UNIVERSIDADE FEDERAL DO PARANÁ
Camila Isoton

Algumas contribuições em Controle Ótimo Discreto

Curitiba, 2017.

UNIVERSIDADE FEDERAL DO PARANÁ
Camila Isoton

Algumas contribuições em Controle Ótimo Discreto

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

Orientadora:
Profª Drª Lucelina Batista dos Santos (UFPR - Brasil)

Coorientador:
Prof. Dr. Marko Antonio Rojas Medar (UTA - Chile)

Curitiba, 2017.

IS85a Isoton, Camila
Algumas contribuições em controle ótimo discreto / Camila Isoton. – Curitiba, 2017.
139 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Matemática, 2017.

Orientadora: Lucelina Batista dos Santos.
Coorientador: Marko Antonio Rojas Medar.

1. Matemática. 2. Teoria de controle ótimo. 3. Princípio de máximo. I. Universidade Federal do Paraná. II. Santos, Lucelina Batista dos. III. Medar, Marko Antonio Rojas. IV. Título.

CDD: 519.4

ATA Nº015

ATA DE SESSÃO PÚBLICA DE DEFESA DE DOUTORADO PARA A OBTENÇÃO DO GRAU DE DOUTOR EM MATEMÁTICA

No dia vinte e cinco de Agosto de dois mil e dezessete às 10:00 horas, na sala de Reuniões do Setor de Ciências Exatas, Centro Politécnico, Universidade Federal do Paraná, foram instalados os trabalhos de arguição da doutoranda **CAMILA ISOTON** para a Defesa Pública de sua tese intitulada **Algumas contribuições em Controle Ótimo Discreto**. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: LUCELINA BATISTA DOS SANTOS (UFPR), VALERIANO ANTUNES DE OLIVEIRA (UNESP/SJRP), ADEMIR ALVES RIBEIRO (UFPR), YURILEV CHALCO CANO (UTA), ADILSON JOSE VIEIRA BRANDAO (UFSCAR), . Dando início à sessão, a presidência passou a palavra a discente, para que a mesma expusesse seu trabalho aos presentes. Em seguida, a presidência passou a palavra a cada um dos Examinadores, para suas respectivas arguições. A aluna respondeu a cada um dos arguidores. A presidência retomou a palavra para suas considerações finais. A Banca Examinadora, então, reuniu-se e, após a discussão de suas avaliações, decidiu-se pela APROVAÇÃO da aluna. A doutoranda foi convidada a ingressar novamente na sala, bem como os demais assistentes, após o que a presidência fez a leitura do Parecer da Banca Examinadora. A aprovação no rito de defesa deverá ser homologada pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais do programa. A outorga do título de doutor está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, LUCELINA BATISTA DOS SANTOS, lavrei a presente ata, que vai assinada por mim e pelos membros da Comissão Examinadora.

Curitiba, 25 de Agosto de 2017.

LUCELINA BATISTA DOS SANTOS
Presidente da Banca Examinadora (UFPR)

VALERIANO ANTUNES DE OLIVEIRA
Avaliador Externo (UNESP/SJRP)

ADEMIR ALVES RIBEIRO
Avaliador Interno (UFPR)

YURILEV CHALCO CANO
Avaliador Externo (UTA)

ADILSON JOSE VIEIRA BRANDAO
Avaliador Externo (UFSCAR)

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
Setor CIÊNCIAS EXATAS
Programa de Pós-Graduação MATEMÁTICA

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **CAMILA ISOTON** intitulada: **Algumas contribuições em Controle Ótimo Discreto**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 25 de Agosto de 2017.

LUCELINA BATISTA DOS SANTOS
Presidente da Banca Examinadora (UFPR)

VALERIANO ANTUNES DE OLIVEIRA
Avaliador Externo (UNESP/SJRP)

ADEMIR ALVES RIBEIRO
Avaliador Interno (UFPR)

YURILEV CHALCO CANO
Avaliador Externo (UTA)

ADILSON JOSE VIEIRA BRANDAO
Avaliador Externo (UFSCAR)

Agradecimentos

Meus sinceros agradecimentos:

Agradeço primeiramente a minha querida orientadora, Lucelina, pelos muitos momentos de discussões, apoio, tempo, esforço, dedicação empregados nesta pesquisa, enfim, pela amizade que construímos ao longo desses seis anos de estudos. E que segue rendendo frutos profissionais e pessoais.

Ao meu co-orientador, Marko, pelas ideias e discussões, pela acolhida em Chillán, pela preocupação, carinho e zelo por mim. A minha querida Professora Violeta, por todas as horas de estudos e discussões nas muitas produtivas idas a Concepción, pela amizade, acolhida e ótimos momentos passados juntas.

Aos meus pais Expedito e Lorena, e a minha irmã, Gisele, que desde sempre me apoiaram, incentivando, ajudando, compreendendo, enfim, por todo o carinho e amor que sempre tiveram comigo. Aos meus familiares que tanto me apoiaram e me incentivaram.

Agradeço a toda a turma do Chile, em especial as minhas amigas Claudia e Caroline pelo carinho, companheirismo e bons momentos passados juntos. A Patrícia Mora que foi praticamente uma mãe em minha estadia em Chillán e a toda família Mora, muito obrigada pela acolhida!

Agradeço aos amigos que o PPGM me presenteou, Ana Paula, Stela, Geovani, Mineiro, Pamela, Teles, Mirela e Monique, por tantos momentos compartilhados, amizade e companheirismo. Este trabalho tem um pouco de cada um de vocês.

Aos professores do programa, que me ajudaram a crescer intelectual e pessoalmente.

Aos membros da banca, professores Ademir, Adilson, Valeriano e Yurilev obrigada por aceitarem o convite e pelas contribuições sugeridas.

A amiga-irmã Patrícia, pelo apoio, compreensão e amizade ao longo destes seis anos no apartamento 3 da rua Buenos Aires. A querida amiga Ana Cristina que esteve presente nos momentos mais marcantes dessa jornada, na chegada a Curitiba e agora ao término do doutorado. A Luana e aos demais amigos que me apoiaram. Muito Obrigada!

Ao Programa de Pós-Graduação em Matemática da UFPR pela oportunidade e formação de qualidade propiciada. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPQ, pelo apoio financeiro.

E finalmente, mas não menos importante, a Deus pela capacitação concedida, pela força espiritual principalmente nos momentos mais difíceis e por seguir ao meu lado.

“ Deus é o geômetra onipotente para quem o mundo é um imenso problema matemático”

Gottfried Wilhelm Leibniz

Aos meus pais e irmã
Expedito, Lorena e Gisele por sempre confiarem, acreditarem, apoiarem e me darem forças para seguir.

Resumo

Neste trabalho consideramos os problemas de controle ótimo discretos com um e com vários objetivos, nos casos regulares e 2 regulares. Este estudo está dividido em três frentes: a primeira trata das condições de otimalidade destes dois tipos de problemas em suas versões mono e multiobjetivo. Nesta parte apresentamos uma versão do Princípio do Máximo Discreto e introduzimos conceitos de invexidade nos quais, os problemas PM-invexos e PM-pseudoinvexos são a chave para garantir a suficiência destas condições para o caso regular. Na segunda, discutimos os conceitos de estabilidade e sensibilidade a certos problemas de controle ótimo discretos escalares, para os quais obtivemos dois resultados importantes envolvendo condições de crescimento quadrático, independência linear e 2 regularidade. Já na última parte, abordamos a otimalidade de um certo problema de controle ótimo discreto multiobjetivo não diferenciável. Através do conceito de diferenciabilidade generalizada de Clarke, apresentamos uma versão do Princípio do Máximo para tal problema.

Palavras-chave: Controle Ótimo Discreto; Princípio do Máximo Discreto; PM-invexidade, 2 Regularidade, Estabilidade e Sensibilidade, Não-suavidade.

Abstract

In this Thesis, we discuss discrete optimal control problems for regular and irregular (2 regular) cases. This study was divided into three fronts: the first deals with the optimality conditions of these two types of problems in their scalar and multiobjective versions. In this part we present a version of the Discrete Maximum Principle and we introduce the concepts of MP-invexity and MP-pseudoinvexity for these problems; these notions were the key to guarantee the adequacy of these conditions for regular problems. In the second part, we discuss the concepts of stability and sensitivity for certain discrete scalar control problems, for which we obtained two important results involving quadratic growth conditions, linear independence and regularity. In this last part, we discuss the optimality of a certain class of nonsmooth discrete multiobjective optimal control problems. Based on Clarke's concept of generalized differentiability, we present a version of the Principle of Maximum.

Keywords: Discrete Optimal Control, Maximum Principle, MP-invexity, 2 regularity, Stability and Sensibility, Nonsmooth.

Lista de Notações

Os principais símbolos utilizados nesta tese são os seguintes:

$\nabla f(x,u)$	Gradiente da função f em (x,u).
$\nabla_x f(x,u)$	Gradiente da função $f(\cdot,u)$ com respeito a primeira variável.
$\nabla^2 f(x,u)$	Hessiana da função f em (x,u).
$\nabla^2_{xx} f(x,u)$	Derivada de segunda ordem da função $f(\cdot,u)$ com respeito à primeira variável.
$\delta^+ f(\hat{\sigma},d)$	Derivada direcional de f no ponto $\hat{\sigma}$ na direção d.
$\phi^0(x;v)$	Derivada direcional de Clarke da função ϕ no ponto x na direção v.
$\partial\phi(x)$	Subdiferencial de Clarke.
$\mathbf{Im} f(x)$	Imagem da função f.
$\mathbf{Ker} f(x)$	Núcleo da função f.
$\mathbf{Fr} A$	Fronteira do conjunto A.
$\mathbf{int} A$	Interior do conjunto A.

Outras notações serão introduzidas ao longo do texto quando se fizerem necessárias.

Sumário

Capítulo 1

Introdução

Vários problemas em planejamento econômico, tecnologia, organização de produção e investigação de operações são descritos por sistemas de equações de diferenças [9, 42, 45, 82]. Além disso, os sistemas discretos são encontrados na resolução de problemas de controle ótimo contínuos por meio de métodos numéricos [82, 96]. A programação dinâmica [9] é um método universal para investigar sistemas de controle ótimo discretos, apesar disso existem algumas desvantagens sobre outros métodos e a aplicação deste modelo em problemas específicos pode ser menos efetiva [42]. Isto é o porquê na maioria dos casos (especialmente nos sistemas discretos de grande porte), o uso das condições de otimalidade se torna preferível ao da programação dinâmica.

O que ocorre na Teoria de Controle Ótimo é um caso um tanto que peculiar mediante o que em geral acontece na história da matemática, pois neste caso a teoria das variantes discretas apareceu depois da teoria das variantes contínuas. Esta última tem como um resultado central o Princípio do Máximo provado por Pontryagin *et al.* [81] e dada a sua significância o que se busca é um resultado análogo em tempo discreto. No artigo de Rozonoer [87] o análogo do princípio do máximo é provado para um problema de controle ótimo discreto linear (com respeito ao espaço das variáveis) com funcional de custo linear. Como bem sabemos o princípio do máximo discreto não é válido em geral sem hipóteses de convexidade impostas *a priori* no sistema de controle, [24]. Depois disso, muitos autores tentaram provar o princípio do máximo em uma versão mais fraca (máximo local, estado de estacionariedade) e juntamente com condições de otimalidade de ordens maiores. Levando em consideração os estudos acima mencionados e os resultados dos artigos [5, 9, 14, 24, 38, 41, 51, 52, 65, 66, 67, 68, 82], dentre outros, podemos dizer que a teoria das condições de otimalidade em sistemas discretos continua sendo um tema de pesquisa atual.

Neste trabalho provamos uma versão do Princípio do Máximo apresentada por [66] a qual é demonstrada com base no formalismo de Dubovitskii-Milyutin, por simplicidade formalismo DM, teoria esta aplicada inicialmente a problemas de otimização e descrita extensivamente no livro de Girsanov [47]. De uma maneira geral a teoria do formalismo envolve o problema de separabilidade de um sistema de cones convexos.

Se o problema é regular ou normal - ou ainda controlável, segundo a terminologia usual em controle ótimo - temos que o multiplicador associado à função objetivo é não nulo. Assim sendo, as condições de otimalidade de primeira ordem (Regra de multiplicadores de Lagrange - Karush-Kuhn-Tucker) no caso da Programação Matemática, (Princípio do

Máximo de Pontryagin) no caso do Controle Ótimo, não são suficientes, em geral, para decidir se um ponto em análise é de fato a solução ótima procurada.

Nos últimos anos surgiu na literatura vinculada à Programação Matemática a noção de invexidade (introduzida por Hanson [49]) e suas generalizações [34], [71], dentre os quais destacamos aqui o conceito de KT-invexidade que foi introduzido por Martin [70] para problemas de programação matemática e posteriormente generalizado por Osuna-Gómez *et al.* [78] para problemas multiobjetivos. O conceito de KT-invexidade é bastante interessante pois é a maior classe de problemas para os quais as condições de Karush-Kuhn-Tucker são necessárias e suficientes para a otimalidade. Seguindo esta linha de raciocínio definimos, de modo similar ao apresentado por [75], os problemas PM-invexos e provamos que esta classe de problemas é a maior classe para a qual as condições do Princípio do Máximo são necessárias e suficientes para otimalidade.

Nosso interesse também está nos problemas em programação multiobjetivo, pois estes surgiram no sentido de contemplar situações mais realistas, nas quais se estudam vários objetivos, em geral conflitantes, que concorrem para uma determinada solução. Informalmente, podemos dizer que um ponto ótimo de um problema multiobjetivo é chamado *eficiente* (Pareto), quando não é possível melhorar nenhum objetivo sem que haja uma piora em algum outro. Essa nomenclatura foi introduzida por Pareto, um dos precursores neste campo da Otimização Multiobjetivo, que tratou do tema em seu célebre trabalho *"Cours d'Economie Politique"* [79].

Em meados da década de 40 (auge da Pesquisa Operacional) muitos trabalhos surgiram no campo da otimização, dentre os quais se destacam os autores Charnes, Cooper, Gale, Kuhn, Tucker e de Karlin. Entre as décadas de 60 e 70 a literatura sobre este tema foi se sofisticando e surgiram as noções de *eficiência fraca* (Pareto fraco) e *eficiência própria* (Pareto próprio); grosso modo, um ponto é fracamente eficiente se não é possível melhorar todos os objetivos simultaneamente e propriamente eficiente se é solução eficiente e se os quocientes entre ganho em um objetivo e perda com respeito aos demais é limitada. A partir disso, inúmeras publicações têm aparecido, visando a desenvolver critérios que permitam decidir quando um ponto factível é ou não eficiente. Nosso objetivo é apresentar condições de otimalidade que garantam a eficiência para os problemas de controle ótimo discretos multiobjetivos de tal forma que obtemos um teorema do Princípio do Máximo para esta classe de problemas e provamos que os problemas PM-pseudosinvexos são a maior classe de problemas para as quais as condições do Princípio do Máximo, sob hipóteses de regularidade fraca, são necessárias e suficientes para a otimalidade.

Por outro lado, se o problema em questão não é regular, ou seja, quando não podemos garantir que o multiplicador associado à função objetivo do problema é não nulo, o Princípio do Máximo proporciona condições degeneradas, que não dependem do funcional a minimizar.

Neste caso se faz necessário caracterizar os cones (definidos em um ponto ótimo) envolvidos no problema por meio de aproximações de segunda ordem, o que nos permite estabelecer condições necessárias de otimalidade de segunda ordem, as quais serão não degeneradas sempre que o problema seja 2-regular. Este conceito foi introduzido por Arutyunov em seu livro, *"Optimality Conditions: Abnormal and Degenerate Problems"*, onde o autor faz um estudo de uma classe de problemas a qual ele define como problemas anormais (ou degenerados). Entende-se por problemas normais, de um modo geral, aqueles que cumprem as condições de regularidade e por isso é comum o uso do termo problemas regulares. Posteriormente o conceito de 2-regularidade foi usado também por Avakov [6], por Izmailov [54] e também por Vivanco-Orellana em [77]. Em nosso trabalho seguimos esta nomenclatura e utilizamos o conceito de 2-regularidade quando o problema não é regular, ou seja, quando as condições necessárias de primeira ordem clássica não valem, em geral. Baseados no trabalho de Marinkovic [69], provamos um teorema do Princípio do Máximo não regular escalar e outro multiobjetivo.

Outra questão notável que podemos observar é que usualmente não conhecemos os valores exatos dos parâmetros de um sistema de controle ou esses valores estão sujeitos a perturbações. Portanto, é importante conhecer como calcular o controle dependendo dos parâmetros do modelo. Assim sendo, o *status* de uma solução não pode ser compreendido sem informações sobre a sensibilidade e a estabilidade.

Nosso interesse é estudar a estabilidade com respeito aos dados de um problema de controle ótimo no sentido das soluções, ou ainda, no sentido do valor ótimo de função. Segundo Bonnans & Shapiro [17],

> ***Estabilidade:*** *refere-se à dependência contínua de soluções ótimas dos dados do problema. Em outras palavras, a estabilidade assegura que o problema é bem posto.*
>
> ***Sensibilidade:*** *refere-se às informações que permitem quantificar a dependência da solução em relação aos dados do problema, utilizando para este fim noções apropriadas de diferenciabilidade.*

Estas propriedades dos conjuntos de soluções têm sido foco de estudo de vários pesquisadores, como por exemplo na programação matemática, [17, 37, 43, 44, 55, 62, 63, 83] onde os autores utilizam esta ferramenta para obter condições de otimalidade, dualidade, estudo de algoritmos: convergência, taxa de convergência e aceleração de convergência, além de suas mais óbvias e imediatas aplicações na estimativa de soluções próximas com dados diferentes. Esse tipo de análise também é feita em [1, 17, 55, 61, 62] no tratamento de problemas de controle ótimo. Além disso, estes conceitos são essenciais em aplicações tais como: validação de modelo e análise custo-benefício, mecânica estática, análise de redes elétricas e projeções métricas ortogonais. Veja Bonnans & Shapiro[17] e suas referências.

Neste contexto estabelecemos um teorema de sensibilidade e estabilidade aos problemas de controle ótimo discretos, estendendo o resultado apresentado por Malanowski em [61] e também aplicamos o Teorema de Estabilidade estabelecido por Arutyunov & Izmailov em [4] para problemas de programação matemática e o aplicaremos a uma certa classe de problemas de controle ótimo discreto.

Por outro lado temos que o papel da Análise não-diferenciável é de fundamental importância na Teoria de Otimização e isto se deve a pelo menos dois motivos fundamentais: Por um lado, em muitas aplicações a hipótese de diferenciabilidade pode ser muito restritiva e, por outro lado, existem inúmeros mecanismos naturais que geram "não suavidades", mesmo quando se parte de problemas diferenciáveis. Isto ocorre, por exemplo, na Teoria de Dualidade, Análise de Estabilidade, Técnicas de Decomposição, Métodos de Penalização, dentre outros. Veja [26, 27] e suas referências.

Daí, nosso interesse no estudo de problemas de controle ótimo discreto não diferenciáveis. Nessa direção, estendemos o teorema do Princípio do Máximo escalar provado por Schavertsman em [91] a problemas de controle ótimo discreto multiobjetivos. Além disto, provamos um resultado de eficiência própria seguindo de modo semelhante ao dado por Benson e Morin em [10], baseando-nos no conceito de *calmness*.

Em suma, esta tese está estruturada em três partes. Na primeira consideramos problemas de controle ótimo discretos regulares. No Capítulo 2 fazemos um resgate dos principais conceitos do formalismo de Dubovistkii-Milyutin e suas extensões a problemas multiobjetivos como feito em [58].

No Capítulo 3 apresentamos o problema de controle ótimo discreto diferenciável regular. Provamos o Teorema do Princípio do Máximo (em sua versão em tempo discreto) via formalismo de Dubovistki-Milyutin, discutimos a questão da regularidade e definimos o conceito de PM-invexidade a problemas de controle ótimo discretos, o qual é importante na obtenção das condições suficientes de otimalidade de primeira ordem.

No Capítulo 4 fazemos um estudo análogo ao caso mono-objetivo, para obter as condições de otimalidade de um problema de controle ótimo discreto multiobjetivo regular. Obtivemos uma versão do Princípio do Máximo Discreto para problemas multiobjetivos e introduzimos o conceito de PM-invexidade para tais problemas, estendendo os resultados obtidos no Capítulo 3 a problemas multiobjetivos.

Na segunda parte vamos tratar de problemas de controle ótimo discreto não-regulares. No Capítulo 5 iniciamos a segunda parte deste trabalho, onde abordamos um problema de controle ótimo discreto diferenciável não regular. Apresentamos o conceitos de 2-regularidade e obtemos um análogo do Teorema do Princípio do Máximo para este caso, caracterizando assim as condições necessárias de otimalidade de segunda ordem para problemas de controle ótimo discreto 2-regulares. Para este fim, inicialmente discutimos o comportamento de um problema de programação matemática e alguns resultados que são imprescindíveis na obtenção dos resultados em controle ótimo discreto. Aproveitando ainda do conceito já definido de PM-invexidade, apresentamos uma versão ao nosso problema de estudo e obtemos as condições suficientes de otimalidade de segunda ordem.

O Capítulo 6 trata do problema de controle ótimo discreto multiobjetivo não-regular através de uma abordagem semelhante à utilizada no Capítulo 5. Adaptamos os conceitos de 2-regularidade para o caso multiobjetivo e obtemos assim condições necessárias de otimalidade.

A Parte 3 desta tese trata de temas correlatos. No Capitulo 7 vamos caracterizar uma classe de problemas de controle ótimo discreto escalares, dependendo do parâmetro, que tem soluções locais localmente isoladas que são localmente Lipschitz contínuas e direcionalmente diferenciáveis em função do parâmetro. Posteriormente trataremos de um problema de controle ótimo discreto escalar não regular, somente com restrições de igualdade, e demonstramos um resultado de estabilidade para este problema.

No Capítulo 8, discutimos a otimalidade para problemas de programação matemática não-diferenciáveis em termos de uma condição de ponto de sela e de dualidade. A partir destes resultados, obtemos uma condição suficiente para que uma solução eficiente de um problema multiobjetivo seja propriamente eficiente. Esta caracterização se dá a partir da propriedade de *calmness* de um certo problema escalarizado. Posteriormente tratamos de um problema de controle ótimo discreto multiobjetivo não suave. Para este problema provamos uma versão do Princípio do Máximo, generalizando os trabalhos de Shvartsman [91] e Mordukhovich e Shvartsman [74].

Por fim fazemos algumas considerações finais e apontamos algumas possibilidades de prosseguimento deste trabalho.

Capítulo 2

Preliminares Teóricos

Neste capítulo apresentaremos definições e resultados que serão muito úteis no decorrer deste trabalho. Na primeira seção trataremos dos conceitos que envolvem a teoria do formalismo de Dubovitskii e Milyutin em problemas de programação matemática escalares (mono-objetivos) e na segunda abordaremos uma extensão destes resultados a problemas de programação matemática multiobjetivos.

2.1 Formalismo de Dubovitskii-Milyutin: Mono-objetivo

Em 1962, Dubovitskii e Milyutin desenvolveram condições necessárias de otimalidade na forma de uma equação baseados na teoria da Análise Funcional. Há uma gama de trabalhos, tanto de programação matemática como de controle ótimo, vetoriais e escalares, de dimensão finita e infinita, diferenciáveis ou não, (veja por exemplo [11], [14], [51], [52]), que fazem uso do formalismo de DM na obtenção das condições necessárias de otimalidade. É por isso que utilizaremos este formalismo em nosso trabalho. No que segue apresentamos definições e resultados necessários para sua aplicação.

Consideremos inicialmente o seguinte problema de otimização,

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

onde $f : X \to \mathbb{R}$, sendo X um espaço de Banach, $Q_i \subset X, i = 1, \dots, k$, conjuntos com interior não vazio e $Q_{k+1} \subset X$ conjunto com interior vazio. Em geral, os conjuntos $Q_i, i = 1, \dots, k$, representam as restrições de desigualdade do problema e Q_{k+1} representa as restrições de igualdade.

Nos trabalhos de Dubovitskii e Milyutin [39] e [40], as condições necessárias de otimalidade local em um ponto $x_0 \in X$, foram obtidas a partir da separação das aproximações cônicas aos conjuntos de restrições $Q_i, i = 1, \dots, k+1$, e do conjunto $\{x \in X : f(x) < f(x_0)\}$. Nos referidos trabalhos as definições foram feitas de tal forma que se tais cones (de direção factível, tangente e de descida) são não vazios e convexos então, x_0 é um mínimo local do problema (OP) se, e somente se, não existe uma direção comum a todas as aproximações cônicas. Os resultados de Dubovitskii e Milyutin provam que esta propriedade geométrica de otimalidade local do ponto x_0 pode ser equivalentemente descrita em termos das formas lineares dos cones duais (ou polares) correspondentes.

2.1.1 Definições e resultados

Apresentamos algumas definições e propriedades necessárias para a aplicação do formalismo de Dubovitskii-Milyutin.

Seja X um espaço de Banach.

Definição 2.1 *Um conjunto $C \subset X$ é chamado **cone** de vértice 0 se para cada $x \in C$ tem-se $\lambda x \in C$ para qualquer $\lambda > 0$, isto é, $C = \lambda C$.*

Se C é um cone com vértice em 0 então $x_0 + C$ é chamado cone de vértice x_0.

Denotamos por X^* o dual topológico do espaço vetorial X, isto é, X^* é o conjunto de todos os funcionais lineares e contínuos, definidos sobre X, um espaço de Banach com a norma

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

Definição 2.2 *Seja C um cone em X com vértice 0. O conjunto C^*, de todos os funcionais $f \in X^*$ tais que $f(x) \leq 0$ para todo $x \in C$, é chamado **cone dual** (ou polar) de C.*

Definição 2.3 *Um cone C em X com vértice 0 é dito ser **convexo** se para quaisquer $x, y \in C$ e $t \in [0, 1]$, tem-se:*

$$tx + (1 - t)y \in C.$$

Observação 2.1 *Se C é um cone com vértice 0 em X, temos que:*

i. *Se $C = \{0\}$ então $C^* = X^*$.*

ii. *C^* é um cone com vértice em 0.*

iii. *Se C é um subespaço arbitrário de um espaço de Banach X então seu cone dual é*

$$C^* = \left\{f \in X^* : f(x) = 0, \forall x \in C\right\}.$$

*e neste caso, C^* é chamado **subespaço anulador** do subespaço C.*

iv. *Sejam $f \in X^*$, $C_1 = \left\{x \in X : f(x) = 0\right\}$, $C_2 = \left\{x \in X : f(x) \geq 0\right\}$, e $C_3 = \left\{x \in X : f(x) > 0\right\}$. Então, $C_1^* = \left\{\lambda f, -\infty < \lambda < \infty\right\}$, $C_2^* = \left\{\lambda f, \lambda \geq 0\right\}$, $C_3^* = X^*$ se, $f = 0$ e $C_3^* = C_2^*$ se $f \neq 0$.*
Veja [47], Teorema 10.2.

Definição 2.4 *Um vetor $w \in X$ é uma **direção de descida** do funcional linear f no ponto x_0 se existem uma vizinhança V de w, $\alpha < 0, \epsilon_0 > 0$, tais que, para todo $0 < \epsilon < \epsilon_0$ e qualquer $\overline{w} \in V$,*

$$f(x_0 + \epsilon\overline{w}) \leq f(x_0) + \epsilon\alpha.$$

As direções de descida geram um cone aberto com vértice 0, $C_D(f, x_0)$, chamado cone de descida do funcional f no ponto x_0. O funcional f é dito ser **regular de descida** no ponto x_0 se o cone $C_D(f, x_0)$ é convexo.

Observação 2.2 *Se f é diferenciável em x_0 então f é regular de descida em x_0 e $C_D(f, x_0) = \{w : \nabla f(x_0)w < 0\}$. Veja [47], Teorema 7.4.*

Definição 2.5 *Sendo K_D um conjunto de restrições de desigualdade, dizemos que o vetor w é uma* ***direção factível*** *para K_D no ponto x_0 se existe uma vizinhança V de w tal que, para todo $0 < \epsilon < \epsilon_0$ e todo $\overline{w} \in V$, $x_0 + \epsilon\overline{w} \in K_D$. Veja Figura 2.1.1.*

O conjunto das direções factíveis gera um cone aberto com vértice 0, $C_F(Q, x_0)$, chamado cone factível para K_D em x_0. Uma restrição K_D de desigualdade é **regular** no ponto x_0 se o cone $C_F(K_D, x_0)$ é convexo.

Se assumirmos $K_D = \{x \in X : f(x) \leq 0\}$, f diferenciável e $\nabla f(x_0) \neq 0$ então $C_D(f, x_0)$ e $C_F(K_D, x_0)$ são convexos e verificam

$$C_D(f, x_0) = C_F(K_D, x_0) = \{w : \nabla f(x_0)w < 0\}.$$

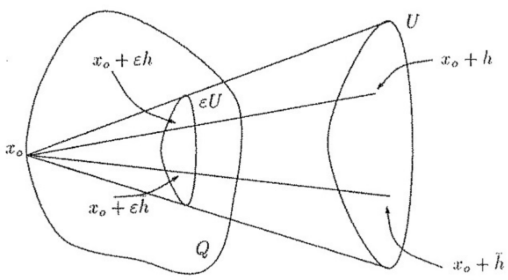

Figura 2.1: Direção factível.

Definição 2.6 *Seja K_I tal que* ***int***$K_I = \emptyset$ *um conjunto de restrições de igualdade. Dizemos que o vetor w é uma* ***direção tangente*** *a K_I no ponto x_0 se para qualquer $0 < \epsilon < \epsilon_0$ existe um ponto $x_\epsilon \in K_I$ tal que, se $x_\epsilon = x_0 + \epsilon w + r(\epsilon)$ então o vetor $r(\epsilon) \in X$ é tal que, para qualquer vizinhança V de zero, $\frac{1}{\epsilon}r(\epsilon) \in V$ para todo $\epsilon > 0$ suficientemente pequeno, ou equivalentemente, $\|r(\epsilon)\| = o(\epsilon)$. Veja Figura 2.1.1.*

O conjunto das direções tangentes é um cone com vértice em 0, $C_T(K_I, x_0)$ que será chamado por cone tangente a K_I no ponto x_0. Além disso, dizemos que a restrição de igualdade K_I é **regular** no ponto x_0 se $C_T(K_I, x_0)$ é convexo. Notemos ainda que o cone $C_T(K_I, x_0)$ não é, em geral, aberto nem fechado.

Observação 2.3 *Em geral se $A_1, \ldots, A_n$ são subconjuntos de X tal que $x_0 \in \bigcap_{i=1}^{n} A_i$ então*

$$\bigcap_{i=1}^{n} C_T(A_i, x_0) \supseteq C_T\left(\bigcap_{i=1}^{n} A_i, x_0\right).$$

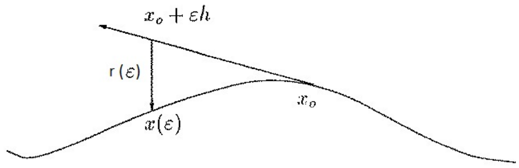

Figura 2.2: Direção tangente.

Temos ainda que w é uma direção tangente ao conjunto K_I no ponto x_0 se, e somente se, para toda vizinhança V de w e qualquer número real $\epsilon_0 > 0$, existe $y \in V$ e $\epsilon \in (0, \epsilon_0)$ tal que $x_0 + \epsilon y \in K_I$.

Apresentamos agora o Teorema de Lyusternik, dado que ele é uma importante ferramenta para determinar os cones de direções tangentes. Sua prova pode ser encontrada em [8].

Teorema 2.7 *[**Lyusternik** [8]] Sejam X e Y dois espaços de Banach, V uma vizinhança do ponto x_0 em X, $f : V \to Y$ e $f(x_0) = 0$. Se f é estritamente diferenciável* [1] *no ponto x_0 e $\nabla f(x_0)$ é sobrejetora, então o conjunto de direções tangentes $C_T(Q, x_0)$ para o conjunto $K_I = \{x \in X : f(x) = 0\}$ no ponto x_0 é o subespaço $C_T(K_I, x_0) = \{w \in X : \nabla f(x_0)w = 0\}$.*

Observemos que não necessitamos que a função f seja de classe C^1, mas apenas estritamente diferenciável.

Observação 2.4 *Se:*

i. *$x_0 \in$ **int**Q então $C_F(Q, x_0) = X$, portanto é de interesse unicamente considerar x_0 na fronteira de Q, isto é, $x_0 \in$ **Fr**(Q).*

ii. *Q é determinado por um funcional $g : X \to \mathbb{R}^m$, isto é, $Q = \{x \in X : g(x) \leq 0\}$ então o cone das direções de descida de g sempre está incluído no cone das direções factíveis de Q.*

iii. *Seja $Q = \{x \in X : g_i(x) \leq 0,\ i = 1, \dots, m\}$ é determinado por um funcional $g_i : X \to \mathbb{R}$ que é diferenciável em x_0 e $\nabla g_i(x_0) \neq 0$ então o cone factível a Q no ponto x_0 em **Fr**(Q) coincide com a interseção dos cones de direções de descida de $g_i, i = 1, \dots, m$ para o ponto $x_0 \in$ **Fr**(Q), isto é,*

$$C_F(Q, x_0) = \{w \in X : \nabla g_i(x_0)w < 0, i \in I(x_0)\} \tag{2.1}$$

onde $I(x_0) = \{i \in I : g_i(x_0) = 0\}$ é o conjunto dos índices das restrições ativas de g, com $I = \{i : i = 1, \dots, m\}$.

---

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

$$\|\phi(x_1) - \phi(x_2) - \Lambda(x_1 - x_2)\| \leq \epsilon \|x_1 - x_2\|.$$

Note que toda função estritamente diferenciável em x_0 é diferenciável em x_0. Com efeito tomando $x_2 = x_0$ e $x_1 = x_0 + h$ na igualdade acima, obtemos que a função é diferenciável e $\Lambda = \nabla\phi(x)$.

iv. $C_F(Q, x_0)$ *é definido como em (2.1) então é aberto e convexo.*

v. $A_1, \dots, A_n$ *são subconjuntos de* X *tal que* $x_0 \in \bigcap_{i=1}^{n} A_i$ *então*

$$\bigcap_{i=1}^{n} C_F(A_i, x_0) = C_F\left(\bigcap_{i=1}^{n} A_i, x_0\right).$$

Para maiores detalhes veja [47].

Para obter uma condição suficiente para garantir que o cone $C_F\left(\bigcap_{i=1}^{n} A_i, x_0\right)$ seja não vazio, necessitamos do seguinte conceito.

Definição 2.8 *[77] Seja* X *um espaço de Banach e* $g : X \to \mathbb{R}^n$ *uma função diferenciável em* $x_0 \in X$*. Dizemos que* x_0 *é um* ***ponto estacionário vetorial*** *de* g *se existe* $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n, \alpha \neq 0, \alpha_j \geq 0, j = 1, \dots, n,$ *tal que* $\nabla g(x_0)^T \alpha = 0$.

O lema seguinte é um caso particular do Teorema de Motzkin (veja Teorema 3.4.2, [33]).

Lema 2.9 ***(Craven [33])****Seja* X *um espaço vetorial normado e* $A : X \to \mathbb{R}^n$ *uma função linear e contínua, então exatamente uma das seguintes alternativas ocorre:*

- $\exists x \in X : Ax < 0$, *ou*
- $\exists p = (p_1, \dots, p_n) \neq 0, p \in \mathbb{R}^n, p_j \geq 0, A^T p = 0.$

E como consequência imediata do Lema 2.9, temos o seguinte Lema.

Lema 2.10 *Seja* $g : X \to \mathbb{R}^n$ *uma função diferenciável em* $x_0 \in X$ *tal que* x_0 *é um ponto não estacionário vetorial de* g*, então existe* $w \in X$*,* $\nabla g(x_0) w < 0$*, isto é,*

$$\exists w \in X : \nabla g_j(x_0) w < 0, \forall j = 1, \dots, n.$$

O seguinte resultado é de fundamental importância no formalismo de DM, e cuja demonstração - baseada no Teorema de Separação de Hahn-Banach - é encontrada em [47], Lema 5.11.

Lema 2.11 ***(Girsanov [47])*** *Sejam* $C_0, \dots, C_{k+1}$ *cones convexos com vértice 0, onde* $C_0, \dots, C_k$ *são abertos. Então* $\bigcap_{i=0}^{k+1} C_i = \emptyset$ *se, e somente se, existem funcionais lineares* $f_i \in C_i^*, i = 0, \dots, k+1,$ *não todos nulos, tais que*

$$f_0 + f_1 + \ldots + f_{k+1} = 0.$$

Teorema 2.12 ***(Dubovitskii-Milyutin,** [47])* *Suponha que o funcional f assume um mínimo local em $Q = \bigcap_{i=1}^{k+1} Q_i$ no ponto $x_0 \in Q$. Suponha ainda que f é regular de descida em x_0, com direções de descida no cone C_0; as restrições de desigualdade Q_i, $i = 1, \dots, k$ são regulares em x_0, com direções factíveis no cone C_i, a restrição de igualdade Q_{k+1} também é regular em x_0, com direções tangentes no cone C_{k+1}. Então existem funcionais lineares contínuos $f_i \in C_i^*$, $i = 0, \dots, k+1$, não todos nulos, que satisfazem a equação de Euler-Lagrange*

$$\sum_{i=0}^{k+1} f_i = 0. \tag{2.2}$$

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

$$\bigcap_{i=0}^{k+1} C_i = \emptyset, \tag{2.3}$$

ou seja, nenhuma direção de descida do funcional f pode ser factível para todas as restrições. Suponha que isto é falso. Então existe $w \in C_i$, $i = 0, \dots, k+1$. Pela definição de C_i, para $i = 0, \dots, k$, existe uma vizinhança V de w tal que, quando $0 < \varepsilon_0 < \varepsilon$ e $\overline{w} \in V$, qualquer vetor

$$x_0 + \varepsilon\overline{w} \in \bigcap_{i=0}^{k} C_i$$

e satisfaz a desigualdade

$$f(x_0 + \varepsilon\overline{w}) \leq f(x_0) + \varepsilon\alpha$$

para algum $\alpha < 0$.

Agora considere o vetor $x(\varepsilon) = x_0 + \varepsilon w + r(\varepsilon) \in Q_{k+1}$ como na definição de direção tangente, e seja ε_1 tal que $\frac{1}{\varepsilon}r(\varepsilon) \in V - w$, ou seja, $\overline{w}(\varepsilon) = w + \frac{1}{\varepsilon}r(\varepsilon) \in V$ para $0 < \varepsilon \leq \varepsilon_1$.

Então quando $0 < \varepsilon < \min\{\varepsilon_0, \varepsilon_1\}$ os vetores

$$x(\varepsilon) = x_0 + \varepsilon\overline{w}(\varepsilon) = x_0 + \varepsilon w + r(\varepsilon)$$

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

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

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

$$\bigcap_{i=0}^{k+1} C_i = \emptyset.$$

Além disto, C_i, para $i = 0, \ldots, k$, são cones abertos convexos e C_{k+1} é um cone convexo. Obtivemos assim as hipóteses necessárias para a utilização do Lema 2.11 e isso implica o resultado desejado. □

Apesar de conhecida, a prova acima foi reproduzida com o intuito de aclarar a Observação 2.5 abaixo, sobretudo Observação 2.5 - item i.

O teorema anterior proporciona condições necessárias de otimalidade, as quais são uma generalização da regra dos multiplicadores de Lagrange para problemas de programação linear e a equação de Euler-Lagrange no cálculo variacional. Devemos notar que para obter tais condições necessárias para algum problema específico, de qualquer natureza, é preciso determinar as direções de descida, factíveis e tangentes e seus respectivos cones duais. Esta caracterização está relacionada com as hipóteses impostas aos funcionais envolvidos no problema, como diferenciabilidade, regularidade, convexidade, etc.

Observação 2.5 *Nas condições do Teorema 2.12 seguem os seguintes resultados:*

i. Uma condição suficiente para assegurar que $f_0 \neq 0$ é que exista ao menos uma direção factível e uma direção tangente, isto é,

$$\bigcap_{i=1}^{k+1} C_i \neq \emptyset.$$

A conclusão é válida para qualquer $f_i, i = 1, \ldots, k+1$.

ii. A condição $\bigcap_{i=0}^{k+1} C_i = \emptyset$ é uma condição necessária para que o ponto x_0 seja uma solução ótima de f e é uma condição necessária e suficiente para obter a equação de Euler-Lagrange, isto é,

$$x_0 \text{ é um ponto de mínimo } \Rightarrow \bigcap_{i=0}^{k+1} C_i = \emptyset \Leftrightarrow \text{ equação de Euler-Lagrange.}$$

Para maiores detalhes consulte [47].

2.2 Formalismo de Dubovitskii-Milyutin: Multiobjetivo

No estudo dos problemas de otimização é comum o surgimento de problemas para os quais se quer otimizar, simultaneamente, múltiplos objetivos que frequentemente encontram-se em conflito entre si, o decrescimento de um dos objetivos pode acarretar no crescimento de algum (ou vários) outro(s) objetivo(s). O estudo deste tipo de problema vem tendo avanços nos últimos anos, tanto no ponto de vista teórico, numérico e de aplicações.

A solução obtida em um problema multiobjetivo dependerá da *"noção de equilíbrio"* utilizada para solucionar os conflitos que surgem da consideração simultânea de vários objetivos. A noção adotada neste trabalho é a noção de equilíbrio de Pareto, introduzida pelo economista *Vilfredo Pareto* (1848-1923).

2.2.1 Definições e resultados

As definições apresentadas nesta seção são válidas a problemas mais gerais. Entretanto optamos por apresentá-las como abaixo, visto que estamos interessados na utilização do formalismo de DM.

Consideremos inicialmente o seguinte problema de otimização multiobjetivo:

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

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

No que segue, apresentamos a notação para comparação de vetores que será utilizada a partir deste momento.

Para $x, y \in \mathbb{R}^n$ com $n \geq 2$, escreveremos

$$\begin{aligned} x &< y \iff x_i < y_i,\ i = 1, \cdots, n\\ x &\leqq y \iff x_i \leq y_i,\ i = 1, \cdots, n\\ x &\leq y \iff x \leqq y \text{ e } x \neq y. \end{aligned}$$

Observemos ainda que em se tratando de desigualdade escalar, isto é, $n = 1$, não faremos distinção entre " $\leq$ " e " $\leqq$ ".

Vários conceitos de solução podem ser associados ao problema (MOP). Neste trabalho abordaremos os seguintes:

- **Solução Pareto fraca:** $x_0 \in \mathbb{R}^n$ é solução Pareto fraca (ou solução fracamente eficiente) de (MOP) se $x_0 \in Q$ e não existe $x \in Q$, $x \neq x_0$ tal que $f_j(x) < f_j(x_0)$ para $j = 1, \dots, p$.

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

- **Solução propriamente Pareto:** $x_0 \in \mathbb{R}^n$ é solução propriamente Pareto de (MOP) se ela é eficiente e se existe $M > 0$ tal que, para cada $1 \leq k \leq p$

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

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

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

É imediato das definições: propriamente Pareto $\Longrightarrow$ Pareto $\Longrightarrow$ Pareto fraca.

As recíprocas são falsas, como mostram os seguintes exemplos:

Exemplo 2.13 *Moulin e Soulié [73].*

Considere o seguinte problema multiobjetivo:

$$\text{Maximizar } f(x) = (f_1(x), f_2(x))$$
$$\text{sujeito a: } x \in X.$$

sendo $x = (x_1, x_2)$, $f_1(x) = x_1, f_2(x) = x_2$ *e* $X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\} \cup [-1, 0] \times [0, 1]$. *(Como é usual, o problema de maximizar* $f(x)$ *sobre* $x \in X$ *é equivalente a minimizar* $-f(x)$ *sobre* $x \in X$.*)*

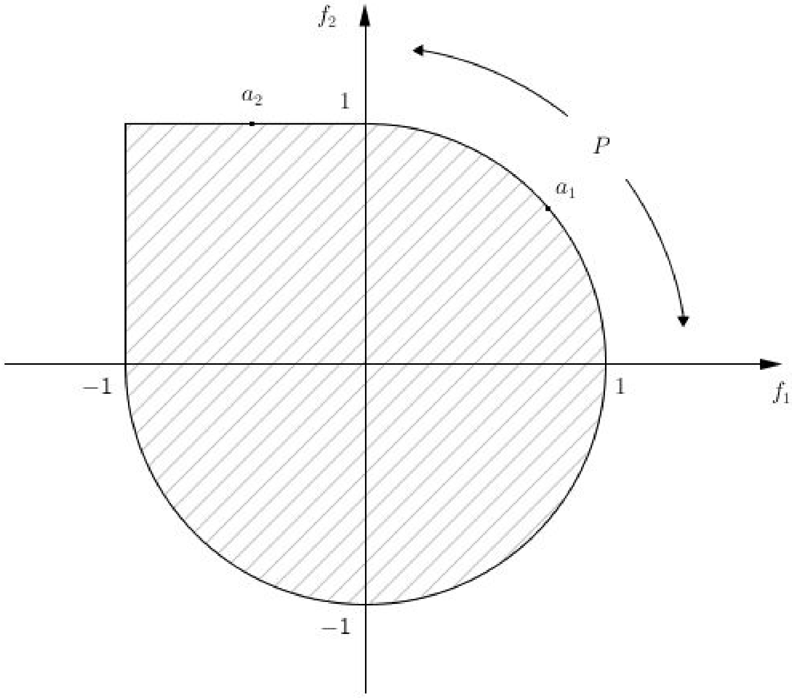

Figura 2.3: Conjunto factível X

Observe que o ponto $a_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ *não pode aumentar (estritamente) o valor de* f_1, *sem diminuir estritamente* f_2. *Assim sendo,* a_1 *é solução Pareto eficiente do problema. Mais geralmente, o conjunto* $\mathcal{P}$ *das soluções Pareto eficientes do problema é o arco de circunferência entre* $(0, 1)$ *e* $(1, 0)$. *Veja figura.*

Note que o ponto $a_2 = (-\frac{1}{2}, 1)$ *não pode aumentar os dois objetivos,* f_1 *e* f_2, *simultaneamente. Deste modo,* a_2 *é solução fracamente Pareto eficiente do problema. Mais geralmente, o conjunto* $\mathcal{P}_f$ *das soluções fracamente Pareto eficientes do problema é* $\mathcal{P}_f = \mathcal{P} \cup \{(\lambda, 1) : -1 \leq \lambda \leq 0\}$.

Exemplo 2.14 *Considere o problema:*

$$\text{Minimizar } f(x) = (f_1(x), f_2(x))$$
$$\text{sujeito a: } x \geq 0, \; x \in \mathbb{R}.$$

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

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

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

2.2.2 Formalismo Dubovitskii-Milyutin

Nesta seção abordaremos as condições de otimalidade para problemas multiobjetivo utilizando o conceito de solução Pareto (e Pareto fraca). Para provas dos resultados e maiores detalhes veja [31], Capítulo 5 e também [25].

Solução Pareto Fraca

O próximo teorema é uma generalização do Teorema de Dubovitskii-Milyutin (Teorema 2.12) ao problema multiobjetivo (MOP) para a obtenção de condições necessárias para que um ponto seja solução Pareto fraca. No caso de $p = 1$, o problema é escalar e retornamos ao Teorema 2.12.

Teorema 2.15 ***(Dubovitskii-Milyutin Multiobjetivo: solução Pareto fraca [31])*** *Sejam* $f : X \to \mathbb{R}^p, C_{D_j}$ *o cone de direções de descida de* $f_j(x)$ *em* x_0 *para* $j = 1, \ldots, p$, Q_i *o conjunto de restrições com* $\textbf{int}Q_i \neq \emptyset$ *para* $i = 1, \ldots, k$, Q_{k+1} *o conjunto de restrições com* $\textbf{int}Q_{k+1} = \emptyset$, *com* $x_0 \in Q = \bigcap_{i=1}^{k+1} Q_i, C_{E_i}$ *o cone de direções factíveis de* Q_i *em* x_0 *para* $i = 1, \ldots, k$ *e* $C_{E_{k+1}}$ *o cone de direções tangentes de* Q_{k+1} *em* x_0. *Suponha que os cones* C_{D_j}, C_{E_i}, *para* $j = 1, \ldots, p$ *e* $i = 1, \ldots, k+1$ *são todos convexos. Se* x_0 *é solução Pareto fraca local para o problema (MOP) então*

$$\sum_{j=1}^{p} h_j + \sum_{i=1}^{k+1} l_i = 0;$$

onde $h_j \in C^*_{D_j}$ *para* $j = 1, \ldots, p$ *e* $l_i \in C^*_{E_i}$ *para* $i = 1, \ldots, k+1$ *são funcionais lineares, não todos identicamente nulos.*

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

Teorema 2.16 *Sejam* $f_j : X \to \mathbb{R}$ *funções convexas para todo* $j = 1, \ldots, p$, C_{D_j} *o cone de direções de descida de* $f_j(x)$ *em* x_0 *para* $j = 1, \ldots, p$, C_{E_i} *o cone de direções factíveis de* Q_i *em* x_0 *para* $i = 1, \ldots, k$ *e* $C_{E_{k+1}}$ *o cone de direções tangentes de* Q_{k+1} *em* x_0. *Suponha que os conjuntos de restrições* Q_i *são convexos para todo* $i = 1, \ldots, k+1$ *e* $Q_{k+1} \bigcap \left(\bigcap_{i=1}^{k} \textbf{int}Q_i \right) \neq \emptyset$. *Então,* x_0 *é solução Pareto fraca para o problema (MOP) se, e somente se,*

$$\sum_{j=1}^{p} h_j + \sum_{i=1}^{k+1} l_i = 0;$$

onde $h_j \in C^*_{D_j}$ para $j = 1, \ldots, p$ e $l_i \in C^*_{E_i}$ para $i = 1, \ldots, k+1$ são funcionais lineares, não todos identicamente nulos.

As provas dos Teoremas 2.15 e 2.16 são análogas à demonstração do Teorema 2.12 e portanto são omitidas.

Solução Pareto

Para abordarmos o problema multiobjetivo considerando a otimalidade de Pareto, utilizando o formalismo de DM, precisamos do conceito de direção de não crescimento.

Definição 2.17 *Um vetor w é dito uma* **direção de não crescimento** *do funcional f no ponto x_0 se existem uma vizinhança V de w e $\overline{\alpha} > 0$ tais que para todo $\alpha \in (0, \overline{\alpha})$ e qualquer $\overline{w} \in V$*

$$f(x_0 + \alpha\overline{w}) \leq f(x_0).$$

Lema 2.18 ***(Censor [25])*** *As direções de não crescimento geram um cone aberto com vértice na origem, que contém o cone de direções de descida no mesmo ponto.*

A demonstração deste resultado pode ser encontrada também em [56].

Denotaremos por C_{D_j} o cone de direções de descida para $f_j(x)$ no ponto x_0, C_{N_j} o cone de direções de não crescimento para $f_j(x)$ em x_0 para $j = 1, \ldots, p$, C_{E_i} o cone de direções factíveis para Q_i, $i = 1, \ldots, k$ em x_0 e $C_{E_{k+1}}$ o cone de direções tangentes para Q_{k+1} em x_0.

Além disso, temos o seguinte:

Lema 2.19 ***(Kotarski [56], Lema 4.1)*** *Se x_0 é uma solução Pareto ótima de (MOP) então*

$$C_{D_r} \cap \left(\bigcap_{\substack{j=1 \\ i \neq r}}^{p} C_{N_j} \right) \cap \left(\bigcap_{i=1}^{k} C_{E_i} \right) \cap C_{E_{k+1}} = \emptyset. \tag{2.4}$$

Veja também [77], Lema 69.

O seguinte teorema nos fornece as condições necessárias de otimalidade local de Pareto para o problema de otimização multiobjetivo (MOP). Quando aplicado ao caso particular $p = 1$, coincide novamente com o Teorema de Dubovitskii-Myliutin (Teorema 2.12).

Teorema 2.20 ***(Dubovitskii-Milyutin Multiobjetivo: solução de Pareto [56])*** *Sejam $f : X \to \mathbb{R}^p$, Q_i o conjunto de restrições com $\textbf{int}Q_i \neq \emptyset$ para $i = 1, \ldots, k$, Q_{k+1} o conjunto de restrições com $\textbf{int}Q_{k+1} = \emptyset$ e $x_0 \in Q = \bigcap_{i=1}^{k+1} Q_i$. Suponha que os cones C_{D_j}, C_{N_j}, C_{E_i}, para $j = 1, \ldots, p$ e $i = 1, \ldots, k+1$ são todos convexos. Se x_0 é Pareto ótimo local para o problema (MOP) então as seguintes p equações devem valer*

$$h_j + \sum_{\substack{l=1 \\ l \neq j}}^{p} h_l^{(j)} + \sum_{i=1}^{k+1} l_i^{(j)} = 0, \qquad j = 1, \ldots, p, \tag{2.5}$$

onde $h_j \in C^*_{D_j}, h_l^{(j)} \in C^*_{N_l}$ *para* $l = 1, \dots, p,\ l \neq j$ *e* $l_i^{(j)} \in C^*_{E_i}$ *para* $i = 1, \dots, k+1$ *são para todo* $j = 1, \dots, p$ *funcionais lineares, não todos identicamente nulos.*

Observação 2.6 *Se para algum índice* j $(1 \leq j \leq p)$, C_{D_j} *é vazio então* $C^*_{D_j} = \mathbb{R}^n$ *e pode-se escolher algum* $h_l^{(j)} \in C^*_{N_l}$ *arbitrário para* $l = 1, \dots, p,\ l \neq j$ *e algum* $l_i^{(j)} \in C^*_{E_i}$ *arbitrário, para* $i = 1, \dots, k+1$ *e colocar*

$$h_j = -\left[\sum_{\substack{l=1 \\ l \neq j}}^{s} h_l^{(j)} + \sum_{i=1}^{k+1} l_i(j) \right].$$

Então $h_j \in C^*_{D_j}$ *e a condição necessária para Pareto ótimo local de* x_0 *é satisfeita para tal* j*. Se para algum índice* l, $C_{N_l} = \emptyset$ *então o mesmo raciocínio é verdadeiro para todo* $j = 1, \dots, p$ *e a condição necessária para Pareto ótimo local de* x_0 *é satisfeita para todo* $j = 1, \dots, p$.

Corolário 2.21 *(**Orellana-Vivanco [77]**) Se além das hipóteses do Teorema 2.20,* f *é tal que para todo* $j = 1, \dots, p$, $C^*_{D_j} = C^*_{N_j}$*, então as* p *equações da conclusão se reduzem à seguinte equação*

$$\sum_{j=1}^{p} h_j + \sum_{i=1}^{k+1} l_i = 0,$$

onde $h_j \in C^*_{D_j}$ *para* $j = 1, \dots, p$ *e* $l_i \in C^*_{E_i}$ *para* $i = 1, \dots, k+1$*, são funcionais lineares, não todos identicamente nulos.*

Uma condição suficiente para assegurar que a igualdade $C^*_{D_j} = C^*_{N_j}$ se cumpra é que para todo $j = 1, \dots, p$, $f_j : X \to \mathbb{R}$ seja Ponstein convexa[1] e que $\inf f_j(x) < f_j(x_0)$, veja Lema 1.4.3., em Kotarski [57]. Ademais, sob hipóteses de diferenciabilidade temos o seguinte lema.

Lema 2.22 *(**Orellana-Vivanco [77]**) Seja* $f(x) : X \to \mathbb{R}^p, f(x) = (f_1(x), \dots, f_p(x))$ *diferenciável em* x_0*, tal que* x_0 *é um ponto não estacionário vetorial de* f*, então os cones duais das direções de descida e de não crescimento dos funcionais* f_i*, no ponto* x_0*, coincidem e além disso é possível determiná-los explicitamente, isto é,*

$$C^*_{D_j} = C^*_{N_j} = \left\{ h_j \in X^* : h_j = -\nu_j \nabla f_j(x_0), \nu_j \geq 0 \right\}, j = 1, \dots, p. \tag{2.6}$$

Para finalizar observemos que:

Observação 2.7 *Seja* $h_j = -\nabla f_j(x_0)^T \nu_j = -\nu_j \nabla f_j(x_0), \nu_j \geq 0,$

1. *se* $h_j = 0$ *então:*

$$h_j \in C^*_{D_j} = C^*_{N_j}, \ \textit{se } \nabla f_j(x_0) \neq 0,$$

$$h_j \in C^*_{N_j} \subsetneqq C^*_{D_j}, \ \textit{se } \nabla f_j(x_0) = 0,$$

[1] A função $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ será chamada **Ponstein-convexa** se

$$f(x_2) \leq f(x_1) \Rightarrow f(\lambda x_1 + \mu x_2) < f(x_1)$$

sempre que $x_1 \neq x_2$, onde λ e μ são números reais positivos sendo $\lambda + \mu = 1$.

2. *se* $h_j \neq 0$ *então* $h_j \in C_{D_j}^* = C_{N_j}^*$, *dado que,*

$$h_j(w) < 0, \forall w \in C_{D_j} \subsetneqq C_{N_j},$$

$$h_j(w) = 0, \forall w \in C_{N_j} \setminus C_{D_j},$$

3. *se* $\exists w \in X, \nabla f_j(x_0)w < 0, \forall j = 1, \ldots, p$ *então*

$$C_D = \bigcap_{j=1}^{p} C_{D_j} = \bigcap_{j=1}^{p} \{w : \nabla f_j(x_0)w < 0\} \neq \emptyset,$$

portanto, como $C_{D_j}, j = 1, \ldots, p$, *são convexos e abertos, cuja interseção é não vazia temos, devido ao Lema 5.10 de [47],*

$$D^* = \left(\bigcap_{j=1}^{p} C_{D_j} \right)^* = \sum_{j=1}^{p} \left(C_{D_j}\right)^*.$$

Do exposto anteriormente e considerando (2.6) obtemos

$$C_D^* = C_N^* = \left\{ h_0 \in X^* : h_0 = -\sum_{j=1}^{p} \nu_j \nabla f_j(x_0), \nu_j \geq 0 \right\}.$$

Parte I

Caso Regular

Capítulo 3

Problema de Controle Ótimo Discreto Mono-objetivo

Neste capítulo formularemos o problema de controle ótimo discreto, com restrições de igualdade e desigualdade tanto no controle quanto no estado, ao qual chamaremos de (PCD) e uma versão deste na qual os controles pertencem a um subconjunto arbitrário do $\mathbb{R}^{n(N)}$ que denotaremos por $(\overline{PCD})$. Então propomos condições de otimalidade, as necessárias por meio de uma versão do Teorema do Princípio do Máximo para problemas de controle ótimo discreto e as suficientes via o conceito de PM-invexidade.

A obtenção do Teorema do Princípio do Máximo é baseada no formalismo de DM. O ganho obtido utilizando esta ferramenta teórica é que se torna possível obter as condições necessárias e suficientes de otimalidade de uma maneira mais clara. Além disso, apresentamos conceitos de regularidade que são imprescindíveis na obtenção das condições de otimalidade.

3.1 Formulação do problema de controle ótimo discreto

De uma forma geral, o problema de controle discreto admite a seguinte formulação:

$$\begin{aligned} &\text{Minimizar } \mathbb{J}(x,u) = \sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i\right) \\ &\text{sujeito a: } x_{i+1} = \varphi_i\left(x_i, u_i\right),\ i = 0, \ldots, N-1, \\ &x_0 = 0 \\ &u_i \in \Omega_i \end{aligned} \tag{3.1}$$

onde $\Phi_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \varphi_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ para $i = 1, \ldots, N-1$, são funções continuamente diferenciáveis, $X = \mathbb{R}^{n(N+1)} = \{(x_0, \ldots, x_N) : x_i \in \mathbb{R}^n, \forall i = 0, \ldots, N\}$, $\Omega_i \subset \mathbb{R}^m, i = 0, \ldots, N-1$ e $U = \mathbb{R}^{m(N)} = \{(u_0, \ldots, u_{N-1}) : u_i \in \mathbb{R}^m, \forall i = 0, \ldots, N-1\}$. A partir de agora, assumimos que

$$\begin{array}{cc} x : [0, N] \to \mathbb{R}^n & u : [0, N-1] \to \mathbb{R}^m \\ x(i) = x_i & u(i) = u_i \end{array}$$

- x_i é a variável de estado;
- u_i é o parâmetro de controle;
- $[0, N]$ é o intervalo discreto da reta real, com $N \geq 0$ o número de etapas (passos) realizados;
- $x \in \mathbb{R}^{n(N+1)}$ é a trajetória;
- $u \in \mathbb{R}^{m(N)}$ é o controle associado à trajetória correspondente.

Para estabelecer as condições de otimalidade via o formalismo DM assumimos, por uma questão de conveniência, que quando nos referirmos ao par (x, u) está subentendido que na realidade estamos nos referindo ao par (x_0, u) já que os demais vetores $x_i, i = 1, \ldots, N$, são obtidos através da relação $x_{i+1} = \varphi_i\,(x_i, u_i)$.

Dizemos que o par (x, u) é um **processo factível** de (3.1) se satisfaz as restrições do referido problema. E dizemos que o processo factível $(\hat{x}, \hat{u})$ é um **processo ótimo** se

$$\mathbb{J}(\hat{x}, \hat{u}) \leq \mathbb{J}(x, u)$$

para todo processo factível (x, u) de (3.1).

