

ANA PAULA CHOROBURA

CONDIÇÕES DE OTIMALIDADE PARA PROBLEMAS  
COM UM E COM VÁRIOS OBJETIVOS: ABORDAGEM  
ATRAVÉS DO FORMALISMO DE  
DUBOVITSKII-MILYUTIN

Curitiba  
26 de Fevereiro de 2015

ANA PAULA CHOROBURA

CONDIÇÕES DE OTIMALIDADE PARA PROBLEMAS  
COM UM E COM VÁRIOS OBJETIVOS: ABORDAGEM  
ATRAVÉS DO FORMALISMO DE  
DUBOVITSKII-MILYUTIN

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Mestre em Matemática.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Lucelina Batista dos Santos.

Curitiba  
26 de Fevereiro de 2015

---

C551c

Chorobura, Ana Paula

Condições de otimalidade para problemas com um e com vários objetivos : abordagem através do formalismo de Dubovitskii-Milyutin/ Ana Paula Chorobura. – Curitiba, 2015.

74 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-graduação em Matemática, 2015.

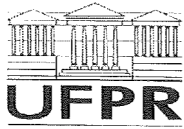
Orientador: Lucelina Batista dos Santos .

Bibliografia: p. 73-74.

1. Otimização matemática. 2. Processo decisório por critério múltiplo. 3. Programação não linear. 4. Domínios convexos. I. Universidade Federal do Paraná. II.Santos, Lucelina Batista dos. III. Título.

CDD: 519.6

---



## ATA DA 63ª DEFESA DE DISSERTAÇÃO DE MESTRADO

Aos vinte e seis dias do mês de fevereiro de 2015, na sala de Reuniões do Setor de Ciências Exatas, foi instalada pela Professora Lucelina Batista dos Santos, a Banca Examinadora para a sexagésima terceira Defesa de Dissertação de Mestrado em Matemática. Estiveram presentes ao Ato, professores, alunos e visitantes.

A banca examinadora, homologada pelo Colegiado do Programa de Pós-Graduação em Matemática, ficou constituída pelos professores: Prof. Dr. Marko Antonio Rojas Medar, da Universidad del Bio-Bio, Prof. Dr. Valeriano Antunes de Oliveira, da Universidade Estadual Paulista, Prof. Lucas Garcia Pedroso, do Programa de Pós-Graduação em Matemática da Universidade Federal do Paraná e a Profa. Dra. Lucelina Batista dos Santos, orientadora da dissertação, a quem coube a presidência dos trabalhos.

Às dez horas, a banca iniciou seus trabalhos, convidando a candidata **ANA PAULA CHOROBURA** a fazer a apresentação do tema da dissertação intitulada "CONDIÇÕES DE OTIMALIDADE PARA PROBLEMAS COM UM E COM VÁRIOS OBJETIVOS: ABORDAGEM ATRAVÉS DO FORMALISMO DE DUBOVITSKII-MILYUTIN". Encerrada a apresentação, iniciou-se a fase de arguição pelos membros participantes. Após a arguição, a banca com pelo menos 03 (três) membros, reuniu-se para apreciação do desempenho do pós-graduando.

A banca considerou que o pós-graduando fez uma apresentação com a necessária concisão. A Dissertação apresenta contribuição à área de estudos e não foram registrados problemas fundamentais de estrutura e redação, resultando em plena e satisfatória compreensão dos objetivos pretendidos.

Tendo em vista a dissertação e a arguição, os membros presentes da banca decidiram pela sua aprovação.

Curitiba, 26 de fevereiro de 2015.

Prof. Dra. Lucelina Batista dos Santos  
Presidente

Prof. Dr. Marko Antonio Rojas-Medar  
Titular

Prof. Dr. Valeriano Antunes de Oliveira  
Titular

Prof. Dr. Lucas Garcia Pedroso  
Titular



Ministério da Educação  
Universidade Federal do Paraná  
Setor de Ciências Exatas/Departamento de Matemática  
Programa de Pós-Graduação em Matemática - PPGM

## PARECER DA BANCA EXAMINADORA

Após a apresentação, a banca deliberou pela aprovação da dissertação da candidata **ANA PAULA CHOROBURA** devendo, para tanto, incorporar as sugestões feitas pelos membros da banca, no prazo estabelecido pelo regimento correspondente.

Curitiba, 26 de fevereiro de 2015.

\_\_\_\_\_  
Prof. Dra. Lucelina Batista dos Santos  
Presidente

\_\_\_\_\_  
Prof. Dr. Marko Antonio Rojas-Medar  
Titular

\_\_\_\_\_  
Prof. Dr. Valeriano Antunes de Oliveira  
Titular

\_\_\_\_\_  
Prof. Dr. Lucas Garcia Pedroso  
Titular

TERMO DE APROVAÇÃO

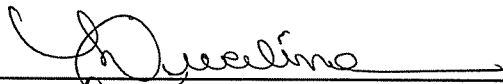
“CONDIÇÕES DE OTIMALIDADE PARA PROBLEMAS COM UM E COM  
VÁRIOS OBJETIVOS: ABORDAGEM ATRAVÉS DO FORMALISMO DE  
DUBOVITSKII-MILYUTIN”

por


**Ana Paula Chorobura**

Dissertação aprovada como requisito parcial para obtenção do grau de  
Mestre no Programa de Pós-Graduação em Matemática,  
pela Comissão Examinadora composta por:


Orientadora:




Profa. Dra. Lucelina Batista dos Santos  
Dep. de Matemática – UFPR

Prof. Dr. Marko Antonio Rojas-Medar  
Universidad del Bio-Bio

Prof. Dr. Valeriano Antunes de Oliveira  
UNESP

Prof. Dr. Lucas Garcia Pedroso  
Dep. de Matemática – UFPR

Curitiba, 26 de fevereiro de 2015.

*Aos meus pais, Luiz e Maria, pelo amor, incentivo e apoio incondicional.*

# Agradecimentos

Agradeço primeiramente a Deus, por estar comigo em todos os momentos de minha vida, me mostrando o caminho a seguir nas horas difíceis.

Aos meus pais Luiz e Maria, e irmãos André e Flávia, por estarem sempre ao meu lado, com muito carinho e apoio, não medindo esforços para que eu chegasse até esta etapa da minha vida.

A minha orientadora pela compreensão, dedicação e incentivo dispensado ao desenvolvimento deste trabalho.

Aos meus amigos do PPGM pela convivência solidária ao compartilhar as apreensões e ansiedades, pelas ricas trocas de experiências e, enfim, pela amizade.

As amigas Ana Claudia, Simone e Madalena pela confiança, apoio e amizade ao longo destes dois anos e aos demais amigos que me apoiaram.

Aos professores do programa que foram grandes incentivadores de meu crescimento pessoal e profissional.

Aos membros da banca, professores Marko, Valeriano e Lucas, obrigada por aceitarem o convite.

Ao Programa de Pós-Graduação em Matemática da UFPR pela oportunidade e formação de qualidade propiciada.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPQ, pelo apoio financeiro.

A todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

# Resumo

Esta dissertação apresenta uma revisão da teoria desenvolvida por Dubovitskii e Milyutin sobre condições de otimalidade para problemas de otimização com um ou mais objetivos. Os resultados de Dubovitskii e Milyutin permitem obter condições de otimalidade para problemas de otimização em espaços de Banach e podem ser aplicados em diferentes áreas. Como aplicação, a teoria é usada para deduzir as condições clássicas de Fritz-John e de Karush-Kuhn-Tucker para problemas de programação não-linear, bem como o Princípio do Máximo de Pontryagin para problemas de controle ótimo mono-objetivo.

**Palavras-chave:** Formalismo de Dubovitskii-Milyutin; Condições de Otimalidade; Otimização Multiobjetivo; Programação Não Linear; Controle Ótimo; Convexidade.

# Abstract

This dissertation presents a review of the theory developed by Dubovitskii and Milyutin about optimality conditions for optimization problems with one or several objectives. The results of Dubovitskii and Milyutin allow to obtain optimality conditions for optimization problems on Banach spaces and can be applied in different areas. As application, the theory is used to deduce the classical conditions of Fritz-John and of Karush-Kuhn-Tucker for nonlinear programming problems, as well the Maximum Principle of Pontryagin for optimal control problems with one objective.

**Keywords:** Dubovitski-Milyutin Formalism; Optimality Conditions; Multiobjective Optimization; Nonlinear Programming Problems; Optimal Control; Convexity.

# Sumário

<b>Introdução</b>	<b>1</b>
<b>1 Preliminares técnicos</b>	<b>3</b>
1.1 Cones e cones duais . . . . .	3
1.2 Cálculo Diferencial . . . . .	8
1.3 Análise Convexa . . . . .	11
<b>I Problema Mono-objetivo</b>	<b>14</b>
<b>2 Formalismo de Dubovitskii-Milyutin</b>	<b>15</b>
2.1 Condições necessárias de otimalidade . . . . .	15
2.2 Cálculo dos cones e de seus duais . . . . .	20
2.2.1 Direções de descida . . . . .	20
2.2.2 Direções factíveis . . . . .	23
2.2.3 Direções tangentes . . . . .	25
2.2.4 Cones Duais . . . . .	27
2.3 Condições suficientes de otimalidade . . . . .	29
<b>3 Aplicações</b>	<b>31</b>
3.1 Problemas de programação não linear . . . . .	31
3.2 Problemas de Controle Ótimo . . . . .	36
<b>II Problema Multiobjetivo</b>	<b>48</b>
<b>4 Formulação do problema, conceitos de solução e escalarização</b>	<b>49</b>
4.1 Formulação do problema e conceitos de solução . . . . .	49
4.2 Escalarização . . . . .	51
<b>5 Formalismo Dubovitskii-Milyutin Multiobjetivo: solução Pareto fraca</b>	<b>53</b>
5.1 Condições Necessárias e Suficientes para otimalidade de Pareto fraca . . . . .	53
5.2 Aplicação a Problemas de Programação Matemática Vetorial . . . . .	55
<b>6 Formalismo de Dubovitskii-Milyutin Multiobjetivo: solução de Pareto</b>	<b>57</b>
6.1 Condições Necessárias para a otimalidade de Pareto . . . . .	57
6.2 Caracterização dos cones duais . . . . .	59
6.3 Condições necessárias e suficientes para otimalidade global de Pareto . . . . .	64
6.4 Escalarização de Problemas de Otimização Multiobjetivo . . . . .	67
6.5 Problemas de programação matemática vetorial . . . . .	70

Conclusão	72
Referências Bibliográficas	73

# Introdução

Problemas com o objetivo de otimizar funções sempre foram objeto de interesse entre os matemáticos. Em 1962, Dubovitskii e Milyutin encontraram condições necessárias de otimalidade, na forma de uma equação estabelecida na linguagem de análise funcional, através da qual podemos determinar condições necessárias de otimalidade para uma ampla classe de problemas. Devido a isso um grande número de trabalhos na área de Otimização fazem uso do formalismo de Dubovitskii-Milyutin para obter condições de otimalidade, tanto para problemas mono-objetivo quanto para multiobjetivo, diferenciáveis ou não diferenciáveis, de dimensão finita ou infinita. Alguns deles são [4, 8, 16, 18, 19, 28].

Neste estudo consideraremos o seguinte problema de Otimização

$$\begin{aligned} &\text{Minimizar } f(x) \\ &\text{sujeito a } x \in Q = \bigcap_{i=1}^{n+1} Q_i \end{aligned}$$

onde  $f : X \rightarrow \mathbb{R}$ , sendo  $X$  espaço de Banach,  $Q_i \subset X$ ,  $i = 1, \dots, n$  são conjuntos com interior não vazio e  $Q_{n+1} \subset X$  possivelmente não possui pontos interiores.

Muitos problemas de Otimização podem ser estudados através deste modelo geral. Alguns exemplos estudados neste trabalho são problemas de programação não linear e de controle ótimo.

Problemas de programação matemática são problemas que exigem calcular o máximo ou mínimo de alguma função objetivo, usualmente uma função de variáveis vetoriais sujeita ou não a restrições.

O problema de controle consiste basicamente em otimizar um funcional, sujeito a um sistema que evolui no tempo. O funcional a ser minimizado ou maximizado depende de uma variável que descreve o estado do sistema em cada momento e de uma variável de controle que intervém no comportamento do sistema. Então procura-se controlar o sistema de maneira ótima ao longo do tempo determinado, de acordo com o objetivo previamente fixado.

Um dos resultados principais da teoria de Controle Ótimo é o chamado Princípio do Máximo de Pontryagin, o qual proporciona condições necessárias de otimalidade para tais problemas.

Por outro lado, nos deparamos com problemas de tomadas de decisões que surgem em diversas áreas como Engenharia, Economia, Administração, entre outros. Tais problemas geralmente possuem mais de um objetivo fixado como meta pelo decisor. Assim surgiu o interesse pelo estudo de Problemas de Otimização com vários objetivos. Formalmente, o problema de Otimização multiobjetivo admite a seguinte formulação:

$$\begin{aligned} &\text{Minimizar } f(x) = (f_1(x), \dots, f_s(x)) \\ &\text{sujeito a } x \in Q \end{aligned} \tag{MP}$$

onde  $f : X \rightarrow \mathbb{R}^s$ , e  $Q$  é um subconjunto não vazio do espaço de Banach  $X$ .

Geralmente, não é possível minimizar simultaneamente todos os objetivos  $f_i$  sujeitos à restrição  $Q$ , o que leva a várias noções de otimalidade para o problema de Otimização (MP).

Um dos precursores no campo da Otimização Multiobjetivo é Pareto, que em seu célebre trabalho “*Cours d’Economie Politique*” [23] introduz o conceito de solução eficiente. Conceitualmente, um ponto é chamado eficiente ou de Pareto quando não é possível melhorar nenhum objetivo sem piorar algum outro.

Desde então, têm aparecido várias publicações, tanto de caráter teórico quanto numérico, nas quais se desenvolvem critérios que permitem decidir quando um ponto viável é ou não eficiente. Merecem destaque as noções de eficiência fraca e de eficiência própria. Intuitivamente, um ponto é chamado fracamente eficiente se não é possível melhorar todos os objetivos simultaneamente e é chamado propriamente eficiente se é solução eficiente e se os quocientes entre o ganho em um objetivo e a perda com respeito aos demais objetivos é limitada.

Um dos métodos de resolver problemas de Otimização multiobjetivo é procurar relacioná-lo com problemas escalares, cuja teoria e aspectos computacionais são bastante desenvolvidos.

Esta Dissertação, se divide em duas partes, a primeira se concentra no estudo da teoria de Dubovitskii-Milyutin para problemas de Otimização com um objetivo e a segunda parte trata de estudar como tal teoria pode ser adequada aos problemas de Otimização com vários objetivos.

No Capítulo 1, apresentamos alguns resultados de Análise Funcional e Análise Convexa fundamentais no desenvolvimento da Dissertação.

No Capítulo 2, estudamos a teoria de Otimização desenvolvida por Dubovitskii e Milyutin, a qual nos fornece condições necessárias de otimalidade, e sob hipóteses adequadas de convexidade e regularidade obtemos também condições suficientes de otimalidade. Para formular tais condições mediante este formalismo caracterizamos o cone de direções de descida do funcional a ser minimizado, o cone de direções factíveis referente aos conjuntos de restrições que possuem pontos interiores e o cone de direções tangentes associado aos conjuntos de restrições que não possuem pontos interiores, no ponto ótimo, e seus respectivos cones duais.

No Capítulo 3, a teoria é utilizada para estabelecer as condições clássicas de Fritz-John e Karush-Kuhn-Tucker para problemas de programação não linear e para obtenção do Princípio do Máximo de Pontryagin para problemas de controle ótimo.

A formulação do problema de Otimização Multiobjetivo, bem como uma discussão sobre os conceitos de solução e alguns resultados conhecidos sobre escalarização são feitas no Capítulo 4.

No Capítulo 5, se faz uma extensão da teoria de Dubovitskii-Milyutin para o estudo de condições de otimalidade para problemas multiobjetivos envolvendo o conceito de solução fracamente eficiente ou Pareto fraca e sua aplicação para obtenção do Teorema de Fritz-John generalizado.

No Capítulo 6 discutimos a otimalidade de Pareto para problemas multiobjetivos, baseada na teoria de Dubovitskii-Milyutin. Para determinar soluções de Pareto além dos cones de direções de descida, factível e tangente necessitamos do cone de direções de não crescimento dos funcionais a serem minimizados.

# Capítulo 1

## Preliminares técnicos

Apresentaremos algumas definições e teoremas de Análise Funcional e Análise Convexa que serão necessários no desenvolvimento da dissertação. Os interessados em aprofundar os temas tratados neste capítulo podem consultar [1, 2, 3, 8, 24, 29].

Neste capítulo, o conjunto  $X$  é um espaço de Banach e  $x_0$  é um ponto de  $X$ , fixado. Quando isto não ocorrer deixaremos explícito no texto. Além disso denotaremos por  $X'$  o espaço dual topológico do espaço de Banach  $X$ , constituído de todos os funcionais lineares contínuos sobre  $X$ . O espaço  $X'$  é um espaço de Banach com a norma

$$\|f\| = \sup\{|f(x)|, x \in X, \|x\| \leq 1\}.$$

### 1.1 Cones e cones duais

Um conjunto  $C \subset X$  é dito **convexo** quando dados  $x, y \in C$  o segmento  $[x, y] = \{(1-t)x + ty : t \in [0, 1]\}$  estiver inteiramente contido em  $C$ .

Na Figura 1.1 ilustramos dois conjuntos, um convexo e outro não.

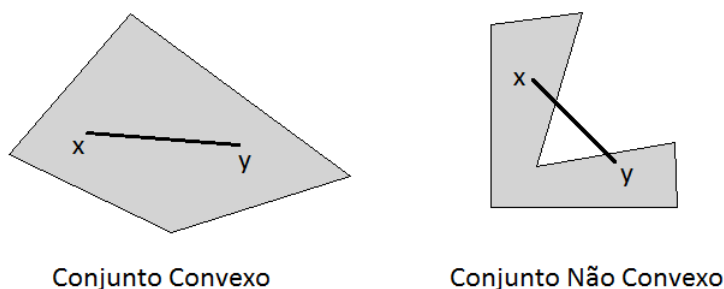


Figura 1.1: Conjuntos convexo e não convexo.

Uma classe de funções que tem ótimas propriedades, particularmente no contexto de otimização, é a classe das **funções convexas**. Seja  $C \subset X$  um conjunto convexo. Dizemos que a função  $f : X \rightarrow \mathbb{R}$  é convexa em  $C$  quando

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

para todos  $x, y \in C$  e  $t \in [0, 1]$ .

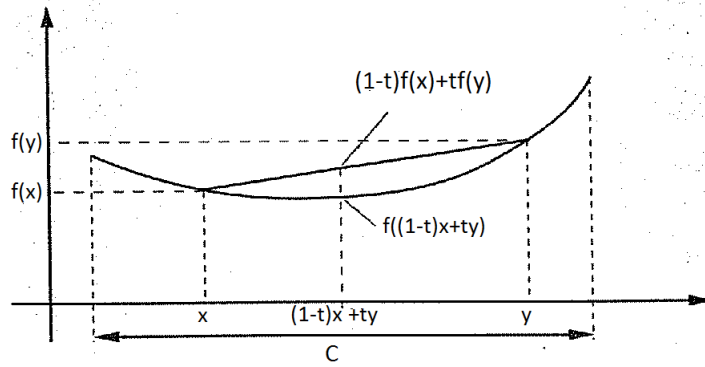


Figura 1.2: Função convexa.

Geometricamente, podemos dizer que qualquer arco no gráfico da função convexa está sempre abaixo do segmento que liga as extremidades. Veja na Figura 1.2 um exemplo de função convexa.

Um subconjunto  $K \subset X$  é um **cone** com vértice na origem quando para todo  $\lambda \geq 0$  e  $h \in K$  tem-se  $\lambda h \in K$ . Se  $K$  é um cone com vértice na origem, então  $x_0 + K$  é um cone com vértice em  $x_0$ . Um cone  $K$  é um **cone convexo**, se dados  $y_1, y_2 \in K$ , então  $y_1 + y_2 \in K$ .

Agora introduziremos o conceito de cone dual, que será utilizado neste trabalho. Seja  $K$  um cone em  $X$ , então o **cone dual**  $K^* \subset X'$  é definido como o conjunto de todos os funcionais lineares contínuos não negativos em  $K$ , ou seja,

$$K^* = \{f \in X' : f(x) \geq 0, \forall x \in K\}.$$

Por definição  $K^*$  é um cone convexo com vértice em 0. Para  $K = \emptyset$ ,  $K^* = X'$ , pois caso contrário, deveria existir um funcional  $f \in X'$  e algum  $x \in K$  tal que  $f(x) < 0$ , o que é um absurdo.

**Exemplo 1.1** Seja  $X = \mathbb{R}^n$  e  $K = \mathbb{R}_+^n$ , ou seja,

$$K = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1, \dots, x_n \geq 0\}.$$

Então  $K$  é um cone fechado em  $X$  com  $K^* = K$ .

**Exemplo 1.2** Considere o Espaço de Banach  $X = \mathbb{R}^2$  e o cone  $K$  dado por

$$K = \{x = (x_1, x_2) : a^T x \geq 0, b^T x \geq 0, \text{ com } a, b \in \mathbb{R}^2 \text{ fixos}\} \subset \mathbb{R}^2$$

então

$$K^* = \{f \in \mathbb{R}^2 : f = \lambda_1 a + \lambda_2 b, \lambda_1, \lambda_2 \geq 0\}.$$

Veja Figura 1.3.

Uma consequência imediata da definição é dada no seguinte lema.

**Lema 1.3** Sejam  $K_1$  e  $K_2$  cones em  $X$ . Se  $K_1 \subset K_2$ , então  $K_2^* \subset K_1^*$ .

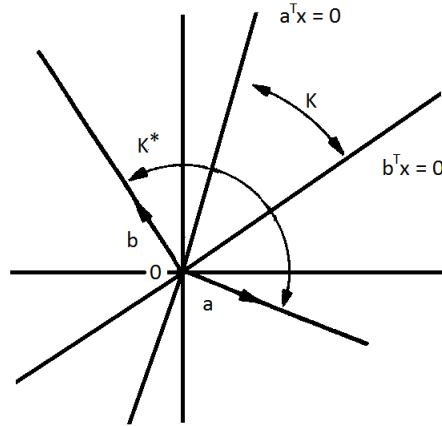


Figura 1.3: Cone  $K$  e seu cone dual  $K^*$  em  $\mathbb{R}^2$ .

*Demonstração.* Seja  $f \in K_2^*$ , então  $f(x) \geq 0, \forall x \in K_2$ . Como  $K_1 \subset K_2$ , temos  $f(x) \geq 0, \forall x \in K_1$  e portanto  $f \in K_1^*$ .  $\square$

Antes de enunciarmos os próximos teoremas recordemos algumas definições.

**Definição 1.4** *Seja  $C$  um subconjunto não vazio de  $X$ .*

- *Se  $C$  é fechado na topologia da norma, isto é, se dada uma sequência  $(x_n) \subset X$  tal que  $\|x_n - x\|_X \rightarrow 0$  quando  $n \rightarrow \infty$ , temos que  $x \in C$ , então diz-se que  $C$  é **fortemente fechado**.*
- *Diz-se que  $C$  é **fracamente fechado**, se ele for fechado na topologia fraca, ou seja, se dada uma sequência  $(x_n) \subset C$ , tal que  $\lim_{n \rightarrow \infty} f(x_n) = f(x), \forall f \in X'$ , então  $x \in C$ .*
- *Diz-se que  $C$  é fechado na topologia fraca\* ou  $C$  é **fracamente\* fechado** se dada uma sequência  $(x_n) \subset C$ , com  $\lim_{n \rightarrow \infty} f(x_n) = f(x)$  para todo  $f \in J(x) \subset X''$ , onde  $J : X \rightarrow X''$  é a imersão canônica, então  $x \in C$ .*

Em espaços normados de dimensão finita a topologia da norma, fraca e fraca\* são equivalentes. Além disso a topologia fraca\* é mais fraca que a topologia fraca de  $X'$  e portanto qualquer conjunto fracamente fechado em  $X'$  é fraco\* fechado; estas topologias são equivalentes quando  $X$  é um espaço de Banach reflexivo.

**Teorema 1.5** *Seja  $C$  um conjunto convexo em  $X$ . Então  $C$  é fortemente fechado se, e somente se, ele é fracamente fechado.*

*Demonstração.* Consulte Brezis [3, Teorema III.7, p.38].  $\square$

**Teorema 1.6** *Sejam  $K_i, i = 1, \dots, n$ , cones convexos, fracamente fechados e  $\sum_{i=1}^n K_i^*$  fraco\* fechado, então*

$$\left( \bigcap_{i=1}^n K_i \right)^* = \sum_{i=1}^n K_i^*.$$

Antes de demonstrar este Teorema enunciemos o Lema abaixo, cuja demonstração pode ser encontrada em Girsanov [8, p. 34 e 35].

**Lema 1.7** *Sejam  $K_i \subset X$  cones, para  $i = 1, \dots, n$ , então*

- $(\bigcup_{i=1}^n K_i)^* = \bigcap_{i=1}^n K_i^*$ ;
- Se  $K_i$  são convexos para  $i = 1, \dots, n$  então  $\bigcup_{i=1}^n K_i^* = \sum_{i=1}^n K_i^*$ ;
- Se  $K$  é fracamente fechado então  $K^{**} = K$ .

*Demonstração do Teorema 1.6*

Seja  $Q = \sum_{i=1}^n K_i^*$  então

$$\begin{aligned} Q^* &= \left( \sum_{i=1}^n K_i^* \right)^* = \left( \bigcup_{i=1}^n K_i^* \right)^* \\ &= \bigcap_{i=1}^n K_i^{**} \\ &= \bigcap_{i=1}^n K_i. \end{aligned}$$

Assim  $Q^{**} = (\bigcap_{i=1}^n K_i)^*$ . Mas  $Q^{**} = Q$  e portanto segue o resultado.

**Teorema 1.8** [Teorema de Krein] *Sejam  $\hat{K}$  um cone convexo com vértice na origem, o qual contém pontos interiores, e  $\hat{L}$  um subespaço tal que  $\text{int}\hat{K} \cap \hat{L} \neq \emptyset$ . Seja  $\bar{g}(x)$  uma forma linear sobre  $\hat{L}$  tal que  $\bar{g}(x) \geq 0$  sobre  $\hat{K} \cap \hat{L}$ . Então existe uma forma linear contínua  $g(x)$  sobre o espaço de Banach  $X$ , de tal maneira que*

$$\begin{cases} g(x) = \bar{g}(x) \text{ para todo } \tilde{x} \in \hat{L}; \\ g(x) \geq 0 \text{ para todo } \tilde{x} \in \hat{K}. \end{cases}$$

*Demonstração.* A demonstração deste teorema pode ser encontra em Girsanov [8, p. 32].

□

**Lema 1.9** *Sejam  $K_i$ ,  $i = 1, \dots, n$  cones abertos convexos,  $\bigcap_{i=1}^n K_i \neq \emptyset$ . Então*

$$\left( \bigcap_{i=1}^n K_i \right)^* = \sum_{i=1}^n K_i^*.$$

*Demonstração.* Definamos o espaço  $Y = \prod_{i=1}^n X$ , onde  $X$  é Banach, isto é, o espaço

$$Y = \{\tilde{x} = (x_1, \dots, x_n) : x_i \in X, i = 1, \dots, n\}.$$

Assim

$$Y' = \{f = (f_1, \dots, f_n) : f_i \in X', i = 1, \dots, n\}.$$

Deste modo, podemos dizer que cada  $F \in Y'$  tem a representação

$$F(\tilde{x}) = \sum_{i=1}^n f_i(x_i), \text{ para todo } \tilde{x} = (x_1, \dots, x_n), \quad (1.1)$$

onde  $f_i \in X'$ , para todo  $i = 1, \dots, n$ .

Consideremos ainda, o conjunto

$$K = \{\tilde{x} = (x_1, \dots, x_n) : x_i \in K_i, i = 1, \dots, n\} \subset Y.$$

Aqui,  $K$  é um cone convexo aberto, porque é o produto direto de cones convexos abertos ( $K = \prod_{i=1}^n K_i$ ).

Definamos  $L = \{\tilde{x} = (x, \dots, x) : x \in X\} \subset Y$ , onde  $L$  é um subespaço vetorial de  $Y$ . Por hipótese, existe  $x_0 \in K_i$ ,  $i = 1, \dots, n$ , logo  $L \cap K \neq \emptyset$ .

Seja  $f \in (\prod_{i=1}^n K_i)^*$ . Tomemos o funcional linear  $\bar{f}$  definido sobre  $L$ , como segue

$$\bar{f}(\tilde{x}) = f(x), \quad (1.2)$$

onde  $\tilde{x} = (x, \dots, x) \in L$ .

Seja  $\tilde{x}_0 = (x_0, \dots, x_0) \in L \cap K$ . Como  $x_0 \in \bigcap_{i=1}^n K_i$ , então  $f(x_0) \geq 0$  e portanto  $\bar{f}(\tilde{x}_0) \geq 0$ , para todo  $\tilde{x}_0 \in L \cap K$ .

Usando o fato de  $\text{int}K = K$ , pois  $K$  é aberto, temos que  $L \cap \text{int}K \neq \emptyset$ . Deste modo, podemos aplicar o Teorema de Krein (Teorema 1.8) e desta forma conseguimos a existência de  $F = (f_1, \dots, f_n) \in Y'$ , tal que

$$F(\tilde{x}) \geq 0, \quad \forall \tilde{x} \in K; \quad (1.3)$$

$$F(\tilde{x}) = \bar{f}(\tilde{x}), \quad \forall \tilde{x} \in L. \quad (1.4)$$

Assim, para todo  $\tilde{x}_0 \in L$ , temos por (1.1) que  $F(\tilde{x}_0) = \sum_{i=1}^n f_i(x_0)$ . Por (1.4)  $F(\tilde{x}_0) = \bar{f}(\tilde{x}_0)$  e por (1.2)  $\bar{f}(\tilde{x}_0) = f(x_0)$ . Logo,

$$\bar{f}(\tilde{x}_0) = f(x_0) = \sum_{i=1}^n f_i(x_0), \quad \forall \tilde{x}_0 \in L.$$

Então, para  $\tilde{x}_0 = (x_0, \dots, x_0)$ ,  $x_0 \in \bigcap_{i=1}^n K_i$ , também teremos

$$f(x_0) = \sum_{i=1}^n f_i(x_0), \quad \text{o que equivale a}$$

$$f = \sum_{i=1}^n f_i, \quad \text{onde } f \in \left( \bigcap_{i=1}^n K_i \right)^*.$$

Devido a (1.3), segue que  $F(\tilde{x}) \geq 0$ , onde  $F(\tilde{x}) = \sum_{i=1}^n f_i(x_i)$ , para todo  $x_i \in K_i$ ,  $i = 1, \dots, n$  e  $\tilde{x} = (x_1, \dots, x_n)$ . Consequentemente,  $f_i(x_i) \geq 0$ ,  $\forall x_i \in K_i$ , ou seja,  $f_i \in K_i^*$ ,  $i = 1, \dots, n$ .

Agora de acordo com a análise feita para  $\tilde{x}_0$  onde  $x_0 \in \bigcap_{i=1}^n K_i$ , podemos usar o resultado obtido acima e concluirmos que  $f(x_0) = \sum_{i=1}^n f_i(x_0)$ , com  $f_i \in K_i^*$ , para  $i = 1, \dots, n$ . Portanto, dado  $f \in (\prod_{i=1}^n K_i)^*$ , existem  $f_i \in K_i^*$ ,  $i = 1, \dots, n$ , tais que  $f = \sum_{i=1}^n f_i$ . Em outras palavras,  $(\prod_{i=1}^n K_i)^* \subset \sum_{i=1}^n K_i^*$ .

Por outro lado, seja  $f \in \sum_{i=1}^n K_i^*$ , isto é,  $f = \sum_{i=1}^n f_i$ , com  $f_i \in K_i^*$  para todo  $i = 1, \dots, n$ . Se  $x_0 \in \bigcap_{i=1}^n K_i$ , segue que  $f_i(x_0) \geq 0$ , para  $i = 1, \dots, n$ . Logo,  $f(x_0) \geq 0$ ,  $\forall x_0 \in \bigcap_{i=1}^n K_i$ . Assim,  $f \in (\prod_{i=1}^n K_i)^*$  e concluímos que  $\sum_{i=1}^n K_i^* \subset (\prod_{i=1}^n K_i)^*$ .  $\square$

**Teorema 1.10** *Seja  $X$  um espaço de Banach de dimensão finita. Então:*

- a)  $K^*$  é fechado.
- b)  $K^* = (\text{cl}K)^*$ , ou seja o cone dual de um cone e de seu fecho são iguais.

