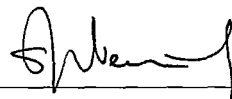


# O PROBLEMA DE PROGRAMAÇÃO MATEMÁTICA DE DOIS NÍVEIS: CONDIÇÕES DE OTIMIDADE E PROPOSTA NUMÉRICA

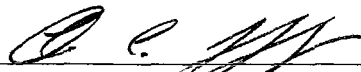
José Ramón Arica Chávez

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

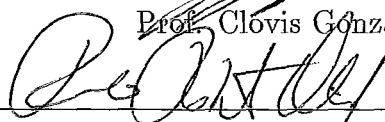
Aprovada por:



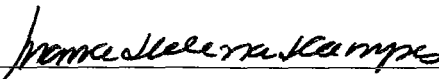
\_\_\_\_\_  
Prof. Susana Scheimberg, D. Sc.  
(Presidente)



\_\_\_\_\_  
Prof. Clovis Gonzaga, D. Sc.



\_\_\_\_\_  
Prof. Paulo Roberto Oliveira, Dr. Ing.



\_\_\_\_\_  
Prof. Ma. Helena Jardim Campos, D. Sc.



\_\_\_\_\_  
Prof. Alfredo Iusem, Ph. D.

RIO DE JANEIRO, RJ - BRASIL  
FEVEREIRO DE 1995

ARICA CHÁVEZ, JOSÉ RAMÓN

O Problema de Programação Matemática de Dois Níveis: Condições de Otimidade e Proposta Numérica [Rio de Janeiro] 1995. III, 116 p. 29.7 cm (COPPE/UFRJ, D. Sc. Engenharia de Sistemas e Computação, 1995)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1. Dois Níveis, Não Diferenciabilidade, Condições de Otimidade.  
I. COPPE/UFRJ II. Título (série).

## AGRADECIMENTOS

Embora seja inecessário sublinhar que a responsabilidade deste trabalho é unicamente do autor, quero insistir em dizer que teria sido muito difícil sua realização sem a estimulante discussão que a Profa. Susana Scheimberg soube conduzir. As muitas horas de paciente escuta que ela dedicou e suas interessantes observações, permitiram-me desenvolver a maior parte das idéias aqui apresentadas. Quero agradecer seu grande interesse por este trabalho.

Agradeço ao CNPq, à Universidade Estadual do Norte Fluminense e à *Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima-Peru*, pelo apóio brindado.

Por último, não posso deixar de mencionar que teria sido impossível terminar este trabalho, sem o constante alento da minha família. Os inumeráveis fins de semana, feriados, aniversários ... Enfim! O grande número de dias de lazer dos que os furtei foram imprescindíveis. Agradeço profundamente a enorme compreensão da minha companheira, Gudélia, e de minhas filhas, Catalina, Natália e Victória.

DEDICO ESTE TRABALHO À MEMÓRIA DE

**EUGEN BLUM R.,**

GRANDE MESTRE, GRANDE AMIGO.  
CATIVANTE COMO MATEMÁTICO,  
EXCEPCIONAL COMO PESSOA.  
NASCIDO NA SUIÇA, EM 1937,  
FALECIDO NO PERU, EM 1993.

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

## O PROBLEMA DE PROGRAMAÇÃO MATEMÁTICA DE DOIS NÍVEIS: CONDIÇÕES DE OTIMIDADE E PROPOSTA NUMÉRICA

José Ramón Arica Chávez  
FEVEREIRO, 1995

Orientadora: Profa. Susana Scheimberg  
Programa: Engenharia de Sistemas e Computação

O problema de Programação Matemática de Dois níveis (PPDN) é um modelo de jogo líder-seguidor, no que as jogadas são seqüenciais e não se permite cooperação. Pode-se formular como uma seqüência de dois problemas de otimização, onde a região de restrições do primeiro está implicitamente determinada pelas soluções do segundo.

O PPDN é, em geral, um problema de programação não diferenciável e não convexo. Vários algoritmos têm sido estabelecidos para casos particulares. No entanto, não existem condições de otimidade para o caso geral.

Neste trabalho, baseados na análise da função marginal do seguidor, no contexto da não diferenciabilidade, propõe-se uma condição necessária de otimidade (do tipo Karush-Kuhn-Tucker). Assume-se a condição de regularidade de *calma* para o PPDN, assim como, condições de segunda ordem para o problema do seguidor.

Por outro lado, formula-se um algoritmo de descida viável para o tratamento numérico do PPDN. Este algoritmo está baseado numa aproximação aqui proposta para o gradiente generalizado da função marginal do seguidor e usa os códigos Fortran NLPQL1 e BTNC, para minimização de programas não lineares (pelo método seqüencial quadrático) e minimização de programas não diferenciáveis sem restrições (misturando o algoritmo de feixes com técnicas de região de confiança), respectivamente. Assume-se convexidade para o problema do seguidor.

Apresentam-se finalmente, antes das conclusões, resultados numéricos de alguns testes feitos usando o algoritmo proposto.

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

## **THE BILEVEL PROGRAMMING PROBLEM: OPTIMALITY CONDITIONS AND A NUMERICAL APPROACH**

**José Ramón Arica Chávez**  
**FEBRUARY, 1995**

Thesis supervisor: Prof. Susana Scheimberg  
Department: Engenharia de Sistemas e Computação

The Bilevel Programming Problem (BLPP) is a model of a leader-follower game in which play is sequential and cooperation is not permitted. It can be formulated as a sequence of two optimization problems, where the constraint region of the first is implicitly determined by the solution to the second.

The BLPP is, in general, a nonconvex and nondifferentiable programming problem and several algorithms have been stated for particular cases. However, there do not exist optimality conditions for the general BLPP.

In this work, based on the analysis of marginal function of the follower, in the context of nondifferentiability, it is proposed a necessary optimality condition ( of Karush-Kuhn-Tucker type), under regularity condition of *calm* for the BLPP and second order conditions for the follower problem.

On the other hand, it is proposed a feasible descent algorithm to numerically treat the BLPP. This algorithm is based on an approximation formulated here for the generalized gradient of the marginal function of the follower and uses the Fortran codes NLPQL1 and BTNC, to minimize nonlinear programs (by the sequential quadratic method) and unconstrained nondifferentiable programs (combining the bundle concept with trust region methods), respectively. Convexity is assumed for the follower problem.

Finally, before the conclusions, results of some numerical experiments using the propose algorithm are shown.

# Índice

	página
<b>Introdução</b>	1
<b>Capítulo I</b>	
<b>Formulação do Problema de Programação</b>	
<b>Matemática de Dois Níveis</b>	3
1.1 Preliminares	3
1.2 Gradiente Generalizado: Conceitos Básicos	7
1.3 Gradiente Generalizado da Função Marginal do seguidor	14
1.4 Resultados Adicionais sobre o Gradiente Generalizado da Função Marginal do seguidor	25
<b>Capítulo II</b>	
<b>Condições Necessárias de Otimidade</b>	
<b>para o Problema de Programação</b>	
<b>Matemática de Dois Níveis</b>	32
2.1 Por Que Não Condições de Qualificação tipo Mangasarian-Fromovitz ?	33
2.2 O conceito de Problema Calmo	40
2.3 Condições de Otimidade Necessárias para (PPDN)	43
2.4 Condições de Otimidade para Alguns Casos Particulares de (PPDN)	50
2.4.1 Conjunto de Reação Racional Unitária	50
2.4.2 O Problema do Seguidor Convexo	61
<b>Capítulo III</b>	
<b>Uma solução Numérica do Problema (PPDN)</b>	65
3.1 Hipóteses para uma Proposta Numérica	66
3.2 Uma Proposta Numérica	74
3.3 Um Algoritmo para o PPDN	99
3.3.1 O Método de Feixe	100
3.3.2 O Método BT	101
3.3.3 O Método BT para o Caso Não Convexo: BTNC	104
3.3.4 BLPP: Um Algoritmo para Solucionar o PPDN	105
<b>Conclusões</b>	110
<b>Referências</b>	111

## Introdução

O Problema de Programação Matemática de Dois Níveis (PPDN) pode ser formulado, informalmente, da seguinte forma

$$\begin{aligned} \min_{(x,y) \in X \times \mathbb{R}^{n_2}} \quad & F(x,y) \\ \text{s.a.} \quad & G(x,y) \leq 0, \end{aligned}$$

onde,  $X \subset \mathbb{R}^{n_1}$  e para  $x \in X$ , fixo,  $y$  é uma solução do programa

$$\begin{aligned} \min_y \quad & f(x,y) \\ \text{s.a.} \quad & g(x,y) \leq 0. \end{aligned}$$

As funções envolvidas supõem-se adequadamente diferenciáveis. O primeiro nível do problema (que envolve as funções  $F$  e  $G$ ) é denominado *problema do líder* e o segundo nível (que envolve as funções  $f$  e  $g$ ), *problema do seguidor*. A variável  $x$  é chamada variável do líder

O PPDN surge naturalmente em diversas áreas relacionadas à Otimização: transportes, problemas de economia, otimização de formas e outras (ver, por exemplo, Anandalingam-Friesz [AF1], Kolstad [KO1], Kolstad-Lasdon [KL1] e Outrata [OT2], [OT4]). Diferentes abordagens e vários algoritmos têm sido propostos para a análise e tratamento numérico do problema. Têm-se, assim, entre outros trabalhos interessantes o de Bard-Falks [BF1], que estuda o caso linear usando uma formulação que permite aproveitar conceitos de separabilidade e técnicas (heurísticas) do tipo *branch and bound*, para determinar soluções globais do PPDN; de Silva-McCormick [dSM1], que usa resultados de sensibilidade não linear para determinar, sob certas hipóteses, derivadas direcionais da função do líder em termos da derivada da resposta do seguidor; Ishizuka-Aiyoshi [IA1] e Loridan-Morgan [LM1], que embora diferentemente, usam dupla penalidade (penalizam os problemas do seguidor e do líder); Kolstad-Lasdon [KL1], que, sob certas hipóteses, usam informação da derivada da resposta do seguidor aplicada ao caso de grande porte; e, Outrata [OT2], [OT4] e Dempe [DE1], que analisam o problema via não diferenciabilidade, usando uma formulação equivalente em termos da função marginal (ou função valor extremo) do problema de seguidor.

Só ultimamente, o estudo de condições de otimidade vem sendo tratado. Alguns trabalhos nesse sentido são os de Chen-Florian [CF1], Dempe [DE2] e Outrata [OT1], [OT3], [OT4].

Em geral, o PPDN é um problema complexo. Resulta ser não diferenciável, não convexo e tem sido provado que, mesmo no caso linear, é NP-hard (ver, por exemplo, Hansen-Jaumard-Savard [HJS1]). A não diferenciabilidade e não convexidade fazem muito difícil estabelecer condições suficientes de otimidade, assim como estudar soluções globais. De fato, neste trabalho, propõem-se condições necessárias para mínimos locais.



O trabalho aqui apresentado tem dois objetivos fundamentais:

- Estudar condições necessárias de otimidade para o PPDN.
- Fazer uma proposta de um algoritmo de descida para a solução desse problema.

No Capítulo I, apresenta-se formalmente o problema, reformulando-o em termos da função marginal do seguidor, que é analisada no contexto da análise não diferenciável. O resultado principal deste capítulo, a Proposição 1.3.7, estabelece uma estimativa do gradiente generalizado da função marginal do seguidor.

O Capítulo II, trata das condições de otimidade. Estuda-se a existência de multiplicadores de Karush-Kuhn-Tucker, sob a condição de qualificação de *calma* (no sentido de Clarke [CL1]). Mostra-se, também, porque não são adequadas condições de qualificação tipo Mangasarian-Fromovitz [MAF1]. O resultado principal deste capítulo, o Teorema 2.3.6, estabelece condições necessárias de otimidade para o PPDN. Um exemplo simples ilustra este resultado. Estudam-se, também, alguns casos particulares do PPDN, sendo que os respectivos resultados são compatíveis com resultados de trabalhos de abordagem análoga (Dempe [DE2] e Outrata [OT1], [OT2], [OT3]), se bem obtidos a partir de um contexto mais geral. Formula-se, por último, uma caracterização do cone de direções viáveis de descida para o PPDN.

No Capítulo III, apresenta-se uma abordagem numérica, para o caso em que o problema do seguidor é convexo. Sendo o PPDN um programa não diferenciável e não convexo, com restrições, e, desde que, os códigos existentes tratam problemas não diferenciáveis, mas sem restrições, reformula-se neste capítulo o PPDN via um problema *min-max*. Neste contexto, considera-se o problema como membro de uma família de programas parametrizados, que são resolvidos numericamente usando uma extensão do Algoritmo dos Centros de Huard [HU1], desenvolvida por Kiwiel [KI1] (aqui, adequadamente modificada). Prova-se que, quando o parâmetro tende a zero, as soluções perturbadas geram pontos viáveis que fornecem uma subsequência convergente a um *ponto estacionário* do PPDN. Finalmente, mostram-se os resultados numéricos obtidos da aplicação do algoritmo de descida proposto a alguns exemplos existentes na literatura. Usa-se, para o tratamento numérico, o código BTNC (Outrata-Schramm-Zowe [OSCHRZ1], Schramm-Zowe [SCHRZ1]), gentilmente fornecido pelo Prof. Jiri Outrata. Este código é uma rotina FORTRAN para o método BT (B de *bundle* e T de *trust*) que é uma modificação do Método de Feixe (ver, por exemplo, Hiriart-Urruty-Lemaréchal [H-ULE1]) para funções não diferenciáveis e não convexas sem restrições.

Encerra-se este trabalho com a apresentação de algumas conclusões.

## Capítulo 1

### Formulação do Problema de Programação Matemática de Dois Níveis

Neste capítulo, na seção 1.1, apresenta-se formalmente o Problema de Programação Matemática de Dois Níveis (PPDN). Introduce-se uma reformulação do PPDN em termos da função marginal (ou função valor extremo) do seguidor que permite considerar o problema como se fosse um programa matemático de *um nível só*. Esta reformulação, com estrutura mais conhecida, tem, naturalmente, um preço a ser pago, pois a função marginal resulta, em geral, sendo não diferenciável. Estabelecendo-se, assim, a necessidade de desenvolver conceitos relacionados ao gradiente generalizado de Clarke para formular algoritmos de descida para a solução do problema. Isto é tratado na seção 1.2. Na seção 1.3, apresenta-se o resultado principal deste capítulo, a Proposição 1.3.7, onde se estabelece uma estimativa do gradiente generalizado da função marginal do segundo nível. Por último, na seção 1.4, sob hipóteses mais exigentes (convexidade do problema do seguidor e a condição de qualificação de Mangasarian-Fromovitz estrita para este) obtem-se resultados adicionais sobre a diferenciabilidade da função marginal do seguidor.

#### 1.1 – Preliminares

O PPDN formula-se da seguinte forma:

$$\begin{aligned}
(PPDN) : \quad & \min_{(x,y) \in X \times \mathbb{R}^{n_2}} F(x,y) \\
& \text{s.a.} \quad G(x,y) \leq 0 \\
& y \in \arg \min_{y \in \mathbb{R}^{n_2}} \{f(x,y)/g(x,y) \leq 0\},
\end{aligned} \tag{1.1.1}$$

onde  $X \subset D \subset \mathbb{R}^{n_1}$ .  $D$  é um conjunto aberto e  $X$  fechado,  $F, f : D \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ ,  $G : D \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_1}$ ,  $g : D \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_2}$ , tal que  $F$  e  $G$  são continuamente diferenciáveis, entanto que,  $f$  e  $g$  duas vezes continuamente diferenciáveis em  $D \times \mathbb{R}^{n_2}$ .

A função  $F$  é chamada *função objetivo do problema de primeiro nível* (ou problema do líder) e  $f$  e  $g$ , respectivamente, a *função objetivo e função das restrições do problema de segundo nível* (ou problema do seguidor).

Neste trabalho serão usadas as seguintes definições e notações:

**Região de restrições do (PPDN):**

$$S := \{(x,y) \in X \times \mathbb{R}^{n_2} / G(x,y) \leq 0, g(x,y) \leq 0\};$$

**Região viável do problema do seguidor:**

Dado  $x \in D$ ,

$$Y(x) := \{y \in \mathbb{R}^{n_2} / g(x,y) \leq 0\};$$

**Conjunto de reação racional do problema do seguidor:**

Dado  $x \in D$ ,

$$O(x) := \arg \min \{f(x, z) / y \in Y(x)\},$$

**Região viável do (PPDN):**

$$\tilde{S} := \{(x, y) \in S / y \in O(x)\}.$$

Notar que se pode escrever (PPDN), de maneira equivalente, como

$$\min_{(x, y) \in \tilde{S}} F(x, y).$$

O problema do seguidor pode ser visto como um problema de programação paramétrica, parametrizado pela variável do líder,  $x$ , i.e., dado  $x \in X$ , pode-se escrever o problema do seguidor como

$$\begin{aligned} (P_x) : \quad & \min_{y \in \mathbb{R}^{n_2}} \quad f(x, y) \\ & \text{s.a.} \quad g(x, y) \leq 0 \end{aligned} \tag{1.1.2}$$

Considerando  $(P_x)$ , pode-se definir a *função marginal do problema do seguidor*  $v : D \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$  como:

$$v(x) := \inf_{y \in Y(x)} f(x, y), \tag{1.1.3}$$

sendo  $v(x) := +\infty$ , se  $Y(x) = \emptyset$ .

Em geral, a função marginal é não diferenciável e, mesmo sendo  $(P_x)$  convexo, pode ser não convexa (poderia inclusive não tomar valores finitos). Para garantir um

comportamento adequado de  $v$ , impõe-se a seguinte *condição de limitação inferior* a (PPDN):

**Hipótese 1.1.1** Para cada  $\bar{x} \in D$  e  $\alpha \in \mathbb{R}$ , existe  $\epsilon > 0$  tal que o conjunto

$$\{(x, y) \in D \times \mathbb{R}^{n_2} / \|x - \bar{x}\| \leq \epsilon, f(x, y) \leq \alpha, g_i(x, y) \leq \epsilon, \forall i\}$$

é limitado (onde  $\|\cdot\|$  denota a norma euclidiana).

Esta hipótese assegura que (ver Rockafellar[R03], [R04]):

- $v(x) > -\infty, \forall x \in D$ ;
- $v$  é semicontínua inferior em  $D$  (embora possa tomar o valor  $+\infty$ );
- $O(x)$  é não vazio e compacto, para  $x \in D$  tal que  $v(x) < +\infty$ .

Além disso, sob a Hipótese 1.1.1, o problema (PPDN) é equivalente ao problema

$$\begin{aligned}
 (P1) : \quad & \min_{(x,y) \in X \times \mathbb{R}^{n_2}} F(x, y) \\
 & \text{s.a} \quad G(x, y) \leq 0 \\
 & f(x, y) - v(x) \leq 0 \\
 & g(x, y) \leq 0.
 \end{aligned} \tag{1.1.4}$$

Como se pode observar, o problema (PPDN) formulado como (P1), tem a forma usual de um problema de programação matemática, mas, é claro, que nesta formulação se tem introduzido não diferenciabilidade. Temos, em consequência, que (PPDN) é, em geral, um problema não diferenciável. Resultados sobre a continuidade da função marginal podem ser encontrados em Hogan [H01] e Gauvin-Dubéau [GD1]. Na próxima seção discutimos a diferenciabilidade da função  $v$ .

**Observação 1.1.2** Alguns problemas relacionados à existência de soluções do problema (PPDN) podem surgir (ver, por exemplo Loridan-Morgan [LM1]). Assumir-se-á, neste trabalho, que (PPDN) tem solução.

## 1.2 – Gradiente Generalizado: Conceitos Básicos

Como se tem visto na seção anterior, a presença da função marginal  $v$  no problema (P1), equivalente ao problema (PPDN), o torna, em geral, não diferenciável e, mesmo no caso em que as funções  $F$ ,  $f$  e  $g$  sejam convexas, não convexo. Na medida em que qualquer método de descida para (P1), envolve necessariamente algum tipo de diferenciabilidade das funções que o compõem, esta seção está dedicada a discussão desta propriedade para a função marginal  $v$ . Apresentam-se, assim, a seguir, alguns conceitos da Análise Não Diferenciável (ou não suave) que ajudarão nesta discussão (ver Clarke [CL1]).

Seja a função  $h : Z \subset \mathbb{R}^n \rightarrow \mathbb{R}$ . Diz-se que  $h$  é *Lipschitz* de módulo  $k > 0$  em  $Z$ , se

$$|h(x_1) - h(x_2)| \leq k \|x_1 - x_2\|, \quad \forall x_1, x_2 \in Z.$$

Analogamente, diz-se que  $h$  é *Lipschitz próxima* de  $x_0$  (de módulo  $k$ ), se existe  $\epsilon > 0$  tal que  $h$  é Lipschitz de módulo  $k$  na  $\epsilon$ -bola aberta centrada em  $x_0$ . Finalmente, diz-se que  $h$  é *localmente Lipschitz* em  $Z$ , se  $h$  é Lipschitz próxima de cada ponto  $x \in Z$ .

Seja  $h$  Lipschitz próxima de  $x_0$  de módulo  $k$ . A *derivada direcional de Clarke* (ou derivada direcional generalizada) de  $h$  em  $x_0$  na direção  $p \in \mathbb{R}^n$ , define-se como

$$h^0(x_0; p) := \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \frac{h(x + tp) - h(x)}{t}.$$

É fácil ver que  $|h^0(x_0; p)| \leq k\|p\|$ ,  $\forall p \in \mathbb{R}^n$ .

**Observação 1.2.1** Desde que as hipóteses feitas sobre (PPDN), em (1.1.1), incluem diferenciabilidade contínua das funções envolvidas, estas são localmente Lipschitz em  $D \times \mathbb{R}^n$  (corolário da Proposição 2.2.1, Clarke [CL1]).

O *gradiente generalizado* de  $h$  em  $x_0$  está definido por

$$\partial h(x_0) := \{\xi \in \mathbb{R}^n / h^0(x_0; p) \geq \langle \xi, p \rangle, \quad \forall p \in \mathbb{R}^n\},$$

onde  $\langle \cdot, \cdot \rangle$  denota o produto interno usual em  $\mathbb{R}^n$ . Prova-se que

$$\partial h(x_0) = \text{Conv}\{\xi \in \mathbb{R}^n / x_i \rightarrow x_0, x_i \notin \Omega, \xi = \lim_{x_i \rightarrow x_0} \nabla h(x_i)\},$$

onde  $\Omega$  é o conjunto onde  $h$  é não diferenciável (que tem medida de Lebesgue nula) e  $\text{Conv}$  denota a envoltória convexa.

No caso em que  $h$  é uma função convexa,  $\partial h(x_0)$  coincide com o conjunto subdiferencial de  $h$  em  $x_0$ .

Se  $h$  é Lipschitz próxima de  $x_0$  de módulo  $k$ , tem-se que  $\partial h(x_0)$  é não vazio, convexo e compacto, com  $\|\xi\| \leq k$ , para todo  $\xi \in \partial h(x_0)$ . Cumpre-se, também, que  $h^0(x_0; p)$  é a função suporte de  $\partial h(x_0)$ , i.e.,

$$h^0(x_0; p) = \max_{\xi \in \partial h(x_0)} \langle \xi, p \rangle.$$

Se  $h$  é uma função convexa,  $h^0(x_0; p)$  coincide com a derivada direcional clássica:

$$h'(x_0; p) := \lim_{t \downarrow 0} \frac{h(x_0 + tp) - h(x_0)}{t}.$$

O seguinte exemplo mostra algumas diferenças entre  $h^0(x_0; p)$  e  $h'(x_0; p)$ .

**Exemplo 1.2.2** (Scheimberg [SCH 1])

(i) Existência de  $h^0$  mas não de  $h'$ :

Seja

$$h(x) = \begin{cases} |x| \operatorname{sen}(\ln |x|), & \text{se } x \neq 0, \\ 0, & \text{se } x = 0. \end{cases}$$

A função  $h$  é Lipschitz próxima de  $x_0 = 0$ , com  $k = \sqrt{2}$ ,  $\partial h(0) = [-\sqrt{2}, \sqrt{2}]$  e  $h^0(0; -1) = \sqrt{2}$ . Mas não existe o limite

$$\lim_{t \downarrow 0} \frac{f(0-t) - f(0)}{t} = \lim_{t \downarrow 0} \operatorname{sen}(\ln t),$$

i.e., não existe  $h'(0; -1)$ .

(ii) Existência de  $h^0$  e  $h'$ , mas  $h^0 < h'$ :

Seja



$$h(x) = 1 - |x|, \quad \forall x \in \mathbb{R}.$$

Então,  $h'(0; 1) = -1 < 1 = h^0(0; 1)$ .

A função  $h$  se diz *regular no sentido de Clarke* em  $x_0$ , se  $h'(x_0; p)$  existe e é igual a  $h^0(x_0; p)$ , para todo  $p \in \mathbb{R}^n$ . Diz-se que a função  $h$  é *estritamente diferenciável* em  $x_0$ , com derivada estrita  $\nabla h(x_0) \in \mathbb{R}^n$ , se para cada  $p \in \mathbb{R}^n$ , tem-se

$$\lim_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \frac{h(x + tp) - h(x)}{t} = \langle \nabla h(x_0), p \rangle,$$

onde a convergência é uniforme em  $p$ , para  $p$  em conjuntos compactos. Segue-se que se  $h$  é continuamente diferenciável em  $x_0$ , então é estritamente diferenciável. Cumpre-se, também, que  $h$  é estritamente diferenciável em  $x_0$  e  $\nabla h(x_0) = \xi$  se, e somente se,  $h$  é Lipschitz próxima de  $x_0$  e  $\partial h(x_0) = \{\xi\}$ .

Dado um conjunto  $C \subset \mathbb{R}^n$ , a função distância de  $x$  a  $C$  se define por

$$d_C(x) := \inf_{c \in C} \|x - c\|, \quad x \in \mathbb{R}^n.$$

A função distância é Lipschitz de módulo 1 em  $\mathbb{R}^n$ . Cumpre-se que  $\partial d_C(x) = \{\nabla d_C(x)\} = \{0\}$ , para  $x \in \text{int}(C)$  (onde  $\text{int}(C)$  denota o interior de  $C$ ). Quando  $C$  é um conjunto convexo,  $d_C(x)$  é uma função convexa.

Dado  $x \in C \subset \mathbb{R}^n$ , o *cone tangente de Clarke* a  $C$  no ponto  $x$ ,  $T_C(x)$  define-se como

$$T_C(x) := \{p \in \mathbb{R}^n / d_C^0(x; p) = 0\}.$$

É claro que zero pertence a  $T_C(x)$ . Prova-se que  $T_C(x)$  é um cone convexo e fechado. Quando  $x \in \text{int}(C)$ , satisfaz-se  $T_C(x) = \mathbb{R}^n$ .

Desde que

$$\partial d_C(x) = \{\xi \in \mathbb{R}^n / d_C^0(x; p) \geq \langle \xi, p \rangle \quad \forall p \in \mathbb{R}^n\},$$

obtéem-se que

$$p \in T_C(x) \iff \langle \xi, p \rangle \leq 0, \quad \forall \xi \in \partial d_C(x). \quad (1.2.1)$$

Outras caracterizações do cone tangente de Clarke podem ser encontradas em Clarke [CL1] e Aubin-Ekeland [AE1].

O *cone normal* a  $C$  no ponto  $x$ ,  $N_C(x)$ , define-se como o cone polar de  $T_C(x)$ , i.e.,

$$N_C(x) := \{\xi \in \mathbb{R}^n / \langle \xi, p \rangle \leq 0, \quad \forall p \in T_C(x)\}. \quad (1.2.2)$$

Prova-se que  $N_C(x)$  é o fecho do cone gerado por  $\partial d_C(x)$ , i.e.,

$$N_C(x) = \overline{\cup_{\lambda \geq 0} \lambda \partial d_C(x)},$$

(onde a barra superior significa fecho).

Observe-se que  $0 \in N_C(x)$  e se  $x \in \text{int}(C)$ , então  $N_C(x) = \{0\}$ .

Por outro lado, diz-se que um vetor  $n$  é *perpendicular* a  $C$  em  $x \in \overline{C}$ , e se denota por  $n \perp C$ , se existe  $x'$ , tal que  $n = x' - x$  e  $\arg \min_{c \in C} \|x - c\| = \{x\}$  (notar que  $d_C(x') = \|n\|$ ).

Verifica-se que

$$N_C(x) = \overline{\text{Conv}}\left\{\lambda \lim_{\substack{n_i \\ \|n_i\|}} \frac{n_i}{\|n_i\|} / \lambda \geq 0, n_i \perp C \text{ em } x_i, x_i \rightarrow x, n_i \rightarrow 0\right\}.$$

Frequentemente, quando se faz referência a  $N_C(x)$ , assume-se que  $C$  é um subconjunto próprio de  $\mathbb{R}^n$  e que  $x$  é um ponto da fronteira de  $C$ . Neste caso, sabe-se que  $\partial d_C(x) \neq \{0\}$ , conseqüentemente  $N_C(x) \neq \{0\}$  (Teorema 2.5.6, Corolário 1 e Corolário 2, Clarke [CL1]).

No caso em que  $C$  é um conjunto convexo,  $T_C(x)$  e  $N_C(x)$  coincidem com os cones tangente e normal no sentido de análise convexa, respectivamente. Na figura 1.1 mostram-se alguns cones tangentes no sentido de Clarke e os respectivos cones normais.

