos existentes e sim tentar mostrar quais são os tipos de abordagens utilizadas e algumas técnicas novas encontradas na literatura.

### 2.2.1 Algoritmo *K-th Best*

O algoritmo *K-th Best* foi introduzido por Bialas e Karwan [BK78] para resolver o BLPP linear. Não é necessário que o conjunto  $\Psi(x)$  seja unitário, mas se ele o for para todo  $x$ , e a região induzida for limitada, então o algoritmo encontra o minimizador global do problema em dois níveis.

Dado que o BLPP é linear, é possível mostrar que sua solução encontra-se em um dos vértices de sua região viável (veja [BK78]). Logo, o funcionamento do algoritmo consiste em caminhar pelos vértices do poliedro em busca do vértice  $(x^K, y^K)$  com menor valor de  $F(x, y)$  e que penceça à região induzida IR.

Para realizar tal tarefa, o primeiro vértice é obtido com a minimização da função objetivo do primeiro nível  $F(x, y)$  sobre a região viável do BLPP. Este é um problema de programação linear que pode ser resolvido com muitas técnicas conhecidas. Se o vértice encontrado pertence a IR o algoritmo pára, caso contrário, seus vértices adjacentes são enumerados e sucessivamente investigados. Para garantir que o algoritmo pára no vértice com menor valor da função objetivo, os vértices são percorridos em ordem crescente de  $F(x, y)$ .

### 2.2.2 Reformulação utilizando condições KKT

Uma técnica bastante conhecida para a resolução de problemas em dois níveis é a substituição do problema do segundo nível por suas condições KKT e conseqüente resolução do problema de programação não linear formado [BF82]. Como será visto nos métodos a seguir, grande parte das abordagens usadas na resolução de problemas em dois níveis se baseia na reformulação do problema do segundo nível por suas condições KKT. O BLPP (2.1) é transformado no seguinte problema de programação não linear:

$$\begin{aligned}
& \underset{x,y,\lambda,\mu}{\text{Minimizar}} && F(x,y) \\
& \text{s. a} && \nabla_y L_f(x,y,\lambda,\mu) = 0 \\
& && H(x,y) = 0 \\
& && G(x,y) \leq 0 \\
& && h(x,y) = 0 \\
& && g(x,y) \leq 0 \\
& && \mu_i g_i(x,y) = 0, \quad i = 1, \dots, p_2 \\
& && \mu \geq 0,
\end{aligned} \tag{2.9}$$

onde  $\lambda \in \mathbb{R}^{m_2}$ ,  $\mu \in \mathbb{R}^{p_2}$  e denotamos por  $L_f$  a função Lagrangiana associada ao problema do segundo nível. Notemos que os multiplicadores de Lagrange tornaram-se variáveis do problema reformulado.

Embora esta abordagem permita a qualquer BLPP ser reformulado como um PNL, o problema reformulado só é equivalente ao original se as condições KKT forem necessárias e suficientes para os minimizadores globais do problema do segundo nível. Isto ocorre, por exemplo, no caso do BLPP linear.

Em [BF82] esta abordagem é discutida e um algoritmo para o caso em que todas as funções do BLPP são lineares é proposto.

### 2.2.3 Branch and bound

As restrições de complementaridade presentes no problema reformulado (2.9) são intrinsecamente combinatoriais o que sugere o uso de algoritmos enumerativos como o *branch and bound*. Em [Bar88] o problema considerado é o BLPP convexo, ou seja, todas as suas funções são convexas e além disso, contínuas e duas vezes diferenciáveis. Também é suposto que o conjunto  $\Psi(x)$  é unitário para todo  $x$  (o que implica que  $y$  pode ser substituído por  $y(x)$ ), o problema do segundo nível é um problema de programação quadrática e as restrições do primeiro nível não dependem da variável  $y$ .

Sob estas hipóteses um método do tipo *branch and bound* é apresentado, no qual o passo de *branching* é realizado nas restrições de complementaridade. A cada nó, ou  $\mu_i = 0$  ou  $g_i(x,y) = 0$  é adicionada ao problema (2.9) que aqui é considerado sem suas restrições de complementaridade. O limitante inferior é calculado resolvendo-se o problema reformulado (2.9) sem as restrições de complementaridade e o limitante superior é obtido, dentre outras maneiras, através de uma determinação bem sucedida de valores para  $\mu_i$  e  $g_i(x,y)$ .

Seguindo a mesma idéia, Bard e Moore [BM90] aplicam a abordagem ao BLPP linear-quadrático, mas retiram a hipótese de  $\Psi(x)$  ser unitário para todo  $x$ . Em [Bar98], Bard modifica este método para lidar com funções mais gerais, supondo convexidade de  $F(x,y)$ , convexidade estrita de  $f(x,y)$  e algumas restrições sobre o conjunto viável do segundo nível.

### 2.2.4 Penalidades

Nos métodos de penalidades, o problema em dois níveis é substituído pelo problema de minimizar uma ou mais funções penalidade. Os trabalhos de Shimizu e Aiyoshi [SA81, AS84] foram os primeiros a resolver problemas em dois níveis gerais através desta abordagem. No

método de penalidade dupla discutido em [AS84] a função penalidade tenta satisfazer a viabilidade tanto do primeiro nível quanto do segundo nível e ainda satisfazer condições de primeira ordem do problema do segundo nível penalizado. De maneira simplificada, a seguinte função é minimizada em  $x$  e  $y$ :

$$Q(x, y; r, s, t) \doteq F(x, y) + t\phi_1(G(x, y)) + r\phi_2(g(x, y)) + s\phi_3(\|\nabla_y p(x, y; r)\|),$$

onde  $\phi(x)$  é a função penalidade e  $p(x, y; r)$  é outra função penalidade apenas para o problema do segundo nível.

Por se tratarem de problemas em dois níveis gerais, diversas hipóteses foram feitas. Primeiramente não é possível trabalhar com restrições de igualdade apenas transformando-as em duas restrições de desigualdade, pois o conjunto  $\{y \mid g(x, y) < 0\}$ ,  $\forall x$ , deve ser não vazio. Além disso, o conjunto viável do seguidor deve ser compacto e  $f(x, y)$  deve ser estritamente convexa, o que implica no conjunto  $\Psi(x)$  ser unitário para todo  $x$ . Com tais hipóteses, é possível mostrar que a sucessiva atualização das penalidades faz o algoritmo convergir a um ponto  $\hat{y}$  que é minimizador do problema do segundo nível para  $\hat{x}$ .

Também usando penalidades, Anandalingam e White [AW90] consideram apenas o BLPP linear. Seu método difere do método de Aiyoshi e Shimizu pois uma função de penalidade exata é utilizada. A nova função a ser minimizada penaliza a folga de dualidade, ou seja, a diferença entre o valor da função objetivo do primal e do dual do problema de programação linear. Como hipóteses principais, além do problema ser linear, pede-se que o problema do segundo nível possua solução (não necessariamente única) para todo  $x$  viável.

### 2.2.5 Direções de descida

Os métodos baseados em direções de descida para programação em dois níveis são muito similares aos encontrados em programação não linear. Basicamente, estes métodos calculam uma direção que faz produto interno negativo com o gradiente da função objetivo do primeiro nível e em seguida escolhem o tamanho do passo que será dado nesta direção para que o novo ponto seja viável. Os passos são repetidos até que alguma condição de otimalidade seja satisfeita.

Savard e Gauvin [SG94] desenvolveram um método de máxima descida para o BLPP geral sem restrições no primeiro nível. Para que o método possa ser definido, supõe-se que o conjunto  $\Psi(x)$  seja unitário para todo  $x$ . Além disso, para todo  $(x, y)$ , os gradientes das restrições do segundo nível ativas em  $(x, y)$  devem ser linearmente independentes e a condição suficiente de segunda ordem (Teorema 1.2.6) deve ser satisfeita no problema do seguidor. Devemos ressaltar que no segundo nível a variável  $x$  é considerada uma constante, logo as derivadas são calculadas apenas em relação a  $y$ .

A direção de máxima descida é a solução de um BLPP quadrático, para o qual existem métodos eficientes de resolução ([BM90], por exemplo). Dada a direção de descida  $d^k = (d_x^k, d_y^k)$ , o cômputo do tamanho do passo consiste em calcular  $(x', y') = (x^k, y^k) + \alpha(d_x^k, d_y^k)$  para algum  $\alpha$ . Esta é uma parte custosa do método visto que, para  $x'$  viável, verificar se  $y'$  é uma solução global do problema do segundo nível significa resolver um problema de otimização. Sabe-se, porém, que esta busca é finita pois resultados na literatura mostram que, sob as hipóteses acima, sempre existe  $\bar{\alpha}$  tal que, para  $0 \leq \alpha < \bar{\alpha}$ , o ponto  $y'$  é solução do segundo nível.

### 2.2.6 Regiões de confiança

Analogamente aos métodos de direções de descida, os métodos de regiões de confiança tentam se aproveitar de conceitos conhecidos em programação não linear para aplicá-los na resolução de problemas em dois níveis.

Recentemente, Colson, Marcotte e Savard [CMS05] apresentaram um método baseado em regiões de confiança para resolver o BLPP geral cujas restrições do primeiro nível não contêm variáveis do segundo nível. Primeiramente o BLPP é transformado em um BLPP linear-quadrático, através da expansão de Taylor de primeira ordem da função  $F(x, y)$  e das restrições do segundo nível em torno de um ponto  $(\hat{x}, \hat{y})$  e de uma expansão de segunda ordem da função  $f(x, y)$  em torno de  $(\hat{x}, \hat{y})$ .

O BLPP linear-quadrático é então reformulado utilizando as condições KKT. A restrição de região de confiança também é adicionada ao problema reformulado. Por fim, as restrições de complementaridade são substituídas por restrições lineares contendo variáveis inteiras, transformando o problema reformulado em um problema de programação inteira mista (MIP), para o qual existem métodos de resolução eficientes.

A cada passo o método avalia a razão entre a melhoria encontrada com a resolução do modelo (o problema MIP) e a melhoria real obtida na função  $F(x, y)$  original, atualizando o tamanho da região de confiança apropriadamente. Apesar de toda a estrutura de regiões de confiança utilizada, não há prova de convergência do método a um minimizador local do problema em dois níveis.

### 2.2.7 Restauração inexata

Os problemas em dois níveis muitas vezes estão associados a problemas de programação matemática com restrições de equilíbrio, motivo pelo qual a viabilidade pode ser muito importante. Com base nestas premissas, em [Cas04, ACC<sup>+</sup>07] uma técnica é proposta para resolver problemas em dois níveis usando algoritmos de restauração inexata.

Os algoritmos de restauração inexata foram introduzidos em [Mar98, MP00] e consideram separadamente duas fases no processo de otimização: a restauração e a minimização. Na fase de restauração o objetivo do método é melhorar a viabilidade do ponto atual. Na segunda fase, denominada fase de minimização, um ponto que diminui o valor da função objetivo é calculado na intersecção de uma região de confiança centrada na solução da fase anterior e uma região que aproxima o conjunto viável do problema. O ponto encontrado na segunda fase é comparado com o ponto atual através de uma função penalidade e se o novo ponto não for aceito, o raio da região de confiança é reduzido.

Em [ACC<sup>+</sup>07], os autores consideram o BLPP geral sem restrições dependentes de  $y$  no primeiro nível e somente com restrições de caixa e de igualdade no segundo nível. Entre as diversas hipóteses utilizadas, pede-se que o conjunto viável do segundo nível seja compacto (isso implica que para cada  $x$  exista uma solução global do segundo nível).

Na fase de restauração, um ponto viável é encontrado resolvendo-se o problema do segundo nível com algum algoritmo de programação não linear. Embora o ideal seja utilizar um algoritmo de otimização global nesta fase, o método necessita apenas de um ponto  $\bar{y}$  que resolva aproximadamente o problema do segundo nível e também encontre uma aproximação dos seus multiplicadores de Lagrange associados. Já na fase de minimização, o sistema KKT associado ao problema do segundo nível é utilizado para construir a função Lagrangiana



que auxiliará na escolha de uma direção que minimize a função objetivo do primeiro nível e mantenha o ponto com uma certa viabilidade.

A vantagem do método é que ele preserva a estrutura do problema do segundo nível, que é resolvido sem reformulações na fase de restauração. Além disso não é necessário usar métodos de otimização não diferenciável. Sob as hipóteses feitas, os autores mostram que o método converge para pontos viáveis que satisfazem a condição de qualificação da projeção aproximada do gradiente (AGP).



## Capítulo 3

# Reformulação através das condições KKT

O uso das condições de Karush-Kuhn-Tucker para transformar um problema de programação em dois níveis em um problema de programação não linear é a forma mais intuitiva para resolver problemas em dois níveis. Como apresentamos no Capítulo 2, a reformulação não aparece apenas como um método por si só, mas também é utilizada como base para diversas abordagens. Abordagens utilizando *branch and bound*, métodos de restauração inexata e até métodos de regiões de confiança têm seu funcionamento baseado nas condições KKT.

Um dos primeiros trabalhos a usar as condições KKT foi [BF82]. No trabalho, uma técnica de *branch and bound* foi aplicada ao problema reformulado, pois as restrições de complementaridade presentes no sistema KKT permitem uma abordagem enumerativa eficiente do problema.

Na teoria também encontram-se aplicações para o problema reformulado. Muitas das condições de otimalidade para problemas em dois níveis são derivadas com o auxílio das condições KKT. Podemos citar, por exemplo, [Bar98] e, principalmente, [Dem02].

Por ser uma abordagem simples, sabemos que pode resolver uma variedade pequena de problemas, pois os problemas reformulados possuem propriedades indesejáveis para os algoritmos de programação não linear. Neste capítulo, nos preocupamos em estudar as possibilidades existentes para resolver problemas em dois níveis através da reformulação. Na Seção 3.1 reapresentamos os problemas reformulados. O foco agora é mostrar as vantagens e desvantagens de utilizá-los. Na Seção 3.2 introduzimos outras formas de representar as restrições de complementaridade do sistema KKT e na Seção 3.3 discutimos alguns métodos que podem ser aplicados à reformulação.

### 3.1 O problema reformulado

Vamos estudar o problema em dois níveis definido por:

$$\begin{aligned}
 & \underset{x,y}{\text{Minimizar}} && F(x,y) \\
 & \text{s. a} && H(x,y) = 0 \\
 & && G(x,y) \leq 0 \\
 & && y \in \underset{y'}{\arg \min} f(x,y') \\
 & && \text{s. a} \quad h(x,y') = 0 \\
 & && g(x,y') \leq 0,
 \end{aligned} \tag{3.1}$$

onde  $x \in \mathbb{R}^{n_1}$ ,  $y \in \mathbb{R}^{n_2}$ ,  $F(x,y) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}$ ,  $G(x,y) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{p_1}$ ,  $H(x,y) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{m_1}$ ,  $f(x,y') : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}$ ,  $g(x,y') : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{p_2}$  e  $h(x,y') : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mapsto \mathbb{R}^{m_2}$ , de acordo com o apresentado no Capítulo 2.

Quando o problema do segundo nível do BLPP (3.1) é substituído por suas condições KKT, temos o seguinte problema de programação não linear, denominado **problema reformulado** (ou reformulação):

$$\begin{aligned}
 & \underset{x,y,\lambda,\mu}{\text{Minimizar}} && F(x,y) \\
 & \text{s. a} && \nabla_y L_f(x,y,\lambda,\mu) = 0 \\
 & && G(x,y) \leq 0 \\
 & && H(x,y) = 0 \\
 & && g(x,y) \leq 0 \\
 & && h(x,y) = 0 \\
 & && \mu_i g_i(x,y) = 0, \quad i = 1, \dots, p_2 \\
 & && \mu \geq 0,
 \end{aligned} \tag{3.2}$$

onde  $\mu \in \mathbb{R}^{p_2}$ ,  $\lambda \in \mathbb{R}^{m_2}$  e  $L_f$  representa a função Lagrangiana associada ao problema do segundo nível:

$$L_f(x,y,\lambda,\mu) = f(x,y) + \sum_{i=1}^{m_2} \lambda_i h_i(x,y) + \sum_{i=1}^{p_2} \mu_i g_i(x,y).$$

Notemos que o número de variáveis aumentou de  $n_1 + n_2$  para  $n_1 + n_2 + m_2 + p_2$ , pois os multiplicadores de Lagrange tornaram-se variáveis do problema. Além disso, o número de restrições também aumentou de  $m_1 + p_1 + m_2 + p_2$  para  $m_1 + p_1 + m_2 + 3p_2 + n_2$ .

É importante destacarmos que no problema em dois níveis o conjunto IR no qual a função  $F(x,y)$  é minimizada está contido no espaço  $\mathbb{R}^{n_1+n_2}$  (espaço das variáveis  $x$  e  $y$ ), enquanto que no problema reformulado o conjunto no qual a função  $F(x,y)$  é minimizada está contido no  $\mathbb{R}^{n_1+n_2+m_2+p_2}$ . Desta forma, é um abuso da definição compararmos os conjuntos de cada problema. Quando mencionamos que os conjuntos são equivalentes, ou que um ponto na região induzida do BLPP pertence ao conjunto viável do problema reformulado, estamos implicitamente dizendo que dado  $(x,y) \in \text{IR}$  no BLPP existem multiplicadores  $\lambda$  e  $\mu$  tais que o ponto  $(x,y,\lambda,\mu)$  satisfaz as restrições de (3.2).

Embora seja possível aplicar um algoritmo de programação não linear para resolver (3.2), tal problema é equivalente ao problema em dois níveis original se e somente se os minimizadores globais do segundo nível satisfazem as condições KKT. Em outras palavras, todos os minimizadores globais devem satisfazer as condições necessárias e suficientes de primeira ordem do problema do segundo nível. O Exemplo 3.1.1 ilustra um caso no qual isto ocorre.

### 3.1.1 EXEMPLO

Considere o BLPP definido a seguir:

$$\begin{array}{ll} \text{Minimizar}_{x,y} & x^2 + y^2 \\ \text{s. a} & y \in \arg \min_{y'} -y'^2 \\ & \text{s. a} \quad 1 \leq y' \leq 2. \end{array}$$

Primeiramente pode-se notar que, de acordo com a Proposição 1.2.4, as condições KKT são necessárias e suficientes para o problema do segundo nível, pois a função  $-y^2$  é pseudo-convexa no intervalo  $[1, 2]$ . O minimizador do problema do segundo nível é o ponto  $\bar{y} = 2$ , independentemente do valor de  $x$ . Logo, a região induzida do BLPP é dada por  $\text{IR} = \{(x, y) \mid x \in \mathbb{R} \text{ e } y = 2\}$  e o valor de mínimo é atingido no ponto  $(\bar{x}, \bar{y}) = (0, 2)$ .

Considere agora a reformulação do problema, formada pela substituição do problema do segundo nível pelas suas condições KKT:

$$\begin{array}{ll} \text{Minimizar}_{x,y,\mu} & x^2 + y^2 \\ \text{s. a} & -2y - \mu_1 + \mu_2 = 0 \\ & \mu_1(1 - y) = 0 \\ & \mu_2(y - 2) = 0 \\ & 1 \leq y \leq 2 \\ & \mu_1, \mu_2 \geq 0. \end{array}$$

Independentemente do valor de  $x$ , o ponto  $(\bar{y}, \bar{\mu}_1, \bar{\mu}_2) = (2, 0, 4)$  é o único que satisfaz as condições KKT do problema do segundo nível e, conseqüentemente, as restrições da reformulação. Logo, o conjunto viável do problema reformulado é dado por  $\Omega = \{(x, y, \mu_1, \mu_2) \mid x \in \mathbb{R}, y = 2, \mu_1 = 0, \mu_2 = 4\}$ . Observe que  $(\tilde{x}, 2) \in \text{IR}$  se e somente se o ponto  $(\tilde{x}, 2, 0, 4) \in \Omega$ . Como a função objetivo de ambos os problemas é a mesma e independe de  $\mu$ ,  $(\bar{x}, \bar{y}) = (0, 2)$  é o minimizador do problema original se e somente se  $(\bar{x}, \bar{y}, \bar{\mu}_1, \bar{\mu}_2) = (0, 2, 0, 4)$  é o minimizador da reformulação.

Quando existem minimizadores globais do segundo nível que não satisfazem as condições necessárias de primeira ordem, isto significa que tais minimizadores podem não satisfazer as condições KKT. Neste caso, existem pontos da região induzida do problema em dois níveis que não pertencem ao conjunto viável de (3.2). O Exemplo 3.1.2 ilustra este caso.

### 3.1.2 EXEMPLO

Considere o problema definido a seguir:

$$\begin{array}{ll} \text{Minimizar}_{x,y} & (x - 1)^2 + y^2 \\ \text{s. a} & y \in \arg \min_{y'} y' \\ & \text{s. a} \quad y'^2 = 0. \end{array}$$

O minimizador global do problema do segundo nível é  $\bar{y} = 0$ , independentemente do valor de  $x$ , como mostra a Figura 3.1. Isto implica que a solução ótima para este BLPP é o ponto  $(\bar{x}, \bar{y}) = (1, 0)$ .

Considere agora o problema de programação não linear obtido substituindo-se o problema do segundo nível por suas condições KKT:

$$\begin{array}{ll} \text{Minimizar}_{x,y,\lambda} & (x-1)^2 + y^2 \\ \text{s. a} & 1 + 2\lambda y = 0 \\ & y^2 = 0. \end{array}$$

Note que este problema é inviável, pois nenhum ponto viável do problema do segundo nível satisfaz as condições KKT. Logo, não existe  $\lambda$  tal que a solução global do BLPP  $(\bar{x}, \bar{y}) = (1, 0)$  seja um ponto viável da reformulação. Neste exemplo, pode-se abusar da definição para afirmar que o conjunto viável da reformulação está contido na região induzida do problema original mas não o contém.

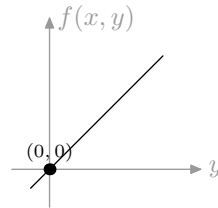


Figura 3.1: O problema do segundo nível só possui um ponto viável, que é sua solução, mas para este ponto não existem multiplicadores de Lagrange.

Quando para todo minimizador global do segundo nível valem as condições necessárias de primeira ordem, mas as condições suficientes não são satisfeitas, então existem pontos que não pertencem à região induzida do BLPP mas estão presentes no conjunto viável do problema reformulado. Neste caso, os minimizadores locais, os pontos de máximo e os pontos de sela do problema do segundo nível são os pontos que podem, eventualmente, satisfazer as condições KKT. Estes pontos, como não são minimizadores globais do segundo nível, não pertencem à região induzida do BLPP, mas tornam-se pontos viáveis da sua reformulação, como mostra o Exemplo 3.1.3.

### 3.1.3 EXEMPLO

Considere o problema em dois níveis definido a seguir.

$$\begin{array}{ll} \text{Minimizar}_{x,y} & x^2 + y^2 \\ \text{s. a} & y \in \arg \min_{y'} -y'^2 \\ & \text{s. a} \quad -2 \leq y' \leq 2. \end{array}$$

As soluções globais do problema do segundo nível são  $\bar{y}^1 = -2$  e  $\bar{y}^2 = 2$ . Logo,  $\text{IR} = \{-2, 2\}$  e as soluções globais do BLPP são  $(\bar{x}^1, \bar{y}^1) = (0, -2)$  e  $(\bar{x}^2, \bar{y}^2) = (0, 2)$ . Veja a Figura 3.2(a). Quando o problema do segundo nível é substituído por suas condições KKT, obtém-se o seguinte problema reformulado:

$$\begin{array}{ll} \text{Minimizar}_{x,y,\mu} & x^2 + y^2 \\ \text{s. a} & -2y - \mu_1 + \mu_2 = 0 \\ & \mu_1(-2 - y) = 0 \\ & \mu_2(y - 2) = 0 \\ & -2 \leq y \leq 2 \\ & \mu_1, \mu_2 \geq 0. \end{array}$$

Note que, equivalentemente às soluções do problema do segundo nível,  $(\bar{y}^1, \bar{\mu}_1^1, \bar{\mu}_2^1) = (-2, 0, 2)$  e  $(\bar{y}^2, \bar{\mu}_1^2, \bar{\mu}_2^2) = (2, 2, 0)$  satisfazem as condições KKT, que neste caso são apenas necessárias, pois o ponto  $(\bar{y}, \bar{\mu}_1, \bar{\mu}_2) = (0, 0, 0)$  também as satisfaz. Desta forma, a solução do problema reformulado é  $(\hat{x}^3, \hat{y}^3, \hat{\mu}_1, \hat{\mu}_2) = (0, 0, 0, 0)$ , ilustrada na Figura 3.2(b). Porém,  $\hat{y}^3 = 0$  corresponde ao maximizador do problema do segundo nível e não pertence à região induzida do BLPP original.

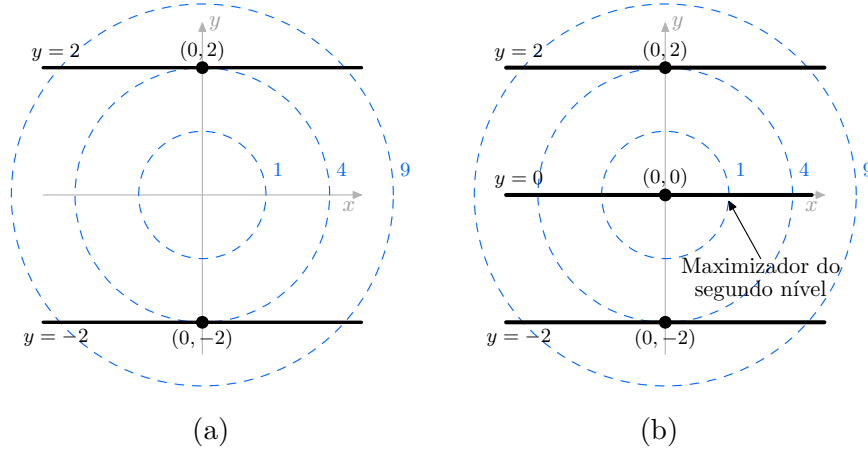


Figura 3.2: (a) O problema original com sua região induzida representada em negrito. (b) O problema reformulado. Nesta figura, o conjunto viável  $\Omega = \{(x, -2, 0, 2)\} \cup \{(x, 2, 2, 0)\} \cup \{(x, 0, 0, 0)\}$  foi projetado no plano  $xy$  e está representado em negrito. As circunferências representam as curvas de nível da função objetivo do primeiro nível.

