UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

Instituto de Biociências, Letras e Ciências Exatas
DEPARTAMENTO DE CIÊNCIAS DE COMPUTAÇÃO E ESTATÍSTICA

Análise Não-Diferenciável e Condições Necessárias de Otimalidade para Problema de Controle Ótimo com Restrições Mistas

Reginaldo César Izelli

Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada

Rua Cristóvão Colombo, 2265
15054-000 - São José do Rio Preto - SP - Brasil
Telefone: (017) 3221-2444
Fax: (017) 3221-2445

Análise Não-Diferenciável e Condições Necessárias de Otimalidade para Problema de Controle Ótimo com Restrições Mistas

Reginaldo César Izelli [1]

Dissertação a ser apresentada no Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto, São Paulo, para a obtenção do título de Mestre em Matemática Aplicada.

Orientador: Prof. Dr. Geraldo Nunes Silva [2]

São José do Rio Preto

2006

[1]contato: reginaldoizelli@yahoo.com.br

[2]contato: gsilva@ibilce.unesp.br

Izelli, Reginaldo César.

Análise não-Diferenciável e condição necessária de otimalidade para problema de controle ótimo com restrições mistas / Reginaldo César Izelli. -São José do Rio Preto: [s.n.], 2006.

114f. : il. ; 30cm.

Orientador: Geraldo Nunes Silva

Dissertação (mestrado) – Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Análise matemática. 2. Otimização matemática. 3. Controle ótimo. 4. Princípio do Máximo. 5. Análise não-diferenciável. 6. Restrições Mistas. I. Silva, Geraldo Nunes. II. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU - 517

Análise Não-Diferenciável e Condições Necessárias de Otimalidade para Problema de Controle Ótimo com Restrições Mistas

Reginaldo César Izelli

Banca Examinadora

Titulares

Prof. Dr. Geraldo Nunes Silva - Orientador (IBILCE/UNESP)

Profa. Dra. Vilma Alves de Oliveira (EESC/USP)

Prof. Dr. Masayoshi Tsuchida (IBILCE/UNESP)

Suplentes

Prof. Dr. Silvio Alexandre de Araujo (IBILCE/UNESP)

Prof. Dr. Roberto Andreani (IMECC/UNICAMP)

São José do Rio Preto

2006

Agradecimentos

A Deus, por tudo.

Aos meus pais Odair e Zilda, que sempre me apoiarão.

À minha irmã Giseli, ao meu cunhado Rogério e a minha afilhada Maria Eduarda.

À minha namorada Mirian e sua filha Ana Cássia, pelo apoio e paciência.

Aos meus professores de graduação e pós-graduação.

Em especial ao Prof. Dr. Geraldo Nunes Silva, pela amizade, orientação, incentivo e paciência na elaboração deste trabalho.

A todas as pessoas e funcionários do IBILCE que, direta ou indiretamente, contribuíram para a elaboração deste trabalho.

Aos meus amigos de república e da minha cidade pelos momentos de descontração.

Aos meus amigos do curso de mestrado.

À CNPq e a CAPES, pelo auxílio financeiro.

Resumo

Estamos interessados em estudar uma generalização do Princípio do Máximo de Pontryagin para problema de controle ótimo com restrições mistas envolvendo funções não-diferenciáveis, pois este princípio não se aplica para todos os tipos de problemas. O principal objetivo deste trabalho é apresentar as condições necessárias de otimalidade na forma do princípio do máximo que serão aplicadas para o problema de controle ótimo com restrições mistas envolvendo funções não-diferenciáveis. Para alcançar este objetivo apresentamos estudos sobre cones normais e cones tangentes os quais são utilizados no desenvolvimento da teoria de subdiferenciais. Após esse embasamento formulamos o problema de controle ótimo envolvendo funções não-diferenciáveis, e apresentamos as condições necessárias de otimalidade.

Palavras-chave: Controle Ótimo, Princípio do Máximo, Análise Não-diferenciável, Restrições Mistas.

Abstract

We are interested in study a generalization of the Pontryagin Maximum Principle for optimal control problems with mixed constraints involving nondifferentiable functions, because this principle can not be applied for all the types of problems. The main objective of this work is to present the necessary conditions of optimality in the form of the maximum principle that will be applied for the optimal control problem with mixed constraints involving nondifferentiable functions. To achieve this objective we present studies above normal cones and tangent cones which are used in the development of the theory of subdifferentials. After this foundation we formulate the optimal control problem involving nondifferentiable functions, and we present the necessary conditions of optimality.

Keywords: optimal control, maximum principle, nonsmooth analysis, mixed constraints.

Sumário

Lista de Figuras

Capítulo 1

Introdução

A teoria de controle ótimo pode ser considerada como uma generalização do Cálculo das Variações, que foi iniciado em 1696 quando John Bernoulli desafiou "as melhores mentes do mundo" com o Problema da Braquistócrona[1](do grego: brachist = menor e chronos = tempo). O cálculo das variações foi desenvolvido por vários matemáticos, dos quais podemos destacar os nomes de Leonard Euler, Jean Louis Lagrange, Adrien Marien Legendre, Gustav Jacobi, Willian Hamilton, Karl Weierstrass, Alfred Clebsch, Oskar Bolza, Gilbert Bliss (veja, por exemplo, [5, 32]).

Em 1957, Richard Bellman desenvolveu o conceito de Programação Dinâmica que pode ser usado para resolver problemas de controle ótimo [3] e no ano 1960, Later Kalman [22] resolveu o problema com dinâmica linear e função de custo integral quadrática (conhecido por Regulator Linear Quadrático), mostrando que o controle ótimo é um "*feedback*" linear. Dois anos mais tarde, em 1962, o russo Lev Semenovich Pontryagin e outros colaboradores [26] formularam condições necessárias de otimalidade para problemas de controle não-linear, conhecida como Princípio do Máximo. Esta condição foi generalizada de muitas maneiras, em particular para permitir a possibilidade de adicionar funções não-diferenciáveis no problema. Algumas destas generalizações são baseadas em Análise Não-diferenciável, iniciadas nos anos 70 por Frank H. Clarke [7].

Em princípio, nos anos 50 e 60, a teoria de controle ótimo foi aplicada na indústria aeroespacial [20]. Atualmente, é aplicada na indústria avançada de robótica [24], na economia

[1]Sejam A e B dois pontos dados sobre um plano vertical, mas não na mesma reta vertical. O problema é de encontrar uma curva unindo estes dois pontos tal que uma massa m, partindo de A, a percorra sob a influência de seu próprio peso e alcance B no menor tempo possível.

[6], na engenharia química [1], dentre outros.

Estimulado pelo trabalho de Pinho, Vinter e Zheng [16], nesta dissertação estudamos as condições necessárias na forma de Princípio do Máximo para problema de controle ótimo com restrições mistas envolvendo funções não-diferenciáveis. Para isto, apresentamos o estudo da análise não-diferenciável.

Assim, a dissertação está organizada em 6 capítulos. No Capítulo 2, apresentamos o Problema de Controle Ótimo Clássico, as condições necessárias de otimalidade e a existência de solução para o problema.

Nos Capítulos 3 e 4, apresentamos a Análise Não-diferenciável, sendo que o Capítulo 3, será dedicado para os estudos dos cones normais, cones tangentes e suas relações e o Capítulo 4 para as subdiferenciais e suas propriedades.

No Capítulo 5, estudamos algumas condições necessárias de otimalidade na forma de Princípio do Máximo para problema de controle ótimo com funções não-diferenciáveis.

O Capítulo 6 é reservado para considerações finais sobre o trabalho desenvolvido e possíveis extensões para o problema de controle ótimo.

O Apêndice A contém conceitos fundamentais que auxiliam a compreensão dos estudos desenvolvidos nos capítulos relacionados acima, e a seguir apresentamos a bibliografia utilizada neste trabalho.

Capítulo 2

Teoria de Controle Ótimo

Neste capítulo, apresentamos o problema de controle ótimo clássico, e também analisamos a existência de solução e as condições necessárias de otimalidade para problema de controle ótimo. Os resultados apresentados são bem conhecidos na literatura da teoria de controle ótimo (veja [8], [25] e [12]).

2.1 Problema de Controle Ótimo

O Problema de Controle Ótimo Clássico (P) pode ser formulado da seguinte forma:

$$(P)\begin{cases} \text{Minimize} \quad l(x(a), x(b)) + \int_a^b L(t, x(t), u(t))dt \\ \text{sujeito a} \\ \begin{array}{rcll} \dot{x}(t) & = & f(t, x(t), u(t)) & \text{q.t. } t \in [a, b] \\ 0 & = & b(t, x(t), u(t)) & \text{q.t. } t \in [a, b] \\ u(t) & \in & U(t) & \text{q.t. } t \in [a, b] \\ (x(a), x(b)) & \in & C & \end{array} \end{cases}$$

sendo dadas as funções $l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $b : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ e os conjuntos $U \subset \mathbb{R}^m$ e $C \subset \mathbb{R}^n \times \mathbb{R}^n$. Também, abreviamos o termo "para quase todo" por q.t..

Então, o problema (P) consiste de:

- Função Custo: $l(x(a), x(b)) + \int_a^b L(t, x(t), u(t))dt$;

- Equação de Estado: é a equação diferencial $\dot{x}(t) = f(t, x(t), u(t))$ para quase todo $t \in [a, b]$;

- Função de Controle: é uma função mensurável $u(t)$ tal que $u(t) \in U(t)$ para quase todo $t \in [a, b]$;

- Função Trajetória de Estado (correspondente a uma função de controle u): é uma função absolutamente contínua $x(t)$, que satisfaz $\dot{x}(t) = f(t, x(t), u(t))$ e $0 = b(t, x(t), u(t))$ para quase todo $t \in [a, b]$.

Chamamos de processo o par (x, u), em que x é uma trajetória correspondente à uma função de controle u. O processo é admissível quando (x, u) satisfaz todas as restrições do problema (P). Dizemos que um processo (x, u) é minimizador, denotado por $(\bar{x}, \bar{u})$, se este processo minimiza a função custo sujeitas todas as restrições do problema (P).

Quando procuramos uma solução do problema de controle ótimo, podemos especificar se estamos interessados em examinar a solução global ou a solução local. No caso da solução local, podemos dividir entre minimizador (forte) local ou minimizador fraco local, pois a solução consiste de uma função de estado x e de uma função de controle u.

Definição 2.1. *Seja (x, u) um processo admissível do problema (P). Então, temos as seguintes definições:*

*Um processo $(\bar{x}, \bar{u})$ é chamado de **minimizador local** se, e somente se, existir algum $\epsilon > 0$ que minimiza o custo sobre todo processo (x, u) de (P) e satisfaz*

$$\|x(t) - \bar{x}(t)\| < \epsilon.$$

para quase todo $t \in [t_0, t_1]$.

*Um processo $(\bar{x}, \bar{u})$ é chamado de **minimizador fraco local** se, e somente se, existir algum $\epsilon > 0$ que minimiza o custo sobre todo processo (x, u) de (P) e satisfaz*

$$\|x(t) - \bar{x}(t)\| < \epsilon \text{ e } \|u(t) - \bar{u}(t)\| < \epsilon,$$

para quase todo $t \in [t_0, t_1]$.

2.2 Existência de Solução

A existência de solução do problema de controle ótimo é muito importante para garantir a coerência das condições de otimalidade (veja [19],[35]).

Seja

$$\tilde{f}(t,x,u) = \begin{pmatrix} f(t,x,u) \\ b(t,x,u) \\ L(t,x,u) \end{pmatrix},$$

para o problema (P).

Consideramos o seguinte conjunto de hipóteses:

1. $\tilde{f}$ é limitado em conjuntos limitados, mensurável em t e contínua em (x,u),

2. C e U são conjuntos compactos,

3. l é uma função contínua,

4. $\tilde{f}$ é uma função lipschitziana com constante Lipschitz K para todo $t \in [a,b]$, isto é,

$$\|\tilde{f}(t,x,u) - \tilde{f}(t,y,u)\| \leq K\|x-y\| \; \forall \; x,y \in \mathbb{R}^n \quad \forall \; u \in U,$$

5. o "conjunto velocidade" $\tilde{f}(t,x,U) = \{\tilde{f}(t,x,u) : u \in U\}$ é convexo para todo $(t,x) \in [a,b] \times \mathbb{R}^n$.

As hipóteses 2 e 3 (condições de compacidade e continuidade) são escolhidas naturalmente para garantir a existência da solução do problema (P), baseado no Teorema de Weiestrass "Toda função contínua definida num conjunto compacto de $\mathbb{R}^n$, atinge seu máximo e seu mínimo neste conjunto". A propriedade de mensurabilidade e de Lipschitz da função velocidade f, garante que a trajetória é definida unicamente pelo controle.

Teorema 2.1. *[12, Teorema 2.3.1] Seja (x,u) um processo admissível. Se todas as hipóteses 1 à 5 são satisfeitas, então existe um processo ótimo.*

2.3 Condições Necessárias de Otimalidade

Nesta seção, apresentamos as condições necessárias de otimalidade para problema de controle ótimo. Este resultado é conhecido na literatura como *Princípio do Máximo de Pontryagin* (veja [26] e [8]).

Substituímos as hipóteses 1 e 3 por:

1'. a função $\tilde{f}(t,x,u)$ é mensurável na variável t, contínua na variável u, continuamente diferenciável na variável x e $\tilde{f}, \tilde{f}_x$ são limitadas no conjunto limitado;

3'. a função l é contínua e continuamente diferenciável.

Defina a função Hamiltoniana da forma

$$H(t,x,p,q,u) := \big\langle p(t), f(t,x,u)\big\rangle + \big\langle q(t), b(t,x,u)\big\rangle - \lambda L(t,x,u).$$

Teorema 2.2. ***(Princípio do Máximo de Pontryagin)*** *Seja $(\bar{x},\bar{u})$ um processo ótimo para o problema (P). Suponha que as hipóteses 2 , 4, 5, $1'$ e $3'$ são satisfeitas. Então, existe uma função absolutamente contínua $p : [a,b] \to \mathbb{R}^n$, uma função $q : [a,b] \to \mathbb{R}^n$ e um escalar $\lambda \geq 0$, que satisfazem:*

(i) Condição não-triviabilidade

$$\lambda + \|p\| = 1,$$

(ii) Equação diferencial de estado (Equação Euler-Lagrange)

$$-\dot{p}(t) = \nabla_x H(t, \bar{x}(t), p(t), q(t), \bar{u}(t)) \text{ para quase todo } t \in [a,b],$$

(iii) Condição de otimalidade (Maximização Hamiltoniana)

$$H(t,\bar{x}(t),p(t),q(t),\bar{u}(t)) = \max_{u\in U(t)} \big\{H(t,\bar{x}(t),p(t),q(t),u)\big\},$$

para quase todo $t \in [a,b]$,

(iv) Condição de Transversalidade

$$-p(b) = \lambda \nabla_x l(\bar{x}(a), \bar{x}(b)).$$

Observe que a função p e o escalar λ são ambos não-nulos, pelo fato da condição não-triviabilidade.

A forma mais geral da demonstração do Princípio do Máximo de Pontryagin é muito longa, porém existem várias interpretações para estes resultados. Alguns autores os consideraram como uma extensão da condição de Euler-Lagrange no cálculo das variações (veja, por exemplo, [25]), outros exploram a relação da Programação Dinâmica (veja [17]). Também, alguns autores aplicam a condição Kuhn-Tucker para reformulação do problema de controle ótimo como problema de programação matemática [33].

Vamos apresentar uma demonstração do Princípio do Máximo de Pontryagin considerando o problema de controle ótimo mais simples possível, da seguinte forma

$$(P')\begin{cases} \text{Minimize} \quad l(x(0),x(1)) \\ \text{sujeito a} \\ \qquad \dot{x}(t) = f(t,x(t),u(t)) \quad \text{q.t. } t\in[0,1] \\ \qquad u(t) \in U(t) \quad \text{q.t. } t\in[0,1] \\ (x(0),x(1)) \in C. \end{cases}$$

Antes de iniciar a demonstração precisamos de alguns resultados.

Proposição 2.1. *[12, Proposição 3.1.2] Seja uma função de controle u. Suponha que as hipóteses 2 , 4, 5, $1'$ e $3'$ são satisfeitas. Para qualquer $t\in[0,1]$ e $\alpha\in\mathbb{R}^n$, defina $x(\cdot;t,\alpha)$ uma única solução para equação diferencial*

$$\begin{aligned} \dot{x}(s) &= f(s,x(s),u(s)) \quad q.t.\ s\in[0,1] \\ x(t) &= \alpha. \end{aligned}$$

Então, para t fixo, a função $\varphi(k) := x(1;t,\alpha)$ é continuamente diferenciável e

$$\nabla\varphi(k) = \Phi(1,t)$$

onde Φ é uma matriz de transição da equação de estado $\dot{x}=f$, linearizada sobre $x(\cdot;t,\alpha)$, isto é, Φ é unicamente definida pela seguinte propriedade: para cada $\tau\in[0,1]$,

$$\begin{aligned} \frac{d}{d\sigma}\Phi(\tau,\sigma) &= -\Phi(\tau,\sigma)f_x(\tau,x(\cdot;t,\alpha),u(\tau)) \quad q.t.\ \sigma\in[0,1] \\ \Phi(\tau,\tau) &= I. \end{aligned}$$

Também necessitamos a noção de um ponto de Lebesgue.

Definição 2.2. *Seja $I\subset\mathbb{R}$ um subintervalo aberto e $g: I\to\mathbb{R}^n$ uma função integrável. Um ponto $t\in I$ é chamado de ponto de Lebesgue de g se, para qualquer seqüencia $\{[s_i,t_i]\}$ de subintervalos não-triviais de I, tal que $s_i\uparrow t$ e $t_i\downarrow t$, temos*

$$\lim_{i\longrightarrow\infty}(t_i-s_i)^{-1}\int_{s_i}^{t_i} g(s)ds = g(t).$$

Um resultado importante sobre pontos de Lebesgue é que o conjunto de tal ponto para g é de medida completa.

Demonstração do Princípio do Máximo de Pontryagin em relação ao problema (P'):

Seja $(\bar{x}, \bar{u})$ um processo ótimo. Tome $u \in U$ um controle arbitrário e defina S um subconjunto de $(0, 1)$ contendo os pontos de Lebesgue de ambas funções $t \to f(t, \bar{x}(t), \bar{u}(t))$ e $t \to f(t, \bar{x}(t), v)$. Observe que S é um conjunto de medida completa, pois o conjunto S é a intersecção de dois conjuntos de medida completa. Escolha $t \in S$ e seja $\{\epsilon_i\}$ uma seqüencia de números em $(0, t)$ tal que $\epsilon_i \downarrow 0$.

Defina uma nova função de controle

$$u_i(s) = \begin{cases} \bar{u}(s) & \text{se } s \in [0,1] \backslash [t - \epsilon_i, t] \\ v & \text{se } t \in [t - \epsilon_i, t], \end{cases}$$

que é uma variação "muito pequena" do controle $\bar{u}$.

Denotamos por x_i a trajetória correspondente à função de controle u_i, isto é, a solução de equação $\dot{x}(s) = f(s, x(s), u_i(s))$ com condição inicial $x(0) = x_0$.

Para cada i, temos

$$\begin{aligned} \frac{1}{\epsilon_i}(x_i(t) - \bar{x}(t)) &= \frac{1}{\epsilon_i}\int_{t-\epsilon_i}^{t} f(s, x_i(s), v) - f(s, \bar{x}(s), \bar{u}(s))ds \\ &= \frac{1}{\epsilon_i}\int_{t-\epsilon_i}^{t} \big[f(s, x_i(s), v) - f(s, \bar{x}(s), v) + f(s, \bar{x}(s), v) - f(s, \bar{x}(s), \bar{u}(s))\big]ds. \end{aligned}$$

Segue da propriedade de f e da definição de x_i, que

$$\lim_{i\to\infty} \frac{1}{\epsilon_i}\int_{t-\epsilon_i}^{t} f(s, x_i(s), v) - f(s, \bar{x}(s), v)ds = 0.$$

Como t é um elemento de S podemos concluir que

$$\lim_{i\to\infty} \frac{1}{\epsilon_i}(x_i(t) - \bar{x}(t)) = f(t, \bar{x}(t), v) - f(t, \bar{x}(t), \bar{u}(t)).$$

Esta última propriedade pode ser expressa de outra maneira: defina

$$y_i := \frac{1}{\epsilon_i}(x_i(t) - \bar{x}(t)) \ i = 1, 2, ...$$

Então,

$$x_i(t) = \bar{x}(t) + \epsilon_i y_i \ i = 1, 2, ...$$

e

$$y_i \to f(t, \bar{x}(t), v) - f(t, \bar{x}(t), \bar{u}(t)).$$

Observe que u_i e $\bar{u}$ coincidem em $[t, 1]$. Temos também que para função continuamente diferenciável $F : \mathbb{R}^n \to \mathbb{R}^k$,

$$\lim_{i \longrightarrow \infty} \frac{1}{\epsilon_i}[F(\alpha + \epsilon_i \eta_i) - F(\alpha)] = \nabla F(\alpha)(\eta)$$

onde α e η são vetores dados e $\{\eta_i\}$ e $\{\epsilon_i\}$ são seqüencias em $\mathbb{R}^n$ e $\mathbb{R}^+$, respectivamente, tais que $\eta_i \longrightarrow \eta$ e $\epsilon_i \downarrow 0$. Substituindo α por $\bar{x}(t)$ e η_i por y_i segue da Proposição 2.1 que

$$\lim_{i \longrightarrow \infty} \frac{x_i(1) - \bar{x}(1)}{\epsilon_i} = \Phi(1, t)[f(t, \bar{x}(t), v) - f(t, \bar{x}(t), \bar{u}(t))]. \tag{2.1}$$

Agora, (x_i, u_i) é um processo admissível e $(\bar{x}, \bar{u})$ é um minimizador. Assim, $l(x_i(1)) - l(\bar{x}(1)) \geq 0, \ i = 1, 2, ...$ Como $\epsilon_i > 0$ também temos

$$\frac{l(x_i(1)) - l(\bar{x}(1))}{\epsilon_i} \geq 0, \ i = 1, 2, ... \tag{2.2}$$

Tome $x_i(t) = \bar{x}(t) + \epsilon_i \dfrac{x_i(1) - \bar{x}(1)}{\epsilon_i}$ e l é continuamente diferenciável numa vizinhança de $\bar{x}(t)$, passamos o limite em (2.2) e usando a equação (2.1) obtemos

$$l_x(\bar{x}(t))\Phi(1, t)[f(t, \bar{x}(t), v) - f(t, \bar{x}(t), \bar{u}(t))] \geq 0, \tag{2.3}$$

para cada $t \in S$. Assim, a desigualdade (2.3) é válida para q.t. $t \in [0, 1]$.

Defina

$$p(t) = -\Phi^T(1, t) l_x(\bar{x}(1)) \text{ para todo } t \in [0, 1].$$

Então,

$$-\dot{p}(t) \quad = \quad -f_x^T(t, \bar{x}(t), \bar{u}(t))\Phi^T(1, t) l_x(\bar{x}(1)) = -f_x^T(t, \bar{x}(t), \bar{u}(t)) p(t).$$

Reescrevendo esta expressão obtemos

$$-\dot{p}(t) = \nabla_x H(t, \bar{x}(t), p(t), \bar{u}(t)) \text{ q.t. } t \in [0, 1]$$

Segue de (2.3) que

$$H(t, \bar{x}(t), p(t), \bar{u}(t)) = \max_{v \in U} H(t, \bar{x}(t), p(t), v) \text{ q.t. } t \in [0, 1],$$

Isto completa a demonstração. ∎

O exemplo a seguir ilustra um caso em que o Princípio do Máximo de Pontryagin não se aplica e ainda mostra que uma generalização direta deste Princípio não é possível.

Exemplo 2.1. *Considera o problema*

$$\begin{cases} \textit{Minimize} \quad x(1) \\ \textit{Sujeita a} \\ \qquad \begin{aligned} \dot{x}(t) &= u_1 \\ 0 &= x - u_2^2 \\ U(t) &= [-1,+1] \times [-1,+1] \\ C &= \{0\} \times \mathbb{R} \end{aligned} \end{cases}$$

Observe que o minimizador deste problema é $(x, u_1, u_2) \equiv (0,0,0)$. *Seja a função Hamiltoniana dada por*

$$H(t, x(t), p(t), q(t), u_1(t), u_2(t)) = p(t)u_1(t) + q(t)(x(t) - u_2^2(t)).$$

Se o Princípio do Máximo de Pontryagin for válido para este problema, então existem os multiplicadores $p(\cdot)$, $q(\cdot)$ *e* $\lambda \geq 0$ *que satisfazem as seguintes condições:*

Equação diferencial de estado:

$$-\dot{p}(t) \;=\; \frac{\partial}{\partial_x} H(t, x(t), p(t), q(t), u_1(t), u_2(t)) = q(t)$$

Condição de otimalidade:

$$H(t, \bar{x}(t), p(t), q(t), \bar{u}_1(t), \bar{u}_2(t)) = \max_{(u_1,u_2)\in U(t)} H(t, \bar{x}(t), p(t), q(t), u_1(t), u_2(t)).$$

Segue de $(\bar{x}, \bar{u}_1, \bar{u}_2) = (0,0,0)$ *que*

$$0 = \max_{(u_1,u_2)\in U(t)} \{p(t)u_1 - q(t)u_2^2\}.$$

Condição de transversalidade: *Como* $C = \{0\} \times \mathbb{R}$, *temos*

$$-p(b) \;=\; \lambda \nabla_x h(\bar{x}(a), \bar{x}(b)) = \lambda.$$

Calculando as equações de Euler-Lagrange, temos

$$\begin{aligned} \frac{\partial}{\partial u_1} H(t, x(t), p(t), q(t), u_1(t), u_2(t)) &= 0 \Rightarrow p(t) = 0 \textit{ para todo } t, \\ \frac{\partial}{\partial u_2} H(t, x(t), p(t), q(t), u_1(t), u_2(t)) &= 0 \Rightarrow -2q(t)u_2(t) = 0. \end{aligned}$$

Assim, a única solução é $(p(t), q(t), \lambda) = (0,0,0)$. *Isto contradiz a condição de não-triviabilidade. Portanto, o Princípio do Máximo de Pontryagin não se aplica para este problema.*

Como as condições necessárias de otimalidade clássicas não são válidas para todo tipo de problema, podemos identificar uma classe de problemas para a qual o Princípio do Máximo seja aplicado. Neste estudo, estamos interessados em problemas de controle ótimo com restrições mistas envolvendo funções não-diferenciáveis no custo e na equação diferencial. Para suprir esta dificuldade, vamos apresentar a forma forte do Princípio do Máximo, que é aplicado ao minimizador local. Mas, para isso será necessário estudar as subdiferenciais para funções semicontínuas inferiores.

Capítulo 3

Cones Normais e Cones Tangentes

Neste capítulo, definimos os cones normais, cones tangentes para um conjunto e suas principais relações. Este estudo será necessário para o entendimento da subdiferencial. Algumas referências podem ser encontradas em Rockafellar & Wets [29] e Vinter [34].

3.1 Cones Normais

Para definir cones normais, precisamos saber o significado de "ponto mais próximo": Seja $C \subset \mathbb{R}^k$ um conjunto fechado e não-vazio e um ponto $z \notin C$, logo existe pelo menos um ponto $x \in C$ que é "mais próximo" de z, isto é, tal que

$$\|z - y\| \geq \|z - x\| \ \forall y \in C. \tag{3.1}$$

O *ponto mais próximo* em C de z é chamado de projeção de z em C e é denotado por $proj_C(z)$. Com isso, podemos chamar o ponto x como *ponto base* de C em relação a z.

Na figura abaixo, temos as seguintes projeções:

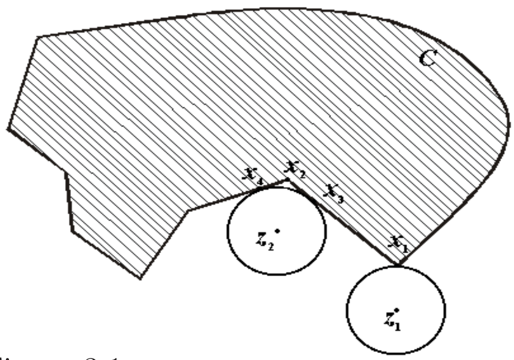

Figura 3.1: Pontos de fronteira do conjunto C.

Temos agora condição de relacionar o conceito de ponto mais próximo com o vetor normal. Elevando ao quadrado em ambos os lados a desigualdade (3.1) e reescrevendo na forma de produto interno, temos

$$2\langle z, y - x\rangle \quad \leq \quad \langle y, y\rangle - \langle x, x\rangle \ \forall y \in C \tag{3.2}$$

Somando e subtraindo o termo $2\langle x, y - x\rangle$, no lado direto na desigualdade (3.2), obtemos

$$\langle w, y - x\rangle \leq \frac{1}{2}\|y - x\|^2 \ \forall y \in C, \tag{3.3}$$

onde $w = z - x$.

O vetor w é chamado de perpendicular ao conjunto C em x. Se $w = z - x$ é perpendicular a C em x, então qualquer múltiplo não-negativo $\xi = t(z - x)$, $t \geq 0$ é chamado de normal proximal a C em x.

De (3.3), podemos generalizar o conceito de vetor normal proximal e juntamente podemos definir o cone normal proximal.

Definição 3.1. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. O vetor $\xi \in \mathbb{R}^k$ é chamado de vetor **normal proximal** a C em x, se existir $M \geq 0$ tal que*

$$\langle \xi, y - x\rangle \leq M\|y - x\|^2 \ \forall \ y \in C. \tag{3.4}$$

*O cone de todos os vetores normais proximais a C em x é chamado de **cone normal proximal** a C em x, denotado por $N_C^p(x)$, e é representado pelo conjunto*

$$N_C^p(x) := \Big\{\xi \in \mathbb{R}^k : \exists M > 0 \ \textit{tal que a Condição (3.4) é satisfeita}\Big\} \tag{3.5}$$

Temos a seguinte proposição:

Proposição 3.1. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e os pontos $x \in C$ e $\eta \in \mathbb{R}^k$. Então, as seguintes afirmações são equivalentes:*

(i) η é um normal proximal a C em x;

(ii) existe um ponto $z \in \mathbb{R}^k$ e um fator multiplicativo $r > 0$ tal que

$$\|z - x\| = \min\big\{\|z - y\| : y \in C\big\} \ \textit{e} \ \eta = r(z - x)$$

Demonstração: Vamos mostrar a implicação da condição (i) para a condição (ii). Se $\eta \in N_C^p(x)$, por definição de normal proximal, existe $M > 0$ tal que

$$\langle \eta, y - x \rangle \leq M||y - x||^2.$$

Dividindo em ambos os lados por $\dfrac{1}{2M}$ e reajustando a desigualdade, temos

$$\frac{1}{2M}\langle \eta, y - x \rangle \leq \frac{1}{2}\left\langle \frac{\eta}{M} - y + x, y - x \right\rangle \tag{3.6}$$

Antes de prosseguir a demonstração, vamos calcular a seguinte norma,

$$\begin{aligned} \|x + \xi - y\|^2 &= \|y - x\|^2 - 2\langle y - x, \xi \rangle + \|\xi\|^2 \\ &= 2\Big(\frac{1}{2}\|y - x\|^2 - \langle \xi, y - x \rangle + \frac{1}{2}\|\xi\|^2\Big), \end{aligned} \tag{3.7}$$

sendo que $\dfrac{1}{2}\|y - x\|^2 - \langle \xi, y - x \rangle \geq 0$ por (3.3). Logo,

$$\|x + \xi - y\| \geq \|\xi\| \ \forall y \in C. \tag{3.8}$$

Substituindo $\xi = \dfrac{1}{2M}\eta$ em (3.8), temos

$$\|x + (2M)^{-1}\eta - y\| \geq (2M)^{-1}\|\eta\| \ \forall y \in C. \tag{3.9}$$

Também podemos substituir $r = 2M$ e $z = x + (2M)^{-1}\eta$ em (3.9), obtendo

$$\|z - y\| \geq r^{-1}\|r(z - x)\| = \|z - x\| \ \forall y \in C.$$

Assim,

$$\|z - x\| = \min\big\{\|z - y\| : y \in C\big\} \text{ e } \eta = r(z - x).$$

Agora, vamos mostrar a implicação inversa. Sejam $z \in \mathbb{R}^k$ e um fator multiplicativo $r > 0$ tal que

$$\|z - x\| = \min\big\{\|z - y\| : y \in C\big\} \text{ e } \eta = r(z - x).$$

De (3.3) sabemos que $z - x \in N_C^p(x)$ e

$$\langle z - x, y - x \rangle \leq \frac{1}{2}\|y - x\|^2 \ \forall y \in C.$$

Como $\eta = r(z - x)$, então

$$\langle r(z - x), y - x \rangle \leq \frac{r}{2}\|y - x\|^2 \ \forall y \in C$$

Segue da Definição 3.5 que $\eta \in N_C^p(x)$. ■

Seguindo nosso estudo, podemos nos perguntar: o que acontece, se o cone normal proximal não existe para um ponto pertencente ao conjunto C? Dados $x \in C$ e $z \in \mathbb{R}^k$ tal que $x \notin proj_C(z)$. Pela definição de normal proximal, temos que $N_C^p(x) = \emptyset$ (veja o ponto x_2 na Figura 3.1, onde encontramos $N_C^p(x_2) = \emptyset$). Mas, podemos notar que nos pontos vizinhos de x_2 temos $N_C^p(x_i) \neq \emptyset$, onde x_i é o ponto na vizinhança de x_2.

Com este fato, temos a motivação de definir um outro tipo de vetor, para explicar este acontecimento, chamado de vetor normal limite, sendo estudado nos pontos vizinhos para os pontos onde não existem dos vetores normais proximais.

A notação $x_i \xrightarrow{C} x$ significa que os elementos da seqüencia x_i pertencem a C e convergem a x.

Definição 3.2. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Um vetor $\xi \in \mathbb{R}^k$ é um vetor **normal limite** a C em x se existir seqüencia $x_i \xrightarrow{C} x$ e $\xi_i \longrightarrow \xi$ tal que*

$$\xi_i \in N_C^p(x_i) \quad \textit{para todo } i$$

*O cone de vetores normais limites a C em x é chamado de **cone normal limite** a C em x, denotado por $N_C(x)$, é o conjunto:*

$$N_C(x) := \{\xi \in \mathbb{R}^k : \exists \ x_i \xrightarrow{C} x \ \ e\ \xi_i \longrightarrow \xi \ \textit{tal que}\ \xi_i \in N_C^p(x_i) \ \ \textit{para todo}\ i\}.$$

Uma representação geométrica do vetor normal limite é apresentada abaixo:

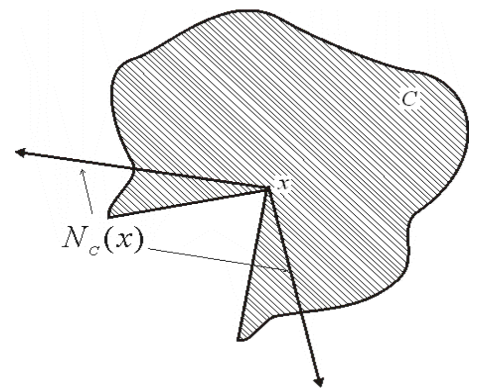

Figura 3.2: Vetores Normais Limite.

Um outro cone normal interessante é o cone normal estrito.

Definição 3.3. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Um vetor $\xi \in \mathbb{R}^k$ é um vetor **normal estrito** a C em x se satisfazer a seguinte desigualdade*

$$\langle \xi, y - x \rangle \leq o(\|y - x\|) \ \textit{para todo}\ y \in C, \tag{3.10}$$

para alguma função $o(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ *tal que* $\lim\limits_{\epsilon \downarrow 0} \big(o(\epsilon)/\epsilon\big) = 0$.
O ***cone normal estrito****, denotado por* $\hat{N}_C(x)$*, é o conjunto*

$$\hat{N}_C(x) := \left\{ \xi \in \mathbb{R}^k : \limsup_{y \xrightarrow{C} x} \frac{\langle \xi, y - x \rangle}{\|y - x\|} \leq 0 \right\}.$$

Aqui $\limsup\limits_{y \xrightarrow{C} x} f(y)$ significa que o limite superior é tomado somente para seqüencias convergentes a x, cujos elementos estão em C.

Na proposição a seguir, temos as relações entre os cones normais. A demonstração é omitida, pois segue diretamente das definições dos cones normais.

Proposição 3.2. *Seja um conjunto fechado* $C \subset \mathbb{R}^k$ *e um ponto* $x \in C$*. Então:*

(i) $N_C^p(x)$*,* $\hat{N}_C(x)$ *e* $N_C(x)$*, são todos cones em* $\mathbb{R}^k$*, contendo* $\{0\}$ *e*

$$N_C^p(x) \subset \hat{N}_C(x) \subset N_C(x);$$

(ii) $N_C^p(x)$ *é convexo (mas, possivelmente não fechado);*

(iii) $\hat{N}_C(x)$ *é fechado e convexo;*

(iv) o mapeamento do conjunto de valores $y \to N_C(y) : C \to \mathbb{R}^k$ *tem um gráfico fechado. Dadas algumas seqüencias* $y_i \xrightarrow{C} y$ *e* $\eta_i \longrightarrow \eta$ *tais que* $\eta_i \in N_C(y_i)$ *para todo* i*, temos* $\eta \in N_C(y)$*.*

Para o cone normal limite vale a seguinte propriedade:

Proposição 3.3. *Seja* $C \subset \mathbb{R}^k$ *um conjunto fechado e os pontos* $x \in C$ *e* $\xi \in \mathbb{R}^k$*. Então, as seguintes afirmações são equivalentes:*

(i) $\xi \in N_C(x)$*;*

(ii) existe uma seqüencia $x_i \xrightarrow{C} x$ *e* $\xi_i \longrightarrow \xi$ *tal que* $\xi_i \in \hat{N}_C(x_i)$ *para todo* i*.*

Demonstração: Vamos mostrar a implicação da condição (i) para a condição (ii). Se $\xi \in N_C(x)$, logo existe uma seqüencia $x_i \xrightarrow{C} x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in N_C^p(x_i)$ para todo i. Como, $N_C^p(x_i) \subset \hat{N}_C(x_i)$, segue que $\xi_i \in \hat{N}_C(x_i)$.

Mostremos a implicação inversa. Fixe $q \in \hat{N}_C(y)$ para algum $y \in C$. É suficiente mostrar que existem seqüencias $y_i \xrightarrow{C} y$ e $q_i \longrightarrow q$ tal que $q_i \in N_C^p(y_i)$ para todo i.

Tome algum $\varepsilon_i \downarrow 0$. Então, para cada i, podemos encontrar o ponto y_i que é um ponto mais proximo de $(y + \varepsilon_i q)$ em C. Como $\varepsilon_i q \longrightarrow 0$, segue da Proposição 3.1 (ii) que $y_i \longrightarrow y$.

