SOBRE ALGUMAS CONTRIBUIÇÕES EM OTIMIZAÇÃO NÃO DIFERENCIÁVEL INVEXA

ADILSON JOSÉ VIEIRA BRANDÃO

Prof. Dr. Marko Antonio Rojas Medar

Orientador

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

IMECC-UNICAMP

MAIO-1998

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Brandão, Adilson José Vieira

B733s Sobre algumas contribuições em otimização não diferenciável invexa / Adilson José Vieira Brandão -- Campinas, [S.P. :s.n.], 1998.

Orientador : Marko A. Rojas Medar

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

1. Otimização não diferenciável. 2. Otimização matemática. 3. Análise funcional não-linear. 4. Programação não-linear. I. Rojas Medar, Marko Antonio. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

SOBRE ALGUMAS CONTRIBUIÇÕES EM OTIMIZAÇÃO NÃO DIFERENCIÁVEL INVEXA

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. Adilson José Vieira Brandão e aprovada pela Comissão Julgadora.

Campinas, 18 de maio de 1998.

Prof. Dr. Marko Antonio Rojas Medar
Orientador

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

Tese de Doutorado defendida e aprovada em 18 de maio de 1998

Pela Banca Examinadora composta pelos Profs. Drs.

____

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

____

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

____

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

____

Prof (a). Dr (a). JOSÉ LUIZ BOLDRINI

____

Prof (a). Dr (a). DOHERTY ANDRADE

Dedicatória

Este trabalho é dedicado aos meus pais Sr. Brandão e D. Neguinha, à minha esposa Maraisa e aos meus filhos Caroline e Thiago.

Agradecimentos

- À instituição CNPq pelo suporte financeiro.
- Aos membros da minha banca examinadora, os professores Geraldo Nunes Silva, Roberto Andreani, Doherty Andrade e José Luis Boldrini, pelas várias sugestões que em muito contribuiram para a redação final da tese.
- Aos professores, Ricardo A. Bacci, Geraldo S. S. Ávila e Rodney C. Bassanezi, pelo incentivo desde a época da graduação até a conclusão do meu doutorado.
- Aos funcionários do IMECC de um forma geral, e à Fátima da Secretaria da Matemática Aplicada em particular, pelo suporte e amizade em todos os momentos.
- A todos os colegas e amigos pela convivência agradável neste período.
- Ao Prof. Marko Antonio Rojas Medar pela orientação segura, pela confiança em minha pessoa e sobretudo pela sua amizade.

A juventude e a velhice

A Juventude não é um período da vida; a juventude é um estado de espírito, um efeito da vontade, uma qualidade da imaginação, uma intensidade emotiva, uma vitória do valor sobre a timidez, do gosto pela aventura sobre o amor ao conforto. Alguém não se torna velho por haver vivido um certo número de anos; torna-se velho porque desertou dos ideais. Os anos enrugam a pele, mas a renúncia a um ideal enruga a alma. As preocupações, as dúvidas, os temores e as desesperanças, são os inimigos que lentamente, nos fazem vergar para o chão e nos convertem em pó antes da morte. Jovem é o que deslumbra e se maravilha ... o que pergunta como menino - E depois? Jovem é o que desafia os acontecimentos e encontra alegrias no jogo da vida. As provas galvanizam-no, os fracassos o tornam mais forte, as vitórias o tornam melhor. Serás tão jovem como tua fé, tão velho como tuas dúvidas, tão jovem como a confiança que tenhas em ti, tão velho como tuas desesperanças, e mais velho ainda como o teu abatimento. Permanecerás jovem, tanto quanto permaneceres verdadeiramente generoso, tanto quanto sentires o entusiasmo de dar alguma coisa de ti: pensamentos, palavras, amor; tanto quanto o fato de dar alguma coisa, te der a impressão de receber; e por conseguinte, se sempre estás devendo e desejando dar mais. Permanecerás jovem enquanto fores receptivo a tudo quanto é belo, bom e grandioso, podendo desfrutar das mensagens da natureza, do homem e do infinito. Se um dia qualquer que seja da tua idade, teu coração for mordido pelo pessimismo, torturado pelo egoísmo, roído pelo cinismo, que Deus tenha piedade de tua alma de velho.