Por fim, é possível que as condições de otimalidade não sejam satisfeitas para alguns minimizadores do problema do segundo nível (isto acontece quando os minimizadores não satisfazem as condições de qualificação) e sim sejam satisfeitas por pontos indesejáveis (minimizadores locais, maximizadores e pontos de sela). Neste caso, ilustrado no Exemplo 3.1.4, o conjunto viável do problema reformulado não contém alguns pontos que pertencem ao conjunto IR do problema original e contém outros que não pertencem a IR.

#### 3.1.4 EXEMPLO

Considere o problema em dois níveis definido como se segue.

$$\begin{aligned} & \underset{x, y}{\text{Minimizar}} && x^2 + y^2 \\ & \text{s. a} && y \in \underset{y'}{\arg \min} -y'^2 \\ & && \text{s. a} && (y' - 1)^3 \leq 0 \\ & && && y' \geq -0.5. \end{aligned}$$

É fácil ver que a restrição  $(y' - 1)^3 \leq 0$  representa implicitamente a restrição  $y' \leq 1$ . Desta forma, as duas restrições do segundo nível podem ser interpretadas como as restrições de caixa:  $-0.5 \leq y' \leq 1$ . O minimizador global da função  $-y'^2$  sujeita a estas restrições é  $\bar{y}' = 1$ , o que implica que, para todo  $x$ ,  $\Psi(x) = 1$ . Como pode ser visto na Figura 3.3(a), o minimizador global do BLPP é  $(\bar{x}, \bar{y}) = (0, 1)$ . O problema do segundo nível ainda possui um maximizador  $\bar{y}' = 0$  e um minimizador local  $\bar{y}' = -0.5$ . Considere agora a reformulação deste BLPP:

$$\begin{array}{ll}
\text{Minimizar}_{x,y,\mu} & x^2 + y^2 \\
\text{s. a} & -2y + 3(y-1)^2\mu_1 - \mu_2 = 0 \\
& \mu_1(y-1)^3 = 0 \\
& \mu_2(-0.5-y) = 0 \\
& (y-1)^3 \leq 0 \\
& y \geq -0.5 \\
& \mu_1, \mu_2 \geq 0.
\end{array}$$

Primeiramente, é fácil verificar que o minimizador global do segundo nível  $\tilde{y}' = 1$  não satisfaz as condições KKT, sendo um ponto inviável da reformulação. Por outro lado, para todo  $x$ , os pontos  $\tilde{y}' = 0$  e  $\tilde{y}' = -0.5$  satisfazem as condições KKT e tornam-se pontos viáveis da reformulação. A função objetivo do problema reformulado atinge valor mínimo no ponto  $(\tilde{x}, \tilde{y}, \tilde{\mu}_1, \tilde{\mu}_2) = (0, 0, 0, 0)$ , como mostra a Figura 3.3(b). Neste exemplo, a região induzida do BLPP e o conjunto viável do problema reformulado são totalmente disjuntos.

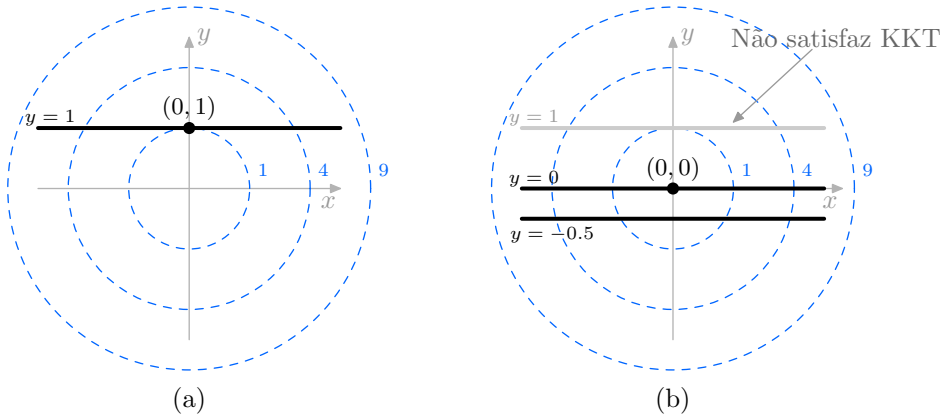


Figura 3.3: A região induzida do BLPP original (a) e o conjunto viável do problema reformulado (b) também podem ser disjuntos. O conjunto viável da reformulação, dado por  $\Omega = \{(x, 0, 0, 0)\} \cup \{(x, -0.5, 0, 1)\}$ , foi projetado no plano  $xy$  para uma melhor comparação.

Para haver equivalência entre o BLPP original e sua reformulação, precisamos garantir que as condições KKT sejam necessárias e suficientes para o problema do segundo nível. Com base na Proposição 1.2.4 apresentada no Capítulo 1, descrevemos um tipo de problema de programação em dois níveis para o qual a substituição do segundo nível pelas suas condições KKT resulta em um problema equivalente. Consideremos o seguinte BLPP:

$$\begin{array}{ll}
\text{Minimizar}_{x,y} & F(x, y) \\
\text{s. a} & G(x, y) \leq 0 \\
& H(x, y) = 0 \\
& y \in \arg \min_{y'} f(x, y') \\
& \text{s. a} \quad g_1(x) + A_1 y' \leq b_1 \\
& \quad \quad h_1(x) + A_2 y' = b_2.
\end{array} \tag{3.3}$$



Como a variável  $x$  é considerada fixa para o problema do segundo nível, se supusermos que a função  $f$  é pseudo-convexa em  $y'$ , para todo  $x$ , podemos aplicar o resultado obtido na Proposição 1.2.4, considerando o problema do segundo nível como um problema de PNL apenas nas variáveis  $y'$ . É fácil notar que na estrutura do problema (3.3) encaixam-se o BLPP linear, o BLPP linear-quadrático e o BLPP quadrático. Para os dois últimos tipos, a função objetivo do segundo nível deve ser pseudo-convexa em  $y'$ . Neste caso, resolver (3.3) é equivalente a resolver o problema:

$$\begin{aligned}
 & \underset{x, y, \lambda, \mu}{\text{Minimizar}} && F(x, y) \\
 & \text{s. a} && \nabla_y f(x, y) + A_1^T \mu + A_2^T \lambda = 0 \\
 & && \mu^T [A_1 y - (b_1 - g_1(x))] = 0 \\
 & && G(x, y) \leq 0 \\
 & && H(x, y) = 0 \\
 & && g_1(x) + A_1 y \leq b_1 \\
 & && h_1(x) + A_2 y = b_2 \\
 & && \mu \geq 0.
 \end{aligned}$$

Como pudemos notar no Exemplo 3.1.3, pontos indesejáveis aparecem no conjunto viável da reformulação devido ao fato das condições KKT serem necessárias, porém não serem suficientes. Utilizando as restrições associadas às condições de segunda ordem apresentadas no Capítulo 1, substituímos o problema do segundo nível por suas condições KKT mais as restrições de segunda ordem (1.5). Logo, o problema do seguidor (2.2) pode ser substituído por:

$$\begin{aligned}
 & \nabla_y L_f(x, y, \lambda, \mu) = 0 \\
 & g(x, y) \leq 0 \\
 & h(x, y) \leq 0 \\
 & \mu_i g_i(x, y) = 0, \quad i = 1, \dots, p_2 \\
 & \nabla_{yy}^2 L_f(x, y, \lambda, \mu) = G G^T \\
 & g_{ii} \geq \varepsilon, \quad i = 1, \dots, n_2 \\
 & \mu \geq 0,
 \end{aligned} \tag{3.4}$$

onde  $\varepsilon > 0$ ,  $G \in \mathbb{R}^{n_2 \times n_2}$  é triangular inferior e  $\nabla_{yy}^2 L_f(x, y, \lambda, \mu)$  é a Hessiana da função Lagrangiana associada ao problema do segundo nível e derivada com relação às variáveis do segundo nível ( $y$ ).

Podemos verificar o efeito de adicionar esta aproximação das condições de segunda ordem no exemplo a seguir.

### 3.1.5 EXEMPLO (UTILIZANDO RESTRIÇÕES DE SEGUNDA ORDEM)

Considere o problema em dois níveis definido como se segue:

$$\begin{aligned}
 & \underset{x, y}{\text{Minimizar}} && x^2 + y^2 \\
 & \text{s. a} && y \in \arg \min_{y'} \frac{y'^4}{4} - \frac{y'^2}{2}.
 \end{aligned}$$

O problema do segundo nível possui soluções  $\bar{y}^1 = -1$  e  $\bar{y}^2 = 1$ . Logo,  $\text{IR} = \{(x, -1)\} \cup \{(x, 1)\}$  e o BLPP possui soluções globais  $(\bar{x}^1, \bar{y}^1) = (0, -1)$  e  $(\bar{x}^2, \bar{y}^2) = (0, 1)$ . Quando o problema do segundo nível é substituído por suas condições KKT, obtém-se o seguinte problema:

$$\begin{array}{ll} \text{Minimizar}_{x,y} & x^2 + y^2 \\ \text{s. a} & y^3 - y = 0. \end{array}$$

A restrição  $y^3 - y = 0$  deste problema equivale a pedir que  $y \in \{-1, 0, 1\}$ . Porém, o ponto  $\tilde{y} = 0$  é um maximizador local da função do segundo nível. Como  $\tilde{y}$  pertence ao conjunto viável do problema reformulado, a solução global  $(\tilde{x}, \tilde{y}) = (0, 0)$  é obtida. Quando o problema do segundo nível é substituído por suas condições de segunda ordem (3.4), obtém-se o seguinte problema:

$$\begin{array}{ll} \text{Minimizar}_{x,y,g_{11}} & x^2 + y^2 \\ \text{s. a} & y^3 - y = 0 \\ & 3y^2 - 1 = g_{11}^2 \\ & g_{11} \geq \varepsilon. \end{array}$$

Tome  $0 < \varepsilon < \sqrt{2}$ . Neste caso,  $\tilde{y} = 0$  não satisfaz o novo conjunto de restrições, enquanto que os minimizadores  $\bar{y}^1 = -1$  e  $\bar{y}^2 = 1$  satisfazem. Desta maneira, é possível encontrar as soluções  $(\bar{x}^1, \bar{y}^1, \bar{g}_{11}^1) = (0, -1, \sqrt{2})$  e  $(\bar{x}^2, \bar{y}^2, \bar{g}_{11}^2) = (0, 1, \sqrt{2})$ , que são equivalentes às soluções do problema em dois níveis original.

Mesmo quando trabalhamos com os BLPPs da forma (3.3), para os quais o problema reformulado é equivalente, na prática encontramos alguns inconvenientes para resolver a reformulação: as restrições de complementaridade e o cálculo das derivadas da função e das restrições do problema.

A grande maioria dos algoritmos de programação não linear utiliza o gradiente (e às vezes as Hessianas) da função objetivo e das restrições que compõem o problema. Isto implica, no caso do problema reformulado, que as derivadas de terceira ordem podem ser necessárias para as funções que fazem parte da função  $L_f$  no problema (3.2). No caso em que usamos as condições de segunda ordem, as derivadas de quarta ordem podem ser necessárias. Além das funções do segundo nível terem que ser quatro vezes diferenciáveis, isto não é usualmente realizado na prática (embora seja válido para as funções lineares e quadráticas).

Para contornar este inconveniente, usamos métodos de diferenciação automática. Um método de diferenciação automática utiliza apenas a função objetivo e as restrições do problema como entrada para devolver funções que avaliam o gradiente, o Jacobiano e as Hessianas. Embora as funções do segundo nível ainda tenham que ser duas, três ou quatro vezes diferenciáveis, todo o trabalho de diferenciação não é mais feito manualmente. Exemplos de ferramentas deste tipo são o ADIFOR [BCC<sup>+</sup>92] e o AMPL [FGK90]. Os detalhes da implementação feita neste trabalho serão discutidos no Capítulo 4.

As restrições de complementaridade são consideradas restrições difíceis por muitos autores. Como veremos na Seção 3.2, dependendo do tipo de restrição de complementaridade usado, é possível provar que nenhum ponto do problema reformulado satisfaz a condição de qualificação de Mangasarian-Fromovitz. Muitos algoritmos de programação não linear têm sua convergência a pontos estacionários garantida desde que tais pontos satisfaçam a condição de qualificação de Mangasarian-Fromovitz. Esses algoritmos podem não resolver os problemas reformulados.

## 3.2 Restrições de complementaridade

Na Seção 1.3 do Capítulo 1 introduzimos brevemente os problemas de complementaridade e como as funções NCP podem gerar outras formulações das restrições de complementaridade encontradas no sistema KKT (1.2). Sabemos que as condições KKT são a idéia central por trás da reformulação de um BLPP em um problema de programação não linear, de modo que diferentes formulações das restrições de complementaridade podem gerar diferentes reformulações de um BLPP.

Neste trabalho, a reformulação de um problema de programação em dois níveis no problema (3.2) foi apresentada usando as restrições suaves como restrições de complementaridade. Sabemos que a vantagem desta formulação é a diferenciabilidade em todo ponto, dado que as funções do problema são diferenciáveis.

Veremos a seguir que esta reformulação possui propriedades teóricas que podem causar um mau funcionamento em muitos algoritmos de programação não linear. O argumento principal desta afirmação está no fato de que, quando as condições KKT são usadas como restrições em um problema de programação não linear, nenhum ponto viável do problema satisfaz a condição de qualificação de Mangasarian-Fromovitz [MF67].

A condição de qualificação de Mangasarian-Fromovitz pode ser definida como se segue.

### 3.2.1 DEFINIÇÃO (MFCQ [MF67])

Um ponto  $\bar{x}$  viável do problema (1.1) satisfaz a condição de qualificação de Mangasarian-Fromovitz (MFCQ) se o conjunto  $\{\nabla h_i(\bar{x}), i = 1, \dots, m\}$  é linearmente independente e existe  $d \in \mathbb{R}^n$  tal que:

$$\begin{aligned}\nabla h_i(\bar{x})^T d &= 0, \quad i = 1, \dots, m \\ \nabla g_i(\bar{x})^T d &< 0, \quad \forall i \in I(\bar{x}).\end{aligned}$$

Em [Ber99] prova-se que MFCQ é uma condição de qualificação, ou seja, um minimizador local de (1.1) que satisfaz MFCQ possui multiplicadores associados que satisfazem as condições KKT.

No próximo teorema, nos preocupamos em mostrar formalmente a afirmação feita nos parágrafos anteriores. Consideramos o problema reformulado como um problema de PNL e mostramos que nenhum ponto viável satisfaz MFCQ. A prova foi baseada em uma demonstração encontrada em [Dem02], adaptada especificamente para programação em dois níveis e para a maneira como as condições KKT estão definidas neste trabalho.

### 3.2.2 TEOREMA

Considere o problema de programação não linear (3.2), formado pelas funções de um BLPP para o qual o problema do segundo nível foi substituído pelas suas respectivas condições KKT. Então nenhum ponto viável deste problema satisfaz MFCQ.

PROVA:

Suponha que existe um ponto  $(\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu})$  viável e que satisfaz MFCQ para o problema (3.2). Logo, como os gradientes das restrições de igualdade avaliados em  $(\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu})$  são linearmente independentes, deve-se ter, entre outros fatos, que o gradiente das restrições de

complementaridade (que são da forma (1.9)) sejam não-nulos:

$$\begin{bmatrix} \hat{\mu}_i \nabla_x g_i(\hat{x}, \hat{y}) \\ \hat{\mu}_i \nabla_y g_i(\hat{x}, \hat{y}) \\ \mathbf{0} \\ 0 \\ \vdots \\ g_i(\hat{x}, \hat{y}) \\ \vdots \\ 0 \end{bmatrix} \neq 0, \quad i = 1, \dots, p_2, \quad (3.5)$$

caso contrário, o conjunto seria linearmente dependente. Lembre-se que a derivada é feita em relação às variáveis  $x$ ,  $y$ ,  $\lambda$  e  $\mu$ . Além disso, de acordo com a Definição 3.2.1, existem  $d_x \in \mathbb{R}^{n_1}$ ,  $d_y \in \mathbb{R}^{n_2}$ ,  $\gamma \in \mathbb{R}^{m_2}$  e  $\nu \in \mathbb{R}^{p_2}$  tais que:

$$\hat{\mu}_i \nabla_x g_i(\hat{x}, \hat{y})^T d_x + \hat{\mu}_i \nabla_y g_i(\hat{x}, \hat{y})^T d_y + \nu_i g_i(\hat{x}, \hat{y}) = 0, \quad i = 1, \dots, p_2 \quad (3.6)$$

$$\nabla_x g_i(\hat{x}, \hat{y})^T d_x + \nabla_y g_i(\hat{x}, \hat{y})^T d_y < 0, \quad \forall i \mid g_i(\hat{x}, \hat{y}) = 0 \quad (3.7)$$

$$\nu_i > 0, \quad \forall i \mid \hat{\mu}_i = 0. \quad (3.8)$$

Seja  $i \in \{1, \dots, p_2\}$  um índice tal que  $\nabla_x g_i(\hat{x}, \hat{y})^T d_x + \nabla_y g_i(\hat{x}, \hat{y})^T d_y \geq 0$ . Então,

$$\begin{aligned} & g_i(\hat{x}, \hat{y}) < 0 \quad \text{por (3.7)} \\ \implies & \hat{\mu}_i = 0 \quad \text{pois } \hat{\mu}_i g_i(\hat{x}, \hat{y}) = 0 \\ \implies & \nu_i > 0 \quad \text{por (3.8)} \end{aligned}$$

e pode-se concluir que

$$\hat{\mu}_i \nabla_x g_i(\hat{x}, \hat{y})^T d_x + \hat{\mu}_i \nabla_y g_i(\hat{x}, \hat{y})^T d_y + \nu_i g_i(\hat{x}, \hat{y}) = \nu_i g_i(\hat{x}, \hat{y}) < 0,$$

pois  $g_i(\hat{x}, \hat{y}) < 0$  e  $\nu_i > 0$ . Logo não existe tal  $i$ , pois (3.6) seria violada. De maneira análoga, usando (3.8) é possível mostrar que não existe  $i$  tal que  $\nu_i \leq 0$ . Logo, para todo  $i = 1, \dots, p_2$ ,

$$\nabla_x g_i(\hat{x}, \hat{y})^T d_x + \nabla_y g_i(\hat{x}, \hat{y})^T d_y < 0 \text{ e } \nu_i > 0.$$

As afirmações acima permitem concluir dois fatos:

(i)  $\hat{\mu}_i = 0 \forall i$ , pois se  $\hat{\mu}_i > 0$  então  $g_i(\hat{x}, \hat{y}) = 0$ , o que implicaria que

$$\hat{\mu}_i \nabla_x g_i(\hat{x}, \hat{y})^T d_x + \hat{\mu}_i \nabla_y g_i(\hat{x}, \hat{y})^T d_y + \nu_i g_i(\hat{x}, \hat{y}) = \hat{\mu}_i \nabla_x g_i(\hat{x}, \hat{y})^T d_x + \hat{\mu}_i \nabla_y g_i(\hat{x}, \hat{y})^T d_y < 0$$

e violaria a restrição (3.6)

(ii)  $g_i(\hat{x}, \hat{y}) = 0 \forall i$ , pois se  $g_i(\hat{x}, \hat{y}) < 0$  então  $\hat{\mu}_i = 0$ , o que implicaria que

$$\hat{\mu}_i \nabla_x g_i(\hat{x}, \hat{y})^T d_x + \hat{\mu}_i \nabla_y g_i(\hat{x}, \hat{y})^T d_y + \nu_i g_i(\hat{x}, \hat{y}) = \nu_i g_i(\hat{x}, \hat{y}) < 0$$

e novamente violaria a restrição (3.6).

Desta maneira,  $\hat{\mu}_i = 0$  e  $g_i(\hat{x}, \hat{y}) = 0$  para todo  $i = 1, \dots, p_2$ . Mas isto é impossível, pois o gradiente das restrições de complementaridade (3.5) será nulo, o que impede que o conjunto dos gradientes das restrições de igualdade seja linearmente independente. Logo, há uma contradição e conclui-se que nenhum ponto viável satisfaz MFCQ.

■

Este resultado é particularmente interessante no presente trabalho porque a técnica utilizada para resolver problemas de programação em dois níveis consiste na substituição do problema do segundo nível pelas suas condições KKT. As restrições associadas às condições KKT são então adicionadas ao problema do primeiro nível, gerando um problema não linear. Se um algoritmo desenvolvido para resolver problemas de programação não linear possui sua convergência assegurada na hipótese de que a sequência de pontos gerados converge para um ponto viável do problema que satisfaz MFCQ, nada pode ser afirmado quando trabalhamos com problemas de programação em dois níveis reformulados.

Vamos agora substituir no problema reformulado (3.2) as restrições suaves (1.7–1.9) pela equação  $\min(\mu_i, -g_i(x, y)) = 0$ , para  $i = 1, \dots, p_2$ , que utiliza a função NCP mínimo. O novo problema reformulado é apresentado a seguir.

$$\begin{aligned}
 &\underset{x, y, \lambda, \mu}{\text{Minimizar}} && F(x, y) \\
 &\text{s. a} && \nabla_y L_f(x, y, \lambda, \mu) = 0 \\
 &&& G(x, y) \leq 0 \\
 &&& H(x, y) = 0 \\
 &&& h(x, y) = 0 \\
 &&& \min(\mu_i, g_i(x, y)) = 0, \quad i = 1, \dots, p_2.
 \end{aligned} \tag{3.9}$$

Por um lado, além do número de restrições ser reduzido, é possível aplicar técnicas de otimização não diferenciável a ele. Em seu livro, Dempe [Dem02] utiliza a teoria de otimização não diferenciável para provar condições de otimalidade para o problema em dois níveis usando a reformulação (3.9). Por outro lado, mesmo que  $g_i(x, y)$  seja diferenciável para todo  $(x, y)$ ,  $\min(\mu_i, -g_i(x, y))$  não é diferenciável para  $\mu_i = -g_i(x, y)$ . Embora nosso foco seja a função mínimo na sua formulação original, em [CM96] apresenta-se uma classe de funções paramétricas suaves que a aproximam. Tais funções não são funções NCP, mas prova-se que, com a variação de seus parâmetros, tendem a ter um comportamento semelhante.

No lugar da função mínimo, também podemos usar a função Fischer-Burmeister perturbada, outra função baseada na filosofia das funções NCP que apresentamos no Capítulo 1. As restrições de complementaridade (1.7–1.9) são substituídas no problema (3.2) pela restrição  $\phi_\varepsilon(\mu_i, -g_i(x, y))$ ,  $i = 1, \dots, p_2$ , gerando o problema de programação não linear dado

por:

$$\begin{aligned}
& \underset{x, y, \lambda, \mu}{\text{Minimizar}} && F(x, y) \\
& \text{s. a} && \nabla_y L_f(x, y, \lambda, \mu) = 0 \\
& && G(x, y) \leq 0 \\
& && H(x, y) = 0 \\
& && h(x, y) = 0 \\
& && \mu_i - g_i(x) - \sqrt{\mu_i^2 + g_i(x)^2} + \varepsilon = 0, \quad i = 1, \dots, p_2.
\end{aligned} \tag{3.10}$$

Com a reformulação dada por (3.10), temos que se as funções que definem o BLPP forem diferenciáveis, então as funções deste problema também o serão. Por outro lado, esta reformulação depende de que um valor de  $\varepsilon$  suficientemente pequeno seja escolhido. Este valor não pode ser muito pequeno, pois a função deixará de ser “numericamente” diferenciável, nem muito grande, pois a função deixará de ter propriedades semelhantes às das funções NCP.

As três reformulações apresentadas neste capítulo, dadas pelos problemas (3.2), (3.9) e (3.10) foram os tipos de reformulação estudados e comparados neste trabalho. No Capítulo 4 cada um dos problemas de teste foi resolvido com os três tipos de reformulação e seus resultados foram apresentados para comparação.

### 3.3 Técnicas de resolução da reformulação

A técnica para resolução de problemas de programação em dois níveis que apresentamos neste trabalho pode não representar corretamente o problema original. Isto ocorre quando o conjunto viável do problema reformulado não é equivalente à região induzida do problema em dois níveis original.

Quando nos deparamos com problemas semelhantes ao Exemplo 3.1.2, sabemos que a reformulação não é a técnica ideal de resolução a ser aplicada. Se existem minimizadores globais do segundo nível que não satisfazem as condições KKT então é possível que a solução do BLPP não seja atingida pelo problema reformulado. Para este inconveniente, porém, nada pode ser feito, pois é uma consequência das características do próprio problema.

Por outro lado, nos casos semelhantes ao Exemplo 3.1.3, nos quais minimizadores locais, maximizadores e pontos de sela satisfazem as condições KKT do problema do segundo nível, encontramos a situação oposta: o conjunto viável do problema reformulado tem mais elementos que a região induzida do BLPP original. É intuitivo tentarmos alguma estratégia que elimine estes elementos excedentes. As duas estratégias apresentadas nesta seção têm por objetivo adicionar restrições no problema reformulado que eliminem pontos que não são minimizadores globais do segundo nível.

A primeira estratégia é descrita a seguir:

#### Estratégia 1

Considere a reformulação de um BLPP como um problema de PNL da forma (3.2). Sejam  $\varepsilon_{\max} > \varepsilon_{\min} > 0$  dados e  $k \leftarrow 1$ .

1. Resolva o problema reformulado com um algoritmo de programação não linear. Seja  $(\bar{x}^k, \bar{y}^k, \bar{\lambda}^k, \bar{\mu}^k)$  o ponto encontrado. Faça  $\varepsilon \leftarrow \varepsilon_{\max}$ .
2. Resolva o problema de viabilidade que consiste em encontrar um ponto que satisfaz as restrições do problema do segundo nível

$$g(\bar{x}^k, y) \leq 0$$

$$h(\bar{x}^k, y) = 0$$

e, além disso, satisfaz a restrição:

$$f(\bar{x}^k, y) \leq f(\bar{x}^k, \bar{y}^k) - \varepsilon \quad (3.11)$$

- 2.1. Se o problema do Passo 2 for viável, então  $\bar{y}^k$  não é um minimizador do problema do segundo nível. Faça  $k \leftarrow k + 1$ , adicione a restrição (3.11) ao problema reformulado e volte para o Passo 1.
- 2.2. Se o problema do Passo 2 for inviável e  $\varepsilon > \varepsilon_{\min}$  escolha  $\varepsilon_{\text{novo}} \in [\varepsilon_{\min}, \varepsilon]$ , faça  $\varepsilon \leftarrow \varepsilon_{\text{novo}}$  e volte para o Passo 2. Caso  $\varepsilon = \varepsilon_{\min}$ , declare  $\bar{y}^k$  um minimizador global do problema do segundo nível.