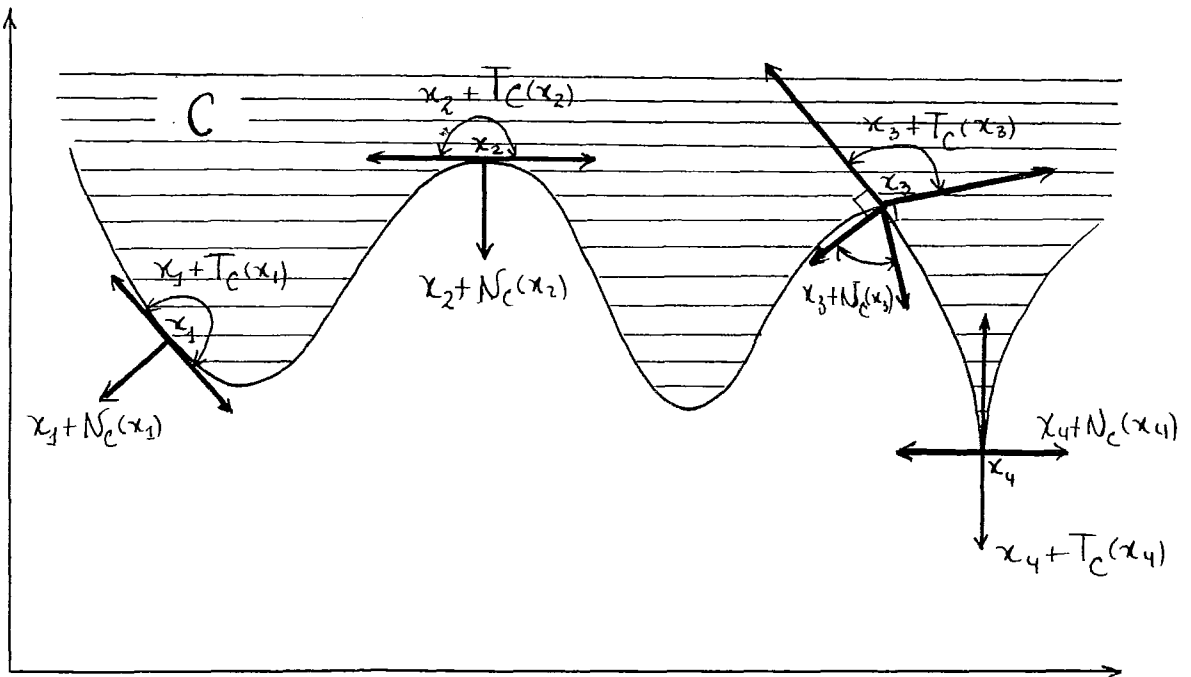


Figura 1.1 - Cones tangentes e Normais no sentido de Clarke

Notar que se  $x$  é um ponto da fronteira de  $C$ , então

$$p \in \text{int}(T_C(x)) \implies \langle \xi, p \rangle < 0, \quad \forall \xi \in \partial d_C(x) \setminus \{0\} \quad (1.2.3)$$

De fato, dado  $p \in \text{int}(T_C(x))$ , existe uma vizinhança  $U(p)$  de  $p$ , tal que  $U(p) \subset T_C(x)$ , de onde, usando (1.2.1), tem-se

$$\langle \xi, w \rangle \leq 0, \quad \forall \xi \in \partial d_C(x), \quad \forall w \in U(p). \quad (1.2.4)$$

Supor que exista  $\xi \in \partial d_C(x) \setminus \{0\}$  tal que  $\langle \xi, p \rangle = 0$ . Considerar  $\epsilon > 0$ , tal que  $p + \epsilon \bar{B} \subset U(p)$ , onde  $B := \{y \in \mathbb{R}^n / \|y\| < 1\}$ .

Tomar  $\tilde{w} = p + \epsilon q$ , com  $q = \xi / \|\xi\|$ , então  $\langle \xi, \tilde{w} \rangle = \epsilon \langle \xi, q \rangle = \epsilon \|\xi\| > 0$ . Mas isto é uma contradição com (1.2.4). Consequentemente, a relação (1.2.3) se satisfaz.

A seguinte proposição prova que a relação (1.2.3) é válida quando  $N_C(x)$  substitui  $\partial d_C(x)$ . Adicionalmente, estabelece que a recíproca também se verifica.

**Proposição 1.2.3** Seja  $x \in C \subset \mathbb{R}^n$  um ponto fronteira. Então

$$p \in \text{int}(T_C(x)) \iff \langle \xi, p \rangle < 0, \quad \forall \xi \in N_C(x) \setminus \{0\}.$$

**Prova:** ( $\implies$ ) A prova é análoga a da relação (1.2.3), acima, usando o fato de que  $N_c(x) = \overline{U_{\lambda \geq 0} \lambda \partial d_c(x)}$ .

( $\impliedby$ ) Seja  $p$  tal que  $\langle \xi, p \rangle < 0, \quad \forall \xi \in N_C(x) \setminus \{0\}$ . Supor que  $p \notin \text{int}(T_C(x))$ . Então, dado  $i \in \mathbb{N}$ , existe  $q_i \in p + \left(\frac{1}{i}\right)B$ , tal que  $\langle \xi_i, q_i \rangle > 0$ , para alguma  $\xi_i \in$

$\partial d_C(x) \subset N_C(x)$ . Desde que  $N_C(x)$  é um cone e  $\xi_i \neq 0$ , para todo  $i$ , tem-se que  $\{\xi_i/\|\xi_i\|\}_{i \in N} \subset N_C(x)$ .

Portanto, desde que a esfera unitária é compacta e  $N_C(x)$  é fechado, tomando subsequências se necessário, conclui-se que

$$\frac{\xi_i}{\|\xi_i\|} \longrightarrow \xi \in N_C(x) \setminus \{0\}.$$

Por outro lado, como  $\langle \xi_i/\|\xi_i\|, q_i \rangle > 0$ , para todo  $i$ , passando ao limite, tem-se que  $\langle \xi, p \rangle \geq 0$ , com  $\xi \in N_C(x) \setminus \{0\}$ . Mas isto é uma contradição com a hipótese. Daqui,  $p \in \text{int}(T_c(x))$ .

### 1.3 – Gradiente Generalizado da Função Marginal do Seguidor

Nesta seção apresentam-se estimativas da derivada direcional de Clarke e do gradiente generalizado da função marginal do seguidor sob condições de primeira e segunda ordem. Segue-se, para isto, os trabalhos desenvolvidos por [CF1], Gauvin-Dubéau [GD1] e Rockafellar [RO1], [RO3], [RO4].

Dado  $x \in X$ , considerando o problema do seguidor  $(P_x)$ , definido pela relação (1.1.2), e a função ordinária de Lagrange associada, dada por

$$l(x, y, \pi) := f(x, y) + \langle \pi, g(x, y) \rangle, \text{ para } (y, \pi) \in \mathbb{R}^{n_2} \times \mathbb{R}_+^{m_2},$$

define-se o conjunto de multiplicadores de primeira ordem de  $(P_x)$  associado a  $y \in Y(x)$

como

$$\prod^1(x, y) := \{\pi \in \mathbb{R}^{m_2} / \nabla_y l(x, y, \pi) = 0, \pi \geq 0, \langle \pi, g(x, y) \rangle = 0\}. \quad (1.3.1)$$

Notar que, desde que  $y \in Y(x)$  e  $\pi \geq 0$ , cumpre-se que

$$\langle \pi, g(x, y) \rangle = 0 \iff \pi_i = 0, \quad \forall i \notin I(x, y),$$

onde  $I(x, y) := \{i / g_i(x, y) = 0\}$ .

Precisar-se-á do conceito de multiplicadores de segunda ordem, analisado por Rockafellar [RO3], [RO4]. Previamente, é necessário introduzir a definição de seqüência de subespaços  $\{M^j\} \subset \mathbb{R}^{n_2}$  convergentes ao subespaço  $M$  quando  $j \rightarrow \infty$ , no sentido que

$$d_{M^j}(z) \rightarrow d_M(z), \quad \forall z \in \mathbb{R}^{n_2}.$$

Toda seqüência de subespaços de  $\mathbb{R}^{n_2}$  tem uma subsequência convergente neste sentido (Rockafellar [RO3]).

Para cada  $y \in Y(x)$ , considerar  $\mathcal{M}(x, y)$  como a coleção de subespaços  $M \subset \mathbb{R}^{n_2}$  que se podem expressar como limites de subespaços  $M^j$ , no sentido anterior, onde

$$M^j = \{w \in \mathbb{R}^{n_2} / \langle w, \nabla_y g_i(x^j, y^j) \rangle = 0, \forall i \in I(x, y)\},$$

com  $(x^j, y^j) \rightarrow (x, y)$ .

É claro que  $\mathcal{M}(x, y)$  contém o subespaço

$$\mathcal{T}(x, y) = \{w \in \mathbb{R}^{n_2} / \langle w, \nabla_y g_i(x, y) \rangle = 0, \forall i \in I(x, y)\}.$$

Pode-se estabelecer que qualquer outro subespaço  $M \in \mathcal{M}(x, y)$  está contido em  $\mathcal{T}(x, y)$  e tem dimensão, ao menos,  $n_2 - |I(x, y)|$ , onde  $|I(x, y)|$  é o número de índices em  $I(x, y)$ .

O conjunto de multiplicadores de segunda ordem de  $(P_x)$  associado com  $y \in Y(x)$ , define-se como

$$\overset{2}{\Pi}(x, y) := \left\{ \pi \in \overset{1}{\Pi}(x, y) / \begin{array}{l} M \in \mathcal{M}(x, y) \text{ tal que} \\ \langle w, \nabla_{yy}^2 l(x, y, \pi)w \rangle \geq 0, \forall w \in M \end{array} \right\}.$$

Os conjuntos de multiplicadores singulares de primeira e segunda ordem,  $\overset{1}{\Pi}_0(x, y)$  e  $\overset{2}{\Pi}_0(x, y)$ , definem-se de maneira análoga a  $\overset{1}{\Pi}(x, y)$  e  $\overset{2}{\Pi}(x, y)$ , respectivamente, mas a função Lagrangeano ordinário é substituída pelo *Lagrangeano singular*:

$$l_0(x, y, \pi) := \langle \pi, g(x, y) \rangle, \quad \text{para } (y, \pi) \in \mathbb{R}^{n_2} \times \mathbb{R}_+^{m_2}.$$

Necessitar-se-á, ainda, de um outro conjunto de multiplicadores relacionado à função *Lagrangeano aumentado*:

$$\begin{aligned} L_a(x, y, \pi, r) := & f(x, y) + \sum_{i=1}^{m_2} \pi_i g_i(x, y) + \frac{r}{2} \sum_{i=1}^{m_2} [g_i(x, y)]^2 \\ & + \frac{1}{2r} \sum_{i=1}^{m_2} \left( [\pi_i + r g_i(x, y)]_+^2 - [\pi_i + r g_i(x, y)]^2 \right) \end{aligned} \tag{1.3.2}$$

para  $(y, \pi, r) \in \mathbb{R}^{n_2} \times \mathbb{R}^{m_2} \times \mathbb{R}_+$ , onde  $[a]_+ := \max\{a, 0\}$ , para  $a \in \mathbb{R}$ .

O conjunto de multiplicadores aumentados associado a  $y \in Y(x)$ , define-se como

$$\overset{a}{\Pi}(x, y) := \left\{ \pi \in \mathbb{R}^{m_2} / \begin{array}{l} (y, \pi) \text{ é um ponto sela de} \\ L_a(x, \cdot, \cdot, r), \text{ para } r \text{ suficientemente grande} \end{array} \right\}. \quad (1.3.3)$$

Prova-se (observação da Proposição 5, Rockafellar [RO3]) que

$$\overset{a}{\Pi}(x, y) := \left\{ \pi \in \overset{1}{\Pi}(x, y) / \begin{array}{l} y \text{ é um mínimo local de} \\ l(x, \cdot, \pi) + \frac{r}{2} \left( \sum_{i \in I_0(x, y, \pi)} [g_i(x, \cdot)]_+^2 + \right. \\ \left. \sum_{i \in I_+(x, y, \pi)} [g_i(x, \cdot)]^2 \right), \\ \text{para } r \text{ suficientemente grande} \end{array} \right.$$

Uma condição suficiente para a otimidade local de  $y$ , no problema  $(P_x)$ , é  $\overset{a}{\Pi}(x, y) \neq \emptyset$  (Rockafellar [RO1]). Outras relações entre os conjuntos de multiplicadores definidos acima e condições de qualificação estão dadas nas proposições seguintes.

**Proposição 1.3.1** (Teorema 1, Proposição 7, Rockafellar [RO2])

Para quaisquer  $x \in X$  e  $y \in Y(x)$ , as seguintes proposições são equivalentes.

- (i)  $\overset{1}{\Pi}_0(x, y) = \{0\}$ .
- (ii) A condição de qualificação de Cottle se cumpre para  $(P_x)$  no ponto  $y$ , i.e., não existe  $\pi \neq 0$  solução do sistema  $\sum_{i \in I(x, y)} \pi_i \nabla_y g_i(x, y) = 0$ ,  $\pi \geq 0$ .



Seja  $y \in Y(x)$  uma solução ótima local de  $(P_x)$ . Então (i) e (ii) são equivalentes a

(iii)  $\prod_1(x, y)$  é não vazio e limitado (e daqui, compacto, pois é fechado).

**Observação 1.3.2** A Proposição 1.3.1 é proposta originalmente por Rockafellar, para programas matemáticos que possuem restrições de igualdade e desigualdade ( $(P_x)$  só tem as últimas restrições). Portanto, em (iii) aparece a condição de qualificação de Mangasarian-Fromovitz (CQMF) em vez da de Cottle (ambas são equivalentes quando não existem restrições de igualdade).

Lembrar que a CQMF se cumpre para o programa

$$(P): \min_{y \in \mathbb{R}^{n_2}} J(y)$$

$$H_i(y) \leq 0, i = 1, \dots, i_1$$

$$L_j(y) = 0, j = 1, \dots, j_1$$

no ponto  $\bar{y}$ , se  $\{\nabla L_j(\bar{y})\}_{j=1}^{j_1}$  é linearmente independente e existe um vetor  $z$ , tal que

$$\langle \nabla H_i(\bar{y}), z \rangle < 0, \quad \forall i \in \{i/H_i(z) = 0\},$$

$$\langle \nabla L_j(\bar{y}), z \rangle = 0, \quad \forall j,$$

onde se assume a diferenciabilidade contínua para todas as funções  $H_i$  e  $L_j$ .

**Proposição 1.3.3** (Teorema 2, Rockafellar [RO4]). As seguintes propriedades se satisfazem:

(i) Os conjuntos de multiplicadores de segunda ordem  $\prod^2(x, y)$  e  $\prod_0^2(x, y)$  são fechados.

(ii) Seja  $y \in Y(x)$  uma solução ótima local de  $(P_x)$ . Se  $\prod_0^2(x, y) = \{0\}$ , então  $\prod^2(x, y)$  é não vazio e limitado (daqui, compacto).

**Observação 1.3.4.** Notar que se a condição de qualificação  $\prod_0^2(x, y) = \{0\}$  se satisfaz, então (ii) estabelece, em particular, que  $\prod^2(x, y) \neq \emptyset$  é uma condição necessária para a otimidade local de  $y \in Y(x)$  para  $(P_x)$ .

A seguinte proposição fornece uma condição suficiente para a existência de  $v^0(x; p)$ .

**Proposição 1.3.5** (Teorema 5, Rockafellar [RO4]). Seja  $x \in X$  e  $Y(x) \neq \emptyset$ . Se para todo  $y \in O(x)$ ,  $\prod_0^2(x, y) = \{0\}$ , então  $v$  é finita, Lipschitz próxima de  $x$  e, para todo  $p \in \mathbb{R}^{n_1}$ , satisfaz-se que

$$v^0(x; p) \leq \max_{y \in O(x)} \max_{\pi \in \prod^2(x, y)} \langle p, \nabla_x l(x, y, \pi) \rangle. \quad (1.3.4)$$

Este resultado é uma generalização do Teorema 5.1, Gauvin-Dubeau [GD1], onde se toma a condição de qualificação de primeira ordem  $\prod_0^1(x, y) = \{0\}$ .

**Proposição 1.3.6** (Teorema 7, Rockafellar [RO4]). Seja  $x \in X$  e  $Y(x) \neq \emptyset$ . Se para todo  $y \in O(x)$ ,  $\prod^a(x, y) \neq \emptyset$ , então, dado  $p \in \mathbb{R}^{n_1}$ , existe  $\bar{y} \in O(x)$  tal que

$$v_+(x; p) \geq \langle p, \nabla_x l(x, \bar{y}, \pi) \rangle, \quad \forall \pi \in \prod^a(x, \bar{y}). \quad (1.3.5)$$

Conseqüentemente,

$$v_+(x; p) \geq \inf_{y \in O(x)} \sup_{\pi \in \prod^a(x, y)} \langle p, \nabla_x l(x, y, \pi) \rangle,$$

onde  $v_+(x; p)$  e  $v^+(x; p)$  denotam as derivadas direcionais inferior e superior da Hadamard, respectivamente, definidas por

$$v_+(x; p) := \liminf_{\substack{p' \rightarrow p \\ t \downarrow 0}} \frac{v(x + tp') - v(x)}{t},$$

$$v^+(x; p) := \limsup_{\substack{p' \rightarrow p \\ t \downarrow 0}} \frac{v(x + tp') - v(x)}{t}.$$

Na seguinte proposição se estabelece uma estimativa do gradiente generalizado  $\partial v(x)$ , análoga à dada no Teorema 5.3, Gauvin-Dubeau [GD1], mas em termos de condições de segunda ordem.

**Proposição 1.3.7** Seja  $x_0 \in X$  e  $Y(x_0) \neq \emptyset$ . Se para todo  $y \in O(x_0)$ ,  $\prod_0^2(x, y) = \{0\}$ , então

$$\partial v(x_0) \subseteq \text{conv} \{ \nabla_x l(x_0, y, \pi) / y \in O(x_0), \pi \in \prod^2(x_0, y) \}. \quad (1.3.6)$$

Adicionalmente, se  $\frac{O(x)}{a}$  é semicontínua inferior em  $x_0$  e para todo  $y \in O(x_0)$  se verifica que  $\prod^2(x_0, y) \subseteq \prod^a(x_0, y)$ , então a igualdade se cumpre em (1.3.6).

**Prova:**

Seja  $\xi \in \partial v(x_0)$  e  $p \in \mathbb{R}^{n_1}$ , então  $\langle \xi, p \rangle \leq v^0(x_0; p)$  e, de (1.3.4),

$$\langle \xi, p \rangle \leq \max_{y \in O(x_0)} \max_{\pi \in \prod^2(x_0, y)} \langle p, \nabla_x l(x_0, y, p) \rangle.$$

Daqui,

$$\langle \xi, p \rangle \leq \max \left\{ \langle \eta, p \rangle / \eta \in \bigcup_{y \in O(x_0)} \bigcup_{\pi \in \prod^2(x_0, y)} \nabla_x l(x_0, y, \pi) \right\}.$$

Portanto, do Teorema 3.2.2, Rockafellar [RO0], obtém-se

$$\langle \xi, p \rangle \leq \max \left\{ \langle \eta, p \rangle / \eta \in \text{Conv} \left( \bigcup_{y \in O(x_0)} \bigcup_{\pi \in \prod^2(x_0, y)} \nabla_x l(x_0, y, \pi) \right) \right\}.$$

Da Hipótese 1.1.1 e da Proposição 1.3.3 - (ii), tem-se, respectivamente, que  $O(x_0)$  e  $\prod^2(x_0, y)$ , para todo  $y \in O(x_0)$ , são compactos. Então, usando o fato de serem as funções  $f$  e  $g$  duas vezes continuamente diferenciáveis, obtém-se que o conjunto

$$\{\nabla_x l(x_0, y, \pi) / y \in O(x_0), \pi \in \prod^2(x_0, y)\}$$

é fechado, pelo que da desigualdade anterior e do Teorema 1.3.1, Rockafellar [RO0], resulta que

$$\xi \in \text{Conv} \left( \{ \nabla_x l(x_0, y, \pi) / y \in O(x_0), \pi \in \prod^2(x_0, y) \} \right).$$

Assim, a primeira parte da proposição está provada.

Para provar a segunda parte, observar que, usando a Proposição 1.3.5, tem-se que  $v$  é finita e Lipschitz próxima de  $x_0$ , portanto,  $v^0(x_0; p)$  é finita para todo  $p \in \mathbb{R}^{n_1}$ .

Então,

$$v_+(x_0; p) \leq v^+(x_0; p) \leq v^0(x_0; p) < +\infty \quad (1.3.7)$$

e, da definição de  $v_+(x_0; p)$ , existem seqüências  $p^j \rightarrow p$  e  $t_j \downarrow 0$ , tal que

$$+\infty > \frac{v(x_0 + t_j p^j) - v(x_0)}{t_j} \rightarrow v_+(x_0; p).$$

Notar que  $x^j := x_0 + t_j p^j \rightarrow x_0$  e  $+\infty > v(x^j) \rightarrow v(x_0)$ .

Considerar, arbitrariamente,  $y_0 \in O(x_0)$ . Da Proposição 1.3.3,  $\prod^2(x_0, y_0) \neq \emptyset$ . Então, desde que  $\prod^2(x_0, y_0) \subseteq \overline{\prod^a(x_0, y_0)}$ , obtém-se que  $\prod^a(x_0, y_0) \neq \emptyset$ .

Seja  $\bar{\pi} \in \prod^a(x_0, y_0)$ . Então, de (1.3.3), existe uma vizinhança  $U(y_0)$  de  $y_0$  e  $\bar{r} > 0$ , tal que

$$v(x_0) = f(x_0, y_0) = L_a(x_0, y_0, \bar{\pi}, \bar{r}) \leq L_a(x_0, y, \bar{\pi}, \bar{r}), \quad \forall y \in U(y_0). \quad (1.3.8)$$

Desde que  $x^j \rightarrow x_0$  e  $y_0 \in O(x_0)$ , usando a semicontinuidade inferior de  $O(\cdot)$ , existe  $\{y^j\} \subseteq \mathbb{R}^{n_2}$ , tal que  $y^j \in O(x^j)$ , para  $j$  suficientemente grande, e  $y^j \rightarrow y_0$ .

Por outro lado, desde que  $y^j \in O(x^j)$ , para  $j$  suficientemente grande, conclui-se que

$$f(x^j, y^j) = v(x^j) \quad \text{e} \quad g_i(x^j, y^j) = 0, \quad \forall i. \quad (1.3.9)$$

Notar que  $\bar{\pi} \in \prod^a(x_0, y_0)$ , implica que  $\bar{\pi} \geq 0$  e  $\bar{\pi}_i = 0$ , para  $g_i(x_0, y_0) < 0$ . Daqui, usando (1.3.9), tem-se que  $\sum_{i=1}^{m_2} \bar{\pi}_i g_i(x^j, y^j) \leq 0$ .

Portanto, de (1.3.2),

$$L_a(x^j, y^j, \bar{\pi}, \bar{r}) \leq f(x_j, y^j) = v(x^j), \quad \text{para } j \text{ suficientemente grande.}$$

Desde que  $y^j \rightarrow y_0$ , pode-se considerar  $y^j \in \mathcal{U}(y_0)$ , para  $j$  suficientemente grande, então, de (1.3.8),

$$v(x_0) \leq L_a(x_0, y^j, \bar{\pi}, \bar{r}).$$

Agora, das duas últimas desigualdades, tem-se que

$$\frac{v(x^j) - v(x_0)}{t_j} \geq \frac{L_a(x^j, y^j, \bar{\pi}, \bar{r}) - L_a(x_0, y^j, \bar{\pi}, \bar{r})}{t_j}$$

e, desde que  $L_a$  é continuamente diferenciável em todas suas variáveis, usando o Teorema do Valor Médio, resulta que

$$\frac{v(x^j) - v(x_0)}{t_j} \geq \langle p^j, \nabla_x L_a(x_0 + \theta_j p^j, y^j, \bar{\pi}, \bar{r}) \rangle,$$

para algum  $\theta_j \in (0, t_j)$ .

Então,

$$\lim_{j \rightarrow \infty} \frac{v(x^j) - v(x_0)}{t_j} = v_+(x_0; p) \geq \langle p, \nabla_x L_a(x_0, y_0, \bar{\pi}, \bar{r}) \rangle = \langle p, \nabla_x l(x_0, y_0, \bar{\pi}) \rangle .$$

Desde que  $\bar{\pi} \in \prod^a(x_0, y_0)$  e  $y_0 \in O(x_0)$  foram escolhidos arbitrariamente, resulta, da desigualdade anterior e (1.3.7), que

$$+\infty > v^0(x_0; p) \geq \sup_{y \in O(x_0)} \sup_{\pi \in \prod^a(x_0, y)} \langle p, \nabla_x l(x_0, y, \pi) \rangle . \quad (1.3.10)$$

Tomar

$$\xi \in \text{Conv} \left( \left\{ \nabla_x l(x_0, y, \pi) / y \in O(x_0), \pi \in \prod^2(x_0, y) \right\} \right),$$

então, desde que  $\prod^2(x_0, y) \subseteq \overline{\prod^a(x_0, y)}$  e de (1.3.10), tem-se

$$\langle p, \xi \rangle \leq v^0(x_0; p), \quad \forall p \in \mathbb{R}^{n_1} .$$

Daqui, resulta que  $\xi \in \partial v(x_0)$ .

Assim,

$$\text{Conv}(\{\nabla_x l(x_0, y, \pi) / y \in O(x_0), \pi \in \prod^2(x_0, y)\}) \subseteq \partial v(x_0). \quad (1.3.11)$$

Finalmente, de (1.3.11) e (1.3.6), cumpre-se que

$$\partial v(x_0) = \text{Conv}(\{\nabla_x l(x_0, y, \pi) / y \in O(x_0), \pi \in \prod_2(x_0, y)\}). \quad (1.3.12)$$

A caracterização de  $\partial v(x_0)$  obtida no Teorema 5.3, Gauvin-Dubeau [GD1], baseia-se na condição de qualificação de independência linear dos gradientes das restrições ativas do seguidor (que implica, por exemplo,  $\prod_1(x_0, y) = \{\pi\}$ ,  $\forall y \in O(x_0)$ ). Observar, pois, que sendo uma condição de primeira ordem, pode ser bastante exigente. Notar, por outro lado, que a condição de semicontinuidade inferior de  $O(\cdot)$  em  $x_0$ , (hipótese da Proposição 1.3.7), limita a família de problemas para a qual é válida a igualdade em (1.3.6).

A Proposição 1.3.7 está baseada num resultado de Chen-Florian [CF1], mas nesta versão tem-se incluído a hipótese  $\prod_2(x_0, y) \subset \overline{\prod_a(x_0, y)}$ ,  $\forall y \in O(x_0)$ , como forma de assegurar a igualdade entre esses conjuntos e fazer com que se satisfaça a relação (1.3.12).

## 1.4 – Resultados Adicionais sobre o Gradiente Generalizado da Função Marginal do Seguidor

Como é de se esperar, sob hipóteses mais exigentes conseguem-se resultados adicionais. Nesta seção, impõem-se, de um lado, a hipótese de convexidade do problema do seguidor, e de outro, uma condição de qualificação mais restritiva que a CQMF (ver Proposição 1.3.1 e Observação 1.3.2), denominada CQMF estrita (Kyparisis [KY1]), para obter igualdade na relação 1.3.6 sem a exigência da semicontinuidade inferior de  $O(\cdot)$ .



Dado  $x_0 \in X$ , diz-se que  $(P_{x_0})$  é convexo, quando  $f(x_0, \cdot)$  e  $g(x_0, \cdot)$  são convexas.

**Teorema 1.4.1** Seja  $(P_{x_0})$  convexo. Se  $O(x_0) \neq \emptyset$  e  $\prod_0^1(x_0, y) = \{0\}$ , para todo  $y \in O(x_0)$ , então a derivada direcional (no sentido de Hadamard) de  $v$  existe, é finita para qualquer  $p \in \mathbb{R}^{n_1}$  e está dada por

$$v'(x_0; p) = \lim_{\substack{p' \rightarrow p \\ t \downarrow 0}} \frac{v(x_0 + tp') - v(x_0)}{t} = \min_{y \in O(x_0)} \max_{\pi \in \prod(x_0)} \langle \nabla_x l(x_0, y, \pi), p \rangle,$$

onde  $\prod(x_0)$  é o conjunto de soluções do programa dual de  $(P_{x_0})$ .

**Prova:**

Desde que  $\prod_0^1(x_0, y) = \{0\}$ , para todo  $y \in O(x_0)$ , usando a Proposição 1.3.1, tem-se que  $\prod^1(x_0, y) \neq \emptyset$  e compacto, para todo  $y \in O(x_0)$ .

Por outro lado,

$$0 \in \prod_0^2(x_0, y) \subseteq \prod_0^1(x_0, y) = \{0\}, \forall y \in O(x_0);$$

implica

$$\prod_0^2(x_0, y) = \prod_0^1(x_0, y) = \{0\}, \forall y \in O(x_0).$$

Além disso, desde que  $(P_{x_0})$  é convexo, obtem-se (Rockafellar [RO4])

$$\emptyset \neq \prod^1(x_0, y) = \prod^a(x_0, y) = \prod^2(x_0, y) = \prod(x_0), \forall y \in 0(x_0).$$

Portanto, as hipóteses do Teorema 3, Rockafellar [RO4], se satisfazem.