Exemplo 3.1 *(Boltyanskii, [14]) Um foguete espacial deve ser lançado do solo. Desejamos projetar os N-estágios (de lançamento) deste foguete cujo peso inicial é G. Sabemos que o peso da espaçonave (acoplada neste foguete) é H e que ela deve ser colocada em órbita após a última etapa de lançamento.*

Cada estágio do foguete tem um fornecimento de combustível e durante a operação de um estágio (do momento em que o motor de ignição é ligado até o deslocamento da fase "queimada") o foguete adquire uma velocidade adicional Δv, a qual é uma função de "payload" (ou carga útil) P, transportada por este estágio e o peso Q do próprio estágio (determinando a quantidade de combustível necessário):

$$\Delta v = f(P, Q).$$

Pretendemos encontrar a distribuição de peso entre os estágios de lançamento para a qual a velocidade da espaçonave (após o desprendimento do último estágio) deve ser máxima.

Se u_i é o peso no estágio i (contados a partir da espaçonave) então o peso do estágio final (o que coloca a espaçonave em órbita) é denotado u_1, logo, o peso do penúltimo estágio é u_2 e assim sucessivamente.

Se tomarmos o peso da espaçonave mais os estágios i ligados a ele temos $x_i, i = 0, \ldots, N$ (veja Figura 3.1) então, é óbvio que:

$$x_i = x_{i-1} + u_i, i = 1, \ldots, N. \tag{3.2}$$

Além disso, temos que o peso da espaçonave e o peso inicial do foguete são especificados pelas seguintes equações:

$$x_0 = H \qquad x_N = G.$$

Tendo em vista que necessitamos da velocidade adicional transmitida no estágio i, enquanto queima, obtemos:

$$\Delta v_i = f(x_{i-1}, u_i),$$

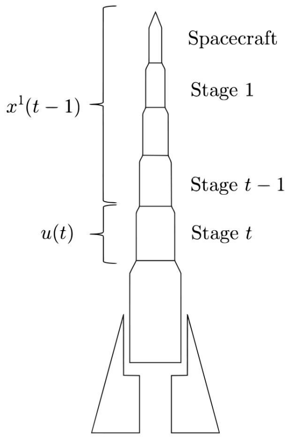

Figura 3.1: Foguete

onde u_i *é o peso do estágio* i *e* x_{i-1} *é o peso da carga que é transportada.*

Assim, a velocidade total transmitida para a espaçonave por todos os estágios de lançamento do foguete é:

$$\mathbb{J}(x,u) = \sum_{i=1}^{N} f(x_{i-1}, u_i). \tag{3.3}$$

Consequentemente, o problema da velocidade máxima da espaçonave pode ser formulado como:

$$\begin{aligned} &\textit{Maximizar } \mathbb{J}(x,u) = \sum_{i=1}^{N} f\left(x_{i-1}, u_i\right) \\ &\textit{sujeito a: } x_i = x_{i-1} + u_i, \;\; i = 1, \ldots, N, \\ &x_0 = H \\ &x_N = G \\ &u_i \geq b_i \end{aligned} \tag{3.4}$$

onde, $x = (x_0, \ldots, x_N)$ *corresponde a uma trajetória, cujo controle associado é* $u = (u_1, \ldots, u_N)$. *A restrição* $u_i \geq b_i$ *sob o parâmetro* u *é naturalmente imposta, embora não a mencionamos anteriormente.*

3.2 Problema (PCD)

Nesta seção formularemos o problema de controle ótimo discreto com restrições de igualdade e desigualdade no estado e no controle e aplicaremos o formalismo DM para obter as condições necessárias de otimalidade que serão apresentadas através de um Teorema do Princípio do Máximo. No que se segue, considere o problema de controle ótimo discreto ao qual denotaremos por Problema de Controle Discreto (PCD):

$$\begin{aligned}
&\text{Minimizar } \mathbb{J}(x,u) = \sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i\right) \\
&\text{sujeito a: } x_{i+1} = \varphi_i\left(x_i, u_i\right),\ i = 0, \ldots, N-1, \\
&h_i(x_i) = 0,\ g_i(x_i) \leq 0, i = 0, \ldots, N, \\
&\widetilde{h}_i(u_i) = 0,\ \widetilde{g}_i(u_i) \leq 0, i = 0, \ldots, N-1, \\
&K(x_0, x_N) = 0,
\end{aligned} \tag{PCD}$$

onde $\Phi_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \varphi_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, h_i : \mathbb{R}^n \to \mathbb{R}^{r_1}, g_i : \mathbb{R}^n \to \mathbb{R}^{s_1}$, $\widetilde{h}_i : \mathbb{R}^m \to \mathbb{R}^{r_2}, \widetilde{g}_i : \mathbb{R}^n \to \mathbb{R}^{s_2}$, $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^k$, são funções continuamente diferenciáveis para todo $i = 1, \ldots, N$. Assumiremos ainda que $r_1 \leq n$, $r_2 \leq m$ e $k \leq 2n$.

Denotamos o conjunto dos processos factíveis do problema (PCD) por

$$Q = Q_1 \bigcap Q_2 \tag{3.5}$$

onde Q_1 é o conjunto com interior vazio definido por

$$Q_1 = Q_1^1 \bigcap Q_1^2 \bigcap Q_1^3 \bigcap Q_1^4 \tag{3.6}$$

sendo

$$\begin{aligned}
Q_1^1 &= \left\{(x,u) \in X \times U : x_{i+1} = \varphi_i\left(x_i, u_i\right), i \in [0, N-1], \right\} \\
Q_1^2 &= \left\{(x,u) \in X \times U : h_i(x_i) = 0, i \in [0, N]\right\} \\
Q_1^3 &= \left\{(x,u) \in X \times U : \widetilde{h}_i(u_i) = 0, i \in [0, N-1]\right\} \\
Q_1^4 &= \left\{(x,u) \in X \times U : K(x_0, x_N) = 0\right\},
\end{aligned}$$

e Q_2 é o conjunto com interior não vazio definido por

$$Q_2 = Q_2^1 \bigcap Q_2^2 \tag{3.7}$$

sendo

$$\begin{aligned}
Q_2^1 &= \left\{(x,u) \in X \times U : g_i(x_i) \leq 0, i \in [0, N]\right\} \\
Q_2^2 &= \left\{(x,u) \in X \times U : \widetilde{g}_i(u_i) \leq 0, i \in [0, N-1]\right\}.
\end{aligned}$$

Observe que (x,u) é um processo factível do problema (PCD) se, e somente se, $(x,u) \in Q$.

Se consideramos

$$\begin{gathered}
\mathbb{J} : X \times U \to \mathbb{R} \\
\mathbb{J}(x,u) = \sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i\right)
\end{gathered} \tag{3.8}$$

$$\begin{gathered}
H : X \times U \to \mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k, \\
H(x,u) = (H^1(x,u), H^2(x,u), H^3(x,u), H^4(x,u))
\end{gathered} \tag{3.9}$$

onde

$$\begin{aligned}
&H^1_{i+1}(x_i,u_i) = \left(x^1_{i+1} - \varphi^1_i\left(x_i,u_i\right),\ldots,x^n_{i+1} - \varphi^n_i\left(x_i,u_i\right)\right)\\
&H^2_i(x_i,u_i) = (h^1_i(x_i),\ldots,h^{r_1}_i(x_i))\\
&H^3_i(x_i,u_i) = (\widetilde{h}^1_i(u_i),\ldots,\widetilde{h}^{r_2}_i(u_i))\\
\\
&H^1(x,u) = \left(x_1 - \varphi_0\left(x_0,u_0\right),\ldots,x_N - \varphi_{N-1}\left(x_{N-1},u_{N-1}\right)\right)\\
&H^2(x,u) = (h_0(x_0),\ldots,h_N(x_N))\\
&H^3(x,u) = (\widetilde{h}_0(u_0),\ldots,\widetilde{h}_{N-1}(u_{N-1})),\\
&H^4(x,u) = K(x_0,x_N),
\end{aligned}$$

$$G : X\times U \to \mathbb{R}^{s_1(N+1)}\times\mathbb{R}^{s_2(N)},\ G(x,u) = (G^1(x,u),G^2(x,u)) \tag{3.10}$$

onde

$$\begin{aligned}
&G^1_i(x_i,u_i) = (g^1_i(x_i),\ldots,g^{s_1}_i(x_i))\\
&G^2_i(x_i,u_i) = (\widetilde{g}^1_i(u_i),\ldots,\widetilde{g}^{s_2}_i(u_i))\\
\\
&G^1(x,u) = (g_0(x_0),\ldots,g_N(x_N)),\\
&G^2(x,u) = (\widetilde{g}_0(u_0),\ldots,\widetilde{g}_{N-1}(u_{N-1})),
\end{aligned}$$

podemos rescrever o problema (PCD) como o seguinte Problema de Otimização Escalar (POE)

$$\begin{gathered}
\text{Minimizar } \mathbb{J}(x,u)\\
\text{sujeito a: } H(x,u) = 0\\
G(x,u) \leqq 0\\
(x,u)\in X\times U.
\end{gathered} \tag{POE}$$

De acordo com as hipóteses impostas sobre as funções envolvidas no problema (PCD), temos que o funcional $\mathbb{J}$ e os operadores H e G são continuamente diferenciáveis, no sentido de Fréchet.

A primeira derivada de $\mathbb{J}$ em $(\hat{x},\hat{u})$ factível, definida para cada $(w,v)\in X\times U$, é dada por

$$\begin{gathered}
\nabla\mathbb{J}(\hat{x},\hat{u}) : X\times U \to \mathbb{R},\\
\nabla\mathbb{J}(\hat{x},\hat{u})(w,v) = \sum_{i=0}^{N-1} \nabla_x\Phi_i\left(x_i,u_i\right)w_i + \nabla_u\Phi_i\left(x_i,u_i,\right)v_i
\end{gathered} \tag{3.11}$$

A primeira derivada de H em $(\hat{x},\hat{u})$ factível, definida para cada $(w,v)\in X\times U$ é o operador

$$\nabla H(\hat{x},\hat{u}) : X\times U \to \mathbb{R}^{n(N)}\times\mathbb{R}^{r_1(N+1)}\times\mathbb{R}^{r_2(N)}\times\mathbb{R}^k$$

$$\nabla H(\hat{x},\hat{u})(w,v) = (\nabla H^1(\hat{x},\hat{u})(w,v),\nabla H^2(\hat{x},\hat{u})(w,v),\nabla H^3(\hat{x},\hat{u})(w,v),\nabla H^4(\hat{x},\hat{u})(w,v)) \tag{3.12}$$

onde

$$\begin{aligned}
&\nabla H^1_{i+1}(\hat{x}_i,\hat{u}_i)(w_i,v_i) = w_{i+1} - \left(\nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)w_i + \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)v_i\right), i = 0,\ldots,N-1\\
&\nabla H^2_i(\hat{x}_i,\hat{u}_i)(w_i,v_i) = \nabla h_i(\hat{x}_i)w_i, i = 0,\ldots,N\\
&\nabla H^3_i(\hat{x}_i,\hat{u}_i)(w_i,v_i) = \nabla\widetilde{h}_i(\hat{u}_i)v_i, i = 0,\ldots,N-1\\
&\nabla H^4(\hat{x},\hat{u})(w,v) = \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)w_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)w_N.
\end{aligned} \tag{3.13}$$

Observemos que para $i \in [0, N-1]$ fixado, $\nabla H^1_{i+1}(x_i, u_i) \in \mathbb{R}^n$ é dado por:

$$\frac{\partial H^1_{i+1}}{\partial x_i}(x_i, u_i) = -\nabla_x \varphi_i(x_i, u_i),$$

$$\frac{\partial H^1_{i+1}}{\partial x_{i+1}}(x_i, u_i) = \mathbf{I}_{n\times n},$$

$$\frac{\partial H^1_{i+1}}{\partial u_i}(x_i, u_i) = -\nabla_u \varphi_i(x_i, u_i),$$

sendo $\mathbf{I}_{n\times n}$ é a matriz identidade de ordem n.

A primeira derivada de G em $(\hat{x}, \hat{u})$ factível, definida para cada $(w, v) \in X \times U$ é o operador

$$\nabla G(\hat{x}, \hat{u}) : X \times U \to \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)}$$

$$\nabla G(\hat{x}, \hat{u})(w, v) = (\nabla G^1(\hat{x}, \hat{u})(w, v), \nabla G^2(\hat{x}, \hat{u})(w, v)) \tag{3.14}$$

onde

$$\begin{array}{l} \nabla G^1_i(\hat{x}_i, \hat{u}_i)(w_i, v_i) = \nabla g_i(\hat{x}_i) w_i, i = 0, \ldots, N \\ \nabla G^2_i(\hat{x}_i, \hat{u}_i)(w_i, v_i) = \nabla \widetilde{g}_i(\hat{u}_i) v_i, i = 0, \ldots, N-1. \end{array} \tag{3.15}$$

3.2.1 Construção dos cones

Para demonstrar o Princípio do Máximo Discreto, utilizando o formalismo DM, devemos determinar o cone das direções de descida do funcional objetivo $\mathbb{J}$, o cone das direções tangentes a Q_1 e o cone das direções factíveis de Q_2, todos calculados no ponto factível $(\hat{x}, \hat{u})$ para o problema (PCD), os quais denotaremos por $C_D(\mathbb{J}, (\hat{x}, \hat{u})), C_T(Q_1, (\hat{x}, \hat{u}))$ e $C_F(Q_2, (\hat{x}, \hat{u}))$ respectivamente e seus cones duais por $C_D(\mathbb{J}, (\hat{x}, \hat{u}))^*$, $C_T(Q_1, (\hat{x}, \hat{u}))^*$ e $C_F(Q_2, (\hat{x}, \hat{u}))^*$.

Construção do cone de descida $C_D(\mathbb{J}, (\hat{x}, \hat{u}))$

O cone das direções de descida de $\mathbb{J}(\hat{x}, \hat{u})$, de acordo com a Observação 2.2, é dado por

$$\begin{aligned} C_D(\mathbb{J}, (\hat{x}, \hat{u})) &= \{(w, v) \in X \times U : \nabla \mathbb{J}(\hat{x}, \hat{u})(w, v) < 0\} \\ &= \Bigg\{(w, v) \in X \times U : \\ &\quad \sum_{i=0}^{N-1} \nabla_x \Phi_i(\hat{x}_i, \hat{u}_i) w_i + \sum_{i=0}^{N-1} \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i) v_i < 0 \Bigg\}. \end{aligned} \tag{3.16}$$

Como $C_D(\mathbb{J}, (\hat{x}, \hat{u})) \neq \emptyset$, de acordo com a Observação 2.1 (iv), temos que o seu cone dual é dado por

$$\begin{aligned} C_D(\mathbb{J}, (\hat{x}, \hat{u}))^* &= \{f_0 \in (X \times U)^* : f_0(w, v) = -\lambda_0 \nabla \mathbb{J}(\hat{x}, \hat{u})(w, v), \lambda_0 \geq 0\} \\ &= \Bigg\{ f_0 \in (X \times U)^* : f_0(w, v) = -\lambda_0 \sum_{i=0}^{N-1} (\nabla_x \Phi_i(\hat{x}_i, \hat{u}_i) w_i + \\ &\quad \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i) v_i), \lambda_0 \geq 0 \Big\}. \end{aligned} \tag{3.17}$$

Construção do cone tangente $C_T(Q_1,(\hat{x},\hat{u}))$

Como o operador H é continuamente diferenciável por hipótese então $\nabla H(\hat{x},\hat{u})$, definido em (3.12), é contínuo. Além disso, assumimos que $\nabla H(\hat{x},\hat{u})$ é sobrejetivo. Assim, de acordo com o Teorema de Lyusternik (Teorema 2.7), o cone das direções tangentes ao conjunto Q_1 coincide com $\mathbf{Ker}\nabla H(\hat{x},\hat{u})$, isto é,

$$C_T(Q_1,(\hat{x},\hat{u})) \quad = \left\{(w,v) \in X \times U : \nabla H(\hat{x},\hat{u})(w,v) = 0\right\}, \tag{3.18}$$

onde $\nabla^T H(\hat{x},\hat{u})(w,v)$ é como em (3.12).

Como $\nabla H(\hat{x},\hat{u})$ é diferenciável e sobrejetivo, de acordo com Lemma 1 de [58],

$$C_T(Q_1,(\hat{x},\hat{u}))^* := \left(\bigcap_{k=1}^{4} C_T(Q_1^k,(\hat{x},\hat{u}))\right)^* = \sum_{k=1}^{4} C_T(Q_1^k,(\hat{x},\hat{u}))^*, \tag{3.19}$$

onde os cones $C_T(Q_1^1,(\hat{x},\hat{u}))^*$, $C_T(Q_1^2,(\hat{x},\hat{u}))^*$, $C_T(Q_1^3,(\hat{x},\hat{u}))^*$ e $C_T(Q_1^4,(\hat{x},\hat{u}))^*$ serão construídos no que se segue.

Construção do cone $C_T(Q_1^1,(\hat{x},\hat{u}))^*$

Segundo [47], para caracterizar o cone dual do cone tangente a $C_T(Q_1^1,(\hat{x},\hat{u}))$, consideremos o operador $\mathscr{F} : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)} \to \mathbb{R}^{n(N)}$, tal que

$$\begin{aligned} &\mathscr{F}(w,v) = (\mathscr{F}_0(w,v),\ldots,\mathscr{F}_{N-1}(w,v)) \\ &\mathscr{F}_i(w,v) = w_{i+1} - \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)w_i - \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)v_i, i = 0,\ldots,N-1. \end{aligned} \tag{3.20}$$

É simples ver que $\mathscr{F}$ é linear e contínuo para todo $i = 0,\ldots,N-1$ e que $C_T(Q_1^1,(\hat{x},\hat{u})) = \mathbf{Ker}\mathscr{F}$. Logo o cone dual é

$$C_T(Q_1^1,(\hat{x},\hat{u}))^* = \left[\mathbf{Ker}\mathscr{F}\right]^*.$$

Como $\left[\mathbf{Ker}\mathscr{F}\right]^* = \mathbf{Im}\mathscr{F}^*$ e $\mathscr{F}^* : (\mathbb{R}^{n(N)})^* \to (\mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)})^*$ é o operador adjunto de $\mathscr{F}$ temos, para todo $f_1^1 \in C_T(Q_1^1,(\hat{x},\hat{u}))^*$, que existe $-\widetilde{p} \in \mathbb{R}^{n(N)^*} \cong R^{n(N)}$, onde $\cong$ representa o isomorfismo canônico entres os espaços, $\widetilde{p} = (p_1,\ldots,p_N)$, tal que $f_1^1 = \mathscr{F}^*(-\widetilde{p})$.

Considerando como se relacionam $\mathscr{F}$ e $\mathscr{F}^*$ e da equação (3.20) temos

$$\begin{aligned} f_1^1(w,v) \quad &= \langle \mathscr{F}^*(-\widetilde{p}),(w,v)\rangle = \langle -\widetilde{p},\mathscr{F}(w,v)\rangle = \sum_{i=0}^{N-1}\langle -p_{i+1},\mathscr{F}_i(w,v)\rangle \\ &= \sum_{i=0}^{N-1}\left\langle -p_{i+1}, w_{i+1} - \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)w_i - \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)v_i\right\rangle. \end{aligned}$$

Pela Observação 2.1 (iv),

$$\begin{aligned} C_T(Q_1^1,(\hat{x},\hat{u}))^* \quad = \Big\{ & f_1^1 \in (X\times U)^* : f_1^1(w,v) = \sum_{i=0}^{N-1}\langle -p_{i+1}, w_{i+1} \\ & -\nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)w_i - \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)v_i\rangle\Big\}. \end{aligned} \tag{3.21}$$

Construção dos cones $C_T(Q_1^k,(\hat{x},\hat{u}))^*, k=2,3,4$
Já para os demais cones observemos que $C_T(Q_1^k,(\hat{x},\hat{u})), k=2,3$ e 4, são subespaços de espaços finitos dimensionais. Novamente pela Observação 2.1 (iv)

$$
\begin{aligned}
C_T(Q_1^2,(\hat{x},\hat{u}))^* &= \Big\{ f_1^2 \in (X\times U)^* : f_1^2(w,v) = \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i)w_i \right\rangle, \\
&\qquad \gamma_i^1 \in \mathbb{R}^{r_1}, i=0,\ldots,N \Big\} \\
C_T(Q_1^3,(\hat{x},\hat{u}))^* &= \Big\{ f_1^3 \in (X\times U)^* : f_1^3(w,v) = \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla\widetilde{h}_i(\hat{u}_i)v_i \right\rangle, \\
&\qquad \gamma_i^2 \in \mathbb{R}^{r_2}, i=0,\ldots,N-1 \Big\} \\
C_T(Q_1^4,(\hat{x},\hat{u}))^* &= \{ f_1^4 \in (X\times U)^* : f_1^4(w,v) = \langle -\gamma^3, \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)w_0 + \\
&\qquad \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)w_N, \rangle, \gamma^3 \in \mathbb{R}^k \big\}.
\end{aligned}
\tag{3.22}
$$

Para maiores detalhes veja [47].

Construção dos cones factíveis $C_F(Q_2^1,(\hat{x},\hat{u}))$ e $C_F(Q_2^2,(\hat{x},\hat{u}))$

Vamos caracterizar os cones factíveis dos conjuntos Q_2^1 e Q_2^1, $C_F(Q_2^1,(\hat{x},\hat{u}))$ e $C_F(Q_2^2,(\hat{x},\hat{u}))$, respectivamente. Para isso assumimos que $(\hat{x},\hat{u})$ é um processo não estacionário de g e $\widetilde{g}$, segundo a Definição 2.8. Então de acordo com a Observação 2.4 (iii)

$$
C_F(Q_2^1,(\hat{x},\hat{u})) = \bigcap_{i=0}^{N} C_{F_i}(Q_2^1,(\hat{x},\hat{u})) \quad \text{e} \quad C_F(Q_2^2,(\hat{x},\hat{u})) = \bigcap_{i=0}^{N-1} C_{F_i}(Q_2^2,(\hat{x},\hat{u})) \tag{3.23}
$$

onde

$$
\begin{aligned}
C_{F_i}(Q_2^1,(\hat{x},\hat{u})) &= \{ (w,v) \in X\times U : \nabla g_i^s(\hat{x}_i)w_i^s < 0, i \in I^1(\hat{x}) \} \\
C_{F_i}(Q_2^2,(\hat{x},\hat{u})) &= \Big\{ (w,v) \in X\times U : \nabla\widetilde{g}_i^t(\hat{u}_i)v_i^t < 0, i \in \widetilde{I}^1(\hat{u}) \Big\}
\end{aligned}
\tag{3.24}
$$

e

$$
I^1(\hat{x}) = \big\{ i \in \{0,\ldots,N\} : g_i^s(\hat{x}_i) = 0, \forall s \in I_i^1(\hat{x}) \big\} \tag{3.25}
$$

$$
\widetilde{I}^1(\hat{u}) = \Big\{ i \in \{0,\ldots,N-1\} : \widetilde{g}_i^t(u_i) = 0, \forall t \in \widetilde{I}_i^1(\hat{u}) \Big\}, \tag{3.26}
$$

$$
I_i^1(\hat{x}) = \{ s \in \{1,\ldots,s_1\} : g_i^s(\hat{x}_i) = 0 \}\; i = 0,\ldots,N, \tag{3.27}
$$

$$
\widetilde{I}_i^1(\hat{u}) = \big\{ t \in \{1,\ldots,s_2\} : \widetilde{g}_i^t(u_i) = 0 \big\}\; i = 0,\ldots,N-1. \tag{3.28}
$$

Pela Observação 2.1 (iv), os respectivos cones duais são caracterizados como segue.

$$
C_F(Q_2^1,(\hat{x},\hat{u}))^* = \left(\bigcap_{i=0}^{N} C_{F_i}(Q_2^1,(\hat{x},\hat{u})) \right)^* = \sum_{i=0}^{N} C_{F_i}(Q_2^1,(\hat{x},\hat{u}))^* \tag{3.29}
$$

onde, para cada $i=0,\ldots,N$,

$$
C_{F_i}(Q_2^1,(\hat{x},\hat{u}))^* = \{ f_i \in (X\times U)^* : f_i(w,v) = \langle \mu_i^{1s}, \nabla g_i^s(\hat{x}_i)w_i \rangle,\ \mu_i^{1s} \geq 0,\ i \in I^1(\hat{x}) \}.
$$

Portanto, se $f_2^1 \in C_F(Q_2^1, (\hat{x}, \hat{u}))^*$ então,

$$f_2^1(w,v) = \sum_{i=0}^{N} f_i(w,v) = \sum_{i=0}^{N} \left\langle \mu_i^1, \nabla g_i(\hat{x}_i) w_i \right\rangle, \ \mu_i^1 \geq 0, \ i = 0, \ldots, N.$$

A construção do cone $C_F(Q_2^2, (\hat{x}, \hat{u}))$ é feita de maneira completamente análoga. Desse modo observemos que, se $f_2^2 \in C_F(Q_2^2, (\hat{x}, \hat{u}))^*$ então

$$f_2^2(w,v) = \sum_{i=0}^{N-1} f_i(w,v) = \sum_{i=0}^{N-1} \left\langle \mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i \right\rangle \ \mu_i^2 \geq 0, i = 0, \ldots, N-1. \tag{3.30}$$

3.2.2 Regularidade

A noção de regularidade é fundamental, tanto para obtenção das condições necessárias de primeira ordem (não degeneradas) como para a derivação das condições suficientes de otimalidade. É importante, portanto, estabelecer condições que assegurem que um problema dado seja regular. Estas condições, denominadas *qualificação de restrição*, são discutidas nesta seção, as quais nos permitem garantir que as condições de otimalidade de primeira ordem para nosso problema, estabelecidas pelo Princípio do Máximo, são não degeneradas e também são suficientes, sob hipóteses de PM-invexidade. Com base nisto, apresentamos a seguinte definição:

Definição 3.2 *Dizemos que $(\hat{x}, \hat{u})$ é um* ***processo regular*** *do problema (PCD) se o cone tangente às restrições de igualdade em $(\hat{x}, \hat{u})$ é determinado por aproximações lineares, isto é,*

$$\begin{aligned} C_T(Q_1, (\hat{x}, \hat{u})) \ = \ \Big\{ & (w,v) \in X \times U : \nabla h_i(\hat{x}_i) w_i = 0, \nabla \widetilde{h}_i(\hat{u}_i, i) v_i = 0, \\ & w_{i+1} = \nabla_x \varphi_i (\hat{x}_i, \hat{u}_i) w_i + \nabla_u \varphi_i (\hat{x}_i, \hat{u}_i) v_i, \\ & \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) w_N = 0 \} \end{aligned} \tag{3.31}$$

e, além disso, existe $(\bar{w}, \bar{v}) \in X \times U$ tal que

$$\begin{aligned} & \bar{w}_{i+1} - \nabla_x \varphi_i (\hat{x}_i, \hat{u}_i) \bar{w}_i - \nabla_u \varphi_i (\hat{x}_i, \hat{u}_i) \bar{v}_i = 0, i = 0, \ldots, N-1 \\ & \nabla h_i(\hat{x}_i) \bar{w}_i = 0, i = 0, \ldots, N \\ & \nabla \widetilde{h}_i(\hat{u}_i) \bar{v}_i = 0, i = 0, \ldots, N-1 \\ & \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) \bar{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \bar{w}_N = 0 \\ & \nabla g_i(\hat{x}_i) \bar{w}_i < 0, i \in I^1(\hat{x}) \\ & \nabla \widetilde{g}_i(\hat{u}_i) \bar{v}_i < 0, i \in \widetilde{I}^1(\hat{u}), \end{aligned} \tag{3.32}$$

onde $I^1(\hat{x})$ e $\widetilde{I}^1(\hat{u})$ estão definidos em (3.25) e (3.26).

Não é difícil mostrar, devido ao Teorema de Lyusternik (Teorema 2.7) que uma condição suficiente para assegurar que (3.31) se cumpra é que o operador $\nabla H(\hat{x}, \hat{u})$, definido em (3.12) seja sobrejetivo, ou ainda, para todo $a \in \mathbb{R}^{n(N)}, b \in \mathbb{R}^{r_1(N+1)}, c \in \mathbb{R}^{r_2(N)}, d \in \mathbb{R}^K$, exista um elemento $(w, v) \in X \times U$ tal que

$$\begin{aligned} & w_{i+1} - \nabla_x \varphi_i (\hat{x}_i, \hat{u}_i) w_i - \nabla_u \varphi_i (\hat{x}_i, \hat{u}_i) v_i = a_i, i = 0, \ldots, N-1 \\ & \nabla h_i(\hat{x}_i) w_i = b_i, i = 0, \ldots, N \\ & \nabla \widetilde{h}_i(\hat{u}_i) v_i = c_i, i = 0, \ldots, N-1 \\ & \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) w_N = d. \end{aligned} \tag{3.33}$$

A condição suficiente (3.32), que garante que $(\hat{x},\hat{u})$ seja regular, é conhecido na literatura como qualificação de restrição de *Mangasarian Fromovitz* (MFCQ), veja [7]. Em programação matemática, (3.33) é equivalente a qualificação de restrição de *Independência Linear* (LICQ), e implica em (MFCQ).

Assim, se as condições (3.32) e (3.33) se cumprem simultaneamente, temos que o processo $(\hat{x},\hat{u})$ é um processo regular do conjunto factível Q.

3.2.3 Condições Necessárias

Apresentamos no que segue uma nova prova do Teorema do Princípio do Máximo Discreto, distinta da apresentada por Marinkovic em [66], onde o autor utiliza ferramentas da Análise Variacional para obter o resultado. A prova aqui apresentada estabelece condições necessárias de otimalidade de primeira ordem para o problema (PCD), através do formalismo de DM que, em nossa opinião, é mais construtiva. Estas condições são uma versão discreta do Princípio do Máximo de Pontryagin.

Teorema 3.3 ***[Princípio do Máximo Discreto]*** *Seja $(\hat{x},\hat{u})$ um processo factível de (PCD). Se $(\hat{x},\hat{u})$ é um processo ótimo de (PCD) então existe $\lambda = (\lambda_0, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, p)$ $\in \mathbb{R} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^{k} \times \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)} \times \mathbb{R}^{n(N+1)}$ tal que $(\lambda_0, p) \neq 0$, $\lambda_0 \geq 0$, $\mu^1 \geqq 0, \mu^2 \geqq 0$ e, tais que, as seguintes condições ocorrem:*

$$p_0 = \nabla_x \mathbb{H}_0(\hat{x},\hat{u},p,\lambda_0) - \nabla_{x_0} K(\hat{x}_0,\hat{x}_N)^T \gamma^3 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 - \nabla g_0(\hat{x}_0)^T \mu_0^1 \tag{3.34}$$

$$p_i = \nabla_x \mathbb{H}_i(\hat{x},\hat{u},p,\lambda_0) - \nabla h_i(\hat{x}_i^T \gamma_i^1) - \nabla g_i(\hat{x}_i)^T \mu_i^1, \;\; i \in [1, N-1] \tag{3.35}$$

$$p_N = -\nabla_{x_N} K(\hat{x}_0,\hat{x}_N)^T \gamma^3 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \nabla g_N(\hat{x}_N)^T \mu_N^1 \tag{3.36}$$

$$\nabla_u \mathbb{H}_i(\hat{x},\hat{u},p,\lambda_0) = \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 + \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2, i \in [0, N-1] \tag{3.37}$$

e a condição de folgas complementares

$$\left\langle \mu_i^1, g_i(\hat{x}_i) \right\rangle = 0, i \in [0, N], \;\; \left\langle \mu_i^2, \widetilde{g}_i(\hat{u}_i) \right\rangle = 0, i \in [0, N-1], \tag{3.38}$$

onde $\mathbb{H}_i(x,u,p,\lambda_0) : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)} \times \mathbb{R}^{n(N+1)} \times \mathbb{R} \to \mathbb{R}$,

$$\mathbb{H}_i(x,u,p,\lambda_0) = \langle p_{i+1}, \varphi_i(x_i,u_i) \rangle - \lambda_0 \Phi_i(x_i,u_i), \;\; \forall i \in [0, N-1],$$

é a função Hamiltoniana do problema (PCD).

Demonstração. Seja $(\hat{x},\hat{u})$ um processo ótimo de (PCD). Então:

(a) Se $\nabla \mathbb{J}(\hat{x},\hat{u}) = 0$ as condições (3.34)-(3.38) são trivialmente satisfeitas, basta considerar $\lambda_0 = 1$, $p = 0$, $\gamma^1 = 0, \gamma^2 = 0$, $\gamma^3 = 0$, $\mu^1 = 0$ e $\mu^2 = 0$.

(b) Se $(\hat{x},\hat{u})$ não é um processo regular, de acordo com a Definição 3.2, então estamos em um dos dois casos abaixo:

1. $C_T(Q_1,(\hat{x},\hat{u})) \neq \mathbf{Ker}\nabla H(\hat{x},\hat{u})$.
Neste caso, não podemos assegurar que o operador $\nabla H(\hat{x},\hat{u})$ é sobrejetivo, porque o núcleo do operador adjunto $\nabla H(\hat{x},\hat{u})$ não é nulo. Então $\exists\psi^* = (\psi_1,\psi_2,\psi_3,\psi_4) \neq 0$ tal que $\nabla H(\hat{x},\hat{u})^T\psi^* = 0$, onde $\psi_1 \in \mathbb{R}^{n(N)}, \psi_2 \in \mathbb{R}^{r_1(N+1)}, \psi_3 \in \mathbb{R}^{r_2(N)}$ e $\psi_4 \in \mathbb{R}^k$ não todos nulos e satisfazendo as condições (3.34)-(3.43). Basta considerar $\lambda_0 = 0$, $p = (p_0,\tilde{p})$ com $p_0 = 0$, $\tilde{p} = \psi_1$, $\gamma^1 = \psi_2, \gamma^2 = \psi_3$, $\gamma^3 = \psi_4$ e $\mu^1 = 0$, $\mu^2 = 0$.
2. $C_T(Q_1,(\hat{x},\hat{u})) = \mathbf{Ker}\nabla H(\hat{x},\hat{u})$ e $\nexists(w,v)$ satisfazendo (3.32).
Como $C_T(Q_1,(\hat{x},\hat{u})) = \mathbf{Ker}\nabla H(\hat{x},\hat{u})$ então para $(w,v) \in C_T(Q_1,(\hat{x},\hat{u}))$ temos válido

$$\begin{aligned} w_{i+1} - \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)w_i - \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)v_i &= 0, i \in [0,N-1] \\ \nabla h_i(\hat{x}_i)w_i &= 0, i \in [0,N] \\ \nabla \widetilde{h}_i(\hat{u}_i)v_i &= 0, i \in [0,N-1] \\ \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)w_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)w_N &= 0. \end{aligned}$$

Como, por hipótese, não existe (w,v) satisfazendo

$$\begin{aligned} \nabla g_i(\hat{x}_i)w_i &< 0, i \in I^1(\hat{x}) \\ \nabla \widetilde{g}_i(\hat{u}_i)v_i &< 0, i \in \widetilde{I}^1(\hat{u}) \end{aligned}$$

isto implica que $(\hat{x},\hat{u})$ é um processo estacionário vetorial de g ou $\widetilde{g}$, segundo Definição 2.8. Então as condições (3.34)-(3.38) são trivialmente satisfeitas, basta considerar $\lambda_0 = 0$, p, γ^1,γ^2, γ^3, como em (3.34)-(3.38) e $\mu^1 \geqq 0$ ou $\mu^2 \geqq 0$.

Para finalizar o caso não regular, observe que de 1. ou 2. temos que

$$C_T(Q_1,(\hat{x},\hat{u})) \cap C_F(Q_2^1,(\hat{x},\hat{u})) \cap C_F(Q_2^2,(\hat{x},\hat{u})) = \emptyset.$$

Logo $f_0 + f_1 + f_2^1 + f_2^2 = 0$ de tal modo que podemos assumir que $f_0 = 0$, ou seja, podemos ter $\lambda_0 = 0$.

(c) Resta considerar o caso regular, isto é, em que $\nabla\mathbb{J}(\hat{x},\hat{u}) \neq 0$, $(\hat{x},\hat{u})$ é um ponto não estacionário de $g_i, i \in I^1(\hat{x})$, e $\widetilde{g}_i, i \in I^1(\hat{u})$, e o operador $\nabla H(\hat{x},\hat{u})$ é sobrejetivo.

Neste caso os cones $C_D(\mathbb{J},(\hat{x},\hat{u}))$, $C_T(Q_1,(\hat{x},\hat{u}))$, $C_F(Q_2^1,(\hat{x},\hat{u}))$, $C_F(Q_2^2,(\hat{x},\hat{u}))$ definidos em (3.16), (3.18) e (3.23), respectivamente são convexos. Logo segue do Teorema 2.12 que existem $f_0 \in C_D(\mathbb{J},(\hat{x},\hat{u}))^*, f_1 = (f_1^1,f_1^2,f_1^3,f_1^4) \in C_T(Q_1,(\hat{x},\hat{u}))^*$, $f_2^1, \in C_F(Q_2^1,(\hat{x},\hat{u}))^*$ e $f_2^2 \in C_F(Q_2^2,(\hat{x},\hat{u}))^*$ não todos nulos, tais que $f_0 + f_1 + f_2^1 + f_2^2 = 0$.

Para cada $(w,v) \in X \times U$, temos:

$$f_0(w,v) + f_1(w,v) + f_2^1(w,v) + f_2^2(w,v) = 0. \tag{3.39}$$

Se $f_0 \in C_D(\mathbb{J},(\hat{x},\hat{u}))^*$ então

$$f_0(w,v) = -\lambda_0\left[\sum_{i=0}^{N-1}\nabla_x\Phi_i\left(\hat{x}_i,\hat{u}_i\right)w_i + \sum_{i=0}^{N-1}\nabla_u\Phi_i\left(\hat{x}_i,\hat{u}_i\right)v_i\right], \lambda_0 \geq 0. \tag{3.40}$$

Se $f_1 \in C_T(Q_1,(\hat{x},\hat{u}))^*$ então

$$\begin{aligned} f_1(w,v) &= \sum_{i=0}^{N-1} \langle -p_{i+1}, w_{i+1} - \nabla_x \varphi_i(\hat{x}_i,\hat{u}_i) w_i - \nabla_u \varphi_i(\hat{x}_i,\hat{u}_i) v_i \rangle \\ &\quad + \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i) w_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i \right\rangle \\ &\quad + \left\langle -\gamma^3, \nabla_{x_0} K(\hat{x}_0,\hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N) w_N \right\rangle, \end{aligned} \tag{3.41}$$

onde $p_i \in \mathbb{R}^n$, $\gamma_i^1 \in \mathbb{R}^{r_1}$, $i \in [0,N]$, $\gamma_i^2 \in \mathbb{R}^{r_2}$, $i \in [0,N-1]$ e $\gamma^3 \in \mathbb{R}^k$.

Se $f_2^1 \in C_F(Q_2^1,(\hat{x},\hat{u}))^*$ e $f_2^2 \in C_F(Q_2^2,(\hat{x},\hat{u}))^*$ então

$$f_2^1(w,v) + f_2^2(w,v) = \sum_{i=0}^{N} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i) w_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i \right\rangle \tag{3.42}$$

e

$$\left\langle \mu_i^1, g_i^1(\hat{x}_i) \right\rangle = 0,\ i \in [0,N] \text{ e } \left\langle \mu_i^2, g_i^2(\hat{u}_i) \right\rangle = 0,\ i = [0,N-1], \tag{3.43}$$

onde $\mu_i^1 \geqq 0$, $i \in [0,N]$ e $\mu_i^2 \geqq 0$, $i \in [0,N-1]$.

Os termos em (3.43) correspondem às condições (3.38).

Substituindo (3.40), (3.41) e (3.42) em (3.39)

$$\begin{aligned} &-\lambda_0 \left(\sum_{i=0}^{N-1} \nabla_x \Phi_i(\hat{x}_i,\hat{u}_i) w_i + \nabla_u \Phi_i(\hat{x}_i,\hat{u}_i) v_i \right) \\ &+ \sum_{i=0}^{N-1} \langle -p_{i+1}, w_{i+1} - \nabla_x \varphi_i(\hat{x}_i,\hat{u}_i) w_i - \nabla_u \varphi_i(\hat{x}_i,\hat{u}_i) v_i \rangle + \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i) w_i \right\rangle \\ &+ \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i \right\rangle + \left\langle -\gamma^3, \nabla_{x_0} K(\hat{x}_0,\hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N) w_N \right\rangle \\ &+ \sum_{i=0}^{N} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i) w_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i \right\rangle = 0. \end{aligned} \tag{3.44}$$

A identidade (3.44) é válida para todo $(w,v) \in X \times U$, em particular para $(w,v) = (w,0)$,

$$\begin{aligned} &\lambda_0 \sum_{i=0}^{N-1} \nabla_x \Phi_i(\hat{x}_i,\hat{u}_i) w_i + \sum_{i=0}^{N-1} \langle p_{i+1}, w_{i+1} - \nabla_x \varphi_i(\hat{x}_i,\hat{u}_i) w_i \rangle + \sum_{i=0}^{N} \left\langle \gamma_i^1, \nabla h_i(\hat{x}_i) w_i \right\rangle \\ &+ \sum_{i=0}^{N} \left\langle \mu_i^1, \nabla g_i(\hat{x}_i) w_i \right\rangle + \left\langle \gamma^3, \nabla_{x_0} K(\hat{x}_0,\hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N) w_N \right\rangle = 0. \end{aligned}$$

Usando as propriedades dos operadores adjuntos e rearranjando os termos da última expressão

$$\begin{aligned} &\Big\langle \nabla_x \Phi_0(\hat{x}_0,\hat{u}_0)^T \lambda_0 + p_0 - \nabla_x \varphi_0(\hat{x}_0,\hat{u}_0)^T p_1 + \nabla_{x_0} K(\hat{x}_0,\hat{x}_N)^T \gamma^3 + \nabla h_0(\hat{x}_0)^T \gamma_0^1 \\ &\qquad + \nabla g_0(\hat{x}_0)^T \mu_0^1, w_0 \Big\rangle \\ &+ \sum_{i=1}^{N-1} \Big\langle p_i + \nabla_x \Phi_i(\hat{x}_i,\hat{u}_i)^T \lambda_0 - \nabla_x \varphi_i(\hat{x}_i,\hat{u}_i)^T p_{i+1} + \nabla h_i(\hat{x}_i)^T \gamma_i^1 + \nabla g_i(\hat{x}_i)^T \mu_i^1, w_i \Big\rangle \\ &+ \left\langle p_N + \nabla h_N(\hat{x}_N)^T \gamma_N^1 + \nabla g_N(\hat{x}_N)^T \mu_N^1 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N)^T \gamma^3, w_N \right\rangle = 0, \forall w \in X, \end{aligned}$$

i.e. $\mathbf{A}w = \langle A_0, w_0\rangle + \sum_{i=1}^{N-1} \langle A_i, w_i\rangle + \langle A_N, w_N\rangle = 0,\ \forall w \in X$ onde

$$\begin{aligned}
A_0 &= \nabla_x \Phi_0\left(\hat{x}_0, \hat{u}_0\right)^T \lambda_0 + p_0 - \nabla_x \varphi_0\left(\hat{x}_0, \hat{u}_0\right)^T p_1 + \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 + \nabla h_0(\hat{x}_0)^T \gamma_0^1 \\
&\quad + \nabla g_0(\hat{x}_0)^T \mu_0^1 \\
A_i &= p_i + \nabla_x \Phi_i\left(\hat{x}_i, \hat{u}_i\right)^T \lambda_0 - \nabla_x \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla h_i(\hat{x}_i)^T \gamma_i^1 + \nabla g_i(\hat{x}_i)^T \mu_i^1, \\
&\quad\ i = 1, \ldots, N-1. \\
A_N &= p_N + \nabla h_N(\hat{x}_N)^T \gamma_N^1 + \nabla g_N(\hat{x}_N)^T \mu_N^1 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3.
\end{aligned}$$

Então necessariamente temos $A_0 = 0, A_i = 0$ e $A_N = 0$, as quais correspondem às condições (3.34), (3.35) e (3.36) respectivamente.

Similarmente considerando $(w, v) = (0, v)$ em (3.44) obtemos

$$-\lambda_0 \sum_{i=0}^{N-1} \nabla_u \Phi_i\left(\hat{x}_i, \hat{u}_i\right) v_i + \sum_{i=0}^{N-1} \langle -p_{i+1}, -\nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right) v_i\rangle + \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i\right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i\right\rangle = 0.$$

Logo,

$$\sum_{i=0}^{N-1} \left\langle \nabla_u \Phi_i\left(\hat{x}_i, \hat{u}_i\right)^T \lambda_0 - \nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 + \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2, v_i\right\rangle = 0,$$

i.e. $\mathbf{B}v = \sum_{i=0}^{N-1} \langle B_i, v_i\rangle = 0,\ \forall v \in U$, onde

$$B_i = \nabla_u \Phi_i\left(\hat{x}_i, \hat{u}_i\right)^T \lambda_0 - \nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 + \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2,\ i = 0, \cdots, N-1.$$

Então necessariamente temos $B_i = 0,\ \forall i = 0, \cdots, N-1$, as quais correspondem às condições (3.37).

Portanto temos estabelecido a existência de $\lambda = (\lambda_0, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, p)$, $p = (p_0, \widetilde{p})$ com $(\lambda_0, p) \neq 0$ para os quais as equações (3.34)-(3.38) são satisfeitas. E com isto o teorema fica demonstrado. □

3.3 Problema (PCD) com restrição abstrata

Neste seção trataremos do problema (PCD) no caso em que, além das restrições de igualdade e desigualdade mencionadas na seção anterior, adicionamos uma restrição abstrata sob o controle u de tal modo que o problema toma a seguinte forma:

$$\begin{aligned}
&\text{Minimizar } \mathbb{J}(x, u) = \sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i\right) \\
&\text{sujeito a: } x_{i+1} = \varphi_i\left(x_i, u_i\right),\ i = 0, \ldots, N-1, \\
&h_i(x_i) = 0,\ g_i(x_i) \leq 0, i = 0, \ldots, N, \\
&\widetilde{h}_i(u_i) = 0,\ \widetilde{g}_i(u_i) \leq 0, i = 0, \ldots, N-1, \\
&K(x_0, x_N) = 0, \\
&u_i \in \Omega_i, i = 0, \ldots, N-1,
\end{aligned} \qquad (\overline{PCD})$$

onde as funções seguem como dadas no problema (PDC) e $\Omega_i \subseteq \mathbb{R}^m, i = 0, \ldots, N-1$, são conjuntos convexos, fechados e com $\mathbf{int}\Omega_i \neq \emptyset$.

Denotamos por $\bar{Q}$ o conjunto dos processos factíveis de $(\overline{PCD})$, sendo

$$\bar{Q} = Q_1 \bigcap Q_2 \bigcap Q_3, \tag{3.45}$$

com Q_1 e Q_2 dados em (3.6), (3.7) e

$$Q_3 = \{(x,u) \in X \times U : u \in \Omega\},$$

onde

$$\Omega = \{u \in U : u_i \in \Omega_i, i = 0, \ldots, N-1\} \tag{3.46}$$

é um conjunto convexo, fechado e com $\mathbf{int}\Omega \neq \emptyset$.

Com esta notação podemos rescrever o problema $(\overline{PCD})$ como o seguinte problema de programação matemática:

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(x,u) \\ \text{sujeito a: } H(x,u) = 0 \\ G(x,u) \leqq 0 \\ (x,u) \in X \times \Omega, \end{array} \tag{$\overline{POE}$}$$

onde $\mathbb{J}$, H e G são definidos por (3.8), (3.9) e (3.10), respectivamente e Ω é definido por (3.46).

Seguindo como na subseção anterior apresentamos agora uma definição de regularidade para o problema $(\overline{PCD})$.

Definição 3.4 *Dizemos que $(\hat{x}, \hat{u})$ é um* ***processo regular*** *do conjunto factível $\bar{Q}$ do problema $(\overline{PCD})$ se o cone tangente às restrições de igualdade em $(\hat{x}, \hat{u})$ é determinado por aproximações lineares, isto é,*

$$\begin{aligned} C_T(Q_1, (\hat{x}, \hat{u})) \; = \; & \{(w,v) \in X \times U : w_{i+1} = \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i) w_i + \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i) v_i, \\ & \nabla h_i(\hat{x}_i) w_i = 0, \nabla \widetilde{h}_i(\hat{u}_i, i) v_i = 0, \\ & \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) w_N = 0\} \end{aligned} \tag{3.47}$$

*e, além disso, existe $(\bar{w}, \bar{v})$, com $\bar{v} = \alpha(u - \hat{u})$ onde $u \in$ **int**Ω e $\alpha > 0$, tal que*

$$\begin{array}{l} \bar{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i) \bar{w}_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i) \bar{v}_i = 0, i = 0, \ldots, N-1 \\ \nabla h_i(\hat{x}_i) \bar{w}_i = 0, i = 0, \ldots, N \\ \nabla \widetilde{h}_i(\hat{u}_i) \bar{v}_i = 0, i = 0, \ldots, N-1 \\ \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) \bar{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \bar{w}_N = 0 \\ \nabla g_i(\hat{x}_i) \bar{w}_i < 0, i \in I^1(\hat{x}), \\ \nabla \widetilde{g}_i(\hat{u}_i, i) \bar{v}_i < 0, i \in \widetilde{I}^1(\hat{u}), \end{array} \tag{3.48}$$

onde $I^1(\hat{x})$ e $\widetilde{I}^1(\hat{u})$ estão definidos em (3.25) e (3.26).

3.3.1 Condicões Necessárias

Vamos estabelecer o Teorema do Princípio do Máximo Discreto para o problema $(\overline{PCD})$. Antes de formalizarmos as condições necessárias de otimalidade para o caso em que o problema possui uma restrição abstrata, vamos apresentar um lema que será útil na prova do referido teorema.

Lema 3.5 *Seja o conjunto $\Omega = \{u \in U : u_i \in \Omega_i, i = 0, \ldots, N-1\}$ com $\Omega_i \subseteq \mathbb{R}^m, \forall i = 0, \ldots, N-1$ e $\hat{u} \in \Omega$. Se o funcional linear definido por:*

$$f(u) = \sum_{i=0}^{N-1} \langle a_i, u_i \rangle, \ a \in U, \tag{3.49}$$

é um suporte para o conjunto Δ no ponto $\hat{u}$, isto é,

$$f(\hat{u}) \leq f(u), \ \forall u \in \Omega \tag{3.50}$$

então, $\langle a_i, u_i - \hat{u}_i \rangle \geq 0, \forall u \in \Omega$ o que implica no vetor $a_i \in \mathbb{R}^m$ ser um suporte para Ω_i no ponto $\hat{u}_i$.

Demonstração. Suponha que a afirmação é falsa. Isto é o mesmo que dizer que existe um subconjunto $R \subseteq \{0, \ldots, N-1\}$, tal que, $\forall i \in R$ existem $\tilde{u}_i \in \Omega_i$ com $\langle a_i, \tilde{u}_i - \hat{u}_i \rangle < 0$. Sem perda de generalidade, podemos afirmar que existem $\gamma_i > 0$, tais que

$$\langle a_i, \tilde{u}_i - \hat{u}_i \rangle = -\gamma_i < 0. \tag{3.51}$$

Agora, definindo

$$\bar{u}_i = \begin{cases} \tilde{u}_i, & \text{se } i \in R, \\ \hat{u}_i, & \text{se } i \in \{0, \ldots, N-1\} - R, \end{cases} \tag{3.52}$$

temos que $\bar{u} \in \Omega$. Então, de (3.51) e (3.52), obtemos que

$$f(\bar{u}) = \sum_{i=0}^{N-1} \langle a_i, \hat{u}_i \rangle + \sum_{i \in R} \langle a_i, \tilde{u}_i - \hat{u}_i \rangle \leq f(\hat{u}) - \gamma < f(\hat{u}),$$

sendo $\gamma = \sum_{i \in R} \gamma_i$. O que implica $f(\bar{u}) < f(\hat{u})$, e isto é uma contradição com o fato de f ser suporte de Ω em $\hat{u}$. □

Com isto, estamos prontos para apresentar o seguinte Teorema do Princípio do Máximo para o Problema $(\overline{PCD})$.

Teorema 3.6 ***[Princípio do Máximo Discreto - com restrição abstrata]*** *Seja $(\hat{x}, \hat{u})$ um processo factível de $(\overline{PCD})$. Se $(\hat{x}, \hat{u})$ é um processo ótimo de $(\overline{PCD})$ então existe $\lambda = (\lambda_0, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, p) \in \mathbb{R} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k \times \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)} \times \mathbb{R}^{n(N+1)}$ tal que $(\lambda_0, p) \neq 0$, $\lambda_0 \geq 0$, $\mu^1 \geqq 0, \mu^2 \geqq 0$ e, tal que, as seguintes condições* *ocorrem:*

$$p_0 = \nabla_x \mathbb{H}_0(\hat{x}, \hat{u}, p, \lambda_0) - \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 - \nabla g_0(\hat{x}_0)^T \mu_0^1 \tag{3.53}$$

$$p_i = \nabla_x \mathbb{H}_i(\hat{x}, \hat{u}, p, \lambda_0) - \nabla h_i(\hat{x}_i)^T \gamma_i^1 - \nabla g_i(\hat{x}_i)^T \mu_i^1, \ i \in [1, N-1] \tag{3.54}$$

$$p_N = -\nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \nabla g_N(\hat{x}_N)^T \mu_N^1 \tag{3.55}$$

$$\begin{aligned} &\nabla_u \mathbb{H}_i(\hat{x}, \hat{u}, p, \lambda_0) - \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 - \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2 \leq \\ &\qquad\qquad \nabla_u \mathbb{H}_i(\hat{x}, u, p, \lambda_0) - \nabla \widetilde{h}_i(u_i)^T \gamma_i^2 - \nabla \widetilde{g}_i(u_i)^T \mu_i^2, \\ &\forall\ (\hat{x}, u) \in X \times \Omega, i \in [0, N-1], \end{aligned} \tag{3.56}$$

e a condição de folgas complementares

$$\left\langle \mu_i^1, g_i^1(\hat{x}_i) \right\rangle = 0, i \in [0, N], \ \left\langle \mu_i^2, \widetilde{g}_i(\hat{u}_i) \right\rangle = 0, i \in [0, N-1], \tag{3.57}$$

onde $\mathbb{H}_i(x, u, p, \lambda_0) : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)} \times \mathbb{R}^{n(N+1)} \times \mathbb{R} \to \mathbb{R}$,

$$\mathbb{H}_i(x, u, p, \lambda_0) = \langle p_{i+1}, \varphi_i(x_i, u_i) \rangle - \lambda_0 \Phi_i(x_i, u_i), \ i \in [0, N-1],$$

é a função Hamiltoniana do problema $(\overline{PCD})$*.*

Demonstração. A prova deste resultado segue de maneira similar a demonstração do Teorema 3.3. Portanto analisaremos aqui apenas o caso em que $(\hat{x}, \hat{u})$ é um processo regular de $(\overline{PCD})$.

Iniciamos analisando a restrição Q_3 do problema $(\overline{PCD})$.
Seja $\Omega = \{u : u_i \in \Omega_i, i = 0, \dots, N-1\}$, pelas hipóteses sobre Ω_i, Ω é um conjunto convexo e fechado, com $\mathbf{int}\Omega \neq \emptyset$. Logo, o conjunto $Q_3 = X \times \Omega$ também é fechado e convexo em $X \times U$ e $\mathbf{int}Q_3 = X \times \Omega \neq \emptyset$.

Seja $C_F(Q_3, (\hat{x}, \hat{u}))$ o cone das direções factíveis de Q_3 em $(\hat{x}, \hat{u})$ é dado por:

$$C_F(Q_3, (\hat{x}, \hat{u})) = \{(w, v) \in X \times U : v := \alpha(u - \hat{u}), (x, u) \in \mathbf{int}Q_3, \alpha > 0\}. \tag{3.58}$$

Veja Teorema 8.2 de [47].

Pelo Teorema 10.5 de [47] se $f_3 \in C_F(Q_3, (\hat{x}, \hat{u}))^*$ então, $f_3 = (0, \hat{f}_3)$ onde $\hat{f}_3 \in U$ é um suporte para Ω.

Sendo $(\hat{x}, \hat{u})$ um processo ótimo de $(\overline{PCD})$ e $C_D(\mathbb{J}, (\hat{x}, \hat{u}))$, $C_T(Q_1, (\hat{x}, \hat{u}))$, $C_F(Q_2^1, (\hat{x}, \hat{u}))$, $C_F(Q_2^2, (\hat{x}, \hat{u}))$ e $C_F(Q_3, (\hat{x}, \hat{u}))$ cones, segue do Teorema 2.12 que existem $f_0 \in C_D(\mathbb{J}, (\hat{x}, \hat{u}))^*$, $f_1 = (f_1^1, f_1^2, f_1^3, f_1^4) \in C_T(Q_1, (\hat{x}, \hat{u}))^*$, $f_2^1 \in C_F(Q_2^1, (\hat{x}, \hat{u}))^*$, $f_2^2 \in C_F(Q_2^2, (\hat{x}, \hat{u}))^*$ e $f_3 \in C_F(Q_3, (\hat{x}, \hat{u}))^*$ não todos nulos tais que, $f_0 + f_1 + f_2^1 + f_2^2 + f_3 = 0$.

Para cada $(w, v) \in X \times U$, temos:

$$f_0(w, v) + f_1(w, v) + f_2^1(w, v) + f_2^2(w, v) + f_3(w, v) = 0. \tag{3.59}$$

onde $f_0(w, v), f_1(w, v)$, $f_2^1(w, v)$ e $f_2^2(w, v)$ são como em (3.40), (3.41), (3.42).

Se $f_3 \in C_F(Q_3, (\hat{x}, \hat{u}))^*$ então

$$f_3(w, v) = (0, \hat{f}_3)(w, v) = \hat{f}_3(v). \tag{3.60}$$

Esta última segue do fato de que $f_3(w,v) = \hat{f}_3(v)$ é um suporte para o conjunto Ω em $\hat{u}$. Além disso, $\lambda_0 \geq 0$, $p_i \in \mathbb{R}^n$, $\gamma_i^1 \in \mathbb{R}^{r_1}$, $i \in [0,N]$, $\gamma_i^2 \in \mathbb{R}^{r_2}$, $i \in [0,N-1]$, $\gamma^3 \in \mathbb{R}^k$ $\mu_i^1 \geqq 0$, $i \in [0,N]$ e $\mu_i^2 \geqq 0$, $i \in [0,N-1]$ e as equações (3.57) seguem de (3.43).

Substituindo (3.40), (3.41), (3.42) e (3.60) em (3.59), obtemos

$$
\begin{aligned}
\hat{f}_3(v) &= -\lambda_0 \left(\sum_{i=0}^{N-1} \nabla_x \Phi_i\left(\hat{x}_i, \hat{u}_i\right) w_i + \nabla_u \Phi_i\left(\hat{x}_i, \hat{u}_i\right) v_i \right) \\
&+ \sum_{i=0}^{N-1} \left\langle -p_{i+1}, -\nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right) v_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i) w_i \right\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i \right\rangle + \left\langle \gamma^3, \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) w_N \right\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i) w_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i \right\rangle,
\end{aligned} \tag{3.61}
$$

com $\hat{f}_3(\hat{u}) \leq \hat{f}_3(u), \forall u \in \Omega$.

A identidade (3.61) é válida para todo $(w,v) \in X \times \Omega$, em particular para $(w,v) = (w,0)$ obtemos as equações (3.53)-(3.55) de modo completamente análogo ao obtido na prova do Teorema 3.3.

Tomando $(w,v) = (0,v)$ em (3.61), temos

$$
\begin{aligned}
\hat{f}_3(v) &= -\lambda_0 \sum_{i=0}^{N-1} \nabla_u \Phi_i\left(\hat{x}_i, \hat{u}_i\right) v_i + \sum_{i=0}^{N-1} \left\langle -p_{i+1}, -\nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right) v_i \right\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i \right\rangle,
\end{aligned}
$$

com $\hat{f}_3(\hat{u}) \leq \hat{f}_3(u), \forall u \in \Omega$.

Assim podemos aplicar o Lema 3.5 para garantir que a condição de máximo (3.56) seja válida, ou seja,

$$
\begin{aligned}
&\nabla_u \mathbb{H}_i(\hat{x}, \hat{u}, p, \lambda_0) - \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 - \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2 \leq \\
&\qquad \nabla_u \mathbb{H}_i(\hat{x}, u, p, \lambda_0) - \nabla \widetilde{h}_i(u_i)^T \gamma_i^2 - \nabla \widetilde{g}_i(u_i)^T \mu_i^2,
\end{aligned}
$$

para todo $u \in \Omega$ com $(x,u) \in \bar{Q}$ e $\lambda_0 = 1$.

Portanto temos estabelecido a existência de $\lambda = (\lambda_0, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, p)$, $p = (p_0, \widetilde{p})$, $(\lambda_0, p) \neq 0$, para os quais as equações (3.53)-(3.57) são satisfeitas. E com isto o teorema fica demonstrado.

$\square$

Salientamos que esta versão do Teorema do Princípio do Máximo em nosso conhecimento é original.

É importante notar que se $(\hat{x}, \hat{u})$ não é um processo regular, então não podemos garantir que o multiplicador associado ao funcional objetivo λ_0 é não nulo nos Teoremas 3.3 e 3.6. Veja prova do Teorema 3.3 item (b).