3. Declare  $(\bar{x}^k, \bar{y}^k)$  uma solução do BLPP.

O Passo 2 da estratégia merece uma atenção especial. Como  $(\bar{x}^k, \bar{y}^k)$  é uma solução da reformulação, sabemos que  $(\bar{x}^k, \bar{y}^k) \in \Omega$  (é viável do BLPP) e além disso  $(\bar{x}^k, \bar{y}^k, \bar{\lambda}^k, \bar{\mu}^k)$  satisfaz as condições KKT do problema do segundo nível. Estamos considerando o caso no qual pontos que não são minimizadores globais do segundo nível satisfazem KKT, portanto é possível que  $\bar{y}^k \notin \Psi(\bar{x}^k)$ . Se isto ocorrer, então  $(\bar{x}^k, \bar{y}^k) \notin \text{IR}$ , o que implica que  $(\bar{x}^k, \bar{y}^k)$  não é uma solução do BLPP original.

O problema de viabilidade que é resolvido consiste em verificar se  $\bar{y}^k \in \Psi(\bar{x}^k)$ . Se existir  $\tilde{y}$  tal que  $f(\bar{x}^k, \tilde{y}) < f(\bar{x}^k, \bar{y}^k)$  então sabemos que  $\bar{y}^k \notin \Psi(\bar{x}^k)$ , ou seja,  $\bar{y}^k$  não é um minimizador global do segundo nível para  $\bar{x}^k$ . Notemos que isto não implica de forma alguma que  $\tilde{y} \in \Psi(\bar{x}^k)$ . Neste caso, adicionamos a restrição (3.11) na reformulação para eliminar  $\bar{y}^k$ .

Dempe [Dem02] afirma que restrições do primeiro nível que são dependentes de variáveis do segundo nível podem desconectar o conjunto viável de um BLPP (quando o mesmo era conexo), além de torná-lo mais difícil de ser resolvido. A simples disposição (no primeiro ou no segundo nível) das restrições que contém variáveis do segundo nível pode gerar dois problemas totalmente distintos. Se uma restrição encontra-se no segundo nível, afeta as soluções globais do seguidor, mas se encontra-se no primeiro nível, influencia as decisões do líder sendo que a viabilidade só poderá ser verificada após uma resposta vinda do problema do segundo nível. O Exemplo 3.3.1 ilustra um caso para o qual a troca de níveis transforma o BLPP em um problema inviável.

### 3.3.1 EXEMPLO

Considere o problema de programação em dois níveis definido como se segue:

$$\begin{array}{ll} \text{Minimizar} & x^2 \\ \text{s. a} & \underset{y}{y} \in \arg \min_{y'} y'^2 \\ & \text{s. a} \quad y' \geq 1. \end{array}$$

Claramente, a solução do problema do segundo nível é  $\bar{y}' = 1$  independentemente da escolha de  $x$ . Com isso,  $\text{IR} = \{(x, 1) \mid x \in \mathbb{R}\}$  e o problema possui minimizador global  $(\bar{x}, \bar{y}) = (0, 1)$ . Se a restrição  $y' \geq 1$  for deslocada do segundo nível para o primeiro, obtém-se o seguinte BLPP:

$$\begin{array}{ll} \underset{x,y}{\text{Minimizar}} & x^2 \\ \text{s. a} & y \geq 1 \\ & y \in \arg \min_{y'} y'^2. \end{array}$$

Observe que este problema é inviável, pois o minimizador do segundo nível é  $\bar{y}' = 0$  para todo  $x$ .

O conjunto no qual a função objetivo do primeiro nível será minimizada é a região induzida do BLPP, formada pela intersecção do conjunto  $\Psi(x)$  com as restrições do primeiro nível. Levando em consideração tal fato, para eliminar um minimizador local, maximizador ou ponto de sela do problema do segundo nível, a adição de uma restrição no nível inferior impõe efeito idêntico ao da adição da mesma restrição no primeiro nível. Isto ocorre porque os pontos eliminados com a nova restrição não pertencem à região induzida do BLPP original, sendo, portanto, irrelevantes. Adicionar restrições no segundo nível, porém, aumenta o número de variáveis do problema reformulado (surgem os multiplicadores de Lagrange para as novas restrições), aumenta o número de restrições (a própria restrição, sua restrição de complementaridade associada e a restrição de caixa do multiplicador, caso esta seja de desigualdade) e altera as condições KKT do problema, como pode ser visto no Exemplo 3.3.2. Deste modo, quando o objetivo é remover pontos que não pertencem ao conjunto viável do BLPP original optamos por incluir a restrição (3.11) no primeiro nível.

### 3.3.2 EXEMPLO

Considere o problema do segundo nível, definido a seguir:

$$\begin{array}{ll} \underset{y'}{\text{Minimizar}} & -y'^2 \\ \text{s. a} & -1 \leq y' \leq 2. \end{array}$$

As condições KKT deste problema formam o seguinte sistema de equações:

$$\begin{aligned} -2y - \mu_1 + \mu_2 &= 0 \\ \mu_1(-y - 1) &= 0 \\ \mu_2(y - 2) &= 0 \\ -1 &\leq y \leq 2 \\ \mu_1, \mu_2 &\geq 0. \end{aligned}$$

Neste problema, para um ponto  $(\bar{x}, \bar{y}, \bar{\mu}_1, \bar{\mu}_2)$  qualquer, a restrição (3.11) é da seguinte forma:

$$-y^2 \leq -\bar{y}^2 - \varepsilon, \quad \varepsilon > 0$$



e, quando é adicionada ao problema, as condições KKT resultam no seguinte conjunto de restrições:

$$\begin{aligned}
 -2y - \mu_1 + \mu_2 - 2\mu_3y &= 0 \\
 \mu_1(-y - 1) &= 0 \\
 \mu_2(y - 2) &= 0 \\
 \mu_3(-y^2 + \bar{y}^2 + \varepsilon) &= 0 \\
 -y^2 + \bar{y}^2 + \varepsilon &\leq 0 \\
 -1 \leq y &\leq 2 \\
 \mu_1, \mu_2, \mu_3 &\geq 0.
 \end{aligned}$$

Note que houve um aumento no número de variáveis e de restrições quando comparado com as condições KKT iniciais.

No próximo exemplo, verificamos que os minimizadores locais e maximizadores são eliminados do conjunto viável da reformulação quando restrições adicionais são incluídas no problema, de acordo com a Estratégia 1 descrita anteriormente.

### 3.3.3 EXEMPLO (ELIMINAÇÃO DO MÍNIMO LOCAL E DO MÁXIMO GLOBAL)

Considere o seguinte problema em dois níveis:

$$\begin{aligned}
 \underset{x,y}{\text{Minimizar}} \quad & x^2 + y^2 \\
 \text{s. a} \quad & y - x \leq 0 \\
 & y \in \underset{y'}{\arg \min} -y'^2 \\
 \text{s. a} \quad & -1 \leq y' \leq 2.
 \end{aligned}$$

É fácil verificar que, para todo  $x$ , o único minimizador do problema do segundo nível é  $\bar{y}' = 2$ , de forma que  $\text{IR} = \{(x, y) \in \mathbb{R}^2 \mid y = 2 \text{ e } y - x \leq 0\}$ . No segundo nível ainda existem um maximizador  $\bar{y}' = 0$  e um minimizador local  $\hat{y}' = -1$  que satisfazem as condições KKT. A substituição do problema do segundo nível por suas condições KKT origina o seguinte problema de programação não linear:

$$\begin{aligned}
 \underset{x,y,\mu}{\text{Minimizar}} \quad & x^2 + y^2 \\
 \text{s. a} \quad & -2y - \mu_1 + \mu_2 = 0 \\
 & \mu_1(-y - 1) = 0 \\
 & \mu_2(y - 2) = 0 \\
 & y - x \leq 0 \\
 & -1 \leq y \leq 2 \\
 & \mu_1, \mu_2 \geq 0.
 \end{aligned}$$

O conjunto viável do problema reformulado é dado por  $\Omega = \{(x, 2, 0, 4)\} \cup \{(x, -1, 1, 0)\} \cup \{(x, 0, 0, 0)\}$ . Este conjunto está representado na Figura 3.4(a). A Figura 3.4(b) representa o problema do segundo nível.

A solução global do problema reformulado é  $(\tilde{x}, \tilde{y}, \tilde{\mu}_1, \tilde{\mu}_2) = (0, 0, 0, 0)$ , como mostra a Figura 3.5(a), mas sabe-se que  $\tilde{y} = 0$  é o maximizador do problema do segundo nível. De acordo com o Passo 2 da estratégia descrita anteriormente, é verificado se, dado  $\tilde{x} = 0$ , existe  $y$  que satisfaça as restrições do problema do segundo nível e ainda satisfaça a seguinte restrição (veja a Figura 3.5(b)):

$$-y^2 \leq -\varepsilon, \quad 0 < \varepsilon \leq \sqrt{2}.$$

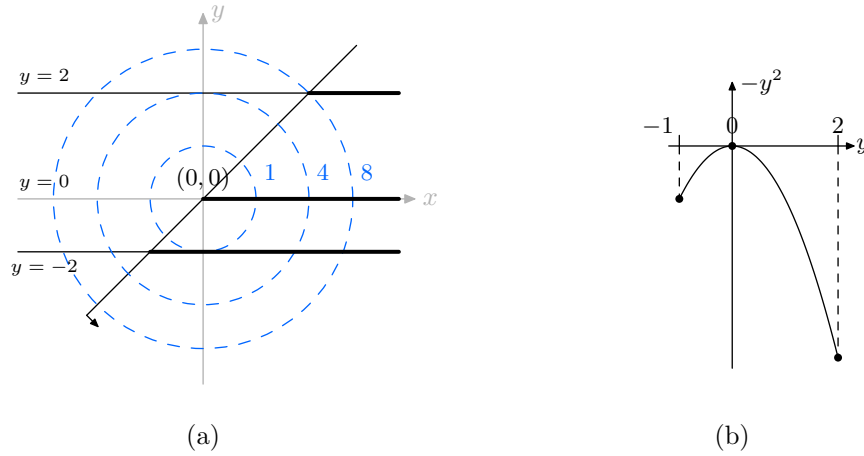


Figura 3.4: (a) O conjunto viável do problema reformulado é composto pela união dos três segmentos de reta em negrito. As circunferências representam as curvas de nível da função objetivo do primeiro nível. (b) Representação do problema do seguidor.

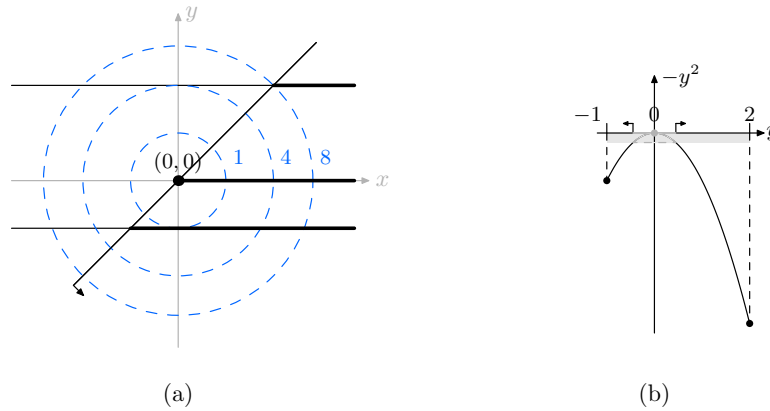


Figura 3.5: (a) A solução encontrada na primeira iteração do algoritmo. (b) Pontos viáveis do segundo nível que satisfazem a restrição  $-y^2 \leq -\varepsilon$ .

Como  $\bar{y} = 2$  é o minimizador global, independentemente de  $x$ ,  $\bar{y} = 2$  satisfaz as desigualdades acima e o problema resolvido no Passo 2 é viável. A restrição é então incluída no problema do primeiro nível, para não alterar o sistema KKT existente, e o ponto  $\tilde{y}$  é eliminado. O novo problema a ser resolvido é como se segue:

$$\begin{aligned}
 &\text{Minimizar}_{x,y,\mu} && x^2 + y^2 \\
 \text{s. a} &&& -2y - \mu_1 + \mu_2 = 0 \\
 &&& \mu_1(-y - 1) = 0 \\
 &&& \mu_2(y - 2) = 0 \\
 &&& y - x \leq 0 \\
 &&& -y^2 \leq -\varepsilon \\
 &&& -1 \leq y \leq 2 \\
 &&& \mu_1, \mu_2 \geq 0,
 \end{aligned}$$

representado na Figura 3.6(a) e cuja solução é  $(\hat{x}, \hat{y}, \hat{\mu}_1, \hat{\mu}_2) = (-1, -1, 1, 0)$ . Sabe-se que  $\hat{y} = -1$  é o minimizador local do problema do segundo nível, de forma que, quando o Passo 2 do algoritmo é aplicado novamente, agora com a restrição

$$-y^2 \leq -1 - \varepsilon, \quad \varepsilon > 0,$$

o problema ainda é viável, como pode ser verificado na Figura 3.6(b). Note que esta restrição engloba a restrição adicionada no passo anterior. Como o problema do Passo 2 é viável, a nova

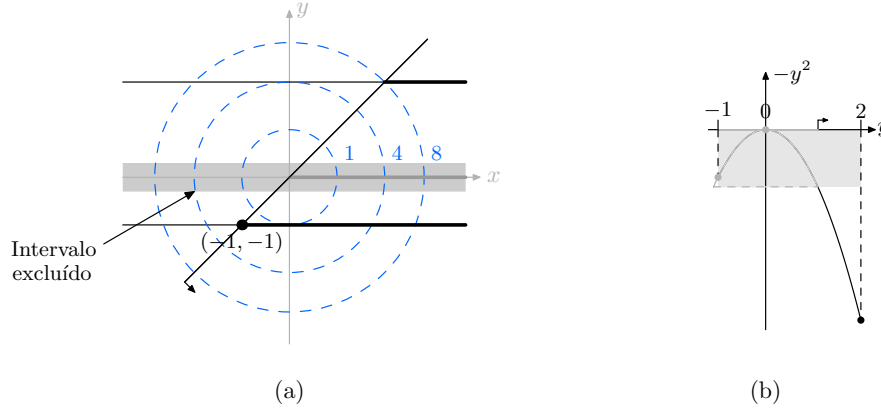


Figura 3.6: (a) A solução encontrada na segunda iteração do algoritmo. (b) Pontos viáveis do segundo nível que satisfazem a restrição  $-y^2 \leq -1 - \varepsilon$ .

restrição é incluída no problema do primeiro nível, o que implica que o novo PNL considerado é como se segue:

$$\begin{aligned} & \underset{x, y, \mu}{\text{Minimizar}} && x^2 + y^2 \\ \text{s. a} &&& -2y - \mu_1 + \mu_2 = 0 \\ &&& \mu_1(-y - 1) = 0 \\ &&& \mu_2(y - 2) = 0 \\ &&& y - x \leq 0 \\ &&& -y^2 \leq -\varepsilon \\ &&& -y^2 \leq -1 - \varepsilon \\ &&& -1 \leq y \leq 2 \\ &&& \mu_1, \mu_2 \geq 0. \end{aligned}$$

Na terceira iteração do algoritmo a solução global encontrada é  $(\bar{x}, \bar{y}, \bar{\mu}_1, \bar{\mu}_2) = (2, 2, 0, 4)$  como mostra a Figura 3.7(a). Esta é a solução global do BLPP original. Quando se aplica o Passo 2 considerando a restrição

$$-y^2 \leq -4 - \varepsilon, \quad \varepsilon > 0,$$

nenhum  $y$  viável é encontrado para todo  $\varepsilon > 0$ , como pode ser visto na Figura 3.7(b), pois  $y = 2$  é o minimizador global do problema do segundo nível. Neste caso,  $(\bar{x}, \bar{y}) = (2, 2)$  é declarada a solução do BLPP.

Esta estratégia não funciona para todos os casos, como podemos ver no Exemplo 3.3.4. Neste caso, o minimizador global do BLPP é eliminado do problema reformulado quando adicionamos uma nova restrição.

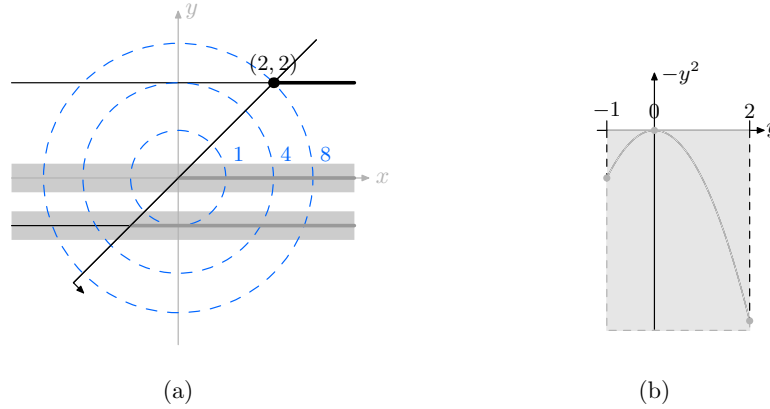


Figura 3.7: (a) A solução global do BLPP é encontrada com sucesso na terceira iteração. Note que, para  $x$  viável, apenas o minimizador global do segundo nível  $\bar{y} = 2$  é viável no problema. (b) Nenhum ponto viável  $y$  do segundo nível satisfaz a restrição  $-y^2 \leq -4 - \varepsilon$ . Sabemos que  $\bar{y} = 2$  é um minimizador global do problema do segundo nível.

### 3.3.4 EXEMPLO (ELIMINAÇÃO DA SOLUÇÃO GLOBAL)

Considere o BLPP definido como se segue:

$$\begin{aligned} & \underset{x,y}{\text{Minimizar}} && (x+1)^2 + y^2 \\ \text{s. a} &&& y \in \arg \min_{y'} -y'^2 \\ &&& \text{s. a} && y'^2 \leq 4 - x^2. \end{aligned}$$

Considere a Figura 3.8(a). A restrição do segundo nível pede que, para  $-2 \leq x \leq 2$ ,  $y' \in [-\sqrt{4-x^2}, \sqrt{4-x^2}]$ . Logo, a função  $-y'^2$  restrita a esta caixa possui valor mínimo em dois pontos:  $\bar{y}' = -\sqrt{4-x^2}$  e  $\bar{y}' = \sqrt{4-x^2}$ , para  $-2 \leq x \leq 2$ . Quando  $x < -2$  ou  $x > 2$  o problema do segundo nível torna-se inviável. Como não há restrições no primeiro nível, a região induzida é formada pela união de  $\Psi(x)$ , para todo  $x$ , e o mínimo global do problema é atingido em  $(\bar{x}, \bar{y}) = (-2, 0)$ . Substituindo o problema do segundo nível por suas condições KKT, o seguinte problema de programação não linear é formado:

$$\begin{aligned} & \underset{x,y,\mu}{\text{Minimizar}} && (x+1)^2 + y^2 \\ \text{s. a} &&& -2y + 2y\mu = 0 \\ &&& \mu(x^2 + y^2 - 4) = 0 \\ &&& y'^2 + x^2 \leq 4 \\ &&& \mu \geq 0. \end{aligned}$$

É fácil verificar que, para todo  $-2 \leq x \leq 2$ ,  $y = \{-\sqrt{4-x^2}, 0, \sqrt{4-x^2}\}$  satisfaz as condições KKT para o problema do segundo nível, formando o conjunto viável representado em negrito na Figura 3.8(b). Para este problema, o conjunto viável da reformulação contém o conjunto viável do BLPP original, mas não está contido no mesmo. A solução do problema reformulado, é  $(\hat{x}, \hat{y}) = (-1, 0)$ , mas  $\hat{y} = 0$  não é um minimizador global do problema do segundo nível para  $\hat{x} = -1$  e sim um maximizador, o que implica que este ponto não é um ponto viável do problema original.

Aplica-se, então, a Estratégia 1. Verifica-se, para  $x = \hat{x}$ , se as restrições do problema do segundo nível mais a restrição

$$f(\hat{x}, y) \leq f(\hat{x}, \hat{y}) - \varepsilon \implies -y'^2 \leq -\varepsilon, \varepsilon > 0$$

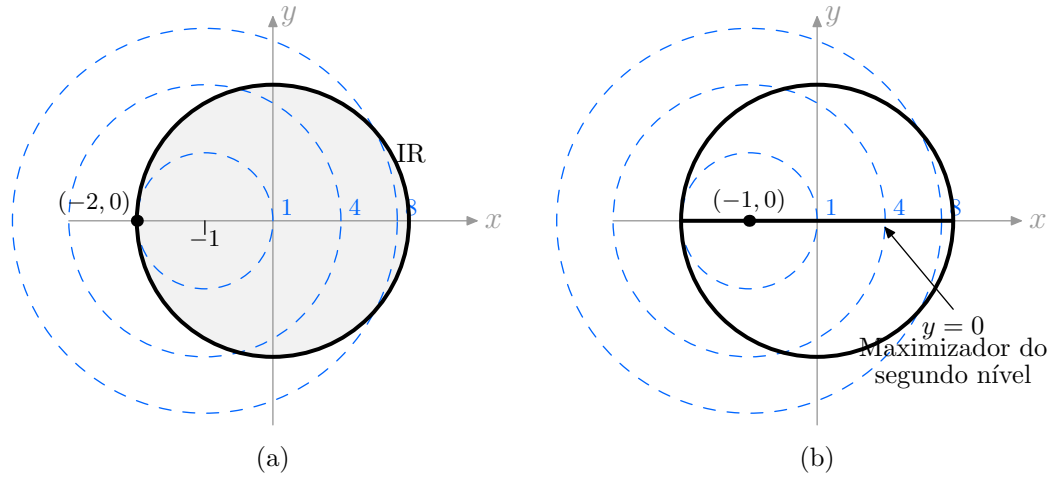


Figura 3.8: (a) Minimizador global do problema do segundo nível e (b) minimizador global da reformulação.

formam um problema viável. Como  $\hat{y} = 0$  é um maximizador, a resposta é afirmativa. Porém, a inclusão desta restrição ao problema reformulado elimina a possibilidade da função objetivo do segundo nível valer 0. Desta forma, o minimizador global do problema, o ponto  $(\bar{x}, \bar{y}) = (-2, 0)$ , foi eliminado, pois a nova restrição impede que  $\bar{y} = 0$ , que é o minimizador do segundo nível quando  $\bar{x} = -2$ .

A Estratégia 1 falha no Exemplo 3.3.4 devido ao seguinte fato: quando a restrição

$$f(\hat{x}, y) \leq f(\hat{x}, \hat{y}) - \varepsilon, \quad \varepsilon > 0$$

é adicionada ao problema reformulado (Passo **2.1**), significa que, para  $x = \hat{x}$  a função  $f(x, y)$  deve valer menos do que o valor  $f(\hat{x}, \hat{y})$  encontrado, onde  $\hat{y}$  não é um minimizador global do segundo nível. Porém, para um  $\tilde{x} \neq \hat{x}$  é possível que o mínimo global do problema do segundo nível seja maior ou igual a  $f(\hat{x}, \hat{y})$ , o que significa que um ponto antes viável agora também está sendo eliminado no novo problema. No Exemplo 3.3.4 isto ocorre porque  $f(\hat{x}, \hat{y}) = f(-1, 0) = 0 = f(-2, 0)$ , onde  $(-2, 0)$  é o minimizador global do BLPP.

Para contornar o obstáculo encontrado, apresentamos outra estratégia. A cada iteração, se o ponto  $(\bar{x}^k, \bar{y}^k)$  é encontrado, mas detectamos que  $\bar{y}^k$  não é minimizador global do problema do segundo nível, o ponto é eliminado da reformulação. Para eliminá-lo, adicionamos ao conjunto de restrições do problema reformulado uma bola de raio  $\delta$  e centro em  $(\bar{x}^k, \bar{y}^k)$ .

## Estratégia 2

Considere a reformulação de um BLPP como um problema de PNL da forma (3.2).

Sejam  $\varepsilon_{\max} > \varepsilon_{\min} > 0$ ,  $\delta_{\max} > \delta_{\min} > 0$  dados e  $j \leftarrow 1$ ,  $k \leftarrow 1$ .

1. Resolva o problema reformulado com um algoritmo de programação não linear. Seja  $(\bar{x}^k, \bar{y}^k, \bar{\lambda}^k, \bar{\mu}^k)$  o ponto encontrado. Faça  $\varepsilon \leftarrow \varepsilon_{\max}$ .

- 2.** Resolva o problema de viabilidade que consiste em encontrar um ponto que satisfaz as restrições do problema do segundo nível

$$\begin{aligned} g(\bar{x}^k, y) &\leq 0 \\ h(\bar{x}^k, y) &= 0 \end{aligned}$$

e, além disso, satisfaz a restrição:

$$f(\bar{x}^k, y) \leq f(\bar{x}^k, \bar{y}^k) - \varepsilon \quad (3.12)$$

- 2.1.** Se o problema do Passo **2** for viável, então  $\bar{y}^k$  não é um minimizador do problema do segundo nível. Adicione a restrição

$$r_j : \|(x, y) - (\bar{x}^k, \bar{y}^k)\|_2^2 \geq \delta_j^2, \quad \delta_{\min} < \delta_j \leq \delta_{\max} \quad (3.13)$$

ao problema reformulado, faça  $j \leftarrow j + 1$ ,  $k \leftarrow k + 1$  e volte para o Passo **1**.

- 2.2.** Se o problema do Passo **2** for inviável e  $\varepsilon > \varepsilon_{\min}$  escolha  $\varepsilon_{\text{nov}} \in [\varepsilon_{\min}, \varepsilon]$ , faça  $\varepsilon \leftarrow \varepsilon_{\text{nov}}$  e volte para o Passo **2**. Caso  $\varepsilon = \varepsilon_{\min}$ , declare  $\bar{y}^k$  um minimizador global do problema do segundo nível.

- 3.** Seja  $\bar{R} = \{i \mid r_i \text{ é ativa no ponto } (\bar{x}^k, \bar{y}^k)\}$ . Escolha  $\delta_{\text{nov}}^i \in [\delta_{\min}, \delta_i]$ , faça  $\delta_i \leftarrow \delta_{\text{nov}}^i$ , para todo  $i \in \bar{R}$ , faça  $k \leftarrow k + 1$  e volte para o Passo **1** (isto significa que o minimizador local/global original pode ter sido eliminado por alguma restrição do tipo (3.13)).

- 4.** Se  $\bar{R} = \emptyset$  ou existe algum  $i \in \bar{R}$  tal que  $\delta_i = \delta_{\min}$ , declare  $(\bar{x}^k, \bar{y}^k)$  a solução do BLPP.