Para cada i, definindo

$$q_i := \varepsilon_i^{-1}(y + \varepsilon_i q - y_i),$$

então $q_i \in N_C^p(y_i)$. Para completar a demonstração, precisamos mostrar que a seqüencia $\varepsilon_i^{-1}(y - y_i) \longrightarrow 0$.

Pela propriedade de ponto mais próximo de y_i, temos que

$$\|(y + \varepsilon_i q) - y_i\|^2 \leq \|(y + \varepsilon_i q) - z\|^2 \text{ para todo } z \in C. \tag{3.11}$$

Escolhendo $z = y$ e substituindo em (3.11), obtemos

$$\|(y + \varepsilon_i q) - y_i\|^2 \leq \varepsilon_i^2 \|q\|^2. \tag{3.12}$$

Por outro lado, calculando a norma $\|(y + \varepsilon_i q) - y_i\|^2$ e substituindo na equação (3.12), segue que

$$\|y_i - y\|^2 \leq 2\varepsilon_i \langle y_i - y, q \rangle.$$

Como $q \in \hat{N}_C(x)$, temos

$$\|y_i - y\|^2 \leq 2\varepsilon_i o(\|y_i - y\|), \tag{3.13}$$

para alguma função $o(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ tal que $\lim_{\epsilon \downarrow 0} \big(o(\epsilon)/\epsilon\big) = 0$. Dividindo por $\dfrac{1}{\varepsilon_i^2}$ a relação (3.13), obtemos

$$\left(\frac{1}{\varepsilon_i}\|y_i - y\|\right)^2 \leq 2\frac{o(\|y_i - y\|)}{\varepsilon_i}.$$

Ajustando, podemos obter

$$\frac{\|y_i - y\|}{\varepsilon_i} \leq \sqrt{2}\sqrt{\frac{o(\|y_i - y\|)}{\|y_i - y\|}}.$$

Passando o limite para $i \longrightarrow \infty$, segue que

$$\lim_{i \longrightarrow \infty} \frac{1}{\varepsilon_i}\|y_i - y\| \leq 0.$$

Portanto, concluímos que $\varepsilon_i^{-1}\|y_i - y\| \longrightarrow 0$ quando $i \longrightarrow \infty$. ∎

Pelas próprias definições de cones normais, temos as seguintes observações:

Observação 3.1. *Em qualquer ponto x de um conjunto $C \subset \mathbb{R}^k$, temos:*

a) O cone normal limite é relacionado ao cone normal proximal, da seguinte forma:

$$N_C(x) := \limsup_{y \xrightarrow{C} x} N_C^p(y).$$

b) A Proposição 3.3 fornece uma relação, do cone normal limite com o cone normal estrito:

$$N_C(x) := \limsup_{y \xrightarrow{C} x} \hat{N}_C(y).$$

Propriedade do cone normal limite em pontos interiores de C ou pontos na fronteira de C.

Proposição 3.4. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então, satisfazem as seguintes afirmações:*

(i) $x \in int\{C\}$ implica que $N_C(x) = \{0\}$ (e portanto $N_C^P(x) = \hat{N}_C(x) = \{0\}$);

(ii) $x \in bdy\{C\}$ implica que $N_C(x)$ contém elementos diferentes de zeros.

Demonstração:

(i) Se $x \in int\{C\}$, então pela Proposição 3.1 temos que todos normais proximal em pontos a C próximo de x são nulos. Assim $N_C(x) = \{0\}$.

(ii) Tomemos um ponto $x \in bdy\{C\}$, então, existe uma seqüencia $x_i \longrightarrow x$ tal que $x_i \notin C$ para todo i.

Para cada i, seleciona c_i o ponto mais próximo de x_i em C, logo

$$x_i - c_i \neq 0 \text{ e } \xi_i := \|x_i - c_i\|^{-1}(x_i - c_i)$$

como o vetor normal proximal a C em c_i de norma unitária.

Então

$$\|x - c_i\| = \|x - x_i + x_i - c_i\| \leq \|x - x_i\| + \|x_i - c_i\|,$$

mas, pela Proposição 3.1, segue que

$$\|x - c_i\| \leq 2\|x - x_i\|.$$

Como $x_i \xrightarrow{C} x$, temos $c_i \xrightarrow{C} x$.

Analogamente, também temos uma subseqüência $\xi_i \longrightarrow \xi$, onde ξ tem norma unitária. Portanto, $\xi \in N_C(x)$, completando assim a demonstração. ■

Temos a seguinte propriedade:

Proposição 3.5. *Sejam os subconjuntos fechados* $C_1 \subset \mathbb{R}^m$*,* $C_2 \subset \mathbb{R}^n$*,e o ponto* $(x_1, x_2) \in C_1 \times C_2$*. Então,*

$$\begin{aligned} N^p_{C_1\times C_2}(x_1,x_2) &= N^p_{C_1}(x_1) \times N^p_{C_2}(x_2) \\ \hat{N}_{C_1\times C_2}(x_1,x_2) &= \hat{N}_{C_1}(x_1) \times \hat{N}_{C_2}(x_2) \\ N_{C_1\times C_2}(x_1,x_2) &= N_{C_1}(x_1) \times N_{C_2}(x_2) \end{aligned}$$

Demonstração: Ver no livro Vinter [34, Proposição 4.2.8]. ∎

Quando consideramos o conjunto C como sendo o conjunto convexo, temos um caso particular da Proposição 3.2.

Proposição 3.6. *Seja* $C \subset \mathbb{R}^k$ *um conjunto fechado convexo e um ponto* $\bar{x} \in C$*. Então,*

$$\begin{aligned} N^p_C(\bar{x}) &= \hat{N}_C(\bar{x}) = N_C(\bar{x}) \\ &= \Big\{\xi : \big\langle \xi, x - \bar{x}\big\rangle \leq 0 \textit{ para todo } x \in C\Big\}. \end{aligned}$$

Demonstração: Tome um ponto $y \in C$ e defina

$$S(y) := \Big\{\xi : \big\langle \xi, y' - y\big\rangle \leq 0 \text{ para todo } y' \in C\Big\}.$$

Dividimos a demonstração em três partes:

1) Nesta primeira parte da demonstração, vamos mostrar a relação $S(\bar{x}) = N^p_C(\bar{x})$. Com as definições de $S(\cdot)$ e $N^p_C(\cdot)$, podemos concluir que para qualquer $\bar{x} \in C$, temos $S(\bar{x}) \subset N^p_C(\bar{x})$.

Por outro lado, se $\xi \in N^p_C(\bar{x})$, então existe $M \geq 0$ tal que

$$\big\langle \xi, x' - \bar{x}\big\rangle \leq M\|x' - \bar{x}\|^2 \text{ para todo } x' \in C. \tag{3.14}$$

Segue da convexidade do conjunto C, que podemos escolher algum ponto $x \in C$ tal que

$$x' = \epsilon x + (1-\epsilon)\bar{x} \in C \text{ para algum } \epsilon \in (0,1). \tag{3.15}$$

Substituindo (3.15) em (3.14), obtemos

$$\big\langle \xi, x - \bar{x}\big\rangle \;\leq\; M\epsilon\|x - \bar{x}\|^2.$$

Passando o limite quando $\epsilon \downarrow 0$, obtemos

$$\left\langle \xi, x - \bar{x} \right\rangle \leq 0.$$

Logo, $\xi \in S(\bar{x})$. Portanto, $S(\bar{x}) \supset N_C^p(\bar{x})$. Assim, concluímos que $S(\bar{x}) = N_C^p(\bar{x})$.

2) Nesta parte, vamos mostrar $S(\bar{x}) = N_C(\bar{x})$. Seja uma multifunção semicontínua inferior $S : C \rightrightarrows \mathbb{R}^k$, onde $y \mapsto S(y)$ para todo $y \in dom(S)$. Logo, para algum $y \in dom(S)$, para qualquer seqüencias convergentes $y_i \xrightarrow{C} y$ e existe $\xi_i \longrightarrow \xi$ tal que $\xi_i \in S(y_i)$ para todo i. Assim $\xi \in S(y)$. Segue da definição de cone normal limite que

$$S(y) \subset N_C(y).$$

Por outro lado, toma $\xi \in N_C(y)$. Então existem seqüencias $y_i \xrightarrow{C} y$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in N_C^p(y_i)$. Segue do item 1) acima que $N_C^p(y_i) = S(y_i)$. Mas, como S é semicontínua inferior, podemos concluir que $\xi \in S(y)$.

3) Para terminar a demonstração falta verificar a igualdade $S(\bar{x}) = \hat{N}_C(\bar{x})$. Segue da Proposição (3.2)i, que

$$N_C^p(\bar{x}) \subset \hat{N}_C(\bar{x}) \subset N_C(\bar{x}) \text{ para todo } \bar{x} \in C.$$

Mas, como $S(\bar{x}) = N_C^p(\bar{x})$ e $S(\bar{x}) = N_C(\bar{x})$, obtemos

$$\hat{N}_C(\bar{x}) = S(\bar{x}).$$

■

3.2 Cones Tangentes

Nesta seção, definimos o vetor tangente e seus cones tangentes (Cone Tangente de Bouligand e Cone Tangente de Clarke).

Definição 3.4 (**Vetor Tangente**). *Seja $C \subset \mathbb{R}^k$ um conjunto não-vazio e $x \in C$ um ponto. Então, um ponto $\xi \in \mathbb{R}^k$ é chamado de **vetor tangente** se existirem seqüencias $\{c_i\} \subset C$ e $\{t_i\} \subset \mathbb{R}^+$ tal que*

$$c_i \longrightarrow x, \; t_i \downarrow 0, \; t_i^{-1}(c_i - x) \longrightarrow \xi.$$

Pela definição de vetor tangente, surgem duas definições de cones tangentes.

Definição 3.5. *Seja $C \subset \mathbb{R}^k$ um conjunto não-vazio e $x \in C$ um ponto.*

i) O ***Cone Tangente de Bouligand*** *a C em x, denotado por $T_C(x)$, é o conjunto*

$$T_C(x) := \limsup_{t \downarrow 0} t^{-1}(C - x).$$

Este conjunto é equivalente a

$$T_C(x) = \left\{ \xi \in \mathbb{R}^k : \exists\, c_i \xrightarrow{C} x \;\; e\; \exists t_i \downarrow 0 \;\, tal\ que\; t_i^{-1}(c_i - x) \longrightarrow \xi \right\}.$$

ii) O ***Cone Tangente de Clarke*** *a C em x, denotado por $\bar{T}_C(x)$, é o conjunto*

$$\bar{T}_C(x) := \liminf_{t \downarrow 0, y \xrightarrow{C} x} t^{-1}(C - y).$$

Este conjunto é equivalente a

$$\bar{T}_C(x) = \left\{ \xi \in \mathbb{R}^k : \forall\, x_i \xrightarrow{C} x \;, \forall t_i \downarrow 0 \;\exists \{c_i\} \in C \;\, tal\ que\; t_i^{-1}(c_i - x_i) \longrightarrow \xi \right\}.$$

Vamos afirmar a seguinte propriedade elementar, porém, sem demonstração (sua demonstração pode ser encontrada em Rockafellar [29, Teorema 6.26]).

Proposição 3.7. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então, $T_C(x)$ e $\bar{T}_C(x)$ são cones fechados contendo a origem e satisfaz a condição $\bar{T}_C(x) \subset T_C(x)$.*

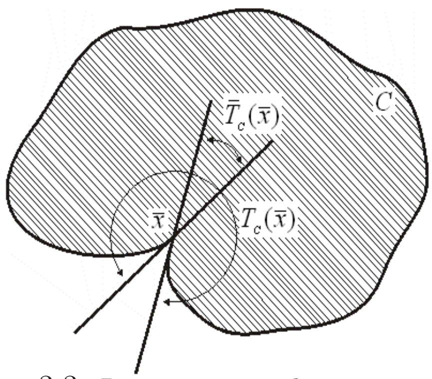

Figura 3.3: Representações dos cones tangentes.

Também, temos a seguinte propriedade.

Proposição 3.8. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então, $\bar{T}_C(x)$ é um conjunto convexo.*

Demonstração: Vamos tomar quaisquer vetores u e v em $\bar{T}_C(x)$. Para mostrar a convexidade, é suficiente mostrar que $u + v \in \bar{T}_C(x)$, pois $\bar{T}(x)$ é um cone.

Tomemos $u \in \bar{T}_C(x)$, logo qualquer seqüencia $x_i \xrightarrow{C} x$ e $t_i \downarrow 0$, existe uma seqüencia $\{x_i + t_i u_i\}$ em C para todo i tal que $t_i^{-1}\Big((x_i + t_i u_i) - x_i\Big) \longrightarrow u$, ou seja, $u_i \longrightarrow u$. Deste modo, se $v \in \bar{T}_C(x)$, então qualquer seqüencia $(x_i + t_i u_i) \xrightarrow{C} x$ e $t_i \downarrow 0$ existe uma seqüencia $\{(x_i + t_i u_i) + t_i v_i\}$ em C para todo i tal que

$$t_i^{-1}\Big((x_i + t_i u_i) + t_i v_i - (x_i + t_i u_i)\Big) \longrightarrow v,$$

ou seja, $v_i \longrightarrow v$.

Tomemos $c_i = x_i + t_i(u_i + v_i)$ em C, temos

$$\frac{1}{t_i}(c_i - x_i) \longrightarrow u + v.$$

Portando, $u + v \in \bar{T}_C(x)$. ■

3.3 Relação entre Cone Normal e Cone Tangente

Denotemos por S^* o cone polar de um conjunto $S \subset \mathbb{R}^k$, isto é,

$$S^* := \Big\{\eta \in \mathbb{R}^k : \big\langle \eta, x \big\rangle \leq 0 \text{ para todo } x \in S\Big\}.$$

Teorema 3.1. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então, o cone normal estrito $\hat{N}_C(x)$ e o cone tangente de Bouligand $T_C(x)$ são relacionados da seguinte maneira*

$$\hat{N}_C(x) = T_C(x)^*.$$

Demonstração: Para facilitar a demonstração, vamos dividir em duas partes:

a) Mostramos a relação $\hat{N}_C(x) \subset T_C(x)^*$.

Tome $\xi \in \hat{N}_C(x)$. Então

$$\big\langle \xi, y - x \big\rangle \leq o(\|y - x\|) \text{ para todo } y \in C, \tag{3.16}$$

para alguma função $o(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ que satisfaz $o(s)/s \longrightarrow 0$ quando $s \downarrow 0$. Escolhe um ponto $v \in T_C(x)$. Segue da definição de Cone Tangente de Bouligand, que existe $x_i \xrightarrow{C} x$ e $t_i \downarrow 0$ tal que, definindo $v_i := t_i^{-1}(x_i - x)$ para cada i, temos que $v_i \longrightarrow v$.

Afirmamos que $\big\langle \xi, v \big\rangle \leq 0$. Então, pela definição de cone polar, temos $\xi \in T_C(x)^*$.

Para completar a prova, precisamos verificar a afirmação. Se $v = 0$ é obvio a afirmação. Assim, suponha que $v \neq 0$. Para cada i, temos

$$\begin{aligned} \big\langle \xi, v_i \big\rangle &= \big\langle \xi, t_i^{-1}(x_i - x) \big\rangle = t_i^{-1} \big\langle \xi, x_i - x \big\rangle \leq t_i^{-1} o(\|x_i - x\|) \\ &\leq \frac{\|v_i\|}{\|x_i - x\|}\, o(\|x_i - x\|) \text{ por (3.16).} \end{aligned}$$

Passando o limite quando $i \longrightarrow \infty$, obtemos

$$\big\langle \xi, v \big\rangle \leq 0.$$

b) Mostremos a relação $\hat{N}_C(x) \supset T_C(x)^*$. Tomemos $\xi \notin \hat{N}_C(x)$, logo existe $\epsilon > 0$ e $x_i \xrightarrow{C} x$ tal que

$$\big\langle \xi, x_i - x \big\rangle > \epsilon \|x_i - x\| \text{ para todo } i.$$

É fácil notar que $x_i \neq x$ para cada i. Podemos considerar $t_i = \|x_i - x\|$ e $v_i := t_i^{-1}(x_i - x)$ (observe que todos os v_i tem comprimento unitário) para cada i. Então, uma subseqüência $v_i \longrightarrow v$ para algum $v \in T_C(x)$ com $\|v\| = 1$, temos

$$\big\langle \xi, v_i \big\rangle = t_i^{-1} \big\langle \xi, x_i - x \big\rangle > t_i^{-1} \epsilon \|x_i - x\| = \|x_i - x\|^{-1} \epsilon \|x_i - x\| = \epsilon.$$

Passando o limite, obtemos $\big\langle \xi, v \big\rangle > \epsilon$. Portanto, $\xi \notin T_C(x)^*$. ■

Também, temos.

Teorema 3.2. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então, o envoltório convexo fechado do cone tangente de Bouligand $T_C(x)$, denotado por $\overline{co} T_C(x)$, e o cone tangente de Clarke $\bar{T}_C(x)$ se relacionam da seguinte maneira:*

$$\liminf_{y \xrightarrow{C} x} \overline{co} T_C(y) = \liminf_{y \xrightarrow{C} x} T_C(y) = \bar{T}_C(x).$$

Demonstração: Como $T_C(y) \subset \overline{co} T_C(y)$ para cada y é suficiente mostrar as seguintes relações:

(i) $\liminf_{y \xrightarrow{C} x} \overline{co} T_C(y) \subset \bar{T}_C(x)$;

(ii) $\bar{T}_C(x) \subset \liminf_{y \xrightarrow{C} x} T_C(y)$

De fato, estas inclusões implicam que

$$\liminf_{y \xrightarrow{C} x} \overline{co} T_C(y) \subset \bar{T}_C(x) \subset \liminf_{y \xrightarrow{C} x} T_C(y) \subset \liminf_{y \xrightarrow{C} x} \overline{co} T_C(y).$$

Logo temos as igualdades desejadas.

Vamos mostrar a relação (i). Tome qualquer $v \in \liminf_{y \xrightarrow{C} x} \overline{co} T_C(y)$. Sejam as seqüencias arbitrárias $x_i \xrightarrow{C} x$ e $t_i \downarrow 0$. Vamos supor que podemos encontrar $\epsilon_i \downarrow 0$ tal que

$$t_i^{-1} d_C(x_i + t_i v) \leq \epsilon_i. \tag{3.17}$$

A demonstração de (3.17) será apresentada após a demonstração de (ii).

Logo,

$$\limsup_i t_i^{-1} d_C(x_i + t_i v) = 0.$$

Então, existem subseqüências de $\{t_i\}$ e $\{x_i\}$ (não reindexamos) tal que

$$\lim_i t_i^{-1} d_C(x_i + t_i v) = 0.$$

Seja $c_i \in C$ que satisfaz

$$\|x_i + t_i v - c_i\| \leq d_C(x_i + t_i v) + \frac{t_i}{i}.$$

Dividindo por t_i e passando o limite quando $i \longrightarrow \infty$, temos

$$\frac{(c_i - x_i)}{t_i} \longrightarrow v.$$

Portanto, pela definição de $\bar{T}_C(x)$, segue que $v \in \bar{T}_C(x)$.

Agora, vamos mostrar a relação (ii). Tomemos $v \in \bar{T}_C(x)$ e seja $x_i \xrightarrow{C} x$ uma seqüencia arbitrária, segue da definição de $\bar{T}_C(x)$, que existem $\epsilon_i \downarrow 0$, $t_i \downarrow 0$ e $N_i \uparrow \infty$ tal que

$$t^{-1} d_C(x_k + tv) \leq \epsilon_i \text{ para todo } 0 \leq t \leq t_i, \text{ e } k \geq N_i.$$

Reindexamos as seqüencias $\{x_{N_i}\}_{i=1}^{\infty}$ como $\{x_i\}$ e $\{t_{N_i}\}_{i=1}^{\infty}$ como $\{t_i\}$, temos em particular que

$$t^{-1} d_C(x_i + tv) \leq \epsilon_i,$$

para cada $i = 1, 2, ...$ e para todo $t \in [0, t_i]$.

Fixamos i e $t \in [0, t_i]$, podemos escolher $z_i(t) \in C$ tal que

$$d_C(x_i + tv) = \|x_i + tv - z_i(t)\|,$$

onde a d_C é a função distância. Definindo

$$v_i(t) := t^{-1}(z_i(t) - x_i) \text{ para todo } t \in (0, t_i], i = 1, 2, ..., \tag{3.18}$$

obtemos

$$\|v_i(t) - v\| \leq \epsilon_i \text{ para todo } t \in (0, t_i], i = 1, 2, ...$$

Fixe i e escolha um ponto de acumulação $\bar{v}_i$ de $\{v_i(t)\}_{t>0}$ em $t = 0$ (esse $\bar{v}_i$ existe uma vez que $v_i(t)$ é limitado em $(0, t_i]$), logo $\|\bar{v}_i - v\| \leq \epsilon_i$. De (3.18) e da definição de $T_C(x_i)$, segue que

$$\bar{v}_i \in T_C(x_i) \ \text{ e } \ \bar{v}_i \longrightarrow v.$$

Vamos retornar no início, quando temos a seqüencia arbitrária $\{x_i\}$ em C convergindo para x, então podemos encontrar um subseqüência $\{x_{i_n}\}_{n=1}^{\infty}$ e a seqüencia $\{\bar{v}_n\}$ tal que

$$\bar{v}_n \in T_C(x_{i_n}) \ \text{ para todo } \ n \ \text{ e } \ \bar{v}_n \longrightarrow v \ \text{ quando } \ n \longrightarrow \infty.$$

Portanto, $v \in \liminf\limits_{y \xrightarrow{C} x} T_C(y)$.

Para completar a demonstração, temos que mostrar a desigualdade (3.17). Para cada i, defina a função

$$g_i(t) := d_C(x_i + tv), \quad 0 \leq t \leq t_i.$$

Escolhe $z_i(t) \in \arg\min\{\|x_i + tv - z\| : z \in C\}$. Observe que, como $x_i \in C$, temos

$$\begin{aligned} \|x - z_i(t)\| &= \|x - z_i(t) + x_i + tv - x_i - tv\| \leq \|x_i + tv - z_i(t)\| + \|x_i - x + tv\| \\ &\leq \|x_i + tv - x_i\| + \|x_i - x + tv\| \text{ (pela propriedade mais próximo)} \\ &\leq 2t\|v\| + \|x_i - x\| \end{aligned}$$

para $0 \leq t \leq t_i$ e $i = 1, 2, ...$ Passando o limite, temos a seguinte equivalência

$$\sup_{0 \leq t \leq t_i} \|x - z_i(t)\| \longrightarrow 0 \ \text{ quando } \ i \longrightarrow \infty.$$

Como $v \in \liminf\limits_{y \xrightarrow{C} x} \overline{co} T_C(y)$, existe seqüencia $\epsilon_i \downarrow 0$ e funções $w_i(\cdot) : [0, t_i] \to \mathbb{R}^n$ tal que

$$w_i(t) \in \overline{co} T_C(z_i(t)) \quad \text{e} \quad \|w_i(t) - v\| \leq \epsilon_i \quad \text{para todo } t \in [0, t_i].$$

Fixe i. Verificamos (3.17) para a escolha acima de ϵ_i.

Vamos, primeiro investigar a propriedade de um ponto $t \in (0, t_i]$ em que $g_i > 0$ e onde g_i é diferenciável.

Para simplificar a notação, denota $z_i(t)$ por z e fazendo $p = x_i + tv$. Temos para $h > 0$ suficientemente pequeno,

$$\begin{aligned} g_i(t+h) - g_i(t) &\leq \|x_i + tv + hv - z\| - \|x_i + tv - z\| \\ &= \|p - z + hv\| - \|p - z\|. \end{aligned}$$

Como $p \neq z$ (relembra que $g_i(t) > 0$) a função $u \to \|p - z + u\|$ é diferenciável em todo ponto suficientemente próximo da origem, em particular em hv (para $h > 0$ suficientemente pequeno). Em vista da convexidade da norma, temos

$$\|p - z + hv\| - \|p - z\| \leq \frac{1}{\|p - z + hv\|}(p - z + hv) \cdot hv,$$

segue que,

$$g_i(t+h) - g_i(t) \leq h||p - z + hv||^{-1}(p - z + hv) \cdot v.$$

Concluímos que

$$\begin{aligned} h^{-1}(g_i(t+h) - g_i(t)) &\leq \left(\frac{p - z + hv}{\|p - z + hv\|} - \frac{p - z}{\|p - z\|} \right) \cdot v + \frac{p - z}{\|p - z\|} \cdot v \\ &\leq \left(\frac{p - z + hv}{\|p - z + hv\|} - \frac{p - z}{\|p - z\|} \right) \cdot v + \frac{p - z}{\|p - z\|} \cdot w_i + \epsilon_i. \end{aligned} \tag{3.19}$$

Como $p - z \in N_C^p(z)$ e $w_i \in \overline{co}T_C(z)$ e em vista do fato que

$$\hat{N}_C(z) = T_C(z)^* = (\overline{co}T_C(z))^*,$$

segue que

$$(p - z) \cdot w_i(t) \leq 0. \tag{3.20}$$

Logo, da desigualdade (3.19) e da (3.20), temos

$$h^{-1}(g_i(t+h) - g_i(t)) \leq \left(\frac{p - z + hv}{\|p - z + hv\|} - \frac{p - z}{\|p - z\|} \right) \cdot v + \epsilon_i,$$

e passando o limite quando $h \downarrow 0$, obtemos

$$\frac{d}{dt} g_i(t) \leq \epsilon_i.$$

Como g_i é uma função Lipschitziana e entretanto diferenciável em quase todo t, esta desigualdade é valida para quase todo ponto $t \in \{s \in [0, t_i) : g_i(s) > 0\}$ eliminado o caso em que $g_i(s) = 0$.

Se $g_i(t_i) = 0$, temos que a condição (3.17) é satisfeita. Suponhamos que $g_i(t_i) > 0$ e definamos

$$r := \inf\{t \in [0, t_i] : g_i(s) > 0 \text{ para todo } s \in [t, t_i]\}.$$

Como g_i são contínuas e $g_i(0) = 0$, concluímos que $g_i(r) = 0$. Segue que

$$g_i(t_i) = g_i(r) + \int_r^{t_i} \dot{g}_i(\sigma)d\sigma \leq 0 + t_i\epsilon_i.$$

Mas, então

$$t_i^{-1}g_i(t_i) = t_i^{-1}d_C(x_i + t_iv) \leq \epsilon_i.$$

Portanto, a condição (3.17) é satisfeita neste caso. Assim, a prova está completa. ∎

Lema 3.1. *Seja $S : \mathbb{R}^k \rightrightarrows \mathbb{R}^n$ uma multifunção, $C \subset \mathbb{R}^k$ um conjunto e $x \in \mathbb{R}^k$ um ponto. Suponha que $S(y)$ é um cone convexo fechado para todo y. Então,*

$$\liminf_{y \xrightarrow{C} x} S(y) = \left[\limsup_{y \xrightarrow{C} x} S(y)^*\right]^*.$$

Demonstração: Dividimos a demonstração em duas partes:

a) Mostremos a relação

$$\liminf_{y \xrightarrow{C} x} S(y) \subset \left[\limsup_{y \xrightarrow{C} x} S(y)^*\right]^*.$$

Tomemos $v \in \liminf_{y \xrightarrow{C} x} S(y)$ arbitrário. Escolha $\xi \in \limsup_{y \xrightarrow{C} x} S(y)^*$. Então existem $x_i \xrightarrow{C} x$ e $\{\xi_i\}$ em $\mathbb{R}^k$ tal que

$$\xi_i \in S(x_i)^* \quad \text{para todo } i \quad \text{e} \quad \xi_i \longrightarrow \xi.$$

Também podemos escolher seqüencia $v_i \longrightarrow v$ tal que $v_i \in S(y_i)$ para todo i. Segue que $\langle \xi_i, v_i \rangle \leq 0$ para cada i. Passando o limite, quando $i \longrightarrow \infty$, obtemos $\langle \xi, v \rangle \leq 0$. Então,

$$v \in \left[\limsup_{y \xrightarrow{C} x} S(y)^*\right]^*.$$

b) Mostremos que

$$\liminf_{y \xrightarrow{C} x} S(y) \supset \left[\limsup_{y \xrightarrow{C} x} S(y)^*\right]^*.$$

Fixe $v \in \left[\limsup_{y \xrightarrow{C} x} S(y)^*\right]^*$ e suponha que $v \notin \liminf_{y \xrightarrow{C} x} S(y)$. Então, existem algum $\epsilon > 0$ e uma seqüencia $x_i \xrightarrow{C} x$ tal que

$$B(v, \epsilon) \cap S(x_i) = \emptyset \text{ para todo } i.$$

Aqui B denota a bola fechada de centro v e raio ϵ.

Pelo Teorema da Separação apropriado, obtemos um vetor ξ_i tal que $\|\xi_i\| = 1$ e

$$\sup_{w \in S(x_i)} \big\langle w, \xi_i \big\rangle \leq \inf_{w \in B(v,\epsilon)} \big\langle w, \xi_i \big\rangle = \big\langle v, \xi_i \big\rangle - \epsilon,$$

para cada i. Como $S(x_i)$ é um cone, então $\sup_{w \in S(x_i)} \big\langle w, \xi_i \big\rangle = 0$ e assim $\xi_i \in S(x_i)^*$. Deste modo, uma subseqüência $\xi_i \longrightarrow \xi$ para algum ξ com $\|\xi\| = 1$. Portanto, $\xi \in \limsup\limits_{y \xrightarrow{C} x} S(y)^*$.

Podemos observar que $0 \leq \big\langle v, \xi_i \big\rangle - \epsilon$. Passando o limite, segue que $\big\langle v, \xi_i \big\rangle \geq \epsilon$. Concluindo que $v \notin \left[\limsup\limits_{y \xrightarrow{C} x} S(y)^* \right]^*$, contradizendo a suposição. ■

Teorema 3.3. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então, existe uma relação entre o cone tangente de Clarke $\bar{T}_C(x)$ e o cone normal limite $N_C(x)$ da seguinte forma*

$$\bar{T}_C(x) = N_C(x)^*.$$

Demonstração: Pelo Teorema 3.2, temos

$$\bar{T}_C(x) = \liminf_{y \xrightarrow{C} x} \overline{co} T_C(y).$$

Segue do Lema 3.1, que

$$\bar{T}_C(x) = \liminf_{y \xrightarrow{C} x} \overline{co} T_C(y) = \left[\limsup_{y \xrightarrow{C} x} \overline{co} T_C(y)^* \right]^*.$$

Sabendo que cone polar de um cone qualquer coincide com o cone polar do fecho de envoltório conexo do cone, isto é, $\overline{co} T_C(y)^* = T_C(y)^*$. Logo do Teorema 3.1, podemos concluir que

$$\bar{T}_C(x) = \left[\limsup_{y \xrightarrow{C} x} \hat{N}_C(y) \right]^*.$$

Então, pela Proposição 3.3, obtemos

$$\bar{T}_C(x) = N_C(x)^*.$$

Completando a demonstração. ■

O diagrama abaixo tem um resumo das relações entre os cones normais e os cones tangentes associados com um conjunto fechado $C \subset \mathbb{R}^k$ e o ponto $x \in C$ que foram encontradas nesta seção.

$$\begin{tikzcd}
T_C(x) \arrow[r, "\liminf"] \arrow[d, "*"] & \bar{T}_C(x) & \\
\hat{N}_C(x) \arrow[r, "\limsup"] & N_C(x) \arrow[u, "*"'] & N_C^p(x); \arrow[l, "\limsup"']
\end{tikzcd}$$

onde * significa que foram tomados os cones polares do conjunto.

Capítulo 4

Subdiferenciais

Seja f uma função de valores reais de classe $\mathcal{C}^1$ (continuamente diferenciável). Então, pelo cálculo clássico, podemos relacionar o gradiente da função com o vetor normal do seu gráfico no ponto $(x, f(x))$, denotado pelo vetor $\lambda(\nabla f(x), -1)$ para $\lambda > 0$.

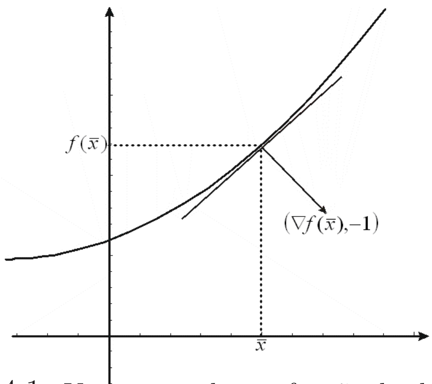

Figura 4.1: Vetor normal para função de classe $\mathcal{C}^1$.

Neste capítulo, apresentamos a noção generalizada de gradiente para funções não necessariamente diferenciáveis e algumas propriedades. Este assunto pode ser encontrado em Rockafellar & Wets [29], Clarke [9], Clarke et al. [10], Loewen [25] e Vinter [34].

4.1 Subdiferencial Proximal, Estrita e Limite

Uma noção de derivada generalizada em termos de cones normais para o conjunto epígrafo de uma função f, não necessariamente diferenciável em todo seu domínio, será apresentada nesta seção. Para isso precisamos de uma classe de funções com o conjunto epígrafo fechado. Assim, a escolha natural é a classe de funções semicontínuas inferiores $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ que possuem o conjunto epígrafo fechado (veja a Proposição A.5).

Definição 4.1. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e $x \in dom(f)$ um ponto. Então:*

(i) A subdiferencial proximal de f em x, denotada por $\partial^p f(x)$, é o conjunto

$$\partial^p f(x) := \Big\{\xi : (\xi, -1) \in N^p_{epi(f)}(x, f(x))\Big\}.$$

Os elementos em $\partial^p f(x)$ são chamados de subgradientes proximais.

(ii) A subdiferencial estrita de f em x, denotada por $\hat{\partial} f(x)$, é o conjunto

$$\hat{\partial} f(x) := \Big\{\xi : (\xi, -1) \in \hat{N}_{epi(f)}(x, f(x))\Big\}.$$

Os elementos em $\hat{\partial} f(x)$ são chamados de subgradientes estritos.

(iii) A subdiferencial limite de f em x, denotada por $\partial f(x)$, é o conjunto

$$\partial f(x) := \Big\{\xi : (\xi, -1) \in N_{epi(f)}(x, f(x))\Big\}.$$

Os elementos em $\partial f(x)$ são chamados de subgradientes limites.

Observação 4.1. *Temos as seguintes observações:*

(i) Pela Proposição 3.2, obtemos

$$\partial^p f(x) \subset \hat{\partial} f(x) \subset \partial f(x).$$

(ii) Estas subdiferenciais fornecem informações sobre uma função f na vizinhança de um ponto x.

Vejamos alguns exemplos:

Exemplo 4.1. *No caso abaixo, considera a função $f : \mathbb{R} \to \mathbb{R}$.*

a) $f(x) = |x|$. Então,

$$\partial^p f(0) = \hat{\partial} f(0) = \partial f(0) = [-1, 1].$$

b) $f(x) = -|x|$. Então,

$$\partial^p f(0) = \hat{\partial} f(0) = \emptyset \text{ e } \partial f(0) = \{-1\} \cup \{1\}.$$

c) $f(x) = |x|^{\frac{1}{2}}$. *Então,*

$$\partial^p f(0) = \hat{\partial} f(0) = \partial f(0) = (-\infty, +\infty).$$

d) $f(x) = sgn\{x\}|x|^{\frac{1}{2}}$. *Então,*

$$\partial^p f(0) = \hat{\partial} f(0) = \partial f(0) = \emptyset.$$

Podemos observar que os exemplos c) e d) possuem subdiferenciais ilimitada e vazia, respectivamente. Este comportamento significa que as inclinações das retas tangentes a função na vizinhança de zero podem assumir valores muito grandes. Com isso, precisamos novas definições para suprir o caso em que a inclinação é ilimitada. Denomina-se por subdiferenciais "assintóticas".

Definição 4.2. *Seja* $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ *uma função semicontínua inferior e* $x \in dom(f)$ *um ponto. Então:*

(i) A subdiferencial proximal assintótica de f *em* x*, denotada por* $\partial_p^\infty f(x)$*, é o conjunto*

$$\partial_p^\infty f(x) := \Big\{ \xi : (\xi, 0) \in N^p_{epi(f)}(x, f(x)) \Big\}.$$

Os elementos em $\partial_p^\infty f(x)$ *são chamados de subgradientes proximais assintóticas.*

(ii) A subdiferencial estrita assintótica de f *em* x*, denotada por* $\hat{\partial}^\infty f(x)$*, é o conjunto*

$$\hat{\partial}^\infty f(x) := \Big\{ \xi : (\xi, 0) \in \hat{N}_{epi(f)}(x, f(x)) \Big\}.$$

Os elementos em $\hat{\partial}^\infty f(x)$ *são chamados de subgradientes estritas assintóticas.*

(iii) A subdiferencial limite assintótica de f *em* x*, denotada por* $\partial^\infty f(x)$*, é o conjunto*

$$\partial^\infty f(x) := \Big\{ \xi : (\xi, 0) \in N_{epi(f)}(x, f(x)) \Big\}.$$

Os elementos em $\partial^\infty f(x)$ *são chamados de subgradientes limites assintóticas.*

Observe no Exemplo 4.1 d), que $\partial f(0) = \emptyset$. No entanto, pela definição de subdiferencial limite assintótica, temos $\partial^\infty f(0) = [0, +\infty)$(veja a Figura 4.1). Assim, o cálculo de $\partial^\infty f(0)$ mostra que a inclinação "infinita" próxima do ponto zero é positiva.

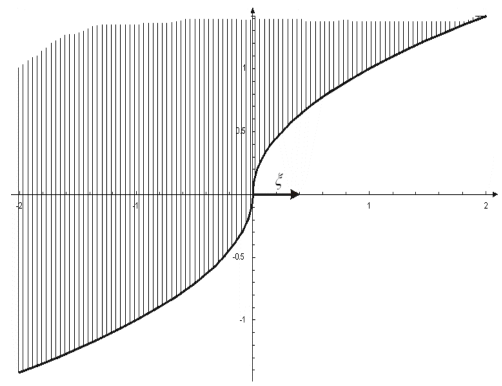

Figura 4.2: Subgradiente limite assintótica para $f(x) = \text{sgn}\{x\}|x|^{\frac{1}{2}}$.

A subdiferencial limite e a subdiferencial limite assintótica juntas fornecem informações completas para o cone normal para o epígrafo da função.

Proposição 4.1. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ uma função semicontínua inferior e $x \in dom(f)$ um ponto. Então:*

$$N_{epi(f)}(x, f(x)) = \Big\{(\lambda\xi, -\lambda) : \lambda > 0, \xi \in \partial f(x)\Big\} \cup \Big(\partial^\infty f(x) \times \{0\}\Big). \tag{4.1}$$

$\partial f(x)$ é um conjunto não-vazio ou $\partial^\infty f(x)$ contém elementos não-nulos.

