

**Universidade Estadual de Campinas  
UNICAMP  
Instituto de Matemática, Estatística e Computação Científica  
IMECC**

**CONDIÇÕES NECESSÁRIAS DE  
OTIMALIDADE PARA PROBLEMAS  
EXTREMAIS E SUAS APLICAÇÕES:  
ABORDAGEM VIA O FORMALISMO ANALÍTICO-  
FUNCIONAL DE DUBOVITSKII E MILYUTIN**

**Autora: Eliane Auxiliadora Rezende**

**Orientador: Prof. Dr. Marko Antonio Rojas Medar**

**Mestrado em Matemática Aplicada  
10 de outubro de 1997  
Campinas/SP**

**CONDIÇÕES NECESSÁRIAS DE OTIMALIDADE PARA  
PROBLEMAS EXTREMAIS E SUAS APLICAÇÕES:  
ABORDAGEM VIA O FORMALISMO DE  
DUBOVITSKII E MILYUTIN**

Este exemplar corresponde à redação final da dissertação devidamente corrigida pela Srta. Eliane Auxiliadora Rezende e aprovada pela Comissão Julgadora.

Campinas, 10 de outubro de 1997



Prof. Dr. Marko Antonio Rojas Medar  
Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do Título de MESTRA em Matemática Aplicada.

**FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DO IMECC DA UNICAMP**

Rezende, Eliane Auxiliadora

R339c      Condições necessárias de otimalidade para problemas extremais e suas aplicações: abordagem via o formalismo analítico-funcional de Dubovitskii e Milyutin / Eliane Auxiliadora Rezende -- Campinas, [S.P. :s.n.], 1997.

Orientador : Marko Antonio Rojas Medar

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Otimização matemática. 2. Teoria do controle. 3. programação não-linear. I. Medar, Marko Antonio Rojas. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

# CONDIÇÕES NECESSÁRIAS DE OTIMALIDADE PARA PROBLEMAS EXTREMAIS E SUAS APLICAÇÕES: ABORDAGEM VIA O FORMALISMO DE DUBOVITSKII E MILYUTIN <sup>1</sup>

**Eliane Auxiliadora Rezende <sup>2</sup>**

Instituto de Matemática Estatística e Computação Científica  
UNICAMP

Banca Examinadora:

- Marko Antonio Rojas Medar (*Orientador*) <sup>3</sup>
- Geraldo Nunes Silva <sup>4</sup>
- Maria do Socorro Nogueira Rangel <sup>4</sup>
- José Luiz Boldrini (*Suplente*) <sup>3</sup>

---

<sup>1</sup> Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da UNICAMP, como requisito parcial para a obtenção do Título de Mestre em Matemática Aplicada.

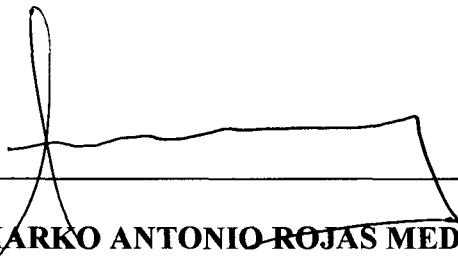
<sup>2</sup> A autora é Licenciada em Matemática pelo departamento de matemática da Universidade Federal de Uberlândia.

<sup>3</sup> Professor Dr. do Instituto de Matemática, Estatística e Computação Científica - UNICAMP.

<sup>4</sup> Professor(a) Dr(a). do Instituto de Matemática - UNESP.

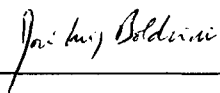
Dissertação de Mestrado defendida e aprovada em 10 de outubro de 1997

pela Banca Examinadora composta pelos Profs. Drs.



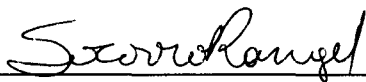
---

Prof (a). Dr (a). **MARKO ANTONIO ROJAS MEDAR**



---

Prof (a). Dr (a). **JOSÉ LUIZ BOLDRINI**



---

Prof (a). Dr (a). **MARIA DO SOCORRO NOGUEIRA RANGEL**

Dedico àqueles sem os quais  
nada seria possível, meus pais,  
João e Elcione.

Às minhas queridas irmãs,  
Adriana e Rosyane, que são  
minhas melhores amigas e  
companheiras.

Aos meus avós:  
Joviniano **f** e Maria;  
Mario e Benita.

E finalmente, ao  
meu tio Viltamar.

# Agradecimentos

Como não poderia deixar de ser, começo com uma oração de louvor porque acredito que nossas vidas, em particular nosso trabalho, deve ser uma contínua oração.

Obrigada meu Deus pela coragem que me destes para realizar o desejo de ser mestra em matemática aplicada pela UNICAMP.

Obrigada Pai porque me dissestes SIM, através dos meus pais, nos momentos de incerteza para seguir este objetivo.

Obrigada Espírito Santo por me instruir em cada exame realizado para que eu conseguisse (e continuasse com) o apoio financeiro da CAPES (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior).

Obrigada Jesus por cada pessoa que passou a fazer parte da minha vida durante todos esses anos aqui em Campinas. Elas foram o sinal eficaz da Tua presença, porque revelavam quem és Tu por meio de: sorrisos, companheirismo, orações, encorajamento, palavras de consolo e de esperança, pedagogia na transmissão de conteúdos, hospedagem, apoio financeiro, enfim, **amizade**. Jesus, quão grande é a minha satisfação por causa dos professores Doutores Maria do Socorro Nogueira Rangel, José Luiz Boldrini e Geraldo Nunes Silva que me ajudaram com suas críticas construtivas e sugestões de mudanças neste trabalho, e também foram escolhidas por Ti para comporem a Banca Examinadora.

Obrigada Virgem Maria, mãe Auxiliadora, por cada momento em que me colocavas em teu colo de Mãe não me deixando desistir nas circunstâncias de medo, ansiedade ou saudade.

Obrigada papai, mamãe, irmãs, avôs e avós, e tios porque sempre intercederam mediante orações pela concretização desta dissertação. Agradeço também pelo exemplo de confiança no Deus de Misericórdia e Libertação.

Obrigada professor Dr. Marko Antonio Rojas Medar pela sua orientação, visto que desde o momento da apresentação deste projeto de tese, confiou em minhas potencialidades e também me ensinou a ter mais paciência comigo mesma por meio do seu próprio exemplo. Realmente, a paciência tudo alcança. Foi através de sua prestatividade que tomei conhecimento de boa parte da teoria aqui encerrada, pois em minha formação acadêmica não tive o prazer de vislumbrar esta face da matemática, cuja aplicação é vasta, mas ainda não totalmente desvendada.

Obrigada àqueles que foram meus companheiros de república e aos amigos do grupo de oração, Charolais e Nossa Senhora Auxiliadora, os quais me ensinaram um pouco da arte de viver, de conviver e de aprender.


Obrigada brasileiros que pagam impostos e, sem saber, contribuem de forma significativa para o progresso tecnológico científico e cultural do país.

Obrigada Deus Pai pela felicidade de ter recebido de Vós a graça de sorrir e de ser humilde nas várias vezes em que recomecei este trabalho por motivos computacionais ou mudanças em sua estrutura.

Senhor Vos louvo pelas inumeráveis obras que fizestes em mim e em minha vida durante o tempo de mestrandia. Aqui não posso anunciá-las e divulgá-las, entretanto reconheço a realização destas suas promessas:

- “ ‘Vou te ensinar, dizeis, vou te mostrar o caminho que deves seguir; vou te instruir, fitando em ti os meus olhos’ .”  
(Salmo 31,8, **Bíblia Sagrada**. 1994, p. 679).
- “ (...) agora o Todo-poderoso lhes concederia a vitória.”  
(II Macabeus 15,8, **Bíblia Sagrada**. 1994, p. 611).





**“A nossa posição  
perante o mundo  
em todas as suas  
expressões,  
dentro do  
caminho cheio  
de desejo rumo  
à totalidade,  
atravessando  
passo a passo  
as situações  
concretas,  
os  
pormenores  
em que a  
existência  
se desenvolve,  
é a de avaliar  
todas as coisas  
na presença  
de Cristo.”**

(GIUSSANI.1996. p. 94).

**“Devemos ao menos ser tão sinceros  
a ponto de colocarmo-nos sempre diante  
desta alternativa: a Presença e a  
companhia de Deus ou é uma  
fábula ou é a verdade da vida,  
a verdade do homem.”**

(GIUSSANI.1996. p. 94).

## Resumo

Neste trabalho apresentamos uma teoria geral devida a Dubovitskii e Milyutin para obtermos condições necessárias de otimalidade em vários problemas de *Otimização*. As ferramentas essenciais usadas são: Análise funcional e Análise convexa. Aplicamos os resultados gerais para obtermos as condições clássicas de Fritz-John e Karush-Kuhn-Tucker da *programação matemática* e as condições necessárias de otimalidade do Princípio do Máximo de Pontryagin da *teoria de controle*. Em seguida fazemos uma aplicação ao clássico problema do Pouso Brando.

# Lista de Figuras

2.1	Cone $K$ e seu dual $K^*$ em $\mathbb{R}^2$ . . . . .	9
2.2	Interpretação geométrica da proposição 2.1.12. . . . .	14
2.3	Direção factível $h$ . . . . .	17
2.4	$h$ é direção tangente unilateral a $M$ em $y^\circ$ . . . . .	18
2.5	Cones, em $\mathbb{R}^2$ , de direções tangentes a $M = \{(y_1, y_2) : y_1 \geq 0\}$ . . . . .	20
3.1	Cone de direção factível $K_a$ . . . . .	36
4.1	$y^\circ$ é ponto interior ao conjunto das restrições $G_1$ , $G_2$ e $G_3$ . . . . .	46
4.2	$y^\circ$ pertence a fronteira da restrição $G_1$ . . . . .	46
4.3	$y^\circ$ pertence a interseção da fronteira da restrição $G_1$ com $G_4$ . . . . .	47
4.4	$y^\circ$ pertence a restrição de igualdade . . . . .	48
5.1	Forças atuando sobre o veículo . . . . .	75

# Índice

<b>1</b>	<b>Introdução</b>	<b>4</b>
1.1	Proposta . . . . .	5
1.2	Organização desta dissertação . . . . .	5
<b>2</b>	<b>Preliminares técnicos</b>	<b>7</b>
2.1	Cones e cones duais . . . . .	7
2.1.1	Cones de direções (descida, factíveis e tangentes) . . . . .	15
2.2	Resultados diversos em análise . . . . .	21
<b>3</b>	<b>Formalismo de Dubovitskii-Milyutin</b>	<b>24</b>
3.1	Condições necessárias de extremo - Equação de Euler-Lagrange	28
3.2	Cálculo de cones e seus duais . . . . .	32
3.2.1	Direções de descida . . . . .	32
3.2.2	Direções factíveis . . . . .	34
3.2.3	Direções tangentes . . . . .	37
<b>4</b>	<b>Aplicação a problemas de otimização</b>	<b>41</b>
4.1	Problema de programação não-linear . . . . .	41
4.1.1	Considerações sobre o problema . . . . .	41
4.1.2	Condições necessárias de Fritz-John . . . . .	42
4.1.3	Interpretações geométricas . . . . .	45
4.2	Problema de controle ótimo . . . . .	49
4.2.1	Considerações sobre o problema . . . . .	49
4.2.2	Princípio do Máximo Local de Pontryagin . . . . .	52
<b>5</b>	<b>Aplicação do Princípio do Máximo</b>	<b>70</b>
5.1	Princípio do Máximo de Pontryagin . . . . .	70
5.1.1	Formulação do Problema . . . . .	70

5.1.2	Condições de otimalidade para problemas de controle .	71
5.2	Problema de aterrissagem moderada de um veículo espacial - problema de pouso brando . . . . .	74
<b>6</b>	<b>Novas Perspectivas</b>	<b>85</b>

# Capítulo 1

## Introdução

Em diversos campos, tais como da Física, Biologia, Economia, Administração, Engenharia, etc, surge a necessidade de resolver problemas os quais exigem otimização. Com este fim, as pessoas dedicadas à solucionar os problemas precisam reformular os mesmos via modelos matemáticos. Nesses modelos deseja-se maximizar (ou minimizar) uma função (às vezes não linear), denominada função objetivo, de várias variáveis escolhidas dentro de alguma região admissível - região definida por restrições impostas às variáveis do problema.

Tais problemas matemáticos que exigem calcular o máximo ou mínimo (a palavra ótimo inclui ambos) de alguma função objetivo, usualmente uma função de variáveis vetoriais, sujeita ou não a restrições, são chamados problemas de **programação matemática**.

Outros tipos de problemas são aqueles cuja função a ser otimizada envolve variáveis em função do tempo (*variável de estado* - descreve o estado do sistema em função do tempo, e *variável de controle* - descreve alguma "entrada" para o sistema) sujeitas a algumas restrições. Tais problemas são chamados problemas de controle ótimo.

Um dos resultados mais importantes da Teoria de Controle Ótimo é o **Princípio do Máximo de Pontryagin**, que é um conjunto de condições necessárias de otimalidade para problemas de controle ótimo, e se demonstra usando principalmente instrumentos de análise funcional e análise convexa.

Os problemas matemáticos podem estar em dimensão finita ou infinita.

## 1.1 Proposta

Nesta dissertação iremos apresentar a teoria desenvolvida por Dubovitskii e Milyutin, em 1965, a qual é geral para problemas extremais sujeitos ou não a certas condições. Dela derivam-se as condições necessárias de otimalidade de Fritz-John e Karush-Kuhn-Tucker da programação matemática. Também esta teoria engloba o Princípio do Máximo de Pontryagin dentro do contexto da regra dos multiplicadores de Lagrange. Este Princípio torna-se fundamental no Cálculo de Variações, visto que as condições necessárias de otimalidade:

- a) Equação de Euler;
- b) Condição de Legendre;
- c) Condição de Weierstrass;
- d) Condição de Weierstrass-Erdmann,

são conseqüências deste Princípio do Máximo de Pontryagin.

Para exemplificar, faremos uma aplicação ao clássico problema do Pouso Brando do campo da Física.

Os instrumentos matemáticos que iremos utilizar na apresentação do formalismo de Dubovitskii e Milyutin serão a Análise funcional e Análise convexa.

## 1.2 Organização desta dissertação

No capítulo 2, foram coletados alguns resultados, os quais serão instrumentos essenciais no decorrer do trabalho, de modo a facilitar a exposição das idéias de Dubovitskii e Milyutin, formalizadas com maiores detalhes na seqüência. No capítulo 4, particularizamos as condições necessárias de otimalidade obtidas no capítulo 3, ao apresentarmos as condições de Fritz-John para problemas de programação não-linear em  $\mathbb{R}^n$  e quando também aplicamos esse mesmo formalismo a um problema específico de controle ótimo,

de onde se chega à condição necessária referida como Princípio do Máximo Local de Pontryagin. No capítulo 5, expomos um problema da Física - Problema de Pouso Brando - cuja análise do ótimo é feita via as condições de extremo de Pontryagin. Encerramos, no capítulo 6, comentando a respeito de trabalhos referentes a problemas de otimização, diferentes daqueles apresentados aqui; bem como citamos questões ainda não resolvidas.

Fazemos uma breve menção de que as figuras, inerentes neste trabalho, são apenas esboços das argumentações explícitas no momento em questão. Algumas notas de rodapé são apenas para definir notações para aquele leitor não envolvido diretamente na área de análise funcional, ou são algumas informações a respeito de alguns matemáticos. Deixamos referências bibliográficas para aquelas demonstrações que poderiam fugir do nosso objetivo porque são extensas ou porque envolvem grande parte de resultados no contexto em que se encontram. Lembramos que foi utilizado o termo operador no sentido de transformação e ainda mais, para intervalo aberto fizemos uso da simbologia  $]a, b[$  ou  $(a, b)$  quando não possibilitasse qualquer tipo de confusão. Algumas vezes aparecerá a abreviação i.e. significando isto é. Fizemos uso da nomenclatura forma linear equivalendo a funcional linear. Na bibliografia constam-se referências as quais ajudariam tanto ao leitor não matemático, como àquele que deseja aperfeiçoar seus conhecimentos no âmbito da matemática exigida aqui. Ordenamos a bibliografia colocando o sobrenome dos autores em ordem alfabética.



# Capítulo 2

## Preliminares técnicos

Neste capítulo, para aquele leitor não familiarizado com assuntos de análise funcional e convexa, apresentamos algumas definições e teoremas os quais usaremos durante o desenvolvimento da dissertação. Para aqueles interessados em aprofundar no assunto sugerimos consultar Girsanov [22], Zeidler [40], Deimling [17], Edwards [19], Brezis [13], Kolmogorov [28], Rudin [35] e Yosida [39, 1978].

### 2.1 Cones e cones duais

Nesta dissertação,  $Y$  designará um espaço de Banach <sup>1</sup> e  $y^o$  um ponto de  $Y$ . Quando isto não ocorrer deixaremos explícito no texto.

**Definição 2.1.1** *Um subconjunto  $K \subset Y$  é chamado um cone com vértice na origem 0 se, e somente se,*

$$h \in K, \text{ implica } \alpha h \in K, \forall \alpha > 0, \text{ i.e., } \alpha K = K, \forall \alpha > 0.$$

**Definição 2.1.2** *Um cone  $K$  é chamado convexo, se dados  $y_1, y_2 \in K$  então  $y_1 + y_2 \in K$ .*

---

<sup>1</sup> *Stefan Banach* (1892-1945), matemático polonês, é considerado um dos maiores heróis da história do desenvolvimento da matemática.

Definimos um espaço de Banach como um espaço normado o qual é completo na métrica provida da sua norma; isto significa que cada seqüência de Cauchy é convergente.

**Exemplo 2.1** Um exemplo de cone convexo em  $\mathbb{R}^n$  é o octante não-negativo,

$$\mathbb{R}_+^n = \{y = (y_1, \dots, y_n) : y_i \geq 0, i = 1, \dots, n\}.$$

Abaixo, introduzimos o importante conceito de cone dual:

**Definição 2.1.3** O cone dual  $K^*$  do cone  $K$  é definido por

$$K^* = \{f \in Y^* : f(y) \geq 0, \forall y \in K\}^2.$$

**Observação 2.1.4** De acordo com a definição acima,  $K^*$  consiste de todos os funcionais lineares contínuos não-negativos em  $K$ . Se  $K = \emptyset$ , então  $K^* = Y^*$ , pois senão, deveria existir algum funcional  $\tilde{f}$  e algum  $\tilde{y} \in K$  tais que  $\tilde{f}(\tilde{y}) < 0$ , o que é absurdo. Já o caso em que  $K = Y$ , temos que  $K^* = \{0\} \neq Y^*$ , como veremos pelo teorema 2.1.5.

A notação  $K^*$  para o cone dual não parece conveniente, de acordo com a observação acima, contudo é usual na literatura. Por isso iremos adotá-la sem que haja confusão.

Facilmente verificamos que o cone dual  $K^*$  é um cone convexo com  $0 \in K^*$ . Além disso, se  $K_1 \subseteq K_2$  sempre teremos  $K_2^* \subseteq K_1^*$ .

**Exemplo 2.2** Sejam  $Y = \mathbb{R}^n$  e  $K = \mathbb{R}_+^n$ , ou seja,

$$K = \{(\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n : \xi_1, \xi_2, \dots, \xi_n \geq 0\}.$$

Então  $K$  é um cone convexo fechado em  $Y$  e  $K^* = K$ .

**Exemplo 2.3** Considere o espaço de Banach  $Y = \mathbb{R}^2$  e o cone  $K$  dado por  $K = \{h = (h_1, h_2) : a \cdot h \geq 0 \text{ e } b \cdot h \geq 0, \text{ com } a, b \in \mathbb{R}^2 \text{ fixos}\} \subset \mathbb{R}^2$ . Então  $K^* = \{f \in \mathbb{R}^2 : f = \lambda_1 a + \lambda_2 b, \lambda_1, \lambda_2 \geq 0\}$ . Veja a figura seguinte.

---

<sup>2</sup> $Y^*$  é o espaço dual topológico do espaço de Banach  $Y$ , consistindo de todos os funcionais lineares contínuos sobre  $Y$ .

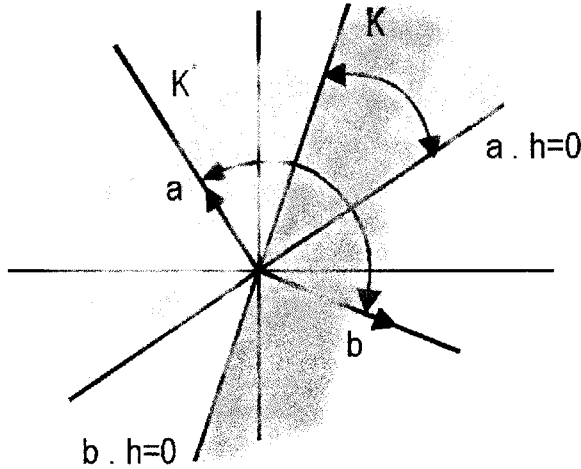


Figura2.1: Cone  $K$  e seu dual  $K^*$  em  $\mathbb{R}^2$ .

Vejamos alguns resultados que nos permitem calcular cones duais de certos conjuntos e que serão fundamentais para nossos futuros argumentos.

**Teorema 2.1.5** *Suponha que o cone  $K$  é um subespaço. Então*

$$K^* = \{f \in Y^* : f(y) = 0, \forall y \in K\}$$

(este conjunto é algumas vezes conhecido como o anulador de  $K$ ).

**Prova:** Provaremos a inclusão  $K^* \subset \{f \in Y^* : f(y) = 0, \forall y \in K\}$ , pois a outra é trivial:

Dado  $f \in K^*$ , suponhamos que exista  $y \in K$ , tal que  $f(y) > 0$ . Como  $K$  é um subespaço,  $-y \in K$ . Então  $f(-y) = -f(y) < 0$ , o que nos leva a uma contradição, pelo fato de que  $f \in K^*$ . Logo  $f(y) = 0, \forall y \in K$ . Portanto,  $K^* = \{f \in Y^* : f(y) = 0, \forall y \in K\}$ . ■

Antes de enunciarmos o teorema 2.1.9 iremos recordar algumas definições.

**Definição 2.1.6** *Seja  $E \subset Y$ . Se  $E$  é fechado na topologia da norma (se é dada uma seqüência  $(x_n) \subseteq E$  tal que  $\|x_n - x\|_Y \xrightarrow{n \rightarrow \infty} 0$ , então  $x \in E$ ), então diz-se que  $E$  é **fortemente fechado**.*

**Definição 2.1.7** Diz-se que  $E \subset Y$  é **fracamente fechado**, se ele for fechado na topologia fraca, i.e., se dada uma seqüência  $(x_n) \subseteq E$ , tal que  $l(x_n) \xrightarrow{n \rightarrow \infty} l(x)$ ,  $\forall l \in Y^*$ , então  $x \in E$ .

**Definição 2.1.8** Dada uma seqüência  $(x_n) \subseteq E \subset Y$ . Se  $l(x_n) \xrightarrow{n \rightarrow \infty} l(x)$ , para todo  $l \in J(x) \subseteq Y^{**}$ , onde  $J(x)$  é isomorfo a  $Y$ , então  $x \in E$ . Diz-se que  $E$  é fechado na topologia fraca\*, de outra forma dito,  $E$  é fracamente\*-fechado (usa-se geralmente a notação  $w^*$ -fechado).

**Teorema 2.1.9** Se os cones  $K_\alpha$  são fracamente fechados e convexos, enquanto  $\sum_{\alpha \in A} K_\alpha^*$  é fraco\*-fechado, então

$$\left( \bigcap_{\alpha \in A} K_\alpha \right)^* = \sum_{\alpha \in A} K_\alpha^*,$$

onde  $\sum_{\alpha \in A} K_\alpha^*$  é o conjunto de todas as somas finitas  $f_{\alpha_1} + \dots + f_{\alpha_n}$ , tal que  $f_{\alpha_i} \in K_{\alpha_i}^*$ ,  $\alpha_i \in A$  sendo  $A$  um conjunto arbitrário de índices.

**Prova:** Para demonstrar este teorema fazemos uso dos resultados abaixo enumerados, os quais envolvem conceito de fecho fraco da envoltória convexa<sup>3</sup> do cone  $K$ ,  $\overline{coK}$ :

- $K^* = \overline{coK}^*$ ;
- $co\left(\bigcup_{\alpha \in A} K_\alpha^*\right) = \sum_{\alpha \in A} K_\alpha^*$ ;
- $\left(\bigcup_{\alpha \in A} K_\alpha\right)^* = \bigcap_{\alpha \in A} K_\alpha^*$ ;

<sup>3</sup>† Dado um espaço vetorial  $X$  e  $E \subset X$ , chama-se *envoltória convexa* de  $E$ , denotada por  $coE$ , à interseção de todos os subconjuntos convexos de  $X$ , os quais contêm  $E$ . Equivalentemente,  $coE$  é o conjunto de todas as combinações convexas finitas de elementos de  $E$ , isto é, o conjunto de todas as somas  $t_1x_1 + t_2x_2 + \dots + t_nx_n$  onde  $x_i \in E$ ,  $t_i \geq 0$ ,  $\sum_{i=1}^n t_i = 1$ , com  $n$  arbitrário.

‡ Se  $X$  é um espaço vetorial topológico e  $E \subset X$ , a *envoltória convexa fechada* de  $E$ , escrita como  $\overline{coE}$ , é a interseção de todos os conjuntos convexos fechados contendo o conjunto  $E$ . Igualmente é definida como o fecho de  $coE$ ,  $\overline{coE}$ .

- $K^{**} = \overline{coK}$ ,

considerando inicialmente  $B = \overline{co\left(\sum_{\alpha \in A} K_{\alpha}^*\right)}$ , onde a barra neste caso quer dizer o fecho fraco\* da envoltória convexa da soma dos cones duais. Os detalhes podem ser encontrados em Girsanov [22, p. 35]. ■

**Teorema 2.1.10** *Sejam  $K_1, \dots, K_n$  cones convexos abertos em  $Y$ , tais que*

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

então

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

**Prova:** Usaremos argumentos geométricos tais como conceitos de produto direto de espaços de Banach e do respectivo espaço dual.

Definamos o espaço  $E = \prod_{i=1}^n Y$ , onde  $Y$  é Banach, isto é, o espaço

$$E = \{\tilde{y} = (y_1, \dots, y_n) : y_i \in Y, i = 1, \dots, n\}.$$

Assim

$$E^* = \{F = (f_1, \dots, f_n) : f_i \in Y^*, i = 1, \dots, n\}.$$

Deste modo, podemos dizer que cada  $F \in E^*$  tem a representação

$$F(\tilde{y}) = \sum_{i=1}^n f_i(y_i), \text{ para todo } \tilde{y} = (y_1, \dots, y_n), \quad (2.1)$$

onde  $f_i \in Y^*$ , para todo  $i = 1, \dots, n$ .

Consideremos ainda, o conjunto

$$K = \{\tilde{y} = (y_1, \dots, y_n) : y_i \in K_i, i = 1, \dots, n\} \subset E.$$

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

Definamos  $L = \{ \tilde{y} = (y, \dots, y) : y \in Y \} \subset E$ , onde  $L$  é um subespaço linear de  $Y$ . Por hipótese, existe  $y^\circ \in K_i$ ,  $i = 1, \dots, n$ , logo  $L \cap K \neq \emptyset$ .