Definição 3.7 *Dizemos que um processo factível* $(\hat{x}, \hat{u})$ *é um* ***extremal regular*** *do problema* (PCD) *(respect. de* $(\overline{PCD})$*) se satisfaz as equações (3.34)-(3.38) (ou (3.53)-(3.57)), dos Teoremas do Princípio do Máximo, com* $\lambda_0 = 1$.

Da discussão acima e da Observação 2.5 (ii), obtemos o seguinte resultado:

Corolário 3.8 *Se* $(\hat{x}, \hat{u})$ *é um processo ótimo de* (PCD) *(respect. de* $(\overline{PCD})$*) satisfazendo a condição de regularidade (Definição 3.2) (ou Definição 3.4) então* $(\hat{x}, \hat{u})$ *é um extremal regular.*

3.4 Condições Suficientes

É bem conhecido que as condições necessárias de otimalidade de primeira ordem, estabelecidas pelo Princípio de Pontryagin, para problemas de controle ótimo, e pela Regra de Lagrange, Karush-Kuhn-Tucker (KKT), para problemas de otimização matemática, não são suficientes em geral, sem hipóteses adicionais sobre as funções envolvidas no problema.

Consequentemente, há duas formas de proceder para decidir se um dado ponto, que satisfaz as condições necessárias de otimalidade, é de fato ótimo. Uma delas é usar informações de segunda ordem e a outra é usar alguma estrutura especial dos funcionais envolvidos na problemática (convexidade, deformação para problemas mais simples, convexidade generalizada, entre outras), veja [47, 53, 64].

As condições necessárias de otimalidade estabelecidas pelas condições de KKT, ou pelo Princípio do Máximo, são também suficientes se as funções envolvidas no problema são convexas. Todavia esta hipótese é muito restritiva, dado que uma ampla classe de problemas são não convexos. Com o intuito de enfraquecer a hipótese de convexidade, surgiu na literatura vinculada a programação matemática a noção de função invexa, que foi introduzida por Hanson em 1981 [49] e suas generalizações (como por exemplo, [19, 34, 70]), as quais tem sido muito frutíferas na obtenção das condições suficientes de otimalidade.

Em 1985, Martin introduziu em [70] uma generalização da noção de invexidade, denominada KT-invexidade e mostrou que a KT-invexidade não é somente uma condição suficiente de otimalidade para problemas de programação matemática, mas também uma condição necessária para assegurar que cada ponto KKT é um mínimo global. Ou seja, um problema é KT-invexo se, e somente se, todo ponto KKT é um ótimo.

Em particular, nesta seção, nos baseamos por um dos trabalhos mais recentes, devido a de Oliveira *et al.* [75], onde os autores estendem a noção de KT-invexidade de problemas de otimização matemática para problemas de controle ótimo, para mostrar que as condições necessárias de otimalidade, estabelecidas pelo Princípio do Máximo, são suficientes se, e somente se, o problema é PM-invexo.

Vamos introduzir como hipótese adicional para garantir que as condições necessárias de otimalidade apresentadas pelo Princípio do Máximo Discreto (Teoremas 3.3 e 3.6) sejam também suficientes para que um dado processo factível $(\hat{x}, \hat{u})$ seja ótimo, o conceito de PM-invexidade. Para tanto passamos à seguinte definição:

Definição 3.9 *Dizemos que o problema* (PCD) *(respect.* $(\overline{PCD})$*) é* ***PM-invexo*** *no* *processo factível $\alpha = (\hat{x}, \hat{u})$ se, existem $\eta \in \mathbb{R}^{n(N+1)}$ e $\chi \in \mathbb{R}^{m(N)}$, tais que as relações*

$$\begin{aligned}
\mathbb{J}(\beta) - \mathbb{J}(\alpha) &\geq \nabla\mathbb{J}(\alpha)(\eta, \chi) \\
\eta_{i+1} &= \nabla_x\varphi_i(\alpha_i)\eta_i + \nabla_u\varphi_i(\alpha_i)\chi_i, i = 0, \ldots, N-1 \\
\nabla h_i(\hat{x}_i)\eta_i &= 0, i = 0, \ldots, N \\
\nabla\widetilde{h}_i(\hat{u}_i)\chi_i &= 0, i = 0, \ldots, N-1 \\
\nabla g_i(\hat{x}_i)\eta_i &\leq 0, i \in I^1(\hat{x}), \\
\nabla\widetilde{g}_i(\hat{u}_i)\chi_i &\leq 0, i \in \widetilde{I}^1_i(\hat{u}), \\
\nabla_{x_0}K(\hat{x}_0, \hat{x}_N)\eta_0 + \nabla_{x_N}K(\hat{x}_0, \hat{x}_N)\eta_N &= 0,
\end{aligned} \tag{3.62}$$

*se verificam para cada processo factível $\beta = (x, u) \in X \times U$, onde $\eta_i := \eta_i(\beta_i, \alpha_i), i = 0, \ldots, N$, $\chi_i := \chi_i(\beta_i, \alpha_i), i = 0, \ldots, N-1$ (no caso do problema $(\overline{PCD})$, $\chi_i := b(\bar{u}_i - \hat{u}_i), b > 0$ e $\bar{u} \in$ **int**Ω, com Ω definido em (3.46)), $I^1(\hat{x})$, e $\widetilde{I}^1(\hat{u})$ são como em (3.25) e (3.26).*

No que segue apresentamos uma condição de qualificação que será necessária para estabelecermos o resultado principal desta seção.

Definição 3.10 *Dizemos que o processo factível $\alpha = (x, u)$ satisfaz (CQ) se existe $\bar{d} = (\bar{w}, \bar{v}) \in X \times U$ tal que o sistema (3.32) é satisfeito, isto é,*

$$\begin{array}{l}
\bar{w}_{i+1} - \nabla_x\varphi_i(\hat{x}_i, \hat{u}_i)\bar{w}_i - \nabla_u\varphi_i(\hat{x}_i, \hat{u}_i)\bar{v}_i = 0, i = 0, \ldots, N-1 \\
\nabla h_i(\hat{x}_i)\bar{w}_i = 0, i = 0, \ldots, N \\
\nabla\widetilde{h}_i(\hat{u}_i)\bar{v}_i = 0, i = 0, \ldots, N-1 \\
\nabla_{x_0}K(\hat{x}_0, \hat{x}_N)\bar{w}_0 + \nabla_{x_N}K(\hat{x}_0, \hat{x}_N)\bar{w}_N = 0 \\
\nabla g_i(\hat{x}_i)\bar{w}_i < 0, i \in I^1(\hat{x}) \\
\nabla\widetilde{g}_i(\hat{u}_i)\bar{v}_i < 0, i \in \widetilde{I}^1(\hat{u}),
\end{array}$$

onde $I^1(\hat{x})$ e $\widetilde{I}^1(\hat{u})$ estão definidos em (3.25) e (3.26).

Vamos utilizar também no lema que se segue o seguinte Teorema da Alternativa:

Teorema 3.11 *[**Teorema de Motzkin [64]**] Sejam A, C e D matrizes dadas, com $A \neq 0$. Então, exatamente uma das seguintes condições deve ocorrer:*

(I) *O sistema*

$$\begin{array}{l} Ax > 0 \\ Cx \geqq 0 \\ Dx = 0 \end{array}$$

tem uma solução x;

(II) *O sistema*

$$\begin{array}{l} A^Ty_1 + C^Ty_2 + D^Ty_3 = 0 \\ y_1 \geq 0, y_2 \geqq 0 \end{array}$$

tem solução (y_1, y_2, y_3).

Lema 3.12 *Seja $\alpha = (\hat{x}, \hat{u}) \in Q$ satisfazendo (CQ). Se α não é um extremal regular então existe $d = (x, u) \in X \times U$ tal que*

$$\begin{aligned}
&\nabla \mathbb{J}(x,u)(w,v) < 0 \\
&w_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i)\, w_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)\, v_i = 0, i = 0, \ldots, N-1 \\
&\nabla h_i(\hat{x}_i) w_i = 0, i = 0, \ldots, N \\
&\nabla \widetilde{h}_i(\hat{u}_i) v_i = 0, i = 0, \ldots, N-1 \\
&\nabla_{x_0} K(\hat{x}_0, \hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) w_N = 0 \\
&\nabla g_i(\hat{x}_i) w_i < 0, i \in I^1(\hat{x}) \\
&\nabla \widetilde{g}_i(\hat{u}_i) v_i < 0, i \in \widetilde{I}^1(\hat{u}),
\end{aligned} \tag{3.63}$$

onde $I^1(\hat{x})$ e $\widetilde{I}^1(\hat{u})$ estão definidos em (3.25) e (3.26).

Demonstração. Suponha por absurdo que o sistema (3.63) não tem solução $d = (w, v) \in X \times U$. Vamos verificar que a condição II do Teorema 3.11 é satisfeita.

Suponha que existe $\widetilde{d} = (\widetilde{w}, \widetilde{v}) \neq 0$ tal que

$$\begin{aligned}
&\nabla \mathbb{J}(\hat{x},\hat{u})(\widetilde{w},\widetilde{v}) = 0 \\
&\widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i)\, \widetilde{w}_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)\, \widetilde{v}_i = 0, i = 0, \ldots, N-1 \\
&\nabla h_i(\hat{x}_i) \widetilde{w}_i = 0, i = 0, \ldots, N \\
&\nabla \widetilde{h}_i(\hat{u}_i) \widetilde{v}_i = 0, i = 0, \ldots, N-1 \\
&\nabla_{x_0} K(\hat{x}_0, \hat{x}_N) \widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \widetilde{w}_N = 0 \\
&\nabla g_i(\hat{x}_i) \widetilde{w}_i = 0, i \in I^1(\hat{x}) \\
&\nabla \widetilde{g}_i(\hat{u}_i) \widetilde{v}_i = 0, i \in \widetilde{I}^1(\hat{u}).
\end{aligned}$$

o que contradiz (CQ). Portanto, segue do Teorema 3.11 que existem $\lambda = \lambda_0 \in \mathbb{R}$, $\gamma = (\gamma^1, \gamma^2, \gamma^3, p) \in \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k \mathbb{R}^{n(N+1)}$ e $\mu = (\mu^1, \mu^2) \in \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)}$ tal que

$$\begin{aligned}
&-\lambda_0 \left(\sum_{i=0}^{N-1} \nabla_x \Phi_i(\hat{x}_i, \hat{u}_i) + \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i) \right) \\
&+ \sum_{i=0}^{N-1} \langle -p_{i+1}, \mathbf{I}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i) - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i) \rangle + \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i) \right\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) \right\rangle + \left\langle -\gamma^3, \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \right\rangle \\
&+ \sum_{i=0}^{N} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i) \right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) \right\rangle = 0.
\end{aligned}$$

com $\lambda_0 > 0$ e $\mu^1 \geqq 0, \mu^2 \geqq 0$ e onde $\mathbf{I}$ é a matriz identidade de ordem n. E isto implica (da prova do Teorema 3.3) que $\alpha = (\hat{x}, \hat{u})$ é um extremal regular o que contradiz a hipótese. Assim, existe $d = (w, v) \in X \times U$ satisfazendo (3.63). □

Com base nas definições e resultados anteriores apresentaremos o resultado principal desta seção.

Teorema 3.13 *Suponha que o problema PCD satisfaz a condição CQ em cada processo admissível. Então, todo extremal regular $\hat{\alpha} = (\hat{x}, \hat{u})$ é processo ótimo do problema (PCD) (respect. $(\overline{PCD})$) se, e somente se, o problema (PCD) (respect. $(\overline{PCD})$) é PM-invexo.*

Demonstração. Suponha que cada extremal regular é um processo ótimo de (PCD). Provaremos que (PCD) é PM-invexo em $\alpha = (\hat{x}, \hat{u})$.

Considere quaisquer processos $\alpha = (\hat{x}, \hat{u}), \beta = (x, u) \in Q$.

1. Se $\mathbb{J}(\beta) \geq \mathbb{J}(\alpha)$ para todo $\beta \in Q$, então as condições (3.62) são trivialmente satisfeitas tomando $\hat{\eta} := (\eta, \chi) \equiv 0$.

2. Se $\mathbb{J}(\beta) < \mathbb{J}(\alpha)$ então α não é um processo ótimo e por hipótese não é um processo extremal regular (PDC). Portanto, pelo Lema 3.12, existe $d = (w, v) \in X \times U$ tal que

$$\begin{array}{rcl} \nabla\mathbb{J}(\alpha)(w,v) & < & 0 \\ w_{i+1} - \nabla_x\varphi_i(\alpha_i)w_i - \nabla_u\varphi_i(\alpha_i)v_i & = & 0, i = 0, \ldots, N-1 \\ \nabla h_i(\hat{x}_i)w_i & = & 0, i = 0, \ldots, N \\ \nabla\widetilde{h}_i(\hat{u}_i)v_i & = & 0, i = 0, \ldots, N-1 \\ \nabla_{x_0}K(\hat{x}_0, \hat{x}_N)w_0 + \nabla_{x_N}K(\hat{x}_0, \hat{x}_N)w_N & = & 0 \\ \nabla g_i(\hat{x}_i)dw_i & < & 0, i \in I^1(\hat{x}), \\ \nabla\widetilde{g}_i(\hat{u}_i)v_i & < & 0, i \in \widetilde{I}^1(\hat{u}). \end{array} \tag{3.64}$$

Tomando $\hat{\eta}(\beta, \alpha) = cd$ tal que $c \geq \frac{\mathbb{J}(\beta)-\mathbb{J}(\alpha)}{\nabla^T\mathbb{J}(\alpha)d} > 0$, obtemos pela condição (3.64)

$$\begin{array}{rcl} \nabla\mathbb{J}(\alpha)c(w,v) & \leq & \mathbb{J}(\beta) - \mathbb{J}(\alpha) \\ cw_{i+1} - \nabla_x\varphi_i(\alpha_i)cw_i - \nabla_u\varphi_i(\alpha_i)cv_i & = & 0, i = 0, \ldots, N-1 \\ \nabla h_i(\hat{x}_i)cw_i & = & 0, i = 0, \ldots, N \\ \nabla\widetilde{h}_i(\hat{u}_i)cv_i & = & 0, i = 0, \ldots, N-1 \\ \nabla g_i(\hat{x}_i)cw_i & < & 0, i \in I^1(\hat{x}), \\ \nabla\widetilde{g}_i(\hat{u}_i)cv_i & < & 0, i \in \widetilde{I}^1(\hat{u}) \\ \nabla_{x_0}K(\hat{x}_0, \hat{x}_N)cw_0 + \nabla_{x_N}K(\hat{x}_0, \hat{x}_N)cw_N & = & 0. \end{array} \tag{3.65}$$

Portanto, $\hat{\eta}(\beta, \alpha) = c(w, v)$ é, de fato, $\hat{\eta} = (\eta, \chi)$ com $\eta = \eta(\beta, \alpha)$ e $\chi = \chi(\beta, \alpha)$ e as condições (3.62) são válidas, ou seja, o problema (PCD) é PM-invexo, de acordo com a Definição 3.9.

Reciprocamente, suponha que o problema (PCD) é PM-invexo em $\alpha = (\hat{x}, \hat{u})$. Mostraremos que se α é um extremal regular de (PCD) então, necessariamente, é um processo ótimo de (PCD).

Seja $\alpha \in Q$ um processo extremal regular. Então de acordo com (3.34)-(3.38) existem $\lambda = (\lambda_0, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, p)$ tais que

$$\begin{array}{rcl} \mathbb{J}(\beta) - \mathbb{J}(\alpha) \geq \nabla\mathbb{J}(\alpha)(\eta,\chi) & & \\ \langle p_{i+1}, \eta_{i+1} - \nabla_x\varphi_i(\alpha_i)\eta_i - \nabla_u\varphi_i(\alpha_i)\chi_i\rangle & = & 0, i = 0, \ldots, N-1 \\ \left\langle \gamma_i^1, \nabla h_i(\hat{x}_i)\eta_i \right\rangle & = & 0, i = 0, \ldots, N \\ \left\langle \gamma_i^2, \nabla\widetilde{h}_i(\hat{u}_i)\chi_i \right\rangle & = & 0, i = 0, \ldots, N-1 \\ \left\langle \mu_i^{1s}, \nabla g_i^s(\hat{x}_i)\eta_i \right\rangle & \leq & 0, s \in I_i^1(\hat{x}), i = 0, \ldots, N \\ \left\langle \mu_i^{2t}, \nabla\widetilde{g}_i^t(\hat{u}_i)\chi_i \right\rangle & \leq & 0, t \in \widetilde{I}_i^1(\hat{u}), i = 0, \ldots, N-1 \\ \left\langle \gamma^3, \nabla_{x_0}K(\hat{x}_0, \hat{x}_N)\eta_0 + \nabla_{x_N}K(\hat{x}_0, \hat{x}_N)\eta_N \right\rangle & = & 0. \end{array} \tag{3.66}$$

para todo $\beta = (w, v) \in Q$.

Somando todos os termos em (3.66)

$$\begin{aligned}
\mathbb{J}(\beta) - \mathbb{J}(\alpha) \;\; \geq \;\; & \nabla \mathbb{J}(\alpha)(\eta, \chi) + \sum_{i=0}^{N-1} \langle p_{i+1}, \eta_{i+1} - \nabla_x \varphi_i(\alpha_i)\eta_i - \nabla_u \varphi_i(\alpha_i)\chi_i \rangle \\
& + \sum_{i=0}^{N} \left\langle \gamma_i^1, \nabla h_i(\hat{x}_i)\eta_i \right\rangle + \sum_{i=0}^{N-1} \left\langle \gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i)\chi_i \right\rangle \\
& + \sum_{i \in I^1(\hat{x})} \left\langle \mu_i^1, \nabla g_i^s(\hat{x}_i)\eta_i \right\rangle + \sum_{i \in I_i^1(\hat{u})} \left\langle \mu_i^2, \nabla \widetilde{g}_i^t(\hat{u}_i)\chi_i \right\rangle \\
& + \left\langle \gamma^3, \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)\eta_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)\eta_N \right\rangle = 0.
\end{aligned}$$

Concluimos que $\mathbb{J}(\beta) - \mathbb{J}(\alpha) \geq 0$ para todo $\beta \in Q$, i.e, α é um processo ótimo de (PCD).

Finalmente, se considerarmos o problema $(\overline{PCD})$ no lugar de (PCD), a demonstração segue a mesma direção, apenas que neste caso, a condição 2. $\mathbb{J}(\beta) < \mathbb{J}(\alpha)$ implica na existência de $d = (w, v)$ com $v = a(u - \hat{u})$, $a > 0$ e $(x, u) \in \mathbf{int}Q_3$ tal que (3.64) ocorre.Definindo-se $\hat{\eta} = (\eta, \chi)$ tal que $\eta = cw$ e $\chi = cv$ com $v = a(u - \hat{u})$, $a > 0$ e $(w, v) \in \mathbf{int}Q_3$ a prova segue como no caso do problema (PCD).
Com isto finalizamos a demonstração.

□

Para finalizar este capítulo apresentamos o seguinte exemplo:

Exemplo 3.14 *Seja o problema de controle ótimo discreto:*

$$\begin{aligned}
& \textit{Minimizar } \mathbb{J}(x, u) = \tfrac{1}{N} \sum_{i=0}^{N-1} \left(1 - e^{-u_i}\right) \\
& \textit{sujeito a: } x_{i+1} = \tfrac{N-1}{N} x_i + \tfrac{1}{N} u_i^2, \;\; i = 0, \ldots, N-1, \\
& x_0 = 1 \\
& u_i \geq 0, i = 0, \ldots, N-1.
\end{aligned} \tag{D}$$

É fácil ver que $(\hat{x}, \hat{u})$ é um processo ótimo com $\hat{x}_i = (\frac{N-1}{N})^i, i = 0, \ldots, N$, e $\hat{u}_i = 0, i = 0, \ldots, N-1$ com $\mathbb{H}_i(x, u, p, \lambda_0) = \left\langle p_{i+1}, \frac{N-1}{N}\hat{x}_i + \frac{1}{N}\hat{u}_i \right\rangle - \lambda_0 \frac{1}{N}(1 - e^{-\hat{u}_i}), i \in [0, N-1]$. Podemos mostrar que este ponto é um extremal regular, além disso é o único processo extremal regular de (D).

Como foi dito, o processo $(\hat{x}, \hat{u})$ é um extremal regular para (D). Temos que (D) é PM-invexo pela Definição 3.9. De fato, para $(\hat{x}, \hat{u})$ com $\hat{x}_i = (\frac{N-1}{N})^i, i = 0, \ldots, N$, e $\hat{u}_i = 0, i = 0, \ldots, N-1$, basta tomar, $\eta_0 = 0$, $\eta_i = \frac{2}{N} \sum_{s=0} i - \hat{u}_s(1 - e^{\hat{u}_s - u_s})$, $i = 1, \ldots, N-1$, e $\chi_i = 1 - e^{\hat{u}_i - u_i}$. Então pelo Teorema 3.13 temos que $(\hat{x}, \hat{u})$ é um processo ótimo de (D).

E uma caracterização completa do processo ótimo é obtida.

Capítulo 4

Problema de Controle Ótimo Discreto Multiobjetivo

Neste capítulo estabeleceremos condições necessárias e suficientes de otimalidade, para problemas de controle ótimo discreto multiobjetivos diferenciáveis com restrições de igualdade e desigualdade no controle e no estado. Trataremos também das condições necessárias de otimalidade para um problema de controle ótimo discreto multiobjetivo ao qual acrescentamos uma restrição abstrata sob os controles. Os resultados deste capítulo generalizam (em certo sentido) os resultados do Capítulo 3 a problemas multiobjetivos. Não temos conhecimento de que os resultados deste tipo tenham sido obtidos por outros autores.

Discutiremos o seguinte Problema de Controle Discreto Multiobjetivo (PCDM):

$$\begin{aligned}
&\text{Minimizar } \mathbb{J}(x,u) = \left(\sum_{i=0}^{N-1} \Phi_i^1 (x_i, u_i), \ldots, \sum_{i=0}^{N-1} \Phi_i^q (x_i, u_i) \right) \\
&\text{sujeito a: } x_{i+1} = \varphi_i (x_i, u_i), \; i = 0, \ldots, N-1, \\
&h_i(x_i) = 0, \; g_i(x_i) \leqq 0, i = 0, \ldots, N, \\
&\widetilde{h}_i(u_i) = 0, \widetilde{g}_i(u_i) \leqq 0, i = 0, \ldots, N-1, \\
&K(x_0, x_N) = 0,
\end{aligned} \tag{PCDM}$$

onde $\Phi_i^j : \mathbb{R}^n \times \mathbb{R}^m \times \to \mathbb{R}$, $\forall j = 1, \ldots, q, \; i = 0, \ldots, N-1$ e as demais funções envolvidas no problema (PCDM) são como na formulação do problema (PCD) (caso mono-objetivo) apresentadas no Capítulo 3, Seção 3.2.

Observemos ainda que

$$\begin{gathered}
\mathbb{J} : X \times U \to \mathbb{R}^q : \\
\mathbb{J}(x,u) = \left(\sum_{i=0}^{N-1} \Phi_i^1 (x_i, u_i), \cdots, \sum_{i=0}^{N-1} \Phi_i^q (x_i, u_i) \right),
\end{gathered} \tag{4.1}$$

$\mathbb{J}^j(x,u) = \sum_{i=0}^{N-1} \Phi_i^j (x_i, u_i)$ para cada $j = 1, \ldots, q$.

Recordemos que o conjunto dos pontos factíveis de (PCDM), Q, é como definido em (3.5).

Seja $\mathbb{J}^j(x,u) = \sum_{i=0}^{N-1} \Phi_i^j(x_i,u_i)$ para cada $j = 1,\ldots,q$. As noções de otimalidade para o problema de controle ótimo discreto multiobjetivo $(PCDM)$ são as seguintes:

- o processo factível $(\hat{x},\hat{u})$ é um processo ótimo **Pareto fraco** se $(\hat{x},\hat{u}) \in Q$ e não existe outro $(x,u) \in Q, (x,u) \neq (\hat{x},\hat{u})$, tal que $\mathbb{J}^j(x,u) < \mathbb{J}^j(\hat{x},\hat{u})$ para $j = 1,\ldots,q$.

- o processo factível $(\hat{x},\hat{u})$ é um processo ótimo **Pareto** se $(\hat{x},\hat{u}) \in Q$ e não existe outro $(x,u) \in Q, (x,u) \neq (\hat{x},\hat{u})$, tal que $\mathbb{J}^j(x,u) \leq \mathbb{J}^j(\hat{x},\hat{u})$ para $j = 1,\ldots,q$ com a desigualdade estrita para pelo menos um j.

- o processo factível $(\hat{x},\hat{u}) \in Q$ é um processo ótimo **propriamente Pareto** se é um processo ótimo Pareto e existe um escalar $M > 0$ tal que para cada j e cada $(x,u) \in Q$ satisfazendo $\mathbb{J}^j(x,u) < \mathbb{J}^j(\hat{x},\hat{u})$, existe um índice $j_0 \neq j$ tal que $\mathbb{J}^{j_0}(x,u) > \mathbb{J}^{j_0}(\hat{x},\hat{u})$, e
$$\frac{\mathbb{J}^j(\hat{x},\hat{u}) - \mathbb{J}^j(x,u)}{\mathbb{J}^{j_0}(x,u) - \mathbb{J}^{j_0}(\hat{x},\hat{u})} \leq M.$$

4.1 Condições Necessárias

Nesta seção apresentamos o Princípio do Máximo Discreto para o problema de controle ótimo discreto multiobjetivo (PCDM), o qual é uma generalização do Teorema do Princípio do Máximo Discreto (Teorema 3.3). Antes porém vamos abordar o conceito de regularidade envolvido.

Podemos rescrever o problema (PCDM) como o seguinte Problema de Otimização Multiobjetivo (POM):

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(x,u) \\ \text{sujeito a: } H(x,u) = 0 \\ G(x,u) \leqq 0 \\ (x,u) \in X \times U \end{array} \tag{POM}$$

onde H, G e $\mathbb{J}$ foram estabelecidos por (3.9), (3.10), (4.1), respectivamente.

Com base no problema (POM) acima podemos enunciar o seguinte resultado clássico.

Teorema 4.1 (Teorema de Fritz-John, [36]) *Seja o problema (POM) tal que todas as funções envolvidas são continuamente diferenciáveis. Se $(\hat{x},\hat{u})$ é um processo Pareto fraco de (POM), então existem os multiplicadores $\nu \in \mathbb{R}^q_+$, $\gamma \in \mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k$, $\mu^1 \mathbb{R}_+^{s_1(N+1)}$ e $\mu^2 \in \mathbb{R}_+^{s_2(N)}$ tais que*

$$\begin{array}{l} \nabla\mathbb{J}(\hat{x},\hat{u})^T\nu + \nabla H(\hat{x},\hat{u})^T\gamma + \nabla G^1(\hat{x},\hat{u})^T\mu^1 + \nabla G^2(\hat{x},\hat{u})^T\mu^2 = 0, \\ \mu^1 G^1(\hat{x},\hat{u}) = 0, \ \ \mu^2 G^2(\hat{x},\hat{u}) = 0, \\ (\nu,\gamma,\mu^1,\mu^2) \neq 0. \end{array} \tag{4.2}$$

Denotaremos por $\Lambda(\hat{x},\hat{u})$ o conjunto de todos os multiplicadores de Fritz John (ν,γ,μ^1,μ^2) que satisfazem as equações (4.2).

Infelizmente as condições de Fritz John não garantem que o multiplicador associado à função objetivo seja não nulo. Para isto, é necessário impor ao problema condições (chamadas condições de regularidade). Seguindo a classificação feita por Bigi e Pappalardo [13] para problemas de programação matemática temos:

Definição 4.2 *Dado $(\hat{x},\hat{u}) \in Q$ tal que $\Lambda(\hat{x},\hat{u}) \neq \emptyset$, dizemos que:*

(a) $(\hat{x},\hat{u})$ é regular fraco se existe $(\nu,\gamma,\mu^1,\mu^2) \in \Lambda(\hat{x},\hat{u})$ com $\nu \neq 0$.

(b) $(\hat{x},\hat{u})$ é totalmente regular fraco se, para todo $(\nu,\gamma,\mu^1,\mu^2) \in \Lambda(\hat{x},\hat{u})$ existe $j \in \{1,\ldots,q\}$ tal que $\nu_j \neq 0$.

(c) $(\hat{x},\hat{u})$ é regular se existe $(\nu,\gamma,\mu^1,\mu^2) \in \Lambda(\hat{x},\hat{u})$ com $\nu_j > 0, \forall j = 1,\ldots,q$.

(d) $(\hat{x},\hat{u})$ é totalmente regular se, para todo $(\nu,\gamma,\mu^1,\mu^2) \in \Lambda(\hat{x},\hat{u})$ tem-se $\nu_j > 0, \forall j = 1,\ldots,q$.

Uma discussão similar pode ser feita para problemas de controle ótimo discreto multiobjetivos. Proporemos uma classificação no sentido de que o multiplicador associado à função objetivo, dado pelo Princípio do Máximo Discreto, seja não nulo ou estritamente positivo. Para este fim, estabelecemos as definições de processo regular fraco e forte.

Definição 4.3 *Dizemos que $(\hat{x},\hat{u})$ é um **processo regular fraco** do problema $(PCDM)$, se o cone tangente às restrições de igualdade em $(\hat{x},\hat{u})$ é determinado por aproximações lineares, isto é,*

$$\begin{aligned} C_T(Q_1,(\hat{x},\hat{u})) \;=\; & \{(w,v) \in X \times U : w_{i+1} = \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right) w_i + \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right) v_i, \\ & \nabla h_i(\hat{x}_i)w_i = 0, \nabla\widetilde{h}_i(\hat{u}_i)v_i = 0, \\ & \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)w_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)w_N = 0\} \end{aligned} \tag{4.3}$$

e, além disso, existe $(\bar{w},\bar{v}) \in X \times U$ tal que

$$\begin{aligned} & \bar{w}_{i+1} - \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)\bar{w}_i - \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)\bar{v}_i = 0, i = 0,\ldots,N-1, \\ & \nabla h_i(\hat{x}_i)\bar{w}_i = 0, i = 0,\ldots,N, \\ & \nabla\widetilde{h}_i(\hat{u}_i)\bar{v}_i = 0, i = 0,\ldots,N-1, \\ & \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)\bar{w}_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)\bar{w}_N = 0 \\ & \nabla g_i(\hat{x}_i)\bar{w}_i < 0, i \in I^1(\hat{x}), \\ & \nabla\widetilde{g}_i(\hat{u}_i)\bar{v}_i < 0, i \in \widetilde{I}^1(\hat{u}), \end{aligned} \tag{4.4}$$

onde $I^1(\hat{x})$ e $\widetilde{I}^1(\hat{u})$ são dados por (3.25) e (3.26).

Definição 4.4 *Dizemos que $(\hat{x},\hat{u})$ é um **processo regular forte** do problema $(PCDM)$ se o cone tangente às restrições de igualdade em $(\hat{x},\hat{u})$ é determinado por aproximações lineares, isto é,*

$$\begin{aligned} C_T(Q_1,(\hat{x},\hat{u})) \;=\; & \{(w,v) \in X \times U : w_{i+1} = \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right) w_i + \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right) v_i, \\ & \nabla h_i(\hat{x}_i)w_i = 0, \nabla\widetilde{h}_i(\hat{u}_i)v_i = 0, \\ & \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)w_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)w_N = 0\} \end{aligned} \tag{4.5}$$

e, além disso, dado um $r \in \{1, \ldots, q\}$ existe $(\bar{w}, \bar{v}) \in X \times U$ tal que

$$\begin{array}{l}
\nabla \mathbb{J}^j(\hat{x}_i, \hat{u}_i)(\bar{w}_i, \bar{v}_i) < 0, j = 1, \ldots, q, j \neq r \\
\bar{w}_{i+1} - \nabla_x \varphi_i\left(\hat{x}_i, \hat{u}_i\right) \bar{w}_i - \nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right) \bar{v}_i = 0, i = 0, \ldots, N-1, \\
\nabla h_i(\hat{x}_i)\bar{w}_i = 0, i = 0, \ldots, N, \\
\nabla \widetilde{h}_i(\hat{u}_i)\bar{v}_i = 0, i = 0, \ldots, N-1, \\
\nabla_{x_0} K(\hat{x}_0, \hat{x}_N)\bar{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)\bar{w}_N = 0 \\
\nabla g_i(\hat{x}_i)\bar{w}_i < 0, i \in I^1(\hat{x}), \\
\nabla \widetilde{g}_i(\hat{u}_i)\bar{v}_i < 0, i \in \widetilde{I}^1(\hat{u}).
\end{array} \tag{4.6}$$

Observação 4.1 ***i.*** *Como feito no caso mono-objetivo, uma condição para assegurar que (4.3) e (4.5) se cumpram, devido ao Teorema de Lyusternik, Teorema 2.7, é que $\nabla H(\hat{x}, \hat{u})$ seja sobrejetivo.*

ii. *Vale ainda comentar que as condições de regularidade fraca não garantem que, no Princípio do Máximo Discreto Multiobjetivo que discutiremos a seguir, o multiplicador associado à função objetivo seja estritamente positivo. Este é o papel desempenhado pelas condições de regularidade forte, como veremos no que se segue.*

4.1.1 Solução Pareto Fraca

Para estabelecer as condições de otimalidade no processo Pareto fraco $(\hat{x}, \hat{u}) \in Q$ do Problema (PCDM), mediante o formalismo DM, Teorema 2.15, necessitamos apenas determinar o cone de direções de descida do operador $\mathbb{J}$ em $(\hat{x}, \hat{u})$ juntamente com seu dual, já que o cone das direções tangente (com interior vazio), $C_T(Q_1, (\hat{x}, \hat{u}))$, os cones das direções factíveis (de interior não vazio), $C_F(Q_2^1, (\hat{x}, \hat{u}))$ e $C_F(Q_2^2, (\hat{x}, \hat{u}))$ e seus respectivos cones duais, $C_T(Q_1, (\hat{x}, \hat{u}))^*$, $C_F(Q_2^1, (\hat{x}, \hat{u}))^*$ e $C_F(Q_2^2, (\hat{x}, \hat{u}))^*$ foram estabelecidos em (3.18), (3.23), (3.19) e (3.29).

Assim para cada $j = 1, \ldots, q$, temos os cones de descida definidos em $(\hat{x}, \hat{u}) \in Q$, dados por:

$$C_D(\mathbb{J}^j, (\hat{x}, \hat{u})) = \left\{(w, v) \in X \times U : \nabla \mathbb{J}^j(\hat{x}, \hat{u})(v, w) < 0\right\}. \tag{4.7}$$

Então podemos denotar

$$C_D(\mathbb{J}, (\hat{x}, \hat{u})) = \bigcap_{j=1}^{q} C_D(\mathbb{J}^j, (\hat{x}, \hat{u})). \tag{4.8}$$

Já para o dual temos, da parte mono-objetivo, que para cada $j = 1, \ldots, q$

$$C_D(\mathbb{J}^j, (\hat{x}, \hat{u}))^* = \left\{f_j \in (X \times U)^* : f_j = -\nu_j \nabla \mathbb{J}^j(\hat{x}, \hat{u}), \nu_j \geqq 0\right\}. \tag{4.9}$$

Logo se assumirmos que $\displaystyle\bigcap_{j=1}^{q} C_D(\mathbb{J}^j, (\hat{x}, \hat{u})) \neq \emptyset$ então,

$$C_D(\mathbb{J}, (\hat{x}, \hat{u}))^* = \left\{\widetilde{f}_0 \in (X \times U)^* : \widetilde{f}_0 \in \left(\bigcap_{j=1}^{q} C_D(\mathbb{J}^j, (\hat{x}, \hat{u}))\right)^*\right\},$$

e por [47],

$$C_D(\mathbb{J},(\hat{x},\hat{u}))^* = \left\{ \widetilde{f}_0 \in (X\times U)^* : \widetilde{f}_0 = \sum_{j=1}^{q} f_j,\ f_j \in C_D(\mathbb{J}^j,(\hat{x},\hat{u}))^*, j=1,\ldots,q \right\}. \tag{4.10}$$

E com isto estamos em condição de enunciar o seguinte Teorema do Princípio do Máximo para processos Pareto fracos.

Teorema 4.5 ***[Princípio do Máximo Discreto - Pareto Fraco]*** *Seja* $(\hat{x},\hat{u})$ *um processo factível de* $(PCDM)$*. Se* $(\hat{x},\hat{u})$ *é um processo ótimo Pareto fraco de* $(PCDM)$ *então existe* $\lambda = (\nu,\gamma^1,\gamma^2,\gamma^3,\mu^1,\mu^2,p) \in \mathbb{R}^q \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k \times \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N+1)} \times \mathbb{R}^{n(N+1)}$*, tal que* $(\nu,p)\neq 0$*,* $\nu \geqq 0$*,* $\mu^1 \geqq 0, \mu^2 \geqq 0$ *e, tal que, as seguintes condições ocorrem:*

$$p_0 = \nabla_x \mathbb{H}_0(\hat{x},\hat{u},p,\nu) - \nabla_{x_0} K(\hat{x}_0,\hat{x}_N)^T\gamma^3 - \nabla h_0(\hat{x}_0)^T\gamma_0^1 - \nabla g_0(\hat{x}_0)^T\mu_0^1 \tag{4.11}$$

$$p_i = \nabla_x \mathbb{H}_i(\hat{x},\hat{u},p,\nu) - \nabla h_i(\hat{x}_i)^T\gamma_i^1 - \nabla g_i(\hat{x}_i)^T\mu_i^1,\ i\in[1,N-1] \tag{4.12}$$

$$p_N = -\nabla_{x_N} K(\hat{x}_0,\hat{x}_N)^T\gamma^3 - \nabla h_N(\hat{x}_N)^T\gamma_N^1 - \nabla g_N(\hat{x}_N)^T\mu_N^1 \tag{4.13}$$

$$\nabla_u \mathbb{H}_i(\hat{x},\hat{u},p,\nu) = \nabla\widetilde{h}_i(\hat{u}_i)^T\gamma_i^2 + \nabla\widetilde{g}_i(\hat{u}_i)^T\mu_i^2, i\in[0,N-1], \tag{4.14}$$

e a condição de folgas complementares

$$\left\langle \mu_i^1, g_i(\hat{x}_i)\right\rangle = 0, i\in[0,N],\ \left\langle \mu_i^2, \widetilde{g}_i(\hat{u}_i)\right\rangle = 0, i\in[0,N-1] \tag{4.15}$$

onde, $\mathbb{H}_i(x,u,p,\nu) : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{(N)m} \times \mathbb{R}^{n(N+1)} \times \mathbb{R}^q \to \mathbb{R}$,

$$\mathbb{H}_i(x,u,p,\nu) = \langle p_{i+1}, \varphi_i(x_i,u_i)\rangle - \sum_{j=1}^{q} \nu_j \Phi_i^j(x_i,u_i),\ i\in[0,N-1],$$

é a função Hamiltoniana do problema $(PCDM)$*.*

Demonstração. Seja $(\hat{x},\hat{u})$ um processo ótimo de $(PCDM)$. Então,

(a) Se existe $j_0 \in \{1,\ldots,q\}$ tal que $\nabla\mathbb{J}^{j_0}(\hat{x},\hat{u}) = 0$ as equações (4.11)-(4.15) são satisfeitas trivialmente, basta considerar $\nu_{j_0} = 1$, $\nu_j = 0, j=1,\ldots,q, j\neq j_0$, $p=0$, $\gamma^1 = 0, \gamma^2 = 0$, $\gamma^3 = 0$, $\mu^1 = 0$, e $\mu^2 = 0$.

(b) Se o processo $(\hat{x},\hat{u})$ não é um processo regular fraco, de acordo com a Definição 4.3 (o que implica que $(\hat{x},\hat{u})$ não é um processo regular forte) então, estamos no caso 1 ou no caso 2 da prova do Teorema 3.3 - parte (b). E não podemos garantir que ν seja não nulo.

(c) Logo, resta considerar o caso em que as condições de regularidade fraca (respectivamente, regularidade forte) são válidas, ou seja, $\nabla\mathbb{J}(\hat{x},\hat{u}) \neq 0$, $(\hat{x},\hat{u})$ é um ponto não estacionário de $g_i, i\in I^1(\hat{x})$, e $\widetilde{g}_i, i\in\widetilde{I}^1(\hat{u})$, e o operador $\nabla H(\hat{x},\hat{u})$ é sobrejetivo.

Neste caso os cones $C_D(\mathbb{J},(\hat{x},\hat{u}))$, $C_T(Q_1,(\hat{x},\hat{u}))$, $C_F(Q_2^1,(\hat{x},\hat{u}))$, $C_F(Q_2^2,(\hat{x},\hat{u}))$ definidos em (4.8), (3.18) e (3.23) são convexos. Então, do Teorema 2.12, existem $\widetilde{f_0} = \sum_{j=1}^{q} f_j \in C_D(\mathbb{J},(\hat{x},\hat{u}))^*$, $f_1 = (f_1^1, f_1^2, f_1^3, f_1^4) \in C_T(Q_1,(\hat{x},\hat{u}))^*$, $f_2^1, \in C_F(Q_2^1,(\hat{x},\hat{u}))^*$ e $f_2^2 \in C_F(Q_2^2,(\hat{x},\hat{u}))^*$ não todos nulos, tais que $\widetilde{f_0} + f_1 + f_2^1 + f_2^2 = 0$.

Para cada $(w,v) \in X \times U$, temos:

$$\widetilde{f_0}(w,v) + f_1(w,v) + f_2^1(w,v) + f_2^2(w,v) = 0. \tag{4.16}$$

Como $\widetilde{f_0} \in C_D(\mathbb{J},(\hat{x},\hat{u}))^*$ então

$$\begin{aligned} \widetilde{f_0}(w,v) &= \sum_{j=1}^{q} f_j(w,v) = \sum_{j=1}^{q} -\nu_j \nabla \mathbb{J}^j(\hat{x},\hat{u})(w,v) \\ &= \sum_{j=1}^{q} -\nu_j \sum_{i=0}^{N-1} \left(\nabla_x \Phi_i^j\left(\hat{x}_i,\hat{u}_i\right) w_i + \nabla_u \Phi_i^j\left(\hat{x}_i,\hat{u}_i\right) v_i\right), \ 0 \leqq \nu \in \mathbb{R}^q. \end{aligned} \tag{4.17}$$

Os demais funcionais de (4.16) são como dados em (3.41), (3.42) e (3.43), onde os termos em (3.43) correspondem às condições (4.15).

Substituindo (4.17), (3.41) e (3.42) em (4.16), obtemos

$$\begin{aligned} &\sum_{j=1}^{q} -\nu_j \left(\sum_{i=0}^{N-1} \nabla_x \Phi_i^j\left(\hat{x}_i,\hat{u}_i\right) w_i + \nabla_u \Phi_i^j\left(\hat{x}_i,\hat{u}_i\right) v_i\right) \\ &+ \sum_{i=0}^{N-1} \langle -p_{i+1}, w_{i+1} - \nabla_x \varphi_i\left(\hat{x}_i,\hat{u}_i\right) w_i - \nabla_u \varphi_i\left(\hat{x}_i,\hat{u}_i\right) v_i\rangle + \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i) w_i\right\rangle \\ &+ \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i\right\rangle + \left\langle -\gamma^3, \nabla_{x_0} K(\hat{x}_0,\hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N) w_N\right\rangle \\ &+ \sum_{i=0}^{N} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i) w_i\right\rangle + \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i\right\rangle = 0. \end{aligned} \tag{4.18}$$

A identidade (4.18) é válida para todo $(w,v) \in X \times U$. Em particular para $(w,v) = (w,0)$, obtemos

$$\begin{aligned} &\sum_{j=1}^{q} -\nu_j \sum_{i=0}^{N-1} \nabla_x \Phi_i^j\left(\hat{x}_i,\hat{u}_i\right) w_i + \sum_{i=0}^{N-1} \langle -p_{i+1}, w_{i+1} - \nabla_x \varphi_i\left(\hat{x}_i,\hat{u}_i\right) w_i\rangle \\ &+ \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i) w_i\right\rangle + \sum_{i=0}^{N} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i) w_i\right\rangle \\ &+ \langle -\gamma^3, \nabla_{x_0} K(\hat{x}_0,\hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N) w_N\rangle = 0. \end{aligned} \tag{4.19}$$

Usando as propriedades dos operadores adjuntos e rearranjando os termos da última expressão,

$$\left\langle \sum_{j=1}^{q} \nu_j \nabla_x \Phi_0^j\left(\hat{x}_0, \hat{u}_0\right) + p_0 - \nabla_x \varphi_0\left(\hat{x}_0, \hat{u}_0\right)^T p_1 + \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 \right.$$
$$\left. + \nabla h_0(\hat{x}_0)^T \gamma_0^1 + \nabla g_0(\hat{x}_0)^T \mu_0^1, w_0 \right\rangle$$
$$+ \sum_{i=1}^{N-1} \left\langle p_i + \sum_{j=1}^{q} \nu_j \nabla_x \Phi_i^j\left(\hat{x}_i, \hat{u}_i\right) - \nabla_x \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla h_i(\hat{x}_i)^T \gamma_i^1 + \nabla g_i(\hat{x}_i)^T \mu_i^1, w_i \right\rangle$$
$$+ \left\langle p_N + \nabla h_N(\hat{x}_N)^T \gamma_N^1 + \nabla g_N(\hat{x}_N)^T \mu_N^1 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3, w_N \right\rangle = 0, \ \forall w \in X,$$

i.e. $\mathbf{A}w = \langle A_0, w_0 \rangle + \sum_{i=1}^{N-1} \langle A_i, w_i \rangle + \langle A_N, w_N \rangle = 0, \ \forall w \in X$ onde

$$\begin{aligned}
A_0 &= \sum_{j=1}^{q} \nu_j \nabla_x \Phi_0^j\left(\hat{x}_0, \hat{u}_0\right) + p_0 - \nabla_x \varphi_0\left(\hat{x}_0, \hat{u}_0\right)^T p_1 + \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 + \nabla h_0(\hat{x}_0)^T \gamma_0^1 \\
&\quad + \nabla g_0(\hat{x}_0)^T \mu_0^1 \\
A_i &= p_i + \sum_{j=1}^{q} \nu_j \nabla_x \Phi_i^j\left(\hat{x}_i, \hat{u}_i\right) - \nabla_x \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla h_i(\hat{x}_i)^T \gamma_i^1 + \nabla g_i(\hat{x}_i)^T \mu_i^1, \\
&\quad i = 0, \ldots, N-1. \\
A_N &= p_N + \nabla h_N(\hat{x}_N)^T \gamma_N^1 + \nabla g_N(\hat{x}_N)^T \mu_N^1 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3.
\end{aligned}$$

Então necessariamente temos $A_0 = 0, A_i = 0$ e $A_N = 0$, as quais correspondem às condições (4.11), (4.12) e (4.13), respectivamente.

Similarmente, considerando $(w, v) = (0, v)$ em (4.18) obtemos

$$\sum_{i=0}^{N-1} \sum_{j=1}^{q} -\nu_j \nabla_u \Phi_i^j\left(\hat{x}_i, \hat{u}_i\right) v_i + \sum_{i=0}^{N-1} \left\langle -p_{i+1}, -\nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right) v_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(\hat{u}_i) v_i \right\rangle$$
$$+ \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i) v_i \right\rangle = 0.$$

Logo,

$$\sum_{i=0}^{N-1} \left\langle \sum_{j=1}^{q} \nu_j \nabla_u \Phi_i^j\left(\hat{x}_i, \hat{u}_i\right) - \nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 + \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2, v_i \right\rangle = 0,$$

i.e. $\mathbf{B}v = \sum_{i=0}^{N-1} \langle B_i, v_i \rangle = 0, \ \forall v \in U$, onde

$$B_i = \sum_{j=1}^{q} \nu_j \nabla_u \Phi_i\left(\hat{x}_i, \hat{u}_i\right) - \nabla_u \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} + \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 + \nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2, \ i = 0, \cdots, N-1.$$

Então necessariamente temos $B_i = 0, \ \forall i = 0, \cdots, N-1$ as quais correspondem às condições (4.14).

Portanto temos estabelecido a existência de $\lambda = (\nu, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, p)$, $p = (p_0, \widetilde{p})$ com $(\nu, p) \neq 0$ para os quais as equações (4.11)-(4.15) são satisfeitas. E com isto o teorema fica demonstrado.

□

4.1.2 Solução Pareto

Para obtermos o Teorema do Princípio do Máximo para processos ótimos de Pareto, via formalismo de DM (Teorema 2.20), necessitamos primeiro caracterizar o cone de direções de não crescimento do funcional objetivo $\mathbb{J}$ do problema $(PCDM)$.

Pelo Lema 2.18 temos que o cone das direções de não crescimento, ao qual denotaremos $C_N(\mathbb{J}^j,(\hat{x},\hat{u}))$, contém o cone $C_D(\mathbb{J}^j,(\hat{x},\hat{u}))$ das direções de descida. Da Definição 2.17 e dado que $\mathbb{J}^j, j=1,\ldots,q$ são diferenciáveis temos que

$$C_N(\mathbb{J}^j,(\hat{x},\hat{u})) = \left\{(w,v)\in X\times U : \nabla\mathbb{J}^j(\hat{x},\hat{u})(w,v)\le 0\right\}, j=1,\ldots,q. \tag{4.20}$$

Além disso, pelo Lema 2.22, se $(\hat{x},\hat{u})$ é um ponto não estacionário de $\mathbb{J}$ temos

$$C_D(\mathbb{J}^j,(\hat{x},\hat{u}))^* = C_N(\mathbb{J}^j,(\hat{x},\hat{u}))^* = \left\{f_j\in(X\times U)^* : f_j = -\nu_j\nabla\mathbb{J}^j(\hat{x},\hat{u}), \nu_j\ge 0\right\}, j=1,\ldots,q,$$

e da Observação 2.7 segue que,

$$C_D(\mathbb{J},(\hat{x},\hat{u}))^* = C_N(\mathbb{J},(\hat{x},\hat{u}))^* = \left\{\overline{f}_0\in(X\times U)^* : \overline{f}_0 = -\sum_{j=1}^{q}\nu_j\nabla\mathbb{J}^j(\hat{x},\hat{u}), \nu_j\ge 0\right\}. \tag{4.21}$$

Com isso, podemos substituir no Teorema do Princípio do Máximo, Teorema 4.5, a hipótese de processo Pareto fraco por processo Pareto e o resultado segue de modo completamente análogo.

Observação 4.2 *Observemos que:*

i. *No caso dos processos ótimos de Pareto, acrescentamos o cone $C_N(\mathbb{J}^j,(\hat{x},\hat{u}))$ das direções de não crescimento e aplicamos o Teorema 2.20 (que é uma generalização do Teorema de Dubovistkii-Milyutin, Teorema 2.12).*

ii. *Podemos assumir que $(\hat{x},\hat{u})$ é um processo não estacionário de $\mathbb{J}$ e pelo Lema 2.22, temos que $C_D(\mathbb{J},(\hat{x},\hat{u}))^* = C_N(\mathbb{J},(\hat{x},\hat{u}))^*$. Isso porque se $(\hat{x},\hat{u})$ é um processo estacionário de $\mathbb{J}$ então, de acordo com a Definição 2.8, existe $\alpha=(\alpha_1,\ldots,\alpha_q)\in\mathbb{R}^q, \alpha_j\ge 0, j=1,\ldots,q,$ tal que $\nabla\mathbb{J}^*(\hat{x},\hat{u})\alpha=0$, onde $\nabla\mathbb{J}^*(\hat{x},\hat{u})$ é o operador adjunto de $\nabla\mathbb{J}(\hat{x},\hat{u})$ e portanto, podemos escolher $\nu=(\nu_1,\ldots,\nu_q)$ com $\nu_j=\alpha_j,\ j=1,\ldots,q$, e as condições do Teorema* **??** *se cumprem trivialmente.*

iii. *Isto implica, devido ao Corolário 2.21, que as q equações determinadas em (2.5), se reduzem a uma única equação como dada em (4.18). Em se tratando de um processo regular forte, apenas substituímos a condição de regularidade fraca, Definição 4.3, pela regularidade forte, Definição 4.4, que garante a positividade estrita do multiplicador ν. O que justifica a terminologia regularidade fraca e forte, em conformidade com a classificação da regularidade dos multiplicadores de Lagrange introduzida por Bigi e Pappalardo em [13].*

4.1.3 Problema $(\overline{PCDM})$ com restrição abstrata

Análogo ao estudo feito na Seção 3.3, passaremos à análise das condições necessárias para o seguinte problema de controle ótimo multiobjetivo:

$$\begin{array}{l}
\text{Minimizar } \mathbb{J}(x,u) = \left(\sum_{i=0}^{N-1} \Phi_i^1(x_i,u_i), \ldots, \sum_{i=0}^{N-1} \Phi_i^q(x_i,u_i)\right) \\
\text{sujeito a: } x_{i+1} = \varphi_i(x_i,u_i),\ i = 0,\ldots,N-1, \\
h_i^1(x_i) = 0,\ g_i(x_i) \leqq 0, i = 0,\ldots,N, \\
\widetilde{h}_i(u_i) = 0,\ \widetilde{g}_i(u_i) \leqq 0, i = 0,\ldots,N-1, \\
K(x_0,x_N) = 0, \\
u_i \in \Omega_i, i = 0,\ldots,N-1,
\end{array} \qquad (\overline{PCDM})$$

onde as funções do problema acima são como dadas no problema (PCDM) e $\Omega_i \subseteq \mathbb{R}^m, i = 0,\ldots,N-1$, são conjuntos convexos, fechados e com $\mathbf{int}\Omega_i \neq \emptyset$.

Por este motivo, o conjunto dos processos factíveis do problema $(\overline{PCDM})$ é o mesmo do problema $(\overline{PCD})$, $\bar{Q}$, dado em (3.45).

Apresentamos duas definições de regularidade para o problema $(\overline{PCDM})$.

Definição 4.6 *Dizemos que $(\hat{x},\hat{u})$ é um **processo regular fraco** do problema $(\overline{PCDM})$ se valem as relações (4.3) e, além disso, existe $(\bar{w},\bar{v})$, com $\bar{v} = a(u-\hat{u})$ onde $u \in$ **int**Ω e $a > 0$, tal que (4.4) se cumpre.*

Definição 4.7 *Dizemos que $(\hat{x},\hat{u})$ é um **processo regular forte** do problema $(\overline{PCDM})$ se vale (4.5), existe $(\bar{w},\bar{v})$, com $\bar{v} = a(u-\hat{u})$ onde $u \in$ **int**Ω e $a > 0$, tal que (4.6) é satisfeita.*

E assim estamos prontos para estabelecer o Teorema do Princípio do Máximo para o problema $(\overline{PCDM})$.

Teorema 4.8 *[**Princípio do Máximo Discreto - multiobjetivo com restrição abstrata**] Seja $(\hat{x},\hat{u})$ um processo factível de $(\overline{PCDM})$. Se $(\hat{x},\hat{u})$ é um processo ótimo Pareto (ou Pareto fraco) de $(\overline{PCDM})$ então existe $\lambda = (\nu,\gamma^1,\gamma^2,\gamma^3,\mu^1,\mu^2,p) \in \mathbb{R}^q \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k \times \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)} \times \mathbb{R}^{n(N+1)}$ tal que $(\nu,p) \neq 0$, $\nu \geqq 0$, $\mu^1 \geq 0, \mu^2 \geq 0$ e tal que as seguintes condições ocorrem:*

$$p_0 = \nabla_x \mathbb{H}_0(\hat{x},\hat{u},p,\nu) - \nabla_{x_0} K(\hat{x}_0,\hat{x}_N)^T \gamma^3 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 - \nabla g_0(\hat{x}_0,)^T \mu_0^1 \qquad (4.22)$$

$$p_i = \nabla_x \mathbb{H}_i(\hat{x},\hat{u},p,\nu) - \nabla h_i(\hat{x}_i)^T \gamma_i^1 - \nabla g_i(\hat{x}_i)^T \mu_i^1,\ i \in [1,N-1] \qquad (4.23)$$

$$p_N = -\nabla_{x_N} K(\hat{x}_0,\hat{x}_N)^T \gamma^3 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \nabla^T g_N(\hat{x}_N)^T \mu_N^1 \qquad (4.24)$$

$$\begin{array}{l}
\nabla_u \mathbb{H}_i(\hat{x},\hat{u},p,\nu) - \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 - \nabla^T \widetilde{g}_i(\hat{u}_i)^T \mu_i^2 \leq \\
\qquad\qquad \nabla_u \mathbb{H}_i(\hat{x},\hat{u},p,\nu) - \nabla \widetilde{h}_i(u_i)^T \gamma_i^2 - \nabla^T \widetilde{g}_i(u_i)^T \mu_i^2, \qquad (4.25) \\
\forall (\hat{x},u) \in X \times \Omega, i \in [0,N-1],
\end{array}$$

e a condição de folgas complementares

$$\left\langle \mu_i^1, g_i(\hat{x}_i) \right\rangle = 0, i \in [0, N], \ \left\langle \mu_i^2, \widetilde{g}_i(\hat{u}_i) \right\rangle = 0, i \in [0, N-1] \tag{4.26}$$

onde, $\mathbb{H}_i(x, u, p, \nu) : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{(N)m} \times \mathbb{R}^{n(N+1)} \times \mathbb{R}^q \to \mathbb{R}$,

$$\mathbb{H}_i(x, u, p, \nu) = \langle p_{i+1}, \varphi_i(x_i, u_i) \rangle - \sum_{j=1}^{q} \nu_j \Phi_i^j(x_i, u_i), \ i \in [0, N-1],$$

é a função Hamiltoniana do problema $(\overline{PCDM})$.

A prova deste resultado segue os passos do Teorema 3.6 do caso escalar e por isso será omitida. Lembramos que, em se tratando de um problema de controle ótimo multiobjetivo, substituímos o cone das direções de descida (3.16) e seu dual (3.17) pelos cones (4.8) e (4.10) e o multiplicador $\lambda_0 \in \mathbb{R}$ por $\nu \in \mathbb{R}^q$. Ainda, para o caso dos processos Pareto ótimos, levamos em consideração a Observação 4.2.

E com base nos Teoremas do Princípio do Máximo e na classificação proposta por Bigi e Pappalardo [13], propomos as seguintes definições:

Definição 4.9 *Dizemos que um processo factível* $(\hat{x}, \hat{u})$ *é um* ***extremal regular fraco*** *do problema* $(PCDM)$ *(respect.* $(\overline{PCDM})$*) se, satisfaz as equações (4.11)-(4.15) do Teorema do Princípio do Máximo (respect. (4.22)-(4.26) para o problema com restrição abstrata), com* $\nu \geq 0$.

Definição 4.10 *Dizemos que um processo factível* $(\hat{x}, \hat{u})$ *é um* ***extremal regular forte*** *do problema* $(PCDM)$ *(respect.* $(\overline{PCDM})$*) se, satisfaz as equações (4.11)-(4.15) (respect. (4.22)-(4.26)), com* $\nu > 0$.

Assim são imediatos os resultados a seguir.

Corolário 4.11 *Se* $(\hat{x}, \hat{u})$ *é um processo Pareto (ou Pareto fraco) de* $(PCDM)$ *(respect.* $(\overline{PCDM})$*) satisfazendo a condição de regularidade (Definição 4.3) (respect. Definição 4.6) então* $(\hat{x}, \hat{u})$ *é um extremal regular fraco.*

Corolário 4.12 *Se* $(\hat{x}, \hat{u})$ *é um processo Pareto (ou Pareto fraco) de* $(PCDM)$ *(respect.* $(\overline{PCDM})$*) satisfazendo a condição de regularidade (Definição 4.4) (respect. Definição4.7), então* $(\hat{x}, \hat{u})$ *é um extremal regular forte.*

Observação 4.3 *Em resumo,*

- *Se* $(\hat{x}, \hat{u})$ *é um extremal regular forte (em particular, um extremal regular fraco) então* $(\hat{x}, \hat{u})$ *é um ponto não estacionário de* G *e o operador* $\nabla H(\hat{x}, \hat{u})$ *é sobrejetivo. Ver demonstração do Teorema 4.5. Portanto os cones de direções tangentes e factíveis e seus respectivos duais são determinados em (3.18), (3.23) e (3.19) e (3.29).*

- *Se $\nu = (\nu_1, \ldots, \nu_q)$ é identicamente nulo então as condições de otimalidade estabelecidas pelos Teoremas 4.5 e 4.8 não envolvem a função objetivo $\mathbb{J}$, e neste caso, estas condições são não regulares (degeneradas). Uma condição suficiente para assegurar que ν não é identicamente nulo é que $(\hat{x}, \hat{u})$ seja um processo regular, no sentido da Definição 4.3. Logo, sob essa hipótese, é possível obter que $(\hat{x}, \hat{u})$ é um extremal regular fraco.*

- *Se $\nu = (\nu_1, \ldots, \nu_q)$ não é identicamente nulo e existe uma componente nula, digamos $\nu_{j_0} = 0$, então as condições de otimalidade estabelecidas pelos Teoremas 4.5, 4.8 não envolvem a componente $\mathbb{J}^r$ da função $\mathbb{J}$. Uma condição suficiente para assegurar que todas as componentes de ν sejam estritamente positivas é que $(\hat{x}, \hat{u})$ seja um processo regular forte, no sentido da Definição 4.4. Logo é possível obter que $(\hat{x}, \hat{u})$ seja um extremal regular forte.*

4.2 Condições Suficientes

Nesta seção obteremos condições suficientes de otimalidade para o problema $(PCDM)$ via convexidade generalizada.