Demonstração: A relação (4.1) segue imediatamene das definições de cone normal limite e de subdiferenciais limite e limite assintótica. Como $(x, f(x))$ é um ponto da fronteira do conjunto $epi(f)$ segue da Proposição 3.4 que $N_{epi(f)}(x, f(x))$ contém elementos não-nulos. O fato de que $\partial f(x)$ é não-vazio ou $\partial^\infty f(x)$ contém elementos não-nulo é uma conseqüência imediata. ∎

A próxima proposição diz que a subdiferencial limite e a subdiferencial limite assintótica possuem gráficos fechados. A demonstração é uma conseqüência direta da Proposição 3.2 (*iv*).

Proposição 4.2. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e $x \in dom(f)$ um ponto. Logo:*

(i) $\partial f(x)$ é um conjunto fechado. Dado uma seqüencia $x_i \xrightarrow{f} x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial f(x_i)$ para todo i, então $\xi \in \partial f(x)$;

(ii) $\partial^\infty f(x)$ é um cone fechado. Dado uma seqüencia $x_i \xrightarrow{f} x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial^\infty f(x_i)$ para todo i, então $\xi \in \partial^\infty f(x)$.

A notação $x_i \xrightarrow{f} x$ significa que x_i convergem a x e $f(x_i)$ convergem a $f(x)$.

Para a função convexa temos a seguinte propriedade.

Proposição 4.3. *Sejam $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e convexa e um ponto $\bar{x} \in dom(f)$. Então*

$$\begin{aligned} \partial^p f(\bar{x}) &= \hat{\partial} f(\bar{x}) = \partial f(\bar{x}) \\ &= \Big\{ \xi : \big\langle \xi, x - \bar{x} \big\rangle \leq f(x) - f(\bar{x}) \text{ para todo } x \in \mathbb{R}^k \Big\}. \end{aligned}$$

Demonstração: Sabemos que

$$\partial^p f(\bar{x}) \subset \hat{\partial} f(\bar{x}) \subset \partial f(\bar{x}).$$

Resta mostrar a inclusão inversa. Tome um vetor $\xi \in \partial f(\bar{x})$, então $(\xi, -1) \in N_{epi(f)}(\bar{x}, f\bar{x})$. Como o epígrafo de f é convexa segue da Proposição 3.6, que

$$N_{epif}(\bar{x}, f\bar{x}) = \hat{N}_{epif}(\bar{x}, f\bar{x}) = N^p_{epif}(\bar{x}, f\bar{x}).$$

Daí segue a inclusão inversa.

Seja $\xi \in \partial^p f(x)$ arbitrário. Então, $(\xi, -1) \in N^p_{epi(f)}(x, f(x))$. Segue da Proposição 3.6, que

$$\Big\langle (\xi, -1), (x, \alpha) - (\bar{x}, f(\bar{x})) \Big\rangle \leq 0 \text{ para todo } (x, \alpha) \in epi(f).$$

Tomando $f(x) = \alpha$ completa a demonstração. ■

4.2 Representação por Quociente de Diferença

Anteriormente as subdiferenciais foram definidas através do vetor normal ao conjunto epígrafo. Nesta seção, vamos representar as subdiferenciais em relação ao quociente de diferença. Esta mudança tem como objetivo facilitar a representação geométrica e simplificar as demonstrações.

Vamos iniciar com a caracterização do subgradiente proximal.

Proposição 4.4. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e os pontos $x \in dom(f)$ e $\xi \in \mathbb{R}^k$. Então as seguintes afirmações são equivalentes.*

(i) $\xi \in \partial^p f(x)$;

(ii) existem $M > 0$ e $\epsilon > 0$ tais que

$$\big\langle \xi, y - x \big\rangle \leq f(y) - f(x) + M\|y - x\|^2, \tag{4.2}$$

para todo $y \in B(x; \epsilon)$.

Demonstração: Mostremos a implicação da condição (i) para a condição (ii). Tome $\xi \in \partial^p f(x)$, segue que $(\xi, -1) \in N^p_{epi(f)}(x, f(x))$. Na figura abaixo, podemos visualizar uma representação geométrica do subgradiente proximal usando a normal proximal a um conjunto epígrafo. Sejam x um ponto fixo e y um ponto arbitrário em $B(x, \eta)$.

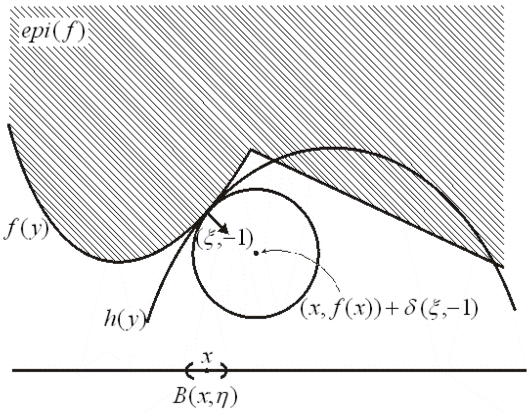

Figura 4.3: Representação geométrica do subgradiente proximal.

Segue da noção de ponto mais próximo (veja a Figura 4.3), que

$$\left\| \Big[(x, f(x)) + \delta(\xi, -1)\Big] - (x, f(x)) \right\|^2 \leq \left\| \Big[(x, f(x)) + \delta(\xi, -1)\Big] - (y, \alpha) \right\|^2, \forall (y, \alpha) \in epi(f) \text{ e } \delta > 0,$$

Tomemos $\alpha = f(y)$, logo

$$\big\|\delta(\xi, -1)\big\|^2 \leq \big\|(x - y + \delta\xi, f(x) - f(y) - \delta)\big\|^2.$$

Reescrevendo esta expressão obtemos

$$\big\|\delta\xi\big\|^2 + \big\|\delta\big\|^2 \leq \big\|x - y + \delta\xi\big\|^2 + \big\|f(y) - f(x) + \delta\big\|^2.$$

Como f é uma função semicontínua inferior, isto é, para todo $\delta > 0$, existe $\eta > 0$ tal que $y \in B(x; \eta)$ e $f(y) \geq f(x) - \delta$, segue que

$$\delta^2\|\xi\|^2 + \delta^2 \leq \|x - y + \delta\xi\|^2 + (f(y) - f(x) + \delta)^2, \tag{4.3}$$

assim

$$\begin{aligned} (f(y) - f(x) + \delta)^2 &\geq \delta^2\|\xi\|^2 + \delta^2 - \|x - y + \delta\xi\|^2 \\ &\geq \delta^2 + 2\delta\big\langle \xi, y - x \big\rangle - \|y - x\|^2. \end{aligned}$$

Isto implica que

$$f(y) \geq f(x) - \delta + \left\{ \delta^2 - 2\delta \big\langle \xi, y - x \big\rangle - \|y - x\|^2 \right\}^{\frac{1}{2}}.$$

Logo,

$$f(y) \geq g(y) := f(x) - \delta + \left\{ \delta^2 - 2\delta \big\langle \xi, y - x \big\rangle - ||y - x||^2 \right\}^{\frac{1}{2}}, \tag{4.4}$$

para todo $y \in B(x, \eta)$.

Por outro lado, para cada $y \in B(x; \eta)$, podemos desenvolver a série de Taylor de segunda ordem com resto da função g. Logo, existe uma vizinhança $B(x; \eta)$ de x, com

$$g(y) = g(x) + \big\langle g'(x), y - x \big\rangle + \frac{1}{2} \big\langle g''(z)(y - x), y - x \big\rangle,$$

para algum $z \in (y, x)$. Considerando que $\|g''(z)\|$ é limitado sobre $y \in B(x; \eta)$ por uma constante $2\sigma > 0$, temos

$$g(y) = g(x) + \big\langle \xi, y - x \big\rangle + \sigma \|y - x\|^2,$$

para qualquer $y \in B(x; \eta)$ e $g'(x) = \xi$. Então, por (4.4), como $g(x) = f(x)$, obtemos

$$f(y) \geq f(x) + \big\langle \xi, y - x \big\rangle + \sigma \|y - x\|^2.$$

para qualquer $y \in B(x; \eta)$.

Para demonstrar a implicação contrária, suponha que

$$\big\langle \xi, y - x \big\rangle \leq \alpha - f(x) + M \|y - x\|^2,$$

para qualquer $(y, \alpha) \in epi(f) \cap B\big((x, f(x)), \epsilon\big)$. Segue que,

$$\big\langle \xi, y - x \big\rangle \leq \alpha - f(x) + M \Big[\|y - x\|^2 + \|\alpha - f(x)\|^2 \Big].$$

Reescrevendo esta expressão obtemos

$$\big\langle \xi, y - x \big\rangle + \big\langle -1, \alpha - f(x) \big\rangle \leq M \Big[\|y - x\|^2 + \|\alpha - f(x)\|^2 \Big].$$

Logo,

$$\big\langle (\xi, -1), (y, \alpha) - (x, f(x)) \big\rangle \leq M \Big[\|(y, \alpha) - (x, f(x))\|^2 \Big].$$

Assim, $(\xi, -1) \in N^p_{epi(f)}(x, f(x))$. Portanto $\xi \in \partial^p f(x)$, completando a demonstração.

■

O próximo resultado, fornece uma característica do quociente de diferença da subdiferencial estrita.

Proposição 4.5. *Seja* $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ *uma função semicontínua inferior e os pontos* $x \in dom(f)$ *e* $\xi \in \mathbb{R}^k$*. Então, as seguintes afirmações são equivalentes:*

(i) $\xi \in \hat{\partial} f(x)$*;*

(ii)

$$\limsup_{y \longrightarrow x} \frac{\langle \xi, y - x \rangle - (f(y) - f(x))}{\|y - x\|} \leq 0. \tag{4.5}$$

Demonstração: Mostremos a implicação da condição (ii) para a condição (i). A Condição (4.5) é equivalente a

$$\langle \xi, y - x \rangle - (f(y) - f(x)) \leq o\big(\|y - x\|\big) \text{ para todo } y \in \mathbb{R}^k,$$

onde $o : \mathbb{R}^+ \to \mathbb{R}^+$ é uma função satisfazendo $\lim_{\varepsilon \downarrow 0} o(\varepsilon)/\varepsilon = 0$. Logo

$$\big\langle (\xi, -1), (y, \alpha) - (x, f(x)) \big\rangle \quad \leq \quad o\big(\|(y, \alpha) - (x, f(x))\|\big),$$

para todo $(y, \alpha) \in epi(f)$. Então, pela definição de cone normal estrito, temos

$$(\xi, -1) \in \hat{N}_{epi(f)}(x, f(x)).$$

Assim,

$$\xi \in \hat{\partial} f(x).$$

Agora, vamos mostrar a implicação contrária. Suponha que a condição (ii) não é satisfeita. Logo, existe um número $\alpha > 0$ e uma seqüencia $y_i \longrightarrow x$ tal que $y_i \neq x$ e

$$\langle \xi, y_i - x \rangle - (f(y_i) - f(x)) \geq \alpha \|y_i - x\|, \text{ para todo } i. \tag{4.6}$$

Reajustado a equação acima, temos

$$f(y_i) \leq f(x) + \langle \xi, y_i - x \rangle - \alpha . \|y_i - x\|, \text{ para todo } i. \tag{4.7}$$

Observe que $\limsup_{i \longrightarrow \infty} f(y_i) \leq f(x)$. Como f é semicontínua inferior, segue que $f(y_i) \longrightarrow f(x)$ quando $i \longrightarrow \infty$.

Defina $z_i := (y_i, f(y_i))$, $t_i := \|z_i - (x, f(x))\|$ e $\eta_i := t_i^{-1}(z_i - (x, f(x)))$. Observe que, $\|\eta_i\| = 1$, para cada i. Conseqüentemente, existe uma subseqüência $\eta_i \longrightarrow \eta$ para algum vetor não-nulo η. Também $z_i \xrightarrow{epi(f)} (x, f(x))$ e $t_i \downarrow 0$ ambos quando $i \longrightarrow \infty$. Segue da definição de cone tangente de Bouligand que $\eta \in T_{epi(f)}(x, f(x))$.

Afirma que

$$\big\langle(\xi,-1),\eta\big\rangle > 0.$$

Segue do Teorema 3.1, que

$$\hat{N}_{epi(f)}(x,f(x)) = T_{epi(f)}(x,f(x))^*.$$

Assim, temos que $\big\langle(\xi,-1),\eta\big\rangle \leq 0$, uma contradição.

Para completar a demonstração, precisamos verificar a afirmação feita acima. Por (4.7), podemos concluir que

$$\limsup_{i\longrightarrow\infty} \frac{f(y_i)-f(x)}{\|y_i-x\|} < \infty.$$

Então, existem duas considerações:

a) se lim sup finito existe $K>0$ tal que

$$f(y_i)-f(x) \leq K\|y_i-x\| \text{ para todo } i \tag{4.8}$$

b) existe uma subseqüência que satisfaz

$$\lim_{i\longrightarrow\infty} \frac{f(y_i)-f(x)}{\|y_i-x\|} = -\infty.$$

Suponha que a condição $a)$ é satisfeita. Segue de (4.7), que

$$\big\langle\xi, y_i-x\big\rangle \geq \alpha||y_i-x|| - (f(y_i)-f(x)).$$

Reescrevendo esta expressão obtemos

$$\big\langle(\xi,-1),(y_i,f(y_i))-(x,f(x))\big\rangle \geq \alpha\|y_i-x\|.$$

Logo,

$$t_i^{-1}\big\langle(\xi,-1),(y_i,f(y_i))-(x,f(x))\big\rangle \geq t_i^{-1}\alpha\|y_i-x\|.$$

Mas $\eta_i = t_i^{-1}(z_i-(x,f(x)))$ e $t_i = \|z_i-(x,f(x))\|$, onde $z_i=(y_i,f(y_i))$, temos

$$\begin{aligned}\big\langle(\xi,-1),\eta_i\big\rangle &= \|z_i-(x,f(x))\|^{-1}\alpha\|y_i-x\| \\ &\geq \frac{\alpha\|y_i-x\|}{\|y_i-x\|+\|f(y_i)-f(x)\|} \\ &= \frac{\alpha\|y_i-x\|}{\|y_i-x\|+\|f(y_i)-f(x)+\epsilon-\epsilon\|}.\end{aligned}$$

Como f é uma função semicontínua inferior, existe $\epsilon > 0$ tal que $f(y_i) > f(x) - \epsilon$, segue que

$$\big\langle (\xi, -1), \eta_i \big\rangle \;\geq\; \frac{\alpha \|y_i - x\|}{\|y_i - x\| + (f(y_i) - f(x) + \epsilon) + \epsilon}.$$

Usando (4.8), e passando o limite quando $\epsilon \downarrow 0$ obtemos

$$\begin{aligned}\big\langle (\xi, -1), \eta_i \big\rangle \;&\geq\; \frac{\alpha \|y_i - x\|}{\|y_i - x\| + K\|y_i - x\|} \\ &\geq\; \frac{\alpha}{(1+K)} > 0.\end{aligned}$$

Passando o limite quando $i \longrightarrow \infty$, segue que

$$\big\langle (\xi, -1), \eta \big\rangle \geq \frac{\alpha}{(1+K)} > 0.$$

Agora, suponha que a condição $b)$ é satisfeita. Então

$$\lim_{i \longrightarrow \infty} \frac{f(y_i) - f(x)}{\|y_i - x\|} = -\infty.$$

Logo,

$$\lim_{i \longrightarrow \infty} \frac{\|y_i - x\|}{f(y_i) - f(x)} = 0.$$

Como $\|y_i - x\| > 0$, temos $f(y_i) - f(x) < 0$. Segue da definição de η_i, que

$$\begin{aligned}\big\langle (\xi, -1), \eta_i \big\rangle \;&=\; \big\langle (\xi, -1), t_i^{-1}(z_i - (x, f(x))) \big\rangle \\ &=\; t_i^{-1} \big\langle \xi, y_i - x \big\rangle - t_i^{-1}(f(y_i) - f(x)). \qquad (4.9)\end{aligned}$$

Assim,

$$\begin{aligned}t_i^{-1}\|y_i - x\| \;&=\; \|z_i - (x, (f(x)))\|^{-1}\|y_i - x\| = \|(y_i, f(y_i)) - (x, f(x))\|^{-1}\|y_i - x\| \\ &=\; \frac{\|y_i - x\|}{\|y_i - x\| + \|f(y_i) - f(x)\|} \leq \frac{\|y_i - x\|}{f(y_i) - f(x)}.\end{aligned}$$

Passando o limite quando $i \longrightarrow \infty$, obtemos

$$t_i^{-1}\|y_i - x\| \longrightarrow 0.$$

Também, temos que

$$t_i^{-1}(f(y_i) - f(x)) = \frac{f(y_i) - f(x)}{\|y_i - x\| + \|f(y_i) - f(x)\|} \leq \frac{f(y_i) - f(x)}{\|f(y_i) - f(x)\|}.$$

Passando o limite quando $i \longrightarrow \infty$, obtemos

$$\lim_{i \longrightarrow \infty} t_i^{-1}(f(y_i) - f(x)) \leq -1.$$

Então, de (4.9), obtemos

$$\lim_{i \longrightarrow \infty} \big\langle (\xi, -1), \eta_i \big\rangle = -\lim_{i \longrightarrow \infty} \frac{1}{t_i}(f(y_i - f(x))) \geq 1.$$

Logo $\big\langle (\xi, -1), \eta \big\rangle > 0$, completando a demonstração. ■

4.3 Desigualdade do Valor Médio

Nesta seção, vamos apresentar e demonstrar uma generalização do Teorema do Valor Médio Clássico. Este resultado é necessário para uma investigação mais detalhada de cones normais, subdiferenciais e outros.

Usaremos as seguintes notações: Seja um ponto $\bar{x} \in \mathbb{R}^k$ e um subconjunto $Y \subset \mathbb{R}^k$, denotamos por $[\bar{x}, Y]$ o conjunto

$$\{x : x = \lambda\bar{x} + (1-\lambda)y \text{ para algum } y \in Y \text{ e } 0 \leq \lambda \leq 1\},$$

e por $(\bar{x}, Y)$ o conjunto

$$\{x : x = \lambda\bar{x} + (1-\lambda)y \text{ para algum } y \in Y \text{ e } 0 \leq \lambda \leq 1\}.$$

No caso em que o conjunto Y tem um único elemento, por exemplo $Y = \{\bar{y}\}$, substituímos $[\bar{x}, Y]$ por $[\bar{x}, \bar{y}]$.

Antes de apresentar a Desigualdade do Valor Médio, vamos recordar o Teorema do Valor Médio: seja $f : \mathbb{R}^k \to \mathbb{R}$ é uma função $\mathcal{C}^1$ num conjunto aberto contendo um segmento de reta $[\bar{x}, \bar{y}]$, então para algum $\bar{z} \in [\bar{x}, \bar{y}]$, temos

$$\Big\langle \nabla f(\bar{z}), \bar{y} - \bar{x} \Big\rangle = f(\bar{y}) - f(\bar{x}).$$

O exemplo a seguir ilustra um caso em que o Teorema do Valor Médio não se aplica e ainda mostra que uma generalização direta envolvendo subgradiente limite não é possível.

Exemplo 4.2. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função definida por $f(x) := 1 - |x|$ no intervalo [-1,1]. Segue da definição de subgradiente limite que*

$$\partial f(x) = \begin{cases} 1 & \textit{se } x < 0, \\ \{-1\} \cup \{1\} & \textit{se } x = 0, \\ -1 & \textit{se } x > 0. \end{cases}$$

Note que, $\partial^p f(0) = \hat{\partial} f(0) = \emptyset$. Então, se fosse possível aplicar o Teorema do Valor Médio "generalizado", existiriam $\bar{z} \in [-1, 1]$ e $\xi \in \partial f(\bar{z})$, tais que

$$f(-1) - f(1) = \Big\langle \xi, -1 - 1 \Big\rangle.$$

Logo, $0 = 2\xi$. Mas a igualdade acima não é satisfeita, pois $0 = \xi \notin \partial f(x) = \{-1\} \cup \{1\}$.

Assim, procura-se uma forma mais fraca para o Teorema do Valor Médio para funções não-diferenciáveis. No exemplo acima,vale a seguinte desigualdade

$$f(\bar{y}) - f(\bar{x}) \leq \langle \xi, \bar{y} - \bar{x} \rangle,$$

com $\xi = -1 \in \partial f(\bar{z})$ e $\bar{z} = 0.5 \in [-1, 1]$.

Apresentamos a seguir a versão da desigualdade do valor médio, para funções não-diferenciáveis, que melhor se aproxima da versão clássica do Teorema do Valor Médio.

Teorema 4.1 (**Desigualdade do Valor Médio Proximal**). *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior , $\bar{x} \in dom(f)$ um ponto e $Y \subset dom(f)$ um conjunto convexo compacto . Define, $\hat{r} \in \mathbb{R} \cup \{+\infty\}$ por*

$$\hat{r} := \inf_{y \in Y} f(y) - f(\bar{x}).$$

Então para algum $r < \hat{r}$ finito e $\epsilon > 0$, podemos escolher $\bar{z} \in [\bar{x}, Y] + \epsilon B$ e $\xi \in \partial^p f(\bar{z})$ tais que

$$r < \langle \xi, y - \bar{x} \rangle \textit{ para todo } y \in Y$$

e

$$f(\bar{z}) - f(\bar{x}) \leq \max\{r, 0\}.$$

Demonstração: Por translação para origem em $\mathbb{R}^n \times \mathbb{R}$ podemos supor que $\bar{x} = 0$ e $f(\bar{x}) = 0$, sem perda de generalidade.

Vamos fixar $r < \hat{r}$ e $\epsilon > 0$. Como Y é conjunto convexo compacto e a função f é semicontínua inferior, podemos escolher constantes positivas δ, M e K tais que

$$\delta \leq \epsilon \text{ e } y \in Y + \delta B \Rightarrow f(y) \geq r + \delta; \tag{4.10}$$

$$f(z) \geq -M \ \forall z \in [0, Y] + \delta B, \tag{4.11}$$

$$k > \frac{1}{\delta^2}(M + |r| + 1). \tag{4.12}$$

Defina $g : [0, 1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ por

$$g(t, y, z) := f(z) + k\|ty - z\|^2 - tr.$$

Usamos a notação

$$S := Y \times ([0, Y] + \delta B).$$

Para seguirmos com a demonstração devemos analisar a função g. Observe que, para qualquer $y \in Y$, temos

$$(0, y, 0) \in [0,1] \times S \text{ e } g(0, y, 0) = 0.$$

Como S é compacto e g é semicontínua inferior segue que o problema de minimização

$$\min g(t, y, z) \text{ sobre } (t, y, z) \in [0,1] \times S,$$

tem uma solução $(\bar{t}, \bar{y}, \bar{z})$ e

$$g(\bar{t}, \bar{y}, \bar{z}) = \min_{(t,y,z) \in [0,1] \times S} g(t, y, z) \leq 0. \tag{4.13}$$

Mostremos que $f(\bar{z})) \leq \max\{r, 0\}$. De (4.13), temos

$$f(\bar{z}) + k\|\bar{t}\bar{y} - \bar{z}\|^2 - \bar{t}r \leq 0 \Rightarrow f(\bar{z}) - \bar{t}r \leq 0,$$

pois $k\|\bar{t}\bar{y} - \bar{z}\|^2 \geq 0$. Agora se $r > 0$ temos que $f(\bar{z}) \leq r$ e se $r \leq 0$ então $f(\bar{z}) \leq 0$. Isto implica que

$$f(\bar{z}) \leq \max\{0, r\} < +\infty.$$

Verifiquemos que $\bar{z} \in [0, Y] + \delta intB$. Para isso, suponha que $\bar{z}$ não satisfaz esta condição. Então $\|\bar{z} - \bar{t}\bar{y}\| \geq \delta$, pois $\bar{t}\bar{y} \in [0, Y]$. Segue que

$$\begin{aligned}
g(\bar{t}, \bar{y}, \bar{z}) &= f(\bar{z}) + k\|\bar{t}\bar{y} - \bar{z}\|^2 - \bar{t}r \\
&> -M + k\delta^2 - \|r\| \text{ (por (4.11))} \\
&> -M + \frac{1}{\delta^2}(M + \|r\| + 1)\delta^2 - \|r\| \text{ (por (4.12))} \\
&> 1.
\end{aligned}$$

Esta desigualdade contradiz (4.13).

Agora mostremos que $\bar{t} < 1$. Por redução ao absurdo, suponha que $\bar{t} = 1$. Logo

$$\begin{aligned}
g(1, \bar{y}, \bar{z}) &= f(\bar{z}) + k\|\bar{y} - \bar{z}\|^2 - r \\
&\geq \begin{cases} f(\bar{z}) + k\|\bar{y} - \bar{z}\|^2 - r & \text{se } \|\bar{y} - \bar{z}\| \leq \delta \\ f(\bar{z}) + k\delta^2 - r & \text{se } \|\bar{y} - \bar{z}\| > \delta \end{cases} \\
&\geq \begin{cases} \delta & \text{se } \|\bar{y} - \bar{z}\| \leq \delta \text{ (por 4.10)} \\ 1 & \text{se } \|\bar{y} - \bar{z}\| > \delta \text{ (segue de 4.11 e 4.12)} \end{cases} \\
&= \min\{\delta, 1\} > 0.
\end{aligned}$$

Esta desigualdade, novamente, contradiz (4.13).

Mostremos que existe $\xi \in \partial^p f(\bar{z})$ satisfazendo as condições do teorema. Sabemos que

$$g(\bar{t}, \bar{y}, z) \geq g(\bar{t}, \bar{y}, \bar{z}), \tag{4.14}$$

para todo z no conjunto aberto de $[0, Y] + \delta intB$. Isto implica que

$$k\big(\|\bar{t}\bar{y} - z\|^2 - \|\bar{t}\bar{y} - \bar{z}\|^2\big) \;\; \geq \;\; f(\bar{z}) - f(z). \tag{4.15}$$

Manipulando os termos do lado esquerdo de (4.15) temos

$$\begin{aligned} \|\bar{t}\bar{y} - z\|^2 - \|\bar{t}\bar{y} - \bar{z}\|^2 &= 2\big(\big\langle \bar{t}\bar{y}, \bar{z}\big\rangle - \big\langle \bar{t}\bar{y}, z\big\rangle\big) + \big\langle z, z\big\rangle - \big\langle \bar{z}, \bar{z}\big\rangle \\ &= -2\big\langle \bar{t}\bar{y} - \bar{z}, z - \bar{z}\big\rangle + \|z - \bar{z}\|^2. \end{aligned} \tag{4.16}$$

Na obtenção da segunda igualdade soma-se e subtrai os termos $\big\langle \bar{z}, \bar{z}\big\rangle$ e $2\big\langle z, \bar{z}\big\rangle$.
Substituindo (4.16) em (4.15), obtemos

$$2k\big\langle \bar{t}\bar{y} - \bar{z}, z - \bar{z}\big\rangle \leq f(z) - f(\bar{z}) + k\|z - \bar{z}\|^2,$$

para todo ponto z em alguma vizinhança de $\bar{z}$. Definindo $\xi := 2k(\bar{t}\bar{y} - \bar{z})$ podemos concluir, através da Proposição 4.4, que

$$\xi \in \partial^p f(\bar{z}).$$

Seja $\xi = 2k(\bar{t}\bar{y} - \bar{z})$. Consideremos dois casos, $\bar{t} = 0$ e $\bar{t} \in (0, 1)$.

1) $\bar{t} = 0$. Neste caso, $\xi = -2k\bar{z}$. Escolha $t \in (0, 1]$. Segue de (4.14), que

$$\begin{aligned} 0 &\leq g(t, y, \bar{z}) - g(0, \bar{t}, \bar{z}) \\ &= t^2 k\|y\|^2 - 2k\big\langle ty, \bar{z}\big\rangle - tr, \end{aligned}$$

para qualquer $y \in Y$. Dividindo por t e passando o limite quando $t \downarrow 0$, obtemos

$$0 \leq -2k\big\langle \bar{z}, y\big\rangle - r = \big\langle \xi, y\big\rangle - r.$$

2) $\bar{t} \in (0, 1)$. $\bar{y}$ minimiza a função convexa quadrática $y \to g(\bar{t}, y, \bar{z})$ sobre o conjunto convexo Y, isto é,

$$g(\bar{t}, y, \bar{z}) \geq g(\bar{t}, \bar{y}, \bar{z}) \text{ para todo } y \in Y.$$

Então,

$$2\bar{t}k\big\langle \bar{t}\bar{y} - \bar{z}, y - \bar{y}\big\rangle \geq 0 \text{ para todo } y \in Y.$$

Como $\bar{t} > 0$, segue que

$$\big\langle \xi, y - \bar{y}\big\rangle \geq 0.$$

Por outro lado, $\bar{t}$ também minimiza a função quadrática $t \to g(t, \bar{y}, \bar{z})$ sobre $(0, 1)$. Isto implica que

$$0 = \frac{\partial}{\partial t} g(t, \bar{y}, \bar{z}) \mid_{t=\bar{t}} = 2k\left\langle \bar{t}\bar{y} - \bar{z}, \bar{y}\right\rangle - r.$$

Logo,$\left\langle \xi, \bar{y}\right\rangle = r$. Assim, podemos concluír que $\left\langle \xi, y\right\rangle \geq r$, completando a demonstração do teorema. ■

No caso Lipschitz, a Desigualdade do Valor Médio é "exata" em termos de subdiferencial limite. Este caso será tratado na Seção 4.4.2

4.4 Algumas Propriedades das Subdiferenciais

Nesta seção, apresentamos algumas propriedades das subdiferenciais de funções diferenciável, de Lipschitz, indicadora e distância. Estas propriedades serão utilizadas posteriormente.

4.4.1 Função Diferenciável

Proposição 4.6. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e $x \in dom(f)$ um ponto. Então, vale as seguintes afirmações:*

(i) Se $f \in C^2$ então

$$\partial^p f(x) = \{\nabla f(x)\} \text{ para todo } x \in dom(f);$$

(ii) Se f é convexa então $\xi \in \partial^p f(x)$ se, e somente se,

$$f(y) \geq f(x) + \left\langle \xi, y - x\right\rangle \text{ para todo } x \in dom(f); \tag{4.17}$$

(iii) Se f tem um mínimo local em x, então $0 \in \partial^p f(x)$;

(iv) Se f é convexa e $0 \in \partial^p f(x)$ então x é um ponto de mínimo global de f.

Demonstração: ***(i)*** Tome $\xi \in \partial^p f(x)$. Então, existem $M > 0$ e $\epsilon > 0$ tais que

$$\left\langle \xi, y - x\right\rangle \leq f(y) - f(x) + M\|y - x\|^2, \ \forall y \in B(x, \epsilon).$$

Substituindo $y = x + tv$ para $t > 0$ suficientemente pequeno, temos

$$t\left\langle \xi, v\right\rangle \leq f(x + tv) - f(x) + Mt^2\|v\|^2 \ \forall v \in \mathbb{R}^k.$$

Dividindo por t e passando o limite quando $t \downarrow 0$, obtemos

$$\langle \xi, v \rangle \leq \lim_{t \downarrow 0} \frac{f(x+tv) - f(x)}{t}.$$

Mas $f \in C^2$, logo

$$\lim_{t \downarrow 0} \frac{f(x+tv) - f(x)}{t} = f'(x, v) = \langle \nabla f(x), v \rangle.$$

Assim,

$$\langle \xi - \nabla f(x), v \rangle \leq 0.$$

Como v é arbitrário, concluímos que $\xi = \nabla f(x)$.

(ii) Seja $(1-t)x + ty \in B(x; \eta)$ para qualquer $t \in (0,1)$ suficientemente pequeno e $\eta > 0$. Segue da convexidade de f, que

$$(1-t)f(x) + tf(y) \geq f((1-t)x + ty)) \ \forall y \in \mathbb{R}^k$$

Tome $\xi \in \partial^p f(x)$. Pela Proposição 4.4, substituindo y por $(1-t)x + ty$, temos

$$\begin{aligned} f((1-t)x + ty)) &\geq f(x) + \langle \xi, (1-t)x + ty - x \rangle - M\|(1-t)x + ty - x\|^2 \\ &= f(x) + t\langle \xi, y - x \rangle - Mt^2\|y - x\|^2. \end{aligned}$$

Então,

$$(1-t)f(x) + tf(y) \geq f(x) + t\langle \xi, y - x \rangle - Mt^2\|y - x\|^2,$$

logo,

$$tf(y) \geq tf(x) + t\langle \xi, y - x \rangle - Mt^2\|y - x\|^2.$$

Dividindo por t e passando o limite quando $t \downarrow 0$, obtemos

$$f(y) \geq f(x) + \langle \xi, y - x \rangle.$$

(iii) Seja $\xi \in \partial^p f(x)$. Então, existem $M > 0$ e $\epsilon > 0$ tais que

$$\langle \xi, y - x \rangle \leq f(y) - f(x) + M\|y - x\|^2 \text{ para todo } y \in B(x, \epsilon).$$

Observe que $\xi = 0$ satisfaz a desigualdade acima até com $M = 0$, devido ao fato de x ser um ponto de mínimo local de f. Logo, segue da Proposição 4.4, que $0 \in \partial^p f(x)$.

(iv) Seja f uma função convexa. Segue de (ii), que

$$f(y) \geq f(x) + \langle \xi, y - x \rangle \ \forall y \in \mathbb{R}^k \ \forall \xi \in \partial f(x).$$

Como $0 \in \partial^p f(x)$, temos

$$f(y) \geq f(x) \forall y \in \mathbb{R}^k.$$

Portanto, x é um mínimo global de f. ■

4.4.2 Função que Satisfaz a Condição de Lipschitz

Definição 4.3. *Uma função $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ é chamada de Lipschitziana se existir um número real $K > 0$, tal que*

$$\|f(y) - f(x)\| \leq K\|y - x\| \text{ para todo } y, x \in \mathbb{R}^k.$$

O número real K é conhecido como constante de Lipschitz da f em $\mathbb{R}^k$.

A próxima proposição é um caso particular da Proposição 4.4 substituindo na hipótese a função semicontínua inferior por uma que satisfaz a condição Lipschitz. Não há necessidade de demonstração, mas vamos apresentar uma demonstração simplificada a título de ilustração.

Proposição 4.7. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz e os pontos $x \in \mathbb{R}^k$, $\xi \in \mathbb{R}^k$. Então, as duas seguintes afirmações são equivalentes:*

(i) $\xi \in \partial^p f(x)$;

(ii) existe $M > 0$ tal que

$$\big\langle \xi, y - x \big\rangle \leq f(y) - f(x) + M\|y - x\|^2, \tag{4.18}$$

para todo $y \in \mathbb{R}^k$.

Demonstração: Tome $\xi \in \partial^p f(x)$. Assim, $(\xi, -1) \in N^p_{epi(f)}(x, f(x))$. Considerando o conjunto,

$$\begin{aligned} epi(f) &:= \big\{(y, \alpha) : y \in dom(f) \text{ e } \alpha \geq f(y)\big\} \\ &= dom(f) \times S_2, \end{aligned}$$

onde $S_2 = \big\{\alpha : \alpha \geq f(y), y \in dom(f)\big\}$. Segue da Proposição (3.5), que

$$(\xi, -1) \in N^p_{epi(f)}(x, f(x)), \text{ implica } (\xi, -1) \in N^p_{dom(f)}(x) \times N^p_{S_2}(f(x)).$$

Então,

$$\xi \in N^p_{epi(f)}(x) \text{ e } (-1) \in N^p_{S_2}(f(x)).$$

Pela definição de normal proximal, obtemos

$$\xi \in N^p_{epi(f)}(x) \text{ implica } \big\langle \xi, y - x \big\rangle \leq R\|y - x\|^2 \;\; \forall\, y \in dom(f), \tag{4.19}$$

e

$$(-1) \in N^p_{S_2}(f(x)) \text{ implica } \big\langle -1, \alpha - f(x) \big\rangle \leq Q\|\alpha - f(x)\|^2 \;\; \forall\, \alpha \in S_2.$$

Tomando, em particular, $\alpha = f(y)$, logo

$$(-1) \in N^p_{S_2}(f(x)) \text{ implica } \big\langle -1, f(y) - f(x) \big\rangle \leq Q\|f(y) - f(x)\|^2 \;\; \forall\, y \in dom(f), \tag{4.20}$$

Somando (4.19) e (4.20), obtemos

$$\big\langle \xi, y - x \big\rangle + \big\langle -1, f(y) - f(x) \big\rangle \leq R\|y - x\|^2 + Q\|f(y) - f(x)\|^2 \;\forall\, y \in dom(f).$$

Como f é uma função Lipschitz, segue que

$$\begin{aligned} \big\langle \xi, y - x \big\rangle + \big\langle -1, f(y) - f(x) \big\rangle &\leq R\|y - x\|^2 + QK^2\|y - x\|^2 \;\forall\, y \in dom(f) \\ &\leq M\|y - x\|^2, \end{aligned}$$

onde $M = \max\{R, QK^2\}$. Portanto,

$$\big\langle \xi, y - x \big\rangle \leq f(y) - f(x) + M\|y - x\|^2, \text{ para todo } y \in dom(f).$$

A prova da implicação contrária, segue da Proposição 4.4, completando a demonstração.

■

Os próximos lemas serão necessários para demonstrar o Teorema 4.2.

Lema 4.1. *Sejam $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior, $x \in \mathbb{R}^k$ e $\beta \geq 0$. Defina*

$$g(y) := \min\big\{ f(y), \beta \big\}.$$

Suponha que ξ é um vetor não-nulo tal que

$$\xi \in \partial^p g(x).$$

Então

$$f(x) \leq \beta \text{ e } \xi \in \partial^p f(x).$$

Demonstração: Tome $\xi \in \partial^p f(x)$. Então existem $M > 0$ e $\alpha > 0$ tais que

$$g(x') - g(x) \geq \big\langle \xi, x' - x \big\rangle - M\|x' - x\|^2 \;\forall x' \in B(x, \alpha). \tag{4.21}$$

Segue de $g(y) = \min\big\{ f(y), \beta \big\}$, que

$$\min\big\{ f(x'), \beta \big\} - \min\big\{ f(x), \beta \big\} \geq \big\langle \xi, x' - x \big\rangle - M\|x' - x\|^2 \;\forall x' \in B(x, \alpha).$$

Se $f(x) > \beta$ podemos supor que $f(x') \geq \beta$ para todo $x' \in B(x, \alpha)$ (ajustando α se necessário). Assim,

$$0 = \beta - \beta \geq \min\big\{f(x'), \beta\big\} - \min\big\{f(x), \beta\big\} \geq \big\langle \xi, x' - x \big\rangle - M\|x' - x\|^2 \ \forall x' \in B(x, \alpha).$$

Escolhendo $x' = x + \delta\xi$ com $\delta > 0$ (para δ suficientemente pequeno), segue que

$$\|\xi\|^2 \quad \leq \quad M\delta\|\xi\|^2.$$

Passando o limite quando $\delta \downarrow 0$, obtemos $\xi = 0$. Mas, isto é impossível, pois pela hipótese $\xi \neq 0$. Portanto, $f(x) \leq \beta$.

Podemos então reescrever (4.21), obtendo

$$f(x') - f(x) \geq \big\langle \xi, x' - x \big\rangle - M||x' - x||^2$$

para todo $x' \in B(x, \alpha)$. Segue da Proposição 4.7 que $\xi \in \partial^p f(x)$.