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

$$\bar{f}(\tilde{y}) = f(y), \quad (2.2)$$

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

Se  $\tilde{y}^\circ = (y^\circ, \dots, y^\circ) \in L \cap K$ , então  $y^\circ \in \bigcap_{i=1}^n K_i$ . E assim,  $f(y^\circ) \geq 0$ . Portanto,

$$\bar{f}(\tilde{y}^\circ) \geq 0, \quad \forall \tilde{y}^\circ \in L \cap K.$$

Usando o fato de  $\text{int } K = K$ ,<sup>4</sup> pois  $K$  é aberto, temos que  $L \cap \text{int } K \neq \emptyset$ . Deste modo, podemos aplicar o **Teorema de Krein** enunciado como a seguir: Sejam  $\hat{K}$  um cone convexo com vértice na origem  $0$ , 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

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

Consequimos dessa forma a existência de  $F = (f_1, \dots, f_n) \in E^*$ , tal que

$$F(\bar{y}) \geq 0, \quad \forall \bar{y} \in K; \quad (2.3)$$

$$F(\bar{y}) = \bar{f}(\bar{y}), \quad \forall \bar{y} \in L. \quad (2.4)$$

Assim, para todo  $\tilde{y}^\circ \in L$ , temos por (2.1) que  $F(\tilde{y}^\circ) = \sum_{i=1}^n f_i(y^\circ)$ ; por (2.4)  $F(\tilde{y}^\circ) = \bar{f}(\tilde{y}^\circ)$  e por (2.2)  $\bar{f}(\tilde{y}^\circ) = f(y^\circ)$ . Logo,

$$f(y^\circ) = \sum_{i=1}^n f_i(y^\circ), \quad \forall \tilde{y}^\circ \in L.$$

Então, para  $\bar{y} = (y^\circ, \dots, y^\circ)$ ;  $y^\circ \in \bigcap_{i=1}^n K_i$ , também teremos

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

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

---

<sup>4</sup>Simbolizamos  $\text{int } A$  quando nos referimos ao interior do conjunto  $A$ .

Devido a (2.3) segue que  $F(\tilde{y}^o) \geq 0$ , onde  $F(\tilde{y}^o) = \sum_{i=1}^n f_i(y^o)$ , para todo  $y_i \in K_i$ ,  $i = 1, \dots, n$  e  $\tilde{y} = (y_1, \dots, y_n)$ . Conseqüentemente,  $f_i(y_i) \geq 0$ ,  $\forall y_i \in K_i$ , ou seja,  $f_i \in K_i^*$ ,  $i = 1, \dots, n$ .

Agora de acordo com a análise feita para  $\bar{y}$  onde  $y^o \in \bigcap_{i=1}^n K_i$ , podemos usar o resultado obtido acima e concluirmos que  $f(y^o) = \sum_{i=1}^n f_i(y^o)$ , com  $f_i \in K_i^*$ ,  $\forall i = 1, \dots, n$ . Portanto, dado  $f \in (\bigcap_{i=1}^n K_i)^*$ , existem  $f_i \in K_i^*$ ,  $i = 1, \dots, n$ , tais que  $f = \sum_{i=1}^n f_i$ . Em outras palavras,  $(\bigcap_{i=1}^n K_i)^* \subset \sum_{i=1}^n K_i^*$ . A recíproca é imediata:

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  $y^o \in \bigcap_{i=1}^n K_i$ , vem que  $f_i(y^o) \geq 0$ ,  $\forall i = 1, \dots, n$ . Logo,  $f(y^o) \geq 0$ ,  $\forall y^o \in \bigcap_{i=1}^n K_i$ . Assim,  $f \in (\bigcap_{i=1}^n K_i)^*$ . Por isso,  $\sum_{i=1}^n K_i^* \subset (\bigcap_{i=1}^n K_i)^*$ . ■

Na demonstração do teorema 2.1.13, enunciado posteriormente, utilizaremos uma ferramenta importante em análise funcional, o chamado **Teorema de Hahn-Banach - forma geométrica**, ou **Teorema da separação**:

**Teorema 2.1.11 (Hahn-Banach)** *Sejam  $Y$  um espaço de Banach e  $A, B \subseteq Y$  conjuntos convexos, não-vazios e disjuntos. Suponha que  $A$  é aberto. Então existe um hiperplano fechado  $H$ ,*

$$H = \{y \in Y : f(y) = \alpha, 0 \neq f \in Y^* \text{ e } \alpha \in \mathbb{R}, \text{ fixo}\},$$

que separa  $A$  e  $B$  no sentido amplo, isto é,

$$\begin{cases} f(\bar{y}) \leq \alpha, \forall \bar{y} \in A, \text{ e} \\ f(\bar{y}) \geq \alpha, \forall \bar{y} \in B. \end{cases}$$

Faremos uso também do próximo resultado, cujo significado geométrico é dado logo em seguida.

**Proposição 2.1.12** *Se  $K$  é um cone com vértice em  $y^o$  e  $f$  é um funcional linear, tal que  $f(y) \geq \alpha$ ,  $\forall y \in K$  e  $\alpha \in \mathbb{R}$ . Então,  $f(y) \geq f(y^o)$ ,  $\forall y \in K$ .*

Geometricamente, se o cone  $K$  pertence a um dos lados de um hiperplano definido por  $f(y) = \alpha$ , então ele deve pertencer a um dos lados de um hiperplano paralelo que passa pelo seu vértice  $y^o$  (observe a figura que se segue).

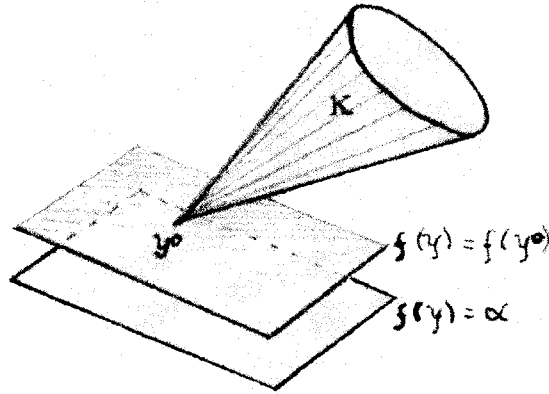


Figura 2.2: Interpretação geométrica da proposição 2.1.12.

**Teorema 2.1.13** *Sejam  $K_0, K_1, \dots, K_n, K_{n+1}$  cones convexos com vértice em 0, onde  $K_0, K_1, \dots, K_n$  são abertos. Então*

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

*se, e somente se, existem funcionais  $f_i \in K_i^*$ ,  $i = 0, \dots, n+1$ , nem todos nulos, tais que*

$$f_0 + f_1 + \dots + f_{n+1} = 0.$$

**Prova:** Começaremos por verificar a **condição necessária**:

Seja  $\tilde{K} = \bigcap_{i=0}^n K_i \neq \emptyset$ . Daí,  $\tilde{K} \cap K_{n+1} = \emptyset$ . Como  $\tilde{K}$  é aberto, podemos aplicar o teorema 2.1.11 de Hahn-Banach e a proposição 2.1.12 acima. Portanto, existe  $f \in Y^*$ ,  $f \neq 0$ , tal que

$$\begin{cases} f(h) \geq 0, & \forall h \in \tilde{K}; \\ f(h) \leq 0, & \forall h \in K_{n+1}. \end{cases}$$

Aplicando o teorema 2.1.10 temos que:

$$\tilde{K}^* = \left( \bigcap_{i=0}^n K_i \right)^* = \sum_{i=0}^n K_i^*.$$



Logo, para  $f \in \tilde{K}^*$ ,

$$f = \sum_{i=0}^n f_i, \text{ com } f_i \in K_i^*, i = 0, \dots, n.$$

Coloquemos  $f_{n+1} = -f \neq 0$ . Então, para  $h \in K_{n+1}$ , obtemos  $f_{n+1}(h) = -f(h) \geq 0$ . Por definição de cone dual,  $f_{n+1} \in \tilde{K}_{n+1}^*$ . Por conseguinte,  $f_0 + f_1 + \dots + f_{n+1} = 0$ .

Agora nos atemos a provar a **condição suficiente**:

Sejam  $f_i \in K_i^*$ ,  $i = 0, \dots, n+1$ , os quais não são simultaneamente nulos, tais que

$$\sum_{i=0}^{n+1} f_i = 0. \quad (2.5)$$

Suponha que exista  $h_0$  de tal maneira que  $h_0 \in \bigcap_{i=0}^{n+1} K_i$ . De (2.5) é impossível ter  $f_0 = \dots = f_n = 0$ , senão  $f_{n+1}$  deveria ser nulo, contradizendo a hipótese. Portanto, suponha  $f_0 \neq 0$  e para algum  $v \in K_0$ ,  $f_0(v) > 0$ . Como o cone  $K_0$  é aberto, qualquer que seja  $\varepsilon$  numa vizinhança de zero,  $h_0 + \varepsilon v \in K_0$  e por isso  $f_0(h_0 + \varepsilon v) \geq 0$ , i.e.,  $f_0(h_0) + \varepsilon f_0(v) \geq 0$ . Assim,  $f_0(h_0) > 0$ . Entretanto, por (2.5), porque  $f_i \in K_i^*$ , chegamos na contradição:

$$0 = f_0(h_0) + \dots + f_{n+1}(h_0) \geq f_0(h_0) > 0.$$

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

### 2.1.1 Cones de direções (descida, factíveis e tangentes)

Apresentaremos definições de vetores diretores, que são importantes nas argumentações do capítulo que se segue.

Seja  $Y$  um espaço de Banach e consideremos  $h$  e  $y^o$  pertencentes a  $Y$ .

**Definição 2.1.14** Diz-se que um vetor  $h$  é uma *direção de descida* do funcional  $F$  no ponto  $y^o$ , se existem uma vizinhança  $V$  do vetor  $h$  e o

número  $\alpha = \alpha(F, y^o, h) < 0$  de tal maneira que dado  $\varepsilon_0 > 0$ ,  $\forall \varepsilon \in ]0, \varepsilon_0[$  e  $\forall \bar{h} \in V$ , tem-se

$$F(y^o + \varepsilon \bar{h}) \leq F(y^o) + \varepsilon \alpha. \quad (2.6)$$

**Lema 2.1.15** *As direções de descida geram um cone aberto  $K$  com vértice na origem  $0$ .*

**Prova:** Seja  $h$  uma direção de descida do funcional  $F$  no ponto  $y^o$ . Então existem uma vizinhança  $V$  do vetor  $h$ ,  $\alpha < 0$  e dado  $\varepsilon_0 > 0$ , tais que para todo  $\varepsilon \in ]0, \varepsilon_0[$  e  $\bar{h} \in V$  temos

$$F(y^o + \varepsilon \bar{h}) \leq F(y^o) + \varepsilon \alpha.$$

Verificaremos se  $\lambda h$ , é uma direção de descida, para  $\lambda > 0$  :

Tomemos como uma vizinhança  $U = \lambda V$ ,  $\tilde{\varepsilon}_0 = \varepsilon_0 / \lambda$  e  $\tilde{\alpha} = \lambda \alpha$ . Então, para todo  $\tilde{\varepsilon} \in ]0, \tilde{\varepsilon}_0[$  e qualquer  $\tilde{h} \in U$  temos que

$$F(y^o + \tilde{\varepsilon} \tilde{h}) = F\left(y^o + \frac{\varepsilon(\lambda \bar{h})}{\lambda}\right) \leq F(y^o) + \varepsilon \alpha = F(y^o) + (\lambda \tilde{\varepsilon}) \frac{\tilde{\alpha}}{\lambda} = F(y^o) + \tilde{\varepsilon} \tilde{\alpha}.$$

De onde obtemos  $\lambda h$  como uma direção de descida. Assim,  $K$  é um cone cujo vértice é  $0$ . Verifiquemos agora se  $K$  é aberto:

De fato, ainda considerando que  $h \in K$  e tomando  $\tilde{h} \in V$ , basta tomarmos o mesmo  $\alpha$  e  $\varepsilon_0$  para obtermos

$$F(y^o + \varepsilon \tilde{h}) \leq F(y^o) + \varepsilon \alpha, \forall \tilde{h} \in V.$$

Portanto,  $\tilde{h} \in K$ . Conseqüentemente  $K$  é aberto. ■

**Definição 2.1.16** *Um funcional  $F(y)$  se diz regular de descida, se o conjunto de suas direções de descida no ponto  $y^o$  é convexo.*

**Definição 2.1.17** *Dado um conjunto  $Q$  com  $\text{int } Q \neq \emptyset$ , geralmente definido por restrições de desigualdades, diz-se que o vetor  $h$  é uma direção factível a  $Q$  no ponto  $y^o$ , se existe uma vizinhança  $V$  do vetor  $h$  e dado  $\varepsilon_0 > 0$ , tais que,  $\forall \varepsilon \in ]0, \varepsilon_0[$  e  $\bar{h} \in V$ , os vetores  $(y^o + \varepsilon \bar{h}) \in Q$  (veja figura).*

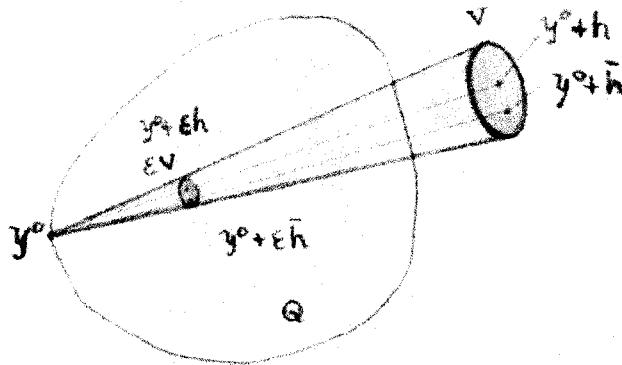


Figura2.3: Direção factível  $h$ .

**Lema 2.1.18** *O conjunto de direções factíveis gera um cone aberto  $K$  com vértice em  $0$ .*

**Prova:** Seja  $K$  o conjunto de direções factíveis para  $Q$  no ponto  $y^o$ . Suponhamos que  $h \in K$ , conforme a definição acima, e tomemos  $U = \lambda V$  e  $\tilde{\epsilon}_0 = \epsilon_0/\lambda$ . Então

$$(y^o + \epsilon \tilde{h}) = (y^o + (\epsilon/\lambda)\lambda \bar{h}) \in Q.$$

Logo,  $\lambda h \in K$ . Portanto  $K$  é um cone com vértice em  $0$ .

Além disso, para qualquer  $\tilde{h} \in V$ , podemos tomar  $\epsilon \in ]0, \epsilon_0[$  e a mesma vizinhança  $V$  tal que  $(y^o + \epsilon \tilde{h}) \in Q$ . Desta forma,  $\tilde{h} \in K$ . O que nos mostra que  $K$  é aberto. ■

**Definição 2.1.19** *Diz-se que uma restrição de desigualdade em  $Q$  é regular no ponto  $y^o$ , se o cone de direções factíveis a  $Q$  em  $y^o$  for convexo.*

Notemos que, se o conjunto  $Q$  fosse dado por restrições de igualdade, isto é, variedade sem pontos interiores, o cone  $K$  de direções factíveis no ponto  $y^o$  seria vazio. Realmente, se existisse  $h \in K$ , então para  $\bar{h} \in V$ ,  $y^o + \epsilon \bar{h} \in Q$ . O que é absurdo, porque  $Q$  não tem pontos interiores.

Por esta razão, introduziremos outros dois conceitos de direção no ponto  $y^o \in Q$ .

**Definição 2.1.20** Seja  $M \subseteq Y$ ,  $M \neq \emptyset$ , dado por restrições de igualdades. O vetor  $h$  é uma **direção tangente** a  $M$  no ponto  $y^o$ , se existem  $\varepsilon_0 > 0$  e uma aplicação  $r : ]-\varepsilon_0, \varepsilon_0[ \rightarrow Y$ , tais que:

- $y_\varepsilon = y^o + \varepsilon h + r(\varepsilon) \in M$ ;
- $\frac{1}{\varepsilon} r(\varepsilon) \in V$ , qualquer vizinhança de zero, para todo  $\varepsilon > 0$  suficientemente pequeno, ou equivalentemente,  $\|r(\varepsilon)\| = o(\varepsilon)$ ,  $\forall \varepsilon \rightarrow 0$ , onde  $o(\varepsilon) = \lim_{\varepsilon \rightarrow 0} \frac{r(\varepsilon)}{\|\varepsilon\|}$ .

**Definição 2.1.21** O vetor  $h \in Y$  é dito **direção tangente unilateral** a  $M$  no ponto  $y^o \in M$ , se existem  $\varepsilon_0 > 0$  e uma aplicação  $r : ]0, \varepsilon_0[ \rightarrow Y$  tais que:

- $y^o + \varepsilon h + r(\varepsilon) \in M$ ;
- $\|r(\varepsilon)\| = o(\varepsilon)$ ,  $\forall \varepsilon \rightarrow 0^+$ .

Abaixo esboçamos uma direção  $h$  unilateral em um ponto arbitrário  $y^o \in M$ :

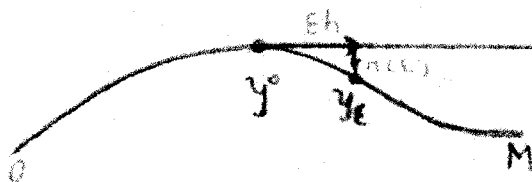


Figura 2.4:  $h$  é direção tangente unilateral a  $M$  em  $y^o$ .

**Observação 2.1.22** • Denotamos o conjunto de todas as direções tangentes a  $M$  no ponto  $y^o$  por  $T_{y^o}M$ , e o conjunto das direções unilaterais por  $T_{y^o}^+M$ .

- É claro que  $T_{y^o}M \subseteq T_{y^o}^+M$  e  $T_{y^o}M = T_{y^o}^+M \cap T_{y^o}^-M$ , onde  $T_{y^o}^-M = -T_{y^o}^+M$  é definido como em 2.1.21, mas com  $\varepsilon_0 < 0$  e  $\varepsilon \rightarrow 0^-$ .

- Se o conjunto  $T_{y^o}M$  é um subespaço vetorial de  $Y$ , então ele se chamará **espaço tangente** a  $M$  no ponto  $y^o$ . Em geometria, usualmente se chama *reta tangente*, *plano tangente*, etc, não a  $T_{y^o}M$  mas a variedade afim  $y^o + T_{y^o}M$ .
- As direções tangentes geram um cone  $K$  com vértice em  $0$  (de acordo com a definição, supondo  $h \in T_{y^o}M$  basta tomar  $\tilde{\varepsilon}_0 = \lambda\varepsilon_0$  e então para  $\lambda > 0$ ,  $\lambda h \in T_{y^o}M$ ). Na maioria dos casos,  $K$  é um semi-espaço.
- Toda direção factível é também uma direção tangente, porém a recíproca é falsa. Desse modo, considerando que existe  $\varepsilon_0 > 0$  tal que  $y^o + \varepsilon\bar{h} \in M$ , com  $\bar{h}$  numa vizinhança  $V$  de  $h$  e  $\varepsilon \in ]0, \varepsilon_0[$ , também será verdade que  $y^o + \varepsilon h \in M$ , onde  $\varepsilon \in ]-\varepsilon_0, \varepsilon_0[$ . Isto implica que  $h$  é uma direção tangente a  $M$  no ponto  $y^o$  admitindo  $r = 0$ .

Como contra exemplo, considere  $Y = \mathbb{R}^n$  e  $M$  um conjunto formado por apenas um ponto arbitrário  $y^o \in \mathbb{R}^n$ . Desta forma,  $T_{y^o}M = \mathbb{R}^n$ . Entretanto, não existem  $\varepsilon_0 > 0$  e uma vizinhança  $V$  de qualquer  $(h_1, \dots, h_n) \in \mathbb{R}^n$  tal que  $y^o + \varepsilon(\bar{h}_1, \dots, \bar{h}_n) \in M$ , para todo  $(\bar{h}_1, \dots, \bar{h}_n) \in V$  e  $\varepsilon \in ]0, \varepsilon_0[$ , isto é, o conjunto de direções factíveis a  $M$  em  $y^o$  é vazio.

**Definição 2.1.23** Diz-se que uma **restrição de igualdade** em  $M$  é **regular** no ponto  $y^o$ , se o cone de direções tangentes a  $M$  em  $y^o$  for convexo.

Neste momento, citaremos alguns exemplos de cones tangentes.

**Exemplo 2.4** Seja  $Y = \mathbb{R}^2$ ,  $M = \{(y_1, y_2) : y_1 \geq 0\}$ . Então  $T_{(0,0)}^+M = M$ ;  $T_{(0,0)}M = \{(0, b) : b \in \mathbb{R}\}$ ;  $T_{(1,0)}M = T_{(1,0)}^+M = \mathbb{R}^2$  (veja adiante o esboço dos cones tangentes  $T_{(0,0)}^+M$ ,  $T_{(0,0)}M$ ,  $T_{(1,0)}M$  e  $T_{(1,0)}^+M$ ).

De fato, tomando  $y^o = (0, 0)$ ,  $h = (h_1, h_2)$ ,  $r(\varepsilon) = (r_1(\varepsilon), r_2(\varepsilon))$  acharemos quais os vetores  $h$  que satisfazem as condições da definição de vetor tangente:

$y^o + \varepsilon h + r(\varepsilon) \in M \implies (0, 0) + \varepsilon(h_1, h_2) + (r_1(\varepsilon), r_2(\varepsilon)) = (\varepsilon h_1 + r_1(\varepsilon), \varepsilon h_2 + r_2(\varepsilon)) \in M$ . Logo  $\varepsilon h_1 + r_1(\varepsilon) \geq 0$ . Quando  $\varepsilon \rightarrow 0^+$ ,  $h_1 \geq -\frac{r_1(\varepsilon)}{\varepsilon}$ . Assim,  $h_1 \geq 0$ . Por isso  $T_{(0,0)}^+M = \{(y_1, y_2) : y_1 \geq 0\}$ . Por outro lado, quando  $\varepsilon \rightarrow 0^-$ ,  $h_1 \leq -\frac{r_1(\varepsilon)}{\varepsilon}$ . Então,  $h_1 \leq 0$  e  $-T_{(0,0)}^+M = \{(y_1, y_2) : y_1 \leq 0\}$ . Portanto,  $T_{(0,0)}M = T_{(0,0)}^+M \cap T_{(0,0)}^-M = \{(0, b) : b \in \mathbb{R}\}$ .

Considerando  $y^o = (1, 0)$ , teremos:

$$(1, 0) + \varepsilon(h_1, h_2) + (r_1(\varepsilon), r_2(\varepsilon)) \in M \Rightarrow \varepsilon h_1 + r_1(\varepsilon) + 1 \geq 0 \Rightarrow \varepsilon h_1 + r_1(\varepsilon) \geq -1.$$

Por conseguinte,

$$\begin{cases} h_1 \geq \frac{-1-r_1(\varepsilon)}{\varepsilon}, & \text{quando } \varepsilon > 0. \text{ De onde segue que } h_1 \geq -\infty, \text{ se } \varepsilon \rightarrow 0^+ \\ h_1 \leq \frac{-1-r_1(\varepsilon)}{\varepsilon}, & \text{quando } \varepsilon < 0. \text{ De onde segue que } h_1 \leq +\infty, \text{ se } \varepsilon \rightarrow 0^-. \end{cases}$$

Conseqüentemente,  $T_{(1,0)}M = T_{(1,0)}^+M = \mathbb{R}^2$ .

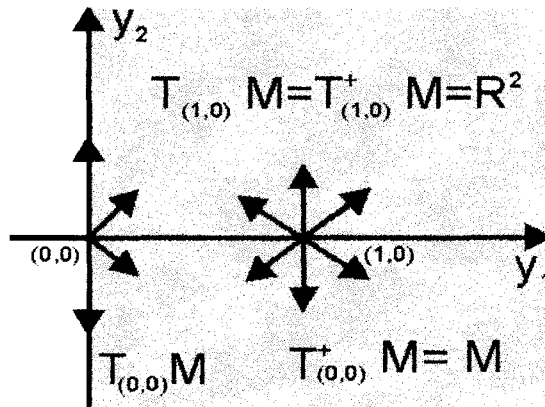


Figura2.5: Cones, em  $\mathbb{R}^2$ , de direções tangentes a  $M = \{(y_1, y_2) : y_1 \geq 0\}$ .

**Exemplo 2.5** Seja  $Z = Y \times \mathbb{R}$ , onde  $Y$  é um espaço normado e  $M = \{(y, t) : t = \|y\|\}$  é um cone. Então,  $T_{(0,0)}M = \{0\}$  e  $T_{(0,0)}^+M = M$ .

Realmente,  $y^o = (0, 0)$  deve satisfazer as condições:

$$\begin{cases} (0, 0) + \varepsilon(h_1, h_2) + (r_1(\varepsilon), r_2(\varepsilon)) \in M; \\ \|(r_1(\varepsilon), r_2(\varepsilon))\|_Z = o(\varepsilon); \varepsilon \rightarrow 0^+. \end{cases}$$

Assim,  $(\varepsilon h_1, \varepsilon h_2) + (r_1(\varepsilon), r_2(\varepsilon)) = (\varepsilon h_1 + r_1(\varepsilon), \varepsilon h_2 + r_2(\varepsilon)) \in M$  se, e somente se,  $\varepsilon h_2 + r_2(\varepsilon) = \|\varepsilon h_1 + r_1(\varepsilon)\|_Y$ . Quando  $\varepsilon \rightarrow 0^+$ , temos que  $o(\varepsilon) \rightarrow 0$ , ou seja,  $\|(r_1(\varepsilon), r_2(\varepsilon))\|_Z \rightarrow 0$ . Então tomamos a norma do máximo e obtemos

$$\left| \frac{r_2(\varepsilon)}{\varepsilon} \right| \rightarrow 0, \text{ se } \varepsilon \rightarrow 0^+ \text{ e } \left\| \frac{r_1(\varepsilon)}{\varepsilon} \right\|_Y \rightarrow 0, \text{ se } \varepsilon \rightarrow 0^+.$$

Deste modo, porque

$$\varepsilon \left( h_2 + \frac{r_2(\varepsilon)}{\varepsilon} \right) = |\varepsilon| \left\| h_1 + \frac{r_1(\varepsilon)}{\varepsilon} \right\|_Y \Rightarrow h_2 + \frac{r_2(\varepsilon)}{\varepsilon} = \frac{|\varepsilon|}{\varepsilon} \left\| h_1 + \frac{r_1(\varepsilon)}{\varepsilon} \right\|_Y,$$

seguem-se que:

a)  $\varepsilon > 0$  implica  $h_2 = \left\| h_1 + \frac{r_1(\varepsilon)}{\varepsilon} \right\|_Y - \frac{r_2(\varepsilon)}{\varepsilon}$ . Então, para  $\varepsilon \rightarrow 0^+$ ,  $h_2 = \|h_1\|_Y$ . Logo,  $T_{(0,0)}^+ M = M$ ;

b)  $\varepsilon < 0$  implica  $h_2 = - \left\| h_1 + \frac{r_1(\varepsilon)}{\varepsilon} \right\|_Y - \frac{r_2(\varepsilon)}{\varepsilon}$ . Assim,  $h_2 = - \|h_1\|_Y$ , quando  $\varepsilon \rightarrow 0^-$ .

Portanto, por a) e b),  $T_{(0,0)} M = \{(h_1, h_2) \in Z : h_1 = h_2 = 0\}$ .

## 2.2 Resultados diversos em análise

**Teorema 2.2.1** *Seja  $Z$  um espaço vetorial linear real. Sejam  $f_1$  e  $f_2 \in Z^*$ ,  $Q_1 = \{z \in Z : f_1(z) = 0\}$ ,  $Q_2 = \{z \in Z : f_2(z) = 0\}$  e  $Q_1 \subset Q_2$ . Então  $g = 0$  ou  $f = \lambda g$ ,  $\lambda \neq 0$ , equivalentemente,  $Q_2 = Z$  ou  $Q_1 = Q_2$ .*

**prova:** Suponhamos que  $f = 0$ . Então  $E = \text{Ker } f$ . Por hipótese,  $\text{Ker } f \subset \text{Ker } g$ , logo  $g = 0$ .