Daí, segue-se o resultado.

Considerar o problema

$$\begin{aligned} (P) : \min_{y \in \mathbb{R}^n} \quad & J(y) \\ \text{s.a.} \quad & H_i(y) \leq 0, i = 1, \dots, i_1 \\ & L_j(y) = 0, j = 1, \dots, j_1, \end{aligned}$$

onde todas as funções são reais e continuamente diferenciáveis.

Diz-se que a *condição de Independência Linear* (CIL) se cumpre para  $(P)$  no ponto  $\bar{y}$ , se

$$\{\nabla H_i(\bar{y})/i \in I(\bar{y})\} \cup \{\nabla L_j(\bar{y})\}_{j=1}^{j_1}$$

é linearmente independente, onde

$$I(\bar{y}) := \{i/H_i(\bar{y}) = 0, 1 \leq i \leq i_1\}.$$

Sabe-se que a CIL implica a unicidade dos multiplicadores de Karush-Kuhn-Tucker  $(K - K - T)$  associados a  $(P)$  no ponto  $\bar{y}$ .

Kyparisis [KY1] introduz uma condição de qualificação que é mais restritiva que a CQMF, mas menos restritiva que a CIL. Esta condição é chamada de *condição de qualificação de Mangasarian-Fromovitz estrita* (CQMFS).

Diz-se que a CQMFS se cumpre para  $(P)$  em  $\bar{y}$ , se

$$\{\nabla H_i(\bar{y})/i \in I_+(\bar{y})\} \cup \{\nabla L_j(\bar{y})\}_{j=1}^{j_1}$$

é linearmente independente e existe  $z \in \mathbb{R}^n$ , tal que

$$\langle \nabla H_i(\bar{y}), z \rangle < 0, i \in I_0(\bar{y}),$$

$$\langle \nabla H_i(\bar{y}), z \rangle = 0, i \in I_+(\bar{y}),$$

$$\langle \nabla L_j(\bar{y}), z \rangle = 0, j = 1, \dots, j_1,$$

onde  $I_+(\bar{y}) := \{i \in I(\bar{y})/\pi_i > 0\}$  e  $I_0(\bar{y}) := \{i \in I(\bar{y})/\pi_i = 0\}$  dependem do multiplicador de  $K - K - T$ ,  $\pi$ , associado as restrições de desigualdade de  $(P)$ .

**Proposição 1.4.2** (Proposição 1, Kyparasis [KY1]) Seja  $y_0$  um ponto viável de  $(P)$  e  $(\pi_0, \lambda_0)$  um multiplicador de  $K - K - T$  associado, i.e.,

$$\nabla J(y_0) + \langle \pi_0, \nabla H(y_0) \rangle + \langle \lambda_0, \nabla L(y_0) \rangle = 0,$$

$$\langle \pi_0, H(y_0) \rangle = 0,$$

$$\pi_0 \geq 0.$$

Então, a CQMFS se cumpre em  $y_0$  se, e somente se, o multiplicador de  $K - K - T$  associado,  $(\pi_0, \lambda_0)$ , é único.

Notar que, desde que a CIL implica unicidade dos multiplicadores de  $K - K - T$ , da Proposição 1.4.2, CIL implica CQMFS. Por outro lado, se  $y_0$  é solução local de  $(P)$ , a unicidade do conjunto de multiplicadores associado, implica, da Proposição 1.3.1 (iii) e Observação 1.3.2, a CQMF. Assim, tem-se que para  $y_0$ , solução local de  $(P)$ ,

$$CIL \implies CQMFS \implies CQMF.$$

Seja  $x \in X$ , diremos que a CQMFS se satisfaz para  $(P_x)$  em  $\bar{y} \in Y(x)$ , se  $\{\nabla_y g_i(x, \bar{y})\}_{i \in I_+(x, \bar{y})}$  é linearmente independente e existe  $z \in \mathbb{R}^{n_2}$ , tal que

$$\langle \nabla_y g_i(x, \bar{y}), z \rangle < 0, \quad i \in I_0(x, \bar{y}),$$

$$\langle \nabla_y g_i(x, \bar{y}), z \rangle = 0, \quad i \in I_+(x, \bar{y}),$$

onde  $I_+(x, \bar{y}) = \{i \in I(x, \bar{y}) / \pi_i > 0\}$  e  $I_0(x, \bar{y}) = I(x, \bar{y}) \setminus I_+(x, \bar{y})$ .

**Teorema 1.4.3** Seja  $(P_{x_0})$  convexo e  $O(x_0) \neq \emptyset$ . Se a CQMFS se verifica para todo  $y \in O(x_0)$ , então se cumpre que

$$\partial v(x_0) = \text{Conv}(\{\nabla_x l(x_0, y, \pi_0) / y \in O(x_0)\}),$$

onde  $\pi_0$  é o único multiplicador associado a qualquer  $y \in O(x_0)$  (i.e.,  $\Pi(x_0) = \{\pi_0\}$ ).

**Prova:**

A CQMFS implica a CQMF, que, pela Proposição 1.3.1, é equivalente a  $\prod_0^1(x_0, y) = \{0\}$ , para todo  $y \in O(x_0)$ . Adicionalmente, a CQMFS também implica que, dado  $y \in O(x_0)$ ,  $\prod^1(x_0, y)$  é unitário.

Desde que  $(P_{x_0})$  é convexo, cumpre-se que existe  $\pi_0 \in \mathbb{R}_+^{m_2}$ , tal que

$$\{\pi_0\} = \prod_0^1(x_0, y) = \prod^a(x_0, y) = \prod^2(x_0, y) = \prod(x_0), \forall y \in O(x_0).$$

Portanto, do Teorema 1.4.1, cumpre-se que

$$v'(x_0; p) = \lim_{t \downarrow 0} \frac{v(x_0 + tp) - v(x_0)}{t} = \min_{y \in O(x_0)} \langle \nabla_x l(x_0, y, \pi_0), p \rangle.$$

Assim, usando a igualdade anterior e o Teorema 5, Rockafellar [RO4], dado  $p \in \mathbb{R}^{n_1}$ , tem-se que

$$\begin{aligned} -v'(x_0; -p) &= \max_{y \in O(x_0)} \langle \nabla_x l(x_0, y, \pi_0), p \rangle = \lim_{t \downarrow 0} \frac{v(x_0) - v(x_0 - tp)}{t} \\ &\leq \limsup_{\substack{x' \rightarrow x_0 \\ t \downarrow 0}} \frac{v(x' + tp) - v(x')}{t} = v^0(x_0; p) \leq \max_{y \in O(x_0)} \langle \nabla_x l(x_0, y, \pi_0), p \rangle, \end{aligned}$$

então,

$$v^0(x_0; p) = \max_{y \in O(x_0)} \langle \nabla_x l(x_0, y, \pi_0), p \rangle.$$

Agora, para  $\xi \in \text{Conv}(\{\langle \nabla_x l(x_0, y, \pi_0) \rangle / y \in O(x_0)\})$ , obtem-se que

$$\langle \xi, p \rangle \leq \max_{y \in O(x_0)} \langle \nabla_x l(x_0, y, \pi_0), p \rangle = v^0(x_0; p), \forall p \in \mathbb{R}^{n_1},$$

e, daí,  $\xi \in \partial v(x_0)$ , i.e.,  $\text{Conv}(\{\nabla_x l(x_0, y, \pi_0) / y \in O(x_0)\}) \subseteq \partial v(x_0)$ .

Por outro lado, desde  $\prod_0^1(x_0, y) = \prod_0^2(x, y) = \{0\}$ , para todo  $y \in O(x_0)$ , as hipóteses da Proposição 1.3.7 se satisfazem. Portanto,

$$\partial v(x_0) = \text{Conv}(\{\nabla_x l(x_0, y, \pi_0) / y \in O(x_0)\}).$$

**Observação 1.4.4** Este resultado é obtido em Gauvin Dubean [GD1] sob a condição mais forte de independência linear substituindo a CQMFS.

## Capítulo 2

### Condições Necessárias de Otimidade para o Problema de Programação de Dois Níveis

Neste capítulo estuda-se a existência de multiplicadores de Karush-Kuhn-Tucker ( $K - K - T$ ) como condição necessária para soluções locais de (PPDN). Em concordância com Chen-Florian [CF1], onde se prova que a condição de qualificação de Mangasarian-Fromovitz (CQMF) não se verifica para o caso do problema com funções objetivo quadráticas, observa-se na seção 2.1, que tal condição também não se satisfaz em qualquer outro caso, para a formulação equivalente (P1). Estabelece-se, assim, a necessidade de propor uma condição menos exigente para a regularidade do problema (PPDN). Isto é feito na seção 2.2 onde se introduz o conceito de *programa calmo* (no sentido de Clarke [CL1], Cap. VI), que, entre outros resultados, é implicado pela condição de Slater e a CQMF (quando é o caso delas se satisfazerem). Na seção 2.3, sob algumas condições de segunda ordem para o problema do seguidor e a de calma para o problema (PPDN), estabelece-se o resultado principal deste capítulo a condição de  $K - K - T$  para (PPDN) (Teorema 2.3.6). Conclui-se esta seção com a apresentação de um exemplo onde se ilustra o referido resultado. Por último, na seção 2.4, estudam-se dois casos particulares de problema (PPDN):

- Conjunto de reação racional unitário, i.e.,  $O(x) = \{y\}$ ;
- Problema do seguidor convexo, i.e.,  $(P_x)$  convexo para  $x \in X$ , fixo.

Estes casos têm sido estudados, de maneira similar, por Dempe [DE2] e Outrata [OT2] (ver também Outrata [OT4], onde programas matemáticos com restrições determinadas por desigualdades variacionais, formulam-se como programas de dois níveis).

Há, no entanto, algumas diferenças entre a abordagem deles e a aqui apresentada, sendo talvez as mais relevantes, de um lado, a condição de segunda ordem por eles imposta sobre o problema do seguidor (condição equivalente à Condição de Regularidade Forte de Robinson [RB1]), para garantir, tanto unicidade de  $O(x)$ , assim como certas propriedades diferenciáveis da (nesse caso) função resposta do seguidor  $y(x)$ ; por outro lado, o fato de aqui se obter resultados análogos aos obtidos por eles, como casos particulares do caso geral estudado.

## 2.1 – Por Que Não Condições de Qualificação Tipo Mangasarian Fromovitz?

Para responder esta pergunta se precisará relembrar alguns conceitos e resultados relativos a condições de otimidade para o caso de problemas de programação não diferenciável. Para um estudo detalhado ver Clarke [CL1].

Considerar o problema

$$\begin{aligned}
 (P) : \quad & \min_{x \in C} J(x) \\
 & \text{s.a.} \quad H_i(x) \leq 0, i = 1, 2, \dots, i_1, \\
 & \quad \quad L_j(x) = 0, j = 1, 2, \dots, j_1,
 \end{aligned} \tag{2.1.1}$$

onde  $C \subset D \subset \mathbb{R}^n$ ,  $C$  é fechado,  $D$  é aberto e todas as funções envolvidas são localmente Lipschitz em  $D$ .

Seguindo Clarke [CL1], Capítulo VI, pode-se definir a função Lagrangeano do programa  $(P)$ , como

$$\mathcal{L}(x, \lambda, r, s, k) := \lambda J(x) + \langle r, H(x) \rangle + \langle s, L(x) \rangle + k \|(\lambda, r, s)\| d_C(x), \tag{2.1.2}$$



para  $(x, \lambda, r, s, k) \in \mathbb{R}^{n+1+i_1+j_1+1}$ .

Tem-se a seguinte proposição:

**Proposição 2.1.1** (Teorema 6.1.1, Clarke [CL1]) . Seja  $x$  uma solução ótima local de  $(P)$ . Então, existem  $\lambda \geq 0, r \geq 0$  e  $s$ , não todos nulos, para  $k$  suficientemente grande, tal que  $\langle r, H(x) \rangle = 0$  e

$$0 \in \lambda \partial J(x) + \sum_{i=1}^{i_1} r_i \partial H_i(x) + \sum_{j=1}^{j_1} s_j \partial L_j(x) + k \|(\lambda, r, s)\| \partial d_C(x). \quad (2.1.3)$$

### Observação 2.1.2

(i) Desde que  $N_C(x) = \overline{\cup_{\lambda \geq 0} \lambda \partial d_C(x)}$ , tem-se, da relação (2.1.3), que

$$0 \in \lambda \partial J(x) + \sum_{i=1}^{i_1} r_i \partial H_i(x) + \sum_{j=1}^{j_1} s_j \partial L_j(x) + N_C(x). \quad (2.1.4)$$

(ii) O vetor  $(\lambda, r, s)$  da relação (2.1.4) será denominado multiplicador (este conceito de multiplicador é uma relaxação daquele considerado por Clarke [CL1], a partir da Proposição 2.1.1).

(iii) Quando  $C = \mathbb{R}^n$  e as funções envolvidas são suaves (i.e., continuamente diferenciáveis), as relações (2.1.3) e (2.1.4) coincidem e reduzem-se ao resultado clássico das condições necessárias de otimidade de Fritz John.



**Definição 2.1.4** Seja  $\bar{x}$  viável para o programa (P), onde  $C \subseteq D \subset \mathbb{R}^n$ ,  $C$  é fechado,  $D$  aberto,  $J, H_i$  e  $L_j$  continuamente diferenciáveis, para todo  $i$  e todo  $j$ , exceto para um índice (i.e., para algum  $k_0 \in \{1, \dots, i_1\} \cup \{1, \dots, j_1\}$ ), onde a respectiva função é localmente Lipschitz em  $D$ . Dir-se-á que a condição de qualificação de Mangasarian-Fromovitz estendida (CQMFEE) se satisfaz no ponto  $\bar{x}$  para o programa (P), se  $\{\xi_j\}_{j=1}^{j_1}$  é linearmente independente, para todo  $(\xi_1, \dots, \xi_{j_1}) \in \partial L_1(x) \times \dots \times \partial L_{j_1}(x)$ , e existe  $z \in \text{int}(T_C(\bar{x}))$  tal que

$$\langle z, \gamma_i \rangle < 0, \quad \forall \gamma_i \in \partial H_i(\bar{x}), \forall i \in I(\bar{x}) := \{i / H_i(\bar{x}) = 0\},$$

$$\langle z, \xi_j \rangle = 0, \quad \forall \xi_j \in \partial L_j(\bar{x}), \forall j.$$

### Observação 2.1.5

- (i) É claro que na definição anterior só se tratará com gradientes generalizados que se reduzem ao respectivo vetor gradiente, exceto para o índice  $k_0$ , onde o respectivo gradiente generalizado não será necessariamente um conjunto unitário.
- (ii) A definição acima se tem restringido ao caso em que as funções envolvidas, exceto uma, são continuamente diferenciáveis; esta última se assume localmente Lipschitz. Esse é o caso do problema (P1), em que a restrição  $v(x) - f(x, y) \leq 0$  é localmente Lipschitz (desde que se satisfaçam as hipóteses da Proposição 1.3.5 para todo  $x \in X$ ), entretanto as outras restrições são pelo menos continuamente diferenciáveis (em particular não existem restrições de igualdade).

Para ilustrar a coerência da CQMFEE, mostra-se a seguir que esta é uma condição suficiente para  $M^0(\bar{x}) = \{0\}$ .

**Proposição 2.1.6** Seja  $\bar{x}$  uma solução ótima local para  $(P)$ , onde se satisfazem as hipóteses da Definição 2.1.4. Se a CQMFEE se verifica em  $\bar{x}$ , então  $M^0(\bar{x}) = \{0\}$ .

**Prova:**

Supor que  $M^0(\bar{x}) \neq \{0\}$ . Então, da Proposição 2.1.1, existe  $(r, s) \in \mathbb{R}_+^{i_1} \times \mathbb{R}^{j_1}$ , tal que  $(r, s) \neq 0$ ,  $\langle r, H(\bar{x}) \rangle = 0$  e  $0 = \sum_{i=1}^{i_1} r_i \gamma_i + \sum_{j=1}^{j_1} s_j \xi_j + \nu$ , para alguns  $\gamma_i \in \partial H_i(\bar{x})$ ,  $\xi_j \in \partial L_j(\bar{x})$  e  $\nu \in N_c(\bar{x})$ .

Supor que  $r = 0$ . Então  $s \neq 0$  e  $0 = \sum_{j=1}^{j_1} s_j \xi_j + \nu$ . Mas como  $\nu \neq 0$ , pois, caso contrário,  $0 = \sum_{j=1}^{j_1} s_j \xi_j$  e  $s \neq 0$ , que é uma contradição com  $\{\xi_j\}_{j=1}^{j_1}$  linearmente independente.

Portanto,  $\nu \in N_c(x) \setminus \{0\}$ . Daqui e do produto interno pelo vetor  $z$  da Definição 2.1.4, na igualdade anterior, resulta

$$0 = \sum_{j=1}^{j_1} s_j \underbrace{\langle z, \xi_j \rangle}_{=0} + \langle z, \nu \rangle = \langle z, \nu \rangle.$$

de onde, usando o fato de que  $z \in \text{int}(T_c(\bar{x}))$  e a Proposição 1.2.3, tem-se

$$0 = \langle z, \nu \rangle < 0,$$

que é uma contradição gerada pela suposição  $r = 0$ .

Então,  $r \neq 0$  (i.e, existe  $r_i > 0$  e  $H_i(\bar{x}) = 0$ ), que implica  $I(\bar{x}) \neq \emptyset$ .

Desde que  $I(\bar{x}) \neq \emptyset$ , da CQMFEE, resulta que

$$0 = \sum_{i \in I(\bar{x})} \underbrace{r_i \langle z, \gamma_i \rangle}_{r_i > 0, \langle z, \gamma_i \rangle < 0} + \sum_{j=1}^{j_1} s_j \underbrace{\langle z, \xi_j \rangle}_{=0} + \underbrace{\langle z, \nu \rangle}_{< 0} < 0,$$

que é uma contradição.

Consequentemente,  $M^0(\bar{x}) = \{0\}$ .

No que segue, formula-se a CQMFE para o programa (PPDN). Sem perda de generalidade e para focar a essência do assunto, considera-se que no programa (PPDN) não aparecem as restrições do tipo  $G(x, y) \leq 0$ . É natural considerar que a CQMFE se satisfaz no ponto  $(x, y)$  para (PPDN), se se satisfaz para (P1) nesse ponto, i.e., se no ponto  $(x, y)$ , viável para (P1), cumpre-se que existe  $z = (z^x, z^y) \in \text{int}(T_X(x)) \times \mathbb{R}^{n_2}$ , tal que

$$\langle z^x, \nabla_x f(x, y) - \xi \rangle + \langle z^y, \nabla_y f(x, y) \rangle < 0, \forall \xi \in \partial v(x),$$

$$\langle z^x, \nabla_x g_i(x, y) \rangle + \langle z^y, \nabla_y g_i(x, y) \rangle < 0, \forall i \in I(x, y).$$

Infelizmente, estas condições não se verificam, como seria natural, ainda nos casos mais simples. Considerar, por exemplo, a seguinte situação: seja  $x \in X$ , fixo, e supor, a partir da Proposição 1.3.7, que

$$\partial v(x) = \text{Conv}(\{\nabla_x l(x, y, \pi) / y \in O(x), \pi \in \prod_{i=1}^2 (x, y)\}).$$

Supor, também, que  $O(x) = \{y\}$  e  $\prod_{i=1}^2 (x, y) = \prod_{i=1}^1 (x, y)$  (que é o caso, por exemplo, quando  $(P_x)$  é convexo e  $f$  estritamente convexa).

Além disso, admita-se que  $\prod^1(x, y) = \{\pi\}$  (que ocorre, por exemplo, quando  $\{\nabla_y g_i(x, y)\}_{i \in I(x, y)}$  é linearmente independente).

Sob essas hipóteses, tem-se que

$$\partial v(x) = \{\nabla_x f(x, y) + \sum_{i=1}^{m_1} \pi_i \nabla_x g_i(x, y)\},$$

onde  $(\pi_1, \dots, \pi_{m_1}) = \pi$ .

Lembrar que

$$\pi \in \prod^1(x, y) \iff \nabla_y f(x, y) + \sum_{i \in I(x, y)} \pi_i \nabla_y g_i(x, y) = 0, \pi \geq 0, \langle \pi, g(x, y) \rangle = 0.$$

Portanto, o sistema em  $z = (z^x, z^y) \in \text{int}(T_X(x) \times \mathbb{R}^{n_2})$

$$\langle z^x, \nabla_x f(x, y) - \xi \rangle + \langle z^y, \nabla_y f(x, y) \rangle < 0, \forall \xi \in \partial v(x),$$

$$\langle z^x, \nabla_x g_i(x, y) \rangle + \langle z^y, \nabla_y g_i(x, y) \rangle < 0, \forall i \in I(x, y),$$

é equivalente ao sistema

$$- \langle z^x, \sum_{i \in I(x, y)} \pi_i \nabla_x g_i(x, y) \rangle - \langle z^y, \sum_{i \in I(x, y)} \pi_i \nabla_y g_i(x, y) \rangle < 0,$$

$$\langle z^x, \sum_{i \in I(x, y)} \pi_i \nabla_x g_i(x, y) \rangle + \langle z^y, \sum_{i \in I(x, y)} \pi_i \nabla_y g_i(x, y) \rangle < 0,$$

que é incompatível.

Portanto, a CQMFEE não se satisfaz.

Para superar as dificuldades assinaladas, é que se propõe, nas próximas seções, uma abordagem do estudo das condições de otimidade do problema (PPDN) através de uma formulação mais geral, fornecida pelo conceito de problema calmo, estudado por Clarke [CL1] e Rockafellar [RO2].

## 2.2 – O Conceito de Problema Calmo

O conceito de problema calmo, aqui discutido, tem sido usado em vários contextos e para diversos fins:

- No contexto da programação paramétrica, por Rockafellar [RO2], para garantir que o conjunto de multiplicadores de segunda ordem  $\prod^2(x, y)$  seja não vazio (fechado, mas não necessariamente compacto – conseguindo, assim, uma formulação mais fraca que a estabelecida pela Proposição 1.3.3 (ii)).
- No contexto do PPDN, por Outrata [OT2], [OT3], em relação à função de penalidade por ele considerada no tratamento numérico do problema (PPDN), para garantir penalidade exata.
- No contexto do estudo de condições de qualificação, por Clarke [CL1], para garantir  $M_k^0(\lambda) = \{0\}$  (esta abordagem é a que se tomará como ponto de partida neste trabalho).

Considerar o problema (P), dado pela relação (2.1.1), como parte da seguinte família de programas matemáticos parametrizados

$$\begin{aligned}
 (P(p, q)) : \min_{x \in C} & \quad J(x) \\
 \text{s.a.} & \quad H(x) + p \leq 0 \\
 & \quad L(x) + q = 0,
 \end{aligned} \tag{2.2.1}$$

onde  $(p, q) \in \mathbb{R}^{i_1+j_1}$ , i.e., (P) é o problema  $(P(0, 0))$ .

**Definição 2.2.1** (Definição 6.4.1, Clarke [CL1]) Seja  $x$  solução de (P). O problema (P) é *calmo* em  $x$ , se existem reais positivos  $\epsilon$  e  $M$ , tais que, para todo  $(p, q)$  em  $\epsilon B$  e para todo  $x'$  em  $x + \epsilon B$ , viáveis para  $(P(p, q))$ , tem-se

$$J(x') - J(x) + M\|(p, q)\| \geq 0.$$

De maneira análoga à definição da função marginal do seguidor, em (1.1.3), define-se a função marginal para o problema  $(P(p, q))$ , como a função  $V : \mathbb{R}^{i_1+j_1} \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$  tal que

$$V(p, q) := \inf_{x \in C} \{J(x) / H(x) + p \leq 0, L(x) + q = 0\}, \tag{2.2.2}$$

onde  $V(p, q) := +\infty$ , se o conjunto viável de  $(P(p, q))$  é vazio.

**Proposição 2.2.2** (Proposição 6.4.2, Clarke [CL1]) Seja  $V(0, 0)$  finito. Supor que



$$\liminf_{(p,q) \rightarrow (0,0)} \frac{V(p,q) - V(0,0)}{\|(p,q)\|} > -\infty$$

(o que é verdadeiro, em particular, se  $V$  é Lipschitz próxima de 0). Então, para qualquer solução  $x$  de (P), (P) é calmo em  $x$ .

Dir-se-á que o problema (P) é calmo sempre que  $V$  satisfaça as hipóteses da Proposição 2.2.2.

Como consequência da definição e do corolário do Teorema 2.9.1, Clarke [CL1], segue-se o seguinte resultado:

**Lema 2.2.3** Se o problema (P) é calmo, então

(i)  $V(\cdot, \cdot)$  é semicontínua inferiormente em  $(0, 0)$  e  $V^0((0, 0); (0, 0)) > -\infty$ .

(ii)  $\partial V(0, 0) \neq \emptyset$ .

O conceito de problema calmo, como será visto na seção 2.3, fornece uma condição de regularidade do problema (P) que é mais fraca que a CQMF (para  $C = \mathbb{R}^n$ ) ou que a condição de Slater (quando  $C$  é convexo,  $J$  Lipschitz em  $C$ ,  $H_i$  convexa para todo  $i$  e não existem restrições de igualdade) – lembrar que a condição de Slater postula a existência de um ponto “estritamente viável”:  $\hat{x} \in C$ , tal que  $H_i(\hat{x}) < 0, \forall i$ . Com efeito, se a condição de Slater se verifica, usando o Corolário 2 do Teorema 6.3.2, Clarke [CL1], tem-se que  $V$  é Lipschitz próxima de 0. Daqui, usando a Proposição 2.2.2, resulta que (P) é calmo (na seção 2.3 prova-se que a CQMF implica calma para (P)).

Como em Clarke [CL1], obtém-se os seguintes resultados:

- Se a condição de Slater é satisfeita para (P), então  $M^0(x) = \{0\}$ .

- Se  $M^0(x) = \{0\}$  para o problema (P), então (P) é calmo em  $x$ .

A seguinte proposição, usada por Outrata [OT2], [OT3], no tratamento numérico do problema (PPDN), relaciona o conceito de calma à penalização exata:

**Proposição 2.2.4** (Proposição 6.4.3, Clarke [CL1]) Seja (P) calmo e  $x$  uma solução de (P). Então existe  $M > 0$ , tal que  $x$  é um mínimo local em  $C$  para a função  $h$ , dada por

$$h(y) := J(y) + M \max_{i,j} \{H_i(y), |L_j(y)|, 0\}.$$

## 2.3 – Condições de Otimidade Necessárias para (PPDN)

O conceito de problema calmo, apresentado na seção anterior, tem sido formulado de maneira geral. Sob hipóteses adicionais, obtém-se resultados que ajudam a estabelecer as condições de otimidade necessárias para o problema (PPDN), apresentadas nesta seção no Teorema 2.3.6.

**Hipótese 2.3.1** (Hipótese 6.5.1, Clarke [CL1])  $V_{(0,0)}$  é finito e existe um subconjunto compacto  $\Omega$  de  $C \subset \mathbb{R}^n$  e um número positivo  $\epsilon_0$ , tais que, para todo  $(p, q)$  em  $\epsilon_0 B$  que verifica  $V(p, q) < V(0, 0) + \epsilon_0$ , o problema (P  $(p, q)$ ) tem uma solução que pertence a  $\Omega$ .

### Observação 2.3.2

- (i) A hipótese anterior não exige que  $(P(p, q))$  admita pontos viáveis para todo  $(p, q)$  próximo de  $(0, 0)$ ; também não exige semicontinuidade inferior de  $V$  em uma vizinhança de  $(0, 0)$ .
- (ii) Uma condição simples que garante a existência de  $\Omega$  e  $\epsilon_0$  é que  $J(\cdot)$  seja compacta inferiormente em  $C$ , i.e., para cada  $\alpha \in \mathbb{R}$ , o conjunto  $\{x \in C / J(x) \leq \alpha\}$  seja compacto.

Denotar-se-á por  $\Sigma$  o conjunto de soluções de  $(P)$  que estão em  $\Omega$  e por  $M_k^\lambda(\Sigma)$  o conjunto  $\cup_{x \in \Sigma} M_k^\lambda(x)$ , onde  $M_k^\lambda(x)$  foi definido na Observação 2.1.3.

Notar que, sob as hipóteses feitas na seção 2.1, existe  $\hat{k} > 0$  para o qual a função  $(J, H, L)(\cdot)$  é Lipschitz, módulo  $\hat{k}$ , em uma vizinhança de  $\Omega$ .

**Proposição 2.3.3** (Corolário 1, Corolário 2 do Teorema 6.5.2, Clarke [CL1]) Sob a hipótese 2.2.5, para  $k > \hat{k}$ , tem-se

- (i) Se  $M_k^0(\Sigma) = \{0\}$ , então  $V$  é Lipschitz próxima de  $(0, 0)$  e  $\partial V(0, 0) \subseteq \text{Conv}(M_k^1(\Sigma))$ .
- (ii) Se  $\partial V(0, 0)$  é não vazio (em particular, se  $(P)$  é calmo), então existe uma solução  $x$  de  $(P)$ , que pertence a  $\Sigma$ , tal que  $M_k^1(x) \cap \partial V(0, 0) \neq \emptyset$ .