■

Lema 4.2. *Nas condições do teorema, as seguintes condições são equivalentes:*

(a) a função f é Lipschitz com constante K em alguma vizinhança de $\bar{x}$;

(b) existe $\epsilon > 0$ tal que para cada $x \in \mathbb{R}^k$ e $\xi \in \partial^p f(x)$ temos

$$\|x - \bar{x}\| \leq \epsilon \ e \ \|f(x) - f(\bar{x})\| \leq \epsilon \Rightarrow \|\xi\| \leq K.$$

Demonstração: A demonstração de que (a) $\Rightarrow$ (b) é por contradição. Suponha que $\forall \epsilon > 0$ existem $x \in \mathbb{R}^k$ e $\xi \in \partial^p f(x)$ tais que

$$\|x - \bar{x}\| \leq \epsilon \ \ \|f(x) - f(\bar{x})\| \leq \epsilon \text{ e } \|\xi\| > K.$$

Isto contradiz o fato já demonstrado de que ξ é limitado por K, pois $\xi \in \partial^p f(x) \subset \partial f(x) \subset KB$.

Mostremos a recíproca. Suponha que f não é Lipschitziana com constante Lipschitz K em alguma vizinhança de $\bar{x}$. Então, é suficiente mostrar que a condição (b) não é satisfeita.

Escolha qualquer $\epsilon > 0$. Como f é semicontínua inferior, podemos encontrar $\delta > 0$, $0 < \delta < \epsilon$ tal que

$$\|z - \bar{x}\| \leq \delta \text{ implica } f(z) > f(\bar{x}) - \epsilon. \tag{4.22}$$

Mas, como f não é função Lipschitziana, existem seqüencias $x_i \longrightarrow \bar{x}, y_i \longrightarrow \bar{x}$ tal que $x_i \neq y_i$ e

$$f(y_i) - f(x_i) > K\|y_i - x_i\| \text{ para todo } i. \tag{4.23}$$

Também, podemos supor que para alguma subseqüência (não reindexamos), temos

$$f(x_i) \longrightarrow f(\bar{x}) \text{ quando } i \longrightarrow \infty. \tag{4.24}$$

De (4.24), podemos escolher uma seqüencia $\epsilon_i \downarrow 0$ tal que

$$f(x_i) - f(\bar{x}) \leq \epsilon_i \tag{4.25}$$

e

$$K\|y_i - x_i\| < f(\bar{x}) - f(x_i) + \epsilon_i. \tag{4.26}$$

Fixe um valor para o indice i tal que

$$\|x_i - \bar{x}\| < \frac{1}{2}\,\delta,\ \|y_i - \bar{x}\| < \frac{1}{2}\,\delta \text{ e } \epsilon_i < \epsilon \tag{4.27}$$

e defina a função

$$g_i(y) := \min\big\{f(\bar{x}) + \epsilon_i, f(y)\big\}.$$

Logo,

$$\begin{aligned} g_i(y_i) - g_i(x_i) &= \min\big\{f(\bar{x}) + \epsilon_i, f(y_i)\big\} - \min\big\{f(\bar{x}) + \epsilon_i, f(x_i)\big\} \\ &= \min\big\{f(\bar{x}) + \epsilon_i - f(x_i), f(y_i) - f(x_i)\big\} \text{ por(4.25)} \\ &> g_i(y_i) - g_i(x_i) > K\|y_i - x_i\|, \text{ por(4.26) e (4.23)} \end{aligned}$$

Podemos observar que g_i é uma função semicontínua inferior, logo pela Desigualdade do Valor Médio (Teorema 4.1) existem

$$z \in B(\bar{x}, \delta) \text{ e } \xi \in \partial^p g_i(z)$$

tal que

$$K\|y_i - x_i\| < \big\langle \xi, y_i - x_i \big\rangle.$$

Como $y_i \neq x_i$, temos $\|\xi\| > K$. Assim, ξ é um vetor não-nulo. Pelo Lema 4.1, obtemos

$$f(z) < f(\bar{x}) + \epsilon \quad \text{e} \quad \xi \in \partial^p f(z). \tag{4.28}$$

No entanto $\delta < \epsilon$, logo $\|z - \bar{x}\| < \epsilon$. Por (4.22), temos

$$f(z) > f(\bar{x}) - \epsilon. \tag{4.29}$$

Segue de (4.28) e (4.29), que

$$\|z - \bar{x}\| < \epsilon, \quad \|f(z) - f(\bar{x})\| < \epsilon, \quad \text{e} \quad \xi \in \partial^p f(z).$$

Como ϵ é um número positivo arbitrário podemos concluir que a Condição (b) não é satisfeita.

■

Teorema 4.2. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ um função semicontínua inferior e $\bar{x} \in \mathbb{R}^k$ um ponto. Seja f uma função Lipschitziana na vizinhança de $\bar{x}$ com constante Lipschitz K. Então*

(i) $\partial f(\bar{x}) \subset KB$;

(ii) $\partial^\infty f(\bar{x}) = \{0\}$.

Reciprocamente, se (ii) é válida, então f é uma função Lipschitziana na vizinhança de $\bar{x}$ Aqui B denota bola fechada unitária de $\mathbb{R}^k$.

Demonstração: $(\Rightarrow)$ Tome um ponto

$$(\xi, -\lambda) \in N_{epi(f)}(\bar{x}, f(\bar{x})).$$

Logo existem seqüencias convergentes

$$(x_i, \alpha_i) \xrightarrow{epi(f)} (\bar{x}, f(\bar{x})) \ \text{ e } (\xi_i, \lambda_i) \longrightarrow (\xi, \lambda)$$

tais que

$$(\xi_i, -\lambda_i) \in N^p_{epi(f)}(x_i, \alpha_i) \ \forall i,$$

Dado $x \in \mathbb{R}^k$ arbitrário, segue da definição de cone normal proximal que

$$\Big\langle (\xi_i, -\lambda_i), (x, \alpha) - (x_i, \alpha_i) \Big\rangle \ \leq \ M\Big[\|x - x_i\|^2 + \|\alpha - \alpha_i)\|^2\Big] \ \forall \alpha \geq f(x).$$

Tomando $\beta_i = \alpha_i - f(x_i)$ e $\beta = \alpha - f(x)$ obtemos

$$\Big\langle \xi_i, x - x_i \Big\rangle + \Big\langle -\lambda_i, f(x) + \beta - f(x_i) - \beta_i \Big\rangle \leq M\Big[\|x - x_i\|^2 + \|f(x) + \beta - \beta_i - f(x_i)\|^2\Big]. \tag{4.30}$$

Substituímos $x = x_i$ em (4.30), temos

$$-\lambda_i(\beta - \beta_i) \leq M|\beta - \beta_i|^2.$$

Considerando $\beta \neq \beta_i$ e passando o limite quando $i \longrightarrow \infty$, obtemos $\lambda \geq 0$.

Agora substituindo $\beta = \beta_i$ em (4.30) obtemos

$$\big\langle \xi_i, x - x_i \big\rangle + \big\langle -\lambda_i, f(x) - f(x_i) \big\rangle \leq M\big[\|x - x_i\|^2 + \|f(x) - f(x_i)\|^2\big].$$

Mas f é Lipschitziana, logo

$$\big\langle \xi_i, x - x_i \big\rangle - \lambda_i K\|x - x_i\| \leq M(1 + K^2)\|x - x_i\|^2, \tag{4.31}$$

para todo x na vizinhança de x_i. Escolhendo $x = x_i + r\xi_i$ para $r > 0$ suficientemente pequeno e substituído na equação(4.31) obtemos

$$r\|\xi_i\|^2 - \lambda_i K r\|\xi_i\| \leq M(1 + K^2)r^2\|\xi_i\|^2.$$

Dividindo por r e passando o limite quando $r \downarrow 0$ segue a desigualdade

$$\|\xi_i\| \leq \lambda_i K.$$

Então, passando o limite quando $i \longrightarrow \infty$, concluímos que

$$\|\xi\| \leq \lambda K, \tag{4.32}$$

onde $\lambda \geq 0$ e $K > 0$.

Daí podemos concluir que:

- se $\lambda = 0$, então $\xi = 0$, ou seja, $\partial^\infty f(x) = \{0\}$, e
- se $\lambda > 0$ podemos supor que $\lambda = 1$ e logo temos que $\|\xi\| \leq K$ i.e., $\partial f(\bar{x}) \subset KB$.

($\Leftarrow$) Suponha que a função não é Lipschitiziana na vizinhança de $\bar{x}$. Pelo Lema 4.2, existem seqüencias $x_i \xrightarrow{f} \bar{x}$ e $\{\xi_i\}$ tal que

$$\|\xi_i\| \in \partial^p f(x_i) \text{ para cada } i$$

e

$$\|\xi_i\| \longrightarrow \infty \text{ quando } i \longrightarrow \infty.$$

Para cada i, temos

$$\big(\|\xi_i\|^{-1}\xi_i, -\|\xi_i\|^{-1}\big) \in N^p_{epi(f)}(x_i, f(x_i)).$$

No entanto, podemos encontrar uma subseqüência (não reindexamos) tal que

$$\big(\|\xi_i\|^{-1}\xi_i, -\|\xi_i\|^{-1}\big) \longrightarrow \big(\bar{\xi}, 0\big) \text{ quando } i \longrightarrow \infty,$$

para algum $\bar{\xi}$, com $\|\bar{\xi}\| = 1$. Segue da definição de cone normal limite que

$$\big(\bar{\xi}, 0\big) \in N_{epi(f)}(\bar{x}, f(\bar{x})).$$

Então, $\bar{\xi}$ é um vetor não-nulo em $\partial^{\infty} f(\bar{x}) = \{0\}$ o que é uma contradição. Portanto f é Lipschitziana na vizinhança de $\bar{x}$. ■

Também, vamos apresentar a Desigualdade do Valor Médio mas envolvendo agora a função Lipschitziana.

Corolário 4.1 (**Caso Lipschitz**). *Sejam $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função que satisfaz a condição de Lipschitz em $[\bar{x}, Y] + \delta B$, $x \in \mathbb{R}^n$ um ponto e $Y \subset \mathbb{R}^n$ um conjunto convexo compacto. Então, para algum $\delta > 0$ existem $\bar{z} \in [\bar{x}, Y]$ e $\xi \in \partial f(\bar{z})$ tais que*

$$\min_{y \in Y} f(y) - f(x) \leq \big\langle \xi, y - \bar{x} \big\rangle \text{ para todo } y \in Y$$

e

$$f(\bar{z}) - f(\bar{x}) \leq \max\Big\{ \min_{y \in Y} f(y) - f(\bar{x}), 0 \Big\}.$$

Demonstração: Tomemos as seqüencias $\varepsilon_i \downarrow 0$ e $r_i \uparrow \hat{r}$, onde

$$\hat{r} := \min_{y \in Y} f(y) - f(x).$$

Aplicando o Teorema 4.1 com $\varepsilon = \varepsilon_i$ e $r = r_i$, para cada i podemos escolher $z_i \in [\bar{x}, Y] + \varepsilon_i B$ e $\xi_i \in \partial^p f(z_i)$ tais que

$$r_i < \big\langle \xi_i, y - \bar{x} \big\rangle \text{ para todo } y \in Y \tag{4.33}$$

e

$$f(z_i) - f(\bar{x}) \leq \max\{r_i, 0\}. \tag{4.34}$$

Observe que seqüencias $\{z_i\}$ e $\{\xi_i\}$ são limitadas, pois o conjunto Y é compacto e a função f satisfaz a condição de Lipschitz em $[\bar{x}, Y] + \delta B$ (veja a Proposição 4.2). Então, podemos encontrar subseqüências convergentes $z_i \longrightarrow \bar{z}$, para algum $\bar{z} \in [\bar{x}, Y]$ e $\xi_i \longrightarrow \xi$, para algum $\xi \in \mathbb{R}^n$. Segue da Proposição 4.2, que $\xi \in \partial f(\bar{z})$. Passando o limite nas equações (4.33) e (4.33) para $i \longrightarrow \infty$, obtemos

$$\hat{r} < \big\langle \xi, y - \bar{x} \big\rangle \text{ para todo } y \in Y$$

e

$$f(z) - f(\bar{x}) \leq \max\{\hat{r}, 0\}.$$

Isto completa a demonstração. ■

4.4.3 Função Indicadora

Definição 4.4. *Seja $C \subset \mathbb{R}^k$ um subconjunto não-vazio. A função indicadora $\Psi_C(x) : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ associada com C é definida por*

$$\Psi_C(x) := \begin{cases} 0 & \text{se } x \in C \\ +\infty & \text{caso contrário} \end{cases}$$

Com a função indicadora temos a seguinte proposição:

Proposição 4.8. *Seja $\Psi_C(x) : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função indicadora e $x \in C$ um ponto. Então,*

$$\partial\Psi_C(x) = N_C(x) = \partial^\infty \Psi_C(x).$$

Demonstração: Tome $\xi \in \partial\Psi_C(x)$. Então, $\xi \in \partial\Psi_C(x)$ se, e somente se, $(\xi, -1) \in N_{epi(\Psi_C)}(x, 0)$. Observe que $epi(\Psi_C) = C \times [0, \infty)$. Logo, a primeira igualdade segue imediatamente da Proposição 3.5.

A demonstração da segunda igualdade é análoga. ■

4.4.4 Função Distância

Definição 4.5. *Seja um conjunto $C \subset \mathbb{R}^k$. A função distância $d_C : \mathbb{R}^k \to \mathbb{R}$ é definida por*

$$d_C(x) := \inf\Big\{\|x - y\| : y \in C\Big\}.$$

Temos as seguintes proposições:

Proposição 4.9. *Seja $d_C : \mathbb{R}^k \to \mathbb{R}$ função distância relacionada com um conjunto qualquer $C \subset \mathbb{R}^k$. Então,*

i) $d_C(\cdot)$ é uma função que satisfaz a condição de Lipschitz com constante igual a 1;

ii) se C é um conjunto convexo, então $d_C(\cdot)$ é uma função convexa.

Demonstração: ***i)*** Sejam $x,\ y \in \mathbb{R}^k$ quaisquer e escolha $\epsilon > 0$. Então pela definição de função distância, existe $z \in C$ tal que $d_C(y) \geq \|y - z\| - \epsilon$. Logo,

$$d_c(x) \leq \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| \leq \|x - y\| + d_C(y) + \epsilon.$$

Mas, como $\epsilon > 0$ é arbitrário, temos que $d_C(x) - d_C(y) \leq \|x - y\|$.

Por outro lado, escolhendo $\delta > 0$ existe $z \in C$ tal que $d_C(x) \geq \|x - z\| - \delta$. Então,

$$d_C(y) \leq \|y - z\| \leq \|x - y\| + d_C(x) + \delta.$$

Como δ é arbitrário, obtemos $d_C(y) - d_C(x) \leq \|x - y\|$. Portanto,

$$\|d_C(x) - d_C(y)\| \leq \|x - y\|.$$

Assim, a função distância d_C é Lipschitziana com constante Lipschitz igual a 1.

ii) Sejam $x, y \in C$ e dado $\epsilon > 0$, então, existem $c_x, c_y \in C$ tais que

$$\|c_x - x\| \leq d_C(x) + \epsilon \quad \text{e} \quad \|c_y - y\| \leq d_C(y) + \epsilon.$$

Sabendo que C é um conjunto convexo, logo existe $c \in C$ tal que $c = \lambda c_x + (1 - \lambda)c_y$ e $\lambda \in (0, 1)$. Então,

$$\begin{aligned} d_C(\lambda x + (1-\lambda)y) &\leq \|(\lambda x + (1-\lambda)y - c\| \\ &= \|\lambda x + (1-\lambda)y - \lambda c_x - (1-\lambda)c_y\| \\ &\leq \lambda\|c_x - x\| + (1-\lambda)\|c_y - y\| \\ &\leq \lambda d_C(x) + (1-\lambda)d_C(y) + \epsilon. \end{aligned}$$

Como ϵ é arbitrário, obtemos $d_C(x)$ é uma função convexa. ■

Proposição 4.10. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então*

$$\partial^p d_C(x) = N_C^p(x) \cap B.$$

Demonstração: Dividimos a demonstração em duas partes:

a) Mostremos a relação $\partial^p d_C(x) \subset N_C^p(x) \cap B$. Suponha que $\xi \in \partial^p d_C(x)$. Como a função $d_C(x)$ é Lipschitz com constante 1, segue da Proposição 4.7, que existe $M > 0$ tal que

$$\langle \xi, y - x \rangle \leq d_C(y) - d_C(x) + M\|y - x\|^2 \text{ para todo } y \in \mathbb{R}^k.$$

Mas $x \in C$. Assim, pela definição de função distância, temos

$$\langle \xi, y - x \rangle \leq M\|y - x\|^2 \text{ para todo } y \in C.$$

Então, $\xi \in N_C^p(x)$. Segue da Proposição 4.9, que $\|\xi\| \leq 1$. Portanto,

$$\xi \in N_C^p(x) \cap B.$$

b) Agora, mostraremos a relação $\partial^p d_C(x) \supset N_C^p(x) \cap B$. Seja $\xi \in N_C^p(x) \cap B$ qualquer. Logo, é suficiente mostrar que $\xi \in \partial^p d_C(x)$. Para $\xi = 0$ é facilmente verificado, assim tomemos $\xi \neq 0$. Segue da definição de cone normal proximal, que existe $M > 0$ tal que

$$\left\langle \xi, y - x \right\rangle \leq M\|y - x\|^2 \text{ para todo } y \in C.$$

Podemos encontrar uma esfera aberta Q (veja, na Seção 3.1) tal que

$$C \cap Q = \emptyset,$$

onde

$$Q := \left\{ y \in \mathbb{R}^k : \|x + (2M)^{-1}\xi - y\| < (2M)^{-1}\|\xi\| \right\}.$$

Para qualquer $y \in Q$, vale

$$\left\langle \xi, y - x \right\rangle \leq d_C(y) - d_C(x) + M\|y - x\|^2,$$

e

$$d_C(y) \quad \geq \quad g(y), \tag{4.35}$$

com

$$g(v) := (2M)^{-1}\|\xi\| - \|v - x - (2M)^{-1}\xi\|.$$

A função g é analítica na vizinhança de x, desde que $\xi \neq 0$. Assim, podemos calcular $\nabla g(x) = \|\xi\|^{-1}\xi$. Segue da definição de subgradiente proximal aplicada na função g, que existem $\alpha > 0$ e $\epsilon > 0$ tais que

$$g(y) - g(x) \geq \left\langle \|\xi\|^{-1}\xi, y - x \right\rangle - \alpha\|y - x\|^2 \text{ para todo } y \in B(x, \epsilon). \tag{4.36}$$

Como $g(x) = d_C(x) = 0$, concluímos de (4.35) que

$$d_C(y) - d_C(x) \geq g(y) - g(x) \geq \left\langle \|\xi\|^{-1}\xi, y - x \right\rangle - \alpha\|y - x\|^2 \text{ para todo } y \in B(x, \epsilon).$$

Mas, como $\|\xi\| \leq 1$, temos

$$1(d_C(y) - d_C(x)) \geq \|\xi\|(d_C(y) - d_C(x)) \geq \left\langle \xi, y - x \right\rangle - \alpha\|\xi\|\|y - x\|^2, \text{ para todo } y \in B(x, \epsilon).$$

Assim, pela Proposição 4.4, temos que $\xi \in \partial^p d_C(x)$. ■

O resultado a seguir tem como objetivo obter informação sobre o cone proximal $\xi \in \partial^p d_C(x)$ num ponto $x \notin C$.

Proposição 4.11. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado, $x \notin C$ um ponto e $\xi \in \partial^p d_C(x)$ um vetor. Então, x tem um único ponto mais próximo $\bar{x}$ em C e*

$$\xi = \|x - \bar{x}\|^{-1}(x - \bar{x}).$$

Concluímos, em particular, que $\xi \in \partial^p d_C(\bar{x})$.

Demonstração: Seja $\xi \in \partial^p d_C(x)$. Então, pela Proposição 4.7, existe $M > 0$ tal que

$$\big\langle \xi, y - x \big\rangle \leq d_C(y) - d_C(x) + M\|y - x\|^2 \text{ para todo } y \in \mathbb{R}^k. \tag{4.37}$$

Seja $\bar{x}$ um ponto mais próximo de x em C. Então,

$$\|x - \bar{x}\| = d_C(x) \ (\neq 0).$$

Observe que $\bar{x}$ é ponto mais próximo de $\bar{x} + (1-\alpha)(x - \bar{x})$ para todo $\alpha \in (0,1)$. Assim

$$\begin{aligned} d_C\big(\bar{x} + (1-\alpha)(x - \bar{x})\big) &= \|\bar{x} + (1-\alpha)(x - \bar{x}) - \bar{x}\| \\ &= (1-\alpha) d_C(x). \end{aligned}$$

Substituindo y por $\bar{x} + (1-\alpha)(x - \bar{x})$ em (4.37) obtemos

$$\begin{aligned} d_C\big(\bar{x} + (1-\alpha)(x - \bar{x})\big) - d_C(x) &\geq \big\langle \xi, \bar{x} + (1-\alpha)(x - \bar{x}) - x \big\rangle - M\|\bar{x} + (1-\alpha)(x - \bar{x}) - x\|^2 \\ &\geq -\alpha\big\langle \xi, x - \bar{x} \big\rangle - \alpha^2 M\|x - \bar{x}\|^2. \end{aligned}$$

Por outro lado, utilizando a definição de função distância encontramos

$$\begin{aligned} d_C\big(\bar{x} + (1-\alpha)(x - \bar{x})\big) - d_C(x) &= (1-\alpha) d_C(x) - d_C(x) \\ &= -\alpha d_C(x) = -\alpha\|x - \bar{x}\|. \end{aligned}$$

Assim,

$$\begin{aligned} -\alpha\|x - \bar{x}\| &= d_C\big(\bar{x} + (1-\alpha)(x - \bar{x})\big) - d_C(x) \geq -\alpha\big\langle \xi, x - \bar{x} \big\rangle - \alpha^2 M\|x - \bar{x}\|^2 \\ &\geq -\alpha\big\langle \xi, x - \bar{x} \big\rangle - \alpha^2 M\|x - \bar{x}\|^2, \text{para todo } \alpha \in (0,1). \end{aligned}$$

Dividimos por α e passando o limite quando $\alpha \downarrow 0$, obtemos

$$\|x - \bar{x}\| \leq \big\langle \xi, x - \bar{x} \big\rangle.$$

Mas, pela Proposição 4.10, $\|\xi\| \leq 1$ e como $x - \bar{x} \neq 0$, obtemos

$$\xi = \|x - \bar{x}\|^{-1}(x - \bar{x}).$$

Agora, vamos mostrar a unicidade do ponto mais próximo. Tomando $\bar{y}$ um outro ponto próximo de x em C, temos

$$d_C(x) = \|x - \bar{y}\| = \|x - \bar{x}\| \neq 0$$

e

$$\xi = \|x - \bar{x}\|^{-1}(x - \bar{x}) = \|x - \bar{y}\|^{-1}(x - \bar{y}).$$

Assim, concluímos que $\bar{x} = \bar{y}$.

O fato que $\xi \in \partial^p d_C(\bar{x})$ segue da Proposição 4.10, completando a demonstração. ■

Denotemos por $\overset{\circ}{B}(x, \delta)$ a bola aberta de centro x e raio δ.

Proposição 4.12. *Seja $C \subset \mathbb{R}^k$ e os pontos $x \notin C$ e $\bar{x} \in C$. Suponha que, para alguma $\alpha > 0$, $y = \bar{x}$ minimiza a função*

$$y \to \|\bar{x} + (1 + \alpha)(x - \bar{x}) - y\|$$

sobre C. Defina

$$\xi := \|x - \bar{x}\|^{-1}(x - \bar{x}).$$

Então, as seguintes afirmações são verdadeiras

(i) $\xi \in \partial^p d_C(x)$,

(ii) se $\{x_i\}$ e $\{\xi_i\}$ são seqüencias tais que $x_i \longrightarrow x$ quando $i \longrightarrow \infty$ e $\xi_i \in \partial^p d_C(x_i)$ para cada i, então existe uma seqüencia de pontos $\{y_i\}$ em C tal que $\xi_i = \|x_i - y_i\|^{-1}(x_i - y_i)$ para todo i, $y_i \longrightarrow \bar{x}$ e $\xi_i \longrightarrow \xi$ quando $i \longrightarrow \infty$.

Demonstração: Por conveniência, podemos supor, que $\bar{x} = 0$.

(i)A propriedade minimizante de $\bar{x}$ pode ser expressado geometricamente (veja a Figura 4.4):

$$\overset{\circ}{B}\big((1 + \alpha)x, (1 + \alpha)\|x\|\big) \cap C = \emptyset.$$

Tomando qualquer ponto $v \in \mathbb{R}^k$ com $\|v - \alpha x\| \leq (1 + \alpha)\|x\|$, temos

$$\begin{aligned} d_C(x + v) &= \|x + v - \bar{y}\| (\text{ } \bar{y} \text{ é o ponto mais próximo de } x + v \text{ em } C) \\ &\geq \|x + \alpha x - \bar{y}\| - \|v - \alpha x\| \\ &\geq \|(1 + \alpha)x\| - \|v - \alpha x\| (\bar{x} = 0 \text{ é o ponto mais próximo de } (1 + \alpha)x) \\ &= \Phi(v), \end{aligned} \tag{4.38}$$

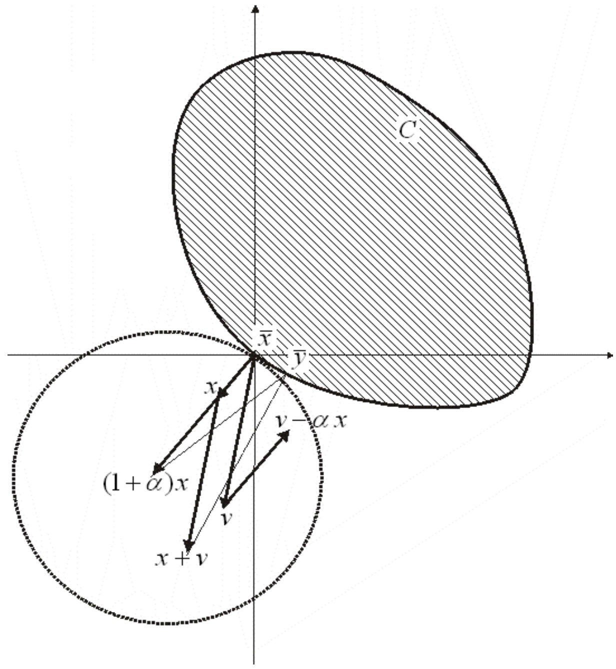

Figura 4.4: Representação geométrica da propriedade minimizante de $\bar{x}$.

onde

$$\Phi(v) := (1+\alpha)\|x\| - \|x + v - (1+\alpha)x\| = (1+\alpha)\|x\| - \|v - \alpha x\|.$$

Observe que a função Φ mede a distância do ponto $x+v$ à fronteira da bola $\overset{\circ}{B}\big((1+\alpha)x, (1+\alpha)\|x\|\big)$. Como $d_C(x) = \|x\| = \Phi(0)$, segue de (4.38), que

$$d_C(x+v) - d_C(x) \geq \Phi(v) - \Phi(0). \tag{4.39}$$

Porém, a função Φ é analítica em $\overset{\circ}{B}\big(0, \alpha\|x\|\big)$ (note que $x \neq 0$, pois $x \notin C$). Assim, podemos calcular o gradiente de Φ em relação a v: $\nabla_v \Phi(0) = \|x\|^{-1}x$.

Portanto, existem $M > 0$ e $\epsilon > 0$ tais que

$$\begin{aligned} \Phi(v) - \Phi(0) &\geq \big\langle \nabla_v \Phi(0), v - 0 \big\rangle - M\|v - 0\|^2 \\ &\geq \big\langle \|x\|^{-1}x, v \big\rangle - M\|v\|^2, \text{ para todo } y \in \epsilon B. \end{aligned}$$

Segue de (4.39), que

$$d_C(x+v) - d_C(x) \geq \big\langle \|x\|^{-1}x, v \big\rangle - M\|v\|^2, \text{ para todo } y \in \epsilon B.$$

Então, pela a Proposição 4.4, temos

$$\|x\|^{-1}x \in \partial^p d_C(x).$$

(ii) Primeiro, vamos mostrar que $\bar{x}(=0)$ é o único minimizador da função $y \to \|x-y\|$ sobre C. Supondo que y' é um minimizador arbitrário, pela propriedade ponto mais próximo segue que

$$\begin{aligned} \|x-y'\|^2 &\leq \|x-\bar{x}\|^2 \\ &\leq \|x-0\|^2, \end{aligned}$$

logo

$$-2\langle x, y'\rangle \leq -\|y'\|^2. \tag{4.40}$$

Mas, pela propriedade minimizante de $\bar{x}(=0)$, temos

$$(1+\alpha)^2\|x\|^2 \leq \|(1+\alpha)x-y'\|^2,$$

desta forma, obtemos

$$0 \leq -2\langle x, y'\rangle(1+\alpha)+\|y'\|^2.$$

Segue de (4.40), que

$$\alpha\|y'\|^2 \leq 0.$$

Mas, como $\alpha > 0$, concluímos que a única solução possível é quando $y' = 0$. Então, $\bar{x}(=0)$ é realmente o único minimizador.

Por hipótese, existem as seqüencias $\{x_i\}$ e $\{\xi_i\}$ tais que $x_i \longrightarrow x$ quando $i \longrightarrow \infty$ e $\xi_i \in \partial^p d_C(x_i)$ para cada i. Como $x \notin C$ e C é um conjunto fechado, podemos considerar que $x_i \notin C$ para todo i. Então, pela Proposição 4.11, existe seqüencia $\{y_i\}$ em C tal que

$$d_C(x_i) = \|x_i - y_i\| \tag{4.41}$$

e

$$\xi_i = \|x_i - y_i\|^{-1}(x_i - y_i). \tag{4.42}$$

Assim, podemos extrair uma subseqüência de $\{y_i\}$ (não reindexamos) tal que $y_i \longrightarrow \bar{y}$ para algum $\bar{y} \in C$. Pela continuidade da função distância, segue de (4.41) e (4.42), que

$$d_C(x) = \|x-\bar{y}\| \text{ e } \xi_i \longrightarrow \|x-\bar{y}\|^{-1}(x-\bar{y}).$$

Assim, $\bar{y} = \bar{x} = 0$. Logo podemos concluir que

$$y_i \longrightarrow 0 \text{ e } \xi_i \longrightarrow \|x\|^{-1}x = \xi. \tag{4.43}$$

Como os limites acima são independentes das subseqüências selecionadas inicialmente, as seqüencias originais satisfazem (4.43). ■

4.5 Descrição Analítica dos Subgradientes Limites

Nesta seção, estamos interessados em analisar o subgradiente limite e o subgradiente limite assintótico em termos de limites de subgradientes proximais. Para enunciar os principais teoremas deste tópico, precisamos do lema abaixo que pode ser resumindo da seguinte forma: os vetores normais proximais horizontais a um conjunto epígrafo podem ser aproximados por normais proximais.

Lema 4.3. *Sejam uma função semicontínua inferior* $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$, *um ponto* $\bar{x} \in dom(f)$ *e um ponto não-nulo* $\xi \in \mathbb{R}^k$ *tal que* $(\xi, 0) \in N^p_{epi(f)}(\bar{x}, f(\bar{x}))$. *Então, existem seqüencias convergentes* $x_i \xrightarrow{f} \bar{x}$, $\xi_i \longrightarrow \xi$ *e uma seqüencia de números positivos* $\lambda_i \downarrow 0$ *tais que*

$$(\xi_i, -\lambda_i) \in N^p_{epi(f)}(x_i, f(x_i)) \text{ para todo } i.$$

Demonstração: Tome $(\xi, 0) \in N^p_{epi(f)}(\bar{x}, f(\bar{x}))$ (um vetor normal proximal "horizontal"). Então, existe um ponto $\big(x, f(\bar{x})\big) \notin epi(f)$ e um número $\sigma > 0$ tal que $\big(\bar{x}, f(\bar{x})\big)$ é um ponto mais próximo de $\big(x, f(\bar{x})\big)$ em $epi(f)$ e $\xi = \sigma(x - \bar{x})$ (veja a Figura 4.5). Como $\big(\bar{x}, f(\bar{x})\big)$ é

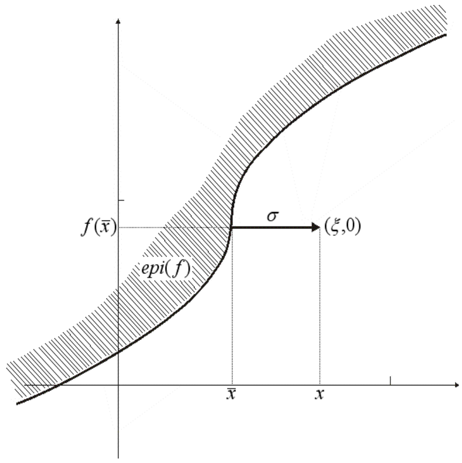

Figura 4.5: Normal proximal horizontal

o ponto mais próximo, logo existe algum $\alpha > 0$ tal que $\big(\bar{x}, f(\bar{x})\big)$ é o ponto mais próximo do ponto $\big(\bar{x} + (1 + \alpha)(x - \bar{x}), f(\bar{x})\big)$ em $epi(f)$.

Afirmamos que

$$d_{epi(f)}\big(x, f(\bar{x}) - t\big) > d_{epi(f)}\big(x, f(\bar{x})\big) \text{ para todo } t > 0. \tag{4.44}$$

Esta afirmação será demonstrada mais tarde. Seja $t_i \downarrow 0$ uma seqüência arbitrária. Aplicamos, para cada i, a Desigualdade do Valor Médio (Teorema 4.1) para função $d_{epi(f)}(\cdot)$ com o ponto base $\big(x, f(\bar{x})\big)$ e o conjunto $Y = \big\{(x, f(\bar{x}) - t_i)\big\}$. Então, existem seqüencias $\big(y_i, r_i\big) \longrightarrow$

$\big(x, f(\bar{x})\big)$ e $\big\{(\xi_i', -\lambda_i')\big\}$ em $\mathbb{R}^k \times \mathbb{R}$ satisfazendo $\big(\xi_i', -\lambda_i'\big) \in \partial^p d_{epi(f)}((y_i, r_i))$ para todo i tal que

$$d_{epi(f)}(x, f(\bar{x}) - t_i) - d_{epi(f)}(x, f(\bar{x})) < \big\langle(\xi_i', -\lambda_i'), (x, f(\bar{x}) - t_i) - (x, f(\bar{x}))\big\rangle \text{ para todo } i.$$

Logo

$$\lambda_i' t_i > d_{epi(f)}(x, f(\bar{x}) - t_i) - d_{epi(f)}(x, f(\bar{x})) > 0 \text{ por (4.44)}.$$

Isto implica que $\lambda_i' > 0$ para todo i.

Utilizando as Proposições 4.12 e 4.11, podemos concluir que

$$\big(\|\xi\|^{-1}\xi, 0\big) \in \partial^p d_{epi(f)}(x, f(\bar{x}))$$

e além disso, que existe uma seqüência (x_i, s_i) de pontos ponto mais próximos a (y_i, r_i) em $epi(f)$ tais que

$$\big(\xi_i', -\lambda_i'\big) \in \partial^p d_{epi(f)}(x_i, s_i) \text{ para todo } i \tag{4.45}$$

$$\big(x_i, s_i\big) \longrightarrow \big(\bar{x}, f(\bar{x})\big), \tag{4.46}$$

e

$$\big(\xi_i', -\lambda_i'\big) \longrightarrow \big(\|\xi\|^{-1}\xi, 0\big). \tag{4.47}$$

Defina

$$\big(\xi_i, -\lambda_i\big) := \big(\|\xi\|\xi_i', -\|\xi\|\lambda_i'\big) \text{ para cada } i.$$

Observe que $\xi \neq 0$ e $\lambda_i > 0$ para cada i. Segue de (4.45) e da Proposição 4.10, que

$$\big(\xi_i, -\lambda_i\big) \in N^p_{epi(f)}(x_i, f(x_i)) \text{ para todo } i.$$

e por (4.47), obtemos

$$\big(\xi_i, -\lambda_i\big) = \big(\|\xi\|\xi_i', -\|\xi\|\lambda_i'\big) \longrightarrow \big(\|\xi\|^{-1}\xi, 0\big) \text{ quando } i \longrightarrow \infty.$$

Assim, encontramos as seqüencias $\{x_i\}$, $\{\xi_i\}$ e $\{\lambda_i\}$ que satisfazem as propriedades declaradas no lema.

Para completar a demonstração, precisamos mostrar a afirmação

$$d_{epi(f)}(x, f(\bar{x}) - t) > d_{epi(f)}(x, f(\bar{x})) \text{ para todo } t > 0.$$

Suponha que a afirmação não seja satisfeita. Logo, existe $t > 0$ tal que

$$d_{epi(f)}(x, f(\bar{x}) - t) \leq d_{epi(f)}(x, f(\bar{x})). \tag{4.48}$$

Podemos encontrar $(y, r) \in \mathbb{R}^n \times \mathbb{R}$ com $r \geq f(y)$ tal que

$$d_{epi(f)}(x, f(\bar{x}) - t) = \|(x, f(\bar{x}) - t) - (y, r)\|.$$

Observa que $y \neq \bar{x}$. De fato, suponha que $y = \bar{x}$, logo

$$\big\|(x, f(\bar{x}) - t) - (\bar{x}, r)\big\| = \big\|(x - \bar{x}, t + r - f(\bar{x}))\big\| > \big\|x - \bar{x}\big\| = d_{epi(f)}(x, f(\bar{x})),$$

contradizendo (4.48). Assim,

$$\begin{aligned} d_{epi(f)}(x, f(\bar{x})) &\geq d_{epi(x)}(x, f(\bar{x}) - t) \\ &= \|(x, f(\bar{x}) - t) - (y, r)\| = \|(x, f(\bar{x})) - (y, r + t)\|. \end{aligned}$$

No entanto, podemos concluir que $(y, r+t) \in epi(f)$ em vista de $(y, r) \in epi(f)$ e $t > 0$. Então, $(y, r + t)$ é um ponto mais próximo de $(x, f(\bar{x}))$ em $epi(f)$ com $y \neq \bar{x}$. Isso é impossível, pois $(\bar{x}, f(\bar{x}))$ é o único ponto mais próximo de $(x, f(\bar{x}))$ em $epi(f)$. Assim temos uma contradição.

■

O subgradiente limite e subgradiente limite assintótico podem ser expressados como limites de subgradientes proximais em pontos vizinhos.

Teorema 4.3. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e os pontos $x \in dom(f)$ e $\xi \in \mathbb{R}^k$.*

(a) As seguintes condições são equivalentes:

(i) $\xi \in \partial f(x)$;

(ii) existe $x_i \xrightarrow{f} x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial^p f(x_i)$ para todo i.