Se  $f \neq 0$ , então existe  $e \in Z$  tal que  $f(e) = 1$ . Agora,  $\forall x \in Z$ ,  $x = f(x)e + (x - f(x)e)$ . Verificamos que  $f(x - f(x)e) = 0$ , de onde definindo  $z = (x - f(x)e)$  segue que  $z \in \text{Ker } f$ . Tomando  $\lambda_1 = f(x)$  escrevemos  $x = \lambda_1 e + z$ ,  $\lambda \in \mathbb{R}$  e  $z \in \text{Ker } f \subset \text{Ker } g$ . Conseqüentemente,  $g(x) = \lambda_1 g(e) + 0 = f(x)g(e)$ . Portanto,  $g = g(e)f$ , ou seja,  $f = \lambda g$  com  $\lambda \neq 0$ . ■

**Teorema 2.2.2** *Seja  $K$  um conjunto convexo em  $Y$ . Então  $K$  é fortemente fechado se, e somente se, ele é fracamente fechado.*

**Prova:** Consulte Brezis [13, p. 38]. ■

**Teorema 2.2.3 (Existência das Funções Implícitas)** *Sejam  $X$  um espaço topológico,  $Y$  e  $Z$  espaços de Banach,  $W$  uma vizinhança do ponto  $(x_0, y_0)$  em  $X \times Y$ . Considere uma aplicação  $\Psi : W \rightarrow Z$  e  $\Psi(x_0, y_0) = z_0$ . Se as seguintes condições são verdadeiras :*

- 1) *A aplicação  $x \mapsto \Psi(x, y_0)$  é contínua em  $x_0$ ;*
- 2) *Existe uma aplicação  $\Lambda \in \mathcal{L}(Y, Z)$ <sup>5</sup> tal que  $\forall \varepsilon > 0, \exists \delta > 0$  e uma vizinhança  $\Sigma$  do ponto  $x_0$ , possuindo a propriedade a seguir:*

- *A condição  $x \in \Sigma$  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 existe um número  $\kappa > 0$ , uma vizinhança  $\mathcal{U}$  do ponto  $(x_0, z_0)$  em  $X \times Z$  e uma aplicação  $\varphi : \mathcal{U} \rightarrow Y$  tais que*

- a)  $\Psi(x, \varphi(x, z)) = z$ ;
- b)  $\|\varphi(x, z) - y_0\| \leq \kappa \|\Psi(x, y_0) - z\|$ .

**Definição 2.2.4** *Denomina-se equação de Volterra de segunda espécie à equação integral*

$$\psi(t) = \int_0^t K(t, s)\psi(s) ds + f(t), \quad (2.7)$$

*onde  $\psi, f \in L_2(0, T)$  e  $K(t, s)$  é uma função mensurável limitada<sup>6</sup>:  $|K(t, s)| \leq M \in \mathbb{R}$ .*

---

<sup>5</sup>  $\mathcal{L}(Y, Z)$  é o conjunto das transformações lineares contínuas de  $Y$  em  $Z$ , onde  $Y$  e  $Z$  são espaços de Banach.

<sup>6</sup>  $L_2(0, T)$  é o espaço de Banach das funções  $x(t)$  definidas de  $(0, T)$  em  $\mathbb{R}$ , quadrados integráveis segundo Lebesgue, cuja norma é dada por:  $\|x\| = \sqrt{\int_0^T |x(t)|^2 dt}$ .



**Proposição 2.2.5** *Qualquer que seja a função  $f \in L_2(0, T)$ , a equação de Volterra (2.7) admite uma única solução.*

**Prova:** A prova deste resultado pode ser encontrada em Kolmogorov [28, p. 462].

Outros tipos de condições estabelecidas sobre as equações integrais de Volterra são apresentados em Neustadt [33] e Yosida [38, 1960]. ■

**Teorema 2.2.6 (Teorema sobre o Anulador)** *Sejam  $E_1$  e  $E_2$  espaços de Banach, e seja  $\Lambda : E_1 \rightarrow E_2$  um operador linear contínuo tal que  $\text{Im } \Lambda = E_2$ . Então  $(\text{Ker } \Lambda)^\perp = \text{Im } \Lambda^*$  (isto é, o anulador do kernel de  $\Lambda$  é igual à imagem do operador adjunto  $\bar{\cdot}$  de  $\Lambda$ , denotado por  $\Lambda^*$ ).*

**Prova:** Veja Alexeév [2, 1982, p. 125 ], Ioffe [26, p. 16] ou Kolmogorov [28, p. 223]. ■

---

<sup>7</sup>Definição de *operador adjunto* pode ser encontrada em Kolmogorov [28, p. 221].

# Capítulo 3

## Formalismo de Dubovitskii-Milyutin

Neste capítulo, apresentaremos o resultado obtido por Dubovitskii e Milyutin<sup>1</sup> para problemas de minimização ou maximização de funcionais em um espaço de Banach - sujeitos ou não a certas restrições. Estes problemas são mencionados como problemas extremais.

Alguns dos resultados que iremos expor, servem também para outros tipos de espaços, como por exemplo, espaços lineares localmente convexos. Entretanto, utilizaremos como hipótese espaços de Banach, pois estes resultados servirão como ferramenta para obtermos o teorema de Fritz-John para programação matemática, e o Princípio do Máximo local de Pontryagin para um problema de controle ótimo<sup>2</sup>.

Exibiremos, a seguir, a forma na qual os problemas extremais podem ser colocados, de modo que a condição necessária de extremo, deduzida por Dubovitskii-Milyutin, seja aplicada a todos eles.

### Formulação do problema (PE):

Sejam dados um espaço de Banach  $Y$ ; um funcional  $F$  (às vezes não-linear) definido numa vizinhança  $V$  de  $y^o$ ; conjuntos  $Q_i \subset Y$ , para todo  $i = 1, \dots, n+1$ , onde o conjunto  $Q_i$  tem interior não vazio, quando  $i = 1, \dots, n$ ,

---

<sup>1</sup>Dubovitskii e Milyutin são matemáticos russos.

<sup>2</sup>Caso o leitor queira saber mais detalhes a respeito dos resultados obtidos em outros tipos de espaços, consulte Girsanov [22], Zeidler [40], Deimling [17], Ioffe [26] e Alexeév [3, 1987].

e o conjunto  $Q_{n+1}$  não tem pontos interiores. Queremos encontrar um ponto  $y^\circ \in Q$ , onde  $Q = \bigcap_{i=1}^{n+1} Q_i$ , tal que

$$F(y^\circ) = \min_{y \in Q \cap V} F(y). \quad (3.1)$$

Geralmente nos problemas práticos os conjuntos  $Q_i$ ,  $i = 1, \dots, n$ , são dados por restrições de desigualdade, e  $Q_{n+1}$  por um sistema de restrições de igualdade, de modo que  $Q_{n+1}$  é uma variedade linear <sup>3</sup> de dimensão menor do que a do espaço  $Y$ .

Abaixo, esboçaremos como expressar alguns problemas na forma geral de um problema extremal.

▷ **Problemas de programação não-linear** (restrições de igualdade e/ou desigualdade):

Seja  $V$  uma vizinhança aberta de  $\mathbb{R}^m$ . Consideremos as aplicações  $F : V \rightarrow \mathbb{R}$ ,  $G_i : V \rightarrow \mathbb{R}$ ,  $i = 1, \dots, r$ ,  $H_j : V \rightarrow \mathbb{R}$ ,  $j = 1, \dots, q$ . Queremos estudar o seguinte problema:

$$\begin{array}{l} \min_{y \in V} F(y) \\ \text{s.a.}^4 \quad \left\{ \begin{array}{l} G_i(y) \leq 0; \quad i = 1, \dots, r; \\ H_j(y) = 0; \quad j = 1, \dots, q. \end{array} \right. \end{array}$$

Neste caso,

$$\begin{aligned} Q_i &= \{y \in Y : G_i(y) \leq 0\}, \quad \forall i = 1, \dots, r; \\ Q_{r+1} &= \{y \in Y : H_j(y) = 0, \quad \forall j = 1, \dots, q\} \quad \text{e} \\ Q &= \bigcap_{i=1}^{r+1} Q_i. \end{aligned}$$

Deste modo, queremos obter

$$F(y^\circ) = \min_{y \in Q \cap V} F(y).$$

<sup>3</sup>Um subconjunto  $\Omega$  de um espaço normado é chamado *variedade linear* se, para quaisquer dois elementos  $y_1, y_2 \in \Omega$ , o elemento  $\omega_1 y_1 + \omega_2 y_2 \in \Omega$ , para todo  $\omega_1, \omega_2 \in \mathbb{R}$ .

<sup>4</sup>A abreviação s.a significa sujeito a.

▷ **Problema de controle ótimo:**

Dado o espaço produto  $E = (C[0, T])^n \times (L^\infty[0, T])^r$ . Sejam  $\varphi(x, u, t)$  uma função com valores vetoriais;  $\mathbb{L}(x, u, t)$  é uma função escalar e  $M \subset \mathbb{R}^r$ .

Queremos determinar em  $E$  as funções  $x(\cdot)$  (trajetória de fase ou de estado) e  $u(\cdot)$  (controle) que minimizem o funcional integral

$$F(x, u) = \int_0^T \mathbb{L}(x(t), u(t), t) dt \quad (3.2)$$

satisfazendo a equação diferencial:

$$\frac{d}{dt}x(t) = \varphi(x(t), u(t), t), \quad (3.3)$$

com as condições de contorno

$$x(0) = c; \quad (3.4)$$

$$x(T) = d, \quad (3.5)$$

onde o controle satisfaz restrições do tipo:

$$u(t) \in M \text{ q.t.p. } 0 \leq t \leq T^5. \quad (3.6)$$

Aqui  $x : [0, T] \longrightarrow \mathbb{R}^n$  e  $u : [0, T] \longrightarrow \mathbb{R}^r$ .

Agora, se colocarmos

$$Q_1 = \{(x, u) \in E : \text{satisfazem (3.6)}\} \text{ e}$$

$$Q_2 = \{(x, u) \in E : \text{satisfazem (3.3), (3.4) e (3.5)}\},$$

nosso problema será, portanto, minimizar a integral

---

<sup>5</sup>O termo q.t.p. significa para quase todo ponto.

Geralmente, usaremos q.t.p. para dizer que uma certa propriedade não vale para um conjunto de medida nula (medida de Lebesgue) em um intervalo qualquer  $(a, b)$ .

Definições a respeito de funções mensuráveis e integráveis segundo Lebesgue estão nas referências: Bartle [7], Fernandez [20], Kolmogorov [28], Rudin [35] e Yosida [38, 1978].

$$F(x, u) = \int_0^T \mathbb{L}(x(t), u(t), t) dt$$

sobre o conjunto  $Q = Q_1 \cap Q_2 \subseteq E$ , ou seja,

$$\min_{(x,u) \in Q} F(x, u).$$

Quando  $n = r = 1$ ,  $\varphi(x, u, t) = u$ , e  $M = \mathbb{R}$ , este problema se reduz ao problema clássico do Cálculo de Variações:

$$\min_{(x,u) \in C[0,T] \times L^\infty[0,T]} \int_0^T \mathbb{L}(x(t), u(t), t) dt$$

s.a

$$\begin{aligned} \frac{d}{dt}x(t) &= u(t); \\ x(0) &= c; \\ x(T) &= d. \end{aligned}$$

Voltemos a analisar o problema extremal na forma geral (3.1):

Para formularmos uma condição necessária para um ótimo iremos considerar o conjunto

$$Q_0 = \{y \in Y : F(y) < F(y^o)\}.$$

Desta forma, uma condição necessária para que  $y^o$  seja um mínimo local do funcional  $F$  é que

$$\bigcap_{i=0}^{n+1} Q_i = \emptyset. \quad (3.7)$$

Esta condição é suficiente também.

Averiguar (3.7), para que consigamos os possíveis candidatos a ótimo, não é muito prático. Visto que, para espaços mais estruturados como os de Banach, deixaríamos de utilizar, nos problemas práticos, resultados através

dos quais nos auxiliariam a encontrar uma forma analítica para esta condição necessária. Por conseguinte, com o fim de obtermos outras considerações a respeito do extremante, sem que percamos a essência geométrica apresentada em (3.7) (ou seja, a interseção entre os conjuntos é vazia), imporemos certas condições às classes dos conjuntos e funcionais a serem considerados, de modo que (3.7) seja expressa analiticamente. Neste contexto, consideraremos cones convexos associados aos conjuntos  $Q_i$ ,  $i = 1, \dots, n+1$ , e ao funcional  $F$ , conforme as idéias de Dubovitskii e Milyutin, as quais recorrem aos recursos das análises funcional e convexa.

### 3.1 Condições necessárias de extremo - Equação de Euler-Lagrange

Trataremos o problema (PE) na página 24, de forma analítica, com base nas definições e resultados expostos até o momento.

De início, apresentamos a condição necessária de extremo, que tanto almejamos neste trabalho. Em seguida, apresentamos o lema que utilizaremos na demonstração da mesma.

**Teorema 3.1.1 (Dubovitskii-Milyutin)** *Seja  $F : V \subseteq Y \rightarrow \mathbb{R}$  um funcional numa vizinhança  $V$  de  $y^o$ , no espaço normado real  $Y$ . Suponha*

$$F(y^o) = \min_{y \in Q \cap V} F(y),$$

onde  $Q = \bigcap_{i=1}^{n+1} Q_i$ . Suponha ainda que:

- (a)  $F(y)$  é regular de descida em  $y^o$  com direções de descida no cone  $K_0$ ;
- (b) as restrições de desigualdades em  $Q_i$ ,  $i = 1, \dots, n$ , são regulares em  $y^o$ , com direções factíveis no cone  $K_i$ ;
- (c) a restrição de igualdade em  $Q_{n+1}$ , é também regular em  $y^o$  com direções tangentes no cone  $K_{n+1}$ .

Então existem funcionais lineares contínuos  $f_i$ ,  $i = 0, \dots, n + 1$ , nem todos nulos, tais que

$$f_i \in K_i^*; i = 0, \dots, n + 1, \quad (3.8)$$

os quais satisfazem a equação de Euler-Lagrange <sup>6</sup>

$$f_0 + f_1 + \dots + f_n + f_{n+1} = 0$$

**Lema 3.1.2** Usando as hipóteses do teorema 3.1.1 acima, temos conseqüentemente que

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

(ou seja, nenhuma direção de descida do funcional  $F$  pode ser admissível para todas as restrições).

**Prova:** Definamos

$$\tilde{Q} = \bigcap_{i=1}^n Q_i.$$

Suponhamos que  $\bigcap_{i=0}^{n+1} K_i \neq \emptyset$ , isto é, que existe  $h \in K_i$ ,  $i = 0, \dots, n + 1$ . Usando a definição de direção de descida (2.1.14) e direção factível (2.1.17) resulta que existe uma vizinhança  $V$  do vetor  $h$  tal que, dado  $\varepsilon_0 > 0$  e  $\alpha < 0$ , para  $\varepsilon \in ]0, \varepsilon_0[$  o vetor

$$y_\varepsilon = y^o + \varepsilon \bar{h} \in \tilde{Q}, \text{ com } \bar{h} \in V, \quad (3.9)$$

satisfaz

$$F(y_\varepsilon) = F(y^o + \varepsilon \bar{h}) \leq F(y^o) + \varepsilon \alpha.$$

---

<sup>6</sup> †Leonardo Euler (1707-1783), filho de Paul Euler e Marguerite Brucker, nasceu na Basiléia. É um dos maiores analistas de todos os tempos e o homem o qual seus contemporâneos chamavam “análise encarnada”. Publicou numerosas obras, tendo estudado inúmeras ciências. No fim da vida ocupou-se do movimento ascensional dos aerolatos e do planeta de Herschel.

‡Luiz José Lagrange (1736-1813) nasceu em Turim, com dezoito anos foi nomeado professor na Escola de Artillaria de Turim. Lagrange publicou, entre outras, as obras: Mecânica Analítica, Teoria das Funções, Tratado da Resolução das Equações numéricas de Todos os Graus. Deixou ainda numerosos trabalhos na Miscelanea Taurinensia e nas Memórias das Academias de Ciências de Turim, de Berlim e Paris. Nos seus trabalhos publicados na Miscelanea Taurinensia, encontra-se o gérmen do Cálculo das Variações.

Agora consideremos o vetor  $y_{\varepsilon_1} = y^o + \varepsilon_1 h + r(\varepsilon_1) \in Q_{n+1}$ , como na definição de direção tangente (2.1.20), contanto que seja dado  $\varepsilon_1$  tal que, para todo  $0 < \varepsilon < \varepsilon_1$ ,  $\frac{r(\varepsilon)}{\varepsilon} \in (V - h)$ , de outra forma dito,

$$\bar{h}(\varepsilon) = (h + \frac{1}{\varepsilon}r(\varepsilon)) \in V.$$

Então, quando  $0 < \varepsilon < \min\{\varepsilon_0, \varepsilon_1\}$ , os vetores  $y(\varepsilon) = y^o + \varepsilon\bar{h}(\varepsilon)$  pertencem a  $\tilde{Q}$ , por um lado, devido a (3.9). Por outro lado,

$$y(\varepsilon) = y^o + \varepsilon\bar{h}(\varepsilon) = y^o + \varepsilon(h + \frac{1}{\varepsilon}r(\varepsilon)) = y^o + \varepsilon h + r(\varepsilon) \in Q_{n+1}.$$

Em outras palavras, estes vetores satisfazem todas as restrições de igualdade e desigualdades. Mas eles também satisfazem:

$$F(y^o + \varepsilon\bar{h}) \leq F(y^o) + \varepsilon\alpha < F(y^o),$$

a qual contradiz o fato de  $y^o$  ser um ponto de mínimo local. Assim,  $\bigcap_{i=0}^{n+1} K_i = \emptyset$ . ■

Apresentaremos neste momento uma demonstração concisa do teorema de Dubovitskii-Milyutin.

**Prova do teorema de Dubovitskii-Milyutin:** Considerando as hipóteses deste teorema, poderemos aplicar o lema 3.1.2 obtendo:

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

De acordo com o lema 2.1.15, a definição 2.1.16, o lema 2.1.18, a definição 2.1.19, o quarto item da observação 2.1.22 da página 19 e a definição 2.1.23 mencionados anteriormente, referentes aos cones de direção de descida  $K_0$ , de direção factível  $K_i$ ,  $i = 1, \dots, n$ , e de direção tangente  $K_{n+1}$ , dois a dois respectivamente, podemos dizer que  $K_0, K_1, \dots, K_n$  são cones convexos abertos e  $K_{n+1}$  um cone convexo. Portanto nos é permitido utilizar o teorema 2.1.13, de onde concluímos a existência de funcionais lineares contínuos  $f_i$ ,  $i = 0, \dots, n + 1$ , não simultaneamente nulos, tais que  $f_0 + f_1 + \dots + f_n + f_{n+1} = 0$ . ■



**Observação 3.1.3** Quando se aplica essa condição de extremo a problemas com extremos condicionados a restrições de igualdade, obtém-se a **regra de multiplicadores de Lagrange**, conhecida na análise clássica; já àqueles problemas com restrições de desigualdade, obtém-se a **condição de Karush-Kuhn-Tucker**. E quando o problema é de controle, obtém-se o **Princípio do Máximo Local de Pontryagin**, o qual é exatamente a equação de Euler-Lagrange deduzida para o problema clássico do Cálculo Variacional:

$$\min \int_0^T \Phi(x(t), u(t), t) dt,$$

s. a

$$\dot{x}(t) = u(t);$$

$$x(0) = c;$$

$$x(T) = d.$$

Isto justifica o nome **Equação de Euler-Lagrange** dado à condição de extremo generalizada por Dubovitskii-Milyutin.

Deixamos, para o leitor interessado na dedução dessas condições, as seguintes referências: Blum [10, p. 78 a 82], Deimling [17, p. 339 e 340]; Alcxécév [2, 1982, p. 373 a 375]; Girsanov [22, p. 79 a 92]; Ioffe [26, p. 142 a 145]; Alcxécév [3, 1987, p. 190 a 196] e Zeidler [40, p. 433 a 435].

**Observação 3.1.4** A característica essencial da condição de extremo dada pela equação (3.8) na página 29 é que cada funcional ou restrição podem ser analisados isoladamente. Assim, precisamos fazer a análise de cada funcional ou conjunto somente uma vez, definitivamente. E os resultados encontrados podem ser utilizados para qualquer problema extremal no qual eles se evidenciam.

De acordo com a observação acima, precisamos saber como calcular os cones e seus respectivos duais para os problemas de programação não-linear e de controle, os quais mencionaremos mais adiante.

## 3.2 Cálculo de cones e seus duais

Nesta seção, faremos uma breve seleção de resultados, os quais nos possibilitam exibir os cones de direções de descida, direções factíveis e direções tangentes e seus respectivos duais, associados aos problemas que apresentaremos no próximo capítulo.

### 3.2.1 Direções de descida

De início iremos recordar a definição de derivada segundo Fréchet, pois a usaremos juntamente com suas propriedades, não mencionadas aqui, mas obtidas em Cartan [14], Flett [21], Alexécév [2, 1982] e Neustadt [33].

**Definição 3.2.1** *Suponhamos que numa vizinhança do ponto  $y^o$  a aplicação  $F$  pode ser representada sob a forma*

$$F(y^o + h) = F(y^o) + \Lambda h + \alpha(h) \|h\|, \quad (3.10)$$

onde  $\Lambda \in \mathcal{L}(Y, E)$  e

$$\lim_{\|h\| \rightarrow 0} \|\alpha(h)\| = \|\alpha(0)\| = 0. \quad (3.11)$$

Então a aplicação  $F(\cdot)$  é dita diferenciável segundo Fréchet no ponto  $y^o$ , ou Fréchet diferenciável em  $y^o$ . O operador  $\Lambda$  é chamado **derivada de Fréchet** (ou simplesmente, **derivada**, ou ainda, **diferencial de Fréchet** e se  $E = \mathbb{R}$ , **gradiente**) da aplicação  $F$  no ponto  $y^o$ . Este operador é denotado por  $F'(y^o)$ ,  $dF(y^o)$  ou  $DF(y^o)$ .

As relações (3.10) e (3.11) podem ser escritas da seguinte forma:

$$F(y^o + h) = F(y^o) + F'(y^o)h + o\|h\|, \quad (3.12)$$

ou na relação limite:

$$\lim_{h \rightarrow 0} \frac{F(y^o + h) - F(y^o) + F'(y^o)h}{\|h\|} = 0.$$

A função  $y \mapsto F'(y)$ , cujo domínio é o conjunto de pontos interiores  $y$  de  $A \subseteq Y$ , para os quais  $F'(y)$  existe, é chamada a diferencial de Fréchet de  $F$ , ou a derivada de  $F$ . Denotaremos-la por  $F'$ . Assim  $F'$  é uma função de um subconjunto  $A \subseteq Y$  em  $\mathcal{L}(Y, E)$ .

Usando a linguagem em  $\varepsilon$  e  $\delta$ , podemos reformular (3.10) e (3.11) do seguinte modo:

Para cada  $\varepsilon > 0$ , existe  $\delta > 0$  tal que, para  $h \in Y$  verificando  $\|h\| < \delta$ , tem-se a desigualdade

$$\|F(y^\circ + h) - F(y^\circ) - F'(y^\circ)h\| \leq \varepsilon \|h\|. \quad (3.13)$$

De acordo com as afirmações anteriormente assinaladas, a derivada  $F'(y^\circ)$  é por definição uma aplicação linear de  $Y$  em  $E$ , e os valores desta aplicação sobre os vetores  $h \in Y$  são usualmente denotados por  $F'(y^\circ)[h]$ , ou  $F'(y^\circ)h$ .

Durante o nosso trabalho faremos uso da notação  $F'(y^\circ)h$ .

Novamente, designemos por  $K_0$  o cone de direções de descida do funcional  $F$  no ponto  $y^\circ$ .

**Teorema 3.2.2** *Seja  $Y$  um espaço de Banach. Se  $F$  é Fréchet-diferenciável em  $y^\circ$ , então  $F$  é regularmente de descida em  $y^\circ$  e*

$$K_0 = \{h \in Y : F'(y^\circ)h < 0\}$$

$$K_0^* = \begin{cases} Y^* & \text{se } F'(y^\circ) \equiv 0 \\ \{-\lambda_0 F'(y^\circ) / \lambda_0 \geq 0\} & \text{se } F'(y^\circ) \neq 0 \end{cases}$$

**Prova:** Pela definição de direção de descida (2.1.14), se  $h \in K_0$ , então  $\frac{1}{\varepsilon} (F(y^\circ + \varepsilon \bar{h}) - F(y^\circ)) \leq \alpha < 0$ . Porque a aplicação  $F$  é Fréchet-diferenciável, ela possui derivada direcional em  $h$ . Logo, tomando  $\varepsilon \rightarrow 0$ , temos  $F'(y^\circ)h < 0$ .

Reciprocamente, se  $F(y)$  é Fréchet-diferenciável em  $y^\circ$  e  $F'(y^\circ)h < 0$ , então quando  $\eta \rightarrow 0$ ,

$$F(y^\circ + \eta) = F(y^\circ) + F'(y^\circ)\eta + o(\|\eta\|).$$

Para  $\eta = \varepsilon \bar{h}$ , onde  $\bar{h}$  pertence a uma vizinhança de  $h$ , segue imediatamente que  $h$  é uma direção de descida, i.e.,  $h \in K_0$ . Como  $\{h \in Y : F'(y^\circ)h < 0\}$

é convexo (pois é um semi- espaço ou o conjunto vazio) teremos que  $F$  é regularmente de descida.

Agora, para calcularmos o cone dual de  $K_0$ , consideremos  $f \in Y^*$  e os cones

$$\begin{aligned}\tilde{K} &= \{h \in Y : f(h) > 0\}, \\ K_1 &= \{h \in Y : f(h) = 0\} \text{ e} \\ K_2 &= \{h \in Y : f(h) \geq 0\}.\end{aligned}$$

Se  $0 \neq g \in K_1^*$ , então  $g(h) = 0$  em  $K_1$ , pelo teorema 2.1.5. Consideremos  $S = \{h \in Y : g(h) = 0\}$ . Deste modo,  $K_1 \subset S$ . Por conseguinte, de acordo com 2.2.1,

$$\begin{cases} g = 0 \text{ ou} \\ g = \lambda f; \lambda \neq 0. \end{cases}$$

A primeira condição não pode ocorrer porque tomamos  $g \neq 0$ . Desde que  $K_1 \subset K_2$ , segue-se que  $K_2^* \subset K_1^*$ . Mas se  $g = \lambda f \in K_2^*$ , então  $g(h) = \lambda f(h) \geq 0, \forall h \in K_2$  com  $\lambda \neq 0$ . Como  $f(h) \geq 0, \forall h \in K_2$ , segue-se que  $g(h) = \lambda f(h) \geq 0, \forall h \in K_2$  e  $\lambda \geq 0$ . Finalmente, se  $f = 0$  então  $\tilde{K} = \emptyset$  e portanto  $\tilde{K}^* = Y^*$ ; senão  $\tilde{K}^* = K_2^*$ , porque o fecho de  $\tilde{K}, \overline{\tilde{K}}$ , coincide com  $K_2$  e  $\tilde{K}^* = (\overline{\tilde{K}})^*$ .