Como consequência da Proposição 2.3.3, obtém-se o seguinte resultado:

**Corolário 2.3.4** Seja  $x$  uma solução de  $(P)$ , onde  $(P)$  é calmo em  $x$ . Então, existe  $(r, s) \in \mathbb{R}_+^{i_1} \times \mathbb{R}^{j_1}$ , com  $\langle r, H(x) \rangle = 0$ , tal que

$$0 \in \partial J(x) + \sum_{i=1}^{i_1} r_i \partial H_i(x) + \sum_{j=1}^{j_1} s_j \partial L_j(x) + N_c(x). \quad (2.3.1)$$

**Prova:**

É consequência imediata da Proposição 6.4.4, Clarke [CL1]. Com efeito, a partir desta proposição, tem-se que para algum  $k$  suficientemente grande, existe  $(r, s) \in \mathbb{R}^{i_1} \times \mathbb{R}^{j_1}$ , tal que o vetor  $(1, r, s) \in \mathbb{R}^{1+i_1+j_1}$  satisfaz

$$0 \in \partial_x \mathcal{L}(x, 1, r, s, k), \langle r, H(x) \rangle = 0, r \geq 0.$$

Da relação (2.1.3) e a Observação 2.1.2(i), obtém-se que

$$\partial \mathcal{L}(x, 1, r, s, k) \subseteq \partial J(x) + \sum_{i=1}^{i_1} r_i \partial H_i(x) + \sum_{j=1}^{j_1} s_j \partial L_j(x) + N_c(x).$$

Portanto, conclui-se que

$$0 \in \partial J(x) + \sum_{i=1}^{i_1} r_i \partial H_i(x) + \sum_{j=1}^{j_1} s_j \partial L_j(x) + N_C(x), \langle r, H(x) \rangle = 0, r \geq 0.$$

### Observação 2.3.5

- (i) Desde que o enunciado do Corolário 2.3.4 pode ser formulado considerando  $x \in C \cap (x + \epsilon \bar{B})$ , para algum  $\epsilon > 0$  pequeno (onde  $\bar{B}$  é o fecho da bola unitária em  $\mathbb{R}^n$ ), é claro que o resultado é igualmente válido se  $x$  é uma solução local de (P).

(ii) A relação (2.3.1) é a condição de  $K - K - T$  para o problema (P). Portanto, a Proposição 2.3.4 estabelece que, sob a hipótese de calma no ponto  $x$ , a condição de  $K - K - T$  é uma condição de otimidade necessária para  $x$  a respeito do problema (P).

Antes de estabelecer o resultado principal deste capítulo, veja que a CQMFE no ponto  $x$  (e, portanto, a CQMF, se for o caso), implica a *condição de qualificação de calma do problema (P) em  $x$* . De fato, se  $x$  satisfaz a CQMFE, da Proposição 2.1.6, tem-se que  $M^0(x) = \{0\}$  e, daqui,  $M_k^0(x) = \{0\}$ , para  $k$  suficientemente grande. Então, da Proposição 2.3.3 (i) e a Proposição 2.2.2, resulta que (P) é calmo em  $x$ .

Em continuação, estabelece-se a condição de  $K - K - T$  como condição de otimidade necessária para o problema (P), sob as hipóteses gerais feitas no capítulo anterior e neste capítulo.

**Teorema 2.3.6** Seja  $(x^*, y^*)$  uma solução local do problema (PPDN), tal que (PPDN) é calmo em  $(x^*, y^*)$ . Se  $\prod_0^2(x^*, y) = \{0\}$  e  $\prod^2(x^*, y) \subseteq \prod^a(x^*, y)$ , para todo  $y \in O(x^*)$ , e se  $O(\cdot)$  é semicontínua inferior em  $x^*$ , então existe um vetor  $r \in \mathbb{R}^{m_1+m_2}$  e um escalar  $s \in \mathbb{R}$ , tais que

$$0 \in \nabla_x F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_x G_i(x^*, y^*) + \sum_{i=1}^{m_2} r_{m_1+i} \nabla_x g_i(x^*, y^*) + s[\nabla_x f(x^*, y^*) - \partial v(x^*)] + N_X(x^*), \quad (2.3.2)$$

$$0 = \nabla_y F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_y G_i(x^*, y^*) + \sum_{i=1}^{m_2} r_{m_1+i} \nabla_y g_i(x^*, y^*) + s \nabla_y f(x^*, y^*), \quad (2.3.3)$$

$$0 = f(x^*, y^*) - v(x^*) \quad (2.3.4)$$

$$0 \leq -(G(x^*, y^*), g(x^*, y^*)) \quad (2.3.5)$$

$$0 = \langle r, (G(x^*, y^*), g(x^*, y^*)) \rangle \quad (2.3.6)$$

$$0 \leq r, \quad 0 \leq s, \quad (2.3.7)$$

onde

$$\partial v(x^*) = \text{Conv}\{\nabla_x l(x^*, y, \pi) / y \in O(x^*), \pi \in \prod^2(x^*, y)\}$$

e, para cada  $x \in X$ ,

$$l(x, y, \pi) = f(x, y) + \langle \pi, g(x, y) \rangle, \text{ para } (y, \pi) \in \mathbb{R}^{n_2} \times \mathbb{R}_+^{m_2}.$$

**Prova:**

Segue-se imediatamente da aplicação do Corolário 2.3.4 ao problema (PPDN), levando em consideração a Proposição 1.3.7.

**Exemplo 2.3.7** (Adaptado de Clark-Westerberg [CW1])

Considerar o problema (PPDN) onde, para  $X = [4, 16]$ , as funções estão definidas por

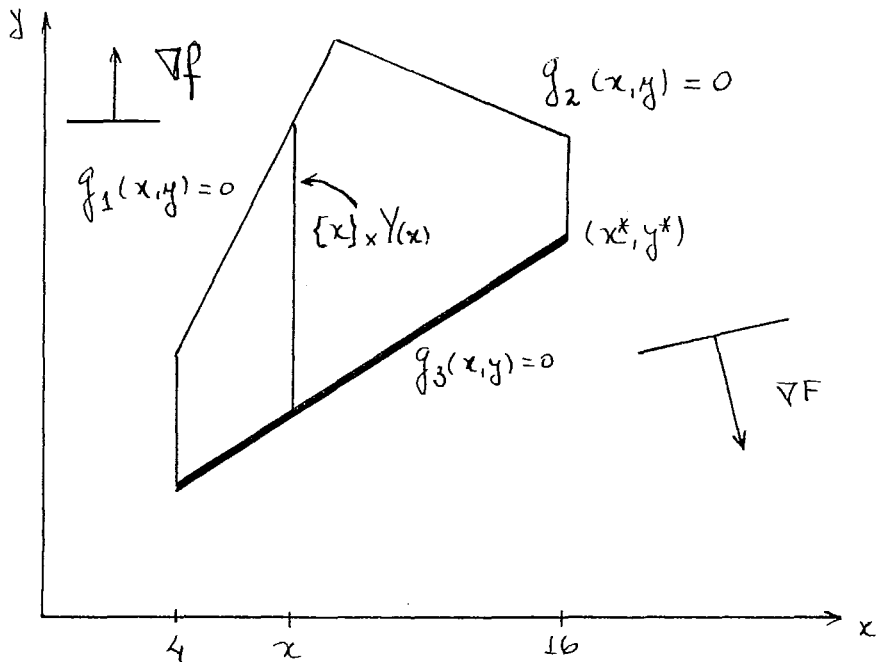
$$F(x, y) = x - 4y, \quad f(x, y) = y,$$

$$g_1(x, y) = -2x + y, \quad g_2(x, y) = 2x + 5y - 108, \quad g_3(x, y) = 2x - 3y + 4.$$

Para este caso, no que  $n_1 = n_2 = 1, m_1 = 0$  e  $m_2 = 3$ , o problema está formulado da seguinte forma

$$\begin{aligned} & \min_{(x,y) \in [4,16] \times \mathbb{R}} && x - 4y \\ & \text{s.a. } \arg \min_y \{y / -2x + y \leq 0, 2x + 5y - 108 \leq 0, 2x - 3y + 4 \leq 0\}. \end{aligned}$$

A geometria deste exemplo está ilustrada na Figura 2.1.



**Figura 2.1 – Geometria do Exemplo 2.3.7**  
**O (·) mostra-se como uma linha grossa.**

$$\begin{aligned} & \text{A solução global é } (x^*, y^*) = (16, 12), \text{ onde } Y(x^*) = [12, 76/5], \prod_1^1(x^*, y^*) \\ & = \prod_2^2(x^*, y^*) = \prod_a^a(x^*, y^*) = \{(0, 0, 1/3)\}, \prod_0^1(x^*, y^*) = \prod_0^2(x^*, y^*) = \{0\}, F(x^*, y^*) \end{aligned}$$

$= -32$ ,  $O(x^*) = \{12\}$  e  $O(x) = \left\{\frac{2x+4}{3}\right\}$ , para  $x \in [4, 16]$ . O  $(\cdot)$  é semicontínua inferiormente em  $x^* = 16$ .

Notar que, desde que  $\{(x, y) \in \mathbb{R}^2 / g_i(x, y) \leq \epsilon, i = 1, 2, 3\}$  é compacto, para todo  $\epsilon \geq 0$ , tem-se que a Hipótese 1.1.1 se satisfaz.

Por outro lado, considerando o problema na forma equivalente (P1) (ver (1.1.5)) e a sua função marginal, dada em (2.2.2), é fácil verificar que, para  $p = (p_1, p_2, p_3, p_4) \in \mathbb{R}^4$ , tem-se que se  $V(p) < +\infty$ , então  $V(p) = F(x^*, y^p)$ , onde  $y^p = (36 + p_4)/3$ . Daqui, um cálculo simples determina que basta considerar  $\epsilon_0$  suficientemente pequeno e  $\Omega = \{x^*\} \times [12 - \epsilon_0/4, 12 + \epsilon_0/4]$ , para garantir que a Hipótese 2.3.1 é válida.

Pode-se ver que o problema é calmo em  $(x^*, y^*)$ . De fato, dado que

$$V(p) - V(0) = F(x^*, y^p) - F(x^*, y^*) = -\frac{4}{3}p_4,$$

tem-se que

$$\liminf_{p \rightarrow 0} \frac{V(p) - V(0)}{\|p\|} = \liminf_{p \rightarrow 0} -\frac{4}{3} \frac{p_4}{\|p\|} > -\infty,$$

de onde, pela Proposição 2.2.2, o problema é calmo em  $(x^*, y^*)$ . Observar que

$$\partial v(x^*) = \left\{ \nabla_x f(x^*, y^*) + \frac{1}{3} \nabla_x g_3(x^*, y^*) \right\} = \left\{ \frac{2}{3} \right\}.$$

Então, em concordância com o Teorema 2.3.6, devem existir  $r_3$  e  $R_+$  e  $s \in \mathbb{R}_+$ , tal que para algum  $\eta \in N_X(x^*) = \mathbb{R}_1$  satisfaz-se que

$$0 = 1 + r_3(2) + s \left( -\frac{2}{3} \right) + \eta$$



$$0 = -4 + r_3(-3) + s(1).$$

Com efeito, para  $r = (0, 0, r_3) \in \mathbb{R}_+^3$ ,  $r_3$  arbitrário,  $s = 4 + 3r_3 \in \mathbb{R}_+$  e  $\eta = 5/3 \in N_X(x^*)$ , o Teorema 2.3.6 se verifica.

## 2.4 – Condições de Otimidade para Alguns Casos Particulares de (PPDN)

Na seção anterior, sob hipóteses de segunda ordem no problema de seguidor e a condição de regularidade de calma sobre o problema (PPDN), estabeleceu-se a condição de  $K - K - T$  como condição necessária de otimidade para este problema (Teorema 2.3.6). Nesta seção se estudam dois casos, que por se apresentarem com freqüência em problemas práticos (ver, por exemplo Outrata [OT2], [OT4]), são de especial interesse:  $O(x) = \{y(x)\}$  e  $(P_x)$  convexo, para  $x \in X$ , fixo.

### 2.4.1 – Conjunto de Reação Racional Unitário

A seguinte propriedade, conseqüência da Proposição 1.3.7, reformula a relação (1.3.6) quando o problema do seguidor tem solução única para um dado  $x_0 \in X$ . Em particular, a igualdade na relação (1.3.6) se satisfaz sem a hipótese de semicontinuidade inferior de  $O(\cdot)$  em  $x_0$ .

**Proposição 2.4.1** Seja  $O(x_0) = \{y_0\}$ ,  $\prod_0^2(x_0, y_0) = \{0\}$  e  $\prod^2(x_0, y_0) \subseteq \overline{\prod^a(x_0, y_0)}$ . Então,  $v$  é regular no sentido de Clarke em  $x_0$  e

$$\partial v(x_0) = \text{Conv}(\{\nabla_x l(x_0, y_0, \pi)/\pi \in \prod^2(x_0, y_0)\}).$$

**Prova:**

Usando o Teorema 3, Rockafellar [RO4], obtem-se que

$$\max_{\pi \in \prod^2(x_0, y_0)} \langle \nabla_x l(x_0, y_0, \pi), p \rangle = v'(x_0; p) \leq v^0(x_0; p), \forall p.$$

Daqui e do Teorema 5, Rockafellar [RO4], conclui-se que

$$\max_{\pi \in \prod^2(x_0, y_0)} \langle \nabla_x l(x_0, y_0, \pi), p \rangle \leq v'(x_0; p) \leq \max_{\pi \in \prod^2(x_0, y_0)} \langle \nabla_x l(x_0, y_0, \pi), p \rangle, \forall p.$$

Portanto,

$$v'(x_0; p) = v^0(x_0; p) = \max_{\pi \in \prod^2(x_0, y_0)} \langle \nabla_x l(x_0, y_0, \pi), p \rangle, \forall p.$$

Assim, tem-se que  $v$  é regular no sentido de Clarke em  $x_0$ .

Considerar, agora,  $\xi \in \text{Conv}(\{\nabla_x l(x_0, y_0, \pi)/\pi \in \prod^2(x_0, y_0)\})$ . Cumpre-se que existe  $k \in \mathbb{N}$ , tal que

$$\langle \xi, p \rangle = \left\langle \sum_{i=1}^k \lambda_i \nabla_x l(x_0, y_0, \pi^i), p \right\rangle, \forall p,$$

onde

$$\sum_{i=1}^k \lambda_i = 1, \lambda_i \geq 0, \pi^i \in \prod^2(x_0, y_0), \forall i.$$

Daqui,

$$\langle \xi, p \rangle = \sum_{i=1}^k \lambda_i \langle \nabla_x l(x_0, y_0, \pi^i), p \rangle \leq \sum_{i=1}^k \lambda_i v^0(x_0; p) = v^0(x_0; p), \forall p.$$

Portanto,  $\xi \in \partial v(x_0)$  e

$$\text{Conv}(\{\nabla_x l(x_0, y_0, \pi) / \pi \in \prod^2(x_0, y_0)\}) \subseteq \partial v(x_0). \quad (2.4.1)$$

Agora, usando (1.3.6) e (2.4.1), obtem-se o resultado procurado.

**Proposição 2.4.2** Seja o problema (PPDN) calmo na solução local  $(x^*, y^*)$ . Se  $O(x^*) = \{y^*\}$ ,  $\prod_0^2(x^*, y^*) = \{0\}$  e  $\prod^2(x^*, y^*) \subseteq \prod^a(x^*, y^*)$ , então existe  $s \in \mathbb{R}$  e  $r \in \mathbb{R}^{m_1+m_2}$  que satisfazem a condição de  $K - K - T$ , dada pelas relações (2.3.2) - (2.3.7), onde

$$\partial v(x^*) = \text{Conv}(\{\nabla_x l(x^*, y^*, \pi) / \pi \in \prod^2(x^*, y^*)\}).$$

**Prova:**

Imediata, da prova do Teorema 2.3.6, considerando a Proposição 2.4.1.

**Corolário 2.4.3** Seja  $O(x_0) = \{y_0\}$ ,  $\prod_0^2(x_0, y_0) = \{0\}$ ,  $\prod^2(x_0, y_0) = \{\pi_0\}$  e  $\prod^a(x_0, y_0) \neq \emptyset$ . Então  $v(\cdot)$  é estritamente diferenciável em  $x_0$  e

$$\nabla v(x_0) = \nabla_x l(x_0, y_0, \pi_0) = \nabla_x f(x_0, y_0) + \sum_i (\pi_0)_i \nabla_x g_i(x_0, y_0). \quad (2.4.2)$$

**Prova:**

Desde  $\overline{\prod^a(x_0, y_0)} \neq \emptyset$  que  $\prod^2(x_0, y_0) \subseteq \prod^a(x_0, y_0) = \{\pi_0\}$ , conclui-se que  $\prod^a(x_0, y_0) = \{\pi_0\} = \prod^2(x_0, y_0)$ . Então as hipóteses da Proposição 2.4.1 se satisfazem e  $\partial v(x_0) = \{\nabla_x l(x_0, y_0, \pi_0)\}$ .

Dado que, da Proposição 1.3.5,  $v$  é Lipschitz próxima de  $x_0$ , tem-se que  $v$  é estritamente diferenciável e

$$\nabla v(x_0) = \nabla_x l(x_0, y_0, \pi_0).$$

**Observação 2.4.4** Notar que, desde que  $\prod^2(x_0, y_0) \subseteq \prod^1(x_0, y_0)$ , sob as hipóteses mais fortes  $\prod^1(x_0, y_0) = \{\pi_0\}$  (por exemplo, quando  $\{\nabla_y g_i(x_0, y_0)\}_{i \in I(x_0, y_0)}$  é linearmente independente) e  $\prod^a(x_0, y_0) \neq \emptyset$  o Corolário 2.4.3 é válido, sendo as conclusões compatíveis com os resultados que obtem Outrata [OT2].

**Corolário 2.4.5** Seja o problema (PPDN) calmo na solução local  $(x^*, y^*)$ . Se  $O(x^*) = \{y^*\}$ ,  $\prod_0^2(x^*, y^*) = \{0\}$ ,  $\prod^2(x^*, y^*) = \{\pi^*\}$  e  $\prod^a(x^*, y^*) \neq \emptyset$ , então existe  $s \in \mathbb{R}$  e  $r \in \mathbb{R}^{m_1+m_2}$ , tal que

$$0 \in \nabla_x F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_x G_i(x^*, y^*) + \sum_{i=1}^{m_2} (r_{m_1+i} - s\pi_i^*) \nabla_x g_i(x^*, y^*) + N_X(x^*), \quad (2.4.3)$$

$$0 = \nabla_y F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_y G_i(x^*, y^*) + \sum_{i=1}^{m_2} (r_{m_1+i} - s\pi_i^*) \nabla_y g_i(x^*, y^*), \quad (2.4.4)$$

$$0 = f(x^*, y^*) - v(x^*), \quad (2.4.5)$$

$$0 \leq -(G(x^*, y^*), g(x^*, y^*)), \quad (2.4.6)$$

$$0 \leq \langle r, (G(x^*, y^*), g(x^*, y^*)) \rangle, \quad (2.4.7)$$

$$0 \leq r, \quad 0 \leq s \quad (2.4.8)$$

**Prova:**

Desde que as hipóteses do Corolário 2.4.3 se satisfazem, usando a Observação 6.1.3 -(iv), Clarke [CL1], resultam as mesmas conclusões que as do Teorema 2.3.6.

De fato, na relação (2.3.2),  $\nabla_x f(x^*, y^*) - \partial v(x_i^*)$  pode ser substituído por  $\nabla_x f(x^*, y^*) - \nabla v(x^*)$ . Então, usando (2.4.2), obtém-se (2.4.3) - (2.4.8).

**Observação 2.4.6** É claro que quando  $X = \mathbb{R}^{n_1}$ , a relação (2.4.3) converte-se na igualdade

$$0 = \nabla_x F(x^*, y^*) + \sum_{i=1}^{m_2} r_i \nabla_x G_i(x^*, y^*) + \sum_{i=1}^{m_2} (r_{m_1+i} - s\pi_i^*) \nabla_x g_i(x^*, y^*). \quad (2.4.9)$$

Sem perda de generalidade, considera-se o seguinte teorema para o problema (PPDN) sem restrições do tipo  $G(x, y)$ , i.e.,  $m_1 = 0$ .

**Teorema 2.4.7** Seja o problema (PPDN) calmo na solução local  $(x^*, y^*)$ . Se  $O(x^*) = \{y^*\}$ ,  $\prod_0^2(x^*, y^*) = \{0\}$  e  $\prod^2(x^*, y^*) \subseteq \prod^a(x^*, y^*)$ , então existe  $\eta \in N_X(x^*)$  tal que não há solução  $z = (z^x, z^y) \in \mathbb{R}^{n_1+n_2}$  para o seguinte sistema de desigualdades

$$\langle z, \nabla F(x^*, y^*) \rangle < 0, \quad (2.4.10)$$

$$\langle z, \nabla g_i(x^*, y^*) \rangle = 0, \quad \forall i \in I(x^*, y^*) \quad (2.4.11)$$

$$\langle z^x, \eta \rangle \leq 0. \quad (2.4.12)$$

**Prova:**

Usa-se, nesta prova, a seguinte notação

$$v^* = v(x^*), f^* = f(x^*, y^*), g_i^* = g_i(x^*, y^*), F^* = F(x^*, y^*), \nabla f^* = \nabla f(x^*, y^*),$$

$$\nabla g_i^* = \nabla g_i(x^*, y^*), \nabla F^* = \nabla F(x^*, y^*) \text{ e } \partial v^* = \partial v(x^*).$$

Da Proposição 2.4.1, tem-se

$$\partial v^* = \text{Conv}(\{\nabla_x f^* + \sum_i \pi_i \nabla_x g_i^* / \pi = (\pi_1, \dots, \pi_{m_2}) \in \prod^2(x^*, y^*)\}).$$

Daqui

$$\nabla_x f^* - \partial v^* = -\text{Conv}(\{\sum_i \pi_i \nabla_x g_i^* / \pi \in \prod^2(x^*, y^*)\}).$$

Da Proposição 2.4.2, sabe-se que as condições do Teorema 2.3.6 se satisfazem. Então, existe  $(\eta, r, s) \in N_X(x^*) \times \mathbb{R}^{m_2} \times \mathbb{R}$ , tal que para algum  $k \in \mathbb{N}$ , tem-se

$$0 = \nabla_x F^* + \sum_{i=1}^{m_2} (r_i - s \sum_{j=1}^k \beta_j \pi_i^j) \nabla_x g_i^* + \eta, \quad (2.4.13)$$

$$0 = \nabla_y F^* + \sum_{i=1}^{m_2} (r_i - s \sum_{j=1}^k \beta_j \pi_i^j) \nabla_y g_i^*, \quad (2.4.14)$$

$$0 = f^* - v^*, \quad (2.4.15)$$

$$0 \leq -g^*, \quad (2.4.16)$$

$$0 \leq \langle r, g^* \rangle, \quad (2.4.17)$$

$$0 \leq r, \quad 0 \leq s, \quad (2.4.18)$$

para alguns  $\beta_j \geq 0, \forall j, \sum_{j=1}^k \beta_j = 1$  e  $\pi^j = (\pi_1^j, \dots, \pi_{m_2}^j) \in \prod^2(x^*, y^*)$ , para  $j = 1, \dots, k$  (notar que em (2.4.14) se tem usado o fato de que  $0 = \nabla_y l(x^*, y^*, \pi^j) = \nabla_y f^* + \sum_i \pi_i^j \nabla_y g_i^*, \forall j$ ).

As relações (2.4.13) - (2.4.14) podem ser escritas como

$$-\begin{pmatrix} \nabla_x F^* \\ \nabla_y F^* \end{pmatrix} = \begin{pmatrix} \nabla_x g_1^* & \cdots & \nabla_x g_{m_2}^* & \eta \\ \nabla_y g_1^* & \cdots & \nabla_y g_{m_2}^* & 0 \end{pmatrix} \begin{pmatrix} r - s \sum_{j=1}^k \beta_j \pi^j \\ e \end{pmatrix},$$

onde  $(r, s) \geq 0$ ,  $e^T = (1, \dots, 1) \in \mathbb{R}^{n_1}$ .

Agora, denotando

$$\nabla_x g^* = (\nabla_x g_1^*, \dots, \nabla_x g_{m_2}^*)$$

$$\nabla_y g^* = (\nabla_y g_1^*, \dots, \nabla_y g_{m_2}^*)$$

$$q = r - s \sum_{j=1}^k \beta_j \pi^j$$

e considerando

$$q_i = q_i^+ - q_i^-, \forall i,$$

onde  $a^+ := \max\{a, 0\}$  e  $a^- := \max\{-a, 0\}$ ,  $\forall a \in \mathbb{R}$ , a igualdade anterior pode ser escrita como

$$-\begin{pmatrix} \nabla_x F^* \\ \nabla_y F^* \end{pmatrix} = \begin{pmatrix} \nabla_x g^* & -\nabla_x g^* & \eta \\ \nabla_y g^* & -\nabla_y g^* & 0 \end{pmatrix} \begin{pmatrix} q^+ \\ q^- \\ e \end{pmatrix}, \quad (2.4.19)$$

onde  $q^+ \geq 0$ ,  $q^- \geq 0$ ,  $e > 0$ .

Observar que, como  $\pi^j \in \prod_1^2(x^*, y^*) \subseteq \prod_1^1(x^*, y^*)$ , cumpre-se que  $\langle \pi^j, g^* \rangle = 0$ , para  $j = 1, \dots, k$ . Assim, usando (2.4.17), tem-se que

$$q_i = 0, \quad \text{para } i \notin I(x^*, y^*).$$



Portanto, denotando por  $\nabla_x g_I^*$ ,  $\nabla_y g_I^*$  e  $q_I$  as submatrizes de  $\nabla_x g^*$ ,  $\nabla_y g^*$  e  $q$ , respectivamente, correspondentes aos índices ativos, o sistema (2.4.19) pode ser expresso como

$$-\begin{pmatrix} \nabla_x F^* \\ \nabla_y F^* \end{pmatrix} = \begin{pmatrix} \nabla_x g_I^* & -\nabla_x g_I^* & \eta \\ \nabla_y g_I^* & -\nabla_y g_I^* & 0 \end{pmatrix} \begin{pmatrix} q_I^+ \\ q_I^- \\ e \end{pmatrix}, \quad (2.4.20)$$

onde  $q_I^+ \geq 0$ ,  $q_I^- \geq 0$ , e  $e > 0$ .

Usando o Lema de Farkas no sistema (2.4.20), tem-se que existe  $\eta \in N_X(x^*)$ , tal que não há solução para

$$\langle (z^x, z^y), -(\nabla_x F^*, \nabla_y F^*) \rangle > 0,$$

$$\begin{pmatrix} (\nabla_x g_I^*)^T & (\nabla_y g_I^*)^T \\ -(\nabla_y g_I^*)^T & -(\nabla_x g_I^*)^T \\ \eta^T & 0 \end{pmatrix} \begin{pmatrix} z^x \\ z^y \end{pmatrix} \leq 0. \quad (2.4.21)$$

Isto é, existe  $\eta \in N_X(x^*)$ , tal que não há solução para

$$\begin{aligned} \langle z, \nabla F(x^*, y^*) \rangle &< 0, \\ \langle z, \nabla g_i(x^*, y^*) \rangle &= 0, \forall i \in I(x^*, y^*) \\ \langle z^*, \eta \rangle &\leq 0. \end{aligned}$$

O Teorema 2.4.7 é um resultado análogo àquele conhecido para programas matemáticos de um nível só, onde se caracteriza o cone de direções viáveis de descida. As relações (2.4.11) e (2.4.12) caracterizam o cone de direções viáveis para o programa (PPDN), sendo que, as relações (2.4.10) - (2.4.12) caracterizam o cone de

*direções viáveis de descida*. O Teorema 2.4.7 também pode ser interpretado da seguinte forma: Se  $(x^*, y^*)$  é uma solução local para (PPDN), sob a condição de qualificação indicada, não existe nenhuma direção de descida para o l'tal que todas as restrições ativas para o seguidor se comportem como igualdade.

No seguinte exemplo, ilustramos o Teorema 2.4.7 para o caso do Exemplo 2.3.7.

**Exemplo 2.4.8** Considerar o programa (PPDN) como no Exemplo (2.3.7). Então, das relações (2.4.11) e (2.4.12), temos que o cone  $K$  de direções viáveis no ponto  $(x^*, y^*) = (16, 12)$  está determinado pelo conjunto de vetores  $z = (z^x, z^y) \in \mathbb{R}^2$ , que satisfazem o sistema

$$\begin{aligned} 2z^x - 3z^y &= 0, \\ z^x \eta &\leq 0, \\ \eta &\in N_X(x^*) = \mathbb{R}^+ \end{aligned}$$

Então,

$$K = \{\alpha(3, 2) / \alpha \in \mathbb{R}^-\}.$$

Observar que o cone  $\tilde{K}$  de direções de descida está determinado, a partir de (2.4.10)-(2.4.12), pelo vetores  $z = (z^x, z^y) \in \mathbb{R}^2$  que satisfazem

$$z^x - 4z^y < 0 \quad , \quad (z^x, z^y) \in K,$$

i.e.,  $\tilde{K} = \emptyset$ .

A figura 2.2 mostra o cone  $K$  para o exemplo anterior.

