UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

JAMIELLI TOMAZ PEREIRA

Condições de Otimalidade para Problemas de Controle Ótimo com Restrições Mistas

Campinas

2016

Jamielli Tomaz Pereira

Condições de Otimalidade para Problemas de Controle Ótimo com Restrições Mistas

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Matemática Aplicada.

Orientador: Roberto Andreani

Coorientador: Valeriano Antunes de Oliveira

O arquivo digital corresponde à versão final da Dissertação defendida pela aluna Jamielli Tomaz Pereira e orientada pelo Prof. Dr. Roberto Andreani.

Campinas

2016

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2014/02028-2; CAPES

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

	Pereira, Jamielli Tomaz, 1991-
P414c	Condições de otimalidade para problemas de controle ótimo com restrições mistas / Jamielli Tomaz Pereira. – Campinas, SP : [s.n.], 2016.
	Orientador: Roberto Andreani.
	Coorientador: Valeriano Antunes de Oliveira.
	Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.
	1. Controle ótimo. 2. Otimização com restrições. 3. Condições de qualificação. I. Andreani, Roberto,1961-. II. Oliveira, Valeriano Antunes de. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Optimality conditions for optimal control problems with mixed constraints
Palavras-chave em inglês:
Optimal control
Constrained optimization
Constraint qualifications
Área de concentração: Matemática Aplicada
Titulação: Mestra em Matemática Aplicada
Banca examinadora:
Roberto Andreani [Orientador]
Paulo José da Silva e Silva
Geraldo Nunes Silva
Data de defesa: 26-02-2016
Programa de Pós-Graduação: Matemática Aplicada

Dissertação de Mestrado defendida em 26 de fevereiro de 2016 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.

Prof.(a). Dr(a). ROBERTO ANDREANI

Prof.(a). Dr(a). PAULO JOSÉ DA SILVA E SILVA

Prof.(a). Dr(a). GERALDO NUNES SILVA

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

Aos Alunos e Alunas das Escolas Públicas

Agradecimentos

Agradeço aos meus pais Carlos e Lucimar e à minha irmã Janaina pelo apoio incondicional, por acreditarem em mim e por mais longe que estivessem sempre me deram força para que eu atingisse meu objetivo. Sou extremamente grata a vocês e faço tudo isso por vocês e para vocês. Obrigada! O agradecimento se estende também para toda a minha família, não dá para citar todos, mas a cada um de vocês o meu muito obrigada.

Agradeço ao meu namorado José Herelis pela ajuda, por estar ao meu lado em todos os momentos deste período, que diga-se de passagem tem vários momentos críticos e desesperadores. Obrigada pela sua companhia, apoio, carinho, e pelos vários momentos de discussão. Não tem nem como retribuir tudo que você fez e faz por mim, muito obrigada de coração. Um agradecimento especial também à toda família Carnaúba que me acolheu em todos os finais de semana deste período.

Um agradecimento especial aos meus queridos orientadores Nino e Valeriano que foram os grandes responsáveis por tudo. O Valeriano que em todos estes anos desde a iniciação científica me orientou e me ajudou a chegar até aqui, foi pela sua indicação que cheguei ao Nino. O Nino que me recebeu extremamente bem e sempre me ajudou em tudo que precisei. Agradeço a vocês os momentos de discussões e aprendizado, vocês são sensasionais.

Sou grata também a todos os meus professores que desde os meus quatro anos contribuíram e muito para a minha formação. Todos foram os principais responsáveis pelas minhas escolhas. Um agradecimento especial aos meus professores da UNICAMP que contribuíram com o meu amadurecimento para que esta etapa fosse cumprida.

Agradeço aos membros da banca, Paulo e Geraldo, que aceitaram o convite e acrescentaram conhecimento e informações de extrema relevância para o desenvolvimento deste trabalho. Agradeço também ao programa de pós do IMECC, aos funcionários da unicamp, em especial aos da secretaria, limpeza, informática e logística, ao IMECC e à UNICAMP.

Agora um parágrafo inteiramente dedicado aos meus amigos que ao longo da minha vida fizeram toda a diferença. À vocês meus amigos de Frutal que por vários momentos foram presentes e importantes, em especial Ingrid, Lucas, Victor, Paulo Henrique, Arthur, Gustavo, todos vocês da minha amada cidadezinha continuam sendo muito importantes. Obrigada! Agradeço aos meus amigos da UNESP de Rio Preto, em especial ao Ricardo, Fernando, Daniela, Bruna, Anderson, a todas as meninas do 1D que moraram comigo, a todos os integrantes do PET da minha época, a todos da moradia da minha época, enfim são muitos, sou muito grata a vocês que neste período importante da minha formação contribuíram com momentos de alegria e companheirismo. Por fim, agradeço aos meus colegas da UNICAMP que estiveram comigo durante o período de disciplinas, em especial à Fernanda. E aos alunos do Ibilce que estão aqui na UNICAMP e sempre estão presentes e dispostos a ajudar, em especial, Fred, Wender, Mayara. Obrigada! Muito obrigada a cada pessoa que fez e faz parte da minha vida, todos contribuíram de alguma maneira para que eu chegasse até aqui.

Agradeço também à minha cunhada Binha que arrumou uma vaga na moradia e me recebeu disposta a me ajudar no que fosse necessário. Agradeço também aos moradores da A6 que me receberam e me deram um teto pelo tempo que eu precisasse. Este ato de vocês foi de extrema importância. Obrigada!

E por último, mas não menos importante, aliás extremamente importante, agradeço a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela bolsa de mestrado sob o processo 2014/02028-2 e também à CAPES pelos primeiros meses de bolsa, através deste incentivo tornaram possível a conclusão deste mestrado.

"I can't go back to yesterday, because I was a different person then."
(Lewis Carroll, Alice in Wonderland)

Resumo

Temos por objetivo principal nesta dissertação de mestrado o estudo de condições necessárias de otimalidade para problemas de controle ótimo com restrições de igualdade e de desigualdade, o celebrado Princípio do Máximo de Pontryagin. As restrições consideradas são do tipo mistas, envolvendo ambas as variáveis de estado e de controle. O princípio do máximo foi estudado no contexto não suave, fizemos uso do subdiferencial limite (de Mordukhovich) e do subdiferencial de Clarke. As condições necessárias de otimalidade são obtidas sob qualificações de restrições do tipo Mangasarian-Fromovitz.

Palavras-chave: controle ótimo. qualificação de restrições. condições necessárias de otimalidade. análise não suave.

Abstract

The main objective in this master's thesis is the study of necessary optimality conditions for optimal control problems with equality and inequality constraints, the celebrated Pontryagin Maximum Principle. The constraints to be considered are mixed type, involving both the state and control variables. The maximum principle has been studied in the nonsmooth setting, the limiting (Mordukhovich) subdifferential and the Clarke's subdifferential has been used. The necessary optimality conditions are obtained under constraint qualifications of Mangasarian-Fromovitz type.

Keywords: optimal control. constraint qualifications. necessary optimality conditions. nonsmooth analysis.

Lista de ilustrações

Lista de abreviaturas e siglas

CQ	Condição de Qualificação (ou Qualificação de Restrição ou *Constraint Qualification*)
KKT	Karush-Kuhn-Tucker
LICQ	Condição de Qualificação de Independência Linear
MFCQ	Condição de Qualificação de Mangasarian-Fromovitz
CRCQ	Condição de Qualificação de Posto Constante
CPLD	Condição de Qualificação de Dependência Linear Positiva Constante
LI	Linearmente Independente

Lista de símbolos

$\lvert \cdot \rvert$	Norma euclidiana
$W^{1,1}(T;\mathbb{R}^p)$	Espaço das funções absolutamente contínuas
$N_C^P(x)$	Cone normal proximal a C em x
$N_C(x)$	Cone normal limite a C em x
$x_i \xrightarrow{C} x$	$x_i \to x$ e $x_i \in C\, \forall\, i$
$epi\, f$	Epígrafo da f
$int\{C\}$	Interior de C
$fr\{C\}$	Fronteira de C
$dom\, f$	domínio (efetivo) da f
$\partial^P f(x)$	Subdiferencial proximal de f em x
$\partial f(x)$	Subdiferencial limite de f em x
$\partial_P^\infty f(x)$	Subdiferencial proximal assintótico de f em x
$\partial^\infty f(x)$	Subdiferencial limite assintótico de f em x
$x_i \xrightarrow{f} x$	$x_i \to x$ e $f(x_i) \to f(x)\, \forall\, i$
B	Bola fechada unitária centrada na origem
$f^0(x,v)$	Derivada direcional generalizada de f em x na direção v
$\bar{\partial} f(x)$	Subdiferencial de Clarke
$co\partial f(x)$	Envoltório convexo do subdiferencial limite
$\nabla f(x)$	Vetor gradiente de f em x
$\Psi_C(x)$	Função indicadora do conjunto C no ponto x
$Gr\, F$	Gráfico de F
$L^\infty(T;\mathbb{R}^p)$	Espaço das funções essencialmente limitadas
$L^1(T;\mathbb{R}^p)$	Espaço das funções integráveis
I	Matriz identidade

Sumário

1 Introdução

As qualificações de restrições (ou condições de qualificação ou constraint qualifications) na Teoria de Otimização são reconhecidamente importantes. Em um problema de programação matemática clássico em que as funções que o define são suaves, sabemos que as soluções ótimas de tais problemas satisfazem regras de multiplicadores de Lagrange. Por exemplo, a de Fritz John, entretanto esta regra permite que o multiplicador associado à função objetivo seja nulo, de modo que perdemos informações importantes do problema. Vê-se então que uma regra em que tal multiplicador seja não-nulo é crucial. Essa regra é conhecida como Condições de Karush-Kuhn-Tucker (ou condições de KKT). As qualificações de restrições desempenham um papel fundamental para garantir a validade das condições de KKT. É importante salientar que quando uma qualificação de restrição é válida, ou seja, quando as condições de KKT são verificadas, algoritmos eficientes podem ser definidos, análise de sensibilidade pode ser levada a cabo, e a teoria de dualidade pode ser desenvolvida. Isso se deve ao fato de que quando uma qualificação de restrição é verificada, é possível captar analiticamente as propriedades geométricas do problema; estas últimas fundamentais no desenvolvimento das condições de otimalidade.

Encontramos na literatura vários tipos de qualificações de restrições, sendo a mais conhecida a de independência linear dos gradientes das restrições ativas. Outras bastante difundidas são a de Abadie, a de Mangasarian-Fromovitz, a de Slater, etc. Há uma vasta bibliografia sobre qualificação de restrição em programação matemática. Ver Andreani et al. [1] [2] [3], Janin [4], Bazaraa et al. [5], Bertsekas [6] e Nocedal e Wright [7], por exemplo.

Naturalmente, as qualificações de restrições são igualmente importantes em outros tipos de problemas de otimização. Em particular, na Teoria de Controle Ótimo. Assim como acontece com as condições KKT que não são, em geral válidas, precisa-se assumir alguma qualificação de restrições. No caso de problemas de controle ótimo com restrições mistas a sua condição necessária também não é em geral válida. Alguma hipótese de regularidade (qualificação de restrições) deve ser assumida sob as restrições. Entretanto, a bibliografia para este assunto em controle ótimo não é tão rica como para problemas de otimização em dimensão finita. Os problemas com restrições mistas foram estudados sistematicamente por Hestenes [8], Neustadt [9] e Dubovickii e Milyutin [10], entre outros, e ainda permanece ativo, ver Dmitruk [11], Páles e Zeidan [12], Stefani e Zezza [13], Milyutin e Osmolovskii [14], Devdariani e Ledyaev [15], Frankowska e Rampazzo [16], Arutyunov [17], de Pinho et al. [18], de Pinho e Ilchmann [19], de Pinho [20], de Pinho e Rosenblueth [21], etc.

Pelo fato da bibliografia em qualificações de restrições em dimensão infinita não ser tão rica, vale a pena conhecer o assunto para possíveis contribuições futuras. Neste trabalho a finalidade é conseguir um bom conhecimento técnico na teoria de controle ótimo. Para isso estudou-se na teoria de otimização em dimensão finita: o problema de programação não linear, as qualificações de restrições e a condição de KKT. Na teoria de otimização em dimensão infinita: o problema de cálculo das variações e a Equação de Euler-Lagrange; o problema de controle ótimo com e sem restrições mistas, as qualificações de restrições e o princípio do máximo. Além disso, análise não suave, já que o problema de controle ótimo considerado pode ter dados não suaves.

A finalidade de estudar o problema de programação não linear é que o problema de controle ótimo considerado se parece um pouco com o de programação não linear no sentido que tem uma função objetivo para ser minimizada, restrições de igualdade e desigualdade. Além disso, para que o princípio do máximo seja válido, assim como acontece com a condição de KKT, precisa-se que seja assumido sob as restrições uma condição de qualificação. As condições de qualificação são muito importantes para provar a convergência de algoritmos, as quais validam a eficiência e a viabilidade dos mesmos, além de permitir desenvolver algumas características importantes. Consegue-se escrever as condições de qualificação clássicas da dimensão finita, independência linear dos gradientes das restrições ativas e Mangasarian-Fromovitz, em dimensão infinita e provar que a condição necessária do controle ótimo vale quando assumido que estas condições de qualificação se verificam. Isto nos faz pensar que dá para escrever as outras condições de qualificação da dimensão finita, que garantem a validade da condição de KKT, em dimensão infinita e provar a validade do princípio do máximo. Por isso, é importante estudar o problema de programação não linear e entender as técnicas utilizadas na teoria de otimização em dimensão finita.

A teoria de controle ótimo não foi uma continuação direta do clássico cálculo das variações, pois uma teoria que levava em consideração as características dos problemas de controle ótimo precisou ser desenvolvida. Entretanto, usa-se o problema de cálculo das variações para introduzir o problema de controle ótimo. Este fato ocorre, pois o problema de cálculo das variações tem uma semelhança extremamente importante com o problema de controle ótimo, ambos problemas minimizam um certo funcional e a solução de ambos é uma função. Assim, as condições de otimalidade terão uma certa equivalência, pois o fato de minimizar um funcional é o mesmo, mas as restrições se diferem e os argumentos do funcional também. Conhecer a condição de otimalidade do problema do cálculo das variações nos ajuda a compreender as condições de otimalidade para o problema de controle ótimo. Compreender e interpretar as condições necessárias de otimalidade do problema de controle ótimo com restrições mistas é o principal objetivo deste trabalho, assim os resultados de cálculo das variações serão úteis nesta tarefa.

Estudado os pré-requisitos ficamos preparados para estudar o problema de controle ótimo. O controle ótimo passou a ser uma área distinta de todas que já existiam quando foi demonstrado o princípio do máximo na década de 1950 e precisou-se de técnicas específicas para o problema. Especificamente neste trabalho as técnicas para demonstrar as condições necessárias de otimalidade para o problema de controle ótimo são baseadas em perturbação, eliminação de restrições e passagem de limite, estas técnicas foram desenvolvidas por Clarke, Ioffe, Loewen, Mordukhovich, Rockafellar, Vinter, e outros.

O problema de controle ótimo considerado neste trabalho tem dados possivelmente não suaves, então se faz necessário, também como pré-requisito, o estudo de análise não suave. Problemas com dados não suaves são considerados, pois segundo Clarke [23], tornou-se claro que eles surgem naturalmente no cenário de controle ótimo.

As qualificações de restrições surgem no contexto de controle ótimo quando é considerado um problema com restrições mistas, precisa-se desta hipótese para garantir a validade do princípio do máximo. Neste trabalho, é visto as técnicas da demonstração do princípio do máximo quando assumimos uma condição do tipo Mangasarian-Fromovitz. Esta hipótese do tipo Mangasarian-Fromovitz tem que ser válida apenas no processo ótimo, para garantir a validade do princípio do máximo. Conhecer o princípio do máximo, bem como as técnicas utilizadas na sua demonstração faz parte do caminho a ser trilhado para possíveis contribuições futuras.

Este trabalho está dividido da seguinte forma: o Capítulo 1 com esta breve introdução; no Capítulo 2 de preliminares temos o problema de programação não linear, o problema de cálculo das variações e suas respectivas condições necessárias de otimalidade, e além disso definimos o problema de controle ótimo; no Capítulo 3 tem-se construções básicas de análise não suave, cones normais, subdiferenciais, cálculo subdiferencial e alguns resultados necessários para este trabalho; no Capítulo 4 há alguns elementos básicos para prosseguimento do trabalho e temos a parte mais importante que são as condições necessárias de otimalidade para o problema de controle ótimo de Pontryagin e com restrições mistas e para finalizar o Capítulo 5 é a conclusão, o fechamento do trabalho.

2 Preliminares

Neste capítulo definiremos o problema principal deste trabalho, o problema de controle ótimo com restrições mistas, além disso o problema de programação não linear e o problema de cálculo das variações.

Antes de definirmos o problema de interesse, o problema de controle ótimo, vamos ver o problema de programação não linear, o problema de cálculo das variações e suas respectivas condições necessárias de otimalidade. Desta forma, vamos nos apropriar do assunto e quando as condições necessárias de otimalidade para o problema de controle ótimo forem apresentadas será mais fácil compreendermos o motivo de serem como são, já que serão análogas às condições das teorias clássicas, programação não linear e cálculo das variações.

Apresentar o problema de programação não linear e sua condição necessária de otimalidade é relevante, mesmo não sendo de dimensão infinita como o problema de controle ótimo, pois a teoria que propomos estudar em controle ótimo é espelhada na teoria de programação não linear. Assim como em programação não linear em que para garantir que vale a condição necessária de otimalidade, a condição de KKT, precisa-se assumir uma hipótese de regularidade sob as restrições, a qualificação de restrição, no controle ótimo também precisa-se assumir tal hipótese (no contexto de dimensão infinita) para garantir que vale a condição necessária de otimalidade, o Princípio do Máximo. Além disso, a condição necessária para o problema de controle ótimo lembra a condição de KKT, só que no contexto de dimensão infinita.

Já o problema de cálculo das variações e sua condição necessária de otimalidade é válido ser apresentado, pois desta forma estaremos nos preparando para um problema em dimensão infinita. O problema de cálculo das variações é clássico e bastante difundido e assim como o problema de controle ótimo ele também é um problema de dimensão infinita, sua solução está em espaços de funções, que são espaços de dimensão infinita. Então, conhecer a condição necessária de otimalidade do problema de cálculo das variações, a Equação de Euler-Lagrange, nos ajudará a compreender as condições necessárias de otimalidade para o problema de controle ótimo.

2.1 O Problema de Programação Não Linear

As principais referências para esta seção são Ribeiro e Karas [24], Andreani, Echagüe e Schuverdt [25] e Andreani, Martínez e Schuverdt [2].

Um problema de programação não linear pode ser posto como:

$$
\begin{array}{lll}
\text{minimizar} & f(x) & \\
\text{sujeito a} & c_{\mathcal{E}}(x) = 0, & \\
 & c_{\mathcal{I}}(x) \leqslant 0, &
\end{array} \tag{C}
$$

onde $f : \mathbb{R}^n \to \mathbb{R}, c_i : \mathbb{R}^n \to \mathbb{R}, i \in \mathcal{E} \cup \mathcal{I}$ são funções de classe C^2. O conjunto

$$\Omega = \{x \in \mathbb{R}^n | c_{\mathcal{E}}(x) = 0, c_{\mathcal{I}}(x) \leqslant 0\}$$

é o conjunto viável.

Definição 2.1. *Seja $\bar{x} \in \Omega$. Uma restrição de desigualdade $c_i, i \in \mathcal{I}$ é dita ativa em $\bar{x}$ se $c_i(\bar{x}) = 0$. Vamos denotar por $I(\bar{x})$ o conjunto de índices das restrições de desigualdade ativas em um ponto viável $\bar{x}$, isto é,*

$$I(\bar{x}) = \{i \in \mathcal{I} | c_i(\bar{x}) = 0\}.$$

Definição 2.2. *Dizemos que $x^* \in \Omega$ é um minimizador local de f em Ω quando existe $\delta > 0$, tal que $f(x^*) \leqslant f(x)$ para todo $x \in B(x^*, \delta) \cap \Omega$.*

Uma condição necessária de otimalidade para o problema (C) é a condição de KKT:

Teorema 2.1. *Seja $x^* \in \Omega$ um minimizador local do problema (C) e suponha que alguma hipótese de regularidade (qualificação de restrição) vale sob as restrições. Então existem vetores λ^* e μ^* tais que*

$$
\begin{aligned}
& -\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in \mathcal{I}} \mu_i^* \nabla c_i(x^*), \\
& \mu_i^* \geqslant 0, \; i \in \mathcal{I}, \\
& \mu_i^* c_i(x^*) = 0, \; i \in \mathcal{I}.
\end{aligned}
$$

2.1.1 Qualificações de Restrições

As qualificações de restrições ou condições de qualificação (CQ) são condições impostas sob as restrições que garantem que um minimizador satisfaz as relações de KKT. As de interesse, neste trabalho, são a Condição de Qualificação de Independência Linear (LICQ), a Condição de Qualificação de Mangasarian-Fromovitz (MFCQ), a Condição de Qualificação de Posto Constante (CRCQ) e a Condição de Qualificação de Dependência Linear Positiva Constante (CPLD).