(b) As seguintes condições são equivalentes:

(iii) $\xi \in \partial^\infty f(x)$;

(iv) existe $x_i \xrightarrow{f} x$, $\xi_i \longrightarrow \xi$ e $t_i \downarrow 0$ tal que $t_i^{-1}\xi_i \in \partial^p f(x_i)$ para todo i.

Demonstração: ***a)*** Mostremos a implicação da condição (i) para (ii). Seja $\xi \in \partial f(x)$. Por definição subdiferencial limite temos que $(\xi, -1) \in N_{epi(f)}(x, f(x))$. Segue da definição de cone normal limite que existem seqüências convergentes

$$(x_i, f(x_i)) \xrightarrow{epi(f)} (x, f(x)) \text{ e } (\xi_i', -t_i) \longrightarrow (\xi, -1) \ i \to \infty,$$

tais que

$$(\xi_i', -t_i) \in N^p_{epi(f)}(x_i, f(x_i)) \text{ para todo } i. \tag{4.49}$$

Observe que $t_i > 0$. Dividindo por t_i e definindo $\xi_i := t_i^{-1}\xi_i'$ podemos concluir que $\xi_i \in \partial^p f(x_i)$, para cada i e que $\xi_i \to \xi$.

Mostremos a implicação contrária. Considerando as seqüencias $x_i \xrightarrow{f} x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial^p f(x_i)$ para todo i. Para cada i, temos $(\xi_i, -1) \in N^p_{epi(f)}(x_i, f(x_i))$ e $(x_i, f(x_i)) \longrightarrow (x, f(x))$ quando $i \longrightarrow \infty$. Segue da definição de cone normal limite, que $(\xi, -1) \in N_{epi(f)}(x, f(x))$. Portanto, $\xi \in \partial f(x)$.

b) Mostremos a implicação (iii) para (iv). Suponha que $\xi \neq 0$. Seja $\xi \in \partial^\infty f(x)$, então $(\xi, 0) \in N_{epi(f)}(x, f(x))$. Segue da definição de cone normal limite, que existem seqüências convergentes

$$(x_i, f(x_i)) \xrightarrow{epi(f)} (x, f(x)) \text{ e } (\xi_i, -t_i) \longrightarrow (\xi, 0),$$

tais que $t_i \geq 0$ e

$$(\xi_i, -t_i) \in N^p_{epi(f)}(x_i, f(x_i)) \text{ para todo } i.$$

Desta maneira, temos dois casos a serem considerados:

- Considerando $t_i > 0$, temos que $(t_i^{-1}\xi_i, -1) \in N^p_{epi(f)}(x_i, f(x_i))$ para todo i. Portanto, $t_i^{-1}\xi_i \in \partial^p f(x_i)$ para todo i.

- considerando $t_i = 0$, segue do Lema 4.3, que existem seqüências convergentes $x_{i_j} \xrightarrow{f} x_i$, $\xi_{i_j} \longrightarrow \xi_i$ e uma subseqüência $t_{i_j} \downarrow 0$ tal que

$$(\xi_{i_j}, -t_{i_j}) \in N^p_{epi(f)}(x_{i_j}, f(x_{i_j})) \text{ para todo } j.$$

Usando o argumento da diagonal de Cantor podemos extrair subsequência $x_j' \xrightarrow{f} x$, $\xi_j' \longrightarrow \xi$ e $t_j' \downarrow 0$ de $\{x_{i_j}\}$, $\{\xi_{i_j}\}$ e $\{t_{i_j}\}$, respectivamente. Observe que

$$(\xi_j', -t_j') \in N^p_{epi(f)}(x_j', f(x_j')) \text{ para todo } j.$$

Logo, podemos concluir que

$$(\xi_j, -1) \in N^p_{epi(f)}(x_j, f(x_j)) \text{ para todo } j,$$

onde $\xi_j = t_j^{-1}\xi_j'$.

Agora, vamos supor que $\xi = 0$. Como $x \in dom(f)$, segue da Proposição 3.4, que existem $(\xi', -\beta) \neq (0,0)$ tal que $(\xi', -\beta) \in N_{epi(f)}(x, f(x))$. Pela definição de cone normal limite, existem seqüencias $(x_i, f(x_i)) \xrightarrow{epi(f)} (x, f(x))$ e $(\xi_i', \beta_i) \longrightarrow (\xi, -\beta)$ tais que $\beta_i \geq 0$ e

$$(\xi_i', -\beta_i) \in N^p_{epi(f)}(x_i, f(x_i)) \text{ para todo } i.$$

Assim, temos dois casos a considerar:

- para $\beta > 0$, temos resultado análogo da parte (a) do Teorema 4.3;
- para $\beta = 0$, segue da parte (b) do Teorema 4.3 (no caso em que $\xi' \neq 0$), que existem seqüencias $x_i \xrightarrow{f} x$, $\xi_i' \longrightarrow \xi'$ e $\beta_i \longrightarrow \beta$ tal que $\beta_i > 0$ e $\beta_i^{-1}\xi_i' \in \partial^p f(x_i)$ para todo i. Escolhendo uma seqüencia $\alpha_i \downarrow 0$ arbitrária e definindo os pontos $\xi_i = \alpha_i \xi_i'$ e $t_i = \alpha_i \beta_i$, para cada i, obtemos

$$\xi_i \longrightarrow \xi(=0),\ t_i \downarrow 0 \text{ e } t_i^{-1}\xi_i \in \partial^p f(x_i) \text{ para todo } i.$$

Mostremos a implicação contrária ($(iv) \Rightarrow (iii)$). Dadas as seqüencias $\{x_i\}$, $\{t_i\}$ e $\{\xi_i\}$, com $t_i^{-1}\xi_i \in \partial^p f(x_i)$ para todo i, temos que $\big(\xi_i, -t_i\big) \in N^p_{epi(f)}(x_i, f(x_i))$. Mas, por hipótese, $(\xi_i, -t_i) \longrightarrow (\xi, 0)$ e $(x_i, f(x_i)) \xrightarrow{epi(f)} (x, f(x))$. Segue da definição de cone normal limite, que $(\xi, 0) \in N_{epi(f)}(x, f(x)$. Então, $\xi \in \partial^\infty f(x)$, completando a demonstração. ■

Agora, temos condição de demonstrar a relação do cone normal limite com a subdiferencial limite para função distância.

Corolário 4.2. *Seja $C \subset \mathbb{R}^k$ um conjunto fechado e $x \in C$ um ponto. Então*

$$\partial d_C(x) = N_C(x) \cap B.$$

Demonstração: Para facilitar a demonstração, vamos dividir em duas partes:

(1) Mostremos a inclusão $\partial d_C(x) \subset N_C(x) \cap B$. Tomando um vetor $\xi \in \partial d_C(x)$, segue do Teorema 4.3, que existe $x_i \longrightarrow x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial^p d_C(x_i)$ para cada i. Mas, pela Proposição 4.10, temos $\xi_i \in N_C^p(x_i) \cap B$, isto é, $\xi_i \in N_C^p(x_i)$ e $\|\xi_i\| \leq 1$ para cada i. Então pela definição de cone proximal limite, temos $\xi \in N_C(x) \cap B$.

(2) Agora, vamos mostrar a inclusão inversa. Tomando $\xi \in N_C(x) \cap B$, logo existem seqüencias $x_i \longrightarrow x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in N_C^p(x_i) \cap B$. Aplicando a Proposição 4.10, temos $\xi_i \in \partial^p d_C(x_i)$. Então, pelo Teorema 4.3, concluímos que $\xi \in \partial d_C(x)$. ■

Também, podemos caracterizar o subgradiente limite e subgradiente limite assintótico em termos de limites de subgradientes estritos.

Teorema 4.4. *Seja $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e os pontos $x \in dom(f)$ e $\xi \in \mathbb{R}^k$.*

(a) As seguintes condições são equivalentes:

(i) $\xi \in \partial f(x)$;

(ii) *existe* $x_i \xrightarrow{f} x$ *e* $\xi_i \longrightarrow \xi$ *tal que* $\xi_i \in \hat{\partial} f(x_i)$ *para todo i.*

(b) As seguintes condições são equivalentes:

(iii) $\xi \in \partial^\infty f(x)$;

(iv) *existe* $x_i \xrightarrow{f} x$, $\xi_i \longrightarrow \xi$ *e* $t_i \downarrow 0$ *tal que* $t_i^{-1}\xi_i \in \hat{\partial} f(x_i)$ *para todo i.*

Demonstração: As demonstrações das implicações $(i) \Longrightarrow (ii)$ e $(iii) \Longrightarrow (iv)$, segue do Teorema 4.3 e sabendo que $\partial^p f(x_i) \subset \hat{\partial} f(x_i)$ para todo i. As implicações inversas são conseqüências da Proposição 3.3. ■

4.6 Subdiferencial de Clarke

Definição 4.6. *Seja* $f : \mathbb{R}^k \to \mathbb{R}$ *uma função Lipschitz na vizinhança de* $x \in dom(f)$. *A derivada direcional de Clarke de* f *em* x *na direção de* $v \in \mathbb{R}^k$, *denotada por* $f^\circ(x; v)$, *é definida por*

$$f^\circ(x; v) := \limsup_{y \longrightarrow x,\ t \downarrow 0} \frac{f(y + tv) - f(y)}{t}.$$

O gradiente generalizado de Clarke de f *em* x, *denotado por* $\bar{\partial} f(x)$, *é o conjunto*

$$\bar{\partial} f(x) = \big\{\xi : \big\langle \xi, v \big\rangle \leq f^\circ(x; v) \text{ para todo } v \in \mathbb{R}^k\big\}.$$

Também, $\bar{\partial} f(x)$ *é chamado de subdiferencial de Clarke de* f *em* x.

A derivada direcional generalizada $f^\circ(x; v)$, possui algumas propriedades básicas.

Proposição 4.13. *Seja* $f : \mathbb{R}^k \to \mathbb{R}$ *uma função Lipschitz na vizinhança de* x *com constante de Lipschitz* K, *sendo que* $x \in dom(f)$. *Então, a função* $v \to f^\circ(x; v)$ *com domínio* $\mathbb{R}^k$ *tem as seguintes propriedades:*

(i) $f^\circ(x; v)$ *é limitada pela constante de Lipschitz* K;

(ii) A derivada direcional generalizada é homogênea positiva, isto é,

$$f^\circ(x; \lambda v) = \lambda f^\circ(x; v) \text{ e } \lambda \geq 0;$$

(iii) $f^\circ(x; v)$ *é convexa;*

(iv) $f^\circ(x;-v) = (-f)^\circ(x;v)$.

Demonstração: ***(i)*** Por hipótese, dado $x \in \mathbb{R}^k$, existem $K > 0$ e $\epsilon > 0$ tais que

$$\|f(z) - f(x)\| \leq K\|z - y\| \text{ para todo } y, z \in B(x,\epsilon).$$

Segue da definição de derivada direcional generalizada, que

$$\begin{aligned} \|f^\circ(x;v)\| &\leq \limsup_{y \longrightarrow x,\ t\downarrow 0} \left\| \frac{f(y+tv) - f(y)}{t} \right\| \leq \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{1}{t}\, K\|y + tv - y\| \\ &\leq K\|v\| \end{aligned}$$

para qualquer que seja $v \in \mathbb{R}^k$. Escolhendo $\|v\| = 1$, temos $\|f^\circ(x;v)\| \leq K$.

(ii) Seja $\lambda > 0$, segue que

$$f^\circ(x;\lambda v) = \lambda \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{f(y + t(\lambda v) - f(y)}{\lambda t}$$

Fazendo a mudança de variável $\tau = \lambda t$, obtemos

$$f^\circ(x;\lambda v) = \lambda \limsup_{y \longrightarrow x,\ \tau\downarrow 0} \frac{f(y + \tau v) - f(y)}{\tau}.$$

Então,

$$f^\circ(x;\lambda v) = \lambda f^\circ(x;v).$$

(iii) Como $f^\circ(x;\cdot)$ é positiva homogênea, resta mostrar que a função $f^\circ(x;\cdot)$ é "sub-aditividade", isto é, $f^\circ(x;v+w) \leq f^\circ(x;v) + f^\circ(x;w)$ para todo $v, w \in \mathbb{R}^k$. Segue da definição, que

$$\begin{aligned} f^\circ(x, v+w) &= \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{f(y + t(v+w)) - f(y)}{t} \\ &= \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{f\big((y+tw) + tv\big) - f(y+tw) + f(y+tw) - f(y)}{t} \\ &\leq \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{f\big((y+tw) + tv\big) - f(y+tw)}{t} + \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{f(y+tw) - f(y)}{t}. \end{aligned}$$

Como $(y + tw) \in B(x;\epsilon)$ e $0 \leq t \leq \epsilon$, logo

$$f^\circ(x, v+w) \leq f^\circ(x;v) + f^\circ(x;w).$$

(iv) Pela definição de derivada direcional generalizada, obtemos

$$\begin{aligned} f^\circ(x;-v) &= \limsup_{y \longrightarrow x,\ t\downarrow 0} \frac{f(y - tv) - f(y)}{t} \\ &= \limsup_{z \longrightarrow x,\ t\downarrow 0} \frac{(-f)(z + tv) - (-f)(z)}{t}, \end{aligned}$$

onde $z = y - tv$. Então,

$$f^\circ(x; -v) = (-f)^\circ(x; v).$$

■

Também, temos as seguintes observações:

Observação 4.2. *As imagens da subdiferencial de Clarke $x \to \bar{\partial} f(x)$ são conjuntos não-vazios, compactos e convexos.*

Observação 4.3. *Os elementos em $\bar{\partial} f(x)$ são uniformemente limitados na norma Euclidiana pela constante de Lipschitz de f na vizinhança de x.*

A derivada direcional generalizada pode ser interpretada como sendo a função suporte de $\overline{\partial} f(x)$:

Proposição 4.14. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz numa vizinhança de um ponto $x \in \mathbb{R}^k$. Então*

$$f^\circ(x; v) = \max\Big\{\big\langle \xi, v \big\rangle : \xi \in \bar{\partial} f(x)\Big\} \text{ para todo } v \in \mathbb{R}^k.$$

Demonstração: Pela definição de $\bar{\partial} f(x)$, temos que $f^\circ(x; v) \geq \max\Big\{\big\langle \xi, v \big\rangle : \xi \in \overline{\partial} f(x)\Big\}$ para todo v. Então, falta mostrar a igualdade para qualquer que seja o vetor v. Sabemos que $v \to f^\circ(x; v)$ é convexa. Fixe v arbitrário e escolha $\xi \in \partial f^\circ(x; v)$ (no sentido de Análise Convexa). Então,

$$\big\langle \xi, w - v \big\rangle \leq f^\circ(x, w) - f^\circ(x, v) \text{ para todo } w \in \mathbb{R}^k.$$

Substituindo $w = \alpha \bar{w}$, onde $\bar{w} \in \mathbb{R}^k$ arbitrário e $\alpha > 0$, temos

$$\big\langle \xi, \alpha \bar{w} - v \big\rangle \leq f^\circ(x, \alpha \bar{w}) - f^\circ(x, v) \text{ para todo } \alpha \bar{w} \in \mathbb{R}^k.$$

Segue da Proposição 4.13(ii), que

$$\alpha \big\langle \xi, \bar{w} - \frac{1}{\alpha} v \big\rangle \leq \alpha f^\circ(x, \bar{w}) - f^\circ(x, v).$$

Dividindo por α e passando o limite quando $\alpha \longrightarrow \infty$, obtemos

$$\big\langle \xi, \bar{w} \big\rangle \leq f^\circ(x, \bar{w}).$$

Logo, $\xi \in \overline{\partial} f(x)$. Por outro lado, tomando $w = 0$, obtemos diretamente $\big\langle \xi, v \big\rangle \geq f^\circ(x, v)$ e portanto,

$$f^\circ(x, v) = \max\Big\{\big\langle \xi, v \big\rangle : \xi \in \overline{\partial} f(x)\Big\}.$$

■

Temos também a seguinte propriedade da Subdiferencial de Clarke.

Proposição 4.15. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz numa vizinhança de um ponto $x \in \mathbb{R}^k$. Então*

$$\overline{\partial}(-f)(x) = -\overline{\partial} f(x).$$

Demonstração: Basta verificar que os dois conjuntos acima têm a mesma função suporte, isto é $(-f)^\circ(x; v) = f^\circ(x; -v)$ para qualquer $v \in \mathbb{R}^k$. Mas isto é o item (iii) da Proposição 4.13. ■

Quando a derivada direcional clássica $f'(x, \cdot)$ existir, ela coincide com a derivada direcional generalizada $f^\circ(x, \cdot)$.

Proposição 4.16. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função convexa e Lipschitz numa vizinhança de um ponto $x \in \mathbb{R}^k$. Então, a derivada direcional $f'(x; v)$ existe e temos $f'(x; v) = f^\circ(x; v)$.*

Demonstração: Sabemos que $f^\circ(x; v) \geq f'(x; v)$. Resta mostrarmos que $f^\circ(x; v) \leq f'(x; v)$.

Afirmamos que a convexidade de f implica que a função

$$h(t) := \frac{f(x' + tv) - f(x')}{t},$$

é não-decrescente para $t > 0$ pequeno e x' arbitrário na vizinhança de x. De fato, tome $s > t > 0$ arbitrário. Então $x' + th = \frac{t}{s}(x' + sh) + (1 - \frac{t}{s})x'$ e

$$\begin{aligned} h(t) &= \frac{f(\frac{t}{s}(x' + sh) + (1 - \frac{t}{s})x') - f(x')}{t} \\ &\leq \frac{\frac{t}{s} f(x' + sh) + (1 - \frac{t}{s}) f(x') - f(x')}{t} \quad \text{(pela convexidade de } f\text{)} \\ &= \frac{1}{s}(f(x' + sh) - f(x')) = h(s). \end{aligned}$$

Segue do fato que f satisfaz a condição de Lipschitz que $h(t) \geq -K\|v\|$ onde K é a constante de Lipschitz de f e v uma direção arbitrária. Isto implica que

$$f'(x; v) = \inf_{t>0} \frac{f(x' + tv) - f(x')}{t}.$$

Fixando $\delta > 0$ e podemos observar que

$$f^\circ(x; v) = \lim_{\varepsilon \downarrow 0} \sup_{\|x' - x\| \leq \varepsilon\delta} \sup_{0<t<\varepsilon} \frac{f(x' + tv) - f(x')}{t}.$$

Logo,

$$f^\circ(x;v) = \lim_{\varepsilon\downarrow 0} \sup_{\|x'-x\|\leq\varepsilon\delta} \frac{f(x'+\epsilon v)-f(x')}{\epsilon}.$$

Mas, para todo $x' \in B(x;\varepsilon\delta)$ com ε suficientemente pequeno, temos

$$\left\|\frac{f(x'+\varepsilon v)-f(x')}{\varepsilon} - \frac{f(x+\varepsilon v)-f(x)}{\varepsilon}\right\| \leq 2\delta K \tag{4.50}$$

e assim

$$f^\circ(x;v) \leq \lim_{\varepsilon\downarrow 0}\Big(\frac{f(x+\varepsilon v)-f(x)}{\varepsilon} + 2\delta K\Big) = f'(x;v) + 2\delta K.$$

Como δ é arbitrário, temos que $f^\circ(x;v) \leq f'(x;v)$. ∎

O resultado seguinte, temos a relação da subdiferencial de Clarke com a subdiferencial limite.

Proposição 4.17. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz na vizinhança de algum ponto $x \in \mathbb{R}^k$. Então:*

$$\bar{\partial} f(x) = co\partial f(x).$$

Demonstração: Antes de iniciarmos a demonstração, observe que os dois conjuntos acima são conjuntos convexos fechados. Logo, é suficiente mostrar que

$$f^\circ(x;v) = \max\Big\{\big\langle v,\xi\big\rangle : \xi \in co\partial f(x)\Big\},$$

para todo $v \in \mathbb{R}^k$.

a) Mostremos que $f^\circ(x;v) \geq \max\{\langle v,\xi\rangle : \xi \in \partial f(x)\}$ para todo $v \in \mathbb{R}^k$. Escolha $\xi \in \partial f(x)$ e $v \in \mathbb{R}^k$ quaisquer. Segue do Teorema 4.3, que existem seqüencias $x_i \xrightarrow{f} x$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial^p f(x_i)$ para todo i. Para cada i, existem $M_i > 0$ e $\delta_i > 0$ tal que

$$f(x) - f(x_i) \geq \big\langle \xi_i, x - x_i\big\rangle - M_i\|x-x_i\|^2 \text{ para todo } x \in B(x_i,\delta_i).$$

Escolhendo $x = x_i + t_i v$ com $t_i \in (0,1)$ e substituindo na desigualdade acima, temos

$$f(x_i+t_iv) - f(x_i) \geq t_i\big\langle\xi_i, v\big\rangle - M_i t_i^2\|v\|^2.$$

Dividindo por t_i e passando o limite quando $t_i \downarrow 0$ obtemos

$$\frac{f(x_i+t_iv)-f(x_i)}{t_i} \geq \big\langle\xi_i, v\big\rangle.$$

Portanto,

$$f^\circ(x;v) \geq \limsup_{i \longrightarrow \infty} t_i^{-1}\big(f(x_i + t_i v) - f(x_i)\big) \geq \limsup_{i \longrightarrow \infty} \big\langle \xi_i, v \big\rangle = \big\langle \xi, v \big\rangle.$$

Mas ξ e v são arbitrários, logo

$$f^\circ(x;v) \geq \max\Big\{ \big\langle v, \xi \big\rangle : \xi \in \partial f(x) \Big\}$$

b) Mostremos que $f^\circ(x;v) \leq \max\Big\{ \big\langle v, \xi \big\rangle : \xi \in \partial f(x) \Big\}$ para todo $v \in \mathbb{R}^k$. Seja $v \in \mathbb{R}^k$ arbitrário. Então, podemos encontrar seqüencias convergentes $x_i \longrightarrow x$ e $t_i \downarrow 0$ tal que

$$f^\circ(x,v) = \lim_{i \longrightarrow \infty} \frac{f(x_i + t_i v) - f(x_i)}{t_i}.$$

Aplicando o Teorema da Desigualdade do Valor Médio (Teorema 4.1) na função f com ponto base x_i e no "conjunto Y:= $\{x_i + t_i v\}$" para cada i. Então, existem $z_i \in B(x_i, t_i\|v\|)$ e $\xi_i \in \partial f(z_i)$ tal que

$$f(x_i + t_i v) - f(x_i) \leq t_i \big\langle \xi_i, v \big\rangle \text{ para todo } z_i \in Y. \tag{4.51}$$

Observe que $z_i \longrightarrow x$. Como f é uma função Lipschitzeana na vizinhança de x, temos que ξ_i são seqüencias uniformemente limitadas. Logo, existe subseqüência $\xi_i \longrightarrow \xi$ para algum $\xi \in \mathbb{R}^k$. Segue do Teorema 4.3, que $\xi \in \partial f(x)$. Dividindo t_i e passando o limite quando $i \longrightarrow \infty$ em (4.51), obtemos

$$f^\circ(x;v) \leq \big\langle \xi, v \big\rangle \text{ para algum } \xi \in \partial f(x).$$

Então,

$$f^\circ(x;v) \leq \max\Big\{ \big\langle \xi, v \big\rangle : \xi \in \partial f(x) \Big\},$$

para todo $v \in \mathbb{R}^k$, completando a demonstração. ■

O lema abaixo, será necessário para demonstrar o Teorema 4.5.

Lema 4.4. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz na vizinha de x.*

(i) Se a função f é diferenciável em x (diferenciável de Gâteaux), então existe a derivada $\nabla f(x)$ e $\nabla f(x) \in co\partial f(x)$;

(ii) Se a função f é continuamente diferenciável em x, então $co\partial f(x) = \{\nabla f(x)\}$.

Demonstração: (i) Por definição de derivada direcional de f, temos

$$f'(x;v) = \langle \nabla f(x), v \rangle \text{ para cada } v \in \mathbb{R}^k.$$

Observe que $f'(x;v) \leq f^\circ(x;v)$. Então, segue da definição de subdiferencial de Clarke que

$$\nabla f(x) \in co\partial f(x).$$

(ii) Suponhamos que $f \in C^1$ numa vizinhança de x e fixe $v \in \mathbb{R}^k$. Então, para cada y próximo de x e $t > 0$ próximo de 0, temos

$$\frac{f(y+tv) - f(y)}{t} = \langle f'(z), v \rangle$$

para algum $z \in (y, y+tv)$, pelo Teorema do Valor Médio Clássico. Como $y \longrightarrow x$ e $t \downarrow 0$ temos que $z \to x$. Tomando o limsup, para $y \longrightarrow x$ e $t \downarrow 0$, na equação acima obtemos

$$\limsup_{y \to x,\ t \downarrow 0} \frac{f(y+tv) - f(y)}{t} = \langle \nabla f(x), v \rangle$$

i.e., $f^\circ(x;v) = \langle \nabla f(x), v \rangle$. Seja $\xi \in co\partial f(x)$ arbitrária. Assim, $\langle \xi, v \rangle \leq \langle \nabla f(x), v \rangle$ para todo $v \in \mathbb{R}^k$. Isto implica que $\langle \xi, v \rangle = \langle \nabla f(x), v \rangle$ para todo $v \in \mathbb{R}^k$, ou seja, $\xi = \{\nabla f(x)\}$.

■

A seguir enunciamos, sem demonstração, uma caracterização da subdiferencial de Clarke para funções que satisfazem a condição de Lipschitz. Esta caracterização diz que para calcular a subdiferencial podemos tomar o limite dos gradientes da função em pontos próximos do ponto de interesse. A demonstração pode ser encontrada, por exemplo, em [34, 10]

Teorema 4.5. *Seja $f : \mathbb{R}^k \to \mathbb{R}$ uma função que satisfaz a condição de Lipschitz na vizinhança de $x \in \mathbb{R}^k$ e qualquer subconjunto $\Omega \subset \mathbb{R}^k$ que tenha medida de Lebesgue nula. Então,*

$$co\partial f(x) = co\{\xi : \exists\ x_i \longrightarrow x, x_i \notin \Omega, \nabla f(x_i) \text{ existe e } \nabla f(x_i) \longrightarrow \xi\}. \tag{4.52}$$

4.7 Regras das Subdiferenciais

Nesta seção, estamos interessado em apresentar generalizações das regras da derivada clássica (regra da soma, da cadeia, do produto) para funções não-diferenciáveis.

Vejamos a Regra da Soma Elementar.

Lema 4.5. *Sejam $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior e $g : \mathbb{R}^k \to \mathbb{R}$ uma função de classe C^2 na vizinhança de x. Suponha que $C \subset \mathbb{R}^k$ seja um conjunto fechado e $x \in \{int(C)\} \cap \{dom(f)\}$ um ponto. Então,* *(i)* $\partial^p(f + g + \Psi_C(x)) = \partial^p f(x) + \{\nabla g(x)\}$;

(ii) $\partial(f + g + \Psi_C(x)) = \partial f(x) + \{\nabla g(x)\}$;

(iii) $\partial^\infty(f + g + \Psi_C(x)) = \partial^\infty f(x)$.

Demonstração: ***(i)*** Por hipótese a função g é de classe de C^2 na vizinhança de x, logo para algum $\epsilon > 0$, sendo que $B(x, \epsilon) \subset int(C)$, existe $m > 0$ tal que

$$\left\| g(y) - g(x) - \left\langle \nabla g(x), y - x \right\rangle \right\| \leq m\|y - x\|^2 \text{ para todo } y \in B(x, \epsilon).$$

Seja $\xi \in \mathbb{R}^k$. Então, por conseqüência desta observação, existem $M > 0$ e $\delta > 0$ tais que

$$\left\langle \xi + \nabla g(x), y - x \right\rangle \leq (f + g + \Psi_C)(y) - (f + g + \Psi_C)(x) + M\|y - x\|^2 \ \forall y \in B(x, \delta)$$

se, e somente se, existem $n > 0$ e $\delta > \epsilon' > 0$ tais que

$$\left\langle \xi, y - x \right\rangle \leq f(y) - f(x) + n\|y - x\|^2 \ \forall y \in B(x, \epsilon').$$

Portanto,

$$(\xi + \nabla g(x)) \in \partial^p(f + g + \Psi_C)(x) \text{ se, e somente se, } \xi \in \partial^p f(x).$$

(ii) $\eta \in \partial(f + g + \Psi_C)(x)$ se, e somente se, existem seqüências $x_i \xrightarrow{f+g+\Psi_C} x$ e $\eta_i \longrightarrow \eta$ tal que $\eta_i \in \partial^p(f + g + \Psi_C)(x_i)$ para todo i. Isto é equivalente, pelo item (i) acima, que $\eta_i = \xi_i + \nabla g(x_i)$, $\xi_i \in \partial f(x_i)$ e $\xi_i \to \xi \in \partial f(x)$. Tudo isso é equivalente a $\eta = \xi + \nabla g(x)$ e $\xi \in \partial f(x)$. Portanto $\partial(f + g + \Psi)(x) = \partial f(x) + \{\nabla g(x)\}$.

(iii) $\eta \in \partial^\infty(f + g + \Psi_C)(x)$, se, e somente se, existem seqüências $x_i \xrightarrow{f+g+\Psi_C} x$, $t_i \downarrow 0$ e $\eta_i \longrightarrow \eta$ tais que $t_i^{-1}\eta_i \in \partial^p(f + g + \Psi_C)(x_i)$ para todo i. Isto é equivalente, pelo item (i) acima, que

$$t_i^{-1}\eta_i = t_i^{-1}(t_i\xi_i) + \nabla g(x_i), \quad \xi_i \in \partial^p f(x_i) \ \forall \ i.$$

Multiplicando por t_i^{-1} e passando o limite quando $i \longrightarrow \infty$, podemos concluir que $t_i\xi_i \to \eta$. Utilizando o Teorema 4.3 obtemos a equivalência: $\eta \in \partial^\infty(f + g + \Psi_C)(x) \Leftrightarrow \eta \in \partial^\infty f(x)$. ■

Os próximos resultados serão necessários para demonstrar a regra da soma.

Teorema 4.6 (Princípio da Função Marginal). *Seja* $F : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R} \cup \{+\infty\}$ *uma função semicontínua inferior e* $(\bar{x}, \bar{u}) \in dom(F)$ *um ponto. Suponha que,*

(i) $\bar{x}$ *é o único minimizador da função* $x \longrightarrow F(x, \bar{u})$*;*

(ii) existe um conjunto limitado $A \subset \mathbb{R}^n$ tal que

$$dom(F(\cdot, u)) \subset A \text{ para todo } u \in \mathbb{R}^m.$$

Defina a função marginal $V : \mathbb{R}^m \longrightarrow \mathbb{R} \cup \{+\infty\}$ por

$$V(u) := \min_{x \in \mathbb{R}^n} F(x, u).$$

Então, temos as seguintes afirmações:

a) Para cada $u \in \mathbb{R}^m$, a função $x \longrightarrow F(x, u)$ possui um minimizador (um ponto de mínimo);

b) A função V é semicontínua inferior;

c) Se existem seqüencias $u_i \xrightarrow{V} \bar{u}$ e $\{x_i\}$ tais que $V(u_i) = F(x_i, u_i)$ para cada i, então $(x_i, u_i) \xrightarrow{F} (\bar{x}, \bar{u})$;

d) O subgradiente de V satisfaz as seguintes relações:

$$\begin{aligned} \partial V(\bar{u}) &\subset \big\{\xi \in \mathbb{R}^m : (0, \xi) \in \partial F(\bar{x}, \bar{u})\big\} \\ \partial^\infty V(\bar{u}) &\subset \big\{\xi \in \mathbb{R}^m : (0, \xi) \in \partial^\infty F(\bar{x}, \bar{u})\big\}; \end{aligned}$$

e) Se $\big\{\xi \in \mathbb{R}^m : (0, \xi) \in \partial^\infty F(\bar{x}, \bar{u})\big\} = \{0\}$, então existe um ponto $\xi \in \mathbb{R}^m$ tal que

$$(0, \xi) \in \partial F(\bar{x}, \bar{u}).$$

Demonstração: ***a)*** Tome $u \in \mathbb{R}^m$ qualquer. Se $dom(F(\cdot, u)) = \emptyset$, então qualquer x é um minimizador de F. Assim, suponha que $dom(F(\cdot, u)) \neq \emptyset$. Como $dom(F(\cdot, u))$ é um conjunto compacto e F é uma função semicontínua inferior, segue do Teorema de Weierstrass, que existe um ponto que minimiza a função $x \longrightarrow F(x, u)$.

b) Fixe $u \in dom(V)$ e defina $\overline{V}(u) := \liminf_{i \longrightarrow \infty} V(u_i)$. Logo, precisamos verificar a seguinte desigualdade $\overline{V}(u) \geq V(u)$. Esta desigualdade é evidente se $\overline{V} = \infty$, assim suponha que $\overline{V}$ é finito. Seja $u_i \longrightarrow u$ uma seqüencia, com $V(u_i) \longrightarrow \overline{V}(u)$. Para cada i, podemos escolher um minimizador x_i da função $x \longrightarrow F(x, u_i)$. Por (a), podemos encontrar uma subseqüência se necessário para garantir que $x_i \longrightarrow x$, para algum x no conjunto compacto $dom(F(\cdot, u))$. Assim, $(x_i, u_i) \longrightarrow (x, u)$. Como F é uma função semicontínua inferior, segue que

$$V(u) \leq F(x, u) \leq \liminf_{i \longrightarrow \infty} F(x_i, u_i) = \lim_{i \longrightarrow \infty} V(u_i) = \overline{V}(u).$$

c) Por hipótese, existem as seqüencias $u_i \xrightarrow{V} \bar{u}$ e $\{x_i\}$ tal que $V(u_i) = F(x_i, u_i)$ para cada i. Observe que $V(u_i)$ é uma função finita para i suficientemente grande. O ponto $x_i \in dom(F(\cdot, u_i))$ para cada i, onde $dom(F(\cdot, u_i))$ é um conjunto compacto. Assim, podemos encontrar uma subseqüência $x_i \longrightarrow x$ para algum $x \in \mathbb{R}^n$. Como F é uma função semicontínua inferior, segue que

$$V(\bar{u}) = \lim_{i \longrightarrow \infty} V(u_i) = \lim_{i \longrightarrow \infty} F(x_i, u_i) \geq F(x, \bar{u}) \geq V(\bar{u}).$$

Logo,

$$F(x_i, u_i) \longrightarrow F(x, \bar{u}).$$

Podemos observar que $x = \bar{x}$, pois a função $F(\cdot, \bar{u})$ tem um único minimizador. Do fato que os limites são independentes da subseqüência extraída, concluímos que a seqüência original satisfaz $(x_i, u_i) \xrightarrow{F} (\bar{x}, \bar{u})$.

d) Tomando $\xi \in \partial V(\bar{u})$ segue do Teorema 4.3, que existem seqüencias $u_i \xrightarrow{V} \bar{u}$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial^p V(u_i)$ para todo i. Para cada i, existem $\epsilon_i > 0$ e $M_i > 0$ tal que

$$\big\langle \xi_i, u - u_i \big\rangle \leq V(u) - V(u_i) + M_i \|u - u_i\|^2 \text{ para todo } u \in B(u_i, \epsilon_i).$$

Como x_i é minimizador único de $x \longrightarrow F(x, u_i)$ e $V(u) \leq F(x, u)$ para todo x e u, obtemos

$$\big\langle (0, \xi_i), (x - x_i, u - u_i \big\rangle \quad \leq \quad F(x, u) - F(x_i, u_i) + M\|(x, u) - (x_i, u_i)\|^2,$$

para todo $(x, u) \in B\big((x_i, u_i), \epsilon_i\big)$. Logo, $(0, \xi_i) \in \partial^p F(x_i, u_i)$. Utilizando a propriedade (c) acima, temos $(x_i, u_i) \xrightarrow{F} (\bar{x}, \bar{u})$, e assim pelo Teorema 4.3, obtemos $(0, \xi) \in \partial F(x, u)$.

Mostremos que $\partial^\infty V(\bar{u}) \subset \big\{\xi \in \mathbb{R}^m : (0, \xi) \in \partial^\infty F(\bar{x}, \bar{u})\big\}$. Tomando $\xi \in \partial^\infty V(\bar{u})$, segue do Teorema 4.3, que existem $u_i \xrightarrow{V} \bar{u}$, $\xi_i \longrightarrow \xi$ e $t_i \downarrow 0$ tal que $t_i^{-1}\xi_i \in \partial^p V(u_i)$ para todo i. Para cada i, temos $t_i^{-1}(0, \xi_i) \in \partial^p F(x_i, u_i)$. Como x_i é o único minimizador da função $x \longrightarrow F(x, u_i)$ temos que $(x_i, u_i) \xrightarrow{F} (\bar{x}, \bar{u})$. Assim, podemos concluir que $(0, \xi) \in \partial^\infty F(\bar{x}, \bar{u})$.

e) Seja $\bar{x}$ o único minimizador da $F(\cdot, \bar{u})$, logo existe $\epsilon > 0$ tal que

$$F(\bar{x}, \bar{u}) = \min\big\{F(x, \bar{u}) : x \in B(\bar{x}, \epsilon)\big\}.$$

Escolha qualquer $r > 0$ e defina a função $\tilde{F} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ por

$$\tilde{F}(x, u) := F(x, u) + r\|x - \bar{x}\|^2 + \Psi_C(x),$$

com $C := B(\bar{x}, \epsilon)$. Observe que o termo de penalidade $r\|x - \bar{x}\|^2$ é necessário para garantir que $\tilde{F}(\cdot, \bar{u})$ possui um único ponto de mínimo. Considerando que $g(x) = r\|x - \bar{x}\|^2$ é uma função de classe C^2 na vizinhança de x, segue do Lema 4.5, que

$$\partial \tilde{F}(\bar{x}, \bar{u}) \;\; = \;\; \partial F(\bar{x}, \bar{u}).$$

Definindo a função marginal $V(\cdot)$ por

$$V(u) := \inf_x \tilde{F}(x, u)$$

segue da propriedade (d) que

$$\partial V(\bar{u}) \subset \big\{\xi : (0, \xi) \in \partial F(\bar{x}, \bar{u})\big\} \quad \text{e} \quad \partial^\infty V(\bar{u}) \subset \big\{\xi : (0, \xi) \in \partial^\infty F(\bar{x}, \bar{u})\big\}.$$

Mas, como $\partial^\infty V(\bar{u}) = \{0\}$, segue do Teorema 4.2 que V é uma função que satisfaz a condição de Lipschitz. Logo, $\partial V(\bar{u})$ não é vazia. Portanto, existe $\xi \in \mathbb{R}^m$ tal que $(0, \xi) \in \partial F(\bar{x}, \bar{u})$.