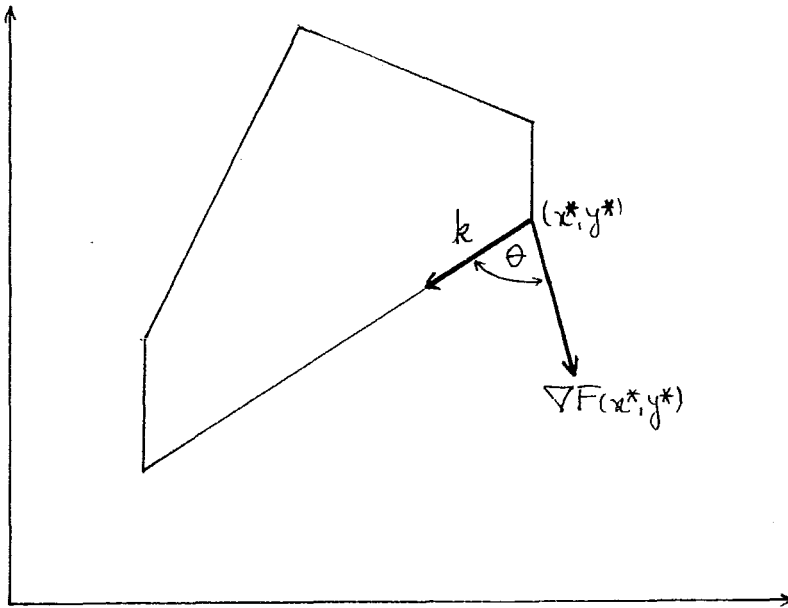


Figura 2.2: O cone de direções viáveis,  $K$ , gerado pelo vetor  $k$ , faz um ângulo agudo,  $\theta$ , com o  $\nabla F(x^*, y^*)$ , i.e., não existem direções viáveis de descida para  $F$ .

## 2.4.2 – O Problema do Seguidor Convexo

Dir-se-á que o problema  $(P_x)$  é convexo em  $x \in X$ , se  $f(x, \cdot)$  e  $g(x, \cdot)$  são funções convexas.

**Teorema 2.4.8** Seja o problema (PPDN) calmo na solução local  $(x^*, y^*)$ , tal que  $(P_{x^*})$  é convexo. Se  $O(x^*) = \{y^*\}$  e a CQMFS se satisfaz para  $(P_{x^*})$  no ponto  $y^*$ , então existe  $(r, s, \pi) \in \mathbb{R}^{m_1+m_2} \times \mathbb{R} \times \mathbb{R}^{m_1}$ , tal que

$$0 \in \nabla_x F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_x G_i(x^*, y^*) + \sum_{i=1}^{m_2} (r_{m_1+i} - s\pi_i) \nabla_x g_i(x^*, y^*) + N_X(x^*) \quad (2.4.22)$$

$$0 = \nabla_y F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_y G_i(x^*, y^*) + \sum_{i=1}^{m_2} (r_{m_1+i} - s\pi_i) \nabla_y g_i(x^*, y^*), \quad (2.4.23)$$

$$0 \leq -(G(x^*, y^*), g(x^*, y^*)), \quad (2.3.24)$$

$$0 = \langle r, (G(x^*, y^*), g(x^*, y^*)) \rangle, \quad (2.4.25)$$

$$0 = \langle \pi, g(x^*, y^*) \rangle, \quad (2.4.26)$$

$$0 \leq r, \quad 0 \leq s, \quad (2.4.27)$$

$$0 \leq \pi. \quad (2.4.28)$$

**Prova:**

Desde que  $(P_{x^*})$  satisfaz a CQMFS no ponto  $y^*$ , tem-se que  $\prod_1(x^*, y^*) = \{\pi^*\}$  e, da convexidade (Rockafellar [RO4]) que  $\{\pi^*\} = \prod_1(x^*, y^*) = \prod_2(x^*, y^*) = \prod_a(x^*, y^*)$ ,  $\prod_1(x^*, y^*) = \prod_2(x^*, y^*)$ . Portanto, da Proposição 1.3.1, verificam-se as hipóteses do Corolário 2.4.5.

Da convexidade de  $(P_{x^*})$ , obtém-se que as relações  $f(x^*, y^*) = v(x^*)$  e  $g(x^*, y^*) \leq 0$  são equivalentes à  $K - K - T$ . Daqui, usando  $\prod_1(x^*, y^*) = \{\pi^*\}$ , conclui-se que

$$\nabla_y f(x^*, y^*) + \sum_{i=1}^{m_2} \pi_i^* \nabla_y g_i(x^*, y^*) = 0,$$

$$\langle \pi^*, g(x^*, y^*) \rangle = 0, g(x^*, y^*) \leq 0, \pi^* \geq 0.$$

Agora, substituindo as primeiras relações por estas últimas em (2.4.3)-(2.4.8), obtém-se (2.4.21)-(2.4.28).

**Corolário 2.4.9** Seja o problema (PPDN) calmo na solução local  $(x^*, y^*)$ , tal que  $(P_{x^*})$  é convexo. Se as hipóteses do Teorema 2.4.8 se satisfazem, então existe  $(r, c) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ , tal que

$$0 \in \nabla_x F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_x G_i(x^*, y^*) + \sum_{i=1}^{m_2} c_i \nabla_x g_i(x^*, y^*) + N_X(x^*), \quad (2.4.29)$$

$$0 = \nabla_y F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_y G_i(x^*, y^*) + \sum_{i=1}^{m_2} c_i \nabla_y g_i(x^*, y^*),$$

$$0 \leq -(G(x^*, y^*), g(x^*, y^*)),$$

$$0 = \langle r, (G(x^*, y^*)) \rangle,$$

$$0 = \langle c_i, g_i(x^*, y^*) \rangle, \quad i = 1, \dots, m_2$$

$$0 \leq r.$$

Quando  $X = \mathbb{R}^{n_1}$ , a relação (2.4.29) é

$$0 = \nabla_x F(x^*, y^*) + \sum_{i=1}^{m_1} r_i \nabla_x G_i(x^*, y^*) + \sum_{i=1}^{m_2} c_i \nabla_x g_i(x^*, y^*). \quad (2.4.30)$$

**Prova:**

Imediata. Basta aplicar o Teorema 2.4.8 com

$$c_i := r_{i+m_1} - s\pi_i, \quad i = 1, \dots, m_2.$$

**Teorema 2.4.10** Seja o problema (PPDN) calmo na solução local  $(x^*, y^*)$ , tal que  $(P_x^*)$  é convexo. Se a CQMFS se satisfaz para todo  $y \in O(x^*)$ , então existe  $(\nu, s) \in \mathbb{R}^{m_1+m_2} \times \mathbb{R}$  que satisfaz as conclusões do Teorema 2.3.6, onde

$$\partial v(x^*) = \text{Conv}(\{\nabla_x l(x^*, y, \pi^*) / y \in O(x^*)\})$$

e  $\pi^*$  é o único multiplicador associado a qualquer  $y \in O(x^*)$ .

**Prova:**

Desde que  $(P_x^*)$  é convexo e a CQMFS é satisfeita para todo  $y \in O(x^*)$ , cumpre-se que  $\prod_0^2(x^*, y^*) = \{0\}$  e

$$\prod^2(x^*, y) = \prod^a(x^*, y) = \prod(x^*) = \{\pi^*\}, \forall y \in O(x^*).$$

Agora, usando a Proposição 2.1.1, a Observação 2.1.2-(i) e o Teorema 1.4.3, tem-se o resultado procurado.

## Capítulo 3

### Uma Solução Numérica do Problema (PPDN)

Neste capítulo apresenta-se um algoritmo de descida para o problema (PPDN). Como já tem-se discutido, o problema (PPDN) é um problema com restrições, em geral, não diferenciável e não convexo. Desde que os códigos existentes, que tratam com problemas não diferenciáveis, estão formulados para problemas sem restrições ou com restrições lineares, tem sido conveniente, aqui, tratar o problema com técnicas não diferenciáveis baseadas em Kiwiel [KI1], que não fazendo uso de penalização formulam o problema original como um problema equivalente irrestrito.

É comum para os algoritmos que tratam com não diferenciabilidade, fazerem a hipótese de possuir um *oráculo* com a capacidade de fornecer duas informações fundamentais a respeito da função a ser minimizada em qualquer ponto do seu domínio: o valor dela e um elemento do gradiente generalizado (um subgradiente). Isto é, assume-se que se tem a capacidade de avaliar a função e conhecer um subgradiente da função. Para o problema (PPDN), na sua versão (P1), a dificuldade que encontrará o *oráculo* será lidar com a função marginal do seguidor. Este problema é resolvido fazendo uso da Proposição 1.3.7, onde apresentou-se uma estimativa de  $\partial v(x)$ . Sob as hipóteses feitas, o trabalho do *oráculo* se reduz a calcular, para  $x_0 \in X$ , um  $y_0 \in O(x_0)$ , um  $\pi_0 \in \prod_1(x_0, y_0)$  e  $\nabla_x l(x_0, y_0, \pi_0)$ . Consegue-se isto, fazendo uso da rotina NLPQL1 (ver Schittkowski [SCH1]), que é um código FORTRAN que implementa o Método Seqüencial Quadrático para resolver programas não lineares com restrições.

Por outro lado, o algoritmo aqui formulado usa a rotina BTNC (ver Schramm [SCHR1], Schramm-Zowe [SCHRZ1], Outrata-Schramm-Zowe [OSCHRZ1] e Arica-Scheimberg [ASCH1]), que é um código FORTRAN para minimização não diferenciável



e não convexa sem restrições. A rotina BTNC, gentilmente cedida pelo Prof. Jiri Outrata, é uma implementação do algoritmo *BT* (*B* de *bundle* e *T* de *trust region*), onde a busca linear do algoritmo de Feixe (ver Hiriart-Urruty-Lemaréchal [H-ULE1]) é substituída por uma versão de região de confiança.

Um tratamento numérico do problema (PPDN), que também faz uso do código BTNC, é proposto por Outrata [OT2], [OT4]; no entanto, de maneira diferente à aqui apresentada. Usam-se nesses trabalhos técnicas de penalização para converter o problema original (restrito), em um problema não restrito. A penalização, sob a hipótese de *calma* para o problema (PPDN), resulta exata, mas gera pontos que sendo aproximações de uma solução local, são inviáveis. O algoritmo apresentado neste capítulo é um algoritmo de descida viável, fornecendo aproximações viáveis de uma solução local.

### 3.1 – Hipóteses para uma Proposta Numérica

Nesta seção propõem-se as hipóteses sob as quais trabalha o algoritmo formulado neste capítulo. Prova-se como resultado principal que, sob essas hipóteses, a função marginal do seguidor é *semisuave fracamente* (Definição 3.1.3) que é uma exigência fundamental para a convergência do algoritmo BTNC (ver Schramm-Zowe [SCHRZ1]).

#### Hipótese 3.1.1 Considerar

$$S' := \{(x, y) \in D \times \mathbb{R}^{n_2} / g(x, y) \leq 0\}.$$

Cumpra-se que

- (i)  $D$  é convexo e a função  $g(\cdot, \cdot)$  é convexa em  $D \times \mathbb{R}^{n_2}$ ;
- (ii) para cada  $x \in D$ , a função  $f(x, \cdot)$  é convexa;

- (iii) existe  $\epsilon \geq 0$  tal que  $Y(x) \neq \emptyset$ , para todo  $x$  da  $\epsilon$ -vizinhança do conjunto  $Proj_{\mathbb{R}^{n_1}}(S')$  (onde  $Proj_{\mathbb{R}^{n_1}}$  é o operador projeção sobre  $\mathbb{R}^{n_1}$ );
- (iv) para cada  $x \in D$ , a CQMFS é válida para  $(P_x)$  em cada  $y \in O(x)$ .

**Observação 3.1.2** Sob a Hipótese 3.1.1, tem-se, da Hipótese 1.1.1, Proposição 1.3.5, Proposição 1.3.7, Teorema 1.4.1 e do Teorema 1.4.3, que

- (i) o conjunto  $Proj_{\mathbb{R}^{n_1}}(S')$  é convexo;
- (ii) os teoremas dos capítulos anteriores são válidos se se substitui  $X$  por qualquer conjunto fechado  $E$  contido em  $D$ ;
- (iii)  $v(\cdot)$  é Lipschitz localmente em  $E$ ;
- (iv)  $v(\cdot)$  é direcionalmente diferenciável para cada  $x \in E$  e

$$v'(x; p) = \min_{y \in O(x)} \langle \nabla_x l(x, y, \pi_x), p \rangle, \forall p \in \mathbb{R}^{n_1},$$

onde  $\pi_x$  é o único multiplicador de  $K - K - T$  associado a qualquer  $y \in O(x)$ ;

- (v) para cada  $x \in E$ ,

$$v^0(x; p) = \max_{y \in O(x)} \langle \nabla_x l(x, y, \pi_x), p \rangle, \forall x \in X, \forall p \in \mathbb{R}^{n_1}.$$

**Definição 3.1.3** (Mifflin [MF1]) Seja  $C \subset \mathbb{R}^n$  aberto. A função  $h : C \rightarrow \mathbb{R}$  é *semisuave fracamente* em  $x_0 \in C$  se

- (i)  $h$  é Lipschitz próxima de  $x_0$ ; e

(ii) dado  $d \in \mathbb{R}^n$ , existe  $h'(x_0; d)$  e, para quaisquer seqüências  $\{t_k\} \subset \mathbb{R}_+$  e  $\{\xi^k\} \subset \mathbb{R}^n$ , tais que  $t_k \downarrow 0$ ,  $x_0 + t_k d \in C$ ,  $\xi^k \in \partial h(x_0 + t_k d)$ , cumpre-se que

$$h'(x_0; d) = \lim_{t_k \downarrow 0} \langle \xi^k, d \rangle .$$

Diz-se que  $h$  é semisuave fracamente em  $B \subset C$ , se é semisuave fracamente em cada  $x \in B$ .

Da Proposição 3, Proposição 4 e Teorema 6, Mifflin [MF2], segue-se o seguinte resultado:

#### Proposição 3.1.4

Seja  $h : \mathbb{R}^n \rightarrow \mathbb{R}$ . Cumpre-se que

- (i) Se  $h$  é convexa ou continuamente diferenciável, então é semisuave fracamente.
- (ii) Se

$$h(x) = \max_{1 \leq i \leq m} h_i(x) \quad , \quad x \in \mathbb{R}^n ,$$

onde as funções  $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$  são semisuaves fracamente em  $C \subset \mathbb{R}^n$ , para todo  $i$ , então  $h$  é semisuave fracamente em  $C$ .

O seguinte teorema, resultado principal desta seção, estabelece que, sob as hipóteses feitas, a função marginal do seguidor é semisuave fracamente.

**Teorema 3.1.5** Sob a Hipótese 3.1.1,  $v(\cdot)$  é semisuave fracamente em  $X$ .

**Prova:** (Esta prova está inspirada na demonstração do Teorema 2, Mifflin [MF2]) Da

Observação 3.1.2 (iii) ,  $v(\cdot)$  é Lipschitz localmente em  $X$ , pelo que a Definição 3.1.3 (i) é satisfeita. Vejamos que a Definição 3.1.3 (ii) se cumpre para  $v(\cdot)$  em qualquer  $x \in X$ .

Seja  $x \in X, p \in \mathbb{R}^{n_1}, \{t_j\} \subset \mathbb{R}_+, x^j := x + t_j p$  e  $\xi^j \in \partial v(x^j)$ , onde  $t_j \downarrow 0$ .

Notar que, do Teorema 1.4.3, cumpre-se que

$$\partial v(x) = \text{Conv}(\{\nabla_x l(x, y, \pi_x) / y \in O(x)\}), \quad (3.1.1)$$

onde

$$\{\pi_x\} = \prod(x) = \prod^1(x, y) = \prod^2(x, y) = \prod^a(x, y).$$

Satisfaz-se, da relação (3.1.1), que

$$\limsup_{j \rightarrow \infty} \langle \xi^j, d \rangle \leq v^0(x; d), \forall d \in \mathbb{R}^{n_1}. \quad (3.1.2)$$

Então, para  $d = -p$ , usando a Observação 3.1.2 (iv) - (v), tem-se que

$$\limsup_{j \rightarrow \infty} \langle \xi^j, -p \rangle \leq v^0(x; -p) = -v'(x; p).$$

Portanto,

$$\liminf_{j \rightarrow \infty} \langle \xi^j, p \rangle \geq v'(x; p). \quad (3.1.3)$$

Supor que

$$\limsup_{j \rightarrow \infty} \langle \xi^j, p \rangle > v'(x; p). \quad (3.1.4)$$

Então, existe uma subsequência de  $\{\xi^j\}_{j=1}^\infty$ , que também será denotada por  $\{\xi^j\}_{j=1}^\infty$ , tal que

$$\lim_{j \rightarrow \infty} \langle \xi^j, p \rangle = v'(x; p) + \epsilon, \quad (3.1.5)$$

para algum  $\epsilon > 0$ .

Para cada  $j$ , escolher  $y^j \in O(x^j)$  e  $\bar{\xi}^j = \nabla_x l(x^j, y^j, \pi^j) \in \partial v(x^j)$  tal que

$$v^0(x^j; p) = \langle \bar{\xi}^j, p \rangle = \max_{y \in O(x^j)} \langle \nabla_x l(x^j, y^j, \pi^j), p \rangle \geq \langle \xi^j, p \rangle, \quad (3.1.6)$$

onde  $\pi^j := \pi_{x^j}$ .

Seja  $\delta > 0$ ,  $\alpha > v(x)$  e

$$Y_{\delta, \alpha}(x) := \{(u, w) \in D \times \mathbb{R}^{n_2} / \|u - x\| \leq \delta, f(u, w) \leq \alpha, g(u, w) \leq \delta\}.$$

Da Hipótese 1.1.1, tem-se que  $Y_{\delta, \alpha}(x)$  é compacto.

Considerar  $y^j \in O(x^j)$ . Então, para  $j$  suficientemente grande, satisfaz-se que

$$\|x^j - x\| \leq \delta, f(x^j, y^j) = v(x^j) < \alpha, g(x^j, y^j) \leq 0 \leq \delta.$$

Daqui,

$$\{x^j\} \times O(x^j) \subset Y_{\delta, \alpha}(x),$$

para  $j$  suficientemente grande.

Portanto, desde que  $Y_{\delta, \alpha}(x)$  é compacto, existe  $\bar{y}$  e uma subsequência  $\{y^{j_k}\} \subset \{y^j\}$ , tal que

$$y^{j_k} \longrightarrow \bar{y}, \quad \text{quando } k \longrightarrow \infty, \quad \text{e } (x, \bar{y}) \in Y_{\delta, \alpha}(x).$$

Por outro lado, da seção 2, Rockafellar [RO4],  $\prod^1(\cdot, \cdot)$  é semicontínua superiormente e limitada localmente, i.e., existe um conjunto limitado  $A \subset \mathbb{R}^m$  tal que  $\prod^1(x', y') \subset A$ , para  $(x', y')$  suficientemente próximo de  $(x, y)$ . Daqui,  $\{\pi^{j_k}\} \subset \bar{A}$ , para  $k$  suficientemente grande. Dado que  $\bar{A}$  é compacto, existe  $\bar{\pi}$  e uma subsequência  $\{\pi^{j_{\tilde{k}}}\} \subset \{\pi^{j_k}\}$  tais que

$$\pi^{j_{\tilde{k}}} \longrightarrow \bar{\pi}, \quad \text{quando} \quad \tilde{k} \longrightarrow \infty.$$

(é claro que  $y^{j_{\tilde{k}}} \longrightarrow \bar{y}$ ).

Desde que  $(P_x)$  é convexo e CQMFS se satisfaz para todo  $y \in O(x)$ , lembrando que, neste caso,  $\prod_0^1(x, y)$  é equivalente a condição de qualificação Slater e usando o Teorema 10 e o Teorema 12 de Hogan [HO1], tem-se que  $Y(\cdot)$  é contínua em  $x$ . Assim, do Teorema 8 de Hogan [HO1], cumpre-se que  $O(\cdot)$  é semicontínua superiormente em  $x$ .

Agora, usando a semicontinuidade superior de  $O(\cdot)$  e  $\prod^1(\cdot, \cdot)$  em  $x$  e  $(x, \bar{y})$ , respectivamente, tem-se, de

$$(x^{j_{\tilde{k}}}, y^{j_{\tilde{k}}}, \pi^{j_{\tilde{k}}}) \quad (x, \bar{y}, \bar{\pi}) \quad \text{e} \quad \pi^{j_{\tilde{k}}} \in \prod^1(x^{j_{\tilde{k}}}, y^{j_{\tilde{k}}}),$$

que

$$\bar{y} \in O(x) \quad \text{e} \quad \bar{\pi} \in \prod^1(x, \bar{y}) = \prod(x) = \{\pi_x\}.$$

Assim,

$$\bar{y} \in O(x) \quad \text{e} \quad \bar{\pi} = \pi_x.$$

Desde que  $\nabla_x l(\cdot, \cdot, \cdot)$  é contínua em  $D \times \mathbb{R}^{n_2} \times \mathbb{R}_+^m$ , das relações (3.1.5) e (3.1.6), tem-se que

$$\langle \underbrace{\nabla_x l(x, \bar{y}, \pi_x)}_{\bar{\xi}}, p \rangle = \langle \bar{\xi}, p \rangle \geq v'(x; p) + \epsilon.$$

Seja  $y^* \in O(x)$  tal que

$$v'(x; p) = \min_{y \in O(x)} \langle \nabla_x l(x, y, \pi_x), p \rangle = \langle \nabla_x l(x, y^*, \pi_x), p \rangle.$$

Então

$$\langle \nabla_x l(x, \bar{y}, \pi_x), p \rangle \geq \langle \nabla_x l(x, y^*, \pi_x), p \rangle + \epsilon$$

e, desde que  $\langle \nabla_x l(\cdot, \cdot, \cdot), \cdot \rangle$  é contínua, existem vizinhanças  $\mathcal{B}(x)$ ,  $\mathcal{V}(\bar{y})$ ,  $\mathcal{P}(\pi_x)$  e  $\mathcal{U}(p)$  tais que

$$\begin{aligned} \langle \nabla_x l(z, y, \pi_x), q \rangle &\geq \langle \nabla_x l(z, y^*, \pi), q \rangle + \frac{\epsilon}{2}, \\ \forall (z, y, \pi, q) &\in \mathcal{B}(x) \times \mathcal{V}(\bar{y}) \times \mathcal{P}(\pi_x) \times \mathcal{U}(p). \end{aligned}$$

Seja  $\tilde{k}$  suficientemente grande tal que  $y^{j_{\tilde{k}}} \in \mathcal{V}(\bar{y})$ ,  $\pi^{j_{\tilde{k}}} \in \mathcal{P}(\pi_x)$  e  $t_{j_{\tilde{k}}} \|p\|$  seja menor que o raio de uma bola centrada em  $x$  contida em  $\mathcal{B}(x)$ . Então, para  $t \in [0, t_{j_{\tilde{k}}}]$ ,  $x(t) := x + tp \in \mathcal{B}(x)$  e  $x'(t) = dx(t)/dt = p \in \mathcal{U}(p)$ .

Assim,

$$\langle \nabla_x l(x(t), y^{j_{\tilde{k}}}, \pi_x), x'(t) \rangle \geq \langle \nabla_x l(x(t), y^*, \pi^{j_{\tilde{k}}}), x'(t) \rangle + \frac{\epsilon}{2},$$

para  $t \in [0, t_{j_{\tilde{k}}}]$ .

Integrando de  $t = 0$  a  $t = t_{j_{\bar{k}}}$ , obtém-se

$$\begin{aligned} l(x^{j_{\bar{k}}}, y^{j_{\bar{k}}}, \pi_x) - l(x, y^{j_{\bar{k}}}, \pi_x) &\geq l(x^{j_{\bar{k}}}, y^*, \pi^{j_{\bar{k}}}) - l(x, y^*, \pi^{j_{\bar{k}}} + t_{j_{\bar{k}}} \frac{\epsilon}{2}) \\ &> l(x^{j_{\bar{k}}}, y^*, \pi^{j_{\bar{k}}}) - l(x, y^*, \pi^{j_{\bar{k}}}) \end{aligned} \quad (3.1.7)$$

Desde que, pela Hipótese 3.1.1,  $(P_x)$  é convexo, para todo  $x \in D$ , e

$$l(x^{j_{\bar{k}}}, y^{j_{\bar{k}}}, \pi) \leq l(x^{j_{\bar{k}}}, y^{j_{\bar{k}}}, \pi^{j_{\bar{k}}}) = v(x^{j_{\bar{k}}}) \leq l(x^{j_{\bar{k}}}, y, \pi^{j_{\bar{k}}}), \forall \pi \in \mathbb{R}_+^m, \forall y, \quad (3.1.8)$$

e

$$l(x, y^*, \pi) \leq l(x, y^*, \pi_x) = v(x) \leq l(x, y, \pi_x), \forall \pi \in \mathbb{R}_+^m, \forall y, \quad (3.1.9)$$

tem-se, da relação (3.1.7), que

$$v(x^{j_{\bar{k}}}) - l(x, y^{j_{\bar{k}}}, \pi_x) > l(x^{j_{\bar{k}}}, y^*, \pi^{j_{\bar{k}}}) - v(x).$$

Daqui,

$$v(x^{j_{\bar{k}}}) + v(x) > l(x^{j_{\bar{k}}}, y^*, \pi^{j_{\bar{k}}}) + l(x, y^{j_{\bar{k}}}, \pi_x). \quad (3.1.10)$$

Por outro lado, das relações (3.1.8) e (3.1.9), obtém-se

$$v(x^{j_{\bar{k}}}) \leq l(x^{j_{\bar{k}}}, y^*, \pi^{j_{\bar{k}}})$$

$$v(x) \leq l(x, y^{j_{\bar{k}}}, \pi_x),$$



que, somadas, geram uma contradição com (3.1.10).

Portanto, a relação (3.1.4) é falsa e

$$\limsup_{j \rightarrow \infty} \langle \xi^j, p \rangle \leq v'(x; p). \quad (3.1.11)$$

Agora, de (3.1.3) e (3.1.11) tem-se que

$$\liminf_{j \rightarrow \infty} \langle \xi^j, p \rangle = \limsup_{j \rightarrow \infty} \langle \xi^j, p \rangle = v'(x; p).$$

Logo, tem-se provado que  $v(\cdot)$  é semisuave fracamente em  $x$ .

## 3.2 – Uma Proposta Numérica para o Problema (PPDN)

Nesta seção formula-se o problema (P1), equivalente ao problema (PPDN), no contexto de uma família de problemas parametrizados, onde a parametrização está determinada por uma perturbação no lado direito das restrições.

Esta família resulta sendo, como o problema original, em geral, não diferenciável e não convexa. Os problemas parametrizados são, como em Kiwiel [KI1], aproximados localmente por problemas convexos e, através da *função de melhoramento* (definida nesta seção), determina-se uma direção de descida viável (se existir) para o problema aproximado resolvendo um problema quadrático. Prova-se que a direção de descida viável determinada para o problema convexo aproximado, é uma direção de descida viável para o próprio problema perturbado.

Determina-se, assim, um algoritmo para encontrar *pontos estacionários* destes prob-

lemas perturbados. Por último, prova-se que quando a perturbação tende a zero, a correspondente seqüência de pontos estacionários perturbados, possui uma subseqüência convergente a um ponto estacionário do problema original, (P1), desde que se admita uma hipótese simples de *inf- limitação* para este.

Os resultados desta seção, excetuando a prova de convergência que é original, são extensões mais ou menos diretas dos resultados de Kiwiel [KI1].

Para começar a discussão, considerar o problema geral

$$\begin{aligned} (Q) : \min_{x \in C} \quad & J(x) \\ \text{s.a.} \quad & K_i(x) \leq 0, \quad i = 1, \dots, i_1 \end{aligned} \tag{3.2.1}$$

onde  $\emptyset \neq C \subset D \subset \mathbb{R}^n$ ,  $C$  é um conjunto convexo e fechado,  $D$  é um conjunto aberto e todas as funções envolvidas são funções reais, Lipschitz localmente em  $D$  (não necessariamente convexas ou diferenciáveis).

Definir a função *total de restrições*, como

$$K(x) := \max_{1 \leq i \leq i_1} K_i(x), \quad \forall x \in D, \tag{3.2.2}$$

e notar que o programa (Q) é equivalente a

$$\begin{aligned} (Q1) : \min_{x \in C} \quad & J(x) \\ \text{s.a.} \quad & K(x) \leq 0. \end{aligned} \tag{3.2.3}$$

Denotar-se-á o conjunto viável de (Q1) por

$$SQ := \{x \in C / K(x) \leq 0\}. \tag{3.2.4}$$

Para evitar discussão não essencial, considerar a seguinte suposição:

### Hipótese 3.2.1

- (i)  $SQ \neq \emptyset$ .
- (ii) Existe uma solução (local) para o programa (Q).