2.1.1.1 Condição de Qualificação de Independência Linear

Definição 2.3. *Dizemos que a condição de qualificação de independência linear é satisfeita em $\bar{x}$ quando o conjunto formado pelos gradientes das restrições de igualdade e das restrições* *de desigualdade ativas é linearmente independente, isto é,*

$$\{\nabla c_i(\bar{x}) | i \in \mathcal{E} \cup I(\bar{x})\} \text{ é LI.}$$

2.1.1.2 Condição de Qualificação de Mangasarian-Fromovitz

Definição 2.4. *A condição de qualificação de Mangasarian-Fromovitz é satisfeita em $\bar{x}$ quando os gradientes das restrições de igualdade são linearmente independentes e existir um vetor $d \in \mathbb{R}^n$ tal que*

$$\nabla c_i(\bar{x})^\top d = 0 \text{ e } \nabla c_j(\bar{x})^\top d < 0,$$

para todos $i \in \mathcal{E}$ e $j \in I(\bar{x})$.

Abaixo definimos, como alternativa, quando um ponto não satisfaz MFCQ (isto é, MF-não regular).

Definição 2.5. *Dizemos que $x \in \mathbb{R}^n$ é MF-não regular se existem multiplicadores $\lambda_1, \ldots, \lambda_m \in \mathbb{R}, \mu_i \geqslant 0 \forall i \in I(x)$, com $\sum_{j \in J} |\lambda_j| + \sum_{i \in I(x)} \mu_i > 0$, tais que*

$$\sum_{j=1}^{m} \lambda_j \nabla h_j(x) + \sum_{i \in I(x)} \mu_i \nabla g_i(x) = 0.$$

Observação 2.1. *A condição de Mangasarian-Fromovitz é também conhecida como condição de independência linear positiva.*

2.1.1.3 Condição de Qualificação de Posto Constante

Definição 2.6. *Dado uma família de funções diferenciáveis $\{f_i(x) : i = 1, \ldots, r\}, f_i : \mathbb{R}^n \to \mathbb{R}$, dizemos que a condição de posto constante vale em x^* se, e somente se, para qualquer subconjunto $K \subset \{i \in \{1, \ldots, r\} : f_i(x^*) = 0\}$, a família de gradientes*

$$\{\nabla f_i(x)\}_{i \in K}$$

permanece com o posto constante na vizinhança do ponto x^*.

Definição 2.7. *Dado a família de funções diferenciáveis $\{c_i(x) : i \in \mathcal{E}; c_j(x) : j \in \mathcal{I}\}$ associadas com o problema (C), dizemos que um ponto viável $x^* \in \Omega$ satisfaz a condição de qualificação de posto constante se, e somente se, a condição de posto constante vale em x^*.*

CRCQ foi definida por Janin [4].

2.1.1.4 Condição de Qualificação de Dependência Linear Positiva Constante

Definição 2.8. *Sejam $x \in \mathbb{R}^n, I_0 \subset I(x), J_0 \subset \mathcal{E}$. Dizemos que o conjunto dos gradientes $\{\nabla c_i(x)\}_{i\in I_0} \cup \{\nabla c_j(x)\}_{j\in J_0}$ é positivo linearmente dependente se existem escalares $\{\alpha_j\}_{j\in J_0}, \{\beta_i\}_{i\in I_0}$ tais que $\beta_i \geqslant 0$ para todo $i \in I_0$, $\sum_{j\in J_0} |\alpha_j| + \sum_{i\in I_0} \beta_i > 0$, e*

$$\sum_{j\in J_0} \alpha_j \nabla c_j(x) + \sum_{i\in I_0} \beta_i \nabla c_i(x) = 0.$$

Caso contrário, dizemos que o conjunto de gradientes $\{\nabla c_i(x)\}_{i\in I_0} \cup \{\nabla c_j(x)\}_{j\in J_0}$ é positivo linearmente independente.

Definição 2.9. *Um ponto viável x é dito satisfazer a condição CPLD se para quaisquer $I_0 \subset I(x), J_0 \subset \mathcal{E}$ tais que o conjunto de gradientes $\{\nabla c_i(x)\}_{i\in I_0} \cup \{\nabla c_j(x)\}_{j\in J_0}$ é positivo linearmente dependente, existe uma vizinhança $N(x)$ de x tal que, para qualquer $y \in N(x)$, o conjunto $\{\nabla c_i(y)\}_{i\in I_0} \cup \{\nabla c_j(y)\}_{j\in J_0}$ é linearmente dependente.*

2.2 O Problema de Cálculo das Variações

A principal referência utilizada nesta seção foi Clarke [26]. O problema clássico de cálculo das variações consiste em encontrar uma função $x(t)$ definida no intervalo $[0,1]$ satisfazendo as condições $x(0) = x_0, x(1) = x_1$ que minimiza o funcional

$$J(x) = \int_0^1 L(t, x(t), \dot{x}(t))dt,$$

onde L é uma função de três variáveis, referida como o Lagrangiano e as três variáveis t, x, v são o tempo, o estado e a velocidade, respectivamente. Tomamos $L : \mathbb{R}^3 \to \mathbb{R}$ sendo uma função duas vezes continuamente diferenciável e limitamos a atenção para funções $x : [0,1] \to \mathbb{R}$ que pertencem a $C^2[0,1]$. Além disso, a integral definindo $J(x)$ está bem definida para cada x. O problema básico é:

$$\begin{array}{ll} \text{minimizar} & J(x) \\ \text{sujeito a} & x \in C^2[0,1], \\ & x(0) = x_0, x(1) = x_1. \end{array} \tag{V}$$

$J(x)$ é referido como o custo correspondente a x. Uma função $x : [0,1] \to \mathbb{R}$ é admissível se satisfaz as restrições de fronteira e está na classe apropriada, nesse caso $C^2[0,1]$. Uma solução x^* de (V) é uma função admissível tal que $J(x^*) \leqslant J(x)$ para todas as outras funções admissíveis x.

Neste problema, sem restrições, não queremos encontrar um ponto que minimiza uma função, como na programação não linear, mas sim uma função que minimiza um funcional. Dessa forma, a condição necessária de otimalidade não é mais a derivada no ponto solução ser igual a zero, mas sim a Equação de Euler-Lagrange. Esta equação é encontrada usando o cálculo das variações e é um resultado análogo a Regra de Fermat de que $f'(x) = 0$ em um mínimo, a condição necessária clássica da derivada na solução ser igual a zero.

A condição necessária de otimalidade para o problema de cálculo das variações, a Equação de Euler-Lagrange, é encontrada da seguinte maneira: tomamos uma perturbação da solução $\bar{x}$ de modo que esta seja viável

$$x = \bar{x} + \lambda\eta$$

onde λ é um parâmetro real e $\eta : [0,1] \to \mathbb{R}^n, \eta \in C^2[0,1]$. Assim, para cada λ,

$$g(\lambda) = J(\bar{x} + \lambda\eta) \geqslant J(\bar{x}) = g(0),$$

segue que g atinge o mínimo em $\lambda = 0$ e por isso $g'(0) = 0$. Usando este fato e o Lema de du Bois-Reymond obtemos a Equação de Euler-Lagrange

$$\frac{d}{dt}L_{\dot{x}}(t, \bar{x}(t), \dot{\bar{x}}(t)) = L_x(t, \bar{x}(t), \dot{\bar{x}}(t)) \ \forall t \in [0,1].$$

A Equação de Euler-Lagrange é uma condição necessária de otimalidade, ou seja, suas soluções são candidatas a solução do problema.

2.3 O Problema de Controle Ótimo

Nesta seção definiremos nosso problema de interesse, o problema de controle ótimo com restrições mistas. O primeiro componente básico de um problema de controle ótimo é o sistema de controle

$$\dot{x}(t) = f(t, x(t), u(t), v(t)), \quad (x(0), x(1)) \in C,$$

que é constituído pela dinâmica que relaciona a variável de estado $x(t) \in \mathbb{R}^n$ e as variáveis de controle $u(t) \in \mathbb{R}^{k_u}$ e $v(t) \in V(t) \subset \mathbb{R}^{k_v}$ e a condição de contorno, onde $C \subset \mathbb{R}^n \times \mathbb{R}^n$. Temos que $t \in T = [0,1]$ é o tempo, com 0 e 1 o tempo inicial e final, respectivamente.

O segundo componente de um problema de controle ótimo é o funcional objetivo, que pode ser dado de três formas

$$\int_0^1 L(t, x(t), u(t), v(t))dt \qquad \text{(Forma de Lagrange)}$$

$$\int_0^1 L(t, x(t), u(t), v(t))dt + l(x(0), x(1)) \qquad \text{(Forma de Bolza)}$$

$$l(x(0), x(1)) \qquad \text{(Forma de Mayer)}$$

onde $l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ e $L : T \times \mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \to \mathbb{R}$. Estas formas do funcional objetivo são todas equivalentes. A forma que trabalharemos será a forma de Mayer, por isso abaixo segue como transformar as formas de Lagrange e Bolza para a forma de Mayer.

Forma de Lagrange $\Rightarrow$ Forma de Mayer

Demonstração. Considere o funcional na forma de Lagrange:

$$\int_0^1 L(t, x(t), u(t), v(t))dt,$$

introduzimos uma nova variável de estado $x_{n+1}(t)$ tal que $\dot{x}_{n+1} = L(t, x(t), u(t), v(t))$. Temos,

$$\int_0^1 L(t, x(t), u(t), v(t))dt = \int_0^1 \dot{x}_{n+1}dt = x_{n+1}(1) - x_{n+1}(0).$$

Considerando $l(x(0), x(1)) = x_{n+1}(1) - x_{n+1}(0)$ temos o funcional na forma de Mayer. □

Forma de Bolza $\Rightarrow$ Forma de Mayer

Demonstração. Considere o funcional na forma de Bolza:

$$\int_0^1 L(t, x(t), u(t), v(t))dt + l(x(0), x(1)),$$

analogamente ao caso anterior temos,

$$\int_0^1 L(t, x(t), u(t), v(t))dt + l(x(0), x(1)) = x_{n+1}(1) - x_{n+1}(0) + l(x(0), x(1)).$$

Considerando $s(x(0), x(1)) = x_{n+1}(1) - x_{n+1}(0) + l(x(0), x(1))$ segue que o funcional está na forma de Mayer. □

A forma de Mayer segundo Clarke [28] é a formulação que se provou ser natural no modelo de uma variedade de problemas de física, economia e engenharia.

Continuando com a formulação do problema de controle ótimo, para o nosso problema de interesse teremos um terceiro componente, que são as restrições mistas, restrições de igualdade e desigualdade,

$$g(t, x(t), u(t), v(t)) \leqslant 0, \quad b(t, x(t), u(t), v(t)) = 0$$

onde $(g, b) : T \times \mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \to \mathbb{R}^{m_g} \times \mathbb{R}^{m_b}$.

O problema de controle ótimo pode ser escrito de várias formas diferentes, com diferentes funcionais objetivo, com controle restrito ou controle irrestrito, com condições de contorno fixas ou não e com ou sem restrições de igualdade e desigualdade, entre outras variações. Mas o problema de interesse principal é o de encontrar um processo (x, u, v) que seja solução do seguinte problema:

$$\begin{array}{ll} \text{minimizar} & l(x(0), x(1)) \\ \text{sujeito a} & \dot{x}(t) = f(t, x(t), u(t), v(t)) \; q.t. \; t \in T, \\ & g(t, x(t), u(t), v(t)) \leqslant 0 \; q.t. \; t \in T, \\ & b(t, x(t), u(t), v(t)) = 0 \; q.t. \; t \in T, \\ & v(t) \in V(t) \; q.t. \; t \in T, \\ & (x(0), x(1)) \in C. \end{array} \tag{P}$$

O problema de controle ótimo difere do problema de cálculo das variações, pois tem, além da variável de estado, a variável de controle, além disso tem a dinâmica $(\dot{x}(t) = f(t, x(t), u(t), v(t))$ que relaciona a variável de estado e as variáveis de controle, este tipo de restrição caracteriza um problema de controle ótimo. Mas o comum dos dois problemas é que estamos em busca de uma função que minimiza um certo funcional e satisfaz as restrições colocadas.

3 Análise Não Suave

O termo análise não suave refere-se à teoria que desenvolve um “cálculo diferencial” para funções que não são diferenciáveis no sentido usual e para conjuntos que não são variedades suaves clássicas. Este é um assunto, dentro de um campo matemático grande, mas tem também desempenhado um papel importante em várias áreas de aplicação como otimização, cálculo das variações, equações diferenciais, mecânica e teoria de controle. Entre aqueles que participam do seu desenvolvimento estão: J. Borwein, A. D. Ioffe, B. Mordukhovich, R. T. Rockafellar, R. B. Vinter e F. H. Clarke.

No final da década de 1950 importantes avanços foram feitos no estudo de problemas de controle ótimo, um deles foi o Princípio do Máximo, um conjunto de condições necessárias para uma função controle ser ótima. Na década seguinte, tornou-se aparente que o progresso foi sendo impedido pela perda da ferramenta analítica adequada para investigar propriedades locais de funções que não são suaves, isto é, não diferenciáveis no sentido tradicional. Quando esforços foram feitos para estender a aplicabilidade das condições necessárias, emergiram as funções não suaves.

A análise não suave teve início com as contribuições de Rockafellar no desenvolvimento da análise convexa, na década de 1960. Na década seguinte iniciou-se um desejo de estender para contextos não convexos. Em 1973, um avanço importante ocorreu, a teoria de F. H. Clarke de gradientes generalizados, começa o campo de análise não suave que fornece uma ponte para as condições necessárias de otimalidade. O livro do Clarke [28] de 1983 cobriu muitos avanços importantes da década anterior e suas aplicações para a derivação de condições necessárias em controle ótimo.

O cone normal proximal, o cone normal limite e seu envoltório convexo (o cone normal de Clarke), junto com seus subdiferenciais associados, a saber, o subdiferencial proximal, subdiferencial limite e o subdiferencial de Clarke (o gradiente generalizado), respectivamente, e a teoria que envolve estes cones e subdiferenciais serão necessários para estabelecer o Princípio do Máximo no contexto não suave. Neste capítulo serão apresentadas construções básicas de análise não suave e resultados de cálculo, que serão relevantes para a teoria que envolve os problemas de controle ótimo. A principal referência utilizada neste capítulo foi Vinter [29], onde também pode ser encontrado as demonstrações de todos os resultados apresentados.

3.1 Cones Normais

Vamos definir o cone normal proximal e o cone normal limite, além disso estabeleceremos algumas de suas propriedades.

Definição 3.1. *Dado um conjunto fechado $C \subset \mathbb{R}^k$ e um ponto $x \in C$, o cone normal proximal a C em x, escrito como $N_C^P(x)$, é o conjunto*

$$N_C^P(x) := \{p \in \mathbb{R}^k : \exists M > 0 \text{ tal que } p \cdot (y - x) \leqslant M|y - x|^2 \ \forall y \in C\}.$$

Os elementos em $N_C^P(x)$ são chamados normais proximais a C em x.

A proposição abaixo dá uma interpretação geométrica de normais proximais.

Proposição 3.1. *Tome um conjunto fechado $C \subset \mathbb{R}^k$ e pontos $x \in C$ e $p \in \mathbb{R}^k$. Então p é um normal proximal a C em x se, e somente se, existe um ponto $z \in \mathbb{R}^k$ e um fator de escala $r > 0$ tais que*

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

Definição 3.2. *Dado um conjunto fechado $C \subset \mathbb{R}^k$ e um ponto $x \in C$, o cone normal limite a C em x, escrito como $N_C(x)$, é o conjunto*

$$N_C(x) := \{p : \text{ existem } x_i \xrightarrow{C} x, p_i \to p \text{ tais que } p_i \in N_C^P(x_i) \ \forall i\}.$$

Os elementos em $N_C(x)$ são chamados normais limites a C em x.

Para prosseguimento precisaremos definir o epígrafo de uma função.

Definição 3.3. *O conjunto*

$$epi\, f := \{(x, \alpha) \in \mathbb{R}^k \times \mathbb{R} : \alpha \geqslant f(x)\}$$

é o epígrafo da f.

Exemplo 3.1. *Sejam $f : \mathbb{R} \to \mathbb{R}$ e $x = 0$, temos:*

(i) $f(x) = |x|$, *então*

$$N^P_{epi\, f}(0,0) = N_{epi\, f}(0,0) = \{(x, y) \in \mathbb{R}^2 : y \leqslant -|x|\}.$$

(ii) $f(x) = -|x|$, *então*

$$N^P_{epi\, f}(0,0) = \{(0,0)\} \text{ e } N_{epi\, f}(0,0) = \{(x, y) \in \mathbb{R}^2 : y = -|x|\}.$$

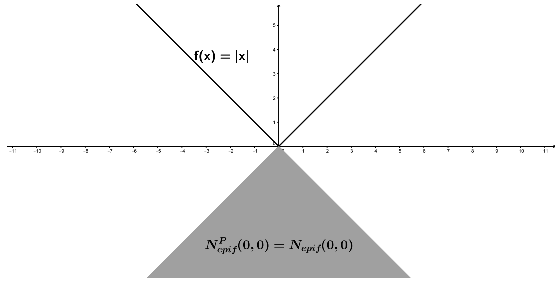

Figura 1 – Exemplo 3.1 (i)

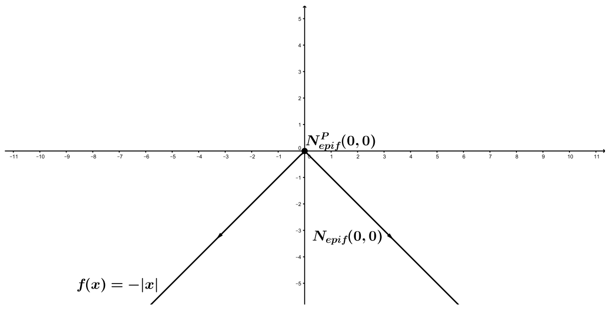

Figura 2 – Exemplo 3.1 (ii)

Segue, agora, algumas propriedades dos cones normais:

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

(i) $N^P_C(x)$ *e* $N_C(x)$ *são cones em* $\mathbb{R}^k$, *contendo o* $\{0\}$ *e* $N^P_C(x) \subset N_C(x)$;

(ii) $N^P_C(x)$ *é convexo;*

(iii) a multifunção $y \to N_C(y) : C \to \mathbb{R}^k$ *tem um gráfico fechado, no sentido que, para quaisquer sequências* $y_i \overset{C}{\to} y$ *e* $p_i \to p$ *tais que* $p_i \in N_C(y_i)$ *para todo* i, *temos* $p \in N_C(y)$.

Proposição 3.3. *Tome um conjunto fechado $C \subset \mathbb{R}^k$ e um ponto $x \in C$. Então:*

(i) $x \in int\{C\}$ *implica* $N_C(x) = \{0\}$.

(ii) $x \in fr\{C\}$ *implica* $N_C(x)$ *contém elementos não nulos.*

Proposição 3.4. *Tome subconjuntos $C_1 \subset \mathbb{R}^m$ e $C_2 \subset \mathbb{R}^n$, e um ponto $(x_1, x_2) \in C_1 \times C_2$. Então*

$$N^P_{C_1 \times C_2}(x_1, x_2) = N^P_{C_1}(x_1) \times N^P_{C_2}(x_2), \tag{3.1}$$

$$N_{C_1 \times C_2}(x_1, x_2) = N_{C_1}(x_1) \times N_{C_2}(x_2). \tag{3.2}$$

No caso convexo, os cones normais coincidem com o cone normal no sentido de Análise Convexa.

Proposição 3.5. *Tome um conjunto convexo fechado $C \subset \mathbb{R}^k$ e um ponto $\bar{x} \in C$. Então*

$$N_C^P(\bar{x}) = N_C(\bar{x}) = \{\xi : \xi \cdot (x - \bar{x}) \leqslant 0 \ \forall x \in C\}.$$

O cone normal limite é o mais utilizado em aplicações na otimização, por conta das suas propriedades analíticas superiores. Muitas dessas propriedades são consequências do fato que o cone normal limite tem um gráfico fechado, quando considerado como uma multifunção do ponto base.