Como será visto no Exemplo 3.3.5 a seguir, a Estratégia 2 tenta eliminar os pares  $(x, y)$  que pertencem ao conjunto viável da reformulação, mas que não estão presentes na região induzida do BLPP original. Como não é possível eliminar somente um ponto através das restrições usuais, eliminamos uma pequena vizinhança ao redor do ponto desejado. Se, por ventura, descobrimos que esta vizinhança é grande demais e há a possibilidade de eliminar uma solução do BLPP, reduzimos seu tamanho (Passo **3**).

As restrições (3.13) são adicionadas diretamente no problema reformulado por dois motivos. O primeiro é que não há a necessidade de alterar as restrições relacionadas às condições KKT do segundo nível, como explicado anteriormente. O segundo motivo é que nesta nova estratégia queremos eliminar o par  $(\bar{x}^k, \bar{y}^k)$  do conjunto viável do problema reformulado, portanto é pouco intuitivo adicionar tais restrições no problema do segundo nível.

Um ponto  $(\bar{x}^k, \bar{y}^k)$  é considerado uma solução do BLPP se ele for uma solução do problema reformulado,  $\bar{y}^k \in \Psi(\bar{x}^k)$  e, além disso, ou nenhuma restrição artificial é ativa no ponto ou então existe uma restrição artificial ativa cujo raio é muito pequeno (Passo **4**). Neste último caso, a solução é aproximada.

Vejamos agora que a nova estratégia consegue eliminar corretamente os pontos viáveis da reformulação que não são viáveis do problema original. O problema em foco é o mesmo utilizado no Exemplo 3.3.4, para o qual a Estratégia 1 não funcionou.

## 3.3.5 EXEMPLO

Considere o BLPP descrito no Exemplo 3.3.4. Já se sabe que seu minimizador global é  $(-2, 0)$ . A Estratégia 2 será aplicada à sua respectiva reformulação, mas como o número de iterações é extenso para ser descrito verbalmente, apenas os primeiros passos serão comentados e os demais podem ser observados na Figura 3.9.

Na primeira iteração, a solução encontrada no Passo 1 é  $(\bar{x}^1, \bar{y}^1) = (-1, 0)$ , de acordo com o Exemplo 3.3.4. O ponto  $\bar{y}^1 = 0$  é o maximizador do segundo nível, logo o problema no Passo 2 de encontrar um ponto que satisfaça as restrições do segundo nível e ainda a restrição (3.12) é factível. Para  $y^2 \geq \varepsilon$ , todos os pontos, na *reta pontilhada*  $x = \bar{x}^1$  (veja a Figura 3.9(a)) satisfazem (3.12). A restrição

$$r_1 : \|(x, y) - (\bar{x}^1, \bar{y}^1)\|_2^2 \geq \delta_1^2$$

é então adicionada à reformulação (Figura 3.9(b)).

A segunda e terceira iterações são parecidas. O problema reformulado, agora com a restrição adicional  $r_1$ , possui dois minimizadores globais:  $(\bar{x}^2, \bar{y}^2) = (-1 - \delta_1, 0)$  e  $(\bar{x}^3, \bar{y}^3) = (-1 + \delta_1, 0)$ . Suponha que  $(\bar{x}^2, \bar{y}^2)$  foi encontrado primeiro pelo algoritmo de otimização não linear. Mais uma vez, o problema de viabilidade do Passo 2 é factível, pois  $\bar{y}^2 = 0$  não é o argumento que minimiza o problema do segundo nível para  $\bar{x}^2 = -1 - \delta_1$ . Adiciona-se a restrição

$$r_2 : \|(x, y) - (\bar{x}^2, \bar{y}^2)\|_2^2 \geq \delta_2^2$$

ao problema reformulado (Figura 3.9(c)). No próximo passo,  $(\bar{x}^3, \bar{y}^3)$  é encontrado, pois agora é o único minimizador global da reformulação. Aplica-se o mesmo procedimento.

Pode-se perceber que, a cada passo, uma parcela da região viável do problema reformulado é eliminada, conforme o algoritmo descobre que o ponto encontrado não pertence à região induzida do problema original. Enquanto houverem soluções globais tais que  $\bar{y}^k = 0$  e  $\bar{x}^k \neq -2$ , os passos executados serão exatamente os mesmos descritos anteriormente, como mostra a Figura 3.9(d).

Na iteração 21, representada pela Figura 3.9(e), ocorre algo importante. Pela primeira vez no processo, um ponto da região induzida do BLPP é um minimizador global da reformulação. Este ponto não é o minimizador global do problema original  $(\bar{x}^*, \bar{y}^*) = (-2, 0)$ , pois como pode-se ver na ampliação apresentada na Figura 3.10(a), a restrição  $r_{20}$  elimina o ponto  $(-2, 0)$ . Para  $\bar{x}^{21}$ , o ponto  $\bar{y}^{21}$  é o argumento que minimiza o segundo nível, o que implica que o problema de viabilidade no Passo 2 será infactível para todo  $\varepsilon > 0$ .

A restrição  $r_{20}$  é ativa em  $(\bar{x}^{21}, \bar{y}^{21})$ , portanto, de acordo com o Passo 3, o raio  $\delta_{20}$  é reduzido e o problema reformulado é resolvido novamente. Observe na Figura 3.10(b) que a solução encontrada  $(\bar{x}^{22}, \bar{y}^{22})$  do novo problema também satisfaz  $r_{20}$  com igualdade, o que implica que  $\delta_{20}$  é reduzido mais uma vez (Figura 3.10(c)). Este processo se repete até a iteração 25.

Na iteração 25, o raio  $\delta_{20}$  da restrição  $r_{20}$  foi reduzido de tal maneira que a solução encontrada  $(\bar{x}^{25}, \bar{y}^{25})$  é tal que  $\bar{y}^{25} \notin \Psi(\bar{x}^{25})$ , como mostra a Figura 3.10(d). Logo, adiciona-se a restrição

$$r_{21} : \|(x, y) - (\bar{x}^{25}, \bar{y}^{25})\|_2^2 \geq \delta_{21}^2$$

ao problema reformulado. Veja, na Figura 3.10(e), que  $\delta_{21}$  possui o mesmo valor de  $\delta_{20}$ , ao invés de valer  $\delta_{\max}$ . Este é um detalhe de implementação que torna mais fácil a ilustração do algoritmo.

Os passos que se seguem são uma variação entre o Passo 2 e o Passo 3. Após 35 iterações, a solução aproximada do BLPP é encontrada, pois o raio de uma restrição ativa no ponto é menor ou igual a  $\delta_{\min}$ . Esta iteração final é mostrada na Figura 3.10(f).

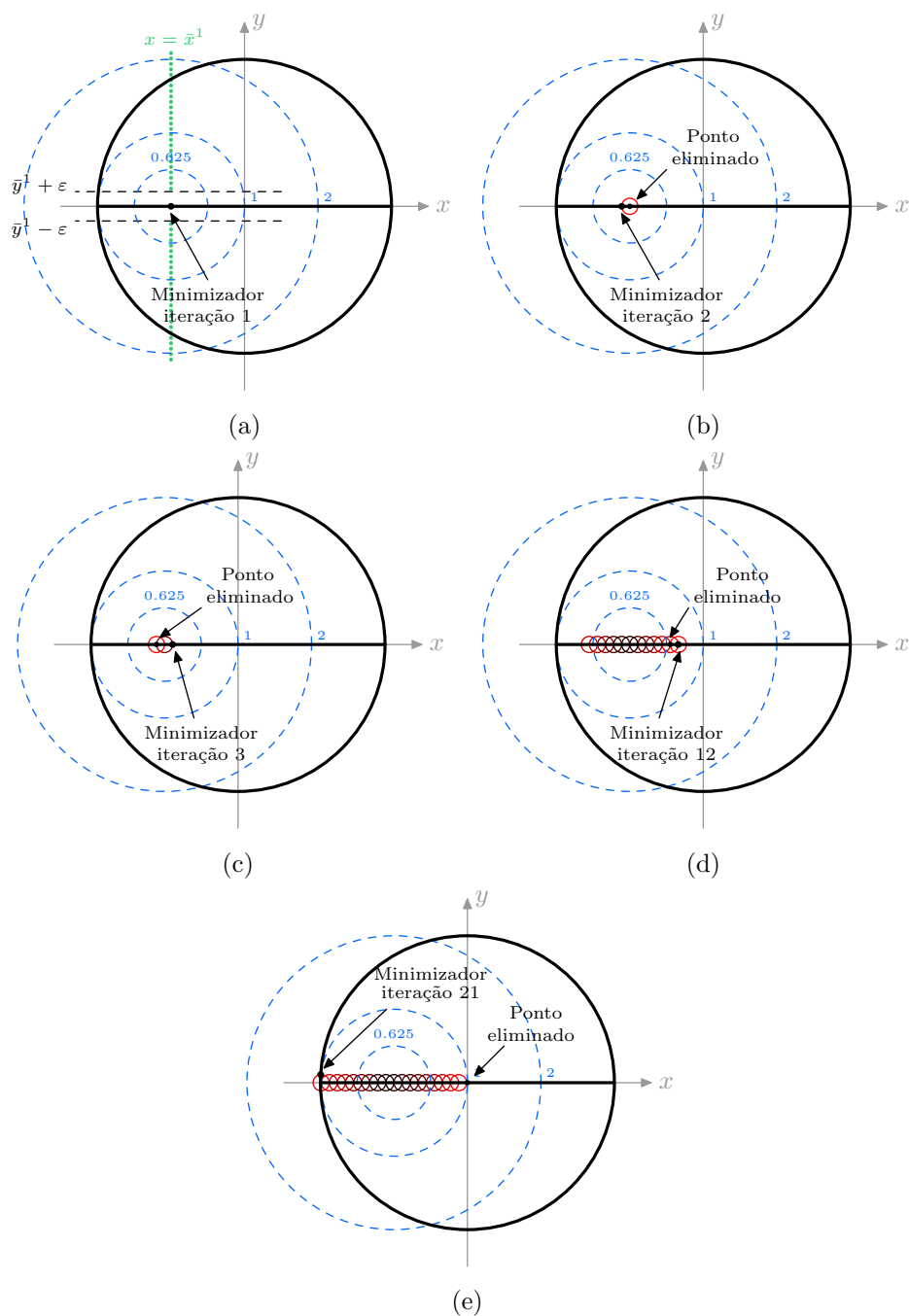


Figura 3.9: A cada iteração da Estratégia 2, uma parte da região viável do problema reformulado que não pertence à região induzida do BLPP é eliminada. Na iteração 21, encontra-se um ponto na região induzida do BLPP.



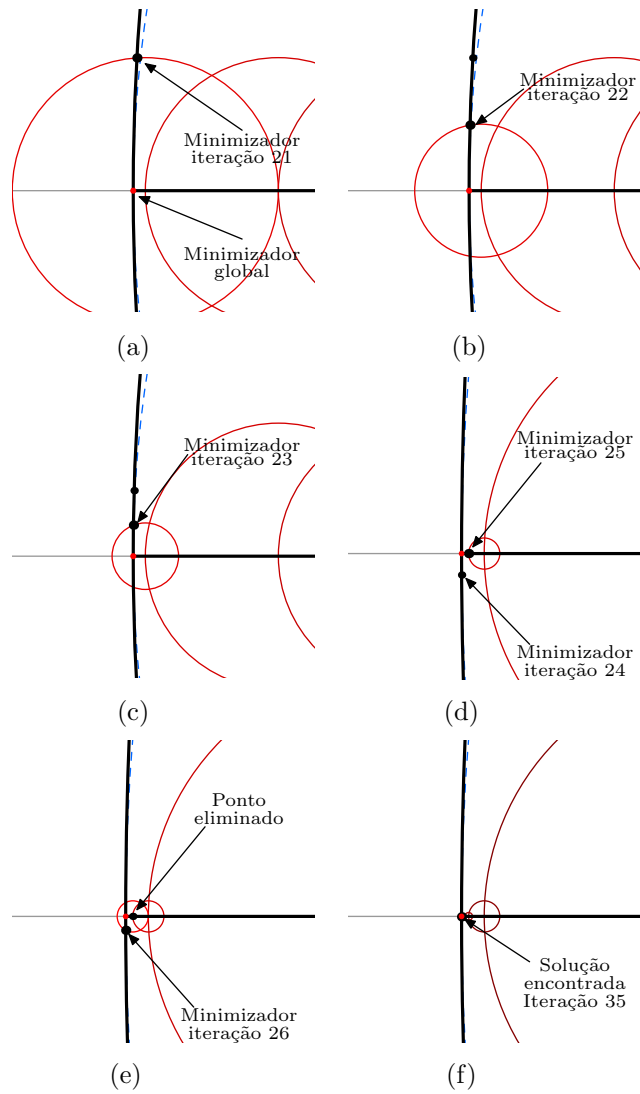


Figura 3.10: Iterações da Estratégia 2 em uma região próxima ao minimizador global do BLPP. Note que nas iterações representadas por (a), (b) e (c) nenhuma restrição foi adicionada pois o ponto encontrado pertence a IR. Somente o raio das restrições ativas é reduzido.



## Capítulo 4

# Resultados numéricos

Após o estudo cuidadoso dos problemas de programação matemática em dois níveis e de suas técnicas de resolução, com especial ênfase na reformulação do problema do segundo nível, nosso próximo passo é observar a eficiência da reformulação em problemas encontrados na literatura. Nutrimos um interesse especial em verificar se os inconvenientes encontrados na teoria se apresentam na prática.

Neste capítulo, fazemos uso de toda o conhecimento adquirido durante o trabalho para analisar o desempenho da técnica da reformulação do problema do segundo nível pelas suas condições KKT. Além disso, comparamos os três modos de representar as restrições de complementaridade, apresentados na Seção 3.2 do Capítulo 3, a saber: a formulação suave, a formulação com a função mínimo e a formulação com a função de Fischer-Burmeister perturbada. Elas serão denominadas ao longo de todo o capítulo por S, M e FBP, respectivamente.

Nos testes, consideramos 20 problemas de programação em dois níveis retirados da literatura. Os problemas são apresentados do modo como foram formulados originalmente, mas, independentemente de serem problemas cuja abordagem é da forma (2.1) ou da forma (2.4), todos serão resolvidos como se sua abordagem fosse a otimista relaxada (2.1).

Na Seção 4.1 descrevemos os problemas considerados nos testes e suas características principais. Quando possível, o minimizador global de cada problema é encontrado analiticamente. Caso não seja possível encontrar analiticamente o minimizador global, aplicamos na reformulação do problema o algoritmo de otimização global denominado BARON [TS04], para otimização não convexa. Na Seção 4.2 reportamos os resultados encontrados com a resolução de cada uma das três reformulações e os comparamos com as soluções retiradas da literatura e com as soluções encontradas analiticamente.

## 4.1 Problemas considerados

**Problema 1** [AS84, Exemplo 2]:

$$\begin{aligned}
 & \underset{x}{\text{Minimizar}} && 2x_1 + 2x_2 - 3y_1 - 3y_2 - 60 \\
 & \text{s. a} && x_1 + x_2 + y_1 - 2y_2 - 40 \leq 0 \\
 & && 0 \leq x_1, x_2 \leq 50 \\
 & y \in \underset{y'}{\arg \min} && (y'_1 - x_1 + 20)^2 + (y'_2 - x_2 + 20)^2 \\
 & \text{s. a} && 2y'_1 - x_1 + 10 \leq 0 \\
 & && 2y'_2 - x_2 + 10 \leq 0 \\
 & && -10 \leq y'_1, y'_2 \leq 20.
 \end{aligned}$$

**Classificação:** BLPP linear-quadrático.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é  $(x_1, x_2, y_1, y_2) = (10, 10, -10, -10)$ .

**Solução reportada:** Em [AS84], Aiyoshi e Shimizu encontraram a solução

$(x_1, x_2, y_1, y_2) = (25, 30, 5, 10)$ , com  $\bar{F} = 5$  e  $\bar{f} = 0$ . Colson, Marcotte e Savard [CMS05] encontraram uma solução com  $\bar{F} = 0$  e  $\bar{f} = 200$ , mas não forneceram o ponto encontrado.

Ao observarmos o problema do segundo nível, podemos notar que a variável do seguidor  $y_1$  depende apenas da variável do líder  $x_1$ . Da mesma maneira, a variável do seguidor  $y_2$  depende apenas da variável do líder  $x_2$ , de forma que estes problemas podem ser separados. O problema do segundo nível, associado a  $x_1$  e a  $y_1$  é dado por:

$$\begin{aligned}
 & \underset{y'_1}{\text{Minimizar}} && (y'_1 - x_1 + 20)^2 \\
 & \text{s. a} && 2y'_1 - x_1 + 10 \leq 0 \\
 & && -10 \leq y'_1 \leq 20.
 \end{aligned} \tag{4.1}$$

O valor de  $y'_1$  que minimiza este problema, desconsiderando suas restrições é  $y'_{\text{irr}} = x_1 - 20$ . Para cada valor de  $x_1$ , minimizamos com relação a  $y'_1$  uma função quadrática com restrições de caixa. Quando  $y'_{\text{irr}}$  estiver dentro da caixa, a solução será o próprio  $y'_{\text{irr}}$  (valor da minimização irrestrita), caso contrário, será o extremo da caixa mais próximo de  $y'_{\text{irr}}$ . Dado  $x_1$ , as restrições de caixa de  $y'_1$  são:

$$-10 \leq y'_1 \leq \min\{20, (x_1 - 10)/2\}.$$

Estas restrições foram construídas a partir das restrições de (4.1). Quando  $y'_1 = y'_{\text{irr}}$  devemos ter que

$$\begin{aligned}
 2(x_1 - 20) - x_1 + 10 &\leq 0 \\
 -10 &\leq x_1 - 20 \leq 20
 \end{aligned}$$

o que implica que, para  $10 \leq x_1 \leq 30$ ,  $y'_1 = y'_{\text{irr}}$  é a solução do problema (4.1).

Quando  $x_1 \leq 10$  temos que  $(x_1 - 10)/2 = 0 \Rightarrow \min\{20, 0\} = 0$ . Então as restrições tornam-se  $-10 \leq y'_1 \leq (x_1 - 10)/2$  para  $x_1 \leq 10$ , pois a função  $(x_1 - 10)/2$  é crescente. Como  $(x_1 - 10)/2$  decresce conforme  $x_1$  decresce, este intervalo de  $y'_1$  fica viável até que seus dois extremos coincidam, ou seja, até que  $(x_1 - 10)/2 = -10 \Rightarrow x_1 = -10$ . Além

disso, para  $x_1 \leq 10$ ,  $y'_{\text{irr}} = x_1 - 20 \leq -10$ . Logo, para  $-10 \leq x_1 \leq 10$  temos que  $y'_1 = -10$  é solução do problema (4.1).

Se  $x_1 \geq 30$ , temos  $(x_1 - 10)/2 = 10$ . Como a função  $(x_1 - 10)/2$  é crescente, temos que  $(x_1 - 10)/2 \geq 20$  somente quando  $x_1 \geq 50$ . Logo, para  $30 \leq x_1 \leq 50$ ,  $\min\{20, (x_1 - 10)/2\} = (x_1 - 10)/2$ , o que implica que a caixa em  $y'_1$  é  $-10 \leq y'_1 \leq (x_1 - 10)/2$  e a solução de (4.1) é  $y'_1 = (x_1 - 10)/2$ , pois  $y'_{\text{irr}} = x_1 - 20 \geq (x_1 - 20)/2$  para  $x_1 \geq 30$ .

Finalmente, para  $x_1 \geq 50$ , temos que  $y'_{\text{irr}} = x_1 - 20 \geq 30$  e  $-10 \leq y'_1 \leq 20$ , o que implica que  $y'_1 = 20$  é a solução de (4.1). Utilizando estas afirmações, podemos construir o conjunto  $\Psi_1(x_1)$ . Como este conjunto é unitário para cada  $x$  viável, pode ser representado pela função  $y_1(x_1)$  a seguir:

$$y_1(x_1) = \begin{cases} -10, & \text{se } -10 \leq x_1 \leq 10 \\ x_1 - 20, & \text{se } 10 \leq x_1 \leq 30 \\ \frac{x_1 - 10}{2}, & \text{se } 30 \leq x_1 \leq 50 \\ 20, & \text{se } x_1 \geq 50 \end{cases}$$

e representado pela Figura 4.1. De maneira análoga (apenas trocando-se os índices), dado  $x_2$ , podemos descrever a função  $y_2(x_2)$  para a variável  $y_2$ .

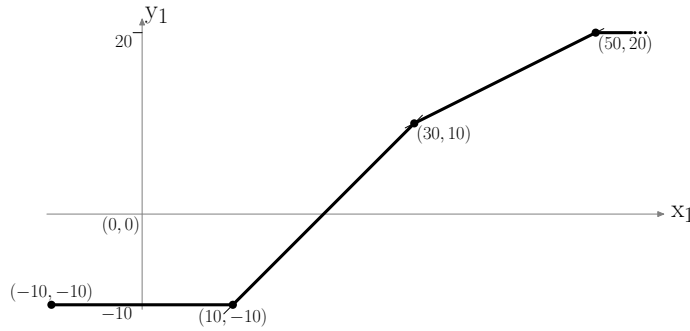


Figura 4.1: A função  $y_1(x_1)$ .

Para encontrarmos as soluções globais do BLPP, analisaremos os subproblemas resultantes de todas as combinações possíveis para os intervalos viáveis das variáveis  $x_1$  e  $x_2$ . Os subproblemas originados podem ser visualizados na Figura 4.2.

**1º caso:**  $0 \leq x_1 \leq 10$  e  $0 \leq x_2 \leq 10$ . Neste caso, temos que  $y_1 = -10$  e  $y_2 = -10$  e o BLPP torna-se o seguinte problema de programação linear:

$$\begin{aligned} &\text{Minimizar}_{x_1, x_2} && 2x_1 + 2x_2 \\ &\text{s. a} && x_1 + x_2 - 30 \leq 0 \\ &&& 0 \leq x_1, x_2 \leq 10, \end{aligned}$$

cujo minimizador é  $(x_1, x_2) = (0, 0)$ , com valor ótimo igual a 0. Logo, a solução é  $(x_1, x_2, y_1, y_2) = (0, 0, -10, -10)$ .

**2º caso:**  $0 \leq x_1 \leq 10$  e  $10 \leq x_2 \leq 30$ . Neste caso,  $y_1 = -10$  e  $y_2 = x_2 - 20$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & 2x_1 - x_2 + 30 \\ \text{s. a} & x_1 - x_2 - 10 \leq 0 \\ & 0 \leq x_1 \leq 10 \\ & 10 \leq x_2 \leq 30, \end{array}$$

cujo minimizador é  $(x_1, x_2) = (0, 30)$  com valor ótimo igual a 0. Isto implica que o BLPP possui solução  $(x_1, x_2, y_1, y_2) = (0, 30, -10, 10)$ .

**3º caso:**  $0 \leq x_1 \leq 10$  e  $30 \leq x_2 \leq 50$ . Neste caso,  $y_1 = -10$  e  $y_2 = x_2/2 - 5$  e o BLPP torna-se o seguinte problema de programação linear:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & 2x_1 + \frac{1}{2}x_2 + 15 \\ \text{s. a} & x_1 - 40 \leq 0 \\ & 0 \leq x_1 \leq 10 \\ & 30 \leq x_2 \leq 50, \end{array}$$

cujo minimizador é  $(x_1, x_2) = (0, 30)$ , com valor ótimo igual a 30. Logo, a solução é  $(x_1, x_2, y_1, y_2) = (0, 30, -10, 15)$ .

**4º caso:**  $10 \leq x_1 \leq 30$  e  $0 \leq x_2 \leq 10$ . Neste caso,  $y_1 = x_1 - 20$  e  $y_2 = -10$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & -x_1 + 2x_2 + 30 \\ \text{s. a} & 2x_1 + x_2 - 40 \leq 0 \\ & 10 \leq x_1 \leq 30 \\ & 0 \leq x_2 \leq 10, \end{array}$$

cujo minimizador é  $(x_1, x_2) = (20, 0)$  com valor ótimo igual a 10. Isto implica que o BLPP possui solução  $(x_1, x_2, y_1, y_2) = (20, 0, 0, -10)$ .

**5º caso:**  $10 \leq x_1 \leq 30$  e  $10 \leq x_2 \leq 30$ . Neste caso,  $y_1 = x_1 - 20$  e  $y_2 = x_2 - 20$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & -x_1 - x_2 + 60 \\ \text{s. a} & 2x_1 - x_2 - 20 \leq 0 \\ & 10 \leq x_1, x_2 \leq 30, \end{array}$$

cujo minimizador é  $(x_1, x_2) = (25, 30)$  com valor ótimo igual a 5. Isto implica que o BLPP possui solução  $(x_1, x_2, y_1, y_2) = (25, 30, 5, 10)$ .

**6º caso:**  $10 \leq x_1 \leq 30$  e  $30 \leq x_2 \leq 50$ . Neste caso,  $y_1 = x_1 - 20$  e  $y_2 = x_2/2 - 5$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & -x_1 + \frac{1}{2}x_2 + 15 \\ \text{s. a} & 2x_1 - 50 \leq 0 \\ & 10 \leq x_1 \leq 30 \\ & 30 \leq x_2 \leq 50, \end{array}$$

cujo minimizador é  $(x_1, x_2) = (25, 30)$  com valor ótimo igual a 5. Isto implica que o BLPP possui solução  $(x_1, x_2, y_1, y_2) = (25, 30, 5, 10)$ .

**7º caso:**  $30 \leq x_1 \leq 50$  e  $0 \leq x_2 \leq 10$ . Neste caso,  $y_1 = x_1/2 - 5$  e  $y_2 = -10$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & \frac{1}{2}x_1 + 2x_2 + 15 \\ \text{s. a} & \frac{3}{2}x_1 + x_2 - 25 \leq 0 \\ & 30 \leq x_1 \leq 50 \\ & 0 \leq x_2 \leq 10. \end{array}$$

Este problema é inviável.

**8º caso:**  $30 \leq x_1 \leq 50$  e  $10 \leq x_2 \leq 30$ . Neste caso,  $y_1 = x_1/2 - 5$  e  $y_2 = x_2 - 20$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & \frac{1}{2}x_1 - x_2 + 45 \\ \text{s. a} & \frac{3}{2}x_1 - x_2 - 5 \leq 0 \\ & 30 \leq x_1 \leq 50 \\ & 10 \leq x_2 \leq 30. \end{array}$$

Este problema é inviável.