Com base no que acabamos de mostrar, verificamos que

$$K_0^* = \{-\lambda_0 F'(y^0) : \lambda_0 \geq 0\}.$$

Mas se  $F'(y^0) \equiv 0$ , então  $K_0 = \emptyset$ . E por conseguinte,  $K_0^* = Y^*$ . ■

### 3.2.2 Direções factíveis

Denotaremos por  $K_a$  o cone de direções factíveis no ponto  $y^0$  para o conjunto  $Q$  em um espaço de Banach  $Y$ .

Como mencionamos anteriormente,  $Q$  pode ser dado por restrições de desigualdades nos problemas extremais.

Primeiramente, o único caso de sumo interesse é aquele onde  $y^0$  é um ponto da fronteira de  $Q$ , visto que se  $y^0 \in \text{int}Q$ , segue-se diretamente da definição 2.1.17 que  $K_a \equiv Y$ , ou seja, qualquer direção é factível.

Caso o conjunto  $Q$  não tenha pontos interiores, obteremos uma resposta trivial:  $K_a \equiv \emptyset$ .

Agora consideremos o caso em que o conjunto  $Q$  é definido por algum funcional:

$$Q = \{y \in Y : F(y) \leq F(y^o)\}$$

e definamos  $K_D$  o cone de direções de descida para  $Q$  em  $y^o$ . Teremos sob certas condições o cone  $K_a$ , como mostra o que se segue:

**Teorema 3.2.3** *Considerando os conjuntos acima definidos e supondo ainda que a aplicação  $F$  é diferenciável em  $y^o$ , tal que  $F'(y^o) \neq 0$ , conclui-se que*

$$K_a = K_D = \{h \in Y : F'(y^o)h < 0\}.$$

**Prova:** Primeiramente iremos verificar que  $K_D \subset K_a$  e em seguida que  $K_a \subset \{h \in Y : F'(y^o)h < 0\}$ . Depois usando o teorema 3.2.2 obteremos o resultado desejado.

De fato, se  $h \in K_D$ , temos que dado  $\alpha < 0$ ,  $F(y^o + \varepsilon \bar{h}) \leq F(y^o) + \varepsilon \alpha < F(y^o)$ , para todo  $\bar{h}$  pertencente a uma vizinhança  $U$  do vetor  $h$  e para todo  $\varepsilon \in ]0, \varepsilon_0[$ . Além disso,  $y^o + \varepsilon \bar{h} \in Q$ ,  $\forall \bar{h} \in U$  e  $\varepsilon \in ]0, \varepsilon_0[$ . Mas isto é a definição de direção factível. Logo  $K_D \subset K_a$ .

Agora suponhamos que  $h \in K_a$ , ou seja, que  $y^o + \varepsilon h \in Q$ ,  $\forall \varepsilon \in ]0, \varepsilon_0[$ . Então  $F(y^o + \varepsilon h) \leq F(y^o)$ .

Como  $F$  é Fréchet-diferenciável,  $F$  é diferenciável em qualquer direção e por isto

$$F'(y^o)h = \lim_{\varepsilon \downarrow 0} \frac{F(y^o + \varepsilon h) - F(y^o)}{\varepsilon} \leq 0^7.$$

Por outro lado, sabemos que o cone  $K_a$  é aberto, portanto existe uma vizinhança  $U$  do vetor  $h$  tal que  $U \subset K_a$ .

Tomemos  $\gamma > 0$  e  $\tilde{h} \in U$ , tais que  $F'(y^o)\tilde{h} < 0$  e  $h_\gamma = h - \gamma(h - \tilde{h}) \in U$ . Devido ao fato de  $h_\gamma \in K_a$ , segue, de acordo com as idéias acima, que  $F'(y^o)h_\gamma \leq 0$ . Porque  $F'(y^o)$  é linear obtemos:

$$\begin{aligned} F'(y^o)h &= F'(y^o)\left[\frac{1}{1+\gamma}h_\gamma + \frac{\gamma}{1+\gamma}\tilde{h}\right] \leq \frac{1}{1+\gamma}F'(y^o)h_\gamma + \frac{\gamma}{1+\gamma}F'(y^o)\tilde{h} \\ &\leq \frac{\gamma}{1+\gamma}F'(y^o)\tilde{h} < 0. \end{aligned}$$

---

<sup>7</sup> $\varepsilon \downarrow 0$  significa que  $\varepsilon$  tende a zero pela direita, ou seja,  $\varepsilon$  tende a zero por valores positivos.

Logo  $K_a \subset \{h \in Y : F'(y^0)h < 0\}$ . Pelo teorema 3.2.2 e inclusões que provamos, segue o resultado desejado. ■

Finalmente mencionaremos o caso especial de um conjunto  $Q$  não definido por um funcional, embora seu cone de direções factíveis tenha uma descrição simples (confira através da figura 3.1, a qual esboça o cone  $K_a$  no ponto  $y^o = (y_1^o, y_2^o)$  de qualquer círculo  $Q$  em  $\mathbb{R}^2$ ).

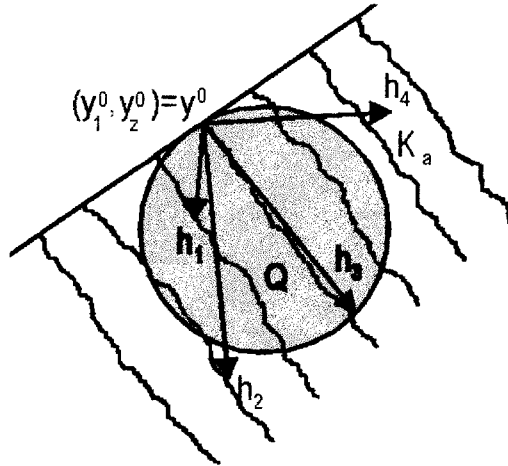


Figura3.1: Cone de direção factível  $K_a$

**Teorema 3.2.4** *Seja  $Q$  um conjunto convexo no espaço de Banach  $Y$ . Então o cone factível,  $K_a$ , no ponto  $y^o$  é dado por*

$$K_a = \{h \in Y : h = \lambda(y - y^o), y \in \text{int}Q, \lambda > 0\}.$$

**Prova:** Seja  $A = \{h \in Y : h = \lambda(y - y^o), y \in \text{int}Q \text{ e } \lambda > 0\}$ . Provaremos que  $K_a = A$ , para o caso em que  $y^o = 0$ .

Inicialmente suponhamos que  $h = \lambda y$ ,  $y \in \text{int}Q$ , e  $\lambda > 0$ . Verificaremos se  $h \in K_a$ :

Dado  $\varepsilon_0 > 0$ , tomemos  $0 < \varepsilon = \frac{\varepsilon_0}{\|\lambda y\|} < \varepsilon_0$ . Sabendo que  $Q$  é convexo,

$$\varepsilon h = \frac{\varepsilon_0(\lambda y)}{\|\lambda y\|} = \left( \frac{\varepsilon_0 \lambda}{\|\lambda y\|} \right) y \in Q.$$

Logo,  $h \in K_a$ . Por outro lado, se  $h \in K_a$ , então dado  $\varepsilon_0 > 0$ , existe  $\varepsilon \in ]0, \varepsilon_0[$ , tal que  $\varepsilon h \in Q$ . Fazendo  $y = \varepsilon h$ , temos que  $h = \frac{y}{\varepsilon}$ . De onde segue que  $h \in A$ , com  $\lambda = \varepsilon^{-1} > 0$ . Desta forma,  $K_a \equiv A$ . ■

### 3.2.3 Direções tangentes

Recordemos a definição de diferenciabilidade estrita tomando os espaços de Banach  $Y$  e  $Z$  (veja novamente Cartan [14], Flett [21] e Alexeév [2, 1982]).

**Definição 3.2.5** *A aplicação  $F$  é chamada estritamente diferenciável no ponto  $y^o$ , se existe um operador  $\Lambda \in \mathcal{L}(Y, Z)$  tal que para todo  $\varepsilon > 0$ , existe  $\delta > 0$ , de modo que para quaisquer  $y_1$  e  $y_2$  satisfazendo  $\|y_1 - y^o\| < \delta$  e  $\|y_2 - y^o\| < \delta$ , tem-se a seguinte desigualdade:*

$$\|F(y_1) - F(y_2) - \Lambda(y_1 - y_2)\| \leq \varepsilon \|y_1 - y_2\|. \quad (3.14)$$

Colocando  $y_2 = y^o$  e  $y_1 = y^o + h$  em (3.14), obtemos (3.13), de modo que uma função estritamente diferenciável é também diferenciável segundo Fréchet, e  $\Lambda = F'(y^o)$ .

Uma ferramenta essencial para determinarmos o cone de direções tangentes é o Teorema de Lyusternik, enunciado como segue:

**Teorema 3.2.6 ( Lyusternik )** *Sejam  $Y, Z$  espaços de Banach;  $V$  uma vizinhança do ponto  $y^o \in Y$  e uma aplicação  $P : V \rightarrow Z$  tal que  $P(y^o) = 0$ . Se  $P$  é estritamente diferenciável em  $y^o$  e  $P'(y^o)Y \equiv Z$  ( i.e.,  $P'(y^o)$  é um epimorfismo), então o conjunto  $M = \{y \in Y : P(y) = 0\}$  possui, no ponto  $y^o$ , um espaço tangente*

$$T_{y^o} M = T_{y^o}^+ M = \ker P'(y^o) = \{h \in Y : P'(y^o)h = 0\}.$$

**Prova:** Primeiramente suponhamos  $h \in T_{y^o}^+ M$ . Assim, se  $r : [0, \varepsilon] \rightarrow Y$  tal que  $y^o + \alpha h + r(\alpha) \in M$ , com  $\|r(\alpha)\| = o(\alpha)$ , para  $\alpha \rightarrow 0^+$ , então, em virtude da diferenciabilidade estrita de  $P$  (de acordo com a definição 3.2.5), temos que

$$0 = P(y^o + \alpha h + r(\alpha)) = P(y^o) + \alpha P'(y^o)h + o(\alpha) = \alpha P'(y^o)h + o(\alpha).$$

Logo,  $P'(y^o)h = 0$ . De onde concluímos que  $h \in \ker P'(y^o)$ .

Agora, para provarmos a outra inclusão, aplicaremos o Teorema das Funções Implícitas 2.2.3 à função  $\Psi(y, h) = P(y + h)$ , com  $h^o = 0$  e  $z^o = 0$ .

Decorre da diferenciabilidade estrita de  $P$  que a aplicação  $y \rightarrow \Psi(y, 0)$  é contínua no ponto  $y^o$  e que

$$\left\| P(y + \tilde{h}) - P(y + \hat{h}) - P'(y^o)(\tilde{h} - \hat{h}) \right\| \leq \varepsilon \left\| \tilde{h} - \hat{h} \right\| \iff$$

$$\left\| \Psi(y, \tilde{h}) - \Psi(y, \hat{h}) - P'(y^o)(\tilde{h} - \hat{h}) \right\| \leq \varepsilon \left\| \tilde{h} - \hat{h} \right\|,$$

quando  $y \in B(y^o, \delta)$  e  $\tilde{h}, \hat{h} \in B(0, \delta)$ <sup>8</sup>. A condição (3), do teorema das Funções Implícitas 2.2.3, também é satisfeita, já que  $P'(y^o)Y \equiv Z$ . Portanto teremos a existência de uma aplicação  $\tilde{\varphi} : V \rightarrow Y$ , onde  $V$  é uma vizinhança do ponto  $y^o$  em  $Y$  e  $\tilde{\varphi}(y) = \varphi(y, z), \forall z \in \mathcal{U}$ , vizinhança de  $(y^o, z^o) = (y^o, 0)$ . Considerando em particular para  $z = 0$ , segue que

$$\Psi(y, \varphi(y, z)) = \Psi(y, \tilde{\varphi}(y)) = 0 \iff P(y + \tilde{\varphi}(y)) = 0; \quad (3.15)$$

$$\begin{aligned} \|\varphi(y, 0) - h^o\| &= \|\tilde{\varphi}(y)\| \leq \kappa \|\Psi(y, h^o) - z\| \\ &= \kappa \|\Psi(y, 0)\| = \kappa \|P(y)\|. \end{aligned} \quad (3.16)$$

Definamos  $r : [-\tilde{\varepsilon}, \tilde{\varepsilon}] \rightarrow Y$  tal que  $r(\alpha) = \tilde{\varphi}(y^o + \alpha h)$ . Daí,

$$P(y^o + \alpha h + r(\alpha)) = P(y^o + \alpha h + \tilde{\varphi}(y^o + \alpha h)) = 0,$$

por (3.15); de onde segue a afirmação  $y^o + \alpha h + r(\alpha) \in M$ . Temos também que

$$\begin{aligned} \|r(\alpha)\| &= \|\tilde{\varphi}(y^o + \alpha h)\|, \text{ por (3.16),} \\ \|\tilde{\varphi}(y^o + \alpha h)\| &\leq \kappa \|P(y^o + \alpha h)\| = \kappa \|P(y^o + \alpha h) - P(y^o)\| = \\ \kappa \|P'(y^o)\alpha h + o(\|\alpha h\|)\| &= \kappa \|\alpha P'(y^o)h + o(\alpha)\|, \text{ para } k > 0. \end{aligned}$$

Se  $h \in \text{Ker } P'(y^o)$ , então  $\|r(\alpha)\| = o(\alpha)$ . Assim, pela definição (2.1.20) de cone tangente, resulta que  $h \in T_{y^o}M \subseteq T_{y^o}^+M$ . Portanto,  $\ker P'(y^o) \subseteq T_{y^o}M \subseteq T_{y^o}^+M$ . Desta maneira mostramos as igualdades:

---

<sup>8</sup> $B(c, \delta) = \{y \in Y : \|y - c\|_Y < \delta\}$ .



$$T_{y^o}M = T_{y^o}^+M = \ker P'(y^o) = \{h \in X; P'(y^o)h = 0\}. \blacksquare$$

Neste momento consideremos a mesma notação para cone factível,  $K_a$ , e definamos  $K_T$  como sendo o cone de direções tangentes unilaterais ao subconjunto  $E \subseteq Y$  no ponto  $y^o \in E$ .

Veremos quando é que os cones duais dos cones tangente e factível coincidem com o conjunto  $\tilde{E}$  dos funcionais lineares suportes para  $E$  em  $y^o$ , queremos dizer, com o conjunto  $\tilde{E} = \{f \in Y^* : f(y) \geq f(y^o), \forall y \in E\}$ .

**Teorema 3.2.7** *Sejam um conjunto  $E$  convexo fechado em  $Y$  e um ponto  $y^o \in E$ . Então,*

$$K_T^* \equiv \tilde{E}.$$

*Se, além disso,  $\text{int } E \neq \emptyset$ , então*

$$K_a^* \equiv \tilde{E}.$$

**Prova:** Sejam dados um funcional  $f \in \tilde{E}$  e uma direção tangente unilateral  $h \in K_T$ . Dado  $0 < \varepsilon < \varepsilon_0$ , existe  $y_\varepsilon \in E$ ,  $y_\varepsilon = y^o + \varepsilon h + r(\varepsilon)$ , tal que  $\frac{r(\varepsilon)}{\varepsilon} \in V$ , para toda vizinhança  $V$  de zero e  $\varepsilon$  suficientemente pequeno. Já que  $y_\varepsilon \in E$  e  $f \in \tilde{E}$ , temos que  $f(y_\varepsilon) \geq f(y^o)$ . Portanto

$$f(h) = \frac{f(y_\varepsilon) - f(y^o)}{\varepsilon} - \frac{f(r(\varepsilon))}{\varepsilon} \geq -f\left(\frac{r(\varepsilon)}{\varepsilon}\right).$$

Mas  $f\left(\frac{r(\varepsilon)}{\varepsilon}\right) \rightarrow 0$ , quando  $\varepsilon \rightarrow 0^+$  (pois, desde que  $f$  é contínua, para cada  $\delta > 0$ , existe uma vizinhança  $V$  de 0, tal que para qualquer  $y \in V$ ,  $|f(y)| \leq \delta$ . Assim,  $\left|f\left(\frac{r(\varepsilon)}{\varepsilon}\right)\right| < \delta$ , para  $\varepsilon < 0$  suficientemente pequeno). Por conseguinte,  $f(h) \geq 0$ , isto é,  $f \in K_T^*$ .

Reciprocamente, sejam  $f \in K_T^*$ ,  $y \in E$ . Então  $h = y - y^o$  é uma direção tangente (para  $0 < \varepsilon < 1$ , tome  $y^o + \varepsilon h \in Y$ , donde obtem que  $y^o + \varepsilon h =$

$(1 - \varepsilon)y^\circ + \varepsilon y \in E$ , já que  $E$  é convexo). Visto que  $f \in K_T^*$ , segue-se que  $f(h) \geq 0$ , ou seja,  $f(y) \geq f(y^\circ)$ . Isto significa que  $f \in E^*$ .

Deste modo, temos provado que  $\tilde{E} = K_T^*$ .

Se  $\text{int } E \neq \emptyset$ , então pelo teorema anterior,

$$K_a = \{h \in Y : h = \lambda(y - y^\circ), y \in \text{int } E \text{ e } \lambda \geq 0\}.$$

Conseqüentemente, se  $f \in K_a^*$ , então, para todo  $y \in \text{int } E$  e todo  $\lambda \geq 0$ ,

$$f(\lambda(y - y^\circ)) \geq 0, \text{ equivalentemente, } \lambda f(y) \geq \lambda f(y^\circ).$$

Em particular, para  $\lambda = 1$ . Entretanto,  $\overline{\text{int } E} = \bar{E}$ , quando  $\text{int } E \neq \emptyset$ . Portanto, pela continuidade do funcional  $f$ , o fato de que  $f(y) \geq f(y^\circ)$ , para todo  $y \in \text{int } E$ , implica que  $f(y) \geq f(y^\circ)$ , para todo  $y \in \bar{E}$ . Podemos então dizer que  $f \in \bar{E}$  e  $K_a^* \subseteq K_T^* = \bar{E}$ . Por outro lado,  $K_a \subset K_T$ , por conseguinte, através da definição de cone dual (2.1.3),  $K_a^* \supset K_T^* = \tilde{E}$ . Logo,  $K_a^* = \tilde{E}$ . ■

Por causa do resultado acima, verifica-se que em muitos casos a determinação dos cones duais é equivalente a determinação dos funcionais suporte para um conjunto convexo fechado em  $Y$ .

# Capítulo 4

## Aplicação a problemas de otimização

Mostraremos, como consequência das idéias desenvolvidas anteriormente, condições necessárias de otimalidade para um problema de programação não-linear no  $\mathbb{R}^m$  e para um problema de controle ótimo a ser especificado na seção 4.2.2.

### 4.1 Problema de programação não-linear

#### 4.1.1 - Considerações sobre o problema

Seja  $V$  uma vizinhança aberta de  $y^o \in \mathbb{R}^m$ . Consideremos as funções  $F : V \rightarrow \mathbb{R}$ ,  $G_i : V \rightarrow \mathbb{R}$ ,  $i = 1, \dots, r$  e  $H_j : V \rightarrow \mathbb{R}$ ,  $j = 1, \dots, q$ . Queremos estudar o seguinte problema não-linear, (PNL):

$$\begin{array}{l} \min_{y \in V} F(y) \\ \text{s.a} \quad \left\{ \begin{array}{ll} G_i(y) \leq 0; & i = 1, \dots, r; \\ H_j(y) = 0; & j = 1, \dots, q. \end{array} \right. \end{array}$$

Ao definirmos  $I = \{1, \dots, r\}$  e  $J = \{1, \dots, q\}$  observemos que:

- se  $I = \emptyset$ , o problema (PNL) apresenta somente restrições de igualdade;
- se  $J = \emptyset$ , o problema (PNL) apresenta somente restrições de desigualdade;
- se  $I = \emptyset$  e  $J = \emptyset$ , o problema (PNL) é sem restrições.

### 4.1.2 Condições necessárias de Fritz-John

No que se segue, obteremos condições para um extremo local em dimensão finita dadas por Fritz-John.

**Teorema 4.1.1 (Fritz-John)** *Considere os funcionais:  $F$  diferenciável em  $y^o$ ,  $G_i$  é diferenciável em  $y^o$ , para todo  $i = 1, \dots, r$  e  $H_j$  estritamente diferenciável em  $y^o$ , para todo  $j = 1, \dots, q$ . Se  $y^o$  é ótimo local de (PNL), então existem escalares  $\lambda_0, \lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_{q-1}$  e  $\mu_q$  nem todos nulos simultaneamente, tais que :*

- i)  $\lambda_0 \nabla F(y^o) + \sum_{i=1}^r \lambda_i \nabla G_i(y^o) + \sum_{j=1}^q \mu_j \nabla H_j(y^o) = 0$ <sup>1</sup>;
- ii)  $\lambda_i \geq 0, i = 0, \dots, r$ ;
- iii)  $\lambda_i G_i(y^o) = 0, i = 1, \dots, r$ .

Antes de provarmos o teorema faremos algumas considerações relativas à definição que se segue:

**Definição 4.1.2** *Diz-se que uma restrição  $G_i$ , para algum  $i$ , do problema (PNL) é **ativa** no ponto  $y^o$  se  $G_i(y^o) = 0$ . Caso contrário, isto é, se  $G_i(y^o) < 0$ , diz-se que a restrição é **inativa**.*

---

<sup>1</sup>  $R'(y^o) = \nabla R(y^o) = \text{grad } R(y^o) = (\frac{\partial}{\partial y_1} R(y^o), \frac{\partial}{\partial y_2} R(y^o), \dots, \frac{\partial}{\partial y_m} R(y^o))$ .

Note que as restrições inativas não têm influência na caracterização dos mínimos locais, já que definindo  $\tilde{I}(y^o) = \{i : G_i(y^o) = 0\}$ , se  $G_s(y^o) < 0$ , para algum  $s \in I$ , ou seja, se  $s \notin \tilde{I}(y^o)$ , então  $\lambda_s = 0$ . Portanto, teremos a condição necessária:

$$\text{i) } \lambda_0 \nabla F(y^o) + \sum_{i \in \tilde{I}(y^o)} \lambda_i \nabla G_i(y^o) + \sum_{j=1}^q \mu_j \nabla H_j(y^o) = 0;$$

$$\text{ii) } \lambda_i \geq 0, \quad i \in \{0, \tilde{I}(y^o)\}.$$

Por isso, basta analisarmos os casos em que as restrições são ativas em  $y^o$ . Isto é o que faremos na demonstração abaixo.

**Prova do Teorema de Fritz-John:** Faremos uma análise dos casos triviais, seguida dos casos em que o formalismo mencionado até agora, facilitará a verificação do resultado desejado.

**1º passo (Casos especiais triviais)**

Se  $\nabla F(y^o) = 0$ , então são satisfeitas (i), (ii) e (iii), com  $\lambda_0 = 1$ ,  $\lambda_i = 0$ , para  $i = 0, \dots, r$  e  $\mu_j = 0$ , para  $j = 1, \dots, q$ .

Se  $\nabla G_s(y^o) = 0$  e  $s \in \tilde{I}(y^o)$ , então (i), (ii) e (iii) são satisfeitas com  $\lambda_s = 1$ ,  $\lambda_i = 0$ , para  $i \neq s$  e  $\mu_j = 0$ , para todo  $j = 1, \dots, q$ .

Se os vetores  $\nabla H_j(y^o)$ , com  $j = 1, \dots, q$ , são linearmente dependentes: existem escalares  $\gamma_j$ ,  $j = 1, \dots, q$  nem todos nulos tais que

$$\sum_{j=1}^q \gamma_j \nabla H_j(y^o) = 0,$$

então tomemos  $\lambda_i = 0$ ,  $i = 0, \dots, r$  e  $\mu_j = \gamma_j$ ,  $j = 1, \dots, q$ , de onde verificamos (i), (ii) e (iii).

**2º passo (Casos não triviais)**

Consideremos de acordo com o capítulo anterior,

$$Q_i = \{y \in \mathbb{R}^m : G_i(y) \leq 0\}, \quad \forall i = 1, \dots, r \quad \text{e}$$

$$Q_{r+1} = \{y \in \mathbb{R}^m : H_j(y) = 0, \quad \forall j = 1, \dots, q\}.$$

Então (PNL) pode ser colocado, sob a forma geral de problemas extremais, equivalentemente a  $(\tilde{P})$ :

$$\min_{y \in V} F(y)$$

s.a

$$y \in Q = \bigcap_{i=1}^{r+1} Q_i.$$

Investiguemos o cone de direções de descida associado a  $F$ . Como  $F$  é diferenciável em  $y^\circ$ , pelo teorema 3.2.2, com  $\nabla F(y^\circ) \neq 0$ ,

$$\begin{aligned} K_D &= \{d \in \mathbb{R}^m : \nabla F(y^\circ) \cdot d < 0\} \text{ e} \\ K_D^* &= \{-\lambda_0 \nabla F(y^\circ) : \lambda_0 \geq 0\}. \end{aligned} \quad (4.1)$$

Agora, definamos  $K_i$  o cone de direções factíveis para  $Q_i$ ,  $i = 1, \dots, r$ . Se  $\nabla G_i(y^\circ) \neq 0$  e  $G_i(y^\circ) = 0$ , para fixado  $i = 1, \dots, r$ , então pelos teoremas 3.2.3 e 3.2.2,

$$\begin{aligned} K_i &= \{d \in \mathbb{R}^m : \nabla G_i(y^\circ) \cdot d < 0\} \text{ e} \\ K_i^* &= \{-\lambda_i \nabla G_i(y^\circ) : \lambda_i \geq 0\} \end{aligned} \quad (4.2)$$

Queremos usufruir do teorema 3.2.6 de Lyusternik, para isto definamos a aplicação  $H : V \rightarrow \mathbb{R}^q$ , tal que  $H(y) = (H_1(y), \dots, H_q(y))$ . Devido aos fatos:  $y^\circ$  é ótimo local,  $H_j$  é estritamente diferenciável em  $y^\circ$  e para todo  $j = 1, \dots, q$ , os vetores  $\nabla H_j(y^\circ)$  são linearmente independentes; teremos as seguintes afirmações respectivas:

a)  $H(y^\circ) = (H_1(y^\circ), \dots, H_q(y^\circ)) = \underbrace{(0, \dots, 0)}_{q \text{ vezes}};$

b)  $H$  é estritamente diferenciável;

c) A matriz Jacobiana  $J(H'(y^\circ))$  tem posto  $q$ . De outra forma dito, a aplicação  $H'(y^\circ) : \mathbb{R}^m \rightarrow \mathbb{R}^q$  é sobrejetora, isto é, existe  $d \in \mathbb{R}^m$ , tal que  $H'(y^\circ)d = v, \forall v \in \mathbb{R}^q$ .