3.2 Subdiferenciais

Definição 3.4. *Uma função $f : \mathbb{R}^k \to \mathbb{R}$ é semicontínua inferior em um ponto $x_0 \in \mathbb{R}^n$ se*

$$f(x_0) \leqslant \liminf_{x \to x_0} f(x).$$

Teorema 3.1. *Uma função $f : \mathbb{R}^k \to \mathbb{R}$ é semicontínua inferior em $x_0 \in \mathbb{R}^n$ se, e somente se, epi f é fechado.*

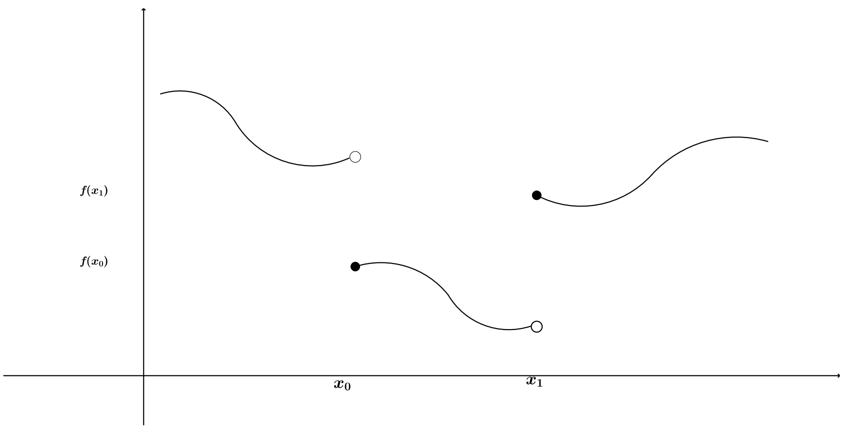

Figura 3 – Exemplo de uma função semicontínua inferior em x_0, mas que não é semicontínua inferior em x_1.

Tomaremos sempre uma função semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$, pois o conjunto epígrafo destas funções é fechado. Isto será importante já que os subgradientes serão definidos em termos do cone normal ao epígrafo da f.

Definição 3.5. *O domínio efetivo de uma função f em $S \subset \mathbb{R}^k$, que denotamos por dom f, é a projeção em $\mathbb{R}^k$ do epígrafo de f:*

$$dom\, f = \{x | \exists \mu, (x, \mu) \in epi\, f\} = \{x | f(x) < \infty\}.$$

Definição 3.6. *Tome uma função semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ e um ponto $x \in dom\, f$.*

(i) O subdiferencial proximal de f em x, escrito como $\partial^P f(x)$, é o conjunto

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

Os elementos do $\partial^P f(x)$ são chamados subgradientes proximais.

(ii) O subdiferencial limite de f em x, escrito como $\partial f(x)$, é o conjunto

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

Os elementos do $\partial f(x)$ são chamados subgradientes limite.

Exemplo 3.2. *Sejam $f : \mathbb{R} \to \mathbb{R}$ e $x = 0$, temos:*

(i) $f(x) = |x|$, então

$$\partial^P f(x) = \partial f(x) = [-1, 1].$$

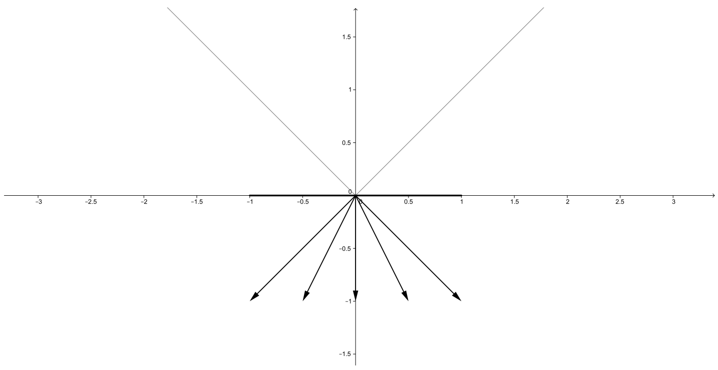

Figura 4 – Exemplo 3.2 (i)

(ii) $f(x) = -|x|$, então

$$\partial^P f(x) = \varnothing \text{ e } \partial f(x) = \{-1\} \cup \{1\}.$$

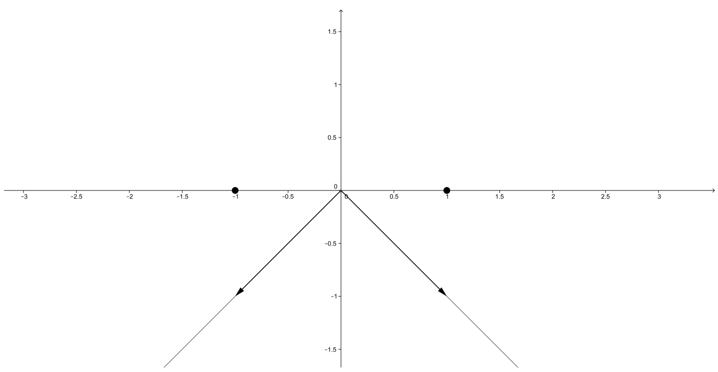

Figura 5 – Exemplo 3.2 (ii)

Subdiferenciais vazios ou ilimitados são um alerta que a inclinação da função em uma bola arbitrariamente pequena sobre o ponto base é ilimitado. Quando isso ocorre gostaríamos de saber a direção dessa inclinação arbitrariamente grande. Esta informação é expressa pela "assintótica" dos subdiferenciais já definidos.

Definição 3.7. *Tome uma função semicontínua inferior* $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ *e um ponto* $x \in dom\, f$*.*

(i) O subdiferencial proximal assintótico de f *em* x*, escrito como* $\partial_P^\infty f(x)$*, é*

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

Os elementos do $\partial_P^\infty f(x)$ *são chamados subgradientes proximais assintóticos.*

(ii) O subdiferencial limite assintótico de f *em* x*, escrito como* $\partial^\infty f(x)$*, é*

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

Os elementos do $\partial^\infty f(x)$ *são chamados subgradientes limite assintóticos.*

Exemplo 3.3. *Considere a função* $f : \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = sgn\{x\}|x|^{1/2}$ *e o ponto* $x = 0$*, temos:*

$$N_{epi\, f}(0,0) = \{(x,0) \in \mathbb{R}^2 : x \geqslant 0\}.$$

Assim

$$\partial f(x) = \varnothing \text{ e } \partial^\infty f(x) = [0, \infty).$$

O cálculo do $\partial^\infty f(x)$ *nos revela que a inclinação "infinita" próximo a* x *é positiva.*

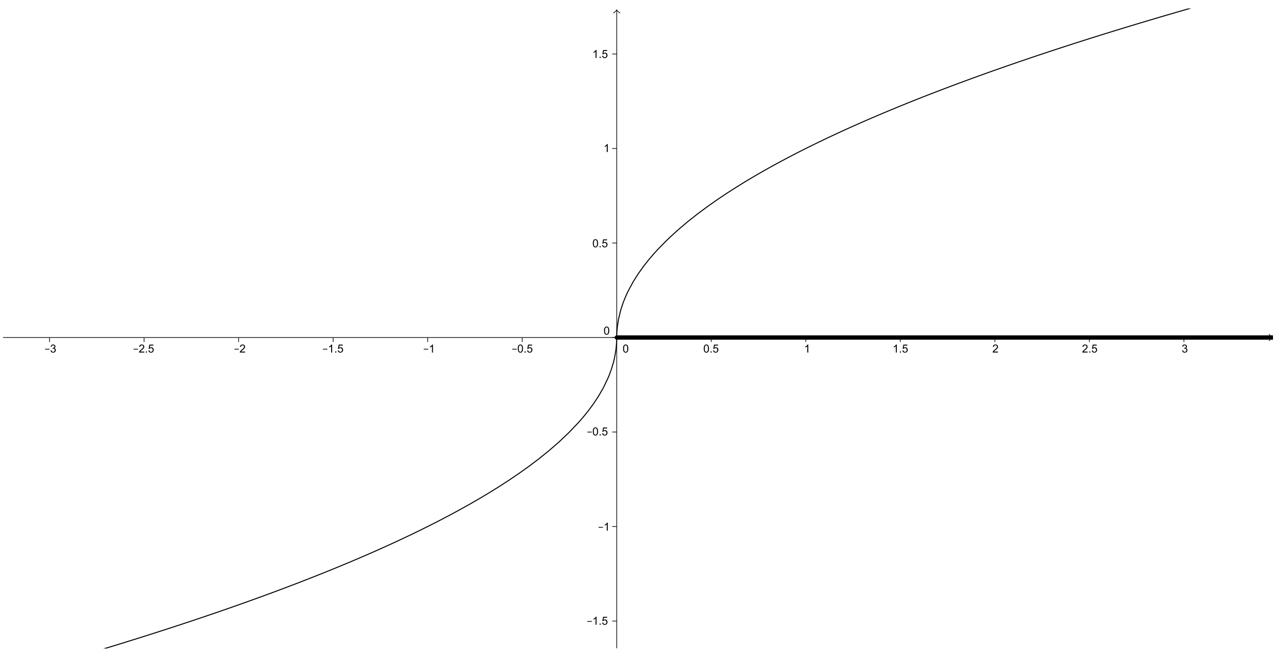

Figura 6 – Exemplo 3.3

Proposição 3.6. *Tome uma função semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ e um ponto $x \in \operatorname{dom} f$. Então:*

(i) $\partial f(x)$ é um conjunto fechado. Dado sequências $x_i \xrightarrow{f} x$ e $\xi_i \to \xi$ tais 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 sequências $x_i \xrightarrow{f} x$ e $\xi_i \to \xi$ tais que $\xi_i \in \partial^\infty f(x_i)$ para todo i, então $\xi \in \partial^\infty f(x)$.

Para funções convexas, as definições de subdiferencial limite e proximal coincidem com o subdiferencial no sentido de Análise Convexa.

Proposição 3.7. *Tome uma função convexa semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ e um ponto $\bar{x} \in \operatorname{dom} f$. Então*

$$\partial^P f(\bar{x}) = \partial f(\bar{x}) = \{\xi : \xi \cdot (x - \bar{x}) \leqslant f(x) - f(\bar{x}) \ \forall x \in \mathbb{R}^k\}.$$

3.2.1 Subgradientes de Funções Lipschitz

Definição 3.8. *Uma função f em um conjunto $S \subset \mathbb{R}^k$ é dita ser Lipschitz se existe um $k \geqslant 0$ para o qual*

$$|f(x') - f(x)| \leqslant k|x - x'| \ \forall \ x, x' \in S.$$

Como funções Lipschitz são um caso especial de função semicontínua inferior, tudo que vimos até aqui vale. Mas funções Lipschitz são encontradas muito frequentemente em aplicações de Análise Não Suave, em particular neste trabalho em que as funções consideradas são Lipschitz. Assim, é importante explorar sua estrutura e abordagens alternativas para aproximação local.

A hipótese de continuidade Lipschitz diz que a função tem uma inclinação limitada. Assim, não é surpresa que os subdiferenciais de funções Lipschitz são conjuntos limitados e o subdiferencial limite assintótico é o conjunto trivial $\{0\}$.

Proposição 3.8. *Tome uma função semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ e um ponto $x \in \mathbb{R}^k$. Assuma que f é Lipschitz em uma vizinhança de x com K a constante de Lipschitz. Então:*

(i) $\partial f(x)$ é não vazio e $\partial f(x) \subset KB$;

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

Agora examinamos uma abordagem alternativa para definir subdiferencial de função Lipschitz baseada em ideias de aproximações convexas e dualidade. Esta abordagem é importante, pois fornecerá novas representações de subdiferenciais que são extremamente úteis nas aplicações.

Iniciamos definindo a derivada direcional generalizada (Clarke) de uma função localmente Lipschitz.

Definição 3.9. *Tome uma função $f : \mathbb{R}^k \to \mathbb{R}$ e pontos $x \in \mathbb{R}^k$ e $v \in \mathbb{R}^k$. Assuma que f é Lipschitz em uma vizinhança de x. A derivada direcional generalizada de f em x na direção v, escrita como $f^0(x, v)$, é o número*

$$f^0(x, v) := \limsup_{y \to x, t \downarrow 0} t^{-1}[f(y + tv) - f(y)].$$

Proposição 3.9. *Tome uma função $f : \mathbb{R}^k \to \mathbb{R}$ e um ponto $x \in \mathbb{R}^k$. Assuma que f é Lipschitz em um vizinhança de x com constante de Lipschitz K. Então a função $v \to f^0(x, v)$ com domínio $\mathbb{R}^k$ tem as seguintes propriedades.*

(i) É valor-finito, Lipschitz com constante de Lipschitz K e positivamente homogênea, no sentido que

$$f^0(x, \alpha v) = \alpha f^0(x, v) \ \forall \ v \in \mathbb{R}^k \text{ e } \alpha \geqslant 0.$$

(ii) É convexa.

Definimos uma função convexa $f^0(x, v)$ que aproxima f "próximo" a x. É, então, natural introduzir um "subdiferencial", escrito como $\bar{\partial} f(x)$, que é o subdiferencial no sentido de análise convexa da função convexa $v \to f^0(x, v)$ em $v = 0$. Como $f^0(x, 0) = 0$, segue que

$$\bar{\partial} f(x) := \{\xi : f^0(x, v) \geqslant \xi \cdot v \ \forall v \in \mathbb{R}^k\}.$$

$\bar{\partial} f(x)$ é chamado de subdiferencial de Clarke de f em x. $\bar{\partial} f(x)$ é um conjunto (para x fixo) não vazio, compacto e convexo. Os elementos do $\bar{\partial} f(x)$ são uniformemente limitados na norma euclidiana pela constante Lipschitz de f em uma vizinhança de x.

A derivada direcional generalizada pode ser interpretada como a função suporte do $\bar{\partial} f(x)$:

Proposição 3.10. *Tome uma função $f : \mathbb{R}^k \to \mathbb{R}$ que é Lipschitz em uma vizinhança de um ponto $x \in \mathbb{R}^k$. Então*

$$f^0(x, v) = \max\{v \cdot \xi : \xi \in \bar{\partial} f(x)\} \text{ para todo } v \in \mathbb{R}^k.$$

O subdiferencial de Clarke comuta com -1:

Proposição 3.11. *Tome uma função $f : \mathbb{R}^k \to \mathbb{R}$ que é Lipschitz em uma vizinhança de um ponto $x \in \mathbb{R}^k$. Então*

$$\bar{\partial}(-f)(x) = -\bar{\partial} f(x).$$

Observação 3.1. *A afirmação da Proposição 3.11 é em geral falsa para o subdiferencial limite. Veja o Exemplo 3.2.*

A proposição seguinte traz a relação entre a subdiferencial limite e o subdiferencial de Clarke.

Proposição 3.12. *Tome uma função semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R}$ que é Lipschitz em uma vizinhança de algum ponto $x \in \mathbb{R}^k$. Então*

$$\bar{\partial} f(x) = co\partial f(x).$$

Onde o $co(\cdot)$ é definido como:

Definição 3.10. *Seja S um subconjunto de E^n. O envoltório convexo de S, denotado por $co(S)$, é o conjunto das interseções de todos os conjuntos convexos contendo S.*

De acordo com o Teorema de Rademacher se $f : \mathbb{R}^k \to \mathbb{R}$ é Lipschitz em uma vizinhança de um ponto $x \in \mathbb{R}^k$, então f é diferenciável em quase todos os pontos nesta vizinhança (em relação a medida k-dimensional de Lebesgue). Pelo teorema seguinte, o envoltório convexo do conjunto de limites de derivadas na vizinhança coincide com o envoltório convexo do subdiferencial limite em x. Este é um teorema de interesse, pois relaciona o conceito moderno e clássico de derivadas, além disso, fornece uma ferramenta computacional para o envoltório convexo do subdiferencial limite de grande poder.

Observação 3.2. *Note que se f é diferenciável em x e Lipschitz em uma vizinhança de x, então*

$$\nabla f(x) \in co\partial f(x).$$

Teorema 3.2. *Tome uma função $f : \mathbb{R}^k \to \mathbb{R}$, um ponto $x \in \mathbb{R}^k$ e qualquer subconjunto $\Omega \subset \mathbb{R}^k$ tendo medida de Lebesgue nula. Assuma que f é Lipschitz em uma vizinhança de x. Então*

$$\bar{\partial} f(x) = co\partial f(x) = co\{\xi : \exists x_i \to x, x_i \notin \Omega, \nabla f(x_i) \text{ existe e } \nabla f(x_i) \to \xi\}.$$

3.3 Cálculo Subdiferencial

Nesta seção apresentamos algumas regras de cálculo de subdiferencial úteis para estimar os subdiferenciais limite. É interessante que estas regras se apliquem a funções semicontínuas inferior gerais, mas isto nem sempre ocorre. Alguma hipótese de regularidade será imposta, que se diferirá de regra para regra, sendo que em cada caso esta hipótese elimina certos tipos de interação do subdiferencial limite assintótico. Para nosso caso de maior interesse, quando as funções são Lipschitz, estas hipóteses de regularidade são automaticamente satisfeitas.

A propriedade, "se f atinge seu valor mínimo em $\bar{x} \in dom\, f$, então $\{0\}$ está contido no subdiferencial de f em $\bar{x}$", é um conceito de subdiferencial útil no campo da otimização e felizmente ele vale.

Proposição 3.13. *Tome uma função semicontínua inferior $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ e um ponto $x \in dom\, f$. Assuma que x atinge o valor mínimo de f sobre uma vizinhança de x; então*

$$0 \in \partial^P f(x).$$

Note que pela Proposição 3.2 (i), temos $N_C^P(x) \subset N_C(x)$, logo como os subdiferenciais proximal e limite são definidos em termos dos cones normais proximal e limite, respectivamente, temos $\partial^P f(x) \subset \partial f(x)$, logo pela proposição anterior $0 \in \partial f(x)$.

Proposição 3.14 (Regra da Soma Básica)**.** *Tome funções $f : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ e $g : \mathbb{R}^k \to \mathbb{R}$, um conjunto fechado $C \subset \mathbb{R}^k$ e um ponto $x \in (int\, C) \cap (dom\, f)$. Assuma que f é semicontínua inferior e g é de classe C^2 em uma vizinhança de x. Então,*

$$\begin{aligned}
&\partial^P(f + g + \Psi_C)(x) = \partial^P f(x) + \{\nabla g(x)\},\\
&\partial(f + g + \Psi_C)(x) = \partial f(x) + \{\nabla g(x)\},\\
&\partial^\infty(f + g + \Psi_C)(x) = \partial^\infty f(x).
\end{aligned}$$

Aqui $\Psi_C : \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ denota a função indicadora do conjunto C, definida por

$$\Psi_C(x) := \begin{cases} 0, & \text{se } x \in C,\\ \infty, & \text{se } x \notin C. \end{cases}$$

Segue, abaixo, um corolário do Teorema do Princípio de Função Marginal, veja [29]. Apresentamos o corolário, pois ele é o mais utilizado. Este teorema e seu corolário são importantes, pois algumas das regras de cálculo apresentadas, a saber, o Subgradiente Parcial Limite, a Regra da Soma, a Regra da Cadeia, são consequências do Princípio de Função Marginal.

Corolário 3.1. *Tome uma função semicontínua inferior $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ e um ponto $(\bar{x}, \bar{u}) \in dom\, f$. Assuma que $\bar{x}$ minimiza $x \to F(x, \bar{u})$ sob alguma vizinhança de $\bar{x}$. Suponha que*

$$\{\eta : (0, \eta) \in \partial^\infty F(\bar{x}, \bar{u})\} = \{0\}.$$

Então existe um ponto $\xi \in \mathbb{R}^m$ tal que

$$(0, \xi) \in \partial F(\bar{x}, \bar{u}).$$

Teorema 3.3 (Subgradiente Limite Parcial)**.** *Tome uma função semicontínua inferior* $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ *e um ponto* $(\bar{x}, \bar{y}) \in dom\, f$. *Assuma que*

$$(0, \eta) \in \partial^\infty f(\bar{x}, \bar{y}) \Rightarrow \eta = 0.$$

Então:

$$\partial_x f(\bar{x}, \bar{y}) \subset \{\xi : \text{ existe } \eta \text{ tal que } (\xi, \eta) \in \partial f(\bar{x}, \bar{y})\}$$
$$\partial_x^\infty f(\bar{x}, \bar{y}) \subset \{\xi : \text{ existe } \eta \text{ tal que } (\xi, \eta) \in \partial^\infty f(\bar{x}, \bar{y})\}.$$

Teorema 3.4 (Regra da Soma)**.** *Tome funções semicontínuas inferior* $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}, i = 1, ..., m,$ *e um ponto* $\bar{x} \in \cap_i dom\, f_i$. *Defina* $f = f_1 + \ldots + f_m$. *Assuma que*

$$v_i \in \partial^\infty f_i(\bar{x}), i = 1, \ldots, m \text{ e } \sum_i v_i = 0 \Rightarrow v_i = 0\ \forall i.$$

Então

$$\partial f(\bar{x}) \subset \partial f_1(\bar{x}) + \ldots + \partial f_m(\bar{x})$$

e

$$\partial^\infty f(\bar{x}) \subset \partial^\infty f_1(\bar{x}) + \ldots + \partial^\infty f_m(\bar{x}).$$

É possível obter uma estimativa do cone normal ao gráfico de uma aplicação Lipschitz por meio de uma aplicação da Regra da Soma.