*Demonstração.*

- a) Seja  $(f_n) \subset K^*$ , com  $\lim_{n \rightarrow \infty} f_n = f$ . Fixando  $x \in K$  arbitrário e passando o limite quando  $n \rightarrow \infty$  em  $f_n(x) \geq 0$ , obtemos  $f(x) \geq 0$ . Portanto, como  $x \in K$  era arbitrário,  $f \in K^*$  o que mostra que  $K^*$  é fechado.
- b) Temos que  $K \subset \text{cl}K$ , logo pelo Lema 1.3  $(\text{cl}K)^* \subset K^*$ . Por outro lado, seja  $f \in K^*$  e  $x \in \text{cl}K$  arbitrários. Então existe uma sequência  $(x_n) \subset K$  tal que  $\lim_{n \rightarrow \infty} x_n = x$ . Agora passando o limite quando  $n \rightarrow \infty$  em  $f(x_n) \geq 0$ , como  $f$  é contínuo, obtemos  $f(x) \geq 0$ . Portanto  $f \in (\text{cl}K)^*$ , o que completa a prova.

□

**Definição 1.11** Dizemos que um funcional linear contínuo não nulo  $f$  **separa** dois conjuntos  $A \subset X$  e  $B \subset X$  se existe  $\alpha \in \mathbb{R}$  tal que

$$f(x) \leq \alpha \leq f(y), \quad \forall x \in A, \quad \forall y \in B.$$

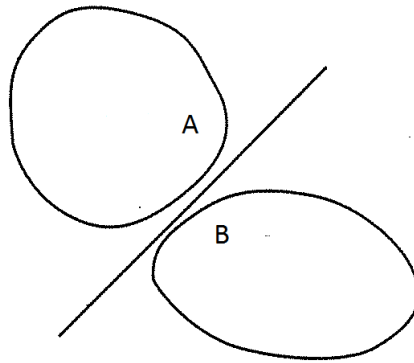


Figura 1.4: Hiperplano separando dois conjuntos  $A$  e  $B$ .

**Teorema 1.12** Se os conjuntos  $A \subset X$  e  $B \subset X$  são convexos, não vazios e disjuntos, e  $A$  é aberto, então existe um funcional linear contínuo não nulo separando  $A$  e  $B$

*Demonstração.* A demonstração deste teorema pode ser encontrada em Alekseev et al. [1, p. 77]. □

## 1.2 Cálculo Diferencial

**Definição 1.13** Um funcional  $f(x)$  em um espaço de Banach  $X$  é dito **derivável** em um ponto  $x_0$  **segundo a direção**  $h$  se

$$\lim_{\varepsilon \rightarrow 0^+} \frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon}$$

existe; neste caso, este limite será denotado por  $f'(x_0, h)$ .

Se  $X = \mathbb{R}$  então  $f(x)$  é uma função de uma variável, assim a existência de  $f'(x_0, h)$  é equivalente a existência de derivada da direita (para  $h > 0$ ) ou esquerda (para  $h < 0$ ) de  $f(x)$  no ponto  $x_0$ .

**Definição 1.14** *Seja  $f : X \rightarrow Y$ , sendo  $X, Y$  espaços de Banach. A função  $f(x)$  é dita **diferenciável** ou **Fréchet-diferenciável** em um ponto  $x_0$  se existe um operador linear  $\Lambda : X \rightarrow Y$  tal que para todo  $h \in X$ ,*

$$f(x_0 + h) = f(x_0) + \Lambda(h) + o(\|h\|).$$

O operador linear  $\Lambda$  é denotado por  $f'(x_0)$  e é chamado de derivada ou derivada de Fréchet da função  $f(x)$  no ponto  $x_0$ .

**Definição 1.15** *Seja  $f : X \rightarrow Y$ , sendo  $X, Y$  espaços de Banach. A função  $f(x)$  é dita **estritamente diferenciável** no ponto  $x_0$ , se existe um operador linear  $\Lambda : X \rightarrow Y$  tal que para todo  $\varepsilon > 0$  existe  $\delta > 0$  tais que para todos  $x_1$  e  $x_2$  satisfazendo  $\|x_1 - x_0\| < \delta$  e  $\|x_2 - x_0\| < \delta$ , tem-se a seguinte desigualdade:*

$$\|f(x_1) - f(x_2) - \Lambda(x_1 - x_2)\| \leq \|x_1 - x_2\|. \quad (1.5)$$

Colocando  $x_2 = x_0$  e  $x_1 = x_0 + h$  em (1.5), obtemos que a função é diferenciável e  $\Lambda = f'(x_0)$ .

**Teorema 1.16** [Teorema do Valor Médio] *Sejam  $X$  e  $Y$  espaços de lineares normados, e seja  $U \subset X$  um conjunto aberto contendo o intervalo  $[a, b]$ . Se  $f : U \rightarrow Y$  é diferenciável no ponto  $x \in [a, b]$  então*

$$\|f(b) - f(a)\| \leq \sup_{c \in [a, b]} \|f'(c)\| \|b - a\|.$$

*Demonstração.* A prova pode ser vista em Alekseev et al. [1, p. 93]. □

**Corolário 1.17** *Suponha que todas as condições do Teorema do Valor Médio são satisfeitas, e seja  $\Lambda$  um operador linear. Então*

$$\|f(b) - f(a) - \Lambda(b - a)\| \leq \sup_{c \in [a, b]} \|f'(c) - \Lambda\| \|b - a\|$$

*Demonstração.* A prova se reduz à aplicação do Teorema do valor médio para a função  $g(x) = f(x) - \Lambda x$ . □

**Corolário 1.18** *Sejam  $X, Y$  espaços de Banach,  $U$  uma vizinhança de  $x_0$  em  $X$  e  $f : U \rightarrow Y$  diferenciável com  $f'(x)$  contínuo no ponto  $x_0$ , então a função  $f$  é estritamente diferenciável em  $x_0$ .*

*Demonstração.* Como  $f'(x)$  é contínua em  $x_0$ , dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$\|x - x_0\| < \delta \Rightarrow \|f'(x) - f'(x_0)\| < \varepsilon. \quad (1.6)$$

Se  $\|x_1 - x_0\| < \delta$  e  $\|x_2 - x_0\| < \delta$ , então para qualquer  $x = x_1 + t(x_2 - x_1) \in [x_1, x_2]$ ,  $0 \leq t \leq 1$ , temos

$$\begin{aligned} \|x - x_0\| &= \|x_1 + t(x_2 - x_1) - x_0\| \\ &= \|t(x_2 - x_0) + (1 - t)(x_1 - x_0)\| \\ &\leq t\|x_2 - x_0\| + (1 - t)\|x_1 - x_0\| \\ &< t\delta + (1 - t)\delta = \delta, \end{aligned}$$

e portanto, em virtude de (1.6), temos  $\|f'(x) - f'(x_0)\| < \varepsilon$ .

Aplicando o Corolário 1.17 a  $\Lambda = f'(x_0)$ , obtemos

$$\|f(x_1) - f(x_2) - f'(x_0)(x_1 - x_2)\| \leq \sup_{x \in [x_1, x_2]} \|f'(x) - f'(x_0)\| \|x_1 - x_2\| \leq \varepsilon \|x_1 - x_2\|$$

que implica que  $f$  é estritamente diferenciável em  $x_0$ . □

**Teorema 1.19** *Sejam  $X$  um espaço topológico,  $Y$  e  $Z$  espaços de Banach,  $W$  uma vizinhança do ponto  $(x_0, y_0)$  em  $X \times Y$ ,  $\psi$  uma função de  $W$  em  $Z$  e  $\psi(x_0, y_0) = z_0$ . Se*

- 1) *a função  $x \rightarrow \psi(x, y_0)$  é contínua no ponto  $x_0$ ;*
- 2) *existe um operador linear contínuo  $\Lambda : Y \rightarrow Z$  tal que, para todo  $\varepsilon > 0$  existe um número  $\delta > 0$  e uma vizinhança  $V$  do ponto  $x_0$  possuindo a propriedade a seguir: a condição  $x \in V$  e as desigualdades*

$$\|y_1 - y_0\| < \delta \text{ e } \|y_2 - y_0\| < \delta$$

*implicam a desigualdade*

$$\|\psi(x, y_1) - \psi(x, y_2) - \Lambda(y_1 - y_2)\| < \varepsilon \|y_1 - y_2\|;$$

- 3)  $\Lambda Y \equiv Z$ ;

*então existem um número  $k > 0$ , uma vizinhança  $U$  do ponto  $(x_0, z_0)$  em  $X \times Z$  e uma aplicação  $\phi : U \rightarrow Y$  tais que*

- a)  $\psi(x, \phi(x, z)) = z$ ;
- b)  $\|\phi(x, z) - y_0\| \leq k \|\psi(x, y_0) - z\|$ .

*Demonstração.* A demonstração deste teorema pode ser encontra em Alekseev et al. [1, p. 103]. □

Lembramos que se  $X$  e  $Y$  são espaços de Banach e se  $A : X \rightarrow Y$  é operador linear contínuo, então o **operador adjunto**  $A^* : Y' \rightarrow X'$  é definido pela igualdade

$$(A^*y, x) = (y, Ax), \quad \forall x \in X \text{ e } y \in Y'.$$

O conjunto

$$L^\perp = \{y \in X' : (y, x) = 0, \quad \forall x \in L\}$$

é chamado de **anulador** do subespaço linear  $L$  de um espaço de Banach  $X$ .

**Lema 1.20** [*Lema da Fatorização*] *Sejam  $X, Y$  e  $Z$  espaços de Banach, e sejam  $A : X \rightarrow Y$  e  $M : X \rightarrow Z$  operadores lineares contínuos tais que  $ImA = Y$  e  $KerA \subset KerM$ . Então existe um operador linear contínuo  $N : Y \rightarrow Z$  tal que  $M = N \circ A$ .*

*Demonstração.* A demonstração deste Lema pode ser encontra em Ioffe et al. [12, p. 15]. □

**Teorema 1.21** *Sejam  $X$  e  $Y$  espaços de Banach, e seja  $A : X \rightarrow Y$  um operador linear contínuo tal que  $ImA = Y$ , então  $(KerA)^\perp = ImA^*$  (isto é, o anulador do núcleo de  $A$  é igual a imagem do operador adjunto de  $A$ ).*

*Demonstração.* Seja  $y \in ImA^*$ , então  $y = A^*z$  para algum  $z \in Y'$ . Logo as seguintes igualdades valem para qualquer  $x \in KerA$

$$(y, x) = (A^*z, x) = (z, Ax) = 0.$$

Segue que  $y \in (KerA)^\perp$ .

Reciprocamente, seja  $y \in (\text{Ker}A)^\perp$ , então  $(y, x) = 0$  para todo  $x \in \text{Ker}A$ . Assim

$$\text{Ker}A \subset \{x \in X : (y, x) = 0\} = \text{Ker}(y).$$

Como  $y$  é operador linear contínuo de  $X$  em  $\mathbb{R}$ , todas as hipóteses do Lema da Fatorização (Lema 1.20) são satisfeitas para  $A : X \rightarrow Y$  e  $y : X \rightarrow \mathbb{R}$ . Portanto existe um funcional linear contínuo  $z \in Y'$  tal que

$$(z, Ax) = (y, x), \forall x \in X.$$

Isso significa que  $y = A^*z$ , isto é  $y \in \text{Im}A^*$ . □

### 1.3 Análise Convexa

Na Seção 1 discutimos algumas definições e propriedades de funções convexas a valores reais. Nesta Seção, discutiremos as propriedades fundamentais das funções convexas, com valores na reta estendida.

Uma função  $f : X \subset \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  é dita **própria** se  $f(x) < +\infty$  para pelo menos um  $x \in X$  e  $f(x) > -\infty$  para todo  $x \in X$ .

Definimos o **epígrafo** de uma função  $f : X \subset \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  como um subconjunto do  $\mathbb{R}^{n+1}$  dado por

$$\text{epi}(f) = \{(x, w) : x \in X, w \in \mathbb{R}, f(x) \leq w\}.$$

Seja  $C$  um subconjunto convexo do  $\mathbb{R}^n$ . Um função  $f : C \rightarrow \overline{\mathbb{R}}$  é dita **convexa** se  $\text{epi}(f)$  é um subconjunto convexo de  $\mathbb{R}^{n+1}$ .

Agora, se  $\text{epi}(f)$  é um conjunto fechado, dizemos que a **função é fechada**.

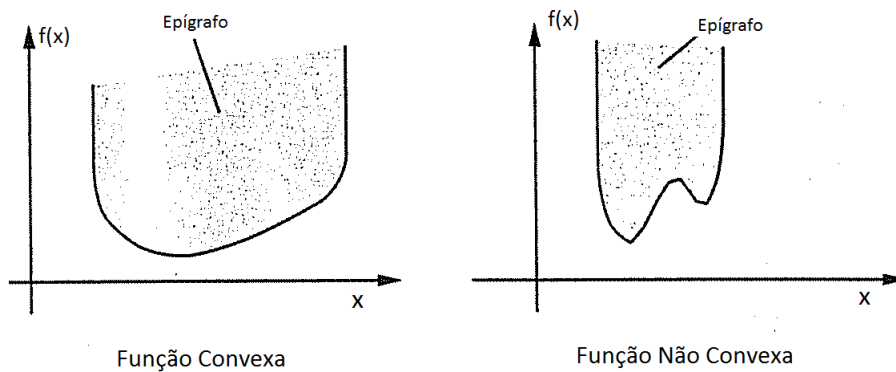


Figura 1.5: Epígrafo de funções convexas e não convexas.

Observamos que para funções convexas a valores reais, esta definição coincide com a definição de função convexa dada na Seção 1.

Um **conjunto afim** em  $\mathbb{R}^n$  é um conjunto  $X$  da forma  $X = \bar{x} + S$ , onde  $\bar{x}$  é um vetor no  $\mathbb{R}^n$  e  $S$  é um subespaço do  $\mathbb{R}^n$ .

Se  $C$  é um subconjunto do  $\mathbb{R}^n$  a **envoltória afim** de  $C$ , denotada por  $\text{aff}(C)$ , é a intersecção de todos conjuntos afim contendo  $C$ .

Seja  $C$  um conjunto convexo não vazio. Dizemos que  $x$  é um **ponto do interior relativo** de  $C$  se  $x \in C$  e existe uma bola aberta  $B$  centrada em  $x$  tal que  $B \cap \text{aff}(C) \subset C$ ,

isto é,  $x$  é um ponto interior de  $C$  relativo a envoltória afim de  $C$ . O conjunto de todos pontos interiores relativo de  $C$  é chamado de **interior relativo** de  $C$  e denotado por  $\text{ri}(C)$ . O conjunto  $C$  é dito relativamente aberto se  $\text{ri}(C) = C$ .

O **cone normal** a um conjunto  $Q$  no ponto  $x_0$  é definido por

$$N(Q; x_0) = \{y \in \mathbb{R}^n; (x - x_0)^T y \leq 0, \forall x \in Q\}.$$

Os resultados abaixo são muito conhecidos em Análise Convexa

**Proposição 1.22**

- a) [Princípio do segmento] Seja  $C$  um conjunto convexo não vazio. Se  $x \in \text{ri}(C)$  e  $\bar{x} \in \text{cl}(C)$  então  $\alpha x + (1 - \alpha)\bar{x} \in \text{ri}(C)$  para  $\alpha \in (0, 1]$ .
- b) Seja  $C$  um conjunto convexo, não vazio em  $\mathbb{R}^n$ . Então  $z \in \text{ri}(C)$  se, e somente se, para todo  $x \in C$ , existe  $\mu > 1$  tal que  $(1 - \mu)x + \mu z \in C$ .
- c) Seja  $C$  um conjunto convexo em  $\mathbb{R}^n$ . Então  $z \in \text{ri}(C)$  se, e somente se, para todo  $y \in \mathbb{R}^n$ , existe  $\varepsilon > 0$  tal que  $z + \varepsilon y \in C$ .
- d) Se  $Q$  é um conjunto convexo e  $\text{int}Q \neq \emptyset$ , então  $\text{cl}(\text{int}Q) = \text{cl}Q$ .
- e) Seja  $C$  um conjunto convexo fechado não vazio que não contém a origem e seja  $K$  um cone convexo gerado por  $C$ . Então

$$\text{cl}K = \cup\{\lambda C : \lambda > 0 \text{ ou } \lambda = 0\}.$$

*Demonstração.* A demonstração dos itens (a) e (b) pode ser encontrada em Bertsekas et al. [2, Proposição 1.4.1, p. 40], (c) em Rockafellar [24, Corolário 6.4.1, p. 47], (d) em Alekseev et al.[1, Proposição 1, p. 139], e (e) em Rockafellar [24, Teorema 9.6, p. 78].  $\square$

A seguir veremos as propriedades de funções convexas não-diferenciáveis. Esta é uma classe de funções muito importante porque a perda da diferenciabilidade acontece com frequência em otimização.

Seja  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  uma função convexa. Dizemos que um vetor  $d \in \mathbb{R}^n$  é um subgradiente de  $f$  no ponto  $x \in \mathbb{R}^n$  se

$$f(z) \geq f(x) + (z - x)^T d, \forall z \in \mathbb{R}^n.$$

O conjunto de todos subgradientes de uma função convexa  $f$  em  $x \in \mathbb{R}^n$  é chamado de subdiferencial de  $f$  em  $x$  e é denotado por  $\partial f(x)$ .

Se  $f$  é diferenciável em  $x_0$  então  $\partial f(x_0) = \{\nabla f(x_0)\}$ , porque o gradiente  $\nabla f(x_0)$  é então o único subgradiente de  $f$  em  $x_0$ .

O subdiferencial é um conjunto convexo e fechado. Além disso  $x_0$  é minimizador de  $f$  em  $\mathbb{R}^n$  se, e somente se,  $0 \in \partial f(x_0)$ . Outra propriedade do subdiferencial é que

$$\partial[\sum_{i=1}^s \lambda_i f_i](x_0) \supset \sum_{i=1}^s \partial[\lambda_i f_i](x_0) = \sum_{i=1}^s \lambda_i \partial f_i(x_0).$$

Para maiores detalhes ver Rockafellar [24] e Bertsekas [2].

**Teorema 1.23** *Seja  $f$  uma função convexa própria. Então*

**a)**  $\partial f(x_0)$  é um conjunto limitado e não vazio se, e somente se  $x_0 \in \text{int}(\text{dom} f)$ .

**b)** para  $\alpha \in \mathbb{R}$ ,  $\alpha > \inf f$ ,

$$\text{cl}\{x \in \text{dom} f : f(x) < \alpha\} = \text{cl}\{x \in \text{dom} f : f(x) \leq \alpha\} = \{x \in \text{dom} f : \text{cl}f(x) \leq \alpha\},$$

$$\text{ri}\{x \in \text{dom} f : f(x) < \alpha\} = \text{ri}\{x \in \text{dom} f : f(x) \leq \alpha\} = \{x \in \text{ri}(\text{dom} f) : f(x) < \alpha\}.$$

**c)** o cone normal para  $C = \{z \in \text{dom} f : f(z) \leq f(x_0)\}$  em  $x_0$  é o fecho do cone convexo gerado por  $\partial f(x_0)$ , onde  $x_0$  é um ponto tal que  $f$  é subdiferenciável em  $x_0$ , e  $\min f < f(x_0)$ .

*Demonstração.* A demonstração pode ser encontrada em Rockafellar [24], item (a) Teorema 23.4, p. 217, item (b) Corolário 7.6.1, p. 59 e item (c) Teorema 23.7, p. 222.

□

**Parte I**

**Problema Mono-objetivo**

# Capítulo 2

## Formalismo de Dubovitskii-Milyutin

Neste capítulo estudaremos a teoria desenvolvida por Dubovitskii e Milyutin para o estudo de condições necessárias de otimalidade. Consideraremos o seguinte problema de otimização:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeito a } x \in Q = \bigcap_{i=1}^{n+1} Q_i \end{aligned} \tag{P}$$

onde  $f : X \rightarrow \mathbb{R}$ , sendo  $X$  espaço de Banach,  $Q_i \subset X$ ,  $i = 1, \dots, n$  são conjuntos com interior não vazio e  $Q_{n+1} \subset X$  possivelmente não possui pontos interiores.

Na formulação usual, os conjuntos  $Q_i$ ,  $i = 1, \dots, n$  representam as restrições de desigualdade do problema e  $Q_{n+1}$  representa as restrições de igualdade do problema.

Neste capítulo, salvo menção em contrário,  $X$  é um espaço de Banach e  $f$  é uma função a valores reais, definida em  $X$ .

### 2.1 Condições necessárias de otimalidade

Primeiramente faremos algumas definições e propriedades que serão necessárias para a obtenção da Equação de Euler-Lagrange através do formalismo de Dubovitskii e Milyutin.

**Definição 2.1** Um vetor  $h$  é dito uma **direção de descida** do funcional  $f(x)$  no ponto  $x_0$  se existe uma vizinhança  $U$  de  $h$ ,  $\alpha < 0$  e  $\varepsilon_0 > 0$ , tais que para todo  $0 < \varepsilon < \varepsilon_0$  e qualquer  $\bar{h} \in U$

$$f(x_0 + \varepsilon \bar{h}) \leq f(x_0) + \varepsilon \alpha. \tag{2.1}$$

O conjunto de tais direções será denotado  $D(f; x_0)$ . O funcional  $f(x)$  é dito **regular de descida** se o cone das direções de descida no ponto  $x_0$  é convexo.

**Lema 2.2** As direções de descida geram um cone aberto  $D(f; x_0)$  com vértice em 0.

*Demonstração.* Se  $h$  é direção de descida então  $\lambda h$ ,  $\lambda > 0$  também é direção de descida, basta trocar  $U$  por  $\lambda U$ ,  $\varepsilon_0$  por  $\frac{\varepsilon_0}{\lambda}$  e  $\alpha$  por  $\alpha \lambda$ . Assim  $D(f; x_0)$  é um cone com vértice em 0. Além disso se  $h \in D(f; x_0)$  e  $\bar{h} \in U$ , então  $\bar{h} \in D(f; x_0)$ , basta tomar o mesmo  $U$ ,  $\alpha$ ,  $\varepsilon_0$  na definição de direção de descida.  $\square$

**Definição 2.3** Sendo  $Q$  um conjunto de restrições de desigualdade, dizemos que o vetor  $h$  é uma **direção factível** para  $Q$  no ponto  $x_0$  se existe uma vizinhança  $U$  de  $h$  tal que para todo  $0 < \varepsilon < \varepsilon_0$  e todo  $\bar{h} \in U$ ,  $x_0 + \varepsilon\bar{h} \in Q$ . Uma restrição de desigualdade  $Q$  é **regular** no ponto  $x_0$  se o conjunto de direções factíveis  $F(Q; x_0)$  para  $Q$  em  $x_0$  é convexo.

Veja Figura 2.1, a seguir.

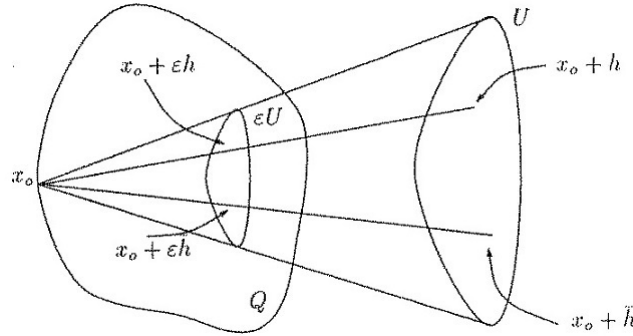


Figura 2.1: Direção factível.

**Lema 2.4** O conjunto de direções factíveis  $F(Q, x_0)$  é um cone aberto com vértice em 0.

*Demonstração.* Se  $h \in F(Q; x_0)$  então  $\lambda h \in F(Q; x_0)$ , para  $\lambda > 0$ , basta trocar  $U$  por  $\lambda U$ ,  $\varepsilon_0$  por  $\frac{\varepsilon_0}{\lambda}$ . Além disso se  $h \in F(Q; x_0)$  então todo  $\bar{h} \in U$  também estará em  $F(Q; x_0)$   $\square$

Para restrições de igualdade o conjunto de direções factíveis é vazio. Assim, torna-se necessária uma ligeira modificação em tal conceito para este tipo de restrição.

**Definição 2.5** Seja  $Q$  um conjunto de restrições de igualdade, dizemos que o vetor  $h$  é uma **direção tangente** a  $Q$  no ponto  $x_0$  se para qualquer  $0 < \varepsilon < \varepsilon_0$  existe um ponto  $x(\varepsilon) \in Q$  tal que se  $x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon)$  então o vetor  $r(\varepsilon) \in X$  é tal que, para qualquer vizinhança  $V$  de zero,  $\frac{1}{\varepsilon}r(\varepsilon) \in V$  para todo  $\varepsilon > 0$  suficientemente pequeno, ou, equivalentemente,  $\|r(\varepsilon)\| = o(\varepsilon)$ . O conjunto de tais direções será denotado  $T(Q; x_0)$ . Além disto, diremos que a restrição de igualdade  $Q$  é regular no ponto  $x_0$  se o conjunto  $T(Q; x_0)$  é convexo.

Veja Figura 2.2, a seguir.

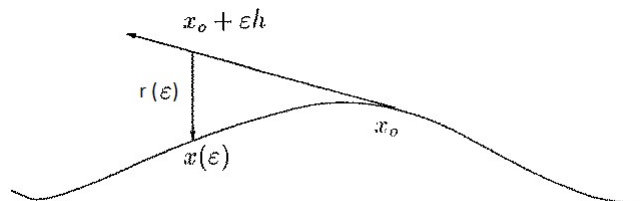


Figura 2.2: Direção tangente.

**Observação 1** Observamos que  $h$  é uma direção tangente ao conjunto  $Q$  no ponto  $x_0$  se e somente se para toda vizinhança  $V$  de  $h$  e qualquer número real  $\varepsilon_0$ , existem  $y \in V$  e  $\varepsilon \in (0, \varepsilon_0)$  tais que  $x_0 + \varepsilon y \in Q$ .

De fato, se  $h$  é uma direção tangente, então para qualquer  $0 < \varepsilon < \varepsilon_0$ , existe  $x(\varepsilon) \in Q$ ,  $x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon)$ , tal que  $\frac{r(\varepsilon)}{\varepsilon} \in U$ , para qualquer vizinhança  $U$  de zero e  $\varepsilon$  suficientemente pequeno. Logo

$$x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon) = x_0 + \varepsilon \left( h + \frac{r(\varepsilon)}{\varepsilon} \right) = x_0 + \varepsilon y,$$

onde  $h + \frac{r(\varepsilon)}{\varepsilon}$  pertence a alguma vizinhança  $V$  de  $h$ . Agora se para toda vizinhança  $V$  de  $h$  e qualquer número real  $\varepsilon_0$ , existe  $y \in V$  e  $\varepsilon \in (0, \varepsilon_0)$  tal que  $x = x_0 + \varepsilon y \in Q$ , tem-se que

$$x = x_0 + \varepsilon y = x_0 + \varepsilon(h + r),$$

onde  $r$  pertence a alguma vizinhança  $U$  de zero, então  $h$  é direção tangente, com  $r(\varepsilon) = \varepsilon r$ .

**Lema 2.6** O conjunto de direções tangentes  $T(Q; x_0)$  é um cone com vértice em 0.

*Demonstração.* Se  $h \in T(Q; x_0)$  então  $\lambda h \in T(Q; x_0)$  para  $\lambda > 0$  e basta trocar  $\varepsilon$  por  $\frac{\varepsilon}{\lambda}$  na definição de direção tangente.  $\square$

Note que toda direção factível é também direção tangente. De fato se  $h$  é direção factível então para todo  $0 < \varepsilon < \varepsilon_0$  e todo  $\bar{h} \in U_h$ ,  $x_0 + \varepsilon \bar{h} \in Q$ . Em particular  $x(\varepsilon) = x_0 + \varepsilon h \in Q$ , assim  $h$  é direção tangente com  $r(\varepsilon) = 0$ .

**Lema 2.7** Seja  $K$  um cone com vértice em  $x_0$  e  $f \in X'$  tal que  $f(x) \geq \alpha$ ,  $\forall x \in K$ , então  $f(x) \geq f(x_0)$ ,  $\forall x \in K$ .

*Demonstração.* Suponha por absurdo que existe  $y \in K$  tal que  $f(y) < f(x_0)$ . Pela definição de cone, para  $t \geq 0$ ,

$$x_0 + t(y - x_0) \in K.$$

Como  $f$  é um funcional linear

$$f(x_0 + t(y - x_0)) = f(x_0) + t(f(y) - f(x_0)),$$

e sendo  $f(y) < f(x_0)$  quando  $t \rightarrow +\infty$  temos  $f(x_0) + t(f(y) - f(x_0)) \rightarrow -\infty$ , o que contradiz o fato de  $f(x) \geq \alpha$ ,  $\forall x \in K$ .  $\square$

O Lema 2.8, enunciado a seguir, será de fundamental importância na obtenção da condição de necessária de otimalidade para o problema (P), conhecida como equação de Euler-Lagrange.

**Lema 2.8** Sejam  $K_i$ ,  $i = 1, \dots, n+1$  cones convexos com vértice em 0, onde  $K_i$ ,  $i = 1, \dots, n$  são abertos. Então  $\bigcap_{i=1}^{n+1} K_i = \emptyset$  se, e somente se, existem funcionais lineares  $g_i \in K_i^*$ ,  $i = 1, \dots, n+1$ , não todos nulos, tais que

$$\sum_{i=1}^{n+1} g_i = 0.$$

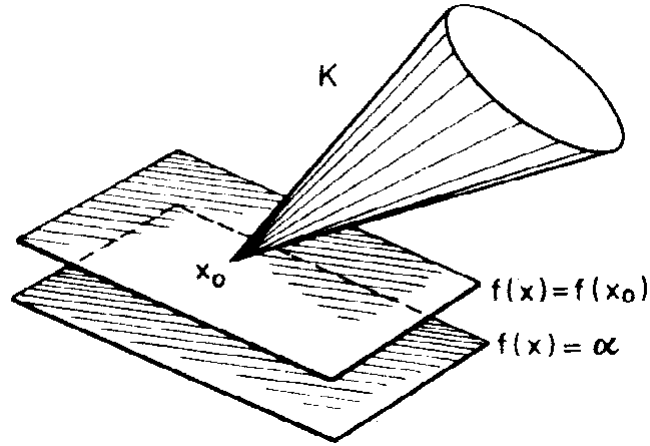


Figura 2.3: Interpretação geométrica do Lema 2.7.

*Demonstração.*

**Necessidade:** Seja  $K = \bigcap_{i=1}^n K_i \neq \emptyset$  e  $K_{n+1} \cap K = \emptyset$ . Como  $K$  é aberto e com vértice em 0, aplicando o Teorema 1.12 (Teorema de Separação) juntamente com o Lema 2.7 obtemos que existe  $g \in X'$ ,  $g \neq 0$ , tal que

$$\begin{aligned} g(x) &\geq 0 \quad \forall x \in K; \\ g(x) &\leq 0 \quad \forall x \in K_{n+1}. \end{aligned}$$

Pelo Lema 1.9,

$$K^* = \left( \bigcap_{i=1}^n K_i \right)^* = \sum_{i=1}^n K_i^*.$$

Portanto  $g \in K^*$  e

$$g = \sum_{i=1}^n g_i; \quad g_i \in K_i^*, \quad i = 1, \dots, n.$$

Defina  $g_{n+1} = -g$ , então  $g_{n+1} \in K_{n+1}^*$ ,  $g_{n+1} \neq 0$  e obtemos  $\sum_{i=1}^{n+1} g_i = 0$ .

Mas se  $\bigcap_{i=1}^n K_i = \emptyset$ , então existe  $1 \leq s \leq n$  tal que

$$K = \bigcap_{i=1}^s K_i \neq \emptyset$$

e

$$\bigcap_{i=1}^{s+1} K_i = \emptyset.$$

Ao aplicar o resultado anterior trocando  $n$  por  $s$ , obtemos  $g_i \in K_i^*$ ,  $g_i \neq 0$ ,  $i = 1, \dots, s+1$ , tais que

$$\sum_{i=1}^{s+1} g_i = 0.$$

Tomando  $g_i = 0$ , para  $i = s + 2, \dots, n + 1$  obtemos os  $n + 1$  funcionais desejados.

**Suficiência:** Suponha

$$\sum_{i=1}^{n+1} g_i = 0; \quad g_i \neq 0,$$

e que

$$\bigcap_{i=1}^{n+1} K_i \neq \emptyset.$$

Então existe  $x_0 \in K_i$ ,  $i = 1, \dots, n + 1$ . Por outro lado, para algum  $1 \leq j \leq n$  deve existir  $g_j \in K_j^*$ ,  $g_j \neq 0$  (pois caso contrário deveríamos ter  $g_{n+1} = -\sum_{i=1}^n g_i = 0$ ), e então  $g_j(x_0) > 0$ , pois  $K_j$  é aberto. Portanto

$$0 = \sum_{i=1}^{n+1} g_i(x_0) \geq g_j(x_0) > 0$$

o que é um absurdo. □

O seguinte Teorema fornece condições necessárias de otimalidade para o problema (P).