Seja  $\epsilon > 0$ . Definir uma  $\epsilon$ - aproximação a  $SQ$  por

$$SQ_\epsilon := \{x \in C / K(x) \leq \epsilon\}. \quad (3.2.5)$$

Observar que

$$0 < \epsilon_2 < \epsilon_1 \implies SQ_{\epsilon_2} \subset SQ_{\epsilon_1}$$

e

$$SQ = \bigcap_{\epsilon > 0} SQ_\epsilon.$$

Considerar o programa perturbado

$$\begin{aligned} (Q1_\epsilon) : \min_{x \in C} \quad & J(x) \\ \text{s.a.} \quad & K(x) \leq \epsilon. \end{aligned} \quad (3.2.6)$$

e, para qualquer  $x \in C$ , fixo, a *função de melhoramento* associada

$$H(y; x) := \max\{J(y) - J(x), K(y) - \epsilon\}, \forall y \in D. \quad (3.2.7)$$

Assim, se  $x_\epsilon$  é uma solução ótima local de  $(Q1_\epsilon)$ , então  $H(\cdot; x_\epsilon)$  atinge um mínimo local em  $x_\epsilon$  no conjunto  $C$ , onde  $H(x_\epsilon; x_\epsilon) = 0$ .

A Figura 3.1 ilustra o comentário anterior:

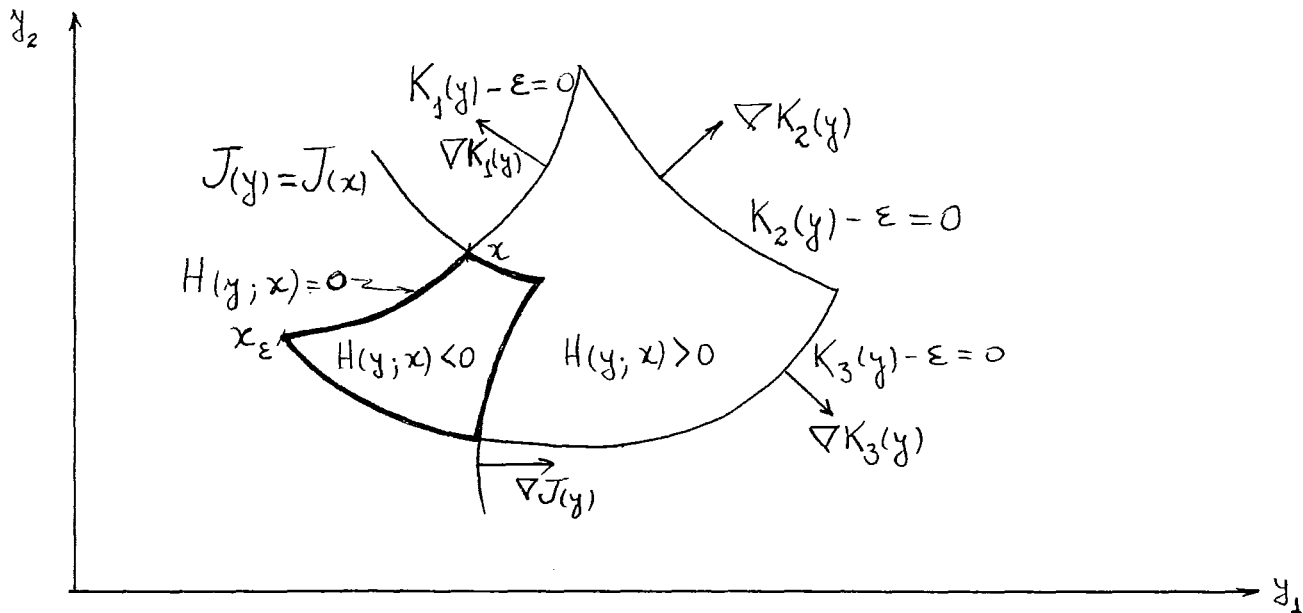


Figura 3.1: Mostram-se as curvas de nível positivo, zero e negativo da função  $H(\cdot; x)$  associada ao problema  $(Q1_\epsilon)$ .

É claro que  $K(\cdot)$  e  $H(\cdot; x)$  são funções Lipschitz localmente em  $D$ . Portanto, se  $x_\epsilon$  é uma solução ótima local de  $(Q1_\epsilon)$ , do Corolário da Proposição 2.4.3, Clarke [CL1], tem-se que

$$0 \in \partial H(x_\epsilon; x_\epsilon) + N_C(x_\epsilon), \tag{3.2.8}$$

onde  $\partial H(\cdot; x)$  denota o gradiente generalizado de  $H(\cdot; x)$ , para  $x \in C$ , fixo.

**Definição 3.2.2** Diz-se que  $x_\epsilon$  é uma solução  $\epsilon$ -local do programa  $(Q)$ , se é uma solução local do programa  $(Q1_\epsilon)$ .

Denotar o conjunto de índices de funções ativas em  $x$ , em (3.2.2), por  $I_k(x)$ , i.e.,

$$I_k(x) := \{i \in \{1, \dots, i_1\} / K_i(x) = K(x)\} \tag{3.2.9}$$

e considerar as seguintes aplicações ponto-conjunto

$$\tilde{\partial}K(x) := \text{Conv}\{\partial K_i(x)/i \in I_K(x)\}, \quad (3.2.10)$$

$$\hat{M}(x) := \begin{cases} \partial J(x), & \text{se } K(x) < \epsilon, \\ \text{Conv}\{\partial J(x) \cup \partial K(x)\}, & \text{se } K(x) = \epsilon, \\ \partial K(x), & \text{se } K(x) > \epsilon, \end{cases} \quad (3.2.11)$$

$$\tilde{M}(x) := \begin{cases} \partial J(x), & \text{se } K(x) < \epsilon, \\ \text{Conv}\{\partial J(x) \cup \tilde{\partial}K(x)\}, & \text{se } K(x) = \epsilon, \\ \tilde{\partial}K(x), & \text{se } K(x) > \epsilon. \end{cases} \quad (3.2.12)$$

Cumpre-se que

$$\partial K(x) \subseteq \tilde{\partial}K(x) \quad , \quad \partial H(x; x) \subseteq \hat{M}(x) \subseteq \tilde{M}(x). \quad (3.2.13)$$

Portanto, se  $x_\epsilon$  é uma  $\epsilon$ -solução local de  $(Q)$ , então, de (3.2.8), tem-se

$$0 \in \tilde{M}(x_\epsilon) + N_C(x_\epsilon). \quad (3.2.14)$$

**Definição 3.2.3** Diz-se que  $x \in SQ_\epsilon$  é um ponto estacionário para  $J$  em  $SQ_\epsilon$ , se satisfaz a relação (3.2.14).

**Proposição 3.2.4** Se  $x_\epsilon$  é uma  $\epsilon$ -solução local de  $(Q)$ , então  $x_\epsilon$  é um ponto estacionário para  $J$  em  $SQ_\epsilon$ . Em particular, existe  $(\lambda, r) \in \mathbb{R}_+^{1+i_1}$ , com  $\lambda + \sum_i r_i = 1$ , tal que

$$0 \in \lambda \partial J(x_\epsilon) + \sum_{i=1}^{i_1} r_i \partial K_i(x_\epsilon) + N_C(x_\epsilon), \quad (3.2.15)$$

$$0 = r_i(K_i(x_\epsilon) - \epsilon), \forall i. \quad (3.2.16)$$

**Prova.**

Imediata, de (3.2.14) e (3.2.12).

**Observação 3.2.5** As relações (3.2.15) e (3.2.16) são a condição necessária de otimidade de Fritz John (estabelecida no Capítulo I, em um contexto mais geral). Quando  $\lambda \neq 0$ , esta passa a ser a condição de  $K - K - T$ .

Considerar, para  $x \in D$ , fixo, e  $z \in D$ , as seguintes funções

$$\bar{J}_\sigma(z) := J(x) + \langle \sigma, z - x \rangle, \quad \text{para cada } \sigma \in \partial J(x), \quad (3.2.17)$$

$$\hat{J}(z) := \max_{\sigma \in \partial J(x)} \bar{J}_\sigma(z), \quad (3.2.18)$$

$$\bar{K}_{i,\kappa_0}(z) := K_i(x) + \langle \kappa_0, z - x \rangle, \quad \text{para cada } \kappa_0 \in \partial K_i(x), \quad (3.2.19)$$

$$i = 1, \dots, i_1,$$

$$\hat{K}_i(z) := \max_{\kappa_0 \in \partial K_i(x)} \bar{K}_{i,\kappa_0}(z), \quad i = 1, \dots, i_1. \quad (3.2.20)$$

Para cada  $\sigma \in \partial J(x)$  e  $\kappa_0 \in \partial K_i(x)$ , as funções  $\bar{J}_\sigma(\cdot)$  e  $\bar{K}_{i,\kappa_0}(\cdot)$  são aproximações lineares a  $J(\cdot)$  e  $K_i(\cdot)$  em  $x$ , respectivamente. Entanto que,  $\hat{J}(\cdot)$  e  $\hat{K}_i(\cdot)$  são aproximações convexas de primeira ordem a  $J(\cdot)$  e  $K_i(\cdot)$ , respectivamente. De fato, são funções convexas em  $C$  (Kiwiel [KI1]).

A Figura 3.2 ilustra o comentário anterior:

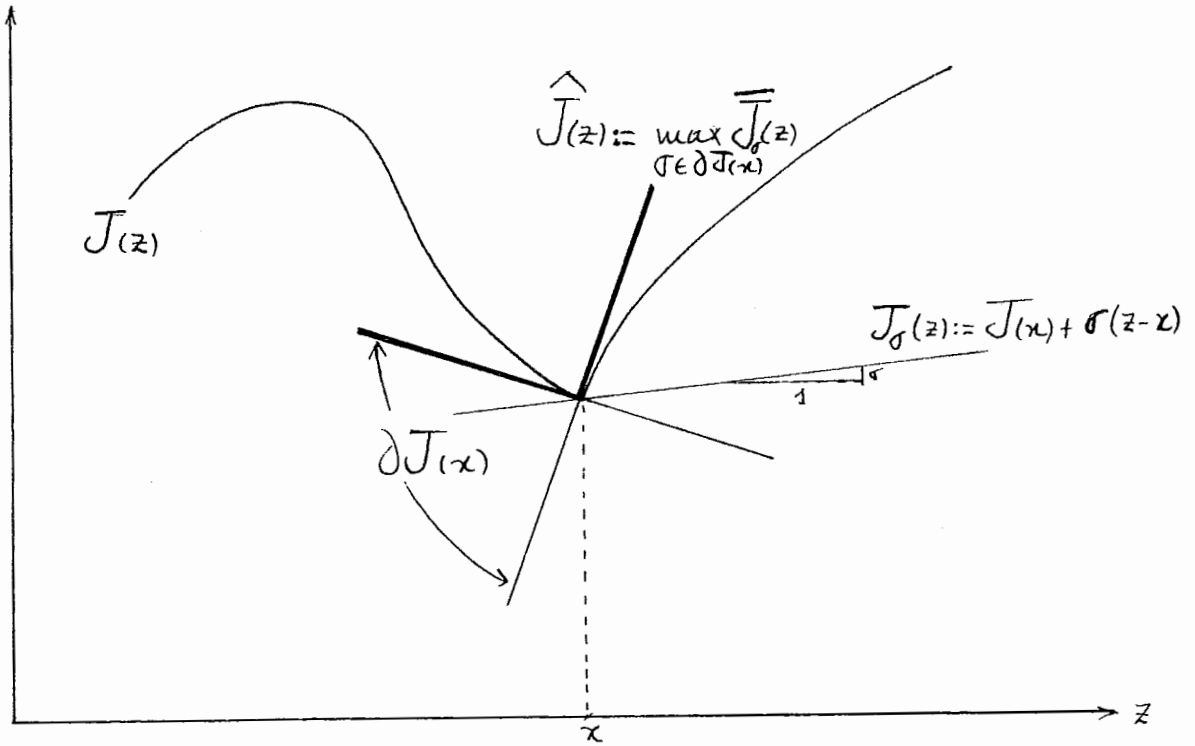


Figura 3.2:  $\bar{J}_\sigma(\cdot)$  e  $\hat{J}(\cdot)$ , uma aproximação linear e a aproximação convexa de primeira ordem à função  $J(\cdot)$  no ponto  $x$ , respectivamente.

É claro que, para  $z \in D$ , também podem-se considerar as aproximações convexas de primeira ordem a  $K(\cdot)$  e  $H(\cdot; x)$  no ponto  $x$ , respectivamente,

$$\hat{K}(z) := \max_{i \in I_K(x)} \hat{K}_i(z), \quad \hat{H}(z) := \max\{\hat{J}(z) - \hat{J}(x), \hat{K}(z) - \epsilon\}. \quad (3.2.21)$$

Considerar, agora, para  $x \in SQ_\epsilon$ , o programa

$$\begin{aligned} (\widehat{Q1}_\epsilon(x)) : \min_{z \in C} & J(x) \\ \text{s.a.} & \hat{K}(z) \leq \epsilon. \end{aligned} \quad (3.2.22)$$

É fácil verificar que  $x$  é viável para  $(\widehat{Q1}_\epsilon(x))$  e, desde que todas as funções envolvidas são convexas no convexo  $C$ , que  $(\widehat{Q1}_\epsilon(x))$  é um programa convexo. Diz-se que  $(\widehat{Q1}_\epsilon(x))$  é uma aproximação convexa ao programa  $(Q1_\epsilon)$  no ponto  $x \in SQ_\epsilon$ .

Usando o Lema 2.9, Capítulo 1, Kiwiel [KI1], satisfaz-se que

$$\partial\hat{J}(x) = \partial J(x) \quad , \quad \partial\hat{K}_i(x) = \partial K_i(x), \forall i. \quad (3.2.23)$$

Portanto, de (3.2.17)-(3.2.21), Corolário 2.6, Capítulo 1, Kiwiel [KI1], (3.2.10) e (3.2.12), obtém-se que

$$\partial\hat{K}(x) = \tilde{\partial}K(x), \quad (3.2.24)$$

$$\partial\hat{H}(x) = \tilde{M}(x). \quad (3.2.25)$$

**Definição 3.2.6** Diz-se que  $d \in \mathbb{R}^n$  é uma direção viável para  $SQ_\epsilon$  em  $x$ , se existe  $\bar{t} > 0$ , tal que

$$x + td \in SQ_\epsilon \quad , \quad \forall t \in (0, \bar{t}]. \quad (3.2.26)$$

**Observação 3.2.7** Desde que  $SQ_\epsilon$  é fechado, (3.2.26) implica que  $x \in SQ_\epsilon$ .

Seja

$$\bar{I}_K(x) := \{i \in \{1, \dots, i_1\} / K_i(x) = K(x) \geq \epsilon\}. \quad (3.2.27)$$

Notar que  $\bar{I}_K(x) = \emptyset$  quando  $K(x) < \epsilon$ . Assim, a Definição 3.2.6 é equivalente à existência de  $\bar{t} > 0$ , tal que

$$x + td \in C, K_i(x + td) \leq \epsilon, \forall t \in (0, \bar{t}], \forall i \in \bar{I}_K(x). \quad (3.2.28)$$



Seja  $x \in SQ_\epsilon$ . Se  $d$  é uma direção viável de descida para  $H(\cdot; x)$  em  $x$ , relativa a  $C$ , i.e., se existe  $\bar{t} > 0$  tal que

$$x + td \in C, \quad H(x + td; x) < H(x; x) = 0, \quad \forall t \in (0, \bar{t}], \quad (3.2.29)$$

então  $d$  também é uma direção viável de descida para o programa  $(Q1_\epsilon)$  em  $x$ , i.e., existe  $\bar{t} > 0$  tal que a relação (3.2.28) se satisfaz e

$$J(x + td) < J(x), \forall t \in (0, \bar{t}]$$

(ver Figura 3.3 para uma ilustração deste fato).

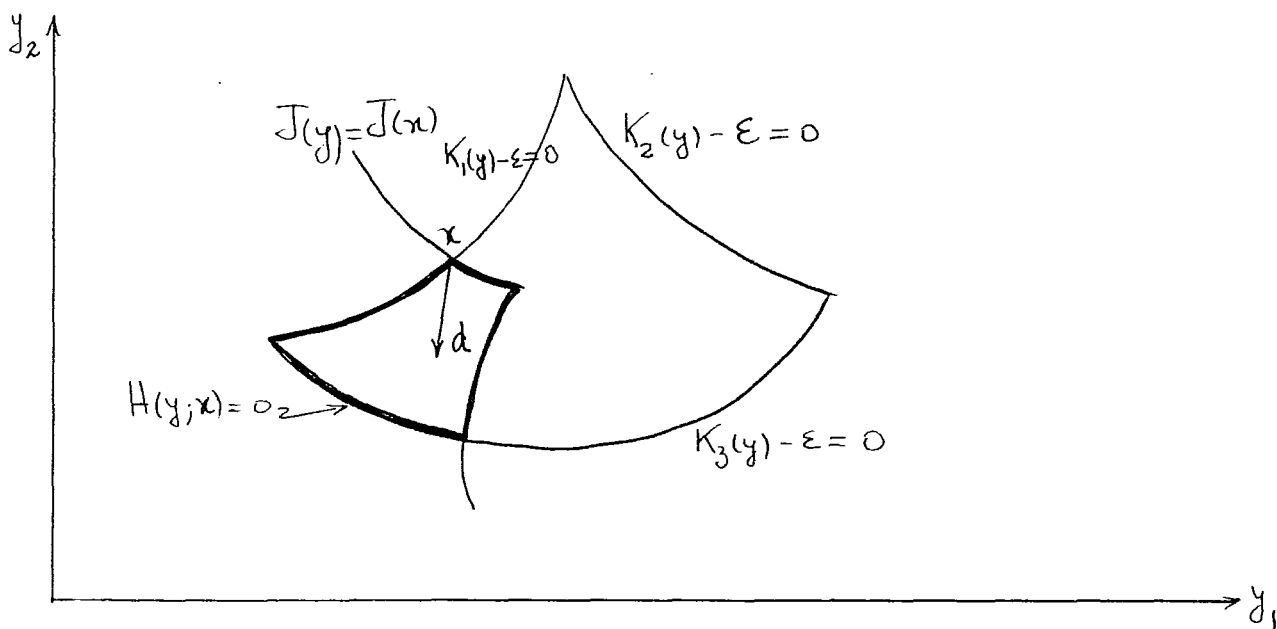


Figura 3.3: Com relação a Figura 3.1, o vetor  $d$  que é uma direção viável de descida para  $H(\cdot; x)$  no ponto  $x$ , também o é para o programa  $(Q1_\epsilon)$  nesse ponto

O Lema seguinte estabelece uma condição suficiente para determinar se um vetor  $d$  é uma direção viável de descida para  $(Q1_\epsilon)$ .

**Lema 3.2.8** Sob as hipóteses anteriores, se  $d$  é uma direção viável de descida para  $\hat{H}(\cdot)$  no ponto  $x \in SQ_\epsilon$ , relativa a  $C$ , i.e., se existe  $\bar{t} > 0$  tal que

$$x + td \in C, \quad \hat{H}(x + td) < \hat{H}(x) = 0, \quad \forall t \in (0, \bar{t}],$$

então  $d$  é uma direção viável de descida para o programa  $(Q1_\epsilon)$  em  $x$ .

**Prova:** É a mesma que a do Lema 2.1.7, Capítulo 1, Kiwiel [KI1].

O seguinte teorema estabelece uma forma computacional para encontrar direções de descida.

**Teorema 3.2.9** Seja  $x \in SQ_\epsilon$ . Considerar

$$\hat{p} + \hat{\eta} := \operatorname{argmin} \left\{ \frac{1}{2} \|\tilde{m} + \eta\|^2 / \tilde{m} \in \tilde{M}(x), \eta \in N_C(x) \right\}, \quad (3.2.30)$$

onde  $\hat{p} + \hat{\eta} \in \tilde{M}(x) + N_C(x)$ , e  $\hat{d}$  é uma solução do problema

$$\min_{d \in C-x} \hat{H}(x + d) + \frac{1}{2} \|d\|^2. \quad (3.2.31)$$

Então

(i)  $\hat{d}$  existe, é único e satisfaz

$$-\hat{d} = \hat{p} + \hat{\eta}, \quad (3.2.32)$$

$$\hat{H}(x + \hat{d}) \leq \hat{H}(x) - \frac{1}{2} \|\hat{d}\|^2, \quad (3.2.33)$$

$$\hat{H}(x + t\hat{d}) \leq \hat{H}(x) - \frac{t}{2}\|\hat{d}\|^2, \forall t \in [0, 1]. \quad (3.2.34)$$

(ii)  $\hat{d} \neq 0 \iff 0 \notin \tilde{M}(x) + N_C(x)$

(iii)  $0 \in \tilde{M}(x) + N_C(x) \iff x$  é mínimo global de  $\hat{H}$  em  $C$ .

(iv) Se adicionalmente,

$$K(x) < \epsilon \quad \text{ou} \quad 0 \notin \tilde{\partial}K(x) + N_C(x), \quad (3.2.35)$$

então  $x$  é um ponto estacionário para  $J$  em  $SQ_\epsilon$  se e somente se  $x$  é solução do programa  $(\widehat{Q1}_\epsilon(x))$ .

**Prova:** Este teorema é uma extensão do Lema 2.1.8, Capítulo 1, Kiwiel [KI1], onde a proposição é formulada para programas não perturbados e  $C = \mathbb{R}^n$ . A prova aqui apresentada segue de perto aquela prova.

Primeiro é conveniente estabelecer que  $\hat{H}$  satisfaz

$$\hat{H}(x + d) = \underbrace{\hat{H}(x)}_{=0} + \max_{\mu \in \partial \hat{H}(x)} \langle \mu, d \rangle, \forall d. \quad (3.2.36)$$

Observar que, admitindo (3.2.36), ter-se-ia, de (3.2.25), que

$$\hat{H}(x + d) = \hat{H}(x) + \max_{\mu \in \tilde{M}(x)} \langle \mu, d \rangle, \forall d. \quad (3.2.37)$$

De (3.2.21), tem-se que

$$\hat{H}(x+d) := \begin{cases} \hat{J}(x+d) - \hat{J}(x), & \text{se } \hat{J}(x+d) - \hat{J}(x) > \hat{K}(x+d) - \epsilon, \\ \hat{J}(x+d) - J(x), & \text{se } \hat{J}(x+d) - \hat{J}(x) = \hat{K}(x+d) - \epsilon, \\ \hat{K}(x+d) - \epsilon, & \text{se } \hat{J}(x+d) - \hat{J}(x) < \hat{K}(x+d) - \epsilon, \end{cases} \quad (3.2.38)$$

e, de (3.2.17) e (3.2.18), que

$$\hat{J}(x+d) = \max_{\sigma \in \partial J(x)} (J(x) + \langle \sigma, d \rangle) = \hat{J}(x) + \max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle, \forall d.$$

Então,

$$\hat{J}(x+d) - \hat{J}(x) = \max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle, \forall d. \quad (3.2.39)$$

Por outro lado, de (3.2.19) e (3.2.20), resulta que

$$\begin{aligned} \hat{K}(x+d) &= \max_{i \in I_K(x)} \hat{K}_i(x+d) = \max_{i \in I_K(x)} \max_{\kappa \in \partial K_i(x)} (K_i(x) + \langle \kappa, d \rangle), \forall d \\ &= \underbrace{K(x)}_{\hat{K}(x)} + \max_{i \in I_K(x)} \max_{\kappa \in \partial K_i(x)} \langle \kappa, d \rangle, \forall d \end{aligned}$$

e, do Teorema 32.2, Rockafellar [RO0],

$$\hat{K}(x+d) = \hat{K}(x) + \max_{\kappa \in \text{Conv}(\cup_{i \in I_K(x)} \partial K_i(x))} \langle \kappa, d \rangle, \forall d.$$

Daqui, usando (3.2.10),

$$\hat{K}(x+d) = \hat{K}(x) + \max_{\kappa \in \bar{\partial} K(x)} \langle \kappa, d \rangle, \forall d. \quad (3.2.40)$$

Supor que

$$\hat{J}(x+d) - \hat{J}(x) > \hat{K}(x+d) - \epsilon \quad \text{e} \quad \hat{K}(x) = \epsilon.$$

Então, de (3.2.39) e (3.2.40), tem-se, equivalentemente, que

$$\langle \sigma_0, d \rangle := \max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle > \max_{\kappa \in \tilde{\partial} K(x)} \langle \kappa, d \rangle \quad (3.2.41)$$

Seja  $\gamma \in \text{Conv}(\partial J(x) \cup \tilde{\partial} K(x))$ , então

$$\exists \{\lambda_i\}_{i \in I_K(x) \cup \{0\}} \subset [0, 1] / \sum_{i \in I_K(x) \cup \{0\}} \lambda_i = 1, \gamma = \lambda_0 \sigma + \sum_{i \in I_K(x)} \lambda_i \kappa_i, \quad (3.2.42)$$

onde  $\sigma \in \partial J(x)$  e  $\kappa_i \in \partial K_i(x), \forall i \in I_K(x)$ .

Assim,

$$\langle \gamma, d \rangle = \lambda_0 \langle \sigma, d \rangle + \sum_{i \in I_K(x)} \lambda_i \langle \kappa_i, d \rangle$$

e, de (3.2.41),

$$\begin{aligned} \langle \gamma, d \rangle &\leq \lambda_0 \langle \sigma_0, d \rangle + \sum_{i \in I_K(x)} \lambda_i \langle \sigma_0, d \rangle \\ &= \langle \sigma_0, d \rangle = \max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle \leq \max_{\gamma \in (\partial J(x) \cup \tilde{\partial} K(x))} \langle \gamma, d \rangle \end{aligned}$$

Então,

$$\max_{\gamma \in (\partial J(x) \cup \tilde{\partial} K(x))} \langle \gamma, d \rangle = \max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle.$$

Portanto, para  $\hat{J}(x+d) - \hat{J}(x) > \hat{K}(x+d) - \epsilon$  e  $\hat{K}(x) \leq \epsilon$ , usando (3.2.38), (3.2.39), (3.2.12) e (3.2.25), obtem-se que

$$\hat{H}(x+d) = \max_{\mu \in \tilde{M}(x)} \langle \mu, d \rangle = \hat{H}(x) + \max_{\mu \in \partial \hat{H}(x)} \langle \mu, d \rangle. \quad (3.2.43)$$

Supor, agora, que

$$\hat{J}(x+d) - \hat{J}(x) < \hat{K}(x+d) - \epsilon \quad \text{e} \quad \hat{K}(x) < \epsilon.$$

Então, de (3.2.39) e (3.2.40), tem-se que

$$\max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle < \underbrace{\max_{\kappa \in \partial \hat{K}(x)} \langle \kappa, d \rangle}_{:= \langle \kappa, d \rangle} + \hat{K}(x) - \epsilon. \quad (3.2.44)$$

Como em (3.2.42), seja  $\gamma \in \text{Conv}(\partial J(x) \cup \partial \hat{K}(x))$ . Então, de (3.2.44),

$$\langle \gamma, d \rangle = \lambda_0 \langle \sigma, d \rangle + \sum_{i \in I_K(x)} \lambda_i \langle \kappa_i, d \rangle \leq \langle \kappa_0, d \rangle + \hat{K}(x) - \epsilon.$$

Portanto, tem-se

$$\begin{aligned} \max_{\gamma \in \text{Conv}(\partial J(x) \cup \partial \hat{K}(x))} \langle \gamma, d \rangle &\leq \langle \kappa_0, d \rangle + \hat{K}(x) - \epsilon < \kappa_0, d \rangle \\ &= \max_{\kappa \in \partial \hat{K}(x)} \langle \kappa, d \rangle \\ &\leq \max_{\gamma \in \text{Conv}(\partial J(x) \cup \partial \hat{K}(x))} \langle \gamma, d \rangle, \end{aligned}$$

que é uma contradição.

Assim, se  $\hat{J}(x+d) - \hat{J}(x) < \hat{K}(x+d) - \epsilon$ , ocorre necessariamente que  $\hat{K}(x) = \epsilon$ , e, como acima, ter-se-á que

$$\max_{\gamma \in \text{Conv}(\partial J(x) \cup \tilde{\partial} K(x))} \langle \gamma, d \rangle \leq \max_{\kappa \in \tilde{\partial} K(x)} \langle \kappa, d \rangle \leq \max_{\gamma \in \text{Conv}(\partial J(x) \cup \tilde{\partial} K(x))} \langle \gamma, d \rangle .$$

Daí que

$$\max_{\kappa \in \tilde{\partial} K(x)} \langle \kappa, d \rangle = \max_{\gamma \in \text{Conv}(\partial J(x) \cup \tilde{\partial} K(x))} \langle \gamma, d \rangle = \max_{\mu \in \hat{M}(x)} \langle \mu, d \rangle$$

e

$$\hat{H}(x + d) = \hat{K}(x + d) - \epsilon = \hat{K}(x + d) - \hat{K}(x).$$

Então, de (3.2.40) e (3.2.25), resulta que

$$\hat{H}(x + d) = \max_{\kappa \in \tilde{\partial} K(x)} \langle \kappa, d \rangle = \hat{H}(x) + \max_{\mu \in \partial \hat{H}(x)} \langle \mu, d \rangle . \quad (3.2.45)$$

Para, finalmente, estabelecer (3.2.37), só falta analisar o caso

$$\hat{J}(x + d) - \hat{J}(x) = \hat{K}(x + d) - \epsilon,$$

que é equivalente a

$$\max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle = \max_{\kappa \in \tilde{\partial} K(x)} \langle \kappa, d \rangle + \hat{K}(x) - \epsilon. \quad (3.2.46)$$

Se  $\hat{K}(x) < \epsilon$ , chegar-se-á a uma contradição como acima. Assim, necessariamente ocorre que  $\hat{K}(x) = \epsilon$ . Então, usando a igualdade (3.2.46), tem-se que

$$\max_{\sigma \in \partial J(x)} \langle \sigma, d \rangle = \max_{\kappa \in \tilde{\partial} K(x)} \langle \kappa, d \rangle = \max_{\gamma \in \text{Conv}(\partial J(x) \cup \tilde{\partial} K(x))} \langle \gamma, d \rangle = \max_{\mu \in \tilde{M}(x)} \langle \mu, d \rangle$$

e

$$\begin{aligned} \hat{H}(x+d) &= \hat{J}(x+d) - \hat{J}(x) = \hat{K}(x+d) - \epsilon = \max_{\mu \in \tilde{M}(x)} \langle \mu, d \rangle \\ &= \hat{H}(x) + \max_{\mu \in \tilde{M}(x)} \langle \mu, d \rangle. \end{aligned} \tag{3.2.47}$$

Portanto, de (3.2.38), (3.2.43), (3.2.45) e (3.2.47), segue-se (3.2.37).