Proposição 3.15. *Tome uma aplicação semicontínua inferior* $G : \mathbb{R}^n \to \mathbb{R}^m$ *e um ponto* $u \in \mathbb{R}^n$. *Assuma que* G *é Lipschitz em uma vizinhança de* u. *Então*

$$N_{Gr\, G}(u, G(u)) \subset \{(\xi, -\eta) : \xi \in \partial(\eta \cdot G)(u), \eta \in \mathbb{R}^m\}.$$

Teorema 3.5 (Regra da Cadeia)**.** *Tome uma função localmente Lipschitz* $G : \mathbb{R}^n \to \mathbb{R}^m$, *uma função semicontínua inferior* $g : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ *e um ponto* $\bar{u} \in \mathbb{R}^n$ *tais que* $G(\bar{u}) \in dom\ g$. *Defina a função semicontínua inferior* $f(u) := g \circ G(u)$. *Assuma que:*

$$\text{O único vetor } \eta \in \partial^\infty g(G(\bar{u})) \text{ tal que } 0 \in \partial(\eta \cdot G)(\bar{u}) \text{ é } \eta = 0.$$

Então

$$\partial f(\bar{u}) \subset \{\xi : \text{ existe } \eta \in \partial g(G(\bar{u})) \text{ tal que } \xi \in \partial(\eta \cdot G)(\bar{u})\}$$

e

$$\partial^\infty f(\bar{u}) \subset \{\xi : \text{ existe } \eta \in \partial^\infty g(G(\bar{u})) \text{ tal que } \xi \in \partial(\eta \cdot G)(\bar{u})\}.$$

Como corolários da Regra da Cadeia, temos a Regra do Máximo e a Regra do Produto, que seguem.

Teorema 3.6 (Regra do Máximo)**.** *Tome funções localmente Lipschitz* $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$ *e um ponto* $\bar{x} \in \mathbb{R}^n$*. Defina* $f(x) = \max_i f_i(x)$ *e* $\Lambda := \{\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m : \lambda_i \geqslant 0, \sum_i \lambda_i = 1\}$*. Então*

$$\partial f(\bar{x}) \subset \{\partial(\sum_{i=1}^{m} \lambda_i f_i)(\bar{x}) : \lambda \in \Lambda,\ \ e\ \lambda_i = 0\ se\ f_i(\bar{x}) < f(\bar{x})\}.$$

Teorema 3.7 (Regra do Produto)**.** *Tome funções localmente Lipschitz* $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$ *e um ponto* $\bar{x} \in \mathbb{R}^n$*. Defina* $f(x) = f_1(x) f_2(x) \ldots f_m(x)$*. Então*

$$\partial f(\bar{x}) \subset \partial(\sum_{i}^{m} \Pi_{j \neq i} f_j(\bar{x}) f_i)(\bar{x}).$$

4 Condições de Otimalidade para Problemas de Controle Ótimo

No início da Guerra Fria, logo após a II Guerra Mundial, as duas superpotências, Estados Unidos e União Soviética se empenharam em resolver vários problemas matemáticos que ficaram conhecidos, mais tarde, como problemas de controle ótimo, por exemplo, o problema de interceptação em tempo mínimo de uma aeronave de caça.

Os matemáticos que trabalharam com o problema de controle ótimo nos Estados Unidos, a saber, Magnus R. Hestenes (1906-1991), Rufus P. Isaacs (1914-1981) e Richard E. Bellman (1920-1984), fizeram avanços na teoria, mas não deram uma demonstração rigorosa do princípio do máximo. Hestenes trabalhou com o problema de interceptação em tempo mínimo, em 1950, ele o reformulou como um problema de cálculo das variações e aplicou os resultados já conhecidos, sendo esta a primeira formulação do princípio do máximo, mas para isso precisou de hipóteses mais fortes como as funções controle serem contínuas e assumirem valores em um domínio aberto. Isaacs e Bellman obtiveram resultados de programação dinâmica em que o princípio do máximo estava oculto, mas eles não reconheceram as consequências e conexões das teorias.

Já os matemáticos que trabalharam com o problema de interceptação em tempo mínimo na União Soviética, a saber, Lev Semyonovich Pontryagin (1908-1988), Vladimir Grigor'evich Boltyanski (1925-) e Revaz Valerianovich Gamkrelidze (1927-), não ficaram atrelados ao cálculo das variações. Pontryagin teve as primeiras ideias, daí Gamkrelidze e Boltyanski completaram. No Congresso Internacional de Matemática em Edinburgh, em agosto de 1958, foi apresentada a demonstração do princípio do máximo e em seguida na Federação Internacional de Controle Automático em 1960 em Moscou foram discutidas as relações entre o princípio do máximo e o cálculo das variações, antes disso de acordo com os matemáticos da União Soviética, as condições do cálculo das variações não eram conhecidas durante o desenvolvimento do princípio do máximo.

O controle ótimo emergiu como um distinto campo de pesquisa na década de 1950 quando L. S. Pontryagin et al. demonstraram o Princípio do Máximo, conjunto de condições necessárias de otimalidade para problemas de controle ótimo. A partir daí, busca-se estender a aplicabilidade do princípio do máximo que é uma ferramenta importante em muitas áreas. Isto é feito, por exemplo, enfraquecendo hipóteses ou estendendo os tipos de problemas para os quais o resultado se aplica. Neste trabalho estudamos um problema mais geral do que o considerado por Pontryagin, com restrições de igualdade e desigualdade, custo e dinâmica não suaves e é mostrado que o Princípio do Máximo vale para este problema com as restrições satisfazendo uma hipótese do tipo Mangasarian-Fromovitz. A principal referência para este capítulo é de Pinho e Rosenblueth [21].

4.1 Definições Básicas

Nesta seção, vamos estabelecer definições e notações que serão usadas posteriormente. Primeiro, relembre que $T = [0, 1]$.

Definição 4.1. *Um processo $(\bar{x}, \bar{u}, \bar{v})$ é uma tripla (x, u, v) compreendendo de uma função $x \in W^{1,1}(T; \mathbb{R}^n)$ e funções mensuráveis $u : T \to \mathbb{R}^{k_u}$ e $v : T \to \mathbb{R}^{k_v}$ satisfazendo as restrições.*

Definição 4.2. *Um processo $(\bar{x}, \bar{u}, \bar{v})$ é um minimizador fraco se, para algum $\epsilon > 0$, o processo minimiza o custo sobre todos os processos (x, u, v) satisfazendo $|x(t) - \bar{x}(t)| \leqslant \epsilon$ para todo $t \in T$ e*

$$|u(t) - \bar{u}(t)| \leqslant \epsilon, \ |v(t) - \bar{v}(t)| \leqslant \epsilon \text{ q.t. } t \in T.$$

A notação $r \geqslant 0$ para $r \in \mathbb{R}^p$ $(p \in \mathbb{N})$ significa que cada componente de r é não negativo e a notação $(b, g)(\cdot)$ significa $(b(\cdot), g(\cdot))$.

$(\bar{x}, \bar{u}, \bar{v})$ é um minimizador local e $\bar{\phi}(t)$ é a função ϕ avaliada em $(\bar{x}, \bar{u}, \bar{v})$. Os conjuntos Ω_ϵ e $\hat{\Omega}_\epsilon$ são definidos por

$$\Omega_\epsilon(t) := (\bar{x}(t) + \epsilon B) \times (\bar{u}(t) + \epsilon B) \times ((\bar{v}(t) + \epsilon B) \cap V(t)),$$

$$\hat{\Omega}_\epsilon(t) := (\bar{x}(t) + \epsilon B) \times ((\bar{v}(t) + \epsilon B) \cap V(t)),$$

respectivamente.

4.2 Problema de Pontryagin

Primeiro consideramos o problema na forma de Pontryagin.

$$\begin{array}{ll} \text{minimizar} & l(x(0), x(1)) \\ \text{sujeito a} & \dot{x}(t) = f(t, x(t), v(t)) \text{ q.t. } t \in T, \\ & v(t) \in V(t) \text{ q.t. } t \in T, \\ & (x(0), x(1)) \in C. \end{array} \qquad (I)$$

Com a dinâmica e o custo possivelmente não suaves. Além disso, mais algumas hipóteses precisam ser feitas para que o problema considerado faça sentido.

(h1) A função $t \mapsto f(t, x, v)$ é Lebesgue mensurável para cada (x, v) e existe uma função integrável k_f tal que, para quase todo $t \in T$,

$$|f(t, x, v) - f(t, x', v')| \leqslant k_f(t)|(x, v) - (x', v')|$$

para todo $(x, v), (x', v') \in \hat{\Omega}_\epsilon(t)$.

(h2) O gráfico de V é um conjunto Borel mensurável e $V_\epsilon(t) := (\bar{v}(t) + \epsilon B) \cap V(t)$ é fechado em quase todo $t \in T$.

(h3) C é fechado e l é localmente Lipschitz em uma vizinhança de $(\bar{x}(0), \bar{x}(1))$.

Agora estamos prontos para enunciar o Princípio do Máximo Não Suave, não é apenas Princípio do Máximo, pois originalmente os dados considerados eram suaves, o resultado apresentado abaixo já é uma extensão do resultado clássico.

Teorema 4.1 (Princípio do Máximo Não Suave)**.** *Seja $(\bar{x}, \bar{v})$ um minimizador fraco para (I). Defina $H(t, x, p, v) := p \cdot f(t, x, v)$. Se, para algum $\epsilon > 0$, (h1)-(h3) são satisfeitas, então existem $p \in W^{1,1}(T; \mathbb{R}^n)$ e $\lambda \geqslant 0$ tais que*

(i) $\|p\|_\infty + \lambda \neq 0$,

(ii) $-\dot{p}(t) \in co\partial_x H(t, \bar{x}(t), p(t), \bar{v}(t))$ *q.t.* $t \in T$,

(iii) $H(t, \bar{x}(t), p(t), \bar{v}(t)) = \max_{v \in V(t)} H(t, \bar{x}(t), p(t), v)$ *q.t.* $t \in T$,

(iv) $(p(0), -p(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \partial l(\bar{x}(0), \bar{x}(1))$.

Os elementos (λ, p), cuja existência é afirmada, são chamados multiplicadores para (I). Os componentes λ e p são referidos como o multiplicador custo e o arco adjunto ou coestado, respectivamente.

As conclusões do teorema anterior são todas esperadas do que já conhecemos das teorias clássicas de programação não linear e cálculo das variações. De fato, a condição (i), a condição de não trivialidade dos multiplicadores, é esperada, pois se os dois multiplicadores forem zero não teremos nenhuma informação do problema, nem das restrições e nem da função objetivo; a condição (ii), a equação de coestado ou equação adjunta, nada mais é do que a Equação de Euler-Lagrange $\dfrac{d}{dt} H_{\dot{x}}(t, \bar{x}(t), p(t), \bar{v}(t)) \in co\partial_x H(t, \bar{x}(t), p(t), \bar{v}(t))$, que já era esperada aparecer por se tratar da condição necessária de um problema que busca uma função e não um ponto como solução, como no problema de cálculo das variações, que é o nosso caso no problema de controle ótimo; a condição (iii), a condição de máximo, é a condição que dá nome ao teorema, Princípio do Máximo, esta condição será apresentada de uma maneira diferente nos resultados que seguem, mas continuam sendo condições necessárias esperadas para um processo ser solução. Finalmente a condição (iv), condição de transversalidade e condição de fronteira, está relacionada com a restrição $(x(0), x(1)) \in C$ e com o funcional objetivo l.

Observação 4.1. *(a) O jacobiano generalizado de Clarke de uma função $L : \mathbb{R}^n \to \mathbb{R}^m$ em um ponto y, tem a seguinte propriedade, para qualquer vetor linha $r \in \mathbb{R}^m$,*

$$co\partial(r \cdot L)(y) = r\, co\partial L(y)$$

onde $\partial(r \cdot L)(y)$ é o subdiferencial limite da função $y \to r \cdot L(y)$. Dessa propriedade segue que a conclusão (ii) do teorema anterior pode ser equivalentemente escrita como

$$-\dot{p}(t) \in p(t)\ co\partial_x f(t, \bar{x}(t), \bar{v}(t)).$$

(b) As condições necessárias do teorema acima são homogêneas em relação aos multiplicadores (p, λ). Isso significa que se (p, λ) servem como um conjunto de multiplicadores, então para qualquer $\alpha > 0, (\alpha p, \alpha\lambda)$ também servem. Como $(p, \lambda) \neq 0$, podemos sempre escolher um α apropriado tal que

$$\|p\|_\infty + \lambda = 1.$$

4.3 Problema com Restrições Mistas

O problema (P) abaixo é o de principal interesse neste trabalho. Comparado com o problema (I) apenas acrescentamos restrições de igualdade, restrições de desigualdade e o controle irrestrito u.

$$\begin{array}{ll} \text{minimizar} & l(x(0), x(1)) \\ \text{sujeito a} & \dot{x}(t) = f(t, x(t), u(t), v(t)) \ q.t.\ t \in T, \\ & g(t, x(t), u(t), v(t)) \leqslant 0 \ q.t.\ t \in T, \\ & b(t, x(t), u(t), v(t)) = 0 \ q.t.\ t \in T, \\ & v(t) \in V(t) \ q.t.\ t \in T, \\ & (x(0), x(1)) \in C. \end{array} \qquad (P)$$

Também o consideramos com a dinâmica e o custo possivelmente não suaves. Assim como fizemos para o problema (I), também precisamos estabelecer algumas hipóteses para que este problema faça sentido.

(H1) A função $t \mapsto f(t, x, u, v)$ é Lebesgue mensurável para cada (x, u, v) e existe uma função integrável k_f tal que, para quase todo $t \in T$,

$$|f(t, x, u, v) - f(t, x', u', v')| \leqslant k_f(t)|(x, u, v) - (x', u', v')|$$

para todo $(x, u, v), (x', u', v') \in \Omega_\epsilon(t)$.

(H2) O gráfico de V é um conjunto Borel mensurável e $V_\epsilon(t) := (\bar{v}(t) + \epsilon B) \cap V(t)$ é fechado em quase todo $t \in T$.

(H3) C é fechado e l é localmente Lipschitz em uma vizinhança de $(\bar{x}(0), \bar{x}(1))$.

(H4) $(b, g)(\cdot, x, u, v)$ é mensurável para cada (x, u, v) e $t \mapsto g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$ é $L^\infty(T, \mathbb{R}^{m_g})$.

(H5) $(b, g)(t, \cdot, \cdot, \cdot)$ é continuamente diferenciável em $(\bar{x}(t), \bar{u}(t), \bar{v}(t)) + \epsilon B$ *q.t.* $t \in T$ e existe uma função integrável $L_{b,g}$ tal que, para quase todo $t \in T$,

$$|(b, g)(t, x, u, v) - (b, g)(t, x', u', v')| \leqslant L_{b,g}(t)|(x, u, v) - (x', u', v')|$$

para todo $(x, u, v), (x', u', v') \in \Omega_\epsilon(t)$.

(H6) Existe $K_{b,g} > 0$ e uma função crescente $\tilde{\theta} : (0, \infty) \to (0, \infty)$ com $\tilde{\theta}(s) \downarrow 0$ quando $s \downarrow 0$ tal que, para quase todo $t \in T$,

$$\begin{aligned} &|\nabla_x(\bar{b}, \bar{g})(t)| + |\nabla_u(\bar{b}, \bar{g})(t)| + |\nabla_v(\bar{b}, \bar{g})(t)| \leqslant K_{b,g}, \\ &|\nabla_{x,u,v}(b, g)(t, x, u, v) - \nabla_{x,u,v}(b, g)(t, x', u', v')| \leqslant \tilde{\theta}(|(x, u, v) - (x', u', v')|) \end{aligned}$$

para todo $(x, u, v) \neq (x', u', v')$ pertencente a $\Omega_\epsilon(t)$.

4.3.1 Qualificações de Restrições

Além das hipóteses (H1)-(H6), para estabelecermos as condições necessárias para o problema (P) precisamos de mais uma hipótese de regularidade nas restrições mistas, igualdade e desigualdade. Esta hipótese de regularidade é chamada de qualificação de restrição (ou *constraint qualification*). Neste trabalho veremos duas qualificações de restrições: a condição de posto completo e a condição do tipo Mangasarian-Fromovitz.

4.3.1.1 Posto Completo

Uma qualificação de restrição, básica e forte, usada na literatura equivalente em problemas de dimensão finita, a hipótese conhecida como independência linear (LICQ), é a hipótese de posto completo

$$\det F(t)F(t)^\top \neq 0 \;\forall\, t \in T,$$

onde

$$F(t) := \begin{pmatrix} \nabla_u b(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) \\ \nabla_u g^{\mathcal{I}_a(t)}(t, \bar{x}(t), \bar{u}(t), \bar{v}(t) \end{pmatrix}, \tag{4.1}$$

com $\mathcal{I}_a(t) = \{i \in \{1, \dots, m_g\} | g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0\}$ denotando o conjunto de índices das restrições ativas. Temos $\nabla_u g^{\mathcal{I}_a(t)}(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) \in \mathbb{R}^{q_a(t) \times k_u}$ a matriz que obtemos removendo de $\nabla_u g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$ todas as linhas de índices $i \notin \mathcal{I}_a(t)$, onde $q_a(t)$ denota a cardinalidade de $\mathcal{I}_a(t)$.

Pelo fato dos dados de (P) serem assumidos apenas mensuráveis com relação a t, a condição de posto completo é substituída pela condição de posto completo uniforme que chamaremos de (H*):

(H*) Existe $K > 0$ tal que $\det F(t)F(t)^\top \geqslant K$ q.t. $t \in T$, onde F é definida em (4.1).

Com as restrições mistas satisfazendo a hipótese de posto completo uniforme (H*) temos que vale o seguinte teorema, um variante do Princípio do Máximo Não Suave, que nos dá as condições necessárias para um processo ser solução do problema (P).

Teorema 4.2. *Seja $(\bar{x}, \bar{u}, \bar{v})$ um minimizador fraco para (P). Defina*

$$H(t, x, p, q, r, u, v) := p \cdot f(t, x, u, v) + q \cdot b(t, x, u, v) + r \cdot g(t, x, u, v).^{1}$$

Se, para algum $\epsilon > 0$, (H1)-(H6) e (H) estão satisfeitas, então existem $p \in W^{1,1}(T; \mathbb{R}^n)$, $q \in L^1(T; \mathbb{R}^{m_b})$, $r \in L^1(T; \mathbb{R}^{m_g})$, $\zeta \in L^1(T; R^{k_v})$ e $\lambda \geqslant 0$ tais que*

(i) $\|p\|_\infty + \lambda \neq 0$,

(ii) $(-\dot{p}(t), 0, \zeta(t)) \in co\partial_{x,u,v} H(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t), \bar{v}(t))$ *q.t.* $t \in T$,

(iii) $\zeta(t) \in coN_{V(t)}(\bar{v}(t))$ *q.t.* $t \in T$,

(iv) $r(t) \cdot g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0$ *e* $r(t) \leqslant 0$ *q.t.* $t \in T$,

(v) $(p(0), -p(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \partial l(\bar{x}(0), \bar{x}(1))$.

Além disso, para alguma função integrável K_m, $|(q(t), r(t))| \leqslant K_m(t)|p(t)|$ q.t. $t \in T$.

Demonstração. Ver [20]. □

Observação 4.2. *(a) A condição (iv) é conhecida como condição de folgas complementares.*

(b) O teorema acima tem uma importância fundamental na demonstração do resultado principal deste trabalho, as condições necessárias para o problema (P) com as restrições satisfazendo a condição do tipo Mangasarian-Fromovitz.

4.3.1.2 Mangasarian-Fromovitz

O principal foco deste trabalho é ver as condições necessárias para problemas de controle ótimo com restrições mistas satisfazendo a hipótese do tipo Mangasarian-Fromovitz e é isso que veremos daqui para frente.

Primeiramente vamos ver do que se trata a condição do tipo Mangasarian-Fromovitz, que chamaremos de (H7):

(H7) Existem constantes $K_1, K_2 > 0$ e funções $h \in L^\infty(T; \mathbb{R}^{k_u})$ e $a \in L^\infty(T; \mathbb{R}^{m_g})$ com $|h(t)| = 1$ *q.t.* $t \in T$ tais que, para quase todo $t \in T$,

[1] Hamiltoniano não maximizado.

(i) $a_i(t) > K_2$ para $i \in \mathcal{I}_a(t)$,

(ii) $\nabla_u \bar{g}^{\mathcal{I}_a(t)}(t) \cdot h(t) = a^{\mathcal{I}_a(t)}(t)$,

(iii) $\nabla_u \bar{b}(t) \cdot h(t) = 0$,

(iv) $\det \nabla_u \bar{b}(t) \nabla_u \bar{b}(t)^T \geqslant K_1$.