**Teorema 2.9** [Dubovitskii-Milyutin] *Suponha que o funcional  $f(x)$  assume um mínimo local em  $Q = \bigcap_{i=1}^{n+1} Q_i$  no ponto  $x_0 \in Q$ . Suponha ainda que  $f(x)$  é regular de descida em  $x_0$ , com direções de descida no cone  $K_0$ ; as restrições de desigualdade  $Q_i$ ,  $i = 1, \dots, n$  são regulares em  $x_0$ , com direções factíveis no cone  $K_i$ , a restrição de igualdade  $Q_{n+1}$  também é regular em  $x_0$ , com direções tangentes no cone  $K_{n+1}$ . Então existem funcionais lineares contínuos  $g_i \in K_i^*$ ,  $i = 0, 1, \dots, n + 1$  que satisfazem a equação de Euler-Lagrange*

$$\sum_{i=0}^{n+1} g_i = 0.$$

*Demonstração.* Vamos primeiramente provar que a condição necessária para o funcional ter um mínimo em  $x_0$  é

$$\bigcap_{i=0}^{n+1} K_i = \emptyset, \tag{2.2}$$

ou seja, nenhuma direção de descida do funcional  $f(x)$  pode ser factível para todas as restrições. Suponha que isto é falso, então existe  $h \in K_i$ ,  $i = 0, \dots, n + 1$ . Pela definição de  $K_i$  para  $i = 0, \dots, n$  existe uma vizinhança  $U$  de  $h$  tal que, quando  $0 < \varepsilon_0 < \varepsilon$  e  $\bar{h} \in U$ , qualquer vetor

$$x_0 + \varepsilon \bar{h} \in \bigcap_{i=0}^n K_i$$

e satisfaz a desigualdade

$$f(x_0 + \varepsilon \bar{h}) \leq f(x_0) + \varepsilon \alpha$$

para algum  $\alpha < 0$ .

Agora considere o vetor  $x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon) \in Q_{n+1}$  como na definição de direção tangente, e seja  $\varepsilon_1$  tal que  $\frac{1}{\varepsilon} r(\varepsilon) \in U - h$ , ou seja  $\bar{h}(\varepsilon) = h + \frac{1}{\varepsilon} r(\varepsilon) \in U$  para  $0 < \varepsilon \leq \varepsilon_1$ . Então quando  $0 < \varepsilon < \min\{\varepsilon_0, \varepsilon_1\}$  os vetores

$$x(\varepsilon) = x_0 + \varepsilon \bar{h}(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon)$$

estão, por um lado, em  $\bigcap_{i=0}^n Q_i$  e, por outro lado, em  $Q_{n+1}$ . Assim eles satisfazem todas as restrições. Mas eles também satisfazem a desigualdade

$$f(x(\varepsilon)) = f(x_0 + \varepsilon \bar{h}(\varepsilon)) \leq f(x_0) + \varepsilon \alpha < f(x_0),$$

o que contradiz o fato de  $x_0$  ser um ponto de mínimo local. Portanto

$$\bigcap_{i=0}^{n+1} K_i = \emptyset.$$

Além disto,  $K_i$  para  $i = 0, \dots, n$  são cones abertos convexos e  $K_{n+1}$  é um cone convexo. Obtivemos assim as hipóteses necessárias para a utilização do Lema 2.8 e isso implica o resultado desejado.  $\square$

**Observação 2** *Analisando a demonstração do teorema podemos afirmar que uma condição suficiente para que  $f_0 \neq 0$  é*

$$\bigcap_{i=1}^{n+1} K_i \neq \emptyset.$$

A conclusão é válida para qualquer  $f_i$ ,  $i = 1, \dots, n + 1$ .

**Observação 3** *Observe que a condição necessária de otimalidade dada pela equação (2.2) é válida, sem quaisquer hipóteses de regularidade nos cones de direções.*

## 2.2 Cálculo dos cones e de seus duais

Segue do Teorema de Dubovitskii-Milyutin que, se quisermos determinar as condições necessárias para um mínimo local em um problema específico, devemos determinar os cones de direções de descida, direções factíveis e de direções tangentes e construir os duais destes cones. Nesta Seção veremos como calcular tais cones.

### 2.2.1 Direções de descida

Derivadas direcionais fornecem uma ferramenta conveniente para construir o cone de direções de descida.

**Teorema 2.10** *Se  $h \in D(f; x_0)$  e  $f'(x_0, h)$  existe, então  $f'(x_0, h) < 0$ .*

*Demonstração.* Como  $h \in D(f; x_0)$ , temos da definição de direção de descida que existem  $\varepsilon > 0$  e  $\alpha < 0$  tais que

$$f(x_0 + \varepsilon h) \leq f(x_0) + \varepsilon \alpha.$$

Logo

$$f'(x_0, h) = \lim_{\varepsilon \rightarrow 0^+} \frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon} \leq \lim_{\varepsilon \rightarrow 0^+} \frac{f(x_0) + \varepsilon \alpha - f(x_0)}{\varepsilon} = \alpha < 0.$$

$\square$

A recíproca é válida somente para algumas classes de funcionais. Estudaremos alguns desses casos.

**Teorema 2.11** *Se  $f(x)$  é diferenciável em  $x_0$ , então  $D(f; x_0) = \{h \in X : f'(x_0)h < 0\}$  e  $f(x)$  é regular de descida em  $x_0$ .*

*Demonstração.* Como  $f(x)$  é diferenciável, possui derivada direcional em qualquer direção e do Teorema 2.10 obtemos que  $D(f; x_0) \subset \{h \in X : f'(x_0)h < 0\}$ .

Reciprocamente, se  $f(x)$  é diferenciável em  $x_0$ , então para todo  $\eta \in X$ ,

$$f(x_0 + \eta) = f(x_0) + f'(x_0)\eta + o(|\eta|).$$

Como por hipótese  $f'(x_0)h < 0$  e tomando  $\eta = \varepsilon \bar{h}$ , onde  $\bar{h}$  pertence a uma vizinhança de  $h$ , segue imediatamente que  $h$  é uma direção de descida, isto é  $h \in D(f; x_0)$ .

Como  $\{h \in X : f'(x_0)h < 0\}$  é convexo, pois é um semi-espaço ou o conjunto vazio, teremos que  $f(x)$  é regular de descida.  $\square$

**Teorema 2.12** *Seja  $f(x)$  um funcional convexo contínuo em um espaço de Banach  $X$ , então:*

- a)  $f(x)$  tem derivada direcional em qualquer ponto e em qualquer direção;
- b)  $f(x_0 + h) \geq f(x_0) + f'(x_0, h)$ ;
- c)  $D(f; x_0) = \{h \in X : f'(x_0, h) < 0\}$ ;
- d)  $f(x)$  é regular de descida em qualquer ponto  $x_0 \in X$ .

*Demonstração.*

- a) Seja  $x_0 \in X$  e considere a função de uma variável real

$$\psi(\varepsilon) = f(x_0 + \varepsilon h).$$

Como  $f(x)$  é convexo e contínuo,  $\psi(\varepsilon)$  também é convexa e contínua.

De fato, para  $\lambda \in [0, 1]$  e  $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$  temos:

$$\begin{aligned} \psi(\lambda\varepsilon_1 + (1 - \lambda)\varepsilon_2) &= f(x_0 + (\lambda\varepsilon_1 + (1 - \lambda)\varepsilon_2)h) \\ &= f(\lambda(x_0 + \varepsilon_1 h) + (1 - \lambda)(x_0 + \varepsilon_2 h)) \\ &\leq \lambda f(x_0 + \varepsilon_1 h) + (1 - \lambda)f(x_0 + \varepsilon_2 h) \\ &= \lambda\psi(\varepsilon_1) + (1 - \lambda)\psi(\varepsilon_2). \end{aligned}$$

Assim, para qualquer  $t \in [0, 1]$  temos

$$\psi(t\varepsilon) = \psi(t\varepsilon + (1 - t)0) \leq t\psi(\varepsilon) + (1 - t)\psi(0),$$

ou seja, para  $\varepsilon > 0$ ,

$$\psi(t\varepsilon) - \psi(0) \leq t\psi(\varepsilon) - t\psi(0).$$

Logo

$$\frac{\psi(t\varepsilon) - \psi(0)}{t\varepsilon} \leq \frac{\psi(\varepsilon) - \psi(0)}{\varepsilon},$$

portanto segue que a função  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  definida como

$$\phi(\varepsilon) = \frac{\psi(\varepsilon) - \psi(0)}{\varepsilon}$$

é monótona decrescente para  $\varepsilon \rightarrow 0^+$ . Além disso para  $\varepsilon > 0$  temos

$$\psi(0) = \psi\left(\frac{-\varepsilon}{\varepsilon+1} + \frac{\varepsilon}{\varepsilon+1}\right) \leq \frac{\varepsilon}{\varepsilon+1}\psi(-1) + \frac{1}{\varepsilon+1}\psi(\varepsilon),$$

e portanto

$$(\varepsilon+1)\psi(0) \leq \varepsilon\psi(-1) + \psi(\varepsilon)$$

ou, equivalentemente,

$$\frac{\psi(\varepsilon) - \psi(0)}{\varepsilon} \geq \psi(0) - \psi(-1).$$

Então  $\phi(\varepsilon)$  é limitada para  $\varepsilon > 0$ . Logo o limite  $\lim_{\varepsilon \rightarrow 0^+} \frac{\psi(\varepsilon) - \psi(0)}{\varepsilon}$  existe. Mas segue da definição de derivada direcional que

$$f'(x_0, h) = \lim_{\varepsilon \rightarrow 0^+} \frac{\psi(\varepsilon) - \psi(0)}{\varepsilon}.$$

Portanto a derivada direcional existe em qualquer ponto e em qualquer direção.

b) Do item a) temos que

$$\phi(\varepsilon) = \frac{\psi(\varepsilon) - \psi(0)}{\varepsilon}$$

é monótona decrescente, o que implica

$$\psi(1) - \psi(0) \geq \lim_{\varepsilon \rightarrow 0^+} \frac{\psi(\varepsilon) - \psi(0)}{\varepsilon},$$

ou seja,

$$f(x_0 + h) \geq f(x_0) + f'(x_0, h).$$

c) Pelo Teorema 2.10 temos que  $D(f; x_0) \subset \{h \in X : f'(x_0, h) < 0\}$ . Resta provar a inclusão contrária. Considere  $h \in X$  tal que  $f'(x_0, h) < 0$ . Segue da definição de derivada direcional que existe  $\varepsilon_0 > 0$  tal que

$$f(x_0 + \varepsilon_0 h) < f(x_0).$$

Defina

$$\delta = f(x_0) - f(x_0 + \varepsilon_0 h) > 0. \quad (2.3)$$

Como  $f(x)$  é contínuo, considere  $U$  uma vizinhança do vetor  $h$  tal que para todo  $\bar{h} \in U$  se verifica

$$|f(x_0 + \varepsilon_0 \bar{h}) - f(x_0 + \varepsilon_0 h)| \leq \frac{\delta}{2}.$$

Então,

$$f(x_0 + \varepsilon_0 \bar{h}) \leq f(x_0 + \varepsilon_0 h) + \frac{\delta}{2},$$

e por (2.3),

$$f(x_0 + \varepsilon_0 \bar{h}) \leq f(x_0) - \delta + \frac{\delta}{2} = f(x_0) - \frac{\delta}{2}. \quad (2.4)$$

Agora da condição de convexidade, para  $0 < \varepsilon < \varepsilon_0$  temos

$$\begin{aligned} f(x_0 + \varepsilon \bar{h}) &= f\left(x_0 - \frac{\varepsilon}{\varepsilon_0}x_0 + \frac{\varepsilon}{\varepsilon_0}x_0 + \frac{\varepsilon}{\varepsilon_0}\varepsilon_0 \bar{h}\right) \\ &= f\left(\left(1 - \frac{\varepsilon}{\varepsilon_0}\right)x_0 + \frac{\varepsilon}{\varepsilon_0}(x_0 + \varepsilon_0 \bar{h})\right) \\ &\leq \left(1 - \frac{\varepsilon}{\varepsilon_0}\right)f(x_0) + \frac{\varepsilon}{\varepsilon_0}f(x_0 + \varepsilon_0 \bar{h}), \end{aligned}$$

e por (2.4),

$$\begin{aligned} f(x_0 + \varepsilon\bar{h}) &\leq f(x_0) - \frac{\varepsilon}{\varepsilon_0}f(x_0) + \frac{\varepsilon}{\varepsilon_0}\left(f(x_0) - \frac{\delta}{2}\right) \\ &= f(x_0) - \frac{\varepsilon\delta}{2\varepsilon_0}. \end{aligned}$$

Ou seja, a condição de direção de descida é satisfeita para  $\alpha = -\frac{\delta}{2\varepsilon_0}$ , portanto  $h \in D(f; x_0)$ .

- d) Para mostrar que  $f(x)$  é regular de descida vejamos que  $D(f; x_0)$  é convexo. Sejam  $h_1, h_2 \in D(f; x_0)$  e  $\lambda \in [0, 1]$ . Como  $f(x)$  é convexo temos que

$$\begin{aligned} \frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon} &= \frac{f(x_0 + \varepsilon(\lambda h_1 + (1 - \lambda)h_2)) - f(x_0)}{\varepsilon} \\ &\leq \frac{\lambda f(x_0 + \varepsilon h_1) + (1 - \lambda)f(x_0 + \varepsilon h_2) - f(x_0)}{\varepsilon} \\ &= \frac{\lambda[f(x_0 + \varepsilon h_1) - f(x_0)] + (1 - \lambda)[f(x_0 + \varepsilon h_2) - f(x_0)]}{\varepsilon}. \end{aligned}$$

Portanto,

$$f'(x_0, h) \leq \lambda f'(x_0, h_1) + (1 - \lambda)f'(x_0, h_2).$$

Como  $h_1, h_2 \in D(f; x_0)$  então  $f'(x_0, h_1) < 0$  e  $f'(x_0, h_2) < 0$ . E assim,

$$f'(x_0, h) < 0,$$

ou seja,  $h = \lambda h_1 + (1 - \lambda)h_2 \in D(f; x_0)$  e portanto  $D(f; x_0)$  é convexo, isto é,  $f(x)$  é regular de descida. □

## 2.2.2 Direções factíveis

Vamos considerar o caso em que o conjunto  $Q$  é definido por algum funcional  $f(x)$ , isto é,

$$Q = \{x \in X : f(x) \leq f(x_0)\}.$$

**Observação 4** Se  $x_0 \in \text{int}Q$  então  $F(Q; x_0) = X$ . Portanto, nos interessa considerar somente o caso em que  $x_0$  está na fronteira de  $Q$ , ou seja  $x \in \text{Fr}(Q)$ .

**Lema 2.13** O cone de direções de descida  $D(f; x_0)$  está contido no cone de direções factíveis  $F(Q; x_0)$ .

*Demonstração.* De fato, se  $h \in D(f; x_0)$ , pela definição de direção de descida, existem  $\alpha < 0$  e  $\varepsilon_0 > 0$  tais que

$$f(x_0 + \varepsilon\bar{h}) \leq f(x_0) + \varepsilon\alpha < f(x_0)$$

para todo  $\bar{h}$  em alguma vizinhança  $U$  do vetor  $h$  e todo  $0 < \varepsilon < \varepsilon_0$ .

Portanto  $x_0 + \varepsilon\bar{h} \in Q$ , para quaisquer  $\bar{h} \in U$ , e  $\varepsilon \in (0, \varepsilon_0)$ . Mas essa é a definição de direção factível, logo  $h \in F(Q; x_0)$ . □

Em alguns casos importantes os cones de direções de descida e de direções factíveis coincidem. Os Teoremas 2.14 e 2.15 se referem a esta questão.

**Teorema 2.14** *Suponha que  $f(x)$  possui derivada direcional em  $x_0$  em qualquer direção,  $f'(x_0, h)$  é convexa em  $h$  e assumamos que existe  $\bar{h}$  tal que  $f'(x_0, \bar{h}) < 0$ . Então  $F(Q; x_0) \subset D(f; x_0) = \{h \in X : f'(x_0, h) < 0\}$ .*

*Demonstração.* Seja  $h \in F(Q; x_0)$ . Pela definição de direção factível

$$x_0 + \varepsilon h \in Q, \quad \forall 0 < \varepsilon < \varepsilon_0,$$

ou seja,

$$f(x_0 + \varepsilon h) \leq f(x_0).$$

Logo,

$$\frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon} \leq 0$$

e portanto,

$$f'(x_0, h) \leq 0.$$

Além disso, o cone  $F(Q; x_0)$  é aberto, (Lema 2.4). Logo existe uma vizinhança  $U$  do vetor  $h$  tal que  $U \subset F(Q; x_0)$ . Seja  $\gamma > 0$  tal que  $h_\gamma = h + \gamma(h - \bar{h}) \in U$ . Como  $h_\gamma \in F(Q; x_0)$  temos que  $f'(x_0, h_\gamma) \leq 0$  e segue da convexidade de  $f'(x_0, h)$  em  $h$  que

$$\begin{aligned} f'(x_0, h) &= f' \left( x_0, \frac{1}{1+\gamma} h_\gamma + \frac{\gamma}{1+\gamma} \bar{h} \right) \\ &\leq \frac{1}{1+\gamma} f'(x_0, h_\gamma) + \frac{\gamma}{1+\gamma} f'(x_0, \bar{h}) \\ &\leq \frac{\gamma}{1+\gamma} f'(x_0, \bar{h}) < 0. \end{aligned}$$

□

**Teorema 2.15** *Assuma que uma das seguintes condições é satisfeita:*

- a)  $f(x)$  é um funcional convexo, contínuo e existe  $\bar{x}$  tal que  $f(\bar{x}) < f(x_0)$ ;
- b)  $f(x)$  é diferenciável em  $x_0$  e  $f'(x_0) \neq 0$ .

Então  $F(Q; x_0) = D(f; x_0) = \{h \in X : f'(x_0, h) < 0\}$ .

*Demonstração.*

- a) Do Teorema 2.12 temos que  $f(x)$  possui derivada direcional em  $x_0$  em qualquer direção, e conseqüentemente,

$$D(f; x_0) = \{h \in X : f'(x_0, h) < 0\}$$

e tomando  $\bar{h} = \bar{x} - x_0$ ,

$$\begin{aligned} f'(x_0, \bar{h}) &\leq f(x_0 + \bar{h}) - f(x_0) \\ &= f(\bar{x}) - f(x_0) < 0. \end{aligned}$$

Assim sendo podemos utilizar o Teorema 2.14 e concluir que  $D(f; x_0) = F(Q; x_0)$ .

- b) Pelo Teorema 2.11,  $D(f; x_0) = \{h \in X : f'(x_0, h) < 0\}$  e  $f'(x_0, h)$  é convexa em  $h$ . Além disso como  $f(x)$  é diferenciável em  $x_0$ , então possui derivada direcional em  $x_0$  em qualquer direção, e por hipótese temos que  $f'(x_0) \neq 0$ , então existe  $\bar{h} \in X$  tal que  $f'(x_0, \bar{h}) < 0$ . Podemos assim aplicar o Teorema 2.14 e concluir a igualdade.

□

**Teorema 2.16** *Seja  $Q$  um conjunto convexo, então o cone de direções factíveis no ponto  $x_0$  é dado por  $F(Q; x_0) = \{h \in X : h = \lambda(x - x_0), x \in \text{int}Q, \lambda > 0\}$  e este conjunto é convexo.*

*Demonstração.* Vamos demonstrar a inclusão  $F(Q; x_0) \subset \{h \in X : h = \lambda(x - x_0), x \in \text{int}Q, \lambda > 0\}$ . Seja  $h \in F(Q; x_0)$ , então pela definição de direção factível existem uma vizinhança  $U$  do vetor  $h$  e  $\varepsilon_0 > 0$  tais que

$$x_0 + \varepsilon \bar{h} \in Q, \forall \bar{h} \in U, \forall \varepsilon \in (0, \varepsilon_0).$$

Considere  $\varepsilon \in (0, \varepsilon_0)$  fixo, então

$$x_0 + \varepsilon U = \{x \in X : x = x_0 + \varepsilon \bar{h}, \bar{h} \in U\}$$

é vizinhança de  $x_0 + \varepsilon h$ , logo  $x_0 + \varepsilon U \subset Q$ . Assim  $x = x_0 + \varepsilon h \in \text{int}Q$  e portanto  $h = \frac{1}{\varepsilon}(x - x_0) \in \{h \in X : h = \lambda(x - x_0), x \in \text{int}Q, \lambda > 0\}$ , com  $\lambda = \frac{1}{\varepsilon}$ .

Agora vamos demonstrar a inclusão contrária. Seja  $h \in \{h \in X : h = \lambda(x - x_0), x \in \text{int}Q, \lambda > 0\}$ . Então existe  $x \in \text{int}Q$  e  $\lambda > 0$  tal que  $h = \lambda(x - x_0)$ . Como  $x \in \text{int}Q$ , existe uma vizinhança  $V$  de  $x$  tal que  $V \subset Q$ . Tome  $U = \lambda V - \lambda x_0$  e  $\varepsilon_0 = \frac{1}{\lambda}$ , então  $U$  é uma vizinhança de  $h$  e para todo  $\bar{h} \in U$  e  $0 < \varepsilon < \frac{1}{\lambda}$  tem-se que  $x_0 + \varepsilon \bar{h} = x_0 + \varepsilon(\lambda v - \lambda x_0)$  onde  $v \in V$ . Assim  $x_0 + \varepsilon \bar{h} = \varepsilon \lambda v + (1 - \varepsilon \lambda)x_0$ . Por outro lado temos  $0 < \varepsilon \lambda < 1$ ,  $v, x_0 \in Q$  e da convexidade de  $Q$ , segue que  $x_0 + \varepsilon \bar{h} \in Q$ . Portanto  $h \in F(Q; x_0)$ .

Agora para mostrar que  $F(Q, x_0)$  é convexo, considere  $h_1, h_2 \in F(Q; x_0)$ , então  $h_1 = \lambda_1(x_1 - x_0)$  e  $h_2 = \lambda_2(x_2 - x_0)$  com  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  e  $x_1, x_2 \in \text{int}Q$ . Além disso sejam  $\gamma_1$  e  $\gamma_2$  escalares não negativos com  $\gamma_1 + \gamma_2 = 1$ .

Vamos mostrar que  $\gamma_1 h_1 + \gamma_2 h_2 \in F(Q; x_0)$ .

De fato,

$$\begin{aligned} \gamma_1 h_1 + \gamma_2 h_2 &= \gamma_1 \lambda_1 (x_1 - x_0) + \gamma_2 \lambda_2 (x_2 - x_0) \\ &= (\gamma_1 \lambda_1 + \gamma_2 \lambda_2) \left( \frac{\gamma_1 \lambda_1}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_1 + \frac{\gamma_2 \lambda_2}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_2 - x_0 \right) \\ &= \lambda (x - x_0), \end{aligned}$$

onde  $\lambda = \gamma_1 \lambda_1 + \gamma_2 \lambda_2 > 0$ ,  $x = \frac{\gamma_1 \lambda_1}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_1 + \frac{\gamma_2 \lambda_2}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_2$ .

Pela convexidade de  $Q$ ,  $x \in \text{int}Q$ , de modo que  $\gamma_1 h_1 + \gamma_2 h_2 = \lambda(x - x_0) \in F(Q; x_0)$ , segue que  $F(Q, x_0)$  é convexo. □

### 2.2.3 Direções tangentes

O Teorema de Lyusternik provado a seguir é uma forte ferramenta para o cálculo de direções tangentes. Observamos que utilizando o Teorema 1.19 não necessitamos que a função seja de classe  $C^1$ , mas apenas estritamente diferenciável.

**Teorema 2.17** [Lyusternik] *Sejam  $X, Y$  espaços de Banach,  $U$  uma vizinhança do ponto  $x_0$  em  $X$ ,  $P : U \rightarrow Y$  e  $P(x_0) = 0$ . Se  $P$  é estritamente diferenciável no ponto  $x_0$  e  $P'(x_0)$  é sobrejetora, então o conjunto de direções tangentes  $T(Q; x_0)$  para o conjunto  $Q = \{x \in X : P(x) = 0\}$  no ponto  $x_0$  é o subespaço  $T(Q; x_0) = \{h \in X : P'(x_0)h = 0\}$ .*

*Demonstração.* Primeiramente suponhamos que  $h \in T(Q; x_0)$ . Se  $r : [0, \varepsilon_0] \rightarrow Y$  é tal que  $x_0 + \varepsilon h + r(\varepsilon) \in Q$ , com  $\|r(\varepsilon)\| = o(\varepsilon)$ , para  $\varepsilon \rightarrow 0^+$ , então, como  $P$  é estritamente diferenciável segue que

$$0 = P(x_0 + \varepsilon h + r(\varepsilon)) = P(x_0) + \varepsilon P'(x_0)h + o(\varepsilon) = \varepsilon P'(x_0)h + o(\varepsilon).$$

Portanto  $P'(x_0)h = 0$ .

Agora para provarmos a outra inclusão aplicamos o Teorema 1.19 à função  $\psi(x, h) = P(x + h)$ , com  $h_0 = 0$  e  $z_0 = 0$ . Decorre da diferenciabilidade estrita de  $P$  que a função  $x \rightarrow \psi(x, 0)$  é contínua no ponto  $x_0$  e que dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que

$$\|P(x + h_1) - P(x + h_2) - P'(x_0)(h_1 - h_2)\| \leq \varepsilon \|h_1 - h_2\|$$

quando  $\|x - x_0\| < \delta$ ,  $\|h_1\| < \delta$  e  $\|h_2\| < \delta$ . A condição 3) do Teorema 1.19 também é satisfeita, uma vez que  $P'(x_0)$  é sobrejetora. Portanto, de acordo com esse teorema, existem  $k > 0$  e uma função  $\phi : U \rightarrow X$ , onde  $U$  é uma vizinhança do ponto  $x_0$  em  $X$  tais que

$$\psi(x, \phi(x)) = 0 \Leftrightarrow P(x + \phi(x)) = 0,$$

$$\|\phi(x)\| \leq k \|\psi(x, 0)\| = k \|P(x)\|.$$

Definimos  $r(\alpha) = \phi(x_0 + \alpha h)$ , então

$$P(x_0 + \alpha h + r(\alpha)) = P(x_0 + \alpha h + \phi(x_0 + \alpha h)) = 0.$$

Assim segue que  $x_0 + \alpha h + r(\alpha) \in Q$ . Temos também que

$$\begin{aligned} \|r(\alpha)\| = \|\phi(x_0 + \alpha h)\| &\leq k \|P(x_0 + \alpha h)\| \\ &= k \|P(x_0 + \alpha h) - P(x_0)\| \\ &= k \|P'(x_0)\alpha h + o(\|\alpha h\|)\| \\ &= k \|\alpha P'(x_0)h + o(\alpha)\|. \end{aligned}$$

Se  $h \in \text{Ker} P'(x_0)$ , então  $\|r(\alpha)\| = o(\alpha)$ . Assim pela definição de direção tangente, resulta que  $h \in T(Q; x_0)$ . Portanto  $T(Q; x_0) = \{h \in X : P'(x_0)h = 0\}$ .  $\square$

**Observação 5** *No caso que  $P'(x_0)$  não é sobrejetora podemos afirmar somente que  $T(Q; x_0) \subset \{h \in X : P'(x_0)h = 0\}$ .*

O Teorema 2.18 a seguir nos fornece uma caracterização do cone tangente a  $Q$  no ponto  $x_0$  para o caso em que o conjunto  $Q$  não é necessariamente definido por uma restrição de igualdade de tipo  $P(x) = 0$ .

**Teorema 2.18** *Seja  $Q$  um conjunto convexo, então o cone de direções tangentes no ponto  $x_0$  é dado por  $T(Q; x_0) = \text{cl}\{h \in X : h = \lambda(x - x_0), x \in Q, \lambda > 0\}$  e este conjunto é convexo.*

*Demonstração.* Considere  $h \in T(Q; x_0)$ , então, pela Observação 1, para toda vizinhança  $V$  de  $h$  existem  $h' \in V$  e  $\varepsilon > 0$  tais que  $x = x_0 + \varepsilon h' \in Q$ . Logo

$$h' = \frac{1}{\varepsilon}(x - x_0) \in \{h \in X : h = \lambda(x - x_0), x \in Q, \lambda > 0\},$$

com  $\lambda = \frac{1}{\varepsilon}$ . Portanto pela definição de fecho de um conjunto temos que

$$h \in \text{cl}\{h \in X : h = \lambda(x - x_0), x \in Q, \lambda > 0\}.$$

Reciprocamente, seja  $h \in \text{cl}\{h \in X : h = \lambda(x - x_0), x \in Q, \lambda > 0\}$ , então pela definição de fecho, para toda vizinhança  $V$  de  $h$  existe  $h' = \lambda(x - x_0) \in V$ , com  $\lambda > 0$  e  $x \in Q$ . Ou seja,  $x = x_0 + \frac{1}{\lambda}h' \in Q$ . Como  $Q$  é convexo para todo  $\theta \in (0, 1)$  temos

$$x_0 + \frac{\theta}{\lambda}h' = x_0 + \theta(x - x_0) \in Q.$$

Assim para toda vizinhança  $V$  de  $h$  e  $\varepsilon_0 > 0$  existe  $h' \in V$  e  $\varepsilon = \frac{\theta}{\lambda} \in (0, \varepsilon_0)$  tal que  $x_0 + \varepsilon h' \in Q$ , ou seja  $h \in T(Q; x_0)$ . Portanto  $T(Q; x_0) = \text{cl}\{h \in X : h = \lambda(x - x_0), x \in Q, \lambda > 0\}$ .

Para mostrar que  $T(Q, x_0)$  é convexo, considere  $h_1, h_2 \in T(Q; x_0)$ , então  $h_1 = \lambda_1(x_1 - x_0)$  e  $h_2 = \lambda_2(x_2 - x_0)$  com  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  e  $x_1, x_2 \in Q$ . Além disso sejam  $\gamma_1$  e  $\gamma_2$  escalares não negativos com  $\gamma_1 + \gamma_2 = 1$ .

Vamos mostrar que  $\gamma_1 h_1 + \gamma_2 h_2 \in T(Q; x_0)$ .

De fato,

$$\begin{aligned} \gamma_1 h_1 + \gamma_2 h_2 &= \gamma_1 \lambda_1 (x_1 - x_0) + \gamma_2 \lambda_2 (x_2 - x_0) \\ &= (\gamma_1 \lambda_1 + \gamma_2 \lambda_2) \left( \frac{\gamma_1 \lambda_1}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_1 + \frac{\gamma_2 \lambda_2}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_2 - x_0 \right). \\ &= \lambda (x - x_0), \end{aligned}$$

onde  $\lambda = \gamma_1 \lambda_1 + \gamma_2 \lambda_2 > 0$ ,  $x = \frac{\gamma_1 \lambda_1}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_1 + \frac{\gamma_2 \lambda_2}{\gamma_1 \lambda_1 + \gamma_2 \lambda_2} x_2$ .

Sendo  $Q$  convexo,  $x \in Q$ , temos que  $\gamma_1 h_1 + \gamma_2 h_2 = \lambda(x - x_0) \in T(Q, x_0)$  e portanto  $T(Q, x_0)$  é convexo.  $\square$

## 2.2.4 Cones Duais

Nas Seções anteriores vimos como calcular os cones de direção de descida, de direções factíveis e o subespaço tangente. Para a aplicação do Teorema de Dubovitskii-Milyutin devemos construir os cones duais. Este será o tema desta seção.

**Teorema 2.19** *Seja  $K$  um subespaço vetorial de  $X$ , então  $K^* = \{f \in X' : f(x) = 0, \forall x \in K\}$ .*

*Demonstração.* Seja  $f(x) > 0$  para algum  $x \in K$  e  $f \in K^*$ . Como  $K$  é subespaço vetorial  $-x \in K$  e  $f(-x) = -f(x) < 0$  o que contradiz o fato de  $f \in K^*$ . Analogamente se  $f(x) \leq 0$ .  $\square$

**Teorema 2.20** *Seja  $f \in X'$  um funcional linear fixado.*

- Se  $K_2 = \{x \in X : f(x) \geq 0\}$ , então  $K_2^* = \{\lambda f : \lambda \geq 0\}$ ;*
- Se  $K_3 = \{x \in X; f(x) > 0\}$ , então  $K_3^* = X'$  se  $f = 0$  e  $K_3^* = K_2^*$  se  $f \neq 0$ .*

*Demonstração.*

- a) Como  $K_1 \subset K_2$ , pelo Lema 1.3,  $K_2^* \subset K_1^*$ . Logo, se  $g \in K_2^*$ , temos  $g(x) = \lambda f(x) \geq 0$  para todo  $x \in K_2$ . Uma vez que  $f(x) \geq 0$  em  $K_2$ , concluímos que  $\lambda \geq 0$ .
- b) Se  $f = 0$ , temos  $K_3 = \emptyset$  e portanto  $K_3^* = X'$ . Sendo  $f \neq 0$  temos  $\text{cl}K_3 = K_2$  e assim  $K_3^* = K_2^*$ .