Seguindo a mesma filosofia dos trabalhos de Martin [70], Osuna-Gómez *et. al.* [78], introduziremos uma noção de invexidade generalizada de modo que seja a maior classe de problemas para as quais todo extremal regular fraco sejam processos ótimos de Pareto.

Definição 4.13 *Dizemos que o problema $(PCDM)$ é **PM-pseudoinvexo** no processo factível $\alpha = (\hat{x}, \hat{u})$ se existem $\eta \in \mathbb{R}^{n(N+1)}$ e $\chi \in \mathbb{R}^{m(N)}$ tais que*

$$\mathbb{J}(\beta) \leq \mathbb{J}(\alpha) \Rightarrow \begin{cases} \nabla \mathbb{J}^j(\alpha)(\eta, \chi) < 0, j = 1, \ldots, q, \\ \eta_{i+1} = \nabla_x \varphi_i(\alpha_i)\eta_i + \nabla_u \varphi_i(\alpha_i)\chi_i, i = 0, \ldots, N-1 \\ \nabla h_i(\hat{x}_i)\eta_i = 0, i = 0, \ldots, N, \\ \nabla \widetilde{h}_i(\hat{u}_i)\chi_i = 0, i = 0, \ldots, N-1, \\ \nabla g_i(\hat{x}_i)\eta_i \leqq 0, i \in I^1(\hat{x}), \\ \nabla \widetilde{g}_i(\hat{u}_i)\chi_i \leqq 0, i \in \widetilde{I}^1(\hat{u}), \\ \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)\eta_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)\eta_N = 0 \end{cases} \tag{4.27}$$

se verificam para cada processo factível $\beta = (x, u) \in X \times U$, onde $\eta_i := \eta_i(\beta_i, \alpha_i), i = 0, \ldots, N$, $\chi_i := \chi_i(\beta_i, \alpha_i), i = 0, \ldots, N-1$, e $I^1(\hat{x}), \widetilde{I}^1(\hat{u})$ são como em (3.25) e (3.26).

O próximo resultado caracteriza os problemas PM-pseudoinvexos para os processos regulares fracos.

Teorema 4.14 *Suponha que o problema $(PCDM)$ satisfaz a condição (CQ) 3.10 para cada $j = 1, \cdots, q$ em cada processo admissível. Então, todo extremal regular fraco é um processo ótimo de Pareto do problema $(PCDM)$ se, e somente se, o problema $(PCDM)$ é PM-pseudoinvexo em Q.*

Demonstração. Suponha que cada extremal regular fraco é um processo ótimo de Pareto de $(PDCM)$. Provaremos que o problema (PCDM) é PM-pseudoinvexo em Q. Considere quaisquer processos admissíveis $\alpha = (\hat{x}, \hat{u}), \beta = (x, u)$.

Suponha que existe ao menos um $\beta = (x, u) \in Q$ tal que $\mathbb{J}^j(\beta) \leq \mathbb{J}^j(\alpha), j = 1, \ldots, q$, com ao menos uma desigualdade estrita para algum $\alpha \in Q$. Então, α não é um processo ótimo de Pareto de (PCDM) e, por hipótese, α não é um extremal regular fraco. Portanto, pelo Lema 3.12 temos que existe $d = (w, v) \in X \times U$ tal que

$$\mathbb{J}(\beta) \leqq \mathbb{J}(\alpha) \Rightarrow \begin{cases} \nabla \mathbb{J}^j(\alpha)(w, v) < 0, j = 1, \ldots, q, \\ w_{i+1} - \nabla_x \varphi_i(\alpha_i) w_i - \nabla_u \varphi_i(\alpha_i) v_i, i = 0, \ldots, N-1, \\ \nabla h_i(\hat{x}_i) w_i = 0, i = 0, \ldots, N, \\ \nabla \widetilde{h}_i(\hat{u}_i) v_i = 0, i = 0, \ldots, N-1, \\ \nabla g_i(\hat{x}_i) w_i \leqq 0, i \in I^1(\hat{x}), \\ \nabla \widetilde{g}_i(\hat{u}_i) v_i \leqq 0, i \in \widetilde{I}^1(\hat{u}), \\ \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) w_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) w_N = 0. \end{cases}$$

Assim, as condições da Definição 4.13 se cumprem para $\eta := w$ e $\chi := v$, dependendo dos pontos factíveis $\beta = (x, u)$ e $\alpha = (\hat{x}, \hat{u})$. Isto implica que (PCDM) é PM-pseudoinvexo.

Reciprocamente, sob a hipótese de que (PCDM) é PM-pseudoinvexo, demonstraremos que cada extremal regular fraco α é um processo ótimo de Pareto do problema.

Seja $\alpha = (\hat{x}, \hat{u}) \in Q$ um extremal regular fraco de (PCDM) e suponha que α não é um processo ótimo de Pareto. Então existe um $\beta \in Q$ tal que $\mathbb{J}(\beta) \leqq \mathbb{J}(\alpha)$ e isto implica que existe $\bar{\eta} = (\eta, \chi) \in X \times U$ tal que

$$\begin{array}{l} \nabla \mathbb{J}^j(\alpha)(\eta, \chi) < 0, j = 1, \ldots, q, \\ w_{i+1} = \nabla_x \varphi_i(\alpha_i) \eta_i + \nabla_u^T \varphi_i(\alpha_i) \chi_i, i = 0, \ldots, N-1, \\ \nabla h_i(\hat{x}_i) \eta_i = 0, i = 0, \ldots, N, \\ \nabla \widetilde{h}_i(\hat{u}_i) \chi_i = 0, i = 0, \ldots, N-1, \\ \nabla g_i(\hat{x}_i) \eta_i \leqq 0, i \in I^1(\hat{x}), \\ \nabla \widetilde{g}_i(\hat{u}_i) \chi_i \leqq 0, i \in \widetilde{I}^1(\hat{u}), \\ \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) \eta_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \eta_N = 0. \end{array}$$

Isto nos permite afirmar, devido à caracterização dos cones, que

$$\left(\bigcap_{j=1}^{q} C_D(\mathbb{J}^j, \alpha) \right) \cap C_T(Q_1, \alpha) \cap \left(\bigcap_{i \in I^1(\hat{x})} C_{F_i}(Q_2^1, \alpha) \right) \cap \left(\bigcap_{i \in \widetilde{I}^1(\hat{u})} C_{F_i}(Q_2^2, \alpha) \right) \neq \emptyset,$$

que é equivalente a dizer, de acordo com o Lema 2.11,

$$\sum_{j=1}^{q} f_j + \sum_{i \in I^1(\hat{x})} f_i + \sum_{i \in \widetilde{I}^1(\hat{u})} f_i + f_1 \neq 0,$$

para todos $f_r \in C_D(\mathbb{J}^r, \alpha)^*, f_j \in C_N(\mathbb{J}^j, \alpha)^*, j = 1, \ldots, p, j \neq r$, $f_i \in C_{F_i}(Q_2^1, \alpha)^*, i \in I^1(\hat{x})$, $f_i \in C_{F_i}(Q_2^2, \alpha)^*, i \in \widetilde{I}^1(\hat{u})$ e $f_1 \in C_T(Q_1, \alpha)^*$ sendo esses cones definidos em (4.7), (3.18) e (3.23).

Portanto, não é possível obter as equações de Euler-Lagrange. Isto implica que α não é um extremal regular fraco, o que contradiz a hipótese inicial. Obtemos então que $\alpha = (\hat{x}, \hat{u})$ é um processo de Pareto. E com isto encerramos a prova. □

Observação 4.4 *Observemos que, sob hipóteses de PM-invexidade, todo extremal regular fraco é um processo ótimo de Pareto, não é necessário que a condição (CQ) se verifique em cada processo admissível para $(PCDM)$.*

4.3 Escalarização

Por "escalarização" entende-se um processo através do qual é possível converter um problema multiobjetivo em um problema de otimização mono-objetivo (ou escalar) equivalente, de tal maneira que as soluções do problema multiobjetivo possam ser obtidas como soluções de um problema clássico de programação não-linear.

Vamos considerar o *Método da Ponderação* e o *Método da ε-restrição.*

4.3.1 Ponderação

Nesta subseção abordaremos a relação existente entre o problema (PCDM) e o problema escalar ponderado ($P\omega$). Definiremos a PM-invexidade para o caso de problemas multiobjetivos e mostraremos que os problemas de controle ótimo discreto pertencentes a esta classe são equivalentes ao problema ponderado associado. Iniciamos com a seguinte definição.

Definição 4.15 *Dizemos que o problema multiobjetivo (PCDM) é um problema **PM-invexo** se para cada par de processos factíveis $\beta = (x,u), \alpha = (\hat{x},\hat{u}) \in X \times U$ existem $\eta \in \mathbb{R}^{n(N+1)}$, $\chi \in \mathbb{R}^{m(N)}$ tais que*

$$\begin{aligned} \mathbb{J}^j(\beta) - \mathbb{J}^j(\alpha) &\geq \nabla \mathbb{J}^j(\alpha)(\eta,\chi),\ j = 1,\dots,q, \\ \eta_{i+1} &= \nabla_x \varphi_i(\alpha_i)\eta_i + \nabla_u \varphi_i(\alpha_i)\chi_i, i = 0,\dots,N-1, \\ \nabla h_i(\hat{x}_i)\eta_i &= 0, i = 0,\dots,N, \\ \nabla \widetilde{h}_i(\hat{u}_i)\chi_i &= 0, i = 0,\dots,N-1, \\ \nabla g_i(\hat{x}_i)\eta_i &\leqq 0, i \in I^1(\hat{x}), \\ \nabla \widetilde{g}_i(\hat{u}_i)\chi_i &\leqq 0, i \in \widetilde{I}^1(\hat{u}), \\ \nabla_{x_0} K(\hat{x}_0,\hat{x}_N)\eta_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N)\eta_N &= 0, \end{aligned} \tag{4.28}$$

onde $\eta_i := \eta_i(\beta_i,\alpha_i), i = 0,\dots,N$, $\chi_i := \chi_i(\beta_i,\alpha_i), i = 0,\dots,N-1$ e $I^1(\hat{x}), \widetilde{I}^1(\hat{u})$ são como em (3.25) e (3.26).

Agora seja $\omega = (\omega_1,\dots,\omega_q)$ um vetor não nulo do $\mathbb{R}^q_+$. O problema escalar ponderado associado ao $(PCDM)$ é dado por

$$\begin{aligned} &\text{Minimizar } \mathbb{L}(x,u) = \sum_{j=1}^{q} \omega_j \mathbb{J}^j(x,u) \\ &\text{sujeito a: } x_{i+1} = \varphi_i(x_i,u_i),\ i = 0,\dots,N-1, \\ &h_i^1(x_i) = 0,\ g_i(x_i) \leq 0, i = 0,\dots,N, \\ &\widetilde{h}_i(u_i) = 0, \widetilde{g}_i(u_i) \leq 0, i = 0,\dots,N-1, \\ &K(x_0,x_N) = 0. \end{aligned} \tag{$P\omega$}$$

Dado $0 \neq \omega \in \mathbb{R}^p_+$, o conjunto dos processos factíveis de $(P\omega)$ é exatamente o mesmo conjunto dos processos factíveis de (PCDM).

Similarmente ao trabalho de Oliveira e Silva [76], vamos definir a seguinte hipótese. Seja $0 \neq \omega \in \mathbb{R}^q_+$.

$$\mathbf{(H(\omega))}\ \{j : \mathbb{J}^j(\hat{x},\hat{u}) < \mathbb{J}^j(x,u)\} \cap \{j : \omega_j > 0\} \neq \emptyset,$$

$\forall\ (\hat{x},\hat{u})$ e (x,u) processos factíveis de (PCDM).

Teorema 4.16 *Seja $0 \neq \omega \in \mathbb{R}^q_+$ e suponha que (H(ω)) é satisfeita. Se $(\hat{x},\hat{u})$ é um processo ótimo de $(P\omega)$, então é um processo ótimo Pareto de (PCDM).*

Demonstração. Suponha que existe um processo factível $(\bar{x},\bar{u})$ de (PCDM) tal que $\mathbb{J}(\bar{x},\bar{u}) \leq \mathbb{J}(\hat{x},\hat{u})$. Como (H(ω)) é satisfeita, segue que $\omega\mathbb{J}(\bar{x},\bar{u}) \leq \omega\mathbb{J}(\hat{x},\hat{u})$ o que contradiz a otimalidade global de $(\hat{x},\hat{u})$ em $(P\omega)$.

□

Quando $\omega \in \mathbb{R}^q, \omega > 0$, então a hipótese (H(ω)) é automaticamente satisfeita. Neste caso se $(\hat{x},\hat{u})$ é uma solução ótima de $(P\omega)$, então ele é um processo ótimo de Pareto de (PCDM), pelo Teorema 4.16.

Já para um processo ótimo Pareto Fraco a hipótese (HA) é desnecessária.

Teorema 4.17 *Seja $0 \neq \omega \in \mathbb{R}^p_+$. Se $(\hat{x},\hat{u})$ é um processo ótimo de $(P\omega)$, então é um processo ótimo Pareto fraco de (PCDM).*

A prova deste resultado é análoga à demonstração do Teorema 4.16 e por isso será omitida.

Observemos que os Teoremas 4.16 e 4.17 são válidos não somente para o problema de controle ótimo em tempo discreto (PCDM), mas sim para os problemas de extremais multiobjetivos em geral.

A recíproca dos Teoremas 4.16 e 4.17 é verdadeira.

Teorema 4.18 *Assuma que (PCDM) é PM-invexo (segundo Definição 4.15). Se $(\hat{x},\hat{u})$ é um processo regular fraco Pareto (respect. Pareto fraco) de (PCDM) então existe um vetor não nulo $\omega \in \mathbb{R}^p_+, \omega \neq 0$, tal que $(\hat{x},\hat{u})$ é uma solução ótima de $(P\omega)$.*

Demonstração. Se $(\hat{x},\hat{u})$ é uma solução Pareto fraco de (PCDM), pelo Teorema do Princípio do Máximo, Teorema 4.5, existe $\lambda = (\nu,\gamma^1,\gamma^2,\gamma^3,\mu^1,\mu^2,p) \in \mathbb{R}^q \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^k \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N+1)} \times \mathbb{R}^{n(N+1)}$ tal que $(\nu,p) \neq 0,\ \nu \geqq 0,\ \mu^1 \geqq 0, \mu^2 \geqq 0$ satisfazendo as equações (4.11)-(4.15).

Suponha agora que não existe $0 \neq \omega \in \mathbb{R}^q_+$ tal que $(\hat{x},\hat{u})$ é uma solução ótima de $(P\omega)$. Em particular, $(\hat{x},\hat{u})$ não é uma solução do problema $(P\nu)$,

$$\sum_{j=1}^{q} \nu_j \left[\mathbb{J}^j(x,u) - \mathbb{J}^j(\hat{x},\hat{u})\right] < 0.$$

Agora da PM-invexidade de (PCDM), obtemos

$$\sum_{i=0}^{N-1}\sum_{j=1}^{q} \nu_j \left(\nabla_x \mathbb{J}^j(\hat{x},\hat{u})\eta_i + \nabla_u \mathbb{J}^j(\hat{x},\hat{u})\chi_i\right) < 0, \tag{4.29}$$

com $0 \neq \nu \in \mathbb{R}^p_+$.

E do Teorema 4.5

$$
\begin{aligned}
&\sum_{j=1}^{q}\sum_{i=0}^{N-1} -\nu_j \left(\nabla_x \mathbb{J}^j\left(\hat{x}_i,\hat{u}_i\right)\eta_i + \nabla_u \mathbb{J}^j\left(\hat{x}_i,\hat{u}_i\right)\chi_i\right) = \\
&\sum_{i=0}^{N-1} \langle p_{i+1}, \eta_{i+1} - \nabla_x\varphi_i\left(\hat{x}_i,\hat{u}_i\right)\eta_i - \nabla_u\varphi_i\left(\hat{x}_i,\hat{u}_i\right)\chi_i\rangle + \sum_{i=0}^{N}\left\langle \gamma_i^1, \nabla h_i(\hat{x}_i)\eta_i\right\rangle \\
&+\sum_{i=0}^{N-1}\left\langle \gamma_i^2, \nabla\widetilde{h}_i(\hat{u}_i)\chi_i\right\rangle + \left\langle \gamma^3, \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)\eta_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)\eta_N\right\rangle \qquad (4.30)\\
&+\sum_{i=0}^{N}\left\langle \mu_i^1, \nabla g_i(\hat{x}_i)\eta_i\right\rangle + \sum_{i=0}^{N-1}\left\langle \mu_i^2, \nabla\widetilde{g}_i(\hat{u}_i)\chi_i\right\rangle.
\end{aligned}
$$

De (4.30) e da PM-invexidade de (PCDM) concluímos que

$$
\sum_{i=0}^{N-1}\sum_{j=1}^{q} \nu_j \left(\nabla_x \mathbb{J}^j(\hat{x},\hat{u})\eta_i + \nabla_u \mathbb{J}^j(\hat{x},\hat{u})\chi_i\right) \geq 0, \tag{4.31}
$$

o que é um absurdo por (4.29).

□

Observemos que dos Teoremas 4.17 e 4.18 temos a equivalência entre os problemas (PCDM) e ($P\omega$), sob hipótese de PM-invexidade.

4.3.2 Outras técnicas de escalarização

Conforme discutido anteriormente, o problema de controle ótimo discreto (PCDM) é equivalente ao problema de programação multiobjetivo (POM). Isto quer dizer que um processo admissível $(\hat{x},\hat{u})$ é processo ótimo de Pareto (Pareto fraco, propriamente Pareto) de (PCDM) se e somente se é solução de Pareto (respect. Pareto fraca, propriamente Pareto) do problema multiobjetivo (POM).

No que se segue, vamos utilizar outras técnicas de escalarização ao problema (POM).

Iniciaremos aplicando o método da ε-restrição ao problema (POM). Sejam $k \in \{1,\cdots,q\}$ e $(\hat{x},\hat{u}) \in \mathbb{X} = \{(x,u) \in X \times U : H(x,u) = 0, G(x,u) \leqq 0\}$ fixados.

$$
\begin{array}{l}
\text{Minimizar } \mathbb{J}^k(x,u)\\
\text{sujeito a:}\\
H(x,u) = 0\\
G(x,u) \leqq 0\\
\mathbb{J}^j(x,u) \leq \mathbb{J}^j(\hat{x},\hat{u}), j \neq k
\end{array}
\tag{$\mathrm{P}_k(\hat{x},\hat{u})$}
$$

Os seguintes resultados são conhecidos:

1. $(\hat{x},\hat{u})$ é solução eficiente de (POM) se e somente se $(\hat{x},\hat{u})$ é solução de $(\mathrm{P}_k(\hat{x},\hat{u}))$, para todo índice k.

2. Se existe um k tal que $(\hat{x},\hat{u})$ é a única solução de $(\mathrm{P}_k(\hat{x},\hat{u}))$, então $(\hat{x},\hat{u})$ é solução eficiente de (POM).

3. Suponha que $\mathbb{X}$ é convexo e que as funções $\mathbb{J}^j$ são explicitamente quaseconvexas[1] em $\mathbb{X}$, para todo j. Se $(\hat{x}, \hat{u})$ é solução Pareto fraca de (POM), então existe um k tal que $(\hat{x}, \hat{u})$ é solução ótima de $(\mathrm{P}_k(\hat{x}, \hat{u}))$.

Os resultados 1 e 2 podem ser encontrados em [30]; o resultado 3 é demonstrado em [88].

Assim sendo, da equivalência dos problemas (PCDM) e (POM) e também destes resultados seguem de maneira imediata:

$1'$. $(\hat{x}, \hat{u})$ é processo Pareto de (PCDM) se e somente se $(\hat{x}, \hat{u})$ é solução de $(\mathrm{P}_k(\hat{x}, \hat{u}))$, para todo índice k.

$2'$. Se existe um k tal que $(\hat{x}, \hat{u})$ é a única solução de $(\mathrm{P}_k(\hat{x}, \hat{u}))$, então $(\hat{x}, \hat{u})$ é processo Pareto de (PCDM).

$3'$. Suponha que $\mathbb{X}$ é convexo e que as funções $\mathbb{J}^j$ são explicitamente quaseconvexas em $\mathbb{X}$, para todo j. Se $(\hat{x}, \hat{u})$ é processo ótimo Pareto de (PCDM), então existe um k tal que $(\hat{x}, \hat{u})$ é solução ótima de $(\mathrm{P}_k(\hat{x}, \hat{u}))$.

- **Solução propriamente Pareto**: $(\hat{x}, \hat{u}) \in Q$ é uma solução propriamente Pareto se existe um escalar $M > 0$ tal que para cada j e cada $(x, u) \in Q$ satisfazendo $\mathbb{J}^j(x, u) < \mathbb{J}^j(\hat{x}, \hat{u})$, existe um índice $j_0 \neq j$ tal que $\mathbb{J}^{j_0}(x, u) > \mathbb{J}^{J_0}(\hat{x}, \hat{u})$ e $\frac{\mathbb{J}^j(x,u) - \mathbb{J}^j(\hat{x},\hat{u})}{\mathbb{J}^{j_0}(\hat{x},\hat{u}) - \mathbb{J}^{j_0}(x,u)} \leq M$.

- **Solução k-Pareto** (Benson e Morin [10]) : Sejam $k \in \{1, \cdots, q\}$ e $(\hat{x}, \hat{u}) \in Q$. Dizemos que $(\hat{x}, \hat{u})$ é solução k- Pareto se para cada $(x, u) \in Q$ com $\mathbb{J}^k(x, u) < \mathbb{J}^k(\hat{x}, \hat{u})$, existe um índice j tal que $\mathbb{J}^j(x, u) > \mathbb{J}^j(\hat{x}, \hat{u})$.

- **Solução k-propriamente Pareto** (Benson e Morin [10]) : Sejam $k \in \{1, \cdots, q\}$ e $(\hat{x}, \hat{u}) \in Q$. Dizemos que $(\hat{x}, \hat{u})$ é solução k-propriamente Pareto se $(\hat{x}, \hat{u})$ é solução k-Pareto e existe um escalar $M_k > 0$ tal que cada $(x, u) \in Q$ com $\mathbb{J}^k(x, u) < \mathbb{J}^k(\hat{x}, \hat{u})$, existe um índice j tal que $\mathbb{J}^j(x, u) > \mathbb{J}^j(\hat{x}, \hat{u})$ e $\frac{\mathbb{J}^k(x,u) - \mathbb{J}^k(\hat{x},\hat{u})}{\mathbb{J}^j(\hat{x},\hat{u}) - \mathbb{J}^j(x,u)} \leq M_k$.

É evidente: $(\hat{x}, \hat{u})$ é solução Pareto (propriamente Pareto) se e somente se $(\hat{x}, \hat{u})$ é solução k-eficiente (respect. k-propriamente Pareto) para todo índice k.

Vamos considerar a seguir uma outra técnica de escalarização que será útil na análise dos processos propriamente Pareto de (PCDM). Para isto, vamos considerar a seguinte escalarização do problema (POM): Para $k \in \{1, \cdots, q\}$, $\omega \in \mathbb{R}_+^{q-1}$ fixado, se define

$$\begin{array}{l} \text{Minimizar } \mathbb{J}^k(x, u) + \sum_{j \neq k} \omega_j \mathbb{J}^j(x, u) \\ \text{sujeito a: } (x, u) \in \mathbb{X}. \end{array} \qquad (\widehat{\mathrm{P}}_k(\omega))$$

[1] Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função e $X \subset \mathbb{R}^n$ um conjunto convexo e não vazio. (i) Dizemos que f é quaseconvexa em X se para quaisquer $x_1, x_2 \in X$, $x_0 \in (x_1, x_2)$, tem-se

$$f(x_0) \leq \max\{f(x_1), f(x_2)\};$$

(ii) Dizemos que f é estritamente quaseconvexa em X se para quaisquer $x_1, x_2 \in X$, $x_0 \in (x_1, x_2)$, $f(x_1) \neq f(x_2)$, tem-se

$$f(x_0) < \max\{f(x_1), f(x_2)\};$$

(iii) Dizemos que f é explicitamente quaseconvexa em X se f é quaseconvexa e estritamente quase convexa em X.

Os seguintes resultados são conhecidos:

1. Se $(\hat{x}, \hat{u})$ é solução de $(\widehat{\mathrm{P}}_k(\omega))$ para algum algum $\omega \in \mathbb{R}_+^{q-1}$, então $(\hat{x}, \hat{u})$ é solução k-propriamente Pareto de (POM);

2. Se $\mathbb{X}$ é convexo e as funções $\mathbb{J}^1, \cdots, \mathbb{J}^q$ são convexas em $\mathbb{X}$, então $(\hat{x}, \hat{u})$ é solução k-propriamente Pareto de (POM) se e somente se existe $\omega \in \mathbb{R}_+^{q-1}$ tal que $(\hat{x}, \hat{u})$ é solução de $(\widehat{\mathrm{P}}_k(\omega))$.

Para demonstração destes resultados, veja Benson e Morin [10].

Novamente, da equivalência dos problemas (PCDM) e (POM) e também destes resultados seguem de maneira imediata:

$1'$. Se $(\hat{x}, \hat{u})$ é solução de $(\widehat{\mathrm{P}}_k(\omega))$ para algum algum $\omega \in \mathbb{R}_+^{q-1}$, então $(\hat{x}, \hat{u})$ é processo k-propriamente Pareto de (PCDM);

$2'$. Se $\mathbb{X}$ é convexo e as funções $\mathbb{J}^1, \cdots, \mathbb{J}^q$ são convexas em $\mathbb{X}$, então $(\hat{x}, \hat{u})$ é processo k-propriamente Pareto de (PCDM) se e somente se existe $\omega \in \mathbb{R}_+^{q-1}$ tal que $(\hat{x}, \hat{u})$ é solução de $(\widehat{\mathrm{P}}_k(\omega))$;

$3'$. Se $\mathbb{X}$ é convexo e as funções $\mathbb{J}^1, \cdots, \mathbb{J}^q$ são convexas em $\mathbb{X}$, então $(\hat{x}, \hat{u})$ é processo propriamente Pareto de (PCDM) se e somente se existe para cada k, existe $\omega \in \mathbb{R}_+^{q-1}$ tal que $(\hat{x}, \hat{u})$ é solução de $(\widehat{\mathrm{P}}_k(\omega))$.

Observação 4.5 *Em [10], Benson e Morin utilizam esta escalarização para demonstrar que, sob hipóteses de convexidade, a propriedade da k-eficiência própria é equivalente à estabilidade de um certo problema escalar associado ao problema multiobjetivo. Retomaremos esta discussão na Terceira Parte desta Tese, onde obteremos um resultado similar a este, sob hipóteses de invexidade.*

Parte II

Caso 2 Regular

Capítulo 5

Problema de Controle Ótimo Discreto Mono-objetivo

Sabemos que as condições de otimalidade estabelecidas pelo Teorema do Princípio do Máximo (Teorema 3.3) são úteis quando o problema é regular, ou seja, quando o multiplicador associado à função objetivo é diferente de zero e as condições suficientes que asseguram a positividade deste são chamadas condições de regularidade ou qualificação de restrição (veja por exemplo, [3, 35]). Neste caso, temos que o processo factível $(\hat{x}, \hat{u})$ é um processo regular, ou ainda, as restrições do problema são regulares no processo $(\hat{x}, \hat{u})$. Por outro lado, se isto não se cumpre dizemos que o processo $(\hat{x}, \hat{u})$ é não regular, ou ainda, degenerado.

Neste capítulo, apresentaremos condições necessárias de otimalidade por meio de um Teorema do Princípio do Máximo Estendido, para um problema de controle ótimo discreto que não é necessariamente regular.

5.1 2-regularidade em Programação Matemática

Consideremos o seguinte Problema Escalar:

$$\begin{array}{c} \text{Minimizar } f(x) \\ \text{sujeito a: } F(x) = 0 \\ g_i(x) \leq 0, i = 1, \ldots, m \\ x \in X \end{array} \tag{PE}$$

onde o funcional $f : X \to \mathbb{R}$ e as funções $g_i : X \to \mathbb{R}, i = 1, \ldots, m$ são diferenciáveis, $F : X \to Y$ é continuamente diferenciável, X e Y são espaços reais finito dimensionais. Além disso denotamos $M = M_F \cap M_g$ o conjunto factível, isto é,

$$M = \{x \in X : F(x) = 0, g_i(x) \leq 0, i \in I\} \tag{5.1}$$

onde

$$\begin{array}{c} M_F = \{x \in X : F(x) = 0\}, \\ M_g = \{x \in X : g_i(x) \leq 0, i \in I\} \text{ e } I = \{i : i = 1, \ldots, m\}. \end{array} \tag{5.2}$$

Temos que um ponto x_0 factível para (PE) é regular quando a seguinte definição é satisfeita.

Definição 5.1 *Dizemos que x_0 é um ponto regular do conjunto factível M de (PE) se*

$$C_T(M_F, x_0) = \{w \in X : \nabla F(x_0)w = 0\}, \tag{5.3}$$

e se existe $\widetilde{w} \in X$ tal que

$$\begin{cases} \nabla F(x_0)\widetilde{w} = 0 \\ \nabla g_i(x_0)\widetilde{w} < 0, i \in I(x_0), \end{cases} \tag{5.4}$$

onde $I(x_0) = \{i \in I : g_i(x_0) = 0\}$ é o conjunto das restrições ativas. Neste caso, dizemos que o problema (PE) é regular em x_0.

Se x_0 não é regular segundo a definição acima, não é possível caracterizar o cone tangente e o cone das direções factíveis por meio das derivadas de primeira ordem dos operadores que determinam as restrições de igualdade e desigualdade do problema em questão. Neste caso a Regra dos Multiplicadores de Lagrange é degenerada, ou seja, não produz informação sobre o extremo em estudo.

Passaremos agora a estabelecer condições necessárias de otimalidade para x_0, mesmo que este seja não regular, mediante uma generalização da Regra dos Multiplicadores de Lagrange que garante que as condições são não degeneradas, sob hipóteses mais fracas que as de um ponto regular.

Apresentaremos o conceito de ponto 2-regular que é fundamental para estabelecer condições necessárias e suficientes de otimalidade não degeneradas de segunda ordem, e caracterizaremos os cones tangente e de direções factíveis juntamente com seus respectivos duais, fazendo uso das derivadas de primeira e segunda ordem dos operadores F e g, que determinam as restrições de igualdade e desigualdade. Para tanto, necessitamos como hipóteses:

1. F e g sejam duas vezes diferenciáveis em uma vizinhança $V(x_0)$ de $x_0 \in M$;

Cone tangente de segunda ordem

A finalidade agora é apresentar o cone tangente para o conjunto com interior vazio M_F, mediante aproximações de segunda ordem.

Para este fim, empregamos a mesma técnica utilizada em [3], [6] e [7].

$$H_F(x_0) = \left\{w \in X : \nabla F(x_0)w = 0, \nabla^2 F(x_0)[w,w] \in \mathbf{Im}\nabla F(x_0)\right\}, \tag{5.5}$$

onde os colchetes denotam a ação de uma forma bilinear sobre w e $\nabla^2 F(x_0)(w)w = \nabla^2 F(x_0)[w,w]$.

Note que a condição $\nabla^2 F(x_0)[w,w] \in \mathbf{Im}\nabla F(x_0)$ é equivalente a dizer que existe algum $x \in X$, que chamaremos x^w tal que

$$\nabla F(x_0)x^w + \nabla^2 F(x_0)[w,w] = 0,$$

ou ainda,

$$\left\langle \psi, \nabla^2 F(x_0)[w,w] \right\rangle = 0, \forall \psi \in [\mathbf{Im}\nabla F(x_0)]^\perp = \mathbf{Ker}\nabla F(x_0)^T.$$

Portanto, podemos escrever

$$\begin{aligned} H_F(x_0) &= \left\{w \in X : \exists x^w, \nabla F(x_0)w = 0, \nabla F(x_0)x^w + \nabla^2 F(x_0)[w,w] = 0\right\} \quad (5.6) \\ &= \left\{w \in X : \nabla F(x_0)w = 0, \left\langle \psi, \nabla^2 F(x_0)[w,w] \right\rangle = 0, \right. \\ &\quad \left. \forall \psi \in [\mathbf{Im}\nabla F(x_0)]^\perp = \mathbf{Ker}\nabla F(x_0)^T \right\}, \end{aligned}$$

onde $[\mathbf{Im}\nabla F(x_0)]^\perp$ é o subespaço ortogonal de $\mathbf{Im}\nabla F(x_0)$ e $\nabla F(x_0)^T : Y^* \to X^*$ é o operador adjunto de $\nabla F(x_0) : X \to Y$.

Para caracterizar o cone tangente a M_F no ponto x_0 apresentamos a seguinte definição de ponto 2-regular do conjunto M_F.

Definição 5.2 ***(Orellana-Vivanco [77])****Dizemos que um ponto x_0 é um ponto 2-regular do conjunto com interior vazio, M_F, se $x_0 \in M_F$ e o cone tangente no ponto x_0, $C_T(M_F, x_0)$, coincide com o cone $H_F(x_0)$,*

$$C_T(M_F, x_0) = \left\{w \in X : \exists x^w, \nabla F(x_0)w = 0, \nabla F(x_0)x^w + \nabla^2 F(x_0)[w, w] = 0\right\}. \quad (5.7)$$

Conforme a definição acima, se x_0 é um ponto 2-regular então w é uma direção tangente de M_F em x_0 se, e somente se, o sistema

$$\begin{array}{lcl} \nabla F(x_0)w & = & 0 \\ \nabla F(x_0)x^w + \nabla^2 F(x_0)[w, w] & = & 0, \end{array} \quad (5.8)$$

admite solução $x^w \in X$, isto é, o cone tangente é determinado por aproximações de segunda ordem.

Cone factível de segunda ordem

Vamos apresentar o cone factível para o conjunto das restrições de interior não vazio, M_g em um ponto não regular.

Sejam $x_0 \in M_g$ e $\nabla g(x_0) : X \to \mathbb{R}^s$, onde s é a cardinalidade do conjunto de índices das restrições ativas $I(x_0)$. O operador $\nabla g(x_0)$ será sobrejetivo se, e somente se, não existe $\mu = (\mu_1, \ldots, \mu_s) \neq 0 \in \mathbb{R}^s$ tal que $\nabla g(x_0)\mu = 0$ (veja [33]), o que implica que x_0 é um ponto regular de M_g.

Se x_0 não é um ponto regular a condição de regularidade

$$C_F(M_g, x_0) = \{w \in X : \nabla g_i(x_0)w < 0, i \in I(x_0)\}, \quad (5.9)$$

onde $I(x_0) = \{i \in I : g_i(x_0) = 0\}$, não se cumpre. O que implica que o sistema de desigualdades

$$\nabla g_i(x_0)w < 0, \forall i \in I(x_0)$$

não tem solução $w \in X$. Então, de acordo com o Lema 2.10, x_0 é um ponto estacionário de g, o que implica que $\nabla g(x_0)$ não é sobrejetiva, logo $\mathbf{Im}\nabla g(x_0)$ é um subconjunto próprio de $\mathbb{R}^s$ e como $\mathbb{R}^s$ é de dimensão finita e um subespaço fechado em $\mathbb{R}^s$.

Portanto, é possível representar $\mathbb{R}^s$ como uma soma direta $\mathbf{Im}\nabla g(x_0) \oplus [\mathbf{Im}\nabla g(x_0)]^\perp$ a qual é isomorfa a $\widetilde{W} = \mathbf{Im}\nabla g(x_0) \times [\mathbf{Im}\nabla g(x_0)]^\perp$, isto é,

$$\mathbb{R}^s = \mathbf{Im}\nabla g(x_0) \oplus [\mathbf{Im}\nabla g(x_0)]^\perp \cong \mathbf{Im}\nabla g(x_0) \times [\mathbf{Im}\nabla g(x_0)]^\perp := \widetilde{W},$$

onde $[\mathbf{Im}\nabla g(x_0)]^\perp$ é o subespaço ortogonal não nulo de $\mathbf{Im}\nabla g(x_0)$ em $\mathbb{R}^s$.

Consideremos agora o seguinte conjunto introduzido por Izamilov em [54],

$$H_g(x_0) = \left\{w \in X : \nabla g_i(x_0)w \leq 0, i \in I(x_0), -\nabla^2 g_i(x_0)[w, w] \in \mathbf{Im}\nabla g_i(x_0) + \mathbb{R}_+, i \in I_w(x_0)\right\}, \quad (5.10)$$

onde

$$I_w(x_0) = \{i \in I(x_0), \nabla g_i(x_0)w = 0\} \tag{5.11}$$

e os colchetes denotam a ação da forma bilinear sobre w, $\nabla^2 g_i(x_0)(w)w = \nabla^2 g_i(x_0)[w, w]$.

O conjunto $H_g(x_0)$ no espaço X é um cone de vértice em 0.

Note ainda que a condição $-\nabla^2 g_i(x_0)[w, w] \in \mathbf{Im}\nabla g_i(x_0) + \mathbb{R}_+, i \in I_w(x_0)$, é equivalente a dizer que existe um elemento em X, o qual denotamos por x^w tal que,

$$\nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w, w] \leq 0, i \in I_w(x_0).$$

Portanto podemos escrever,

$$\begin{aligned} H_g(x_0) = \{&w \in X : \exists x^w, \nabla g_i(x_0)w \leq 0, i \in I(x_0), \nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w, w] \leq 0, \\ &i \in I_w(x_0)\}. \end{aligned} \tag{5.12}$$

No que segue apresentaremos a definição de ponto 2-regular do conjunto M_g.

Definição 5.3 ***(Orellana-Vivanco [77])*** *Dizemos que um ponto $x_0 \in M_g$ é um ponto 2-regular de M_g se o cone factível ao conjunto M_g no ponto x_0, $C_F(M_g, x_0)$ pode ser caracterizado por aproximações de segunda ordem, isto é,*

$$\begin{aligned} C_F(M_g, x_0) \quad = \quad &\{w \in X : \exists x^w, \nabla g_i(x_0)w < 0, i \in I(x_0), \nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w, w] \leq 0, \\ &i \in I_w(x_0)\} \neq \emptyset \end{aligned} \tag{5.13}$$

com $I_w(x_0)$ definido em (5.11).

Observemos que se $I_w(x_0) = \emptyset$ em (5.13) então, $\nabla g_i(x_0)w < 0, \forall i \in I(x_0)$ e isto é equivalente a

$$C_F(M_g, x_0) = \{w \in X : \nabla g_i(x_0)w \leq 0, i \in I(x_0)\} \neq \emptyset,$$

o que implica, de acordo com a Definição 5.1, que x_0 é um ponto regular de M_g.

Da definição anterior segue que se x_0 é um ponto 2-regular então w é uma direção factível para M_g no ponto x_0 se, e somente se, o sistema de desigualdades

$$\begin{array}{lll} \nabla g_i(x_0)w & \leq & 0, i \in I(x_0) \\ \nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w, w] & < & 0, i \in I_w(x_0) \end{array} \tag{5.14}$$

admite solução $x^w \in X$.

Apresentaremos agora a definição de ponto 2-regular dada em [7] para um ponto factível de (PE).

Definição 5.4 ***(Avakov, Arutyunov e Izmailov [7])*** *Dizemos que um ponto x_0 é um* ***ponto 2-regular*** *do problema (PE), se para todo $w \in H_M(x_0)$, a condição de 2-regularidade: para todo $y = y_1 + y_2$, existem $q_1, q_2 \in X$ tais que*

$$\begin{array}{lll} \nabla F(x_0)q_2 & = & y_1 \\ \nabla F(x_0)q_1 + \nabla^2 F(x_0)[w, q_2] & = & y_2 \end{array} \tag{5.15}$$

se verifica e o sistema

$$\begin{array}{lcl} \nabla F(x_0)q_2 & = & 0 \\ \nabla F(x_0)q_1 + \nabla^2 F(x_0)[w, q_2] & = & 0 \\ \nabla g_i(x_0)q_2 & \leq & 0, \forall i \in I_w(x_0) \\ \nabla g_i(x_0)q_1 + \nabla^2 g_i(x_0)[w, q_2] & < & 0, \forall i \in \hat{I}_w(x_0) \end{array} \tag{5.16}$$

tem solução $q_1, q_2 \in X$, *onde*

$$\begin{aligned} H_M(x_0) = \{ & w \in X : \nabla F(x_0)w = 0, \nabla g_i(x_0)w \leq 0, i \in I(x_0), \exists x^w \in X; \\ & \nabla F(x_0)x^w + \nabla^2 F(x_0)[w,w] = 0, \nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w,w] < 0, i \in I_w(x_0) \big\}, \end{aligned} \tag{5.17}$$

$$\hat{I}_w(x_0) = \big\{ i \in I_w(x_0) : \nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w,w] = 0 \big\}. \tag{5.18}$$

Diremos que o problema (PE) é 2-regular em x_0 ou ainda que o conjunto factível M é 2-regular em x_0.

Observação 5.1 *É importante ressaltar que* $H_M(x_0) \subseteq H_F(x_0) \cap H_g(x_0)$ *e que sempre se verifica*

$$C_T(M_F, x_0) \subseteq H_F(x_0), \;\; C_F(M_g, x_0) \subseteq H_g(x_0),$$

onde

$$\begin{array}{rcl} H_F & = & \big\{ w \in X : \exists x^w \text{ com } \nabla F(x_0)w = 0, \nabla F(x_0)x^w + \nabla^2 F(x_0)[w,w] = 0 \big\}, \\ H_g & = & \big\{ w \in X : \exists x^w \text{ com } \nabla g_i(x_0)w \leq 0, i \in I(x_0), \nabla g_i(x_0)x^w + \nabla^2 g_i(x_0)[w,w] \leq 0, \\ & & i \in I_w(x_0) \big\}. \end{array}$$

Em particular, se x_0 *é um ponto 2-regular de* M *então as inclusões anteriores se cumprem na igualdade e portanto tem-se que*

$$H_M(x_0) \subseteq C_T(M_F, x_0) \cap C_F(M_g, x_0) \neq \emptyset.$$

Isto nos permite determinar condições de otimalidade de segunda ordem (não degeneradas) para (PE) *em um ponto 2-regular.*

Condições Necessárias em Programação Matemática

Para formular as condições necessárias de otimalidade de problemas com restrições mistas, de igualdade e desigualdade, precisamos caracterizar os cones de direções tangentes e factíveis no ponto ótimo. Se é possível determinar estes cones por aproximações de segunda ordem tais que a interseção dos correspondentes cones é não vazia, então o problema em questão é 2-regular e portanto podemos obter condições necessárias de segunda ordem não degeneradas. Nesta subseção estabelecemos condições necessárias de otimalidade para problemas não regulares mediante o formalismo de DM.

Se x_0 é uma solução ótima de (PE) então, de acordo com a Observação 2.5, $C_D(f, x_0) \cap C_F(M_g, x_0) \cap C_T(M_F, x_0) = \emptyset$, onde $C_D(f, x_0)$ é o cone de direções de descida de f, $\bigcap_{i=1}^{m} C_F(M_{g_i}, x_0) = C_F(M_g, x_0)$ é o cone de direções factíveis de $M_g(x_0)$ e $C_T(M_F, x_0)$ é cone de direções tangentes de $M_F(x_0)$. Logo a seguinte proposição vale.

Proposição 5.5 *Se x_0 é uma solução ótima local de (PE) então*

$$\nabla f(x_0)w \geq 0, \forall w \in C_T(M_F, x_0) \cap C_F(M_g, x_0). \tag{5.19}$$

No que segue apresentamos a definição de ponto Karush-Kuhn-Tucker-Avakov (KKTA) em uma certa direção do problema (PE). Este conceito é uma generalização do conceito de ponto KKT, foi introduzido em [7] e nomeado desta forma por [77].

Definição 5.6 *Dizemos que um ponto factível x_0 é um* ***ponto Karush-Kuhn-Tucker-Avakov*** *(KKTA) de (PE), na direção $w \in X$, se existem as funções $\mu = \mu(w), \hat{\mu} = \hat{\mu}(w) \in \mathbb{R}^m$, $\psi = \psi(w), \hat{\psi} = \hat{\psi}(w) \in Y^*$ não todas nulas, tais que*

$$\begin{aligned} \nabla f(x_0) + \sum_{i \in I_w(x_0)} \nabla g_i(x_0)^T \mu_i + \sum_{i \in \hat{I}_w(x_0)} \nabla^2 g_i(x_0)(w)^T \hat{\mu}_i + \nabla F(x_0)^T \psi + \nabla^2 F(x_0)(w)^T \hat{\psi} &= 0 \\ \sum_{i \in \hat{I}_w(x_0)} \nabla g_i(x_0)^T \hat{\mu}_i + \nabla F(x_0)^T \hat{\psi} &= 0 \end{aligned} \tag{5.20}$$

onde $\mu_i \geq 0, \hat{\mu}_i \geq 0, \psi$, e $\hat{\psi}$ são chamados os ***multiplicadores de Lagrange-Avakov****, $I_w(x_0)$ e $\hat{I}_w(x_0)$ são dado por (5.11) e (5.18) respectivamente.*

Com base no formalismo DM, temos o seguinte resultado:

Teorema 5.7 *(**Avakov, Arutyunov e Izmailov [7]**) Se um ponto $x_0 \in M$ é um ponto 2-regular segundo Definição 5.4 e uma solução ótima local de (PE), tal que $\nabla f(x_0)w = 0$, $w \in H_M(x_0)$, então x_0 é um ponto KKTA na direção de $w \in H_M(x_0)$, com $H_M(x_0)$ dado por (5.17).*

Para prova e maiores detalhes veja também [77].

Por fim, podemos concluir que:

Observação 5.2 *Se x_0 é um ótimo, de acordo com o Teorema 5.7, uma das afirmações a seguir ocorre:*

- *se x_0 é um ponto regular então as equações de Euler-Lagrange são válidas com $\hat{\mu}, \hat{\psi}$ identicamente nulos.*

- *se x_0 é um ponto 2-regular então, para cada $w \in H_M(x_0)$, exite uma família de equações de Euler-Langrange que são válidas com $\hat{\psi} = \hat{\psi}(w)$ e $\hat{\mu} = \hat{\mu}(w)$, não todos identicamente nulos.*

5.2 Condições necessárias

Nesta seção apresentaremos uma versão do Teorema do Princípio do Máximo para o Problema de Controle Ótimo Discreto não regular, $(PCD2)$, que será formulado no que se segue. Tais condições são obtidas a partir das condições de Karush-Kuhn-Tucker-Avakov (KKTA) apresentadas na Definição 5.6.

Inicialmente daremos a definição de processo 2-regular para as restrições do problema (PCD), definido no Capítulo 3. Para tanto, admitiremos que $\Phi_i : \mathbb{R}^n \times \mathbb{R}^m \to$

$\mathbb{R}, i = 1, \ldots, N-1$, são funções continuamente diferenciáveis e as demais funções do problema são duas vezes continuamente diferenciáveis. A este novo problema denotaremos por (PCD2).

Dessa maneira, temos que o conjunto dos processos factíveis do problema (PCD2) é o mesmo do problema (PCD), Q, dado em (3.5).

Dados $\mathbb{J}$, H e G como em (3.8), (3.9) e (3.10) temos que $(PCD2)$ pode ser escrito como o seguinte problema de programação matemática:

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(x,u) \\ \text{sujeito a: } H(x,u) = 0 \\ G(x,u) \leqq 0 \\ (x,u) \in X \times U. \end{array} \tag{POE2}$$

O conjunto factível de (POE2) é o conjunto $M = M_F \cap M_G$, onde

$$M_F = \{(x,u) \in X \times U : H(x,u) = 0\}, \ M_G = \{(x,u) \in X \times U : G(x,u) \leqq 0\}.$$

Assumimos ainda que $\mathbf{Im}\nabla H^l(\hat{x}, \hat{u})$ é um subespaço fechado de $V^l, l = 1, \ldots, 4$, onde $V^1 = \mathbb{R}^{n(N)}, V^2 = \mathbb{R}^{r_1(N+1)}, V^3 = \mathbb{R}^{r_2(N)}$ e $V^4 = \mathbb{R}^k$, logo $\mathbf{Im}\nabla H(\hat{x}, \hat{u})$ é um subespaço fechado em $V = V^1 \times V^2 \times V^3 \times V^4$.

Em [77] a autora propõe a noção de ponto 2-regular para problemas de controle ótimo contínuos, a definição que apresentaremos no que se segue é a extensão deste conceito aos problemas de controle ótimo discretos.

Definição 5.8 *Dizemos que o par $(\hat{x}, \hat{u})$ é um **processo 2-regular** do problema (PCD2), se e somente se o cone tangente as restrições de igualdade, no processo factível $(\hat{x}, \hat{u})$, é*

$$\begin{aligned} C_T(Q_1, (\hat{x}, \hat{u})) = \{(w,v) \in X \times U, \ \exists\ (x^w, u^v) \in X \times U : & \\ & w_{i+1} = \nabla\varphi_i\left(\hat{x}_i, \hat{u}_i\right)(w_i, v_i), \\ & \nabla h_i(\hat{x}_i) w_i = 0, \nabla\widetilde{h}_i(\hat{u}_i) v_i = 0, \\ & \nabla K(\hat{x}_0, \hat{x}_N)(w_0, w_N) = 0, \\ & x^w_{i+1} = \nabla\varphi_i\left(\hat{x}_i, \hat{u}_i\right)(x^w_i, u^v_i) + \nabla^2\varphi_i\left(\hat{x}_i, \hat{u}_i\right)[(w_i, v_i), (w_i, v_i)] \\ & \nabla h_i(\hat{x}_i) x^w_i + \nabla h_i(\hat{x}_i)[w_i, w_i] = 0, \nabla\widetilde{h}_i(\hat{u}_i) u^v_i + \nabla\widetilde{h}_i(\hat{u}_i)[v_i, v_i] = 0 \\ & \nabla K(\hat{x}_0, \hat{x}_N)(w_0, w_N) + \nabla^2 K(\hat{x}_0, \hat{x}_N)[(w_0, w_0), (w_N, w_N)] = 0. \end{aligned} \tag{5.21}$$

Além disso, para todo $(w,v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ existem $(w^1, v^1), (w^2, v^2)$ tais que

$$\begin{aligned} & w^2_{i+1} - \nabla\varphi_i\left(\hat{x}_i, \hat{u}_i\right)(w^2_i, v^2_i) = 0, i \in [0, N-1] \\ & \nabla h_i(\hat{x}_i) w^2_i = 0, i \in [0, N] \\ & \nabla\widetilde{h}_i(\hat{u}_i) v^2_i = 0, i \in [0, N-1] \\ & \nabla K(\hat{x}_0, \hat{x}_N)(w^2_0, w^2_N) = 0 \\ & w^1_{i+1} - \nabla\varphi_i\left(\hat{x}_i, \hat{u}_i\right)[(w_i, v_i), (w^1_i, v^1_i)] = 0, i \in [0, N-1] \\ & \nabla h_i(\hat{x}_i) w^1_i + \nabla h_i(\hat{x}_i)[w_i, w^1_i] = 0, i \in [0, N] \\ & \nabla\widetilde{h}_i(\hat{u}_i) v^1_i + \nabla\widetilde{h}_i(\hat{u}_i)[v_i, v^1_i] = 0, i \in [0, N-1] \\ & \nabla K(\hat{x}_0, \hat{x}_N)(w^1_0, w^1_N) + \nabla^2 K(\hat{x}_0, \hat{x}_N)[(w_0, w_N), (w^1_0, w^1_N] = 0, \end{aligned} \tag{5.22}$$

e

$$\begin{aligned}
&\nabla g_i(\hat{x}_i)w_i^2 < 0, i \in I^1(\hat{x}),\\
&\nabla \widetilde{g}_i(\hat{u}_i)v_i^2 < 0, i \in \widetilde{I}^1(\hat{u}),\\
&\nabla g_i(\hat{x}_i)w_i^1 + \nabla^2 g_i(\hat{x}_i)[w_i, w_i^2] < 0, i \in I^2(\hat{x}),\\
&\nabla \widetilde{g}_i(\hat{u}_i)v_i^1 + \widetilde{g}_i(\hat{u}_i)[v_i, v_i^2] < 0, i \in \widetilde{I}^2(\hat{u}),
\end{aligned} \tag{5.23}$$

com as seguintes condições de contorno

$$\begin{aligned}
&\nabla K(\hat{x}_0, \hat{x}_N)(w_0^2, w_N^2) = 0,\\
&\nabla K(\hat{x}_0, \hat{x}_N)(w_0^1, w_N^1) + \nabla^2 K(\hat{x}_0, \hat{x}_N)[(w_0, w_N), (w_0^2, w_0^2)] = 0,
\end{aligned} \tag{5.24}$$

onde,

$$\begin{aligned}
\mathbb{T}(Q, (\hat{x}, \hat{u})) \ = \ &\big\{(w, v) \in C_T(Q_1, (\hat{x}, \hat{u})), \nabla g_i(\hat{x}_i)w_i \leq 0, i \in I^1(\hat{x}),\\
&\nabla \widetilde{g}_i(\hat{u}_i)v_i \leq 0, i \in \widetilde{I}^1(\hat{u}),\\
&\nabla g_i(\hat{x}_i)x_i^w + \nabla^2 g_i(\hat{x}_i)[w_i, w_i] < 0, i \in I^2(\hat{x}),\\
&\nabla \widetilde{g}_i(\hat{u}_i)u_i^v + \nabla^2 \widetilde{g}_i(\hat{u}_i)[v_i, v_i] < 0, i \in \widetilde{I}^2(\hat{u})\Big\},
\end{aligned} \tag{5.25}$$

$$I^2(\hat{x}) = \big\{i \in I^1(\hat{x}) : \nabla g_i^s(\hat{x}_i)w_i + \nabla^2 g_i^s(\hat{x}_i)[w_i, w_i] = 0, \forall s \in I_i^2(\hat{x})\big\} \tag{5.26}$$

$$\widetilde{I}^2(\hat{u}) = \Big\{i \in \widetilde{I}^1(\hat{u}) : \nabla \widetilde{g}_i^t(\hat{u}_i)v_i + \nabla^2 \widetilde{g}_i^t(\hat{u}_i)[v_i, v_i] = 0, \forall t \in \widetilde{I}_i^2(\hat{u})\Big\}, \tag{5.27}$$

$$I_i^2(\hat{x}) = \big\{s \in I_i^1(\hat{x}) : \nabla g_i^s(\hat{x}_i)w_i + \nabla^2 g_i^s(\hat{x}_i)[w_i, w_i] = 0\big\}\ i = 0, \ldots, N, \tag{5.28}$$

$$\widetilde{I}_i^2(\hat{u}) = \Big\{t \in \widetilde{I}_i^1(\hat{u}) : \nabla \widetilde{g}_i^t(\hat{u}_i)v_i + \nabla^2 g_i^t(\hat{u}_i)[v_i, v_i] = 0\Big\}\ i = 0, \ldots, N-1, \tag{5.29}$$

$I^1(\hat{x})$ *e* $\widetilde{I}^1(\hat{u})$ *definidos em (3.25), (3.26), respectivamente.*

O lema a seguir será útil na demonstração do Teorema do Princípio do Máximo Estendido.

Lema 5.9 *O processo* $(\hat{x}, \hat{u})$ *é 2-regular do problema de controle (PCD2) se, e somente se,* $(\hat{x}, \hat{u})$ *é um ponto 2-regular do problema de programação matemática (POE2).*

Demonstração. É imediato que as equações abaixo apresentadas de forma explícita,

$$w_{i+1} = \nabla \varphi_i\left(\hat{x}_i, \hat{u}_i\right)(w_i, v_i), \nabla h_i(\hat{x}_i)w_i = 0, \nabla \widetilde{h}_i(\hat{u}_i)v_i = 0, \nabla K(\hat{x}_0, \hat{x}_N)(w_0, w_N) = 0,$$

são equivalentes a $\nabla H(\hat{x}, \hat{u})(w, v) = 0$, calculada em (3.12).

Analogamente, é fácil ver que a condição de que existe (x^w, u^v) tal que

$$\begin{aligned}
&x_{i+1}^w = \nabla \varphi_i\left(\hat{x}_i, \hat{u}_i\right)(x_i^w, u_i^v) + \nabla^2 \varphi_i\left(\hat{x}_i, \hat{u}_i\right)[(w_i, v_i), (w_i, v_i)]\\
&\nabla h_i(\hat{x}_i)x_i^w + \nabla h_i(\hat{x}_i)[w_i, w_i] = 0\\
&\nabla \widetilde{h}_i(\hat{u}_i)u_i^v + \nabla \widetilde{h}_i(\hat{u}_i)[v_i, v_i] = 0\\
&\nabla K(\hat{x}_0, \hat{x}_N)(w_0, w_N) + \nabla^2 K(\hat{x}_0, \hat{x}_N)[(w_0, w_N), (w_0, w_N)] = 0,
\end{aligned}$$

é equivalente à condição que existe $(x^w, u^v) \in X \times U$ tal que

$$\nabla H(\hat{x}, \hat{u})(x^w, u^v) + \nabla^2 H(\hat{x}, \hat{u})[d, d] = 0,$$

sendo

$$\begin{aligned} \nabla^2 H(\hat{x}, \hat{u})(w, v) &: X \times U \to \mathbb{R}^{n(N+1)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k; \\ \nabla^2 H(\hat{x}, \hat{u})[(w, v), (\bar{w}, \bar{v})] &= \big(\nabla^2 H^1(\hat{x}, \hat{u})[(w, v), (\bar{w}, \bar{v})], \nabla^2 H^2(\hat{x}, \hat{u})[(w, v), (\bar{w}, \bar{v})], \\ &\quad \nabla^2 H^3(\hat{x}, \hat{u})[(w, v), (\bar{w}, \bar{v})], \nabla^2 H^4(\hat{x}, \hat{u})[(w, v), (\bar{w}, \bar{v})]\big) \end{aligned} \tag{5.30}$$

onde

$$\begin{aligned} \nabla^2 H^1_{i+1}(\hat{x}_i, \hat{u}_i)[(w_i, v_i), (\bar{w}_i, \bar{v}_i)] &= -\nabla^2_{xx}\varphi_i\,(\hat{x}_i, \hat{u}_i)\, w_i\bar{w}_i - \nabla^2_{xu}\varphi_i\,(\hat{x}_i, \hat{u}_i)\, w_i\bar{v}_i \\ &\quad -\nabla^2_{ux}\varphi_i\,(\hat{x}_i, \hat{u}_i)\, v_i\bar{w}_i - \nabla^2_{uu}\varphi_i\,(\hat{x}_i, \hat{u}_i)\, v_i\bar{v}_i,\ \ i = 0, \ldots, N-1, \\ \nabla^2 H^2_i(\hat{x}_i, \hat{u}_i)[(w_i, v_i), (\bar{w}_i, \bar{v}_i)] &= \nabla^2 h_i(\hat{x}_i) w_i \bar{w}_i,\ \ i = 0, \ldots, N, \\ \nabla^2 H^3_i(\hat{x}_i, \hat{u}_i)[(w_i, v_i), (\bar{w}_i, \bar{v}_i)] &= \nabla^2 \widetilde{h}_i(\hat{u}_i) v_i \bar{v}_i,\ \ i = 0, \ldots, N-1, \\ \nabla^2 H^4(\hat{x}, \hat{u})[(w, v), (\bar{w}, \bar{v})] &= \nabla^2_{x_0x_0} K(\hat{x}_0, \hat{x}_N) w_0, \bar{w}_0 + \nabla^2_{x_0x_N} K(\hat{x}_0, \hat{x}_N) w_0, \bar{w}_N \\ &\quad +\nabla^2_{x_Nx_0} K(\hat{x}_0, \hat{x}_N) w_N, \bar{w}_0 + \nabla^2_{x_Nx_N} K(\hat{x}_0, \hat{x}_N) w_N, \bar{w}_N. \end{aligned}$$

e $d = (w, v) \in X \times U$.

Portanto podemos escrever o cone tangente às restrições de igualdade no ponto $(\hat{x}, \hat{u})$ como

$$\begin{aligned} C_T(Q_1, (\hat{x}, \hat{u})) &= \{(w, v) \in X \times U,\ \exists (x^w, u^v) \in X \times U : \\ &\quad \nabla H(\hat{x}, \hat{u}) d = 0, \nabla H(\hat{x}, \hat{u})(x^w, u^v) + \nabla^2 H(\hat{x}, \hat{u})[d, d] = 0\}. \end{aligned} \tag{5.31}$$

Por outro lado observemos que $I^2(\hat{x})$ é formado pelos $i \in I^1(\hat{x})$, definido em (3.25), tal que $\nabla g^s_i(\hat{x}_i) x^w_i + \nabla^2 g^s_i(\hat{x}_i)[w_i, w_i] = 0, \forall s \in I^2_i(\hat{x})$ que equivale a

$$\begin{cases} \nabla G^{1s}_i(\hat{x}, \hat{u}) d = 0, \\ \nabla G^{1s}_i(\hat{x}, \hat{u})(x^w, u^v) + \nabla^2 G^{1s}_i(\hat{x}, \hat{u})[d, d] = 0, \end{cases} \tag{5.32}$$

onde $d = (w, v) \in X \times U$. O mesmo ocorre para $\widetilde{I}^2(\hat{u})$.