Gal. Douglas MacArthur

Conteúdo

Resumo

Nosso objetivo neste trabalho de tese é estudar alguns problemas de otimização onde estabelecemos, entre outros resultados, condições suficientes de otimalidade global sem nenhuma hipótese de convexidade ou diferenciabilidade. As técnicas para se atacar tais problemas são a análise não diferenciável devida ao matemático canadense Clarke e o conceito de convexidade generalizada, chamado invexidade, introduzido pelo matemático americano Hanson, as quais são detalhadas no capítulo 1. No capítulo 2 estudamos alguns problemas de programação matemática estabelecendo condições suficientes de otimalidade global e dualidade. De posse desses resultados estabelecemos nosso principal resultado na seção: um teorema de alternativa invexo do tipo Gordan, onde as funções envolvidas são localmente Lipschitz e invexas. No capítulo 3 obtemos condições suficientes de otimalidade global na forma de uma regra de multiplicadores para um problema de otimização entre espaços de Banach. No capítulo 4 obtemos condições suficientes de otimalidade global na forma de uma regra de multiplicadores para um problema de programação matemática com tempo contínuo o qual estende os resultados obtidos pelo matemático americano Zalmai para o mesmo problema no caso diferenciável. Também estabelecemos condições suficientes de 2a. ordem utilizando a noção de Hessiano generalizado introduzida pelos matemáticos chilenos Cominetti e Correa. No último capítulo damos algumas direções de pesquisa futura dentro da área de otimização não diferenciável.

Capítulo 1

Introdução e resultados preliminares

1.1 Introdução

Nosso objetivo nessa tese é estudar alguns problemas de otimização que, numa forma geral, podem ser colocados da seguinte maneira: Sejam E, F e G espaços normados, $\Omega \subset E$ um aberto de E, $C \subset \Omega$ um subconjunto não vazio de Ω e $K \subset G$ um cone em G. Consideremos o seguinte problema geral de otimização:

$$\left.\begin{array}{lr} \text{Minimizar} & f(x) \\ \\ \text{sujeito a} & -g(x) \in K \\ \\ & x \in C, \end{array}\right\} \quad (PG)$$

onde $f : \Omega \to F$ e $g : \Omega \to G$ são aplicações sem, a priori, nenhuma hipótese de diferenciabilidade ou convexidade. Quando $F = \Re$ o problema acima toma a forma de um problema usual de programação matemática. Quando $F = \Re^m$ ou F é infinito dimensional temos um problema de programação multiobjetivo.

Uma das razões para estudarmos tal tipo de problema é fato que, nos dias de hoje, existe uma vasta produção científica na área de otimização no caso em que os problemas tratados envolvem funções diferenciáveis e/ou convexas. No entanto, em se tratando de problemas onde as funções utilizadas não possuem tais agradáveis propriedades, a situação muda de figura. Somente nos últimos trinta anos foi desenvolvido algum ferramental matemático no sentido de se tentar atacar tais problemas de maneira sistemática. Dessa forma, essa área vem se tornando muito atrativa e promissora, pois ainda existe muita coisa por se fazer. Outra razão para direcionarmos algum esforço matemático nessa linha de pesquisa é a necessidade das aplicações pois, em muitos modelos matemáticos, hipóteses de diferenciabilidade e ou convexidade vem se tornando excessivamente exigentes e restritivas. Em vista disso tudo, procuraremos dar algumas contribuições no estudo de tais problemas. Passemos a discutir como pretendemos fazer isso.

O termo *não diferenciável* é muito genérico e pode envolver uma classe muito grande de funções. Em nosso trabalho, nos restringiremos a funções que possuam algum tipo de propriedade *"Lipschitz"*, sem, a priori, nenhuma hipótese de diferenciabilidade (pelo menos não nos sentidos clássicos de diferenciabilidade como Fréchet, Gâteaux, etc). Por exemplo, seja E um espaço de Banach, $\Omega \subset E$ um aberto não vazio de E e $f : \Omega \to \Re$ uma função. Dizemos que f é uma função Lipschitz em Ω, se existe uma constante $K > 0$ tal que

$$|f(y) - f(x)| \leq K\|y - x\|,$$

para todo $x, y \in \Omega$. Esse tipo de função, além de ser abundante nas aplicações, possui boas propriedades quando se associa a ela um tipo de derivada generalizada introduzida por Clarke em [16]: a derivada direcional generalizada de f em $x \in \Omega$ na direção $v \in E$, denotada $f^o(x; v)$, é dada por

$$f^o(x; v) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{1}{\lambda}[f(y + \lambda v) - f(y)],$$

onde $y \in \Omega, \lambda \in (0, +\infty)$. Em dualidade com essa derivada generalizada, Clarke introduziu um gradiente generalizado que é uma boa extensão tanto do gradiente clássico, no caso de funções continuamente diferenciáveis, quanto do subdiferencial da análise convexa, no caso de funções convexas.

É fato conhecido que para se obter condições globais de otimalidade em problemas de otimização, é necessário algum tipo de hipótese de convexidade.