A partir disso, podemos exibir  $T_{y^\circ}Q_{r+1}$ , o cone tangente a  $Q_{r+1}$  em  $y^\circ$ , como a seguir:

$$\begin{aligned} T_{y^\circ}Q_{r+1} &= \text{Ker}H'(y^\circ) = \{d \in \mathbb{R}^m : H'(y^\circ)d = 0\} = \\ &= \{d \in \mathbb{R}^m : \nabla H_j(y^\circ)^T \cdot d = 0, \forall j = 1, \dots, q\} \end{aligned}$$

Para simplificar, estabeleçamos  $A = H'(y^\circ)$ . Dessa maneira, o cone dual de  $T_{y^\circ}Q_{r+1}$  será constituído de todos os funcionais lineares contínuos  $f \in (\mathbb{R}^m)^* = \mathbb{R}^m$ , tais que  $f(d) = 0$ , para todo  $d \in T_{y^\circ}Q_{r+1} = \text{Ker}A$ , pois  $\text{Ker}A$  é subespaço. Logo,  $f \in (\text{Ker}A)^\perp$ . Usando a hipótese (c), segue do teorema 2.2.6 que

$$\text{Im}A^* = (\text{Ker}A)^\perp,$$

onde  $A^*$  é o operador adjunto de  $A$  e no caso em questão,  $A^*$  é a transposta da matriz jacobiana  $J(A)$ . Portanto,  $f \in \text{Im}A^*$ , ou seja,  $f = A^*(-z^*)$ , para  $z^* \in (\mathbb{R}^q)^*$  fixo. Conseqüentemente,  $\forall d \in \mathbb{R}^m$  e  $z^* \in (\mathbb{R}^q)^*$  fixo, temos que  $f \in (\mathbb{R}^m)^*$  pode ser representado como

$$f(d) = [A^*(-z^*)]^T \cdot d = -(z^*)^T \cdot Ad = -(z^*)^T \cdot H'(y^\circ)d.$$

Desta forma,

$$(T_{y^\circ}Q_{r+1})^* = \left\{ - \sum_{j=1}^q \mu_j \nabla H_j(y^\circ) : \mu_j \in \mathbb{R}, j = 1, \dots, q \right\}. \quad (4.3)$$

Através de (4.1), (4.2) e (4.3) conseguiremos a existência dos multiplicadores  $\lambda_i \geq 0$ , para  $i = 0, \dots, r$ , e  $\mu_j \in \mathbb{R}$ , para  $j = 1, \dots, q$ , nem todos nulos, tais que (ii) e (iii) são verdadeiras, e por causa da aplicação do teorema 3.1.1 página 28, é válida a igualdade (i). ■

### 4.1.3 Interpretações geométricas

Faremos uma interpretação geométrica da condição necessária de Fritz-John para mínimos locais, supondo que  $\lambda_0 > 0$ ; que o domínio de restrições  $Q$  de (PNL) está definido em  $\mathbb{R}^2$  somente por restrições de desigualdades, e definido em  $\mathbb{R}^3$  por restrições de igualdades.

Suporemos também que os vetores gradientes:

$\nabla H_1(y^o), \dots, \nabla H_q(y^o)$  e  $\nabla G_i(y^o)$ , para todo  $i \in \tilde{I}(y^o)$ ; são linearmente independentes, o que quer dizer que  $y^o \in Q$  é um ponto regular das restrições.

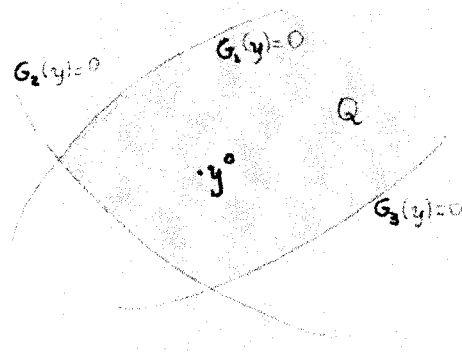


Figura4.1:  $y^o$  é ponto interior ao conjunto das restrições  $G_1$ ,  $G_2$  e  $G_3$ .

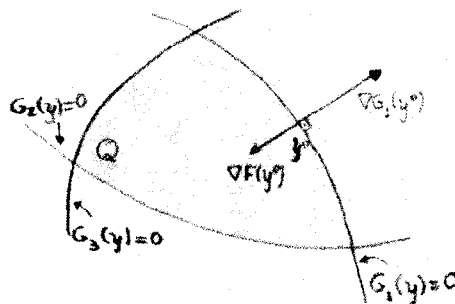


Figura4.2:  $y^o$  pertence a fronteira da restrição  $G_1$ .



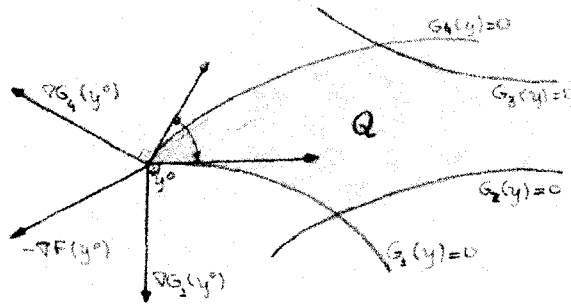


Figura 4.3:  $y^o$  pertence a interseção da fronteira da restrição  $G_1$  com  $G_4$ .

Nas figuras (4.1), (4.2) e (4.3), esboçamos as equipotenciais (curvas de nível)  $E_0(G_i) = \{y \in \mathbb{R}^2 : G_i(y) = 0\}$   $i = 1, 2, 3$  e/ou 4 associadas a cada uma das restrições  $G_i$ . O conjunto de índices das restrições ativas em  $y^o$  é dado respectivamente por  $\tilde{I}(y^o) = \emptyset$ ,  $\tilde{I}(y^o) = \{1\}$ ,  $\tilde{I}(y^o) = \{1, 4\}$ .

Os gradientes  $\nabla G_i(y^o)$  são naturalmente ortogonais a seus respectivos equipotenciais  $E_0(G_i)$  no ponto  $y^o$ , e apontam na direção para onde  $G_i(y)$  é positiva (direção de máximo crescimento de  $G_i(y)$  a partir de  $y^o$ , onde  $G_i(y^o) = 0$ ).

Se  $y^o$  é mínimo local do problema, na primeira situação teremos

$$\nabla F(y^o) = 0.$$

Pois como  $y^o$  é ponto interior de  $Q$ , o cone tangente e factível é o próprio  $\mathbb{R}^2$  e conseqüentemente o cone dual tangente e factível é o funcional nulo  $(0, 0)$ . Pelo teorema 3.1.1, existem funcionais nem todos nulos, tais que  $\lambda_0 \nabla F(y^o) = 0$ . Tomando  $\lambda_0 = 1$ , segue a igualdade acima.

Na segunda situação, existe  $\lambda_1 \geq 0$  tal que

$$\nabla F(y^o) = -\lambda_1 \nabla G_1(y^o).$$

Finalmente na terceira, o teorema 4.1.1 (denominado **Teorema de Karush-Kuhn-Tucker**<sup>2</sup>) garante a existência de  $\lambda_1$  e  $\lambda_4 \geq 0$ , tais que

$$-\nabla F(y^o) = \lambda_1 \nabla G_1(y^o) + \lambda_4 \nabla G_4(y^o),$$

<sup>2</sup>A generalização do Teorema de Karush-Kuhn-Tucker pode ser encontrada em Zeidler [40, p. 417-422]. Sugerimos consultar Contesse [8, p. 39], para o caso em que o espaço é o  $\mathbb{R}^m$ .

ou seja,  $-\nabla F(y^o)$  pode ser expressado como combinação linear, com coeficientes não negativos, dos gradientes das restrições ativas em  $y^o$ . E esta condição é equivalente a projetar  $\nabla F(y^o)$  sobre o cone tangente a  $Q$  em  $y^o$  e obter o vértice deste cone, ou melhor, o vetor nulo de  $\mathbb{R}^2$  (poderíamos substituir  $\mathbb{R}^2$  por  $\mathbb{R}^m$  e a projeção seria euclidiana). Portanto, a condição de Karush-Kuhn-Tucker pode ser considerada como uma condição de equilíbrio segundo a qual  $-\nabla F(y^o)$  (força aplicada a uma partícula situada no ponto  $y^o \in Q$ ) não admite componentes tangenciais, sobre o domínio  $Q$ , que tirem a partícula de sua posição  $y^o$  levando em consideração as restrições de deslocamento para esta partícula (ou melhor, as restrições ativas em  $y^o$ ).

Neste momento, passemos a interpretar o caso em que as restrições do problema (PNL) são de igualdades - as condições necessárias de otimalidade são conhecidas como **regra dos multiplicadores de Lagrange**.

O domínio de restrição (de factibilidade)  $Q$  de (PNL) é uma certa hiper-superfície em  $\mathbb{R}^m$  (no caso ilustrado a seguir, é uma superfície em  $\mathbb{R}^3$ ).

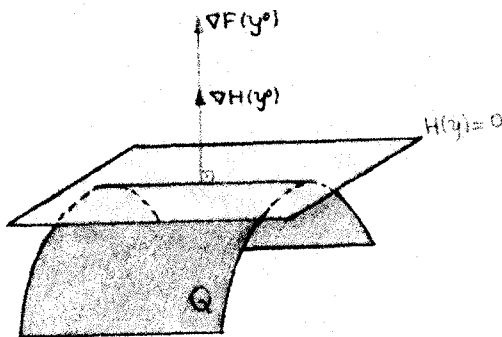


Figura4.4:  $y^o$  pertence a restrição de igualdade

Se o ponto  $y^o$ , pertencente a esta hiper-superfície, é um mínimo local de (PNL), a condição de Lagrange <sup>3</sup> garante que o gradiente da função objetivo avaliado em  $y^o$ ,  $\nabla F(y^o)$ , pode ser expressado como combinação linear dos gradientes das restrições de igualdades avaliados em  $y^o$  :

$$\nabla F(y^o) = \mu_1 \nabla H_1(y^o) + \dots + \mu_q \nabla H_q(y^o).$$

<sup>3</sup>Consulte em Contesse [8, p. 56].

Os escalares  $\mu_1, \dots, \mu_q$  são usualmente chamados **multiplicadores de Lagrange**.

Esta regra é equivalente a que o gradiente da função objetivo avaliado em  $y^\circ$  pertencá ao subespaço ortogonal ao subespaço tangente à superfície  $Q$  em  $y^\circ$  (na figura (4.4), o subespaço tangente é o semi-espaço que contém  $Q$  e é determinado pelo plano tangente a  $Q$  em  $y^\circ$ ). Esta condição é ainda equivalente a que o vetor nulo de  $\mathbb{R}^m$  seja a projeção euclidiana do  $\nabla F(y^\circ)$  sobre o subespaço tangente a  $Q$  em  $y^\circ$ .

## 4.2 Problema de controle ótimo

### 4.2.1 Considerações sobre o problema

No espaço  $KC^1(\Delta, \mathbb{R}^n) \times L_\infty(\Delta, \mathbb{R}^r) \times \mathbb{R}^{2-4}$ , chamamos **problema de controle ótimo** o seguinte problema (PC):

$$\min \mathcal{B}_0(x(\cdot), u(\cdot), t_0, t_1); \quad (4.4)$$

$$\dot{x}(t) = \varphi(t, x(\cdot), u(\cdot), t_0, t_1), \quad (4.5)$$

$$u(t) \in \mathcal{U}, \forall t \in [t_0, t_1], \quad (4.6)$$

---

<sup>4</sup>†  $KC^1(\Delta, \mathbb{R}^n)$  denota o espaço das funções de  $\Delta$  em  $\mathbb{R}^n$  cuja derivada é contínua por partes no intervalo  $\Delta$ , ou seja, a derivada apresenta um número finito de descontinuidades de primeira espécie (pontos de descontinuidade possuem os dois limites - à esquerda e à direita).

‡  $L_\infty(\Delta, \mathbb{R}^r)$  é o espaço das funções  $u = (u_1, u_2, \dots, u_r) : \Delta \rightarrow \mathbb{R}^r$  mensuráveis segundo Lebesgue, essencialmente limitadas, definidas sobre o intervalo compacto  $\Delta \subseteq \mathbb{R}$ , com a norma  $\|\cdot\|_\infty$  definida por

$$\|u\|_\infty = \max_{1 \leq j \leq r} \text{ess sup}_{t \in \Delta} |u_j(t)|$$

onde  $\text{ess sup}$  denota o menor número  $\lambda$  tal que  $|u_j(t)| \leq \lambda$  para quase todo  $t \in \Delta$ .

Às vezes, durante o decorrer da dissertação, usaremos a notação  $L_\infty^r(\Delta)$  ou  $(L_\infty(\Delta))^r$ .

$$\mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) \leq 0, \quad i = 1, \dots, m', \quad (4.7)$$

$$\mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) = 0, \quad i = m' + 1, \dots, m, \quad \text{onde} \quad (4.8)$$

$$\mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) = \int_{t_0}^{t_1} \mathbb{L}_i(t, x(t), u(t)) dt + \psi_i(t_0, x(t_0), t_1, x(t_1));$$

$$i = 0, \dots, m.$$

Aqui é dado um intervalo  $\Delta$  finito;  $t_0, t_1 \in \Delta$ ;  $\mathbb{L}_i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$  são funções de  $n + r + 1$  variáveis;  $\psi_i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$  são funções de  $2n + 2$  variáveis;  $\varphi : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$  é uma função vetorial de  $n + r + 1$  variáveis;  $\mathcal{U}$  é um conjunto arbitrário tomado em  $\mathbb{R}^r$ .

Um caso particular do problema (PC) é apresentado quando  $t_0$  e/ou  $t_1$  são fixos.

A função  $\mathcal{B}_0(x(\cdot), u(\cdot), t_0, t_1)$  é chamada **função objetivo**. Já a função vetorial  $x(\cdot)$  é denominada **variável de estado** (*trajetória de fase*) e a função  $u(\cdot)$  é denominada **controle**. A equação (4.5), a qual chamamos **equação de movimento** (*equação controlada; relação diferencial ou restrição em forma de solução*) deve ser satisfeita em todos os pontos de continuidade do controle  $u(\cdot)$  sobre o intervalo  $[t_0, t_1]$  (designaremos por  $\tilde{T}$  o intervalo  $[t_0, t_1]$ , apenas para simplificar). A condição (4.6) é chamada **restrição de controle** (*restrição do tipo inclusão*).  $\mathcal{U}$  se chama **domínio de controle**. A quádrupla  $(x(\cdot), u(\cdot), t_0, t_1)$  no problema de controle ótimo é denominada **processo controlável**, se  $x(\cdot) \in KC^1(\Delta, \mathbb{R}^n)$ ,  $u(\cdot) \in L_\infty(\Delta, \mathbb{R}^r)$  e se são satisfeitas a equação de movimento (4.5) e a restrição de controle (4.6). Um **processo controlável** é considerado **admissível** se também são verificáveis as relações (4.7) e (4.8). Dizemos que um processo controlável admissível  $\xi^o = (x^o(\cdot), u^o(\cdot), t_0^o, t_1^o)$  é ótimo se existe  $\delta > 0$ , tal que a desigualdade  $\mathcal{B}_0(\xi) \geq \mathcal{B}_0(\xi^o)$  é satisfeita para qualquer processo admissível  $\xi = (x(\cdot), u(\cdot), t_0, t_1)$ , para o qual verifica-se

$$\|(x(\cdot), t_0, t_1) - (x^o(\cdot), t_0^o, t_1^o)\|_{C^n(\Delta) \times \mathbb{R}^2} < \delta \quad (\text{para } t \in \tilde{T} \cap [t_0^o, t_1^o]) \quad ^5.$$

Em decorrência disso, dizemos que  $x^o(\cdot)$  é o estado ótimo e  $u^o(\cdot)$  é o controle ótimo do problema (PC).

<sup>5</sup> Usaremos também a notação  $C^m(a, b)$  ou  $(C(a, b))^n$  para o espaço  $C((a, b), \mathbb{R}^n)$ , das funções contínuas de  $(a, b)$  em  $\mathbb{R}^n$ .

Quando definimos o conjunto  $A$  como o conjunto formado pelos processos admissíveis  $\xi = (x(\cdot), u(\cdot), t_0, t_1)$ , o problema (PC) representará um **programa matemático** (PMT):

$$\min \{B_0(\xi) : \xi \in A\}.$$

**Observação 4.2.1** *O tempo inicial  $t_0$  (respectivamente final  $t_1$ ) bem como o estado inicial  $x_0$  do processo de controle são variáveis.*

**Observação 4.2.2** *O conjunto  $\mathcal{U}$  pode ser de diversas naturezas, isto é, ele pode ser um conjunto discreto. Por esta razão não é natural considerar somente controles contínuos ou diferenciáveis nos problemas de controle.*

**Observação 4.2.3** *Uma hipótese comum nos problemas de controle ótimo é a continuidade por partes em todas as variáveis e suavidade (no sentido de existir diferenciabilidade pelo menos de primeira ordem) com respeito as variáveis  $t$  e  $x$ .*

O problema (PC), que descrevemos anteriormente, chama-se **Problema de Bolza** (PB). Com frequência, a função objetivo  $B_0(x(\cdot), u(\cdot), t_0, t_1)$  tem a seguinte forma:

$$B_0(x(\cdot), u(\cdot), t_0, t_1) = \int_{t_0}^{t_1} \mathbb{L}(t, x(t), u(t)) dt.$$

Neste caso chamaremos (PC) de **problema de Lagrange** (PLG). Se a função objetivo  $B_0(x(\cdot), u(\cdot), t_0, t_1)$  é dada por

$$B_0(x(\cdot), u(\cdot), t_0, t_1) = \psi_0(t_0, x(t_0), t_1, x(t_1))$$

teremos um **Problema de Mayer** (PMY).

Cada um destes problemas pode ser transformado nos outros dois, fazendo algumas modificações na forma com que o problema é apresentado. Por isto os resultados obtidos para problemas na forma de Bolza, por exemplo, são válidos para os respectivos problemas de Lagrange e de Mayer.

Mostraremos, a título de exemplificação, que (PB)  $\Rightarrow$  (PMY) e (PB)  $\Rightarrow$  (PLG):

Fixemos  $\ell = (t_0, x(t_0), t_1, x(t_1))$ , para simplificar nossa exposição.

- (PB)  $\Rightarrow$ (PMY)

Considerando um problema de Bolza como na página 77, o problema de Mayer equivalente será dado por:

$$\min \mathcal{B}_0(x(\cdot), u(\cdot), t_0, t_1) = \min[x_{n+1}(t_1) + \psi_0(\ell)]$$

s.a

$$\begin{aligned} \dot{x}(t) &= \varphi(t, x(\cdot), u(\cdot), t_0, t_1), \\ \dot{x}_{n+1}(t) &= \mathbb{L}(t, x(t), u(t)), \\ u(t) &\in \mathcal{U}, \forall t \in \tilde{T}, \\ \mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) &\leq 0, \quad i = 1, \dots, m', \\ \mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) &= 0, \quad i = m' + 1, \dots, m, \\ x_{n+1}(t_0) &= 0; \end{aligned}$$

- (PB)  $\Rightarrow$ (PLG)

Considerando que, para o problema de Bolza, queremos

$$\min \int_{t_0}^{t_1} \mathbb{L}(t, x(t), u(t)) dt + \psi_0(\ell).$$

Para transformá-lo em um problema de Lagrange acrescentamos uma igualdade específica. Isto é o que mostra a forma de Lagrange abaixo apresentada:

$$\min \mathcal{B}_0(x(\cdot), u(\cdot), t_0, t_1) = \min \int_{t_0}^{t_1} \mathbb{L}(t, x(t), u(t)) dt$$

s.a

$$\begin{aligned} \dot{x}(t) &= \varphi(t, x(\cdot), u(\cdot), t_0, t_1), \\ u(t) &\in \mathcal{U}, \forall t \in \tilde{T}, \\ \mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) &\leq 0, \quad i = 1, \dots, m', \\ \mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) &= 0, \quad i = m' + 1, \dots, m, \\ \mathcal{B}_{m+1}(x(\cdot), u(\cdot), t_0, t_1) &= \psi_0(\ell) = 0; \end{aligned}$$

Maiores detalhes são comentados em Blum [10, p. 62 e 63].

## 4.2.2 Princípio do Máximo Local de Pontryagin

Trabalharemos, de maneira particular, com o seguinte problema de controle, salvo algumas exigências impostas imediatamente:

Encontrar  $x : [0, T] \longrightarrow \mathbb{R}^n$ ,  $u : [0, T] \longrightarrow \mathbb{R}^r$  tais que satisfaçam a seguinte equação diferencial:

$$\frac{d}{dt}x(t) = \varphi(x(t), u(t), t),$$

com as condições de contorno

$$\begin{aligned}x(0) &= c \\x(T) &= d,\end{aligned}$$

onde  $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times [0, T] \longrightarrow \mathbb{R}^n$ , de modo que minimizem o funcional integral

$$\int_0^T \mathbb{L}(x(t), u(t), t) dt,$$

cujos controle satisfaz restrições do tipo

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

com  $\mathbb{L} : \mathbb{R}^n \times \mathbb{R}^r \times [0, T] \longrightarrow \mathbb{R}$ .

Como conjunto de controles admissíveis tomamos a classe de todas as funções mensuráveis limitadas (isto é,  $u \in (L_\infty(0, T))^r$ ). Em vez da equação (4.10) consideramos a equação integral equivalente, ou seja, a solução será um par  $(x, u) \in (C[0, T])^n \times (L_\infty(0, T))^r$  satisfazendo a equação integral

$$x(t) = c + \int_0^t \varphi(x(\tau), u(\tau), \tau) d\tau.$$

O problema que acabamos de apresentar pode ser reescrito da seguinte forma:

**Problema (P):**

$$\min_{(x,u) \in C^n[0,T] \times L_\infty^r(0,T)} \int_0^T \mathbb{L}(x(t), u(t), t) dt, \quad (4.9)$$

s.a.

$$\frac{d}{dt}x(t) = \varphi(x(t), u(t), t); \quad (4.10)$$

$$x(0) = c; \quad (4.11)$$

$$x(T) = d; \quad (4.12)$$

$$u(t) \in M, \text{ q.t.p. } 0 \leq t \leq T. \quad (4.13)$$

**Teorema 4.2.4 (Princípio do Máximo Local de Pontryagin)** *Sejam  $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times [0, T] \rightarrow \mathbb{R}^n$  e  $\mathbb{L} : \mathbb{R}^n \times \mathbb{R}^r \times [0, T] \rightarrow \mathbb{R}$  tais que  $\varphi(x(t), u(t), t) \in \mathbb{L}(x(t), u(t), t)$  são funções contínuas em  $x$  e  $u$ , mensuráveis em  $t$ , continuamente diferenciáveis com respeito a  $x$  e  $u$ . Além disso, sejam  $\varphi_x(x(t), u(t), t)$ ,  $\varphi_u(x(t), u(t), t)$ <sup>6</sup>,  $\mathbb{L}_x(x(t), u(t), t)$ ,  $\mathbb{L}_u(x(t), u(t), t)$  funções limitadas para todas  $x$  e  $u$  limitadas. Seja  $M \subseteq \mathbb{R}^r$  convexo e fechado, tal que  $\text{int } M \neq \emptyset$ . Suponha que  $(x^o(t), u^o(t))$  é uma solução do problema (P). Então existem um número  $\lambda_0 \geq 0$  e uma função  $p(\cdot) : [0, T] \rightarrow \mathbb{R}^n$ , satisfazendo a equação*

$$\frac{d}{dt}p(t) = -\varphi_x^T(x^o(t), u^o(t), t)p(t) + \lambda_0 \mathbb{L}_x(x^o(t), u^o(t), t), \quad (4.14)$$

tais que  $\lambda_0$  e  $p(\cdot)$  não são simultaneamente iguais a zero, além disso

$$(-\varphi_u^T(x^o(t), u^o(t), t)p(t) + \lambda_0 \mathbb{L}_u(x^o(t), u^o(t), t))^T \cdot (u(t) - u^o(t)) \geq 0^7, \quad (4.15)$$

para quase todo  $t \in [0, T]$  e para todo  $u \in M$ .

**Prova:** Consideraremos o espaço  $E = (C[0, T])^n \times (L_\infty(0, T))^r$ , o qual abreviaremos por  $E = C \times L_\infty$ . Denotemos por

$$Q_2 = \{(x, u) \in E : \text{satisfazem (4.10), (4.11) e (4.12)}\} \text{ e}$$

$$Q_1 = \{(x, u) \in E : \text{satisfazem (4.13)}\}.$$

Nosso problema é, assim, minimizar a integral

$$F(x, u) = \int_0^T \mathbb{L}(x(t), u(t), t) dt$$

sobre o conjunto  $Q = Q_1 \cap Q_2$ .

---

<sup>6</sup>  $\varphi_x(x, u, t)$  e  $\varphi_u(x, u, t)$  denotam as matrizes cujos elementos são as derivadas parciais

$$\frac{\partial \varphi_i}{\partial x_j}(x, u, t), \quad i = 1, \dots, n; \quad j = 1, \dots, n$$

$$\frac{\partial \varphi_i}{\partial u_j}(x, u, t), \quad i = 1, \dots, n; \quad j = 1, \dots, r.$$

<sup>7</sup>  $\varphi_x^T$  e  $\varphi_u^T$  denotam as transpostas das matrizes  $\varphi_x$  e  $\varphi_u$  respectivamente.



Analisaremos o problema de acordo com nosso esquema geral.

Primeiramente iremos determinar o cone associado ao funcional  $F$ , aos conjuntos  $Q_1$  e  $Q_2$ , e por fim calcularemos os respectivos cones duais.

**a) Análise do funcional  $F$ .**

Visto que  $F(x, u) = \int_0^T \mathbb{L}(x(t), u(t), t) dt$ , onde  $\mathbb{L}(x(t), u(t), t)$  é contínua em  $x, u, t$  e diferenciável com respeito a  $x, u$ ;  $\mathbb{L}_x(x(t), u(t), t)$  e  $\mathbb{L}_u(x(t), u(t), t)$  são contínuas em  $x, u, t$ , obteremos

$$F'(x^\circ, u^\circ)(\bar{x}, \bar{u}) = \int_0^T \mathbb{L}_x(x^\circ(t), u^\circ(t), t) \bar{x}(t) + \mathbb{L}_u(x^\circ(t), u^\circ(t), t) \bar{u}(t) dt.$$

Pelo teorema 3.2.2, o cone de direções de descida é dado por

$$K_0 = \{(\bar{x}, \bar{u}) \in E : F'(x^\circ, u^\circ)(\bar{x}, \bar{u}) < 0\}.$$

Se  $K_0 \neq \emptyset$ , então para qualquer  $f_0 \in K_0^*$ ,

$$f_0(\bar{x}, \bar{u}) = -\lambda_0 \int_0^T [\mathbb{L}_x(x^\circ(t), u^\circ(t), t) \bar{x} + \mathbb{L}_u(x^\circ(t), u^\circ(t), t) \bar{u}] dt; \lambda_0 \geq 0, \tag{4.16}$$

ou seja,  $K^* = \{-\lambda_0 F'(x^\circ, u^\circ) : \lambda_0 \geq 0\}$ .

**b) Análise da restrição em  $Q_1$ .**

O conjunto  $U = \{u \in L_\infty^r(0, T) : u(t) \in M \subset \mathbb{R}^r, \text{ q.t.p. } t \in [0, T]\}$  é convexo fechado no espaço  $L_\infty$ . Também,  $\text{int } U \neq \emptyset$ , devido ao fato de  $u^\circ(t) \in M$  e  $\text{int } M \neq \emptyset$ . Além disso, o conjunto  $Q_1 = C \times U$  é fechado e convexo em  $E$ , e

$$\text{int } Q_1 = C \times \text{int } U \neq \emptyset.$$

Seja  $K_1$  o cone de direções factíveis para  $Q_1$  no ponto  $(x^\circ, u^\circ)$ . Então, se  $f_1 \in K_1^*$ , segue-se que  $f_1 = (0, \tilde{f}_1)$ , onde  $\tilde{f}_1 \in L_\infty^*$  é um suporte para  $U$  no ponto  $u^\circ$  (conforme os teoremas 2.1.5 e ??).