Portanto a condição: para todo $d = (w, v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ existem $(w^1, v^1), (w^2, v^2)$ satisfazendo (5.22), (5.23) e (5.24) simultaneamente, é equivalente a dizer: para todo $d = (w, v) \in \mathbb{T}(M, (\hat{x}, \hat{u}))$ o sistema

$$\begin{aligned} \nabla H(\hat{x}, \hat{u}) q_2 &= 0, \\ \nabla H(\hat{x}, \hat{u}) q_1 + \nabla^2 H(\hat{x}, \hat{u})[d, q_2] &= 0, \\ \nabla G^{1s}_i(\hat{x}, \hat{u}) q_2 &< 0, \forall s \in I^1_i(\hat{x}), i = 0, \ldots, N, \\ \nabla G^{2t}_i(\hat{x}, \hat{u}) q_2 &< 0, \forall t \in \widetilde{I}^1_i(\hat{u}), i = 0, \ldots, N-1, \\ \nabla G^{1s}_i(\hat{x}, \hat{u}) q_1 + \nabla^2 G^{1s}_i(\hat{x}, \hat{u})[d, q_2] &< 0, \forall s \in I^2_i(\hat{x}), i = 0, \ldots, N, \\ \nabla G^{2t}_i(\hat{x}, \hat{u}) q_1 + \nabla^2 G^{2t}_i(\hat{x}, \hat{u})[d, q_2] &< 0, \forall t \in \widetilde{I}^2_i(\hat{u}), i = 0, \ldots, N-1, \end{aligned} \tag{5.33}$$

tem solução $q_1, q_2 \in X \times U$, $q_1 = (w^1, v^1)$ e $q_2 = (w^2, v^2)$, onde

$$\begin{aligned} \mathbb{T}(M, (\hat{x}, \hat{u})) = \big\{ & d = (w, v) \in X \times U : \nabla H(\hat{x}, \hat{u}) d = 0, \nabla G^{1s}_i(\hat{x}, \hat{u}) d \leqq 0, \forall s \in I^1_i(\hat{x}), \\ & i = 0, \ldots, N; \nabla G^{2t}_i(\hat{x}, \hat{u}) d \leqq 0, \forall t \in \widetilde{I}^1_i(\hat{u}), i = 0, \ldots, N-1; \\ & \exists (x^w, u^v) \in X \times U : \nabla H(\hat{x}, \hat{u})(x^w, u^v) + \nabla^2 H(\hat{x}, \hat{u})[d, d] = 0, \\ & \nabla G^{1s}_i(\hat{x}, \hat{u})(x^w, u^v) + \nabla G^{1s}_i(\hat{x}, \hat{u})[d, d] < 0, \forall s \in I^2_i(\hat{x}), i = 0, \ldots, N, \\ & \nabla G^{2t}_i(\hat{x}, \hat{u})(x^w, u^v) + \nabla G^{2t}_i(\hat{x}, \hat{u})[d, d] < 0, \forall t \in \widetilde{I}^2_i(\hat{u}), i = 0, \ldots, N-1 \big\} \end{aligned} \tag{5.34}$$

e M é o conjunto factível de (POE2).

Portanto, de acordo com a Definição 5.4, $(\hat{x},\hat{u})$ é um ponto 2-regular de (POE2).

□

No que se segue apresentamos o Teorema do Princípio do Máximo Estendido para o problema (PCD2).

Teorema 5.10 (Princípio do Máximo Discreto Estendido) *Seja o processo factível, $(\hat{x},\hat{u}) \in Q$. Se $(\hat{x},\hat{u})$ é um processo ótimo de $(PCD2)$ então, para todo $d=(w,v) \in \mathbb{T}(Q,(\hat{x},\hat{u}))$ tal que $\nabla\mathbb{J}(\hat{x},\hat{u})d=0$, existem:*
$\lambda_0=\lambda_0(d) \in \mathbb{R}$, com $\lambda_0 \geq 0$, $p=p(d) \in \mathbb{R}^{n(N+1)}$, $(\lambda_0,p) \neq 0$, $\hat{p}=\hat{p}(d), \hat{p}_i \in R^{n(N+1)}$, $\gamma^1=\gamma^1(d)$ e $\hat{\gamma}^1=\hat{\gamma}^1(d), \in \mathbb{R}^{r_1(N+1)}$, $\gamma^2=\gamma^2(d)$ e $\hat{\gamma}^2=\hat{\gamma}^2(d) \in \mathbb{R}^{r_2(N)}$, γ^3 e $\hat{\gamma}^3 \in \mathbb{R}^k$, $\mu^1=\mu^1(d)$ e $\hat{\mu}^1=\hat{\mu}^1(d) \in \mathbb{R}^{s_1}$, $\mu^1_i \geq 0, \hat{\mu}^1_i \geq 0$, $\forall i \in [0,N]$, $\mu^2=\mu^2(d)$ e $\hat{\mu}^2=\hat{\mu}^2(d) \in \mathbb{R}^{s_2(N)}$, $\mu^2_i \geq 0, \hat{\mu}^2_i \geq 0$, $\forall i \in [0,N-1]$, não todos simultaneamente nulos, tais que

$$\begin{aligned} p_0 \;=\;& \lambda_0\nabla_x\Phi_0(\hat{x}_0,\hat{u}_0) - \nabla_x\varphi_0(\hat{x}_0,\hat{u}_0)^T p_1 \\ & - \left[\nabla^2_{xx}\varphi_0(\hat{x}_0,\hat{u}_0)w_0\right]^T \hat{p}_1 - \left[\nabla^2_{ux}\varphi_0(\hat{x}_0,\hat{u}_0)v_0\right]^T \hat{p}_1 \\ & -\nabla_{x_0}K(\hat{x}_0,\hat{x}_N)^T\gamma^3 - \left[\nabla^2_{x_0x_0}K(\hat{x}_0,\hat{x}_N)w_0\right]^T \hat{\gamma^3} - \left[\nabla^2_{x_Nx_0}K(\hat{x}_0,\hat{x}_N)w_N\right]^T \hat{\gamma^3} \\ & -\nabla h_0(\hat{x}_0)^T\gamma^1_0 - \left[\nabla^2 h_0(\hat{x}_0)w_0\right]^T \hat{\gamma}^1_0 - \nabla g_0(\hat{x}_0)^T\mu^1_0 - \left[\nabla^2 g_0(\hat{x}_0)w_0\right]^T \hat{\mu}^1_0, \end{aligned} \tag{5.35}$$

$$\begin{aligned} p_i \;=\;& \lambda_0\nabla_x\Phi_i(\hat{x}_i,\hat{u}_i) - \nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)^T p_{i+1} \\ & - \left[\nabla^2_{xx}\varphi_i(\hat{x}_i,\hat{u}_i)w_i\right]^T \hat{p}_{i+1} - \left[\nabla^2_{ux}\varphi_i(\hat{x}_i,\hat{u}_i)v_i\right]^T \hat{p}_{i+1} \\ & -\nabla h_i(\hat{x}_i)^T\gamma^1_i - \left[\nabla^2 h_i(\hat{x}_i)w_i\right]^T \hat{\gamma}^1_i - \nabla g_i(\hat{x}_i)^T\mu^1_i - \left[\nabla^2 g_i(\hat{x}_i)w_i\right]^T \hat{\mu}^1_i, \\ & i=1,\ldots,N-1, \end{aligned} \tag{5.36}$$

$$\begin{aligned} p_N \;=\;& -\nabla_{x_N}K(\hat{x}_0,\hat{x}_N)^T\gamma^3 - \left[\nabla^2_{x_Nx_N}K(\hat{x}_0,\hat{x}_N)w_N\right]^T \hat{\gamma^3} - \left[\nabla^2_{x_0x_N}K(\hat{x}_0,\hat{x}_N)w_0\right]^T \hat{\gamma^3} \\ & -\nabla h_N(\hat{x}_N)^T\gamma^1_N - \left[\nabla^2 h_N(\hat{x}_N)w_N\right]^T \hat{\gamma}^1_N \\ & -\nabla g_N(\hat{x}_N)^T\mu^1_N - \left[\nabla^2 g_N(\hat{x}_N)w_N\right]^T \hat{\mu}^1_N, \end{aligned} \tag{5.37}$$

$$\begin{aligned} \hat{p}_0 \;=\;& -\nabla_x\varphi_0(\hat{x}_0,\hat{u}_0)^T\hat{p}_1 - \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)^T\hat{\gamma^3} - \nabla h_0(\hat{x}_0)^T\hat{\gamma}^1_0 \\ & -\nabla g_0(\hat{x}_0)^T\hat{\mu}^1_0, \end{aligned} \tag{5.38}$$

$$\hat{p}_i \;=\; -\nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)^T\hat{p}_{i+1} - \nabla h_i(\hat{x}_i)^T\hat{\gamma}^1_i - \nabla g_i(\hat{x}_i)^T\hat{\mu}^1_i, i=1,\ldots,N-1, \tag{5.39}$$

$$\hat{p}_N \;=\; -\nabla h_N(x_N)^T\hat{\gamma}^1_N - \nabla g_N(x_N)^T\hat{\mu}^1_N - \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)^T\hat{\gamma}^3. \tag{5.40}$$

Além disso, satisfaz a condição de máximo

$$\begin{aligned} & \lambda_0\nabla_u\Phi_i(\hat{x}_i,\hat{u}_i) - \nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)^T p_{i+1} - \left[\nabla^2_{uu}\varphi_i(\hat{x}_i,\hat{u}_i)v_i\right]^T \hat{p}_{i+1} - \left[\nabla^2_{xu}\varphi_i(\hat{x}_i,\hat{u}_i)w_i\right]^T \hat{p}_{i+1} \\ & \quad -\nabla\widetilde{h}_i(\hat{u}_i)^T\gamma^2_i - \left[\nabla^2\widetilde{h}_i(\hat{u}_i)v_i\right]^T \hat{\gamma}^2_i - \nabla\widetilde{g}_i(\hat{u}_i)^T\mu^2_i - \left[\nabla^2\widetilde{g}_i(\hat{u}_i)v_i\right]^T \hat{\mu}^2_i = 0, \\ & \qquad i=0,\ldots,N-1, \end{aligned} \tag{5.41}$$

e a condição

$$-\nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)^T\hat{p}_{i+1} - \nabla\widetilde{h}_i(\hat{u}_i)^T\hat{\gamma}_i^2 - \nabla\widetilde{g}_i(\hat{u}_i)^T\hat{\mu}_i^2 = 0, i = 1,\ldots,N-1. \qquad (5.42)$$

Demonstração. Iniciamos a demonstração observando que:

(a) Se $\nabla\mathbb{J}(\hat{x},\hat{u}) = 0$, então o teorema é satisfeito trivialmente, basta assumir $\lambda_0 = 1$ e $p, \hat{p}$, $\gamma^1, \hat{\gamma}^1$, $\gamma^2, \hat{\gamma}^2$, $\gamma^3, \hat{\gamma}^3$ μ^1, $\hat{\mu}^1$, μ^2 e $\hat{\mu}^2$ identicamente nulos.

(b) Se $(\hat{x},\hat{u})$ não é um processo 2-regular de Q do problema (PCD2), de acordo com a Definição 5.8, então nos encontramos em um dos dois casos abaixo:

1. $C_T(Q_1,(\hat{x},\hat{u}))$ não satisfaz (5.21).
   Neste caso não existe $(x^w,v^u) \in X\times U$ satisfazendo (5.21). Então, o operador

   $$\nabla H(\hat{x},\hat{u})((x^w,u^v),(w,v)) : X\times U \to V := \mathbf{Im}\nabla H(\hat{x},\hat{u}) \oplus [\mathbf{Im}\nabla H(\hat{x},\hat{u})]^{\perp},$$

   do problema equivalente (POE2), não é sobrejetivo. Isto é equivalente, sempre que o subespaço $\mathbf{Im}\nabla H(\hat{x},\hat{u})((x^w,u^v),(w,v))$ seja fechado em V, que o subespaço anulador de $\mathbf{Im}\nabla H(\hat{x},\hat{u})$ é não nulo. Logo, existe um elemento não nulo $\tilde{\psi}^* = (\psi^*,\hat{\psi}^*) \in V^*$ tal que,

   $$\nabla H(\hat{x},\hat{u})((x^w,u^v),(w,v))\tilde{\psi}^T = 0$$

   onde $\psi^* = (\psi_1^*,\psi_2^*,\psi_3^*,\psi_4^*) \in \mathbf{Im}\nabla H(\hat{x},\hat{u})^T$ e $\hat{\psi}^* = (\hat{\psi}_1^*,\hat{\psi}_2^*,\hat{\psi}_3^*,\hat{\psi}_4^*) \in [\mathbf{Im}\nabla H(\hat{x},\hat{u})]^{\perp}$ $= \mathbf{Ker}\nabla H(\hat{x},\hat{u})^T$. Portanto, as equações (5.35)-(5.42) são satisfeitas, basta considerar $\lambda_0 = 0$, $p = (p_0,\tilde{p})$ com $p_0 = 0, \tilde{p} = \psi_1^*$, $\gamma^1 = \psi_2^*, \gamma^2 = \psi_3^*$, $\gamma^3 = \psi_4^*$, $\hat{p} = (\hat{p}_0,\tilde{p}_1)$ com $\hat{p}_0 = 0, \tilde{p}_1 = \hat{\psi}_1^*$, $\hat{\gamma}^1 = \hat{\psi}_2^*, \hat{\gamma}^2 = \hat{\psi}_3^*$, $\hat{\gamma}^3 = \hat{\psi}_4^*$ e $\mu^1 = 0, \hat{\mu}^1 = 0$, $\mu^2 = 0, \hat{\mu}^2 = 0$.

2. $C_T(Q_1,(\hat{x},\hat{u}))$ satisfaz (5.21) porém, não existem $(w^1,v^1),(w^2,v^2)$ tais que, para todo $(w,v) \in \mathbb{T}(Q,(\hat{x},\hat{u}))$, sejam válidas (5.22), (5.23).
   Como $C_T(Q_1,(\hat{x},\hat{u}))$ satisfaz (5.21) então temos que existem $(w^1,v^1),(w^2,v^2)$ tais que, para todo $(w,v) \in \mathbb{T}(Q,(\hat{x},\hat{u}))$ as equações (5.22) ocorrem. Portanto, neste caso, não existem $(w^1,v^1),(w^2,v^2)$ tais que, para todo $(w,v) \in \mathbb{T}(Q,(\hat{x},\hat{u}))$ vale (5.23). Isto implica que $((x^w,u^v),(w,v))$ é um ponto estacionário de g ou $\widetilde{g}$, segundo a Definição 2.8. Portanto pelo Lema 2.10, existe um elemento não nulo $(\mu_i^{1*},\hat{\mu}^{1*}), i \in I^2(\hat{x})$, $(\mu_i^{1*},\hat{\mu}_i^{1*}) \geqq 0$, em $\mathbb{R}^{s_1} := \Big(\mathbf{Im}\nabla g_i(\hat{x}) \times [\mathbf{Im}\nabla g_i(\hat{x})]^{\perp}\Big)^T$ ou $(\mu_i^{2^*},\hat{\mu}_i^{2*}), i \in \widetilde{I}^2(\hat{u})$, $(\mu_i^{2*},\hat{\mu}_i^{2*}) \geqq 0$ em $\mathbb{R}^{s_2} := \Big(\mathbf{Im}\nabla\widetilde{g}_i(\hat{u}) \times [\mathbf{Im}\nabla\widetilde{g}_i(\hat{u})]^{\perp}\Big)^T$ onde $[\mathbf{Im}\nabla g_i(\hat{x})]^{\perp} = \mathbf{Ker}\nabla g_i(\hat{x})^T$ e $[\mathbf{Im}\nabla\widetilde{g}_i(\hat{u})]^{\perp} = \mathbf{Ker}\nabla\widetilde{g}_i(\hat{u})^T$. Então as condições (5.35)-(5.42) são trivialmente satisfeitas, basta considerar $\lambda_0 = 0$, $p,\hat{p}$ $\gamma^1,\hat{\gamma}^1$, $\gamma^2,\hat{\gamma}^2$ $\gamma^3,\hat{\gamma}^3$ como em (5.35)-(5.42) e $(\mu^1,\hat{\mu}^1) \geqq 0$ ou $(\mu^2,\hat{\mu}^2) \geqq 0$, com os dois últimos não ambos nulos.

Portanto se o processo ótimo $(\hat{x},\hat{u})$ não é 2-regular, podemos tomar $\lambda_0 = 0$, $\psi_1 = p, \hat{\psi}_1 = \hat{p}$, $\psi_2 = \gamma^1, \hat{\psi}_2 = \hat{\gamma}^1$, $\psi_3 = \gamma^2, \hat{\psi}_3 = \hat{\gamma}^2$, $\psi_4 = \gamma^3, \hat{\psi}_4 = \hat{\gamma}^3$, $\mu^1,\hat{\mu}^1,\mu^2,\hat{\mu}^2 \geqq 0$, não todos nulos, satisfazendo as condições (5.35)-(5.42) do teorema.

(c) Resta então provar o teorema para o caso em que o processo ótimo $(\hat{x},\hat{u})$ do problema $(PCD2)$ é um processo 2-regular. Neste caso, temos que $(\hat{x},\hat{u})$ é um ponto 2-regular de (POE2), acordo com o Lema 5.9.

A demonstração desta parte se baseia no Teorema 5.7, portanto, vamos verificar que as hipóteses do mesmo se cumprem.

De acordo com as hipóteses impostas às funções que definem o problema $(PCD2)$, sendo (PCD2) equivalente ao problema de programação matemática (POE2), é imediato que $\mathbb{J}$ (definido em (3.8)) é diferenciável e que os operadores H e G (dados por 3.9 e 3.10) são duas vezes continuamente diferenciáveis, com respeito a $(x,u) \in Q$.

Como o ponto factível $(\hat{x},\hat{u})$ é uma solução ótima e um ponto 2-regular de $(POE2)$ tal que $\nabla\mathbb{J}(\hat{x},\hat{u})(w,v) \geq 0$, para todo $d=(w,v) \in \mathbb{T}(M,(\hat{x},\hat{u}))$ segue, de acordo com o Teorema 5.7, que $(\hat{x},\hat{u})$ é um ponto KKTA de (POE2) na direção $(w,v) \in \mathbb{T}(M,(\hat{x},\hat{u}))$. Então, de acordo com a Definição 5.6, existem $\psi=(\psi_1,\psi_2,\psi_3,\psi_4)$ e $\hat{\psi}=(\hat{\psi}_1,\hat{\psi}_2,\hat{\psi}_3,\hat{\psi}_4) \in (\mathbb{R}^{n(N+1)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k)^T$, $\psi_l=\psi_l(d), \hat{\psi}_l=\hat{\psi}_l(d), l=1,\ldots,4$, $\mu^1,\hat{\mu}^1 \in (\mathbb{R}^{s_1(N+1)})^T$, $\mu^2,\hat{\mu}^2 \in (\mathbb{R}^{s_2(N)})^T$, $\mu^1=\mu^1(d) \geqq 0, \mu^2=\mu^2(d) \geqq 0$, $\hat{\mu}^1=\hat{\mu}^1(d) \geqq 0$, e $\hat{\mu}^2=\hat{\mu}^2(d) \geqq 0$ não todos identicamente nulos, tais que

$$\begin{aligned}
\nabla\mathbb{J}(\hat{x},\hat{u}) + \sum_{i\in I^1(\hat{x})} \nabla G_i^1(\hat{x},\hat{u})^T\mu_i^1 + \sum_{i\in I^1(\hat{u})} \nabla G_i^2(\hat{x},\hat{u})^T\mu_i^2 + \sum_{i\in I^2(\hat{x})} \left[\nabla^2 G_i^1(\hat{x},\hat{u})d\right]^T\hat{\mu}_i^1 & \\
+ \sum_{i\in \widetilde{I}^2(\hat{u})} \left[\nabla^2 G_i^2(\hat{x},\hat{u})d\right]^T\hat{\mu}_i^2 + \sum_{l=1}^4 \nabla H^l(\hat{x},\hat{u})^T\psi_k + \sum_{l=1}^4 \left[\nabla^2 H^l(\hat{x},\hat{u})d\right]^T\hat{\psi}_k &= 0 \\
\sum_{i\in I^2(\hat{x})} \nabla G_i^1(\hat{x},\hat{u})^T\hat{\mu}_i^1 + \sum_{i\in \widetilde{I}^2(\hat{u})} \nabla G_i^2(\hat{x},\hat{u})^T\hat{\mu}_i^2 + \sum_{l=1}^4 \nabla H^k(\hat{x},\hat{u})^T\hat{\psi}_k &= 0,
\end{aligned}$$

onde $\mu_i^1(z)$ e $\hat{\mu}_i^1(\bar{z}_i)$ são não negativos para todo $z_i \in \mathbf{Im}\nabla g_i(\hat{x})$ (subespaço fechado de $\mathbb{R}^{s_1}$) e $\bar{z} \in (\mathbb{R}^{s_1} \setminus \mathbf{Im}\nabla g_i(\hat{x},\hat{u})), \forall i=0,\ldots,N$, e analogamente para $\mu_i^2(t)$ e $\hat{\mu}_i^2(\bar{t}_i)$.

Notemos que por $\mathbf{Im}\nabla G_i^1(\hat{x},\hat{u})$, $\mathbf{Im}\nabla G_i^2(\hat{x},\hat{u})$ serem subespaços fechados de $\mathbb{R}^{s_1}$ e $\mathbb{R}^{s_2}$, respectivamente,

$$\begin{aligned}
&\mu_i^1 \in (\mathbf{Im}\nabla G_i^1(\hat{x},\hat{u}))^T \cong \mathbb{R}^{s_1} \setminus [\mathbf{Im}\nabla G_i^1(\hat{x},\hat{u})]^\perp \\
&\mu_i^2 \in (\mathbf{Im}\nabla G_i^2(\hat{x},\hat{u}))^T \cong \mathbb{R}^{s_2} \setminus [\mathbf{Im}\nabla G_i^2(\hat{x},\hat{u})]^\perp \\
&\hat{\mu}_i^1 \in (\mathbb{R}^{s_1} \setminus \mathbf{Im}\nabla G_i^1(\hat{x},\hat{u})^T) \cong [\mathbf{Im}\nabla G_i^1(\hat{x},\hat{u})]^\perp \cong \mathbf{Ker}\nabla G_i^1(\hat{x},\hat{u})^T, \\
&\hat{\mu}_i^2 \in (\mathbb{R}^{s_2} \setminus \mathbf{Im}\nabla G_i^2(\hat{x},\hat{u})^T) \cong [\mathbf{Im}\nabla G_i^2(\hat{x},\hat{u})]^\perp \cong \mathbf{Ker}\nabla G_i^2(\hat{x},\hat{u})^T,
\end{aligned} \tag{5.43}$$

onde $\cong$ indica a relação de isomorfismo isométrico.

Considerando agora a definição de operador adjunto obtemos para todo $\widetilde{d} = (\widetilde{w},\widetilde{v}) \in X \times U = \mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)}$,

$$\begin{aligned}
\nabla\mathbb{J}(\hat{x},\hat{u})(\widetilde{d}) + \sum_{i\in I^1(\hat{x})} \left\langle \mu_i^1, \nabla G_i^1(\hat{x},\hat{u})(\widetilde{d})\right\rangle + \sum_{i\in \widetilde{I}^1(\hat{u})} \left\langle \mu_i^2, \nabla G_i^2(\hat{x},\hat{u})(\widetilde{d})\right\rangle \\
+ \sum_{i\in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla^2 G_i^1(\hat{x},\hat{u})(d)(\widetilde{d})\right\rangle + \sum_{i\in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla^2 G_i^2(\hat{x},\hat{u})(d)(\widetilde{d})\right\rangle \\
+ \sum_{l=1}^4 \left\langle \psi_l, \nabla H^l(\hat{x},\hat{u})(\widetilde{d})\right\rangle + \sum_{l=1}^4 \left\langle \hat{\psi}_l, (\nabla^2 H^l(\hat{x},\hat{u})(d)(\widetilde{d})\right\rangle = 0,
\end{aligned} \tag{5.44}$$

$$\sum_{i\in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla G_i^1(\hat{x},\hat{u})(\widetilde{d})\right\rangle + \sum_{i\in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla G_i^2(\hat{x},\hat{u})(\widetilde{d})\right\rangle + \sum_{l=1}^4 \left\langle \hat{\psi}_l, \nabla H^l(\hat{x},\hat{u})(\widetilde{d})\right\rangle = 0, \tag{5.45}$$

onde $I^1(\hat{x})$, $\widetilde{I}^1(\hat{u})$, $I^2(\hat{x})$ e $\widetilde{I}^2(\hat{u})$ são definidos em (3.25), (3.26), (5.26) e (5.27), respectivamente.

No que segue vamos determinar explicitamente o termos da equação (5.44):

1. A primeira derivada do funcional $\mathbb{J}$ em $(\hat{x}, \hat{u})$ definida para cada $(\widetilde{w}, \widetilde{v})$ em $X \times U$, é

$$\nabla \mathbb{J}(\hat{x}, \hat{u})(\widetilde{w}, \widetilde{v}) = \sum_{i=0}^{N-1} \left(\nabla_x \Phi_i(\hat{x}_i, \hat{u}_i)\widetilde{w}_i + \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i \right). \tag{5.46}$$

2. Dado que $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$ e $\hat{\psi} = (\hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3, \hat{\psi}_4) \in (\mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k)^*$ podemos garantir, pela representação única de ψ_1 e $\hat{\psi}_1$ em $(\mathbb{R}^{n(N)})^T$ (veja [?]), a existência de p e $\hat{p} \in \mathbb{R}^{n(N)}$ tais que

$$\begin{aligned}
\left\langle \psi_1, \nabla H^1(\hat{x}, \hat{u})(\widetilde{w}, \widetilde{v}) \right\rangle &= \sum_{i=0}^{N-1} \left\langle -p_{i+1}, \nabla H^1_{i+1}(\hat{x}_i, \hat{u}_i)(\widetilde{w}_i, \widetilde{v}_i) \right\rangle \\
&= \sum_{i=0}^{N-1} \left\langle -p_{i+1}, \widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i)\widetilde{w}_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i \right\rangle \\
\left\langle \hat{\psi}_1, \nabla^2 H^1(\hat{x}, \hat{u})[(w, v), (\widetilde{w}, \widetilde{v})] \right\rangle &= \sum_{i=0}^{N-1} \left\langle -\hat{p}_{i+1} (\nabla^2 H^1_{i+1}(\hat{x}_i, \hat{u}_i)(w_i, v_i))(\widetilde{w}_i, \widetilde{v}_i) \right\rangle \\
&= \sum_{i=0}^{N-1} \Big\langle -\hat{p}_{i+1}, \nabla^2_{xx} \varphi_i(\hat{x}_i, \hat{u}_i)[w_i, \widetilde{w}_i] - \nabla^2_{xu} \varphi_i(\hat{x}_i, \hat{u}_i)[w_i, \widetilde{v}_i] \\
&\quad - \nabla^2_{ux} \varphi_i(\hat{x}_i, \hat{u}_i)[v_i, \widetilde{w}_i] - \nabla^2_{uu} \varphi_i(\hat{x}_i, \hat{u}_i)[v_i, \widetilde{v}_i] \Big\rangle .
\end{aligned}$$

Além disso, como $\psi_2, \hat{\psi}_2 \in (\mathbb{R}^{r_1(N+1)})^T = \mathbb{R}^{r_1(N+1)}$, existem vetores $\gamma^1, \hat{\gamma}^1 \in \mathbb{R}^{r_1(N+1)}$ tais que

$$\begin{aligned}
\left\langle \psi_2, \nabla H^2(\hat{x}, \hat{u})(\widetilde{w}, \widetilde{v}) \right\rangle &= \sum_{i=0}^{N} \left\langle -\gamma^1_i, \nabla H^2_i(\hat{x}_i, \hat{u}_i)(\widetilde{w}_i, \widetilde{v}_i) \right\rangle \\
&= \sum_{i=0}^{N} \left\langle -\gamma^1_i, \nabla h_i(\hat{x}_i)\widetilde{w}_i \right\rangle , \\
\left\langle \hat{\psi}_2, \nabla^2 H^2(\hat{x}, \hat{u})[(w, v), (\widetilde{w}, \widetilde{v})] \right\rangle &= \sum_{i=0}^{N} \left\langle -\hat{\gamma}^1_i, \nabla^2 H^2_i(\hat{x}_i, \hat{u}_i)(w_i, v_i))(\widetilde{w}_i, \widetilde{v}_i \right\rangle \\
&= \sum_{i=0}^{N} \left\langle -\hat{\gamma}^1_i, \nabla^2 h_i(\hat{x}_i)[w_i, \widetilde{w}_i] \right\rangle .
\end{aligned}$$

O mesmo ocorre para $\psi_3, \hat{\psi}_3 \in (\mathbb{R}^{r_2(N)})^T = \mathbb{R}^{r_2(N)}$ e então existem vetores $\gamma^2, \hat{\gamma}^2 \in \mathbb{R}^{r_2(N)}$ tais que

$$
\begin{aligned}
\left\langle \psi_3, \nabla H^3(\hat{x},\hat{u})(\widetilde{w},\widetilde{v})\right\rangle &= \sum_{i=0}^{N-1}\left\langle -\gamma_i^2, \nabla H_i^3(\hat{x}_i,\hat{u}_i)(\widetilde{w}_i,\widetilde{v}_i)\right\rangle \\
&= \sum_{i=0}^{N-1}\left\langle -\gamma_i^2, \nabla \widetilde{h}_i(u_i)\widetilde{v}_i\right\rangle \\
\left\langle \hat{\gamma}_2, \nabla^2 H_i^3(\hat{x},\hat{u})[(w,v),(\widetilde{w},\widetilde{v})]\right\rangle &= \sum_{i=0}^{N-1}\left\langle -\hat{\gamma}_i^2, \nabla^2 H_i^3(\hat{x}_i,\hat{u}_i)(w_i,v_i))(\widetilde{w}_i,\widetilde{v}_i)\right\rangle \\
&= \sum_{i=0}^{N-1}\left\langle -\hat{\gamma}_i^2, \nabla^2 \widetilde{h}_i(u_i)[v_i,\widetilde{v}_i]\right\rangle .
\end{aligned}
$$

Por fim, como $\psi_4, \hat{\psi}_4 \in (\mathbb{R}^k)^T = \mathbb{R}^k$, existem vetores $\gamma^3, \hat{\gamma^3} \in \mathbb{R}^k$ tais que

$$
\begin{aligned}
\left\langle \psi_4, \nabla H^4(\hat{x},\hat{u})(\widetilde{w},\widetilde{v})\right\rangle &= \left\langle -\gamma^3, \nabla_{x_0} k(\hat{x}_0,\hat{x}_N)\widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N\right\rangle \\
\left\langle \hat{\psi}_4, \nabla^2 H^4(\hat{x},\hat{u})[(w,v),(\widetilde{w},\widetilde{v})]\right\rangle &= \left\langle -\hat{\gamma}^3, \nabla^2_{x_0x_0} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_0] \right. \\
&\quad \left. +\nabla^2_{x_Nx_N} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_N]\right\rangle \\
&\quad + \left\langle -\hat{\gamma}^3, \nabla^2_{x_0x_N} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_N] \right. \\
&\quad \left. +\nabla^2_{x_Nx_0} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_0]\right\rangle .
\end{aligned}
$$

Portanto,

$$
\begin{aligned}
\sum_{l=1}^{4}\left\langle \psi_l, \nabla H^l(\hat{x},\hat{u})(\widetilde{w},\widetilde{v})\right\rangle &= \sum_{i=0}^{N-1}\langle -p_{i+1}, \widetilde{w}_{i+1} - \nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i - \nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{v}_i\rangle \\
&\quad + \sum_{i=0}^{N}\left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i)\widetilde{w}_i\right\rangle + \sum_{i=0}^{N-1}\left\langle -\gamma_i^2, \nabla\widetilde{h}_i(u_i)\widetilde{v}_i\right\rangle \\
&\quad + \left\langle -\gamma^3, \nabla_{x_0} k(\hat{x}_0,\hat{x}_N)\widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N\right\rangle , \qquad (5.47)
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{l=1}^{4}\left\langle \hat{\psi}_l, \nabla^2 H^l(\hat{x},\hat{u})(\widetilde{w},\widetilde{v})\right\rangle = \\
&\qquad \sum_{i=0}^{N-1}\left\langle -\hat{p}_{i+1}, -\nabla^2_{xx}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{w}_i] - \nabla^2_{xu}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{v}_i] \right. \\
&\qquad \left. -\nabla^2_{ux}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{w}_i] - \nabla^2_{uu}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{v}_i]\right\rangle \\
&\qquad + \sum_{i=0}^{N}\left\langle -\hat{\gamma}_i^1, \nabla^2 h_i(\hat{x}_i)[w_i,\widetilde{w}_i]\right\rangle + \sum_{i=0}^{N-1}\left\langle -\hat{\gamma}_i^2, \nabla^2\widetilde{h}_i(u_i)[v_i,\widetilde{v}_i]\right\rangle \\
&\qquad + \left\langle -\hat{\gamma}^3, \nabla^2_{x_0x_0} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_0] + \nabla^2_{x_Nx_N} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_N]\right\rangle \\
&\qquad + \left\langle -\hat{\gamma}^3, \nabla^2_{x_0x_N} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_N] + \nabla^2_{x_Nx_0} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_0]\right\rangle . \qquad (5.48)
\end{aligned}
$$

3. Analogamente ao item 2, considerando (5.43), podemos assegurar, dados $\mu_i^1, i \in I^1(\hat{x})$, $\hat{\mu}_i^1, i \in I^2(\hat{x})$, pertencem a $(\mathbb{R}^{s_1})^T \equiv \mathbb{R}^{s_1}$, $i = 0, \ldots, N$ e $\mu_i^2, i \in \widetilde{I}^1(\hat{u})$, $\hat{\mu}_i^2, i \in \widetilde{I}^2(\hat{u})$ pertencem a $(\mathbb{R}^{s_2})^T \equiv \mathbb{R}^{s_2}$,$i = 0, \ldots, N-1$, a existência de $\mu^1, \hat{\mu}^1 \in \mathbb{R}^{s_1(N+1)}$ e $\mu^2, \hat{\mu}^2 \in \mathbb{R}^{s_2(N)}$ tais que

$$\sum_{i \in I^1(\hat{x})} \left\langle \mu_i^1, \nabla G_i^1(\hat{x}, \hat{u})(\widetilde{w}, \widetilde{v}) \right\rangle = \sum_{i \in I^1(\hat{x})} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i)\widetilde{w}_i \right\rangle \tag{5.49}$$

$$\sum_{i \in \widetilde{I}^1(\hat{u})} \left\langle \mu_i^2, \nabla G_i^2(\hat{x}, \hat{u})(\widetilde{w}, \widetilde{v}) \right\rangle = \sum_{i \in \widetilde{I}^1(\hat{u})} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i)\widetilde{v}_i \right\rangle. \tag{5.50}$$

$$\sum_{i \in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla^2 G_i^1(\hat{x}, \hat{u})[(w,v),(\widetilde{w}, \widetilde{v})] \right\rangle = \sum_{i \in I^2(\hat{x})} \left\langle -\hat{\mu}_i^1, \nabla^2 g_i(\hat{x}_i)[w_i, \widetilde{w}_i] \right\rangle, \tag{5.51}$$

$$\sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla^2 G_i^2(\hat{u})[(w,v),(\widetilde{w}, \widetilde{v})] \right\rangle = \sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle -\hat{\mu}_i^2, \nabla^2 \widetilde{g}_i(\hat{u}_i)[v_i, \widetilde{v}_i] \right\rangle. \tag{5.52}$$

Substituindo (5.46)-(5.52) em (5.44) obtemos

$$\begin{aligned}
&\sum_{i=0}^{N-1} (\nabla_x \Phi_i(\hat{x}_i, \hat{u}_i)\widetilde{w}_i + \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i) \\
&+ \sum_{i=0}^{N-1} \langle -p_{i+1}, \widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i)\widetilde{w}_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i \rangle \\
&+ \sum_{i=0}^{N-1} \langle -\hat{p}_{i+1}, \nabla^2_{xx}\varphi_i(\hat{x}_i, \hat{u}_i)[w_i, \widetilde{w}_i] - \nabla^2_{xu}\varphi_i(\hat{x}_i, \hat{u}_i)[w_i, \widetilde{v}_i] \\
&\qquad - \nabla^2_{ux}\varphi_i(\hat{x}_i, \hat{u}_i)[v_i, \widetilde{w}_i] - \nabla^2_{uu}\varphi_i(\hat{x}_i, \hat{u}_i)[v_i, \widetilde{v}_i] \rangle \\
&+ \sum_{i \in I^1(\hat{x})} \left\langle -\mu_i^1, \nabla g_i(\hat{x}_i)\widetilde{w}_i \right\rangle + \sum_{i \in \widetilde{I}^1(\hat{u})} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i)\widetilde{v}_i \right\rangle \\
&+ \sum_{i \in I^2(\hat{x})} \left\langle -\hat{\mu}_i^1, \nabla^2 g_i(\hat{x}_i)[w_i, \widetilde{w}_i] \right\rangle + \sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle -\hat{\mu}_i^2, \nabla^2 \widetilde{g}_i(\hat{u}_i)[v_i, \widetilde{v}_i] \right\rangle \\
&+ \sum_{i=0}^{N} \left\langle -\gamma_i^1, \nabla h_i(\hat{x}_i)\widetilde{w}_i \right\rangle + \sum_{i=0}^{N} \left\langle -\hat{\gamma}_i^1, \nabla^2 h_i(\hat{x}_i)[w_i, \widetilde{w}_i] \right\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\gamma_i^2, \nabla \widetilde{h}_i(u_i)\widetilde{v}_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\hat{\gamma}_i^2, \nabla^2 \widetilde{h}_i(u_i)[v_i, \widetilde{v}_i] \right\rangle \\
&+ \left\langle -\gamma^3, \nabla_{x_0} k(\hat{x}_0, \hat{x}_N)\widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)\widetilde{w}_N \right\rangle \\
&+ \left\langle -\hat{\gamma}^3, \nabla^2_{x_0x_0} K(\hat{x}_0, \hat{x}_N)[w_0, \widetilde{w}_0] + \nabla^2_{x_Nx_N} K(\hat{x}_0, \hat{x}_N)[w_N, \widetilde{w}_N] \right\rangle \\
&+ \left\langle -\hat{\gamma}^3, \nabla^2_{x_0x_N} K(\hat{x}_0, \hat{x}_N)[w_0, \widetilde{w}_N] + \nabla^2_{x_Nx_0} K(\hat{x}_0, \hat{x}_N)[w_N, \widetilde{w}_0] \right\rangle = 0.
\end{aligned} \tag{5.53}$$

Sem perda de generalidade podemos assumir em (5.53)

$$\sum_{i \in I^1(\hat{x})} \left\langle \mu_i^1, \nabla g_i(\hat{x}_i)\widetilde{w}_i \right\rangle = \sum_{i=0}^{N} \left\langle \mu_i^1, \nabla g_i(\hat{x}_i)\widetilde{w}_i \right\rangle, \tag{5.54}$$

pois $\mu_i^1 = 0, i \in [\{1,\dots,N\} \setminus I^1(\hat{x})]$. O mesmo argumento é válido para as demais somas em $\widetilde{I}^1(\hat{u})$, $I^2(\hat{x})$ e $\widetilde{I}^2(\hat{u})$. Portanto podemos rescrever (5.53) como

$$\begin{aligned}
&\sum_{i=0}^{N-1} (\nabla_x \Phi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i + \nabla_u \Phi_i(\hat{x}_i,\hat{u}_i)\widetilde{v}_i) \\
&+ \sum_{i=0}^{N-1} \langle -p_{i+1}, \widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i - \nabla_u \varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{v}_i \rangle \\
&+ \sum_{i=0}^{N-1} \big\langle -\hat{p}_{i+1}, \nabla^2_{xx}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{w}_i] - \nabla^2_{xu}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{v}_i] \\
&\qquad - \nabla^2_{ux}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{w}_i] - \nabla^2_{uu}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{v}_i] \big\rangle \\
&+ \sum_{i=0}^{N} \big\langle -\mu_i^1, \nabla g_i(\hat{x}_i)\widetilde{w}_i \big\rangle + \sum_{i=0}^{N-1} \big\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i)\widetilde{v}_i \big\rangle \\
&+ \sum_{i=0}^{N} \big\langle -\hat{\mu}_i^1, \nabla^2 g_i(\hat{x}_i)[w_i,\widetilde{w}_i] \big\rangle + \sum_{i=0}^{N-1} \big\langle -\hat{\mu}_i^2, \nabla^2 \widetilde{g}_i(\hat{u}_i)[v_i,\widetilde{v}_i] \big\rangle \\
&+ \sum_{i=0}^{N} \big\langle -\gamma_i^1, \nabla h_i(\hat{x}_i)\widetilde{w}_i \big\rangle + \sum_{i=0}^{N} \big\langle -\hat{\gamma}_i^1, \nabla^2 h_i(\hat{x}_i)[w_i,\widetilde{w}_i] \big\rangle \\
&+ \sum_{i=0}^{N-1} \Big\langle -\gamma_i^2, \nabla \widetilde{h}_i(u_i)\widetilde{v}_i \Big\rangle + \sum_{i=0}^{N-1} \Big\langle -\hat{\gamma}_i^2, \nabla^2 \widetilde{h}_i(u_i)[v_i,\widetilde{v}_i] \Big\rangle \\
&+ \big\langle -\gamma^3, \nabla_{x_0} k(\hat{x}_0,\hat{x}_N)\widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N \big\rangle \\
&+ \big\langle -\hat{\gamma}^3, \nabla^2_{x_0x_0} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_0] + \nabla^2_{x_Nx_N} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_N] \big\rangle \\
&+ \big\langle -\hat{\gamma}^3, \nabla^2_{x_0x_N} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_N] + \nabla^2_{x_Nx_0} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_0] \big\rangle \;=\; 0. \qquad (5.55)
\end{aligned}$$

Como a identidade (5.55) é válida para todo $(\widetilde{w},\widetilde{v}) \in X \times U$, em particular para $(\widetilde{w},\widetilde{v}) = (\widetilde{w},0)$,

$$\begin{aligned}
&\sum_{i=0}^{N-1} \nabla_x \Phi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i + \sum_{i=0}^{N-1} \langle -p_{i+1}, \widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i \rangle \\
&+ \sum_{i=0}^{N-1} \big\langle -\hat{p}_{i+1}, \nabla^2_{xx}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{w}_i] - \nabla^2_{ux}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{w}_i] \big\rangle \\
&+ \sum_{i=0}^{N} \big\langle -\mu_i^1, \nabla g_i(\hat{x}_i)\widetilde{w}_i \big\rangle + \sum_{i=0}^{N} \big\langle -\hat{\mu}_i^1, \nabla^2 g_i(\hat{x}_i)[w_i,\widetilde{w}_i] \big\rangle \\
&+ \sum_{i=0}^{N} \big\langle -\gamma_i^1, \nabla h_i(\hat{x}_i)\widetilde{w}_i \big\rangle + \sum_{i=0}^{N} \big\langle -\hat{\gamma}_i^1, \nabla^2 h_i(\hat{x}_i)[w_i,\widetilde{w}_i] \big\rangle \\
&+ \big\langle -\gamma^3, \nabla_{x_0} k(\hat{x}_0,\hat{x}_N)\widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N \big\rangle \\
&+ \big\langle -\hat{\gamma}^3, \nabla^2_{x_0x_0} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_0] + \nabla^2_{x_Nx_N} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_N] \big\rangle \\
&+ \big\langle -\hat{\gamma}^3, \nabla^2_{x_0x_N} K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_N] + \nabla^2_{x_Nx_0} K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_0] \big\rangle \;=\; 0. \qquad (5.56)
\end{aligned}$$

Usando as propriedades dos operadores adjuntos e rearranjando os termos desta última expressão temos

$$\Big\langle \nabla_x \Phi_0(\hat{u}_0, \hat{u}_0) + p_0 - \nabla_x \varphi_0 \left(\hat{x}_0, \hat{u}_0\right)^T p_1 - \left[\nabla^2_{xx}\varphi_0(\hat{x}_0, \hat{u}_0) w_0\right]^T p_1 - \left[\nabla^2_{ux}\varphi_0(\hat{x}_0, \hat{u}_0) v_0\right]^T p_1$$
$$-\nabla g_0(\hat{x}_0)^T \mu_0^1 - \left[\nabla^2 g_0(\hat{x}_0) w_0\right]^T \hat{\mu}_0^1 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 - \left[\nabla^2 h_0(\hat{x}_0) w_0\right]^T \hat{\gamma}_0^1$$
$$-\nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_0} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 - \left[\nabla^2_{x_N x_0} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3, \widetilde{w}_0 \Big\rangle$$
$$+ \sum_{i=1}^{N-1} \Big\langle p_i - \nabla_x \Phi_i \left(\hat{x}_i, \hat{u}_i\right) - \nabla_x \varphi_i \left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} - \left[\nabla^2_{xx}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T p_{i+1}$$
$$- \left[\nabla^2_{ux}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T p_{i+1} - \nabla g_i(\hat{x}_i)^T \mu_i^1 - \left[\nabla^2 g_i(\hat{x}_i) w_i\right]^T \hat{\mu}_i^1 - \nabla h_i(\hat{x}_i)^T \gamma_i^1 - \left[\nabla^2 h_i(\hat{x}_i) w_i\right]^T \hat{\gamma}_i^1, \widetilde{w}_i \Big\rangle$$
$$+ \Big\langle p_N - \nabla g_N(\hat{x}_N)^T \mu_N^1 + - \left[\nabla^2 g_N(\hat{x}_N) w_N\right]^T \hat{\mu}_N^1 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \left[\nabla^2 h_N(\hat{x}_N) w_N\right]^T \hat{\gamma}_N^1$$
$$-\nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_N} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 - \left[\nabla^2_{x_N x_N} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3, \widetilde{w}_N \Big\rangle = 0, \forall \widetilde{w} \in X,$$

i.e., $\mathbf{A}\widetilde{w} = \langle A_0, \widetilde{w}_0\rangle + \displaystyle\sum_{i=1}^{N-1} \langle A_i, \widetilde{w}_i\rangle + \langle A_N, \widetilde{w}_N\rangle = 0, \ \forall \widetilde{w} \in X$, onde

$$\begin{aligned}
A_0 \quad &= \quad \nabla_x \Phi_0, (\hat{u}_0, \hat{u}_0) + p_0 - \nabla_x \varphi_0 \left(\hat{x}_0, \hat{u}_0\right)^T p_1 - \left[\nabla^2_{xx}\varphi_0(\hat{x}_0, \hat{u}_0) w_0\right]^T p_1 \\
&\quad - \left[\nabla^2_{ux}\varphi_0(\hat{x}_0, \hat{u}_0) v_0\right]^T p_1 - \nabla g_0(\hat{x}_0)^T \mu_0^1 - \left[\nabla^2 g_0(\hat{x}_0) w_0\right]^T \hat{\mu}_0^1 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 \\
&\quad - \left[\nabla^2 h_0(\hat{x}_0) w_0\right]^T \hat{\gamma}_0^1 - \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_0} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 \\
&\quad - \left[\nabla^2_{x_N x_0} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3 \\
A_i \quad &= \quad p_i - \nabla_x \Phi_i \left(\hat{x}_i, \hat{u}_i\right) - \nabla_x \varphi_i \left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} - \left[\nabla^2_{xx}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T p_{i+1} \\
&\quad - \left[\nabla^2_{ux}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T p_{i+1}, \ i = 1, \ldots, N-1. \\
&\quad -\nabla g_i(\hat{x}_i)^T \mu_i^1 - \left[\nabla^2 g_i(\hat{x}_i) w_i\right]^T \hat{\mu}_i^1 - \nabla h_i(\hat{x}_i)^T \gamma_i^1 - \left[\nabla^2 h_i(\hat{x}_i) w_i\right]^T \hat{\gamma}_i^1 \\
A_N \quad &= \quad p_N - \nabla g_N(\hat{x}_N)^T \mu_N^1 - \left[\nabla^2 g_N(\hat{x}_N) w_N\right]^T \hat{\mu}_N^1 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \left[\nabla^2 h_N(\hat{x}_N) w_N\right]^T \hat{\gamma}_N^1 \\
&\quad -\nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_N} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 - \left[\nabla^2_{x_N x_N} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3.
\end{aligned}$$

Então necessariamente temos $A_0 = 0, A_i = 0$ e $A_N = 0$, as quais correspondem às condições (5.35), (5.36) e (5.37) respectivamente.

Similarmente, considerando $(\widetilde{w}, \widetilde{v}) = (0, \widetilde{v})$ em (5.55) temos

$$\begin{aligned}
&\sum_{i=0}^{N-1} \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i + \sum_{i=0}^{N-1} \langle -p_{i+1}, -\nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\hat{p}_{i+1}, -\nabla^2_{xu}\varphi_i(\hat{x}_i, \hat{u}_i)[w_i, \widetilde{v}_i] - \nabla^2_{uu}\varphi_i(\hat{x}_i, \hat{u}_i)[v_i, \widetilde{v}_i]\right\rangle \\
&+ \sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i)\widetilde{v}_i\right\rangle + \sum_{i=0}^{N-1} \left\langle -\hat{\mu}_i^2, \nabla^2 \widetilde{g}_i(\hat{u}_i)[v_i, \widetilde{v}_i]\right\rangle \\
&+ \sum_{i=0}^{N-1} \Big\langle -\gamma_i^2, \nabla \widetilde{h}_i(u_i)\widetilde{v}_i \Big\rangle + \sum_{i=0}^{N-1} \Big\langle -\hat{\gamma}_i^2, \nabla^2 \widetilde{h}_i(u_i)[v_i, \widetilde{v}_i] \Big\rangle \quad = \quad 0.
\end{aligned}$$

Logo,

$$\sum_{i=0}^{N-1} \Big\langle \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i) - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)^T p_{i+1} - \left[\nabla^2_{xu}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T \hat{p}_{i+1} - \left[\nabla^2_{uu}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T \hat{p}_{i+1}$$
$$-\nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2 - \left[\nabla^2 \widetilde{g}_i(\hat{u}_i) v_i\right]^T \hat{\mu}_i^2 - \nabla \widetilde{h}_i(u_i)^T \gamma_i^2 - \left[\nabla^2 \widetilde{h}_i(u_i) v_i\right]^T \hat{\gamma}_i^2, \widetilde{v}_i \Big\rangle = 0,$$

i.e., $\mathbf{B}\widetilde{v} = \sum_{i=0}^{N-1} \langle B_i, \widetilde{v}_i \rangle = 0, \ \forall \widetilde{v} \in U$, onde

$$\begin{aligned} B_i &= \nabla_u \Phi_i(\hat{x}_i, \hat{u}_i) - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)^T p_{i+1} - \left[\nabla^2_{xu}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T \hat{p}_{i+1} - \left[\nabla^2_{uu}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T \hat{p}_{i+1} \\ &\quad -\nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2 - \left[\nabla^2 \widetilde{g}_i(\hat{u}_i) v_i\right]^T \hat{\mu}_i^2 - \nabla \widetilde{h}_i(u_i)^T \gamma_i^2 - \left[\nabla^2 \widetilde{h}_i(u_i) v_i\right]^T \hat{\gamma}_i^2, \ i = 0, \cdots, N-1. \end{aligned}$$

Então necessariamente temos $B_i = 0, \ \forall i = 0, \cdots, N-1$ as quais correspondem às condições (5.41).

Determinaremos agora explicitamente os termos da expressão (5.45):

Considerando agora (5.47), (5.48), (5.49) e (5.50) temos, para $\hat{\psi} = (\hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3, \hat{\psi}_4) \in (\mathbb{R}^{n(N+1)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k)^*$, $\hat{\mu}^1 \in (\mathbb{R}^{s_1(N+1)})^*$ e $\hat{\mu}^2 \in (\mathbb{R}^{s_2(N)})^*$, que existem $\hat{p} \in \mathbb{R}^{n(N+1)}$, $\hat{\gamma}^1 \in \mathbb{R}^{r_1(N+1)}$, $\hat{\gamma}^2 \in \mathbb{R}^{r_2(N+1)}$, $\hat{\gamma}_3 \in \mathbb{R}^k$, $\hat{\mu}^1 \in \mathbb{R}^{s_1(N+1)}$, $\hat{\mu}^2 \in \mathbb{R}^{s_2(N)}$, $\hat{\mu}^1 \geqq 0$, $\hat{\mu}^2 \geqq 0$, tais que

$$\begin{aligned} 0 &= \sum_{i \in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla g_i(\hat{x})(\widetilde{w}) \right\rangle + \sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla \widetilde{g}_i(\hat{u})(\widetilde{v}) \right\rangle + \sum_{k=1}^{4} \left\langle \hat{\psi}_k, \nabla H^k(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle = \\ &\quad \sum_{i \in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla g_i(\hat{x}) \widetilde{w}_i \right\rangle + \sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla \widetilde{g}_i(\hat{u}) \widetilde{v}_i \right\rangle \\ &\quad + \sum_{i=0}^{N-1} \langle -\hat{p}_{i+1}, \widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i) \widetilde{w}_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i) \widetilde{v}_i \rangle + \sum_{i=0}^{N} \left\langle -\hat{\gamma}_i^1, \nabla h_i(\hat{x}_i) \widetilde{w}_i \right\rangle \\ &\quad + \sum_{i=0}^{N-1} \left\langle -\hat{\gamma}_i^2, \nabla \widetilde{h}_i(\hat{u}_i) \widetilde{v}_i \right\rangle + \left\langle \hat{\gamma}^3, \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) \widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \widetilde{w}_N \right\rangle. \end{aligned} \tag{5.57}$$

E pelos mesmos argumentos discutidos em (5.54), podemos rescrever a equação acima como

$$\begin{aligned} &\sum_{i=0}^{N} \left\langle \hat{\mu}_i^1, \nabla g_i(\hat{x}) \widetilde{w}_i \right\rangle + \sum_{i=0}^{N-1} \left\langle \hat{\mu}_i^2, \nabla \widetilde{g}_i(\hat{u}) \widetilde{v}_i \right\rangle \\ &+ \sum_{i=0}^{N-1} \langle -\hat{p}_{i+1}, \widetilde{w}_{i+1} - \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i) \widetilde{w}_i - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i) \widetilde{v}_i \rangle + \sum_{i=0}^{N} \left\langle -\hat{\gamma}_i^1, \nabla h_i(\hat{x}_i) \widetilde{w}_i \right\rangle \\ &+ \sum_{i=0}^{N-1} \left\langle -\hat{\gamma}_i^2, \nabla \widetilde{h}_i(\hat{u}_i) \widetilde{v}_i \right\rangle + \left\langle \hat{\gamma}^3, \nabla_{x_0} K(\hat{x}_0, \hat{x}_N) \widetilde{w}_0 + \nabla_{x_N} K(\hat{x}_0, \hat{x}_N) \widetilde{w}_N \right\rangle = 0. \end{aligned} \tag{5.58}$$

Como (5.58) vale para todo $(\widetilde{w}, \widetilde{v}) \in X \times U$. Em particular para $(\widetilde{w}, \widetilde{v}) = (\widetilde{w}, 0)$,

$$\begin{aligned}
&\sum_{i=0}^{N}\left\langle \hat{\mu}_i^1, \nabla g_i(\hat{x})\widetilde{w}_i\right\rangle + \sum_{i=0}^{N-1}\langle -\hat{p}_{i+1}, \widetilde{w}_{i+1} - \nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i\rangle \\
+&\sum_{i=0}^{N}\left\langle -\hat{\gamma}_i^1, \nabla h_i(\hat{x}_i)\widetilde{w}_i\right\rangle + \left\langle \hat{\gamma}^3, \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)\widetilde{w}_0 + \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N\right\rangle \;=\; 0. \quad (5.59)
\end{aligned}$$

Usando as propriedades dos operadores adjuntos e rearranjando os termos da última expressão,

$$\begin{aligned}
&\left\langle \nabla g_0(\hat{x}_0)^T\hat{\mu}_0^1 - \hat{p}_0 + \nabla_x\varphi_0(\hat{x}_0,\hat{u}_0)^T\hat{p}_1 - \nabla h_0(\hat{x}_0)^T\hat{\gamma}_0^1 - \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)^T\hat{\gamma}^3, \widetilde{w}_0\right\rangle \\
&+\sum_{i=1}^{N-1}\left\langle \nabla g_i(\hat{x}_i)^T\hat{\mu}_i^1 - \hat{p}_i + \nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)^T\hat{p}_{i+1} - \nabla h_i(\hat{x}_i)^T\hat{\gamma}_i^1, \widetilde{w}_i\right\rangle \\
&+\left\langle \nabla g_N(\hat{x})^T\hat{\mu}_N^1 - \hat{p}_N - \nabla h_N(\hat{x}_N)^T\hat{\gamma}_N^1 - \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)^T\hat{\gamma}^3, \widetilde{w}_N\right\rangle \;=\; 0,
\end{aligned}$$

i.e., $\hat{\mathbf{A}}\widetilde{w} = \left\langle \hat{A}_0, \widetilde{w}_0\right\rangle + \sum_{i=1}^{N-1}\left\langle \hat{A}_i, \widetilde{w}_i\right\rangle + \left\langle \hat{A}_N, \widetilde{w}_N\right\rangle = 0,\ \forall \widetilde{w} \in X$, onde

$$\begin{aligned}
\hat{A}_0 &= \nabla g_0(\hat{x}_0)^T\hat{\mu}_0^1 - \hat{p}_0 + \nabla_x\varphi_0(\hat{x}_0,\hat{u}_0)^T\hat{p}_1 - \nabla h_0(\hat{x}_0)^T\hat{\gamma}_0^1 - \nabla_{x_0}K(\hat{x}_0,\hat{x}_N)^T\hat{\gamma}^3. \\
\hat{A}_i &= \nabla g_i(\hat{x}_i)^T\hat{\mu}_i^1 - \hat{p}_i + \nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)^T\hat{p}_{i+1} - \nabla h_i(\hat{x}_i)^T\hat{\gamma}_i^1,\ i = 1,\ldots,N-1. \\
\hat{A}_N &= \nabla g_N(\hat{x})^T\hat{\mu}_N^1 - \hat{p}_N - \nabla h_N(\hat{x}_N)^T\hat{\gamma}_N^1 - \nabla_{x_N}K(\hat{x}_0,\hat{x}_N)^T\hat{\gamma}^3.
\end{aligned}$$

Então necessariamente temos $\hat{A}_0 = 0, \hat{A}_i = 0$ e $\hat{A}_N = 0$, as quais correspondem às condições (5.38), (5.39) e (5.40) respectivamente.

Agora tomando $(\widetilde{w},\widetilde{v}) = (0,\widetilde{v})$ em (5.58), levando em consideração (5.45), obtemos

$$\sum_{i=0}^{N-1}\left\langle \hat{\mu}_i^2, \nabla\widetilde{g}_i(\hat{u})\widetilde{v}_i\right\rangle + \sum_{i=0}^{N-1}\langle -\hat{p}_{i+1}, -\nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{v}_i\rangle + \sum_{i=0}^{N-1}\left\langle -\hat{\gamma}_i^2, \nabla\widetilde{h}_i(\hat{u}_i)\widetilde{v}_i\right\rangle \;=\; 0.$$

Então,

$$\sum_{i=0}^{N-1}\left\langle \nabla\widetilde{g}_i(\hat{u})^T\hat{\mu}_i^2 + \nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)^T\hat{p}_{i+1} - \nabla\widetilde{h}_i(\hat{u}_i)^T\hat{\gamma}_i^2, \widetilde{v}_i\right\rangle \;=\; 0,$$

i.e., $\hat{\mathbf{B}}\widetilde{v} = \sum_{i=0}^{N-1}\langle \hat{B}_i, \widetilde{v}_i\rangle = 0,\ \forall \widetilde{v} \in U$, onde

$$\hat{B}_i = \nabla\widetilde{g}_i(\hat{u})^T\hat{\mu}_i^2 + \nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)^T\hat{p}_{i+1} - \nabla\widetilde{h}_i(\hat{u}_i)^T\hat{\gamma}_i^2.$$

Portanto, necessariamente temos $B_i = 0,\ \forall i = 0,\cdots,N-1$ as quais correspondem às condições (5.42).

E com isto encerramos a prova e o teorema esta demonstrado. □

Se no Teorema do Princípio do Máximo Estendido (Teorema 5.10) considerarmos como hipótese que o processo $(\hat{x},\hat{u})$ é 2-regular de $(PCD2)$ então, necessariamente λ_0 é não nulo e, neste caso, podemos assumir que $\lambda_0 = 1$.

Com base nisto daremos a seguinte definição:

Definição 5.11 *Diremos que um processo factível $(\hat{x}, \hat{u})$ é um processo **extremal 2-regular**, na direção de $(w, v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ do problema $(PCD2)$, se satisfaz as condições (5.35)-(5.42), do Teorema do Princípio do Máximo Estendido, com $\lambda_0 = 1$.*

E obtemos o seguinte resultado:

Corolário 5.12 *Se $(\hat{x}, \hat{u})$ é um processo ótimo de $(PCD2)$ satisfazendo a condição de 2-regularidade (Definição 5.8) então $(\hat{x}, \hat{u})$ é um extremal 2-regular.*

Capítulo 6

Problema de Controle Ótimo Discreto Multiobjetivo

Neste capítulo estabeleceremos condições de otimalidade necessárias de segunda ordem não degeneradas para problemas de controle ótimo discretos multiobjetivos, ainda quando o problema não for regular fraco ou forte, segundo Definições 4.3 e 4.4.