**9º caso:**  $30 \leq x_1 \leq 50$  e  $30 \leq x_2 \leq 50$ . Neste caso,  $y_1 = x_1/2 - 5$  e  $y_2 = x_2/2 - 5$  e o BLPP torna-se:

$$\begin{array}{ll} \underset{x_1, x_2}{\text{Minimizar}} & \frac{1}{2}x_1 + \frac{1}{2}x_2 + 30 \\ \text{s. a} & \frac{3}{2}x_1 - 35 \leq 0 \\ & 30 \leq x_1, x_2 \leq 50. \end{array}$$

Este problema é inviável.

Ao analisarmos todos os casos, concluímos que os minimizadores globais do BLPP são:  $(x_1^1, x_2^1, y_1^1, y_2^1) = (0, 0, -10, -10)$  e  $(x_1^2, x_2^2, y_1^2, y_2^2) = (0, 30, -10, 10)$ . A função objetivo do primeiro nível possui valor  $F = 0$  para ambas as soluções, mas para a função objetivo do segundo nível temos  $f^1 = 200$  e  $f^2 = 100$ .

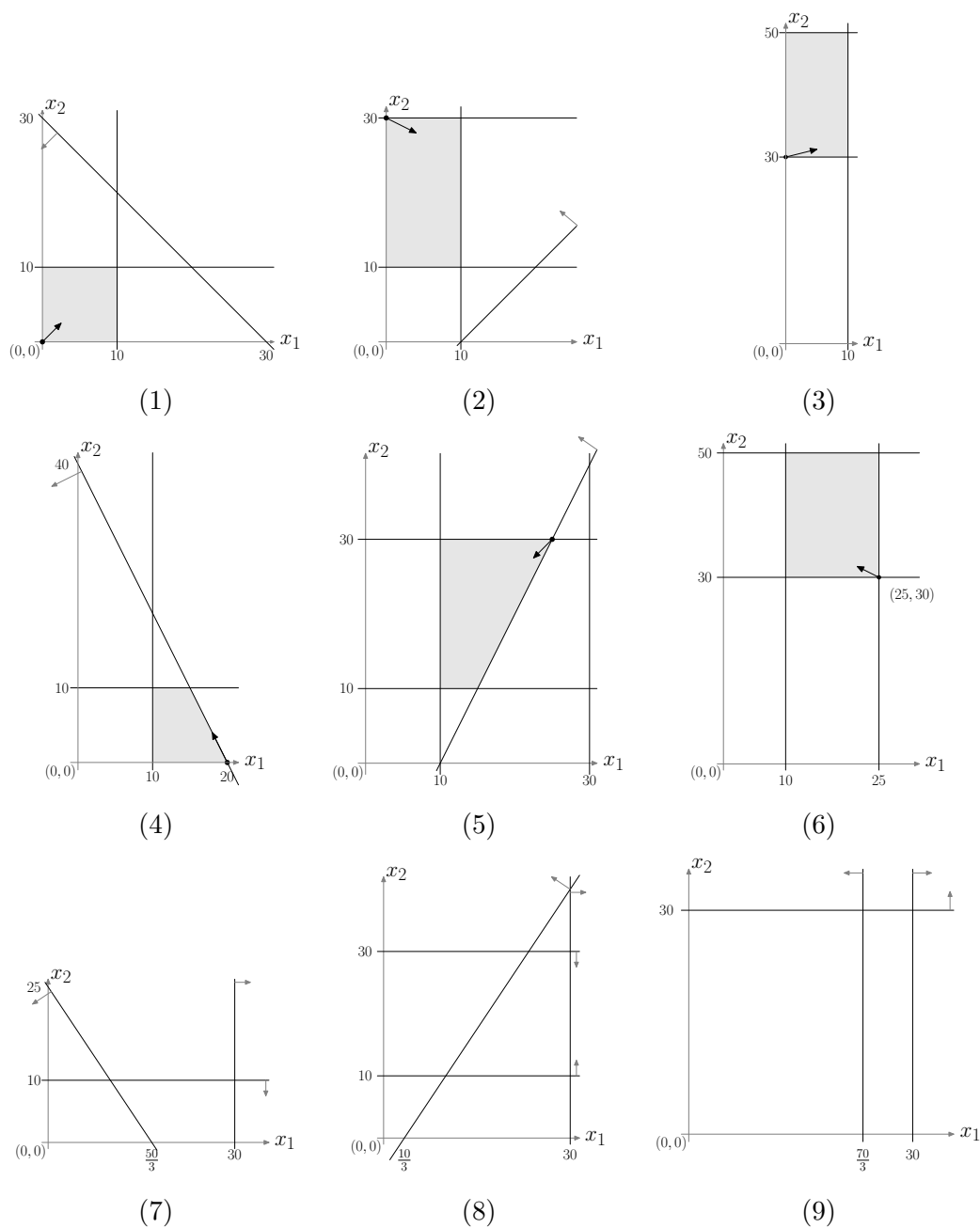


Figura 4.2: Os 9 casos para os intervalos das variáveis do primeiro nível e as soluções encontradas em cada caso (quando possível).



**Problema 2 [Bar84, Exemplo 2]:**

$$\begin{aligned}
& \underset{x}{\text{Minimizar}} && -2x_1^2 - 3x_2 - 4y_1 + y_2^2 \\
& \text{s. a} && y \in \underset{y'}{\arg \min} \quad 2x_1^2 + y_1'^2 - 5y_2' \\
& && \text{s. a} && x_1^2 - 2x_1 + x_2^2 - 2y_1' + y_2' \geq -3 \\
& && && x_2 + 3y_1' - 4y_2' \geq 4 \\
& && && x_1^2 + 2x_2 \leq 4 \\
& && && x_1, x_2, y_1', y_2' \geq 0.
\end{aligned}$$

**Classificação:** BLPP quadrático.

**Ponto inicial:** Não fornecido.

**Solução reportada:** Em [Bar84] utiliza-se uma técnica semelhante à reformulação. A solução encontrada é  $(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (2, 0, 1.6, 0.2)$  e  $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l) = (0.44, 1.36, 0, 0)$ , com  $\bar{F} = -14.36$  e  $\bar{f} = 9.56$ .

Para encontrarmos o minimizador global deste problema aplicamos à sua reformulação o algoritmo de otimização global BARON. O minimizador global pode ser encontrado porque o conjunto viável da reformulação é equivalente ao conjunto viável do problema original, devido à função objetivo do segundo nível ser convexa em  $y$  e às restrições do segundo nível serem lineares em  $y$  para todo  $x$ . A solução encontrada foi  $(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (2, 0, 1.6, 0.2)$ , com  $\bar{F} = -14.36$  e  $\bar{f} = 9.56$ .

**Problema 3 [Bar88, Exemplo 1]:**

$$\begin{aligned}
& \underset{x}{\text{Minimizar}} && (x - 5)^2 + (2y + 1)^2 \\
& \text{s. a} && x \geq 0 \\
& && y \in \underset{y'}{\arg \min} \quad (y' - 1)^2 - 1.5xy' \\
& && \text{s. a} && 3x - y' \geq 3 \\
& && && -x + 0.5y' \geq -4 \\
& && && -x - y' \geq -7 \\
& && && y' \geq 0.
\end{aligned}$$

**Classificação:** BLPP quadrático.

**Ponto inicial** Fornecido em [Col02]. O ponto utilizado é  $(x, y) = (4, 0)$ .

**Solução reportada** Em [CMS05], os autores encontram uma solução com  $\bar{F} = 17$  e  $\bar{f} = 1$ , mas não apresentam o ponto associado.

Para encontrarmos analiticamente o minimizador global do problema, primeiramente construiremos o conjunto  $\Psi(x)$  dos minimizadores do problema do segundo nível. O problema do segundo nível é um problema de programação quadrática onde a variável  $y'$  está

sujeita somente a restrições de caixa. Logo, podemos descrevê-lo da seguinte maneira:

$$\begin{aligned} & \underset{y'}{\text{Minimizar}} && (y' - 1)^2 - 1.5xy' \\ \text{s. a} &&& y' \leq 3x - 3 \\ &&& y' \geq 2x - 8 \\ &&& y' \leq 7 - x \\ &&& y' \geq 0. \end{aligned} \tag{4.2}$$

Reescrevendo as restrições de caixa em  $y'$ , temos:

$$\max\{0, 2x - 8\} \leq y' \leq \min\{3x - 3, 7 - x\}.$$

A solução de um problema de programação quadrática e com uma variável sujeita apenas a restrições de caixa é o minimizador global do problema irrestrito, quando este estiver dentro da caixa, ou então é o limitante da caixa mais próximo do minimizador irrestrito. A solução irrestrita do problema (4.2) é  $y'_{\text{irr}} = (1.5x + 2)/2$ . Desejamos encontrar o intervalo de  $x$  para o qual a solução do problema irrestrito está dentro da caixa:

$$\begin{aligned} (1.5x + 2)/2 &\geq 2x - 8 \Rightarrow x \leq 18/2.5 \\ (1.5x + 2)/2 &\geq 0 \Rightarrow x \geq -2/1.5 \\ (1.5x + 2)/2 &\leq 3x - 3 \Rightarrow x \geq 8/4.5 \\ (1.5x + 2)/2 &\leq 7 - x \Rightarrow x \leq 12/3.5. \end{aligned}$$

Fazendo a intersecção dos intervalos temos que, para  $\frac{8}{4.5} \leq x \leq \frac{12}{3.5}$  a solução de (4.2) é  $y' = y'_{\text{irr}} = (1.5x + 2)/2$ .

Se  $x \geq \frac{12}{3.5}$  temos que  $y'_{\text{irr}} = (1.5x + 2)/2 \geq 7 - x$ . Neste caso, podemos verificar que  $\min\{3x - 3, 7 - x\} = 7 - x$ , pois em  $x = \frac{12}{3.5}$  temos  $3x - 3 \geq 7 - x$  e  $7 - x$  é uma função decrescente. No limitante inferior, temos que  $\max\{0, 2x - 8\} = 0$  para  $x \leq 4$ . Logo, para  $\frac{12}{3.5} \leq x \leq 4$  as restrições de caixa tornam-se  $0 \leq y' \leq 7 - x$  e a solução de (4.2) é  $\bar{y}' = 7 - x$ , que é o limitante da caixa mais próximo de  $y'_{\text{irr}}$ .

Se  $x \geq 4$ , temos que  $\max\{0, 2x - 8\} = 2x - 8$ , o que implica que  $2x - 8 \leq y' \leq 7 - x$ , pois já sabemos que  $\min\{3x - 3, 7 - x\} = 7 - x$  para  $x \geq \frac{12}{3.5}$ . O intervalo fica inviável quando  $2x - 8 = 7 - x \Rightarrow x = 5$ . Logo, para  $4 \leq x \leq 5$  a solução de (4.2) é  $\bar{y}' = 7 - x$ . Para  $x > 5$  o problema do segundo nível torna-se inviável.

Se  $x \leq \frac{8}{4.5}$  podemos verificar que  $\min\{3x - 3, 7 - x\} = 3x - 3$  e  $\max\{0, 2x - 8\} = 0$ . Quando  $x = 1$  temos  $3x - 3 = 0$ , ou seja, os limitantes inferior e superior da caixa possuem o mesmo valor, de maneira que, para  $1 \leq x \leq \frac{8}{4.5}$ ,  $0 \leq y' \leq 3x - 3$  e a solução de (4.2) é  $\bar{y}' = 3x - 3$ , pois para  $x \leq \frac{8}{4.5}$  sabemos que  $y'_{\text{irr}} = (1.5x + 2)/2 \geq 3x - 3$ . Para  $x < 1$  o problema é inviável.

Com estas observações, sabemos que para cada valor de  $x$ , a variável do segundo nível  $y$  é unicamente determinada e podemos representar o conjunto  $\Psi(x)$  através da função  $y(x)$ :

$$y(x) = \begin{cases} 3x - 3, & \text{se } 1 \leq x \leq \frac{8}{4.5} \\ \frac{1.5x + 2}{2}, & \text{se } \frac{8}{4.5} \leq x \leq \frac{12}{3.5} \\ 7 - x, & \text{se } \frac{12}{3.5} \leq x \leq 5. \end{cases}$$

Esta função pode ser visualizada na Figura 4.3.

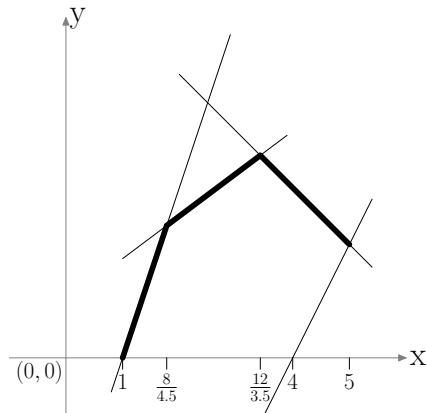


Figura 4.3: A função  $y(x)$ .

Agora substituímos o problema do segundo nível pela função  $y(x)$  e obtemos o seguinte problema de otimização:

$$\begin{aligned} & \underset{x,y}{\text{Minimizar}} && (x-5)^2 + (2y+1)^2 \\ & \text{s. a} && y = y(x) \\ & && x \geq 0, \end{aligned}$$

representado na Figura 4.4. Para encontrar o minimizador global, minimizamos a função objetivo do primeiro nível em cada uma das faces que compõem a função  $y(x)$ , para  $x$  viável e no qual  $y(x)$  está bem definida. Escolhemos a solução com menor valor, em um processo análogo ao aplicado ao Problema 1.

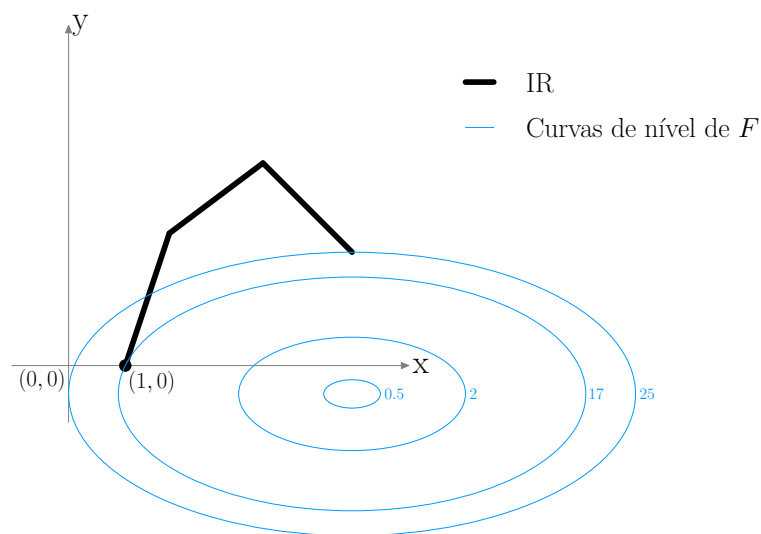


Figura 4.4: Região induzida (em negrito) e minimizador global do Problema 3.

**Problema 4 [Bar88, Exemplo 2]:**

Originalmente, este exemplo considera o seguinte problema em dois níveis:

$$\begin{aligned}
 & \underset{x}{\text{Minimizar}} && -(200 - y_1 - y_3)(y_1 + y_3) - (160 - y_2 - y_4)(y_2 + y_4) \\
 & \text{s. a} && x_1 + x_2 + x_3 + x_4 \leq 40 \\
 & && 0 \leq x_1 \leq 10 \\
 & && 0 \leq x_2 \leq 5 \\
 & && 0 \leq x_3 \leq 15 \\
 & && 0 \leq x_4 \leq 20 \\
 & (y_1, y_2) \in && \underset{y'_1, y'_2}{\arg \min} (y'_1 - 4)^2 + (y'_2 - 13)^2 \\
 & \text{s. a} && 0.4y'_1 + 0.7y'_2 \leq x_1 \\
 & && 0.6y'_1 + 0.3y'_2 \leq x_2 \\
 & && 0 \leq y'_1, y'_2 \leq 20 \\
 & (y_3, y_4) \in && \underset{y'_3, y'_4}{\arg \min} (y'_3 - 35)^2 + (y'_4 - 2)^2 \\
 & \text{s. a} && 0.4y'_3 + 0.7y'_4 \leq x_3 \\
 & && 0.6y'_3 + 0.3y'_4 \leq x_4 \\
 & && 0 \leq y'_3, y'_4 \leq 40.
 \end{aligned}$$

O problema possui dois seguidores, que desejam minimizar suas respectivas funções objetivo. Como suas variáveis não estão relacionadas, podemos transformar o problema acima em um BLPP com apenas um problema do segundo nível. O problema que resolvemos é como segue:

$$\begin{aligned}
 & \underset{x}{\text{Minimizar}} && -(200 - y_1 - y_3)(y_1 + y_3) - (160 - y_2 - y_4)(y_2 + y_4) \\
 & \text{s. a} && x_1 + x_2 + x_3 + x_4 \leq 40 \\
 & && 0 \leq x_1 \leq 10 \\
 & && 0 \leq x_2 \leq 5 \\
 & && 0 \leq x_3 \leq 15 \\
 & && 0 \leq x_4 \leq 20 \\
 & y \in && \underset{y'}{\arg \min} (y'_1 - 4)^2 + (y'_2 - 13)^2 + (y'_3 - 35)^2 + (y'_4 - 2)^2 \\
 & \text{s. a} && 0.4y'_1 + 0.7y'_2 \leq x_1 \\
 & && 0.6y'_1 + 0.3y'_2 \leq x_2 \\
 & && 0.4y'_3 + 0.7y'_4 \leq x_3 \\
 & && 0.6y'_3 + 0.3y'_4 \leq x_4 \\
 & && 0 \leq y'_1, y'_2 \leq 20 \\
 & && 0 \leq y'_3, y'_4 \leq 40.
 \end{aligned}$$

**Classificação:** BLPP quadrático.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é

$$(x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4) = (5, 5, 15, 15, 0, 0, 0, 0).$$

**Solução reportada:** Em [Bar88] a solução encontrada é

$$\begin{aligned}(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) &= (7.91, 4.37, 11.08, 16.63) \\ (\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4) &= (2.29, 10, 27.21, 0),\end{aligned}$$

com a função objetivo do primeiro nível igual a  $-6600$  e as funções do segundo nível  $\bar{f}_1 = 11.92$  e  $\bar{f}_2 = 55$ . Em [CMS05] é encontrado o mesmo valor para a função objetivo.

O algoritmo utilizado em [Bar88] é do tipo *branch and bound*, o que significa que um minimizador global do BLPP foi encontrado.

**Problema 5 [BK84, Exemplo 7.1]:**

$$\begin{aligned}\text{Minimizar } & -y \\ \text{s. a } & x \in \arg \min_{y'} y' \\ & \text{s. a } \quad \begin{aligned} -x - 2y' &\leq -10 \\ x - 2y' &\leq 6 \\ 2x - y' &\leq 21 \\ x + 2y' &\leq 38 \\ -x + 2y' &\leq 18 \\ x, y' &\geq 0. \end{aligned}\end{aligned}$$

**Classificação:** BLPP linear.

**Ponto inicial:** Não fornecido.

**Solução reportada:** Em [BK84] a solução encontrada é  $(\bar{x}, \bar{y}) = (16, 11)$ , com  $\bar{F} = -11$  e  $\bar{f} = 11$ .

Analisando o problema do segundo nível de modo semelhante ao realizado em problemas anteriores (veja o Problema 1 e o Problema 3) podemos construir o conjunto  $\Psi(x)$  que, neste caso, é dado pela função

$$y(x) = \begin{cases} \frac{10-x}{2}, & \text{se } 0 \leq x \leq 8 \\ \frac{x-6}{2}, & \text{se } 8 \leq x \leq 12 \\ 2x-21, & \text{se } 12 \leq x \leq 16. \end{cases}$$

Esta função está ilustrada na Figura 4.5 e representa a região induzida do BLPP, conjunto no qual  $F(x, y)$  será minimizada.

Ao minimizarmos a função  $F(x, y)$  em cada uma das faces que formam a função  $y(x)$  obtemos o minimizador global do problema  $(\bar{x}, \bar{y}) = (16, 11)$ , apresentado na Figura 4.5.

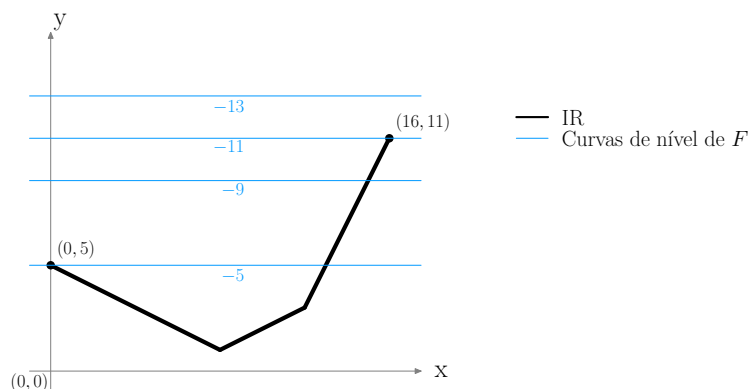


Figura 4.5: Ilustração do Problema 5.

**Problema 6 [CW90, Problema (10)]:**

$$\begin{aligned}
 &\underset{x}{\text{Minimizar}} && -x - 3y_1 + 2y_2 \\
 &\text{s. a} && 0 \leq x \leq 8 \\
 &&& y \in \arg \min_{y'} -y'_1 \\
 &&& \text{s. a} && 2x - y'_1 - 4y'_2 \geq -16 \\
 &&& && -8x - 3y'_1 + 2y'_2 \geq -48 \\
 &&& && 2x - y'_1 + 3y'_2 \geq 12 \\
 &&& && 0 \leq y'_1 \leq 4.
 \end{aligned}$$

**Classificação:** BLPP linear.**Ponto inicial:** Não fornecido.**Solução reportada:** Em [CW90] a solução encontrada é  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (5, 4, 2)$ , com  $\bar{F} = -13$  e  $\bar{f} = -4$ .Clark e Westerberg [CW90] mostram com detalhes o conjunto viável e a região induzida do problema e apresentam o seu minimizador global  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (5, 4, 2)$ .**Problema 7 [CW90, Problema (11)]:**

$$\begin{aligned}
 &\underset{x}{\text{Minimizar}} && (x - 3)^2 + (y - 2)^2 \\
 &\text{s. a} && 0 \leq x \leq 8 \\
 &&& y \in \arg \min_{y'} (y' - 5)^2 \\
 &&& \text{s. a} && 2x - y' \geq -1 \\
 &&& && -x + 2y' \geq 2 \\
 &&& && -x - 2y' \geq -14.
 \end{aligned}$$

**Classificação:** BLPP quadrático.**Ponto inicial:** Não fornecido.

**Solução reportada:** Em [CW90] os autores apresentam o minimizador  $(\bar{x}, \bar{y}) = (1, 3)$ , com  $\bar{F} = 5$  e  $\bar{f} = 4$ .

Para verificarmos que  $(1, 3)$  é o minimizador global do problema, aplicamos o mesmo processo usado no Problema 3. O conjunto viável da reformulação é equivalente ao conjunto viável do BLPP original, pois a função do segundo nível é convexa e as restrições são lineares. O conjunto  $\Psi(x)$  deste problema é unitário para todo  $x$ , de forma que pode ser substituído pela função  $y(x)$ :

$$y(x) = \begin{cases} 2x + 1, & \text{se } 0 \leq x \leq 2 \\ 5, & \text{se } 2 \leq x \leq 4 \\ 7 - \frac{x}{2}, & \text{se } 4 \leq x \leq 6. \end{cases}$$

A intersecção do gráfico desta função com as restrições do primeiro nível forma IR, representada em negrito na Figura 4.6. Basta, então, minimizarmos a função objetivo do primeiro nível em cada uma das faces de  $y(x)$  e escolher a solução (viável) com o menor valor de  $F$ . A solução encontrada é o ponto  $(\bar{x}, \bar{y}) = (1, 3)$ . Ainda encontramos dois minimizadores locais:  $(3, 5)$  e  $(\frac{22}{5}, \frac{48}{10})$ .

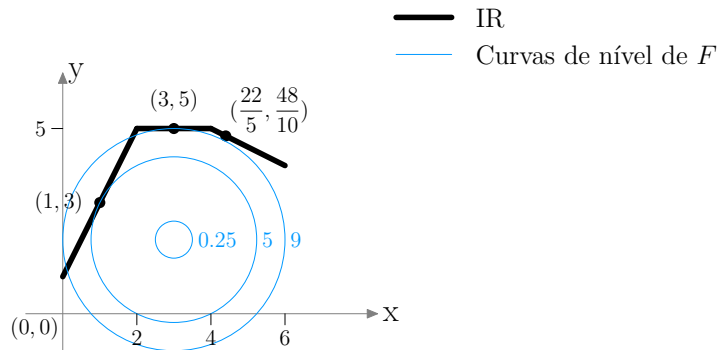


Figura 4.6: Ilustração do Problema 7. Os pontos  $(3, 5)$  e  $(\frac{22}{5}, \frac{48}{10})$  são os minimizadores locais do problema.

#### Problema 8 [Col02, BIPA1]:

$$\begin{aligned} & \underset{x}{\text{Minimizar}} && (10 - x)^3 + (10 - y)^3 \\ & \text{s. a} && -x + y \leq 0 \\ & && 0 \leq x \leq 15 \\ & && y \in \underset{y'}{\arg \min} (x + 2y' - 15)^4 \\ & && \text{s. a} && x + y' \leq 20 \\ & && && 0 \leq y' \leq 20. \end{aligned}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é  $(x, y) = (10, 10)$ .

**Solução reportada:** Não foi reportada solução.

Para encontrar o minimizador global deste BLPP vamos analisar a estrutura do problema, de forma análoga à feita em problemas anteriores. Levando em consideração apenas o problema do segundo nível, a solução do problema irrestrito é  $y' = (15 - x)/2$ , ponto no qual o gradiente com relação a  $y'$  vale 0. A restrição de caixa em  $y'$  é:

$$0 \leq y' \leq \min\{20 - x, 20\}.$$

Podemos verificar que, para  $-25 \leq x \leq 15$ , temos que  $0 \leq (15 - x)/2 \leq \min\{20 - x, 20\} = 20 - x$ , logo, a solução do problema irrestrito está dentro da caixa neste intervalo.

Se  $x \geq 15$ , temos  $(15 - x)/2 \leq 0$ . Sabemos ainda que, quando  $x > 20$  o problema em  $y'$  torna-se inviável. Logo, para  $15 \leq x \leq 20$ , devemos ter  $0 \leq y' \leq 20 - x \Rightarrow \bar{y}' = 0$  é a solução do problema do segundo nível.