**c) Análise das restrições em  $Q_2$ .**

Queremos determinar o subespaço  $K_2$  tangente a

$$Q_2 = \left\{ (x, u) \in E : \begin{aligned} x(t) &= c + \int_0^t \varphi(x(t), u(t), \tau) d\tau, \quad \forall 0 \leq t \leq T, \quad c \\ x(T) &= d \in \mathbb{R}^n \end{aligned} \right\}$$

no ponto  $(x^o, u^o)$ . Para isto, definamos o operador  $P : E \rightarrow E_1$ , onde  $E_1 = C^n[0, T] \times \mathbb{R}^n$  e

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

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

De acordo com as hipóteses relativas às derivadas  $\varphi_x(x, u, \tau)$  e  $\varphi_u(x, u, \tau)$  teremos que

$$P(x^o + \bar{x}, u^o + \bar{u}) - P(x^o, u^o) = \left( \bar{x}(t) - \int_0^t [\varphi_x(x^o, u^o, \tau)\bar{x}(\tau) + \varphi_u(x^o, u^o, \tau)\bar{u}(\tau)] d\tau, \bar{x}(T) \right) + \delta,$$

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

$\left(\bar{x}(t) - \int_0^t [\varphi_x(x^o, u^o, \tau)\bar{x}(\tau) + \varphi_u(x^o, u^o, \tau)\bar{u}(\tau)] d\tau, \bar{x}(T)\right)$  é um operador linear de  $E$  em  $E_1$ . Logo,  $P$  é Fréchet-diferenciável (veja a definição 3.2.1), em  $(x^o, u^o)$  e temos que

$$P'(x^o, u^o)(\bar{x}, \bar{u}) = \left( \bar{x}(t) - \int_0^t \varphi_x(x^o, u^o, \tau)\bar{x}(\tau) + \varphi_u(x^o, u^o, \tau)\bar{u}(\tau) d\tau, \bar{x}(T) \right).$$

Ainda mais,  $P'(x^o, u^o)$  é contínua numa vizinhança de  $(x^o, u^o)$ . Portanto,  $P'(x^o, u^o)$  é estritamente diferenciável (usamos resultados de análise diferencial, os quais garantem que se um operador é continuamente diferenciável, ele é estritamente diferenciável).

Desejamos determinar quando esse operador linear é sobrejetivo, pois assim conseguiremos exibir o cone  $K_2$ .

Afirmamos que, se a condição de não degeneração  $\mathbb{A}$ , definida na página 63 depois desta prova, é satisfeita para o sistema

$$\begin{cases} \frac{d}{dt}\bar{x}(t) = A(t)\bar{x}(t) + B(t)\bar{u}(t); \\ \bar{x}(0) = 0, \end{cases} \quad (4.17)$$

com  $A(t) = \varphi_x(x^o(t), u^o(t), t)$ ,  $B(t) = \varphi_u(x^o(t), u^o(t), t)$ , então  $P'(x^o, u^o)$  é sobrejetor, ou seja,  $P'(x^o, u^o)E = E_1$ .

Por conseguinte, todas as asserções do teorema 3.2.6 de Lyusternik, na página 37, são satisfeitas para o operador  $P$  e assim, o subespaço tangente  $K_2$  consiste de todos os pares  $(\bar{x}, \bar{u}) \in E$  os quais satisfazem a equação integral

$$\bar{x}(t) = \int_0^t [\varphi_x(x^o, u^o, \tau)\bar{x}(\tau) + \varphi_u(x^o, u^o, \tau)\bar{u}(\tau)] d\tau; \quad \bar{x}(T) = 0$$

ou equivalentemente, a equação diferencial linear

$$\dot{\bar{x}}(t) = \varphi_x(x^o(t), u^o(t), t)\bar{x} + \varphi_u(x^o(t), u^o(t), t)\bar{u}, \quad \bar{x}(0) = 0 \quad (4.18)$$

$$\bar{x}(T) = 0 \quad (4.19)$$

Denotemos por  $L_1 \subset E$ , o conjunto de todos os pares  $(\bar{x}, \bar{u})$  satisfazendo a equação (4.18), e  $L_2 \subset E$ , o conjunto dos pares satisfazendo (4.19). Então  $L_1$  e  $L_2$  são subespaços e  $K_2 = L_1 \cap L_2$ .

Definamos  $L_2^*(T) = \tilde{L}_2^* = \{(\bar{x}(T), \bar{u}(T)) \in \mathbb{R}^n \times \mathbb{R}^r : \bar{x}(T) = 0\}$ . Então, se  $\hat{f} \in L_2^*$ ,  $\hat{f}(\bar{x}, \bar{u}) = 0$ , para todo  $(\bar{x}, \bar{u}) \in L_2$ . Equivalentemente,  $\hat{f}(\bar{x}(T), \bar{u}(T)) = 0$ , para todo  $(\bar{x}, \bar{u}) \in \tilde{L}_2$ . Assim,  $\hat{f} \in \tilde{L}_2^*$  se, e somente se,  $\exists \hat{a} = (a, b) \in \mathbb{R}^{n+r}$  tal que  $\hat{f}(\bar{x}(T), \bar{u}(T)) = \hat{a} \cdot (\bar{x}(T), \bar{u}(T)) = a \cdot \bar{x}(T) + b \cdot \bar{u}(T)$ . É claro que  $b = 0$ , por isso  $\hat{f}(\bar{x}(T), \bar{u}(T)) = a \cdot \bar{x}(T)$ . Logo, se  $\hat{f} \in L_2^*$ , teremos que  $\hat{f}(\bar{x}, \bar{u}) = a \cdot \bar{x}(T)$ , com  $a \in \mathbb{R}^n$ . Desde que  $L_2$  é definido por  $n$  funcionais lineares  $\bar{x}_i$ ,  $L_2^*$  é  $n$ -dimensional. Portanto  $L_1^* + L_2^*$  é fechado, e por conseguinte,  $L_1^* + L_2^*$  é  $w^*$ -fechado. Isto decorre do seguinte teorema:

Se  $L_1$  e  $L_2$  são subespaços fechados de um espaço de Banach e  $L_2$  é finito-dimensional, então  $L_1 + L_2$  é fechado.

É decorre também do teorema (2.2.2). Pela proposição (2.1.9), segue-se que  $K_2^* = L_1^* + L_2^*$ .

Desde que  $L_1$  é um subespaço, segue-se do teorema (2.1.5) que, para qualquer  $f_3 \in L_1^*$ , nós temos  $f_3(\bar{x}, \bar{u}) = 0$ , para todo  $(\bar{x}, \bar{u})$  satisfazendo

(4.18). Finalmente, como mencionado anteriormente, se  $f_2 \in L_1^* + L_2^*$ , então  $f_2(\bar{x}, \bar{u}) = \hat{f}(\bar{x}, \bar{u}) + f_3(\bar{x}, \bar{u}) = a \cdot \bar{x}(T)$ ;  $a \in \mathbb{R}^n$ .

**d) Equação de Euler:**

Agora podemos aplicar o teorema de Dubovitskii-Milyutin (3.1.1). Portanto obteremos que existem funcionais  $f_0, f_1$  e  $f_2 \in E^*$ , nem todos nulos, tais que

$$f_0(\bar{x}, \bar{u}) + f_1(\bar{x}, \bar{u}) + f_2(\bar{x}, \bar{u}) = 0, \text{ para todo } (\bar{x}, \bar{u}) \in E, \quad (4.20)$$

onde  $f_0(\bar{x}, \bar{u})$  é dado por (4.16),  $f_1(\bar{x}, \bar{u}) = \tilde{f}_1(\bar{u})$  é suporte para  $U$  em  $u^\circ$ ,  $f_2(\bar{x}, \bar{u}) = \hat{f}(\bar{x}, \bar{u}) + f_3(\bar{x}, \bar{u})$ , com  $f_3(\bar{x}, \bar{u})$  nulo, para  $(\bar{x}, \bar{u})$  satisfazendo a equação (4.18), e  $\hat{f}(\bar{x}, \bar{u}) = a \cdot \bar{x}(T)$ ;  $a \in \mathbb{R}^n$ .

**e) Análise da equação de Euler:**

A equação (4.20) deve ser satisfeita para  $(\bar{x}, \bar{u})$  arbitrários. Considere  $\bar{u}$  qualquer, e determine  $\bar{x} = \bar{x}(\bar{u})$  da equação (4.18). Com esta escolha de  $\bar{x}$  e  $\bar{u}$ , nós temos  $f_3(\bar{x}, \bar{u}) = 0$ , assim a condição (4.20) torna-se

$$\tilde{f}_1(\bar{u}) = \lambda_0 \int_0^T [\mathbb{L}_x(x^\circ(t), u^\circ(t), t) \bar{x} + \mathbb{L}_u(x^\circ(t), u^\circ(t), t) \bar{u}] dt - a \cdot \bar{x}(T). \quad (4.21)$$

Agora transformaremos a expressão do lado direito de (4.21) de tal maneira que substituamos  $\bar{x}$  por  $\bar{u}$ . Suponhamos que  $p$  é solução do sistema (4.14) com a condição de fronteira:  $p(T) = a$ . Então,

$$\begin{aligned} & \lambda_0 \int_0^T \mathbb{L}_x(x^\circ, u^\circ, t) \bar{x} dt - a \cdot \bar{x}(T) = \\ & \int_0^T (\dot{p}(t) + \varphi_x^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{x} dt - a \cdot \bar{x}(T) = \\ & \int_0^T \dot{p}^T(t) \cdot \bar{x} dt + \int_0^T (\varphi_x^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{x} dt - a \cdot \bar{x}(T). \end{aligned}$$

Integrando por partes vem que:

$$\int_0^T \dot{p}^T(t) \cdot \bar{x} dt + \int_0^T (\varphi_x^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{x} dt - a \cdot \bar{x}(T) =$$

$$p \bar{x} \Big|_0^T - \int_0^T p^T(t) \cdot \dot{\bar{x}} dt + \int_0^T p^T(t) \cdot (\varphi_x(x^\circ, u^\circ, t)\bar{x}) dt - a \cdot \bar{x}(T).$$

Usando o fato de que  $(\bar{x}, \bar{u})$  satisfaz a equação (4.18), temos que

$$\varphi_u(x^\circ, u^\circ, t) \bar{u} = \frac{d}{dt} x(t) - \varphi_x(x^\circ, u^\circ, t)\bar{x}.$$

De onde obtemos

$$- \int_0^T p^T(t) \cdot \dot{\bar{x}} dt + \int_0^T p^T(t) \cdot (\varphi_x(x^\circ, u^\circ, t)\bar{x}) dt =$$

$$- \int_0^T p^T(t) \cdot (\varphi_u(x^\circ, u^\circ, t) \bar{u}) dt = - \int_0^T (\varphi_u^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{u} dt.$$

Assim,

$$\lambda_0 \int_0^T \mathbb{L}_x(x^\circ, u^\circ, t) \bar{x} dt - a \cdot \bar{x}(T) = - \int_0^T (\varphi_u^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{u} dt.$$

Portanto a condição (4.20) torna-se

$$\tilde{f}_1(\bar{u}) = \int_0^T (-\varphi_u^T(x^\circ, u^\circ, t)p(t) + \lambda_0 \mathbb{L}_u(x^\circ, u^\circ, t))^T \cdot \bar{u} dt,$$

onde  $\bar{u}$  é arbitrário e  $\tilde{f}_1$  é um suporte para  $U$  no ponto  $u^\circ$ . Conseqüentemente,

$$(-\varphi_u^T(x^\circ, u^\circ, t)p(t) + \lambda_0 \mathbb{L}_u(x^\circ, u^\circ, t))^T \cdot (u(t) - u^\circ(t)) \geq 0,$$

para quase todo  $0 \leq t \leq T$  e para todo  $u(t) \in M$ . Isto decorre do seguinte fato:

Se o funcional linear  $f$  definido por

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

é um suporte para  $U$  no ponto  $u^\circ$ , com  $g(t) \in L_1^r(0, T)$ <sup>8</sup>. então  $g(t) \cdot (u - u^\circ)(t) \geq 0$ , para todo  $u(t) \in M$  e quase todo  $t \in [0, T]$  (i.e., para quase todo  $t \in [0, T]$ , o vetor  $g(t) \in \mathbb{R}^r$  é um suporte para  $M$  no ponto  $u^\circ(t)$ ).

De fato, suponhamos que a afirmação é falsa, isto é, existe um subconjunto  $R_1 \subset [0, T]$ ,  $\mu(R_1) \neq 0$  (onde  $\mu(R_1)$  denota a medida de Lebesgue<sup>9</sup> de  $R_1$ ), tal que para todo  $t \in R_1$ , existe  $\tilde{u}(t) \in M$  de modo que  $g(t) \cdot (\tilde{u} - u^\circ)(t) < 0$ . De acordo com Lusin, dado um intervalo  $J$ , se uma função  $\Upsilon : J \rightarrow \mathbb{R}^r$  é mensurável, então, dado  $\varepsilon > 0$ , existe um conjunto fechado  $J_\varepsilon \subset J$  tal que  $J \setminus J_\varepsilon$  tem medida  $\leq \varepsilon$  e  $\Upsilon$  é contínua em  $J_\varepsilon$ <sup>10</sup>. 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  $[0, T] - R_2$ ,  $u^\circ$  é contínua em  $[0, T] - R_3$ . 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, já que  $g$  e  $u^\circ$  são contínuas em  $t_0$  e  $g(t_0) \cdot (\tilde{u}(t_0) - u^\circ(t_0)) = \gamma < 0$ , existe  $R_4 \subset [0, T]$ ,  $\mu(R_4) > 0$ , tal que  $g(t) \cdot (\tilde{u}(t_0) - u^\circ(t)) \leq \frac{\gamma}{2}$ , para todo  $t \in R_4$ .

Consideremos a função

$$u_1(t) = \begin{cases} \tilde{u}(t_0), & \text{se } t \in R_4 \\ u^\circ(t), & \text{se } t \in [0, T] - R_4. \end{cases}$$

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

$$\begin{aligned} f(u_1) &= \int_0^T g(t) \cdot u^\circ(t) dt + \int_{R_4} g(t) \cdot (\tilde{u}(t_0) - u^\circ(t)) dt \leq \\ &\leq f(u^\circ) + \frac{\gamma}{2} \mu(R_4), \text{ ou seja, } f(u_1) < f(u^\circ). \end{aligned}$$

O que contradiz o fato de  $f$  ser um funcional suporte para  $U$  no ponto  $u^\circ$ .

Sob estas suposições, o caso  $\lambda_0 = 0$ ,  $p(t) = 0$  não pode ocorrer, desde que deveríamos ter  $f_0(\bar{x}, \bar{u}) = 0 = \lambda_0$ ,  $a = p(T) = 0$ . Mas assim  $\tilde{f}_1(\bar{u}) = 0$ ,

<sup>8</sup>  $L_1^r(0, T)$  é o espaço das funções  $g(t) = (g_1(t), \dots, g_r(t))$  definidas de  $(0, T)$  em  $\mathbb{R}^r$ , tais que  $\int_0^T |g(t)| dt < +\infty$ .

<sup>9</sup> Lebesgue matemático francês.

No que diz respeito a teoria da medida, em especial medida de Lebesgue consulte Bartle [7], Fernandez [20], Halmos [25] e Yosida [39, 1978].

<sup>10</sup> Para melhores esclarecimentos sobre o **Teorema de Lusin**, consulte Deimling [17, p. 103], Edwards [19, p. 151-153] e Kolmogorov [28, p. 281].

por causa de (4.21), de onde  $f_0 = 0$ ,  $f_1 = 0$ ,  $\dot{f} = 0$ . A equação (4.20) deveria implicar também que  $f_2 = 0$ , o qual contradiz a hipótese de que existem funcionais  $f_i$ ,  $i = 0, 1, 2$ , nem todos nulos.

### f) Análise dos casos excepcionais - casos triviais:

Iremos analisar o caso em que o cone  $K_0$  de direções de descida do funcional  $F$  no ponto  $(x^\circ, u^\circ)$  é vazio. Em seguida, vamos assumir que o sistema (4.18) é degenerado. Portanto, mostraremos que aquelas afirmações contrárias feitas no decorrer da prova são supérfluas.

De fato, se  $K_0 \equiv \emptyset$ , então

$$\int_0^T \mathbb{L}_x(x^\circ(t), u^\circ(t), t) \bar{x} + \mathbb{L}_u(x^\circ(t), u^\circ(t), t) \bar{u} dt = 0 \quad (4.22)$$

para todo  $(\bar{x}, \bar{u})$ . Tomemos  $p$  solução do sistema (4.14), com  $p(T) = 0$  e  $\lambda_0 = 1$ .

Visto que

$$\begin{aligned} \int_0^T \mathbb{L}_x(x^\circ, u^\circ, t) \bar{x} dt &= \int_0^T (\dot{p}(t) + \varphi_x^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{x} dt = \\ &= \int_0^T \dot{p}^T(t) \cdot \bar{x} dt + \int_0^T (\varphi_x^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{x} dt = \\ &= p\bar{x}|_0^T - \int_0^T p^T(t) \cdot \bar{x} dt + \int_0^T p^T(t) \cdot \varphi_x(x^\circ, u^\circ, t)\bar{x} dt \end{aligned}$$

Novamente, pelo fato de  $(\bar{x}, \bar{u})$  satisfazer a equação (4.18), temos que

$$\varphi_u(x^\circ, u^\circ, t) \bar{u} = \frac{d}{dt} x(t) - \varphi_x(x^\circ, u^\circ, t)\bar{x}.$$

Portanto,

$$\begin{aligned} & - \int_0^T [p^T(t) \cdot \dot{\bar{x}} - p^T(t) \cdot (\varphi_x(x^\circ, u^\circ, t)\bar{x})] dt \\ &= - \int_0^T p^T(t) \cdot (\varphi_u(x^\circ, u^\circ, t) \bar{u}) dt \\ &= - \int_0^T (\varphi_u^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{u} dt. \end{aligned}$$

Por conseguinte,

$$\int_0^T \mathbb{L}_x(x^\circ, u^\circ, t) \bar{x} dt = - \int_0^T (\varphi_u^T(x^\circ, u^\circ, t)p(t))^T \cdot \bar{u} dt.$$

Assim,

$$\int_0^T (-\varphi_u^T(x^\circ, u^\circ, t)p(t) + \mathbb{L}_u(x^\circ(t), u^\circ(t), t))^T \cdot \bar{u} dt = 0,$$

para todo  $\bar{u}$ , por (4.22). Logo,  $-\varphi_u^T(x^\circ, u^\circ, t)p(t) + \mathbb{L}_u(x^\circ(t), u^\circ(t), t) = 0$  para quase todo  $t \in [0, T]$ . Verificamos desta forma a desigualdade (4.15).

Agora, digamos que o sistema (4.18) é degenerado, então existe uma solução não nula  $p(\cdot)$ , da equação (4.14), com  $\lambda_0 = 0$ , tal que

$$-\varphi_u^T(x^\circ, u^\circ, t)p(t) \equiv 0.$$

O que nos mostra a veracidade da inequação (4.15) onde  $\lambda_0 = 0$ . ■

#### Observação 4.2.5 Definamos a função de Pontryagin $\mathbf{H}$ :

$$\mathbf{H}(x, u, p, t) = \varphi^T(x, u, t) \cdot p(t) - \lambda_0 \mathbb{L}(x(t), u(t), t).$$

Se  $(x^\circ, u^\circ)$  é solução do problema (P) e as afirmações do teorema do Máximo local de Pontryagin (4.2.4) são satisfeitas, então  $\mathbf{H}(x^\circ, u, p, t)$ , como função de  $u$  sobre  $M$ , satisfaz as condições necessárias para um máximo, para quase todo  $0 \leq t \leq T$ , no ponto  $u = u^\circ(t)$  (esta é outra maneira de expressar a equação (4.15)). Expliquemos melhor:

Considerando que

$$\mathbf{H}_u(x^\circ, u^\circ, p, t) = \varphi_u^T(x^\circ, u^\circ, t) \cdot p(t) - \lambda_0 \mathbb{L}_u(x^\circ(t), u^\circ(t), t),$$

uma condição necessária para que  $\mathbf{H}(x^\circ, u, p, t)$  tenha um máximo em  $M$ , como função de  $u$ , é que  $-\mathbf{H}_u(x^\circ, u^\circ, p, t)$  seja um suporte de  $M$  no ponto  $u^\circ(t)$ .

Se  $\varphi$  é diferenciável com respeito a  $u$  e no tempo  $t$ ,  $u(t) \in \text{int } M$ , então a condição  $\mathbf{H}(x^\circ(t), u^\circ(t), p(t), t) \geq \mathbf{H}(x^\circ(t), u(t), p(t), t)$ , para todo  $u(t) \in M$ , implica que

$$\frac{d}{dt} \mathbf{H}(x^\circ(t), u^\circ(t), p(t), t) = 0.$$



**Observação 4.2.6** *Encontra-se usualmente na literatura o termo função Hamiltoniana ao invés de função de Pontryagin, devido ao fato de o  $\max H(x(t), u(t), p(t), t) = H(x^o(t), u^o(t), p(t), t)$  ser chamado Hamiltoniano  $\mathcal{H}$ . Contudo, neste trabalho, preferimos usar o termo função de Pontryagin, porque justifica melhor o título do teorema 4.2.4 na página 54.*

Agora discutiremos a respeito da condição de não degeneração do sistema (4.17), a qual usamos para obtermos a sobrejetividade do operador  $P'(x^o, u^o)$ .

**Definição 4.2.7 ( Condição de não degeneração  $\mathbb{A}$  )** *O sistema (4.17) se diz não degenerado, se qualquer solução não nula  $p(\cdot)$  da equação*

$$\frac{d}{dt}p(t) = -A^T(t)p(t), \quad (4.23)$$

*satisfaz a condição*

$$B^T(t)p(t) \neq 0, \text{ em um conjunto de medida positiva.} \quad (4.24)$$

*Aqui,  $A(t)$  e  $B(t)$  são matrizes  $n \times n$  e  $n \times r$  respectivamente e  $p(t) \in \mathbb{R}^n$ .*

**Observação 4.2.8** *Temos outra condição de não degeneração, a condição de não degeneração  $\mathbb{B}$ . Ela é uma condição necessária para a condição de não degeneração descrita acima. Por isto concluímos o que afirmamos anteriormente.*

*Para melhor esclarecimento disto que acabamos de mencionar, iremos dar maiores detalhes a respeito:*

**Definição 4.2.9 ( Condição de não degeneração  $\mathbb{B}$  )** *Seja  $D \subset \mathbb{R}^n$  um conjunto de todos os vetores  $\bar{x}(T) \in \mathbb{R}^n$ , os quais satisfazem a equação diferencial*

$$\begin{cases} \frac{d}{dt}\bar{x}(t) = A(t)\bar{x}(t) + B(t)\bar{u}(t); \\ \bar{x}(0) = 0, \end{cases} \quad (4.25)$$

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

*Se  $D \equiv \mathbb{R}^n$ , então se diz que a equação (4.25) é não degenerada.*

Continuando as nossas observações, às vezes expressa-se esta condição dizendo que o sistema é **completamente controlável**, em virtude do fato de que ao se considerar o sistema

$$\dot{x}(t) = \Phi(x(t), u(t), t), \text{ para } t \geq t_0, \quad (4.26)$$

com estado inicial  $x(t_0) = x_0$ , se existe um tempo finito  $t_1 \geq t_0$  e um controle  $u(t)$ , para  $t \in [t_0, t_1]$ , o qual transfere o estado inicial  $x_0$  para outro estado no tempo  $t_1$ , dizemos que o estado  $x_0$  é controlável em  $t_0$ . Ainda mais, se todos os valores de  $x_0$  são controláveis para todo  $t_0$ , o sistema (4.26) é completamente controlável, e qualitativamente isto significa que ele é controlável se cada variável de estado de (4.26) pode ser influenciada pelo controle  $u$  ("sinal de entrada"  $u$ ). R. E. Kalman mostrou que um sistema linear invariante no tempo (isto é, as matrizes  $A(t)$  e  $B(t)$  são constantes) é controlável se, e somente se, a matriz  $n \times r n$

$$\Theta = [B \ AB \ A^2B \ \dots \ A^{n-1}B] \text{ tem posto } n^{11}.$$

Se  $r = 1$ , uma condição necessária e suficiente para controlabilidade é que a matriz  $\Theta_{n \times n}$  seja não singular.

**Lema 4.2.10** Se a condição de não degeneração  $\mathbb{A}$  é verdadeira, então vale a condição de não degeneração  $\mathbb{B}$ .

**Prova:** Suponha que a condição de não degeneração  $\mathbb{B}$  é inválida, i.e.,  $D \neq \mathbb{R}^n$ . Como uma consequência direta da linearidade da equação diferencial,  $D$  é um subespaço e disto segue-se que existe um vetor não nulo  $\omega \in \mathbb{R}^n$  ortogonal a  $D$ . Equivalentemente,  $\omega \cdot \bar{x}(T) = 0$ , para qualquer  $\bar{x}(T) \in D$ , onde  $\bar{x}(t)$  é solução da equação (4.17).

Consideremos agora o sistema

$$\begin{aligned} \frac{d}{dt}p(t) &= -A^T(t)p(t); \\ p(T) &= \omega. \end{aligned}$$

---

<sup>11</sup>Para melhores informações a respeito de controlabilidade de um sistema, veja Barnett [6, p. 97-102], Chen [15, p. 179-192] e Shinnars [36, p. 463-466].

Então  $p(t) \neq 0$ , desde que  $\omega \neq 0$ . Conseqüentemente, para qualquer solução  $\bar{x}(t)$  de (4.17),

$$\begin{aligned} 0 &= \int_0^T (\dot{p}(t) + A^T(t)p(t))^T \cdot \bar{x}(t) dt = \\ &= p^T(T) \cdot \bar{x}(T) - p^T(0) \cdot \bar{x}(0) - \int_0^T p^T(t) \cdot \dot{\bar{x}}(t) dt + \\ &\quad + \int_0^T p^T(t) \cdot (A(t)\bar{x}(t)) dt \\ &= \omega \cdot \bar{x}(T) - \int_0^T p^T(t) \cdot (B(t)\bar{u}(t)) dt = \\ &= - \int_0^T (B^T(t)p(t))^T \cdot \bar{u}(t) dt, \end{aligned}$$

isto é,

$$\int_0^T (B^T(t)p(t))^T \cdot \bar{u}(t) dt = 0, \forall \bar{u} \in L_\infty^r(0, T).$$

Mas isto é possível somente se  $B^T(t)p(t) = 0$ , em quase toda parte de  $(0, T)$ , de onde chegamos numa contradição com a condição (4.24) da definição de não degeneração  $\mathbb{A}$ . Temos, portanto, o resultado desejado. ■

Agora, suponhamos que a condição de não degeneração  $\mathbb{B}$  é satisfeita para o sistema (4.17), novamente com  $A(t) = \varphi_x(x^o(t), u^o(t), t)$  e  $B(t) = \varphi_u(x^o(t), u^o(t), t)$ . Consideremos  $(\bar{a}(t), b) \in E_1$  e tomemos  $z \in C^n[0, T]$  tal que

$$z(t) = \bar{a}(t) + \int_0^t \varphi_x(x^o(\tau), u^o(\tau), \tau)z(\tau) d\tau, \forall t \in [0, T], \quad (4.27)$$

visto que (4.27) é uma equação linear integral de Volterra <sup>12</sup> de segunda espécie (pela definição 2.2.4), devido à proposição 2.2.5, ela possuirá uma única solução  $z \in C^n(0, T)$ , qualquer que seja  $\bar{a} \in C^n(0, T)$ .

Notemos que a equação

$$\hat{x}(t) - \int_0^t \varphi_x(x^o(\tau), u^o(\tau), \tau)\hat{x}(\tau) + \varphi_u(x^o(t), u^o(t), t)\hat{u}(t) d\tau = \bar{a}(t), \forall t \in [0, T],$$

---

<sup>12</sup>Vito Volterra, matemático italiano, com trabalhos em análise funcional e equações integrais.

tem em particular a solução  $\tilde{u}(t) = 0$  e  $\tilde{x}(t) = z(t)$ , para qualquer  $\bar{a} \in C^n(0, T)$ .