□

**Definição 2.21** *O conjunto de funcionais suporte para um conjunto  $Q$  no ponto  $x_0$  é definido por  $Q^* = \{f \in X' : f(x) \geq f(x_0), \forall x \in Q\}$ .*

Em alguns casos os duais dos cones de direções factíveis e de direções tangentes coincidem com o conjunto dos funcionais suporte para  $Q$  no ponto  $x_0$ .

**Teorema 2.22** *Seja  $Q$  um conjunto convexo. Então  $T^*(Q; x_0) = Q^*$ . Além disso se  $\text{int}Q \neq \emptyset$ , então  $F^*(Q; x_0) = Q^*$ .*

*Demonstração.* Seja  $f \in Q^*$ ,  $h \in T(Q; x_0)$ . Pela definição de direção tangente, dado  $0 < \varepsilon < \varepsilon_0$ , existe  $x(\varepsilon) \in Q$ ,  $x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon)$ , tal que  $\frac{r(\varepsilon)}{\varepsilon} \in U$ , para qualquer vizinhança  $U$  de zero e  $\varepsilon$  suficientemente pequeno. Como  $x(\varepsilon) \in Q$  e  $f \in Q^*$ , segue que  $f(x(\varepsilon)) \geq f(x_0)$  e portanto,

$$f(h) = \frac{f(x(\varepsilon)) - f(x_0) - f(r(\varepsilon))}{\varepsilon} \geq -f\left(\frac{r(\varepsilon)}{\varepsilon}\right).$$

Mas  $f\left(\frac{r(\varepsilon)}{\varepsilon}\right) \rightarrow 0$  quando  $\varepsilon \rightarrow 0^+$ , pois como  $f$  é contínua, para todo  $\delta > 0$  existe uma vizinhança  $U$  de zero tal que  $|f(x)| \leq \delta$  para  $x \in U$  e portanto  $\left|f\left(\frac{r(\varepsilon)}{\varepsilon}\right)\right| \leq \delta$ , para  $\varepsilon > 0$  suficientemente pequeno. Logo  $f(h) \geq 0$ , ou seja  $f \in T^*(Q; x_0)$ .

Reciprocamente, seja  $f \in T^*(Q; x_0)$ ,  $x \in Q$ . Então  $h = x - x_0$  é uma direção tangente (para  $0 < \varepsilon < 1$ , tome  $x_0 + \varepsilon h \in Q$ , donde obtém-se que  $x_0 + \varepsilon h = (1 - \varepsilon)x_0 + \varepsilon x \in Q$ , uma vez que  $Q$  é convexo). Como  $f \in T^*(Q; x_0)$ , segue que  $f(h) \geq 0$ , ou seja  $f(x) \geq f(x_0)$ , e isso significa que  $f \in Q^*$ . Deste modo temos provado que  $T^*(Q; x_0) = Q^*$ .

Agora se  $\text{int}Q \neq \emptyset$ , então pelo Teorema 2.16,

$$F(Q; x_0) = \{h \in X : h = \lambda(x - x_0), \lambda \geq 0, x \in \text{int}Q\}.$$

Portanto, se  $f \in F^*(Q; x_0)$ , então

$$f(\lambda(x - x_0)) \geq 0, \lambda > 0, x \in \text{int}Q.$$

Logo, pela linearidade de  $f$ ,

$$\lambda f(x) \geq \lambda f(x_0)$$

para todo  $x \in \text{int}Q$  e todo  $\lambda \geq 0$ , em particular para  $\lambda = 1$ . Mas como  $\text{int}Q \neq \emptyset$  e  $Q$  é convexo, pela Proposição 1.22 (d), temos  $\text{cl}(\text{int}Q) = \text{cl}Q$ . E assim, pela continuidade do funcional  $f$ , o fato que  $f(x) \geq f(x_0)$  para todo  $x \in \text{int}Q$  implica que  $f(x) \geq f(x_0)$  para todo  $x \in \text{cl}Q$ , ou seja,  $f \in Q^*$  e assim  $F^*(Q; x_0) \subset T^*(Q; x_0) = Q^*$ .

Por outro lado  $F(Q; x_0) \subset T(Q; x_0)$ , e portanto, pelo Lema 1.3,  $F^*(Q; x_0) \supset T^*(Q; x_0) = Q^*$ . Donde se conclui que  $F^*(Q; x_0) = Q^*$ .

□

## 2.3 Condições suficientes de otimalidade

Agora vamos determinar quais condições necessárias para mínimo são também suficientes. Naturalmente, exemplos elementares nos mostram que em geral isso não é verdadeiro. No entanto, veremos que as condições necessárias de mínimo são também suficientes em uma importante classe de problemas, os problemas convexos.

Um problema é convexo quando o funcional objetivo a ser minimizado e as restrições são convexas. Esses problemas possuem uma propriedade importante: o mínimo local e global coincidem.

**Teorema 2.23** *Sejam  $Q$  um subconjunto convexo e não vazio de  $X$  e  $f : Q \rightarrow \mathbb{R}$  um funcional convexo. Se  $x_0 \in Q$  é minimizador local de  $f$ , então  $x_0$  é minimizador global de  $f$ .*

*Demonstração.* Seja  $V$  uma vizinhança de  $x_0$  tal que  $f(x_0) \leq f(x)$  para todo  $x \in V \cap Q$ . Considere  $y \in Q$  arbitrário e tome  $\lambda \in (0, 1)$  tal que  $x_\lambda = (1 - \lambda)x_0 + \lambda y \in V$ . Como  $Q$  é convexo segue que  $x_\lambda \in Q$ , portanto  $x_\lambda \in V \cap Q$  e assim  $f(x_0) \leq f(x_\lambda)$ . Pela convexidade do funcional  $f$ ,

$$f(x_0) \leq f(x_\lambda) \leq (1 - \lambda)f(x_0) + \lambda f(y),$$

donde segue que  $f(x_0) \leq f(y)$ , e isso significa que  $x_0$  é ponto de mínimo global.  $\square$

**Teorema 2.24** *Sejam  $f(x)$  funcional convexo e contínuo,  $Q_i, i = 1, \dots, n + 1$  conjuntos convexos e assumamos que existe um ponto  $\bar{x} \in \text{int}Q_i$ , para  $i = 1, \dots, n$  e  $\bar{x} \in Q_{n+1}$ . Sejam  $x_0 \in Q = \bigcap_{i=1}^{n+1} Q_i$ ,  $K_0$  cone de direções de descida de  $f(x)$  em  $x_0$ ,  $K_i, i = 1, \dots, n$  cones de direções factíveis para  $Q_1, \dots, Q_n$  e  $K_{n+1}$  cone de direções tangentes para  $Q_{n+1}$  em  $x_0$ . Então  $x_0$  é ponto de mínimo para  $f(x)$  em  $Q$  se e somente se existem,  $g_i \in K_i^*$ , para  $i = 0, 1, \dots, n + 1$ , não todos nulos, tais que*

$$\sum_{i=0}^{n+1} g_i = 0.$$

*Demonstração.*

- **Necessidade** Como  $f(x)$  é funcional convexo e contínuo, pelo Teorema 2.12 é regular de descida. Além disso como  $Q_i$  são conjuntos convexos para  $i = 1, \dots, n$  então os cones factíveis  $K_i, i = 1, \dots, n$  são regulares, pelo Teorema 2.16. Além disto como  $Q_{n+1}$  é convexo, pelo Teorema 2.18 o cone  $K_{n+1}$  é regular. Agora o resultado segue do Teorema de Dubovitski-Milyutin, 2.9.
- **Suficiência** Suponha que existe  $x_1 \in Q$  tal que  $f(x_1) < f(x_0)$ . Defina  $x_\lambda = \lambda \bar{x} + (1 - \lambda)x_1$  para  $0 \leq \lambda \leq 1$ . Como os conjuntos  $Q_i$  são convexos e  $\bar{x} \in Q_i$ , para  $i = 1, \dots, n + 1$ , segue que  $x_\lambda \in Q_i$  para  $i = 1, \dots, n + 1$ . Além disto como  $\bar{x} \in \text{int}Q_i$  para  $i = 1, \dots, n$ , então  $x_\lambda \in \text{int}Q_i$  para  $i = 1, \dots, n$ .

Tome  $\lambda > 0$  suficientemente pequeno de modo que  $f(x_\lambda) < f(x_0)$  (isso pode ser feito pela continuidade de  $f(x)$  e pelo fato de  $f(x_1) < f(x_0)$ ).

Defina  $h = x_\lambda - x_0$ . Então para  $0 < \varepsilon < 1$ ,

$$f(x_0 + \varepsilon h) \leq \varepsilon f(x_\lambda) + (1 - \varepsilon)f(x_0),$$

e assim

$$\begin{aligned}
 f'(x_0, h) &= \lim_{\varepsilon \rightarrow 0^+} \frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon} \\
 &\leq \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon f(x_\lambda) + (1 - \varepsilon)f(x_0) - f(x_0)}{\varepsilon} \\
 &= \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon(f(x_\lambda) - f(x_0))}{\varepsilon} \\
 &= f(x_\lambda) - f(x_0) < 0.
 \end{aligned}$$

Portanto, pelo Teorema 2.12,  $h \in K_0$ .

Além disso,  $h \in K_i$ ,  $i = 1, \dots, n$ , pois  $x_0 + \varepsilon h = x_0 + \varepsilon(x_\lambda - x_0) \in \text{int}Q_i$  para  $0 < \varepsilon < 1$ , e  $h \in K_{n+1}$ , pois  $x_0 + \varepsilon h \in Q_{n+1}$ , para  $0 < \varepsilon < 1$ .

Logo  $h \in \bigcap_{i=0}^{n+1} K_i$ . Além disso os cones  $K_i$ ,  $i = 1, \dots, n$  são abertos e convexos. Agora pelo Lema 2.8, isso contradiz o fato de  $\sum_{i=0}^{n+1} g_i = 0$ .

Assim a hipótese de  $f(x_1) < f(x_0)$  deve ser falsa e  $x_0$  é um ponto de mínimo global.

□

# Capítulo 3

## Aplicações

Neste capítulo veremos como a teoria desenvolvida por Dubovitskii e Myliutin pode ser utilizada para obter condições necessárias de otimalidade para problemas de programação não linear e de controle ótimo com um único objetivo. Além disto vemos que sob certas hipóteses de convexidade e regularidade tais condições são suficientes.

### 3.1 Problemas de programação não linear

Consideraremos o seguinte problema de otimização:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeito a: } \begin{aligned} & g_i(x) \leq 0, \quad i = 1, \dots, r \\ & h_j(x) = 0, \quad j = 1, \dots, m \end{aligned} \end{aligned} \quad (\text{P})$$

onde  $f : X \rightarrow \mathbb{R}$ ,  $g_i : X \rightarrow \mathbb{R}$ ,  $i = 1, \dots, r$  e  $h_j : X \rightarrow \mathbb{R}$ ,  $j = 1, \dots, m$ , sendo  $X$  um espaço de Banach.

Vejamos agora as condições necessárias para que um ponto seja solução do problema anterior, conhecido na literatura por teorema de Fritz-John.

**Teorema 3.1** *Considere os funcionais  $f(x)$ ,  $g_i(x)$ ,  $i = 1, \dots, r$ , diferenciáveis e  $h_j(x)$ ,  $j = 1, \dots, m$  estritamente diferenciáveis em uma vizinhança de  $x_0$ . Se  $x_0$  é ótimo local do problema (P), então existem números reais  $\lambda_i$ ,  $i = 0, 1, \dots, r$  e  $\mu_j$ ,  $j = 1, \dots, m$ , não todos nulos, tais que:*

$$\lambda_0 f'(x_0) + \sum_{i=1}^r \lambda_i g'_i(x_0) + \sum_{j=1}^m \mu_j h'_j(x_0) = 0, \quad (3.1)$$

$$\lambda_i \geq 0, \quad i = 0, 1, \dots, r, \quad (3.2)$$

$$\lambda_i g_i(x_0) = 0, \quad i = 1, \dots, r. \quad (3.3)$$

*Demonstração.* Dividiremos a demonstração em 2 passos. No primeiro faremos uma análise dos casos triviais e no segundo passo, a análise dos casos não triviais baseada no formalismo de Dubovitski-Milyutin.

1° passo (Casos triviais)

Se  $f'(x_0) = 0$ , então as equações (3.1), (3.2) e (3.3) são satisfeitas, com  $\lambda_0 = 1$ ,  $\lambda_i = 0$ , para  $i = 1, \dots, r$  e  $\mu_j = 0$ , para  $j = 1, \dots, m$ .

Se  $g'_s(x_0) = 0$  com  $g_s(x) = 0$ , para algum  $s \in [1, r]$  então (3.1), (3.2) e (3.3) são satisfeitas com  $\lambda_s = 1$ ,  $\lambda_i = 0$  para  $i = 0, \dots, r$ ,  $i \neq s$  e  $\mu_j = 0$  para  $j = 1, \dots, m$ .

Se os vetores  $h'_j(x_0)$  com  $j = 1, \dots, m$  são linearmente dependentes, então existem  $\gamma_j$ ,  $j = 1, \dots, m$  não todos nulos tais que

$$\sum_{j=1}^m \gamma_j h'_j(x_0) = 0.$$

Assim podemos tomar  $\lambda_i = 0$ ,  $i = 0, 1, \dots, r$  e  $\mu_j = \gamma_j$ ,  $j = 1, \dots, m$  e serão satisfeitas as equações (3.1), (3.2) e (3.3).

## 2º passo (Casos não triviais)

Consideremos conforme no capítulo anterior,

$$Q_i = \{x \in X; g_i(x) \leq 0\}, \quad \forall i = 1, \dots, r$$

e

$$Q_{r+1} = \{x \in X; h_j(x) = 0, \quad \forall j = 1, \dots, m\}.$$

Então nosso problema será

$$\begin{aligned} &\text{Minimizar } f(x) \\ &\text{sujeito a } x \in Q = \bigcap_{i=1}^{r+1} Q_i. \end{aligned}$$

Analisemos agora os cones de direções de descida, factível, tangente e seus respectivos duais para este problema.

Como  $f$  é diferenciável em  $x_0$  e  $f'(x_0) \neq 0$ , pelo Teorema 2.11, o cone de direções de descida é dado por

$$D(f; x_0) = \{d \in X : f'(x_0)d < 0\}$$

e pelo Teorema 2.20,

$$D^*(f; x_0) = \{-\lambda_0 f'(x_0) : \lambda_0 \geq 0\}. \quad (3.4)$$

Agora, determinemos  $K_i$  o cone de direções factíveis para  $Q_i$ ,  $i = 1, \dots, r$ . Se  $g_i(x_0) < 0$ , segue diretamente da definição que  $K_i = E$ , logo  $K_i^* \equiv \{0\}$ . Agora se  $g_i(x_0) = 0$  e  $g'_i(x_0) \neq 0$ , como  $g_i$  são diferenciáveis em  $x_0$  para  $i = 1, \dots, r$ , pelo Teorema 2.15, temos

$$K_i = \{d \in X : g'_i(x_0)d < 0\}$$

e

$$K_i^* = \{-\lambda_i g'_i(x_0) : \lambda_i \geq 0\}. \quad (3.5)$$

Para determinar o cone de direções tangentes queremos utilizar o Teorema 2.17 (Teorema de Lyusternik). Para isto, definamos a aplicação  $H : X \rightarrow \mathbb{R}^m$ , tal que

$$H(x) = (h_1(x), \dots, h_m(x)).$$

Por hipótese o ponto  $x_0$  é ótimo local,  $h_j$  é estritamente diferenciável em  $x_0$  para todo  $j = 1, \dots, m$ , os vetores  $h'_j(x_0)$  são linearmente independentes, então

- a)  $H(x_0) = (h_1(x_0), \dots, h_m(x_0)) = (0, \dots, 0)$ ;
- b)  $H$  é estritamente diferenciável;
- c) A aplicação  $H'(x_0) : X \rightarrow \mathbb{R}^m$  é sobrejetora.

Assim as hipóteses do Teorema de Lyusternik são satisfeitas e

$$K_{r+1} = \text{Ker}H'(x_0) = \{d \in X : H'(x_0)d = 0\} = \{d \in X : h'_j(x_0)d = 0, \forall j = 1, \dots, m\}$$

Como  $\text{Ker}H'(x_0)$  é subespaço, pelo Teorema 2.19

$$K_{r+1}^* = \{y \in X' : (y, d) = 0, \forall d \in K_{r+1}\}.$$

Logo  $y \in (\text{Ker}H'(x_0))^\perp$ . Segue do Teorema 1.21 que

$$\text{Im}(H'(x_0))^* = (\text{Ker}H'(x_0))^\perp$$

onde  $(H'(x_0))^*$  é o operador adjunto de  $H(x_0)$ . Portanto,  $y \in \text{Im}(H'(x_0))^*$ , ou seja,  $y = H'(x_0)^*(-z)$ , para  $z \in \mathbb{R}^m$  fixo. Consequentemente, para todo  $d \in X$  e  $z = (\mu_1, \dots, \mu_m) \in \mathbb{R}^m$  fixo, temos que  $y \in X'$  pode ser representado como

$$(y, d) = (H'(x_0)^*(-z), d) = (-z, H'(x_0)d).$$

Desta forma

$$K_{r+1}^* = \left\{ -\sum_{j=1}^m \mu_j h'_j(x_0) : \mu_j \in \mathbb{R}, j = 1, \dots, m \right\}. \quad (3.6)$$

Através de (3.4), (3.5) e (3.6) temos a existência dos multiplicadores  $\lambda_i \geq 0, i = 1, \dots, r$  e  $\mu_j \in \mathbb{R}$ , para  $j = 1, \dots, m$ , não todos nulos tais que (3.2) e (3.3) são verdadeiras, e usando o Teorema 2.9 (Teorema de Dubovitskii-Milyutin) é válida a igualdade (3.1).  $\square$

Vamos agora analisar este resultado, indicando quando podemos assumir que  $\lambda_0 \neq 0$  (problemas desse tipo são conhecidos como regulares ou não degenerados).

**Teorema 3.2** *Sob as hipóteses do Teorema 3.1 assuma que os funcionais  $g_i, i = 1, \dots, r$  são convexos,  $h_j, j = 1, \dots, m$  são lineares e existe  $\bar{x}$  tal que  $g_i(\bar{x}) < 0, i = 1, \dots, r$  e  $h_j(\bar{x}) = 0, j = 1, \dots, m$ . Então, na equação 3.1,  $\lambda_0 \neq 0$ , ou seja a condição necessária para mínimo em  $x_0$  é:*

$$f'(x_0) + \sum_{i=1}^r \lambda_i g'_i(x_0) + \sum_{j=1}^m \mu_j h'_j(x_0) = 0, \quad (3.7)$$

$$\lambda_i \geq 0, i = 1, \dots, r, \quad (3.8)$$

$$\lambda_i g_i(x_0) = 0, i = 1, \dots, r. \quad (3.9)$$

*Demonstração.* Seja  $Q_i = \{x \in X : g_i(x) \leq 0\}, i = 1, \dots, r$ . Como os funcionais  $g_i, i = 1, \dots, r$  são convexos, então, pelo Teorema 2.16, os cones de direções factíveis  $K_i$  são convexos e pelo mesmo Teorema temos que  $\bar{x} - x_0$  é uma direção factível no ponto  $x_0$  para todo  $i$ . Além disso, para o conjunto convexo  $Q_{r+1} = \{x \in X : (h_1, \dots, h_m) = 0\}$ , pelo Teorema 2.18,  $\bar{x} - x_0$  é uma direção tangente no ponto  $x_0$  e o cone de direções tangentes é convexo. Portanto

$$\bigcap_{i=1}^{n+1} K_i \neq \emptyset$$

e pela Observação 2 esta é uma condição suficiente para que  $f_0 \neq 0$  e portanto  $\lambda_0 \neq 0$ .  $\square$

As equações (3.7), (3.8) e (3.9) são conhecidas como condições necessárias de Karush-Kuhn-Tucker (KKT). A condição utilizada anteriormente de as restrições de desigualdade serem convexas, as de igualdade lineares e existir  $\bar{x}$  tal que  $g_i(\bar{x}) < 0$ ,  $i = 1, \dots, r$  e  $h_j(\bar{x}) = 0$ ,  $j = 1, \dots, m$  é conhecida como condição de qualificação de Slater, a qual garante que as condições de KKT são necessárias para a otimalidade local. Existem muitas outras condições de qualificação, tais como Guignard, Mangasarian-Fromovitz, independência linear dos gradientes, mas não trataremos delas neste trabalho. Para maiores detalhes ver [7, 9, 14]

Além das restrições de desigualdade e igualdade podemos ter uma condição adicional que não é necessariamente determinada por um funcional. Neste caso nosso problema é expresso da seguinte forma:

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \quad g_i(x) \leq 0, \quad i = 1, \dots, r \\ \text{sujeito a: } & \quad h_j(x) = 0, \quad j = 1, \dots, m \\ & \quad x \in Q, \end{aligned} \tag{P'}$$

onde  $f : X \rightarrow \mathbb{R}$ ,  $g_i : X \rightarrow \mathbb{R}$ ,  $i = 1, \dots, r$  e  $h_j : X \rightarrow \mathbb{R}$ ,  $j = 1, \dots, m$ , sendo  $Q$  um subconjunto não vazio do espaço de Banach  $X$ .

Vejamos agora como ficam as condições necessárias para que um ponto seja ótimo local do problema anterior.

**Teorema 3.3** *Seja  $x_0$  ótimo local do problema (P') e assuma que os funcionais  $f(x)$ ,  $g_i(x)$ ,  $i = 1, \dots, r$ ,  $h_j(x)$ ,  $j = 1, \dots, m$  são diferenciáveis em uma vizinhança de  $x_0$ ,  $Q$  é convexo e  $\text{int}Q \neq \emptyset$ . Então existem números reais  $\lambda_i$ ,  $i = 0, 1, \dots, r$  e  $\mu_j$ ,  $j = 1, \dots, m$ , não todos nulos, tais que:*

$$\lambda_i \geq 0, \quad i = 0, 1, \dots, r, \tag{3.10}$$

$$\lambda_i g_i(x_0) = 0, \quad i = 1, \dots, r, \tag{3.11}$$

$$f = \lambda_0 f'(x_0) + \sum_{i=1}^r \lambda_i g'_i(x_0) + \sum_{j=1}^m \mu_j h'_j(x_0) \tag{3.12}$$

é um funcional suporte para  $Q$  no ponto  $x_0$ .

*Demonstração.* A demonstração segue a mesma linha da demonstração do Teorema 3.1. No entanto agora temos mais uma restrição dada por  $x \in Q$ . Como  $Q$  é um conjunto convexo e  $\text{int}Q \neq \emptyset$ , pelo Teorema 2.22 temos

$$F^*(Q; x_0) = Q^* = \{f \in X : f(x) \geq f(x_0), \forall x \in Q\} \tag{3.13}$$

ou seja o cone dual do cone de direções factíveis para  $Q$  no ponto  $x_0$  coincide com o conjunto dos funcionais lineares que são suporte para  $Q$  no ponto  $x_0$ .

Do Teorema 3.1 temos que as equações (3.10) e (3.11) são satisfeitas. E através de (3.4), (3.5) e (3.6) obtidas no teorema anterior e (3.13), obtemos como consequência do Teorema 2.9 (de Dubovitskii-Milyutin) que (3.12) é válida.  $\square$

Vamos reformular o Teorema 2.24 para o caso em que os conjuntos  $Q_i$ , para  $i = 1, \dots, n + 1$  são definidos por funções e assim determinar quais são as condições suficientes para o problema de programação não linear (P).

**Teorema 3.4** *Sejam  $f, g_i : X \rightarrow \mathbb{R}, i = 1, \dots, r$  funções convexas diferenciáveis e  $h_j : X \rightarrow \mathbb{R}, j = 1, \dots, m$  funções lineares. Assuma que existe um ponto  $\bar{x} \in X$  tal que  $g_i(\bar{x}) < 0, i = 1, \dots, r, h_j(\bar{x}) = 0, j = 1, \dots, m$ . Então  $f(x)$  possui um mínimo no ponto  $x_0$  se, e somente se, existem números  $\lambda_i, i = 1, \dots, r$  e  $\mu_j, j = 1, \dots, m$  tais que*

$$f'(x_0) + \sum_{i=1}^r \lambda_i g'_i(x_0) + \sum_{j=1}^m \mu_j h'_j(x_0) = 0, \quad (3.14)$$

$$\lambda_i \geq 0, \quad i = 1, \dots, r, \quad (3.15)$$

$$\lambda_i g_i(x_0) = 0, \quad i = 1, \dots, r. \quad (3.16)$$

*Demonstração.*

- **Necessidade** Segue do Teorema 3.2.
- **Suficiência** Seja  $x_1$  um ponto viável arbitrário, isto é,  $g_i(x_1) \leq 0, i = 1, \dots, r$  e  $h_j(x_1) = 0, j = 1, \dots, m$ . Como as funções  $f, g_i$ , para  $i = 1, \dots, r$  são convexas temos

$$f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0), \quad (3.17)$$

$$g_i(x_1) \geq g_i(x_0) + g'_i(x_0)(x_1 - x_0), \text{ para } i = 1, \dots, r. \quad (3.18)$$

E como as funções  $h_j$  são lineares

$$h_j(x_1) = h_j(x_0) + h'_j(x_0)(x_1 - x_0), \text{ para } j = 1, \dots, m. \quad (3.19)$$

Multiplicando as desigualdades (3.17) e (3.18) por números não negativos  $\lambda_i, i = 1, \dots, r$  e as igualdades (3.19) por  $\mu_j \in \mathbb{R}, j = 1, \dots, m$  e somando os resultados obtemos

$$\begin{aligned} f(x_1) + \sum_{i=1}^r \lambda_i g_i(x_1) &\geq f(x_0) + \sum_{i=1}^r \lambda_i g_i(x_0) + (f'(x_0) + \sum_{i=1}^r \lambda_i g'_i(x_0) \\ &\quad + \sum_{j=1}^m \mu_j h'_j(x_0))(x_1 - x_0). \end{aligned}$$

Como por hipótese  $\lambda_i \geq 0$  e  $g_i(x_1) \leq 0$  então  $\lambda_i g_i(x_1) \leq 0$  para  $i = 1, \dots, r$ . Além disso por (3.14), segue que

$$f(x_1) \geq f(x_0) + \sum_{i=1}^r \lambda_i g_i(x_1) \geq f(x_0),$$

ou seja,  $f(x_1) \geq f(x_0)$  e portanto  $x_0$  é um ponto de mínimo. □

Note que a prova de suficiência não fez uso da hipótese que  $\bar{x}$  existe para garantir que  $\lambda_0 \neq 0$ .

A suficiência das condições (3.14), (3.15) e (3.16), ou seja, das condições de KKT, é válida para uma classe mais ampla de funções, chamadas convexas generalizadas. Os interessados em aprofundar no tema podem consultar [10, 11, 13, 26].

## 3.2 Problemas de Controle Ótimo

Nesta Seção estudaremos o problema básico de controle ótimo em tempo contínuo e o resultado fundamental para obter a solução ótima do problema de controle ótimo: o princípio do máximo de Pontryagin. Mais informações sobre o problema de controle Ótimo podem ser encontradas em [1, 19, 27, 28].

O problema de controle ótimo trata de otimizar sistemas que evoluem ao longo do tempo. Dado um sistema que evolui ao longo de um período de tempo dado  $[t_0, t_1]$ , cuja situação inicial é dada pelo vetor  $n$ -dimensional  $x_0$ , e que evolui no tempo, procura-se controlar o sistema de maneira ótima ao longo de um período de tempo determinado, de acordo com o objetivo previamente fixado. Tal evolução depende do valor que se dá a certas variáveis, chamadas variáveis de controle, que permitem influenciar no sistema.

Consideramos o seguinte problema de controle ótimo:

$$\text{Minimizar } \int_0^T F(x(t), u(t), t) dt \quad (3.20)$$

$$\text{sujeito a: } \frac{d}{dt} x(t) = f(x(t), u(t), t) \quad (3.21)$$

$$x(0) = c \quad (3.22)$$

$$x(T) = d \quad (3.23)$$

$$u(t) \in M \text{ q.t.p } 0 \leq t \leq T, \quad (3.24)$$

onde  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^r$ ,  $f(x(t), u(t), t)$  é uma função vetorial,  $F(x(t), u(t), t)$  é uma função escalar,  $x \in C^{(n)}[0, T]$ ,  $u \in L_\infty^{(r)}[0, T]$  e  $M \subset \mathbb{R}^r$ . Sendo

- $C^{(n)}[0, T]$  o espaço das funções contínuas de  $[0, T]$  em  $\mathbb{R}^n$ , com a norma  $\|x\| = \max_{0 \leq t \leq T} \|x(t)\|$ .
- $L_\infty^{(r)}[0, T]$  o espaço das funções de  $[0, T]$  em  $\mathbb{R}^r$ , mensuráveis segundo Lebesgue, essencialmente limitadas, com norma  $\|u\| = \text{ess sup}_{0 \leq t \leq T} \|u(t)\|$ , onde  $\text{ess sup}$  denota o menor número  $\lambda$  tal que  $\|u(t)\| \leq \lambda$  para quase todo  $t \in [0, T]$ .
- $u(t)$  o vetor  $r$ -dimensional de variáveis de controle no instante  $t$  e  $x(t)$ , para cada  $t \in [0, T]$ , o vetor  $n$ -dimensional chamado de vetor de variáveis de estado, que nos indica a situação do sistema no instante  $t$ .

Existem outras versões do problema de controle diferindo na forma das condições de contorno, nas restrições impostas no controle, e em alguns problemas o tempo final não é fixo. O caso em que a solução é estipulada para ter domínio  $[t_0, T]$ ,  $t_0 \neq 0$ , pode ser reduzida ao caso anterior por uma simples substituição de variáveis.

Além disso existem diferentes formas que podemos encontrar o funcional objetivo. Anteriormente definimos o funcional objetivo como

$$\int_0^T F(x(t), u(t), t) dt. \quad (3.25)$$

Este funcional é dito que está na **forma de Lagrange**. Se o funcional é do tipo

$$\int_0^T F(x(t), u(t), t) dt + S[x(T)], \quad (3.26)$$

dizemos que está na **forma de Bolza**. Caso tenhamos somente

$$S[x(T)], \quad (3.27)$$

dizemos que o funcional está na **forma de Mayer**.

No entanto podemos formular cada um dos problemas em cada uma das outras formas, conforme o seguinte teorema.

**Teorema 3.5** *Os três problemas de controle, (3.25), (3.26), (3.27), são equivalentes, no sentido que qualquer um deles pode ser formulado em cada uma das outras formas.*

*Demonstração.*

- Lagrange  $\Rightarrow$  Mayer

Consideraremos o problema de controle ótimo, com funcional objetivo tipo Lagrange

$$\text{Minimizar } \int_{t_0}^{t_1} F(x(t), u(t), t) dt$$

$$\text{sujeito a: } x'(t) = f(x(t), u(t), t)$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1. \end{cases}$$

Vejamos que pode-se expressar na forma de Mayer. Para isso introduz-se uma nova variável de estado  $x_{n+1}(t)$  e considera-se

$$x'(t) = f(x(t), u(t), t), \text{ com } x(t_0) = x_0$$

$$x'_{n+1}(t) = F(x(t), u(t), t), \text{ com } x_{n+1}(t_0) = 0.$$

Assim o objetivo é

$$\text{Minimizar } \int_{t_0}^{t_1} x'_{n+1}(t) dt = x_{n+1}(t_1)$$

$$\text{sujeito a: } \begin{cases} x'(t) = f(x(t), u(t), t) \\ x'_{n+1}(t) = F(x(t), u(t), t) \end{cases}$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ x_{n+1}(t_0) = 0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1 \end{cases}$$

que é um problema na forma de Mayer.