Observação 4.3. *Dizemos condição do tipo Mangasarian-Fromovitz, pois a condição acima é equivalente a condição de Mangasarian-Fromovitz clássica dos problemas de dimensão finita.*

A hipótese (H7) é uma condição do tipo Mangasarian-Fromovitz uniforme. Esta hipótese é mais fraca que a condição de posto completo uniforme quando o problema tem restrições de igualdade e desigualdade ou só de desigualdade, quando o problema tem apenas restrições de igualdade ela se resume a condição de posto completo. Depois veremos exemplos em que fica claro que a condição do tipo Mangasarian-Fromovitz é mais fraca do que posto completo.

Antes do resultado principal, que é o resultado que dá as condições necessárias para (P) com as restrições satisfazendo a condição do tipo Mangasarian-Fromovitz, vamos ver as condições necessárias para um problema (Q) apenas com restrições de desigualdade.

$$\begin{array}{ll} \text{minimizar} & l(x(0), x(1)) \\ \text{sujeito a} & \dot{x}(t) = f(t, x(t), u(t), v(t)) \ q.t. \ t \in T, \\ & g(t, x(t), u(t), v(t)) \leqslant 0 \ q.t. \ t \in T, \\ & v(t) \in V(t) \ q.t. \ t \in T, \\ & (x(0), x(1)) \in C. \end{array} \tag{Q}$$

Note que, para este caso, as hipóteses (H7)(iii) e (H7)(iv) são ignoradas, pois se tratam de hipóteses assumidas para restrições de igualdade. O resultado que segue será crucial na demonstração.

Teorema 4.3 (Princípio Variacional de Ekeland)**.** *Tome um espaço métrico completo* $(X, d(\cdot,\cdot))$*, uma função semicontínua inferior* $f : X \to \mathbb{R} \cup \{\infty\}$*, um ponto* $x_0 \in \operatorname{dom} f$ *e números* $\alpha > 0$ *e* $\lambda > 0$*. Assuma que*

$$f(x_0) \leqslant \inf_{x \in X} f(x) + \lambda\alpha.$$

Então, existe $\bar{x} \in X$ *tal que*

(i) $f(\bar{x}) \leqslant f(x_0)$,

(ii) $d(x_0, \bar{x}) \leqslant \lambda$,

(iii) $f(\bar{x}) \leqslant f(x) + \alpha d(x, \bar{x})$ *para todo* $x \in X$.

Demonstração. Ver [29] □

Observação 4.4. *A demonstração do teorema abaixo é longa, por isso alguns detalhes serão feitos no apêndice.*

Teorema 4.4. *Seja $(\bar{x}, \bar{u}, \bar{v})$ um minimizador fraco para (Q) com $m_g \geqslant 1$. Defina*

$$H(t, x, p, r, u, v) := p \cdot f(t, x, u, v) + r \cdot g(t, x, u, v).$$

Se, para algum $\epsilon > 0$, as hipóteses (H1)-(H7) são satisfeitas, então existem $p \in W^{1,1}(T; \mathbb{R}^n), r \in L^1(T; \mathbb{R}^{m_g}), \zeta \in L^1(T; R^{k_v})$ e $\lambda \geqslant 0$ tais que

(i) $\|p\|_\infty + \lambda \neq 0$,

(ii) $(-\dot{p}(t), 0, \zeta(t)) \in co\partial_{x,u,v} H(t, \bar{x}(t), p(t), r(t), \bar{u}(t), \bar{v}(t))$ *q.t.* $t \in T$,

(iii) $\zeta(t) \in coN_{V(t)}(\bar{v}(t))$ *q.t.* $t \in T$,

(iv) $r(t) \cdot g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0$ *e* $r(t) \leqslant 0$ *q.t.* $t \in T$,

(v) $(p(0), -p(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \partial l(\bar{x}(0), \bar{x}(1))$.

Além disso, para alguma função integrável R, $|r(t)| \leqslant R(t)|p(t)|$ q.t. $t \in T$.

Demonstração. A demonstração consiste em definir uma sequência de problemas de otimização (R_k) e aplicar o Princípio Variacional de Ekeland para tal sequência, obtendo-se uma sequência de problemas de controle ótimo com restrições mistas, que satisfazem a condição de posto completo, de modo que o Teorema 4.2 se aplica. O limite da solução desta sequência de problemas de controle ótimo é solução do problema (Q), por isso, tomando limites das condições dadas pelo Teorema 4.2, obtemos as condições necessárias para (Q) satisfazendo (ii)-(v) do Teorema 4.4. Finalmente, usando (H7), provamos que a condição de não trivialidade (i) também é verificada. Procedemos em seis passos.

Passo 1. Introdução de novos conjuntos e funções. É conveniente introduzirmos funções auxiliares $\bar{\alpha}, \bar{z}, \bar{w} : T \to \mathbb{R}^{m_g}$ tais que $\bar{\alpha}(t) = \bar{z}(t) = \bar{w}(t) = 0$ *q.t.* $t \in T$. Também defina

$$U_\epsilon(t) := \bar{u}(t) + \epsilon B, \; V_\epsilon(t) := V(t) \cap (\bar{v}(t) + \epsilon B), \; \mathcal{A}_\epsilon(t) := \bar{\alpha}(t) + \epsilon B,$$

e considere as funções

$$F_z(w, \alpha) := w - \alpha, \; F_w(\alpha) := \alpha^2, \; G(t, x, z, u, v, \alpha) := g(t, x, u, v) + z + \alpha$$

que tomam valores em $\mathbb{R}^{m_g}$. Componente a componente, estas funções são definidas como

$$F_{z_i}(w, \alpha) := w_i - \alpha_i, \; F_{w_i}(\alpha) := \alpha_i^2, \; G_i(t, x, z, u, v, \alpha) := g_i(t, x, u, v) + z_i + \alpha_i$$

para todo $i \in \{1, \ldots, m_g\}$. Sob as hipóteses, F_z, F_w e G são mensuráveis com relação a t e Lipschitz com relação às variáveis restantes próximo a

$$(\bar{x}(t), \bar{z}(t), \bar{w}(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t)).$$

Também existem funções integráveis C_f, C_{F_z}, C_{F_w} e C_G tais que

$$|f(t, x, u, v)| \leqslant C_f(t),\ |F_z(w, \alpha)| \leqslant C_{F_z}(t),\ |F_w(\alpha)| \leqslant C_{F_w}(t),\ |G(t, x, z, u, v, \alpha)| \leqslant C_G(t)$$

para todo $(x, z, w, u, v, \alpha) \in (\bar{x}(t), \bar{z}(t), \bar{w}(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t)) + \epsilon B$ *q.t.* $t \in T$.

Passo 2. Definição de uma sequência de problemas de otimização e verificação de que o Princípio Variacional de Ekeland se aplica para tal sequência. Defina W como o conjunto de todas as funções mensuráveis (u, v, α) e todos os vetores $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$ tais que, para quase todo $t \in T, (u(t), v(t), \alpha(t)) \in U_\epsilon(t) \times V_\epsilon(t) \times \mathcal{A}_\epsilon(t)$ e $(a, b) \in C$ e para o qual existem funções absolutamente contínuas x, y, z e w tais que

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t), v(t)),\ \dot{y}(t) = 0,\ \dot{z}(t) = F_z(w(t), \alpha(t)),\ \dot{w}(t) = F_w(\alpha(t))\ q.s., \\ G(t, x(t), z(t), u(t), v(t), \alpha(t)) \leqslant 0\ q.s., \\ (x(t), y(t), z(t), w(t)) \in (\bar{x}(t), \bar{x}(1), \bar{z}(t), \bar{w}(t)) + \epsilon B\ q.s., \\ (x(0), y(0), z(0), w(0)) = (a, b, 0, 0). \end{cases}$$

Aqui y, z e w são variáveis de estado adicionais e α um controle adicional.

Seja $\{\epsilon_k\}_{k\in\mathbb{N}}$ uma sequência de escalares positivos tal que $\epsilon_k \to 0$ quando $k \to \infty$, e defina

$$\Psi_k(x, y, x', y', z, w) := \max\{l(x, y) - l(\bar{x}(0), \bar{x}(1)) + \epsilon_k^2, \epsilon_k^2|w|, |w - z|, |x' - y'|\}.$$

Para simplificar a notação defina $E = (a, b) \in \mathbb{R}^{2n}$. Seja $|E - E'| = |a - a'| + |b - b'|$ e

$$\nu((u, v, \alpha), (u', v', \alpha')) := \int_0^1 |u(t) - u'(t)|dt + \int_0^1 |v(t) - v'(t)|dt + \int_0^1 |\alpha(t) - \alpha'(t)|dt.$$

Defina $\delta_W : W \times W \to \mathbb{R}$ por

$$\delta_W((u, v, \alpha, E), (u', v', \alpha', E')) = \nu((u, v, \alpha), (u', v', \alpha')) + |E - E'|.$$

Considere a sequência de problemas de otimização

$$\min J_k(u, v, \alpha, E) \text{ s.a } (u, v, \alpha, E) \in W, \tag{R_k}$$

onde

$$J_k(u, v, \alpha, E) = \Psi_k(x(0), y(0), x(1), y(1), z(1), w(1)).$$

Seja $\bar{E} = (\bar{x}(0), \bar{x}(1))$. Como $(\bar{u}, \bar{v}, \bar{\alpha}, \bar{E}) \in W$, W é não vazio. Além disso, δ_W define uma métrica em W, o conjunto W é um espaço métrico completo em relação a esta métrica e a função $(u, v, \alpha, E) \mapsto J_k(u, v, \alpha, E)$ é contínua em (W, δ_W). Agora, para todo $k \in \mathbb{N}$, temos

$$J_k(\bar{u}, \bar{v}, \bar{\alpha}, \bar{E}) = \Psi(\bar{x}(0), \bar{x}(1), \bar{x}(1), \bar{x}(1), \bar{z}(1), \bar{w}(1)) = \epsilon_k^2.$$

Como, para todo $k \in \mathbb{N}$, $J_k(u,v,\alpha,E) \geqslant 0$, $(\bar{u},\bar{v},\bar{\alpha},\bar{E})$ é um "ϵ_k^2-minimizador" para (R_k). Pelo Princípio Variacional de Ekeland, existe uma sequência $(u_k,v_k,\alpha_k,E_k) \in W$ tal que, para cada $k \in \mathbb{N}$,

$$\delta_W((u_k,v_k,\alpha_k,E_k),(\bar{u},\bar{v},\bar{\alpha},\bar{E})) \leqslant \epsilon_k \tag{4.2}$$

e (u_k,v_k,α_k,E_k) minimiza a função custo perturbada

$$J_k(u,v,\alpha,E) + \epsilon_k \delta_W((u_k,v_k,\alpha_k,E_k),(u,v,\alpha,E))$$

para todo $(u,v,\alpha,E) \in W$.

Passo 3. Derivação de uma sequência de problemas de controle ótimo padrão reescrevendo as conclusões do Teorema de Ekeland em termos teóricos de controle. Seja (x_k,y_k,z_k,w_k) a trajetória correspondente a (u_k,v_k,α_k,E_k). Para cada $k \in \mathbb{N}$, o processo

$$(x_k,y_k,z_k,w_k,\omega_1,\omega_2,\omega_3,u_k,v_k,\alpha_k) \text{ com } (\omega_1,\omega_2,\omega_3) \equiv (0,0,0)$$

resolve o problema de controle (C_k) posto a seguir:

$$\begin{cases} \text{minimizar} & \Phi_k(\gamma(0),\gamma(1)) \\ \text{sujeito a} & \dot{x}(t) = f(t,x(t),u(t),v(t)),\ \dot{y}(t) = 0,\ \dot{z}(t) = F_z(w(t),\alpha(t)),\ \dot{w}(t) = F_w(\alpha(t)), \\ & \dot{\omega}_1 = |u(t) - u_k(t)|,\ \dot{\omega}_2 = |v(t) - v_k(t)|,\ \dot{\omega}_3 = |\alpha(t) - \alpha_k(t)|, \\ & G(t,x(t),z(t),u(t),v(t),\alpha(t)) \leqslant 0, \\ & (x(t),y(t),z(t),w(t)) \in (\bar{x}(t),\bar{x}(1),\bar{z}(t),\bar{w}(t)) + \epsilon B, \\ & (u(t),v(t),\alpha(t)) \in U_\epsilon(t) \times V_\epsilon(t) \times \mathcal{A}_\epsilon(t), \\ & (x(0),y(0),z(0),w(0)) \in C \times \{(0,0)\}, (\omega_1(0),\omega_2(0),\omega_3(0)) = (0,0,0), \end{cases}$$

onde

$$\begin{aligned} \Phi_k(\gamma(0),\gamma(1)) :=& \Psi_k(x(0),y(0),x(1),y(1),z(1),w(1)) \\ &+ \epsilon_k|x(0) - x_k(0)| + \epsilon_k|y(0) - y_k(0)| + \sum_{i=1}^{3} \epsilon_k \omega_i(1), \end{aligned}$$

$\gamma(t) = (x(t),y(t),z(t),w(t),\omega_1(t),\omega_2(t),\omega_3(t))$ e todas as relações, exceto as últimas duas, sendo entendidas no sentido *q.s.*. Como $\epsilon_k \downarrow 0$, podemos arranjar por extração de subsequência, se necessário, que $\sum \epsilon_k < \infty$. Por (4.2) deduzimos que $(u_k,v_k,\alpha_k) \to (\bar{u},\bar{v},\bar{\alpha})$ fortemente em L^1 e $E_k \to (\bar{x}(0),\bar{x}(1))$. Com a extração de subsequência temos $(u_k(t),v_k(t),\alpha_k(t)) \to (\bar{u}(t),\bar{v}(t),\bar{\alpha}(t))$ *q.s.*[2]. Deduzimos então que $(x_k,y_k,z_k,w_k) \to (\bar{x},\bar{x}(1),\bar{z},\bar{w})$ uniformemente. Descartando termos iniciais da sequência, se necessário, temos

$$(x_k(t),y_k(t),z_k(t),w_k(t)) \in (\bar{x}(t),\bar{x}(1),\bar{z}(t),\bar{w}(t)) + (\epsilon/2)B\ \forall k.$$

[2] Ver resultado na p. 61 no apêndice.

Então $S_k := (x_k, y_k, z_k, w_k, 0, 0, 0, u_k, v_k, \alpha_k)$ é um minimizador fraco de $(\tilde{C}_k)$ (veja a demonstração na p. 61 no apêndice), uma variante de (C_k) obtida tirando a seguinte restrição de estado

$$(x(t), y(t), z(t), w(t)) \in (\bar{x}(t), \bar{x}(1), \bar{z}(t), \bar{w}(t)) + \epsilon B.$$

Esta restrição é retirada, para que o problema fique do formato adequado para aplicarmos o Teorema 4.2.

Passo 4. Derivação de condições necessárias para $(\tilde{C}_k)$. Seja $\mathcal{R} := \{(\rho_1, \rho_2, \rho_3, \rho_4) \in \mathbb{R}^4 \mid \rho_i \geqslant 0, \rho_1 + \ldots + \rho_4 = 1\}$ e defina $\tilde{H}$ como

$$\tilde{H}(t, x, z, w, p, q_z, q_w, r, u, v, \alpha) := p \cdot f(t, x, u, v) + q_z \cdot F_z(w, \alpha) + q_w \cdot F_w(\alpha) + r \cdot G(t, x, z, u, v, \alpha).$$

Lema 4.1. *Existem escalares $\lambda_k, \rho_{1_k}, \rho_{2_k}, \rho_{3_k}$ e ρ_{4_k}, vetores $e_{2_k}, e_{3_k} \in \mathbb{R}^{m_g}$ e $e_{4_k} \in \mathbb{R}^n$, funções integráveis $\zeta_k : T \to \mathbb{R}^{k_v}$ e $r_k : T \to \mathbb{R}^{m_g}$ e funções absolutamente contínuas p_k, q_{z_k} e q_{w_k}, tais que*

(a) $\lambda_k \rho_{1_k} + \|p_k\|_\infty + \|q_{z_k}\|_\infty + \|q_{w_k}\|_\infty + 3\lambda_k \epsilon_k = 1,$

(b) $\lambda_k \geqslant 0,\ |e_{i_k}| = 1$ *para* $i = 2, 3, 4$ *e* $\rho_k = (\rho_{1_k}, \rho_{2_k}, \rho_{3_k}, \rho_{4_k}) \in \mathcal{R},$

(c) $p_k(1) = -\lambda_k \rho_{4_k} e_{4_k},\ q_{w_k}(1) = -q_{z_k}(1) + \lambda_k \rho_{2_k} \epsilon_k^2 e_{2_k},\ q_{z_k}(1) = -\lambda_k \rho_{3_k} e_{3_k},$

(d) $(p_k(0), -p_k(1)) \in N_C(x_k(0), y_k(0)) + \lambda_k \rho_{1_k} \partial l(x_k(0), y_k(0)) + \lambda_k \epsilon_k (B \times B),$

(e) $\zeta_k(t) \in co N_{V(t)}(v_k(t))$ *q.s.,*

(f) $$\begin{aligned}&(-\dot{p}_k(t), -\dot{q}_{z_k}(t), -\dot{q}_{w_k}(t), 0, \zeta_k(t), 0)\\ &\quad \in co\partial \tilde{H}(t, x_k(t), z_k(t), w_k(t), p_k(t), q_{z_k}(t), q_{w_k}(t), r_k(t), u_k(t), v_k(t), \alpha_k(t))\\ &\qquad + \lambda_k \epsilon_k (\{(0,0,0)\} \times B \times B \times B) \text{ q.s.,}\end{aligned}$$

onde $\partial \tilde{H}$ refere-se ao subdiferencial de $\tilde{H}$ nas variáveis (x, z, w, u, v, α),

(g) $r_k(t) \cdot G(t, x_k(t), z_k(t), u_k(t), v_k(t), \alpha_k(t)) = 0$ *e* $r_k(t) \leqslant 0$ *q.s..*

Demonstração. Ver Apêndice A. □

Passo 5. Derivação de condições necessárias para (Q) considerando $\epsilon_k \to 0$ e tomando limites. As sequências $\{e_{2_k}\}, \{e_{3_k}\}, \{e_{4_k}\}$ e $\{\rho_k\}$ são uniformemente limitadas pelo Lema 4.1(b). Do Lema 4.1(a), as sequências $\{\lambda_k\}, \{\|p_k\|_\infty\}, \{\|q_{z_k}\|_\infty\}$ e $\{\|q_{w_k}\|_\infty\}$ são também uniformemente limitadas. Podemos portanto arranjar, por extração de subsequência, se necessário, que

$$e_{2_k} \to e_2, \quad e_{3_k} \to e_3, \quad e_{4_k} \to e_4, \quad \lambda_k \to \hat{\lambda} \quad \text{e} \quad \rho_k \to \rho = (\rho_1, \rho_2, \rho_3, \rho_4),$$

onde $|e_2| = |e_3| = |e_4| = 1, \hat{\lambda} \geqslant 0$ e $\rho \in \mathcal{R}$.

Agora, usando teoremas de seleção mensurável e levando em consideração as propriedades de diferenciabilidade de g, F_z e F_w, deduzimos, pelo Lema 4.1(f), a existência de uma função integrável K_r tal que $|r_k(t)| \leqslant K_r(t)$ $q.t.$ $t \in T$. Também deduzimos que as sequências $\{\dot{p}_k\}, \{\dot{q}_{z_k}\}$ e $\{\dot{q}_{w_k}\}$ são uniformemente integravelmente limitadas e existe uma função integrável K_ζ tal que $|\zeta_k(t)| \leqslant K_\zeta(t)$ $q.t.$ $t \in T$. (Os detalhamentos estão no apêndice p. 65).