Por outro lado, já que o sistema (4.17) é não degenerado, podemos encontrar  $\bar{u} \in L_\infty^r(0, T)$  e  $y \in C^n(0, T)$  de modo que

$$\left. \begin{aligned} \dot{y}(t) &= \varphi_x(x^o(t), u^o(t), t)y(t) + \varphi_u(x^o(t), u^o(t), t)\bar{u}(t); \\ y(0) &= 0; \\ y(T) &= b - z(T). \end{aligned} \right\} \quad (4.28)$$

Finalmente, coloquemos  $\bar{x}(t) = y(t) + z(t)$  e para simplificar,  $s = (x^o(t), u^o(t), \tau)$ . Logo,

$$\begin{aligned} P'(x^o, u^o)(\bar{x}, \bar{u}) &= (\bar{x}(t) - \int_0^t [\varphi_x(s)\bar{x}(\tau) + \varphi_u(s)\bar{u}(\tau)] d\tau, \bar{x}(T)) = \\ &= (z(t) + y(t) - \int_0^t \{\varphi_x(s)[z(\tau) + y(\tau)] + \varphi_u(s)\bar{u}(\tau)\} d\tau, \bar{x}(T)). \end{aligned}$$

Assim, pelas equações expressas por (4.28), temos que

$$P'(x^o, u^o)(\bar{x}, \bar{u}) = (z(t) + y(t) - \int_0^t \varphi_x(x^o(t), u^o(t), \tau)z(\tau) d\tau - y(t), b).$$

E por (4.27) obtemos  $P'(x^o, u^o)(\bar{x}, \bar{u}) = (\bar{a}(t), b)$ , ou seja,  $P'(x^o, u^o)E = E_1$ .

Podemos dizer então, que a sobrejetividade do operador  $P'(x^o, u^o)$  pode ser verificada através da condição de não degeneração  $\Delta$ , de onde basta encontrar a matriz fundamental  $\Lambda$  (matriz transição de estado) da equação

$$\frac{d}{dt}p(t) = -A^T(t)p(t).$$

(de outra forma dito, basta resolvermos a equação matricial

$$\begin{aligned} \frac{d}{dt}\Lambda(t) &= -A^T(t)\Lambda(t); \\ \Lambda(0) &= I, \end{aligned}$$

onde  $I$  é a matriz identidade e  $\Lambda(t)$  é uma matriz  $n \times n$  e posteriormente encontrar a matriz  $B^T(t)\Lambda(t)$ .

Recalmente, como qualquer solução não nula  $p$  do sistema (4.23) tem a forma  $\Lambda(t)\lambda$ , com  $\lambda = (\lambda_1, \dots, \lambda_n)^T \neq 0$ , se as colunas  $\xi_i(t)$  da matriz  $B^T(t)\Lambda(t)$  são linearmente independentes, então

$$B^T(t)p(t) = B^T(t)\Lambda(t)\lambda = \sum_{i=1}^n \lambda_i \xi_i(t) \neq 0.$$

E como consequência, o sistema (4.17) será não degenerado.

**Observação 4.2.11** Se de antemão sabemos que o sistema linearizado (4.18) é não degenerado, então poderemos assumir  $\lambda_0 = 1$ .

A seguir, discutimos algumas modificações das condições de extremo para algumas versões do problema de controle ótimo:

A condição final (4.12) é dada por uma restrição mais geral

$$G_i(x(T)) = 0, \quad i = 1, \dots, q,$$

onde  $G_i$  são funcionais diferenciáveis em  $\mathbb{R}^n$ . Então para calcularmos o subespaço tangente  $K_2$ , assumiremos que a condição de não degeneração é satisfeita e que o operador  $P$  é dado por

$$P(x, u) = \left( x(t) - c - \int_0^t \varphi(x(t), u(t), \tau) d\tau, G_1(x(T)), \dots, G_q(x(T)) \right).$$

De maneira análoga,  $K_2$  consiste de todos os pares  $(\bar{x}, \bar{u}) \in E$  tais que

$$\frac{d}{dt} \bar{x} = \varphi_x(x^o(t), u^o(t), t) \bar{x} + \varphi_u(x^o(t), u^o(t), t) \bar{u}, \quad \bar{x}(0) = 0,$$

$$G'_i(x^o(T)) \cdot \bar{x}(T) = 0, \quad i = 1, \dots, q.$$

Agora assumiremos que  $p(\cdot)$  satisfaz a equação diferencial (4.14) com a condição de fronteira

$$p(T) = \sum_{i=1}^q \mu_i G'_i(x^o(T)), \quad (4.29)$$

onde  $\mu_i$ ,  $i = 1, \dots, q$ , são escalares arbitrários.

Assim, podemos usar os mesmos argumentos para provar (4.15) levando em conta o fato de que  $f_2(\bar{x}, \bar{u}) = \sum_{i=1}^q \mu_i G'_i(x^\circ(T)) \cdot \bar{x}(T)$ ;  $\mu_i \in \mathbb{R}$ . Portanto a formulação da condição necessária de extremo pode ser considerada na forma (4.15), com a única diferença que a função  $p$  nela figurada deve satisfazer a condição de fronteira (4.29).

Em geral, se a função de estado  $x$  é sujeita a condições de fronteira na forma geral,  $x(0) \in S_1$ ,  $x(T) \in S_2$ , onde  $S_1$  e  $S_2$  são variedades suaves, então adicionamos hipóteses referentes a  $p(0)$  e  $p(T)$ , a saber,  $p(0)$  deve ser transversal (i.e., ortogonal ao subespaço tangente) a  $S_1$  no ponto  $x^\circ(0)$ , e  $p(T)$  transversal a  $S_2$  no ponto  $x^\circ(T)$ .

Particularmente, se é dada a posição de estado inicial,  $x(0) = c$ , onde  $c$  é fixo, enquanto que para o tempo final  $T$ ,  $x(T)$  é livre, então serão necessárias condições de contorno apenas sobre  $p(T)$  (em virtude do fato de que não teríamos, durante a prova, o conjunto  $L_2$  e por isto na equação (4.20) não apareceria o termo  $f_3(\bar{x}, \bar{u})$ ). Assim, ao continuar o mesmo raciocínio na demonstração, deveríamos impor  $p(T) = 0$ , por causa do termo  $p\bar{x}|_0^T$ ). Portanto, podemos dizer que a condição de não degeneração é desprezível. Conseqüentemente nos é permitido considerar  $\lambda_0 = 1$ .

São feitas modificações semelhantes para o caso em que  $S_1$  e  $S_2$  são conjuntos convexos. Basta lembrarmos das justificativas concernentes ao cone dual das direções factíveis a  $Q_1$  e então impormos que  $p(0)$  é suporte de  $S_1$  no ponto  $x^\circ(0)$  e  $p(T)$  é suporte de  $S_2$  em  $x^\circ(T)$ .

Finalmente, se o funcional a ser minimizado não é o funcional integral (4.9), mas o funcional  $\mathbb{L}_0(x(T))$ , onde  $\mathbb{L}_0(x(\cdot))$  é uma função diferenciável em  $\mathbb{R}^n$ , ou seja, o problema é de Mayer, então este problema se reduzirá ao problema (P), página 53, desde que

$$\mathbb{L}_0(x(T)) = \mathbb{L}_0(x(0)) + \int_0^T \frac{d}{dt} \mathbb{L}_0(x(t)) \cdot \varphi(x(t), u(t), t) dt.$$

Entretanto, problemas deste tipo podem também ser considerados diretamente, uma vez que o cone de direções de descida para este funcional possa ser determinado usando o teorema 3.2.2 na página 33. Disto segue-se que as condições de extremo tem a mesma forma; uma vez que é preciso apenas tomar  $\mathbb{L}(x, u, t) = 0$  e  $p(x(T)) = -\frac{d}{dt} \mathbb{L}_0(x^\circ(t))$ .

Outra classe de problemas de controle é aquela em que aparecem restrições nas coordenadas de fase do tipo  $G(x(t), t) \leq 0, \forall t \in [0, T]$  e então  $p(T) = a$ , e no caso em que também  $T$  é livre, tomamos  $p(T) = 0$ . No que se refere a estes problemas,  $p(\cdot)$  é uma função em geral descontínua e satisfaz a equação diferencial

$$\frac{d}{dt}p(t) = -\varphi_x^T(x^o(t), u^o(t), t)p(t) + G_x(x^o, t)\frac{d}{dt}\mu(t),$$

onde a medida não-negativa,  $d\mu(t)$ , tem por suporte o conjunto

$$R = \{t \in [0, T] : G(x^o(t), t) = 0\}.$$

Para estas duas situações as quais acabamos de mencionar, as condições de extremo são deduzidas em Girsanov [22, capítulo 14] e em Ioffe [26, capítulo 15].

No contexto de problemas de controle ótimo com tempo livre e condições de estado final dadas, sugerimos uma leitura mais detalhada em Girsanov [22, capítulo 13]; também para tempo livre e condições impostas aos estados inicial e final, veja Fomine [2, capítulo 4], Ioffe [26, capítulo 2] e Zeidler [40, capítulo 48].

# Capítulo 5

## Aplicação do Princípio do Máximo

Neste capítulo apresentaremos um problema prático, o problema do pouso brando, sobre o qual aplicaremos o Princípio do Máximo. Entretanto, primeiramente, faremos algumas considerações a respeito da forma com que o problema de controle pode ser considerado, seguidas do Princípio do Máximo de Pontryagin correspondente a essa forma.

### 5.1 Princípio do Máximo de Pontryagin

#### 5.1.1 Formulação do Problema

Consideremos no espaço  $KC^1(\Delta, \mathbb{R}^n) \times KC(\Delta, \mathbb{R}^m) \times \mathbb{R}^2$ , o problema de controle:

$$\min B_0(x(\cdot), u(\cdot), t_0, t_1);$$

$$\begin{aligned}\dot{x}(t) &= \varphi(t, x(\cdot), u(\cdot), t_0, t_1), \\ u(t) &\in \mathcal{U}, \forall t \in [t_0, t_1],\end{aligned}$$



$$\begin{aligned} \mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) &\leq 0, \quad i = 1, \dots, m', \\ \mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) &= 0, \quad i = m' + 1, \dots, m, \end{aligned}$$

onde

$$\mathcal{B}_i(x(\cdot), u(\cdot), t_0, t_1) = \int_{t_0}^{t_1} \mathbb{L}_i(t, x(t), u(t)) dt + \psi_i(t_0, x(t_0), t_1, x(t_1)),$$

$$i = 0, 1, \dots, m.$$

Aqui,  $KC(\Delta, \mathbb{R}^n)$  é o espaço das funções contínuas por partes em  $\Delta$ , ou seja, das funções que apresentam um número finito de descontinuidades de primeira espécie (nos pontos de descontinuidade, terá os limites finitos à direita e à esquerda), onde  $\Delta$  é um intervalo finito dado;  $t_0, t_1 \in \Delta$ ;

$\mathbb{L}_i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}$  são funções de  $n + r + 1$  variáveis;  $\psi_i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}$  são funções de  $2n + 2$  variáveis;  $\varphi : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \longrightarrow \mathbb{R}^n$  é uma função vetorial de  $n + r + 1$  variáveis;  $\mathcal{U}$  é um conjunto arbitrário tomado em  $\mathbb{R}^r$ .

Utilizaremos, neste contexto, a *função de Lagrange*

$$\begin{aligned} \mathcal{L}(x, u, t_0, t_1; p, \lambda) &= \int_{t_0}^{t_1} (\sum_{i=0}^m \lambda_i \mathbb{L}_i(t, x, u) + p(t)(\dot{x} - \varphi(t, x, u))) dt \\ &\quad + \sum_{i=0}^m \lambda_i \psi_i(t_0, x(t_0), t_1, x(t_1)), \end{aligned}$$

onde  $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m)$  e  $p(\cdot) \in KC^1([t_0, t_1], \mathbb{R}^{n*})$ .

### 5.1.2 Condições de otimalidade para problemas de controle

**Teorema 5.1.1 (Princípio do Máximo Local de Pontryagin)** *Suponhamos que  $\hat{\xi} = (\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$  seja o processo ótimo do problema de controle ótimo, dado acima, para o qual  $\mathbb{L}_i, i = 0, 1, \dots, m$ , são funções contínuas com suas derivadas parciais com respeito a  $x$  no conjunto  $\mathcal{V} \times \mathcal{U}$ , onde  $\mathcal{V}$  é uma certa vizinhança do conjunto  $M = \{(t, \hat{x}(t)) : t \in [\hat{t}_0, \hat{t}_1]\}$ ; e*

$\psi_i, i = 0, 1, \dots, m$ , são funções continuamente deriváveis na vizinhança do ponto  $(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1))$  (condição de diferenciabilidade). Portanto, existem multiplicadores de Lagrange  $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in p(\cdot) \in KC^1([\hat{t}_0, \hat{t}_1], \mathbb{R}^{n^*})$ , que não são simultaneamente iguais a zero, e tais que, para a função de Lagrange  $\mathcal{L}$ , sejam satisfeitas as condições:

a) de estacionaridade com respeito à  $x$  (Equação de Euler):

$$-\frac{d}{dt} \hat{L}_{\dot{x}}(t) + \hat{L}_x(t) = 0^1,$$

para o Lagrangiano

$$L(t, x, \dot{x}, u) = \sum_{i=0}^m [\lambda_i \mathbb{L}_i(t, x, u)] + p(t) (\dot{x} - \varphi(t, x, u)),$$

sc. e somente se,

$$\dot{p}(t) = \mathbb{L}_x(t) - p(t) \hat{\varphi}_x(t), \quad \forall t \in [t_0, t_1],$$

onde definiremos

$$\mathbb{L}(t, x, u) = \sum_{i=0}^m \lambda_i \mathbb{L}_i(t, x, u);$$

b) de transversalidade com respeito a  $x$ :

$$\hat{L}_{\dot{x}}(\hat{t}_k) = (-1)^k \hat{\ell}_{x_k} \iff p(\hat{t}_k) = (-1)^k \sum_{i=0}^m \lambda_i \hat{\psi}_{ix_k}, \quad k = 0, 1$$

para  $\ell = \ell(t_0, x_0, t_1, x_1) = \sum_{i=0}^m \lambda_i \psi_i(t_0, x_0, t_1, x_1)$ <sup>2</sup>, ou seja,

$$\begin{aligned} p(\hat{t}_0) &= \hat{\ell}_{x_0}, \\ p(\hat{t}_1) &= \hat{\ell}_{x_1}; \end{aligned}$$

<sup>1</sup>Usaremos as notações  $\hat{L}_x(t) = L_x(t, x(t), \frac{d}{dt}x(t), u(t))$ ;  $\hat{\varphi}(t) = \varphi(t, x(t), u(t))$ ;

$\hat{\ell}_{x_k} = \ell_{x_k}(t_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1))$ , etc.  $\mathcal{L}_{t_k}$  será avaliada no ponto ótimo. Por problemas de estética, não usaremos circunflexo em  $\mathbb{L}$ . Entretanto está fácil de detectar no texto quando isso acontece.

<sup>2</sup> $x(t_0) = x_0$  e  $x(t_1) = x_1$ .

c) de otimalidade com respeito a  $u$ , consistindo no **Princípio do Mínimo** sob a forma de **Lagrange**:

$$\min_{u \in \mathcal{U}} \dot{L}(t, \hat{x}(t), \frac{d}{dt} \hat{x}(t), u) = \dot{L}(t, \hat{x}(t), \frac{d}{dt} \hat{x}(t), \hat{u}(t)) \text{ se, e somente se,}$$

$$\min_{u \in \mathcal{U}} \mathbb{L}(t, \hat{x}, u) - p(t)\varphi(t, \hat{x}, u) = \mathbb{L}(t, \hat{x}, \hat{u}) - p(t)\varphi(t, \hat{x}, \hat{u}), \forall t \in [t_0, t_1];$$

ou no **Princípio do Máximo** sob a forma do **Hamiltoniano**  $\mathcal{H}$  (forma de Pontryagin):

$$\max_{u \in \mathcal{U}} \mathbb{H}(t, \hat{x}(t), u, p(t)) = \mathcal{H}(t, \hat{x}(t), \hat{u}(t), p(t)),$$

onde  $\mathbb{H}(t, x, u, p) = p\varphi(t, x, u) - \mathbb{L}(t, x, u)$  é a **função de Pontryagin**:

d) de estacionaridade com respeito a  $t_k$ , para  $k = 0, 1$ :

$$\mathcal{L}_{t_k} = 0 \text{ se, e somente se,}$$

$$(-1)^{k+1} \sum_{i=0}^m \lambda_i \mathbb{L}_i(\hat{t}_k) + \sum_{i=0}^m \lambda_i (\hat{\psi}_{i t_k} + \hat{\psi}_{i x_k} \frac{d}{dt} \hat{x}(t_k)) = 0 \text{ se, e somente se,}$$

$$\mathcal{H}(\hat{t}_k) = (-1)^{k+1} \hat{\ell}_{t_k},$$

equivalentemente,

$$-\mathbb{L}(\hat{t}_0) + \hat{\ell}_{t_0} + \hat{\ell}_{x_0} \hat{\varphi}(\hat{t}_0) = 0,$$

$$-\mathbb{L}(\hat{t}_1) + \hat{\ell}_{t_1} + \hat{\ell}_{x_1} \hat{\varphi}(\hat{t}_1) = 0;$$

e) de não rigidcz complementar:

$$\lambda_i \mathcal{B}_i(\hat{\xi}) = 0, \quad i = 1, \dots, m';$$

f) de não negatividade:

$$\lambda_i \geq 0, \quad i = 0, 1, \dots, m'.$$

**Prova:** A demonstração deste teorema pode ser encontrada em Alexeév [3, 1987, p. 169 a 178], para problemas com estado final livre e tempo fixo. Também podem ser encontradas diferentes provas, mas com a mesma lógica na exibição, utilizando outros espaços na formulação do problema, salvo algumas considerações abaixo citadas:

- estados inicial e final fixos e tempo final livre, em Girsanov [22, capítulo 13];
- condições de fronteira:  $B_i(x(\cdot), u(\cdot), t_0, t_1) = \psi_i(x(t_1), t_1) = 0$ , para  $i = 1, \dots, m$ , e tempo final livre, em Zeidler [40, p. 426 a 433];
- condições de fronteira:

$$B_i(x(\cdot), u(\cdot), t_0, t_1) = \psi_i(x(t_0), t_0) = 0, \quad i = 1, \dots, m';$$

$$B_i(x(\cdot), u(\cdot), t_0, t_1) = \psi_i(x(t_1), t_1) = 0, \quad i = m' + 1, \dots, m,$$

e tempo final livre, em Ioffe [26, p. 147 a 159]. ■

**Observação 5.1.2** *Apresentamos o Princípio do Máximo Local de Pontryagin. para problemas com tempo final livre, com uma aparência diferente daquela apresentada na página 54, porque aparece com mais freqüência na li-*

*teratura. Contudo, em cada condição do teorema 5.1.1, página 71, fizemos a relação com a função de Pontryagin  $\epsilon$ /ou a função  $p(\cdot) \in KC^1([t_0, t_1], \mathbb{R}^{n*})$ , a qual aparece no Princípio dado na página 54.*

## 5.2 Problema de aterrissagem moderada de um veículo espacial - problema de pouso brando

Um veículo espacial deve realizar um pouso brando ( aterrissagem moderada), isto é, chegar à superfície plana de corpos celestes privados de atmosfera, por exemplo, a Lua, com velocidade vertical nula, minimizando o consumo de combustível.

**Observação 5.2.1** *Em toda essa seção consideraremos a superfície da Lua, por isso usaremos também o termo alunagem.*

Estamos supondo que o veículo espacial realize uma trajetória retilínea perpendicular a superfície da Lua. Designemos por:

- $h(t)$  a altura do veículo acima da superfície da Lua no tempo  $t$ ;
- $v(t)$  sua velocidade vertical no tempo  $t$ ;
- $m(t)$  a massa do veículo no instante  $t$ .

Então  $x(t) = (h(t), v(t), m(t))$  descreve o estado do veículo como função do tempo.

Suponhamos que durante a manobra de freagem, o veículo espacial esteja sob a ação da força de gravidade  $F_g(t) = -gm(t)$ , onde o fator  $g$  é uma constante positiva (aceleração da gravidade) e sob a ação da força de impulso (força propulsora)  $F_I(t) = k u(t)$ , onde  $k$  é uma constante positiva (aceleração de impulso) e  $u(t)$  é a força oposta a velocidade de consumo do combustível:  $u(t) = -\dot{m}(t)$ . Assumamos finalmente que a função  $u$  satisfaz a restrição  $0 \leq u \leq U, \forall t \geq 0$  (observe a figura 5.1).

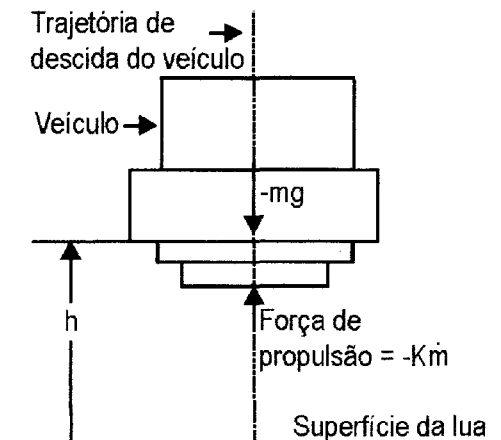


Figura5.1: Forças atuando sobre o veículo

A partir disso, o mecanismo de vôo será descrito, de acordo com a 2ª Lei de Newton, pelas equações

$$\left. \begin{aligned} m\ddot{h} &= ku - gm, \\ \dot{m} &= -u, \quad 0 \leq u \leq U. \end{aligned} \right\} \quad (5.1)$$

Assumamos que o pouso se inicia no tempo  $t_0 = 0$  e o estado inicial é dado por

$$x(0) = (h_0, v_0, m_0),$$

onde  $h_0 = h(0) > 0$ ;  $v_0 = v(0)$ ;  $m_0 = m(0)$  é a massa de combustível disponível,  $M_d$ , mais a massa do veículo sem combustível,  $M_v$ , no instante inicial.

Podemos portanto, juntamente com (5.1), descrever o comportamento dinâmico do veículo através das seguintes equações de movimento:

$$\left. \begin{aligned} \dot{h} &= v; \\ \dot{v} &= k\frac{u}{m} - g; \\ \dot{m} &= -u; \\ x(0) &= (h_0, v_0, m_0). \end{aligned} \right\} \quad (5.2)$$

Se  $t_1 = T$  é o momento final da manobra de freagem, então para que a alunagem se efetue com suavidade (brandura) deve-se cumprir

$$h(T) = v(T) = 0 \quad (5.3)$$

Porque  $T > 0$ , dizemos que a função  $u : [0, T] \rightarrow \mathcal{U} = [0, U]$  é um controle admissível, se  $u \in KC([0, T], \mathbb{R})$ , e se a solução  $x(u, t)$  de (5.2), associada a  $u$ , é definida em  $[0, T]$  e satisfaz (5.3), ou melhor dizendo,

$$h(u, T) = v(u, T) = 0.$$

Supondo que  $A$  é o conjunto de controles admissíveis e considerando que a massa  $M$  do combustível consumido é igual à diferença entre a massa inicial e a massa final, então o problema pode ser expressado como o seguinte **programa matemático (PM)**:

$$\min \{M(u, T) : u \in A\}. \quad (5.4)$$

Observemos que:

- O tempo necessário para o pouso brando não é fixo;
- As condições finais  $h(T) = v(T) = 0$  deveriam ser completadas pelas restrições

$$\left. \begin{aligned} h(t) &\geq 0, & \forall t \in [0, T]; \\ m(t) &\geq M_v, & \forall t \in [0, T], \end{aligned} \right\} \quad (5.5)$$

onde  $M_v$  é a massa do veículo sem combustível. Entretanto, veremos que é possível resolver o problema sem levar em conta, de maneira explícita, as restrições (5.5).

A seguir, colocamos o problema (PM) como um problema de controle ótimo (PC).

(1) **Formalização do problema:**

$$\mathbf{PC} : \min_{m_0 \geq m \geq UT} m_0 - m(T)$$

s.a

$$\left. \begin{aligned} \dot{h} &= v; \\ \dot{v} &= k \frac{u}{m} - g; \\ \dot{m} &= -u; \end{aligned} \right\} \quad (5.6)$$

$$\begin{aligned} x(0) &= (h_0, v_0, m_0); \\ h(T) &= v(T) = 0; \\ u(t) &\in \mathcal{U} = [0, U], \quad \forall t \in [0, T]. \end{aligned}$$

Aplicaremos agora o Princípio do Máximo (página 71):

Seja  $u \in KC([0, T], \mathbb{R})$  um controle ótimo para o problema (PC) e  $x(t)$  a trajetória ótima associada a  $u(t)$ . De acordo com aquele Princípio, devem existir uma função  $p : [0, T] \rightarrow \mathbb{R}^3$  e um vetor  $\lambda = (\lambda_0, \mu_1, \mu_2)$  com  $\lambda_0 \geq 0$ , tais que as condições (a), (b), (c), (d), (e) e (f) especificadas abaixo, no item (2), sejam satisfeitas.