- Bolza  $\Rightarrow$  Mayer

Partimos agora do problema

$$\text{Minimizar } \int_{t_0}^{t_1} F(x(t), u(t), t) dt + S[x(t_1)]$$

$$\text{sujeito a: } x'(t) = f(x(t), u(t), t)$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1. \end{cases}$$

Fazendo o mesmo processo que no caso anterior, obtém-se que o problema dado pode-se expressar como

$$\text{Minimizar } \int_{t_0}^{t_1} x'_{n+1}(t) dt + S[x(t_1)] = x_{n+1}(t_1) + S[x(t_1)]$$

$$\text{sujeito a: } \begin{cases} x'(t) = f(x(t), u(t), t) \\ x'_{n+1}(t) = F(x(t), u(t), t) \end{cases}$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ x_{n+1}(t_0) = 0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1. \end{cases}$$

- Bolza  $\Rightarrow$  Lagrange

Consideramos o seguinte problema:

$$\text{Minimizar } \int_{t_0}^{t_1} F(x(t), u(t), t) dt + S[x(t_1)]$$

$$\text{sujeito a: } x'(t) = f(x(t), u(t), t)$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1. \end{cases}$$

Verifica-se

$$S[x(t_1)] - S[x(t_0)] = \int_{t_0}^{t_1} \frac{d}{dt} S[x(t)] dt = \int_{t_0}^{t_1} \nabla S(x(t)) x'(t) dt.$$

Para cada função admissível do nosso problema

$$S[x(t_0)] = S[x_0].$$

Portanto,  $u^*(t)$  é solução ótima do problema considerado se, e somente se  $u^*(t)$  e solução ótima do problema

$$\text{Minimizar } \int_{t_0}^{t_1} [F(x(t), u(t), t) + \nabla S(x(t)) f(x(t), u(t), t)] dt$$

$$\text{sujeito a: } x'(t) = f(x(t), u(t), t)$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1. \end{cases}$$

- Mayer  $\Rightarrow$  Lagrange

Repetimos a mesma argumentação do caso anterior, porém, com  $F = 0$

$$\text{Minimizar } \int_{t_0}^{t_1} \nabla S(x(t)) f(x(t), u(t), t) dt$$

$$\text{sujeito a: } x'(t) = f(x(t), u(t), t)$$

$$\text{com: } \begin{cases} x(t_0) = x_0 \\ u(t) \in M \text{ q.t.p } t_0 \leq t \leq t_1. \end{cases}$$

- Lagrange  $\Rightarrow$  Bolza

O problema com funcional objetivo na forma de Lagrange já está com funcional objetivo na forma de Bolza, porém sendo  $S = 0$ .

- Mayer  $\Rightarrow$  Bolza

O problema na forma de Mayer já está na forma de Bolza, considerando-se que  $F = 0$ .

□

A seguir apresentamos o conceito de sistema não degenerado, e demonstraremos dois lemas que serão necessários na demonstração do teorema do Princípio de Máximo de Pontryagin.

Considere o problema de valor inicial

$$\begin{cases} \frac{d}{dt}\bar{x}(t) = A(t)\bar{x}(t) + B(t)\bar{u}(t) \\ \bar{x}(0) = 0 \end{cases} \quad (3.28)$$

onde  $A(t)$  e  $B(t)$  são matrizes  $n \times n$  e  $n \times r$  respectivamente, e  $\bar{u}(t)$  é uma função arbitrária em  $L_\infty^{(r)}[0, T]$ .

**Definição 3.6** (i) Dizemos que o sistema dado por (3.28) satisfaz a condição de não degeneração (A) se qualquer solução não nula  $\psi(t)$  da equação

$$\frac{d}{dt}\psi(t) = -A^\top(t)\psi(t) \quad (3.29)$$

satisfaz a condição

$$B^\top(t)\psi(t) \neq 0, \text{ em um conjunto de medida positiva.} \quad (3.30)$$

(ii) Seja  $D \subset \mathbb{R}^n$  o conjunto de todos os vetores  $\bar{x}(T)$ , onde  $x(t)$  satisfaz o problema de valor inicial (3.28). Dizemos que o sistema 3.28 satisfaz a condição de não degeneração (B) se  $D \equiv \mathbb{R}^n$ .

**Lema 3.7** A condição de não degeneração (A) implica a condição de não degeneração (B).

*Demonstração.* Suponha que a condição de não degeneração (B) não é válida, isto é,  $D \neq \mathbb{R}^n$ . Como uma consequência da linearidade da equação diferencial,  $D$  é um subespaço e disto segue que existe um vetor não nulo  $v \in \mathbb{R}^n$  ortogonal a  $D$ , ou seja  $a^\top \bar{x}(T) = 0$  para todo  $\bar{x}(t)$  que é solução da equação (3.28).

Considere o sistema

$$\begin{cases} \frac{d}{dt}\psi(t) = -A^\top(t)\psi(t) \\ \psi(T) = a. \end{cases}$$

Como  $a \neq 0$  segue que  $\psi(t) \neq 0$ . Então para qualquer solução  $\bar{x}(t)$  de (3.28)

$$\begin{aligned}
0 &= \int_0^T (\psi'(t) + A^\top(t) \psi(t))^\top \bar{x}(t) dt \\
&= \psi^\top(T) \bar{x}(T) - \psi^\top(0) \bar{x}(0) - \int_0^T \psi^\top(t) \bar{x}'(t) dt + \int_0^T \psi^\top(t) (A(t) \bar{x}(t)) dt \\
&= a^\top \bar{x}(T) - \int_0^T \psi^\top(t) (\bar{x}'(t) - A(t) \bar{x}(t)) dt \\
&= - \int_0^T \psi^\top(t) (B(t) \bar{u}(t)) dt \\
&= - \int_0^T (B^\top(t) \psi(t))^\top \bar{u}(t) dt.
\end{aligned}$$

Logo para todo  $\bar{u}(t) \in L_\infty^{(r)}$  temos

$$\int_0^T (B^\top(t) \psi(t))^\top \bar{u}(t) dt = 0.$$

Mas isso é possível somente se  $B^\top(t)\psi(t) = 0$  em quase toda parte de  $[0, T]$ , onde chegamos a uma contradição com a condição (3.30). Temos portanto o resultado desejado.  $\square$

**Lema 3.8** *Seja  $U = \{u \in L_\infty^{(r)}[0, T]; u(t) \in M \text{ q.t.p } 0 \leq t \leq T\}$ , onde  $M \subset \mathbb{R}^r$  e  $u_0 \in U$ . Então se o funcional linear definido por*

$$f(x) = \int_0^T g(t) \cdot u(t) dt$$

*é suporte para  $U$  no ponto  $u_0$ , com  $g(t) \in L_1^{(r)}[0, T]$  então  $g(t)^\top (u - u_0)(t) \geq 0$ , para todo  $u(t) \in M$  e quase todo  $t \in [0, T]$  (isto é para quase todo  $t \in [0, T]$  o vetor  $g(t) \in \mathbb{R}^r$  é suporte de  $M$  no ponto  $u_0$ ).*

*Demonstração.* Suponhamos que a afirmação é falsa, ou seja, existe um subconjunto  $R_1 \subset [0, T]$ ,  $\mu(R_1) \neq 0$  (onde  $\mu(R_1)$  denota a medida de Lebesgue de  $R_1$ ), tal que para todo  $t \in R_1$ , existe  $\bar{u}(t) \in M$  tal que  $g(t)^\top (\bar{u} - u_0)(t) < 0$ . De acordo com o Teorema de Lusin<sup>1</sup>, dado um intervalo  $J$ , se uma função  $\phi : J \rightarrow \mathbb{R}^r$  é mensurável, então, dado  $\varepsilon > 0$ , existe um conjunto fechado  $J_\varepsilon \subset J$  tal que  $\mu(J \setminus J_\varepsilon) \leq \varepsilon$  e  $\phi$  é contínua em  $J_\varepsilon$ . Logo existem subconjuntos  $R_2 \subset [0, T]$ ,  $\mu(R_2) < \frac{\varepsilon}{2}$ ,  $R_3 \subset [0, T]$ ,  $\mu(R_3) < \frac{\varepsilon}{2}$ , tais que  $g$  é contínua sobre  $[0, T] \setminus R_2$  e  $u_0$  é contínua em  $[0, T] \setminus R_3$ . Desde que  $\mu(R_2) + \mu(R_3) < \mu(R_1)$ , existe um ponto  $t_0 \in R_1$ ,  $t_0 \notin R_2 \cup R_3$ . Agora, como  $g$  e  $u_0$  são contínuas em  $t_0$  e  $g(t_0)(\bar{u}(t_0) - u_0(t_0)) = \gamma < 0$ , existe  $R_4 \subset [0, T]$ ,  $\mu(R_4) > 0$  tal que  $g(t)(\bar{u}(t_0) - u_0(t)) \leq \frac{\gamma}{2}$ , para todo  $t \in R_4$ . Consideremos a função

$$u_1(t) = \begin{cases} \bar{u}(t_0), & \text{se } t \in R_4 \\ u_0(t), & \text{se } t \in [0, T] \setminus R_4. \end{cases}$$

Deste modo,  $u_1 \in U$  e ao mesmo tempo,

$$\begin{aligned}
f(u_1) &= \int_0^T g(t) u_0(t) dt + \int_{R_4} g(t) (\bar{u}(t_0) - u_0(t)) dt \\
&\leq f(u_0) + \frac{\gamma}{2} \mu(R_4),
\end{aligned}$$

[1] Para melhores esclarecimentos sobre o Teorema de Lusin consulte [6, 15, 20]

ou seja,  $f(u_1) < f(u_0)$ . Isto contradiz o fato de  $f$  ser um funcional suporte para  $U$  no ponto  $u_0$ .  $\square$

Vejamos agora quais são as condições necessárias de otimalidade para problemas de controle ótimo.

**Teorema 3.9** [*Princípio de Máximo Local de Pontryagin*] *Sejam  $f(x(t), u(t), t)$ ,  $F(x(t), u(t), t)$  contínuas em  $x$ ,  $u$ , mensuráveis em  $t$ , continuamente diferenciáveis com respeito a  $x$  e  $u$ . Além disso sejam  $f_x(x(t), u(t), t)$ ,  $f_u(x(t), u(t), t)$ ,  $F_x(x(t), u(t), t)$ ,  $F_u(x(t), u(t), t)$  limitadas para todos  $x, u$  limitados. Seja  $M$  um conjunto convexo fechado em  $\mathbb{R}^r$ , com  $\text{int}M \neq \emptyset$ . Sejam  $x_0(t), u_0(t)$  soluções do problema (PC), então existem um número  $\lambda_0 \geq 0$  e uma função  $y(t)$ , satisfazendo a equação*

$$\frac{d}{dt}y(t) = -f_x^\top(x_0(t), u_0(t), t)y(t) + \lambda_0 F_x(x_0(t), u_0(t), t), \quad (3.31)$$

tais que  $\lambda_0$  e  $y(t)$  não podem ser ambos nulos, e além disso

$$[-f_u^\top(x_0(t), u_0(t), t)y(t) + \lambda_0 F_u(x_0(t), u_0(t), t)]^\top (u(t) - u_0(t)) \geq 0 \quad (3.32)$$

para quase todo  $0 \leq t \leq T$  e todo  $u \in M$ .

*Demonstração.* Consideremos o espaço  $X = C^{(n)}[0, T] \times L_\infty^{(r)}[0, T]$ . Denotemos por

$$Q_1 = \{(x, u) \in X : \text{satisfazem (3.24)}\}$$

e

$$Q_2 = \{(x, u) \in X : \text{satisfazem (3.21), (3.22), (3.23)}\}.$$

Assim nosso problema é

$$\text{Minimizar } J(x, u) = \int_0^T F(x(t), u(t), t) dt$$

$$\text{sujeito a: } (x, u) \in Q = Q_1 \cap Q_2.$$

Vamos agora proceder à análise do problema de acordo com nosso esquema geral, onde  $Q_1$  é o conjunto de restrições de desigualdade e  $Q_2$  o conjunto de restrições de igualdade. Primeiramente vamos determinar os cones correspondentes e então calcular os respectivos cones duais.

#### a. Análise do funcional objetivo:

Visto que  $J(x, u) = \int_0^T F(x(t), u(t), t) dt$ , onde  $F(x(t), u(t), t)$  é contínua em  $x$ ,  $u$ , mensurável em  $t$ , continuamente diferenciável com respeito a  $x$  e  $u$ ,  $F_x(x(t), u(t), t)$ ,  $F_u(x(t), u(t), t)$  limitadas para todos  $x$ ,  $u$  limitados, ou seja contínuas em  $x$ ,  $u$ ,  $t$ , obtemos que:

$$J'(x_0, u_0)(\bar{x}, \bar{u}) = \int_0^T [F_x(x_0(t), u_0(t), t), \bar{x}(t) + F_u(x_0(t), u_0(t), t), \bar{u}(t))] dt.$$

Pelo Teorema 2.11,

$$\begin{aligned} D &= \{(\bar{x}, \bar{u}) \in X : J'(x_0, u_0)(\bar{x}, \bar{u}) < 0\} \\ &= \{(\bar{x}, \bar{u}) \in X : \int_0^T [F_x(x_0(t), u_0(t), t), \bar{x}(t) + F_u(x_0(t), u_0(t), t), \bar{u}(t))] dt < 0\}. \end{aligned}$$

Pelo Teorema 2.20 item b), se  $D \neq \emptyset$ , então, para qualquer  $g_0 \in D^*$ ,

$$g_0(\bar{x}, \bar{u}) = -\lambda_0 \int_0^T [F_x(x_0(t), u_0(t), t), \bar{x}(t) + F_u(x_0(t), u_0(t), t), \bar{u}(t))] dt; \quad \lambda_0 \geq 0, \quad (3.33)$$

ou seja

$$D^* = \{-\lambda_0 J'(x_0, u_0) : \lambda_0 \geq 0\}.$$

**b. Análise da restrição de desigualdade  $Q_1$ :** O conjunto de funções

$$U = \{u \in L_\infty^{(r)} : u(t) \in M \text{ q.t.p } t \in [0, T]\}$$

é convexo fechado no espaço  $L_\infty^{(r)}$  e  $\text{int}U \neq \emptyset$ , devido ao fato de  $M$  ser convexo fechado e  $\text{int}M \neq \emptyset$ . Portanto o conjunto  $Q_1 = C^{(n)} \times U$  é também um conjunto convexo e fechado em  $X$ , e

$$\text{int}Q_1 = C^{(n)} \times \text{int}U \neq \emptyset.$$

Seja  $F(Q_1; (x_0, u_0))$  o cone de direções factíveis para  $Q_1$  no ponto  $(x_0, u_0)$ . Então se  $g_1 \in F^*(Q_1; (x_0, u_0))$  segue que  $g_1 = (0, g'_1)$  onde  $g'_1 \in L_\infty^{(r)}$  é um funcional suporte para  $U$  no ponto  $u_0$  conforme Teorema 2.22.

**c. Análise da restrição de igualdade  $Q_2$ :** Queremos calcular o subespaço tangente  $T(Q_2; (x_0, u_0))$  ao conjunto

$$Q_2 = \left\{ (x, u) \in X : x(t) = c + \int_0^t f(x(\tau), u(\tau), \tau) d\tau, \quad \forall 0 \leq t \leq T, \text{ e } x(T) = d \right\}$$

no ponto  $(x_0, u_0)$ . Para isto, definimos o operador  $P : X \rightarrow X_1$ , onde  $X_1 = C^{(n)}[0, T] \times \mathbb{R}^n$  e

$$P(x, u) = \left( x(t) - c - \int_0^t f(x(\tau), u(\tau), \tau) d\tau, x(T) - d \right).$$

Então

$$Q_2 = \{(x, u) \in X : P(x, u) = (0, 0)\}.$$

De acordo com as hipóteses relativas às derivadas  $f_x(x(t), u(t), t)$  e  $f_u(x(t), u(t), t)$  teremos pela expansão de Taylor de primeira ordem que

$$P(x_0 + \bar{x}, u_0 + \bar{u}) - P(x_0, u_0) =$$

$$\left( \bar{x}(t) - \int_0^t [f_x(x_0(\tau), u_0(\tau), \tau)\bar{x}(\tau) + f_u(x_0(\tau), u_0(\tau), \tau)\bar{u}(\tau)] d\tau, \bar{x}(T) \right) + \delta$$

onde para o resto  $\delta$  podemos encontrar a estimativa:  $\delta = o\left(\sqrt{\|\bar{x}\|_C^2 + \|\bar{u}\|_{L_\infty}^2}\right)$ . Também notamos que

$$\left( \bar{x}(t) - \int_0^t [f_x(x_0(\tau), u_0(\tau), \tau)\bar{x}(\tau) + f_u(x_0(\tau), u_0(\tau), \tau)\bar{u}(\tau)] d\tau, \bar{x}(T) \right)$$

é um operador linear de  $X$  em  $X_1$ . Logo  $P(x, u)$  é diferenciável em  $(x_0, u_0)$  e

$$P'(x_0, u_0)(\bar{x}, \bar{u}) = \left( \bar{x}(t) - \int_0^t [f_x(x_0(\tau), u_0(\tau), \tau)\bar{x}(\tau) + f_u(x_0(\tau), u_0(\tau), \tau)\bar{u}(\tau)] d\tau, \bar{x}(T) \right)$$

Além disso,  $P'(x_0, u_0)$  é contínua em uma vizinhança de  $(x_0, u_0)$ , portanto é estritamente diferenciável (Corolário 1.18).

Queremos determinar quando esse operador linear é sobrejetor, pois assim podemos utilizar o Teorema de Lyusternik (Teorema 2.17) para exibir o cone tangente  $T(Q_2; (x_0, u_0))$ .

Afirmamos que se a condição de não degeneração (A) é satisfeita para o sistema

$$\begin{cases} \frac{d}{dt}\bar{x}(t) = A(t)\bar{x}(t) + B(t)\bar{u}(t) \\ \bar{x}(0) = 0 \end{cases} \quad (3.34)$$

com  $A(t) = f_x(x_0(t), u_0(t), t)$  e  $B(t) = f_u(x_0(t), u_0(t), t)$ , então  $P'(x_0, u_0)$  é sobrejetor, ou seja,  $P'(x_0, u_0)X = X_1$ .

Assim todas as condições do Teorema de Lyusternik são satisfeitas para o operador  $P$  e então o cone tangente  $T(Q_2; (x_0, u_0))$  consiste de todos os pares  $(\bar{x}, \bar{u}) \in X$  tais que

$$\bar{x}(t) = \int_0^t [f_x(x_0, u_0, \tau)\bar{x}(\tau) + f_u(x_0, u_0, \tau)\bar{u}(\tau)]d\tau, \quad \bar{x}(T) = 0.$$

Ou em termos da equação diferencial linear

$$\frac{d}{dt}\bar{x}(t) = f_x(x_0(t), u_0(t), t)\bar{x}(t) + f_u(x_0(t), u_0(t), t)\bar{u}(t), \quad \bar{x}(0) = 0, \quad (3.35)$$

$$\bar{x}(T) = 0. \quad (3.36)$$

Seja  $V \subset X$  o conjunto de todos os pares  $(\bar{x}, \bar{u})$  satisfazendo (3.35) e  $W \subset X$  o conjunto de todos os pares satisfazendo (3.36). Então  $V$  e  $W$  são subespaços e  $T(Q_2; (x_0, u_0)) = V \cap W$ .

Como  $V$  é um subespaço vetorial, segue do Teorema 2.10 que para qualquer  $g_2 \in V$  temos  $g_2(\bar{x}, \bar{u}) = 0$  para todo  $(\bar{x}, \bar{u}) \in V$ .

Como  $W$  é subespaço, se  $g_3 \in W^*$ ,  $g_3(\bar{x}, \bar{u}) = 0$ , para todo  $(\bar{x}, \bar{u}) \in W$ . Assim,  $g_3 \in W^*$  se, e somente se, existe  $c = (a, b) \in \mathbb{R}^{n+r}$  tal que

$$g_3(\bar{x}(T), \bar{u}(T)) = c^\top(\bar{x}(T), \bar{u}(T)) = a^\top\bar{x}(T) + b^\top\bar{u}(T).$$

É claro que  $b = 0$ , por isso  $g_3(\bar{x}(T), \bar{u}(T)) = a^\top\bar{x}(T)$ . Logo se  $g_3 \in W^*$ , teremos que  $g_3(\bar{x}, \bar{u}) = a^\top\bar{x}(T)$ , com  $a \in \mathbb{R}^n$ . Como  $W$  é definido por  $n$  funcionais lineares  $\bar{x}_i$ ,  $W^*$  é  $n$ -dimensional. Portanto  $V^* + W^*$  é fechado, logo fracamente\* fechado. De fato, se  $V$  e  $W$  são subespaços fechados de um espaço de Banach e  $W$  é finito dimensional, então  $V+W$  é fechado. Como  $V$  e  $W$  são subespaços, logo são convexos, e pelo Teorema 1.5 são fechados se, e somente se, são fracamente fechados. Além disso todo conjunto fechado na topologia fraca é fechado na topologia fraca\*. Logo podemos aplicar o Teorema 1.6 e segue que  $T^*(Q_2; (x_0, u_0)) = V^* + W^*$ . Portanto se  $g \in T^*(Q_2; (x_0, u_0))$ , então  $g(\bar{x}, \bar{u}) = g_2(\bar{x}, \bar{u}) + g_3(\bar{x}, \bar{u}) = a^\top\bar{x}(T)$ ;  $a \in \mathbb{R}^n$ .

**d. Equação de Euler** Aplicando o Teorema de Dubovitskii-Milyutin (Teorema 2.9) ao nosso problema segue que existem  $g_0, g_1, g_2, g_3 \in X'$ , não todos nulos, tais que

$$g_0(\bar{x}, \bar{u}) + g_1(\bar{x}, \bar{u}) + g_2(\bar{x}, \bar{u}) + g_3(\bar{x}, \bar{u}) = 0, \quad \forall(\bar{x}, \bar{u}) \in X. \quad (3.37)$$

onde  $g_0(\bar{x}, \bar{u})$  é dado por (3.33),  $g_1(\bar{x}, \bar{u}) = g'_1(\bar{u})$  é suporte para  $Q'_1$  no ponto  $u_0$ ,  $g_2(\bar{x}, \bar{u}) = 0$  para todo  $(\bar{x}, \bar{u})$  satisfazendo (3.35) e  $g_3(\bar{x}, \bar{u}) = a^\top\bar{x}(T)$ ;  $a \in \mathbb{R}^n$ .

**e. Análise da Equação de Euler** A equação (3.37) deve valer para todos  $(\bar{x}, \bar{u}) \in X$ . Seja  $\bar{u}$  arbitrário e determine  $\bar{x} = \bar{x}(\bar{u})$  solução da equação (3.35). Com essa escolha de  $(\bar{x}, \bar{u})$  temos  $g_2(\bar{x}, \bar{u}) = 0$  e então a equação (3.37) torna-se

$$g'_1(\bar{u}) = \lambda_0 \int_0^T [F_x(x_0(t), u_0(t), t)\bar{x}(t) + F_u(x_0(t), u_0(t), t)\bar{u}(t)]dt - a^\top \bar{x}(T). \quad (3.38)$$

Suponhamos que  $\psi(t)$  é solução do sistema (3.31) com a condição de fronteira  $\psi(T) = a$ . Então

$$\begin{aligned} & \lambda_0 \int_0^T F_x(x_0(t), u_0(t), t)\bar{x}(t)dt - a^\top \bar{x}(T) = \\ & \int_0^T (\psi'(t) + f_x^\top(x_0(t), u_0(t), t)\psi(t))^\top \bar{x}(t)dt - a^\top \bar{x}(T) = \\ & \int_0^t (\psi'(t))^\top \bar{x}(t)dt + \int_0^T (f_x^\top(x_0(t), u_0(t), t)\psi(t))^\top \bar{x}(t)dt - a^\top \bar{x}(T). \end{aligned}$$

Integrando por partes obtemos que

$$\begin{aligned} & \int_0^t (\psi'(t))^\top \bar{x}(t)dt + \int_0^T (f_x^\top(x_0(t), u_0(t), t)\psi(t))^\top \bar{x}(t)dt - a^\top \bar{x}(T) = \\ & \psi(t)\bar{x}(t)|_0^\top - \int_0^t (\psi(t))^\top \bar{x}'(t)dt + \int_0^T (\psi(t))^\top (f_x(x_0(t), u_0(t), t))\bar{x}(t)dt - a^\top \bar{x}(T). \end{aligned}$$

e usando o fato que  $(\bar{x}, \bar{u})$  satisfaz as equações (3.35) e (3.36), obtemos

$$\begin{aligned} & - \int_0^T \psi(t)^\top [\bar{x}'(t) - f_x(x_0(t), u_0(t), t)\bar{x}(t)]dt = \\ & - \int_0^T \psi(t)^\top (f_u(x_0(t), u_0(t), t)\bar{u}(t))dt = \\ & - \int_0^T (f_u^\top(x_0(t), u_0(t), t)\psi(t))^\top \bar{u}(t)dt. \end{aligned}$$

Assim sendo,

$$\lambda_0 \int_0^T F_x(x_0(t), u_0(t), t)\bar{x}(t)dt - a^\top \bar{x}(T) = - \int_0^T (f_u^\top(x_0(t), u_0(t), t)\psi(t))^\top \bar{u}(t)dt.$$

Portanto a condição (3.38) torna-se

$$g'_1(\bar{u}) = \int_0^T (-f_u^\top(x_0(t), u_0(t), t)\psi(t) + \lambda_0 F_u(x_0(t), u_0(t), t))^\top \bar{u}(t)dt$$

onde  $\bar{u}$  é arbitrário e  $g'_1$  é funcional suporte para  $U$  no ponto  $u_0$ . Logo, pelo Lema 3.8,

$$[-f_u^\top(x_0(t), u_0(t), t)\psi(t) + \lambda_0 F_u(x_0(t), u_0(t), t)]^\top (u(t) - u_0(t)) \geq 0,$$

para quase todo  $0 \leq t \leq T$  e para todo  $u(t) \in M$ .

Sob essas hipóteses, o caso  $\lambda_0 = 0$ ,  $\psi(t) = 0$  não pode ocorrer, pois então deveríamos ter  $g_0(\bar{x}, \bar{u}) = 0$ ,  $a = \psi(T) = 0$ . Mas assim, por (3.38),  $g'_1(\bar{u}) = 0$ , logo

$$g_0 = g_1 = g_3 = 0.$$

A equação (3.37) implica também que  $g_2 = 0$  o que contradiz a hipótese de que os funcionais  $g_i$ ,  $i = 0, 1, 2, 3$  não são todos nulos.

**f. Análise dos casos excepcionais - casos triviais** No decorrer da prova fizemos duas hipóteses adicionais:

- $D \neq \emptyset$ ;
- o sistema (3.34) é não degenerado.

Vamos mostrar agora que essas afirmações são supérfluas.

De fato, se  $D = \emptyset$  então

$$\int_0^T [F_x(x(t), u(t), t) \bar{x}(t) + F_u(x(t), u(t), t) \bar{u}(t)] dt = 0$$

para todo  $(\bar{x}, \bar{u})$ . Tomando  $\lambda_0 = 1$  e  $\psi$  solução do sistema (3.31), com  $\psi(T) = 0$ , então

$$\begin{aligned} & \int_0^T F_x(x(t), u(t), t) \bar{x}(t) dt = \\ & \int_0^T (\psi'(t) + f_x^\top(x_0(t), u_0(t), t) \psi(t))^\top \bar{x}(t) dt = \\ & \psi(t) \bar{x}(t) \Big|_0^T - \int_0^T \psi^\top(t) \bar{x}'(t) dt + \int_0^T \psi^\top(t) f_x(x_0(t), u_0(t), t) \bar{x}(t) dt. \end{aligned}$$

Novamente pelo fato de  $(\bar{x}, \bar{u})$  satisfazer (3.35) temos

$$\begin{aligned} & - \int_0^T \psi^\top(t) [\bar{x}'(t) - f_x(x_0(t), u_0(t), t) \bar{x}(t)] dt = \\ & - \int_0^T \psi^\top(t) (f_u(x_0(t), u_0(t), t) \bar{u}(t)) dt = \\ & - \int_0^T (f_u(x_0(t), u_0(t), t) \psi(t))^\top \bar{u}(t) dt. \end{aligned}$$

Portanto

$$\int_0^T F_x(x(t), u(t), t) \bar{x}(t) dt = - \int_0^T (f_u(x_0(t), u_0(t), t) \psi(t))^\top \bar{u}(t) dt.$$

Assim

$$\int_0^T (-f_u^\top(x_0(t), u_0(t), t) \psi(t) + F_u(x_0(t), u_0(t), t))^\top \bar{u}(t) dt = 0$$

para todo  $\bar{u}$ . Logo

$$-f_u^\top(x_0(t), u_0(t), t) \psi(t) + F_u(x_0(t), u_0(t), t) = 0$$

para quase todo  $t \in [0, T]$ , ou seja a desigualdade (3.32) é satisfeita.

Agora se o sistema (3.34) é degenerado, então existe  $\psi(t) \neq 0$  que é solução da equação (3.31), com  $\lambda_0 = 0$ , tal que

$$-f_u^\top(x_0(t), u_0(t), t) \psi(t) \equiv 0.$$

Isto mostra a veracidade da equação (3.32), com  $\lambda_0 = 0$ .

□

**Observação 6** Se soubermos que o sistema linearizado (3.34) é não degenerado, então podemos assumir  $\lambda_0 = 1$ .

Vejamos agora sob quais hipóteses as condições necessárias são também suficientes para o Problema de Controle Ótimo.

$$\text{Minimizar } \int_0^T F(x(t), u(t), t) dt \quad (3.39)$$

$$\text{sujeito a: } \frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t) \quad (3.40)$$

$$x(0) = c \quad (3.41)$$

$$x(T) = d \quad (3.42)$$

$$u(t) \in M \text{ q.t.p } 0 \leq t \leq T. \quad (3.43)$$

**Teorema 3.10** Assuma que  $F(x(t), u(t), t)$  é convexa e continuamente diferenciável com respeito a  $x$ ,  $u$  e contínua em  $t$ ,  $A(t)$ ,  $B(t)$  são matrizes que dependem continuamente de  $t$  e  $M$  é conjunto convexo fechado em  $\mathbb{R}^r$ , com interior não vazio. Assuma além disso que a Condição de Não Degeneração (A) é satisfeita, ou seja  $B^\top(t)\psi(t) \neq 0$  para toda solução não nula  $\psi(t)$  do sistema

$$\frac{d\psi(t)}{dt} = -A^\top(t)\psi(t).$$

Então, o Princípio do Máximo Local 3.9 é condição suficiente para mínimo.

*Demonstração.* Como assumimos a condição de não degeneração temos que  $\lambda_0 > 0$  e  $\frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t)$ . Assim as equações (3.31) e (3.32) tornam-se

$$\frac{d}{dt}y(t) = -A(t)^\top y(t) + F_x(x_0(t), u_0(t), t), \quad (3.44)$$

$$[-B(t)^\top y(t) + F_u(x_0(t), u_0(t), t)]^\top (u(t) - u_0(t)) \geq 0$$

ou seja,

$$F_u(x_0(t), u_0(t), t)^\top (u_0(t) - u(t)) \leq y(t)^\top B(t)(u_0(t) - u(t)) \quad (3.45)$$

para quase todo  $0 \leq t \leq T$  e todo  $u \in M$ .

Como a função  $F(x(t), u(t), t)$  é convexa, temos que

$$\begin{aligned} F(x_0(t), u_0(t), t) - F(x(t), u(t), t) &\leq F_x(x_0(t), u_0(t), t)(x_0(t) - x(t)) \\ &\quad + F_u(x_0(t), u_0(t), t)(u_0(t) - u(t)). \end{aligned}$$

Seja  $(x, u) \in Q$  arbitrário. Vejamos que

$$J_0 = \int_0^T F(x_0(t), u_0(t), t) dt \leq \int_0^T F(x(t), u(t), t) dt = J.$$

Usando a desigualdade anterior obtemos

$$\begin{aligned} J_0 - J &= \int_0^T [F(x_0(t), u_0(t), t) - F(x(t), u(t), t)] dt \\ &\leq \int_0^T [F_x(x_0(t), u_0(t), t)(x_0(t) - x(t)) + F_u(x_0(t), u_0(t), t)(u_0(t) - u(t))] dt. \end{aligned}$$

Aplicando as condições (3.44) e (3.45) do Princípio do Máximo obtemos

$$J_0 - J \leq \int_0^T \{[y'(t) + A(t)^\top y(t)]^\top (x_0(t) - x(t)) + y(t)^\top B(t)(u_0(t) - u(t))\} dt.$$

Resolvendo por partes a seguinte integral

$$\begin{aligned} \int_0^T y'(t)^\top (x_0(t) - x(t)) dt &= y(t)^\top [x_0(t) - x(t)]_0^\top - \int_0^T y(t)^\top (x'_0(t) - x'(t)) dt \\ &= - \int_0^T y(t)^\top (x'_0(t) - x'(t)) dt \\ &= - \int_0^T y(t)^\top (A(t)x_0(t) + B(t)u_0(t) - A(t)x(t) - B(t)u(t)), \end{aligned}$$

e substituindo, obtém-se finalmente que

$$\begin{aligned} J_0 - J &\leq \int_0^T y(t)^\top [-A(t)x_0(t) - B(t)u_0(t) + A(t)x(t) + B(t)u(t) + A(t)(x_0(t) - x(t)) \\ &\quad + B(t)(u_0(t) - u(t))] dt = 0. \end{aligned}$$

Portanto  $J_0 - J \leq 0$  e o teorema está demonstrado. □

**Parte II**

**Problema Multiobjetivo**

# Capítulo 4

## Formulação do problema, conceitos de solução e escalarização

### 4.1 Formulação do problema e conceitos de solução

Consideraremos o seguinte problema de otimização multiobjetivo:

$$\begin{aligned} &\text{Minimizar } f(x) = (f_1(x), \dots, f_s(x)) \\ &\text{sujeito a } x \in Q = \bigcap_{i=1}^{p+1} Q_i \end{aligned} \tag{MOP}$$

onde  $f : X \rightarrow \mathbb{R}^s$ ,  $Q_i \subset X$ ,  $i = 1, \dots, p$  são conjuntos com interior não vazio, e  $Q_{p+1} \subset X$  não possui pontos interiores, sendo  $X$  espaço de Banach.