É conveniente introduzir a seguinte versão escalada $\tilde{r}_k$ de r_k, dada por $\tilde{r}_k(t) = (1 + K_r(t))^{-1} r_k(t)$. Note que a sequência $\{\|\tilde{r}_k\|_\infty\}$ é uniformemente limitada e $\{t \mapsto \int_0^t \zeta_k ds\}$ é equicontínua e uniformemente limitada. Por extração de subsequências[3] temos que, para funções absolutamente contínuas p, q_z, q_w, uma função integrável ζ e $\tilde{r} \in L^\infty$,

$$p_k \to p, \quad q_{z_k} \to q_z, \quad q_{w_k} \to q_w, \quad \int_0^t \zeta_k ds \to \int_0^t \zeta ds, \text{ uniformemente,}$$

$$\dot{p}_k \to \dot{p}, \quad \dot{q}_{z_k} \to \dot{q}_z, \quad \dot{q}_{w_k} \to \dot{q}_w, \quad \zeta_k \to \zeta, \text{ fracamente em } L^1$$

$$\text{e } \tilde{r}_k \to \tilde{r} \text{ fracamente* em } L^\infty.$$

Levando em consideração o Lema 4.1(g) deduzimos que, para qualquer conjunto mensurável $B \subset T$,

$$0 = \int_B \tilde{r}_k(t) \cdot G_k(t) dt = \int_B \tilde{r}_k(t) \cdot \{G_k(t) - \bar{G}(t)\} dt + \int_B \tilde{r}_k(t) \cdot \bar{G}(t) dt$$

e $\int_B \tilde{r}_k(t) dt \leqslant 0$, onde

$$G_k(t) = G(t, x_k(t), z_k(t), u_k(t), v_k(t), \alpha_k(t)), \quad \bar{G}(t) = G(t, \bar{x}(t), \bar{z}(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t)).$$

Tomando limites, concluímos que $\tilde{r}_k(t) \leqslant 0$ e $\tilde{r}(t) \cdot \bar{G}(t) = 0$ $q.t.$ $t \in T$. Também, por uma modificação simples na demonstração do Teorema 3.1.7 [28] e por propriedades de semicontinuidade superior de cones normal e subdiferenciais limite, podemos passar o limite de (a)-(g) do Lema 4.1. O que resulta em

(A') $\lambda + \|p\|_\infty + \|q_z\|_\infty + \|q_w\|_\infty > 0$,

(B') $(-\dot{p}(t), -\dot{q}_z(t), -\dot{q}_w(t), 0, \zeta(t), 0)$
$\in co\partial \tilde{H}(t, \bar{x}(t), \bar{z}(t), \bar{w}(t), p(t), q_z(t), q_w(t), r(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t))$ $q.s.$,

(C') $\zeta(t) \in coN_{V(t)}(\bar{v}(t))$ $q.s.$,

(D') $(p(0), -p(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda\partial l(\bar{x}(0), \bar{x}(1))$,

[3] Veja estes resultados no apêndice p. 67

(E') $|q_w(1)| = |q_z(1)|, r(t) \leqslant 0$ e $r(t) \cdot G(t, \bar{x}(t), \bar{z}(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t)) = 0$ $q.s.$,
onde $r(t) = (1 + K_r(t))\tilde{r}(t)$.

Passo 6. Reescrevendo as relações (A')-(E') na forma requerida. Relembre que
$\tilde{H}(t, x, z, w, p, q_z, q_w, r, u, v, \alpha) = p \cdot f(t, x, u, v) + q_z \cdot (w - \alpha) + q_w \cdot \alpha^2 + r \cdot (g(t, x, u, v) + z + \alpha)$.
Agora vamos estimar o $co\partial\tilde{H}$. Temos

$$\begin{aligned} co\partial\tilde{H}(t, x, z, w, p, q_z, q_w, r, u, v, \alpha) \subset \{(&\theta_1 + r\nabla_x g, r, q_z, \theta_2 + r\nabla_u g, \\ &\theta_3 + r\nabla_v g, -q_z + 2q_w\alpha + r) \mid (\theta_1, \theta_2, \theta_3) \in co\partial_{x,u,v} p \cdot f\}. \end{aligned}$$

Como $\bar{\alpha}(t) = 0$ e $\bar{w}(t) = 0$ $q.t.$ $t \in T$, por um teorema de seleção apropriado, deduzimos a existência de funções mensuráveis $\theta_1, \theta_2, \theta_3$ e ζ satisfazendo

$$\begin{aligned} &(\theta_1(t), \theta_2(t), \theta_3(t)) \in co\partial_{x,u,v} p(t) \cdot f(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) \ q.t.\ t \in T, \\ &\zeta(t) \in coN_{V(t)}(\bar{v}(t)) \ q.t.\ t \in T. \end{aligned} \tag{4.3}$$

Concluímos de cima que

$$\begin{aligned} (-\dot{p}(t), -\dot{q}_z(t), -\dot{q}_w(t), 0, \zeta(t), 0) = (&\theta_1(t) + r(t)\nabla_x \bar{g}(t), r(t), q_z(t), \\ &\theta_2(t) + r(t)\nabla_u \bar{g}(t), \theta_3(t) + r(t)\nabla_v \bar{g}(t), r(t) - q_z(t)). \end{aligned} \tag{4.4}$$

De (4.4) e (E') deduzimos que, para quase todo $t \in T$,

$$q_z(t) = r(t), \quad \dot{q}_z(t) = -r(t), \quad \dot{q}_w(t) = -q_z(t) \tag{4.5}$$

e $|q_w(1)| = |q_z(1)|$. Defina

$$H(t, x, p, r, u, v) := p \cdot f(t, x, u, v) + r \cdot g(t, x, u, v).$$

De cima e (4.4) temos

$$(-\dot{p}(t), 0, \zeta(t)) = (\theta_1(t) + r(t)\nabla_x \bar{g}(t), \theta_2(t) + r(t)\nabla_u \bar{g}(t), \theta_3(t) + r(t)\nabla_v \bar{g}(t)) \tag{4.6}$$

para algum $(\theta_1(t), \theta_2(t), \theta_3(t)) \in co\partial_{x,u,v} p(t) \cdot f(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$ $q.t.$ $t \in T$. Segue que

$$(-\dot{p}(t), 0, \zeta(t)) \in co\partial_{x,u,v} H(t, \bar{x}(t), p(t), r(t), \bar{u}(t), \bar{v}(t)) \ q.t.\ t \in T. \tag{4.7}$$

Levando (E') em consideração, também temos

$$r(t) \cdot g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0 \text{ e } r(t) \leqslant 0 \ q.t.\ t \in T. \tag{4.8}$$

Deduzimos que (4.7), (C'), (4.8) e (D') são, respectivamente, (ii), (iii), (iv) e (v) do teorema.

Resta provar (i). Afirmamos que

$$\lambda + |p(t)| > 0 \quad \forall\ t \in T. \tag{4.9}$$

Primeiro observe que se $q_z \equiv 0$ e $q_w \equiv 0$, então (4.9) vale. Buscando uma contradição, assuma que $\lambda = 0$ e $p(\tau) = 0$ para algum $\tau \in T$. Existem dois casos a serem considerados.

Caso 1. Suponha que $r(t) = 0$ $q.s.$. Então por (4.5) temos $q_z \equiv 0, q_w \equiv 0$. Deduzimos de (4.6) e (H1) a existência de K_1 integrável com $K_1(t) \geqslant 0$ $q.s.$ tal que $|\dot{p}(t)| \leqslant K_1(t)|p(t)|$ $q.s.$ a Desigualdade de Gronwall e o fato que $p(\tau) = 0$ para algum $\tau \in T$ implicam que $p \equiv 0$ (veja os detalhamentos na p. 68 no apêndice). Mas isto, junto com o fato que $\lambda = 0$ e $q_z \equiv 0, q_w \equiv 0$ e $p \equiv 0$, contradiz (A').

Caso 2. Suponha agora que $r(t) \neq 0$ em um subconjunto de T de medida de Lebesgue diferente de zero. Deduzimos de (4.6) que $0 = \theta_2(t) + r(t) \cdot \nabla_u \bar{g}(t)$ $q.t.$ $t \in T$, onde $\theta_2(t)$ é definido em (4.3). Tomando o produto interno do lado esquerdo da equação anterior e o vetor h (definido em (H7)) obtemos, de (H7) e (iv) do teorema, que $\theta_2 \cdot h(t) = -r(t) \cdot a(t)$ $q.t.$ $t \in T$. Pelas hipóteses, existe K_2 integrável com $K_2(t) \geqslant 0$ $q.s.$ tal que

$$|r(t)| \leqslant K_2(t)|p(t)| \quad q.t.\ t \in T. \tag{4.10}$$

Deduzimos de (4.6), (4.10) e (H1) a existência de K_3 integrável com $K_3(t) \geqslant 0$ $q.t.$ tal que $|\dot{p}(t)| \leqslant K_3(t)|p(t)|$ $q.t.$ $t \in T$. Novamente pela Desigualdade de Gronwall e o fato que $p(\tau) = 0$ para algum $\tau \in T$ implica que $p \equiv 0$. Mas então, por (4.10), temos $r(t) = 0$ $q.t.$, uma contradição.

Provamos que $\lambda + |p(t)| > 0\ \forall\, t \in T$, uma condição que assegura que $\lambda + \|p\|_\infty > 0$. A demonstração está completa. □

Agora sim, estamos quase prontos para enunciar o principal resultado do trabalho, mas antes disso será enunciado o Teorema da Função Implícita Uniforme, que desempenha um papel crucial na demonstração das condições necessárias para o problema (P) com as restrições satisfazendo a condição do tipo Mangasarian-Fromovitz.

Teorema 4.5 (Teorema da Função Implícita Uniforme)**.** *Considere um conjunto $A \subset \mathbb{R}^k$, um número $a > 0$, uma família de funções $\{\psi_a : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n\}_{a \in A}$, e um ponto $(u_0, v_0) \in \mathbb{R}^m \times \mathbb{R}^n$ tal que $\psi_a(u_0, v_0) = 0$ para todo $a \in A$. Suponha que*

(i) ψ_a *é continuamente diferenciável em* $(u_0, v_0) + \alpha B$ *para todo* $a \in A$*,*

(ii) existe uma função monótona crescente $\theta : (0, \infty) \to (0, \infty)$ *com* $\theta(s) \downarrow 0$ *quando* $s \downarrow 0$ *tal que, para todos* $a \in A$ *e* $(u, v) \neq (u', v') \in (u_0, v_0) + \alpha B$*,*

$$|\nabla \psi_a(u, v) - \nabla \psi_a(u', v')| \leqslant \theta(|(u, v) - (u', v')|),$$

(iii) $\nabla_v \psi_a(u_0, v_0)$ *é não singular para cada* $a \in A$ *e existe* $c > 0$ *tal que, para todo* $a \in A$, $|[\nabla_v \psi_a(u_0, v_0)]^{-1}| \leqslant c$*.*

Então existem $\delta \geqslant 0$ e uma família de funções continuamente diferenciáveis $\{\phi_a : u_0+\delta B \to v_0 + \alpha B\}_{a\in A}$ que são Lipschitz com constante de Lipschitz k comum tais que, para todo $a \in A$,

(a) $v_0 = \phi_a(u_0)$,

(b) $\psi_a(u, \phi_a(u)) = 0$ *para todo* $u \in u_0 + \delta B$,

(c) $\nabla_u \phi(u_0) = -[\nabla_v \psi_a(u_0, v_0)]^{-1} \nabla_u \psi_a(u_0, v_0)$.

Os números δ e k dependem de θ, c e α somente. Além disso, se A é um conjunto Borel e $a \mapsto \psi_a(u, v)$ é uma função mensurável Borel para cada $(u, v) \in (u_0, v_0) + \alpha B$, então $a \mapsto \phi_a(u)$ é uma função mensurável Borel para cada $u \in u_0 + \delta B$.

O teorema que segue é o principal resultado deste trabalho, a técnica utilizada na demonstração consistirá em montar um problema auxiliar a partir do problema (P) apenas com restrições de desigualdade e para isso usaremos o Teorema da Função Implícita Uniforme, que "embute" as restrições de igualdade de (P) neste problema auxiliar e em seguida utilizamos o Teorema 4.4 para concluir as condições necessárias para (P).

Teorema 4.6. *O Teorema 4.2 permanece válido se substituímos (H*) por (H7).*

Demonstração. Considere a função $\mu : T \times (\mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v}) \times \mathbb{R}^{m_b} \to \mathbb{R}^{m_b}$ dada por

$$\mu(t, (\xi, u, v), \eta) := b(t, \bar{x}(t) + \xi, \bar{u}(t) + u + \nabla_u \bar{b}(t)^\top \eta, \bar{v}(t) + v).$$

Temos $\mu(t, (0, 0, 0), 0) = 0$ $q.t.$ $t \in T$. Escolha $S_0 \subset T$ sendo o maior subconjunto tal que esta relação e cada uma das hipóteses não valem para cada $t \in S_0$. Por hipóteses, S_0 tem medida de Lebesgue nula. Assim, existe um conjunto Borel S_1, que é a interseção de uma coleção enumerável de conjuntos abertos, tal que $S_0 \subset S_1$ e $S_1 \backslash S_0$ tem medida nula. Assim S_1 é um conjunto Borel de medida nula. Definimos $S := T \backslash S_1$, um conjunto Borel de medida total. Temos

$$\frac{\partial \mu}{\partial \eta}(t, (0, 0, 0), 0) = \Gamma(t) := \nabla_u \bar{b}(t) \nabla_u \bar{b}(t)^\top \quad \forall\ t \in S.$$

Por (H6) e (H7)(iv) existe uma constante $M > 0$ tal que $|[\Gamma(t)]^{-1}| \leqslant M$ para todo $t \in S$. Assim o Teorema 4.5 se aplica para a função μ e assegura a existência de $\epsilon_1 \in (0, \epsilon)$, $\delta_1 \in (0, \epsilon)$ e uma aplicação implícita $d : T \times \epsilon_1 B \times \epsilon_1 B \times \epsilon_1 B \to \delta_1 B$ tal que $d(\cdot, \xi, u, v)$ é uma função mensurável para (ξ, u, v) fixo, as funções $\{d(t, \cdot, \cdot, \cdot) \mid t \in S\}$ são Lipschitz com constante Lipschitz $K_d > 0$ comum, e $d(t, \cdot, \cdot, \cdot)$ é continuamente diferenciável para $t \in S$ fixo. Escolha $\sigma_1, \delta > 0$ tais que

$$\sigma_1 \in (0, \min\{\epsilon_1, \epsilon/2\}), \quad \delta \in (0, \min\{\delta_1, \epsilon/2\}), \quad \sigma_1 + K_{b,g}\delta \in (0, \epsilon/2), \tag{4.11}$$

onde ϵ_1 e δ_1 (que não dependem de t) são como acima e $K_{b,g}$ é dado por (H6).

No que segue e sem perda de generalidade, consideramos a função implícita d definida em $T \times \sigma_1 B \times \sigma_1 B \times \sigma_1 B$ e tomando valores em δB. A função $d(t, \cdot, \cdot, \cdot)$ é continuamente diferenciável e Lipschitz em $\sigma_1 B \times \sigma_1 B \times \sigma_1 B$ com K_d a constante de Lipschitz e satisfaz $d(t, 0, 0, 0) = 0$ $q.t.$ $t \in T$:

$$\mu(t, (\xi, u, v), d(t, \xi, u, v)) = 0 \ q.t. \ t \in T, \ \forall \ (\xi, u, v) \in \sigma_1 B \times \sigma_1 B \times \sigma_1 B,$$

$$(d_\xi, d_u, d_v)(t, 0, 0, 0) = -\Gamma(t)^{-1}(\nabla_x \bar{b}(t), \nabla_u \bar{b}(t), \nabla_v \bar{b}(t)) \ q.t. \ t \in T.$$

Defina as funções

$$D(t, x, u, v) := d(t, x - \bar{x}(t), u - \bar{u}(t), v - \bar{v}(t)),$$

$$F(t, x, u, v) := f(t, x, u + \nabla_u \bar{b}(t)^\top D(t, x, u, v), v),$$

$$G(t, x, u, v) := g(t, x, u + \nabla_u \bar{b}(t)^\top D(t, x, u, v), v)$$

e os conjuntos $\mathcal{U}(t) := \bar{u}(t) + \sigma_1 B, \mathcal{V}(t) := V(t) \cap (\bar{v}(t) + \sigma_1 B)$. Considere o problema

$$\begin{array}{ll} \text{minimizar} & l(x(0), x(1)) \\ \text{sujeito a} & \dot{x}(t) = F(t, x(t), u(t), v(t)) \ q.t. \ t \in T, \\ & G(t, x(t), u(t), v(t)) \leqslant 0 \ q.t. \ t \in T, \\ & (u(t), v(t)) \in \mathcal{U}(t) \times \mathcal{V}(t) \ q.t. \ t \in T, \\ & (x(0), x(1)) \in C. \end{array} \tag{P_{aux}}$$

O processo $(\bar{x}, \bar{u}, \bar{v})$ é admissível para (P_{aux}) (ver detalhamentos no apêndice p. 68). Assuma que $(\tilde{x}, \tilde{u}, \tilde{v})$ é uma solução com custo menor. Defina

$$\hat{u}(t) := \tilde{u}(t) + \nabla_u \bar{b}(t)^\top D(t, \tilde{x}(t), \tilde{u}(t), \tilde{v}(t)),$$

$$\xi(t) := \tilde{x}(t) - \bar{x}(t), \quad u_1(t) := \tilde{u}(t) - \bar{u}(t), \quad v_1(t) := \tilde{v}(t) - \bar{v}(t).$$

De (4.11) e da definição de d segue que

$$|\hat{u}(t) - \bar{u}(t)| \leqslant |\tilde{u}(t) - \bar{u}(t)| + K_{b,g}\delta \leqslant \sigma_1 + K_{b,g}\delta < \epsilon, \quad |\tilde{v}(t) - \bar{v}(t)| \leqslant \sigma_1 < \epsilon.$$

Pela definição de d, para quase todo $t \in T$ temos

$$\mu(t, (\xi(t), u_1(t), v_1(t)), d(t, \xi(t), u_1(t), v_1(t))) = b(t, \tilde{x}(t), \hat{u}(t), \tilde{v}(t)) = 0.$$

Concluímos que $(\tilde{x}, \hat{u}, \tilde{v})$ é uma solução para (P) com menor custo (veja os detalhes na p. 69 no apêndice), contradizendo a otimalidade de $(\bar{x}, \bar{u}, \bar{v})$. Segue que $(\bar{x}, \bar{u}, \bar{v})$ é um minimizador para (P_{aux}).

Vamos agora checar que (P_{aux}) satisfaz as condições sob as quais o Teorema 4.4 vale. Precisamos somente verificar que

$$\nabla_u \bar{G}^{\mathcal{I}_a(t)}(t) \cdot h(t) = a^{\mathcal{I}_a(t)}(t) \ q.t. \ t \in T, \tag{4.12}$$

onde h e a são as funções cuja existência é postulada em (H7) e satisfaz (H7)(i) e (H7)(ii). Observe que

$$\nabla_u G(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = \nabla_u \bar{g}(t) - \nabla_u \bar{g}(t) \nabla_u \bar{b}(t)^\top \Gamma(t)^{-1} \nabla_u \bar{b}(t).$$

Levando em consideração (H7), o produto interno do lado esquerdo da equação anterior e o vetor $h(t)$ implica em

$$\nabla_u \bar{G}^{\mathcal{I}_a(t)}(t) \cdot h(t) = \nabla_u \bar{g}^{\mathcal{I}_a(t)}(t) \cdot h(t) = a^{\mathcal{I}_a(t)}(t) \ q.t. \ t \in T$$

fornecendo (4.12). Agora aplicamos o Teorema 4.3 para (P_{aux}). Que garante a existência de $\lambda \geqslant 0, p \in W^{1,1}(T; \mathbb{R}^n), r \in L^1(T; \mathbb{R}^{m_g})$ e $\zeta \in L^1(T; \mathbb{R}^{k_v})$ tais que

(i) $\|p\|_\infty + \lambda \neq 0$,

(ii) $(-\dot{p}(t), 0, \zeta(t)) \in co\partial_{x,u,v} \tilde{H}(t, \bar{x}(t), p(t), r(t), \bar{u}(t), \bar{v}(t)) \ q.t. \ t \in T$,

(iii) $\zeta(t) \in coN_{V(t)}(\bar{v}(t)) \ q.t. \ t \in T$,

(iv) $r(t) \cdot G(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0$ e $r(t) \leqslant 0 \ q.t. \ t \in T$,

(v) $(p(0), -p(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \partial l(\bar{x}(0), \bar{x}(1))$,

onde $\tilde{H}(t, x, p, r, u, v) = p \cdot F(t, x, u, v) + r \cdot G(t, x, u, v)$. Da regra da cadeia não suave (Teorema 3.5), das propriedades de diferenciabilidade de d e um teorema de seleção apropriado, existem funções mensuráveis $\theta_1, \theta_2, \theta_3$ e ζ satisfazendo, $q.t. \ t \in T$,

$$(\theta_1(t), \theta_2(t), \theta_3(t)) \in co\partial_{x,u,v} p(t) \cdot f(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)), \quad \zeta(t) \in coN_{V(t)}(\bar{v}(t)),$$

$$\begin{aligned}(-\dot{p}(t), 0, \zeta(t)) = (&\theta_1(t) + q(t)\nabla_x \bar{b}(t) + r(t)\nabla_x \bar{g}(t), \\ &\theta_2(t) + q(t)\nabla_u \bar{b}(t) + r(t)\nabla_u \bar{g}(t), \theta_3(t) + q(t)\nabla_v \bar{b}(t) + r(t)\nabla_v \bar{g}(t)),\end{aligned}$$

onde

$$q(t) = -(\theta_2(t) + r(t)\nabla_u \bar{g}(t))\nabla_u \bar{b}(t)^\top \Gamma(t)^{-1}. \tag{4.13}$$

Sob as hipóteses, $\theta_1, \theta_2, \theta_3, r$ e ζ são todas funções integráveis e assim q o é. Isto prova que λ, p, q, r e ζ satisfazem (i)-(v) do teorema. Como existe R integrável tal que $|r(t)| \leqslant R(t)|p(t)| \ q.t. \ t \in T$, deduzimos de (4.13) a existência de K_m integrável tal que a última conclusão do teorema vale (detalhamento no apêndice p. 69). Isso completa a demonstração.