Quando  $x \leq -25$  temos  $(15 - x)/2 \geq 20$ . Além disso, o intervalo viável de  $y'$  é  $[0, 20]$ , o que implica que a solução do problema é  $\bar{y}' = 20$ .

Dessa maneira, o conjunto  $\Psi(x)$  é unitário em todo  $x$  para o qual está bem definida e pode ser descrito pela função:

$$y(x) = \begin{cases} 20, & \text{se } x \leq -25 \\ \frac{15 - x}{2}, & \text{se } -25 \leq x \leq 15 \\ 0, & \text{se } 15 \leq x \leq 20. \end{cases}$$

Quando substituimos o problema do segundo nível pela função  $y(x)$ , é fácil verificar que o seguinte problema de otimização é formado:

$$\begin{aligned} & \underset{x, y}{\text{Minimizar}} && (10 - x)^3 + (10 - y)^3 \\ & \text{s. a} && y = (15 - x)/2 \\ & && -x + y \leq 0 \\ & && 0 \leq x \leq 15, \end{aligned}$$

pois para o intervalo viável de  $x$  no primeiro nível, o único valor de  $y(x)$  é  $(15 - x)/2$ . Este problema pode ser transformado (eliminando-se a variável  $y$ ) em um problema de uma variável, cuja única solução é  $\bar{x} = \frac{85-30\sqrt{2}}{7}$ . Isto implica que  $\bar{y} = \frac{20+30\sqrt{2}}{14}$ . Logo, o minimizador global do problema é  $(\bar{x}, \bar{y}) = (\frac{85-30\sqrt{2}}{7}, \frac{20+30\sqrt{2}}{14}) \cong (6.08194, 4.45902)$  e que pode ser visto na Figura 4.7.

#### Problema 9 [Col02, BIPA2]:

$$\begin{aligned} & \underset{x}{\text{Minimizar}} && (x - 5)^2 + (2y + 1)^2 \\ & \text{s. a} && x \geq 0 \\ & && y \in \underset{y'}{\arg \min} (y' - 1)^2 - 1.5xy' + x^3 \\ & && \text{s. a} && -3x + y' + 3 \leq 0 \\ & && && x - 0.5y' - 4 \leq 0 \\ & && && x + y' - 7 \leq 0 \\ & && && y' \geq 0. \end{aligned}$$



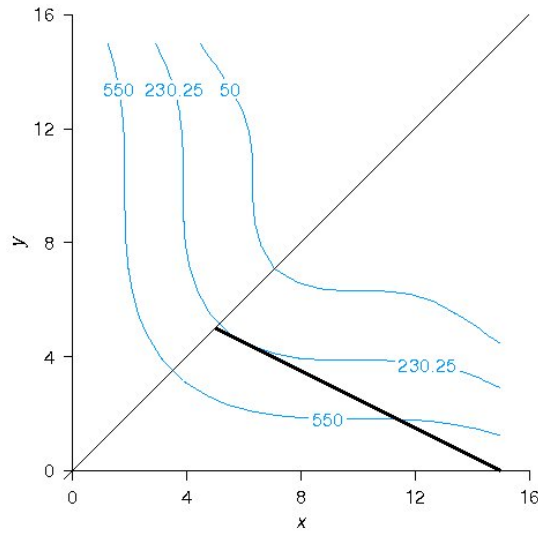


Figura 4.7: Ilustração da solução do Problema 8. As curvas numeradas são as curvas de nível da função  $F(x, y)$  e em negrito está a região induzida do BLPP.

**Classificação:** BLPP quadrático.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é  $(x, y) = (4, 0)$ .

**Solução reportada:** Em [CMS05] é reportada uma solução com  $\bar{F} = 17$  e  $\bar{f} = 2$ , porém o ponto não é apresentado.

Este problema é semelhante ao Problema 3, a menos da função objetivo do segundo nível, que possui o termo  $x^3$  a mais. Como a variável do primeiro nível é considerada constante no segundo nível, este termo não altera o conjunto  $\Psi(x)$ , de forma que o problema pode ser visualizado na Figura 4.4 e possui minimizador global  $(1, 0)$ .

**Problema 10 [Col02, BIPA3]:**

$$\begin{aligned}
 &\underset{x}{\text{Minimizar}} && (x - 5)^4 + (2y + 1)^4 \\
 \text{s. a} &&& x + y - 4 \leq 0 \\
 &&& x \geq 0 \\
 &&& y \in \underset{y'}{\arg \min} \exp(-x + y') + x^2 + 2xy' + y'^2 + 2x + 6y' \\
 &&& \text{s. a} && -x + y' - 2 \leq 0 \\
 &&& && y' \geq 0.
 \end{aligned}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é  $(x, y) = (4, 0)$ .

**Solução reportada:** Em [CMS05] o minimizador do problema não foi apresentado e sim os valores da função objetivo do primeiro e do segundo nível  $\bar{F} = 2.0$  e  $\bar{f} = 24.02$ , respectivamente.

Quando observamos apenas o problema do segundo nível e o substituímos por suas condições KKT, obtemos o seguinte sistema:

$$\begin{aligned} -x + y - 2 &\leq 0 \\ \exp(-x + y) + 2x + 2y + 6 + \mu - \mu^l &= 0 \\ \mu(-x + y - 2) &= 0 \\ \mu^l y &= 0 \\ y, \mu, \mu^l &\geq 0. \end{aligned}$$

Notemos que o problema do segundo nível enquadra-se nas condições da Proposição 1.2.4, pois suas restrições são lineares em  $y$  e para todo  $x$ ,  $f(x, y)$  é convexa em  $y$ . Para verificar tal fato, basta notar que a segunda derivada de  $f(x, y)$  com relação a  $y$  é estritamente positiva. Desta forma, as condições KKT são necessárias e suficientes para o problema do segundo nível.

Para que a restrição

$$\exp(-x + y') + 2x + 2y' + 6 + \mu - \mu^l = 0$$

seja válida, devemos ter  $\mu^l > 0$ , pois o termo  $\exp(-x + y') + 2x + 2y' + 6 + \mu$  é estritamente positivo para  $x \geq 0$  (restrição do primeiro nível). Isto implica que  $y' = 0$ . Logo,  $\Psi(x) = \{0\}$  para todo  $x \geq 0$ . Com isso, temos que o problema em dois níveis original é equivalente ao seguinte problema de otimização:

$$\begin{aligned} &\underset{x}{\text{Minimizar}} && (x - 5)^4 + 1 \\ &\text{s. a} && 0 \leq x \leq 4, \end{aligned}$$

cujo minimizador global é  $\bar{x} = 4$ . Logo,  $(\bar{x}, \bar{y}) = (4, 0)$  é o minimizador global do problema.

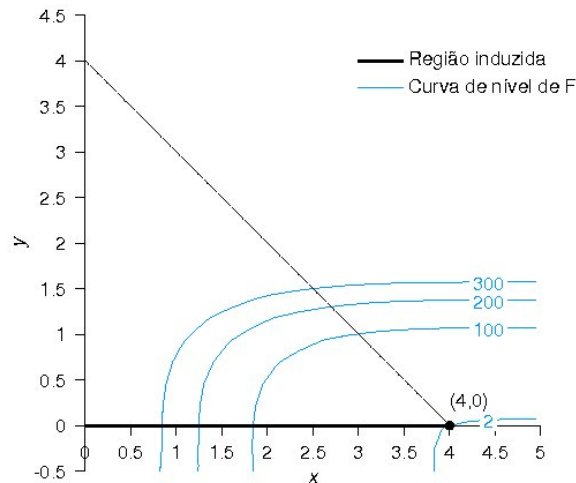


Figura 4.8: Ilustração do Problema 10 e seu minimizador global  $(\bar{x}, \bar{y}) = (4, 0)$ .

**Problema 11 [Col02, BIPA4]:**

$$\begin{aligned}
& \underset{x}{\text{Minimizar}} && x^2 + (y - 10)^2 \\
& \text{s. a} && x + 2y - 6 \leq 0 \\
& && x \geq 0 \\
& y \in \underset{y'}{\arg \min} && x^3 + 2y'^3 + x - 2y' - x^2 \\
& \text{s. a} && -x + 2y' - 3 \leq 0 \\
& && y' \geq 0.
\end{aligned}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é  $(x, y) = (1.5, 2.25)$ .

**Solução reportada:** Em [CMS05] o minimizador do problema não foi apresentado e sim os valores da função objetivo do primeiro e do segundo nível  $\bar{F} = 88.79$  e  $\bar{f} = -0.77$ , respectivamente.

Vamos verificar que  $(\bar{x}, \bar{y}) = (0, 0.57735)$  é o minimizador global do BLPP. Para  $y' \geq 0$  temos que  $f(x, y')$  é convexa em  $y'$  para todo  $x$ , pois  $\frac{\partial^2 f}{\partial y'^2} = 12y' \geq 0$ . Desta forma pode-se verificar que  $\bar{y}' = \frac{\sqrt{3}}{3}$  é o minimizador de  $f(x, y')$ , para  $y' \geq 0$  e um  $x$  fixado. Ao considerarmos a outra restrição do segundo nível, temos que para  $x \geq \frac{2\sqrt{3}-9}{3}$  o ponto  $\bar{y}' = \frac{\sqrt{3}}{3}$  pertence ao conjunto viável do segundo nível e é, portanto, o seu minimizador global. Como uma restrição do primeiro nível pede que  $x \geq 0$ , temos que, para todo  $x$  viável,  $\Psi(x) = \left\{ \frac{\sqrt{3}}{3} \right\}$ . Com isso, o problema em dois níveis pode ser reescrito como um problema de otimização com uma variável, basta substituímos  $y$  por  $\frac{\sqrt{3}}{3}$ :

$$\begin{aligned}
& \underset{x}{\text{Minimizar}} && x^2 \\
& \text{s. a} && 0 \leq x \leq 6 - \frac{2\sqrt{3}}{3},
\end{aligned}$$

cujo minimizador é  $\bar{x} = 0$ . Logo, o minimizador do BLPP é  $(\bar{x}, \bar{y}) = (0, \frac{\sqrt{3}}{3}) \cong (0, 0.57735)$ . A Figura 4.9 apresenta o problema e o seu minimizador global.

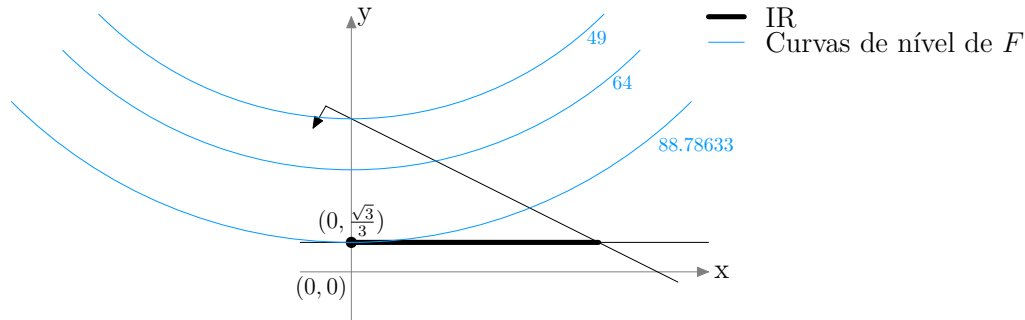


Figura 4.9: Ilustração do Problema 11.

**Problema 12 [Col02, BIPA5]:**

$$\begin{array}{ll}
\text{Minimizar} & (x - y_2)^4 + (y_1 - 1)^2 + (y_1 - y_2)^2 \\
\text{s. a} & x \geq 0 \\
& y \in \arg \min_{y'} 2x + \exp(y'_1) + y_1'^2 + 4y'_1 + 2y_2'^2 - 6y'_2 \\
& \text{s. a} \quad 6x + y_1'^2 + \exp(y'_2) - 15 \leq 0 \\
& \quad 5x + y_1'^4 - y'_2 - 25 \leq 0 \\
& \quad 0 \leq y'_1 \leq 4 \\
& \quad 0 \leq y'_2 \leq 2.
\end{array}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é  $(x, y) = (1, 1, 1)$ .

**Solução reportada:** Em [CMS05] encontrou-se uma solução com  $\bar{F} = 2.75$  e  $\bar{f} = 0.57$ , mas o ponto não foi apresentado.

O problema do segundo nível não satisfaz as condições da Proposição 1.2.4, mas a função objetivo do segundo nível é convexa em  $y'$ , para  $x \geq 0$ , pois sua Hessiana é definida positiva, o que significa que não existem maximizadores nem pontos de sela. Primeiramente vamos analisar as condições KKT do problema do segundo nível:

$$\begin{aligned}
6x + y_1'^2 + \exp(y'_2) - 15 &\leq 0 \\
5x + y_1'^4 - y'_2 - 25 &\leq 0 \\
\exp(y'_1) + 2y'_1 + 4 + 2y'_1\mu_1 + 4y_1'^3\mu_2 - \mu_1^l + \mu_1^u &= 0 \\
4y'_2 - 6 + \exp(y'_2)\mu_1 - \mu_2 - \mu_2^l + \mu_2^u &= 0 \\
&\vdots \\
\mu_1^l y'_1 &= 0 \\
\mu_1^u (y'_1 - 4) &= 0 \\
&\vdots \\
0 &\leq y'_1 \leq 4 \\
0 &\leq y'_2 \leq 2 \\
\mu_1, \mu_2, \mu_1^l, \mu_2^l, \mu_1^u, \mu_2^u &\geq 0.
\end{aligned}$$

Observando a equação

$$\exp(y'_1) + 2y'_1 + 4 + 2y'_1\mu_1 + 4y_1'^3\mu_2 - \mu_1^l + \mu_1^u = 0,$$

podemos deduzir que  $\mu_1^l > 0$ , pois sabemos que  $y'_1$ ,  $\mu_1$ ,  $\mu_2$  e  $\mu_1^u$  são maiores ou iguais a zero. Isso implica que devemos ter  $\bar{y}'_1 = 0$  e podemos eliminá-la do problema original. Desta

maneira, o problema do segundo nível torna-se:

$$\begin{aligned} & \underset{y'_2}{\text{Minimizar}} && 2x + 1 + 2y'^2_2 - 6y'_2 \\ & \text{s. a} && 6x + \exp(y'_2) - 15 \leq 0 \\ & && 5x - y'_2 - 25 \leq 0 \\ & && 0 \leq y'_2 \leq 2. \end{aligned}$$

Notemos que a solução do problema irrestrito é  $y'_2 = 1.5$ . As restrições de caixa em  $y'_2$  são:

$$\max\{0, 5x - 25\} \leq y'_2 \leq \min\{2, \ln(15 - 6x)\}.$$

Podemos concluir que, para  $x \leq \frac{15 - \exp(1.5)}{6}$ , a solução do problema será  $\bar{y}'_2 = 1.5$ , basta substituir  $y'_2$  por 1.5 nas desigualdades acima e resolver com relação a  $x$ . Quando  $x \geq \frac{15 - \exp(1.5)}{6}$ , temos que  $1.5 \geq \min\{2, \ln(15 - 6x)\} = \ln(15 - 6x)$ . Também podemos verificar que  $y_2$  deve pertencer à caixa  $[0, \ln(15 - 6x)]$  até o momento que o intervalo fica inviável, ou seja, quando  $x > \frac{14}{6}$ . Logo, para  $\frac{15 - \exp(1.5)}{6} \leq x \leq \frac{14}{6}$  temos que a solução do problema é  $y_2 = \ln(15 - 6x)$ . A Figura 4.10(b) apresenta o conjunto  $\Psi(x)$  que, por ser unitário para todo  $x$ , pode ser representado pela função:

$$y(x) = \begin{cases} (0, 1.5), & \text{se } x \leq \frac{15 - \exp(1.5)}{6} \\ (0, \ln(15 - 6x)), & \text{se } \frac{15 - \exp(1.5)}{6} \leq x \leq \frac{14}{6}. \end{cases}$$

Para encontrar o minimizador global, minimizamos a função objetivo do primeiro nível em cada uma das duas restrições que compõem  $y(x)$ . A Figura 4.10(a) ilustra as curvas de nível de  $F(x, y)$  para  $y_1 = 0$  e a Figura 4.10(b) ilustra a região induzida do BLPP, onde  $F(x, y)$  será minimizada. Há dois casos a considerar:

**Caso 1.**  $0 \leq x \leq \frac{15 - \exp(1.5)}{6}$ . Isto implica que  $(y_1, y_2) = (0, 1.5)$  e obtemos o seguinte problema:

$$\begin{aligned} & \underset{x}{\text{Minimizar}} && (x - 1.5)^4 + 3.25 \\ & \text{s. a} && 0 \leq x \leq \frac{15 - \exp(1.5)}{6}. \end{aligned}$$

O minimizador deste problema é  $\bar{x}^1 = 1.5$ . A função objetivo do primeiro nível possui valor  $\bar{F}^1 = 3.25$ .

**Caso 2.**  $\frac{15 - \exp(1.5)}{6} \leq x \leq \frac{14}{6}$ . Isto implica que  $(y_1, y_2) = (0, \ln(15 - 6x))$ . Neste caso resolvemos o seguinte problema, ilustrado na Figura 4.11:

$$\begin{aligned} & \underset{x}{\text{Minimizar}} && (x - \ln(15 - 6x))^4 + \ln^2(15 - 6x) + 1 \\ & \text{s. a} && \frac{15 - \exp(1.5)}{6} \leq x \leq \frac{14}{6}. \end{aligned}$$

Não conseguimos resolver este problema analiticamente, mas o seu minimizador (encontrado por ALGENCAN) é  $\bar{x}^2 \cong 1.94053$ . Logo,  $\bar{y}_2^2 \cong 1.211$ . Neste ponto, a função objetivo do primeiro nível possui valor  $\bar{F}^2 \cong 2.74977$ .

Logo, temos que o minimizador global do problema é o ponto  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (1.94053, 0, 1.211)$ .

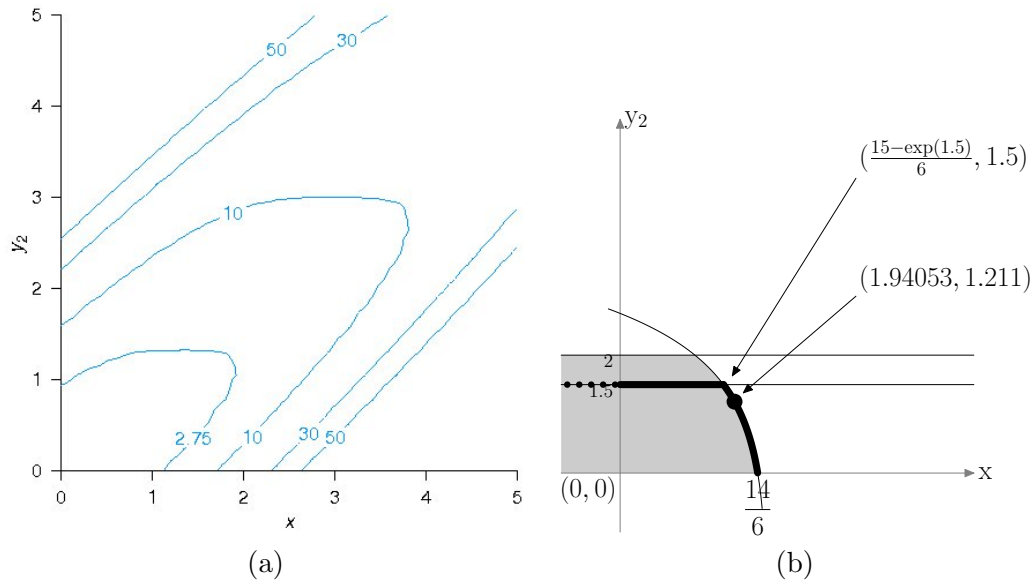


Figura 4.10: Curvas de nível (a) e região induzida (b) do Problema 12.

### Problema 13 [GF01, Exemplo 1]:

$$\begin{aligned}
 &\underset{x,y}{\text{Minimizar}} && 16x^2 + 9y^2 \\
 &\text{s. a} && -4x + y \leq 0 \\
 &&& x \geq 0 \\
 &y \in \arg \min_{y'} && (x + y' - 20)^4 \\
 &\text{s. a} && 4x + y' - 50 \leq 0 \\
 &&& y' \geq 0.
 \end{aligned}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Não fornecido.

**Solução reportada:** As soluções encontradas por [GF01] são  $(\bar{x}^1, \bar{y}^1) = (7.2, 12.8)$  (minimizador local) e  $(\bar{x}^2, \bar{y}^2) = (11.25, 5)$  (minimizador global).

**Observações:** Este problema foi retirado do trabalho de Gümüş e Floudas [GF01], mas foi originalmente formulado em [SIB97].

Utilizando o mesmo método usado para analisar o Problema 3, construímos o conjunto  $\Psi(x)$ , que pode ser representado pela função  $y(x)$ :

$$y(x) = \begin{cases} 20 - x, & \text{se } x \leq 10 \\ 50 - 4x, & \text{se } 10 \leq x \leq 12.5. \end{cases}$$

Para encontrar o minimizador global, minimizamos a função objetivo do primeiro nível em cada uma das faces de  $y(x)$ , associadas ao intervalo viável de  $x$  e às restrições do primeiro nível. O minimizador global é  $(\bar{x}, \bar{y}) = (11.25, 5)$ . O problema está ilustrado na Figura 4.12.

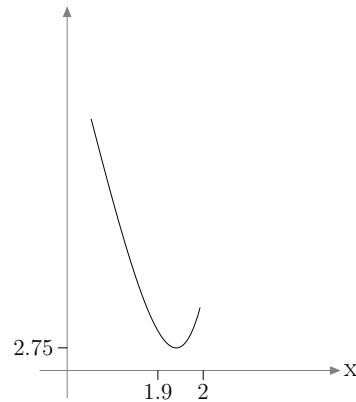


Figura 4.11: Representação da função objetivo para o problema considerado no caso **2**, encontrado durante a resolução analítica do Problema 12.

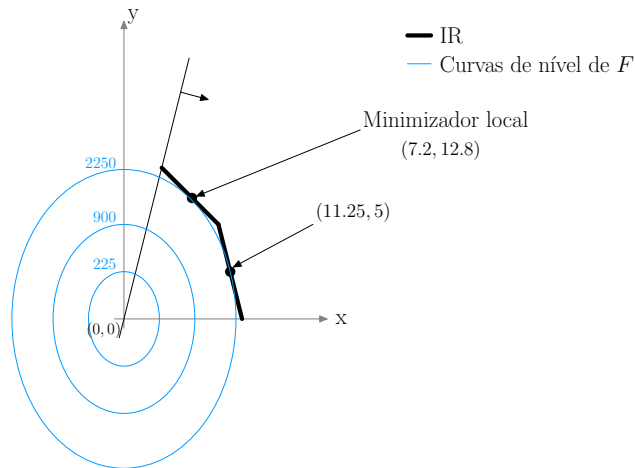


Figura 4.12: Representação gráfica do Problema 13 e seu minimizador global.

#### Problema 14 [GF01, Exemplo 2]:

$$\begin{aligned}
 &\text{Minimizar}_{x,y} && x^3 y_1 + y_2 \\
 \text{s. a} &&& 0 \leq x \leq 1 \\
 &&& y \in \arg \min_{y'} -y_2' \\
 &&& \text{s. a} && x y_1' \leq 10 \\
 &&& && y_1'^2 + x y_2' \leq 1 \\
 &&& && y_2' \geq 0.
 \end{aligned}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Não fornecido.

**Solução reportada:** A solução encontrada em [GF01] é  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (1, 0, 1)$ .

**Observações:** Este problema foi retirado do trabalho de Gümüş e Floudas [GF01], mas foi originalmente formulado em [FZ98].

Utilizamos o algoritmo de otimização global BARON para resolver o problema reformulado, porém, o problema do segundo nível não satisfaz as condições da Proposição 1.2.4, pois suas restrições não são lineares em  $y$ . Portanto, não sabemos se o minimizador encontrado por BARON é o minimizador global do BLPP original. Independentemente deste fato, BARON encontrou o minimizador  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (1, 0, 1)$ . Este ponto está na região induzida do problema, pois para  $x = 1$  temos que  $(y_1, y_2) = (0, 1)$  é o minimizador global do segundo nível.

**Problema 15 [Sav89]:**

$$\begin{aligned}
 &\underset{x}{\text{Minimizar}} && -8x_1 - 4x_2 + 4y_1 - 40y_2 - 4y_3 \\
 &\text{s. a} && x_1 + 2x_2 - y_3 - 1.3 \leq 0 \\
 &&& x_1, x_2 \geq 0 \\
 &y \in \underset{y'}{\arg \min} && 2y'_1 + y'_2 + 2y'_3 \\
 &\text{s. a} && -y'_1 + y'_2 + y'_3 - 1 \leq 0 \\
 &&& 4x_1 - 2y'_1 + 4y'_2 - y'_3 - 2 \leq 0 \\
 &&& 4x_2 + 4y'_1 - 2y'_2 - y'_3 - 2 \leq 0 \\
 &&& y'_1, y'_2, y'_3 \geq 0.
 \end{aligned}$$

**Classificação:** BLPP linear.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é

$$(x_1, x_2, y_1, y_2, y_3) = (0.75, 0.75, 0, 0, 1)$$

**Solução reportada:** Não foi reportada solução.

**Observações:** Este problema foi introduzido na tese de doutorado de Savard [Sav89], mas foi retirado de [Col02].

Utilizamos o algoritmo de otimização global BARON para encontrar o minimizador global do problema reformulado, cujo conjunto viável é equivalente à região induzida do BLPP original, pela Proposição 1.2.4. A solução encontrada foi  $(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2, \bar{y}_3) = (0.5, 0.8, 0, 0.2, 0.8)$ .

**Problema 16 [SC98, Exemplo 3]:**

$$\begin{aligned}
 &\underset{x}{\text{Minimizar}} && (x - 3)^2 + (y - 2)^2 \\
 &\text{s. a} && -2x + y - 1 \leq 0 \\
 &&& x - 2y + 2 \leq 0 \\
 &&& x + 2y - 14 \leq 0 \\
 &&& 0 \leq x \leq 8 \\
 &y \in \underset{y'}{\arg \min} && (y' - 5)^2.
 \end{aligned}$$

**Classificação:** BLPP quadrático.



**Ponto inicial:** Fornecido em [SC98]. O ponto utilizado é:  $(x, y) = (3, 1)$ .

**Solução reportada:** A solução encontrada em [SC98] é  $(\hat{x}, \hat{y}) = (3.132, 5.003)$ , com  $\bar{F} = 9.03543$  e  $\bar{f} = 0.00001$ . O minimizador global apresentado é  $(\bar{x}, \bar{y}) = (3, 5)$ .