Vários conceitos de solução podem ser associados ao problema (MOP). Talvez, os mais conhecidos sejam os seguintes:

- **Solução Pareto fraca:**  $x_0 \in \mathbb{R}^n$  é solução Pareto fraca (ou solução fracamente eficiente) de (MOP) se  $x_0 \in Q$  e não existe  $x \in Q$ ,  $x \neq x_0$  tal que  $f_j(x) < f_j(x_0)$  para  $j = 1, \dots, s$ .
- **Solução de Pareto:**  $x_0 \in \mathbb{R}^n$  é solução de Pareto (ou solução eficiente) de (MOP) se  $x_0 \in Q$  e não existe  $x \in Q$ ,  $x \neq x_0$  tal que  $f_j(x) \leq f_j(x_0)$  para  $j = 1, \dots, s$ , com a desigualdade estrita para pelo menos um  $j$ .
- **Solução propriamente eficiente:**  $x_0 \in \mathbb{R}^n$  é solução propriamente eficiente de (MOP) se ela é eficiente e se existe  $M > 0$  tal que, para cada  $1 \leq k \leq s$

$$\frac{f_k(x_0) - f_k(x)}{f_j(x) - f_j(x_0)} \leq M$$

para algum  $j$  tal que  $f_j(x_0) < f_j(x)$  quando  $x \in Q$  e  $f_k(x) < f_k(x_0)$ .

Para problemas de maximização multiobjetivo, os correspondentes conceitos de solução são obtidos revertendo-se os sentidos das desigualdades.

É imediato das definições: eficiência própria  $\implies$  Pareto  $\implies$  Pareto fraca.

As recíprocas são falsas, como mostram os seguintes exemplos:

**Exemplo 4.1** *Moulin e Soulié [21]*

*Seja o problema:*

$$\begin{aligned} &\text{Maximizar } f(x) = (f_1(x), f_2(x)) \\ &\text{sujeito a: } x \in Q \end{aligned} \tag{4.1}$$

onde  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  com  $f_1(x) = x_1$  e  $f_2(x) = x_2$ , e  $Q = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\} \cup [-1, 0] \times [0, 1] \subseteq \mathbb{R}^2$ .

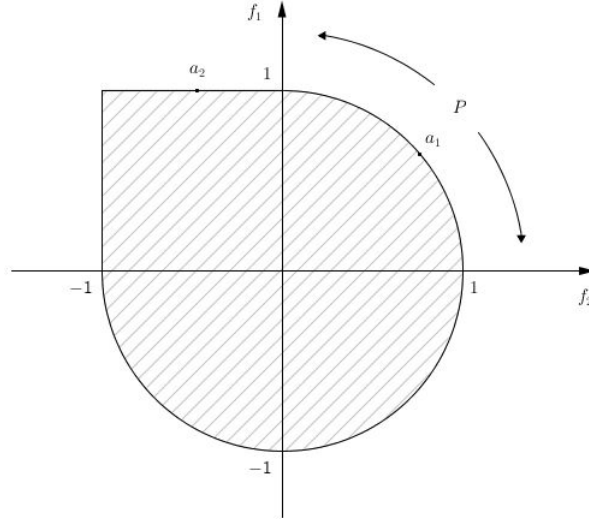


Figura 4.1: Conjunto factível Q.

- O ponto  $a_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$  não pode melhorar estritamente  $f_1$ , sem diminuir estritamente  $f_2$ . O ponto  $a_1$  é solução de Pareto pertencente ao conjunto B, onde B é o arco compreendido entre  $(1, 0)$  e  $(0, 1)$ .
- O ponto  $a_2 = (-\frac{1}{2}, 1)$  pode melhorar  $f_1$  sem diminuir estritamente  $f_2$ , (consideremos por exemplo, o ponto  $(0, 1)$ ), mas não pode melhorar os dois critérios ao mesmo tempo. O ponto  $a_2 \in B_f$  entretanto  $a_2$  pode não pertencer a B. Com:

$$B_f = \{B \cup (\lambda, 1) : -1 \leq \lambda \leq 0\}$$

**Exemplo 4.2** Considere o problema:

$$\begin{aligned} & \text{Minimizar } f(x) = (f_1(x), f_2(x)) \\ & \text{sujeito a: } x \geq 0, x \in \mathbb{R}, \end{aligned}$$

onde  $f_1(x) = -x^2$  e  $f_2(x) = x^3$ . Temos que o ponto  $x = 0$  é Pareto, porém não é propriamente eficiente. De fato, a razão

$$\frac{f_1(0) - f_1(x)}{f_2(x) - f_2(0)} = \frac{1}{x}$$

é ilimitada quando  $x$  tende a zero pela direita.

## 4.2 Escalarização

O termo escalarização apresenta para o problema multiobjetivo uma família de problemas de otimização escalar, de tal maneira que as soluções do problema multiobjetivo podem ser obtidas como soluções de um problema clássico de programação não linear. Consideraremos o Método da Ponderação e o Método da  $\varepsilon$ -restrição.

1. **Método da Ponderação:** Seja  $W = \{w \in \mathbb{R}^s : w_j \geq 0, 1 \leq j \leq s \text{ e } \sum_{j=1}^s w_j = 1\}$ . Para cada  $w \in W$ , consideraremos o seguinte problema (ponderado)

$$\begin{aligned} & \text{Minimizar } \sum_{j=1}^s w_j f_j(x) \\ & \text{sujeito a: } x \in Q. \end{aligned} \tag{P}_w$$

2. **Método da  $\varepsilon$ -restrição:** Para cada vetor  $\varepsilon \in \mathbb{R}^s$  e  $1 \leq k \leq s$ , definimos o problema

$$\begin{aligned} & \text{Minimizar } f_k(x) \\ & \text{sujeito a:} \\ & f_j(x) \leq \varepsilon_j \quad 1 \leq j \leq s, \quad j \neq k \\ & x \in Q. \end{aligned} \tag{P}_k(\varepsilon)$$

Agora descreveremos algumas relações existentes entre as soluções do problema multiobjetivo (MOP) e as soluções dos problemas escalarizados  $(P_w)$  e  $(P_k(\varepsilon))$ .

**Teorema 4.3** *Se existe  $w \in W$  tal que  $x_0 \in Q$  é solução de  $(P_w)$ , então  $x_0$  é solução Pareto fraca de (MOP).*

*Demonstração.* Suponha que  $x_0$  não é solução Pareto fraca de (MOP), então existe  $x \in Q$  tal que  $f_j(x) < f_j(x_0)$ , para cada  $1 \leq j \leq s$ . Se multiplicamos estas desigualdades por  $w_j \geq 0$  e somamos sobre  $1 \leq j \leq s$  obtemos  $\sum_{j=1}^s w_j f_j(x) < \sum_{j=1}^s w_j f_j(x_0)$  o que é absurdo, pois  $x_0 \in Q$  é solução de  $(P_w)$ .  $\square$

**Teorema 4.4** *Suponha que todas as funções de (MOP) sejam convexas. Se  $x_0$  é uma solução Pareto fraca de (MOP) então existe  $w \in W$  tal que  $x_0$  é solução de  $(P_w)$ .*

*Demonstração.* Suponha que  $x_0 \in Q$  é solução Pareto fraca de (MOP). Seja  $x \in Q$  e defina:

$$\begin{aligned} U &= \{u \in \mathbb{R}^s : u > 0\} \\ V &= \{v \in \mathbb{R}^s : v \leq f(x_0) - f(x)\}. \end{aligned}$$

Então,  $U \cap V = \emptyset$ . Além disso,  $U$  e  $V$  são convexos e não-vazios. Logo, pelo Teorema de Separação de Hahn-Banach (Teorema 1.12) existe um vetor  $w \in \mathbb{R}^s$ ,  $w \neq 0$  tal que

$$w^T v \leq 0 < w^T u, \forall u \in U, \forall v \in V.$$

Portanto,  $w \geq 0$ . É claro que  $v = f(x_0) - f(x) \in V$ . Assim,

$$w^T (f(x_0) - f(x)) \leq 0$$

e, portanto

$$w^T f(x_0) \leq w^T f(x), \forall x \in Q$$

ou, equivalentemente,  $x_0$  é solução de  $(P_w)$ .  $\square$

**Teorema 4.5** *Se existem  $w \in W$  tais que  $w > 0$  e  $x_0 \in Q$  é solução de  $(P_w)$ , então  $x_0$  é solução propriamente eficiente de (MOP).*

*Demonstração.* Vamos supor que  $w > 0$  e que  $\sum_{j=1}^s w_j f_j(x_0) \leq \sum_{j=1}^s w_j f_j(x), \forall x \in Q$  e, por absurdo, suporemos que  $x_0$  não é solução de Pareto de (MOP). Neste caso, existe  $x_1 \in Q$  tal que  $f_j(x_1) \leq f_j(x_0)$  para  $j = 1, \dots, s$ , com a desigualdade estrita para pelo menos um  $j$ , e como  $w > 0$ , teríamos  $\sum_{j=1}^s w_j f_j(x_0) > \sum_{j=1}^s w_j f_j(x_1)$ , o que é absurdo.

Agora, definimos  $M = (s - 1) \max_{j,k} \frac{w_j}{w_k} > 0$ . Vamos demonstrar que  $x_0$  é solução propriamente eficiente de (MOP), com respeito a este  $M$ .

Suponha por absurdo que  $x_0$  não é solução propriamente eficiente de (MOP) para este  $M$ . Então, para algum  $1 \leq k \leq s$ , e algum  $x \in Q$ , temos

$$f_k(x_0) - f_k(x) > M(f_j(x) - f_j(x_0))$$

para cada  $j$  com  $f_j(x_0) < f_j(x)$  e  $f_k(x) < f_k(x_0)$ . Logo, para cada  $j \neq k$  temos

$$f_k(x_0) - f_k(x) > \frac{s-1}{w_k} w_j (f_j(x) - f_j(x_0))$$

e, multiplicando a desigualdade acima por  $\frac{w_k}{s-1}$  e somando todas as desigualdades obtidas sobre  $j \neq k$ , obtemos

$$w_k (f_k(x_0) - f_k(x)) > \sum_{j \neq k} w_j (f_j(x) - f_j(x_0))$$

donde

$$\sum_{j=1}^s w_j f_j(x_0) > \sum_{j=1}^s w_j f_j(x)$$

o que contraria o fato de  $x_0$  ser solução ótima de  $(P_w)$ . □

**Teorema 4.6** **a)** *Se  $x_0 \in Q$  é solução de Pareto de (MOP), então  $x_0$  é solução ótima de  $(P_k(\varepsilon))$ , para todo  $1 \leq k \leq s$  e algum  $\varepsilon \in \mathbb{R}^s$ .*

**b)** *Seja  $x_0 \in Q$ . Defina  $\varepsilon_j = f_j(x_0)$ , para cada  $1 \leq j \leq s$ . Se  $x_0$  é solução única de  $(P_k(\varepsilon))$ , então  $x_0 \in Q$  é solução de Pareto de (MOP).*

*Demonstração.*

**a)** Se  $x_0$  é solução de Pareto de (MOP), então não existe  $x \in Q$  tal que  $f(x) \leq f(x_0)$ . Definindo-se  $\varepsilon_j = f_j(x_0)$ , para cada  $1 \leq j \leq s$ , temos que  $f_j(x_0) \leq \varepsilon_j$ , para  $1 \leq j \leq s$ . Fixado  $1 \leq k \leq s$  temos que  $x_0$  é factível para  $(P_k(\varepsilon))$ . Além disto,  $x_0$  é solução ótima de  $(P_k(\varepsilon))$ . (Caso contrário, existiria um  $x \in Q$  tal que  $f_j(x) \leq \varepsilon_j = f_j(x_0), j \neq k$  e  $f_k(x) < f_k(x_0)$ , o que contraria a eficiência de  $x_0$ .)

**b)** Suponha que  $x_0$  é solução única de  $(P_k(\varepsilon))$  e, por contradição, que  $x_0$  não é solução de Pareto de (MOP). Então, existe  $x \in Q$  tal que  $f_j(x) \leq f_j(x_0)$  para  $1 \leq j \leq s$ , com a desigualdade estrita para pelo menos uma  $j \in \{1, \dots, s\}$ .

Então  $x$  é factível para  $(P_k(\varepsilon))$  e  $f_k(x) \leq f_k(x_0)$ . Logo,  $f_k(x) = f_k(x_0)$ , e, por unicidade, teríamos  $x = x_0$ , o que é absurdo, pois  $f_j(x) \leq f_j(x_0)$  para  $j = 1, \dots, s$  com a desigualdade estrita para pelo menos um  $j$ . □

## Capítulo 5

# Formalismo Dubovitskii-Milyutin Multiobjetivo: solução Pareto fraca

Neste capítulo estudaremos as condições de otimalidade para problemas multiobjetivo utilizando o conceito de solução Pareto fraca.

### 5.1 Condições Necessárias e Suficientes para otimalidade de Pareto fraca

O próximo teorema é uma generalização do Teorema de Dubovitskii-Milyutin (Teorema 2.9) para o problema multiobjetivo (MOP), para a obtenção de condições necessárias para a Pareto eficiência fraca do mesmo. No caso de  $s = 1$ , o problema é escalar e retorna ao Teorema 2.9.

**Teorema 5.1** [Dubovitskii-Milyutin Multiobjetivo: solução Pareto fraca] *Sejam  $f : X \rightarrow \mathbb{R}^s$ ,  $K_i$  o cone de direções de descida de  $f_i(x)$  em  $x_0$  para  $i = 1, \dots, s$ ,  $Q_i$  o conjunto de restrições com  $\text{int}Q_i \neq \emptyset$  para  $i = 1, \dots, p$ ,  $Q_{p+1}$  o conjunto de restrições com  $\text{int}Q_{p+1} = \emptyset$ , com  $x_0 \in Q = \bigcap_{i=1}^{p+1} Q_i$ ,  $M_i$  o cone de direções factíveis de  $Q_i$  em  $x_0$  para  $i = 1, \dots, p$  e  $M_{p+1}$  o cone de direções tangentes de  $Q_{p+1}$  em  $x_0$ . Os cones  $K_i$ ,  $M_j$ , para  $i = 1, \dots, s$  e  $j = 1, \dots, p+1$  são todos convexos. Se  $x_0$  é solução fracamente eficiente local para o problema MOP então*

$$\sum_{i=1}^s h_i + \sum_{k=1}^{p+1} l_k = 0;$$

onde  $h_i \in K_i^*$  para  $i = 1, \dots, s$  e  $l_k \in M_k^*$  para  $k = 1, \dots, p+1$  são funcionais lineares, não todos identicamente nulos.

*Demonstração.* Uma condição necessária para  $x_0$  ser solução fracamente eficiente local para o problema MOP é

$$\left( \bigcap_{i=1}^s K_i \right) \cap \left( \bigcap_{k=1}^{p+1} M_k \right) = \emptyset. \quad (5.1)$$

Para ver isso assuma o contrário, ou seja, que existe  $z \in X$  que pertence a  $\left( \bigcap_{i=1}^s K_i \right) \cap \left( \bigcap_{k=1}^{p+1} M_k \right)$  e portanto são válidas

- $\exists U_z^i, \bar{\alpha}^i > 0 : f_i(x_0 + \alpha y) < f_i(x_0) \forall y \in U_z^i, \forall \alpha \in (0, \bar{\alpha}^i)$ ; para todo  $j = 1, \dots, s$ ;
- $\exists U_z^k, \bar{\alpha}^k > 0 : x_0 + \alpha y \in Q_k, \forall y \in U_z^k, \forall \alpha \in (0, \bar{\alpha}^k)$ ; para todo  $k = 1, \dots, p$ ;

onde  $U_z^i$  e  $U_z^k$  são vizinhanças de  $z$ . Seja  $U_z$  a vizinhança de  $z$  definida por

$$U_z = \left[ \bigcap_{i=1}^s U_z^i \right] \cap \left[ \bigcap_{k=1}^p U_z^k \right],$$

e seja  $\bar{\alpha}$  o número real definido por

$$\bar{\alpha} = \min(\bar{\alpha}^i (i = 1, \dots, s), \bar{\alpha}^k (k = 1, \dots, p)).$$

Como  $z$  também pertence a  $M_{p+1}$ , existe  $\hat{\alpha} \in (0, \bar{\alpha})$  e  $\hat{y} \in U_z$  tal que  $x_0 + \hat{\alpha}\hat{y} \in Q_{p+1}$ . O vetor  $\hat{x}_0 = x_0 + \hat{\alpha}\hat{y}$  satisfaz  $x_0 \neq \hat{x}_0 \in Q$ , onde  $Q = \bigcap_{i=1}^{p+1} Q_i$  e  $f_i(\hat{x}_0) < f_i(x_0)$  para todo  $i = 1, \dots, s$  e isso contradiz o fato de  $x_0$  ser solução fracamente eficiente local. A conclusão agora é obtida aplicado o Lema 2.8 à equação (5.1).  $\square$

Vejamos agora sob quais hipóteses adicionais as condições necessárias para otimalidade de Pareto fraca são suficientes.

**Teorema 5.2** *Sejam  $f_i : X \rightarrow \mathbb{R}$  funções convexas para todo  $i = 1, \dots, s$ .  $K_i$  o cone de direções de descida de  $f_i(x)$  em  $x_0$  para  $i = 1, \dots, s$ ,  $M_i$  o cone de direções factíveis de  $Q_i$  em  $x_0$  para  $i = 1, \dots, p$  e  $M_{p+1}$  o cone de direções tangentes de  $Q_{p+1}$  em  $x_0$ . Os conjuntos de restrições  $Q_i$  são convexos para todo  $i = 1, \dots, p+1$  e  $Q_{p+1} \cap (\bigcap_{i=1}^p \text{int}Q_i) \neq \emptyset$ . Então,  $x_0$  é solução fracamente eficiente para o problema MOP se, e somente se,*

$$\sum_{i=1}^s h_i + \sum_{k=1}^{p+1} l_k = 0;$$

onde  $h_i \in K_i^*$  para  $j = 1, \dots, s$  e  $l_k \in M_k^*$  para  $k = 1, \dots, p+1$  são funcionais lineares, não todos identicamente nulos.

*Demonstração.*

- **Necessidade** Como  $f_i(x)$  é função convexa, pelo Teorema 2.12, é regular de descida, para  $i = 1, \dots, s$ . Como  $Q_i$  são conjuntos convexos para  $i = 1, \dots, n$ , pelo Teorema 2.16 os cones  $M_i$ ,  $i = 1, \dots, n$  são convexos. Além disso o cone  $M_{p+1}$  é convexo porque o conjunto  $Q_{p+1}$  é convexo, Teorema 2.18. Assim todas as hipóteses do Teorema 5.1 são satisfeitas e a condição segue.
- **Suficiência** Suponha que  $x_0$  não é solução fracamente eficiente, então existe  $x_1 \in Q$ ,  $x_1 \neq x_0$  tal que  $f_i(x_1) < f_i(x_0)$  para todo  $i = 1, \dots, s$ . Por hipótese existe  $\bar{x} \in Q_{p+1} \cap (\bigcap_{i=1}^p \text{int}Q_i)$ . Defina, para  $0 < \lambda < 1$ ,  $x_\lambda = \lambda\bar{x} + (1-\lambda)x_1$ , então a convexidade implica que  $x_\lambda \in Q_i$  para  $i = 1, \dots, p+1$ . Além disso, pela Proposição 1.22,  $x_\lambda \in \text{int}Q_i$  para  $i = 1, \dots, p$ . Como  $f_i$  é contínua para todo  $i$ , podemos tomar  $\lambda \in (0, 1)$  suficientemente pequeno de modo que  $f_i(x_\lambda) < f_i(x_0)$  para  $i = 1, \dots, s$ . Defina  $h = x_\lambda - x_0$ . Por definição  $h \neq 0$ . Vamos mostrar que:

- (i)  $h \in M_k$  para todo  $k = 1, \dots, p$ ;
- (ii)  $h \in M_{p+1}$ ;
- (iii)  $h \in K_i$  para todo  $i = 1, \dots, s$ .

Dessas três afirmações temos que  $(\bigcap_{i=1}^s K_i) \cap (\bigcap_{k=1}^{p+1} M_k) \neq \emptyset$  o que contradiz a equação mencionada no Lema 2.8.

- (i) Como  $x_0 \in Q_i$  e  $x_\lambda \in \text{int}Q_i$ , para todo  $i = 1, \dots, p$ , pela Proposição 1.22 (a) temos

$$x_0 + \varepsilon h = x_0 + \varepsilon(x_\lambda - x_0) = (1 - \varepsilon)x_0 + \varepsilon x_\lambda \in \text{int}Q_i$$

para todo  $0 < \varepsilon < 1$ . Mas  $x_0 + \varepsilon h \in \text{int}Q_i$  para todo  $0 < \varepsilon < 1$  significa que  $h \in M_i$  para todo  $i = 1, \dots, p$ .

- (ii) Para todo  $0 < \varepsilon < 1$ ,  $x_0 + \varepsilon h \in Q_{p+1}$ , pois  $x_0$  e  $x_\lambda$  pertencem a  $Q_{p+1}$  e assim  $h \in M_{p+1}$ .
- (iii) Para  $i = 1, \dots, s$ ,  $f_i(x_\lambda) < f_i(x_0)$ . Para  $0 < \varepsilon < 1$  temos da convexidade de  $f_i$  que

$$f_i(x_0 + \varepsilon h) = f_i((1 - \varepsilon)x_0 + \varepsilon x_\lambda) \leq (1 - \varepsilon)f_i(x_0) + \varepsilon f_i(x_\lambda).$$

Então a derivada direcional de  $f_i$ ,  $i = 1, \dots, s$  em  $x_0$  na direção de  $h$  é

$$\begin{aligned} f'_i(x_0, h) &= \lim_{\varepsilon \rightarrow 0^+} \frac{f_i(x_0 + \varepsilon h) - f_i(x_0)}{\varepsilon} \\ &\leq \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon[f_i(x_\lambda) - f_i(x_0)]}{\varepsilon} \\ &= f_i(x_\lambda) - f_i(x_0) < 0. \end{aligned}$$

Assim, pelo Teorema 2.12,  $h \in K_i$ .

□

## 5.2 Aplicação a Problemas de Programação Matemática Vetorial

Considere o seguinte problema de programação matemática vetorial:

$$\begin{aligned} &\text{Minimizar } f(x) = (f_1(x), \dots, f_s(x)) \\ &\text{sujeito a: } \begin{aligned} &g_j(x) \leq 0, \quad j = 1, \dots, p \\ &h_k(x) = 0, \quad k = 1, \dots, m, \end{aligned} \end{aligned} \quad (\text{VOP})$$

onde  $f : X \rightarrow \mathbb{R}^s$ ,  $g_j : X \rightarrow \mathbb{R}$ ,  $j = 1, \dots, p$  e  $h_k : X \rightarrow \mathbb{R}$ ,  $k = 1, \dots, m$ , sendo  $X$  um espaço de Banach.

**Teorema 5.3** *Sejam  $f_i$ ,  $g_j$ ,  $i = 1, \dots, s$  e  $j = 1, \dots, p$  funções diferenciáveis em uma vizinhança de  $x_0$ ,  $h_k$ ,  $k = 1, \dots, m$  estritamente diferenciáveis em uma vizinhança de  $x_0$ . Se  $x_0$  é solução fracamente eficiente local para o problema (VOP) então existem  $\lambda \in \mathbb{R}^s$ ,  $\mu \in \mathbb{R}^p$  e  $\alpha \in \mathbb{R}^m$  com  $(\lambda, \mu, \alpha) \neq 0$ ,  $\lambda_i \geq 0$  para  $i = 1, \dots, s$  e  $\mu_j \geq 0$  para  $j = 1, \dots, p$ , tais que*

$$\sum_{i=1}^s \lambda_i f'_i(x_0) + \sum_{j=1}^p \mu_j g'_j(x_0) + \sum_{k=1}^m \alpha_k h'_k(x_0) = 0, \quad (5.2)$$

$$\sum_{j=1}^p \mu_j g_j(x_0) = 0. \quad (5.3)$$

*Demonstração.* Dividiremos a demonstração em 2 passos. No primeiro faremos uma análise dos casos triviais e no segundo passo a análise dos casos não triviais baseada no formalismo de Dubovitski-Milyutin.

**1° passo** (Casos triviais)

Se  $f'_l(x_0) = 0$ , para algum  $l \in \{1, \dots, s\}$ , então as equações (5.2), (5.3) são satisfeitas, com  $\lambda_l = 1$ ,  $\lambda_i = 0$  para  $i \neq l$ ,  $\mu = 0$  e  $\alpha = 0$ .

Se  $g'_r(x_0) = 0$  com  $g_r(x_0) = 0$  para algum  $r \in \{1, \dots, p\}$ , então (5.2), (5.3) são satisfeitas com  $\mu_r = 1$ ,  $\mu_j = 0$ , para  $j \neq r$ ,  $\lambda = 0$  e  $\alpha = 0$ .

Se os gradientes  $\{h'_k(x_0)\}_{k=1}^m$  são linearmente dependentes então

$$\sum_{k=1}^m \gamma_k h'_k(x_0) = 0 \text{ com } \sum_{k=1}^m \gamma_k^2 \neq 0$$

e (5.2), (5.3) valem com a escolha de  $\lambda = 0$ ,  $\mu = 0$  e  $\alpha_k = \gamma_k$  para  $k = 1, \dots, m$ .

**2° passo** (Casos não triviais)

Para os funcionais objetivo  $\{f_i(x)\}_{i=1}^s$ , de acordo com o Teorema 2.11, o cone de direções de descida é dado por

$$D_i = \{d \in X : f'_i(x_0)d < 0\}$$

e pelo Teorema 2.20

$$D^* = \{-\lambda_i f'_i(x_0) : \lambda_i \geq 0\} \quad (5.4)$$

para  $i = 1, \dots, s$ .

Agora, determinemos  $M_j$  o cone de direções factíveis para  $Q_j = \{x \in X : g_j(x) \leq 0\}$ ,  $j = 1, \dots, p$ . Se  $j$  é tal que  $g_j(x_0) < 0$ , então  $M_j = X$  e  $M_j^* = \emptyset$ . Agora se  $j$  é tal que  $g_j(x_0) = 0$  e  $g'_j(x_0) \neq 0$ , como  $g_j$  são diferenciáveis em  $x_0$  para  $j = 1, \dots, p$  pelo Teorema 2.15 temos

$$M_j = \{d \in X : g'_j(x_0)d < 0\}$$

e

$$M_j^* = \{-\mu_j g'_j(x_0) : \mu_j \geq 0\}. \quad (5.5)$$

Assuma que  $\{h'_k(x_0)\}_{k=1}^m$  são linearmente independentes. Neste caso o cone dual do cone de direções tangentes de  $Q_{p+1}$  no ponto  $x_0$  é  $M_{p+1}^* = \{\sum_{k=1}^m \alpha_k h'_k(x_0) : \alpha_k \geq 0\}$  (conforme obtido na demonstração do Teorema 3.1).

Aplicando o Teorema 5.1 segue o resultado.

□

Observamos que a argumentação utilizada na prova do Teorema 5.3 é bastante similar à utilizada na demonstração do Teorema 3.1, referente às condições necessárias de Fritz-John para problemas com um único objetivo.

# Capítulo 6

## Formalismo de Dubovitskii-Milyutin Multiobjetivo: solução de Pareto

Neste estudo desenvolveremos condições de otimalidade para problemas multiobjetivo utilizando o conceito de solução de Pareto.

### 6.1 Condições Necessárias para a otimalidade de Pareto

Para abordarmos o problema multiobjetivo, considerando a otimalidade de Pareto, através do formalismo de Dubovitskii-Milyutin precisamos do conceito de direção de não crescimento.

**Definição 6.1** Um vetor  $h$  é dito uma **direção de não crescimento** do funcional  $f(x)$  no ponto  $x_0$  se existem uma vizinhança  $U$  de  $h$  e  $\bar{\alpha} > 0$  tais que para todo  $\alpha \in (0, \bar{\alpha})$  e qualquer  $\bar{h} \in U$

$$f(x_0 + \alpha\bar{h}) \leq f(x_0).$$

**Lema 6.2** As direções de não crescimento geram um cone aberto com vértice na origem, que contém o cone de direções de descida no mesmo ponto.

*Demonstração.* Seja  $z$  uma direção de não crescimento e  $\lambda > 0$  arbitrário, então  $\lambda z$  também é uma direção de não crescimento onde, ao invés de uma vizinhança  $U$  e um número real  $\bar{\alpha} > 0$ , a vizinhança  $\lambda U$  e  $\hat{\alpha} = \frac{\bar{\alpha}}{\lambda} > 0$  são tomados. Logo as direções de não crescimento geram um cone com vértice na origem. Se  $z$  é uma direção de não crescimento então existe uma vizinhança  $U$  de  $z$ , tal que todo  $y \in U$  é uma direção de não crescimento e portanto o cone é aberto. A inclusão é clara.  $\square$

Denotaremos por  $K_i$  o cone de direções de descida para  $f_i(x)$  no ponto  $x_0$ ,  $L_i$  o cone de direções de não crescimento para  $f_i(x)$  em  $x_0$  para  $i = 1, \dots, s$ ,  $M_j$  o cone de direções factíveis para  $Q_j$ ,  $j = 1, \dots, p$  em  $x_0$  e  $M_{p+1}$  o cone de direções tangentes para  $Q_{p+1}$  em  $x_0$ .

O seguinte teorema nos fornece as condições necessárias de otimalidade local de Pareto para o problema de otimização multiobjetivo (MOP). Quando aplicado ao caso particular  $s = 1$ , coincide com o Teorema de Dubovitskii-Milyutin (Teorema 2.9).

**Teorema 6.3** [Dubovitskii-Milyutin Multiobjetivo: solução de Pareto]

Sejam  $f : X \rightarrow \mathbb{R}^s$ ,  $Q_j$  o conjunto de restrições com  $\text{int}Q_j \neq \emptyset$  para  $j = 1, \dots, p$ ,  $Q_{p+1}$  o conjunto de restrições com  $\text{int}Q_{p+1} = \emptyset$  e  $x_0 \in Q = \bigcap_{j=1}^{p+1} Q_j$ . Suponha que os cones  $K_i$ ,  $L_j$ ,  $M_j$ , para  $i = 1, \dots, s$  e  $j = 1, \dots, p+1$  são todos convexos. Se  $x_0$  é Pareto ótimo local para o problema (MOP) então as seguintes  $s$  equações devem valer

$$h_i + \sum_{\substack{j=1 \\ j \neq i}}^s h_j^{(i)} + \sum_{k=1}^{p+1} l_k^{(i)} = 0, \quad i = 1, \dots, s,$$

onde  $h_i \in K_i^*$ ,  $h_j^{(i)} \in L_j^*$  para  $j = 1, \dots, s$ ,  $j \neq i$  e  $l_k^{(i)} \in M_k^*$  para  $k = 1, \dots, p+1$  são para todo  $i = 1, \dots, s$  funcionais lineares, não todos identicamente nulos.