□

4.3.2 Exemplos

Dos exemplos que seguem, para o primeiro (H7) é satisfeita, mas (H*) não e vale o princípio do máximo, e para o segundo (H7) não é satisfeita e o princípio do máximo não é válido.

Exemplo 4.1. *Considere o problema:*

$$\begin{array}{ll} \min & x(1) \\ \textit{sujeito a} & \dot{x}(t) = u_1^2(t) + u_3^2(t), \\ & u_1(t) + u_2^2(t) + u_3^2(t) \leqslant 0, \\ & u_1(t) - u_2^3(t) \leqslant 0, \\ & x(0) = 0. \end{array}$$

Aqui as variáveis de controle são (u_1, u_2, u_3)*. Este é um problema sem restrições de igualdade. Um minimizador é* $(0,0,0,0)$ *e* $\mathcal{I}_a(t) = \{1,2\}\ \forall\ t \in T$*. O Teorema 4.4 vale com* $p(t) = -1$ *e* $\lambda = 1$*. Para este problema, a matriz*

$$F(t) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

satisfaz (H7), mas não (H): defina* $h(t) = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^\top$*,* $a(t) = \begin{pmatrix} 1 & 1 \end{pmatrix}^\top$ *e* $K_2 = 1/2$*.*

Em situações quando $m_g \neq 0$ e $m_b \neq 0$, é crucial para $h \in L^\infty(T;\mathbb{R}^{k_u})$ satisfazer simultaneamente (H7)(ii) e (H7)(iii). O exemplo que segue, ilustra o fato onde não existe tal h.

Exemplo 4.2. *Considere o problema:*

$$\begin{array}{ll} \min & x(1) \\ \textit{sujeito a} & \dot{x}(t) = u_1(t), \\ & u_2(t) - x^2(t) = 0, \\ & u_2(t) - u_1^3(t) \leqslant 0, \\ & x(0) = 0. \end{array}$$

Um minimizador é $(\bar{x}, \bar{u}_1, \bar{u}_2) \equiv (0,0,0)$*. Neste minimizador a restrição de desigualdade é ativa para todo* $t \in T$*. Para este problema temos*

$$H(t, x, p, q, r, u_1, u_2) = pu_1 + q(u_2 - x^2) + r(u_2 - u_1^3).$$

O Teorema 4.6 não vale para este problema. De fato, aplicando as conclusões deste teorema, temos

$$(-\dot{p}(t), 0, 0) = (0, p(t), q(t) + r(t))$$

e $p(1) = -\lambda$*. Segue que* $p \equiv 0$*. Consequentemente* $\lambda = 0$*, contradizendo a condição de não-trivialidade (i) do Teorema 4.6.*

Neste problema (H7) não é satisfeita. De fato, como

$$\nabla_{u_1,u_2} b(t,0,0,0) = \nabla_{u_1,u_2} g(t,0,0,0) = (0,1) \ \forall \ t \in T,$$

é impossível definir um vetor $h \in \mathbb{R}^2$ para o qual

$$\nabla_{u_1,u_2} b(t,0,0,0) \cdot h = 0 \ \ e \ \nabla_{u_1,u_2} g(t,0,0,0) \cdot h \neq 0.$$

5 Conclusão

Nesta dissertação, estudamos as condições necessárias de otimalidade para o problema de controle ótimo com restrições mistas de igualdade e desigualdade sob a qualificação de restrição do tipo Mangasarian-Fromovitz, bem como as técnicas utilizadas na demonstração. Para o entendimento pleno foram estudados os seguintes pré-requisitos: a teoria de otimização no contexto de dimensão finita, as qualificações de restrições mais clássicas e a condição de KKT para o problema de programação não linear; o problema de cálculo das variações e a Equação de Euler-Lagrange; as construções básica de análise não suave; o problema de controle ótimo de Pontryagin e o princípio do máximo não suave clássico, com a condição do máximo.

O próximo passo deste trabalho é generalizar para o contexto de controle ótimo as qualificações de restrições recentes da literatura que se mostraram ser de grande importância na área de programação não linear (ver Andreani et al. [1], [3]) e foram muito bem recebidas pela comunidade científica, com impacto considerável. A partir desta dissertação a ideia é generalizar a condição de qualificação de posto constante (CRCQ), que é uma hipótese alternativa à condição de qualificação de Mangasarian-Fromovitz (MFCQ), e a condição de qualificação de dependência linear positiva constante (CPLD), que é mais fraca que MFCQ e CRCQ. Em seguida, provar o princípio do máximo assumindo estas hipóteses, CRCQ e CPLD, dessa forma aumentamos a quantidade de problemas para os quais o princípio do máximo é aplicável.

Referências

1 ANDREANI, R.; HAESER, G.; SCHUVERDT, M. L.; SILVA, P. J. S. A relaxed constant positive linear dependence constraint qualification and applications. *Math. Program., Ser. A*, v. 135, p. 255–273, 2012.

2 ANDREANI, R.; MARTíNEZ, J. M.; SCHUVERDT, M. L. On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *J. Optim. Theory Appl.*, v. 125, p. 473–485, 2005.

3 ANDREANI, R.; HAESER, G.; SCHUVERDT, M. L.; SILVA, P. J. S. Two new weak constraint qualifications and applications. *SIAM J. Optim.*, v. 22, p. 1109–1135, 2012.

4 JANIN, R. Directional derivative of the marginal function in nonlinear programming. *Math. Program. Stud.*, v. 21, p. 110–126, 1984.

5 BAZARAA, M.; SHERALI, H. D.; SHETTY, C. M. *Nonlinear Programming - Theory and Algorithms*. New Jersey: John Wiley and Sons, 2006.

6 BERTSEKAS, D. P. *Nonlinear Programming*. Massachusetts: Athena Scientific, 1995.

7 NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. New York: Springer, 2006.

8 HESTENES, M. R. *Calculus of Variations and Optimal Control Theory*. New York: John Wiley, 1966.

9 NEUSTADT, L. W. *Optimization*. Princeton: Princeton University Press, 1976.

10 DUBOVITSKII, A. Y.; MILYUTIN, A. A. Theory of the principle of the maximum. In: ______. *Methods of the theory of extremal problems in economics*. Moscow: Nauka, 1981.

11 DMITRUK, A. V. Maximum principle for the general optimal control problem with phase and regular mixed constraints. *Comput. Math. Model.*, v. 4, p. 364–377, 1993.

12 PáLES, Z.; ZEIDAN, V. First and second order necessary conditions for optimal control problems with constraints. *Trans. Amer. Math. Soc.*, v. 346, p. 421–453, 1994.

13 STEFANI, G.; ZEZZA, P. Optimality conditions for a constrained control problem. *SIAM J. Control Optim.*, v. 34, p. 635–359, 1996.

14 MILYUTIN, A. A.; OSMOLOVSKII, N. P. *Calculus of Variations and Optimal Control*. Providence: American Math. Soc., 1998.

15 DEVDARIANI, E. N.; LEDYAEV, Y. S. Maximum principle for implicit control systems. *Appl. Math. Optim.*, v. 40, p. 79–103, 1999.

16 FRANKOWSKA, H.; RAMPAZZO, F. Relaxation of control systems under state constraints. *SIAM J. Control Optim.*, v. 37, p. 1291–1309, 1999.

17 ARUTYUNOV, A. V. *Optimality Conditions: Abnormal and Degenerate Problems*. Dordrecht: Kluwer Academic Publishers, 2000.

18 DE PINHO, M. R.; VINTER, R. B.; ZHENG, H. A maximum principle for optimal control problems with mixed constraints. *IMA J. Math. Control Inform.*, v. 18, p. 189–205, 2001.

19 DE PINHO, M. R.; ILCHMANN, A. Weak maximum principle for optimal control problems with mixed constraints. *Nonlinear Anal.*, v. 48, p. 1179–1196, 2002.

20 DE PINHO, M. R. Mixed constrained control problems. *J. Math. Anal. Appl.*, v. 278, p. 293–307, 2003.

21 DE PINHO, M. R.; ROSENBLUETH, J. F. Necessary conditions for constrained problems under mangasarian-fromowitz conditions. *SIAM J. Control Optim.*, v. 47, p. 535–552, 2008.

22 CLARKE, F. H.; PINHO, M. R. de. Optimal control problems with mixed constraints. *SIAM J. Control Optim.*, v. 48, p. 4500–4524, 2010.

23 CLARKE, F. H. Nonsmooth analysis in systems and control theory. *Encyclopedia of Complexity and System Science, Springer*, p. 6271–6285, 2009.

24 RIBEIRO, A. A.; KARAS, E. W. *Otimização Contínua: aspectos teóricos e computacionais*. São Paulo: Cengage Learning, 2013.

25 ANDREANI, R.; ECHAGüE, C. E.; SCHUVERDT, M. L. Constant-rank condition and second-order constraint qualification. *J. Optim. Theory Appl.*, v. 146, p. 255–266, 2010.

26 CLARKE, F. H. *Functional Analysis, Calculus of Variations and Optimal Control*. London: Springer-Verlag, 2013.

27 LIBERZON, D. *Calculus of Variations and Optimal Control Theory: a concise introduction*. Illinois: Princeton University Press, 2011.

28 CLARKE, F. H. *Optimization and Nonsmooth Analysis*. New York: Wiley-Interscience, 1983.

29 VINTER, R. *Optimal Control*. Boston: Birkhäuser, 2000.

30 KURDILA, A. J.; ZABARANKIN, M. *Convex Functional Analysis*. Basel: Birkhäuser, 2005.

31 ROCKAFELLAR, R. T. *Convex Analysis*. New Jersey: Princeton University Press, 1970.

32 ROYDEN, H. L.; FITZPATRICK, P. M. *Real Analysis*. Asia: Pearson Education, 2010.

33 LUENBERGER, D. G.; YE, Y. *Linear and Nonlinear Programming*. New York: Springer, 2008.

34 PESCH, H. J.; PLAIL, M. The maximum principle of optimal control: A history of ingenious ideas and missed opportunities. *J. Control and Cybernetics*, v. 38, p. 973–995, 2009.

35 CLARKE, F. H. The maximum principle in optimal control then and now. *J. Control and Cybernetics*, v. 34, p. 709–722, 2005.

36 FOLLAND, G. B. *Real analysis: modern techniques and their applications*. New York: John Wiley, 1999.

37 RUDIN, W. *Principles of mathematical analysis*. New York: McGraw-Hill, 1964.

Apêndices

APÊNDICE A – Definições, resultados e demonstrações utilizados no capítulo 4

Proposição A.1. *[36] Suponha $1 \leqslant p < \infty$. Se $\|f_n - f\|_p \to 0$, então $f_n \to f$ na medida e por isso alguma subsequência converge para f q.s..*

Demonstração que S_k é um minimizador fraco de $(\tilde{C}_k)$

Demonstração. Tomamos um processo factível S arbitrário de $(\tilde{C}_k)$ na vizinhança fraca de S_k, seja

$$S = (x, y, z, w, \omega_1, \omega_2, \omega_3, u, v, \alpha).$$

Sabemos que S_k é mínimo fraco para (C_k), logo se o processo S que tomamos é factível para (C_k) temos S_k mínimo fraco para $(\tilde{C}_k)$. Então, temos que verificar se

$$(x(t), y(t), z(t), w(t)) \in (\bar{x}(t), \bar{x}(1), \bar{z}(t), \bar{w}(t)) + \epsilon B, \tag{A.1}$$

pois esta é a única restrição que tem no problema (C_k) e não tem no problema $(\tilde{C}_k)$.

Como S está na vizinhança fraca de S_k, temos

$$|(x, y, z, w, \omega_1, \omega_2, \omega_3) - (x_k, y_k, z_k, w_k, 0, 0, 0)| < \epsilon/2.$$

Além disso, temos $(x_k, y_k, z_k, w_k) \to (\bar{x}, \bar{x}(1), \bar{z}, \bar{w})$ uniformemente e assim $(x_k(t), y_k(t), z_k(t), w_k(t)) \in (\bar{x}(t), \bar{x}(1), \bar{z}(t), \bar{w}(t)) + \epsilon/2B, \ \forall k$. Logo,

$$|(x(t), y(t), z(t), w(t)) - (\bar{x}(t), \bar{x}(1), \bar{z}(t), \bar{w}(t))| \leqslant |(x(t), y(t), z(t), w(t) - (x_k(t), y_k(t), z_k(t), w_k(t))| \\ + |(x_k(t), y_k(t), z_k(t), w_k(t)) - (\bar{x}(t), \bar{x}(1), \bar{z}(t), \bar{w}(t))| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Portanto, a equação (A.1) está satisfeita, então S_k é um minimizador fraco para $(\tilde{C}_k)$, como queríamos. □

Demonstração do Lema 4.1

Demonstração. [21] Sem perda de generalidade assuma que $|w_k(1)| < 1$ para todo $k \in \mathbb{N}$, isto pode ser feito, pois $w_k \to \bar{w}$ uniformemente e $\bar{w}(t) = 0$ $q.t.$ $t \in T$. Defina

$$\begin{aligned} &h(t, x, y, z, w, \omega_1, \omega_2, \omega_3, p, s, q_z, q_w, r, \Pi, u, v, \alpha) \\ &\quad := p \cdot f(t, x, u, v) + s \cdot 0 + q_z \cdot F_z(w, \alpha) + q_w \cdot F_w(\alpha) + r \cdot G(t, x, z, u, v, \alpha) \\ &\quad\quad + \pi_1 |u - u_k(t)| + \pi_2 |v - v_k(t)| + \pi_3 |\alpha - \alpha_k(t)|, \end{aligned}$$

onde $\Pi = (\pi_1, \pi_2, \pi_3)$. Para simplificar a notação, defina $P_k := (p_k, s_k, q_{z_k}, q_{w_k})$. É uma questão simples ver que

$$\frac{\partial}{\partial \alpha} G(t, x_k(t), z_k(t), u_k(t), v_k(t), \alpha_k(t)) = I \ q.t. \ t \in T,$$

onde I é a matriz identidade. Assim o problema $(\tilde{C}_k)$ satisfaz as condições sob as quais o Teorema 4.2 se aplica. Observando que u_k e α_k tomam valores no interior do conjunto controle apropriado, uma aplicação do Teorema 4.2 dá funções absolutamente contínuas $P_k(t), \Pi_k(t)$, um escalar $\lambda_k \geqslant 0$, e funções integráveis r_k, ζ_k tais que

(A) $\lambda_k + \|p_k\|_\infty + \|s_k\|_\infty + \|q_{z_k}\|_\infty + \|q_{w_k}\|_\infty + \sum_{i=1}^{3} \|\pi_{i_k}\|_\infty > 0,$

(B) $(-\dot{P}_k(t), -\dot{\Pi}_k(t), 0, \zeta_k(t), 0) \in co\partial h_k(t) \ q.s.,$

(C) $\zeta_k(t) \in coN_{V(t)}(v_k(t)) \ q.s.,$

(D) $(P_k(0), \Pi_k(0), -P_k(1), -\Pi_k(1))$
$\in (N_C(x_k(0), y_k(0)) \times \mathbb{R}^{m_g} \times \mathbb{R}^{m_g}) \times \mathbb{R}^3 \times \{(0,0,0,0)\} \times \{(0,0,0)\}$
$+\lambda_k \partial \Phi_k(\gamma_k(0), \gamma_k(1)),$

(E) $r_k(t) \cdot G(t, x_k(t), z_k(t), u_k(t), v_k(t), \alpha_k(t)) = 0$ e $r_k(t) \leqslant 0 \ q.s.,$

onde $\partial h_k(t)$ denota o subdiferencial de h com relação a

$$(x, y, z, w, \omega_1, \omega_2, \omega_3, u, v, \alpha)$$

avaliado em S_k. Como h não depende de y, ω_1, ω_2 e ω_3, deduzimos de (B) que

$$-\dot{P}_k(t) = (-\dot{p}_k(t), 0, -\dot{q}_{z_k}(t), -\dot{q}_{w_k}(t)) \ \text{e} \ -\dot{\Pi}_k(t) = (0,0,0). \tag{A.2}$$

Agora nos voltamos a Ψ_k. Observe que, por (H3), $\Psi_k(x, y, x', y', z, w)$ é Lipschitz em torno de

$$(\bar{x}(0), \bar{x}(1), \bar{x}(1), \bar{x}(1), \bar{z}(1), \bar{w}(1)).$$

Afirmamos que, para k suficientemente grande,

$$\Psi_k(x_k(0), y_k(0), x_k(1), y_k(1), z_k(1), w_k(1)) > 0. \tag{A.3}$$

Por definição esta relação é não negativa. Suponha que é igual a zero. Então

$$y_k(1) = x_k(1), \quad l(x_k(0), y_k(0)) < l(\bar{x}(0), \bar{x}(1)), \tag{A.4}$$

$$(x_k(0), x_k(1)) \in C, \quad z_k(1) = 0, \quad w_k(1) = 0. \tag{A.5}$$

Como $\dot{w}_k(t) \geqslant 0$ $q.s.$ e $w_k(0) = w_k(1) = 0$ temos $w_k(t) = 0$ e $\alpha_k(t) = 0$ $q.s.$. Assim $\dot{z}_k(t) = 0$ $q.s.$, que, junto com $z_k(0) = 0$, implica que $z_k(t) = 0$ $q.s.$. Segue dos fatos acima que

$$G_i(t, x_k(t), z_k(t), u_k(t), v_k(t), \alpha_k(t)) = g_i(t, x_k(t), u_k(t), v_k(t)) \leqslant 0 \; q.s. \tag{A.6}$$

para todo $i \in \{1, \ldots, m_g\}$. Assim, se $\Psi_k(x_k(0), y_k(0), x_k(1), y_k(1), z_k(1), w_k(1)) = 0$, deduzimos de (A.4)-(A.6) que (x_k, u_k, v_k) é um processo admissível para (Q), contradizendo a otimalidade de $(\bar{x}, \bar{u}, \bar{v})$. Isto prova a afirmação.

Vamos agora provar que o cálculo do subdiferencial limite (Regra do Máximo, Teorema 3.6) e (A.3) dão a seguinte estimativa para o subdiferencial de Ψ_k com relação a (x, y, x', y', z, w):

$$\begin{aligned} \partial\Psi_k(x_k(0), y_k(0), x_k(1), y_k(1), z_k(1), w_k(1)) \subset \{&\rho_1(\theta_1, \theta_2, 0, 0, 0, 0) \\ &+ \rho_2(0, 0, 0, 0, 0, \epsilon_k^2 e_2) + \rho_3(0, 0, 0, 0, e_3, -e_3) + \rho_4(0, 0, e_4, -e_4, 0, 0) | \\ &(\theta_1, \theta_2) \in \partial l(x_k(0), y_k(0)), |e_2| = |e_3| = |e_4| = 1, \rho \in \mathcal{R}, \rho_i = 0 \text{ se } \bar{\psi}_{i_k} < \bar{\Psi}_k\}, \end{aligned} \tag{A.7}$$

onde e_2, e_3 e e_4 são vetores tais que $e_2, e_3 \in \mathbb{R}^{m_g}$, $e_4 \in \mathbb{R}^n$,

$$\Psi_k = \max\{\psi_{1_k}, \psi_{2_k}, \psi_{3_k}, \psi_{4_k}\},$$

$$\psi_{1_k} = l(x, y) - l(\bar{x}(0), \bar{x}(1)) + \epsilon_k^2, \quad \psi_{2_k} = \epsilon_k^2 |w|, \quad \psi_{3_k} = |w - z|, \quad \psi_{4_k} = |x' - y'|,$$

e $\bar{\Psi}_k$ e $\bar{\psi}_{i_k}$ denotam, respectivamente, as funções Ψ_k e ψ_{i_k} avaliadas em

$$(x_k(0), y_k(0), x_k(1), y_k(1), z_k(1), w_k(1)).$$

Suponha que $x_k(1) \neq y_k(1)$. Então a inclusão (A.7) vale com $e_4 \in \mathbb{R}^n$ e $|e_4| = 1$. Analogamente, se $w_k(1) \neq 0$, então (A.7) vale com $e_3 \in \mathbb{R}^{m_g}$, $|e_3| = 1$. Finalmente, se $w_k(1) \neq 0$, então (A.7) vale com $e_2 \in \mathbb{R}^{m_g}$, $|e_2| = 1$. Por outro lado, se $\psi_{i_k} = 0$ para algum $i \in \{2, 3, 4\}$, de (A.3) temos $\rho_{i_k} = 0$. Assim (A.7) vale.