Para este problema, temos que  $\Psi(x) = \{5\}$  para todo  $x$ . Logo, podemos eliminar o problema do segundo nível e substituir  $y$  por 5 no problema do primeiro nível, obtendo um problema equivalente. O problema formado é como se segue:

$$\begin{array}{ll} \text{Minimizar} & (x - 3)^2 + 9 \\ \text{s. a} & 2 \leq x \leq 4, \end{array}$$

cujas solução é  $\bar{x} = 3$ . O problema pode ser visualizado na Figura 4.13.

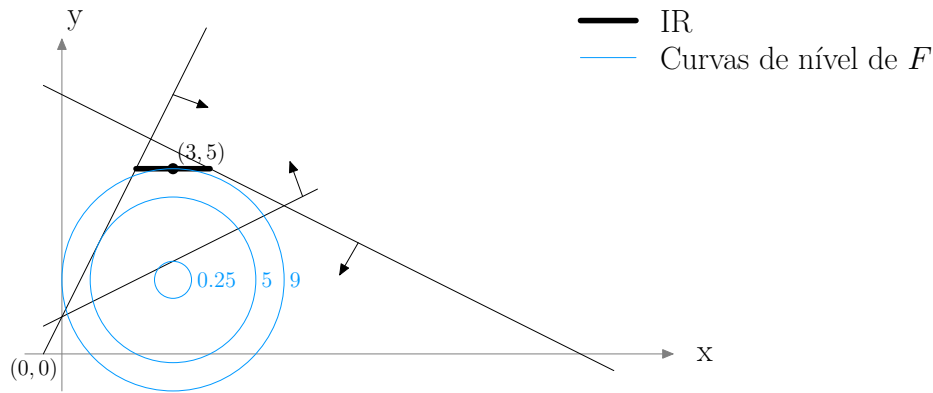


Figura 4.13: Ilustração da solução global do Problema 16.

#### Problema 17 [SA81, Exemplo 1]:

$$\begin{array}{ll} \text{Minimizar} & x^2 + (y - 10)^2 \\ \text{s. a} & -x + y \leq 0 \\ & 0 \leq x \leq 15 \\ & y \in \arg \min_{y'} (x + 2y' - 30)^2 \\ & \text{s. a} \quad x + y' - 20 \leq 0 \\ & 0 \leq y' \leq 20. \end{array}$$

**Classificação:** BLPP quadrático.

**Ponto inicial:** Fornecido em [Col02]. O ponto utilizado é:  $(x, y) = (5, 5)$ .

**Solução reportada:** A solução do problema, apresentada em [SA81], é:  $(\bar{x}, \bar{y}) = (10, 10)$ , com  $\bar{F} = 100$  e  $\bar{f} = 0$ .

Com um processo semelhante ao usado no Problema 3 para encontrar o minimizador global, primeiramente encontramos o conjunto  $\Psi(x)$  que, neste caso, pode ser substituído

pela função:

$$y(x) = \begin{cases} \frac{30-x}{2}, & \text{se } x \leq 10 \\ 20-x, & \text{se } 10 \leq x \leq 20. \end{cases}$$

Como pode ser observado na Figura 4.14, quando minimizamos o problema do primeiro nível em cada uma das faces de  $y(x)$  e consideramos as restrições do primeiro nível, temos que o BLPP é equivalente ao seguinte PNL:

$$\begin{aligned} & \underset{x,y}{\text{Minimizar}} && x^2 + (y-10)^2 \\ & \text{s. a} && -x + y \leq 0 \\ & && x + y - 20 = 0 \\ & && 0 \leq x \leq 15, \end{aligned}$$

pois apenas uma das faces que compõem  $y(x)$  é viável no problema do primeiro nível. O minimizador global é  $(\bar{x}, \bar{y}) = (10, 10)$ .

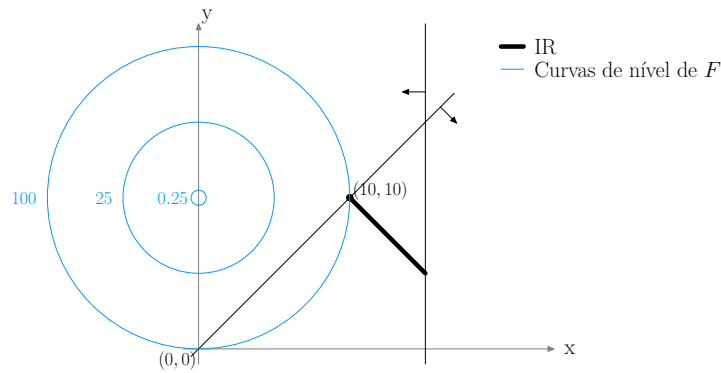


Figura 4.14: Ilustração do Problema 17.

#### Problema 18 [SA81, Exemplo 2]:

$$\begin{aligned} & \underset{x,y}{\text{Minimizar}} && (x_1 - 30)^2 + (x_2 - 20)^2 - 20y_1 + 20y_2 \\ & \text{s. a} && x_1 + 2x_2 \geq 30 \\ & && x_1 + x_2 \leq 25 \\ & && x_2 \leq 15 \\ & && y \in \underset{y'}{\arg \min} (x_1 - y'_1)^2 + (x_2 - y'_2)^2 \\ & && \text{s. a} && 0 \leq y'_1, y'_2 \leq 10. \end{aligned}$$

**Classificação:** BLPP quadrático.

**Ponto inicial:** Não fornecido.

**Solução reportada:** Em [SA81] a solução reportada é  $(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (20, 5, 10, 5)$ , com  $\bar{F} = 225$  e  $\bar{f} = 100$ .

Para encontrar o minimizador global, utilizamos a mesma técnica com a qual encontramos o minimizador do Problema 1, pois as variáveis  $y_1$  e  $y_2$  não estão relacionadas entre si. Temos que:

$$y_1(x_1) = \begin{cases} 0, & \text{se } x_1 \leq 0 \\ x_1, & \text{se } 0 \leq x_1 \leq 10 \\ 10, & \text{se } x_1 \geq 10 \end{cases} \quad \text{e} \quad y_2(x_2) = \begin{cases} 0, & \text{se } x_2 \leq 0 \\ x_2, & \text{se } 0 \leq x_2 \leq 10 \\ 10, & \text{se } x_2 \geq 10. \end{cases}$$

Podemos, então minimizar a função objetivo do primeiro nível, sujeito a  $x_1$  e  $x_2$  pertencerem a cada um dos intervalos e ainda satisfazerem as restrições do primeiro nível. Assim temos a seguinte tabela, que mostra o mínimo encontrado para cada um dos problemas gerados pela combinação dos intervalos:

$F(x, y)$	$x_1 \leq 0$	$0 \leq x_1 \leq 10$	$x_1 \geq 10$
$x_2 \leq 0$	Inviável	Inviável	Inviável
$0 \leq x_2 \leq 10$	Inviável	500	225
$x_2 \geq 10$	1225	425	325

Considerando os intervalos  $x_1 \geq 10$  e  $0 \leq x_2 \leq 10$ , temos que  $y_1 = 10$  e  $y_2 = x_2$ . Logo, resolvemos o seguinte problema de otimização:

$$\begin{aligned} &\underset{x}{\text{Minimizar}} && (x_1 - 30)^2 + (x_2 - 20)^2 + 20x_2 - 200 \\ \text{s. a} &&& x_1 + 2x_2 \geq 30 \\ &&& x_1 + x_2 \leq 25 \\ &&& x_1 \geq 10 \\ &&& 0 \leq x_2 \leq 10, \end{aligned}$$

cujo minimizador global é  $(\bar{x}_1, \bar{x}_2) = (20, 5)$ . Logo, obtemos a solução  $(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (20, 5, 10, 5)$ , com valor  $\bar{F} = 225$ . Este ponto é o minimizador global do problema em dois níveis original.

#### Problema 19 [ZXL04, Exemplo 5.1]:

$$\begin{aligned} &\underset{x, y}{\text{Minimizar}} && x + 5y \\ \text{s. a} &&& x \geq 0 \\ &&& y \in \underset{y'}{\arg \min} -y' \\ &&& \text{s. a} && x + y' \geq 8 \\ &&& && -3x + 2y' \leq 6 \\ &&& && x + 4y' \leq 48 \\ &&& && x - 5y' \leq 9 \\ &&& && y' \geq 0. \end{aligned}$$

**Classificação:** BLPP linear.

**Ponto inicial:** Não fornecido.

**Solução reportada:** Dois minimizadores são reportados em [ZXL04]:  $(2, 6)$  (minimizador global) e  $(30\frac{2}{3}, 4\frac{1}{3})$  (minimizador local).

Para verificarmos que a solução encontrada pelo método em estudo é o minimizador global do problema, primeiramente podemos notar que o problema do segundo nível é um problema de programação linear apenas com restrições de caixa em  $y$ . Construímos, então, a função  $y(x)$  (pois  $\Psi(x)$  é unitário para todo  $x$ ) da mesma maneira feita com o Problema 3:

$$y(x) = \begin{cases} \frac{6 + 3x}{2}, & \text{se } 2 \leq x \leq \frac{36}{7} \\ \frac{48 - x}{4}, & \text{se } \frac{36}{7} \leq x \leq \frac{92}{3}. \end{cases}$$

Quando minimizamos a função objetivo do primeiro nível em cada uma das faces que compõem  $y(x)$  encontramos o minimizador global  $(2, 6)$ . O problema está representado através da Figura 4.15.

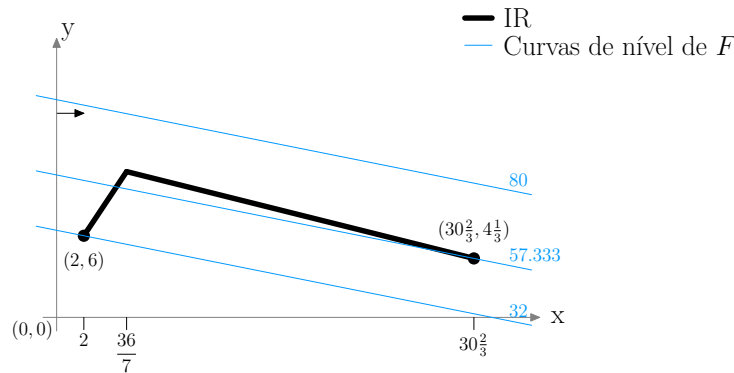


Figura 4.15: Ilustração da solução do Problema 19.

**Problema 20 [ZXL04, Exemplo 5.2]:**

$$\begin{aligned} &\text{Minimizar}_{x,y} \quad (x - 3.5)^2 + (y + 4)^2 \\ &\text{s. a} \quad y \in \arg \min_{y'} \quad (y' - 3)^2 \\ &\quad \text{s. a} \quad y'^2 - x \leq 0. \end{aligned}$$

**Classificação:** BLPP geral.

**Ponto inicial:** Não fornecido.

**Solução reportada:** A solução encontrada em [ZXL04] foi  $(\bar{x}, \bar{y}) = (1, 1)$ , com valores  $F = 31.25$  e  $f = 4$ .

Para encontrar o minimizador global, vamos primeiramente construir o conjunto  $\Psi(x)$ . A solução irrestrita do problema do segundo nível é  $y' = 3$ . A única restrição do segundo nível pede que, para cada  $x \geq 0$ ,  $y'$  pertença à caixa  $[-\sqrt{x}, \sqrt{x}]$ . Podemos verificar que se  $x \geq 9$  a solução irrestrita  $\bar{y}' = 3$  está dentro da caixa. Para  $0 \leq x \leq 9$  temos que  $\sqrt{x} \leq 3$

e, pelo fato da função do segundo nível ser uma quadrática em  $y$ , é fácil ver que a solução é o extremo da caixa mais próximo de  $y' = 3$ . Logo, a solução é  $\bar{y}' = \sqrt{x}$ . Para  $x < 0$  o problema do segundo nível torna-se inviável. Como, para cada  $x$ ,  $y'$  está univocamente determinado, podemos utilizar  $y(x)$  no lugar do conjunto  $\Psi(x)$ . Assim temos:

$$y(x) = \begin{cases} \sqrt{x}, & \text{se } 0 \leq x \leq 9 \\ 3, & \text{se } x \geq 9, \end{cases}$$

como pode ser visto na Figura 4.17(a).

Ao considerarmos o intervalo  $x \geq 9$ , temos que  $y = 3$ , o que implica que o BLPP é equivalente ao seguinte PNL:

$$\begin{aligned} &\underset{x}{\text{Minimizar}} && (x - 3.5)^2 + 49 \\ &\text{s. a} && x \geq 9, \end{aligned}$$

cuja solução é  $\bar{x} = 9$ , com valor 79.25. Ao considerarmos  $0 \leq x \leq 9$ , temos que  $y = \sqrt{x}$ , ou então, como será utilizado,  $x = y^2$ , para  $y \geq 0$ . Neste caso, resolvemos o seguinte problema de otimização em  $y$ :

$$\begin{aligned} &\underset{y}{\text{Minimizar}} && (y^2 - 3.5)^2 + (y + 4)^2 \\ &\text{s. a} && y \geq 0. \end{aligned}$$

A solução deste problema é  $y = 0$  ( $\Rightarrow x = 0$ ), com valor 28.25, que pode ser encontrada resolvendo o sistema KKT:

$$\begin{aligned} 4y^3 - 12y + 8 - \mu &= 0 \\ \mu y &= 0 \\ \mu, y &\geq 0. \end{aligned}$$

Notemos que  $y = 1$  ( $\Rightarrow x = 1$ ), a solução encontrada em [ZXL04], também satisfaz as condições KKT, mas, como podemos ver na Figura 4.16, é um ponto de sela do problema.

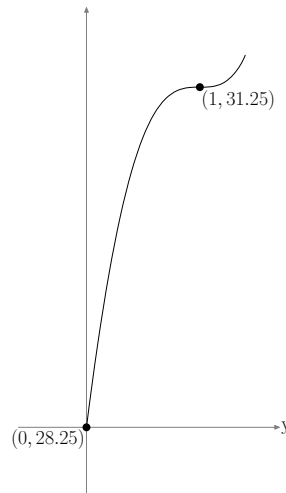


Figura 4.16: Gráfico de  $(y^2 - 3.5)^2 + (y + 4)^2$  no intervalo  $[0, 1.5]$ . Note que a origem está deslocada para  $(0, 28.25)$ .

Assim temos que o minimizador global do BLPP é  $(\bar{x}, \bar{y}) = (0, 0)$ , apresentado na Figura 4.17(b). Para este ponto, porém, não existem multiplicadores de Lagrange que satisfaçam o sistema KKT associado ao problema do segundo nível:

$$\begin{aligned} y^2 - x &\leq 0 \\ 2y - 6 + 2\mu y &= 0 \\ \mu(y^2 - x) &= 0 \\ \mu &\geq 0. \end{aligned}$$

Desta forma, a solução não deveria ser encontrada pela técnica estudada neste trabalho. De fato, ALGENCAN não encontra exatamente o minimizador do problema, mas como tal ponto está em uma vizinhança onde os pontos possuem multiplicadores de Lagrange, consegue encontrar uma solução que satisfaz a precisão requisitada.

Embora nosso ponto inicial  $(x, y) = (0, 0)$  seja a solução do BLPP, devemos notar que  $(x, y, \mu) = (0, 0, 0)$  não é solução da reformulação. Em um outro teste, usamos  $(x, y, \mu) = (1, 1, 0)$  – solução encontrada por [ZXL04] – como ponto inicial, obtendo novamente o minimizador global.

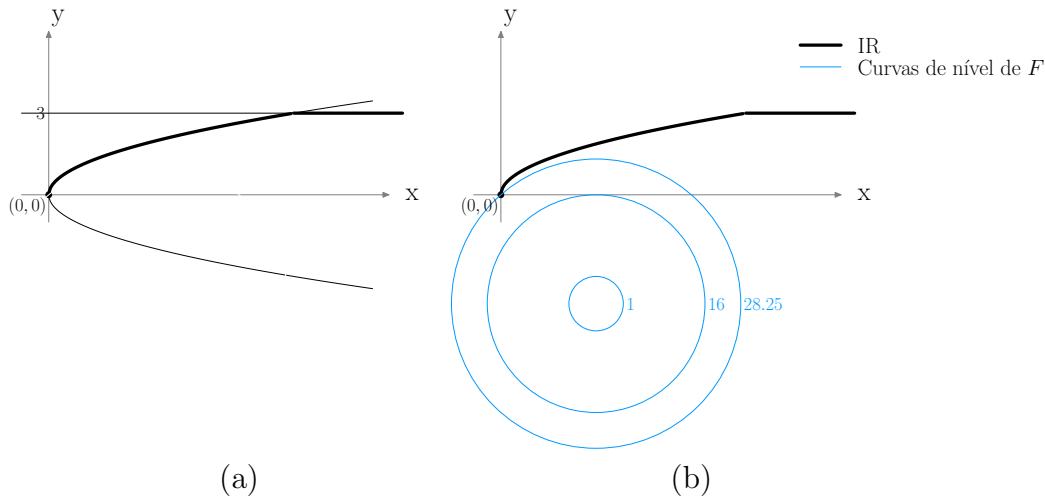


Figura 4.17: Ilustração do Problema 20. (a) A função  $y(x)$  representada em negrito. (b) Representação do BLPP: a região induzida em negrito e as curvas de nível da função objetivo do primeiro nível.

## 4.2 Resolução das reformulações

Os problemas foram reformulados com cada uma das três formas de representar as condições de complementaridade: S, M e FBP, e posteriormente resolvidos com o algoritmo de programação não linear ALGENCAN [ABMS07]. ALGENCAN é um algoritmo baseado em Lagrangianos Aumentados para problemas de programação não linear gerais que, em suas iterações internas utiliza um algoritmo de minimização em caixas denominado GENCAN [BM02] para resolver os subproblemas associados à função de Lagrangianos Aumentados.

ALGENCAN possui uma interface que nos permite resolver problemas formulados na linguagem de modelagem matemática AMPL [FGK90]. Esta linguagem utiliza técnicas de diferenciação automática para calcular o gradiente da função objetivo, o Jacobiano das restrições e a Hessiana da função Lagrangiana, todos usados por ALGENCAN para resolver um problema de programação não linear. Através desta linguagem de formulação de problemas de PNL pudemos contornar o inconveniente de calcular manualmente as derivadas de segunda ordem das funções do sistema KKT.

Quando fornecidos, os pontos iniciais do trabalho original são usados e, no caso em que não encontramos nenhum ponto inicial de referência, inicializamos todas as variáveis com zero (que chamaremos ponto inicial usual). É necessário ressaltar que na literatura, nenhum ponto inicial é fornecido especificamente para o problema reformulado, de forma que o valor inicial dos multiplicadores de Lagrange ( $\lambda$  e  $\mu$ ) é sempre zero.

Todos os problemas foram resolvidos com um computador Intel Core2 Quad 2.40GHz, com 4 GB de memória RAM e sistema operacional Linux (Ubuntu) de 64 bits. A precisão com que os valores são mostrados é de 5 casas decimais. As soluções encontradas foram comparadas com as soluções retiradas da literatura (quando disponíveis) e com as soluções analíticas dos problemas (apresentadas na seção anterior).

De acordo com a característica de cada problema podemos provar, ou não, que o conjunto viável da reformulação é equivalente à região induzida do BLPP associado. Em todos os casos nos quais há equivalência, os problemas se encaixam nos requisitos da Proposição 1.2.4. Para os outros casos, o que ocorre principalmente para os BLPPs gerais, nada podemos afirmar. Logo, não há garantia que a resolução da reformulação encontrará sequer um ponto na região induzida. Através da análise dos problemas, temos que, exceto nos problemas 12, 14 e 20, em todos os problemas o conjunto viável da reformulação e a sua região induzida são equivalentes. Como podemos notar, na maioria dos problemas temos esperança de encontrar um ponto estacionário do problema original.

A Tabela 4.1 apresenta a comparação do valor de mínimo encontrado para cada problema. Os pontos iniciais usados, minimizadores encontrados e informações adicionais encontram-se na Seção 4.1. Na tabela,  $\bar{F}_S$ ,  $\bar{F}_M$  e  $\bar{F}_{FBP}$  denominam o valor da função objetivo do primeiro nível encontrado para as reformulações S, M e FBP, respectivamente.  $\bar{F}_{lit}$  representa os resultados encontrados na literatura e  $\bar{F}_{an}$  é o valor da função objetivo avaliada no minimizador global do BLPP, encontrado analiticamente.

Para auxiliar na análise dos resultados, a Tabela 4.2 mostra a classificação do ponto encontrado por ALGENCAN ao resolver cada problema em cada uma das três reformulações.

Quando a reformulação é um problema equivalente ao respectivo BLPP, então, para todo ponto viável  $(x, y)$  do problema reformulado, temos que  $y \in \Psi(x)$ . Portanto  $y$  é um minimizador global do problema do segundo nível. ALGENCAN garante convergência a um ponto estacionário do problema de programação não linear, o que implica que neste caso podemos encontrar um ponto estacionário do BLPP (minimizador global, minimizador local, maximizador ou ponto de sela). Denotamos por **GG** o minimizador global, **LG** o minimizador local e **EG** os demais pontos estacionários. A teoria sobre a qual ALGENCAN está estruturado ainda prevê convergência para um ponto estacionário de inviabilidade, considerado, neste caso, uma falha e denotado por **NN**.

Quando não temos a equivalência entre o conjunto viável do problema reformulado e seu respectivo BLPP, não podemos garantir que  $y$  seja um minimizador global do problema

Problema	$F_S$	$F_M$	$F_{FBP}$	$F_{lit}$	$F_{an}$
1	0.0	0.0	5.0	0.0	0.0
2	-12.67871	-12.67871	-14.36	-14.36	-14.36
3	17.0	17.0	17.0	17.0	17.0
4	-	-6600.0	-6600.0	-6600.0	-6600.0
5	-11.0	-5.0	-11.0	-11.0	-11.0
6	-13.0	-13.0	-13.0	-13.0	-13.0
7	5.0	5.0	5.0	5.0	5.0
8	230.21980	230.21289	230.2350	-	230.26769
9	17.0	17.0	17.0	17.0	17.0
10	2.0	2.0	2.0	2.0	2.0
11	88.78633	88.78633	88.78633	88.79	88.78633
12	2.74977	3.25	2.74977	2.75	2.74977
13	2303.72583	2250.0	2250.0	2250.0	2250.0
14	1.0	1.0	1.0	1.0	1.0
15	-18.4	-18.4	-18.4	-	-18.4
16	9.0	9.0	9.0	9.03543	9.0
17	99.9998	100.0	100.0	100.0	100.0
18	225.0	225.0	225.0	225.0	225.0
19	32.0	-	32.0	32.0	32.0
20	28.25008	28.25040	28.25040	31.25	28.25

Tabela 4.1: Comparação das três reformulações realizadas neste trabalho com os resultados encontrados na literatura e analiticamente. Valores não disponíveis significam que o algoritmo não encontrou a solução para o ponto inicial fornecido (nas reformulações S, M e FBP) ou que não foram encontrados resultados na literatura.

do segundo nível, pois um minimizador local, maximizador ou ponto de sela também pode satisfazer as condições KKT. Nessa situação, além da possibilidade de acontecer um dos casos descritos anteriormente, ALGENCAN podem encontrar  $(x, y)$  tal que  $y$  é um minimizador local do segundo nível (denotado **NL**), um maximizador do segundo nível (denotado **NM**) ou um ponto de sela (denotado **NS**). Denotamos também por **NN** o caso no qual o problema é inviável (por exemplo, se houver  $y$  minimizador global do segundo nível, para o qual não existem multiplicadores de Lagrange).

De acordo com a Tabela 4.2, com a reformulação S, que utiliza as restrições de complementaridade suaves, ALGENCAN resolveu 19 problemas e em 17 deles encontrou o minimizador global do problema original. No Problema 13 ALGENCAN encontrou um minimizador local conhecido e no Problema 2 encontrou um ponto da região induzida do BLPP, como mostrado na Seção 4.1. No Problema 4, ALGENCAN falhou em encontrar uma solução viável utilizando o ponto inicial fornecido.

Ao resolver a implementação M, que utiliza a função NCP mínimo, ALGENCAN encontrou o minimizador global em 16 dos 20 problemas considerados. No Problema 5 encontrou um minimizador local do BLPP original, apresentado na Seção 4.1. No problema 2 a solução encontrada pertence à região induzida do BLPP original, pois o problema do segundo nível satisfaz as condições da Proposição 1.2.4 e é um ponto estacionário da reformulação. No Problema 12 nada podemos afirmar, pois, para este BLPP, a reformulação não é um pro-



Problema	$F_S$	$F_M$	$F_{FBP}$	Problema	$F_S$	$F_M$	$F_{FBP}$
1	GG	GG	EG	11	GG	GG	GG
2	EG	EG	GG	12	GG	EG	GG
3	GG	GG	GG	13	LG	GG	GG
4	-	GG	GG	14	GG	GG	GG
5	GG	LG	GG	15	GG	GG	GG
6	GG	GG	GG	16	GG	GG	GG
7	GG	GG	GG	17	GG	GG	GG
8	GG	GG	GG	18	GG	GG	GG
9	GG	GG	GG	19	GG	-	GG
10	GG	GG	GG	20	GG	GG	GG

Tabela 4.2: Classificação das soluções encontradas com a resolução de cada reformulação.

blema equivalente ao original. Por fim, no Problema 19 ALGENCAN falhou em encontrar uma solução com o ponto inicial fornecido.

Com a reformulação que usa a versão perturbada da função NCP de Fischer-Burmeister, denominada FBP, ALGENCAN obteve os melhores resultados dentre as três reformulações. Dos 20 problemas resolvidos, ALGENCAN encontrou o minimizador global em 19 deles. No Problema 1, a solução encontrada pertence à região induzida do BLPP original (pois a reformulação é um problema equivalente), mas não sabemos afirmar se é um minimizador local do problema.

Notemos que a formulação M, que utiliza a função mínimo, obteve os piores resultados dentre as três formulações. Não podemos, porém, fazer uma comparação muito severa, dado que é a única não diferenciável das formulações. De fato, existem aproximações suaves da função mínimo (veja [CM96]) que poderiam ser usadas para contornar o inconveniente da falta de diferenciabilidade, da mesma forma que utilizamos a versão suave da função de Fischer-Burmeister para semelhante fim.