As condições necessárias que serão estudadas na Seção 6.2 generalizam os Teoremas do Princípio do Máximo com hipóteses debilitadas de regularidade. Isto será feito através do conceito de 2-regularidade introduzido por Arutyunov, Avakov e Izmailov em [7] para problemas de programação matemática.

6.1 2-regularidade em Programação Matemática

Considere agora o Problema de Otimização Vetorial:

$$\begin{array}{ll} \text{Minimizar} & f(x) = (f_1(x), \ldots, f_p(x)) \\ \text{sujeito a:} & \begin{array}{l} F(x) = 0 \\ g_i(x) \leq 0, \;\; i = 1, \ldots, m \end{array} \end{array} \tag{VOP}$$

onde $f : X \to \mathbb{R}^p$ é diferenciável e $g_i : X \to \mathbb{R}$, $i = 1, \ldots, m$, $F : X \to Y$ são duas vezes diferenciáveis em uma vizinhança $V(x_0)$, $x_0 \in X$ e tal que para cada $x \in X, F(x) = (F_1(x), \ldots, F_l(x))$ com $F_k : X \to Y_k$, X e $Y = Y_1 \times \ldots \times Y_l$ espaços de Banach (reais). Além disso assumimos que o subespaço $\mathbf{Im}\nabla F_k(\hat{x}, \hat{u})$ é fechado em Y_k para todo $k = 1, \ldots, l$.

Assim temos as seguintes definições relacionadas com o conceito de 2-regularidade.

Definição 6.1 *Dizemos que x_0 é um **ponto 2-regular fraco** do conjunto factível M de (VOP), se para a direção $w \in H_M(x_0)$, F satisfaz a seguinte condição:*
Seja $x_0 \in M_F$, se para todo $y \in Y, y = y_1 + y_2$, existem $q_1, q_2 \in X$ tais que

$$\begin{array}{rcl} \nabla F(x_0)q_2 & = & y_1 \\ \nabla F(x_0)q_1 + \nabla^2 F(x_0)[w, q_2] & = & y_2. \end{array} \tag{6.1}$$

Além disso, o sistema

$$\begin{array}{rcl} \nabla F(x_0)q_2 & = & 0 \\ \nabla F(x_0)q_1 + \nabla^2 F(x_0)[w, q_2] & = & 0 \\ \nabla g_i(x_0)q_2 & < & 0, \forall i \in I_w(x_0) \\ \nabla g_i(x_0)q_1 + \nabla^2 g_i(x_0)[w, q_2] & < & 0, \forall i \in \hat{I}_w(x_0) \end{array} \tag{6.2}$$

tem solução $q_1, q_2 \in X$, *com* $H_M(x_0)$ *definido em (5.17).*

Definição 6.2 *Dizemos que* x_0 *é um* ***ponto 2-regular forte*** *de (VOP), na direção* $w \in H_M(x_0)$ *se* F *satisfaz a condição (6.1) e dado um* $r \in \{1, \ldots, p\}$ *o sistema*

$$\begin{array}{lcl} \nabla f_j(x_0)q_2 & < & 0,\ j = 1, \ldots, p, j \neq r, \\ \nabla F(x_0)q_2 & = & 0 \\ \nabla F(x_0)q_1 + \nabla^2 F(x_0)[w, q_2] & = & 0 \\ \nabla g_i(x_0)q_2 & < & 0, \forall i \in I_w(x_0) \\ \nabla g_i(x_0)q_1 + \nabla^2 g_i(x_0)[w, q_2] & < & 0, \forall i \in \hat{I}_w(x_0) \end{array} \tag{6.3}$$

tem solução $q_1, q_2 \in X$, *com* $H_M(x_0)$ *definido em (5.17).*

E de acordo com as definições acima é fácil ver que $\boxed{\text{2-regularidade forte}} \Rightarrow \boxed{\text{2-regularidade fraca}}$.

Condições Necessárias em Programação Matemática

Obteremos aqui condições necessárias de otimalidade de segunda ordem não degeneradas para problemas vetoriais não regulares, através do formalismo de DM.

Observemos que se x_0 é uma solução Pareto ótima de (VOP) então, de acordo com o Lema 2.19,

$$C_D(f_r, x_0) \cap \left(\bigcap_{\substack{j=1 \\ j \neq r}}^{p} C_N(f_j, x_0) \right) \cap C_F(M_g, x_0) \cap C_T(M_F, x_0) = \emptyset, r = 1, \ldots, p, \tag{6.4}$$

onde $C_D(f_j, x_0), C_N(f_j, x_0)$ são os cones de direções de descida e de não crescimento do funcional f_j, $C_F(M_g, x_0), C_T(M_F, x_0)$ são os cones das direções factíveis ao conjunto de restrições de desigualdades M_g e das direções tangentes ao conjunto das restrições de igualdade M_F, todos calculados no ponto x_0.

Então se $\left(\bigcap_{\substack{j=1 \\ j \neq r}}^{p} C_N(f_j, x_0) \right) \cap C_F(M_g, x_0) \cap C_T(M_F, x_0) \neq \emptyset$, obtemos o seguinte resultado que corresponde a Proposição 5.5 do caso escalar.

Proposição 6.3 *Seja* x_0 *uma solução Pareto ótima do problema (VOP) então o seguinte é verdadeiro*

$$\nabla f_r(x_0)z \geq 0, \forall z \in \left(\left(\bigcap_{\substack{j=1 \\ j \neq r}}^{p} C_N(f_j, x_0) \right) \cap C_F(M_g, x_0) \cap C_T(M_F, x_0) \right), r = 1, \ldots, p.$$

No que segue daremos a definição de ponto Karush Kuhn Tucker Avakov Fraco vetorial (KKTA1) e ponto Karush Kuhn Tucker Avakov Forte vetorial (KKTA2) do problema (VOP).

Definição 6.4 *Dizemos que um ponto factível x_0 é um* ***ponto KKTA1*** *na direção $w \in X$ de (VOP) se existem $\nu = \nu(w) \in \mathbb{R}^p$, $\mu = \mu(w)$, $\hat{\mu} = \hat{\mu}(w) \in \mathbb{R}^m$, $\psi = \psi(w)$, $\hat{\psi} = \hat{\psi}(w) \in Y^*$, não todos identicamente nulos, tais que*

$$\begin{aligned}
&\nabla f(x_0)\nu + \sum_{i \in I_w(x_0)} \nabla g_i(x_0)^T \mu_i + \sum_{i \in \hat{I}_w(x_0)} \left[\nabla^2 g_i(x_0)w\right]^T \hat{\mu}_i + \nabla F(x_0)^T \psi + \left[\nabla^2 F(x_0)w\right]^T \hat{\psi} = 0 \\
&\sum_{i \in \hat{I}_w(x_0)} \nabla g_i(x_0)^T \hat{\mu}_i + \nabla F(x_0)^T \hat{\psi} = 0
\end{aligned} \tag{6.5}$$

onde $\nu = (\nu_1, \ldots, \nu_p) \neq 0$, $\nu_j \geq 0, j = 1 \ldots, p$, $\mu_i \geq 0, \hat{\mu}_i \geq 0, \psi$, e $\hat{\psi}$ são chamados os ***multiplicadores de Lagrange-Avakov****, os conjuntos de índices $I_w(x_0), \hat{I}_w(x_0)$ são definidos em (5.11) e (5.18).*

Definição 6.5 *Dizemos que um ponto factível x_0 é um* ***ponto KKTA2*** *na direção $w \in X$ de (VOP) se existem $\nu = \nu(w) \in \mathbb{R}^p$, $\mu = \mu(w)$, $\hat{\mu} = \hat{\mu}(w) \in \mathbb{R}^m$, $\psi = \psi(w)$, $\hat{\psi} = \hat{\psi}(w) \in Y^*$, não todos identicamente nulos, satisfazendo (6.5), onde todas as componentes de ν são estritamente positivas, isto é, $\nu = (\nu_1, \ldots, \nu_p) > 0$, $\nu_j > 0, j = 1 \ldots, p$, $\mu_i \geq 0, \hat{\mu}_i \geq 0, \psi$, e $\hat{\psi}$ são chamados os* ***multiplicadores de Lagrange-Avakov****.*

$$\boxed{\text{Karush Kuhn Tucker Forte (KKTA2)}} \Rightarrow \boxed{\text{Karush Kuhn Tucker Fraco (KKTA1)}}$$

Estamos prontos para apresentar os teoremas das condições necessárias de otimalidade.

Teorema 6.6 ***(Orellana-Vivanco [77])*** *Seja $x_0 \in M$ um ponto 2-regular forte na direção $w \in H_M(x_0)$. Se x_0 é uma solução Pareto do problema (VOP) e tal que $\nabla f(x_0)w = 0$ então x_0 é um ponto KKTA2 na direção de $w \in H_M(x_0)$.*

Teorema 6.7 ***(Orellana-Vivanco [77])*** *Seja $x_0 \in M$ um ponto 2-regular fraco na direção $w \in H_M(x_0)$. Se x_0 é uma solução Pareto do problema (VOP) e tal que $\nabla f(x_0)w \leq 0$ então x_0 é um ponto KKTA1 na direção de $w \in H_M(x_0)$.*

Observação 6.1 *Seja x_0 uma solução Pareto. Então, de acordo com os Teoremas 6.6 e 6.7, vale uma das afirmações abaixo:*

- *se x_0 é um ponto regular forte (fraco) as equações de Euler-Lagrange 6.5 são válidas com $\hat{\psi}, \hat{\mu}_i$ identicamente nulos.*
- *se x_0 não é um ponto regular forte (fraco), mas um ponto 2-regular forte (fraco) na direção $w \in H_M(x_0)$ então, há uma família de equações de Euler-Lagrange, 6.5, que são válidas com $\nu \neq 0, \nu_j > 0, j = 1, \ldots, p$ ($\nu \neq 0, \nu_j \geq 0, j = 1, \ldots, p$), $\hat{\psi} = \hat{\psi}(w)$ e $\hat{\mu}_i = \hat{\mu}_i(w)$, não todos identicamente nulos, dependendo do parâmetro $w \in H_M(x_0)$.*

6.2 Condições Necessárias

Nesta seção estabeleceremos condições necessárias de otimalidade de segunda ordem, para problemas de controle ótimo discretos multiobjetivos. As condições aqui discutidas são uma generalização das obtidas na Seção 5.2, para problemas fortemente e fracamente regulares.

Denotaremos por (PCDM2) o problema multiobjetivo (PCDM) definido no Capítulo 4 com a hipótese de que os operadores H e G do problema de programação matemática equivalente, (POM), sejam duas vezes diferenciáveis e os subespaços $\mathbf{Im}\nabla H^l(\hat{x},\hat{u})$ são fechados em $V^l, l = 1,\ldots,4$, onde $V^1 = \mathbb{R}^{n(N)}, V^2 = \mathbb{R}^{r_1(N+1)}, V^3 = \mathbb{R}^{r_2(N)}$ e $V^4 = \mathbb{R}^k$. Assim $\mathbf{Im}\nabla H(\hat{x},\hat{u})$ é um subespaço fechado de $V = V^1 \times V^2 \times V^3 \times V^4$ e denotaremos a este problema por (POM2).

Recordemos que no Capítulo 4 propomos os conceitos de processo regular fraco e forte. Para isto, nos baseamos na classificação de Bigi e Pappalardo [13] e esta se relaciona com o Princípio do Máximo Discreto. Uma análise similar será feita para problemas de controle ótimo discreto multiobjetivos 2-regulares. Mais precisamente:

Definição 6.8 *Dizemos que $(\hat{x},\hat{u})$ é um **processo 2-regular fraco** do problema (PCDM2), se para $d = (w,v) \in \mathbb{T}(Q,(\hat{x},\hat{u}))$, definido em (5.25), temos que ocorrem as condições (5.21)-(5.24) da Definição 5.8.*

Definição 6.9 *Dizemos que $(\hat{x},\hat{u})$ é um **processo 2-regular forte** do problema (PCDM2), se para $d = (w,v) \in \mathbb{T}(Q,(\hat{x},\hat{u}))$, valem as condições (5.21)-(5.24) e além disso, dado um $r \in \{1,\ldots,q\}$,*

$$\sum_{i=0}^{N-1} \nabla_x \Phi_i^j(\hat{x}_i,\hat{u}_i)w_i^2 + \nabla_u \Phi_i^j(\hat{x}_i,\hat{u}_i)v_i^2 < 0, \forall j = 1,\ldots,q, j \neq r. \tag{6.6}$$

O resultado seguinte relaciona a 2-regularidade forte (fraca) do problema de controle (PCDM2) com a 2-regularidade forte (fraca) do problema de programação matemática (POM2).

Lema 6.10 *O processo $(\hat{x},\hat{u})$ é um processo 2-regular forte (fraco) do problema (PCDM2) se, e somente se, $(\hat{x},\hat{u})$ é um ponto 2-regular forte (fraco) do problema (POM2).*

Demonstração. A demonstração para o caso de um processo 2-regular fraco segue de maneira análoga a prova do Lema 5.9. Já para o caso de um processo 2-regular forte resta simplesmente analisar o seguinte:
Seja $(\hat{x},\hat{u})$ um processo 2-regular forte de (PCDM2). Então temos por definição que

$$\sum_{i=0}^{N-1} \nabla_x \Phi_i^j(\hat{x}_i,\hat{u}_i)w_i^2 + \nabla_u \Phi_i^j(\hat{x}_i,\hat{u}_i)v_i^2 < 0, \forall j = 1,\ldots,q, j \neq r,$$

o que equivale a dizer que

$$\nabla \mathbb{J}^j(\hat{x},\hat{u})(w^2,v^2) < 0, \forall j = 1,\ldots,q, j \neq r$$

e desta segunda equação podemos concluir que $(\hat{x}, \hat{u})$ é um ponto 2-regular forte de (POM2), segundo a Definição 6.2. Temos provado assim o resultado.

□

Similarmente ao caso mono-objetivo, Teorema 5.10, proporemos as condições necessárias de otimalidade para o problema (PCDM2) por meio de um Teorema do Princípio do Máximo.

Teorema 6.11 ***(Princípio do Máximo Discreto Estendido Multiobjetivo: regular fraco)*** *Seja $(\hat{x}, \hat{u}) \in Q$. Se $(\hat{x}, \hat{u})$ é um processo ótimo Pareto de $(PCDM2)$ tal que $\nabla \mathbb{J}(\hat{x}, \hat{u})d \leqq 0$ para $d = (w, v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ existem: $\nu \in \mathbb{R}^q$, com $\nu_j \geq 0, j = 1, \ldots, q$, $p = p(d) \in \mathbb{R}^{n(N+1)}$, $(\nu, p) \neq 0$, $\hat{p} = \hat{p}(d), \hat{p}_i \in R^{n(N+1)}$, $\gamma^1 = \gamma^1(d)$ e $\hat{\gamma}^1 = \hat{\gamma}^1(d), \in \mathbb{R}^{r_1(N+1)}$, $\gamma^2 = \gamma^2(d)$ e $\hat{\gamma}^2 = \hat{\gamma}^2(d) \in \mathbb{R}^{r_2(N)}$, γ^3 e $\hat{\gamma}^3 \in \mathbb{R}^k$, $\mu^1 = \mu^1(d)$ e $\hat{\mu}^1 = \hat{\mu}^1(d) \in \mathbb{R}^{s_1}$, $\mu_i^1 \geq 0, \hat{\mu}_i^1 \geq 0, i \in [0, N]$, $\mu^2 = \mu^2(d)$ e $\hat{\mu}^2 = \hat{\mu}^2(d) \in \mathbb{R}^{s_2(N)}$, $\mu_i^2 \geq 0, \hat{\mu}_i^2 \geq 0, i \in [0, N-1]$, não todos simultaneamente nulos, tais que*

$$\begin{aligned} p_0 = \sum_{j=1}^{p} \nu_j \nabla_x \Phi_0^j(\hat{x}_0, \hat{u}_0) &- \nabla_x \varphi_0(\hat{x}_0, \hat{u}_0)^T p_1 - \left[\nabla_{xx}^2 \varphi_0(\hat{x}_0, \hat{u}_0) w_0\right]^T \hat{p}_1 \\ &- \left[\nabla_{ux}^2 \varphi_0(\hat{x}_0, \hat{u}_0) v_0\right]^T \hat{p}_1 - \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^{T\gamma^3} \\ &- \left[\nabla_{x_0,x_0}^2 K(\hat{x}_0, \hat{x}_N) w_0\right]^T \hat{\gamma}^3 - \left[\nabla_{x_N,x_0}^2 K(\hat{x}_0, \hat{x}_N) w_N\right]^T \hat{\gamma}^3 \\ &- \nabla h_0(\hat{x}_0)^T \gamma_0^1 - \left[\nabla^2 h_0(\hat{x}_0) w_0\right]^T \hat{\gamma}_0^1 - \nabla g_0(\hat{x}_0)^T \mu_0^1 - \left[\nabla^2 g_0(\hat{x}_0) w_0\right]^T \hat{\mu}_0^1, \end{aligned} \tag{6.7}$$

$$\begin{aligned} p_i = \sum_{j=1}^{p} \nu_j \nabla_x \Phi_i^j(\hat{x}_i, \hat{u}_i) &- \nabla_x \varphi_i(\hat{x}_i, \hat{u}_i)^T p_{i+1} - \left[\nabla_{xx}^2 \varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T \hat{p}_{i+1} \\ &- \left[\nabla_{ux}^2 \varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T \hat{p}_{i+1} - \nabla h_i(\hat{x}_i)^T \gamma_i^1 - \left[\nabla^2 h_i(\hat{x}_i) w_i\right]^T \hat{\gamma}_i^1 \\ &- \nabla g_i(\hat{x}_i)^T \mu_i^1 - \left[\nabla^2 g_i(\hat{x}_i) w_i\right]^T \hat{\mu}_i^1, i = 1, \ldots, N-1, \end{aligned} \tag{6.8}$$

$$\begin{aligned} p_N = -\nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 &- \left[\nabla_{x_N,x_N}^2 K(\hat{x}_0, \hat{x}_N) w_N\right]^T \hat{\gamma}^3 - \left[\nabla_{x_0,x_N}^2 K(\hat{x}_0, \hat{x}_N) w_0\right]^T \hat{\gamma}^3 \\ &- \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \left[\nabla^2 h_N(\hat{x}_N) w_N\right]^T \hat{\gamma}_N^1 - \nabla g_N(\hat{x}_N)^T \mu_N^1 \\ &- \left[\nabla^2 g_N(\hat{x}_N) w_N\right]^T \hat{\mu}_N^1, \end{aligned} \tag{6.9}$$

$$\hat{p}_0 = -\nabla_x \varphi_0(\hat{x}_0, \hat{u}_0)^T \hat{p}_1 - \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \hat{\gamma}^3 - \nabla h_0(\hat{x}_0)^T \hat{\gamma}_0^1 - \nabla g_0(\hat{x}_0)^T \hat{\mu}_0^1 \tag{6.10}$$

$$\hat{p}_i = -\nabla_x \varphi_i(\hat{x}_i, \hat{u}_i)^T \hat{p}_{i+1} - \nabla h_i(\hat{x}_i)^T \hat{\gamma}_i^1 - \nabla g_i(\hat{x}_i)^T \hat{\mu}_i^1, \quad i = 1, \ldots, N-1, \tag{6.11}$$

$$\hat{p}_N = -\nabla h_N(x_N)^T \hat{\gamma}_N^1 - \nabla g_N(x_N)^T \hat{\mu}_N^1 - \nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \hat{\gamma}^3. \tag{6.12}$$

Além disso, satisfaz a condição de máximo

$$\sum_{j=1}^{p} \nu_j \nabla \Phi_i^j(\hat{x}_i, \hat{u}_i) - \nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)^T p_{i+1} - \left[\nabla^2_{uu}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T \hat{p}_{i+1}$$
$$- \left[\nabla^2_{xu}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T \hat{p}_{i+1} - \nabla \widetilde{h}_i(\hat{u}_i)^T \gamma_i^2 - \left[\nabla^2 \widetilde{h}_i(\hat{u}_i) v_i\right]^T \hat{\gamma}_i^2$$
$$-\nabla \widetilde{g}_i(\hat{u}_i)^T \mu_i^2 - \left[\nabla^2 \widetilde{g}_i(\hat{u}_i) v_i\right]^T \hat{\mu}_i^2 = 0, \quad i = 0, \ldots, N-1, \tag{6.13}$$

e a condição,

$$-\nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)^T \hat{p}_{i+1} - \nabla \widetilde{h}_i(u_i)^T \hat{\gamma}_i^2 - \nabla \widetilde{g}_i(u_i)^T \hat{\mu}_i^2 = 0, \quad i = 1, \ldots, N-1. \tag{6.14}$$

Demonstração. A demonstração deste resultado é similar à prova do Teorema 5.10. Seja $(\hat{x}, \hat{u})$ um processo ótimo Pareto de $(PCDM2)$, iniciamos analisando os seguintes casos:

(a) Se existe $j_0 \in \{1, \ldots, q\}$ tal que $\nabla \mathbb{J}^{j_0}(\hat{x}, \hat{u}) = 0$ as equações (6.7)-(6.14) são satisfeitas trivialmente, basta considerar $\nu_{j_0} = 1$, $\nu_j = 0, j = 1, \ldots, q, j \neq j_0$, $p = \hat{p} = 0$, $\gamma^1 = \hat{\gamma}^1 = 0, \gamma^2 = \hat{\gamma}^2 = 0$, $\gamma^3 = \hat{\gamma}^3 = 0$, $\mu^1 = \hat{\mu}^1 = 0$, e $\mu^2 = \hat{\mu}^2 = 0$.

(b) Se $(\hat{x}, \hat{u})$ não é um processo regular fraco então, estamos nas hipóteses da parte (b) da prova do Teorema 5.10. E por conseguinte, isto implica que podemos ter ν identicamente nulo.

(c) Assim, resta considerar o caso em que processo ótimo $(\hat{x}, \hat{u})$ do problema $(PCDM2)$ é um processo 2-regular fraco segundo a Definição 6.8.

A prova se baseia no Teorema 6.7 e portanto vamos verificar que as hipóteses do mesmo se cumprem.

De acordo com as hipóteses impostas às funções que definem o problema $(PCDM2)$ (sendo (PCDM2) equivalente ao problema de programação matemática (POM2) é imediato que $\mathbb{J}$ (definido em (3.8)) é diferenciável e que os operadores H, G^1 e G^2 são duas vezes continuamente diferenciáveis com respeito a $(x, u) \in M$.

Sendo $(\hat{x}, \hat{u})$ um processo 2-regular fraco de (PCDM2), de acordo com o Lema 6.10, ele é um ponto 2-regular fraco de (POM2).

Como o ponto factível $(\hat{x}, \hat{u})$ é uma solução ótima e um ponto 2-regular fraco de $(POM2)$ tal que $\nabla \mathbb{J}(\hat{x}, \hat{u})(w, v) \leq 0$, para $d = (w, v) \in H_M(\hat{x}, \hat{u})$ segue, de acordo com o Teorema 6.7, que $(\hat{x}, \hat{u})$ é um ponto KKTA1 na direção $d = (w, v) \in H_M(\hat{x}, \hat{u})$. Então, de acordo com a Definição 6.4, existem $\nu = \nu(d) \in \mathbb{R}^q$,,$\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$ e $\hat{\psi} = (\hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3, \hat{\psi}_4) \in (\mathbb{R}^{n(N+1)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k)^*$, $\psi_k = \psi_k(d), \hat{\psi}_k = \hat{\psi}_k(d), k = 1, \ldots, 4$, $\mu^1, \hat{\mu}^1 \in (\mathbb{R}^{s_1(N+1)})^*$, $\mu^2, \hat{\mu}^2 \in (\mathbb{R}^{s_2(N)})^*$, $\mu^1 = \mu^1(d) \geqq 0, \mu^2 = \mu^2(h) \geqq 0$, $\hat{\mu}^1 = \hat{\mu}^1(d) \geqq 0$, e $\hat{\mu}^2 = \hat{\mu}^2(d) \geqq 0$, não todos identicamente nulos, tais que

$$\begin{aligned}
\nabla \mathbb{J}(\hat{x}, \hat{u})^T \nu + \sum_{i \in I^1(\hat{x})} \nabla G_i^1(\hat{x}, \hat{u})^T \mu_i^1 + \sum_{i \in I^1(\hat{u})} \nabla G_i^2(\hat{x}, \hat{u})^T \mu_i^2 + \sum_{i \in I^2(\hat{x})} \left[\nabla^2 G_i^1(\hat{x}, \hat{u}) d\right]^T \hat{\mu}_i^1 & \\
+ \sum_{i \in \widetilde{I}^2(\hat{u})} \left[\nabla^2 G_i^2(\hat{x}, \hat{u}) d\right]^T \hat{\mu}_i^2 + \sum_{l=1}^{4} \nabla H^l(\hat{x}, \hat{u})^T \psi_l + \sum_{l=1}^{4} \left[\nabla^2 H^l(\hat{x}, \hat{u}) d\right]^T \hat{\psi}_l &= 0 \\
\sum_{i \in I^2(\hat{x})} \nabla G_i^1(\hat{x}, \hat{u})^T \hat{\mu}_i^1 + \sum_{i \in \widetilde{I}^2(\hat{u})} \nabla G_i^2(\hat{x}, \hat{u})^T \hat{\mu}_i^2 + \sum_{l=1}^{4} \nabla H^l(\hat{x}, \hat{u})^T \hat{\psi}_l &= 0,
\end{aligned}$$

onde $\nu = (\nu_1, \ldots, \nu_q), \nu_j \geq 0, j = 1, \ldots, q$, μ_l e $\mu_i^1(z_i)$ e $\hat{\mu}_i^1(\bar{z}_i)$ são não negativos para todo $z_i \in \mathbf{Im}\nabla g_i(\hat{x})$ e $\bar{z}_i \in (\mathbb{R}^{s_1} \setminus \mathbf{Im}\nabla g_i(\hat{x}, \hat{u})), \forall i = 0, \ldots, N$, e analogamente para $\mu_i^2(t_i)$ e $\hat{\mu}_i^2(\bar{t}_i)$. Novamente observemos que valem as relações explicitadas em (5.43).

Considerando agora a definição de operador adjunto obtemos para todo $\widetilde{d} = (\widetilde{w}, \widetilde{v}) \in X \times U = \mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)}$,

$$\begin{aligned}\sum_{j=1}^{p} \left\langle \nu_j, \nabla \mathbb{J}^j(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle + \sum_{i \in I^1(\hat{x})} \left\langle \mu_i^1, \nabla G_i^1(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle + \sum_{i \in \widetilde{I}^1(\hat{u})} \left\langle \mu_i^2, \nabla G_i^2(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle \\ + \sum_{i \in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla^2 G_i^1(\hat{x}, \hat{u})(d)(\widetilde{d}) \right\rangle + \sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla^2 G_i^2(\hat{x}, \hat{u})(d)(\widetilde{d}) \right\rangle \\ + \sum_{l=1}^{4} \left\langle \psi_l, \nabla H^l(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle + \sum_{l=1}^{4} \left\langle \hat{\psi}_l, (\nabla^2 H^l(\hat{x}, \hat{u})(d)(\widetilde{d}) \right\rangle = 0, \end{aligned} \tag{6.15}$$

$$\sum_{i \in I^2(\hat{x})} \left\langle \hat{\mu}_i^1, \nabla G_i^1(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle + \sum_{i \in \widetilde{I}^2(\hat{u})} \left\langle \hat{\mu}_i^2, \nabla G_i^2(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle + \sum_{l=1}^{4} \left\langle \hat{\psi}_l, \nabla H^l(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle = 0, \tag{6.16}$$

onde $I^1(\hat{x})$, $\widetilde{I}^1(\hat{u})$, $I^2(\hat{x})$ e $\widetilde{I}^2(\hat{u})$ são dados por (3.25), (3.26), (5.26) e (5.27).

Vamos determinar explicitamente os termos de (6.15).

Como $(\mathbb{R}^q)^* = \mathbb{R}^q$ e $(\hat{x}, \hat{u})$ é um ponto 2-regular fraco, podemos garantir a existência de $\nu \in \mathbb{R}^q$, com $\nu_j \geq 0, j = 1 \ldots, q$ tal que

$$\sum_{j=1}^{q} \left\langle \nu_j, \nabla \mathbb{J}^j(\hat{x}, \hat{u})(\widetilde{d}) \right\rangle = \sum_{j=1}^{q} \nu_j \left(\sum_{i=0}^{N-1} \nabla_x \Phi_i^j(\hat{x}_i, \hat{u}_i)\widetilde{w}_i + \nabla_u \Phi_i^j(\hat{x}_i, \hat{u}_i)\widetilde{v}_i \right) \tag{6.17}$$

Já os demais termos desta equação foram determinados em (5.47)-(5.52).

Agora, substituindo (6.17), (5.47)-(5.52) em (6.15), segue de (5.54) que

$$\begin{aligned}
&\sum_{j=1}^{q}\nu_j\left(\sum_{i=0}^{N-1}\nabla_x\Phi_i^j(\hat{x}_i,\hat{u}_i)\widetilde{w}_i+\nabla_u\Phi_i^j(\hat{x}_i,\hat{u}_i)\widetilde{v}_i\right)\\
&+\sum_{i=0}^{N-1}\langle -p_{i+1},\widetilde{w}_{i+1}-\nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i-\nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{v}_i\rangle\\
&+\sum_{i=0}^{N-1}\big\langle -\hat{p}_{i+1},\nabla^2_{xx}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{w}_i]-\nabla^2_{xu}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{v}_i]\\
&\qquad-\nabla^2_{ux}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{w}_i]-\nabla^2_{uu}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{v}_i]\big\rangle\\
&+\sum_{i=0}^{N}\big\langle -\mu_i^1,\nabla g_i(\hat{x}_i)\widetilde{w}_i\big\rangle+\sum_{i=0}^{N-1}\big\langle -\mu_i^2,\nabla\widetilde{g}_i(\hat{u}_i)\widetilde{v}_i\big\rangle\\
&+\sum_{i=0}^{N}\big\langle -\hat{\mu}_i^1,\nabla^2 g_i(\hat{x}_i)[w_i,\widetilde{w}_i]\big\rangle+\sum_{i=0}^{N-1}\big\langle -\hat{\mu}_i^2,\nabla^2\widetilde{g}_i(\hat{u}_i)[v_i,\widetilde{v}_i]\big\rangle\\
&+\sum_{i=0}^{N}\big\langle -\gamma_i^1,\nabla h_i(\hat{x}_i)\widetilde{w}_i\big\rangle+\sum_{i=0}^{N}\big\langle -\hat{\gamma}_i^1,\nabla^2 h_i(\hat{x}_i)[w_i,\widetilde{w}_i]\big\rangle\\
&+\sum_{i=0}^{N-1}\Big\langle -\gamma_i^2,\nabla\widetilde{h}_i(u_i)\widetilde{v}_i\Big\rangle+\sum_{i=0}^{N-1}\Big\langle -\hat{\gamma}_i^2,\nabla^2\widetilde{h}_i(u_i)[v_i,\widetilde{v}_i]\Big\rangle\\
&+\big\langle -\gamma^3,\nabla_{x_0}k(\hat{x}_0,\hat{x}_N)\widetilde{w}_0+\nabla_{x_N}K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N\big\rangle\\
&+\big\langle -\hat{\gamma}^3,\nabla^2_{x_0x_0}K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_0]+\nabla^2_{x_Nx_N}K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_N]\big\rangle\\
&+\big\langle -\hat{\gamma}^3,\nabla^2_{x_0x_N}K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_N]+\nabla^2_{x_Nx_0}K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_0]\big\rangle = 0. \qquad (6.18)
\end{aligned}$$

Como a identidade (6.18) é válida para todo $(\widetilde{w},\widetilde{v})\in X\times U$, em particular para $(\widetilde{w},\widetilde{v})=(\widetilde{w},0)$ temos

$$\begin{aligned}
&\sum_{j=1}^{q}\nu_j\sum_{i=0}^{N-1}\nabla_x\Phi_i^j(\hat{x}_i,\hat{u}_i)\widetilde{w}_i+\sum_{i=0}^{N-1}\langle -p_{i+1},\widetilde{w}_{i+1}-\nabla_x\varphi_i(\hat{x}_i,\hat{u}_i)\widetilde{w}_i\rangle\\
&+\sum_{i=0}^{N-1}\big\langle -\hat{p}_{i+1},\nabla^2_{xx}\varphi_i(\hat{x}_i,\hat{u}_i)[w_i,\widetilde{w}_i]-\nabla^2_{ux}\varphi_i(\hat{x}_i,\hat{u}_i)[v_i,\widetilde{w}_i]\big\rangle\\
&+\sum_{i=0}^{N}\big\langle -\mu_i^1,\nabla g_i(\hat{x}_i)\widetilde{w}_i\big\rangle+\sum_{i=0}^{N}\big\langle -\hat{\mu}_i^1,\nabla^2 g_i(\hat{x}_i)[w_i,\widetilde{w}_i]\big\rangle\\
&+\sum_{i=0}^{N}\big\langle -\gamma_i^1,\nabla h_i(\hat{x}_i)\widetilde{w}_i\big\rangle+\sum_{i=0}^{N}\big\langle -\hat{\gamma}_i^1,\nabla^2 h_i(\hat{x}_i)[w_i,\widetilde{w}_i]\big\rangle\\
&+\big\langle -\gamma^3,\nabla_{x_0}k(\hat{x}_0,\hat{x}_N)\widetilde{w}_0+\nabla_{x_N}K(\hat{x}_0,\hat{x}_N)\widetilde{w}_N\big\rangle\\
&+\big\langle -\hat{\gamma}^3,\nabla^2_{x_0x_0}K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_0]+\nabla^2_{x_Nx_N}K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_N]\big\rangle\\
&+\big\langle -\hat{\gamma}^3,\nabla^2_{x_0x_N}K(\hat{x}_0,\hat{x}_N)[w_0,\widetilde{w}_N]+\nabla^2_{x_Nx_0}K(\hat{x}_0,\hat{x}_N)[w_N,\widetilde{w}_0]\big\rangle = 0. \qquad (6.19)
\end{aligned}$$

Usando as propriedades dos operadores adjuntos e rearranjando os termos desta última expressão,

$$\Big\langle \sum_{j=1}^{q} \nu_j \nabla_x \Phi_0^j\left(\hat{x}_0, \hat{u}_0\right) + p_0 - \nabla_x \varphi_0\left(\hat{x}_0, \hat{u}_0\right)^T p_1 - \left[\nabla^2_{xx}\varphi_0(\hat{x}_0, \hat{u}_0) w_0\right]^T p_1 - \left[\nabla^2_{ux}\varphi_0(\hat{x}_0, \hat{u}_0) v_0\right]^T p_1$$
$$-\nabla g_0(\hat{x}_0)^T \mu_0^1 - \left[\nabla^2 g_0(\hat{x}_0) w_0\right]^T \hat{\mu}_0^1 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 - \left[\nabla^2 h_0(\hat{x}_0) w_0\right]^T \hat{\gamma}_0^1$$
$$-\nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_0} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 - \left[\nabla^2_{x_N x_0} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3, \widetilde{w}_0 \Big\rangle$$
$$+\sum_{i=1}^{N-1} \Big\langle p_i + \sum_{j=1}^{q} \nu_j \nabla_x \Phi_i^j\left(\hat{x}_i, \hat{u}_i\right) - \nabla_x \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} - \left[\nabla^2_{xx}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T p_{i+1}$$
$$-\left[\nabla^2_{ux}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T p_{i+1} - \nabla g_i(\hat{x}_i)^T \mu_i^1 - \left[\nabla^2 g_i(\hat{x}_i) w_i\right]^T \hat{\mu}_i^1 - \nabla h_i(\hat{x}_i)^T \gamma_i^1 - \left[\nabla^2 h_i(\hat{x}_i) w_i\right]^T \hat{\gamma}_i^1, \widetilde{w}_i \Big\rangle$$
$$+\Big\langle p_N - \nabla g_N(\hat{x}_N)^T \mu_N^1 + - \left[\nabla^2 g_N(\hat{x}_N) w_N\right]^T \hat{\mu}_N^1 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \left[\nabla^2 h_N(\hat{x}_N) w_N\right]^T \hat{\gamma}_N^1$$
$$-\nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_N} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 - \left[\nabla^2_{x_N x_N} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3, \widetilde{w}_N \Big\rangle = 0, \forall \widetilde{w} \in X,$$

i.e., $\mathbf{A}\widetilde{w} = \langle A_0, \widetilde{w}_0\rangle + \sum_{i=1}^{N-1} \langle A_i, \widetilde{w}_i\rangle + \langle A_N, \widetilde{w}_N\rangle = 0,\ \forall \widetilde{w} \in X$, onde

$$\begin{aligned}
A_0 &= \sum_{j=1}^{q} \nu_j \nabla_x \Phi_0^j\left(\hat{x}_0, \hat{u}_0\right) \nu_j + p_0 - \nabla_x \varphi_0\left(\hat{x}_0, \hat{u}_0\right)^T p_1 - \left[\nabla^2_{xx}\varphi_0(\hat{x}_0, \hat{u}_0) w_0\right]^T p_1 \\
&\quad - \left[\nabla^2_{ux}\varphi_0(\hat{x}_0, \hat{u}_0) v_0\right]^T p_1 - \nabla g_0(\hat{x}_0)^T \mu_0^1 - \left[\nabla^2 g_0(\hat{x}_0) w_0\right]^T \hat{\mu}_0^1 - \nabla h_0(\hat{x}_0)^T \gamma_0^1 \\
&\quad - \left[\nabla^2 h_0(\hat{x}_0) w_0\right]^T \hat{\gamma}_0^1 - \nabla_{x_0} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_0} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 \\
&\quad - \left[\nabla^2_{x_N x_0} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3 \\
A_i &= p_i - \sum_{j=1}^{q} \nu_j \nabla_x \Phi_i^j\left(\hat{x}_i, \hat{u}_i\right) - \nabla_x \varphi_i\left(\hat{x}_i, \hat{u}_i\right)^T p_{i+1} - \left[\nabla^2_{xx}\varphi_i(\hat{x}_i, \hat{u}_i) w_i\right]^T p_{i+1} \\
&\quad - \left[\nabla^2_{ux}\varphi_i(\hat{x}_i, \hat{u}_i) v_i\right]^T p_{i+1} - \nabla g_i(\hat{x}_i)^T \mu_i^1 - \left[\nabla^2 g_i(\hat{x}_i) w_i\right]^T \hat{\mu}_i^1 \\
&\quad -\nabla h_i(\hat{x}_i)^T \gamma_i^1 - \left[\nabla^2 h_i(\hat{x}_i) w_i\right]^T \hat{\gamma}_i^1,\ i = 1, \ldots, N-1. \\
A_N &= p_N - \nabla g_N(\hat{x}_N)^T \mu_N^1 - \left[\nabla^2 g_N(\hat{x}_N) w_N\right]^T \hat{\mu}_N^1 - \nabla h_N(\hat{x}_N)^T \gamma_N^1 - \left[\nabla^2 h_N(\hat{x}_N) w_N\right]^T \hat{\gamma}_N^1 \\
&\quad -\nabla_{x_N} K(\hat{x}_0, \hat{x}_N)^T \gamma^3 - \left[\nabla^2_{x_0 x_N} K(\hat{x}_0, \hat{x}_N) w_0\right]^T \gamma^3 - \left[\nabla^2_{x_N x_N} K(\hat{x}_0, \hat{x}_N) w_N\right]^T \gamma^3.
\end{aligned}$$

Então necessariamente temos $A_0 = 0, A_i = 0$ e $A_N = 0$, as quais correspondem às condições (6.7), (6.8) e (6.9) respectivamente.

Agora, se considerarmos $(\widetilde{w}, \widetilde{v}) = (0, \widetilde{v})$ em (6.18), temos

$$\sum_{j=1}^{q} \nu_j \sum_{i=0}^{N-1} \nabla_u \Phi_i^j(\hat{x}_i, \hat{u}_i)\widetilde{v}_i + \sum_{i=0}^{N-1} \left\langle -p_{i+1}, -\nabla_u \varphi_i(\hat{x}_i, \hat{u}_i)\widetilde{v}_i \right\rangle$$
$$+\sum_{i=0}^{N-1} \left\langle -\hat{p}_{i+1}, -\nabla^2_{xu}\varphi_i(\hat{x}_i, \hat{u}_i)[w_i, \widetilde{v}_i] - \nabla^2_{uu}\varphi_i(\hat{x}_i, \hat{u}_i)[v_i, \widetilde{v}_i] \right\rangle$$
$$+\sum_{i=0}^{N-1} \left\langle -\mu_i^2, \nabla \widetilde{g}_i(\hat{u}_i)\widetilde{v}_i \right\rangle + \sum_{i=0}^{N-1} \left\langle -\hat{\mu}_i^2, \nabla^2 \widetilde{g}_i(\hat{u}_i)[v_i, \widetilde{v}_i] \right\rangle$$

$$+\sum_{i=0}^{N-1}\left\langle -\gamma_i^2, \nabla\widetilde{h}_i(u_i)\widetilde{v}_i\right\rangle + \sum_{i=0}^{N-1}\left\langle -\hat{\gamma}_i^2, \nabla^2\widetilde{h}_i(u_i)[v_i,\widetilde{v}_i]\right\rangle \ = \ 0.$$

Logo,

$$\sum_{i=0}^{N-1}\left\langle \sum_{j=1}^{q}\nu_j\nabla_u\Phi_i^j(\hat{x}_i,\hat{u}_i) - \nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)^T p_{i+1} - \left[\nabla^2_{xu}\varphi_i(\hat{x}_i,\hat{u}_i)w_i\right]^T\hat{p}_{i+1} - \left[\nabla^2_{uu}\varphi_i(\hat{x}_i,\hat{u}_i)v_i\right]^T\hat{p}_{i+1}\right.$$
$$\left. -\nabla\widetilde{g}_i(\hat{u}_i)^T\mu_i^2 - \left[\nabla^2\widetilde{g}_i(\hat{u}_i)v_i\right]^T\hat{\mu}_i^2 - \nabla\widetilde{h}_i(u_i)^T\gamma_i^2 - \left[\nabla^2\widetilde{h}_i(u_i)v_i\right]^T\hat{\gamma}_i^2, \widetilde{v}_i\right\rangle = 0.$$

i.e., $\mathbf{B}\widetilde{v} = \sum_{i=0}^{N-1}\langle B_i,\widetilde{v}_i\rangle = 0,\ \forall\widetilde{v}\in U$, onde

$$B_i = \ \sum_{j=1}^{q}\nabla_u\Phi_i^{j*}(\hat{x}_i,\hat{u}_i) - \nabla_u\varphi_i(\hat{x}_i,\hat{u}_i)^T p_{i+1} - \left[\nabla^2_{xu}\varphi_i(\hat{x}_i,\hat{u}_i)w_i\right]^T\hat{p}_{i+1} - \left[\nabla^2_{uu}\varphi_i(\hat{x}_i,\hat{u}_i)v_i\right]^T\hat{p}_{i+1}$$
$$-\nabla\widetilde{g}_i(\hat{u}_i)^T\mu_i^2 - [\nabla^2\widetilde{g}_i(\hat{u}_i)v_i]^T\hat{\mu}_i^2 - \nabla\widetilde{h}_i(u_i)^T\gamma_i^2 - \left[\nabla^2\widetilde{h}_i(u_i)v_i\right]^T\hat{\gamma}_i^2,\ i=0,\cdots,N-1.$$

Então necessariamente temos $B_i = 0,\ \forall i = 0,\cdots,N-1$ as quais correspondem às condições (6.13).

Observe que o termos da expressão (6.16) são os mesmos da expressão (5.45) e portanto são obtidos de forma completamente análoga à realizada na prova do Teorema 5.10, de modo que obtemos (6.10)- (6.12) e (6.14).

E com isto encerramos a demostração.

$\square$

Através de uma pequena modificação nas hipóteses do teorema anterior podemos garantir que o multiplicador associado a função objetivo $\mathbb{J}$ seja estritamente positivo, isto é, $\nu_j > 0,\ j = 1,\ldots,q$. Mais precisamente temos:

Teorema 6.12 ***(Princípio do Máximo Discreto Estendido Multiobjetivo: regular forte)*** *Seja $(\hat{x},\hat{u})\in Q$. Se $(\hat{x},\hat{u})$ é um processo ótimo Pareto de $(PCDM2)$ tal que $\nabla\mathbb{J}(\hat{x},\hat{u})d = 0$ para $d=(w,v)\in\mathbb{T}(Q,(\hat{x},\hat{u}))$ existem: $\nu\in\mathbb{R}^q$, com $\nu_j>0, j=1,\ldots,q$, $p=p(d)\in\mathbb{R}^{n(N+1)}$, $(\nu,p)\neq 0$, $\hat{p}=\hat{p}(d), \hat{p}_i\in R^{n(N+1)}$, $\gamma^1=\gamma^1(d)$ e $\hat{\gamma}^1=\hat{\gamma}^1(d), \in \mathbb{R}^{r_1(N+1)}$, $\gamma^2=\gamma^2(d)$ e $\hat{\gamma}^2=\hat{\gamma}^2(d)\in\mathbb{R}^{r_2(N)}$, γ^3 e $\hat{\gamma}^3\in\mathbb{R}^k$, $\mu^1=\mu^1(d)$ e $\hat{\mu}^1=\hat{\mu}^1(d)\in\mathbb{R}^{s_1}$, $\mu_i^1\geq 0, \hat{\mu}_i^1\geq 0, i\in[0,N]$, $\mu^2=\mu^2(d)$ e $\hat{\mu}^2=\hat{\mu}^2(d)\in\mathbb{R}^{s_2(N)}$, $\mu_i^2\geq 0,\hat{\mu}_i^2\geq 0, i\in[0,N-1]$, não todos simultaneamente nulos, tais que (6.7)-(6.14) ocorrem.*

A prova deste teorema é totalmente análogo à prova do Teorema 6.11. A diferença consiste no fato de que, sob a hipótese de 2-regularidade forte, nos baseamos pelo Teorema 6.6 (enquanto a demonstração do Teorema 6.11 foi baseada no Teorema 6.7).

Observação 6.2 *Seja $(\hat{x},\hat{u})$ uma solução Pareto. Então, de acordo com os Teoremas 6.11 e 6.12, uma das afirmações abaixo ocorre:*

- *se $(\hat{x},\hat{u})$ é um ponto regular forte (fraco) as equações de Euler-Lagrange (6.7) - (6.14) são válidas com $\hat{\gamma}^1,\hat{\gamma}^2,\hat{\gamma}^3,\hat{\mu}^1,\hat{\mu}^2$ identicamente nulos.*

- *se* $(\hat{x}, \hat{u})$ *não é um ponto regular forte (fraco), mas um ponto 2-regular forte (fraco) na direção* $d = (w, v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ *então, há uma família de equações de Euler-Lagrange, (6.7) - (6.14), que são válidas com* $\nu \neq 0, \nu_j > 0, j = 1, \ldots, q$ $(\nu \neq 0, \nu_j \geq 0, j = 1, \ldots, q)$, $\hat{\gamma}^1 = \hat{\gamma}^1(d)$, $\hat{\gamma}^2 = \hat{\gamma}^2(d)$, $\hat{\gamma}^3 = \hat{\gamma}_1^3(d)$, $\hat{\mu}^1 = \hat{\mu}^1(d)$, $\hat{\mu}^2 = \hat{\mu}^2(d)$, *não todos identicamente nulos, dependendo do parâmetro* $d \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$.

E com base no resultados acima temos as seguintes definições.

Definição 6.13 *Diremos que um processo factível* $(\hat{x}, \hat{u})$ *é um processo* ***extremal 2-regular fraco****, na direção de* $d = (w, v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ *do problema* $(PCDM2)$*, se satisfaz as condições (6.7) - (6.14), do Teorema do Princípio do Máximo Estendido: regular fraco, com* $\nu \geq 0$.

Definição 6.14 *Diremos que um processo factível* $(\hat{x}, \hat{u})$ *é um processo* ***extremal 2-regular forte****, na direção de* $d = (w, v) \in \mathbb{T}(Q, (\hat{x}, \hat{u}))$ *do problema* $(PCDM2)$*, se satisfaz as condições (6.7) - (6.14), do Teorema do Princípio do Máximo Estendido: regular forte, com* $\nu > 0$.

São imediatos os seguintes resultados:

Corolário 6.15 *Se* $(\hat{x}, \hat{u})$ *é um processo Pareto ótimo de* $(PCDM2)$ *satisfazendo a condição de 2-regularidade fraca (Definição 6.8) então* $(\hat{x}, \hat{u})$ *é um extremal 2-regular fraco.*

Corolário 6.16 *Se* $(\hat{x}, \hat{u})$ *é um processo Pareto ótimo de* $(PCDM2)$ *satisfazendo a condição de 2-regularidade forte (Definição 6.9) então* $(\hat{x}, \hat{u})$ *é um extremal 2-regular forte.*

Parte III

Tópicos Correlatos

Capítulo 7

Análise de Estabilidade e Sensibilidade

A análise de estabilidade e sensibilidade em problemas de otimização tem sido estudada por vários pesquisadores, [1, 17, 37, 43, 55, 62, 63, 83], visto que sua relevância se dá devido ao fato de, em geral, não se conhecem os valores exatos dos parâmetros de um problema ou esses valores estão sujeitos a perturbações. Intuitivamente , o que se deseja é caracterizar os problemas para os quais se possa garantir que pequenas perturbações nos parâmetros acarretem em pequenas alterações no valor ótimo do problema.

A ideia central é, sabendo que o problema de controle ótimo discreto pode ser reformulado como um problema de programação matemática com uma estrutura específica, transpor os resultados de estabilidade e sensibilidade da programação matemática para o controle ótimo discreto.

7.1 Resultados Preliminares

Especificamente, vamos caracterizar uma classe de problemas de controle ótimo discreto, dependendo de um parâmetro, que tem soluções locais localmente isoladas[1] que são localmente Lipschitz contínuas e direcionalmente diferenciáveis em função do parâmetro.

Se considerarmos um problema de programação matemática somente com restrições de igualdade ao qual denotaremos (PH)

$$\begin{array}{l} \text{Minimizar } f(x,\sigma) \\ \text{sujeito a: } x \in D(\sigma) \cap B \end{array} \tag{PH}$$

com $D(\sigma) := \{x \in X : H(x,\sigma) = 0\}$, onde X, Σ e Y são espaços de Banach reais, $\sigma \in \Sigma$ é um parâmetro, $f : X \times \Sigma \to \mathbb{R}$ é uma função suave, $H : X \times \Sigma \to Y$ é uma aplicação suave e $B = \{x \in X : \|x - \hat{x}\| \leq r\}$ para algum $r > 0$ tal que $\hat{x}$ é solução do problema não perturbado, isto é, para $\hat{\sigma} \in \Sigma$ fixado restrito a B.

Definimos ainda, $S(\sigma)$ como o conjunto solução do problema (PH) e o valor ótimo

[1] Se $\bar{x}$ é um minimizador local do problema existe uma vizinhança B de $\bar{x}$ tal que em B, $\bar{x}$ é o único minimizador do problema.

do problema (local) de (PH) como:

$$\omega(\sigma) := \inf_{x \in D(\sigma) \cap B} f(x, \sigma), \ \sigma \in \Sigma.$$

Seja $\hat{x}$ um ponto regular de (PH), segundo a Definição 3.2 isto significa que o cone tangente ao conjunto com interior vazio $D(\sigma)$ no ponto $\hat{x}$ pode ser caracterizado por aproximações lineares, ou seja,

$$C_T(D(\sigma), \hat{x}) = \{w \in X : \nabla_x H(\hat{x}, \sigma) w = 0\}. \tag{7.1}$$

Neste caso, essa condição corresponde a condição de qualificação LICQ, dada por exemplo em [53]. Com isso segue o seguinte resultado de sensibilidade:

Teorema 7.1 (Levitin [59]) *Sob as hipóteses assumidas em (PH) segue que ω é contínuo em $\hat{\sigma}$, $S(\sigma) \neq \emptyset, \forall \sigma \in \Sigma$ suficientemente próximo de $\hat{\sigma}$ e,*

$$\sup_{x \in S(\sigma)} dist(x, S(\hat{\sigma})) \to 0, \ \textit{quando } \sigma \to \hat{\sigma}$$

onde $dist(x, S(\hat{\sigma})) = \inf_{z \in S(\sigma)} \|x - z\|$, $x \in X$ e $S(\sigma) \subset X$. Em particular, se $\hat{x}$ é um minimizador local estrito do problema não perturbado e B é escolhido de tal forma que $\hat{x}$ é a única solução deste mesmo problema, então

$$\sup_{x \in S(\sigma)} \|x - \hat{x}\| \to 0, \ \textit{quando } \sigma \to \hat{\sigma}.$$

Além disso, temos que:

Proposição 7.2 (Levitin [59]) *Para $\sigma \in \Sigma$, o seguinte limite superior linear é válido:*

$$\omega(\sigma) \leq \omega(\hat{\sigma}) + O(\|\sigma - \hat{\sigma}\|).$$

Teorema 7.3 (Bonnans e Shapiro [16]) *Seja $\mathcal{L}(x, \gamma, \sigma) = f(x, \sigma) + \langle \gamma, H(x, \sigma) \rangle$, $\sigma \in \Sigma, x \in X$ e $\gamma \in Y^*$, a Lagrangeana do problema (PH). Se $\hat{x}$ é a única solução do problema não perturbado, então temos que*

$$\omega'(\hat{\sigma}, d) = \langle \nabla_\sigma \mathcal{L}(\hat{x}, \hat{\gamma}, \hat{\sigma}), d \rangle$$

é a derivada direcional de ω em $\hat{\sigma}$ com respeito a qualquer direção $d \in \Sigma$. Em particular, ω é Gâteaux diferenciável em $\hat{\sigma}$.

Entretanto, nos casos mais gerais, hipóteses adicionais são necessárias para garantir a existência do limitante inferior de ω e dos limitantes nas soluções dos problemas perturbados. A hipótese mais popular e natural neste contexto é a *condição de crescimento quadrático* (CCQ) a qual consiste na existência de $C > 0$ e de uma vizinhança V de $\hat{x}$ tal que

$$f(x, \hat{\sigma}) \geq f(\hat{x}, \hat{\sigma}) + C\|x - \hat{x}\|^2, \ \forall x \in D(\hat{\sigma}) \cap V,$$

ou alguma condição que garanta (CCQ). No caso de um problema com restrições de igualdade e desigualdade:

$$\begin{array}{l} \text{Minimizar } f(x, \sigma) \\ \text{sujeito a: } x \in D(\sigma) \cap B. \end{array} \tag{PHg}$$

com $D(\sigma) = \{x \in X : H(x,\sigma) = 0, g(x,\sigma) \leqq 0\}$ onde $g : X \times \Sigma \to \mathbb{R}^m$ é uma aplicação suave e as demais funções são como definidas no problema (PH).

Neste caso a Lagrangeana assume a forma: $\mathcal{L}(x,\gamma,\mu,\sigma) = f(x,\sigma) + \langle\gamma, H(x,\sigma)\rangle + \langle\mu, g(x,\sigma)\rangle, \sigma \in \Sigma, x \in X, \gamma \in Y^*$, e $\mu \in \mathbb{R}^m$.

Assim, temos a seguinte definição:

Definição 7.4 *Sejam $f(\cdot,\sigma), H(\cdot,\sigma)$ e $g(\cdot,\sigma)$ de classe C^2 numa vizinhança de $\hat{x}(\sigma)$. Dizemos que a condição suficiente de segunda ordem forte (FSC) é satisfeita em $\hat{x}(\sigma)$, se*

$$\left\langle d, \nabla^2_{xx}\mathcal{L}(\hat{x}(\sigma),\gamma(\sigma),\mu(\sigma),\sigma)d\right\rangle > 0$$

para todo $d \neq 0$ tal que,

$$d \in \{d \in X : \nabla_x H(\hat{x}(\sigma),\sigma)d = 0,\ \nabla_x g^s(\hat{x}(\sigma),\sigma)d = 0, s \in I^+(\sigma)\} \tag{7.2}$$

onde,

$$\begin{array}{l} I^+(\sigma) := \{s \in I(\sigma) : \mu^s(\sigma) > 0\} \\ I(\sigma) := \{s \in \{1,\ldots,m\} : g^s(\hat{x}(\sigma),\sigma) = 0\} \end{array} \tag{7.3}$$

Observe que esta definição implica (CCQ).

Com isso, temos que se o problema (PHg) satisfaz as condições de independência linear (LICQ) e suficiência de segunda ordem forte (FSC), nos trabalhos [1, 55, 83] os autores provam que: se $\hat{x}$ é minimizador local do problema não perturbado, existe uma vizinhança V de $\hat{x}$ tal que em V, $\hat{x}$ é o único minimizador e que as soluções locais localmente isoladas são localmente Lipschitz contínuas e direcionalmente diferenciáveis. Se além disso vale a condição de folgas complementares, ou seja, $I^+(\sigma) = I(\sigma)$ então a solução torna-se diferenciável em função do parâmetro, veja [43]. Mais precisamente, temos o seguinte resultado clássico de estabilidade e sensibilidade.

Teorema 7.5 (Malanowski [61]) *Assuma que as funções $f(\cdot,\sigma)$, $H(\cdot,\sigma)$ e $g(\cdot,\sigma)$ do problema (PHg) são C^2 numa vizinhança de $\hat{x}(\hat{\sigma})$. Além disso, (LICQ) e (FSC) são válidas em $\hat{x}(\sigma)$. Então existem as vizinhanças V_1 de $\hat{\sigma}$ em Σ e V_2 de $\hat{x}(\sigma)$ em X, tais que para cada $\sigma \in V_1$, $\hat{x}(\sigma)$ é o único minimizador de (PHg) em V_2 e $(\gamma(\sigma),\mu(\sigma))$ são os únicos multiplicadores de Lagrange associados. As funções $\hat{x}(\cdot),\gamma(\cdot),\mu(\cdot)$ são Lipschitz continuas em V_1 e direcionalmente diferenciáveis em $\hat{\sigma}$. Além disso em cada direção $\widetilde{d}$, $\hat{x}(\hat{\sigma}) = \delta^+\hat{x}(\hat{\sigma},\widetilde{d})$ é a única solução do seguinte problema de programação quadrático, cujos únicos multiplicadores associados são $\gamma(\hat{\sigma}) = \delta^+\gamma(\hat{\sigma},\widetilde{d})$ e $\mu(\hat{\sigma}) = \delta^+\mu(\hat{\sigma},\widetilde{d})$,*

$$\begin{array}{ll} \textit{Minimizar} & \frac{1}{2}\langle d, \nabla^2_{xx}\mathcal{L}(\hat{\sigma})d\rangle + \left\langle d, \nabla^2_{\alpha\sigma}\mathcal{L}(\hat{\sigma})\widetilde{d}\right\rangle \\ \multicolumn{2}{l}{\textit{sujeito a: } \nabla_x H(\hat{x}(\hat{\sigma}),\hat{\sigma}) + \nabla_\sigma H(\hat{x}(\hat{\sigma}),\hat{\sigma})\widetilde{d} = 0} \\ & \nabla_x G^s(\hat{x}(\hat{\sigma}),\hat{\sigma})d + \nabla_\sigma G^s(\hat{x}(\hat{\sigma}),\hat{\sigma})\widetilde{d} \left\{ \begin{array}{l} = 0, s \in I^+(\hat{\sigma}) \\ \leqq 0, s \in I(\hat{\sigma})\backslash I^+(\hat{\sigma}), \end{array} \right. \end{array} \tag{$PQ_{\hat{\sigma},d_0}$}$$

onde $\mathcal{L}(\hat{\sigma}) := \mathcal{L}(\hat{x}(\hat{\sigma}),\gamma(\hat{\sigma}),\mu(\hat{\sigma}),\sigma)$.

Para prova e maiores detalhes veja também [55] e [83].

7.2 Problema (PCD_σ)

Sejam Σ um conjunto aberto em um espaço de Banach (real), o qual representa o conjunto de parâmetros factíveis e $\hat{\sigma} \in \Sigma$ um parâmetro fixado. Para cada σ pertencente a uma vizinhança $V \subset \Sigma$ de $\hat{\sigma}$, consideramos o seguinte problema de controle ótimo discreto:

$$\begin{array}{l} \text{Minimizar } \mathbb{J}(x_i, u_i, \sigma) = \displaystyle\sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i, \sigma\right) \\ \text{sujeito a: } x_{i+1} = \varphi_i\left(x_i, u_i, \sigma\right), \ i = 0, \ldots, N-1, \\ h_i(x_i, \sigma) = 0, \ g_i(x_i, \sigma) \leq 0, i = 0, \ldots, N, \\ \widetilde{h}_i(u_i, \sigma) = 0, \ \widetilde{g}_i(u_i, \sigma) \leq 0, i = 0, \ldots, N-1, \\ K(x_0, x_N, \sigma) = 0, \end{array} \qquad (PCD_\sigma)$$

onde $\Phi_i : \mathbb{R}^n \times \mathbb{R}^m \times V \to \mathbb{R}$, $\varphi_i : \mathbb{R}^n \times \mathbb{R}^m \times V \to \mathbb{R}^n$, $h_i : \mathbb{R}^n \times V \to \mathbb{R}^{r_1}$, $g_i : \mathbb{R}^n \times V \to \mathbb{R}^{s_1}$, $\widetilde{h}_i : \mathbb{R}^m \times V \to \mathbb{R}^{r_2}$, $\widetilde{g}_i : \mathbb{R}^n \times V \to \mathbb{R}^{s_2}$, $K : \mathbb{R}^n \times \mathbb{R}^n \times V \to \mathbb{R}^k$, são funções continuamente diferenciáveis para todo $i = 1, \ldots, N$. Assumiremos ainda que $r_1 \leq n$, $r_2 \leq m$ e $k \leq 2n$.