*Demonstração.* Uma condição necessária para que  $x_0$  seja mínimo local de Pareto para o problema (MOP) é

$$K_i \cap \left( \bigcap_{\substack{j=1 \\ j \neq i}}^s L_j \right) \cap \left( \bigcap_{k=1}^{p+1} M_k \right) = \emptyset, \quad \text{para todo } i = 1, \dots, s. \quad (6.1)$$

Para ver isso assuma o contrário, ou seja, que existe  $i$ ,  $1 \leq i \leq s$ , tal que

$$K_i \cap \left( \bigcap_{\substack{j=1 \\ j \neq i}}^s L_j \right) \cap \left( \bigcap_{k=1}^{p+1} M_k \right) \neq \emptyset.$$

Então existe  $z \in \mathbb{R}^n$  que pertence a  $K_i \cap \left( \bigcap_{\substack{j=1 \\ j \neq i}}^s L_j \right) \cap \left( \bigcap_{k=1}^{p+1} M_k \right)$  e portanto são válidas:

- $\exists U_z^i, \bar{\alpha}^i > 0 : f_i(x_0 + \alpha y) < f_i(x_0) \forall y \in U_z^i, \forall \alpha \in (0, \bar{\alpha}^i)$ ;
- $\exists U_z^j, \bar{\alpha}^j > 0 : f_j(x_0 + \alpha y) \leq f_j(x_0) \forall y \in U_z^j, \forall \alpha \in (0, \bar{\alpha}^j)$ ; para todo  $j = 1, \dots, s$ ,  $j \neq i$ ;
- $\exists U_z^k, \bar{\alpha}^k > 0 : x_0 + \alpha y \in Q_k, \forall y \in U_z^k, \forall \alpha \in (0, \bar{\alpha}^k)$ ; para todo  $k = 1, \dots, p$ ;

onde  $U_z^i$ ,  $U_z^j$  e  $U_z^k$  são vizinhanças de  $z$ . Seja  $U_z$  a vizinhança de  $z$  definida por

$$U_z = U_z^i \cap \left[ \bigcap_{\substack{j=1 \\ j \neq i}}^s U_z^j \right] \cap \left[ \bigcap_{k=1}^p U_z^k \right],$$

e seja  $\bar{\alpha}$  o número real definido por

$$\bar{\alpha} = \min\{\bar{\alpha}^i, \bar{\alpha}^j (j = 1, \dots, s; j \neq i), \bar{\alpha}^k, (k = 1, \dots, p)\}.$$

Como  $z$  também pertence a  $M_{p+1}$ , existe  $\hat{\alpha} \in (0, \bar{\alpha})$  e  $\hat{y} \in U_z$  tal que  $x_0 + \hat{\alpha}\hat{y} \in Q_{p+1}$ . O vetor  $\hat{x}_0 = x_0 + \hat{\alpha}\hat{y}$  satisfaz  $x_0 \neq \hat{x}_0 \in Q$ , onde  $Q = \bigcap_{i=1}^{p+1} Q_i$  e  $f_j(\hat{x}_0) \leq f_j(x_0)$  para todo  $j = 1, \dots, s$  com pelo menos uma desigualdade estrita e isso contradiz a otimalidade de Pareto de  $x_0$ . A conclusão agora é obtida aplicado o Lema 2.8 à equação (6.1) para  $i = 1, \dots, s$ .  $\square$

**Observação 7** Se para algum índice  $i$  ( $1 \leq i \leq s$ ),  $K_i$  é vazio então  $K_i^* = \mathbb{R}^n$  e pode-se escolher algum  $h_j^{(i)} \in L_j^*$  arbitrário para  $j = 1, \dots, s$ ,  $j \neq i$  e algum  $l_k^{(i)} \in M_k^*$  arbitrário, para  $k = 1, \dots, p+1$  e colocar

$$h_i = - \left[ \sum_{\substack{j=1 \\ j \neq i}}^s h_j^{(i)} + \sum_{k=1}^{p+1} l_k^{(i)} \right].$$

Então  $h_i \in K_i^*$  e a condição necessária para Pareto ótimo local de  $x_0$  é satisfeita para tal  $i$ . Se para algum índice  $j$ ,  $L_j = \emptyset$  então o mesmo raciocínio é verdadeiro para todo  $i = 1, \dots, s$  e a condição necessária para Pareto ótimo local de  $x_0$  é satisfeita para todo  $i = 1, \dots, s$ .

**Corolário 6.4** Se além das hipóteses do Teorema 6.3,  $f$  é tal que para todo  $i = 1, \dots, s$ ,  $K_i^* = L_i^*$  então as  $s$  equações da conclusão se reduzem à seguinte equação

$$\sum_{i=1}^s h_i + \sum_{k=1}^{p+1} l_k = 0,$$

onde  $h_i \in K_i^*$  para  $i = 1, \dots, s$  e  $l_k \in M_k^*$  para  $k = 1, \dots, p+1$ , são funcionais lineares, não todos identicamente nulos.

## 6.2 Caracterização dos cones duais

Na sequência verificaremos que sob hipóteses de convexidade e subdiferenciabilidade dos funcionais  $f_i$ , a hipótese adicional  $K_i^* = L_i^*$  é satisfeita, sem a necessidade de  $K_i = L_i$ .

**Definição 6.5** Seja  $Q \subset \mathbb{R}^n$  um conjunto não vazio e  $x_0 \in Q$ . Temos três diferentes definições de cone de direções factíveis de  $Q$  no ponto  $x_0$ :

- $F_1 \equiv F_1(Q; x_0) = \{z \in \mathbb{R}^n / \exists \alpha > 0, x_0 + \alpha z \in Q\}$ ;
- $F_2 \equiv F_2(Q; x_0) = \{z \in \mathbb{R}^n / \exists \bar{\alpha} > 0, x_0 + \alpha z \in Q \forall \alpha \in (0, \bar{\alpha})\}$ ;
- $F_3 \equiv F_3(Q; x_0) = \{z \in \mathbb{R}^n / \exists U(z), \exists \bar{\alpha} > 0, x_0 + \alpha y \in Q, \forall y \in U(z) \text{ e } \forall \alpha \in (0, \bar{\alpha})\}$ , onde  $U(z)$  é vizinhança de  $z$ .

Os três cones são bastante usados na literatura e  $F_3(S; x_0)$  é o cone de direções factíveis como definido no Capítulo 2.

Conforme demonstrado no Capítulo 2, Lema (2.4), o cone  $F_3(S; x_0)$  é aberto. Quanto aos cones  $F_1(S; x_0)$  e  $F_2(S; x_0)$  não podemos garantir que eles são abertos, assim para podermos utilizar o Teorema de Dubovitskii-Milyutin devemos utilizar o cone  $F_3(S; x_0)$ .

**Lema 6.6**  $F_1(Q; x_0)$  é o cone mínimo que contém  $Q - x_0$ .

*Demonstração.*  $F_1(Q; x_0)$  é o cone mínimo que contém  $Q - x_0$  se, e somente se,  $F_1(Q; x_0)$  é a interseção de todos os cones  $K_\alpha$  que contém  $Q - x_0$ .

Sejam  $A$  um conjunto de índices e  $K_\alpha$  cones, tais que  $Q - x_0 \subset K_\alpha$ ,  $\forall \alpha \in A$ . Como  $F_1(Q; x_0)$  é um cone e contém  $Q - x_0$  então  $\cap_{\alpha \in A} K_\alpha \subset F_1(Q; x_0)$ .

Por outro lado seja  $z \in F_1(Q; x_0)$ , então existe  $\beta > 0$  tal que  $x_0 + \beta z \in Q$ . Logo  $\beta z \in Q - x_0 \subset K_\alpha$ , ou seja  $\beta z \in K_\alpha$  para  $\alpha \in A$ . Como  $K_\alpha$  é cone segue que  $z \in K_\alpha$  para todo  $\alpha \in A$  e portanto  $z \in \cap_{\alpha \in A} K_\alpha$ . □

**Lema 6.7** *Se  $Q$  é convexo e  $F_3 \neq \emptyset$  então  $\text{cl}F_3 = \text{cl}F_1$ .*

*Demonstração.* Da definição dos cones temos que  $F_3(Q; x_0) \subset F_1(Q; x_0)$ . Pelo Lema 6.6,  $Q - x_0 \subset F_1(Q; x_0)$ . Além disso, como  $F_3$  é aberto e não vazio por hipótese então  $\text{int}F_1 \neq \emptyset$  e assim  $\text{int}Q \neq \emptyset$ . Vamos mostrar que  $\text{int}Q - x_0 \subset F_3$ . Seja  $x \in \text{int}Q$ , então existe uma vizinhança  $V$  de zero tal que  $x + V \subset \text{int}Q$ .

Se  $y \in x - x_0 + V$  e  $\delta \in (0, 1)$  temos

$$x_0 + \delta y = (1 - \delta)x_0 + \delta(x_0 + y).$$

Mas  $x_0 + y \in x + V \subset \text{int}Q$  e  $x_0 \in \text{cl}Q$ , então como  $Q$  é convexo, pela Proposição 1.22 (a), temos que  $x_0 + \delta y \in \text{int}Q$  para  $\delta \in (0, 1)$ . Isto mostra que  $x - x_0 \in F_3$ , ou seja  $\text{int}Q - x_0 \subset F_3$ .

Pela convexidade de  $Q$  e como  $\text{int}Q \neq \emptyset$ , pela Proposição 1.22 (d), temos  $\text{cl}(\text{int}Q) - x_0 = \text{cl}Q - x_0$ , logo

$$Q - x_0 \subset \text{cl}Q - x_0 \subset \text{cl}F_3.$$

Mas como  $F_1$  é cone mínimo que contém  $Q - x_0$  então  $F_1 \subset \text{cl}F_3$ . Isso juntamente como o fato de  $F_3 \subset F_1$  implica que  $\text{cl}F_1 = \text{cl}F_3$ .  $\square$

**Teorema 6.8** *Se  $Q$  é convexo e  $F_3(Q; x_0) \neq \emptyset$  então  $F_1^*(Q, x_0) = F_2^*(Q, x_0) = F_3^*(Q, x_0)$ .*

*Demonstração.* Pela definição dos cones temos que  $F_3 \subset F_2 \subset F_1$ . O Lema 6.7 implica que  $\text{cl}F_1 = \text{cl}F_2 = \text{cl}F_3$ . O resultado segue do fato que o cone e de seu fecho são iguais, Teorema 1.10 (b).  $\square$

**Definição 6.9** *Os seguintes conjuntos são denominados conjuntos de nível de um funcional  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , onde  $x_0 \in \text{dom}f$ :*

- $S_f^{\leq}(x_0) = \{x \in \mathbb{R}^n; f(x) \leq f(x_0)\};$
- $S_f^{<}(x_0) = \{x \in \mathbb{R}^n; f(x) < f(x_0)\};$
- $S_f^{\bar{}}(x_0) = \{x \in \mathbb{R}^n; f(x) = f(x_0)\}.$

O cone  $F_3(S_f^{\leq}(x_0); x_0)$  é o cone de direções de não crescimento do funcional  $f$  no ponto  $x_0$ . Considerando o Teorema 6.8, seu cone dual será caracterizado pelo cálculo de  $F_1^*(S_f^{\leq}(x_0); x_0)$ . Os seguintes teoremas e lemas serão necessários.

**Teorema 6.10** *Seja  $x_0 \in \mathbb{R}^n$  e  $Q \subset \mathbb{R}^n$ , então  $N(Q; x_0) = -F_1^*(Q; x_0)$ .*

*Demonstração.* Seja  $y \in -F_1^*(Q; x_0)$ , então  $z^T y \leq 0, \forall z \in F_1(Q; x_0)$ . Como  $Q - x_0 \subset F_1(Q; x_0)$ , de acordo com o Lema 6.6, em particular  $z^T y \leq 0, \forall z \in Q - x_0$ . Logo  $(x - x_0)^T y \leq 0, \forall x \in Q$ , ou seja  $y \in N(Q; x_0)$ . Para provar a outra inclusão considere  $y \in N(Q; x_0)$ , então  $(x - x_0)^T y \leq 0, \forall x \in Q$ . Em particular para todo  $z \in F_1(Q; x_0)$  e  $\alpha > 0$  tal que  $x_0 + \alpha z \in Q$  temos  $(x_0 + \alpha z - x_0)^T y \leq 0, \forall z \in F_1(Q; x_0)$ . Logo  $z^T y \leq 0, \forall z \in F_1(Q; x_0)$ , ou seja  $y \in -F_1^*(Q; x_0)$ .  $\square$

**Teorema 6.11** *Seja  $f$  uma função convexa própria,  $x_0 \in \text{int}(\text{dom}f)$  e  $\min f < f(x_0)$ . Um vetor  $y \in N(S_f^{\leq}(x_0); x_0)$  se, e somente se, existe  $\lambda \geq 0$  tal que  $y \in \lambda \partial f(x_0)$ .*

*Demonstração.* Do Teorema 1.23 (c) temos que  $N(S_f^{\leq}(x_0); x_0)$  é o fecho do cone convexo gerado por  $\partial f(x_0)$ . Como  $f$  é função convexa própria,  $x_0 \in \text{int}(\text{dom} f)$  e  $\min f < f(x_0)$ , então o subdiferencial é um conjunto convexo e fechado. Além disso pelo Teorema 1.23 (a) o conjunto  $\partial f(x_0)$  é não vazio e limitado, e como  $x_0$  não é minimizador de  $f$  então,  $\partial f(x_0)$  não contém a origem. Agora aplicando o Teorema 1.22 (e), obtemos que o fecho do cone convexo gerado por  $\partial f(x_0)$  é a união dos conjunto  $\lambda \partial f(x_0)$  para  $\lambda \geq 0$ .  $\square$

### Lema 6.12

- a) Se  $Q \subset \mathbb{R}^n$  é convexo e  $x_0 \in \text{cl}(Q)$ , então  $\text{int} Q \neq \emptyset$  se, e somente se,  $F_3(Q; x_0) \neq \emptyset$ .
- b) Se  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  é uma função convexa própria e  $x_0 \in \text{int}(\text{dom} f)$ , então

$$S_f^{\leq}(x_0) \neq \emptyset \Rightarrow \text{int} S_f^{\leq}(x_0) \neq \emptyset.$$

*Demonstração.*

- a) Da definição de  $F_3$  temos que se  $\text{int} Q = \emptyset$  então  $F_3(Q; x_0) = \emptyset$ . Agora se  $x \in \text{int} Q$ , então  $y = x - x_0 \in F_3$ , pois  $x_0 + \alpha y = (1 - \alpha)x_0 + \alpha(x_0 + y)$ , onde  $x_0 \in \text{cl}(Q)$  e  $x_0 + y = x \in \text{int} Q$ . Pelo Princípio do Segmento (Proposição 1.22 (a)),  $x_0 + \alpha y \in \text{int} Q \neq \emptyset$  para todo  $\alpha \in (0, 1)$ . Assim,  $\text{int} Q - x_0 \subset F_3(Q; x_0)$  e segue a conclusão desejada.
- b) Temos que  $S_f^{\leq}(x_0) = S_f^{\leq}(x_0) \cup S_f^{\overline{\leq}}(x_0)$  e como  $S_f^{\overline{\leq}}(x_0)$  não tem pontos interiores então  $\text{int} S_f^{\leq}(x_0) \subset S_f^{\leq}(x_0)$ . Para  $x_0 \in \text{int}(\text{dom} f)$  a função  $f$  é contínua, logo  $S_f^{\leq}(x_0)$  é um conjunto aberto, sendo assim é igual ao seu interior e segue a conclusão.  $\square$

O próximo teorema nos fornece a caracterização do cone dual do cone de direções de não-crescimento de um funcional convexo não diferenciável.

**Teorema 6.13** *Seja  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  uma função convexa própria,  $x_0 \in \text{int}(\text{dom} f)$  e assumamos  $S_f^{\leq}(x_0) \neq \emptyset$ . Então*

$$F_i^*(S_f^{\leq}(x_0); x_0) = \{\lambda x : \lambda \leq 0, x \in \partial f(x_0)\}, \text{ para } i = 1, 2, 3.$$

*Demonstração.* Temos que  $x_0 \in S_f^{\leq}(x_0)$  que é um conjunto convexo. Pelo Lema 6.12,  $\text{int} S_f^{\leq}(x_0) \neq \emptyset$  implica que  $F_3(S_f^{\leq}(x_0); x_0) \neq \emptyset$ . Então tome  $Q = S_f^{\leq}(x_0)$  no Teorema 6.8, assim  $F_1^* = F_2^* = F_3^*$ . Além disso,  $S_f^{\leq}(x_0) \neq \emptyset$  equivale a  $\min f < f(x_0)$  como necessitamos no Teorema 6.11, e juntamente com o Teorema 6.10 obtemos

$$-F_i^*(S_f^{\leq}(x_0); x_0) = N(S_f^{\leq}(x_0); x_0) = \{\lambda x : x \in \partial f(x_0) \text{ e } \lambda \geq 0\}, \text{ para } i = 1, 2, 3,$$

ou seja,

$$F_i^*(S_f^{\leq}(x_0); x_0) = \{\lambda x : x \in \partial f(x_0) \text{ e } \lambda \leq 0\}, \text{ para } i = 1, 2, 3.$$

$\square$

A condição  $x_0 \in \text{int}(\text{dom} f)$  equivale a  $\partial f(x_0)$  limitado e não vazio, pelo Teorema 1.23 (a), e assim assegura a subdiferenciabilidade de  $f$  em  $x_0$ .

À luz do Teorema 6.13, o problema de encontrar o cone dual do cone de direções de não crescimento é equivalente a calcular o subdiferencial de  $f$  em  $x_0$ .

O próximo teorema que relaciona  $S_f^{\leq}(x_0)$  e  $S_f^{\overline{\leq}}(x_0)$  nos permite caracterizar o cone dual do cones de direções de descida.

**Teorema 6.14** *Seja  $f$  uma função própria convexa,  $x_0 \in \text{dom}f$  e assuma que  $\inf f < f(x_0)$ , então*

- a)  $f$  fechada  $\Rightarrow \text{cl}S_f^<(x_0) = S_f^<(x_0)$ .
- b)  $\text{dom}f$  relativamente aberto  $\Rightarrow \text{ri}S_f^<(x_0) = S_f^<(x_0)$ .

*Demonstração.*

- a) Como  $f$  é fechada  $\text{cl}f = f$ , logo do Teorema 1.23 (b) obtemos

$$\text{cl}\{x \in \text{dom}f : f(x) < \alpha\} = \{x \in \text{dom}f : \text{cl}f(x) \leq \alpha\} = \{x \in \text{dom}f : f(x) \leq \alpha\},$$

$$\text{ou seja, } \text{cl}S_f^<(x_0) = S_f^<(x_0).$$

- b) Por hipótese  $\text{ri}(\text{dom}f) = \text{dom}f$ , então do Teorema 1.23 (b)

$$\text{ri}\{x \in \text{dom}f : f(x) \leq \alpha\} = \{x \in \text{ri}(\text{dom}f) : f(x) < \alpha\} = \text{ri}\{x \in \text{dom}f : f(x) < \alpha\},$$

$$\text{ou seja, } \text{ri}S_f^<(x_0) = S_f^<(x_0).$$

□

**Teorema 6.15** *Se  $Q_1 \subset Q_2$  então  $F_3(Q_1; x_0) \subset F_3(Q_2; x_0)$ .*

*Demonstração.* Seja  $z \in F_3(Q_1; x_0)$ , então existe uma vizinhança  $U$  de  $z$  e  $\bar{\alpha} > 0$  tal que  $x_0 + \alpha y \in Q_1 \subset Q_2$ , para todo  $y \in U$  e todo  $\alpha \in (0, \bar{\alpha})$ . Portanto  $z \in F_3(Q_2; x_0)$ . □

**Teorema 6.16** *Se  $\emptyset \neq Q \subset \mathbb{R}^n$  então  $F_3(Q; x_0) = F_3(\text{int}Q; x_0)$ .*

*Demonstração.* Pelo Teorema 6.15,  $F_3(\text{int}Q; x_0) \subset F_3(Q; x_0)$ . Para provar a outra inclusão considere  $z \in F_3(Q; x_0)$ , então existe uma vizinhança  $V$  de  $z$  e  $\bar{\alpha} > 0$  tal que  $\forall y \in V$  e  $\alpha \in (0, \bar{\alpha})$ , temos que  $x_0 + \alpha y \in Q$ , ou seja  $A = x_0 + \bigcup_{\alpha \in (0, \bar{\alpha})} \alpha V \subset Q$ . Como  $A$  é aberto,  $A \subset \text{int}Q$  e portanto  $z \in F_3(\text{int}Q; x_0)$ . □

**Teorema 6.17** *Se  $\emptyset \neq Q \subset \mathbb{R}^n$  então  $\text{cl}F_1(\text{cl}Q; x_0) = \text{cl}F_1(Q; x_0)$ .*

*Demonstração.* Como  $Q \subset \text{cl}Q$ , pelo Teorema 6.15 temos  $F_1(Q; x_0) \subset F_1(\text{cl}Q; x_0)$ , logo  $\text{cl}F_1(Q; x_0) \subset \text{cl}F_1(\text{cl}Q; x_0)$ .

Para ver a outra inclusão tome  $y \in \text{cl}F_1(\text{cl}Q; x_0)$ .

- (a)  $y \in \text{cl}F_1(\text{cl}Q; x_0) \Rightarrow \exists \{z_n\}_{n=1}^{\infty} \subset F_1(\text{cl}Q; x_0)$  tal que  $z_n \rightarrow y$ , quando  $n \rightarrow \infty$ .
- (b)  $z_n \in F_1(\text{cl}Q; x_0)$ ,  $\forall n \geq 1 \Rightarrow \exists \alpha_n > 0$  tal que  $x_0 + \alpha_n z_n \in \text{cl}Q$ ,  $\forall n \geq 1$ .  
Denote  $x_0 + \alpha_n z_n = w_n$  e observe que
- (c)  $w_n \in \text{cl}Q \Rightarrow \forall n \geq 1 \exists \{w_n^k\}_{k=1}^{\infty} \subset Q$  tal que  $w_n^k \rightarrow w_n$ , quando  $k \rightarrow \infty$ .

Temos que mostrar que existe uma sequência  $\{u_i\}_{i=1}^{\infty} \subset F_1(Q; x_0)$  tal que  $u_i \rightarrow y$  quando  $i \rightarrow \infty$ .

Defina

$$u_n^k = \frac{w_n^k - x_0}{\alpha_n}, \quad \forall n \geq 1, \quad \forall k \geq 1.$$

Para todo  $n$  natural fixado, quando fazemos  $k \rightarrow \infty$  temos

$$u_n^k = \frac{w_n^k - x_0}{\alpha_n} \rightarrow \frac{w_n - x_0}{\alpha_n} = z_n$$

por causa de (c) e das definições e  $u_n^k$  e  $w_n$ . Assim para  $\varepsilon_i = 2^{-i}$  existe um número natural  $N_i$  tal que para todo  $k \geq N_i$ ,  $|u_n^k - z_i| < 2^{-i}$ . Agora para todo  $i \geq 1$  defina  $u_i = u_i^{N_i}$ .

Para todo  $n \geq 1$  e todo  $k \geq 1$ ,  $u_n^k \in F_1(Q; x_0)$ , pois

$$x_0 + \alpha_n u_n^k = x_0 + \alpha_n \left( \frac{w_n^k - x_0}{\alpha_n} \right) = w_n^k \in Q,$$

onde a existência de  $\alpha_n > 0$  é garantida por (b) e  $w_n^k \in Q$  de acordo com (c). Portanto  $u_i \in F_1(Q; x_0)$ ,  $\forall i \geq 1$ .

Para provar que a sequência  $\{u_i\}_{i=1}^{\infty}$  converge para  $y$  tome  $\varepsilon > 0$  arbitrário, (a) implica que existe um número natural  $M$  tal que para todo  $k \geq M$ ,  $|z_k - y| < \frac{\varepsilon}{2}$ .

Como  $\lim_{m \rightarrow \infty} 2^{-m} = 0$ , existe  $J \in \mathbb{N}$  tal que  $2^{-J} < \frac{\varepsilon}{2}$ . De acordo com a construção da sequência  $\{u_i\}_{i=1}^{\infty}$  temos que para todo  $i \geq J$ ,  $|u_i - z_i| < 2^{-k} \leq 2^{-J} < \frac{\varepsilon}{2}$ . Agora seja  $N = \max\{M, J\}$ , então para todo  $n \geq N$ ,  $|u_n - y| \leq |u_n - z_n| + |z_n - y| < \varepsilon$ .

□

**Teorema 6.18** *Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma função convexa própria e assuma que  $x_0 \in \text{int}(\text{dom}f)$  e  $S_f^<(x_0) \neq \emptyset$ . Se  $f$  é fechada ou tem o domínio efetivo relativamente aberto, então*

$$F_i^*(S_f^<(x_0); x_0) = F_j^*(S_f^<(x_0); x_0) \text{ para todo } i, j = 1, 2, 3, \quad (6.2)$$

e todos esses cones são iguais a  $\{\lambda x : \lambda \leq 0, x \in \partial f(x_0)\}$ .

*Demonstração.*  $S_f^<(x_0) \neq \emptyset$  implica que  $\min f < f(x_0)$ .

Sob a hipótese que  $f$  é um função fechada, temos do Teorema 6.14, item (a) que  $\text{cl}S_f^<(x_0) = S_f^<(x_0)$ . Escolha  $Q = S_f^<(x_0)$  no Teorema 6.17, então  $\text{cl}F_1(S_f^<(x_0); x_0) = \text{cl}F_1(S_f^<(x_0); x_0)$ . Agora pelo Teorema 1.10(b)  $F_1^*(S_f^<(x_0); x_0) = F_1^*(S_f^<(x_0); x_0)$  e  $S_f^<(x_0) \neq \emptyset$  implica que  $F_3(S_f^<(x_0); x_0) \neq \emptyset$  (Lema 6.12(a)), portanto podemos aplicar o Teorema 6.8, com  $S = S_f^<(x_0)$  para obter  $F_i^*(S_f^<(x_0); x_0) = F_1^*(S_f^<(x_0); x_0)$  para  $i = 1, 2, 3$ . Além disso,  $S_f^<(x_0)$  é aberto, convexo e não vazio (pois como  $f$  é convexa e  $x_0 \in \text{int}(\text{dom}f)$  por hipótese, então  $f$  é contínua em  $x_0$ ), assim  $\text{int}S_f^<(x_0) \neq \emptyset$  e o Lema 6.12(a) garante que  $F_3(S_f^<(x_0); x_0) \neq \emptyset$  e o Teorema 6.8 pode ser aplicado novamente, agora com  $Q = S_f^<(x_0)$  para obter (6.2).

Sob a hipótese que  $\text{dom}f$  é relativamente aberto temos do Teorema 6.14(b) que  $\text{ri}S_f^<(x_0) = S_f^<(x_0)$ . Aqui escolha  $Q = S_f^<(x_0)$  no Teorema 6.16, então

$$F_3(S_f^<(x_0); x_0) = F_3(S_f^<(x_0); x_0) \text{ e assim } F_3^*(S_f^<(x_0); x_0) = F_3^*(S_f^<(x_0); x_0).$$

Aplicando o Teorema 6.8 obtemos  $F_i^*(S_f^<(x_0); x_0) = F_j^*(S_f^<(x_0); x_0)$  para  $i, j = 1, 2, 3$ . A última afirmação segue o Teorema 6.13.

□

### 6.3 Condições necessárias e suficientes para otimalidade global de Pareto

Nessa Seção vamos mostrar que sob hipóteses adicionais o conjunto de  $s$  equações do Teorema 6.3 é condição necessária e suficiente para otimalidade global de Pareto para o problema multiobjetivo (MOP). Em problemas multiobjetivo, como em problemas escalares, convexidade das funções objetivo e dos conjuntos de restrições garante que ótimo local e global coincidem.

**Teorema 6.19** *Sejam  $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  funcionais convexos para  $i = 1, \dots, s$  e  $Q \subset \mathbb{R}^n$  conjunto convexo. Então todo ponto ótimo local de Pareto para o seguinte problema*

$$\begin{aligned} & \text{Minimizar } f(x) = (f_1(x), \dots, f_s(x)) \\ & \text{sujeito a } x \in Q \end{aligned} \tag{6.3}$$

*é também ponto ótimo global de Pareto.*

*Demonstração.* Seja  $x_0$  um ponto ótimo local de Pareto, isto é, existe uma vizinhança  $U$  de  $x_0$  tal que não existe  $x \neq x_0$  em  $Q \cap U$  tal que  $f_i(x) \leq f_i(x_0)$  para  $i = 1, \dots, s$  com pelo menos uma desigualdade estrita. Tome  $x_1 \in Q$  arbitrário, defina  $x_\lambda = x_0 + \lambda(x_1 - x_0)$  e escolha  $\lambda \in (0, 1)$  tal que  $x_\lambda \in U$ . A convexidade de  $Q$  assegura que  $x_\lambda \in Q$  para todo  $\lambda \in (0, 1)$ .

A otimalidade local de  $x_0$  significa

$$f_i(x_\lambda) \leq f_i(x_0), \quad i = 1, \dots, s \Rightarrow f_i(x_\lambda) = f_i(x_0), \quad i = 1, \dots, s.$$

Para provar a otimalidade global de  $x_0$  sobre  $Q$  vamos mostrar que

$$f_i(x_1) \leq f_i(x_0), \quad i = 1, \dots, s \Rightarrow f_i(x_1) = f_i(x_0), \quad i = 1, \dots, s.$$

De fato, a convexidade de  $f_i(x)$  implica que  $f_i(x_\lambda) \leq \lambda f_i(x_1) + (1 - \lambda)f_i(x_0)$  para todo  $i$ , e com  $f_i(x_1) \leq f_i(x_0)$  temos  $f_i(x_\lambda) \leq f_i(x_0)$  que por sua vez implica que  $f_i(x_\lambda) = f_i(x_0)$  e assim  $f_i(x_0) \leq \lambda f_i(x_1) + (1 - \lambda)f_i(x_0)$ . Como  $\lambda > 0$  obtemos que  $f_i(x_0) \leq f_i(x_1)$ . Juntamente com a hipótese que  $f_i(x_1) \leq f_i(x_0)$  temos que  $f_i(x_0) = f_i(x_1)$  para  $i = 1, \dots, s$ .  $\square$

Definiremos um tipo particular de convexidade, que será necessária na sequência.

**Definição 6.20** *A função  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  será chamada **Ponstein-convexa** se*

$$f(x_2) \leq f(x_1) \Rightarrow f(\lambda x_1 + \mu x_2) < f(x_1)$$

*sempre que  $x_1 \neq x_2$ , onde  $\lambda$  e  $\mu$  são números reais positivos sendo  $\lambda + \mu = 1$ .*

Esse tipo de convexidade foi chamado de “unnamed convexit” por Ponstein em [22] e abreviada como  $XC$ . Ressaltamos que a classe das funções Ponstein-convexas contém a classe das funções estritamente convexas (isto é,  $f(x_2) < f(x_1) \Rightarrow f(\lambda x_1 + \mu x_2) < f(x_1)$ ,  $\lambda, \mu > 0$ ,  $\lambda + \mu = 1$ ), logo tem uma interseção não vazia com as funções convexas. A relação entre estas classes de funções pode ser vista na Figura 6.1.

**Lema 6.21** *Se  $f$  é Ponstein-convexa então  $S_f^{\leq}(x_0)$  é um conjunto convexo.*

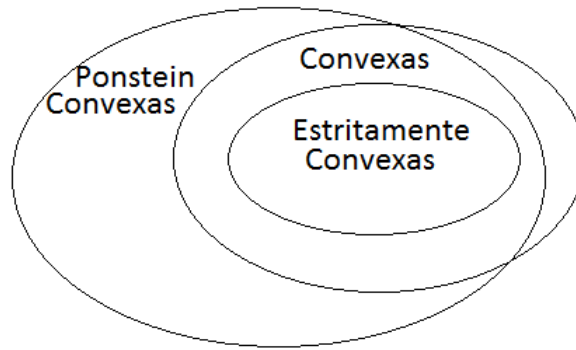


Figura 6.1: Comparação tipos de convexidade.

*Demonstração.* Sejam  $x_1, x_2 \in S_f^{\leq}(x_0)$ , com  $f(x_2) \leq f(x_1)$ , então como  $f$  é Ponstein-convexa para todo  $0 < \lambda < 1$  temos  $f((1 - \lambda)x_1 + \lambda x_2) < f(x_1) \leq f(x_0)$ . Assim  $(1 - \lambda)x_1 + \lambda x_2 \in S_f^{\leq}(x_0)$  e portanto  $S_f^{\leq}(x_0)$  é convexo.  $\square$

Para funções Ponstein-convexas temos o seguinte lema.

**Lema 6.22** *Seja  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  contínua e Ponstein-convexa. Se  $x_0 \in \text{dom} f$  e  $\inf f < f(x_0)$  então  $\text{int} S_f^{\leq}(x_0) \neq \emptyset$  e  $S_f^{\leq}(x_0) = \text{int} S_f^{\leq}(x_0)$ .*

*Demonstração.*  $x_0 \in S_f^{\leq}(x_0)$  e como  $\inf f < f(x_0)$  existe pelo menos um  $y_0 \in S_f^{\leq}(x_0)$ ,  $y_0 \neq x_0$ , isto é  $f(y_0) \leq f(x_0)$ . A Ponstein-convexidade implica que  $f(x_\lambda) < f(x_0)$  para todo  $0 < \lambda < 1$ , onde  $x_\lambda = \lambda x_0 + (1 - \lambda)y_0$  e a continuidade de  $f$  implica que existe alguma vizinhança de  $x_\lambda$  que está contida em  $S_f^{\leq}(x_0)$  e assim  $\text{int} S_f^{\leq}(x_0) \neq \emptyset$ .

Como  $f$  é Ponstein-convexa, pelo Lema 6.21,  $S_f^{\leq}(x_0)$  é convexo. Agora pela Proposição 1.22 (b),  $z \in \text{int} S_f^{\leq}(x_0)$  implica que existe  $\mu > 1$  tal que  $(1 - \mu)x_0 + \mu z \in S_f^{\leq}(x_0)$ , isto é,  $f((1 - \mu)x_0 + \mu z) \leq f(x_0)$ . Escolha  $\lambda = 1 - \frac{1}{\mu}$  e, aplicando a Ponstein-convexidade, obtemos que

$$f(\lambda x_0 + (1 - \lambda)((1 - \mu)x_0 + \mu z)) = f\left(\left(1 - \frac{1}{\mu}\right)x_0 + \frac{1}{\mu}((1 - \mu)x_0 + \mu z)\right) = f(z) < f(x_0),$$

assim  $z \in S_f^{\leq}(x_0)$ .