O motivo de utilizarmos a versão perturbada da função de Fischer-Burmeister é porque, como foi observado em testes preliminares, com a versão original da função, ALGENCAN não é capaz de resolver o problema reformulado, pois não consegue avaliar as derivadas na origem quando necessárias. Já no caso da função mínimo isso não ocorre pois é possível calculá-las, como mostramos na Seção 1.3.

As soluções encontradas com a resolução de cada problema com cada uma das reformulações são mostradas em detalhes nas Tabelas 4.3, 4.4 e 4.5.

Uma observação deve ser feita no que se refere aos testes realizados. ALGENCAN possui um grande número de parâmetros que podem ser alterados pelo usuário para cada problema. Observamos que com um tipo de configuração alguns problemas eram resolvidos com sucesso, enquanto outros problemas não. Escolhemos, então, duas configurações distintas e resolvemos todos os problemas com ambas. Os resultados aqui apresentados refletem os melhores dentre as duas configurações.

PROB.	REFORM.	SOLUÇÃO ENCONTRADA	$F(x, y)$	$f(x, y)$
1	S, M	$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (0, 0, -10, -10)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_1^u, \bar{\mu}_2^u) = (0, 0, 20, 20, 0, 0)$	0.0	200.0
	FBP	$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (25, 30, 5, 10)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_1^u, \bar{\mu}_2^u) = (0, 0, 0, 0, 0, 0)$	5.0	0.0
2	S, M	$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (0, 2, 1.875, 0.90625)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l) = (0, 1.25, 0, 0)$	-12.67871	-1.01562
	FBP	$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (2, 0, 1.6, 0.2)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l) = (0.44, 1.36, 0, 0)$	-14.36	9.56
3	S	$(\bar{x}, \bar{y}) = (1, 0)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l) = (3.5, 0, 0, 0)$	17.0	1.0
	M	$(\bar{x}, \bar{y}) = (1, 0)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l) = (3.57162, 0, 0, 0.07162)$	17.0	1.0
	FBP	$(\bar{x}, \bar{y}) = (1, 0)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l) = (9173.95203, 0, 0, 9170.45203)$	17.0	1.0
4	S	-	-	-
	M	$(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = (7.13919, 3.20878, 11.86081, 17.79122)$ $(\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4) = (0.34797, 10, 29.65203, 0)$ $(\bar{\mu}_1, \dots, \bar{\mu}_4) = (4.69595, 9.04279, 5.56755, 14.11488)$ $(\bar{\mu}_1^l, \dots, \bar{\mu}_4^l) = (0, 0, 0, 4.13175)$ $(\bar{\mu}_1^u, \dots, \bar{\mu}_4^u) = (0, 0, 0, 0)$	-6600.0	54.93811 (22.33732 + 32.60078)
	FBP	$(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = (7.37999, 3.56998, 11.62001, 17.43002)$ $(\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4) = (0.94997, 10, 29.05003, 0)$ $(\bar{\mu}_1, \dots, \bar{\mu}_4) = (5.89993, 6.23350, 12.40551, 11.56288)$ $(\bar{\mu}_1^l, \dots, \bar{\mu}_4^l) = (0, 0, 0, 8.15272)$ $(\bar{\mu}_1^u, \dots, \bar{\mu}_4^u) = (0, 0, 0, 0)$	-6600.0	57.70483 (18.30268 + 39.40214)

Tabela 4.3: Soluções encontradas por ALGENCAN para os problemas de 1 a 4.

PROB.	REFORM.	SOLUÇÃO ENCONTRADA	$F(x, y)$	$f(x, y)$
5	S	$(\bar{x}, \bar{y}) = (16, 11)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}_4, \bar{\mu}_5, \bar{\mu}^l) = (0, 0, 539.89445, 269.44722, 0, 0)$	-11.0	11.0
	M	$(\bar{x}, \bar{y}) = (0, 5)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}_4, \bar{\mu}_5, \bar{\mu}^l) = (0.5, 0, 0, 0, 0, 0)$	-5.0	5.0
	FBP	$(\bar{x}, \bar{y}) = (16, 11)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}_4, \bar{\mu}_5, \bar{\mu}^l) = (0, 0, 81.83565, 40.41782, 0, 0)$	-11.0	11.0
6	S, M, FBP	$(\bar{x}, \bar{y}_1, \bar{y}_2) = (5, 4, 2)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l, \bar{\mu}^u) = (0, 0, 0, 0, 1)$	-13.0	-4.0
7	S, M, FBP	$(\bar{x}, \bar{y}) = (1, 3)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3) = (4, 0, 0)$	5.0	4.0
8	S	$(\bar{x}, \bar{y}) = (6.08222, 4.45941)$ $(\bar{\mu}, \bar{\mu}^l, \bar{\mu}^u) = (0, 0, 0)$	230.21980	0.0
	M	$(\bar{x}, \bar{y}) = (6.08225, 4.45947)$ $(\bar{\mu}, \bar{\mu}^l, \bar{\mu}^u) = (0, 0, 0)$	230.21289	0.0
	FBP	$(\bar{x}, \bar{y}) = (6.08213, 4.45929)$ $(\bar{\mu}, \bar{\mu}^l, \bar{\mu}^u) = (0, 0, 0)$	230.2350	0.0
9	S	$(\bar{x}, \bar{y}) = (1, 0)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l) = (3.5, 0, 0, 0)$	17.0	2.0
	M	$(\bar{x}, \bar{y}) = (1, 0)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l) = (3.57162, 0, 0, 0.07162)$	17.0	2.0
	FBP	$(\bar{x}, \bar{y}) = (1, 0)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}^l) = (9173.95203, 0, 0, 9170.45203)$	17.0	2.0
10	S, M, FBP	$(\bar{x}, \bar{y}) = (4, 0)$ $(\bar{\mu}, \bar{\mu}^l) = (0, 14.01832)$	2.0	24.01832
11	S, M, FBP	$(\bar{x}, \bar{y}) = (0, 0.57735)$ $(\bar{\mu}, \bar{\mu}^l) = (0, 0)$	88.78633	-0.76980
12	M	$(\bar{x}, \bar{y}) = (1.49995, 0, 1.5)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_1^u, \bar{\mu}_2^u) = (0, 0, 5, 0, 0, 0)$	3.25	-0.50010
	S, FBP	$(\bar{x}, \bar{y}_1, \bar{y}_2) = (1.94053, 0, 1.2110)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_1^u, \bar{\mu}_2^u) = (0.34438, 0, 5, 0, 0, 0)$	2.74977	0.54810
13	S	$(\bar{x}, \bar{y}) = (7.19958, 12.79923)$ $(\bar{\mu}, \bar{\mu}^l) = (0, 0)$	2303.72583	0.0
	M, FBP	$(\bar{x}, \bar{y}) = (11.25, 5)$ $(\bar{\mu}, \bar{\mu}^l) = (210.9375, 0)$	2250.0	197.75391
14	S, M, FBP	$(\bar{x}, \bar{y}_1, \bar{y}_2) = (1, 0, 1)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}^l) = (0, 1, 0)$	1	-1

Tabela 4.4: Soluções encontradas por ALGENCAN para os problemas de 5 a 14.

PROB.	REFORM.	SOLUÇÃO ENCONTRADA	$F(x, y)$	$f(x, y)$
15	S	$(\bar{x}_1, \bar{x}_2) = (0.5, 0.8)$ $(\bar{y}_1, \bar{y}_2, \bar{y}_3) = (0, 0.2, 0.8)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3) = (85049.50158, 14175.41693, 70876.08465)$ $(\bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_3^l) = (170106.00316, 0, 0)$	-18.4	1.8
	M	$(\bar{x}_1, \bar{x}_2) = (0.5, 0.8)$ $(\bar{y}_1, \bar{y}_2, \bar{y}_3) = (0, 0.2, 0.8)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3) = (0.52468, 0.58745, 1.93723)$ $(\bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_3^l) = (8.04936, 0, 0)$	-18.4	1.8
	FBP	$(\bar{x}_1, \bar{x}_2) = (0.5, 0.8)$ $(\bar{y}_1, \bar{y}_2, \bar{y}_3) = (0, 0.2, 0.8)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3) = (33189.63025, 5532.10504, 27659.52521)$ $(\bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_3^l) = (66386.26050, 0, 0)$	-18.4	1.8
16	S, M, FBP	$(\bar{x}, \bar{y}) = (3, 5)$	2303.72583	0.0
17	S	$(\bar{x}, \bar{y}) = (9.99999, 9.99999)$ $(\bar{\mu}, \bar{\mu}^l, \bar{\mu}^u) = (0.00007, 0, 0)$	99.99980	0.0
	M, FBP	$(\bar{x}, \bar{y}) = (10, 10)$ $(\bar{\mu}, \bar{\mu}^l, \bar{\mu}^u) = (0, 0, 0)$	100.0	0.0
18	S, M, FBP	$(\bar{x}_1, \bar{x}_2, \bar{y}_1, \bar{y}_2) = (20, 5, 10, 5)$ $(\bar{\mu}_1^l, \bar{\mu}_2^l, \bar{\mu}_1^u, \bar{\mu}_2^u) = (0, 0, 20, 0)$ $(\bar{x}, \bar{y}) = (2, 6)$	225.0	100.0
19	S	$(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}_4, \bar{\mu}^l) = (0, 4248.90337, 0, 0, 8496.80674)$	32.0	-6.0
	M	-	-	-
	FBP	$(\bar{x}, \bar{y}) = (2, 6)$ $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \bar{\mu}_4, \bar{\mu}^l) = (2.62447, 1.81223, 0, 0, 0)$	32.0	-6.0
20	S	$(\bar{x}, \bar{y}, \bar{\mu}) = (0, 0.00001, 343006.56239)$	28.25008	8.99994
	M	$(\bar{x}, \bar{y}, \bar{\mu}) = (0, 0.00005, 65869.4180)$	28.25040	8.99997
	FBP	$(\bar{x}, \bar{y}, \bar{\mu}) = (0, 0.00005, 64648.59593)$	28.25040	8.99997

Tabela 4.5: Soluções encontradas por ALGENCAN para os problemas de 15 a 20.

## Capítulo 5

# Considerações finais

Problemas de programação matemática em dois níveis tornaram-se importantes principalmente porque são capazes de modelar com bastante realismo situações de natureza hierárquica. Iniciando com problemas militares e de economia, rapidamente se espalharam para as mais diversas áreas do conhecimento, tais como ecologia, química e engenharia.

Essa avalanche de problemas a serem resolvidos fez com que uma grande quantidade de métodos fosse desenvolvida para resolvê-los. No presente trabalho, nos propusemos a estudar profundamente o mais intuitivo de todos esses métodos, que consiste em substituir o problema do segundo nível pelas suas condições KKT. O novo problema é um problema de programação não linear, para o qual existem algoritmos clássicos de resolução.

Discutimos intensamente sobre as possíveis abordagens que um problema em dois níveis pode ter. O objetivo era saber em quais casos a reformulação poderia ser aplicada com coerência e se a solução encontrada para determinada abordagem poderia ser comparada com as demais. Nesse aspecto, descobrimos que a abordagem otimista relaxada é a única para o qual o método da reformulação faz sentido. Também definimos conceitos sobre os minimizadores e maximizadores locais e globais de problemas em dois níveis gerais sob as abordagens otimista e pessimista. Estas definições haviam sido apresentadas em trabalhos anteriores de forma incorreta. Reescrevemos as definições e mostramos que elas realmente se aplicam aos minimizadores (ou maximizadores) dos problemas para os quais foram criadas.

Definida a abordagem que garante a coerência do método, nos dedicamos a estudar as possibilidades, vantagens e inconvenientes que aparecem ao resolvermos o problema reformulado. Apresentamos uma classe de problemas em dois níveis cuja reformulação é um problema equivalente ao original, o que implica que a solução global do BLPP pode ser recuperada.

Através de vários exemplos, concluímos que dois grandes inconvenientes ocorrem quando: (i) para um minimizador global do segundo nível não existem multiplicadores de Lagrange; (ii) existem pontos que não são minimizadores globais do segundo nível, mas satisfazem as condições KKT (os minimizadores locais, maximizadores e pontos de sela). Para o caso (i) nada pudemos fazer, pois as características do problema do segundo nível não permitem que utilizemos uma técnica baseada nas condições KKT. Porém, como pudemos observar no Problema 20, tal inconveniente não se apresentou na prática. Para o caso (ii) apresentamos uma estratégia que a cada passo tenta eliminar os pontos que julga não serem minimizadores globais do segundo nível. Quando aplicada a um problema com estas características, a

estratégia conseguiu com sucesso recuperar o minimizador global do BLPP.

Por fim, é um fato conhecido que os problemas reformulados possuem propriedades muito ruins no que se refere à regularidade e à suavidade das funções que os compõem. Grande parte desta afirmação se refere às restrições de complementaridade. Por este motivo, apresentamos outras maneiras de representá-las, usando as funções NCP. Estudamos duas funções em particular: a função mínimo e a versão perturbada da função de Fischer-Burmeister.

Estas duas funções, em conjunto com as restrições de complementaridade usuais (aqui denominadas suaves), formaram os três tipos de reformulação com os quais testamos o método em 20 problemas encontrados na literatura. Também estávamos particularmente interessados em verificar o desempenho do algoritmo para problemas de programação não linear ALGENCAN, motivo pelo qual, este foi o algoritmo usado para resolver todos os problemas reformulados.

Os resultados obtidos foram positivos. O pior desempenho de ALGENCAN ocorreu ao resolver a reformulação que usa a função mínimo: dos 20 problemas encontrou o minimizador global em 16. Por outro lado, com a reformulação que usa a função de Fischer-Burmeister perturbada, ALGENCAN encontrou o minimizador global em 19 dos 20 problemas resolvidos.

Sabemos que as restrições de complementaridade suaves (usuais) possuem a propriedade indesejável de que todo ponto viável do problema reformulado não satisfaz MFCQ. Porém, não sabemos se o mesmo resultado vale para a função de Fischer-Burmeister perturbada. Este pode ser um dos motivos que fez com que esta última formulação tivesse os melhores resultados, visto que, assim como as restrições suaves, a função de Fischer-Burmeister perturbada também é diferenciável em todos os pontos.

É importante ressaltarmos que o objetivo principal dos testes era o de observar o comportamento do método da reformulação em problemas da literatura, e não compará-lo com outros métodos existentes. O fato do método ter encontrado resultados iguais e/ou melhores do que os outros métodos não indica que seu desempenho é superior, mas pode ser um indício de que os problemas-teste não sirvam para o seu real propósito: mostrar a vantagem de aplicar determinados métodos para resolvê-lo.

## 5.1 Linhas para futuros estudos

A área de programação em dois níveis está em pleno desenvolvimento, tanto no aspecto prático, com desenvolvimento de novos métodos, quanto no teórico, com o estudo das condições de otimalidade, por exemplo.

Apesar de ser um método bastante simples, a reformulação mostrou um bom desempenho nos testes realizados, o que pode incentivar um estudo mais abrangente. Um aspecto interessante para ser aprofundado são as funções NCP. Na teoria de problemas de complementaridade não linear, existem muitas outras funções que possuem a mesma propriedade que as funções mínimo e de Fischer-Burmeister discutidas neste trabalho. Estudar o comportamento da reformulação, quando as restrições de complementaridade usuais são substituídas por estas outras restrições pode levar a resultados ainda melhores.

Além do aspecto prático, verificar as propriedades que os problemas reformulados possuem utilizando estas funções NCP é uma outra linha de pesquisa promissora. Não sabemos, por exemplo, se a reformulação que usa a função de Fischer-Burmeister perturbada possui

as mesmas propriedades ruins que a reformulação usual possui. Como já foi mencionado, Dempe [Dem02] usou a reformulação com a função mínimo para apresentar condições de otimalidade para os problemas em dois níveis, aplicando técnicas de otimização não diferenciável.

Uma outra abordagem a ser levada em consideração está relacionada com a Estratégia 2 apresentada no Capítulo 3. Testes mostraram que de fato os pontos que não são minimizadores globais do BLPP são eliminados do conjunto viável do problema reformulado. Porém, os problemas testados são muito simples em sua grande maioria, de forma que nada pode ser concluído. A convergência da estratégia também merece uma atenção especial, e deve ser o foco de estudos futuros.





# Referências Bibliográficas

- [ABMS07] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On Augmented Lagrangian methods with general lower-level constraints. *SIAM Journal on Optimization*, 18:1286–1309, 2007. Disponível em [www.ime.usp.br/~egbirgin/tango](http://www.ime.usp.br/~egbirgin/tango). 4, 88
- [ACC<sup>+</sup>07] R. Andreani, S. L. C. Castro, J. L. Chela, A. Friedlander, and S. A. Santos. An inexact-restoration method for nonlinear bilevel programming problems. *Computational Optimization and Applications*, 2007. DOI 10.1007/s10589-007-9147-4. 2, 19, 34
- [AM99] M. A. Amouzegar and K. Moshirvaziri. Determining optimal pollution control policies: An application of bilevel programming. *European Journal of Operational Research*, 119:100–120, 1999. 2
- [AMS05] R. Andreani, J. M. Martínez, and M. L. Schuverdt. On the relation between Constant Positive Linear Dependence condition and Quasinormality Constraint Qualification. *Journal of Optimization Theory and Applications*, 125:473–485, 2005. 7
- [AS84] E. Aiyoshi and K. Shimizu. A solution method for the static constrained stackelberg problem via penalty method. *IEEE Transactions on Automatic Control*, 12:1111–1114, 1984. 2, 19, 32, 33, 62
- [AW90] G. Anandalingam and D. J. White. A solution method for the linear static stackelberg problem using penalty functions. *IEEE Transactions on Automatic Control*, 35(10):1170–1173, 1990. 2, 33
- [BAB90] O. Ben-Ayed and C. E. Blair. Computational difficulties of bilevel linear programming. *Operations Research*, 38:556–560, 1990. 31
- [Bar84] J. F. Bard. Optimality conditions for the bilevel programming problem. *Naval Research Logistics Quarterly*, 31:13–26, 1984. 67
- [Bar88] J. F. Bard. Convex two-level optimization. *Mathematical Programming*, 40:15–27, 1988. 2, 19, 32, 67, 70, 71
- [Bar98] J. F. Bard. *Practical Bilevel Optimization : Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, 1998. 22, 32, 37

- [BCC<sup>+</sup>92] C. Bischof, A. Carle, G. Corliss, A. Griewank, and P. Hovland. ADIFOR – Generating derivative codes from Fortran programs. *Scientific Programming*, 1:1–29, 1992. 44
- [Ber99] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2 edition, 1999. 8, 10, 45
- [BF82] J. F. Bard and J. E. Falk. An explicit solution to the multi-level programming problem. *Computers and Operations Research*, 9:77–100, 1982. 2, 3, 31, 32, 37
- [BK78] W. F. Bialas and M. H. Karwan. *Multilevel linear programming*, May 1978. Research Report 78-1. 19, 31
- [BK82] W. F. Bialas and M. H. Karwan. On two-level optimization. *IEEE Transactions on Automatic Control*, 27(1):211–214, 1982. 19
- [BK84] W. F. Bialas and M. H. Karwan. Two-level linear programming. *Management Science*, 30:1004–1020, 1984. 71
- [BM73] J. Bracken and T. McGill. Mathematical programs with optimization problems in the constraints. *Operations Research*, 21:37–44, 1973. 1, 19
- [BM90] J. F. Bard and J. T. Moore. A branch and bound algorithm for the bilevel programming problem. *SIAM Journal on Scientific Computing*, 11(2):281–292, 1990. 32, 33
- [BM02] E. G. Birgin and J. M. Martínez. Large-scale active-set box-constrained optimization method with spectral projected gradients. *Computational Optimization and Applications*, 23:101–125, 2002. 88
- [BPS00] J. F. Bard, J. Plummer, and J. C. Sourie. A bilevel programming approach to determining tax credits for biofuel production. *European Journal of Operational Research*, 120:30–46, 2000. 2
- [BS08] E. G. Birgin and F. N. C. Sobral. Minimizing the object dimensions in circle and sphere packing problems. *Computers and Operations Research*, 35:2357–2375, 2008. 2
- [BSS93] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming – Theory and Algorithms*. John Wiley & Sons, 2 edition, 1993. 11, 12
- [Cas04] S. L. C. Castro. *Algoritmo de restauração inexata aplicado à resolução de problemas de programação matemática em dois níveis*. PhD thesis, Instituto de Matemática, Estatística e Computação Científica – Universidade Estadual de Campinas, 2004. 34
- [CM96] C. Chen and O. L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Computational Optimization and Applications*, 5:97–138, 1996. 47, 91

- [CMS05] B. Colson, P. Marcotte, and G. Savard. A trust-region method for nonlinear bilevel programming: Algorithm and computational experience. *Computational Optimization and Applications*, 30:211–227, 2005. 2, 19, 34, 62, 67, 71, 75, 77, 78
- [CMS07] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153:235–256, 2007. 23, 24, 25, 26, 27, 28, 29
- [CN77] W. Candler and R. Norton. *Multilevel programming*, 1977. Technical Report 20. 2, 19
- [Col02] B. Colson. *BIPA (Bilevel programming with approximation methods): Software guide and test problems*, October 2002. Technical Report CRT-2002-38. 62, 67, 70, 73, 74, 75, 77, 78, 82, 83
- [CW90] P. A. Clark and A. W. Westerberg. Bilevel programming for steady-state chemical process design - I. Fundamentals and algorithms. *Computers and Chemical Engineering*, 14:87–97, 1990. 72, 73
- [Dem02] S. Dempe. *Foundations of Bilevel Programming*. Kluwer Academic Publishers, 2002. 22, 24, 25, 26, 28, 30, 37, 45, 47, 49, 97
- [Dem03] S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52(3):333–359, 2003. 2
- [DK02] S. Dempe and V. Kalashnikov. *Discrete Bilevel Programming: Application to a Gas Shipper's Problem*, 2002. Technical Report 2002-02. 2
- [FGK90] R. Fourer, D. M. Gay, and B. W. Kernighan. A modeling language for mathematical programming. *Management Science*, 36:519–554, 1990. 44, 89
- [Fis92] A. Fischer. A special newton-type optimization method. *Optimization*, 24:269–284, 1992. 16
- [FP97] M. C. Ferris and J. S. Pang. Engineering and economic applications of complementarity problems. *SIAM Review*, 39(4):669–713, 1997. 14
- [FZ98] C. A. Floudas and S. Zlobec. Optimality and duality in parametric convex lexicographic programming. In A. Migdalas, P. M. Pardalos, and P. Varbrand, editors, *Multilevel Optimization: Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998. 81
- [GF01] Z. H. Gümüş and C. A. Floudas. Global optimization of nonlinear bilevel programming problems. *Journal of Global Optimization*, 20:1–31, 2001. 80, 81
- [Isa92] G. Isac. *Complementarity problems*, volume 1528 of *Lecture Notes in Mathematics*. Springer-Verly, 1992. 14
- [Kan96] C. Kanzow. Some noninterior continuation methods for linear complementarity problems. *SIAM Journal on Matrix Analysis and Applications*, 17(4):851–868, 1996. 17

- [LLS00] D. Lavigne, R. Loulou, and G. Savard. Pure competition, regulated and stackelberg equilibria: Application to the energy system of québec. *European Journal of Operational Research*, 125:1–17, 2000. 2
- [LPR96] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints*. Cambridge University Press, 1996. 19
- [Mar98] J. M. Martínez. Two-phase model algorithm with global convergence for nonlinear programming. *Journal of Optimization Theory and Applications*, 96(2):397–436, 1998. 34
- [MF67] O. L. Mangasarian and S. Fromovitz. The Fritz-John necessary optimality conditions in presence of equality and inequality constraints. *Journal of Mathematical Analysis and Applications*, 17:37–47, 1967. 7, 45
- [MP00] J. M. Martínez and E. A. Pilotta. Inexact restoration algorithms for constrained optimization. *Journal of Optimization Theory and Applications*, 104(1):135–163, 2000. 34
- [NW99] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, 1999. 14
- [ODJ95] H. Önal, D. H. Darmawan, and S. H. Johnson. A multilevel analysis of agricultural credit distribution in East Java, Indonesia. *Computers and Operations Research*, 22:227–236, 1995. 2
- [QW00] L. Qi and Z. Wei. On the constant positive linear dependence condition and its application to SQP methods. *SIAM Journal on Optimization*, 10:963–981, 2000. 7
- [SA81] K. Shimizu and E. Aiyoshi. A new computational method for Stackelberg and min-max problems by use of a penalty method. *IEEE Transactions on Automatic Control*, 26:460–466, 1981. 19, 32, 83, 84
- [Sav89] G. Savard. *Contribution a la programmation mathématique à deux niveaux*. PhD thesis, Ecole Polytechnique de Montréal, Université de Montréal, 1989. 82
- [SC98] K. H. Sahin and A. R. Ciric. A dual temperature simulated annealing approach for solving bilevel programming problems. *Computers and Chemical Engineering*, 23:11–25, 1998. 21, 82, 83
- [Sch06] Maria Laura Schuverdt. *Métodos de Lagrangiano Aumentado com convergência utilizando a condição de dependência linear positiva constante*. PhD thesis, Instituto de Matemática, Estatística e Computação Científica - Universidade Estadual de Campinas, 2006. 6
- [SG94] G. Savard and J. Gauvin. The steepest descent direction for the nonlinear bilevel programming problem. *Operations Research Letters*, 15:265–272, 1994. 2, 19, 33

- 
- [SIB97] K. Shimizu, Y. Ishizuka, and J. F. Bard. *Nondifferentiable and Two-level Mathematical Programming*. Kluwer Academic Publishers, 1997. 80
- [Sla50] M. Slater. Lagrange multipliers revisited: a contribution to non-linear programming. Cowles Comission Discussion paper, 1950. 7
- [Sta52] H. Stackelberg. *The Theory of Market Economy*. Oxford University Press, Oxford, 1952. 1, 19
- [TS04] M. Tawarmalani and N. V. Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical Programming*, 99:563–591, 2004. 61
- [VC94] L. N. Vicente and P. H. Calamai. Bilevel and multilevel programming: A bibliographic review. *Journal of Global Optimization*, 5(3):291–306, 1994. 2, 22
- [Wat91] D. S. Watkins. *Fundamentals of Matrix Computations*. John Wiley & Sons, 1991. 14
- [ZXL04] D. L. Zhu, Q. Xu, and Z. Lin. A homotopy method for solving bilevel programming problem. *Nonlinear Analysis*, 57:917–928, 2004. 85, 86, 87, 88