■

Teorema 4.7 (**Subgradiente Limite Parcial**). *Seja* $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ *uma função semicontínua inferior e* $(\bar{x}, \bar{y})$ *um ponto. Suponha que*

$$(0, \eta) \in \partial^\infty f(\bar{x}, \bar{y}) \text{ implica } \eta = 0. \tag{4.53}$$

Então,

$$\begin{aligned} \partial_x f(\bar{x}, \bar{y}) \;\; &\subset \;\; \big\{\xi : \exists\, \eta \text{ tal que } (\xi, \eta) \in \partial f(\bar{x}, \bar{y})\big\} \\ \partial_x^\infty f(\bar{x}, \bar{y}) \;\; &\subset \;\; \big\{\xi : \exists\, \eta \text{ tal que } (\xi, \eta) \in \partial^\infty f(\bar{x}, \bar{y})\big\}. \end{aligned}$$

Demonstração: Tomando $\xi \in \partial_x f(\bar{x}, \bar{y})$, segue do Teorema 4.3, que existem as seqüencias $x_i \xrightarrow{f(\cdot,\bar{y})} \bar{x}$ e $\xi_i \longrightarrow \xi$ tal que $\xi_i \in \partial_x^p f(x_i, \bar{y})$ para todo i. Para cada i, podemos escolher $\epsilon_i > 0$ e $M_i > 0$ tal que x_i é um único minimizador da função $F_i(\cdot, \bar{y})$, onde

$$F_i(x, y) := f(x, y) - \big\langle \xi_i, x - x_i \big\rangle + M_i\|x - x_i\|^2 + \Psi_{B(x_i, \epsilon_i)}(x).$$

Aplicando o Teorema 4.6(e) e o Lema 4.5 podemos encontrar η_i tal que

$$(0, \eta_i) \in \partial F_i(x_i, \bar{y}) = \partial f(x_i, \bar{y}) - (\xi_i, 0).$$

Logo, $(\xi_i, \eta_i) \in \partial f(x_i, \bar{y})$. Suponha que $\{\eta_i\}$ é uma seqüencia limitada para todo i. Assim, podemos extrair uma subseqüência $\{\eta_i\} \longrightarrow \eta$, para algum $\eta \in \mathbb{R}^n$. Também temos

$(x_i, \bar{y}) \xrightarrow{f} (\bar{x}, \bar{y})$ e $(\xi_i, \eta_i) \longrightarrow (\xi, \eta)$ (subseqüências convergentes). Logo, segue do Teorema 4.3 que $(\xi, \eta) \in \partial f(\bar{x}, \bar{y})$. Assim, podemos concluir que

$$\partial_x f(\bar{x}, \bar{y}) \subset \big\{\xi : \exists\ \eta \text{ tal que } (\xi, \eta) \in \partial f(\bar{x}, \bar{y})\big\}.$$

Para completar a demonstração, temos que verificar a suposição acima. Suponha que $\{\eta_i\}$ não é uma seqüência limitada, logo podemos extrair outras subseqüências $|\eta_i| \longrightarrow \infty$ e $\|\eta_i\|^{-1}\eta_i \longrightarrow \nu$ para algum ν com $\|\nu\| = 1$. Assim, definindo $t_i := \|\eta_i\|^{-1}$, temos

$$t_i \downarrow 0 \text{ e } t_i(\xi_i, \eta_i) \in \partial^\infty(\bar{x}, \bar{u}).$$

Como $(x_i, \bar{y}) \longrightarrow (\bar{x}, \bar{y})$ segue do Teorema 4.3 (b), que

$$(0, \nu) \in \partial^\infty f(\bar{x}, \bar{y}).$$

Assim, temos uma contradição à hipótese (4.53).

Mostremos agora a relação

$$\partial_x^\infty f(\bar{x}, \bar{y}) \subset \big\{\xi : \exists\ \eta \text{ tal que } (\xi, \eta) \in \partial^\infty f(\bar{x}, \bar{y})\big\}.$$

Tomando $\xi \in \partial_x^\infty f(\bar{x}, \bar{y})$, segue do Teorema 4.3, que existem $x_i \xrightarrow{f(\cdot,\bar{y})} \bar{x}$, $t_i \downarrow 0$ e $\xi_i \longrightarrow \xi$ tal que $t_i^{-1}\xi_i \in \partial_x^p f(x_i, \bar{y})$ para todo i. Pelo Teorema 4.6 (d) e pelo Lema 4.5, para cada i, existe η_i tal que

$$t_i^{-1}(\xi_i, t_i\eta_i) \in \partial f(x_i, \bar{y}).$$

Como $t_i \downarrow 0$ e η_i é uma seqüência limita podemos extrair uma subseqüência convergente de $t_i\eta_i$ convergindo a η'. Pelo Teorema 4.3 concluímos que $(\xi, \eta') \in \partial^\infty f(\bar{x}, \bar{y})$. ■

A regra da soma é considerada como uma conseqüência do Princípio da Função Marginal.

Teorema 4.8 (Regra da Soma). *Sejam $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ funções semicontínuas inferiores e $\bar{x} \in \cap_i dom(f_i)$ para $i = 1, 2, ..., m$. Defina a função $f(\bar{x}) := f_1(\bar{x}) + f_2(\bar{x}) + ... + f_m(\bar{x})$. Suponha que $\upsilon_i \in \partial^\infty f_i(\bar{x})$ para $i = 1, 2, ..., m$ e $\sum_i \upsilon_i = 0$ implica que $\upsilon_i = 0$ para todo i (hipótese de não-degeneração). Então*

$$\partial f(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}) + ... + \partial f_m(\bar{x})$$

e

$$\partial^\infty f(\bar{x}) \subset \partial^\infty f_1(\bar{x}) + \partial^\infty f_2(\bar{x}) + ... + \partial^\infty f_m(\bar{x}).$$

Demonstração: Define a função $F : \mathbb{R}^n \times (\mathbb{R}^n)^m \to \mathbb{R} \cup \{+\infty\}$ por

$$F(x,y) := \sum_{i=1}^{m} f_i(x+y_i) \text{ para } y = (y_1,...,y_m).$$

Observe que $f(x) = F(x,0)$, logo $\partial f(\bar{x}) = \partial_x F(\bar{x},0)$. Tomando $(\xi,\eta) \in \partial^p F(x,y)$ para algum ponto $(x,y) \in dom(f)$, segue da Proposição 4.4, que existe $M > 0$ tal que

$$\big\langle (\xi,\eta),(x',y') - (x,y)\big\rangle \leq F(x',y') - F(x,y) + M\|(x',y') - (x,y)\|^2,$$

para todo (x',y') na vizinhança de (x,y). Expressando em termo de f_i, temos

$$\big\langle \xi, x'-x\big\rangle + \sum_{i=1}^{m}\langle \eta_i, y_i' - y_i\rangle \leq \sum_{i=1}^{m}\big[f_i(x'+y_i') - f_i(x+y_i)\big] + M\|(x',y') - (x,y)\|^2. \quad (4.54)$$

Fixando um indice $j \in \{1,...,m\}$ e escolhendo $x' = x$ e $y_i' = y_i$ para $i \neq j$, obtemos

$$\big\langle \eta_j, y_j' - y_j\big\rangle \leq f_j(x+y_j') - f_j(x+y_j) + M\|y_j' - y_j\|^2$$

para todos os y_j' próximo de y_j. Logo, $\eta_j \in \partial^p f_j(x+y_j)$.

Novamente, escolhendo $y_i' = x + y_i - x'$ para $i = 1,...,m$ e substituindo em (4.54), obtemos

$$\big\langle \xi, x'-x\big\rangle - \big\langle \sum_{i=1}^{m}\eta_i, x'-x\big\rangle \;\; \leq \;\; 0 + M\Big[\|x'-x\|^2 + \sum_{i=1}^{m}\|x-x'\|^2\Big],$$

para todo x' na vizinhança x. Logo,

$$\big\langle \xi - \sum_{i=1}^{m}\eta_i, x'-x\big\rangle \leq (m+1)M\|x'-x\|^2,$$

para todo x' na vizinhança x. Tome $x' = x + tv$ para algum $v \in \mathbb{R}^n$ e $t > 0$ suficientemente pequeno. Substituindo na desigualdade acima obtemos

$$t\big\langle \xi - \sum_{i=1}^{m}\eta_i, v\big\rangle \leq (m+1)Mt^2\|v\|^2.$$

Dividindo por t e passando o limite quando $t \downarrow 0$ obtemos

$$\big\langle \xi - \sum_{i=1}^{m}\eta_i, v\big\rangle \leq 0 \;\Rightarrow \xi = \sum_{i=1}^{m}\eta_i.$$

Portanto,

$$\partial^p F(x,y) \subset \big\{(\eta_1+\eta_2+...+\eta_m,\eta) : \eta = (\eta_1,\eta_2,...,\eta_m) \text{ e } \eta_i \in \partial^p f_i(x+y_i) \text{ para } i \in \{1,2,...,m\}\big\}.$$

Agora, tomando $(\xi, \eta) \in \partial F(\bar{x}, 0)$ qualquer, segue do Teorema 4.3 que existem seqüências convergentes $(x_k, y^k) \xrightarrow{F} (\bar{x}, 0)$ e $(\xi_k, \eta^k) \longrightarrow (\xi, \eta)$ tais que $(\xi_k, \eta^k) \in \partial^p F(x_k, y^k)$ para todo k. Pelos resultados anteriores temos

$$\xi_k = \sum_{i=1}^{m} \eta_i^k \quad \text{e} \quad \eta_i^k \in \partial^p f_i(x^k + y_i^k).$$

para todo i. Afirmando que

$$f_i(x_k + y_i^k) \longrightarrow f_i(\bar{x}), \text{para cada } i,$$

(esta afirmação será provado no final da segunda parte da demonstração), podemos escrever

$$\begin{aligned} x_k + y_i^k \xrightarrow{f} \bar{x} \;\text{ e } \eta_i^k \longrightarrow \eta_i \quad & \text{quando } k \longrightarrow \infty, \\ \eta_i^k \in \partial^p f_i(x_k + y_i^k) \quad & \text{para todo } k, \end{aligned}$$

Segue da definição de subdiferencial limite que $\eta_i \in \partial f_i(\bar{x})$ para cada i. Então,

$$\partial F(\bar{x}, 0) \subset \big\{(\eta_1 + \eta_2 + ... + \eta_m, \eta) : \eta = (\eta_1, \eta_2, ..., \eta_m) \text{ e } \eta_i \in \partial f_i(\bar{x}) \text{ para } i \in \{1, 2, ..., m\}\big\}.$$

Mostremos que $\partial^\infty f(\bar{x}) \subset \partial^\infty f_1(\bar{x}) + \partial^\infty f_2(\bar{x}) + ... + \partial^\infty f_m(\bar{x})$. Com argumento análogo baseado na representação de $\partial^\infty F$ em termos de subgradientes proximal em pontos vizinhos, obtemos

$$\partial^\infty F(\bar{x}, 0) \subset \big\{(\eta_1 + \eta_2 + ... + \eta_m, \eta) : \eta = (\eta_1, \eta_2, ..., \eta_m) \text{ e } \eta_i \in \partial^\infty f_i(\bar{x}) \text{ para } i \in \{1, 2, ..., m\}\big\}.$$

Observe que a hipótese de não-degeneração implica que em (4.53) do Teorema 4.7 é satisfeita. Isto implica que

$$\partial f(\bar{x}) = \partial_x F(\bar{x}, 0) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}) + ... + \partial f_m(\bar{x})$$

e

$$\partial^\infty f(\bar{x}) = \partial_x^\infty F(\bar{x}, 0) \subset \partial^\infty f_1(\bar{x}) + \partial^\infty f_2(\bar{x}) + ... + \partial^\infty f_m(\bar{x}).$$

Para finalizar a demonstração, precisamos verificar a afirmação

$$f_i(x_k + y_i^k) \longrightarrow f_i(\bar{x}) \text{ para todo } i.$$

Defina

$$\gamma_i := \limsup_{k \longrightarrow \infty} \big(f_i(x_k + y_i^k) - f_i(\bar{x})\big) \text{ para } i = 1, 2, ..., m.$$

Como f_i são funções semicontinuas inferior, podemos mostrar que $\gamma_i = 0$ para cada i. Extraindo uma subseqüência se necessária, podemos encontrar que

$$\gamma_i := \lim_{k \longrightarrow \infty} \big(f_i(x_k + y_i^k) - f_i(\bar{x}) \big) \text{ para } i = 1, 2, ..., m.$$

(aqui "lim"tem substituído por "lim sup"). Porém $(x_k, y^k) \xrightarrow{F} (\bar{x}, 0)$, logo

$$\lim_{k \longrightarrow \infty} F(x_k, y^k) \;=\; F(\bar{x}, 0),$$

ou seja,

$$\lim_{k \longrightarrow \infty} \sum_{i=1}^{m} f_i(x_k + y_i^k) \;=\; \sum_{i=1}^{m} f_i(\bar{x}),$$

Utilizando o Lema de Fatou, obtemos

$$\sum_{i=1}^{m} \liminf_{k \longrightarrow \infty} \big(f_i(x_k + y_i^k) \big) \leq \lim_{k \longrightarrow \infty} \sum_{i=1}^{m} f_i(x_k + y_i^k).$$

Assim

$$\sum_{i=1}^{m} \liminf_{k \longrightarrow \infty} \big(f_i(x_k + y_i^k) - f_i(\bar{x}) \big) \leq 0.$$

Como as funções são semicontínuas inferiormente, cada termo do somatório é não-negativo. Portanto, para cada i,

$$0 = \liminf_{k \longrightarrow \infty} \big(f_i(x_k + y_i^k) - f_i(\bar{x}) \big) = \lim_{k \longrightarrow \infty} \big(f_i(x_k + y_i^k) - f_i(\bar{x}) \big) = \gamma_i,$$

completando a demonstração. ■

Note que as inclusões apresentadas no teorema não podem ser substituídas por igualdades de conjuntos. O exemplo a seguir ilustra este fato para a primeira inclusão. Para ser verificada a segunda inclusão ainda é necessária a hipótese de não-degeneração para as subdiferenciais limites assintóticas, a qual não pode ser descartada (ver Exemplo 4.4).

Exemplo 4.3. *Sejam as funções $f : \mathbb{R} \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$ definidas por*

$$f(x) := |x| \quad e \quad g(x) := -|x|$$

Observe que,

$$\partial(f + g)(0) = \{0\}.$$

Por outro lado, temos $\partial f(0) = [-1, 1]$ e $\partial g(0) = \{-1\} \cup \{+1\}$. Logo,

$$\partial(f + g)(0) \neq \partial f(0) + \partial g(0).$$

Mas, vale a seguinte relação

$$\partial(f + g)(0) \subset \partial f(0) + \partial g(0).$$

O próximo exemplo, verifica que a hipótese de não-degeneração não pode ser descartada.

Exemplo 4.4. *Sejam as funções $f : \mathbb{R} \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$ definidas por*

$$f(x) := sgn\{x\}|x|^{1/2} \quad e \quad g(x) := -sgn\{x\}|x|^{1/2}$$

Observe que

$$\partial f(0) = \partial g(0) = \emptyset \quad e \quad \partial(f+g)(0) = \{0\}.$$

Logo,

$$\{0\} = \partial(f+g)(0) \not\subset \partial f(0) + \partial g(0) = \emptyset.$$

Mas, para subdiferencial limite assintótica, temos

$$\partial^\infty f(0) = [0, \infty) \quad e \quad \partial^\infty g(0) = (-\infty, 0].$$

Então existem números não-nulos a e b tal que $a \in \partial^\infty f(0)$, $b \in \partial^\infty g(0)$ e $a + b = 0$ (tome $a = 1$ e $b = -1$, por exemplo), contradizendo a hipótese de não-degeneração apropriada para regra da soma.

O próximo resultado é uma aplicação da Regra da Soma.

Proposição 4.18. *Sejam $G : \mathbb{R}^n \to \mathbb{R}^m$ uma função semicontínua inferior e $u \in \mathbb{R}^n$ um ponto. Suponha que a função G satisfaz a condição de Lipchitz na vizinhança de u. Então*

$$N_{gr(G)}(u, G(u)) = \big\{(\xi, -\eta) : \xi \in \partial\big\langle \eta, G(u)\big\rangle, \eta \in \mathbb{R}^m\big\}.$$

Demonstração: Dividimos a demonstração em duas partes:

1) Mostremos a inclusão $N_{gr(G)}(u, G(u)) \subset \big\{(\xi, -\eta) : \xi \in \partial\big\langle \eta_i, G(u)\big\rangle, \eta \in \mathbb{R}^m\big\}$. Tomando $(\xi, -\eta) \in N_{grG}(u, G(u))$, segue da definição de cone normal limite que existem $u_i \xrightarrow{G} u$ e $(\xi_i, -\eta_i) \longrightarrow (\xi, -\eta)$ tal que $(\xi_i, -\eta_i) \in N^p_{gr(G)}(u_i, G(u_i))$ para todo i. Para cada i, existem $M_i > 0$ e $\delta_i > 0$ tais que

$$\big\langle(\xi_i, -\eta_i), (u', G(u') - (u_i, G(u_i)))\big\rangle \leq M_i \|(u', G(u')) - (u_i, G(u_i))\|^2 \; \forall \; u' \in B(u_i, \delta_i).$$

Mas, como G é uma função Lipschitzeana em $B(u, \delta)$, obtemos

$$\big\langle \xi_i, , u' - u_i\big\rangle \leq \big\langle \eta_i, G(u')\big\rangle - \big\langle \eta_i, G(u_i)\big\rangle + M\|u' - u_i\|^2,$$

onde $M = (1+K^2)M_i$ sendo K a constante de Lipschitz. Então, $\xi_i \in \partial^p\big\langle \eta_i, G\big\rangle(u_i)$. Aplicando a Regra da Soma (Teorema 4.8) e o Teorema 4.2, obtemos

$$\begin{aligned}
\xi_i &\in \partial^p\big\langle \eta_i, G\big\rangle(u_i)\\
&\subset \partial^p\big[\big\langle \eta, G\big\rangle + \big\langle \eta_i - \eta, G\big\rangle\big](u_i)\\
&\subset \partial\big\langle \eta, G\big\rangle(u_i) + \partial\big\langle \eta_i - \eta, G\big\rangle(u_i)\\
&\subset \partial\big\langle \eta, G\big\rangle(u_i) + K\|\eta_i - \eta\|B.
\end{aligned}$$

Passando o limite quando $i \longrightarrow \infty$, obtemos que $\xi \in \partial\big\langle \eta, G\big\rangle(u)$.

*2)*Mostremos, agora, a inclusão inversa. Tome u arbitrário na vizinhança de $\bar{u}$ e ξ e η tais que $\xi \in \partial^p(\eta \cdot G)(u)$. Então, para todo $u' \in B(x,\delta)$, existem $M > 0$ e $\delta > 0$ tais que

$$\big\langle \xi, u' - u\big\rangle \leq \big\langle \eta, G\big\rangle(u') - \big\langle \eta, G\big\rangle(u) + M\|u' - u\|^2.$$

Reescrevendo esta expressão temos

$$\big\langle \xi, u' - u\big\rangle - \big\langle \eta, G(u') - G(u)\big\rangle \leq M\|u' - u\|^2 + M\|G(u') - G(u)\|^2.$$

Logo,

$$\big\langle (\xi, -\eta), (u', G(u') - (u, G(u))\big\rangle \leq M\|(u', G(u')) - (u, G(u))\|^2.$$

Pela definição de cone normal proximal obtemos $(\xi, -\eta) \in N^p_{grG}(u, G(u))$. Segue da Proposição 3.2(i), que $(\xi, -\eta) \in N_{grG}(u, G(u))$. ■

Agora, vamos apresentar a Regra da Cadeia:

Teorema 4.9 (Regra da Cadeia). *Seja $f(u) := g(G(u))$, onde a função $G : \mathbb{R}^n \to \mathbb{R}^m$ satisfaz a condição de Lipschitz na vizinhança de $\bar{u} \in \mathbb{R}^n$ e $g : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ é uma função semicontínua inferior com $G(\bar{u}) \in dom(g)$. Suponha que, $\eta \in \partial^\infty g(G(\bar{u}))$ e $0 \in \partial\big\langle \eta, G\big\rangle(\bar{u})$ implica $\eta = 0$. Então,*

$$\partial f(\bar{u}) \subset \big\{\xi : \exists\ \eta \in \partial g(G(\bar{u}))\ \textit{tal que}\ \xi \in \partial\big\langle \eta, G\big\rangle(\bar{u})\big\}$$

e

$$\partial^\infty f(\bar{u}) \subset \big\{\xi : \exists\ \eta \in \partial^\infty g(G(\bar{u}))\ \textit{tal que}\ \xi \in \partial\big\langle \eta, G\big\rangle(\bar{u})\big\}$$

Demonstração: Define uma função semicontínua inferior $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ por

$$F(x,u) := g(x) + \Psi_{gr(G)}(x,u) = \begin{cases} g(G(u)) & \text{se } x = G(u) \\ +\infty & \text{caso contrário} \end{cases} \tag{4.55}$$

De (4.55), podemos concluir que

$$f(u) := \min\{F(x,u) : x \in \mathbb{R}^m\}.$$

Observe $(\bar{u}, G(\bar{u})) \in dom(F)$, $dom(F)$ é limitado e $G(\bar{u})$ é um único minimizador da função $F(\cdot, \bar{u})$. Então, pelo Teorema 4.6 *d)* obtemos

$$\partial f(\bar{u}) \subset \big\{\xi : (0,\xi) \in \partial F(G(\bar{u}), \bar{u})\big\}$$

e

$$\partial^\infty f(\bar{u}) \subset \big\{\xi : (0,\xi) \in \partial^\infty F(G(\bar{u}), \bar{u})\big\}.$$

Segue do Teorema 4.8 (Regra da Soma) em (4.55) que

$$\partial F(\bar{x}, \bar{u}) \subset \partial g(\bar{x}) \times \{0\} + N_{gr(G)}(G(\bar{u}), \bar{u})$$

e

$$\partial^\infty F(\bar{x}, \bar{u}) \subset \partial^\infty g(\bar{x}) \times \{0\} + N_{gr(G)}(G(\bar{u}), \bar{u})$$

Aplicando a Proposição 4.18, podemos deduzir que se $\xi \in \partial f(\bar{u})$, então

$$(0,\xi) \in \partial g(\bar{x}) \times \{0\} + \big\{(-\eta', \xi') : \xi' \in \partial\big\langle \eta', G\big\rangle(\bar{u})\big\}.$$

Implica que,

$$\xi = \xi', \eta' \in \partial g(\bar{x}) \text{ para algum } (-\eta', \xi') \text{ satisfazendo } \xi' \in \partial\big\langle \eta', G\big\rangle(\bar{u}).$$

Logo,

$$\xi \in \partial\big\langle \eta', G\big\rangle(\bar{x}) \text{ para qualquer ponto } \eta' \in \partial g(\bar{x}).$$

Vamos verificar se a condição de não-degeneração da Regra da Soma é satisfeita. Como $\partial^\infty F(\bar{x}, \bar{u}) \subset \partial^\infty g(\bar{x}) \times \{0\} + N_{gr(G)}(\bar{u}, \bar{x})$, logo existem os pontos $(\xi_1, \eta_1) \in \partial^\infty g(\bar{x}) \times \{0\}$ e $(\xi_2, \eta_2) \in N_{gr(G)}(\bar{x}, \bar{u})$, ambos não-nulos, tais que

$$(\xi_1, \eta_1) + (\xi_2, \eta_2) = (0, 0).$$

Isto implica que $\xi_1 = 0 = \xi_2$ e $\eta_2 = -\eta_1$. Observe que η_1 pode ser não-nulo e satisfaz as seguintes relações

$$\eta_1 \in \partial^\infty g(\bar{x}) \;\text{ e }\; (-\eta_1, 0) \in N_{gr(G)}(G(\bar{u}), \bar{u}).$$

Segue da Proposição 4.18, que

$$(-\eta_1, 0) \in N_{gr(G)}(G(\bar{u}), \bar{u}) = \{(-\eta_1, \xi) : \xi \in \partial\big\langle \eta_1, G\big\rangle(\bar{u})\}.$$

Logo, $0 \in \partial\langle \eta_1, G\rangle(\bar{u})$ com $\eta_1 \in \partial^\infty g(G(\bar{u}))$. Assim, pela suposição inicial do teorema, obtemos $\eta_1 = 0$. Assim, $(\xi_1, \eta_1) + (\xi_2, \eta_2) = (0,0)$ implica que $(\xi_1, \eta_1) = (\xi_2, \eta_2) = (0,0)$.

Considerando os resultados anteriores e aplicando Teorema 4.8 obtemos

$$\begin{aligned} \partial f(\bar{u}) &\subset \big\{\xi : (0,\xi) = (\lambda, 0) + (-\eta, \xi), \lambda \in \partial g(G(\bar{u})), \xi \in \partial\langle \eta, G\rangle(\bar{u})\big\} \\ &\subset \big\{\xi : \exists\ \eta \in \partial g(G(\bar{u})) \text{ tal que } \xi \in \partial\langle \eta, G\rangle(\bar{u})\big\}. \end{aligned}$$

Analogamente, podemos obter o resultado abaixo

$$\partial^\infty f(\bar{u}) \subset \big\{\xi : \exists\ \eta \in \partial^\infty g(G(\bar{u})) \text{ tal que } \xi \in \partial\langle \eta, G\rangle(\bar{u})\big\}.$$

■

Se consideramos o conjunto $Y = \{\bar{y}\}$ e substituímos a subdiferencial limite ∂f por envoltório convexo da subdiferencial limite $co\partial f$, obtemos o Teorema do Valor Médio análogo ao resultado Clássico. Este resultado é conhecido por Teorema de Valor Médio Lebourg.

Teorema 4.10 (Teorema do Valor Médio de Lebourg). *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função localmente Lipschitz no conjunto aberto contendo o segmento linear $[\bar{x}, \bar{y}] \subset \mathbb{R}^n$. Então, existem $z \in \{\epsilon\bar{x} + (1-\epsilon)\bar{y} : 0 < \epsilon < 1\}$ e $\xi \in co\partial f(z)$ tal que*

$$f(\bar{y}) - f(\bar{x}) = \big\langle \xi, \bar{y} - \bar{x}\big\rangle.$$

Demonstração: Definimos a função $g : [0,1] \to \mathbb{R} \cup \{+\infty\}$ por

$$g(t) := f(\bar{x} + t(\bar{y} - \bar{x})) - t(f(\bar{y}) - f(\bar{x})).$$

Observe que g satisfaz a condição de Lipschitz e $g(0) = g(1) = f(\bar{x})$. Assim, uma das condições é satisfeita:

(a) g possui ponto de mínimo $\tau \in (0,1)$;

(b) $-g$ possui ponto de mínimo $\sigma \in (0,1)$.

Considerando o caso (a). Segue da desigualdade do cone normal proximal que $0 \in \partial^p g(\tau)$. Então, pelo Teorema da Regra da Cadeia, temos

$$\begin{aligned} 0 \in \partial g(\tau) &\subset \big\langle co\partial f(\bar{x} + \tau(\bar{y} - \bar{x})), \bar{y} - \bar{x}\big\rangle - (f(\bar{y}) - f(\bar{x})) \Rightarrow \\ 0 &= \big\langle \xi, \bar{y} - \bar{x}\big\rangle - (f(\bar{y}) - f(\bar{x})), \end{aligned}$$

para algum $\xi \in co\partial f(\bar{x} + \tau(\bar{y} - \bar{x}))$.

Agora, considerando o caso (b), segue da Regra da Cadeia, que

$$0 \in \partial(-g)(t) = \left\langle co\partial(-f)(\bar{x} + \sigma(\bar{y} - \bar{x})), \bar{y} - \bar{x} \right\rangle + (f(\bar{y}) - f(\bar{x})).$$

Mas, como

$$co\partial f(x) = -co\partial(-f(x)),$$

(veja Proposição 4.15), podemos concluir que

$$0 = \left\langle \xi', \bar{y} - \bar{x} \right\rangle - (f(\bar{y}) - f(\bar{x})),$$

para algum $\xi' \in co\partial f(\bar{x} + \sigma(\bar{y} - \bar{x}))$, completando a demonstração. ■

Como conseqüência da Regra da Cadeia, temos a Regra do Máximo e a Regra do Produto.

Proposição 4.19 (**Regra do Máximo**). *Sejam $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, ..., m$, as funções satisfazendo a condição de Lipschitz na vizinhança de $\bar{x} \in \mathbb{R}^n$. Defina $f(x) := \max_i f_i(x)$. Então*

$$\partial f(\bar{x}) \subset co\left\{\partial f_i(\bar{x}) : i = 1, 2, ..., m\right\}.$$

Demonstração: Vamos primeiro determinar o gradiente generalizado da função $g : \mathbb{R}^m \to \mathbb{R}$ definida por

$$g(y) := \max_i\{g_i(y) : i = 1, 2, ..., m\} \text{ para } y = (y_1, y_2, ..., y_m),$$

onde $g_i : \mathbb{R}^m \to \mathbb{R}$ é dada por $g_i(y) = y_i$. Observe que g é uma função convexa. De fato, o $epi(g) = \cap_i^m epi(g_i)$. Aplicando as Proposições A.1 e A.3 obtemos a convexidade de g. Denota por $I(y)$ o conjunto de indices i tal que $g_i(y) = g(y)$. Calculando a derivada direcional, obtemos

$$\begin{aligned} g'(y; v) &:= \lim_{t \downarrow 0} \frac{g(y + tv) - g(y)}{t} \\ &= \lim_{t \downarrow 0} \max_i \frac{(y_i + tv_i) - g(y)}{t}. \end{aligned}$$

Como t é suficientemente pequeno, podemos considerar apenas os indices $i \in I(y)$. Logo

$$\begin{aligned} g'(y; v) &= \lim_{t \downarrow 0} \max_{i \in I(y)} \frac{(y_i + tv_i - y_i)}{t} \\ &= \max_{i \in I(y)} v_i. \end{aligned}$$

Pela Proposição 4.16 obtemos

$$\max_{i\in I(y)} v_i = g'(y;v) = g^\circ(y;v) \geq \langle \eta, v\rangle \text{ para todo } v \in \mathbb{R}^m, \eta \in \partial g(y).$$

Assim $\partial g(y)$ consiste de todos vetores $\eta_1, \eta_2, ..., \eta_m$ tal que $\eta_i \geq 0$, $\sum \eta_i = 1$, $\eta_i = 0$ se $i \notin I(y)$. Então,

$$\partial g(y) = \Big\{\eta \in \mathbb{R}^m : \big\langle \eta, v\big\rangle \leq \max_{i\in I(y)} v_i \text{ para todo } \eta_i \geq 0, \sum_i \eta_i = 1 \text{ e } \eta_i = 0 \text{ se } i \notin I(y)\Big\}.$$

Defina $G(x) := (f_1, f_2, ..., f_m)(x)$ e observe que $f = g \circ G$. Segue do Teorema 4.9, que

$$\partial f(\bar{x}) = \partial(g \circ G)(\bar{x}) \subset \Big\{\xi : \exists \eta \in \partial g(G(\bar{x})) \text{ tal que } \xi \in \partial\big\langle \eta, G\big\rangle(\bar{x})\Big\}.$$

Logo,

$$\partial f(\bar{x}) \subset \Big\{\partial(\sum_{i=i}^{m} \eta_i f_i)(\bar{x}) \text{para todo } \eta_i \geq 0, \sum_i \eta_i = 1 \text{ e } \eta_i = 0 \text{ se } i \notin I(G(\bar{x})).\Big\}$$

Nota que a hipótese de não-degeneração é satisfeita pelo Teorema 4.2. ∎

Teorema 4.11 (Regra do Produto). *Sejam $f_i : \mathbb{R}^n \to \mathbb{R}$, para $i = 1, 2, ..., m$, as funções que satisfazem a condição de Lipschitz na vizinhança de $\bar{x} \in \mathbb{R}^n$. Defina $f(x) := f_1(x)f_2(x)...f_m(x)$. Então,*

$$\partial f(\bar{x}) \subset \partial\Bigg(\sum_i^m \prod_{j\neq i} f_j(\bar{x}) f_i\Bigg)(\bar{x}).$$

Demonstração: Sejam $g(y) := \prod_i y_i$ para $y = (y_1, y_2, ..., y_m)$ e $G(x) := (f_1, f_2, ..., f_m)(x)$ para $x \in \mathbb{R}^n$. Segue do Teorema 4.9 que

$$\partial f(x) = \partial(g \circ G)(x) \subset \Big\{\xi : \exists\ \eta \in \partial g(G(x)) \text{ tal que } \xi \in \partial\big\langle \eta, G\big\rangle(x)\Big\}.$$

Como a função $g \in C^\infty$, podemos verificar $\eta = \Big(\prod_{j\neq 1} f_j(\bar{x}), \prod_{j\neq 2} f_j(\bar{x}), ..., \prod_{j\neq m} f_j(\bar{x})\Big)$. Isto implica facilmente que

$$\partial f(\bar{x}) \subset \partial\Bigg(\sum_i^m \prod_{j\neq i} f_j(\bar{x}) f_i\Bigg)(\bar{x}).$$

∎

Capítulo 5

Problema de Controle Ótimo Envolvendo Funções Não-diferenciáveis

Neste capítulo, estudamos a forma forte do Princípio do Máximo para problema de controle ótimo com restrições mistas envolvendo funções não-diferenciáveis. Os resultados aqui apresentados podem ser encontrados em [14], [15] e [16].

5.1 Formulação do Problema

Consideramos o problema de controle ótimo do tipo:

$$
(Q)\begin{cases}
\text{Minimize} \quad h(x(a), x(b)) \\
\text{sujeita a} \\
\begin{array}{rcll}
\dot{x}(t) & = & f(t, x(t), u(t)) & \text{q.t. } t \in [a,b] \\
0 & = & b(t, x(t), u(t)) & \text{q.t. } t \in [a,b] \\
u(t) & \in & U(t) & \text{q.t } t \in [a,b] \\
(x(a), x(b)) & \in & C &
\end{array}
\end{cases}
$$

sendo dadas as funções $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $b : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$, uma multifunção $U : [a,b] \rightrightarrows \mathbb{R}^m$ e o conjunto $C \subset \mathbb{R}^n \times \mathbb{R}^n$.

Neste problema vamos considerar que a função custo h e a equação de estado $\dot{x} = f$ como funções não-diferenciáveis. Denominamos restrições mistas pelo fato de que as funções $f(t, x(t), u(t))$ e $b(t, x(t), u(t))$ do problema (Q) são compostas de uma função de controle $u : [a,b] \to \mathbb{R}^m$ e a função absolutamente contínua $x : [a,b] \to \mathbb{R}^n$ (correspondente a $u(t)$).

5.2 Condições Necessárias de Otimalidade

Para mostrar a forma forte do Princípio do Máximo, precisamos primeiro enunciar a Inclusão de Euler-Lagrange. Sua demonstração pode ser encontrada em Pinho & Vinter [14].

Denotamos por $AC([a,b];\mathbb{R}^n)$ o espaço de funções absolutamente contínuas de $[a,b]$ para $\mathbb{R}^n$ e $L^1([a,b];\mathbb{R}^k)$ o espaço de funções integráveis de $[a,b]$ para $\mathbb{R}^n$.

Teorema 5.1. *[14, Teorema 3.2] Tome $\epsilon > 0$. Seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ um minimizador local para o problema*

$$\begin{cases} \textit{Minimize} \quad h(x(a), x(b)) \\ \textit{Sujeito a} \\ \begin{array}{rcll} \dot{x}(t) & = & f(t, x(t), u(t)) & q.t.\ t \in [a,b] \\ u(t) & \in & U(t) & q.t.\ t \in [a,b] \\ (x(a), x(b)) & \in & C & \end{array} \end{cases}$$

Suponha que as seguintes hipóteses sejam satisfeitas:

(i) $f(\cdot, x, u)$ é mensurável no sentido de Lebesgue e existe $K \in L^1$ tal que

$$\|f(t,x,u) - f(t,x',u')\| \leq K(t)\|(x,u) - (x',u')\|,$$

para todo $x, x' \in B(\bar{x}(t), \epsilon)$, $u, u' \in B(\bar{u}(t), \epsilon)$ e quase todo $t \in [a,b]$.

(ii) O gráfico de U é um conjunto mensurável no sentido de Borel e $B(\bar{u}(t), \epsilon) \cap U(t)$ é um conjunto fechado para quase todo $t \in [a,b]$.

(iii) a função h satisfaz a condição de Lipschitz localmente e C é um conjunto fechado.

Então existem $\lambda \geq 0$, $p(\cdot) \in AC([a,b];\mathbb{R}^n)$ e $\zeta(\cdot) \in L^1([a,b];\mathbb{R}^n)$ tais que

$$\begin{array}{rcll} \lambda + \|p(\cdot)\|_{L^\infty} & = & 1 & \\ (-\dot{p}(t), \dot{\bar{x}}(t), \zeta(t)) & \in & co\ \partial\big\langle p(t), f(t, \bar{x}, \bar{u}(t))\big\rangle & q.t.\ t \in [a,b] \\ \zeta(t) & \in & \overline{co} N_{U(t)}(\bar{u}(t)) & q.t.\ t \in [a,b] \\ (p(a), -p(b)) & \in & N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial h(\bar{x}(a), \bar{x}(b)) & \end{array}$$

onde $\partial\big\langle p(t), f(t, \bar{x}, \bar{u}(t))\big\rangle$ denota a subdiferencial limite de $(x,p,u) \to \big\langle p, f(t,x,u)\big\rangle$ para t fixo.

Também na demonstração será utilizanda a propriedade de M-matriz, que podem ser encontrados em Horn & Johnson [21] e Berman & Plemmons [4].

Definição 5.1. *Seja A uma matriz real $m \times m$. Dizemos que a matriz A é uma M-matriz se $a_{ij} \leq 0, i \neq j$ e as partes reais dos autovalores da matriz A são positivas.*

Lema 5.1. *[4] As seguintes afirmações são equivalentes:*

(i) A é uma M-matriz,

(ii) Existe A^{-1} e $A^{-1} \geq 0$;

(iii) $a_{ii} > 0$ e $a_{ii} > \sum_{j \neq i} |a_{ij}|$ para $i = 1, 2, ..., m$, ou seja, a matriz A é estritamente diagonalmente dominante por linha.

Aqui $A^{-1} \geq 0$ significa que todos os elementos da matriz A são não-negativos.