Agora seja  $z \in S_f^{\leq}(x_0)$ , então  $f(z) < f(x_0)$ . Suponha que  $z \notin \text{int} S_f^{\leq}(x_0)$ . De acordo com a Proposição 1.22 (c), existe  $y \in \mathbb{R}^n$  tal que para todo  $\varepsilon > 0$ ,  $z + \varepsilon y \notin S_f^{\leq}(x_0)$ , isto é,  $f(x_0) < f(z + \varepsilon y)$ . Tomando  $\varepsilon \rightarrow 0$ , temos da continuidade de  $f$  que  $f(x_0) \leq f(z)$ , o que é um absurdo.  $\square$

Considere o problema

$$\begin{aligned} &\text{Minimizar } f(x) = (f_1(x), \dots, f_s(x)) \\ &\text{sujeito a } x \in Q = \bigcap_{i=1}^{p+1} Q_i \end{aligned} \tag{6.4}$$

onde  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}^s$ ,  $Q_i \subset \mathbb{R}^n$ ,  $i = 1, \dots, p$  são conjuntos convexos com interior não vazio, e  $Q_{p+1} \subset \mathbb{R}^n$  conjunto convexo que possivelmente não possui pontos interiores.

O próximo teorema nos fornece condições necessárias e suficientes para otimalidade global de Pareto do problema multiobjetivo (MOP). Assemelha-se ao Teorema 2.24 para o caso escalar, exceto que aqui nós requeremos que as funções devem ser não somente convexas, mas também Ponstein-convexas.

**Teorema 6.23** *Sejam  $f_i : X \rightarrow \overline{\mathbb{R}}$ , funções convexas próprias, contínuas e Ponstein convexas com  $x_0 \in \text{dom} f_i$  e  $\inf f_i < f_i(x_0)$  para todo  $i = 1, \dots, s$ . Suponha que os conjuntos de restrições  $Q_i$  são convexos para todo  $i = 1, \dots, p+1$  e  $Q_{p+1} \cap (\bigcap_{i=1}^p \text{int} Q_i) \neq \emptyset$ .  $x_0$  é Pareto ótimo local para o problema (MOP) se, e somente se, as seguintes s equações são satisfeitas*

$$h_i + \sum_{\substack{j=1 \\ j \neq i}}^s h_j^{(i)} + \sum_{k=1}^{p+1} l_k^{(i)} = 0, \quad i = 1, \dots, s;$$

onde  $h_i \in K_i^*$ ,  $h_j^{(i)} \in L_j^*$  para  $j = 1, \dots, s$ ,  $j \neq i$  e  $l_k^{(i)} \in M_k^*$  para  $k = 1, \dots, p+1$  são, para todo  $i = 1, \dots, s$ , funcionais lineares, não todos identicamente nulos.

*Demonstração.*

- **Necessidade** A convexidade das funções  $f_i$  assegura a convexidade de  $S_{f_i}^{\leq}(x_0)$  e  $S_{f_i}^{\leq}(x_0)$ . Portanto  $K_i = F_3(S_{f_i}^{\leq}(x_0); x_0)$  e  $L_i = F_3(S_{f_i}^{\leq}(x_0); x_0)$  são cones convexos pelo Teorema 2.16. O mesmo Teorema implica a convexidade de  $M_i = F_3(Q_i; x_0)$  para todo  $i = 1, \dots, p$ . O cone  $M_{p+1}$  é convexo porque o conjunto  $Q_{p+1}$  é convexo, pelo Teorema 2.18. Assim todas as hipóteses do Teorema 6.3 são satisfeitas e a condição segue.
- **Suficiência** Suponha que  $x_0$  não é ótimo de Pareto, então existe  $x_1 \in Q$ ,  $x_1 \neq x_0$  tal que  $f_i(x_1) \leq f_i(x_0)$  para todo  $i = 1, \dots, s$  com pelo menos uma desigualdade estrita. Por hipótese existe  $\bar{x} \in Q_{p+1} \cap (\bigcap_{i=1}^p \text{int} Q_i)$ . Defina, para  $0 < \lambda < 1$ ,  $x_\lambda = \lambda \bar{x} + (1 - \lambda)x_1$ , então a convexidade implica que  $x_\lambda \in Q_i$  para  $i = 1, \dots, p+1$ . Além disso, pela Proposição 1.22 (a),  $x_\lambda \in \text{int} Q_i$  para  $i = 1, \dots, p$ . Como  $f_i$  é contínua para todo  $i$ , podemos tomar  $\lambda \in (0, 1)$  suficientemente pequeno de modo que  $f_i(x_\lambda) \leq f_i(x_0)$  para  $i = 1, \dots, s$  com desigualdade estrita na mesma componente como em  $f_i(x_1) \leq f_i(x_0)$ . Defina  $h = x_\lambda - x_0$ . Por definição  $h \neq 0$ , vamos mostrar que:

- (i)  $h \in M_i$  para todo  $i = 1, \dots, p$ ;
- (ii)  $h \in M_{p+1}$ ;
- (iii)  $h \in K_j$  para pelo menos um  $j$ ,  $1 \leq j \leq s$ ;
- (iv)  $h \in L_i$  para todo  $i = 1, \dots, s$ ,  $i \neq j$ .

Dessas quatro afirmações temos que  $K_j \cap (\bigcap_{i=1, i \neq j}^s L_i) \cap (\bigcap_{k=1}^{p+1} M_k) \neq \emptyset$  para pelo menos um  $j$  o que contradiz o conjunto de  $s$  equações mencionadas no Teorema (veja Lema 2.8).

- (i) Como  $x_0 \in Q_i$  e  $x_\lambda \in \text{int} Q_i$ , para todo  $i = 1, \dots, p$ , pela Proposição 1.22 (a) temos

$$x_0 + \varepsilon h = x_0 + \varepsilon(x_\lambda - x_0) = (1 - \varepsilon)x_0 + \varepsilon x_\lambda \in \text{int} Q_i$$

para todo  $0 < \varepsilon < 1$ . Mas  $x_0 + \varepsilon h \in \text{int}Q_i$  para todo  $0 < \varepsilon < 1$  significa que  $h \in M_i$  para todo  $i = 1, \dots, p$ .

- (ii) Para todo  $0 < \varepsilon < 1$ ,  $x_0 + \varepsilon h \in Q_{p+1}$ , pois  $x_0$  e  $x_\lambda$  pertencem a  $Q_{p+1}$  e assim  $h \in M_{p+1}$ .
- (iii) Seja  $j$  tal que  $f_j(x_\lambda) < f_j(x_0)$ . Para  $0 < \varepsilon < 1$ , temos da convexidade de  $f_j$  que

$$f_j(x_0 + \varepsilon h) = f_j((1 - \varepsilon)x_0 + \varepsilon x_\lambda) \leq (1 - \varepsilon)f_j(x_0) + \varepsilon f_j(x_\lambda).$$

Então a derivada direcional de  $f_j$  em  $x_0$  na direção de  $h$  é

$$\begin{aligned} f'_j(x_0, h) &= \lim_{\varepsilon \rightarrow 0^+} \frac{f_j(x_0 + \varepsilon h) - f_j(x_0)}{\varepsilon} \\ &\leq \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon[f_j(x_\lambda) - f_j(x_0)]}{\varepsilon} \\ &= f_j(x_\lambda) - f_j(x_0) < 0. \end{aligned}$$

Assim pelo Teorema 2.12  $h \in K_j$ .

- (iv) Para  $i = 1, \dots, s$ ,  $i \neq j$ ,  $f_i(x_\lambda) \leq f_i(x_0)$ . Se a desigualdade é estrita então  $h \in K_i \subset L_i$ . Agora se  $f_i(x_\lambda) = f_i(x_0)$  então para todo  $0 < \varepsilon < 1$  temos, pela Ponstein convexidade de  $f_i$  que

$$f_i(x_0 + \varepsilon h) = f_i(\varepsilon x_\lambda + (1 - \varepsilon)x_0) < f_i(x_0).$$

Assim  $x_0 + \varepsilon h \in S_{f_i}^<(x_0) = \text{int}S_{f_i}^<(x_0) \neq \emptyset$  para todo  $0 < \varepsilon < 1$ , pelo Teorema 6.22, e portanto  $h \in D_3(S_{f_i}^<(x_0); x_0) = L_i$ .

□

## 6.4 Escalarização de Problemas de Otimização Multiobjetivo

O objetivo é encontrar relações de inclusão entre o conjunto de soluções de Pareto para o problema multiobjetivo e o conjunto de soluções ótimas para a família de problemas escalares. A abordagem aqui é utilizar a teoria desenvolvida por Dubovitskii e Milyutin para problemas escalares e a caracterização do problema multiobjetivo de acordo com o Teorema 6.3 e então utilizar esses teoremas de caracterização para obter relações de inclusão entre os conjuntos de soluções.

Considere o seguinte problema multiobjetivo

$$\begin{aligned} &\text{Minimizar } f(x) = (f_1(x), \dots, f_s(x)) \\ &\text{sujeito a } x \in Q = \bigcap_{i=1}^{n+1} Q_i \end{aligned} \tag{MOP}$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}^s$ ,  $Q_i \subset \mathbb{R}^n$ ,  $i = 1, \dots, p$  são conjuntos com interior não vazio, e  $Q_{p+1} \subset \mathbb{R}^n$  não possui pontos interiores.

Este problema se relaciona com a seguinte família de problemas escalares

$$\begin{aligned} &\text{Minimizar } \sum_{i=1}^s \lambda_i f_i(x) \\ &\text{sujeito a } x \in Q = \bigcap_{i=1}^{p+1} Q_i \end{aligned} \tag{P_\lambda}$$

onde  $\lambda = (\lambda_1, \dots, \lambda_s) \in \mathbb{R}^s$  é um vetor parâmetro.

Defina os seguintes conjuntos de parâmetros:

$$\Lambda = \{\lambda \in \mathbb{R}^s : \sum_{i=1}^s \lambda_i = 1, \lambda_i \geq 0, \forall i = 1, \dots, s\},$$

$$\Lambda^+ = \{\lambda \in \mathbb{R}^s : \sum_{i=1}^s \lambda_i = 1, \lambda_i > 0, \forall i = 1, \dots, s\}.$$

Defina também os seguintes conjuntos de soluções:

$$M = \{x \in \mathbb{R}^n : x \text{ é ótimo de Pareto para (MOP)}\},$$

$$E = \{x \in \mathbb{R}^n : x \text{ resolve (P}_\lambda\text{) para algum } \lambda \in \Lambda^+\}.$$

Os conjuntos  $\Lambda$  e  $\Lambda^+$  estão relacionados através do seguinte lema.

**Lema 6.24**  $\text{cl}\Lambda^+ = \Lambda$ .

O próximo teorema especifica uma relação de inclusão entre o cone dual e direções de descida de uma função objetivo de  $(P_\lambda)$  e a soma dos cones duais dos cones de direções de descida dos funcionais objetivos de (MOP). A demonstração utiliza a caracterização desses cones duais e propriedades do subdiferencial.

**Teorema 6.25** *Sejam  $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  funções convexas próprias e fechadas para  $i = 1, \dots, s$ . Assuma também que para todo  $i$ ,  $x_0 \in \text{int}(\text{dom} f_i)$  e  $S_{f_i}^<(x_0) \neq \emptyset$ ,  $K_i$  e  $K$  denotam o cone de direções de descida de  $f_i$  e  $\sum_{i=1}^s \lambda_i f_i$  em  $x_0$ , respectivamente. Então  $\sum_{i=1}^s K_i^* \subset K^*$  para todo  $\lambda \in \Lambda^+$ .*

*Demonstração.* Primeiramente observe que para todo  $\lambda \in \Lambda^+$

$$\partial[\sum_{i=1}^s \lambda_i f_i](x_0) \supset \sum_{i=1}^s \partial[\lambda_i f_i](x_0) = \sum_{i=1}^s \lambda_i \partial f_i(x_0). \quad (6.5)$$

Agora de acordo com o Teorema 6.18

$$K^* = F_3^*(S_{\sum \lambda_i f_i}^<(x_0); x_0) = \{\theta y : \theta \leq 0, y \in \partial[\sum_{i=1}^s \lambda_i f_i](x_0)\}.$$

Vamos provar a inclusão do teorema mostrando que

$$\sum_{i=1}^s K_i^* \subset \text{cl}L \subset K^*, \text{ onde } L = \{\theta y : \theta \leq 0, y \in \sum_{i=1}^s \lambda_i \partial f_i(x_0), \lambda \in \Lambda^+\}.$$

Seja  $z \in \text{cl}L$ , então existe uma sequência  $\{z_k\}_{k=1}^\infty$  com  $z_k \in L$ ,  $\forall k$  e  $\lim_{k \rightarrow \infty} z_k = z$ . Pela definição de  $L$  temos que  $z_k = \theta_k y_k$ , onde  $\theta_k \leq 0$  e  $y_k = \sum_{i=1}^s \lambda_i^{(k)} y_i^{(k)}$ , com  $y_i^{(k)} \in \partial f_i(x_0)$  e  $\lambda^{(k)} \in \Lambda^+$ , para todo  $k$ .

De (6.5) temos que  $L \subset K^*$  e assim  $z_k \in K^*$  para todo  $k$ . Como o cone dual  $K^*$  é sempre fechado, pelo Teorema 1.10(a), então  $\lim_{k \rightarrow \infty} z_k = z \in K^*$ . Portanto  $\text{cl}L \subset K^*$ . Para provar que  $\sum_{i=1}^s K_i^* \subset \text{cl}L$  lembre que

$$\sum_{i=1}^s K_i^* = \sum_{i=1}^s D_3^*(S_{f_i}^<(x_0); x_0).$$

Assim se  $z \in \sum_{i=1}^s K_i^*$  então  $z = \sum_{i=1}^s z_i$  tal que  $z_i \in F_3^*(S_{f_i}^<(x_0); x_0)$ .

Do Teorema 6.18 temos que  $z_i = \theta_i y_i$ ,  $\theta_i \leq 0$  e  $y_i \in \partial f_i(x_0)$ , que assegura que  $z = \sum_{i=1}^s \theta_i y_i$ , com  $\theta_i \leq 0$  e  $y_i \in \partial f_i(x_0)$ . Defina  $\theta = \sum_{i=1}^s \theta_i$  e  $\hat{\lambda}_i = \frac{\theta_i}{\theta}$ . Assuma sem perda de generalidade que  $z \neq 0$  (para  $z = 0$  a inclusão é claramente válida).  $\theta = 0$  implicaria que  $\theta_i = 0$ , para todo  $i$ , ou seja  $z = 0$ , logo  $\theta < 0$ , assim  $\hat{\lambda}_i \geq 0$  para todo  $1 \leq i \leq s$  e  $\sum_{i=1}^s \hat{\lambda}_i = 1$ . Portanto  $z = \sum_{i=1}^s \theta_i y_i = \sum_{i=1}^s \hat{\lambda}_i \theta y_i = \theta \sum_{i=1}^s \hat{\lambda}_i y_i$ , isto é  $z = \theta y$ , com  $\theta < 0$  e  $y = \sum_{i=1}^s \hat{\lambda}_i y_i$ ,  $\hat{\lambda} \in \Lambda = \text{cl}\Lambda^+$ .

Agora vamos provar que existe uma seqüência  $\{v_k\}_{k=1}^\infty$  de elementos de  $L$  que converge para  $z$ .

Temos que  $\partial f_i(x_0)$  é fechado. Então para todo  $y_i \in \partial f_i(x_0)$  existe uma seqüência  $\{y_i^{(k)}\}_{k=1}^\infty$  de elementos de  $\partial f_i(x_0)$  tal que  $\lim_{k \rightarrow \infty} y_i^{(k)} = y_i$ . Agora escolha uma seqüência real  $\{\alpha_k\}_{k=1}^\infty$  convergindo para  $\theta$  com  $\alpha_k \leq 0$  para todo  $k$ . Além disso, tome uma seqüência  $\{\hat{\lambda}^{(k)}\}_{k=1}^\infty$  de elementos de  $\Lambda^+$  convergindo para  $\lambda \in \text{cl}\Lambda^+$ . Agora defina  $v_k = \alpha_k \sum_{i=1}^s \hat{\lambda}_i^{(k)} y_i^{(k)}$ . Da construção de  $\{v_k\}_{k=1}^\infty$  é claro que  $v_k \in L$  para todo  $k$  e  $\lim_{k \rightarrow \infty} v_k = \theta \sum_{i=1}^s \hat{\lambda}_i y_i = z$  e portanto  $z \in \text{cl}L$  o que completa a prova.  $\square$

Pelo Teorema a seguir obtemos a seguinte relação entre os conjuntos soluções  $M$  e  $E$ , definidos anteriormente.

**Teorema 6.26** *Sejam  $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  funções convexas próprias e contínuas para todo  $i = 1, \dots, s$ . Assuma que todo  $x_0$  que é ótimo de Pareto do problema (MOP) cumpre  $x_0 \in \text{int}(\text{dom} f_i)$  para todo  $i$  e  $\bigcap_{i=1}^s S_{f_i}^<(x_0) \neq \emptyset$ . Sejam  $Q_i$  conjuntos de restrições como definidos em (MOP) para  $i = 1, \dots, p+1$ , todos convexos e  $Q_{p+1} \cap (\bigcap_{i=1}^s \text{int} Q_i) \neq \emptyset$ . Então  $M \subset E$ .*

*Demonstração.* Tome  $x_0 \in M$ . Utilizando o Teorema 2.9 (a convexidade dos cones de direções segue da convexidade dos conjuntos de restrições e das  $f_i$ 's como explicado na parte da prova de necessidade do Teorema 6.23), juntamente com o Teorema 6.18 obtemos que  $K_i^* = L_i^*$ , uma vez que  $f_i$  é convexa própria e contínua, então é fechada e como por hipótese  $\bigcap_{i=1}^s S_{f_i}^<(x_0) \neq \emptyset$  então  $S_{f_i}^<(x_0) \neq \emptyset$  para todo  $i$  e assim pelo Corolário 6.4 temos

$$\sum_{i=1}^s h_i + \sum_{k=1}^{p+1} l_k = 0,$$

onde  $l_k \in M_k^*$  para  $k = 1, \dots, p+1$  e  $h_i \in K_i^* = L_i^* = \{\theta y_i : \theta \leq 0, y_i \in \partial f_i(x_0)\}$  para  $i = 1, \dots, s$  são funcionais lineares não todos identicamente nulos.

Seja  $\phi = \sum_{i=1}^s h_i$ , vamos primeiramente mostrar que  $\phi \neq 0$ . Pois se  $\phi = 0$  então  $\sum_{k=1}^{p+1} l_k = 0$  e então  $l_k$  deve ser identicamente nula para todo  $k = 1, \dots, p+1$ , pois caso contrário, pelo Lema 2.8 devemos ter  $\bigcap_{k=1}^{p+1} M_k = \emptyset$  contradizendo a hipótese que todos os conjuntos de restrições convexos cumprem  $Q_{p+1} \cap (\bigcap_{i=1}^s \text{int} Q_i) \neq \emptyset$ . Com todos  $l_k$  identicamente nulos devemos admitir que  $h_i$ ,  $i = 1, \dots, s$  não são todos identicamente nulos, mas juntamente com  $\phi = 0$  e novamente pelo Lema 2.8, isso significa que

$$\emptyset = \bigcap_{i=1}^s K_i = \bigcap_{i=1}^s F_3(S_{f_i}^<(x_0); x_0) = F_3\left(\bigcap_{i=1}^s S_{f_i}^<(x_0); x_0\right).$$

Agora pelo Lema 6.12(a) temos que  $\text{int} \bigcap_{i=1}^s S_{f_i}^<(x_0) = \emptyset$  que implica em  $\bigcap_{i=1}^s S_{f_i}^<(x_0) = \emptyset$ , uma vez que  $S_{f_i}^<(x_0)$  são conjuntos abertos e isto contradiz a hipótese, assim a possibilidade

de  $\phi = 0$  é descartada.

Utilizando o Teorema 6.25 concluímos que  $\phi \in K^*$ , o cone dual do cone de direções de descida de  $\sum_{i=1}^s \lambda_i f_i$  em  $x_0$ ,  $\lambda \in \Lambda^+$ . E o Teorema 2.24 garante que  $x_0 \in E$ .  $\square$

Nas hipóteses feitas, toda solução de Pareto é obtida por escalarização. Este resultado também vale para funções invexas não diferenciáveis, no entanto, finito-valoradas. Para maiores detalhes consultar Brandão [25].

A inclusão  $E \subset M$  é válida sem nenhuma hipótese sob os problemas (MOP) e  $(P_\lambda)$ , conforme demonstrado no Teorema 4.5. Portanto, sob as hipóteses do Teorema 6.26, temos  $E = M$ , e o problema de encontrar solução de  $(P)$  reduz-se ao problema de otimização paramétrica de resolver  $(P_\lambda)$  enquanto percorre todo  $\Lambda^+$ .

O próximo teorema mostra que no Teorema 6.25 a inclusão é realmente uma igualdade, porque a observação anterior não adiciona nenhuma hipótese para escalarização e é dado aqui por uma questão de rigor somente.

**Teorema 6.27** *Nas notações e sob as hipóteses do Teorema 6.25,  $K^* \subset \sum_{i=1}^s K_i^*$  para todo  $\lambda \in \Lambda^+$ .*

*Demonstração.* Como os cones  $K_1, \dots, K_s$  são abertos e convexos e  $\bigcap_{i=1}^s K_i \neq \emptyset$  pelo Lema 1.9,

$$\left( \bigcap_{i=1}^s K_i \right)^* = \sum_{i=1}^s K_i^*.$$

Vamos mostrar que  $\bigcap_{i=1}^s K_i \subset K$  e assim, pelo Lema 1.3,  $K^* \subset (\bigcap_{i=1}^s K_i)^*$ . Para isso usaremos os seguintes fatos que são consequências da definição:

- $F_3(A; x_0) \cap F_3(B; x_0) = F_3(A \cap B; x_0)$ ;
- $\bigcap_{i=1}^s S_{f_i}^<(x_0) \subset S_{\sum \lambda_i f_i}^<(x_0)$ ,  $\lambda \in \Lambda^+$ ;
- $A \subset B \Rightarrow F_3(A; x_0) \subset F_3(B; x_0)$ .

Assim

$$\begin{aligned} \bigcap_{i=1}^s K_i &= \bigcap_{i=1}^s [F_3(S_{f_i}^<(x_0); x_0)] \\ &= F_3(\bigcap_{i=1}^s S_{f_i}^<(x_0); x_0) \\ &\subset F_3(S_{\sum \lambda_i f_i}^<(x_0); x_0) = K, \quad \forall \lambda \in \Lambda^+. \end{aligned}$$

$\square$

## 6.5 Problemas de programação matemática vetorial

Da mesma forma que a teoria original de otimização escalar de Dubovitskii e Milyutin, a teoria apresentada nas seções anteriores pode ser aplicada a uma variedade de problemas de otimização multiobjetivo. Nesta seção o problema de programação não linear multiobjetivo é estudado. O Teorema 6.3, em conjunto com os resultados sobre a construção dos cones de direções de descida, não crescimento, factível e tangente e seus respectivos duais, são aplicados para obter uma generalização do Teorema de Fritz-John para tais problemas.

Considere o seguinte problema de programação não linear multiobjetivo

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{sujeito a: } \begin{cases} g_j(x) \leq 0, & j = 1, \dots, p \\ h_k(x) = 0, & k = 1, \dots, m \end{cases} \end{aligned} \quad (\text{VOP}') \end{aligned}$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}^s$ ,  $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, \dots, p$  e  $h_k : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $k = 1, \dots, m$ .

**Teorema 6.28** *Sejam  $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  funções próprias convexas, fechadas ou com domínio efetivo relativamente aberto, assuma  $x_0 \in \text{int}(\text{dom} f_i)$  e  $S_{f_i}^<(x_0) \neq \emptyset$  para todo  $i = 1, \dots, s$ . As funções  $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, \dots, p$ , satisfazem as hipóteses anteriores,  $h_k : \mathbb{R}^n \rightarrow \mathbb{R}$ , são estritamente diferenciáveis em uma vizinhança de  $x_0$  para todo  $k = 1, \dots, m$ . Se  $x_0$  é ótimo local de Pareto para o problema (VOP') então existem  $\lambda \in \mathbb{R}^s$ ,  $\mu \in \mathbb{R}^p$  e  $\alpha \in \mathbb{R}^m$  com  $(\lambda, \mu, \alpha) \neq 0$ ,  $\lambda_i \geq 0$  para  $i = 1, \dots, s$  e  $\mu \geq 0$  para  $j = 1, \dots, p$ , tais que*

$$\sum_{i=1}^s \lambda_i y_i + \sum_{j=1}^p \mu_j z_j + \sum_{k=1}^m \alpha_k \nabla h_k(x_0) = 0 \quad (6.6)$$

$$\sum_{j=1}^p \mu_j g_j(x_0) = 0 \quad (6.7)$$

onde,  $y_i \in \partial f_i(x_0)$  para  $i = 1, \dots, s$  e  $z_j \in \partial g_j(x_0)$  para  $j = 1, \dots, p$ .

*Demonstração.* Se os gradientes  $\{\nabla h_k(x_0)\}_{k=1}^m$  são linearmente dependentes então  $\sum_{k=1}^m \gamma_k \nabla h_k(x_0) = 0$  com  $\sum_{k=1}^m \gamma_k^2 \neq 0$  e (6.6) e (6.7) valem com a escolha de  $\lambda = 0$ ,  $\mu = 0$  e  $\alpha_k = \gamma_k$  para  $k = 1, \dots, m$ .

Portanto assumamos que  $\{\nabla h_k(x_0)\}_{k=1}^m$  são linearmente independentes. Neste caso o cone dual do cone de direções tangentes de  $Q_{p+1}$  no ponto  $x_0$  é  $M_{p+1} = \{\sum_{k=1}^m \alpha_k \nabla h_k(x_0) : p \in \mathbb{R}^m\}$  (conforme obtido na demonstração do Teorema 3.1).

Para os funcionais objetivo  $\{f_i(x)\}_{i=1}^s$ , de acordo com o Teorema 6.18 temos

$$K_i^* = L_i^* = \{\lambda_i y_i : \lambda_i \leq 0, y_i \in \partial f_i(x_0)\}, \text{ para todo } i = 1, \dots, s.$$

Para as restrições de desigualdade, se  $j$  é tal que  $g_j(x_0) < 0$ , então  $M_j = \mathbb{R}^n$  e  $M_j^* = \emptyset$ . Agora se  $j$  é tal que  $g_j(x_0) = 0$  então  $M_j^* = F_3^*(Q_j; x_0) = \{\mu_j z_j : \mu_j \leq 0, z_j \in \partial g_j(x_0)\}$ , pelo Teorema 6.18.

Aplicando o Teorema 6.3 e o Corolário 6.4, segue o resultado.  $\square$

# Conclusão

Nesta Dissertação, estudamos o formalismo de Dubovitskii-Milyutin, o qual é uma ferramenta poderosa para tratar vários problemas de otimização que aparecem em diferentes áreas. Em particular, estudamos como esta teoria pode ser utilizada no tratamento de problemas de programação não linear e de controle ótimo, para problema com um objetivo e para problemas de programação não linear com vários objetivos.

Nesta teoria, o conhecimento do cone tangente e de seu dual (obtidos via Teorema de Lyusternik) é fundamental. O prosseguimento natural deste trabalho é caracterizar este cone quando as hipóteses do Teorema de Lyusternik não se aplicam. Por exemplo, relaxando-se as hipóteses de diferenciabilidade ou quando o operador  $P'(x_0)$  não for sobrejetor.

Outra possibilidade é utilizar este formalismo a outros problemas de otimização como Problemas de Controle Ótimo envolvendo equações em derivadas parciais, multiobjetivos ou ainda em Teoria de Jogos.

Estudamos as condições necessárias e suficientes de otimalidade para Problemas Multiobjetivo, analisando os conceitos de solução de Pareto e Pareto fraca, espera-se estender essa teoria para analisar os Problemas Multiobjetivo através do conceito de solução propriamente eficiente.

# Referências Bibliográficas

- [1] V. M. Alekseev, V. M. Tikhomirov e S. V. Fomin. *Optimal Control*. Springer, Moscow, 1987.
- [2] D.P. Bertsekas, A. Nedic, e A.E. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, Belmont, Massachusetts, 2003.
- [3] H. Brezis. *Analyse Fonctionnelle: Théorie et Applications*. Masson, Paris, 1987.
- [4] Y. Censor. *Pareto Optimality in Multiobjective Problems*. Applied Mathematics and Optimization. 4, 1977, pg 41-59
- [5] V. Chankong e Y. Y. Haimes. *Multiobjective Decision Making Theory and Methodology*. North-Holland, Amsterdam, 1983.
- [6] R.E. Edwards. *Functional Analysis: Theory and Applications*. Dover, New York, 1965.
- [7] R. G. Eustáquio. *Condições de Otimalidade e de Qualificação para Problemas de Programação Não-Linear*. Dissertação de Mestrado, Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Curitiba, 2007.
- [8] I.V. Girsanov. *Lectures on Mathematical Theory of Extremum Problems*. Springer, New York, 1972.
- [9] F.J. Gould e J.W. Tolle. *A necessary and sufficient qualification for constrained optimization*. SIAM J. Appl. Math. 20, 1971, pg 164-172.
- [10] M. A. Hanson e B. Mond. *Further Generalizations of Convexity in Mathematical Programming*. Journal of Information and Optimization Sciences, vol.3, 1982, pg 25-32.
- [11] B. Hernández-Jiménez, M. Rojas-Medar, R. Osuna-Gómez and A. Beato-Moreno. *Generalized convexity in non-regular programming problems with inequality-type constraints*. Vol 352, no. 2, 2009, pg 604-613.
- [12] A.D. Ioffe e V.M. Tihomirov. *Theory of Extremal Problems*. North-Hollan, Amsterdam, 1979.
- [13] C. Isoton. *Condições necessárias e suficientes de otimalidade para problemas com um e com vários objetivos*. Dissertação de Mestrado, Programa de Pós-Graduação em Matemática Aplicada, Curitiba, 2013.
- [14] E. W. Karas e A. A. Ribeiro. *Um curso de Otimização*, Cengage Learning, Brasil, 2013.

- [15] A.N. Kolmogorov e S.V. Fomin. *Elementos de la Teoria de Funciones y del Analisis Funcional*. MIR, Moscou, 1975.
- [16] W. Kotarski. *On Some Specification of the Dubovitskii-Milyutin Theorem for Pareto Optimal Problems*. *Nonlinear Analysis: Theory, Methods and Applications*, vol 14, no. 3, 1990, pg 287-291.
- [17] P.J. Laurent. *Approximation et Optimisation*. Hermann, Paris, 1972.
- [18] U. Ledzewicz-Kowalewska. *On Some Specification of the Dubovitskii-Milyutin Method*. *Nonlinear Analysis: Theory, Methods and Applications*, vol 10, no.12, 1986, pg. 1367-1371.
- [19] U. Ledzewicz-Kowalewska. *Application of Some Specification of the Dubovitskii-Milyutin Method to Problems of Optimal Control*. *Nonlinear Analysis: Theory, Methods and Applications*, vol 12, no. 2, 1988, pg. 101-108.
- [20] L.A.J. Medeiros e E.A. Mello. *A Integral de Lebesgue*. Instituto de Matemática - UFRJ, Rio de Janeiro, 2008.
- [21] H. Moulin et F. F. Soulié. *La convexité dans les mathématiques de la décision*. Hermann, Paris, 1979.
- [22] J. Ponstein. *Seven Kinds of Convexity*. *SIAM J. Optim*, vol 9, no. 1, 1967, pg 115-119.
- [23] V. Pareto. *Cours d'economie politique*. Rouge, Lausanni, Suíça, 1896.
- [24] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, 1970.
- [25] M.A. Rojas-Medar, G.N. Silva e A.J.V. Brandão. *InveX Nonsmooth Alternative Theorem and Applications*. *Optimization*, 48, 2000, pg 239-253.
- [26] L. B. Santos. *Algumas contribuições em otimização multiobjetivo*. Tese de Doutorado. Departamento de Matemática Aplicada. Universidade Estadual de Campinas. 2004
- [27] E.M. Tena. *Optimización Dinámica*. Prentice Hall, Madrid, 2001.
- [28] V. Vivanco-Orellana. *Problemas de Extremos Regulares y No Regulares, via Formalismo de Dubovitskii-Milyutin. Aplicación a Problemas de Control Óptimo*. Tese de Doutorado. Universidad de Sevilla, Espanha, 2013.
- [29] E. Zeidler. *Nonlinear functional analysis and its applications: Variational methods and Optimization*, vol. 3, Springer, Berlin, 1989.