Agora, usando (3.2.37), observa-se que (3.2.31) pode ser escrito como

$$\min_{d \in C-x} \Phi(x),$$

onde

$$\Phi(d) := u(d) + \frac{\|d\|^2}{2}$$

e

$$u(d) := \max_{\mu \in \tilde{M}(x)} \langle \mu, d \rangle$$

(lembrar que  $\hat{H}(x) = 0$ ).

Fixar  $\alpha \in \mathbb{R}$  e, para  $\tilde{\mu} \in \tilde{M}(x)$ , considerar



$$\mathcal{D}_\alpha := \{d \in C - x / \Phi(d) \leq \alpha\}.$$

Desde que

$$\alpha \geq \Phi(d) \geq \langle \tilde{\mu}, d \rangle + \frac{\|d\|^2}{2} \geq -\|\tilde{\mu}\| \|d\| + \frac{\|d\|^2}{2}, \quad (3.2.48)$$

tem-se que  $\mathcal{D}_\alpha$  é limitado, caso contrário, a compacidade de  $\tilde{M}(x)$  e (3.2.48) implicariam que  $\Phi(d) \rightarrow +\infty$ ; o que seria uma contradição com  $\Phi(d) \leq \alpha$ , para todo  $d \in \mathcal{D}_\alpha$ .

Portanto, tem-se que  $\mathcal{D}_\alpha$  é compacto, i.e.,  $\Phi(\cdot)$  é inf-compacta em  $C - x$ , e, daqui, conclui-se que existe solução  $\hat{d}$  para (3.2.31). Desde que  $\Phi(\cdot)$  é estritamente convexa, esta solução é única e a condição

$$0 \in \partial\Phi(\hat{d}) + N_C(x + \hat{d})$$

é necessária e suficiente para que  $\hat{d}$  seja solução de (3.2.31).

Então,

$$\hat{d} = 0 \iff 0 \in \underbrace{\partial\Phi(0)}_{\partial u(0) = \tilde{M}(x)} + N_C(x) \iff 0 \in \tilde{M}(x) + N_C(x).$$

Assim, provam-se as proposições (ii) e (iii).

Para completar a prova de (i), notar que se  $0 \in \partial\Phi(\hat{d}) + N_C(x + \hat{d})$  é equivalente à existência de vetores

$$\bar{p} \in \partial u(\hat{d}) \subset \tilde{M}(x) \quad \text{e} \quad \bar{\eta} \in N_C(x + \hat{d}),$$

tal que

$$0 = \bar{p} + \hat{d} + \bar{\eta}. \quad (3.2.49)$$

Dado que

$$\bar{p} \in \partial u(\hat{d}) = \{\mu \in \tilde{M}(x) / \langle \mu, \hat{d} \rangle = u(\hat{d})\},$$

tem-se que

$$u(\hat{d}) = \langle \bar{p}, \hat{d} \rangle \geq \langle \mu, \hat{d} \rangle, \forall \mu \in \tilde{M}(x),$$

que é equivalente a

$$\langle \mu, -\hat{d} \rangle \geq \langle \bar{p}, -\hat{d} \rangle, \forall \mu \in \tilde{M}(x),$$

e, de (3.2.49), equivalente a

$$\langle \mu, \bar{p} + \bar{\eta} \rangle \geq \langle \bar{p}, \bar{p} + \bar{\eta} \rangle, \forall \mu \in \tilde{M}(x). \quad (3.2.50)$$

Desde que

$$\bar{\eta} \in N_C(x + \hat{d}) \iff \langle \bar{\eta}, c - x - \hat{d} \rangle \leq 0, \forall c \in C,$$

para  $c = x$ , resulta  $\langle \bar{\eta}, -\hat{d} \rangle \leq 0$  e, de (3.2.49),

$$\langle \bar{\eta}, \bar{p} + \bar{\eta} \rangle \leq 0. \quad (3.2.51)$$

Agora, para qualquer  $\eta \in N_C(x)$ , tem-se

$$\langle \eta, c - x \rangle \leq 0, \forall c \in C,$$

e, para  $c = x + \hat{d}$ ,

$$\langle \eta, \hat{d} \rangle \leq 0, \forall \eta \in N_C(x).$$

Daqui, usando (3.2.49), obtem-se que

$$\langle \eta, \bar{p} + \bar{\eta} \rangle \geq 0, \forall \eta \in N_C(x) \quad (3.2.52)$$

Assim, de (3.2.50), (3.2.51) e (3.2.52), resulta que

$$\langle \mu + \eta, \bar{p} + \bar{\eta} \rangle \geq \|\bar{p} + \bar{\eta}\|^2, \forall \mu + \eta \in \tilde{M}(x) + N_C(x),$$

e, do Lema 2.12, Capítulo 1, Kiwiel [KI1],

$$\bar{p} + \bar{\eta} = \hat{p} + \hat{\eta}. \quad (3.2.53)$$

Portanto, de (3.2.49) e (3.2.53), segue-se (3.2.32).

Por outro lado, de (3.2.37) e (3.2.49),

$$\begin{aligned} \hat{H}(x + \hat{d}) &= \langle \bar{p}, \hat{d} \rangle + \frac{\|\hat{d}\|^2}{2} = - \langle \hat{d} + \bar{\eta}, \hat{d} \rangle + \frac{\|\hat{d}\|^2}{2} \\ &= - \langle \bar{\eta}, \hat{d} \rangle - \frac{\|\hat{d}\|^2}{2}. \end{aligned}$$

e, desde que  $\langle \bar{\eta}, -\hat{d} \rangle \leq 0$ ,

$$\hat{H}(x + \hat{d}) \leq -\frac{\|\hat{d}\|^2}{2}.$$

Lembrando que  $\hat{H}(x) = 0$ , segue-se (3.2.33).

Para provar (3.2.34), notar que  $\hat{H}(\cdot)$  é convexa em  $C$ , pelo que, de (3.2.33), tem-se, para  $t \in [0, 1]$ ,

$$\begin{aligned} \hat{H}(x + t\hat{d}) &= \hat{H}(t(x + \hat{d}) + (1-t)x) \leq t\hat{H}(x + \hat{d}) + (1-t)\hat{H}(x) \\ &\leq t\hat{H}(x) - \frac{t\|\hat{d}\|^2}{2} + (1-t)\hat{H}(x). \end{aligned}$$

Assim,

$$\hat{H}(x + \hat{d}) \leq \hat{H}(x) - \frac{t}{2}\|\hat{d}\|^2, \forall t \in [0, 1].$$

e (3.2.34) está provado.

Finalmente, para provar (iv), notar que quando  $x$  é um ponto estacionário para  $J(\cdot)$  em  $SQ_\epsilon$ , a condição (3.2.35) elimina o caso  $\lambda = 0$  em (3.2.15). Pelo que, neste caso, a partir da Observação 3.2.5, obtem-se a condição de  $K - K - T$  (a relação (3.2.35) é conhecida como a condição de qualificação de Cottle para  $(Q1_\epsilon)$  no ponto  $x$ ).

Desde que  $\hat{K}(x) = K(x)$  e  $\partial\hat{K}(x) = \tilde{\partial}K(x)$ , de (3.2.35), obtem-se que

$$\hat{K}(x) < \epsilon \quad \text{ou} \quad 0 \notin \partial\hat{K}(x) + N_C(x). \quad (3.2.53')$$

Observar que, dado  $(\widehat{Q1_\epsilon}(x))$  é um programa convexo, ocorre que

$$0 \in \partial \hat{K}(x) + N_C(x) \iff \hat{K}(x) \leq \hat{K}(y), \forall y \in C.$$

Assim,

$$0 \notin \partial \hat{K}(x) + N_C(x) \iff \exists \bar{x} \in C / \hat{K}(\bar{x}) < \hat{K}(x) \leq \epsilon.$$

Daqui, (3.2.53') é equivalente a

$$\exists \bar{x} \in C / \hat{K}(\bar{x}) < \epsilon,$$

que é a condição de qualificação de Slater para  $(\widehat{Q1}_\epsilon(x))$ .

Portanto, desde que para programas convexos que satisfazem a condição de qualificação de Slater, a condição de  $K - K - T$  no ponto  $x$  é necessária e suficiente para a otimidade deste ponto, usando a relação (3.2.25), obtém-se que  $x \in SQ_\epsilon$  é solução de  $(\widehat{Q1}_\epsilon(x))$  se, e somente se,  $0 \in \hat{H}(x) + N_C(x) = \tilde{M}(x) + N_C(x)$ , que, de (iii), constitui a prova de (iv).

O Teorema 3.2.9, através da solução do programa quadrático (3.2.31) e da relação (3.2.34), respectivamente, gera direções de descida e pontos de uma seqüência de minimização viável para  $(Q1_\epsilon)$ , por exemplo, através do *método dos centros* de Huard [HU1]:

### Algoritmo 3.2.10 (Método dos centros)

Passo 0: Seja  $x \in SQ_\epsilon$ . Fazer

$$j \leftarrow 1 \quad e \quad x^1 \leftarrow x \in SQ_\epsilon$$

Passo 1: Encontrar  $d^j$ , resolvendo o programa

$$\min_{d \in C-x^j} \hat{H}(x^j + d) + \frac{1}{2}\|d\|^2.$$

Se  $d^j = 0$ ,  $x_\epsilon \leftarrow x^j$ , PARAR (pois,  $0 \in \tilde{M}(x_\epsilon) + N_C(x_\epsilon)$ ).

Se  $d^j \neq 0$ , continuar ( $d^j$  é direção de descida viável).

Passo 2: Determinar  $t_j \in [0, 1]$  /  $\hat{H}(x^j + t_j d^j) < \hat{H}(x^j)$  e fazer

$$x^{j+1} \leftarrow x^j + t_j d^j$$

Passo 3:  $j \leftarrow j + 1$ , e ir para o Passo 1.

Se a seqüência  $\{x^j\} \subset SQ_\epsilon$  gerada pelo Algoritmo 3.2.10 tem um ponto de acumulação  $x_\epsilon$ , então  $0 \in \tilde{M}(x_\epsilon) + N_C(x_\epsilon)$  (Capítulo 5, Kiwiel [KI1]).

Observar que (3.2.31) é um programa convexo não diferenciável, que normalmente não pode ser resolvido diretamente. Uma versão implimentável do Algoritmo 3.2.10, na que o subproblema é substituído por um subproblema mais simples (uma aproximação local adequada), que é usada neste trabalho, é o algoritmo BTNC (Schram-Zowe [SCHRZ1]).

Portanto, dado  $\epsilon > 0$ , obtém-se desta maneira um ponto estacionário  $x_\epsilon$  para o problema  $(Q1_\epsilon)$ . No que se segue, prova-se que a seqüência  $\{x_\epsilon\}$ , sob uma hipótese simples, tem uma subseqüência convergente a um ponto estacionário do problema  $(Q1)$  (que é equivalente ao problema original  $(Q)$ ).

Considerar  $X^*$  como o conjunto de pontos estacionários do problema  $(Q)$ , dado em (3.2.1),  $x_\epsilon$  um ponto estacionário de  $(Q1_\epsilon)$  e a hipótese de inf-limitação seguinte:

**Hipótese 3.2.11** Para cada  $\alpha \in \mathbb{R}$ , existe  $\beta > 0$  tal que

$$\mathcal{J}_\alpha := \{x \in D/J(x) \leq \alpha, K(x) \leq \beta\}$$

é limitado (então, compacto).

**Teorema 3.2.12** Seja  $x_{\epsilon_k}$  ponto estacionário de  $(Q1_{\epsilon_k})$  e a seqüência  $\{x_{\epsilon_k}\}_k$ , onde  $\epsilon_k \downarrow 0$ . Então, sob a Hipótese 3.2.11, cumpre-se que existe uma subsequência de  $\{x_{\epsilon_k}\}_k$  convergente a um ponto de  $X^*$ .

**Prova:**

Da Hipótese 3.2.1 e de  $SQ \subset SQ_{\epsilon_k}$ , para  $\epsilon_k > 0$ , tem-se que existe  $x^*$  solução de  $(Q)$  (i.e.,  $X^* \neq \emptyset$ ), que  $SQ_{\epsilon_k} \neq \emptyset$  e que existe solução para  $(Q1_{\epsilon_k})$ . De forma que, de fato, existe a seqüência  $\{x_{\epsilon_k}\}_k$ .

Então desde que  $SQ \subset SQ_{\epsilon_k}$ , resulta que  $\{x_{\epsilon_k}\}_{k \in \{k/\epsilon_k \leq \beta\}} \subset \mathcal{J}_\alpha$ , para  $\alpha = J(x^*)$ ; e, da Hipótese 3.2.11, tem-se que existe  $x$  tal que a seqüência  $\{x_{\epsilon_k}\}_{k \in \{k/\epsilon_k \leq \beta\}}$ , considerando uma subsequência se necessário, e convergente a  $x$ .

Notar que

$$0 \in \partial \tilde{M}_k(x_{\epsilon_k}) + N_C(x_{\epsilon_k}),$$

onde  $\tilde{M}_k(\cdot)$  está determinada pela relação (3.2.12), para  $\epsilon = \epsilon_k$ .

Esta última relação, pela Proposição 3.2.4, implica

$$0 \in \lambda^k \partial J(x_{\epsilon_k}) + \sum_i r_i^k \partial K_i(x_{\epsilon_k}) + N_C(x_{\epsilon_k}), \quad (3.2.54)$$

$$0 = r_i^k (K_i(x_{\epsilon_k} - \epsilon_k), \forall i, \quad (3.2.55)$$

$$0 \leq r_i^k, \quad (3.2.56)$$

$$0 \leq \lambda^k, \quad (3.2.57)$$

$$1 = \lambda^k + \sum_i r_i^k. \quad (3.2.58)$$

Por outro lado, desde que  $x_{\epsilon_k} \in SQ_{\epsilon_k}$ , para todo  $k$ , tem-se que  $K(x_{\epsilon_k}) \leq \epsilon_k$ . Daqui, passando ao limite, obtém-se que  $K(x) \leq 0$  e  $x \in C$  (pois  $C$  é fechado), i.e.,  $x \in SQ$ . É claro que, (3.2.54) pode ser escrita, equivalentemente, como a existência de  $\xi^k \in \partial J(x_{\epsilon_k})$ ,  $\kappa_i^k \in \partial K_i(x_{\epsilon_k})$ ,  $\forall i$  e  $\eta^k \in N_C(x_{\epsilon_k})$  tal que

$$0 = \lambda^k \xi^k + \sum_i r_i^k \kappa_i^k + \eta^k, \quad (3.2.54')$$

Desde que  $\partial J(\cdot)$  e  $\partial K_i(\cdot)$  são aplicações ponto-conjunto localmente limitadas (i.e., limitadas em conjuntos limitados de  $\mathbb{R}^n$ ), tem-se, de  $x_{\epsilon_k} \rightarrow x$ , que existem pontos de acumulação  $\xi$  e  $\kappa_i$ , para cada seqüência  $\{\xi^k\}_k$  e  $\{\kappa_i^k\}_k, \forall i$ , respectivamente.

Da semicontinuidade superior de  $\partial J(\cdot)$  e  $\partial K_i(\cdot), \forall i$ , resulta que  $\xi \in \partial J(x)$  e  $\kappa_i \in \partial K_i(x), \forall i$ .

Observar que para as seqüências de multiplicadores  $\{\lambda^k\}_k$  e  $\{r_i^k\}_k, \forall i$ , considerando subsequências se necessário, tem-se que existem  $\lambda \geq 0$  e  $r_i \geq 0, \forall i$ , tal que  $\lambda^k \rightarrow \lambda$  e  $r_i^k \rightarrow r_i, \forall i$ , onde  $\lambda + \sum_i r_i = 1$  e  $r_i K_i(x) = 0, \forall i$ . Finalmente, este processo gera também um ponto de acumulação  $\eta \in N_C(x)$ , para a seqüência  $\{\eta^k\}$ . Assim, tem-se obtido que existem  $\lambda \in \mathbb{R}, \{r_i\}_{i=1}^{i_1} \subset \mathbb{R}, \xi \in \partial J(x), \kappa_i \in \partial K_i(x), \forall i$ , e  $\eta \in N_C(x)$ , tal que



$$0 = \lambda\xi + \sum_i r_i \kappa_i + \eta$$

$$0 = r_i K_i(x), \forall i$$

$$1 = \lambda + \sum_i r_i$$

$$0 \leq \lambda$$

$$0 \leq r_i, \forall i,$$

Daqui, obtem-se que  $x \in X^*$ .

### 3.3 Um algoritmo para o PPDN

Lembrar que o problema a ser tratado numericamente é

$$(P1) : \min_{(x,y) \in X \times \mathbb{R}^{n_2}} F(x, y) \\ \text{s.a} \quad G_i(x, y) \leq 0, i = 1, \dots, m_1, \\ f(x, y) - v(x) \leq 0, \\ g_i(x, y) \leq 0, i = 1, \dots, m_2,$$

sob as hipóteses feitas nos capítulos anteriores.

Desde que, para cada  $x \in X$ ,  $v(x) = \min_{y \in Y(x)} f(x, y)$ , se  $Y(x) \neq \emptyset$ , e  $v(x) = +\infty$ , se  $Y(x) = \emptyset$ , tem-se que a restrição  $f(x, y) - v(x) \leq 0$ , em  $(P1)$  é equivalente à igualdade  $f(x, y) - v(x) = 0$ . Esta igualdade, por motivos algorítmicos e computacionais, é conveniente perturbá-la, considerando a relação

$$0 \leq f(x, y) - v(x) \leq \epsilon,$$

para  $\epsilon \geq 0$ .

Obtem-se, assim, a família parametrizada em  $\epsilon$

$$(P1_\epsilon) : \min_{(x,y) \in X \times \mathbb{R}^{n_2}} F(x, y) \tag{3.3.1} \\ \text{s.a} \quad G_i(x, y) \leq 0, i = 1, \dots, m_1, \\ g_i(x, y) \leq 0, i = 1, \dots, m_2, \\ v(x) - f(x, y) \leq 0, \\ f(x, y) - v(x) \leq \epsilon.$$

Sob a Hipótese 1.1.1 e a Hipótese 3.1.1, tem-se que o problema  $(P1_\epsilon)$  tem solução.  $(P1)$  é membro dessa família.

O algoritmo a propor trabalhará com os programas perturbados  $(P1_\epsilon)$ , diminuindo  $\epsilon$  convenientemente, até atingir uma certa tolerância. É claro que uma solução local de  $(P1_\epsilon)$  satisfará as restrições  $G(x, y) \leq 0$  e  $g(x, y) \leq 0$ . Pelo que fornecerá uma aproximação *viável* a uma solução do problema  $(P1)$ .

Dado que o algoritmo se formulará tomando como base o algoritmo BTNC, extensão do algoritmo BT, que por sua vez é uma modificação do método de feixe, no que segue descrevem-se, brevemente, estes métodos. Um estudo detalhado do algoritmo BT pode ser encontrado em Arica-Scheimberg [ASCH1], Schramm [SCHR1] e Schramm-Zowe [SCHRZ1], enquanto que para o método de feixe pode-se consultar Hiriart-Urruty-Lemaréchal [H-ULE1]. A exposição que segue, sobre os métodos mencionados, foi literalmente tomada de Outrata-Schramm-Zowe [OSCHRZ1].

### 3.3.1 O método de feixe

Considerar, por simplicidade, o problema irrestrito

$$(IR) : \min_{x \in \mathbb{R}} h(x),$$

onde  $h$  é uma função Lipschitz localmente em  $\mathbb{R}$ .

É comum, para o problema (IR), fazer a hipótese de que se tem a capacidade de calcular, para  $x \in \mathbb{R}$ , um elemento  $\xi \in \partial h(x)$ , i.e., possui-se uma subrotina (um *oráculo*) que além de calcular  $h(x)$ , calcula simultaneamente  $\xi \in \partial h(x)$ .

Os principais pontos do método de feixe são:

- Armazenamento de subgradientes  $\xi_i$  avaliados durante iterações anteriores em pontos  $y_i$  no feixe.  $\mathcal{J}_k$  denotará o conjunto de índices dos subgradientes no feixe.
- Ponderação da informação. Para  $h$  convexa, investiga-se os erros de linearização

$$\alpha(x_k, y_i) := \alpha_i^k := h(x_k) - [h(y_i) + \langle \xi_i, x_k - y_i \rangle],$$

para  $\xi_i \in \partial h(y_i)$  e o ponto iterado atual  $x_k$ , que fornece uma medida da distância de  $\xi_i$  ao conjunto  $\partial h(x_k)$ .

- Cálculo da direção de busca  $d_k = -z_k$ , onde  $z_k$  é uma combinação convexa dos subgradientes no feixe. Quanto maior o peso  $\alpha_i^k$  de um subgradiente, menor será sua influência sobre  $d_k$ .
- Busca linear ao longo de  $d_k$  para calcular o próximo ponto iterado e enriquecer a informação dada pelos subgradientes; para isto, dois passos diferentes são possíveis:

(i) *passo sério*:

$$x_{k+1} \leftarrow y_{k+1} \leftarrow x_k + td_k, \quad t > 0,$$

no caso em que um decrescimento suficiente seja conseguido em  $x_{k+1}$ .

(ii) *passo nulo*:

$$x_{k+1} \leftarrow x_k, \quad y_{k+1} \leftarrow x_k + td_k, \quad t > 0,$$

no caso em que não seja conseguido um decrescimento suficiente, mas  $\xi_{k+1} \in \partial h(y_{k+1})$ , enriquece a informação dada pelos subgradientes significativamente.

#### Observação 3.3.1

(i) A implementação mais conhecida do conceito de feixe é o código FORTRAN *M1FC1* (Lemaréchal-Bancora [LEBA1]). Em cada iteração de *M1FC1*, calcula-se uma solução  $\lambda_k^i$ ,  $i \in \mathcal{J}_k$ , do programa quadrático

$$(PQ) : \min \frac{1}{2} \left\| \sum_{i \in \mathcal{J}_k} \lambda_i \xi_i \right\|^2$$

$$\begin{aligned} \text{s.a. } \quad & \sum_{i \in \mathcal{J}_k} \lambda_i = 1, \\ & \sum_{i \in \mathcal{J}_k} \lambda_i \alpha_i^k \leq \nu, \\ & \lambda_i \geq 0, \quad i \in \mathcal{J}_k. \end{aligned}$$

A direção de busca se define como  $d_k = -\sum_{i \in \mathcal{J}_k} \lambda_i^k \xi_i$ . Isto está motivado pela consideração do  $\nu_k$ -subdiferencial

$$\partial_{\nu_k} h(x) := \{ \xi / h(y) \geq h(x) + \langle \xi, y - x \rangle - \nu_k \}$$

e a correspondente derivada direcional

$$h'_{\nu_k}(x; d) := \inf_{t > 0} \frac{h(x + td) - h(x) + \nu_k}{t} = \max_{\xi \in \partial_{\nu_k}} \langle \xi, d \rangle.$$

A  $\nu_k$ -direção de descida de  $h$  em  $x_k$  é a solução de  $\min_{\|d\| \leq 1} h'_{\nu_k}(x_k; d)$ . Prova-se que  $z_k := -\sum_{i \in \mathcal{J}_k} \lambda_i^k \xi_i$  é a projeção da origem a uma aproximação interior de  $\partial_{\nu_k} h(x_k)$  e, assim,  $d_k = z_k$  é uma aproximação da  $\nu_k$ -direção de descida.

- (ii) Existe uma íntima relação entre os métodos de feixe e o método de planos de corte, pois uma outra implementação (Kiwiel [KI1]) encontra a direção de busca resolvendo

$$\begin{aligned} (PQ_1) : \quad & \min \quad \frac{1}{2} \left\| \sum_{i \in \mathcal{J}_k} \lambda_i \xi_i \right\|^2 + \sum_{i \in \mathcal{J}_k} \lambda_i \alpha_i^k \\ \text{s.a. } \quad & \sum_{i \in \mathcal{J}_k} \lambda_i = 1, \\ & \lambda_i \geq 0, \quad i \in \mathcal{J}_k, \end{aligned}$$

cujos dual é

$$(PC) : \min_d \hat{h}_k(x_k + d) + \frac{1}{2} \|d\|^2,$$

onde

$$\hat{h}_k(x_k + d) := \max_{i \in \mathcal{J}_k} \{ h(y_i) + \langle \xi_i, x_k + d - y_i \rangle \}$$

é uma aproximação poliedral inferior de  $h$  em  $x_k$ , i.e, um modelo gerado por planos de corte de  $h$  ao redor de  $x_k$ .

### 3.3.2 O método BT

O método BT combina características dos métodos de regiões de confiança e de feixe. Esta baseado no fato de o método de feixe usar uma aproximação

poliedral de  $h$  para calcular a direção de busca. Entre tanto, a busca linear do método de feixe é substituída por uma estratégia equivalente à de regiões de confiança.

Por simplicidade, supor ainda que  $h$  é convexa. Usam-se os subgradientes armazenados para construir a aproximação poliedral  $\hat{h}_k$  de  $h$  ao redor de  $x_k$ . Considerando os erros de linearização  $\alpha_i^k$ , antes definidos, pode-se escrever

$$\hat{h}_k(x_k + d) = h(x_k) + \max_{i \in \mathcal{J}_k} \{-\alpha_i^k + \langle \xi_i, d \rangle\} .$$

Para calcular o próximo ponto iterado a formulação de regiões de confiança consiste em minizar o modelo  $\hat{h}_k$  em uma bola centrada em  $x_k$ , i.e., dado  $\rho_k$ , resolver

$$(RC) : \min_d \hat{h}_k(x_k + d) \\ \text{s.a.} \quad \frac{1}{2} \|d\|^2 \leq \rho_k .$$

A solução  $d_k$  de  $(RC)$  gera o próximo ponto de procura,  $x_k + d_k$ . Logo  $\rho_k$  é ajustado para o seguinte passo. A partir do fato de ser  $(RC)$  um programa com uma restrição quadrática, prefere-se, por questões computacionais, tratar com o programa

$$\min_d \hat{h}_k(x_k + d) + \frac{1}{2t_k} \|d\|^2 ,$$

que pode ser escrito como

$$(BT) : \min_d v + \frac{1}{2t_k} \|d\|^2 \\ \text{s.a.} \quad -\alpha_i^k + \langle \xi_i, d \rangle \leq v, \quad i \in \mathcal{J}_k .$$

Então, modifica-se  $t_k$  em vez de  $\rho_k$ . Justifica-se isto pela seguinte relação:

(i) Se  $d^*$  é solução de  $(RC)$  e  $1/t_k$  é um multiplicador de Lagrange da restrição quadrática em  $(RC)$ , então  $d^*$  é solução de  $(BT)$ .

(ii) Se  $d^*$  é solução de  $(BT)$  e  $\rho_k := \|d^*\|^2/2$ , então  $d^*$  é solução de  $(RC)$ .

Para atingir um passo sério ou um passo nulo como no método de feixe, substitui-se a busca linear pela formulação de região de confiança. Durante este processo, chamado de *iteração interna*, ajusta-se  $t_k$ .

### Observação 3.3.2

- (i) *Estabelece-se o seguinte critério de parada :*  
 Seja  $\nu \geq 0$ . Supor que  $\|\sum_{i \in \mathcal{J}_k} \lambda_i^k \xi_i\| \leq \nu$  e  $\sum_{i \in \mathcal{J}_k} \lambda_i^k \alpha_i \leq \nu$ . Então  $h(y) \leq h(x_k) - \nu \|y - x_k\| - \nu$ , para todo  $y \in \mathbb{R}^n$ , i.e.,  $x_k$  é aproximadamente  $\nu$ -ótimo.

(ii) Os critérios para passo sério e passo nulo do método de feixe se modificam da seguinte forma:

Sejam  $0 < m_1 < m_2$  e  $(d(t_k), v(t_k))$  a solução de (BT) correspondente a  $t_k$ . Considerar  $\xi_{k+1} \in \partial h(x_k + d(t_k))$  e  $\alpha_{k+1}^{k+1} := \alpha(x_k, x_k + d(t_k))$   
**critério de passo sério:**

$$\begin{aligned} h(x_k + d(t_k)) - h(x_k) &< m_1 v(t_k), \\ < \xi_{k+1}, d(t_k) &\geq m_2 v(t_k); \end{aligned}$$

**critério de passo nulo:**

$$\begin{aligned} h(x_k + d(t_k)) - h(x_k) &\geq m_1 v(t_k), \\ -\alpha_{k+1}^{k+1} + < \xi_{k+1}, d(t_k) > &\geq m_2 v(t_k); \\ \alpha_{k+1}^{k+1} &\leq m_3 \sum_{i \in \mathcal{J}_k} \lambda_i^k \alpha_i. \end{aligned}$$

Na iteração interna adapta-se  $t_k$  (via bissecção, por exemplo). Prova-se que em um número finito de passos se encontra um  $t_k$  adequado, tal que o critério de passo sério ou passo nulo se satisfaz.