**Observação 5.2.2** Não estamos considerando a condição de contorno em  $t = 0$ , pois não contribuirá para a resolução. As funções são consideradas somente para parâmetros variáveis e por isso  $\lambda \in \mathbb{R}^3$ . E para escrevermos a função de Lagrange  $\mathcal{L}$ , iremos fazer uma analogia com o problema geral de controle descrito na seção anterior:

$$\begin{aligned}
\mathcal{B}_0(x(\cdot), u(\cdot), T) &= \psi_0(x(T)) = m_0 - m(T); \\
\dot{x} &= (\dot{h}, \dot{v}, \dot{m})^T; \\
\varphi &= (v, k\frac{u}{m} - g, -u)^T; \\
\mathbb{L}_i &= 0, \quad i = 0, 1, 2. \\
\psi_1(x(T)) &= h(T); \\
\psi_2(x(T)) &= v(T).
\end{aligned}$$

A função de Lagrange será escrita como

$$\begin{aligned}
\mathcal{L}(x(\cdot), u(\cdot), T; p(\cdot), \lambda) &= \int_0^T p(t)(\dot{x} - \varphi)dt + \sum_{i=1}^3 \lambda_i \psi_i(T, x(T)) \\
&= \lambda_0 (m_0 - m(T)) + \int_0^T [p_1(\dot{h} - \dot{v}) + p_2(\dot{v} - \frac{ku}{m} + g) + p_3(\dot{m} + u)]dt \\
&\quad + \mu_1 h(T) + \mu_2 v(T),
\end{aligned}$$

onde o Lagrangeano é dado por

$$L(x, u, t, p) = p(t)(\dot{x} - \varphi) = p_1(\dot{h} - \dot{v}) + p_2(\dot{v} - \frac{ku}{m} + g) + p_3(\dot{m} + u).$$

## (2) Condições necessárias:

a) equação de Euler:

$$\dot{p}^T = -p^T \varphi_x.$$

Logo,

$$-\dot{p}_1 = 0, \quad -\dot{p}_2 - p_1 = 0 \quad \text{e} \quad -\dot{p}_3 + p_2 \frac{ku}{m^2} = 0;$$

b) transversalidade com respeito à  $x$ :

$$\begin{aligned}
p(T) &= -\lambda_0 \psi_{0x(T)}(T) - \mu_1 \psi_{1x(T)}(T) - \mu_2 \psi_{2x(T)}(T) \\
&= -\lambda_0(0, 0, -1)^T - \mu_1(1, 0, 0)^T - \mu_2(0, 1, 0)^T.
\end{aligned}$$



Portanto,

$$\begin{aligned}
 p_1(T) &= -\mu_1, \\
 p_2(T) &= -\mu_2, \\
 p_3(T) &= -\lambda_0;
 \end{aligned} \tag{5.7}$$

c) *otimalidade com respeito a u :*

$$\begin{aligned}
 \min_{c \in \mathcal{U}} L(x, c, t, p) &= \max_{c \in \mathcal{U}} H(x, c, t, p) = \\
 \max_{c \in \mathcal{U}} p^T(t) \varphi(t, x, c) &= \max_{c \in \mathcal{U}} [p_1(t)v + p_2(t) \frac{kc}{m(t)} - p_2(t)g - p_3(t)c] = \\
 \max_{c \in \mathcal{U}} [p_2(t) \frac{kc}{m(t)} - p_3(t)c] &= p_2(t) \frac{ku}{m(t)} - p_3(t)u.
 \end{aligned}$$

Assim,

$$\max_{c \in \mathcal{U}} H(x, c, t, p) = p_2(t) \frac{ku}{m(t)} - p_3(t)u, \quad \forall t \in [0, T]. \tag{5.8}$$

d) *estacionaridade com respeito a T :*

Para  $\ell = \ell(x(T)) = \lambda_0 \psi_0(x(T)) - \mu_1 \psi_1(x(T)) - \mu_2 \psi_2(x(T))$ , temos  $\ell_T = 0$  e portanto,  $-\ell_T - \ell_{x(T)} \varphi(T) = 0$  se, e somente se,

$$\begin{aligned}
 0 &= -\lambda_0(0, 0, -1)(\dot{h}(T), \dot{v}(T), \dot{m}(T))^T - \\
 &\quad -\mu_1(1, 0, 0)(\dot{h}(T), \dot{v}(T), \dot{m}(T))^T - \\
 &\quad -\mu_2(0, 1, 0)(\dot{h}(T), \dot{v}(T), \dot{m}(T))^T.
 \end{aligned}$$

$$\lambda_0 \dot{m}(T) + \mu_1 \dot{h}(T) + \mu_2 \dot{v}(T) = 0 \tag{5.9}$$

se, e somente se,

$$\lambda_0 u(T) - p_2(T) \left( \frac{ku(t)}{m(t)} - g \right) = 0.$$

e) *não rigidez complementar:*

Como não temos restrições de desigualdades, essa condição não existe.

f) não negatividade:

$$\lambda_0 \geq 0.$$

### (3) Controles admissíveis:

Resulta do ítem a) que

$$\left. \begin{aligned} p_1(t) &= z = cte \\ p_2(t) &= -zt + q; \quad q = cte. \end{aligned} \right\} \quad (5.10)$$

Fazendo

$$w(t) = -p_3(t) - k \frac{p_2(t)}{m(t)}, \quad (5.11)$$

obtemos  $\dot{w}(t) = -\dot{p}_3(t) - \frac{k}{m^2(t)}[-\dot{p}_2(t)m(t) - p_2(t)\dot{m}(t)]$ . Agora de (a), (5.10) e (5.2) teremos

$$\dot{w}(t) = -\frac{kz}{m(t)}. \quad (5.12)$$

Na condição (5.9), de estacionaridade com respeito a  $T$ , use (5.7) e (5.11). Então

$$\begin{aligned} p_3(T)u(T) - k \frac{p_2(T)u(T)}{m(T)} + gp_2(T) &= 0 \Leftrightarrow \\ \left[ p_3(T) - k \frac{p_2(T)}{m(T)} \right] u(T) + gp_2(T) &= 0 \Leftrightarrow \\ -w(T)u(T) + gp_2(T) &= 0. \end{aligned}$$

Logo,

$$w(T)u(T) = gp_2(T). \quad (5.13)$$

Agora, obtemos pela otimalidade (5.8) com respeito a  $u$ , considerando  $w(t)u$  linear em  $[0, U]$ , a seguinte solução para o controle:

$$u = \begin{cases} 0, & \text{se } w(t) < 0; \\ U, & \text{se } w(t) > 0. \end{cases} \quad (5.14)$$

Se  $z = 0$ , de (5.12) vem que

$$\dot{w}(t) = -\frac{k \cdot 0}{m(t)} = 0 \Rightarrow w(t) = w_0 = cte.$$

Neste caso, se  $w_0 = 0$ , então em (5.10)  $p_2(t) = q$  e de (5.13),  $p_2(T) = \frac{w(T)u(T)}{g} = q$ . Mas  $w(T) = w_0 = 0 \Rightarrow p_2(T) = 0 = q$ . Logo,  $p_2(t) = 0, \forall t$ .

De onde  $p_2(T) = -\mu_2 = 0$ . Também  $p_1(t) = z = 0 = p_1(T) = -\mu_1 = 0$ . Visto que  $-\dot{p}_3 + p_2 \frac{kv}{m^2} = 0$ , então  $-\dot{p}_3 = 0$ . Por conseguinte,  $p_3(t) = cte$ . Da equação (5.7),  $p_3(t) = p_3(T) = \lambda_0$ . Entretanto  $0 = w_0 = w(t) = -p_3(t) = -\lambda_0$ , ou seja,  $\lambda_0 = 0$ . Deste modo provamos que todos os multiplicadores de Lagrange são nulos. Portanto, se  $z = 0$ , então  $w_0 \neq 0$ . Por esta razão, em (5.14),

$$\begin{cases} u(t) = 0 \\ \text{ou} \\ u(t) = U. \end{cases}$$

No entanto, se  $z \neq 0$ , concluímos que a função  $w$  é estritamente monótona. Porque pode existir algum  $\bar{t} \in [0, T]$  tal que  $w(\bar{t}) = 0$ , teremos uma única mudança do valor de  $u$ . Podemos expressar isso da seguinte forma:

$$u(t) = \begin{cases} U, & \text{em } [0, \bar{t}]; \\ 0, & \text{em } [\bar{t}, T], \end{cases} \quad (5.15)$$

ou

$$u(t) = \begin{cases} 0, & \text{em } [0, \bar{t}]; \\ U, & \text{em } [\bar{t}, T]. \end{cases}$$

Agora por (5.6) e (5.15), para  $t \in [\bar{t}, T]$  o veículo estaria em queda livre. Assim, quando o veículo atinge a superfície da Lua, sua velocidade é diferente de zero, portanto um pouso brando é impossível. Assim, o veículo se desloca utilizando  $u(t) = U$  durante todo o pouso, ou muda de  $u(t) = 0$  para  $u(t) = U$  num certo instante  $t \in [0, T]$ .

#### (4) Solução:

Definamos por  $\tau$  o instante de mudança do valor de  $u$ ; assumamos que  $u(t) = 0, \forall t \in [0, \tau[$  e  $u(t) = U, \forall t \in [\tau, T]$ .

Agora busquemos o lugar geométrico dos pontos  $(h, v) = (h(\tau), v(\tau))$  para os quais se verifica:

$$h(T) = v(T) = 0.$$

Consideremos o intervalo  $[0, \tau[$ , onde  $u(t) = 0$ . Então o deslocamento do veículo em queda livre é, com  $\hat{h}(0) = h_0, \hat{v}(0) = v_0, \hat{m}(0) = m_0$ , definido

por:

$$\tilde{h}(t) = h_0 + v_0 t - g \frac{t^2}{2}; \quad (5.16)$$

$$\tilde{v}(t) = v_0 - gt; \quad (5.17)$$

$$\tilde{m}(t) = m_0.$$

(Resolvemos o sistema (5.6) com  $u = 0$ ).

Isolemos  $t$  em (5.17) e substituamos em (5.16):

$$t = \frac{\tilde{v} - v_0}{-g}.$$

$\tilde{h} = h_0 + v_0 \left( \frac{\tilde{v} - v_0}{-g} \right) - \frac{g}{2} \left( \frac{\tilde{v} - v_0}{-g} \right)^2 \Rightarrow \tilde{h} = h_0 - v_0 \frac{\tilde{v}}{g} + \frac{\tilde{v}_0^2}{g} - \frac{\tilde{v}^2}{2g} + \frac{2\tilde{v}v_0}{2g} - \frac{\tilde{v}_0^2}{2g}$ . Por conseguinte,

$$\tilde{h} = -\frac{1}{2g}\tilde{v}^2 + \frac{1}{2g}\tilde{v}_0^2 + h_0. \quad (5.18)$$

Dessa forma, a trajetória  $(\tilde{h}(t), \tilde{v}(t))$ ,  $t \in [0, \tau]$  é em conseqüência uma parábola, chamada **parábola de queda livre**.

Por outro lado, quando  $u(t) = U$ ,  $\forall t \in [\tau, T]$ , o deslocamento do veículo é definido pelas equações (5.6), com  $h(\tau) = \tilde{h}$ ,  $v(\tau) = \tilde{v}$ ,  $m(\tau) = m_0$ .

Determinemos a solução função de estado do veículo:

$$\dot{m}(t) = -U \Rightarrow \int_{\tau}^t \dot{m}(r) dr = - \int_{\tau}^t U dr \Rightarrow m(t) - m(\tau) = -U(t - \tau) \Rightarrow$$

$$m(t) = m_0 - U(t - \tau). \quad (5.19)$$

$$\dot{v}(t) = k \frac{U}{m(t)} - g \Rightarrow \int_{\tau}^t \dot{v}(r) dr = \int_{\tau}^t k \frac{U}{m(r)} dr - g \int_{\tau}^t dr \Rightarrow v(t) - v(\tau) = -k \ln \left[ \frac{m(t)}{m(\tau)} \right] - g(t - \tau). \text{ Logo,}$$

$$v(t) = \tilde{v} - g(t - \tau) - k \ln \left[ 1 - \frac{U}{m_0}(t - \tau) \right]. \quad (5.20)$$

$$\dot{h}(t) = v(t) = \tilde{v} - g(t - \tau) - k \ln \left[ 1 - \frac{U}{m_0}(t - \tau) \right] \text{ de onde segue que,}$$

$$h(t) = \tilde{h} + \tilde{v}(t - \tau) - \frac{g}{2}(t - \tau)^2 - k \int_{\tau}^t \ln \left[ 1 + \frac{U\tau}{m_0} - \frac{Ur}{m_0} \right] dr. \text{ Sabendo que}$$

$$\int \ln[\gamma - \beta y] dy = -\frac{1}{\beta}(\gamma - \beta y) \ln[\gamma - \beta y] + \frac{\gamma - \beta y}{\beta} + cte,$$

teremos:

$$\int_{\tau}^t \ln \left[ 1 + \frac{U\tau}{m_0} - \frac{U\tau}{m_0} \right] dr = -\frac{m_0}{U} \left[ \frac{m_0+U\tau}{m_0} - \frac{Ut}{m_0} \right] \ln \left[ \frac{m_0+U\tau}{m_0} - \frac{Ut}{m_0} \right] +$$

$$+ \frac{m_0}{U} \left[ \frac{m_0+U\tau}{m_0} - \frac{Ut}{m_0} \right] + \frac{m_0}{U} \left[ \frac{m_0+U\tau}{m_0} - \frac{U\tau}{m_0} \right] \ln \left[ \frac{m_0+U\tau}{m_0} - \frac{U\tau}{m_0} \right] -$$

$$- \frac{m_0}{U} \left[ \frac{m_0+U\tau}{m_0} - \frac{U\tau}{m_0} \right] = \left[ -\frac{m_0}{U} + (t-\tau) \right] \ln \left[ 1 - \frac{U}{m_0}(t-\tau) \right] - (t-\tau).$$

Assim,

$$h(t) = \bar{h} + \bar{v}(t-\tau) - \frac{g}{2}(t-\tau)^2 + k(t-\tau) +$$

$$+ \frac{km_0}{U} \left[ 1 - \frac{U}{m_0}(t-\tau) \right] \ln \left[ 1 - \frac{U}{m_0}(t-\tau) \right]. \quad (5.21)$$

Suponhamos que o tempo durante o qual se usa o impulso máximo seja  $s$ , ou seja,  $s = T - \tau$ . Portanto em (5.19), (5.20) e (5.21) juntamente com (5.3) obtemos:  $m(T) = m_0 - Us$ ;  $v(T) = \bar{v} - gs - k \ln \left[ 1 - \frac{U}{m_0}s \right]$ . De onde

$$\bar{v} = gs + k \ln \left[ 1 - \frac{U}{m_0}s \right]; \quad (5.22)$$

na seqüência,  $h(T) = \bar{h} + \bar{v}s - \frac{g}{2}s^2 + ks + \frac{km_0}{U} \left[ 1 - \frac{U}{m_0}s \right] \ln \left[ 1 - \frac{U}{m_0}s \right]$ . Portanto,

$$\bar{h} = -\bar{v}s + \frac{g}{2}s^2 - ks - \frac{km_0}{U} \left[ 1 - \frac{U}{m_0}s \right] \ln \left[ 1 - \frac{U}{m_0}s \right]. \quad (5.23)$$

As equações (5.23) e (5.22) nos indicam a altura e a velocidade iniciais, para as quais é possível realizar um pouso brando, usando a força máxima de freio durante o tempo  $s$ .

Para que o problema seja realista, deve-se cumprir:  $U > \frac{m_0}{k}g$ . Isto é, a força máxima de freio (força de propulsão) deve superar a força da gravidade inicial. Devido a (5.5),  $m(T) \geq M_v$ , e se supusermos que todo combustível é consumido no tempo  $\hat{T}$ , teremos  $M_v = M_v + M_d - Us$ . Portanto, o maior tempo durante o qual se pode frear com força máxima e pousar brandamente é dado por  $\hat{T} = \frac{M_d}{U}$ .

Seja  $\Gamma = \left\{ (h(s), v(s)) : 0 \leq s \leq \hat{T} = \frac{M_d}{U} \right\}$ . Verificamos para os pontos  $(h(s), v(s)) \in \Gamma$  o seguinte:

Se o estado do veículo é  $(h(s), v(s), m_0)$ , então usando a força máxima de freio, o veículo poussa brandamente depois do tempo  $s$ . Denominaremos  $\Gamma$  por

**curva comutadora.** Esta terminologia nos lembra o fato de fazermos uma mudança no valor do controle e acarretarmos uma mudança na curva plana que descreve o estado do veículo.

Podemos eliminar o parâmetro  $s$  das equações (5.22) e (5.23), e obtermos uma curva  $\Theta(\bar{h}, \bar{v}) \subset \Gamma$ . Desta forma, dado  $(h_0, v_0)$  e considerando que a trajetória  $(\tilde{h}(t), \tilde{v}(t))$  intercepta  $\Gamma$  no ponto  $(\tilde{h}(\tau), \tilde{v}(\tau)) = (h(s), v(s)) = (\bar{h}, \bar{v})$ , obtemos o tempo  $\tau$  definido como a primeira raiz positiva da equação

$$\Theta\left(h_0 + v_0\tau - g\frac{\tau^2}{2}, v_0 - g\tau\right) = 0. \quad (5.24)$$

Podemos dar portanto, a seguinte resolução:

Durante o tempo  $\tau$  deixamos o veículo se deslocar em queda-livre, em seguida propomos uma força de freio máxima. Se para  $(h_0, v_0)$  fixo, a equação (5.24) não admite solução, a alunagem branda será impossível.

# Capítulo 6

## Novas Perspectivas

Pretendemos, neste momento, fazer alguns comentários a respeito de suposições feitas sobre as hipóteses do teorema de Lyusternik. Deixamo-las para o final, porque foram motivos para novos rumos tomados nesta área de otimização, mediante o formalismo até aqui discutido. Ainda mais, essas considerações, que faremos em algumas hipóteses do teorema de Lyusternik, são no momento novas perspectivas para o surgimento de condições de extremais.

Observemos que o teorema de Dubovitskii-Milyutin pode ser satisfeito para problemas de otimização com restrições arbitrárias de igualdade em  $Q_{n+1}$ . Se  $Q_{n+1}$  é dado por  $\{y \in Y : \mathbb{F}(y) = 0\}$ , onde  $\mathbb{F}$  é um operador do espaço de Banach  $Y$  no espaço de Banach  $Z$ , então o cone tangente associado a  $Q_{n+1}$  pode ser determinado usando o teorema de Lyusternik (página 37). Contudo, este teorema requer a regularidade do operador  $\mathbb{F}$  em pontos candidatos a serem ponto de ótimo, pois, caso contrário, nestes pontos  $\tilde{y}$  o multiplicador  $\lambda_0$  associado à função objetivo, pode ser nulo e assim impossibilitar a ligação da função a ser otimizada com as restrições do problema. E neste caso dizemos que a equação de Euler-Lagrange é degenerada.

Devido a isto, Avakov [4, 1985], tendo essa visão analítico-funcional para problemas com restrições de igualdade, generaliza o teorema de Lyusternik sobre a hipótese de duas vezes Fréchet diferenciabilidade do operador não regular  $\mathbb{F}$ . No que se refere aos problemas de otimização não regulares citamos o trabalho de Silva [37] que ilustra de maneira didática essa questão.

Todavia nos restam perguntas referentes às hipóteses de Lyusternik, cujas respostas ainda não sabemos. Quem sabe o leitor, que após ler nosso

trabalho, interessou-se no assunto corrente, poderá contribuir com soluções para as questões abaixo explicitadas:

- Se  $\text{Im } \mathbb{F}'(\tilde{y}) \equiv Z$  e  $\mathbb{F}$  não é estritamente diferenciável em  $\tilde{y}$ , será que ainda se pode caracterizar o cone tangente?
- Se  $\text{Im } \mathbb{F}'(\tilde{y}) \neq Z$  e  $\mathbb{F}$  não é estritamente diferenciável em  $\tilde{y}$ , como se pode tipificar o cone tangente?

Além disso, se  $\mathbb{F}'(\tilde{y})$  não é um epimorfismo, também não se sabe sobre a veracidade da caracterização do cone tangente:  $T_{\tilde{y}}Q_{n+1} = \text{Ker}\mathbb{F}'(\tilde{y})$ . Entretanto, nada impediu (ou impede) o uso de outros métodos para calcular o cone tangente, por causa da definição outrora especificada, a qual independe do operador  $\mathbb{F}$ .

A seguir citaremos alguns dos diversos trabalhos desenvolvidos no campo da otimização, que eventualmente servirão para o leitor tanto a título de curiosidade quanto de pesquisa, visto que uns se utilizam dos argumentos matemáticos os quais usamos e outros abordam os problemas extremais diferentemente a forma aqui exibida. De repente, com as idéias formalizadas neste trabalho, poder-se-á fazer outra análise dos resultados inerentes em cada um deles.

Começemos por mencionar o artigo de Ledzewicz [30], que trata da extensão do Princípio do Máximo Local para problemas de controle ótimo com tempo fixo, considerados no caso anormal, ou seja, problemas cujo clássico Princípio do Máximo de Pontryagin tem a forma degenerada, a qual independe do funcional a ser minimizado. Ledzewicz utilizou uma extensão do método de Dubovitskii-Milyutin para o caso não regular, decorrente da generalização por Avakov do teorema de Lyusternik. É interessante também o artigo de Avakov [5, 1988] dentro deste contexto.

Sobre problemas de controle nos referimos, além de [30] e [5, 1988], aos livros de Barnett [6], Shinnars [36] e o clássico de Pontryagin [34].

No que se refere aos problemas de otimização que não são de controle, cabe-nos apresentar alguns dos resultados recentemente publicados por Brandão [11, 1996] e [12, 1997]. O primeiro trabalho trata de condições necessárias e suficientes de otimalidade para problemas de programação não-linear com tempo contínuo, usando subdiferenciabilidade de análise convexa; o segundo é constituído por condições necessárias e suficientes do tipo



Karush-Tucker para programação não-suave multi-objetiva, onde é assumido que as funções são somente Lipschitz e invexas. Já na dissertação de mestrado de Primo [24] encontram-se condições necessárias e suficientes de otimalidade para problemas de programação matemática em dimensão finita e infinita, utilizando teoremas de alternativa do tipo Gordan, tanto para o caso finito quanto para o caso infinito dimensional, obtendo-se teoremas de Fritz-John e Karush-Kuhn-Tucker para ambos os casos. Temos como outra referência a abordagem feita por Mateus [32] para problemas de programação não-linear.

No contexto de algoritmos para problemas de programação linear ou não-linear e problemas de controle, indicamos, respectivamente, Martínez [31] e Craven [16]. Juntamente lembramos a pesquisa de Adami [1], na qual encontram-se ferramentas numéricas para a determinação do controle ótimo utilizando o algoritmo BOX de programação não-linear.

A nível de exemplificação, sugerimos ao leitor o livro de Craven [16], no qual constam-se aplicações do Princípio do Máximo a outros problemas concretos, como por exemplo, controle de uma epidemia.

No âmbito da história da matemática fizemos nos rodapés apenas uma breve citação biográfica de alguns matemáticos, no entanto na bibliografia constam-se os livros [9], [27] e [29] que são de agradável leitura para quem se interessar. Agora, indicamos Dieudonne [18] por se tratar da história da análise funcional.

Sem mais a acrescentar, acredito que o leitor dar-me-á a honra de saber que, pelo menos via uma rápida leitura desta dissertação, também se encantara com a elegância da prova do Princípio do Máximo Local de Pontryagin, em virtude da estrutura matemática intrínseca a Dubovitskii e Milyutin, ou ainda, com os ramos da matemática através dos quais nos referimos como instrumento de trabalho e os quais deram-lhe o desejo de enveredar neles.

Termo compartilhando *a alegria que há em meu coração por vencer mais esta etapa em minha vida, por causa da Sabedoria e Misericórdia recebidas de Deus...*

# Bibliografia

- [1] ADAMI, A. **Resolução de problemas de controle ótimo utilizando o algoritmo BOX**. Campinas, IMECC-UNICAMP, 1996. (Tese Mestrado).
- [2] ALEXEÉV, V.; FOMINE, S; TIKHOMIROV, V. **Commande optimale**. Moscou, Mir, 1982.
- [3] ALEXEÉV, V.; GALEÉV, E.; TIKHOMIROV, V. **Recueil de problèmes d'optimisation** Moscou, Mir, 1987.
- [4] AVAKOV, E. R. Extremum conditions for smooth problems with equality: Type constraints. **USSR Computational Mathematics and Mathematical Physics**. 25 (5): 680-693, 1985.
- [5] AVAKOV, E. R. Necessary conditions for minimum of nonregular problems in Banach spaces: Maximum principle for abnormal problems of optimal control. **Trudy Matematicheskogo Instituta, Akademii Nauk SSSR**. 185: 3-29, 1988.
- [6] BARNETT, S.; CAMERON, R. G. **Introduction Mathematical Control Theory**. Oxford, Clarendon Press, 1984. (Oxford applied mathematics and computing science series).
- [7] BARTLE, R. G. **The elements of integration**. New York, John Wiley, 1966.
- [8] BECKER-CONTESSÉ, L. **Introducción a la optimización con restricciones**. Santiago, Departamento de Ingeniería Matemática, 1994.
- [9] BELL, E. T. **Men of mathematics**. New York, Simon and Schuster, 1937.

- [10] BLUM-RUCKSTUHL, E. Teoría control óptimo: El principio del máximo. In: **V Colóquio da sociedade matemática peruana**. Trujillo, Sociedad Matemática Peruana, 1987.
- [11] BRANDÃO, A. J. V.; ROJAS-MEDAR, M. A.; SILVA, G. N. **Condições necessárias y suficientes de optimalidade en programación no-diferenciável con tiempo contínuo: Caso convexo**. Campinas, IMECC-UNICAMP, 1996. (Relatório de Pesquisa, 65).
- [12] BRANDÃO, A. J. V.; ROJAS-MEDAR, M. A.; SILVA, G. N. **Optimality conditions for Pareto nonsmooth nonconvex programming in Banach spaces**. Campinas, IMECC-UNICAMP, 1997. (Relatório de Pesquisa, 17).
- [13] BREZIS, H. **Analyse fonctionnelle: Théorie et applications**. Paris, Masson, 1987. (Collection mathématiques appliquées pour la maîtrise).
- [14] CARTAN, H. **Calcul différentiel**. Paris, Hermann, 1967.
- [15] CHEN, C. T. **Linear system theory and design**. New York, Holt, Rinehart and Winston, 1984.
- [16] CRAVEN, B. D. **Control and optimization**. London, Chapman & Hall, 1995.
- [17] DEIMLING, K. **Nonlinear functional analysis**. Berlin, Springer, 1985.
- [18] DIEUDONNE, J. A. **History of functional analysis**. Amsterdam, North-Holland, 1981.
- [19] EDWARDS, R.E. **Functional analysis: Theory and applications**. New York, Dover, 1995.
- [20] FERNANDEZ, P. J. **Medida e integração**. Rio de Janeiro. IMPA, 1976.
- [21] FLETT, T.M. **Differential analysis**. Cambridge, University Press, 1980.

- [22] GIRSANOV, I. V. **Lectures on mathematical theory of extremum problems.** New York, Springer, 1972. (Lectures notes in economics and mathematical systems, 67).
- [23] GIUSSANI, L. **Em busca do rosto do homem.** São Paulo, Companhia Ilimitada, 1996.
- [24] PRIMO, P. G. **Condições de otimalidade para problemas finito e infinito dimensionais: Abordagem via teorema de alternativa do tipo Gordan.** Campinas, IMECC-UNICAMP, 1997. (Tese Mestrado).
- [25] HALMOS, P. R. **Measure theory.** Princeton, D. Van Nostrand, 1965.
- [26] IOFFE, A. D.; TIHOMIROV, V. M. **Theory of extremal problems.** Amsterdam, North-Holland, 1979.
- [27] KAC, M. **Enigmas of chance.** New York, Harper & Row, 1985.
- [28] KOLMOGOROV, A. N.; FOMIN, S.V. **Elementos da teoria das funções e de análise funcional.** Moscou, Mir, 1982.
- [29] LACAZ NETTO **Pequena história de grandes matemáticos.** São José dos Campos, Equilab, 1973.
- [30] LEDZEWICZ, U. Extension of the local maximum principle to abnormal optimal control problems. **Journal of optimization theory and applications.** 77 (3): 661-681, 1993.
- [31] MARTÍNEZ, J. M.; SANTOS, S.A. Métodos computacionais de otimização. In: **20º Colóquio Brasileiro de Matemática.** Rio de Janeiro, IMPA, 1995.
- [32] MATEUS, G. R.; LUNA, H. P. L. **Programação não-linear.** Belo Horizonte, UFMG, 1986.
- [33] NEUSTADT, L. W. **Optimization a theory of necessary conditions.** New Jersey, Princeton University Press, 1976.

- [34] PONTRYAGIN, L. W.; BOLTYANSKII, V. G.; GAMKRELIDZE, R. V.; MISHENKO, E. F. **The mathematical theory of optimal process.** New York, Interscience, 1962.
- [35] RUDIN, W. **Functional Analysis.** New York, McGraw-Hill, 1991.
- [36] SHINNERS, M. S. **Modern control system theory and application.** Massachusetts, Addison-Wesley, 1978.
- [37] SILVA, A. L. C. R. da **Caracterização funcional de pontos ótimos para problemas de otimização, regulares e não-regulares.** Campinas, IMECC-UNICAMP, 1996. (Tese Mestrado).
- [38] YOSIDA, K. **Lectures on differential and integral equations.** New York, Interscience, 1960.
- [39] YOSIDA, K. **Functional analysis.** Berlim, Springer, 1978.
- [40] ZEIDLER, E. **Nonlinear functional analysis and its applications: Variational methods and Optimization.** Berlin, Springer, 1989. Vol III.