5.2.1 Problema com Restrições de Igualdade

Consideramos o seguinte conjunto de hipóteses com referência a um processo $(\bar{x}(\cdot), \bar{u}(\cdot))$ de (Q) e algum escalar $\epsilon > 0$:

H1. (**Condição de mensurabilidade**) Para cada $x \in \mathbb{R}^n$, as funções $f(\cdot, x, \cdot)$ e $b(\cdot, x, \cdot)$ são mensuráveis com relação à sigma-álgebra do produto $\mathcal{L} \times \mathcal{B}$, onde $\mathcal{L}$ e $\mathcal{B}$ denotam a sigma-álgebra de Lebesgue de subconjuntos de $[a, b]$ e a sigma-álgebra de Borel de subconjuntos de $\mathbb{R}^m$, respectivamente;

H2. Para quase todo $t \in [a, b]$, existe uma função integrável $K_f(t)$ tal que

$$\|f(t, x'(t), u(t)) - f(t, x(t), u(t))\| \leq K_f(t)\|x'(t) - x(t)\|,$$

para todo $x(t), x'(t) \in B(\bar{x}(t), \epsilon)$ e todo $u(t) \in U(t)$;

H3. Para quase todo $t \in [a, b]$, $b(t, \cdot, u)$ é uma função continuamente diferenciável e existe uma constante $K_b > 0$ tal que

$$\|b(t, x'(t), u(t)) - b(t, x(t), u(t))\| \leq K_b\|x'(t) - x(t)\|,$$

para todo $x(t), x'(t) \in B(\bar{x}(t), \epsilon)$ e todo $u(t) \in U(t)$;

H4. (**Condição de interioridade**) existe um número positivo δ tal que

$$\delta B \subset b(t, \bar{x}(t), U(t)),$$

para quase todo $t \in [a, b]$;

H5. A função h satisfaz a condição de Lipschitz na vizinhança de $(\bar{x}(a), \bar{x}(b))$ e o conjunto C é fechado;

H6. O gráfico de $U(\cdot)$ é um conjunto mensurável no sentido de Borel;

H7. (**Condição de convexidade**) O conjunto $\{f(t, x(t), u(t)), b(t, x(t), u(t)) : u(t) \in U(t)\}$ é convexo para cada $t \in [a, b]$.

Assim, podemos apresentar as condições necessárias de otimalidade na forma de Princípio do Máximo para o problema (Q).

Teorema 5.2 (Princípio do Máximo). *Seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ um processo minimizador do problema (Q). Suponha que, para algum $\epsilon > 0$, as hipóteses $H1 - H7$ são satisfeitas. Então existem $p(\cdot) \in AC([a, b]; \mathbb{R}^n)$, $q(\cdot) \in L^1([a, b]; \mathbb{R}^k)$ e $\lambda \geq 0$ tais que:*

(i) $\lambda + \|p\|_{L^\infty} = 1$,

(ii) $-\dot{p}(t) \in co\partial_x H(t, \bar{x}(t), p(t), q(t), \bar{u}(t))$ *para quase todo* $t \in [a, b]$,

(iii) $H(t, \bar{x}(t), p(t), q(t), \bar{u}(t)) = \max_{u \in U(t)}\{H(t, \bar{x}(t), p(t), q(t), u)\}$ *para quase todo* $t \in [a, b]$,

(iv) $(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial h(\bar{x}(a), \bar{u}(b))$.

Demonstração: Para demonstrar este teorema, precisamos adicionar uma outra hipótese:
H* Existem uma função $c_f \in L^1$ e um número c_b tais que

$$\|f(t, \bar{x}(t), u)\| \leq c_f(t) \quad \text{e} \quad \|b(t, \bar{x}(t), u)\| \leq c_b,$$

para todo $u \in U(t)$ e quase todo $t \in [a, b]$. Esta hipótese será eliminada no final da demonstração.

Para facilitar a demonstração, vamos dividir em cinco etapas.

Etapa 1: Defina o vetor linha $e_i = (e_{i1}, e_{i2}, ..., e_{ik}) \in \mathbb{R}^k$ da forma

$$e_{ij} = \begin{cases} \delta_1 & \text{se } i = j, \\ -\frac{\delta_1}{k} & \text{se } i \neq j \end{cases}$$

onde $\delta_1 \in (0, \delta)$, sendo que δ é dado pela hipótese H4. Então, por H4 e pelo Teorema da Seleção Mensurável, podemos encontrar funções de controle $\{u_1, u_2, ..., u_k\}$ e $\{v_1, v_2, ..., v_k\}$ tais que

$$\begin{aligned} b(t, \bar{x}(t), u_i(t)) &= e_i, \\ b(t, \bar{x}(t), v_i(t)) &= -e_i. \end{aligned}$$

Fixe i. Segue da hipótese H3, que para quase todo $t \in [a,b]$ temos

$$\begin{aligned} b_i(t,x,u_i(t)) &\geq b_i(t,\bar{x}(t),u_i(t)) - K_b\|x-\bar{x}(t)\| \\ &= e_{ii} - K_b\|x-\bar{x}(t)\| \\ &= \delta_1 - K_b\|x-\bar{x}(t)\| \end{aligned} \tag{5.1}$$

e também

$$b_i(t,x,\bar{u}(t)) \leq K_b\|x-\bar{x}(t)\|. \tag{5.2}$$

Para $j \neq i$, temos

$$\begin{aligned} b_j(t,x,u_i(t)) &\leq b_j(t,\bar{x}(t),u_i(t)) + K_b\|x-\bar{x}(t)\| \\ &= e_{ij} + K_b\|x-\bar{x}(t)\| \\ &= -\frac{\delta_1}{k} + K_b\|x-\bar{x}(t)\| \end{aligned} \tag{5.3}$$

e também

$$b_j(t,x,u_i(t)) \geq -\frac{\delta_1}{k} - K_b\|x-\bar{x}(t)\|, \tag{5.4}$$

$$\|b_j(t,x,\bar{u}(t))\| \leq K_b\|x-\bar{x}(t)\|. \tag{5.5}$$

Defina a variação da função b por

$$\Delta b(t,x,u) := b(t,x,u) - b(t,x,\bar{u}(t)). \tag{5.6}$$

Segue de (5.1) e (5.2), que

$$\Delta b_i(t,x,u_i(t)) \geq \delta_1 - 2K_b\|x-\bar{x}(t)\| \text{ para } i \text{ fixo.} \tag{5.7}$$

De (5.3) e (5.5), obtemos

$$\Delta b_j(t,x,u_i(t)) \leq -\frac{\delta_1}{k} + 2K_b\|x-\bar{x}(t)\| \text{ para } j \neq i. \tag{5.8}$$

e de (5.4) e (5.5) obtemos

$$\Delta b_j(t,x,u_i(t)) \geq -\frac{\delta_1}{k} - 2K_b\|x-\bar{x}(t)\| \text{ para } j \neq i. \tag{5.9}$$

Como $x \in B(\bar{x},\epsilon_1)$ podemos escolher $\epsilon_1 < \min\left\{\frac{\delta_1}{2kK_b}, \epsilon\right\}$ com $\epsilon > 0$, para quase todo $t \in [a,b]$. Segue de (5.7) e (5.8) que

$$\Delta b_i(t,x,u_i(t)) > \delta_1 - 2K_b\epsilon_1 > 0 \text{ para } i \text{ fixo e } \Delta b_j(t,x,u_i(t)) < 0 \text{ para } j \neq i.$$

A mesma análise aplicada à $b(t, \bar{x}(t), v_i(t)) = -e_i$ acarreta

$$\begin{aligned} \|b_i(t, x, v_i(t)) + \delta_1\| &\leq K_b\|x - \bar{x}(t)\| \text{ para } i \text{ fixa} \\ \|b_j(t, x, v_i(t)) - \frac{\delta_1}{k}\| &\leq K_b\|x - \bar{x}(t)\| \text{ para } j \neq i, \end{aligned}$$

e

$$\begin{aligned} \|b_i(t, x, \bar{v}(t))\| &\leq K_b\|x - \bar{x}(t)\| \text{ para } i \text{ fixa} \\ \|b_j(t, x, \bar{v}(t))\| &\leq K_b\|x - \bar{x}(t)\| \text{ para } j \neq i. \end{aligned}$$

Utilizando, novamente, o fato que $x \in B(\bar{x}(t), \epsilon_1)$ para quase todo $t \in [a, b]$, obtemos

- para $j = i$
$$\Delta b_i(t, x, v_i(t)) \leq -\delta_1 + 2K_b\|x - \bar{x}(t)\| < 0; \tag{5.10}$$

- para $j \neq i$
$$\Delta b_j(t, x, v_i(t)) \geq \frac{\delta_1}{k} - 2K_b\|x - \bar{x}(t)\| > 0 \tag{5.11}$$

e por outro lado, temos

$$\Delta b_j(t, x, v_i(t)) \leq \frac{\delta_1}{k} + 2K_b\|x - \bar{x}(t)\|. \tag{5.12}$$

Defina as matrizes $k \times k$ por

$$B_1(t, x) := \big(\Delta b(t, x, u_1(t)), \Delta b(t, x, u_2(t)), ..., \Delta b(t, x, u_k(t))\big),$$

e

$$B_2(t, x) := \big(\Delta b(t, x, v_1(t)), \Delta b(t, x, v_2(t)), ..., \Delta b(t, x, v_k(t))\big).$$

Para continuar a demonstração, precisamos provar o seguinte lema:

Lema 5.2. *Se existir um $\epsilon' > 0$ tal que para todo $x \in B(\bar{x}(t), \epsilon')$, então as matrizes $B_1(t, x)$ e $-B_2(t, x)$ são M-matrizes, para quase todo $t \in [a, b]$.*

Demonstração: Mostremos que a matriz $B_1(t, x)$ é uma M-matriz. Pelas desigualdades (5.8) e (5.9), temos

$$\big\|\Delta b_j(t, x, u_i(t))\big\| - \frac{\delta_1}{k} \leq \big\|\Delta b_j(t, x, u_i(t)) - \big(-\frac{\delta_1}{k}\big)\big\| \leq 2K_b\|x - \bar{x}(t)\|.$$

Isto implica que

$$\big\|\Delta b_j(t, x, u_i(t))\big\| \leq \frac{\delta_1}{k} + 2K_b\|x - \bar{x}(t)\|$$

Segue que

$$\begin{aligned}
\sum_{\substack{j=1\\ j\neq i}}^{k}\big\|\Delta b_j(t,x,u_i(t))\big\| &\leq \sum_{\substack{j=1\\ j\neq i}}^{k}\big(\frac{\delta_1}{k}+2K_b\|x-\bar{x}(t)\|\big)\\
&= \frac{\delta_1(k-1)}{k}+2K_b\|x-\bar{x}(t)\|(k-1)\\
&\leq \delta_1+2kK_b\|x-\bar{x}(t)\|-\frac{\delta_1}{k}-2K_b\|x-\bar{x}\|.
\end{aligned}$$

Mas $-\dfrac{\delta_1}{k}+2kK_b\|x-\bar{x}\|<0$ para todo $x\in B(\bar{x}(t),\epsilon_1)$ e para quase todo $t\in[a,b]$, logo

$$\begin{aligned}
\sum_{\substack{j=1\\ j\neq i}}^{k}\big\|\Delta b_j(t,x,u_i(t))\big\| &< \delta_1-2K_b\|x-\bar{x}(t)\|\\
&< \Delta b_i(t,x,u_i(t)) \quad \text{por } (5.7),
\end{aligned}$$

para todo $x\in B(\bar{x}(t),\epsilon_1)$ e quase todo $t\in[a,b]$. Então, pelo Lema 5.1, obtemos que $B_1(t,x)$ é uma M-matriz para todo $x\in B(\bar{x}(t),\epsilon')$ e para quase todo $t\in[a,b]$.

Mostremos que a matriz $-B_2(t,x)$ é uma M-matriz. Das relações (5.11) e (5.12), obtemos

$$\begin{aligned}
\|\Delta b_j(t,x,v_i(t))\|-\frac{\delta_1}{k} &\leq 2K_b\|x-\bar{x}(t)\|\\
\sum_{\substack{j=1\\ j\neq i}}^{k}\|\Delta b_j(t,x,v_i(t))\| &\leq \sum_{\substack{j=1\\ j\neq i}}^{k}\left(\frac{\delta_1}{k}+2K_b\|x-\bar{x}(t)\|\right)\\
&= \frac{\delta_1(k-1)}{k}+2K_b\|x-\bar{x}(t)\|(k-1)\\
&< \delta_1-2K_b\|x-\bar{x}(t)\|\\
&\leq -\Delta b_i(t,x,v_i(t)) \text{ por } (5.10),
\end{aligned}$$

para todo $x\in B(\bar{x}(t),\epsilon_1)$ e quase todo $t\in[a,b]$. Então, pelo Lema 5.1 temos que $-B_2(t,x)$ é uma M-matriz para todo $x\in B(\bar{x}(t),\epsilon_1)$ e para quase todo $t\in[a,b]$. ■

Utilizando os Lemas 5.2 e 5.1, obtemos

$$B_1^{-1}(t,x)\geq 0 \quad \text{e} \quad B_2^{-1}(t,x)\leq 0,$$

para todo $x\in B(\bar{x}(t),\epsilon_1)$ e para quase todo $t\in[a,b]$.

Como os elementos das matrizes $B_1(t,\bar{x}(t))$ e $B_2(t,\bar{x}(t))$ são uniformemente limitados e os elementos de matrizes inversas são produtos e somas, segue que existe um escalar $m_b>0$, que depende de δ e m, tal que

$$\|B_1^{-1}(t,\bar{x}(t))\|_\infty\leq m_b \text{ e } \|B_2^{-1}(t,\bar{x}(t))\|_\infty\leq m_b. \tag{5.13}$$

Parte 2: Escolha uma coleção finita de funções de controle $\{w_i\}_{i=1}^M$ (também contendo o controle $\bar{u}(t)$). Seja $\beta(t) = (\beta_1(t), \beta_2(t), ..., \beta_M(t))$ qualquer função mensurável tomando valores do conjunto simplex S, onde

$$S = \left\{ (\beta_1(t), \beta_2(t), ..., \beta_M(t)) \in \mathbb{R}^M : \sum_{i=1}^{M} \beta_i \leq 1, \beta_i \geq 0, i = 1, 2, ..., M \right\}.$$

Defina as funções com valores vetoriais $\gamma^1, \gamma^2 : [a, b] \times \mathbb{R}^n \times S \to \mathbb{R}^k$ por

$$\begin{aligned} \gamma^1(t, x, \beta) &:= \max\left\{ -b(t, x, \bar{u}(t)) - \sum_{l=1}^{M} \beta_l \Delta b(t, x, w_l(t)), 0 \right\} \\ \gamma^2(t, x, \beta) &:= \min\left\{ -b(t, x, \bar{u}(t)) - \sum_{l=1}^{M} \beta_l \Delta b(t, x, w_l(t)), 0 \right\} \end{aligned}$$

onde o máximo e mínimo são tomados componente à componente. Observe que

(i) $\gamma^1(t, x, \beta) \geq 0$ e $\gamma^2(t, x, \beta) \leq 0$ para todo $(t, x, \beta) \in [a, b] \times \mathbb{R}^n \times S$,

(ii) as funções $\gamma^i(\cdot, x, \beta)$ são mensuráveis $\mathcal{L}$, $\gamma^i(t, \cdot, \cdot)$ são funções que satisfazem condição de Lipschitz em $B(\bar{x}(t), \epsilon') \times S$ para $i = 1, 2$, e

(iii) $\gamma^1(t, \bar{x}(t), 0) = \gamma^2(t, \bar{x}(t), 0) = 0$ para quase todo $t \in [a, b]$

Também, defina as funções com valores vetoriais $\alpha^1, \alpha^2 : [a, b] \times \mathbb{R}^n \times S \to \mathbb{R}^k$ por

$$\begin{aligned} \alpha^1(t, x, \beta) &:= B_1^{-1}(t, x)\gamma^1(t, x, \beta), \\ \alpha^2(t, x, \beta) &:= B_2^{-1}(t, x)\gamma^2(t, x, \beta). \end{aligned}$$

Como $B_1^{-1}(t, x) \geq 0$, $\gamma^1(t, x, \beta) \geq 0$, $B_2^{-1}(t, x) \leq 0$ e $\gamma^2(t, x, \beta) \leq 0$, obtemos

$$\begin{aligned} \alpha^1(t, x, \beta) &\geq 0; \\ \alpha^2(t, x, \beta) &\geq 0, \end{aligned}$$

para todo $B(\bar{x}(t), \epsilon')$, para todo $\beta \in S$ e para quase todo $t \in [a, b]$. Então, pelas definições das funções α^1, α^2, γ^1 e γ^2 obtemos

$$b(t, x, \bar{u}(t)) + B_1(t, x)\alpha^1(t, x, \beta) + B_2(t, x)\alpha^2(t, x, \beta) + \sum_{l=1}^{M} \beta_l \Delta b(t, x, w_l(t)) \equiv 0. \qquad (5.14)$$

Agora, defina as matrizes $n \times k$:

$$\begin{aligned} F_1(t, x) &:= \big(\Delta f(t, x, u_1(t)), ..., \Delta f(t, x, u_k(t))\big), \\ F_2(t, x) &:= \big(\Delta f(t, x, v_1(t)), ..., \Delta f(t, x, v_k(t))\big), \end{aligned}$$

onde

$$\Delta f(t, x, u_1(t) = f(t, x, u) - f(t, x, \bar{u}(t)).$$

Também defina a função

$$\phi(t, x, \beta) := f(t, x, \bar{u}(t)) + F_1(t, x)\alpha^1(t, x, \beta) + F_2(t, x)\alpha^2(t, x, \beta) + \sum_{l=1}^{M} \beta_l \Delta f(t, x, w_l(t)). \quad (5.15)$$

Segue das hipóteses H1 e H2, que a função $\phi(\cdot, x, \beta)$ é o mensurável no sentido de Lebesgue e a função $\phi(t, \cdot, \cdot)$ satisfaz a condição de Lipschitz na vizinhança de $(\bar{x}(t), 0)$.

Parte 3: Para algum $\eta > 0$, defina o conjunto $S_\eta := S \cap (\eta B)$ e considere o seguinte problema

$$(R) \begin{cases} \text{Minimize} \quad h(x(a), x(b)) \\ \text{sujeito a} \\ \qquad \dot{x}(t) \;=\; \phi(t, x(t), \beta(t)) \quad \text{q.t. } t \in [a, b] \\ \qquad \beta(t) \;\in\; S_\eta \quad \text{q.t. } t \in [a, b] \\ (x(a), x(b)) \;\in\; C \end{cases}$$

Observe que $(\bar{x}, \bar{\beta} \equiv 0)$ é um processo ótimo para o problema (R). De fato, suponha que existe um outro processo viável (x, β) de (R) com o custo menor, isto é,

$$h(x(a), x(b)) < h(\bar{x}(a), \bar{x}(b)).$$

Segue das relações (5.14), (5.15) e da hipótese H7, que

$$(\dot{x}, 0) \in \Big\{ \big(f(t, x, u), b(t, x, u)\big) : u \in U(t) \Big\}.$$

Utilizando o Teorema de Seleção Mensurável, podemos encontrar uma função de controle $\tilde{u}$ tal que

$$\dot{x} = f(t, x, \tilde{u}) \text{ e } 0 = b(t, x, \tilde{u}).$$

Logo, $(x, \tilde{u})$ é um processo viável de (Q), contradizendo a otimalidade de $(\bar{x}, \bar{u})$ para o problema (Q). Portanto, $(\bar{x}, \bar{\beta} \equiv 0)$ é uma solução para o problema (R).

Defina a função hamiltoniana para o problema (R)

$$\begin{aligned} H_R(t, x, p, \beta) &:= p(t) \cdot \phi(t, x, \beta) \\ &= p(t) \cdot f(t, x, \bar{u}) + p(t) \cdot F_1(t, x)\alpha^1(t, x, \beta) \\ &\quad + p(t) \cdot F_2(t, x)\alpha^2(t, x, \beta) + p(t) \cdot \sum_{l=1}^{m} \beta_l \Delta f(t, x, w_l(t)). \end{aligned}$$

Segue do Teorema 5.1, que existem $\lambda \geq 0$, $p(\cdot) \in AC([a,b];\mathbb{R}^n)$ e $\zeta(\cdot) \in L^1([a,b];\mathbb{R}^M)$ tais que

$$\lambda + \|p(\cdot)\|_{L^\infty} = 1, \tag{5.16}$$

$$(-\dot{p}(t), \dot{\bar{x}}(t), \zeta(t)) \in co\partial_{x,p,\beta} H_R(t, \bar{x}(t), p(t), 0) \quad \text{q.t. } t \in [a,b], \tag{5.17}$$

$$\zeta(t) \in \bar{co} N_{S_\eta}(0) \quad \text{q.t. } t \in [a,b], \tag{5.18}$$

$$(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial h(\bar{x}(a), \bar{x}(b)). \tag{5.19}$$

Aplicando a subdiferencial para função hamiltoniana H_R em relação a x e β no ponto $(t, \bar{x}, p, 0)$, obtemos

$$\begin{aligned}
\partial_{x,\beta} H_R(t, \bar{x}, p, 0) \;\subset\; & \big\{\partial_x p(t) \cdot f(t, \bar{x}(t), \bar{u}(t))\big\} \\
& \times\big\{p(t) \cdot \Delta f(t, \bar{x}(t), w_1(t)), ..., p(t) \cdot \Delta f(t, \bar{x}(t), w_M(t))\big\} \\
& +\big\{\Phi_1(t, \bar{x}(t)) \cdot \partial_x \gamma^1(t, \bar{x}(t), 0)\big\}\times\big\{\Phi_1(t, \bar{x}(t)) \cdot \partial_\beta \gamma^1(t, \bar{x}(t), 0)\big\} \\
& +\big\{\Phi_2(t, \bar{x}(t)) \cdot \partial_x \gamma^2(t, \bar{x}(t), 0)\big\} \times \big\{\Phi_2(t, \bar{x}(t)) \cdot \partial_\beta \gamma^2(t, \bar{x}(t), 0)\big\} \\
= \;& \big\{\partial_x p(t) \cdot f(t, \bar{x}(t), \bar{u}(t))\big\} \\
& \times\big\{p(t) \cdot \Delta f(t, \bar{x}(t), w_1(t)), ..., p(t) \cdot \Delta f(t, \bar{x}(t), w_M(t))\big\} \\
& +\Phi_1(t, \bar{x}(t)) \cdot \partial_{x,\beta} \gamma^1(t, \bar{x}(t), 0) + \Phi_2(t, \bar{x}(t)) \cdot \partial_{x,\beta} \gamma^2(t, \bar{x}(t), 0),
\end{aligned}$$

onde

$$\Phi_1(t, \bar{x}(t)) = (B_1^{-1})^T(t, \bar{x}(t)) \; F_1^T(t, \bar{x}(t)) \; p(t) \text{ e} \tag{5.20}$$

$$\Phi_2(t, \bar{x}(t)) = (B_2^{-1})^T(t, \bar{x}(t)) \; F_2^T(t, \bar{x}(t)) \; p(t). \tag{5.21}$$

Pela regra do máximo (Proposição 4.19), temos a seguinte estimativa da Jacobiana generalizada,

$$\begin{aligned}
\partial_{x,\beta} \gamma^1(t, x, \beta) \;&\subset\; co\Big\{\partial_{x,\beta}\Big(-b(t, x, \bar{u}(t)) - \sum_{l=1}^{M} \beta_l \Delta b(t, x, w_l(t))\Big), 0\Big\} \\
&\subset\; -co\Big(\nabla_x b(t, x, \bar{u}(t)), \Big(\Delta b(t, x, w_1(t)), ..., \Delta b(t, x, w_M(t))\Big)\Big).
\end{aligned}$$

Utilizando o teorema de seleção mensurável, existem $\mu_1(t), \mu_2(t), ..., \mu_k(t)$ funções de classe L^∞ tais que

$$\Big\{(\mu_1(t), \mu_2(t), ..., \mu_M(t)) \in \mathbb{R}^M : \sum_{i=1}^{M} \mu_i \leq 1, \mu_i \geq 0, i = 1, 2, ..., M\Big\}.$$

Logo

$$\begin{aligned}\partial_{x,\beta}\gamma^1(t,\bar{x},0) \quad \subset \quad & \Big\{-\mathcal{M}(t)\Big(\nabla_x b(t,\bar{x}(t),\bar{u}(t)),\Big(\Delta b(t,\bar{x},w_1(t)),...,\Delta b(t,\bar{x},w_M(t))\Big)\Big) : \\ & \mathcal{M}(t) = \mathrm{diag}(\mu_1(t),\mu_2(t),...,\mu_k(t)), \mu_i \in [0,1] \text{ mensurável}\Big\}.\end{aligned}$$

Também, temos

$$\begin{aligned}\partial_{x,\beta}\gamma^2(t,\bar{x},0) \quad \subset \quad & \Big\{-\mathcal{N}(t)\Big(\nabla_x b(t,\bar{x}(t),\bar{u}(t)),\Big(\Delta b(t,\bar{x},w_1(t)),...,\Delta b(t,\bar{x},w_M(t))\Big)\Big) : \\ & \mathcal{N}(t) = \mathrm{diag}(\nu_1(t),\nu_2(t),...,\nu_k(t)), \nu_i \in [0,1] \text{ mensurável}\Big\}.\end{aligned}$$

Segue de (5.17) que

$$\begin{aligned}(-\dot{p}(t),\zeta(t)) \quad \in \quad & \big\{co\partial_x p(t)\cdot f(t,\bar{x}(t),\bar{u}(t))\big\} \\ & \times\big\{p(t)\cdot\Delta f(t,\bar{x}(t),w_1(t)),...,p(t)\cdot\Delta f(t,\bar{x}(t),w_M(t))\big\} \\ & +\big\{q(t)\cdot\nabla_x b(t,\bar{x}(t),\bar{u}(t))\big\} \\ & \times\big\{q(t)\cdot\Delta b(t,\bar{x}(t),w_1(t)),...,q(t)\cdot\Delta b(t,\bar{x}(t),w_M(t))\big\}, \end{aligned} \tag{5.22}$$

para quase todo $t \in [a,b]$, onde

$$q(t) := -\mathcal{M}(t)\Phi_1(t,\bar{x}(t)) - \mathcal{N}(t)\Phi_2(t,\bar{x}(t)). \tag{5.23}$$

Utilizando as relações (5.13) e a hipótese H* podemos concluir que existe uma função integrável K_q tal que

$$\|q(t)\| \leq K_q(t)\|p(t)\|. \tag{5.24}$$

Como o conjunto S_η é convexo, temos que $N_{S_\eta}(0)$ é um conjunto convexo. Segue de (5.18), que

$$\zeta(t) \in N_{S_\eta}(0) = \big\{(\eta_1,\eta_2,...,\eta_M) \in \mathbb{R}^M : \eta_i \leq 0,\ i \in \{1,2,...,M\}\big\}.$$

Redefina a função hamiltoniana

$$H(t,x,p,q,u) := p\cdot f(t,x,u) + q\cdot b(t,x,u).$$

De (5.22) temos

$$-\dot{p}(t) \in co\partial_x H(t,\bar{x}(t),p(t),q(t),\bar{u}(t)) \tag{5.25}$$

e (como $\eta_i \leq 0$)

$$p\cdot\Delta f(t,\bar{x}(t),w_i) + q\cdot\Delta b(t,\bar{x}(t),w_i) \leq 0 \text{ para } i \in 1,2,...,M. \tag{5.26}$$

Segue das definições de Δf e Δb, que

$$p \cdot f(t, \bar{x}(t), w_i) + q \cdot b(t, \bar{x}(t), w_i) \leq p \cdot f(t, \bar{x}(t), \bar{u}(t)) + q \cdot b(t, \bar{x}(t), \bar{u}(t)). \tag{5.27}$$

Isto implica que

$$\operatorname*{Max}_{u \in \hat{U}(t)} H(t, \bar{x}(t), p(t), q(t), u) = H(t, \bar{x}(t), p(t), q(t), \bar{u}(t)) \text{ q.t. } t \in [a, b] \tag{5.28}$$

onde

$$\hat{U}(t) = \bigcup_{i}^{M} \{w_i(t)\}.$$

Parte 4: Agora, vamos mostrar que a relação (5.28) será satisfeita quando substituímos $\hat{U}(t)$ por $U(t)$. Para fazer esta subtituição, usaremos o raciocínio análogo da prova do Teorema 5.1.1 em [9]. Por conveniência, toma uma função $\tilde{q}(\cdot)$ dado por

$$\tilde{q}(t) = K_q(t)^{-1} q(t). \tag{5.29}$$

Das equações (5.16) e (5.25), e das hipóteses H2 e H3, temos que $\|p(\cdot)\|_{L^\infty}$ é limitado por uma constante (não depende da escolha de w_i). Segue de (5.24) que $\|\tilde{q}(\cdot)\|_{L^\infty}$ é limitado por uma constante (novamente independe da escolha de w_i). Substituindo $q(t)$ por $c_f(t)\tilde{q}(t)$ em (5.27) obtemos

$$\big(p(t), \tilde{q}(t)\big) \cdot \big(f(t, \bar{x}(t), u), c_f(t) b(t, \bar{x}(t), u)\big) \leq \big(p(t), \tilde{q}(t)\big) \cdot \big(f(t, \bar{x}(t), \bar{u}(t)), c_f(t) b(t, \bar{x}(t), \bar{u}(t))\big).$$

Defina as funções

$$r(t) := \big(p(t), \tilde{q}(t)\big) \text{ e } \beta(t, x, u) = \big(f(t, x, u), c_f(t) b(t, x, u)\big).$$

Observe que $\|r(\cdot)\|_{L^\infty} \leq L$, onde L é uma constante (independente de w_i). Seja $\{Z_j\}_{j=1}^{\infty}$ uma família crescente de subconjunto finito LB tal que

$$Z_j + j^{-1} B \supset LB \text{ para cada } j.$$

Fixe j e considere $z \in Z_j$. Então, pelo Teorema da Seleção apropriado (veja Rieder [27, Teorema 3.3]), podemos encontrar uma função mensurável $u_z(\cdot), u_z(t) \in U(t)$ para quase todo $t \in [a, b]$ tal que

$$\sup_{u \in U(t)} \big\{z \cdot \beta(t, \bar{x}(t), u)\big\} \leq z \cdot \beta(t, \bar{x}(t), u_z(t)) + j^{-1} \text{ q.t. } t \in [a, b]. \tag{5.30}$$

Defina a multifunção mensurável por $U_j(t) := \{u_z(t) : z \in Z_j\}$ para cada j e $t \in [a,b]$. Considere o problema

$$\begin{cases} \text{Minimize} \quad h(x(a),x(b)) \\ \text{sujeito a} \\ \qquad \dot{x}(t) = \beta(t,x(t),u(t)) \quad \text{q.t. } t \in [a,b] \\ \qquad u(t) \in U_j(t) \quad \text{q.t. } t \in [a,b] \\ (x(a),x(b)) \in C \end{cases}$$

Fazendo uma análise análoga da parte 3, podemos concluir que existem $p_j(\cdot) \in AC$, $\tilde{q}_j(\cdot) \in L^\infty$ e $\lambda_j \geq 0$ tais que

$$\lambda_j + \|p_j(\cdot)\|_{L^\infty} = 1, \tag{5.31}$$

$$-\dot{p}_j(t) \in co\partial_x p_j \cdot f + \{c_f \tilde{q}_j \cdot \nabla_x b\} \text{ q.t. } t \in [a,b], \tag{5.32}$$

$$(p_j(t), \tilde{q}_j(t)) \cdot \beta(t,\bar{x}(t),u) \leq (p_j(t),\tilde{q}_j(t)) \cdot \beta(t,\bar{x}(t),\bar{u}) \ \forall u \in U_j(t) \text{ q.t. } t \in [a,b], \tag{5.33}$$

$$(p_j(a), -p_j(b)) \in N_C(\bar{x}(a),\bar{x}(b)) + \lambda_j \partial h(\bar{x}(a),\bar{x}(b)). \tag{5.34}$$

Observe que as seqüencias $\{\|p_j(\cdot)\|_{L^\infty}\}$, $\{\|\tilde{q}_j(\cdot)\|_{L^\infty}\}$ e $\{\lambda_j\}$ são uniformemente limitadas e $\{\dot{p}_j\}$ é uma seqüencia uniformemente limitada em L^1. Logo, obtemos as seguintes subseqüências convergentes (não reindexamos)

$$p_j(\cdot) \longrightarrow p(\cdot) \text{ uniformemente}, \quad \tilde{q}_j(\cdot) \longrightarrow \tilde{q}(\cdot) \text{ fracamente}^* \text{ em } L^\infty,$$

para algum $p \in AC$, $\tilde{q} \in L^\infty$, e $\{\dot{p}_j\}$ tem um limite fraco em L^1. Passando o limite quando $j \longrightarrow \infty$ em (5.32) e fazendo uma simples mudança[1] na demonstração de Clarke ([9, Teorema 3.1.7]), obtemos

$$-\dot{p}(t) \in co\partial_x p(t) \cdot f(t,\bar{x},\bar{u}(t)) + q(t) \cdot \nabla_x b(t,\bar{x},\bar{u}(t)),$$

para quase todo $t \in [a,b]$ e onde $q(t) := K_q(t)\tilde{q}(t)$. Também passando o limite em (5.34) e (5.31) quando $j \longrightarrow \infty$, obtemos

$$(p(a), -p(b)) \in N_C(\bar{x}(a),\bar{x}(b)) + \lambda \partial h(\bar{x}(a),\bar{x}(b)) \text{ (pela Proposição 3.2)},$$

e

$$\lambda + \|p(\cdot)\|_{L^\infty} = 1.$$

[1] Considera a seqüencia convergente $c_f \tilde{q}_j \longrightarrow c_f \tilde{q}$ na demonstração de Clarke.

Falta mostrar a condição da "maximização da Hamiltoniana". Seja $\tilde{u}(\cdot)$ uma função de controle arbitrária. Fixa j. Então, para quase todo $t \in [a,b]$, podemos encolher $z \in Z_j$ tal que $\|(p_j(t), \tilde{q}_j(t)) - z\| \leq j^{-1}$. Segue, para quase todo $t \in [a,b]$, que

$$\begin{aligned}
\big(p_j(t), \tilde{q}_j(t)\big) \cdot \beta(t, \bar{x}(t), \tilde{u}(t)) &= \big((p_j(t), \tilde{q}_j(t)) - z + z\big) \cdot \beta(t, \bar{x}(t), \tilde{u}(t)) \\
&= \big((p_j(t), \tilde{q}_j(t)) - z\big) \cdot \beta(t, \bar{x}(t), \tilde{u}(t)) + z \cdot \beta(t, \tilde{u}(t)) \\
&\leq z \cdot \beta(t, \bar{x}(t), \tilde{u}(t)) + j^{-1} \cdot \|\beta(t, \bar{x}(t), \tilde{u}(t))\| \\
&= z \cdot \beta(t, \bar{x}(t), \tilde{u}(t)) + j^{-1} \cdot \|\big(f(t, \bar{x}(t), \tilde{u}(t)), c_f(t) b(t, \bar{x}(t), \tilde{u}(t))\big)\| \\
&\leq \big(p_j(t), \tilde{q}_j(t)\big) \cdot \beta(t, \bar{x}(t), u_z(t)) + j^{-1}\big(1 + 2[c_j(t) + c_b c_f(t)]\big) \\
&\leq \big(p_j(t), \tilde{q}_j(t)\big) \cdot \beta(t, \bar{x}(t), \bar{u}(t)) + j^{-1}\big(1 + 2[c_j(t) + c_b c_f(t)]\big).
\end{aligned}$$

A penúltima desigualdade é conseqüência da hipótese H* e do uso de (5.30) e a última da equação (5.33). Usando a notação $f(\bar{u})$ para $f(t, \bar{x}(t), \bar{u}(t))$ e $b(\bar{u})$ para $b(t, \bar{x}(t), \bar{u}(t))$, e manipulando algebricamente a última desigualdade, podemos reescrevê-la da seguinte forma

$$p_j(t)\big[f(\bar{u}) - f(\tilde{u})\big] + \tilde{q}_j(t) c_f(t) [b(\bar{u}) - b(\tilde{u})] \;\geq\; -j^{-1}\big(1 + 2[c_j(t) + c_b c_f(t)]\big).$$

Isto implica que

$$\int_a^b \Big\{p_j(t)\big[f(\bar{u}) - f(\tilde{u})\big] + \tilde{q}_j(t) c_f(t)[b(\bar{u}) - b(\tilde{u})]\Big\} dt \geq -\int_a^b j^{-1}\big(1 + 2[c_j(t) + c_b c_f(t)]\big) dt.$$

Como $c_f(t) \in L^1$, temos

$$\int_a^b \Big\{p_j(t)\big[f(\bar{u}) - f(\tilde{u})\big] + \tilde{q}_j(t) c_f(t)[b(\bar{u}) - b(\tilde{u})]\Big\} dt \geq -j^{-1}\Big[\|b - a\| + 2\|c_f(t)\|_{L^1} + 2c_b\|c_f(t)\|_{L^1}\Big].$$

Passando o limite quando $j \longrightarrow \infty$, segue que

$$\int_a^b \Big\{\big[p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) + q(t) \cdot b(t, \bar{x}(t), \bar{u}(t))\big] - \big[p(t) \cdot f(t, \bar{x}(t), \tilde{u}(t)) + q(t) \cdot b(t, \bar{x}(t), \tilde{u}(t))\big]\Big\} dt \geq 0. \tag{5.35}$$

Esta desigualdade é valida para qualquer função de controle $\tilde{u}(t)$ e é uma forma "integral" de maximização da Hamiltoniana.

Parte 5: Sob a hipótese H*, pudemos mostrar um caso especial do princípio do máximo. Agora, supomos que a hipótese H* não esteja presente.