De (D), (A.2) e (A.7) deduzimos a existência de vetores

$$\rho_k = (\rho_{1_k}, \rho_{2_k}, \rho_{3_k}, \rho_{4_k}), \quad e_{2_k}, e_{3_k} \in \mathbb{R}^{m_g}, \quad e_{4_k} \in \mathbb{R}^n$$

tais que $\rho_k \in \mathcal{R}$, $\rho_{i_k} = 0$ se $\bar{\psi}_{i_k} < \bar{\Psi}_k$,

$$|e_{2_k}| = |e_{3_k}| = |e_{4_k}| = 1, \tag{A.8}$$

$$\Pi_k(t) \equiv \lambda_k \epsilon_k (1, 1, 1). \tag{A.9}$$

Também temos que s_k é constante, e

$$q_{z_k}(1) = -\lambda_k \rho_{3_k} e_{3_k}, \tag{A.10}$$

$$q_{w_k}(1) = -q_{z_k}(1) + \lambda_k \rho_{2_k} \epsilon_k^2 e_{2_k}, \tag{A.11}$$

$$p_k(1) = -s_k = -\lambda_k \rho_{4_k} e_{4_k}, \tag{A.12}$$

$$|p_k(1)| = \lambda_k \rho_{4_k}, \tag{A.13}$$

$$(p_k(0), s_k) \in N_C(x_k(0), y_k(0)) + \lambda_k \rho_{1_k} \partial l(x_k(0), y_k(0)) + \lambda_k \epsilon_k (B \times B). \tag{A.14}$$

As equações (A.8), (A.12) e (A.13) implicam $|s_k| = |p_k(1)|$. Como $\rho_k \in \mathcal{R}$ temos $|s_k| = \lambda_k(1-\rho_{1_k}-\rho_{2_k}-\rho_{3_k})$, que, junto com (A.10), implica que $\lambda_k = |s_k|+|q_{z_k}(1)|+\lambda_k\rho_{1_k}+\lambda_k\rho_{2_k}$. Segue de (A) e de cima que

$$\lambda_k \rho_{1_k} + \lambda_k \rho_{2_k} + \|p_k\|_\infty + 2|p_k(1)| + \|q_{z_k}\|_\infty + |q_{z_k}(1)| + \|q_{w_k}\|_\infty + 3\lambda_k \epsilon_k > 0. \tag{A.15}$$

Afirmamos que esta última desigualdade implica que

$$\lambda_k \rho_{1_k} + \|p_k\|_\infty + 2|p_k(1)| + \|q_{z_k}\|_\infty + |q_{z_k}(1)| + \|q_{w_k}\|_\infty + 3\lambda_k \epsilon_k > 0. \tag{A.16}$$

Buscando uma contradição, suponha que (A.15) vale e

$$\lambda_k \rho_{1_k} + \|p_k\|_\infty + 2|p_k(1)| + \|q_{z_k}\|_\infty + |q_{z_k}(1)| + \|q_{w_k}\|_\infty + 3\lambda_k \epsilon_k = 0. \tag{A.17}$$

Então $\lambda_k \rho_{2_k} \neq 0$, isto é, $\lambda_k \neq 0$ e $\rho_{2_k} \neq 0$. Segue de (A.17) que $\rho_{1_k} = \rho_{3_k} = \rho_{4_k} = 0$ e consequentemente, $\rho_{2_k} = 1$. Por (A.17) também temos $q_{z_k}(1) = 0$ e $q_{w_k}(t) \equiv 0$. Contudo, como $\lambda_k \rho_{2_k} \neq 0$, segue de (A.11) que $q_{w_k}(1) \neq 0$, uma contradição. Assim (A.16) vale, uma condição que assegura que

$$\lambda_k \rho_{1_k} + \|p_k\|_\infty + \|q_{z_k}\|_\infty + \|q_{w_k}\|_\infty + 3\lambda_k \epsilon_k > 0.$$

As equações (A.10)-(A.13) implicam (c). Por outro lado, o requerimento que $\lambda_k \geqslant 0$, $\rho_k \in \mathcal{R}$, e (A.8) corresponde a (b). A inclusão (A.14), junto com (A.12), implica (d), e (C) e (E) são, respectivamente, (e) e (g). Para provar (f), note que

$$\begin{aligned} &h(t, x, y, z, w, \omega_1, \omega_2, \omega_3, p, s, q_z, q_w, r_k, \Pi, u, v, \alpha) \\ &\quad = \tilde{H}(t, x, z, w, p, q_z, q_w, r_k, u, v, \alpha) + \pi_1|u - u_k| + \pi_2|v - v_k| + \pi_3|\alpha - \alpha_k|. \end{aligned}$$

Estimando o sudiferencial $co\partial h_k$ com a ajuda da regra da soma e (A.9), temos

$$\begin{aligned} co\partial h_k \subset \{&(\varrho_1, 0, \varrho_2, \varrho_3, 0, 0, 0, \varrho_4, \varrho_5, \varrho_6) | (\varrho_1, \varrho_2, \varrho_3, \varrho_4, \varrho_5, \varrho_6) \\ &\in co\partial_{x,z,w,u,v,\alpha} \tilde{H}(t, x_k, z_k, w_k, p_k, q_{z_k}, q_{w_k}, r_k, u_k, v_k, \alpha_k)\} \\ &+ \lambda_k \epsilon_k (\{0, 0, 0, 0, 0, 0, 0\} \times B \times B \times B), \end{aligned}$$

e assim (f) vale em virtude dos fatos acima, (B) e (A.2). Finalmente, por um simples argumento de escala, podemos normalizar a soma dos multiplicadores sem afetar as outras conclusões do lema, implicando que (a) também vale. Isto completa a demonstração do lema. □

A referência dos próximos quatro resultados é Vinter [29]

Definição A.1. *Tome uma multifunção* $\Gamma : I \rightsquigarrow \mathbb{R}^k$. *Dizemos que uma função* $x : I \to \mathbb{R}^k$ *é uma seleção mensurável para* Γ *se* *(i) x é mensurável Lebesgue, e*

(ii) $x(t) \in \Gamma(t)$ q.s..

Teorema A.1. *Seja $\Gamma : I \rightsquigarrow \mathbb{R}^k$ uma multifunção não vazia. Assuma que Γ é fechada e mensurável. Então Γ tem uma seleção mensurável.*

Teorema A.2 (Teorema de Seleção Mensurável de Aumann)**.** *Seja $\Gamma \rightsquigarrow \mathbb{R}^k$ uma multifunção não vazia. Assuma que*

$$Gr\,\Gamma \text{ é } \mathcal{L} \times \mathcal{B}^k \text{ mensurável.}$$

Então Γ tem uma seleção mensurável.

Teorema A.3 (Teorema de Seleção Generalizado de Filippov)**.** *Considere uma multifunção não vazia $U : I \rightsquigarrow \mathbb{R}^m$ e uma função $g : I \times \mathbb{R}^m \to \mathbb{R}^n$ satisfazendo*

(a) o conjunto $Gr\,U$ é $\mathcal{L} \times \mathcal{B}^m$ mensurável;

(b) a função g é $\mathcal{L} \times \mathcal{B}^m$ mensurável.

Então para qualquer função mensurável $v : I \to \mathbb{R}^n$, a multifunção $U' : I \rightsquigarrow \mathbb{R}^m$ definida por

$$U'(t) := \{u \in U(t) : g(t, u) = v(t)\}$$

tem um gráfico $\mathcal{L} \times \mathcal{B}^m$ mensurável. Além disso, se

$$v(t) \in \{g(t, u) : u \in U(t)\} \ q.s.$$

então existe uma função mensurável $u : I \to \mathbb{R}^m$ satisfazendo

$$u(t) \in U(t) \ q.s.$$

$$g(t, u(t)) = v(t) \ q.s.$$

Demonstração de que existe uma função integrável K_r tal que $|r_k(t)| \leqslant K_r(t)\, q.t.\, t \in T$

Demonstração. Pelo Lema 4.1 (f), temos

$$0 \in co\partial_\alpha \tilde{H}(t, x_k(t), z_k(t), w_k(t), p_k(t), q_{z_k}(t), q_{w_k}(t), r_k(t), u_k(t), v_k(t), \alpha_k(t)) + \lambda_k \epsilon_k B,$$

onde

$$co\partial_\alpha \tilde{H}_k = q_{z_k}(-I) + q_{w_k}(2diag(\alpha_{k_i})) + r_k I.$$

Logo,

$$r_k(t) \in q_{z_k}(t) - 2q_{w_k}(t)diag(\alpha_{k_i}(t)) - \lambda_k\epsilon_k B\, q.t.\, t \in T.$$

Assim, pelos teoremas de seleção mensurável $\exists\, b(t)$, tal que

$$r_k(t) = q_{z_k}(t) - 2q_{w_k}(t)diag(\alpha_{k_i}(t)) - \lambda_k\epsilon_k b(t)\, q.t.\, t \in T.$$

Calculando a norma de r_k:

$$|r_k| \leqslant |q_{z_k}| + 2|q_{w_k}||diag(\alpha_{k_i})| + \lambda_k\epsilon_k|b|,$$

temos $|q_{z_k}| \leqslant 1$ e $|q_{w_k}| \leqslant 1$, pois são multiplicadores (ver Observação 4.1 (b)); $|diag(\alpha_{k_i})| \leqslant l_1$ para algum l_1 constante, pois $\alpha_k \in \mathcal{A}_\epsilon$ e $\lambda_k\epsilon_k|b| \leqslant l_2$ para algum l_2 constante, pois $b \in B$ e $\lambda_k\epsilon_k \leqslant l_2$. Então, existe uma função integrável K_r tal que

$$|r_k(t)| \leqslant K_r(t)\, q.t.\, t \in T,$$

onde $K_r = 1 + l_1 + l_2$. □

Definição A.2. *[29] Uma família de funções $\mathcal{S}$ é uniformemente integravelmente limitada se existe uma função integrável $\alpha \in L^1$ tal que*

$$|x(t)| \leqslant \alpha(t)\ q.s.$$

para todo $x \in \mathcal{S}$.

Demonstração que as sequências $\{\dot{p}_k\}$, $\{\dot{q}_{z_k}\}$ e $\{\dot{q}_{w_k}\}$ são uniformemente integravelmente limitadas

Demonstração. Pelo Lema 4.1 (f), temos

$$(-\dot{p}_k(t), -\dot{q}_{z_k}(t), -\dot{q}_{w_k}(t)) \in co\partial_{x,z,w}\tilde{H}_k(t).$$

Pela Proposição 3.8, $\partial\tilde{H}_k \subset KB$, onde K é a constante de Lipschitz de $\tilde{H}$, temos K integrável por (H1) e (H5). Logo, usando a Proposição 3.8, teoremas de seleção mensurável com ideias análogas à demonstração anterior, concluímos que as sequências são de fato uniformemente integravelmente limitadas. □

Demonstração que existe uma função integrável K_ζ tal que $|\zeta_k(t)| \leqslant K_\zeta(t)\, q.t.\, t \in T$

Demonstração. Segue análogo às duas últimas demonstrações. □

A referência dos resultados que seguem é Rudin [37]

Definição A.3. *Uma família $\mathcal{F}$ de funções f definidas em um conjunto E em um espaço métrico X é dita ser equicontínua em E se para cada $\epsilon > 0$ existe um $\delta > 0$ tal que*

$$|f(x) - f(y)| < \epsilon$$

sempre que $d(x, y) < \delta, x \in E, y \in E$ e $f \in \mathcal{F}$. Aqui d denota a métrica de X.

Teorema A.4. *Seja K um conjunto compacto.*

(a) Se $\{f_n\}$ é uma sequência uniformemente convergente de funções contínuas em K, então $\{f_n\}$ é equicontínua em K.

(b) Se $\{f_n\}$ é limitada ponto a ponto e equicontínua em K, então $\{f_n\}$ contém uma subsequência uniformemente convergente, e $\{f_n\}$ é uniformemente limitada em K.

Teorema A.5 (Compacidade das Trajetórias)**.** *[29] Tome um subconjunto relativamente aberto $\Omega \subset [S, T] \times \mathbb{R}^n$ e uma multifunção $F : \Omega \rightsquigarrow \mathbb{R}^n$. Assuma que, para alguma multifunção fechada $X : [S, T] \rightsquigarrow \mathbb{R}^n$ tal que $Gr\, X \subset \Omega$, as seguintes hipóteses são satisfeitas:*

(i) F é uma multifunção fechada, convexa, não vazia.

(ii) F é $\mathcal{L} \times \mathcal{B}^n$ mensurável.

(iii) Para cada $t \in [S, T]$, o gráfico de $F(t, \cdot)$ restrito a $X(t)$ é fechado.

Considere uma sequência $\{x_i\}$ de funções $W^{1,1}([S, T]; \mathbb{R}^n)$, uma sequência $\{r_i\}$ em $L^1([S, T]; \mathbb{R})$ tal que $\|r_i\|_{L^1} \to 0$ quando $i \to \infty$, e uma sequência $\{A_i\}$ de subconjuntos mensuráveis de $[S, T]$ tal que $meas A_i \to |T - S|$ quando $i \to \infty$. Suponha que:

(iv) $Gr\, x_i \subset Gr\, X$ para todo i;

(v) $\{\dot{x}_i\}$ é uma sequência de funções uniformemente integravelmente limitadas em $[S, T]$ e $\{x_i(S)\}$ é uma sequência limitada;

(vi) existe $c \in L^1$ tal que

$$F(t, x_i(t)) \subset c(t)B$$

para quase todo $t \in A_i$ e para $i = 1, 2, \ldots$.

Suponha, além disso, que

$$\dot{x}_i(t) \in F(t, x_i(t)) + r_i(t)B \text{ q.t. } t \in A_i.$$

Então ao longo de alguma subsequência

$$x_i \to x \text{ uniformemente e } \dot{x}_i \to \dot{x} \text{ fracamente em } L^1$$

para algum $x \in W^{1,1}([S,T];\mathbb{R}^n)$ *satisfazendo*

$$\dot{x}(t) \in F(t, x(t)) \text{ q.t. } t \in [S,T].$$

Teorema A.6 (Desigualdade de Gronwall)**.** *[29] Tome uma função absolutamente contínua* $z : [S,T] \to \mathbb{R}^n$. *Assuma que existem funções integráveis não negativas* k *e* v *tais que*

$$|\frac{d}{dt}z(t)| \leqslant k(t)|z(t)| + v(t) \text{ q.t. } t \in [S,T].$$

Então

$$|z(t)| \leqslant \exp\left(\int_S^t k(\sigma)d\sigma\right)\left[|z(S)| + \int_S^t \exp\left(-\int_S^\tau k(\sigma)d\sigma\right)v(\tau)d\tau\right]$$

para todo $t \in [S,T]$.

Demonstração de que a Desigualdade de Gronwall e o fato que $p(\tau) = 0$ para algum $\tau \in T$ implicam que $p \equiv 0$.

Demonstração. Considere a seguinte mudança de variável

$$s = \tau - t.$$

Defina $q : [0,\tau] \to \mathbb{R}^n, q(s) := p(\tau - s)$. Assim, $q(0) = p(\tau) = 0$. Temos

$$|\dot{q}(s)| = |-\dot{p}(\tau - s)| = |\dot{p}(\tau - s)| \leqslant K_1(\tau - s)|p(\tau - s)| = \tilde{K}_1(s)|q(s)|.$$

Pela Desigualdade de Gronwall,

$$|q(s)| \leqslant \exp\left(\int_0^\tau \tilde{K}_1(\sigma)d\sigma\right)|q(0)| = 0 \Rightarrow q(s) = 0, s \in [0,\tau].$$

Assim, $p(t) = p(\tau - s) = q(s) = 0, t \in [0,\tau]$.

Analogamente, mostra-se que $p(t) = 0, t \in [\tau, 1]$. Basta definir $s = t - \tau$ e $q : [0, 1-\tau] \to \mathbb{R}^n, q(s) := p(s+\tau)$.

Portanto, $p \equiv 0$. □

Demonstração que o processo $(\bar{x}, \bar{u}, \bar{v})$ é admissível para (P_{aux})

Demonstração. De fato,

$$\begin{aligned} F(t,\bar{x}(t),\bar{u}(t),\bar{v}(t)) &= f(t,\bar{x}(t),\bar{u}(t) + \nabla_u\bar{b}(t)^\top D(t,\bar{x},\bar{u},\bar{v}),\bar{v}(t)) \\ &= f(t,\bar{x}(t),\bar{u}(t) + \nabla_u\bar{b}(t)^\top d(t,0,0,0),\bar{v}(t)) = f(t,\bar{x}(t),\bar{u}(t),\bar{v}(t)) = \dot{\bar{x}}(t), \end{aligned}$$

$$G(t,\bar{x}(t),\bar{u}(t),\bar{v}(t)) = g(t,\bar{x}(t),\bar{u}(t)+\nabla_u\bar{b}(t)^\top D(t,\bar{x},\bar{u},\bar{v}),\bar{v}(t)) = g(t,\bar{x}(t),\bar{u}(t),\bar{v}(t)) \leqslant 0,$$

$$(\bar{u}(t),\bar{v}(t)) \in \mathcal{U}(t)\times\mathcal{V}(t),$$

$$(\bar{x}(0),\bar{x}(1)) \in C.$$

Como $(\bar{x},\bar{u},\bar{v})$ satisfaz as restrições de (P_{aux}), temos $(\bar{x},\bar{u},\bar{v})$ admissível para (P_{aux}). □

Assumindo que $(\tilde{x},\tilde{u},\tilde{v})$ é uma solução para (P_{aux}), concluímos que $(\tilde{x},\hat{u},\tilde{v})$ é uma solução para (P)

Demonstração. De fato,

$$\mu(t,(\xi(t),u_1(t),v_1(t)),d(t,\xi(t),u_1(t),v_1(t))) = b(t,\tilde{x}(t),\hat{u}(t),\tilde{v}(t)),$$

como $(\xi(t),u_1(t),v_1(t)) \in \sigma_1 B\times\sigma_1 B\times\sigma_1 B$, temos $\mu(t,(\xi(t),u_1(t),v_1(t)),d(t,\xi(t),u_1(t),v_1(t))) = 0$, logo

$$b(t,\tilde{x}(t),\hat{u}(t),\tilde{v}(t)) = 0.$$

Além disso, temos

$$g(t,\tilde{x}(t),\hat{u}(t),\tilde{v}(t)) = G(\tilde{x}(t),\tilde{u}(t),\tilde{v}(t)) \leqslant 0,$$

$$f(t,\tilde{x}(t),\hat{u}(t),\tilde{v}(t)) = F(t,\tilde{x}(t),\tilde{u}(t),\tilde{v}(t)) = \dot{\tilde{x}}(t),$$

$$\tilde{v}(t) \in \mathcal{V} \Rightarrow \tilde{v}(t) \in V(t),$$

$$(\tilde{x}(0),\tilde{x}(1)) \in C$$

pois $(\tilde{x},\tilde{u},\tilde{v})$ é uma solução para (P_{aux}).

Como as restrições de (P) são satisfeitas temos o que queríamos. □

Demonstração de que existe K_m integrável tal que $|q(t)| \leqslant K_m(t)|p(t)|$

Demonstração. Temos $q(t) = -(\theta_2(t)+r(t)\nabla_u\bar{g}(t))\nabla_u\bar{b}(t)^\top\Gamma(t)^{-1}$. Logo,

$$\begin{aligned}
|q(t)| &= |(\theta_2(t)+r(t)\nabla_u\bar{g}(t))\nabla_u\bar{b}(t)^\top\Gamma(t)^{-1}| \\
&= |\theta_2(t)+r(t)\nabla_u\bar{g}(t)||\nabla_u\bar{b}(t)^\top||\Gamma(t)^{-1}| \\
&\leqslant (|\theta_2(t)|+|r(t)||\nabla_u\bar{g}(t)|)|\nabla_u\bar{b}(t)^\top|M \\
&\overset{(H6)}{\leqslant} (|\theta_2(t)|+|r(t)|K_{b,g})K_{b,g}M \\
&\overset{(H1)}{\leqslant} (k_f(t)|p(t)|+R(t)|p(t)|K_{b,g})K_{b,g}M = K_m(t)|p(t)|,
\end{aligned}$$

onde $K_m(t) = k_f(t)K_{b,g}M + R(t)K_{b,g}K_{b,g}M$. Portanto, $|q(t)| \leqslant K_m(t)|p(t)|$, como queríamos. □

A referência para os resultados que seguem é Clarke [28]

Proposição A.2 (Somas Finitas)**.**

$$\partial \left(\sum f_i \right)(x) \subset \sum \partial f_i(x).$$

Corolário A.1. *Vale a igualdade na Proposição A.2 se todas as funções f_i são, ou no máximo uma não é, estritamente diferenciáveis em x.*