Para uma implementação eficiente, devem-se levar em consideração vários aspectos. Assim, por exemplo, o conjunto de índices  $\mathcal{J}_k$  (o número de subgradientes considerado) deve ser mantido de um tamanho razoável quando  $k \rightarrow \infty$ . Portanto, na implementação, cada certo número de iterações será reiniciado o feixe. A análise de convergência requer  $|\mathcal{J}_k| \geq 3$  mais uma certa estratégia de reiniciação.

### Algoritmo 3.3.3 (Algoritmo BT)

Escolher um ponto inicial  $x_1 \in \mathbb{R}^n$  e os parâmetros  $T > 0$ ,  $0 < m_1 < m_2 < 1$ ,  $0 < m_3 < 1$ ,  $\nu \geq 0$ . Estabelecer uma cota superior  $\mathcal{J}_{max} \geq 3$  para  $|\mathcal{J}_k|$ .

**Passo 0:** Calcular  $h(x_1)$ ,  $\xi_1 \in \partial h(x_1)$  e fazer

$$y_1 \leftarrow x_1, \mathcal{J}_1 \leftarrow \{1\}, k \leftarrow 1.$$

**Passo 1: (ITERAÇÃO INTERNA)** Calcular  $x_{k+1}$  e  $\xi_{k+1}$  tal que o critério de passo sério ou o critério de passo nulo se satisfaça ou seja atingido  $x_k$  aproximadamente  $\nu$ -ótimo (neste último caso PARAR).

**Passo 2:** Se  $|\mathcal{J}_k| = \mathcal{J}_{max}$  ir ao Passo 3; caso contrário

$$\mathcal{J} \leftarrow \mathcal{J}_k \text{ e ir ao Passo 4.}$$

**Passo 3: (REINÍCIO)** Escolher  $\mathcal{J} \subset \mathcal{J}_k$  tal que  $|\mathcal{J}_k| \leq \mathcal{J}_{max} - 2$  e  $\max\{i/i \in \mathcal{J}_k, \alpha_i^k = 0\} \in \mathcal{J}$ .

Introduzir um índice adicional  $\tilde{k}$  e definir

$$z_k \leftarrow -\frac{1}{t_k} d_k = \sum_{i \in \mathcal{J}_k} \lambda_i^k \xi_i, \sigma_k \leftarrow \sum_{i \in \mathcal{J}_k} \lambda_i^k \alpha_i^k$$

$$\xi_k \leftarrow z_k, \alpha_k^k \leftarrow \sigma_k, \mathcal{J} \leftarrow \mathcal{J} \cup \{\tilde{k}\}.$$

**Passo 4 (ATUALIZAÇÃO)** Se o resultado da iteração interna é passo sério, fazer

$$\alpha_i^{k+1} \leftarrow \alpha_i^k + h(x_{k+1}) - h(x_k) - \langle \xi_i, d_k \rangle, \forall i \in \mathcal{J},$$

$$\alpha_{k+1}^{k+1} \leftarrow 0.$$

Se o resultado da iteração interna é passo nulo, fazer

$$\alpha_i^{k+1} \leftarrow \alpha_i^k \quad \forall i \in \mathcal{J},$$

$$\alpha_{k+1}^{k+1} \leftarrow \alpha(x_k, y_{k+1}).$$

Fazer

$$\mathcal{J}_{k+1} \leftarrow \mathcal{J} \cup \{k+1\} \text{ e ir ao Passo 1.}$$

**Observação 3.3.4** A combinação convexa  $\xi_k$ , no Passo 3, corresponde à agregação do subgradiente introduzida por Kiwiel [KI1].

Considerando  $X^* := \{x^* \in \mathbb{R}^n / h(x^*) = \inf_{x \in \mathbb{R}^n} h(x)\}$ , têm-se os seguintes resultados de convergência

- Se  $X^* \neq \emptyset$ , então  $x_k \rightarrow x^* \in X^*$ , quando  $k \rightarrow \infty$ .
- Se  $X^* = \emptyset$ , então  $h(x_k) \rightarrow \inf_{x \in \mathbb{R}^n} h(x) \in [-\infty, +\infty)$ .

### 3.3.3 O método BT para o caso não convexo: BTNC

Apresentam-se aqui as modificações necessárias para que o método BT possa ser aplicado ao caso em que  $h$  não é convexa. O método modificado é denominado BTNC.

Supor que  $h$  é Lipschitz localmente e semisuave fracamente (Definição 3.1.3).

No caso convexo, tem-se que

$$\partial h(x) = \{\xi \in \mathbb{R}^n / h(y) \geq h(x) + \langle \xi, y - x \rangle, \forall y \in \mathbb{R}^n\}.$$

No caso não convexo, a relação anterior não se satisfaz. Portanto,  $\alpha_i^k$  poderia ser negativo e, nesse caso, o modelo  $\hat{h}_k$ , gerado pelos planos de corte não é mais uma aproximação inferior de  $h(x_k + \cdot) - h(x_k)$ .

Para evitar essa dificuldade, procede-se a aplicar a mesma estratégia que em outros métodos de feixe, substituindo  $\alpha_i^k$  por

$$\beta_i^k := \beta(x_k, y_i) := \max\{\alpha_i^k, c_0 \|x_k - y_i\|^2\},$$

onde  $c_0$  é um real positivo pequeno (com  $c_0 = 0$  se  $h$  for convexa). É claro que  $\beta_i^k \geq 0$ . Considera-se, logo, o modelo modificado

$$\hat{h}_k(d) := \max_{\mathcal{J}_k} \{\xi_i, d > -\beta_i^k\}. \quad (3.3.2)$$

Os  $\beta_i^k$  copiam parte do papel dos  $\alpha_i^k$ : sempre que  $y_i$  esteja longe do ponto atual  $x_k$ , então  $\beta_i^k$  é grande e, portanto,  $\xi_i$  influencia pouco o modelo.

Substitui-se, agora,  $\alpha_i^k$  por  $\beta_i^k$  no algoritmo BT. Isto não muda o caracter do programa (BT). Entretanto, para  $h$  não convexa, a condição

$$\left\| \sum_{i \in \mathcal{J}_k} \lambda_i \xi_i \right\| \leq \nu \text{ e } \sum_{i \in \mathcal{J}_k} \lambda_i \beta_i^k \leq \nu$$

só indica que 0 está próximo, com respecto a  $\nu$ , da envoltória convexa de certos  $\xi_i \in \partial h(y_i)$  para os que  $y_i$  não está longe de  $x_k$ .

A iteração interna requer duas modificações: eliminar a segunda parte do critério de passo sério e adicionar a seguinte condição de passo nulo

$$\langle \xi_{k+1}, d_k \rangle - \beta_{k+1}^{k+1} \geq m_2 v_k.$$

Isto garante que depois do passo nulo, o modelo gera uma direção suficientemente diferente da insatisfatória direção anterior.

Para o caso não convexo, tem-se o seguinte resultado de convergência

**Teorema 3.3.5** (Teorema 3.1, Schramm-Zowe [SCHRZ1]) *Se  $h$  é semisuave fracamente, limitada inferiormente e a seqüência de pontos iterados é limitada, então existe um ponto de acumulação  $\bar{x}$  da seqüência  $\{x_k\}_{k \in \mathbb{N}}$  tal que  $0 \in \partial h(\bar{x})$ .*

### 3.3.4 BLPP: Um algoritmo para solucionar o PPDN

Considerar a família de programas parametrizados em  $\epsilon \geq 0$ , definida na relação (3.3.1)

$$(P1_\epsilon) : \begin{array}{ll} \min_{(x,y) \in X \times \mathbb{R}^{m_2}} & F(x,y) \\ \text{s.a} & G_i(x,y) \leq 0, \quad i = 1, \dots, m_1, \\ & g_i(x,y) \leq 0, \quad i = 1, \dots, m_2, \\ & v(x) - f(x,y) \leq 0, \\ & f(x,y) - v(x) \leq \epsilon. \end{array}$$

É claro que o problema (P1), equivalente a (PPDN), pertence a essa família, pois (P1<sub>0</sub>) é (P1).



Seja  $\epsilon > 0$  e  $K_\epsilon(x, y)$  a função total de restrições para o programa  $(P1_\epsilon)$ . Então, denotando as restrições deste programa por

$$\tilde{g}_i(x, y) := G_i(x, y), \quad \forall i \in \{1, \dots, m_1\},$$

$$\tilde{g}_{m_1+i}(x, y) := g_i(x, y), \quad \forall i \in \{1, \dots, m_2\},$$

$$\tilde{g}_{m_1+m_2+1}(x, y) := v(x) - f(x, y),$$

e

$$\tilde{g}_{m_1+m_2+2}(x, y) := f(x, y) - v(x) - \epsilon,$$

tem-se

$$K_\epsilon(x, y) = \max_{1 \leq i \leq m_1+m_2+2} \{\tilde{g}_i(x, y)\}. \quad (3.3.3)$$

O seguinte programa é equivalente a  $(P1_\epsilon)$

$$(P2_\epsilon): \quad \min_{(x,y) \in X \times \mathbb{R}^{n_2}} \quad F(x, y) \quad (3.3.4)$$

*s.a*  $K_\epsilon(x, y) \leq 0.$

Considerar  $(x^k, y^k)$  viável para  $(P2_\epsilon)$  e a função de melhoramento associada

$$H_\epsilon^k := H_\epsilon(x, y; x^k, y^k) = \max\{F(x, y) - F(x^k, y^k), K_\epsilon(x, y)\}. \quad (3.3.5)$$

Sabe-se, de (3.2.14), que se  $(x^k, y^k)$  é solução de  $(P2_\epsilon)$ , então

$$0 \in \tilde{M}_\epsilon(x^k, y^k) + \begin{pmatrix} N_X(x^k) \\ 0 \end{pmatrix}, \quad (3.3.6)$$

onde, considerando (3.3.4) e (3.2.12), tem-se que

$$\tilde{M}_\epsilon(x, y) = \begin{cases} \nabla F(x, y) & , \text{ se } K_\epsilon(x, y) < F(x, y) - F(x^k, y^k) \\ \text{Conv}(\nabla F(x, y) \cup \tilde{\partial} K_\epsilon(x, y)) & , \text{ se } K_\epsilon(x, y) = F(x, y) - F(x^k, y^k) \\ \tilde{\partial} K_\epsilon(x, y) & , \text{ se } K_\epsilon(x, y) > F(x, y) - F(x^k, y^k) \end{cases}$$

e

$$\tilde{\partial} K_\epsilon(x, y) = \text{Conv}\{\partial g_i / g_i(x, y) = K_\epsilon(x, y)\}.$$

Assim, a partir da Def. 3.2.6, os comentários que seguem a Obs. 3.2.7, o Lema 3.2.8 e o Teorema 3.2.9, tem-se que para encontrar uma solução do programa  $(P2_{\epsilon_k})$ , para  $\epsilon_k > 0$ , pode-se proceder segundo o seguinte **algoritmo conceitual**:

(i) Encontrar uma direção de descida viável  $d^k$  para  $H_{\epsilon_k}^k$  no ponto  $(x^k, y^k)$ . Se  $\|d^k\| = 0$ , ir a (iii). Caso contrário, continuar.

(ii) Considerar o novo ponto iterado  $(x^{k+1}, y^{k+1}) := (x^k, y^k) + d^k$  e a função  $H_{\epsilon_k}^{k+1}$ . Voltar para (i).

(iii) O ponto  $(x_{\epsilon_k}, y_{\epsilon_k}) := (x^k, y^k)$  é um ponto viável para o programa  $(P2_{\epsilon_k})$ , estacionário para a função  $F$ . PARAR.

**Observação 3.3.6** Desde que, pela Prop. 3.1.4, a função  $H_{\epsilon_k}^k$  é semisuave fracamente para qualquer  $k$ , usa-se neste trabalho o Algoritmo BTNC para determinar  $(x^{k+1}, y^{k+1})$ , tomando como ponto inicial  $(x^k, y^k)$ . Segundo o Teorema 3.3.5 o algoritmo é convergente.

Determinado  $(x_{\epsilon_k}, y_{\epsilon_k})$ , pode-se encontrar  $y'_{\epsilon_k} \in O(x_{\epsilon_k})$  e, diminuindo adequadamente  $\epsilon_k$  para  $\epsilon_{k+1}$ , reiniciar o algoritmo conceitual para encontrar uma solução de  $(P2_{\epsilon_{k+1}})$ .

**Observação 3.3.7**

(i) Para encontrar  $y'_{\epsilon_k} \in O(x_{\epsilon_k})$ , usa-se neste trabalho a rotina NLPQL1 (Schittkowski [SCH1]).

(ii) O algoritmo conceitual aplicado ao programa  $(P2_{\epsilon_{k+1}})$ , pode-se aplicar tomando como ponto inicial  $(x^1, y^1) := (x_{\epsilon_k}, y'_{\epsilon_k})$  como ponto inicial.

(iii) Sob as hipóteses do Teorema 3.2.12, a seqüência de pontos  $\{(x_{\epsilon_k}, y_{\epsilon_k})\}$  converge a um ponto estacionário do problema (PPDN).

No que segue formula-se o algoritmo proposto neste trabalho, denominado Algoritmo BLPP (das siglas do PPDN em inglês: *Bilevel Programming Problem*), e alguns resultados numéricos.

**Algoritmo 3.3.8** (Algoritmo BLPP)

Escolher  $\epsilon > 0$ ,  $TOL > 0$  (pequeno),  $KONS > 0$ ,  $(x^0, y^0)$  ponto inicial viável para  $(P1_\epsilon)$ , tal que  $y^0 \in O(x^0)$ , e os parâmetros para o algoritmo BTNC.

Iniciar o contador de iterações  $k$ , fazendo

$$k \leftarrow 0.$$

**Passo 1 (Iteração maior)** Fazer

$$(x^1, y^1) \leftarrow (x^k, y^k), k \leftarrow k + 1, \epsilon_k \leftarrow \epsilon$$

**Passo 2 (iteração menor)** Encontrar uma solução  $(x_{\epsilon_k}, y_{\epsilon_k})$  do programa  $(P2_{\epsilon_k})$ , usando repetidamente o algoritmo BTNC, como indicado no algoritmo conceitual anterior (**fim da iteração menor**).

**Passo 3** Encontrar  $y'_{\epsilon_k} \in O(x_k)$ , usando o algoritmo NLPQL1 com ponto inicial  $y_{\epsilon_k}$ .

Fazer

$$(x^k, y^k) \leftarrow (x_{\epsilon_k}, y_{\epsilon_k}).$$

Se  $\epsilon_k \leq TOL$ , ir ao Passo 4. Se  $\epsilon_k > TOL$ , fazer  $\epsilon \leftarrow \epsilon_k / KONS$ . Ir ao Passo 1.

**Passo 4** PARAR,  $(x^k, y^k)$  é um ponto viável próximo de um ponto estacionário do problema (PPDN) (**fim da iteração maior**).

O algoritmo BLPP foi testado para resolver os seguintes problemas:

**Problema 3.3.9** (Adaptado de Clark-Westerberg [CW1]) O programa (*PPDN*), onde  $X = [4, 16]$ ,  $Y = \mathbb{R}$ ,  $F(x, y) = x - 4y$ ,  $f(x, y) = y$ ,  $g_1(x, y) = -2x + y$ ,  $g_2(x, y) = 2x + 5y - 108$ ,  $g_3(x, y) = 2x - 3y + 4$ . Não existem restrições do tipo  $G_i(x, y) \leq 0$ .

**Problema 3.3.10** (Exemplo 1, Outrata [OT2]) O programa (*PPDN*), onde  $X = \mathbb{R}_+$ ,  $Y = \mathbb{R}_+^2$ ,  $F(x, y) = \frac{1}{2}[(y_1 - 3)^2 + (y_2 - 4)^2]$ ,  $f(x, y) = \frac{1}{2}(y_1^2 + y_2^2) - (3 + 1.333x)y_1 - xy_2$ ,  $g_1(x, y) = -0.333y_1 + y_2 - 2$ ,  $g_2(x, y) = y_1 - 0.333y_2 - 2$ . Não existem restrições do tipo  $G_i(x, y) \leq 0$ .

**Problema 3.3.11** (Exemplo 3, Outrata [OT2]) O programa (*PPDN*), onde os dados são os mesmos que no Problema 3.3.10, exceto que

$$f(x, y) = \frac{1}{2}[(1 + x)y_1^2 + (1 + 0.1x)y_2^2] - (3 + 1.333x)y_1 - xy_2 .$$

Os resultados conseguidos, aplicando algoritmo BLPP a estes problemas, mostram-se na seguinte tabela

PROB.	PONTO INICIAL	SOLUÇÃO DO ALGRT. BLPP	SOLUÇÃO DE [OT2]
3.3.9	$x = 4$ $y = 4$ $\epsilon = 8/3$ $F(x, y) = -12$	$x = 15.999856449176$ $y = 11.999904299451$ $\epsilon = (8/3)10^{-2}$ $F(x, y) =$ $-31.999760748627$	<i>não foi testado</i>
3.3.10	$x = 1$ $y_1 = 2.5326401080576$ $y_2 = 1.5995198440168$ $\epsilon = 1$ $F(x, y) =$ $2.9903651239327$	$x = 2.0577778281451$ $y_1 = 2.9904675647858$ $y_2 = 2.9743770714289$ $\epsilon = 0.125$ $F(x, y) =$ $0.52599662946596$	$x = 2.07$ $y_1 = 3$ $y_2 = 3$  $F(x, y) =$ $0.5$
3.3.11	$x = 1$ $y_1 = 2.16649663016667$ $y_2 = 0.90908168342285$ $\epsilon = 0.3$ $F(x, y) =$ $5.1242522274437$	$x = 5.457978256261$ $y_1 = 1.7046718730350$ $y_2 = 2.5676557337207$ $\epsilon = (75)10^{-5}$ $F(x, y) =$ $1.8647425268249$	$x = 3.456$ $y_1 = 1.707$ $y_2 = 2.569$  $F(x, y) =$ $1.859$

**Observação 3.3.12** Na tabela acima, tem-se que:

(i) Na coluna *PONTO INICIAL*, só foi fornecido o ponto inicial para  $x$ , pois  $y \in O(x)$  foi obtido aplicando a rotina *NLPQL1*.

(ii) Na coluna *SOLUÇÃO DE [OT2]*, fornecem-se os resultados dos respectivos exemplos obtidos por *Outrata [OT2]*, usando penalização. Como é de esperar os resultados são pontos inviáveis; entanto que, na coluna *SOLUÇÃO DO ALGRT. BLPP*, que fornece os resultados obtidos usando o algoritmo proposto, exibem-se pontos viáveis.

## Conclusões

Neste trabalho tem-se discutido, principalmente, dois fatos relativos ao Problema de Programação de Dois Níveis (PPDN):

- Condições de Otimalidade.
- Tratamento numérico.

Em relação às condições de otimalidade, tem-se estabelecido, sob a condição de qualificação de *programa calmo* para o PPDN e no contexto da análise não diferenciável, condições necessárias do tipo K-K-T para mínimos locais (Teorema 2.3.6). É importante lembrar que, em geral, o PPDN é um problema não convexo, o que faz difícil o estudo de soluções globais.

As condições necessárias aqui apresentadas, até onde o autor conhece, são os primeiros resultados obtidos para o caso geral. É conveniente assinalar que se requerem condições de segunda ordem para o problema do seguidor e que, tanto a semicontinuidade do conjunto de *Reação Racional* (determinada pela aplicação ponto-conjunto  $O(\cdot)$ ), assim como a condição de calma para o PPDN, são hipóteses que não são simples de serem testadas, embora sob hipóteses adicionais não seja tão complicado verificá-las. De fato, apresenta-se aqui um exemplo onde todas as hipóteses solicitadas se satisfazem e, portanto, o Teorema 2.3.6 (que estabelece as condições de K-K-T) se verifica.

Em relação ao tratamento numérico, deve-se dizer que tem tido dois objetivos básicos. De um lado, propor um algoritmo de descida viável; por outro lado, verificar as propostas teóricas. Deve-se indicar que a implementação numérica, em termos da *função de melhoramento* (Kiwiel [KI1]), usando códigos computacionais existentes (NLPQL1, Schittkowski [SCH1], e BTNC, Outrata-Schramm-Zowe [OSCHRZ1]) é uma primeira versão, onde não foram exploradas todas as possibilidades teóricas. Assim, os resultados, em termos de tempo de processamento, número de iterações, exemplos testados, etc. podem ser aprimorados. Isto requer uma adequada modificação dos códigos usados. O algoritmo BLPP, aqui proposto, tem sido codificado em FORTRAN e rodado em um microcomputador 486.

## Referências

- [AE1] AUBIN, J.P.; EKELAND, I. *Applied Nonlinear Analysis*. Wiley-Interscience, New York (1984).
- [AF1] ANANDALINGAM, G.; FRIESZ, T. *Hierarchical Optimization*. Annals of Operations Research, Vol. 34 (1992).
- [AS1] AIYOSHI, E.; SHIMIZU, K.A. *Solution Method for the Static Constrained Stackelberg Problem Via Penalty Method*. IEEE Trans. Auto. Control AC-29, pp. 1111-1114. (1982).
- [ASCH1] ARICA, J.; SCHEIMBERG, S. *Regiões de Confiança e o Método de Feixe: Uma Extensão do Algoritmo de Schramm-Zowe para Restrições Lineares*. Publicações Técnicas ES-249/91. Programa de Engenharia de Sistemas e Computação, COPPE/UFRJ (1991).
- [EF1] BARD, J.; FALK, J. *An Explicit Solution to the Multi-level Programming Problem*. Compt. & Ops. Res., Vol. 9, N. 1, pp. 77-100. (1982).
- [CF1] CHEN, Y.; FLORIAN, M. *The Nonlinear Bilevel Programming Problem: Formulations, Regularity and Optimality Conditions*. CRT-794. Centre de Recherche sur les Transports, Université de Montréal, C.P. 6128, Succursale A, Montréal, Canada, H3C3J7. (July, 1991 - Revised July, 1993).

- [CL1] CLARKE, F.H. *Nonsmooth Analysis and Optimization*. Wiley-Interscience, New York (1983).
- [CW1] CLARK, P.A.; WESTERBERG, A.W. *A Note on the Optimality Conditions for the Bilevel Programming Problem*. Naval Research Logistics, Vol. 35, pp. 413-418. (1988).
- [DE1] DEMPE, S. *A Simple Algorithm for the Linear Bilevel Programming Problem*. Optimization, 18, 3, pp. 373-385. (1987).
- [DE2] DEMPE, S. *A Necessary and a Sufficient Optimality Condition for Bilevel Programming Problems*, Optimization, 25, pp. 344-354. (1992).
- [dSM1] de SILVA, A.H.; McCORMICK, G.P. *Implicitly Defined Optimization Problems*. Annals of Operations Research 34, pp. 107-124. (1992).
- [GD1] GAUVIN, J.; DUBEAU, F. *Differential Properties of the Marginal Function in Mathematical Programming*. Mathematical Programming Study 19, pp. 101-119. (1982).
- [HJS1] HANSEN, P.; JAUMARD, B.; SAVARD, G. *A Variable Elimination Algorithm for Bilevel Linear Programming*. SIAM J. Scientific and. Statist. Comput., 13, pp. 1194-1217 (1992).
- [H-ULE1] HIRIART-URRUTY, J.B.; LEMARÉCHAL, C. *Convex Analysis and Minimization Algorithms II, Advanced Theory and Bundle Methods*. Springer-Verlag Berlin, Heidelberg, New York (1993).
- [HO1] HOGAN, W.W.. *Point-to Set Maps in Mathematical Programming*. SIAM Review 15, N. 3, pp. 591-603. (July, 1973).

- [HU1] HUARD, P. *Resolution of Mathematical Programming with Non-Linear Constraints by the Method of Centers*. Nonlinear Programming, J. Abadie, Ed. Academic Press, New York (1967).
- [IA1] ISHIZUKA, Y.; AIYOSHI, E. *Double Penalty Method for Bilevel Optimization Problems*. Annals of Operations Research 34, pp. 73-88. (1992).
- [KI1] KIWIEL, K.C. *Methods of Descent for Nondifferentiable Optimization*. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo. (1985).
- [KY1] KYPARISIS, J. *On Uniqueness of Kuhn-Tucker Multipliers in Nonlinear Programming*. Mathematical Programming 32, pp. 242-246. (1985).
- [KO1] KOLSTAD, C.D. *A Review of the Literature on Bilevel Mathematical Programming*. Report N. LA-10284-MS, Los Alamos National Laboratory, Los Alamos, New Mexico, U.S.A. (1985).
- [KL1] KOLSTAD, C.D.; LASDON, L.S. *Derivative Evaluation and Computational Experience with Large Bilevel Mathematical Programs*, Journal of Optimization Theory and Appl., Vol. 65, N. 3, pp. 485-499. (June, 1990).
- [LEBA1] LEMARÉCHAL, C.; BANCORA IMBERT, M. *Le Module M1FC1*. INRIA, Le Chesnay (1985).
- [LM1] LORIDAN, R.; MORGAN, J. *Approximate Solutions for Two-level Optimization Problems*, International Series of Numerical Mathematics, Vol. 84 (Trends in Mathematical Optimization). Birkhäuser, Basel, pp. 181-196. (1988).



- [MAF1] MANGASARIAN, O.L.; FROMOVITZ, S. *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints*. Journal of Mathematical Analysis and Applications, 17, pp. 37-47. (1967).
- [MF1] MIFFLIN, R. *An Algorithm for Constraint Optimization with Semismooth Functions*. Math. of Operations Research, Vol. 2, No. 2, pp. 191-206 (May, 1977).
- [MF2] MIFFLIN, R. *Semismooth and Semiconvex Functions in Constrained Optimization*. SIAM J. Control and Optimization, Vol. 15, No. 6 (November, 1977).
- [OT1] OUTRATA, J.V. *On Generalized Gradients in Optimization Problems with Set-Valued Constraints*. Math. of Oper. Res., Vol. 15, N. 4, pp. 626-639. (1990).
- [OT2] OUTRATA, J.V. *On the Numerical Solution of a Class of Stackelberg Problems*. Zeit. Oper. Res. 34, pp. 255-277. (1990).
- [OT3] OUTRATA, J.V. *Necessary Optimality Conditions for Stackelberg Problems*. Journal of Optim. Theory and Appl., Vol. 76, N. 2, pp. 305-320. (February, 1993).
- [OT4] OUTRATA, J.V. *On Optimization Problems with Variational Inequalities Constraints*. SIAM J. Optimization 4, No. 2, (1994).
- [OSCHRZ1] OUTRATA, J., SCHRAMM, H.; ZOWE, J. *Bundle Trust Methods: Fortran Codes for Nondifferentiable Optimization User's Guide*. Report N. 269. Mathematisches Institut, Universität Bayreuth, Postfach 10 12 51, D-8580, Bayreuth, Germany (1991).
- [RB1] ROBINSON, S.M. *Strongly regular Generalized Equations*. Mathematical of Operations Research. 5, pp. 43-62. (1980).

- [RO0] ROCKAFELLAR, R.T. *Convex Analysis*. Princeton University Press. (1970).
- [RO1] ROCKAFELLAR, R.T. *Augmented Lagrange Multiplier Functions and Duality in Non-convex Programming*. SIAM Journal on Control 12, pp. 268-285. (1971).
- [RO2] ROCKAFELLAR, R.T. *Lagrange Multipliers and Subderivatives of Optimal Value Functions in Nonlinear Programming*. Mathematical Programming Study 17, pp. 28-66. (1982).
- [RO3] ROCKAFELLAR, R.T. *Marginal Values and Second Order Necessary Conditions for Optimality*. Mathematical Programming 26, pp. 245-286. (1983).
- [RO4] ROCKAFELLAR, R.T. *Directional Differentiability of the Optimal Value Function in a Nonlinear Programming Problem*. Mathematical Programming Study 21, pp. 213-226. (1984).
- [SCH1] SCHEIMBERG, S. *Algoritmos com Informações de Segunda Ordem e Métrica Variável para Funções Convexas Não Diferenciáveis*. Ph.D. Th. Instituto de Matemática, Universidade Federal do Rio de Janeiro (UFRJ). Rio de Janeiro. Brazil. (1985).
- [SCH11] SCHITTKOWSKI, K. *A Fortran Subroutine Solving Constrained Nonlinear Programming Problems*. Annals of Operations Research, Vol. 5, pp. 485-500. (1985/6).
- [SCHR1] SCHRAMM, H. *Eine Kombination von Bundle-und Trust Region-Verfahren zur Lösung nichtdifferenzierbarer Optimierungsprobleme*. Bayreuther Mathematischen Schriften, Heft 30, Bayreuth. Germany, 1989.

[SCHRZ1] SCHRAMM, H.; ZOWE, J. *A Version of the Bundle Idea for Minimizing a Nonsmooth Function: Conceptual Idea, Convergence Analysis, Numeric Results*. SIAM J. Optimization, Vol. 2, N. 1, pp. 121-152. (February, 1992).