Podemos rescrever (PCD_σ) como o seguinte problema de programação matemática:

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(\alpha, \sigma) \\ \text{sujeito a: } H(\alpha, \sigma) = 0 \\ G(\alpha, \sigma) \leqq 0 \\ (\alpha, \sigma) \in \mathbb{X} := X \times U \times V. \end{array} \qquad (POE_\sigma)$$

onde $\alpha = (x, u) \in X \times U$,

$$\mathbb{J} : \mathbb{X} \to \mathbb{R}; \ \mathbb{J}(\alpha, \sigma) = \sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i, \sigma\right) \tag{7.4}$$

$$\begin{array}{c} H : \mathbb{X} \to \mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k; \\ H(\alpha, \sigma) = (H^1(\alpha, \sigma), H^2(\alpha, \sigma), H^3(\alpha, \sigma), H^4(\alpha, \sigma)) \end{array} \tag{7.5}$$

sendo

$$\begin{array}{l} H^1_{i+1}(\alpha_i, \sigma) = (x^1_{i+1} - \varphi^1_i\left(x_i, u_i, \sigma\right), \ldots, x^n_{i+1} - \varphi^n_i\left(x_i, u_i, \sigma\right)) \\ H^2_i(\alpha_i, \sigma) = (h^1_i(x_i, \sigma), \ldots, h^{r_1}_i(x_i, \sigma)) \\ H^3_i(\alpha_i, \sigma) = (\widetilde{h}^1_i(u_i, \sigma), \ldots, \widetilde{h}^{r_2}_i(u_i, \sigma)) \end{array}$$

$$\begin{array}{l} H^1(\alpha, \sigma) = (x_1 - \varphi_0\left(x_0, u_0, \sigma\right), \ldots, x_N - \varphi_{N-1}\left(x_{N-1}, u_{N-1}, \sigma\right)) \\ H^2(\alpha, \sigma) = (h_0(x_0, \sigma), \ldots, h_N(x_N, \sigma)) \\ H^3(\alpha, \sigma) = (\widetilde{h}_0(u_0, \sigma), \ldots, \widetilde{h}_{N-1}(u_{N-1}, \sigma)), \ H^4(\alpha, \sigma) = K(x_0, x_N, \sigma). \end{array}$$

$$\begin{array}{c} G : \mathbb{X} \to \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)}; \\ G(\alpha, \sigma) = (G^1(\alpha, \sigma), G^2(\alpha, \sigma)) \end{array} \tag{7.6}$$

onde

$$G_i^1(\alpha_i,\sigma) = (g_i^1(x_i,\sigma),\ldots,g_i^{s_1}(x_i,\sigma))$$
$$G_i^2(\alpha_i,\sigma) = (\widetilde{g}_i^1(u_i,\sigma),\ldots,\widetilde{g}_i^{s_2}(u_i,\sigma))$$

$$G^1(\alpha,\sigma) = (g_0(x_0,\sigma),\ldots,g_N(x_N,\sigma)),$$
$$G^2(\alpha,\sigma) = (\widetilde{g}_0(u_0,\sigma),\ldots,\widetilde{g}_{N-1}(u_{N-1},\sigma)).$$

Assumiremos que para cada $\sigma \in V$ o conjunto factível de (PCD_σ) é não vazio e existe um minimizador local $(x(\sigma),u(\sigma))$. Nosso interesse é fazer uma análise da estabilidade e sensibilidade de um processo $(x(\sigma),u(\sigma))$ na vizinhança V de $\hat{\sigma}$, ou seja, investigar as propriedades de continuidade e diferenciabilidade de $(x(\sigma),u(\sigma))$, tratado como uma função de parâmetro σ.

Em nossa análise vamos utilizar os resultados de estabilidade e sensibilidade da programação matemática obtidos em [43, 55, 83]. Usando uma estrutura específica das funções $\mathbb{J}$, H e G em (POE_σ) reformularemos estes resultados nos termos dos dados do problema original (PCD_σ).

Iniciamos introduzindo os seguintes conjuntos de índices das restrições ativas de G_i^1 e G_i^2, respectivamente.

$$I_i^1(\sigma) = \{s \in \{1,\ldots,s_1\} : g_i^s(x_i(\sigma),\sigma) = 0\}\,, i = 0,\ldots,N, \tag{7.7}$$
$$\widetilde{I}_i^1(\sigma) = \left\{t \in \{1,\ldots,s_2\} : \widetilde{g}_i^t(u_i(\sigma),\sigma) = 0\right\}, i = 0,\ldots,N-1. \tag{7.8}$$

7.2.1 Condição de Independência Linear

Analisaremos nesta seção a condição de independência linear, como apresentada na definição abaixo.

Definição 7.6 *Sejam $H(\cdot,\cdot,\sigma)$ e $G(\cdot,\cdot,\sigma)$ de classe C^1 para todo $\sigma \in V$. Dizemos que a condição de qualificação de independência linear (LICQ) é satisfeita em $(x(\sigma),u(\sigma))$ se o conjunto dos gradientes das restrições de desigualdades ativas e de restrição de igualdade são linearmente independentes.*

É conhecido (veja [50]) que a condição (LICQ) é equivalente a que o operador

$$\begin{gathered} M(\alpha(\sigma),\sigma) : \mathbb{R}^{n(N+1)+m(N)} \times \mathbb{R}^{s_1(N+1)+s_2(N)} \to \mathbb{R}^{n(N)+k+s_1(N+1)+s_2(N)}, \\ M(\alpha(\sigma),\sigma) = \begin{bmatrix} \nabla_\alpha H(\alpha(\sigma),\sigma) & 0 \\ \nabla_\alpha G(\alpha(\sigma),\sigma) & G(\alpha(\sigma),\sigma) \end{bmatrix} \end{gathered} \tag{7.9}$$

seja sobrejetivo.

Aqui $G(\alpha(\sigma),\sigma)$ denota a matriz diagonal $(s_1(N+1)+s_2(N))\times(s_1(N+1)+s_2(N))$ cujos elementos diagonais são $g_i^s(x_i(\sigma),\sigma)$ e $\widetilde{g}_i^t(u_i(\sigma),\sigma)$, nesta ordem.

Para simplificar a notação tomaremos $\varphi_i(\sigma) = \varphi_i(x_i(\sigma),u_i(\sigma),\sigma)$, $h_i(\sigma) = h_i(x_i(\sigma),\sigma)$, $\widetilde{h}_i(\sigma) = \widetilde{h}_i(u_i(\sigma),\sigma)$, $K(\sigma) = K(x_0(\sigma),x_N(\sigma),\sigma)$, $g_i(\sigma) = g_i(x_i(\sigma),\sigma)$ e $\widetilde{g}_i(\sigma) = \widetilde{g}_i(u_i(\sigma),\sigma)$.

Graças à estrutura do problema de controle ótimo discreto (PCD_σ) a condição de independência linear pode ser checada independentemente em cada estágio $i = 0,\ldots,$ $N-1$, como é formulado a seguir:

Lema 7.7 *Assuma que*

$$g_0^s(x_0,\sigma) < 0,\ \ s = 1,\ldots,s_1. \tag{7.10}$$

Então (LICQ) é satisfeita em $(x(\sigma), u(\sigma))$ de (PCD_σ) se e somente se os seguintes operadores

$$A_i(\sigma) : \mathbb{R}^{m+s_2+s_1} \to \mathbb{R}^{s_2+s_1},$$
$$A_i(\sigma) = \begin{bmatrix} \nabla\widetilde{g}_i(\sigma) & \widetilde{g}_i(\sigma) & 0 \\ \nabla g_{i+1}\nabla_u\varphi_i(\sigma) & 0 & g_{i+1}(\sigma) \end{bmatrix} \tag{7.11}$$

são sobrejetivos.

Aqui $\widetilde{g}_i(\sigma)$ e $g_i(\sigma)$ são matrizes diagonais $(s_2 \times s_2)$ e $(s_1 \times s_1)$, cujos elementos diagonais são $\widetilde{g}_i^t(\sigma)$ e $g_i^s(\sigma)$, respectivamente.

Demonstração. Segue de (7.9) que (LICQ) é satisfeita para (PCD_σ) se e somente se o seguinte sistema:

$$\begin{aligned}
w_{i+1} - \nabla_x\varphi_i(\sigma)\, w_i + \nabla_u\varphi_i(\sigma)\, v_i &= a_i & (7.12)\\
\nabla_{x_0}K(\sigma)w_0 + \nabla_{x_N}K(\sigma)w_N &= b_0 & (7.13)\\
\nabla h_i(\sigma)w_i &= c_i & (7.14)\\
\nabla\widetilde{h}_i(\sigma)v_i &= d_i & (7.15)\\
\nabla g_i(\sigma)w_i + g_i(\sigma)z_i &= e_i,\ i = 0, \ldots, N-1, & (7.16)\\
\nabla\widetilde{g}_i(\sigma)v_i + \widetilde{g}_i(\sigma)y_i &= f_i, & (7.17)
\end{aligned}$$

tem solução para quaisquer $a_i \in \mathbb{R}^n, b_0 \in \mathbb{R}^k$, $c_i \in \mathbb{R}^{r_1}$, $d_i \in \mathbb{R}^{r_2}$, $e_i \in \mathbb{R}^{s_1}$ e $f_i \in \mathbb{R}^{s_2}$.

De (7.16) temos

$$\nabla g_0(\sigma)w_0 + g_0(\sigma)z_0 = e_0 \tag{7.18}$$
$$\nabla g_{i+1}(\sigma)w_{i+1} + g_{i+1}(\sigma)z_{i+1} - \nabla g_i(\sigma)w_i - g_i(\sigma)z_i = e_{i+1} - e_i,\ i = 1, \ldots, N-1, \tag{7.19}$$

ou ainda,

$$\nabla g_{i+1}(\sigma)w_{i+1} + g_{i+1}(\sigma)(z_{i+1} - z_i) = e_{i+1} - e_i + \nabla g_i(\sigma)w_i + (g_i(\sigma) - g_{i+1}(\sigma))z_i, \quad i = 1, \ldots, N-1. \tag{7.20}$$

Multiplicando (7.12) por $\nabla g_{i+1}(\sigma)$ e subtraindo o resultado obtido de (7.20) obtemos

$$\begin{gathered}
\nabla g_{i+1}(\sigma)\nabla_u\varphi_i(\sigma)\, v_i + g_{i+1}(\sigma)(z_{i+1} - z_i) = \\
(\nabla g_i(\sigma) - \nabla_x\varphi_i(\sigma))w_i + (g_i(\sigma) - g_{i+1}(\sigma))z_i + e_{i+1} - e_i - \nabla g_{i+1}(\sigma)a_i = \\
B_i^1(\sigma)w_i + B_i^2(\sigma)z_i + \hat{e}_i,
\end{gathered} \tag{7.21}$$

onde,

$$\begin{aligned}
B_i^1(\sigma) &:= \nabla g_i(\sigma) - \nabla_x\varphi_i(\sigma)\\
B_i^2(\sigma) &:= g_i(\sigma) - g_{i+1}(\sigma)\\
\hat{e}_i &:= e_{i+1} - e_i - \nabla g_{i+1}(\sigma)a_i.
\end{aligned}$$

Logo, podemos rescrever as equações (7.17) e (7.21) da seguinte forma

$$\begin{gathered}
A_i(\sigma)\begin{bmatrix} v_i \\ y_i \\ z_{i+1} - z_i \end{bmatrix} := \begin{bmatrix} \nabla\widetilde{g}_i(\sigma) & \widetilde{g}_i(\sigma) & 0 \\ \nabla g_{i+1}\nabla_u\varphi(\sigma) & 0 & g_{i+1}(\sigma) \end{bmatrix}\begin{bmatrix} v_i \\ y_i \\ z_{i+1} - z_i \end{bmatrix} = \\
\begin{bmatrix} f_i \\ B_i^1(\sigma)w_i + B_i^2(\sigma)z_i + \hat{e}_i \end{bmatrix}.
\end{gathered} \tag{7.22}$$

É fácil ver que (7.22) tem uma solução para qualquer $(f_i, \hat{e}_i)$ se, e somente se, $A_i(\sigma)$ é sobrejetiva. Neste caso, (7.22) é satisfeita se tomarmos

$$\begin{aligned} \begin{bmatrix} v_i \\ y_i \\ z_{i+1} - z_i \end{bmatrix} &= A_i(\sigma)(A_i(\sigma)A_i^T(\sigma))^{-1} \begin{bmatrix} f_i \\ B_i^1(\sigma)w_i + B_i^2(\sigma)z_i + \hat{e}_i \end{bmatrix} := \\ \Theta_i(\sigma) \begin{bmatrix} f_i \\ B_i^1(\sigma)w_i + B_i^2(\sigma)z_i + \hat{e}_i \end{bmatrix} &:= \begin{bmatrix} \hat{\theta}_i^{11}(\sigma) + \hat{\theta}_i^{12}(\sigma)w_i + \hat{\theta}_i^{13}(\sigma)z_i \\ \hat{\theta}_i^{21}(\sigma) + \hat{\theta}_i^{22}(\sigma)w_i + \hat{\theta}_i^{23}(\sigma)z_i \\ \hat{\theta}_i^{31}(\sigma) + \hat{\theta}_i^{32}(\sigma)w_i + \hat{\theta}_i^{33}(\sigma)z_i \end{bmatrix}, \end{aligned} \tag{7.23}$$

onde $A_i(\sigma)(A_i(\sigma)A_i^T(\sigma))^{-1} := \Theta_i(\sigma)$ é a matriz por blocos

$$\Theta_i(\sigma) = \begin{bmatrix} \theta_i^{11}(\sigma) & \theta_i^{12}(\sigma) \\ \theta_i^{21}(\sigma) & \theta_i^{22}(\sigma) \\ \theta_i^{31}(\sigma) & \theta_i^{32}(\sigma) \end{bmatrix}.$$

Agora de (7.12), (7.14), (7.15) e (7.23) obtemos

$$\begin{gathered} w_{i+1} - (\nabla_x \varphi_i(\sigma) + \nabla_u \varphi_i(\sigma)\, \hat{\theta}_i^{12}(\sigma))w_i - \nabla_u \varphi_i(\sigma)\, \hat{\theta}_i^{13}(\sigma) v_i = \\ a_i + \nabla_u \varphi_i(\sigma)\, \hat{\theta}_i^{11}(\sigma) := \hat{a}_i, \end{gathered} \tag{7.24}$$

$$z_{i+1} - z_i - \hat{\theta}_i^{32}(\sigma)w_i - \hat{\theta}_i^{33}(\sigma)z_i = \hat{\theta}_i^{31}(\sigma) \;:=\; \hat{\theta}_i \tag{7.25}$$

$$\nabla h_i(\sigma)w_i \;:=\; c_i \tag{7.26}$$

$$\nabla \widetilde{h}_i(\sigma)\hat{\theta}_i^{12}(\sigma)w_i + \nabla \widetilde{h}_i(\sigma)\hat{\theta}_i^{13}(\sigma)z_i = d_i - \hat{\theta}_i^{11}(\sigma) \;:=\; \hat{d}_i. \tag{7.27}$$

Por outro lado, (7.13) juntamente com (7.18) e a hipótese (7.10) fornecem

$$\nabla_{x_0} K(\sigma)w_0 + \nabla_{x_N} K(\sigma)w_N \;:=\; b_0 \tag{7.28}$$

$$z_0 = g_0^{-1}(\sigma)(e_0 - \nabla g_0(\sigma)w_0) \;:=\; \hat{e}_0. \tag{7.29}$$

É obvio que (7.24)-(7.29) são unicamente determinados por (w_i, z_i) para qualquer $\hat{a}_i, \hat{\delta}_i, \hat{d}_i, \hat{e}_0$, isto é, para qualquer $a_i, b_0, c_i, d_i, e_i, f_i$. Em posse de (w_i, z_i) encontramos v_i e y_i por (7.23).

E com isto finalizamos a prova.

□

7.2.2 Condição Suficiente de Segunda Ordem Forte

No que se segue procederemos à análise da condição suficiente de segunda ordem forte (FSC) para o problema (PCD_σ). Para isso consideraremos o Lagrangeano
$\mathcal{L} : \mathbb{R}^{n(N+1)} \times \mathbb{R}^{m(N)} \times \mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^{k} \times \mathbb{R}^{s_1(N+1)} \times \mathbb{R}^{s_2(N)} \times V \to \mathbb{R}$

$$\begin{aligned} \mathcal{L}(x, u, p, \gamma^1, \gamma^2, \gamma^3, \mu^1, \mu^2, \sigma) \;:=\; & \mathbb{J}(x, u, \sigma) + K(x_0, x_N, \sigma)^T \gamma^3 + \sum_{i=0}^{N} h_i(x_i, \sigma)^T \gamma_i^1 + \\ & \sum_{i=0}^{N-1} \left(x_{i+1}^T p_i - \varphi_i(x_i, u_i, \sigma)^T p_i \right) + \sum_{i=0}^{N-1} \widetilde{h}_i(u_i, \sigma)^T \gamma_i^2 + \\ & \sum_{i=0}^{N} g_i(x_i, \sigma)^T \mu_i^1 + \sum_{i=0}^{N-1} \widetilde{g}_i(u_i, \sigma)^T \mu_i^2. \end{aligned} \tag{7.30}$$

Assumiremos que (LICQ) vale para (PCD_σ). Então, em particular, os multiplicadores de Lagrange $p(\sigma), \gamma^1(\sigma)$, $\gamma^2(\sigma), \gamma^3(\sigma), \mu^1(\sigma)$ e $\mu^2(\sigma)$ são unicamente determinados.

Por simplicidade denotaremos

$$\mathcal{L}(x(\sigma), u(\sigma), p(\sigma), \gamma^1(\sigma), \gamma^2(\sigma), \gamma^3(\sigma), \mu^1(\sigma), \mu^2(\sigma), \sigma) := \mathcal{L}(\sigma). \tag{7.31}$$

Segue da Definição 7.4 que a condição suficiente de segunda ordem forte (FSC) para (PCD_σ) equivale a

$$\sum_{i=0}^{N-1} \left[\left\langle w_i, \nabla^2_{xx}\mathcal{L}(\sigma)w_i\right\rangle + 2\left\langle w_i, \nabla^2_{xu}\mathcal{L}(\sigma)v_i\right\rangle + \left\langle v_i, \nabla^2_{uu}\mathcal{L}(\sigma)v_i\right\rangle\right] + \left\langle w_N, \nabla^2_{xx}\mathcal{L}(\sigma)w_N\right\rangle > 0$$

para todo $(w, v) \neq 0$ tal que

$$\begin{aligned}
w_{i+1} &= \nabla_x\varphi_i(\sigma)\,w_i + \nabla_u\varphi_i(\sigma)\,v_i, i = 0, \ldots, N-1,\\
\nabla_{x_0}K(\sigma)w_0 + \nabla_{x_N}K(\sigma)w_N &= 0,\\
\nabla h_i(\sigma)w_i &= 0, i = 0, \ldots, N,\\
\nabla\widetilde{h}_i(\sigma)v_i &= 0, i = 0, \ldots, N-1,\\
\nabla g_i^s(\sigma)w_i &= 0, s \in I_i^{1+}(\sigma), i = 0, \ldots, N,\\
\nabla\widetilde{g}_i^t(\sigma)v_i &= 0, t \in \widetilde{I}_i^{1+}(\sigma), i = 0, \ldots, N-1.
\end{aligned}$$

onde

$$I_i^{1+}(\sigma) := \left\{s \in I_i^1(\sigma) : \mu_i^{1s}(\sigma) > 0\right\} \tag{7.32}$$

$$\widetilde{I}_i^{1+}(\sigma) := \left\{t \in \widetilde{I}_i^1(\sigma) : \mu_i^{2t}(\sigma) > 0\right\}, \tag{7.33}$$

e $I_i^1(\sigma)$, $\widetilde{I}_i^1(\sigma)$ são como dados em (7.7) e (7.8).

Lema 7.8 *Suponha que (LICQ) vale para (PCD_σ). Então (FSC) é satisfeita se e somente se o seguinte problema de controle ótimo quadrático, $(QC_{\sigma,\boldsymbol{a}})$:*

Minimizar

$$\sum_{i=0}^{N-1} \left[\left\langle w_i, \nabla^2_{xx}\mathcal{L}(\sigma)w_i\right\rangle + 2\left\langle w_i, \nabla^2_{xu}\mathcal{L}(\sigma)v_i\right\rangle + \left\langle v_i, \nabla^2_{uu}\mathcal{L}(\sigma)v_i\right\rangle\right] + \left\langle w_N, \nabla^2_{xx}\mathcal{L}(\sigma)w_N\right\rangle$$

sujeito a:

$$\begin{aligned}
w_{i+1} &= \nabla_x\varphi_i(\sigma)\,w_i + \nabla_u\varphi_i(\sigma)\,v_i, i = 0, \ldots, N-1,\\
\nabla_{x_0}K(\sigma)w_0 + \nabla_{x_N}K(\sigma)w_N &= \boldsymbol{a},\\
\nabla h_i(\sigma)w_i &= 0, i = 0, \ldots, N,\\
\nabla\widetilde{h}_i(\sigma)v_i &= 0, i = 0, \ldots, N-1,\\
\nabla g_i^s(\sigma)w_i &= 0, s \in I_i^{1+}(\sigma), i = 0, \ldots, N,\\
\nabla\widetilde{g}_i^t(\sigma)v_i &= 0, t \in \widetilde{I}_i^{1+}(\sigma), i = 0, \ldots, N-1.
\end{aligned}$$

tem uma única solução para qualquer $\boldsymbol{a} \in \mathbb{R}^k$.

Demonstração. Inicialmente consideramos o caso homogêneo, isto é, quando $\mathbf{a} = 0$. O problema $(QC_{\sigma,\mathbf{0}})$ tem solução se e somente se o funcional objetivo é não negativo para todo processo factível.

De fato, suponha que exista um $(\mathbf{w}, \mathbf{v})$ factível de $(QC_{\sigma,\mathbf{a}})$ para o qual o funcional objetivo é negativo, escolhendo se necessário $(\mathbf{w}, \mathbf{v}) = \beta(\mathbf{w}, \mathbf{v})$, com $\beta > 0$ arbitrariamente grande, obtemos que o funcional objetivo é arbitrariamente negativo, ou seja, $(QC_{\sigma,\mathbf{0}})$ não tem solução.

Se $(QC_{\sigma,\mathbf{0}})$ tem solução, então $(\mathbf{w}, \mathbf{v}) \equiv (\mathbf{0}, \mathbf{0})$ é uma solução. Além disso, se a solução é única então (FSC) ocorre. Certamente, se (FSC) ocorre então $(QC_{\sigma,\mathbf{0}})$ tem única solução $(\mathbf{w}, \mathbf{v}) \equiv (\mathbf{0}, \mathbf{0})$.

Portanto, para completar a prova é suficiente mostrar que (LICQ) e (FSC) implicam a existência e unicidade da solução de $(QC_{\sigma,\mathbf{a}})$ para qualquer $\mathbf{a} \in \mathbb{R}^{\mathbf{k}}$.

Notemos que, usando o mesmo argumento utilizado na prova do Lema 7.7, encontramos que para qualquer $\mathbf{a} \in \mathbb{R}^{\mathbf{k}}$ existe um par $(\bar{w}(\sigma), \bar{v}(\sigma))$ satisfazendo as restrições do problema $(QC_{\sigma,\mathbf{a}})$. Introduzindo novas variáveis: $z = w - \bar{w}(\sigma)$ e $y = v - \bar{v}(\sigma)$ temos que $(QC_{\sigma,\mathbf{a}})$ formulado em termos de (z, y) possui restrições homogêneas e consequentemente (de (FSC)) tem solução única. E assim encerramos a demonstração.

□

Observemos que as condições suficientes de segunda ordem forte (FSC) podem ser verificadas através da existência e unicidade de solução do problema auxiliar de controle ótimo quadrático $(QC_{\sigma,\mathbf{a}})$.

7.2.3 Resultado de Estabilidade e Sensibilidade

Agora estamos em condições de estabelecer os resultados de estabilidade e sensibilidade para o problema de controle ótimo discreto (PCD_σ). É obvio que para garantir que as condições (LICQ) e (FSC) sejam satisfeitas necessitamos que todas as funções envolvidas no problema sejam de classe C^2 numa vizinhança de $(x(\hat{\sigma}), u(\hat{\sigma}), \hat{\sigma})$. Como consequência dos Lemas 7.7 e 7.8, obtemos:

Teorema 7.9 *Assuma que:*

i. *As funções de (PCD_σ) são de classe C^2 numa vizinhança de $(x(\hat{\sigma}), u(\hat{\sigma}), \hat{\sigma})$,*

ii. $g_0^s(x_0(\hat{\sigma}), \hat{\sigma}) < 0, \; s = 1, \ldots, s_1$,

iii. *para $\sigma = \hat{\sigma}$ as matrizes $A_i(\sigma)$ em (7.11) tem posto completo,*

iv. *o problema de controle ótimo quadrático $(QC_{\hat{\sigma},\boldsymbol{a}})$ tem uma única solução para qualquer $\boldsymbol{a} \in \mathbb{R}^k$.*

(a) *Então existem vizinhanças $V(\hat{\sigma}) \subset \Sigma$ e $U(x(\hat{\sigma}), u(\hat{\sigma})) \subset \mathbb{R}^{n(N+1)+m(N)}$, tais que para cada $\sigma \in V(\hat{\sigma})$, $(x(\hat{\sigma}), u(\hat{\sigma}))$ é o único minimizador de (PCD_σ) em $U(x(\hat{\sigma}), u(\hat{\sigma}))$ e $p(\sigma)$, $\gamma^1(\sigma), \gamma^2(\sigma)$, $\gamma^3(\sigma), \mu^1(\sigma), \mu^2(\sigma)$ são os únicos multiplicadores de Lagrange associados.*

(b) *As funções $x(\cdot), u(\cdot), p(\cdot), \gamma^1(\cdot), \gamma^2(\cdot), \gamma^3(\cdot), \mu^1(\cdot), \mu^2(\cdot)$ são Lipschitz contínuas em $V(\hat{\sigma})$ e direcionalmente diferenciáveis em $\hat{\sigma}$.*

(c) *Além disso, para cada direção d, $(\hat{x}, \hat{u})$ é processo ótimo do problema de controle ótimo quadrático abaixo, onde $\hat{x}(\hat{\sigma}) = \delta_\sigma^+ x(\hat{\sigma}, d)$ e $\hat{u}(\hat{\sigma}) = \delta_\sigma^+ u(\hat{\sigma}, d)$, cujos únicos* *multiplicadores associados são* $p(\hat{\sigma}) = \delta_\sigma^+ p(\hat{\sigma}, d)$, $\gamma^1(\hat{\sigma}) = \delta_\sigma^+ \gamma^1(\hat{\sigma}, d)$, $\gamma^2(\hat{\sigma}) = \delta_\sigma^+ \gamma^2(\hat{\sigma}, d)$, $\gamma^3(\hat{\sigma}) = \delta_\sigma^+ \gamma^3(\hat{\sigma}, d)$, $\mu^1(\hat{\sigma}) = \delta_\sigma^+ \mu^1(\hat{\sigma}, d)$, $\mu^2(\hat{\sigma}) = \delta_\sigma^+ \mu^2(\hat{\sigma}, d)$,

$$\begin{array}{l}
\textit{Minimizar} \\
\displaystyle\sum_{i=0}^{N-1} \left[\frac{1}{2} \left\langle w_i, \nabla^2_{xx}\mathcal{L}(\hat{\sigma}) w_i \right\rangle + \left\langle w_i, \nabla^2_{xu}\mathcal{L}(\hat{\sigma}) v_i \right\rangle + \frac{1}{2} \left\langle v_i, \nabla^2_{uu}\mathcal{L}(\hat{\sigma}) v_i \right\rangle + \right. \\
\left. \left\langle w_i, \nabla^2_{x\sigma}\mathcal{L}(\hat{\sigma}) d \right\rangle + \left\langle v_i, \nabla^2_{u\sigma}\mathcal{L}(\hat{\sigma}) d \right\rangle \right] + \left\langle w_N, \nabla^2_{xx}\mathcal{L}(\hat{\sigma}) w_N \right\rangle + \left\langle w_N, \nabla^2_{x\sigma}\mathcal{L}(\hat{\sigma}) d \right\rangle \\
\textit{sujeito a:} \\
w_{i+1} = \nabla_x \varphi_i(\hat{\sigma})\, w_i + \nabla_u \varphi_i(\hat{\sigma})\, v_i + \nabla_\sigma \varphi_i(\hat{\sigma})\, d, i = 0, \ldots, N-1, \\
\nabla_{x_0} K(\hat{\sigma}) w_0 + \nabla_{x_N} K(\hat{\sigma}) w_N + \nabla_\sigma K(\hat{\sigma}) d = 0, \\
\nabla_x h_i(\hat{\sigma}) w_i + \nabla_\sigma h_i(\hat{\sigma}) d = 0, i = 0, \ldots, N, \\
\nabla_u \widetilde{h}_i(\hat{\sigma}) v_i + \nabla_\sigma \widetilde{h}_i(\hat{\sigma}) d = 0, i = 0, \ldots, N-1,
\end{array} \qquad (QC_{\hat{\sigma},d})$$

$$\begin{array}{ll}
\nabla_x g_i^s(\hat{\sigma}) w_i + \nabla_\sigma g_i^s(\hat{\sigma}) d \begin{cases} = 0, s \in I_i^{1+}(\hat{\sigma}) \\ \leqq 0, s \in I_i^1(\hat{\sigma}) \backslash I_i^{1+}(\hat{\sigma}), \end{cases} & i = 0, \ldots, N \\
\nabla_u \widetilde{g}_i^t(\hat{\sigma}) w_i + \nabla_\sigma \widetilde{g}_i^t(\hat{\sigma}) d \begin{cases} = 0, t \in \widetilde{I}_i^{1+}(\hat{\sigma}) \\ \leqq 0, t \in \widetilde{I}_i^1(\hat{\sigma}) \backslash \widetilde{I}_i^{1+}(\hat{\sigma}), \end{cases} & i = 0, \ldots, N-1.
\end{array}$$

Note que no Teorema 7.9, a condição (FSC) nos permite garantir que o problema auxiliar $(QC_{\hat{\sigma},d})$ possui solução única enquanto à condição (LICQ) está associado a unicidade dos multiplicadores de Lagrange.

Ainda da equivalência dos problemas (PCD_σ) e (POE_σ) outros resultados conhecidos da programação matemática podem ser adaptados aos problemas de controle ótimo discreto. Em particular, usando o resultado de estabilidade estabelecido por Fiacco [43] com respeito à diferenciabilidade contínua das soluções (relativas ao parâmetro) de problemas de programação matemática, obtemos:

Corolário 7.10 *Se no Teorema 7.9 em adição às hipóteses (i)-(iv), a condição de complementaridade estrita é válida em* $\hat{\sigma}$*, isto é,*

$$(v) \quad \begin{array}{l} I_i^1(\hat{\sigma}) = I_i^{1+}(\hat{\sigma}), \ i = 0, \ldots, N \\ \widetilde{I}_i^1(\hat{\sigma}) = \widetilde{I}_i^{1+}(\hat{\sigma}), \ i = 0, \ldots, N-1, \end{array} \tag{7.34}$$

sendo $I_i^{1+}(\hat{\sigma})$ *e* $\widetilde{I}_i^{1+}(\hat{\sigma})$ *dados por (7.32) e (7.33), respectivamente, então* $(x(\cdot), u(\cdot))$ *é diferenciável numa vizinhança de* $\hat{\sigma}$*. A derivada é dada como solução de* $(QC_{\hat{\sigma},d})$ *onde as restrições de desigualdade se anulam.*

7.3 Problema (PCH_σ)

No início deste capítulo apontamos alguns aspectos a respeito do problema de programação matemática (PH). Este tipo de problema merece uma atenção especial porque esta configuração clássica remete aos fundamentos da análise não linear e otimização. Grosso modo, como veremos abaixo, algumas propriedades de problemas com restrições de igualdade não regulares tornam-se muito similares àquelas dos problemas regulares.

Seja o seguinte problema de controle ótimo discreto, definido para cada $\sigma \in V \subset \Sigma$:

$$\begin{array}{l} \text{Minimizar } \mathbb{J}(x_i, u_i, \sigma) = \sum_{i=0}^{N-1} \Phi_i\left(x_i, u_i, \sigma\right) \\ \text{sujeito a: } x_{i+1} = \varphi_i\left(x_i, u_i, \sigma\right), \ i = 0, \ldots, N-1, \\ h_i(x_i, \sigma) = 0, i = 0, \ldots, N, \\ \widetilde{h}_i(u_i, \sigma) = 0, i = 0, \ldots, N-1, \\ K(x_0, x_N, \sigma) = 0, \end{array} \qquad (PCH_\sigma)$$

onde as funções são dadas como no problema (PDC_σ).

Naturalmente o problema de programação matemática equivalente a (PCH_σ) é:

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(\alpha, \sigma) \\ \text{sujeito a: } H(\alpha, \sigma) = 0 \end{array}$$

onde J e H são como (7.4) e (7.5), $\alpha = (x, u) \in X \times U$.

Neste caso, existe uma vizinhança $\bar{V}$ de $\hat{\alpha} = (\hat{x}, \hat{u})$ tal que $\hat{\alpha}$ é solução global do problema não perturbado restrito a $\bar{V}$:

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(\alpha, \hat{\sigma}) \\ \text{sujeito a: } H(\alpha, \hat{\sigma}) = 0 \\ \alpha \in \bar{V} \end{array} \qquad (POH_{\hat{\sigma}})$$

onde assumimos $\bar{V}$ um conjunto compacto em $X \times U$. E definimos o problema perturbado como o seguinte problema de programação matemática:

$$\begin{array}{c} \text{Minimizar } \mathbb{J}(\alpha, \sigma) \\ \text{sujeito a: } H(\alpha, \sigma) = 0 \\ \alpha \in \bar{V} \end{array} \qquad (POH_\sigma)$$

De tal maneira que $S(\sigma)$ é conjunto de soluções do problema (POH_σ),

$$\omega(\sigma) = \inf_{(\hat{x}, \hat{u}) \in D(\sigma) \cap \bar{V}} \mathbb{J}(\alpha, \sigma), \ \sigma \in \Sigma$$

o valor ótimo da função do problema (POH_σ), onde $D(\sigma) := \{\alpha \in X \times U : H(\alpha, \sigma) = 0\}$, o Lagrangeano é dado em (7.30) e

$$\begin{aligned} \Lambda(\hat{\alpha}, \hat{\sigma}) \ = \ & \Big\{\gamma = (p, \gamma^1, \gamma^2, \gamma^3) \in Y^* = (\mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k)^* : \\ & \nabla_x \mathcal{L}(\hat{\sigma}) + \nabla_u \mathcal{L}(\hat{\sigma}) = 0\} \end{aligned} \qquad (7.35)$$

é o conjunto dos multiplicadores de Lagrange do problema (POH_σ) onde, $\mathcal{L}(\sigma)$ é como em (7.31).

Além disso, temos a seguinte definição de condição suficiente de segunda ordem.

Definição 7.11 *(**Arutyunov e Izmailov [4]**) A condição suficiente de segunda ordem (SOSC) é válida em $\hat{\alpha}$ se $\Lambda(\hat{\alpha}, \hat{\sigma}) \neq \emptyset$ e $\forall d \in$ **Ker**$\nabla_\alpha H(\hat{\alpha}, \hat{\sigma}),\ d \neq 0, \exists \gamma \in \Lambda(\hat{\alpha}, \hat{\sigma})$ tal que $\nabla^2_{\alpha\alpha} \mathcal{L}(\hat{\sigma})[d, d] > 0$.*

Assumimos *a priori* que a seguinte condição não é válida:

$$\mathbf{Im}\nabla_\alpha H(\hat{\alpha}, \hat{\sigma}) = Y$$

com $Y = (\mathbb{R}^{n(N)} \times \mathbb{R}^{r_1(N+1)} \times \mathbb{R}^{r_2(N)} \times \mathbb{R}^k)$. Isto implica que a condição de regularidade 7.1, não é satisfeita para o problema $(POH)_\sigma$.

Seja $\hat{\alpha} \in D(\hat{\sigma})$. Assuma que $Y_1 = \mathbf{Im}\nabla_\alpha H(\hat{\alpha}, \hat{\sigma})$ é um subespaço fechado em Y e tem $Y_1^\perp$ como seu subespaço complementar. Considere o seguinte operador:

$$\Psi(P, d) : X \times U \to Y;\ \Psi(P, d) = \nabla_\alpha H(\hat{\alpha}, \hat{\sigma}) + P\nabla^2_{\alpha\alpha} H(\hat{\alpha}, \hat{\sigma})d,\ d \in X \times U, \tag{7.36}$$

onde P é o projetor ortogonal de $Y_1^\perp$ em Y.

Então, temos o seguinte resultado:

Teorema 7.12 (Arutyunov e Izmailov [4]) *Seja $\hat{\alpha}$ uma solução do problema não perturbado $(POH_{\hat{\sigma}})$. Se $\hat{\alpha}$ é um ponto 2-regular no sentido da Definição 5.4 e além disso, vale a condição (SOSC) e* $\boldsymbol{Im}\Psi(P, d) = Y$*, então, existe uma vizinhança V_1 de $\hat{\alpha}$ tal que para $\sigma \in \Sigma$*

$$\sup_{\alpha \in S(\sigma)} \|\alpha - \hat{\alpha}\| \to 0, \quad \textit{quando } \sigma \to \hat{\sigma}.$$

Em particular, se $\hat{\alpha}$ é a única solução de $POH_{\hat{\sigma}}$, então

$$\sup_{\alpha \in S(\sigma)} \|\alpha - \hat{\alpha}\| = O(\|\sigma - \hat{\sigma}\|^{\frac{1}{2}}),\ \ e\ \ \omega(\sigma) = \omega(\hat{\sigma}) + O(\|\sigma - \hat{\sigma}\|).$$

A prova deste teorema pode ser encontrada em [4] ou ainda, [65].

Observação 7.1 *1. Observemos que a hipótese de que $\hat{\alpha}$ é um ponto 2-regular, segundo a Definição 5.4 juntamente com a hipótese de que* $\boldsymbol{Im}\Psi(P, d) = Y$ *implica na hipótese (H1) do Teorema 12 de [4].*

2. *É importante ressaltar ainda que da equivalência entre a 2-regularidade da Definição 5.4 e da Definição 5.8 (provada no Lema 5.9), basta assumir que o problema (PCH_σ) seja 2-regular no sentido da Definição 5.8.*

Capítulo 8

Problemas Não-Suaves

Neste capítulo abordaremos problemas não-suaves tanto no contexto da programação matemática quanto sob a perspectiva do controle ótimo discreto.

Na Seção 8.1 recordaremos algumas definições e resultados básicos da Análise Não-Suave e apresentamos alguns resultados a um problema de programação matemática específico. Na Seção 8.2 demonstraremos o Princípio do Máximo Discreto para uma certa classe de problemas de controle ótimo discreto com restrições de igualdade e de desigualdade.

8.1 Problemas de Otimização Não-Diferenciáveis

Nesta seção, consideraremos o seguinte problema de Otimização Não-Diferenciável:

$$\begin{array}{l} \text{Minimizar } f(x) := (f_1(x), \cdots, f_p(x)) \\ \text{sujeito a: } g_i(x) \leq 0,\ i = 1, \cdots, m \\ h_j(x) = 0,\ j = 1, \cdots, r \\ x \in C \end{array} \tag{MOP}$$

onde $f_k, g_i, h_j : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, Ω é um conjunto aberto que contém o conjunto C e C é um subconjunto fechado e não vazio de $\mathbb{R}^n$. As funções envolvidas são todas localmente Lipschitz em Ω, $p \geq 1$.

Utilizaremos aqui as ferramentas da Análise Não-Suave de Clarke - derivadas e gradientes de Clarke para estudarmos este problema e também utilizaremos noções adequadas de convexidade generalizada.

Discutiremos a relação existente entre *calmness*[1] e eficiência própria para uma certa classe de problemas multiobjetivos não diferenciáveis. Estes resultados aparentemente estão desvinculados do tema desta Tese. Entretanto, esta desvinculação é apenas aparente, uma vez que todo problema de controle ótimo discreto pode ser reformulado como um problema de programação matemática - conforme fora longamente explanado ao longo deste trabalho.

É conhecido na literatura especializada que a propriedade de *calmness* é uma condição de regularidade básica para o estudo do comportamento de sensibilidade de problemas variacionais sujeitos a perturbações. Veja por exemplo, as referências [22, 60].

[1] Preferimos não traduzir o termo "calmness" para o português.

Também é conhecida a relação existente entre a propriedade de *calmness* para problemas de otimização (escalar) com restrições e a validade da penalização exata. Veja [26, 86] e [98], por exemplo. Assim sendo, o estudo desta propriedade é de relevância tanto teórica quanto abstrata. Daí, nosso interesse em estudá-la. Por outro lado, para problemas de programação multiobjetivo não diferenciáveis, Benson e Morin [10] estabelecem algumas relações existentes entre as soluções propriamente Pareto e a propriedade de estabilidade para um certo problema (escalar) relacionado ao problema original. Mais precisamente, demonstram uma condição necessária e suficiente para que uma solução Pareto seja propriamente Pareto. Tal condição se expressa em termos da estabilidade de um certo problema escalar. Para isto, os autores assumem que os dados do problema são convexos.

Nesta Seção, obteremos um resultado semelhante a este, para uma certa classe de problemas multiobjetivos não convexos, utilizaremos uma noção de convexidade generalizada e também o conceito de *calmness*.

Esta Seção tem a seguinte estrutura: Na Subseção 8.1.1, apresentamos os resultados básicos de Análise Não Suave que serão necessários e as definições de convexidade generalizada que serão utilizadas na discussão subsequente; na Subseção 8.1.2 discutiremos a propriedade de *calmness*, teoremas de dualidade e pontos de sela (para o caso $p = 1$) e, na Subseção 8.1.3 discutiremos as relações existentes entre as soluções propriamente Pareto de (MOP) e a propriedade de calmness de uma certa classe de problemas escalarizados.

8.1.1 Resultados Básicos de Análise não-suave

Neste Capítulo, assumiremos que o espaço $\mathbb{R}^n$ está munido do produto interno usual o qual denotaremos $\langle \cdot, \cdot \rangle$ e a norma associada a este produto será denotada $|| \cdot ||$. Admitiremos que este espaço está munido da topologia induzida pela norma $|| \cdot ||$.

Sejam Ω um subconjunto aberto e não vazio de $\mathbb{R}^n$ e $\phi : \Omega \to \mathbb{R}$ uma função dada. Diremos que ϕ é **Lipschitz próximo** de $x \in \Omega$ se existem $\delta > 0$ e $k = k(x, \delta) > 0$ tais que

$$|\phi(x_1) - \phi(x_2)| \leq k||x_1 - x_2||$$

para quaisquer $x_1, x_2 \in \Omega \cap B(x, \delta)$, onde $B(x, \delta)$ é a bola aberta de centro em x e raio δ. Além disto, se ϕ é Lipschitz próximo de x, para todo $x \in \Omega$, diremos simplesmente que ϕ é **localmente Lipschitz** em Ω.

Suponha agora que a função ϕ é localmente Lipschitz em Ω. Sejam fixados $x \in \Omega$ um ponto e $v \in \mathbb{R}^n$ uma direção. A **derivada direcional generalizada** (no sentido de Clarke) de ϕ em x, na direção v, é dada por

$$\phi^0(x; v) = \lim \sup_{\substack{y \to x \\ \lambda \to 0^+}} \frac{\phi(y + \lambda v) - \phi(y)}{\lambda}$$

e o **subdiferencial** (de Clarke) de ϕ em $x \in \Omega$ é o conjunto

$$\partial\phi(x) = \{\xi \in \mathbb{R}^n : \phi^0(x; v) \leq \langle \xi, v \rangle, \forall v \in \mathbb{R}^n\}.$$

Algumas das principais propriedades da derivada e do subdiferencial de Clarke são dadas na seguinte proposição.

Proposição 8.1 *Seja $\phi : \Omega \to \mathbb{R}$ uma função localmente Lipschitz em Ω, com constante k. Então:*

1. *A função $v \longmapsto \phi^0(x; v)$ é finita, sublinear e satisfaz $|\phi^0(x; v)| \leq k||v||, \forall v \in \mathbb{R}^n$;*

2. *Para cada $x \in \Omega$, o conjunto $\partial\phi(x)$ é convexo, compacto e não vazio. Além disto, $||\xi|| \leq k$, para todo $\xi \in \partial\phi(x)$;*

3. *Para todo $v \in \mathbb{R}^n$, tem-se $\phi^0(x; v) = \max\{\langle \xi, v \rangle : \xi \in \partial\phi(x)\}$;*

4. *$\xi \in \partial\phi(x)$ se, e somente se, $\phi^0(x; v) \geq \langle \xi, v \rangle$, para todo $v \in \mathbb{R}^n$.*

Para obtermos condições de otimalidade em termos destes conceitos, necessitaremos relembrar as noções de vetor tangente e de vetor normal.

Inicialmente observamos que se C é um subconjunto não vazio de $\mathbb{R}^n$, então a função distância $d_C : \mathbb{R}^n \to \mathbb{R}$ definida por $d_C(x) = \inf_{y \in C} ||y - x||$ não é diferenciável, mas é globalmente Lipschitz - neste caso, com constante de Lipschitz igual a 1.

Sejam $C \subset \mathbb{R}^n$ e $x \in C$. Dizemos que $v \in \mathbb{R}^n$ é um vetor tangente a C em x se $d_C^0(x; v) = 0$. O conjunto dos vetores tangentes será denotado $T_C(x)$ e é chamado **cone tangente** (de Clarke). Por polaridade, se define o **cone normal** (de Clarke),

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

Pode-se demonstrar que estes cones coincidem com os cones tangente e normal (usuais, da Análise Convexa), quando C é um conjunto convexo. Além disto, para cada $x \in C$, estes conjuntos são, ambos, cones convexos e fechados.

Por fim, recordamos um resultado que estabelece uma condição necessária de otimalidade em termos de uma condição estacionária.

Proposição 8.2 *Sejam $\phi : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ uma função definida no conjunto aberto Ω e C um subconjunto não vazio de Ω. Se x^* é um minimizador de ϕ em C, então vale a inclusão*

$$0 \in \partial\phi(x^*) + N_C(x^*). \tag{8.1}$$

(Além disto, um ponto $x^* \in C$ que satisfaz (8.1) é chamado ponto estacionário de ϕ em C)

Observação 8.1 1. *Das considerações feitas, a equação (8.1) é equivalente a*

$$\phi^0(x^*; v) \geq 0, \forall v \in T_C(x^*).$$

2. *Para o caso mais simples em que $C = \Omega = \mathbb{R}^n$ (ou mais geralmente, se C é um subconjunto aberto e não vazio de $\mathbb{R}^n$) tem-se $T_C(x^*) = \mathbb{R}^n$ e portanto $N_C(x^*) = \{0\}$ e (8.1) se torna*
$$0 \in \partial\phi(x^*).$$

3. *Se ϕ é continuamente diferenciável em C e C é aberto, tem-se $\partial\phi(x^*) = \{\nabla\phi(x^*)\}$ (e, neste caso, (8.1) se torna $\nabla\phi(x^*) = 0$).*

Para mais detalhes sobre Análise Não-diferenciável de Clarke, consulte por exemplo [26].

Noções de convexidade generalizada

Utilizaremos neste trabalho as seguintes noções de convexidade generalizadas introduzidas por Phoung, Sach e Yen [80]:

Definição 8.3 *Sejam Ω um subconjunto aberto e não vazio de $\mathbb{R}^n$, $\phi : \Omega \to \mathbb{R}$ uma função localmente Lipschitz em Ω e C um subconjunto não vazio de Ω. (i) Dizemos que ϕ é* ***invexa*** *em x_0 no conjunto C se para todo $y \in C$, existe um vetor $\eta(y, x_0) \in T_C(x_0)$ satisfazendo*

$$\phi(y) - \phi(x_0) \geq \phi^0(x_0; \eta(y, x_0)).$$

(ii) Dizemos que ϕ é ***infine***[2] *em x_0 no conjunto C se para todo $y \in C$, existe um vetor $\eta(y, x_0) \in T_C(x_0)$ satisfazendo*

$$\phi(y) - \phi(x_0) = \phi^0(x_0; \eta(y, x_0)).$$

Definição 8.4 *Sejam Ω um subconjunto aberto e não vazio de $\mathbb{R}^n$, $f_k : \Omega \to \mathbb{R}$, $g_i : \Omega \to \mathbb{R}$ $(k = 1, \cdots m,\ i = 1, \cdots, r)$ funções localmente Lipschitz em Ω e C um subconjunto não vazio de Ω. Sejam, ainda, $f(x) = (f_1(x), \cdots, f_m(x))$ e $g(x) = (g_1(x), \cdots, g_r(x))$. Diremos que (f, g) é* ***invex-infine*** *em x_0 no conjunto C se as funções f_k $(k = 1, \cdots, m)$ são invexas em x_0, as funções g_i $(i = 1, \cdots, r)$ são infine em x_0 no conjunto C, com respeito a uma mesma η.*

Uma interessante propriedade desta classe de funções é que elas satisfazem uma generalização do Teorema de Alternativa de Motzkin, a qual utilizaremos neste trabalho na obtenção de alguns de nossos resultados, como se verá nas próximas subseções. Antes de enunciá-lo, fixaremos a notação utilizada neste enunciado.

Considere os conjuntos de índices:

$$K = \{1, \cdots, p\}, I = \{1, \cdots, m\}, J = \{1, \cdots, r\},$$

e a função

$$s(\cdot) = \max\{\max_{k \in K} f_k(\cdot), \max_{i \in I} g_i(\cdot), \max_{j \in J} |h_j(\cdot)|\}.$$

Para cada $x_0 \in C$ fixado, definem-se:

$$\begin{aligned} K_0 &: = \{k \in K : f_k(x_0) = s(x_0)\} \\ I_0 &: = \{i \in I : g_i(x_0) = s(x_0)\} \\ J_0 &: = \{j \in J : |h_j(x_0)| = s(x_0)\} \\ I_0' &= I_0 \cup J_0. \end{aligned}$$

Definição 8.5 *Diremos que a* ***condição (CQ)*** *vale em $x_0 \in C$ se não existem escalares $\beta_j \in \mathbb{R}$, $j \in I_0'$, não todos nulos e tais que $\beta_j \geq 0$, $j \in I_0$ e*

$$0 \in \sum_{j \in I_0} \beta_j \partial g_j(x_0) + \sum_{j \in J_0} \beta_j \partial h_j(x_0) + N_C(x_0).$$

Por fim, podemos enunciar o Teorema de Motzkin Generalizado:

[2] Preferimos não traduzir o termo "infine" para o português e igualmente para o termo "invex-infine".

Teorema 8.6 (Teorema de Motzkin Generalizado) *(Sach, Lee e Kim [89]) Assuma que a função s atinge seu mínimo em C no ponto $x_0 \in C$ e que $((f_{K_0}, g_{I_0}); h_{J_0})$ é invex-infine em x_0 no conjunto C. Suponha também que a condição (CQ) é válida em x_0 e que o sistema*

$$f(x) \leqq 0, g(x) \leqq 0, h(x) = 0,\ x \in C$$

possui pelo menos uma solução $x \in \mathbb{R}^n$. Então, uma e somente uma das seguintes condições é satisfeita:

(i) O sistema $f(x) < 0, g(x) \leqq 0, h(x) = 0,\ x \in C$ tem uma solução;

(ii) Existem vetores $\alpha_K \geq 0,\ \beta_I \geqq 0, \gamma_J$ tais que $\sum_{k \in K} \alpha_k f_k(x) + \sum_{i \in I} \beta_i g_i(x) + \sum_{j \in J} \gamma_j h_j(x) \geq 0, \forall x \in C$.

Sobre a notação utilizada aqui: $f_{K_0}(x) = (f_k(x))_{k \in K_0}$, $\alpha_K = (\alpha_k)_{k \in K}$, etc.

Vale observar que, na ausência de restrições de igualdade, este Teorema se colapsa no Teorema de Gordan Generalizado, obtido por Brandão *et al.* [23].

8.1.2 Condições de Ponto de Sela, Dualidade e *calmness*

O conceito de calmness está muito relacionado à análise de estabilidade de problemas de otimização (escalares).

Vamos considerar a seguinte família de problemas: Para cada $p = (p_1, \cdots, p_m) \in \mathbb{R}^m$, $q = (q_1, \cdots, q_r) \in \mathbb{R}^r$ se define

$$\begin{array}{l} \text{Minimizar } f(x) \\ \text{sujeito a: } g_j(x) \leq p_j,\ j = 1, \cdots, m \\ h_k(x) = q_k,\ k = 1, \cdots, r \\ x \in C \end{array} \qquad (\mathrm{P}(p,q))$$

onde onde $f, g_i, h_j : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, Ω é um conjunto aberto que contém o conjunto C e C é um subconjunto fechado e não vazio de $\mathbb{R}^n$. Por simplicidade, nesta subseção denotaremos o problema $(\mathrm{P}(0,0))$ por (P).

Além disto, definiremos também a **função valor** associada ao problema $(\mathrm{P}(p,q))$, $V : \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R} \cup \{+\infty, -\infty\}$,

$$V(p,q) = \inf\{f(x) : x \in C, g(x) \leqq p,\ h(x) = q\}.$$

Definição 8.7 *(Clarke [29]) (i) Seja x uma solução de (P). Dizemos que o problema (P) é **calmo em** x se existem escalares $\varepsilon > 0$, $M > 0$ tais que $p \in \mathbb{R}^m, q \in \mathbb{R}^r$, $||(p,q)|| < \varepsilon$ e $x' \in B(x, \varepsilon)$ factível para (P(p, q)) tem-se*

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

*(ii) Dizemos que o problema (P) é **calmo** se $V(0,0)$ é finito e suponha que a função valor satisfaz*

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

Observação 8.2 *Pode-se demonstrar que se (P) é calmo, então (P) é calmo em cada solução x de (P). (Veja Proposição 6.4.2, de [26]). Além disto, é imediato da definição: Se V é Lipschitz perto da origem, então (P) é calmo.*

Teorema 8.8 *(Fritz-John [29]) Se x^* é solução local de (P), então existem escalares λ, μ_i, ν_j $(i = 1, \cdots, m,\ j = 1, \cdots, r)$ não todos nulos e tais que*

$$\begin{aligned} 0 \in \lambda\partial f(x^*) + \sum_{i=1}^{m} \mu_i \partial g_i(x^*) + \sum_{j=1}^{r} \nu_j \partial h_j(x^*) + N_C(x^*) \\ \lambda, \mu_i \geq 0,\ i = 1, \cdots, m \\ \mu_i g_i(x^*) = 0. \end{aligned} \tag{8.2}$$

Se (8.2) se cumpre com $\lambda = 1$, dizemos que x^* é um ponto KKT de (P).

Observação 8.3 *A propriedade de calmness está muito relacionada à regularidade do problema (P). De fato, se x^* é uma solução de (P) e o problema (P) é calmo em x^*, então $\lambda \neq 0$ em (8.2). (Veja Proposição 6.4.4 de [26]).*

É bastante conhecida na literatura as relações existentes entre convexidade, otimalidade e pontos de sela. Esses conceitos são relacionados através de alguma condição de qualificação e alguma noção de convexidade generalizada. Veja [64], [90] e [92], por exemplo. Nossos resultados foram obtidos utilizando os conceitos de invexidade e de *calmness*.

Recordemos que uma terna $(x^*, \mu^*, \gamma^*) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^r$ é chamado **ponto de sela** do problema (P) se satisfaz

$$\mathcal{L}(x^*, \mu, \gamma) \leq \mathcal{L}(x^*, \mu^*, \gamma^*) \leq \mathcal{L}(x, \mu^*, \gamma^*),\ \forall \mu \in \mathbb{R}^m_+, \gamma \in \mathbb{R}^r, x \in C$$

onde $\mathcal{L}(x, \mu, \gamma) = f(x) + \langle \mu, g(x) \rangle + \langle \gamma, h(x) \rangle$ $\quad (x \in \mathbb{R}^n, \mu \in \mathbb{R}^m_+, \gamma \in \mathbb{R}^r)$ é a função Lagrangeana.

Se $((f, g), h)$ é invex-infine em x_0 no conjunto C e $x_0 \in C$ é um ponto KKT de (P), então x_0 é solução de (P). (De fato, sob estas hipóteses, x_0 é um ponto estacionário de $\mathcal{L}(\cdot, \mu^*, \gamma^*)$ e a função Lagrangeana é invexa com respeito a η em x_0 no conjunto C. Logo, x_0 minimiza $\mathcal{L}(\cdot, \mu^*, \gamma^*)$ sobre C. Deste fato e das condições de factibilidade segue que x_0 é solução de (P)).

No que se segue, estabeleceremos as relações existentes entre as soluções de (P) e a condição de ponto de sela.

Proposição 8.9 *Seja x^* uma solução de (P). Suponha que (P) é calmo em x^* e que $((f, g); h)$ é invex-infine em x^* no conjunto C. Então, existem $\mu^* \in \mathbb{R}^m_+$, $\gamma^* \in \mathbb{R}^r$ tais que (x^*, μ^*, γ^*) é um ponto de sela de (P).*

Demonstração. Segue as mesmas linhas da prova do Teorema 2.5 de [19] (veja também [23]), onde o autor não considera as restrições de igualdade em (P) e utiliza uma condição de qualificação de tipo Slater. Mas o argumento por ele utilizado se baseia nas condições KKT e no fato da Lagrangeana ser invexa em x^* no conjunto C. Assim, o mesmo argumento utilizado por Brandão em [19] pode ser utilizado sob as nossas hipóteses. □

A seguir estabeleceremos algumas relações de dualidade.

Associado ao problema (primal) (P), definiremos o problema dual (D):

$$\max_{\substack{u \in \mathbb{R}^m_+ \\ v \in \mathbb{R}^r}} \inf_{x \in C} \mathcal{L}(x, u, v). \tag{D}$$

As relações de dualidade entre (P) e (D) são especificadas na seguinte proposição:

Proposição 8.10 *(i)* ***Dualidade fraca:*** *Se x é factível para (P) e (u, v) é factível para (D), então*

$$f(x) \geq \inf_{y \in C} \mathcal{L}(y, u, v).$$

(ii) ***Dualidade forte:*** *Suponha que x^* é solução de (P) e que (P) é calmo em x^*. Suponha também que $((f, g), h)$ é invex-infine em x^* no conjunto C. Então o problema (D) também tem solução e os valores ótimos de (P) e de (D) coincidem.*

Demonstração. (i) É evidente (segue da factibilidade de x e de (u, v)).

(ii) Suponha que x^* é uma solução de (P) e que (P) é calmo em x^*. Pela Proposição 8.9, existem u^*, v^* tais que (x^*, u^*, v^*) é um ponto de sela de (P). Provaremos que (u^*, v^*) é solução de (D) e $\langle u^*, g(x^*)\rangle = 0$. Da condição de ponto de sela, segue

$$f(x^*) + \langle u, g(x^*)\rangle + \langle v, h(x^*)\rangle \leq f(x^*) + \langle u^*, g(x^*)\rangle + \langle v^*, h(x^*)\rangle, \forall u \in \mathbb{R}^m_+, \forall v \in \mathbb{R}^r$$

e, portanto,

$$\langle u, g(x^*)\rangle + \langle v, h(x^*)\rangle \leq \langle u^*, g(x^*)\rangle, \forall u \in \mathbb{R}^m_+, \forall v \in \mathbb{R}^r \tag{8.3}$$

e, tomando-se $u = 0, v = 0$ em (8.3), obtemos $0 \leq \langle u^*, g(x^*)\rangle$. Mas $g(x^*) \leqq 0$ e $u^* \geqq 0$, logo, temos $\langle u^*, g(x^*)\rangle = 0$. Por absurdo, suponha que (u^*, v^*) não é solução de (D). Então, existem $\widetilde{u} \in \mathbb{R}^m_+, \widetilde{v} \in \mathbb{R}^r$ tais que

$$\inf_{x \in C}[f(x) + \langle \widetilde{u}, g(x)\rangle + \langle \widetilde{v}, h(x)\rangle] > \inf_{x \in C}[f(x) + \langle u^*, g(x)\rangle + \langle v^*, h(x)\rangle].$$

Com maior razão, para cada $x \in C$ fixado,

$$f(x) + \langle \widetilde{u}, g(x)\rangle + \langle \widetilde{v}, h(x)\rangle > \inf_{x \in C}[f(x) + \langle u^*, g(x)\rangle + \langle v^*, h(x)\rangle]. \tag{8.4}$$

Da condição de ponto de sela, segue

$$f(x^*) + \langle u^*, g(x^*)\rangle + \langle v^*, h(x^*)\rangle \leq \inf_{x \in C}[f(x) + \langle u^*, g(x)\rangle + \langle v^*, h(x)\rangle]. \tag{8.5}$$

De (8.4) e (8.5), segue

$$f(x) + \langle \widetilde{u}, g(x)\rangle + \langle \widetilde{v}, h(x)\rangle > f(x^*) + \langle u^*, g(x^*)\rangle + \langle v^*, h(x^*)\rangle, \forall x \in C. \tag{8.6}$$

Em particular, tomando-se $x = x^*$ em (8.6), $\langle \widetilde{u}, g(x^*)\rangle > \langle u^*, g(x^*)\rangle = 0$, o que é absurdo, pois $\widetilde{u} \geqq 0$ e $g(x^*) \leqq 0$. Portanto, (u^*, v^*) é solução de (D). Resta provar que os valores ótimos de (P) e (D) coincidem. É claro que

$$f(x^*) = f(x^*) + \langle u^*, g(x^*)\rangle + \langle v^*, h(x^*)\rangle \geq \inf_{x \in C}[f(x) + \langle u^*, g(x)\rangle + \langle v^*, h(x)\rangle] \tag{8.7}$$

e da condição ponto de sela,

$$f(x^*) = f(x^*) + \langle u^*, g(x^*)\rangle + \langle v^*, h(x^*)\rangle \leq f(x) + \langle u^*, g(x)\rangle + \langle v^*, h(x)\rangle, \forall x \in C$$

e, tomando-se o ínfimo sobre $x \in C$,

$$\inf_{x \in C}[f(x) + \langle u^*, g(x)\rangle + \langle v^*, h(x)\rangle] \geq f(x^*). \tag{8.8}$$

Logo, (8.7) e (8.8) implicam que $f(x^*)$ é o valor ótimo de (D). □

Observação 8.4 *(i) Um resultado similar à Proposição 8.10 foi estabelecido por Brandão em [19], na ausência de restrições de igualdade. Entretanto o autor considera um dual de tipo Wolfe.*

(ii) A argumentação utilizada na prova da Proposição 8.10 é essencialmente a mesma utilizada por Geoffrion em [46], onde o autor considera problemas estáveis, convexos, não diferenciáveis e não considera restrições de igualdade.