Começamos pela elaboração de um novo problema em que a hipótese H* é satisfeita, porém envolve-se restrições no conjunto $U(\cdot)$ e convexidade no conjunto velocidade resultante, para garantir as hipóteses H1-H7. Defina

$$\tilde{U}(t) := \Big\{u \in U(t) : \|f(t,\bar{x}(t),u) - f(t,\bar{x}(t),\bar{u}(t))\| \leq 1 \text{ e } \|b(t,\bar{x}(t),u)\| \leq 1\Big\}$$

e considere uma perturbação para o problema (Q) da seguinte forma

$$(Q')\left\{\begin{array}{l} \text{Minimize} \quad h(x(a),x(b)) \\ \text{sujeito a} \\ \qquad \begin{array}{rcl} \dot{x}(t) & = & \displaystyle\sum_{i=0}^{n+k} \alpha_i f(t,x,v_i)) \\ 0 & = & \displaystyle\sum_{i=0}^{n+k} \alpha_i b(t,x,v_i) \\ \alpha(t) & \in & P^{n+k} \\ v(t) & \in & \tilde{U}(t) \times ... \times \tilde{U}(t) \\ (x(a),x(b)) & \in & C \times C \end{array} \end{array}\right.$$

onde $\alpha = (\alpha_0, ..., \alpha_{n+k})$ e $v = (v_0, ..., v_{n+k})$ são consideradas as variáveis de controle e

$$P^{n+k} := \Big\{(\alpha'_0, ..., \alpha'_{n+k}) : \sum_{i=0}^{n+k} \alpha'_i = 1, \alpha'_i \geq 0, \text{para todo } i\Big\}.$$

Observe que o problema (Q') satisfaz as hipóteses H1-H7 e também (H*). Além disso

$$(\bar{x}, \alpha(t), v(t)) \equiv \big(\bar{x}, (1,0,...,0), (\bar{u}(t), ..., \bar{u}(t))\big)$$

é um minimizador do problema Q'. Utilizando a análise anterior, existem multiplicadores $p \in AC$, $q \in L^1$ e $\lambda \geq 0$, que satisfazem as condições (5.31), (5.32) e (5.34). Também satisfaz a condição (5.35) para cada $\tilde{u}$ em $\tilde{U}(t)$. Precisamos mostrar que desigualdade (5.35) implica que

$$\max_{u \in U(t)} \big\{p(t) \cdot f(t,\bar{x}(t),u) + q(t) \cdot b(t,\bar{x}(t),u)\big\} = p(t) \cdot f(t,\bar{x}(t),\bar{u}(t)) + q(t) \cdot b(t,\bar{x}(t),\bar{u}(t)), \quad (5.36)$$

para quase todo $t \in [a,b]$. Suponha que a condição (5.36) não seja satisfeita. Então podemos encontrar um controle $u'(\cdot) \in \tilde{U}(\cdot)$ tal que

$$p(t) \cdot \big[f(t,\bar{x}(t),u'(t)) - f(t,\bar{x}(t),\bar{u}(t))\big] + q(t) \cdot \big[b(t,\bar{x}(t),u'(t)) - b(t,\bar{x}(t),\bar{u}(t))\big] \geq 0 \quad (5.37)$$

para quase todo $t \in [a,b]$. Observe que a desigualdade acima é estrita em um subconjunto de medida positiva. Defina a função mensurável $\gamma : [a,b] \to [0,1)$ por

$$\gamma(t) := \big(1 + \|f(t,\bar{x}(t),u'(t)) - f(t,\bar{x}(t),\bar{u}(t))\| + \|b(t,\bar{x}(t),u'(t)) - b(t,\bar{x}(t),\bar{u}(t))\|\big)^{-1}.$$

Segue de (5.37) que

$$\gamma(t)\Big[p(t)\cdot\big(f(t,\bar{x}(t),u'(t))-f(t,\bar{x}(t),\bar{u}(t))\big)+q(t)\cdot\big(b(t,\bar{x}(t),u'(t))-b(t,\bar{x}(t),\bar{u}(t))\big)\Big]\geq 0,$$

para quase todo $t\in[a,b]$. Somando e subtraindo $p(t)\cdot f(t,\bar{x}(t),\bar{u}(t))+q(t)\cdot b(t,\bar{x}(t),\bar{u}(t))$, obtemos

$$\begin{aligned} p(t)\ &\cdot\big[f(t,\bar{x}(t),\bar{u}(t))+\gamma(t)\big(f(t,\bar{x}(t),u'(t))-f(t,\bar{x}(t),\bar{u}(t))\big)\big]\\ &+q(t)\cdot\big[b(t,\bar{x}(t),\bar{u}(t))+\gamma(t)\big(b(t,\bar{x}(t),u'(t))-b(t,\bar{x}(t),\bar{u}(t))\big)\big]\\ &-\big[p(t)\cdot f(t,\bar{x}(t),\bar{u}(t))+q(t)\cdot b(t,\bar{x}(t),\bar{u}(t))\big]\geq 0, \end{aligned}$$

e novamente esta desigualdade é estrita em um subconjunto de medida positiva. Utilizando a hipótese da convexidade e o Teorema de Seleção Mensurável, podemos encontrar uma função de controle $\tilde{u}(\cdot)\in U(\cdot)$ tal que

$$\begin{aligned} f(t,\bar{x}(t),\tilde{u}(t)) &= f(t,\bar{x}(t),\bar{u}(t))+\gamma(t)\big(f(t,\bar{x}(t),u'(t))-f(t,\bar{x}(t),\bar{u}(t))\big)\ \text{q.t. } t\in[a,b]\ ,\\ b(t,\bar{x}(t),\tilde{u}(t)) &= b(t,\bar{x}(t),\bar{u}(t))+\gamma(t)\big(b(t,\bar{x}(t),u'(t))-b(t,\bar{x}(t),\bar{u}(t))\big)\ \text{q.t. } t\in[a,b]\ . \end{aligned}$$

Logo,

$$\int_a^b\big([p(t)\cdot f(t,\bar{x}(t),\tilde{u}(t))+q(t)\cdot b(t,\bar{x}(t),\tilde{u}(t))]-[p(t)\cdot f(t,\bar{x}(t),\bar{u}(t))+q(t)\cdot b(t,\bar{x}(t),\bar{u}(t))]\big)dt>0.$$

Mas, para nossa escolha de γ, $\tilde{u}$ é realmente um membro de $\tilde{U}(t)$, violando (5.35). Portanto a equação (5.36) é verdadeira. ■

5.2.2 Problema com Restrições de Igualdade e Desigualdade

Considere agora o problema do tipo

$$(R)\begin{cases} \text{Minimize}\quad h(x(a),x(b)) \\ \text{Sujeito a} \\ \qquad \begin{aligned} \dot{x}(t) &= f(t,x(t),u(t)) && \text{q.t. } t\in[a,b]\\ 0 &= b(t,x(t),u(t)) && \text{q.t. } t\in[a,b]\\ 0 &\geq g(t,x(t),u(t)) && \text{q.t. } t\in[a,b]\\ u(t) &\in U(t) && \text{q.t. } t\in[a,b]\\ (x(a),x(b)) &\in C \end{aligned} \end{cases}$$

sendo dadas as funções $h:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$, $f:[a,b]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, $b:[a,b]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^k$ e $g:[a,b]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^l$, a multifunção $U:[a,b]\rightrightarrows\mathbb{R}^m$ e o conjunto $C\subset\mathbb{R}^n\times\mathbb{R}^n$.

Seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ uma solução ótima do problema (R) e que as hipóteses H1-H3 e H5-H7 são satisfeitas, quando substituímos a função b por $\tilde{b} : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^s$ $(s = k + l)$ definida por

$$\tilde{b}(t, x(t), u(t)) := \big(b(t, x(t), u(t)), g(t, x(t), u(t))\big).$$

Para demonstrar o princípio do máximo para o problema (R) precisamos adicionar a seguinte hipótese ao problema:

H4'. Existe um escalar $\delta > 0$ tal que

$$\delta B \subset \Big\{\big(b(t, \bar{x}(t), U(t)), g(t, \bar{x}(t), U(t)) + v\big) : v \in \mathbb{R}^l, v \geq 0\Big\},$$

para quase todo $t \in [a, b]$.

Defina a função Hamiltoniana

$$H_R(t, x(t), p(t), q(t), r(t), u(t)) := p(t) \cdot f(t, x(t), u(t)) + q(t) \cdot b(t, x(t), u(t)) + r(t) \cdot g(t, x(t), u(t)).$$

Assim, podemos apresentar as condições necessárias em forma de Princípio do Máximo para o problema (R).

Corolário 5.1. *Seja $(\bar{x}(\cdot), \bar{u}(\cdot))$ um processo minimizador para o problema (R). Suponha que as hipóteses $H1 - H3$, $H5 - H7$ (onde a função b é substituído por $\tilde{b}$) e $H4'$ são todas satisfeitas. Então, existem funções $p \in AC([a, b], \mathbb{R}^k)$, $q \in L^1([a, b]; \mathbb{R}^k)$ e $r \in L^1([a, b]; \mathbb{R}^l)$ e um escalar $\lambda \geq 0$ tais que*

(i) $\lambda + \|p\|_{L^\infty} = 1,$

(ii) $-\dot{p}(t) \in co\partial_x H_R(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t))$ *q.t.* $t \in [a, b],$

(iii) $H_R(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t)) = \max\limits_{u \in U(t)} \big\{H_R(t, \bar{x}(t), p(t), q(t), r(t), u)\big\}$ *q.t.* $t \in [a, b],$

(iv) $(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial h(\bar{x}(a), \bar{u}(b)),$

(v) $r(t) \leq 0$, $r(t) \cdot g(t, \bar{x}(t), \bar{u}(t)) = 0$ *q.t.* $t \in [a, b].$

Demonstração: Vamos incluir uma nova função de controle v (variável de folga) e reescrever o problema (R) da seguinte forma

$$(R')\begin{cases} \text{Minimize} \quad h(x(a), x(b)) \\ \text{sujeito a} \\ \qquad \dot{x}(t) = f(t, x(t), u(t)) & \text{q.t. } t \in [a, b] \\ \qquad 0 = b(t, x(t), u(t)) & \text{q.t. } t \in [a, b] \\ \qquad 0 = g(t, x(t), u(t)) + v(t) & \text{q.t. } t \in [a, b] \\ \qquad (u(t), v(t)) \in U(t) \times \{v \in \mathbb{R}^l : v \geq 0\} & \text{q.t. } t \in [a, b] \\ \qquad (x(a), x(b)) \in C. \end{cases}$$

Observe que o problema (R') está na mesma forma do problema (Q), com a nova variável de controle $w = (u, v)$ e com o conjunto de controle $W(t) = U(t) \times \{v \in \mathbb{R}^l : v \geq 0\}$. Logo, $(\bar{x}, \bar{u}, \bar{v})$ é um processo viável para (R'), onde

$$\bar{v}(t) = -g(t, \bar{x}, \bar{u}).$$

Assim, $(\bar{x}, \bar{u}, \bar{v})$ é um processo ótimo para (R').

Também, podemos observar que a hipótese $H4'$ implica

$$B_\delta(0, 0) \subset \hat{b}(t, \bar{x}(t), W(t)) \text{ para quase todo } t \in [a, b],$$

onde $\hat{b}(t, x, w) = (b(t, x, u), g(t, x, u) + v)$.

Aplicando o Teorema 5.2 a (R'), obtemos os multiplicadores p, q, r e $\lambda \geq 0$ que satisfazem os ítens (i), (ii) e (iv) do corolário e

$$p \cdot f(t, \bar{x}, \bar{u}) = \max_{(u,v) \in W(t)} \big\{p \cdot f(t, \bar{x}, u) + q \cdot b(t, \bar{x}, u) + r \cdot g(t, \bar{x}, u) + r \cdot v\big\} \tag{5.38}$$

para quase todo $t \in [a, b]$. Fixe $v = \bar{v}$. Segue de (5.38), que

$$p \cdot f(t, \bar{x}, \bar{u}) + r \cdot g(t, \bar{x}(t), \bar{u}(t)) = \max_{u \in U(t)} \big\{p \cdot f(t, \bar{x}, u) + q \cdot b(t, \bar{x}, u) + r \cdot g(t, \bar{x}, u)\big\} \tag{5.39}$$

para quase todo $t \in [a, b]$, isto é, satisfaz o item (iii).

Resta mostrar o item (v). Fixe $u = \bar{u}$. Segue de (5.38), que

$$p \cdot f(t, \bar{x}, \bar{u}) = \max_{v \geq 0} \big\{p \cdot f(t, \bar{x}, \bar{u}) + q \cdot b(t, \bar{x}, \bar{u}) + r \cdot g(t, \bar{x}, u) + r \cdot v\big\}.$$

Isto implica que

$$0 = \max_{v \geq 0} \big\{r \cdot (g(t, \bar{x}, \bar{u}) + v)\big\}.$$

Equivalentemente temos

$$-r \cdot g(t, \bar{x}, \bar{u}) = \max_{v \geq 0} r \cdot v.$$

Note que $r \leq 0$. De fato, suponha que $r > 0$. Então $-r \cdot g(t, \bar{x}(t), \bar{u}(t)) < \infty$. Mas, por outro lado, temos $\max_{v \geq 0} r \cdot v \longrightarrow \infty$, o que é um absurdo.

Por outro lado, se $g(t, \bar{x}(t), \bar{u}(t)) < 0$, então $r = 0$. De fato, suponha que $r < 0$. Isto implica que $-r \cdot g(t, \bar{x}(t), \bar{u}(t)) < 0$ e $\max_{v \geq 0} r \cdot v = 0$, o que é um absurdo.

Resumindo, temos a condição conhecida como condição de folgas complementares

$$r(t) \leq 0, \quad r(t) \cdot g(t, \bar{x}(t), \bar{u}(t)) = 0 \;\; \text{para quase todo } t \in [a, b],$$

completando a demonstração. ∎

Capítulo 6

Considerações Finais

O objetivo principal desta dissertação foi discutir condições necessárias de otimalidade para o problema de controle ótimo com restrições mistas envolvendo hipóteses mínimas de diferenciabilidade das funções do problema. Assim, após um estudo da literatura vigente, verificamos que se pode tratar o problema com funções não-diferenciáveis no custo e na equação diferencial, entretanto, a restrição mista no estado e no controle deve ser de classe C^1.

Para simplificar a apresentação, no estudo das condições necessárias de otimilidade, optamos por utilizar hipótese de interioridade H4. Em [11] mostra-se que se $\det[\Upsilon(t)\Upsilon^T(t)] \geq K$, para algum $K > 0$ onde $\Upsilon(t) = \nabla_u b(t, \bar{x}(t), \bar{u}(t))$, então, a hipótese de interioridade é satisfeita.

Para estudar o problema sem diferenciabilidade, propusemo-nos a apresentar nesta dissertação um estudo da teoria de análise não-diferenciável, que ocupou a sua maior parte. Este estudo sistematizado, com certeza, deverá servir de base para introduzir novos estudantes na área.

No Exemplo 2.1, observa-se que o princípio do máximo (Teorema 5.2) não pode ser aplicado, pois não satisfaz a condição de interioridade (H4). Também, existem classes de problemas em que a condição de convexidade no conjunto velocidade (H7) não é satisfeita (veja o Exemplo 3 em [16]). Para resolver estes tipos de problemas, encontramos na literatura as condições necessárias de otimalidade na forma fraca do princípio do máximo, que é aplicado ao minimizador fraco local (veja [13]).

Como trabalho futuro, um grande desafio é eliminar a hipótese de diferenciabilidade da função de restrição conjunta na variável de estado e no controle.

Apêndice A

Conceitos Preliminares

Neste apêndice apresentamos alguns resultados básicos e importantes, que serão necessários para o desenvolvimento e para o entendimento dos capítulos desta dissertação. Em cada seção, serão apresentadas as definições consideradas essenciais e os enunciados dos teoremas sem demonstração, que podem ser facilmente encontrados na literatura.

A.1 Análise Convexa

Apresentamos, nesta seção, as noções básicas de conjuntos convexos, funções convexas e suas propriedades. A maioria dos resultados aqui apresentados pode ser encontrada em Rockafellar [28].

Conjuntos Convexos

Definição A.1. *Um subconjunto $C \subset \mathbb{R}^n$ é chamado convexo se*

$$\alpha x + (1-\alpha)y \in C,$$

para qualquer $x, y \in C$ e $\alpha \in [0,1]$ (ou equivalente para $\alpha \in (0,1)$).

Para dar uma noção geométrica de conjunto convexo precisamos primeiro definir o conceito de segmento linear.

Definição A.2. *Para $x, y \in \mathbb{R}^n$, o segmento linear $[x,y]$ (ou (x,y)) é definido por $[x,y] := \{\alpha x + (1-\alpha)y : 0 \leq \alpha \leq 1\}$ (ou, $(x,y) := \{\alpha x + (1-\alpha)y : 0 < \alpha < 1\}$), ou seja, os pontos x e y são conectados por um segmento de reta.*

Definição A.3 (Definição Geométrica de um Conjunto Convexo)**.** *Um conjunto C é chamado convexo, se o segmento linear $[x,y] \subset C$ (ou $(x,y) \subset C$) quando $x, y \in C$.*

Conjuntos convexos possuem as seguintes propriedades:

Proposição A.1.

a) Seja uma família arbitrária $\{C_j\}_{j \in \mathcal{J}}$, $C_j \subset \mathbb{R}^n$ de conjuntos convexos. Então, $C := \bigcap_{j \in \mathcal{J}} C_j$ é convexo.

b) Sejam $C_1, ..., C_k$, com $C_i \subset \mathbb{R}^{n_i}$. Então, $C_1, ..., C_k$ são conjuntos convexos se, somente se, $C_1 \times ... \times C_k$ é um conjunto convexo em $\mathbb{R}^{n_1} \times ... \times \mathbb{R}^{n_k}$.

c) Se $C \subset \mathbb{R}^n$ é convexo, então, o interior e o fecho de C, denotados por int C e cl C respectivamente, são convexos.

Em análise convexa, temos as definições de envoltório convexo e o fecho do envoltório convexo:

Definição A.4. *Seja $S \subset \mathbb{R}^n$ um subconjunto não-vazio. O* ***envoltório convexo*** *de S, denotado por co S, é definido por*

$$co\ S := \bigcap \{C : S \subset C,\ C \text{ é convexo}\}.$$

Teorema A.1. *O envoltório convexo de um subconjunto finito $\{x_1, ..., x_k\}$ de $\mathbb{R}^n$ consiste de todos os vetores da forma $\sum_{i=1}^{k} \alpha_i x_i$ com $\alpha_1 \geq 0, ..., \alpha_k \geq 0$ e $\sum_{i=1}^{k} \alpha_i = 1$.*

Definição A.5. *O* ***fecho do envoltório convexo*** *de $S \subset \mathbb{R}^n$, não-vazio, denotado por $\overline{co}$ S, é definido como a intersecção de todos conjuntos convexos fechados contendo S.*

Temos as seguintes propriedades:

Teorema A.2 (Teorema de Caratheodory). *Seja um subconjunto $S \subset \mathbb{R}^n$. Qualquer $x \in co\ S$ pode ser representado como uma combinação convexa de $n+1$ elementos de S.*

Proposição A.2. *Seja um subconjunto $S \subset \mathbb{R}^n$ não-vazio. Então $\overline{co}\ S = cl(coS)$.*

Teorema A.3. *Se um conjunto $S \subset \mathbb{R}^n$ é limitado (respectivamente compacto), então co S é limitado (respectivamente compacto).*

Um resultado importante dos conjuntos convexos é o Teorema da separação.

Teorema A.4 (Teorema da separação). *Sejam C_1 e C_2 conjuntos convexos e não-vazios em $\mathbb{R}^n$. Existe um hiperplano apropriado separando C_1 e C_2 se, somente se, existir um vetor s tal que*

$$\sup_{y_1 \in C_1} \big\langle s, y_1 \big\rangle \leq \inf_{y_2 \in C_2} \big\langle s, y_2 \big\rangle$$

e

$$\inf_{y_1 \in C_1} \big\langle s, y_1 \big\rangle < \sup_{y_2 \in C_2} \big\langle s, y_2 \big\rangle.$$

Cone Convexo e Cone Polar

Com análise convexa, temos as definições de cone convexo e cone polar:

Definição A.6. *Um conjunto $k \in \mathbb{R}^n$ é chamado de **cone** se $\lambda\, x \in k$, para qualquer $x \in k$ e para qualquer $\lambda > 0$. Um cone k é chamado de **cone convexo** se k é convexo.*

Definição A.7. *O **cone polar** de k, denotado por k° é definido pelo conjunto*

$$k^\circ := \Big\{ s \in \mathbb{R}^n : \big\langle s, x \big\rangle \leq 0 \text{ para todo } x \in k \Big\}.$$

Funções Convexas

Temos as seguintes definições:

Definição A.8. *Considere uma função $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Então:*

(i) f é chamado convexo se, e somente se,

$$f\big(\alpha x + (1-\alpha)y\big) \leq \alpha f(x) + (1-\alpha) f(y),$$

para todo $x, y \in \mathbb{R}^n$ e $\alpha \in (0,1)$.

*(ii) O **domínio** de f é o conjunto*

$$dom(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}.$$

*(iii) O **epígrafo** da função f é o conjunto*

$$epi(f) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : \alpha \geq f(x)\}.$$

*(iv) Dizemos que a função convexa f é **próprio**, se $f(x) < +\infty$ para algum x e $f(x) > -\infty$ para cada $x \in \mathbb{R}^n$.*

Com as definições acima, temos as seguintes propriedades:

Proposição A.3. *Seja uma função $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Temos as seguintes afirmações:*

(i) f é convexa se, e somente se, epi(f) é convexo.

(ii) se f é convexa então dom(f) é convexo.

Para funções diferenciáveis, existe uma caracterização alternativa de convexidade dada pela seguinte proposição:

Proposição A.4. *Sejam $C \subset \mathbb{R}^n$ um conjunto convexo e $f : \mathbb{R}^n \to \mathbb{R}$ é diferenciável sobre C. Então,*

i) a função f é convexa sobre C se, e somente se,

$$f(z) \geq f(x) + \big\langle \nabla f(x), (z - x) \big\rangle \ \forall x, z \in C. \tag{A.1}$$

ii) se a desigualdade (A.1) é estrita para $x \neq z$, então f é estritamente convexa.

Funções Convexas Fechadas

Definição A.9. *Um função $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é semicontínua inferior em $\mathbb{R}^n$ se para cada $x \in \mathbb{R}^n$, temos*

$$\liminf_{y \to x} f(y) \geq f(x).$$

A importância da função semicontínua inferior no estudo de função convexa é o aparecimento do seguinte resultado:

Proposição A.5. *Seja a função $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Então as seguintes afirmações são equivalentes:*

i) a função f é semicontínua inferior em todo $\mathbb{R}^n$;

ii) o conjunto $\{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ é um conjunto fechado para cada $\alpha \in \mathbb{R}$;

iii) o conjunto epi$f$ é um conjunto fechado em $\mathbb{R}^{n+1}$.

Definição A.10. *Dizemos que a função convexa $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é fechada se seu conjunto epígrafo for fechado.*

Observe que a função semicontínua inferior para todo $x \in dom(f)$ é uma função convexa fechada.

A.2 Noções Básicas de Teoria da Medida

Nesta seção, apresentamos algumas definições e resultados básicos da Teoria da Medida. As referências básicas podem ser encontradas em Rudin [30, 31], Fernandez [18], Lieb & Loss [23] ou Bartle [2].

Sigma-Álgebra

Definição A.11 (Sigma-Álgebra). *Uma família $\mathcal{S}$ de subconjuntos de X é uma Sigma-Álgebra se*

(i) $\emptyset, X \in \mathcal{S}$*;*

(ii) se $A \in \mathcal{S}$*, então o complemento de* A*, denotado por* A^C*, pertence a* $\mathcal{S}$*;*

(iii) se $\{A_n\}_{n=1}^{\infty}$ *é uma seqüencia de conjuntos em* S*, então a união* $\displaystyle\bigcup_{n=1}^{\infty} A_n$ *pertence à* $\mathcal{S}$*.*

Definição A.12 (Sigma-álgebra de Borel-$\mathcal{B}$). *A Sigma- álgebra de Borel é sigma-álgebra gerada por subconjuntos abertos de* $\mathbb{R}^n$*. Também é gerada por esferas abertas de* $\mathbb{R}^n$*,* $(B_{x,R} = \{y \in \mathbb{R}^n : \|x - y\| < R\})$*.*

Definição A.13 (Conjunto Mensurável). *O par* $(X, \mathcal{S})$*, sendo* X *um conjunto não-vazio e* $\mathcal{S}$ *uma sigma-álgebra de* X*, é denominado de espaço mensurável. Qualquer conjunto em* $\mathcal{S}$ *é chamado de conjunto* $\mathcal{S}$*-mensurável, mas quando a sigma-álgebra é fixa (como geralmente é o caso), o conjunto será chamado de mensurável.*

Observação A.1. *Seja* $X = \mathbb{R}$*. A* ***Sigma-álgebra de Borel*** *é a sigma-álgebra gerada por todos os intervalos abertos* (a, b) *em* $\mathbb{R}$*. Observe que a sigma-álgebra de Borel* $\mathcal{B}$ *é também a sigma-álgebra gerada por todos os intervalos fechados* $[a, b]$ *em* $\mathbb{R}$*. Qualquer conjunto em* $\mathcal{B}$ *é chamado um* ***conjunto de Borel****.*

Função Mensurável

Consideramos um espaço mensurável fixo $(X, \mathcal{S})$, logo temos a seguinte definição:

Definição A.14 (Função Mensurável). *Seja* $f : X \to \mathbb{R}$ *uma função real em* X*. Dada uma sigma-álgebra* $\mathcal{S}$*, diremos que a função* f *é mensurável (com respeito a* $\mathcal{S}$*) se para cada número* α*, o conjunto*

$$\{x \in X : f(x) > \alpha\}, \tag{A.2}$$

é mensurável.

São conseqüências da definição de função mensurável:

1. Os conjuntos $\{x \in X : f(x) = \alpha\}$, $\{x \in X : f(x) \geq \alpha\}$, $\{x \in X : f(x) < \alpha\}$ e $\{x \in X : f(x) \leq \alpha\}$ são mensuráveis.

2. Se X é o conjunto $\mathbb{R}$ e $\mathcal{S}$ é a sigma-algebra de Borel $\mathcal{S}$, então qualquer função contínua $f : \mathbb{R} \to \mathbb{R}$ é mensurável de Borel.

3. Se $X = \mathbb{R}$ e $\mathcal{S} = \mathcal{B}$, então qualquer função monótona é mensurável de Borel.

Medida

Introduzimos a noção de espaço mensurável $(X, \mathcal{S})$ consistindo de um conjunto X e uma sigma-álgebra $\mathcal{S}$ de subconjuntos de X. Vamos considerar certas funções definidas em $\mathcal{S}$ e tomando valores no conjunto do números reais ou reais estendidos. Estas funções generalizam a idéia de comprimento, área, massa, assim em diante.

Definição A.15. *Uma função $\mu : \mathcal{S} \to [0, +\infty]$ é chamada de medida em $(X, \mathcal{S})$ se*

(i) $\mu(A) \geq 0$ *para todo* $A \in \mathcal{S}$*;*

(ii) $\mu(\emptyset) = 0$*;*

(iii) se $\{A_n\}_{n=1}^{\infty}$ *é qualquer seqüencia disjunta de conjuntos em* $\mathcal{S}$*, então*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Vamos apresentar um exemplo considerado importante:

Exemplo A.1. *Se $X = \mathbb{R}$ e $\mathcal{S} = \mathcal{B}$, a sigma-álgebra de Borel, então podemos mostrar que existe uma única medida $\lambda : \mathcal{B} \to [0, +\infty]$ a qual coincide com o comprimento dos intervalos abertos, isto significa que se $A = (a, b) \neq \emptyset$, então $\lambda(a, b) = b - a$. Esta única medida é geralmente chamada* ***medida de Lesbegue (ou Borel)****. Esta não é uma medida finita, mas sim σ-finita.*

Podemos apresentar alguns resultados simples da teoria da medida:

Lema A.1. *Seja μ um medida definida em uma sigma-álgebra $\mathcal{S}$. Se $E, F \in \mathcal{S}$ e $E \subset F$, então $\mu(E) \leq \mu(F)$. Se $\mu < +\infty$, então $\mu(F - E) = \mu(F) - \mu(E)$.*

Também temos:

Lema A.2. *Seja $\mu : \mathcal{S} \to [0, +\infty]$ uma medida. Temos as seguintes afirmações:*

(a) Se $\{A_n\}$ é uma seqüencia crescente em $\mathcal{S}$, então:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

(b) Se $\{B_n\}$ é uma seqüencia decrescente em $\mathcal{S}$ e $\mu(B_1) < +\infty$, então

$$\mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n).$$

Espaço de Medida

Definição A.16 (Espaço de Medida). *É uma tripla $(X, \mathcal{S}, \mu)$ consistindo de um conjunto X, uma sigma-álgebra $\mathcal{S}$ de subconjuntos de X e uma medida μ definida em $\mathcal{S}$.*

Definição A.17. *Se μ é uma medida definida em $\mathcal{S}$, então uma relação envolvendo os elementos de X é dita **quase sempre** se o conjunto A de todos pontos para os quais a relação falha é um conjunto de medida nula, isto é, $\mu(A) = 0$.*

Integral de Lebesgue

Seja $\mathcal{S}$ uma sigma-álgebra em um conjunto X e μ uma medida positiva em $\mathcal{S}$.

Se $s : X \to \mathbb{R}^+$ é uma função simples mensurável, da forma

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

onde $\alpha_1, ..., \alpha_n$ são valores distintos de s e a χ_{A_i} é a função característica do conjunto A_i. Se $E \in \mathcal{S}$ definimos a integral de s em E com relação μ por:

$$\int_E s \; d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E).$$

Definição A.18. *Seja $f : X \to \mathbb{R}^+$ uma função mensurável e $E \in \mathcal{S}$. Definimos*

$$\int_E f \; d(\mu) = \sup_{s \leq f} \int_E s \; d(\mu). \tag{A.3}$$

A integral do lado direito de (A.3) é chamada integral de Lebesgue de f sobre E, com relação à medida μ.

Da definição de integral de Lebesgue temos as seguintes propriedades:

Teorema A.5 (**Teorema da Convergência Monótona**). *Seja $\{f_n\}$ uma seqüencia de funções mensuráveis em X, e suponha que*

(a) $0 \leq f_1(x) \leq f_2(x) \leq ... \leq \infty$ *para todo* $x \in X$*;*

(b) $f_n(x) \to f(x)$ *quando* $n \to \infty$*, para cada* $x \in X$*.*

Então, f é mensurável e

$$\int_X f_n d(\mu) \to \int_X f d(\mu).$$

Lema A.3 (**Lema de Fatou**). *Seja $\{f_n\}$ uma seqüencia de funções não-negativas e somáveis em $(X, \mathcal{S}, \mu)$. Então*

$$f(x) := \liminf_{n\to\infty} f_n(x)$$

é mensurável e

$$\int_X \left(\liminf_{n\to\infty} f_n\right) d(\mu) \leq \liminf_{n\to\infty} \int_X f_n d(\mu).$$

Teorema A.6 (**Teorema da Convergência Dominante de Lebesgue**). *Suponha uma seqüencia de funções integrais $\{f_n\}$ que converge para uma função f de valores reais mensurável, quase sempre. Se existir uma função integrável g tal que $|f_n| \leq g$ para todo n, então f é integrável e*

$$\lim_{n\to\infty} \int_X f_n(x)\ d(\mu) = \int_X f(x)\ d(\mu).$$

Teorema A.7 (**Teorema de Fubini**). *Sejam $(X_1, \mathcal{S}_1, \mu_1)$ e $(X_2, \mathcal{S}_2, \mu_2)$ dois espaços mensuráveis de sigma-finitos e $\mu_1 \times \mu_2$. Se f é uma função mensurável $\mathcal{S}_1 \times \mathcal{S}_2$ em $X_1 \times X_2$ e não-negativa, então a seguintes integrais são iguais*

$$\int_{X_1 \times X_2} f(x,y)\ d(\mu_1 \times \mu_2),$$

$$\int_{X_1} \left(\int_{X_2} f(x,y)\ d(\mu_2)\right) d(\mu_1),$$

$$\int_{X_2} \left(\int_{X_1} f(x,y)\ d(\mu_1)\right) d(\mu_2).$$

A.3 Multifunção Mensurável

Nesta seção, apresentamos limites de conjuntos, multifunção mensurável e o teorema da seleção mensurável. Estes resultados serão muitos utilizados em nossos estudos. As referências básicas podem ser encontradas em Clarke [9], Loewen [25] ou Vinter [34].

Limites de Conjuntos

Seja $\{A_i\}_{i\in\mathbb{N}}$ uma seqüencia de subconjuntos de $\mathbb{R}^n$. Definimos os seguintes conjuntos:

$$\limsup_{i\longrightarrow\infty} A_i \;:=\; \Big\{x\in\mathbb{R}^n : \exists\{A_{i_j}\}\subset\{A_i\} \text{ e } x_{i_j}\longrightarrow x \text{ tal que } x_{i_j}\in A_{i_j} \text{ para todo } j\Big\},$$

$$\liminf_{i\longrightarrow\infty} A_i \;:=\; \Big\{x\in\mathbb{R}^n : \exists\; x_i\longrightarrow x \text{ tal que } x_i\in A_i \text{ para todo } i\Big\}.$$

Podemos observar que os dois conjuntos $\limsup\limits_{n\longrightarrow\infty} A_n$ e $\liminf\limits_{n\longrightarrow\infty} A_n$ são fechados (possivelmente vazios) e que

$$\liminf_{n\longrightarrow\infty} A_n \subset \limsup_{n\longrightarrow\infty} A_n.$$

Se

$$\lim_{n\longrightarrow\infty} A_n := \liminf_{n\longrightarrow\infty} A_n (= \limsup_{n\longrightarrow\infty} A_n),$$

então, dizemos que $\{A_n\}$ tem limite.

Um contexto mais geral: Seja um conjunto $D\subset\mathbb{R}^n$ e uma família de conjuntos $\{S(y)\subset\mathbb{R}^n : y\in D\}$. Fixe um ponto $x\in\mathbb{R}^k$. Definimos os seguintes conjuntos:

$$\limsup_{y\xrightarrow{D}x} S(y) \;=\; \Big\{\xi : \exists\; y_i \xrightarrow{D} x \;\text{ e } \xi_i\longrightarrow\xi \text{ tal que } \xi_i\in S(y_i) \text{ para todo } i\Big\},$$

$$\liminf_{y\xrightarrow{D}x} S(y) \;=\; \Big\{\xi : \forall\; y_i \xrightarrow{D} x\;, \exists\; \xi_i\longrightarrow\xi \text{ tal que } \xi_i\in S(y_i) \text{ para todo } i\Big\}.$$

A notação $y_i \xrightarrow{D} x$ significa que $y_i\longrightarrow x$ e $y_i\in D$ para todo i.

Se D é uma vizinhança de x, podemos escrever $\liminf\limits_{y\longrightarrow x} S(y)$ no lugar de $\limsup\limits_{y\xrightarrow{D}x} S(y)$, etc. Também, observe que

$$\liminf_{y\xrightarrow{D}x} S(y) \subset \limsup_{y\xrightarrow{D}x} S(y)$$

e ambos conjuntos são fechados (possivelmente vazios).

Multifunção Mensurável

Seja $\Gamma : X \rightrightarrows \mathbb{R}^n$ uma multifunção de um conjunto X para subconjunto de $\mathbb{R}^n$, sendo que, para cada $x \in X$ atribui um conjunto $\Gamma(x) \subset \mathbb{R}^n$.

Considere $(X, \mathcal{S})$ um espaço mensurável. A multifunção $\Gamma : X \rightrightarrows \mathbb{R}^n$ é mensurável quando o conjunto $\{x \in X : \Gamma(x) \cap C \neq 0\}$ é mensurável em $\mathcal{S}$ para cada conjunto aberto $C \subset \mathbb{R}^n$.

Seleção Mensurável

Seja Γ uma multifunção com valores fechadas, compactas, convexas e não-vazia em X. Então, existe uma função mensurável $\gamma : X \rightarrow \mathbb{R}^n$ tal que $\gamma(x)$ pertence a $\Gamma(x)$ para todo $x \in X$.

Referências Bibliográficas

[1] Athans, M., and Falb, P. *Optimal Control: An Introduction to the Theory and Its Applications*. McGraw-Hill Book Company, New York, 1966.

[2] Bartle, R. G. *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, INC, New York, 1995.

[3] Bellman, R. *Dynamic Programming*. Princeton University Press, Princeton, 1957.

[4] Berman, A., and Plemmons, R. J. *Nonnegative Matrices in the Mathematical Sciences*. SIAM Classics in Applied Mathematical, Philadelphia, 1994.

[5] Bryson, A. E. Optimal control - 1650 to 1985. *IEEE Control Systems* (June 1996), 26–33.

[6] Clark, C. W., Clarke, F. H., and Munro, G. R. The optimal exploitation of renewable resource stocks: Problems of irreversible investment. *Econometrica 47*, (1979), 25–48.

[7] Clarke, F. H. *Necessary Conditions for Nonsmooth Problems in Optimal Control and the Calculus of Variations*. PhD thesis, Univ. of Washington, 1973.

[8] Clarke, F. H. *Methods of Dynamic and Nonsmooth Optimization*. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 57, SIAM Publications, Philadelphia, 1989.

[9] Clarke, F. H. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1990.

[10] Clarke, F. H., Stern, R., Ledyaev, Y. S., and Wolenski, P. R. *Nonsmooth Analysis and Control Theory*. New York, Springer, 1998.

[11] De Pinho, M. d. R. Maximum principle for mixed constrained control problem. *Nonlinear Analysis 47*, 1 (2001), 387–398.

[12] De Pinho, M. d. R., and Ferreira, M. M. A. Notes on nonsmooth analysis and optimal control. Disponível em: <http://cantor1.mathematik.uni-halle.de/institute/optimierung/reports/report02-25/Principal.ps> Acesso em: 12 setembro 2006.

[13] De Pinho, M. d. R., and Ilchmann, A. Weak maximum principle for optimal control problems with mixed constraints. *Nonlinear Analysis 48*, 8 (2002), 1179–1196.

[14] De Pinho, M. d. R., and Vinter, R. B. An Euler-Lagrange inclusion for optimal control problems. *IEEE Transactions Automatic Control 40*, **7** (1995), 1191–1198.

[15] De Pinho, M. d. R., and Vinter, R. B. Necessary conditions for optimal control problems involving nonlinear differential algebraic equations. *Journal of Mathematical Analysis and Applications 212* (1997), 493–516.

[16] De Pinho, M. d. R., Vinter, R. B., and Zheng, H. A maximum principle for optimal control problems with mixed constraints. *IMA Journal of Mathematical Control and Information 18*, 2 (2001), 189–205.

[17] Dreyfus, S. E. *Dynamic Programming and the Calculus of Variations*. Academic Press, New York, 1965.

[18] Fernandez, P. J. *Medida e Integração*. Instituto de Matemática Aplicada Pura e Aplicada, CNPq, Rio de Janeiro, 1976.

[19] Fleming, W. H., and Rishel, R. W. *Deterministic and Stochastic Optimal Control*. Springer Verlag, New Yourk, 1975.

[20] Holley, G. E., and Bryson, A. E. Wind modelling and lateral control for automatic landing. *Jour. Spacecraft and Aircraft 14*, **2** (1977), 65–72.

[21] Horn, R. A., and Johnson, C. R. *Topics in Matrix Analysis*. Cambridge University Press, New York, 1994.

[22] Kalman, R. *Contributions to the theory of optimal control*. Bol. de Soc. Math. Mexicana, 1960.

[23] Lieb, E. H., and Loss, M. *Analysis.* American Mathematical Society, Providence, 1997.

[24] Lin, F., and Brandt, R. D. An optimal control approach to robust control of robot manipulators. *IEEE Transactions on Robotics and Automation 14*, **1** (1998), 69–77.

[25] Loewen, P. *Optimal Control via Nonsmooth Analysis*, vol. Vol. 2. CRM Procedings & Lectures Notes, American Mathematical Society ,Providence, 1993.

[26] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F. *The Mathematical Theory of Optimal Processes.* Wiley Interscience, New York, 1962.

[27] Rieder, U. Measurable selection theorems for optimization problems. *Manuscripta Math 24* (1978), 115–131.

[28] Rockafellar, R. T. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1972.

[29] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis.* Berlin, Springer, 1998.

[30] Rudin, W. *Principles of Mathematical Analysis*, third ed. Mcgraw-Hill, New York, 1976.

[31] Rudin, W. *Real and Complex Analysis.* Mcgraw-Hill, New York, 1987.

[32] Sussman, H. J., and Willems, J. C. 300 years of optimal control: From the brachystochrone to the maximum principle. *IEEE Control Systems* (June 1997), 32–44.

[33] Varaya, P. *Notes on Optimization.* Van Nostrand Reinhold, 1972.

[34] Vinter, F. H. *Optimal Control.* Boston, Birkhauser, 2000.

[35] Vinter, R. B. *Optimal Control Theory.* Lectures Notes, Summer School "Optimization, Control and Related Subjects", Nice, 1987.