8.1.3 Soluções eficientes e *calmness*

Nesta subseção, discutiremos as relações existentes entre as soluções Pareto de (MOP) e a propriedade de *calmness* de uma certa classe de problemas escalarizados.

Denotaremos por $X = \{x \in C : g(x) \leqq 0, h(x) = 0\}$ i.e, X é o conjunto factível de (MOP).

Novamente, $K = \{1, \cdots, p\}$, $I = \{1, \cdots, m\}$ e $J = \{1, \cdots, r\}$. Para cada $k_0 \in K$ e $u_0 \in \mathbb{R}_+^{p-1}$ fixados, definiremos o problema:

$$\begin{aligned} &\text{Minimizar } f_{k_0}(x) + \sum_{\substack{k \in K \\ k \neq k_0}} u_k f_k(x) \\ &\text{sujeito a: } x \in X. \end{aligned} \tag{$\mathrm{P}_{u_0}^{k_0}$}$$

Recordemos também que:

- **Solução propriamente Pareto**: $x^0 \in Q$ é uma solução propriamente Pareto se existe um escalar $M > 0$ tal que para cada i e cada $x \in Q$ satisfazendo $f_i(x) < f_i(x^0)$, existe um índice $i_0 \neq i$ tal que $f_{i_0}(x) > f_{i_0}(x^0)$ e $\frac{f_i(x) - f_i(x^0)}{f_{i_0}(x^0) - f_{i_0}(x)} \leq M$.

- **Solução k-Pareto** (Benson e Morin [10]) : Sejam $k \in \{1, \cdots, q\}$ e $x^0 \in Q$. Dizemos que x^0 é solução k- Pareto se para cada $x \in Q$ com $f_k(x) < f_k(x^0)$, existe um índice j tal que $f_j(x) > f_j(x^0)$.

- **Solução k-propriamente Pareto** (Benson e Morin [10]) : Sejam $k \in \{1, \cdots, q\}$ e $x^0 \in Q$. Dizemos que x^0 é solução k-propriamente Pareto se x^0 é solução k-Pareto e existe um escalar $M_k > 0$ tal que cada $x \in Q$ com $f_k(x) < f_k(x^0)$, existe um índice j tal que $f_j(x) > f_j(x^0)$ e $\frac{f_k(x) - f_k(x^0)}{f_j(x^0) - f_j(x)} \leq M_k$.

É evidente: x^0 é solução Pareto (propriamente Pareto) se e somente se x^0 é solução k-eficiente (respect. k-propriamente Pareto) para todo índice k.

Proposição 8.11 *Se existe $u_0 \in \mathbb{R}_+^{p-1}$ tal que x_0 é solução de ($P_{u_0}^{k_0}$) então x_0 é solução k_0-propriamente Pareto de (MOP).*

Demonstração. A demonstração segue a mesma argumentação feita por Benson e Morin ([10], Lemma 1 (i)). Entretanto, como o autor considera apenas restrições de desigualdade e não considera a restrição abstrata, nós apresentaremos a prova. Inicialmente, provaremos que x_0 é solução k_0-Pareto de (MOP). Sejam $U_{k_0} := \{k : u_k^0 > 0, k \neq k_0\}$ e x_0 uma solução de ($\mathrm{P}_{u_0}^{k_0}$). Se $U_{k_0} = \emptyset$, então $u_k = 0$ para todo $k \neq k_0$, logo x_0 minimiza f_{k_0} em X. Assim, não existe $x \in X$ tal que $f_{k_0}(x) < f_{k_0}(x_0)$. Por vacuidade, x_0 é solução k_0-Pareto de (MOP). Agora, suponha que $U_{k_0} \neq \emptyset$. Tome $x \in X$ tal que $f_{k_0}(x) < f_{k_0}(x_0)$. Sendo x_0 solução de ($\mathrm{P}_{u_0}^{k_0}$), temos

$$f_{k_0}(x) + \sum_{j \neq k_0} u_j^0 f_j(x) \geq f_{k_0}(x_0) + \sum_{j \neq k_0} u_j^0 f_j(x_0). \tag{8.9}$$

Mas: $f_{k_0}(x) - f_{k_0}(x_0) < 0$. Disto e de (8.9),

$$0 > \sum_{j \in U_{k_0}} u_j^0 [f_j(x^0) - f_j(x)]. \tag{8.10}$$

Temos também $u_j > 0, \forall j \in U_{k_0}$; assim sendo, (8.10) implica $f_j(x^0) - f_j(x) < 0$ para algum $j \in U_{k_0}$, e portanto, x^0 é solução k_0-Pareto de (MOP). □

A recíproca desta proposição é válida, sob hipóteses adequadas de convexidade generalizada. De fato:

Proposição 8.12 *Suponha que* $((f, g_{I_0}); h)$ *é invex-infine em* x_0 *no conjunto* C *e que a condição (CQ) vale em* x_0*. Se* $x_0 \in X$ *uma solução* k_0*-propriamente Pareto de (MOP), então existe* $u_0 \in \mathbb{R}_+^{p-1}$ *tal que* x_0 *é solução de* $(P_{u_0}^{k_0})$.

Demonstração. Seja x_0 uma solução k_0-propriamente Pareto de (MOP). Então, existe um escalar $M_{k_0} > 0$ tal que o seguinte sistema

$$\begin{cases} f_{k_0}(x) - f_{k_0}(x_0) < 0 \\ f_{k_0}(x) + M_{k_0} f_k(x) - f_{k_0}(x_0) - M_{k_0} f_k(x_0) < 0, \ k \in K, k \neq k_0 \\ g(x) \leqq 0 \\ h(x) = 0 \\ x \in C \end{cases}$$

não tem solução. Considere as funções

$$\begin{aligned} w_{k_0}(x) \ &: \ = f_{k_0}(x) - f_{k_0}(x_0) \\ w_k(x) \ &: \ = f_{k_0}(x) + M_{k_0} f_k(x) - f_{k_0}(x_0) - M_{k_0} f_k(x_0), \ k \neq k_0 \\ \rho(x) \ &: \ = \max\{\max_{k \in K} w_k(x), \max_{i \in I} g_i(x), \max_{j \in J} |h_j(x)|\}. \end{aligned}$$

Então: $\rho(x_0) = 0$, $K_0 = \{k \in K : w_i(x_0) = 0\}$, $I_0 = \{i \in I : g_i(x_0) = 0\}$, $J_0 = \{j \in J : |h_j(x_0)| = 0\} = J$. Das hipóteses feitas, $((w_K, g_{I_0}); h)$ é invex-infine em x_0 no conjunto C e o sistema $w(x) \leqq 0$, $g(x) \leqq 0$, $h(x) = 0$, $x \in C$ tem pelo menos uma solução (a saber, o próprio x_0). Aplicando-se o Teorema de Motzkin Generalizado (Teorema 8.6) ao sistema $w(x) < 0$, $g(x) \leqq 0$, $h(x) = 0$, $x \in C$ temos que existem escalares $\overline{\lambda}_k \geq 0$, $k \in K$, $\beta_i \geq 0$, $i \in I$ e γ_j, $j \in J$ tais que $\sum_{k \in K} \overline{\lambda}_k = 1$ e

$$\sum_{k \in K} \overline{\lambda}_k w_k(x) + \sum_{i \in I} \beta_i g_j(x) + \sum_{j \in J} \gamma_j h_j(x) \geq 0, \forall x \in C. \tag{8.11}$$

Mas, utilizando-se a mesma argumentação utilizada em [89] (na demonstração do Teorema 4.2) pode-se provar que $\beta_i g_i(x_0) = 0$, $\forall i \in I$. Assim, podemos rescrever (8.11),

$$\begin{aligned} \overline{\lambda}_{k_0}[f_{k_0}(x) - f_{k_0}(x_0)] + \sum_{k \neq k_0} \overline{\lambda}_k [f_{k_0}(x) + M_{k_0} f_k(x) - f_{k_0}(x_0) - M_{k_0} f_k(x_0)] \\ + \sum_{i \in I} \beta_i [g_i(x) - g_i(x_0)] + \sum_{j \in J} \gamma_j [h_j(x) - h_j(x_0)] \geq 0, \forall x \in C \end{aligned}$$

ou ainda (após rearranjo dos termos e utilizando-nos de $\beta_i g_i(x_0) = 0$, $\forall i \in I$ e $h_j(x_0) = 0$),

$$f_{k_0}(x) - f_{k_0}(x_0) + \sum_{k \neq k_0} M_{k_0} \overline{\lambda}_k [f_k(x) - f_k(x_0)] \geq - \sum_{i \in I} \beta_i g_i(x) - \sum_{j \in J} \gamma_j h_j(x), \ \forall x \in C.$$

Em particular, tomando-se $x \in X$ na desigualdade acima,

$$f_{k_0}(x) + \sum_{k \neq k_0} M_{k_0}\overline{\lambda}_k f_k(x) \geq f_{k_0}(x_0) + \sum_{k \neq k_0} M_{k_0}\overline{\lambda}_k f_k(x_0), \ \forall x \in X.$$

Basta, pois, escolher: $u_0 = (u_k^0)_{k \neq k_0}$, $u_k^0 = M_{k_0}\overline{\lambda}_k$ e o teorema está demonstrado. $\square$

Para $k_0 \in K$ e $b \in \mathbb{R}^{p-1}$ fixados, definiremos o problema:

$$\begin{array}{l} \text{Minimizar } f_{k_0}(x) \\ \text{sujeito a: } f_k(x) \leq b_k, \ k \neq k_0 \\ x \in X. \end{array} \qquad (\mathrm{P}_{b,k_0})$$

O seguinte resultado é quase imediato:

Lema 8.13 *Sejam $x_0 \in X$ uma solução k_0-Pareto de (MOP) e $\overline{b}_k = f_k(x_0)$, $k \neq k_0$. Então, x_0 é solução de $(P_{\overline{b},k_0})$.*

Demonstração. Se $x_0 \in X$ é solução k_0-Pareto de (MOP) e $\overline{b}_k = f_k(x_0)$, $k \neq k_0$, em particular, x_0 é factível para $(\mathrm{P}_{\overline{b},k_0})$. Por absurdo, suponha que x_0 não é solução de $(\mathrm{P}_{\overline{b},k_0})$. Então, existe $x \in X$ tal que $f_k(x) \leq f_k(x_0)$, $k \neq k_0$ e $f_{k_0}(x) < f_{k_0}(x_0)$, o que contraria a k_0-eficiência de x_0. $\square$

O seguinte teorema nos fornece uma condição suficiente para que uma solução k_0-Pareto de (MOP) seja uma solução k_0-propriamente Pareto de (MOP).

Teorema 8.14 *Suponha que $((f,g);h)$ é invex-infine em $x_0 \in X$ no conjunto C. Suponha também que x_0 é uma solução k_0-Pareto de (MOP). Se $(P_{\overline{b},k_0})$ é calmo em x_0, então x_0 é solução k_0-propriamente Pareto de (MOP).*

Demonstração. Pela Proposição 8.10 (ii), sob nossas hipóteses, o problema dual de $(\mathrm{P}_{\overline{b},k_0})$ também admite uma solução $(u^0, v^0) \in (\mathbb{R}_+^{p-1} \times \mathbb{R}_+^m) \times \mathbb{R}^r$ e os valores ótimos coincidem, isto é,

$$f_{k_0}(x_0) = \max_{\substack{u=(u_1,u_2) \in \mathbb{R}_+^{p-1} \times \mathbb{R}_+^m \\ v \in \mathbb{R}^r}} \inf_{x \in C} [f_{k_0}(x) + \sum_{k \neq k_0} (u_1)_k (f_k(x) - f_k(x_0)) + \sum_{i \in I} (u_2)_i g_i(x) + \sum_{j \in J} v_j h_j(x)]$$

(onde $(u_1)_k$ denota a k-ésima componente de u_1 e similarmente para $(u_2)_i$ e v_j). Assim sendo, existem $u_0 = (u_1^0, u_2^0) \in \mathbb{R}_+^{p-1} \times \mathbb{R}_+^m$, $v_0 \in \mathbb{R}^r$, tais que

$$f_{k_0}(x_0) = \inf_{x \in C} [f_{k_0}(x) + \sum_{k \neq k_0} (u_1^0)_k (f_k(x) - f_k(x_0)) + \sum_{i \in I} (u_2^0)_i g_i(x) + \sum_{j \in J} v_j^0 h_j(x)]$$

e, com maior razão,

$$f_{k_0}(x_0) \leq f_{k_0}(x) + \sum_{k \neq k_0} (u_1^0)_k (f_k(x) - f_k(x_0)) + \sum_{i \in I} (u_2^0)_i g_i(x) + \sum_{j \in J} v_j^0 h_j(x), \forall x \in X.$$

Mas $g(x) \leqq 0$ e $h(x) = 0$, para todo $x \in X$, donde segue

$$f_{k_0}(x_0) = \min_{x \in X} f_{k_0}(x) + \sum_{k \neq k_0} (u_1^0)_k f_k(x).$$

Logo, pela Proposição 8.11, x_0 é solução k_0-propriamente Pareto de (MOP).

□

Uma espécie de recíproca deste resultado é válido, na ausência de restrições de igualdade e de restrições abstratas.

Para demonstrá-lo, necessitaremos do seguinte resultado auxiliar:

Lema 8.15 *Suponha que as funções $f, g_1, \cdots, g_m : \mathbb{R}^n \to \mathbb{R}$ são localmente Lipschitz, f é inf-compacta*[3] *e as funções g_i, $i \in I$ são invexas em x_0, com respeito a uma mesma η. Se existe $\overline{y} \in \mathbb{R}^n$ tal que $g_i(\overline{y}) < 0, i \in I(x_0) = \{i : g_i(x_0) = 0\}$, então o problema parametrizado*

$$\begin{array}{l} \text{Minimizar } f(y) \\ \text{sujeito a: } g_i(y) \leq p_i, \ i = 1, \cdots, m \end{array} \qquad (\widetilde{\mathrm{P}}(p))$$

é calmo.

Demonstração. Sob as hipóteses do Teorema, a função valor $V(p)$ é finita perto de $p = 0$. A ideia da demonstração é mostrar que $\dfrac{V(p) - V(0)}{||p||}$ é limitada inferiormente quando $p \to 0$. Para isto, basta considerarmos p tal que $V(p) < V(0)$ e é suficiente considerar o caso $p \geqq 0$. Tome uma sequencia $(p_k)_{k \in N}$, com $p_k \to 0^+$. Podemos escolher $(y_k)_{k \in N}$, tal que $g_i(y_k) \leqq (x_k)_i$ e $f(y_k) \leq V(p_k) + ||p_k||^2$. Além disto, (y_k) é limitada e, sem perda de generalidade, podemos assumir $y_k \to y_\infty$. Também temos

$$f(y_k) \leq V(p_k) + ||p_k||^2 \leq V(0) + ||p_k||^2. \tag{8.12}$$

Sejam $\widehat{I}(y_k) := \{i : g_i(y_k) \geq 0\}$, $\eta_k = \eta(\overline{y}, y_k)$. Segue da invexidade e da condição de Slater, $0 > g_i(\overline{y}) - g_i(y_k) \geq (g_i)^0(y_k; \eta(\overline{y}, y_k)) = \limsup_{\substack{y \to y_k \\ t \to 0^+}} \dfrac{g_i(y + t\eta_k) - g_i(y_k)}{t}, \forall i \in \widehat{I}(y_k)$. Das propriedades do limsup, temos que existe $\delta_k > 0$ tal que $g_i(y_k + t\eta_k) < 0$, $\forall t \in (0, \delta_k)$, $i \in \widehat{I}(y_k)$. Da continuidade das g_i, segue que podemos escolher $\varepsilon_k \in (0, \delta_k)$ tal que $g_i(y_k + t\eta_k) \leq 0$, $\forall t \in (0, \varepsilon_k)$, $i = 1, \cdots, m$. Seja $t_k = \min\{\varepsilon_k, \frac{||p_k||}{||\eta_k||}\}$. Então,

$$V(0) \leq f(y_k + t_k\eta_k). \tag{8.13}$$

Seja K a constante de Lipschitz de f perto de y_∞. Temos,

$$f(y_k + t_k\eta_k) - f(y_k) \leq K t_k ||\eta_k|| \leq K||p_k||. \tag{8.14}$$

De (8.13)-(8.14),

$$V(p_k) - V(0) \geq f(y_k) - f(y_k + t_k\eta_k) - ||p_k||^2 \geq -K||p_k|| - ||p_k||^2$$

e portanto,

$$\frac{V(p_k) - V(0)}{||p_k||} \geq -K - ||p_k||$$

e, sendo $p_k \to 0^+$ arbitrária, temos que

$$\liminf_{p \to 0} \frac{V(p) - V(0)}{||p||} > -\infty$$

[3] f é inf-compacta se $\{x : f(x) \leq r\}$ é compacto para todo $r \in \mathbb{R}$.

e portanto o problema é calmo.

□

Considere o problema multiobjetivo

$$\begin{array}{l} \text{Minimizar } f(x) := (f_1(x), \cdots, f_p(x)) \\ \text{sujeito a: } g_i(x) \leq 0,\ i = 1, \cdots, m \end{array} \qquad (\widetilde{\text{MOP}})$$

e o problema

$$\begin{array}{l} \text{Minimizar } f_{k_0}(x) \\ \text{sujeito a: } f_k(x) \leq b_k,\ k \neq k_0 \\ x \in \widetilde{X} := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i = 1, \cdots, m\} \end{array} \qquad (\widetilde{\text{P}}_{b,k_0})$$

Teorema 8.16 *Suponha que as funções f_k, g_i $(k = 1, \cdots, p, i = 1, \cdots, m)$ são localmente Lipschitz em C e invexas em x_0 no conjunto C com respeito a uma mesma η e que a função f_{k_0} é inf-compacta. Seja $x_0 \in \widetilde{X}$ uma solução k_0-Pareto de $(\widetilde{MOP})$. Suponha que o problema $(\widetilde{P}_{\overline{b},k_0})$ satisfaz a condição de Slater, isto é, existe $\overline{x} \in C$ tal que*

$$\begin{array}{c} f_k(\overline{x}) - f_k(x_0) < 0, k \neq k_0 \\ g_i(\overline{x}) < 0,\ i \in I(x_0). \end{array}$$

Então, x_0 é solução k_0-propriamente Pareto se e somente se $(\widetilde{P}_{\overline{b},k_0})$ é calmo em x_0.

Demonstração. É consequencia imediata do Teorema 8.14 e do Lema 8.15. □

Observação 8.5 *Seria interessante investigar se Lema 8.15 poderia ser estendido ao caso geral, onde o conjunto factível de (MOP) contém também restrições de igualdade e restrições abstratas. Deste modo, não conseguimos através do Lema 8.15 estudar o problema de controle ótimo discreto, uma vez que tais problemas, quando reformulados como problemas de programação matemática, contém também restrições de igualdade - dadas pela equação de movimento.*

8.2 Problemas de Controle Ótimo Discreto Não-Diferenciáveis

Em muitas aplicações a hipótese de diferenciabilidade nos dados do problema pode ser muito restritiva e além disso existem inúmeros mecanismos naturais que geram não suavidade mesmo quando se parte de problemas diferenciáveis.

Por este motivo, esta seção aborda o estudo de problemas multiobjetivos de controle ótimo discreto não diferenciáveis. Nossos resultados aqui estabelecidos generalizam os obtidos por Shvartsman [91] e Mordukhovich e Shvartsman [74] para problemas multiobjetivos.

Nesta Seção obteremos uma prova simplificada para o Princípio do Máximo Discreto para uma certa classe de problemas de controle ótimo discretos multiobjetivos não-diferenciáveis.

Consideramos o seguinte problema:

$$\begin{array}{l} \text{Minimizar } (\Phi_0(x_1,\cdots,x_N),\cdots,\Phi_{r_0}(x_1,\cdots,x_N)) \\ \text{sujeito a:} \\ x_{i+1}=\varphi_i(x_i,u_i),\ i=0,\cdots,N-1 \\ x_0=\mathbf{x}_0 \\ \Phi_j(x_1,\cdots,x_N)\le 0,\ j=r_0+1,\cdots,r_0+r_1 \\ \Phi_j(x_1,\cdots,x_N)=0,\ j=r_0+r_1+1,\cdots,r_0+r_1+r_2 \\ u_i\in U_i,\ i=0,\cdots,N-1 \end{array} \tag{PND}$$

onde $\varphi : [0, N-1]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $\Phi_j:\mathbb{R}^{n(N)}\to\mathbb{R}$, $j=0,\cdots,r_0+r_1+r_2$ são funções dadas e $U_i\subset\mathbb{R}^m, i\in[0,N-1]$, são subconjuntos não vazios de $\mathbb{R}^m$.

Neste caso, a otimalidade será entendida no sentido da eficiência fraca. Mais precisamente, um processo admissível para (PND), $(\hat{x},\hat{u})$ é um processo Pareto fraco se não existe (x,u) admissível para (PND) tal que $\Phi_j(x_1,\cdots,x_N)<\Phi_j(\hat{x}_1,\cdots,\hat{x}_N)$, para todo $j=0,\cdots,r_0$.

O Princípio do Máximo Discreto para o problema não diferenciável (PND) é estabecido no seguinte teorema:

Teorema 8.17 *Suponha que $\varphi_i(\cdot,\cdot)$ é contínua ao redor do processo Pareto fraco $(\hat{x},\hat{u})$ de (PND); as funções $\varphi_i(\cdot,\hat{u}_i)$ são diferenciáveis em $\hat{x}_i$, para todo $i\in[0,N-1]$; as funções Φ_j, $j=0,\cdots,r_0+r_1+r_2$ são Lipschitz em $(\hat{x}_1,\cdots\hat{x}_N)$ e os conjuntos $\varphi_i(\hat{x}_i,U_i)$ são convexos, para cada $i\in[0,N-1]$. Então, existem*

1. *$\lambda_j\in\mathbb{R}$, $j=0,\cdots,r_0+r_1+r_2$, não todos nulos e tais que*

$$\begin{array}{l} \lambda_j\ge 0,\ j=0,\cdots,r_0+r_1; \\ \lambda_j\Phi_j(\hat{x}_1,\cdots,\hat{x}_N)=0,\ j=r_0+1,\cdots,r_0+r_1. \end{array} \tag{8.15}$$

2. *$\bar{z}^j=(\bar{z}_1^j,\cdots,\bar{z}_N^j)\in\mathbb{R}^N$, $\bar{z}^j\in\partial(\lambda_j\Phi_j)(\hat{x}_1,\cdots\hat{x}_N)$, $j=0,\cdots,r_0+r_1+r_2$;*

3. *$p\in R^{n(N+1)}$ satisfaz a equação adjunta*

$$\begin{array}{l} p_i=\nabla_x\mathbb{H}_i(p_{i+1},\hat{x}_i,\hat{u}_i)-\displaystyle\sum_{j=0}^{r_0+r_1+r_2}\bar{z}_i^j,\ i=0,\cdots,N-1 \\ p_N=-\displaystyle\sum_{j=0}^{r_0+r_1+r_2}\bar{z}_N^j \end{array} \tag{8.16}$$

(sendo $\mathbb{H}_i(p,x,u):=\langle p_{i+1},\varphi_i(x_i,u_i)\rangle$). Além disto,

$$\mathbb{H}_i(p_{i+1},\hat{x}_i,\hat{u}_i)=\max_{u_i\in U_i}\mathbb{H}_i(p_{i+1},\hat{x}_i,u_i), \tag{8.17}$$

para cada $i\in[0,N-1]$.

A demonstração do Teorema 8.17 é feita mediante a reformulação de (PND) como um problema de programação matemática multiobjetivo, ao qual aplicaremos a Regra de Multiplicadores de Lagrange; também utilizaremos a versão da Regra da Cadeia, para funções localmente Lipschitz, em termos dos gradientes generalizados de Clarke.

Recordemos tais resultados.

Lema 8.18 ***(Regra da Cadeia)*** *Sejam $g : \mathbb{R}^n \to \mathbb{R}^k$ uma função estritamente diferenciável em $\hat{y}$ e $f : \mathbb{R}^k \to \mathbb{R}$ uma função Lipschitz perto de $\hat{z} = g(\hat{y})$. Então,*

$$\partial(f \circ g)(\hat{y}) \subset \bigcup_{\bar{z} \in \partial f(\hat{z})} \langle \bar{z}, \nabla g(\hat{y}) \rangle.$$

A demonstração do Lema 8.18 pode ser encontrada em [26] (v. Teorema 2.3.10, p. 45).

Consideremos o problema de programação matemática multiobjetivo:

$$\begin{array}{l} \text{Minimizar } (f_0(y), \cdots, f_{r_0}(y)) \\ \text{sujeito a:} \\ f_j(y) \leq 0, \; j = r_0 + 1, \cdots, r_0 + r_1 \\ f_j(y) = 0, \; j = r_0 + r_1 + 1, \cdots, r_0 + r_1 + r_2 \\ y \in \Omega \end{array} \tag{8.18}$$

onde $f_j : \mathbb{R}^n \to \mathbb{R}$, $j = 0, \cdots, r_0 + r_1 + r_2$ e Ω é um subconjunto não vazio de $\mathbb{R}^n$.

As condições necessárias de Fritz-John para o problema não diferenciável (8.18) via gradientes generalizados de Clarke são dadas no seguinte lema:

Lema 8.19 *Suponha que $\hat{y}$ é uma solução Pareto fraca do problema (8.18). Suponha ainda que Ω é um subconjunto convexo, fechado e não vazio de $\mathbb{R}^n$ e que as funções f_j, $j = 0, \cdots, r_0 + r_1 + r_2$ são Lipschitz perto de $\overline{y}$. Então existem números λ_j, $i = 0, \cdots, r_0 + r_1 + r_2$, não todos nulos e tais que*

$$\begin{array}{l} 0 \in \displaystyle\sum_{i=0}^{r_0+r_1+r_2} \partial(\lambda_j f_j)(\hat{y}) + N_\Omega(\hat{y}) \\ \lambda_j \geq 0, \; j = 0, \cdots, r_0 + r_1 \\ \lambda_j f_j(\hat{y}) = 0, \; j = r_0 + 1, \cdots, r_0 + r_1 \end{array}$$

(onde $N_\Omega(\hat{y})$ é o cone normal (usual da Análise Convexa)).

O Lema 8.19 é um caso particular do Teorema 6.1.3, p. 230 de [26]. De fato, conforme mencionamos na Observação 8.1, o cone normal usual da Análise Convexa e o cone normal de Clarke, coincidem quando o conjunto é convexo.

No que segue, apresentamos a demonstração do Teorema 8.17.

Demonstração. [**do Teorema 8.17**] Fixado $\theta \in [0, N-1]$ e $v \in U_\theta$, considere o controle

$$u_i = \begin{cases} \hat{u}_i, \; i \neq \theta \\ v, \; i = \theta. \end{cases}$$

Sejam: $y = \varphi_\theta(\hat{x}_\theta, v)$; $\hat{y} = \varphi_\theta(\hat{x}_\theta, \hat{u}_\theta)$ e x^y a trajetória correspondente ao controle u.

Como $(\hat{x}, \hat{u})$ é um processo ótimo Pareto fraco (PND), então $y = \hat{y}$ é solução Pareto fraca do seguinte problema multiobjetivo:

$$\begin{array}{l} \text{Minimizar } (\Phi_0(x_1^y, \cdots, x_N^y), \cdots, \Phi_{r_0}(x_1^y, \cdots, x_N^y)) \\ \text{sujeito a:} \\ \Phi_j(x_1^y, \cdots, x_N^y) \leq 0, \; j = r_0 + 1, \cdots, r_0 + r_1 \\ \Phi_j(x_1^y, \cdots, x_N^y) = 0, \; j = r_0 + r_1 + 1, \cdots, r_0 + r_1 + r_2 \\ y \in \varphi_\theta(\hat{x}_\theta, U_\theta). \end{array} \tag{8.19}$$

Aplicando o Lema 8.19 ao problema (8.19), obtemos que existem $\lambda_j \in \mathbb{R}$, $j = 0, \cdots, r_0 + r_1 + r_2$, tais que $\lambda_j \geq 0$, $j = 0, \cdots, r_0 + r_1$ e

$$0 \in \sum_{j=0}^{r_0+r_1+r_2} \partial_y(\lambda_j \Phi_j)(x_1^{\hat{y}}, \cdots, x_N^{\hat{y}}) + N_{\varphi_\theta(\hat{x}_\theta, U_\theta)}(\hat{y}) \tag{8.20}$$

(onde, é claro, $x^{\hat{y}} = \hat{x}$).

A função $y \mapsto x_i^y$ $(i = 0, \cdots, N)$ é diferenciável em $\hat{y}$. Neste caso, a derivada é dada por

$$\nabla_y x_i^{\hat{y}} = \begin{cases} 0, \ i \leq \theta \\ f(\theta+1, i), \ i \geq \theta + 1 \end{cases}$$

sendo

$$f(s,i) = \begin{cases} \mathbf{I}, \text{ se } s = i \\ \nabla_x \varphi_{i-1}(\hat{x}_{i-1}, \hat{u}_{i-1}) \cdot \ldots \cdot \nabla_x \varphi_s(\hat{x}_s, \hat{u}_s), \ s < i \end{cases}$$

onde $\mathbf{I}$ é a matriz identidade $n \times n$. Para demonstrá-lo, basta provar que para $y \in \varphi_i(\hat{x}_i, U_i)$, tem-se

$$x_i^y - x_i^{\hat{y}} = x_i^y - \hat{x}_i = \nabla_y x_i^{\hat{y}}(y - \hat{y}) + o(||y - \hat{y}||) \tag{8.21}$$

para todo $i \geq \theta + 1$ e isto pode ser demonstrado indutivamente. Com efeito:

Para $i = \theta + 1$:

$$x_{\theta+1}^y - \hat{x}_{\theta+1} = \varphi_\theta(\hat{x}_\theta, v) - \varphi_\theta(\hat{x}_\theta, \hat{u}_\theta) = y - \hat{y} = \mathbf{I}(y - \hat{y}) = f(\theta+1, \theta+1)(y - \hat{y})$$

e portanto (8.21) se verifica trivialmente em $i = \theta + 1$.

Para $i = \theta + 2$: Sendo $\varphi_{\theta+1}(\cdot, \hat{u}_{\theta+1})$ diferenciável em $\hat{x}_{\theta+1}$, temos

$$\begin{aligned} x_{\theta+2}^y - \hat{x}_{\theta+2} &= \varphi_{\theta+1}(x_{\theta+1}^y, \hat{u}_{\theta+1}) - \varphi_{\theta+1}(\hat{x}_{\theta+1}, \hat{u}_{\theta+1}) \\ &= \underbrace{\nabla_x \varphi_{\theta+1}(\hat{x}_{\theta+1}, \hat{u}_{\theta+1})}_{f(\theta+1,\theta+2)}(y - \hat{y}) + o(||y - \hat{y}||) \\ &= f(\theta+1, \theta+2)(y - \hat{y}) + o(||y - \hat{y}||) \end{aligned}$$

e portanto (8.21) se verifica em $i = \theta + 2$.

Em $i = \theta + 3$: Sendo $\varphi_{\theta+2}(\cdot, \hat{u}_{\theta+2})$ diferenciável em $\hat{x}_{\theta+2}$, temos

$$\begin{aligned} x_{\theta+3}^y - \hat{x}_{\theta+3} &= \varphi_{\theta+2}(x_{\theta+2}^y, \hat{u}_{\theta+2}) - \varphi_{\theta+2}(\hat{x}_{\theta+2}, \hat{u}_{\theta+2}) \\ &= \nabla_x \varphi_{\theta+2}(\hat{x}_{\theta+2}, \hat{u}_{\theta+2})(y - \hat{y}) + o(||y - \hat{y}||) \\ &= f(\theta+2, \theta+3)(y - \hat{y}) + o(||y - \hat{y}||) \end{aligned}$$

e novamente, (8.21) se verifica em $i = \theta + 3$. (E, indutivamente, (8.21) se verifica em $i \geq \theta + 1$.) De (8.20) segue que existem vetores $\xi \in N_{\varphi_\theta(\hat{x}_\theta, U_\theta)}(\hat{y})$ e $\bar{y}^j \in \partial_y(\lambda_j \Phi_j)(x_1^{\hat{y}}, \cdots, x_N^{\hat{y}})$, $j = 0, \cdots, r_0 + r_1 + r_2$ e tais que

$$0 = \bar{\xi} + \sum_{j=0}^{r_0+r_1+r_2} \bar{y}^j \tag{8.22}$$

e, pela definição de cone normal,

$$\langle \bar{\xi}, y - \hat{y} \rangle \leq 0, \forall y \in \varphi_\theta(\hat{x}_\theta, U_\theta) \tag{8.23}$$

e portanto, de (8.22) e (8.23), temos

$$\sum_{j=0}^{r_0+r_1+r_2} \langle \bar{y}^j, y-\hat{y}\rangle \geq 0,\ \forall y \in \varphi_\theta(\hat{x}_\theta, U_\theta). \tag{8.24}$$

Aplicando-se a Regra da Cadeia (Lema 8.18) à função composta $g_j(y) = \Phi_j(x^y)$:

$$\begin{aligned}
\langle \bar{y}^j, y-\hat{y}\rangle &= \langle \bar{z}^j, (\nabla_y x_1^{\hat{y}}, \cdots, \nabla_y x_N^{\hat{y}})(y-\hat{y})\rangle \\
&= \bar{z}_1^j \nabla_y x_1^{\hat{y}}(y-\hat{y}) + \cdots + \bar{z}_N^j \nabla_y x_N^{\hat{y}}(y-\hat{y}) \\
&= \sum_{s=\theta+1}^{N} \langle \bar{z}_s^j, f(\theta+1, s)(y-\hat{y})\rangle,\ j = 0, \cdots, r_0+r_1+r_2.
\end{aligned} \tag{8.25}$$

Defina:

$$p_i := -\sum_{j=0}^{r_0+r_1+r_2} \sum_{s=i}^{N} \bar{z}_s^j f(i,s). \tag{8.26}$$

Segue de (8.24) e (8.25),

$$\langle p_{\theta+1}, y-\hat{y}\rangle = \langle p_{\theta+1}, \varphi_\theta(\overline{x}_\theta, v) - \varphi_\theta(\hat{x}_\theta, \hat{u}_\theta)\rangle \leq 0$$

que é a afirmação do Princípio de Máximo dada por (8.17). Resta provar que p_i definido em (8.26) satisfaz a equação adjunta (8.16). Isto será verificado indutivamente onde admitiremos, por simplicidade, que $\varphi_i(\hat{x}_i, \hat{u}_i) = \varphi_i$.

De fato, tomando-se $i = N$ em (8.26):

$$p_N = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_N^j \underbrace{f(N,N)}_{=\mathbf{I}} = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_N^j.$$

Agora, faça $i = N-1$ em (8.26):

$$\begin{aligned}
p_{N-1} &= -\sum_{j=0}^{r_0+r_1+r_2} \left[\bar{z}_{N-1}^j \underbrace{f(N-1,N-1)}_{=\mathbf{I}} + \bar{z}_N^j \underbrace{f(N-1,N)}_{=\nabla_x \varphi_{N-1}} \right] \\
&= -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-1}^j - \underbrace{\sum_{j=0}^{r_0+r_1+r_2}}_{=-p_N} \bar{z}_N^j \nabla_x \varphi_{N-1} \\
&= -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-1}^j + \langle p_N, \nabla_x \varphi_{N-1}\rangle = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-1}^j + \nabla_x H_{N-1}(p_N, \hat{x}_{N-1}, \hat{u}_{N-1})
\end{aligned}$$

e portanto (8.16) vale para $i = N-1$.

Tomemos agora $i = N-2$ em (8.26):

$$\begin{aligned}
p_{N-2} &= -\sum_{j=0}^{r_0+r_1+r_2} \left[\bar{z}_{N-2}^j \underbrace{f(N-2,N-2)}_{=\mathbf{I}} + \bar{z}_{N-1}^j \underbrace{f(N-2,N-1)}_{=\nabla_x \varphi_{N-2}} + \bar{z}_N^j \underbrace{f(N-2,N)}_{=\nabla_x \varphi_{N-1} \nabla_x \varphi_{N-2}} \right] \\
&= -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}_{N-2}^j + \nabla_x \varphi_{N-2} \left[\bar{z}_{N-1}^j + \bar{z}_N^j \nabla_x \varphi_{N-1} \right].
\end{aligned} \tag{8.27}$$

Mas:

$$-\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-2}+\nabla_x\varphi_{N-2}\left[\bar{z}^j_{N-1}+\bar{z}^j_N\nabla_x\varphi_{N-1}\right] = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-1}+\langle p_{N-1}\nabla_x\varphi_{N-2}\rangle\,. \quad (8.28)$$

De fato, segue de (8.26)

$$p_{N-1} = -\sum_{j=0}^{r_0+r_1+r_2}\sum_{s=N-1}^{N} \bar{z}^j_s f(i,s) = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-1}+\bar{z}^j_N\nabla_x\varphi_{N-1} \quad (8.29)$$

e de (8.29),

$$\begin{aligned} &-\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-1}\nabla_x\varphi_{N-2}-\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_N\nabla_x\varphi_{N-1}\nabla_x\varphi_{N-2} \\ = &\left[-\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-1}-\sum_{j=0}^{r_0+r_1+r_2}\nabla_x\varphi_{N-1}\right]\nabla_x\varphi_{N-2} \\ = &\ \langle p_{N-1},\nabla_x\varphi_{N-2}\rangle\,. \end{aligned} \quad (8.30)$$

Comparando (8.27) e (8.30), obtemos

$$p_{N-2} = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-2}+\langle p_{N-1}\nabla_x\varphi_{N-2}\rangle = -\sum_{j=0}^{r_0+r_1+r_2} \bar{z}^j_{N-2}+\nabla_x H_{N-2}(p_{N-1},\hat{x}_{N-2},\hat{u}_{N-2}).$$

E, indutivamente, p_i definida em (8.26) satisfaz a equação adjunta (8.16), para todo $i\in[0,N-1]$. □

Capítulo 9

Conclusões e possibilidades de prosseguimento deste trabalho

Neste trabalho, estudamos os problemas de controle ótimo discreto com um e com vários objetivos. Apresentamos uma prova alternativa para o Princípio do Máximo Discreto para o caso mono-objetivo bem como uma generalização deste para o caso multi-objetivo, utilizando para isto o Formalismo de Dubovitskii-Milyutin. Também estudamos o problema não-regular, através do conceito da 2-regularidade. Investigamos ainda a suficiência das condições de otimalidade obtida, através de noções adequadas de convexidade. Nossos resultados referentes à suficiência das condições de otimalidade foram obtidos no sentido de se caracterizar a maior classe de problemas de controle discreto para as quais o Princípío do Máximo Discreto é necessário e também suficiente para a otimalidade. Além disto, analisamos a estabilidade de problemas de controle ótimo discreto. Ainda, obtivemos uma demonstração para o Princípío do Máximo Discreto para uma certa classe de problemas de controle ótimo discreto não diferenciável, utilizando a noção de gradientes e derivadas generalizadas de Clarke.

No que se segue, apontamos algumas possibilidades para o prosseguimento deste trabalho.

- **Estudo de problemas de controle ótimo discreto irregulares**: No artigo [93], os autores descrevem a teoria da p-regularidade, a qual generaliza a noção de 2-regularidade tratada nesta Tese. O resultado fundamental desta teoria é a descrição detalhada do conjunto $\{x : F(x) = 0\}$, onde F é um operador não linear, não regular, definido entre espaços de Banach. Um interessante resultado é que, sob hipóteses de p-regularidade, é possível caracterizar o cone tangente a este conjunto, através de uma generalização do Teorema de Lysternik. Esta ferramenta tem sido utilizada com sucesso na Programação Não Linear - tanto do ponto de vista teórico [94] quanto numérico [20], no estudo de problemas de Cálculo Variacional [95], em Sistemas Dinâmicos [21], dentre outras áreas. Uma possibilidade seria estudar problemas de controle ótimo discretos, utilizando a teoria da p-regularidade.

- **Estudo de problemas de controle ótimo discreto sob incerteza**: A Análise Intervalar (isto é, o estudo analítico de funções a valores intervalares) é uma importante área de estudo, seja do ponto de vista teórico ou aplicado. Foi introduzida como uma ferramenta utilizada no tratamento de incertezas em modelos matemáticos e computacionais. Veja Moore [72], por exemplo. Na Otimização, também tem sido empregada com sucesso. Veja [48, 97]. Recentemente, Costa *et al.* [32] introduziram um método através do qual é possível dotar o espaço de intervalos generalizados com algumas estruturas, como por exemplo, operações algébricas (com as quais este espaço se torna espaço vetorial), relações de ordem e cálculo algébrico. Com estes conceitos, os autores formulam problemas de otimização intervalares, relacionando-os com problemas (clássicos) de otimização multiobjetivo. Uma possibilidade seria tratar de problemas de controle ótimo discreto com incerteza, através de funções intervalares, utilizando para isto os resultados do artigo [32].

- **Estudo de estabilidade para problemas de controle discreto não diferenciáveis**: As relações existentes entre estabilidade e qualificações de restrição são bastante conhecidas e foram elucidadas por Robinson em [84, 85] e por Geoffrion [46] no âmbito da programação matemática convexa. O conceito de *calmness* foi introduzido por Clarke em [28] é uma interessante generalização do conceito de estabilidade, em contexto não convexo. Veja também [29]. O estudo das propriedades de *calmness* para problemas de controle ótimo discreto é uma questão totalmente em aberto.

- **Estudo de problemas de controle discreto com horizonte de tempo infinito**: Os primeiros trabalhos em controle ótimo com horizonte de tempo infinito são devidos a Pontryagin *et al.* [81] e publicados em 1974. Desde então, tais problemas têm sido extensivamente estudados devido à sua importância, sobretudo, na Teoria de Crescimento Econômico. Veja [2], por exemplo. As suas variantes em controle ótimo discreto também têm sido investigadas. Boltyanskii demonstra em [15] que, para problemas de controle discreto com horizonte de tempo infinito, o Princípio de Máximo Discreto é válido sob hipóteses bastante restritivas de convexidade e de invertibilidade de um certo operador. Mais recentemente, Blot e Hayek [18] obtêm uma versão do Princípio do Máximo Discreto para tais problemas, com hipóteses enfraquecidas de convexidade e sem a hipótese de invertibilidade do operador. Neste trabalho, os autores consideram problemas cuja dinâmica é dada pela equação de movimento e uma condição no estado inicial. Esta abordagem é descrita de maneira detalhada na dissertação de J. Beran [12]. Uma possibilidade seria a de estender os resultados de Blot e Hayek para problemas de controle discreto multiobjetivos; outra possibilidade seria a de adaptar a abordagem destes autores para problemas com restrições mais gerais - por exemplo, incorporando-se restrições de igualdades e/ou desigualdades a estes problemas.

Referências Bibliográficas

[1] Alt, W. "Stability of solutions to control constrained nonlinear optimal control problems." *Applied Mathematics & Optimization*, v. 21, n.1, p. 53-68, 1990.

[2] Araujo, A., e Scheinkman, J. A. " Smoothness, comparative dynamics, and the turnpike property." *Econometrica*, v. 45, n. 3, p. 601–620, 1977.

[3] Arutyunov, A. V. *Optimality conditions: Abnormal and degenerate problems.* Springer Science & Business Media, v. 526, 2000.

[4] Arutyunov, A V. e Alexey F. Izmailov. "Abnormal equality-constrained optimization problems: sensitivity theory." *Mathematical programming*, v. 100, n. 3, p. 485-515, 2004.

[5] Arutyunov, A. V. e Marinkovic, B. Necessary conditions of optimality for discrete optimal control problems. *University Computational Mathematics and Cybernetics*, Moscow, v.1, p. 38-44, 2005.

[6] Avakov, E. R. "Conditions for an extremum for smooth problems with constraints of equality type." *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, v.25, n.5, p. 680-693, 1985.

[7] Avakov, E. R, Arutyunov, A. V. e A. F. Izmailov. "Necessary conditions for an extremum in a mathematical programming problem." *Proceedings of the Steklov institute of mathematics*, v. 256, n. 1, p. 2-25, 2007.

[8] Alekseev, V. M. *Optimal control.* 1987.

[9] Bellman, R. E. *Dynamic programming.* Mineola, NY: Dover Publications, 2003.

[10] Benson, H. P, Morin, T. L." The Vector Maximization Problem: Proper Efficiency and Stability". *SIAM Journal on Applied Mathematics*, v. 32, n. 1, p. 64-72, 1977.

[11] Ben-Tal, A., e Zowe J. A unified theory of first and second order conditions for extremum problems in topological vector spaces. *Springer Berlin Heidelberg*, 1982.

[12] Beran, J. *Maximum principle for infinite horizon discrete time optimal control problems.* Dissertação de Mestrado. Comenius University in Bratislava (Eslováquia), 2011.

[13] Bigi, G., e Pappalardo, M. "Regularity conditions in vector optimization." *Journal of Optimization Theory and Applications*, v. 102, n.1, p. 83-96, 1999.

[14] Boltyanskii, V.G. *Optimal control of discrete systems*. Halsted Press, 1978.

[15] Boltianski, V. *Commande Optimale des Systemes Discrets*. Traduit du Russe par A. Sossinsk, Mir, Moscow, Russia, 1976.

[16] Bonnans, J. Frédéric, e Alexander Shapiro. "Optimization problems with perturbations: A guided tour." *SIAM review*, v. 40, n. 2, p. 228-264, 1998.

[17] Bonnans, J. F., e Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2000.

[18] Blot, J., e Hayek, N. "Infinite Horizon Discrete Time Control Problems for Bounded Processes." *Advances in Difference Equations*, v. 2008, Article ID 654267, 14 pages, 2009.

[19] Brandão, A. J. V. *Sobre Algumas Contribuições em Otimização Não Diferenciável Invexa*. 1998. Tese de Doutorado. IMECC- Universidade de Campinas, Campinas, 1998.

[20] Brezhneva, O. A., e Tretyakov, A."On the choice of a method for solving a general system of nonlinear equations." *Comput. Math.Math.Phys.*, v. 41, p. 633-637, 2001.

[21] Buchner, M., Marsden, J. E., e Schecter, S." Applications of the blowing-up construction and algebraic geometry to bifurcation problems." *Journal Differential Equations*, v. 48, p. 729-745, 1983.

[22] Burke, J.V." An exact penalization viewpoint of constrained optimization." *SIAM Journal Control Optim.* v. 29, p. 968-998, 1991.

[23] Brandão, A. J. V., M. A. Rojas-Medar, and G. N. Silva. "Invex nonsmooth alternative theorem and applications." *Optimization* v. 48, p. 239-253, 2000.

[24] Butkovskii, A.G." On necessary and sufficient optimality conditions for impulse control systems." *Avtomatika i telemekhanika*, v. 24, n. 8, p. 1056-1064, 1963.

[25] Censor, Y. "Pareto Optimality in Multiobjetive Problems." *Applied Mathematics and Optimization*, v. 4, p. 41-59, 1977.

[26] Clarke, F. H. *Optimization and Nonsmooth Analysis. Wiley*, Nova York, Nova York, 1983.

[27] Clarke, F. H., Ledyaev, Y. S., Stern, R. J., e Wolenski, P. R. (2008). *Nonsmooth analysis and control theory*. Springer Science & Business Media, v. 178, 2008.

[28] Clarke, F. H.*Necessary conditions for nonsmooth problems in Optimal Control and the Calculus of Variations.* Tese de Doutorado. Univ. of Washington, Washington, 1973.

[29] Clarke, F. H. "A new approach to Lagrange multipliers." *Math. Oper. Res.*, v. 1, p. 165-174, 1976.

[30] Chankong, V., e Haimes, Y. Y. *Multiobjective decision making: Theory and methodology.* North-Holland, Amsterdam, 1983.

[31] Chorobura, A. P. *Condições de otimalidade para problemas com um e com vários objeticos: abordagem através do formalismo de Dubovitskii-Milyutin.* Dissertação de Mestrado, Programa de Pós-Graduação em Matemática Aplicada, Curitiba, 2014.

[32] Costa, T. M., Chalco-Cano, Y., Lodwick, W. A., Silva, G. N." Generalized interval vector spaces and interval optimization." *Information Sciences*, v. 311, p. 74-85, 2015.

[33] Craven,B. D. *Mathematical programming and control problems.* Chapman&Hall, Londres, 1978.

[34] Craven, B. D., e B. M. Glover. "Invex functions and duality." *Journal of the Australian Mathematical Society (Series A)*, v. 39, n. 01, p. 1-20, 1985.

[35] Craven, B. D. *Control and optimization.* CRC Press, v. 16, 1998.

[36] Da Cunha, N.O., e Polak, E." Constraint minimization under vector-value criteria in finite dimensional spaces." *Journal of Mathematical Analysis and Applications*, v. 19, p. 103-124, 1967.

[37] Dontchev, A. L."Perturbations, approximations and sensitivity analysis of optima control systems."*Lecture Notes in Control and Information Sciences - Springer-Verlag*, Berlin, v. 53, 1983.

[38] Dubovitskii, A.Ya." Discrete maximum principle." *Automation and Remote Control*,v. 10, n. 1, p. 55-71, 1978.

[39] Dubovitskii, A. Ya, e A. A. Milyutin. "Extremum problems in the presence of restrictions."*Zh. Vychisl. Mat. Mat. Fiz*, v. 5, n. 3, p. 395-453, 1965.

[40] Dubovitskii, A. Ya, e A. A. Milyutin. "Second variations in extremal problems with constraints."*Dokl. Akad. Nauk SSSR.*, v. 160, n. 1, 1965.

[41] Dzyuba, N. N., and B. N. Pshenichnyi. "Discrete maximum principle." *Cybernetics and Systems Analysis*, v. 11, n. 2, p. 226-230, 1975.

[42] Fan, L. T., e C. S. Wang. *The Discrete MAximum Principle.* John Wiley, Nova York, 1964.

[43] Fiacco, A. V. *Introduction to sensitivity and stability analysis in nonlinear programming.* Academic Press, New York, 1983.

[44] Fiacco, A. V. Mathematical programming with data perturbations.*CRC Press*, v. 195, 1997.

[45] Gaishun, I. V. *Multiparameter control systems.* Minsk: Navuka i Tekhnika, 1996.

[46] Geoffrion, A. M." Duality in nonlinear programming: a simplified applications oriented development." *SIAM Review.*, v. 13, n. 1, p. 1-37, 1971.

[47] Girsanov, I. V. Lecture Notes in Economics and Mathematical Systems. *Springer-Verlag*, Berlin - Heidelberg - Nova York, v. 67, 1972.

[48] Hansen, E.*Global Optimization using Interval Analysis.* Marcel Dekker, 1992.

[49] Hanson,M. A., e Mond, B."Further Generalizations of Convexity in Mathematical Programming." *Journal of Information anf Optimization Sciences*, v. 3, p. 25-32, 1982.

[50] Hestenes, M. R. *Calculus of variations and optimal control theory.* J Wiley and Sons, Nova York, 1969.

[51] Hilscher, R., e Zeidan, V." Second order sufficiency criteria for a discrete optimal control problem." *The Journal of Difference Equations and Applications*, v. 8, n. 6, p. 573-603, 2002.

[52] Hilscher, R., e Zeidan, V. " Discrete optimal control: second order optimality conditions." *The Journal of Difference Equations and Applications*, v. 8, n. 10, p. 875-896, 2002.

[53] Izmailov, A., e Solodov, M. *Otimização, volume I.* IMPA, Rio de Janeiro, 2009.

[54] Izmailov, A. F."Optimality conditions in extremal problems with nonoregular inequality constraints a translated from Matematieheskie Zametki."v. 66, n. 1,p. 89-101, julho, 1999. Artigo original submetido em 22 de setembro de 1997.

[55] Jittorntrum, K. "Solution point differentiability without strict complementarity in nonlinear programming."*Sensitivity, Stability and Parametric Analysis*, p. 127-138, 1984.

[56] Kotarski, W."On some specification of the Dubovitskii-Milyutin Theorem for Pareto optimal problems." *Nonlinear Analysis*, v. 14, p. 287-291, 1990.

[57] Kotarski, W. *Some problems of optimal and Pareto optimal control for distributed parameter systems.* Wydawnictwo Uniwersytetu Slskiego, 1997.

[58] Ledzewicz-Kowalewska, U. "Optimality and Pareto optimality conditions for the problems with nonregular operator equality constraints." *Nonlinear Analysis: Theory, Methods & Applications*, v. 17, n. 4, p. 347-360, 1991.

[59] Levitin, E. S. *Perturbation theory in mathematical programming and its applications.* Wiley, v. 38, 1994.

[60] Luo, Z.Q., Pang, J.S., e Ralph, D. *Mathematical Programs with Equilibrium Constraints.*Cambridge University Press, Nova York 1997.

[61] Malanowski, K. "Stability and Sensitivity Analysis of Discrete Optimal-Control Problems." *Problems of Control and Information Theory-Problemy Upravleniya I Teorri Informatsii*, v. 20, n. 3, p. 187-200, 1991.

[62] Malanowski, K. "Differential sensitivity of solutions to convex constrained optimal control problems for discrete systems." *Journal of Optimization Theory and Applications*, v. 53, p. 429-449, 1987.

[63] Malanowski, K. "Stability of solutions to convex problems of optimization." *Lecture and Notes - Springer-Verlag*, Nova York, v. 93, 1987.

[64] Mangasarian, O. L. *Nonlinear Programming*. SIAM, v. 10, 1994.

[65] Marinkovic, B. "Sensitivity analysis for discrete optimal control problems." *Mathematical Methods of Operations Research*, v. 63, n. 3, p. 513-524, 2006.

[66] Marinkovic, B. " Optimality conditions in discrete optimal control problems with state constraints." *Numerical Functional Analysis and Optimization*, v. 28, n. 7-8, p. 945-955, 2007.

[67] Marinkovic, B. " Optimality conditions for discrete optimal control problems." *Optimizaton Methods and Software*, v. 22, n. 6, p.959-969, 2007.

[68] Marinkovic, B. "Second-order optimality conditions in a discrete optimal control problem." *Optimization*, v. 57, n. 4, p. 539-548, 2008.

[69] Marinkovi?, B. "2-regularity and 2-normality conditions in discrete optimal control problems. Numerical Functional Analysis and Optimization." v. 29, n.11-12, p. 1286-1298, 2008.

[70] Martin, D. H. " The essence of invexity." *Journal of Optimization Theory and Applications*, v. 47, n. 1, p. 65-76, 1985.

[71] Mishra, S. K., e Giorgio, G. *Invexity and optimization*. Springer Science & Business Media, v. 88, 2008.

[72] Moore, R. E. *Methods and Applications of Interval Analysis*. SIAM, 1979.

[73] Moulin, H., e Soulié, F. F. *La convexité dans les mathématiques de la décision*. HERMANN, Paris, 1979.

[74] Mordukhovich B. S.; I. Shvartsman: Discrete maximimum principle for nonsmooth optimal control problems with delays. Cybernetics and Systems Analysis (2002) 38: 255–264

[75] de Oliveira, V. A., e Silva, G. N. " New optimality conditions for nonsmooth control problems." *J. of Global Optim.*, v. 57, n. 4, p. 1465-1484, 2013.

[76] de Oliveira, V. A., e Silva, G. N. "On sufficient optimality conditions for multiobjective control problems." *Journal of Global Optimization*, v. 64, n. 4, p. 721, 2016.

[77] Orellana-Vivanco, V. N. *Problemas de extremos regulares y no regulares, vía formalismo de Dubovitskii-Milyutin: aplicación a problemas de control óptimo*. Tese de Doutorado. Universidad de Sevilla, 2013.

[78] Osuna-Gómez, R., Rufián-Lizana, A., e Ruiz-Canales, P.“ Invex functions and generalized convexity in multiobjective programming.” *Journal of Optimization Theory and Applications*, v. 98, n. 3, p. 651-661, 1998.

[79] Pareto, V. *Cours d'economie politique.* Rouge, Lausanni, Suíça, 1896.

[80] T. D. Phuong, T. D., Sach, P. H., e Yen, N. D. “Strict lower semicontinuity of the level sets and invexity of a locally Lipschitz function.” *Journal of Optimization Theory and Applications*, v. 87, p. 579-594, 1995.

[81] Pontryagin, L., V. Boltianski, V., Gramkrelidze, R., e Mitchenko, E. *Theorie Mathématique des Processus Optimaux,* Mir, Moscow, Russia, 1974.

[82] Propoi, A. I. *Elements of the theory of optimal discrete processes.* M.: Nauka, 1973.

[83] Robinson, S. M. “Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity.”*Nonlinear Analysis and Optimization- Springer*, Berlin, Heidelberg, p. 45-66, 1987.

[84] Robinson, S. M. “Regularity and stability for convex multivalued functions.” *Mathematics of Operations Research*, v. 1, p. 130-145, 1976.

[85] Robinson, S. M. “Stability theory for systems of inequalities, part II: Differentiable nonlinear systems.”*SIAM Journal of Numerical Analysis*, v. 13, p. 497-513, 1976.

[86] Rosenberg, E.“Exact penalty functions and stability in locally Lipschitz programming.” *Mathematical Programming*, v. 30, p. 340-356, 1984.

[87] Rozonoer, L. I. “Pontryagin's maximum principle in optimal system theory.”*Automatika i telemekhan*, v. 20, n. 10, p. 1320-1334, 1959.

[88] Ruiz-Canales, P., e A. Rufián-Lizana, A. “ A characterization of weakly efficient points.” *Mathematical Programming*, v. 68, p. 205-212, 1995.

[89] Sach, P. H., Lee, M. G, e Kim, D. S.“ Infine functions, nonsmooth alternative theorems and vector problems.”textit Journal of Global Optimization, v. 27, p. 51-81, 2003.

[90] Santos, L. B., Rojas-Medar, M., e de Oliveira, V. A. “ Saddle Point and Second Order Optimality in Nondifferentiable Nonlinear Abstract Multiobjective Optimization.” *TEMA Tend. Mat. Apl. Comput.*, v. 13, n. 2, p. 179-191, 2012.

[91] Shvartsman I.“ Necessary Optimality Conditions in Discrete Nonsmooth Optimal Control.”*Journal of Optimization Theory and Applications*, v. 153, p. 578-586, 2012.

[92] Tind J.,e Wolsey L.“ An elementary survey of general duality theory in mathematical programming”*Mathematical Programming*, v. 21, p. 241-261, 1981.

[93] Tretyakov, A. A., e Marsden, J. E.“ Factor analysis of nonlinear mappings: p-regularity theory.” *Communications on Pure and Applied Analysis.*, v. 2, n. 4, p. 425-445, 2003.

[94] Tretyakov, A. A., e Brezhneva, O. A." The pth-order optimality conditions for inequality constrained optimization problems." *Nonlinear Analysis*, v. 63, p. 1357-1366, 2005.

[95] Tretyakov, A. A., e Szczepanik, E. "Irregular optimization models and p-order Kuhn Kucker optimality conditions. *Journal of Computer and Systems Sciences International*, v. 53, p. 86-93, 2014.

[96] Vasilev, F.P. *Methods for solving extremal problems* [in Russian]. Moscow: Nauka. 1981.

[97] Voller, R. L., e Ratchek, H." What can interval analysis do for global optimization." *Journal of Global Optimization*, v. 1, p. 111-130, 1990.

[98] Zhai, J., e Huang, X. X." Calmness and Exact Penalization in Vector Optimization under Nonlinear Perturbations." *Journal of Optimization Theory and Applications*, v. 162, p. 856-872, 2014.