Para substituirmos a ausência de convexidade nos nossos problemas, e ainda assim obtermos resultados globais, utilizaremos um conceito de convexidade generalizada introduzido por Hanson [28] no caso diferenciável e extendida para o caso de funções Lipschitz por Phuong, Sach e Yen [47]. Craven [19] chamou tais funções de *"invex functions"* numa alusão ao termo invariante convexo, e nós, numa livre tradução do inglês, chamaremos tais funções de funções invexas. A definição de função invexa no caso Lipschitz pode ser colocada, grosso modo, da seguinte forma: Dizemos que f é invexa em Ω se, para todo $x, y \in \Omega$, existe $\eta(y,x) \in E$ tal que

$$f(y) - f(x) \geq f^o(x, \eta(y,x)).$$

Essa classe de funções é estritamente maior do que a classe das funções convexas, possuindo muitas das boas propriedade que funções convexas tem em otimização, tais como garantia de otimalidade global, bons resultados de dualidade, etc.

Na próxima seção falaremos um pouco mais dessas duas ferramentas básicas: derivada generalizada de Clarke e invexidade. Na última seção daremos uma prévia dos resultados que iremos apresentar na tese, discutindo a relevância e originalidade dos mesmos e comparando com outros trabalhos similares já existentes na literatura.

1.2 Análise não diferenciável e invexidade

1.2.1 Introdução

Nesta seção introduziremos as definições e resultados *básicos* que serão necessários para uma boa compreensão dos capítulos seguintes. Resultados específicos de cada capítulo serão apresentados nos mesmos. Como dissemos na introdução, as duas ferramentas essenciais que utilizaremos para obter nossos resultados em otimização não diferenciável e não convexa são a teoria de derivadas e gradientes generalizados de Clarke e a noção de convexidade generalizada, chamada *invexidade*. Para maiores detalhes e informações sobre análise não diferenciável, derivadas e gradientes generalizados o leitor pode consultar o excelente livro de Clarke [16]. As referências sobre invexidade serão citadas no transcorrer do texto.

1.2.2 Análise não diferenciável

No que se segue E denotará um espaço de Banach real com a norma $||.||$, E^* será seu espaço dual com a norma dada por

$$||\zeta||_* = \sup\{\langle \zeta, v\rangle : v \in E, ||v|| \leq 1\},$$

onde $\langle \cdot, \cdot \rangle$ é a aplicação dualidade canônica entre E^* e E. $\Omega \subset E$ denotará um aberto não vazio de E.

Definição 1.1. *Dizemos que uma função $f : \Omega \longrightarrow \Re$ é Lipschitz próxima de $x \in \Omega$, se existe um $\varepsilon > 0$ uma constante positiva $K = K(x, \varepsilon)$ tal que*

$$|f(x_1) - f(x_2)| \leq K||x_1 - x_2||,$$

para todo $x_1, x_2 \in \{x + \varepsilon B\} \cap \Omega$, onde B é a bola unitária aberta em E. A constante K é chamada a constante de Lipschitz de f. Se f é Lipschitz próxima de todo ponto de Ω dizemos que f é localmente Lipschitz em Ω.

Definição 1.2. *Definimos a derivada direcional generalizada de f em $x \in \Omega$ na direção $v \in E$, denotada $f^o(x; v)$, por*

$$f^o(x; v) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{1}{\lambda}[f(y + \lambda v) - f(y)],$$

onde $y \in \Omega, \lambda \in (0, +\infty)$.

Definição 1.3. *Definimos o gradiente generalizado de f em x, denotado $\partial f(x)$, como sendo o subconjunto de E^* dado por*

$$\partial f(x) = \{\xi \in E^* : f^o(x; v) \geq \langle \xi, v\rangle, \ \ \forall v \in E\}.$$

Teorema 1.4. *Suponha que $f : \Omega \to \Re$ é uma função localmente Lipschitz. Para cada $x \in \Omega$ seja K a constante de Lipschitz de f. Então,*

(i) A função $v \to f^o(x;v)$ é finita, sublinear e vale $|f^o(x;v)| \leq K\|v\|$;
(ii) $\partial f(x)$ é um subconjunto não vazio, convexo e w^*-compacto de E^* e para cada $\xi \in \partial f(x)$, temos que $\|\xi\|_* \leq K$;
(iii) Para todo $v \in E$,

$$f^o(x;v) = \max\{\langle v, \xi\rangle : \xi \in \partial f(x)\};$$

(iv) $\xi \in \partial f(x)$ se e somente se $f^o(x;v) \geq \langle \xi, v\rangle\ \forall v \in E$.

Como mostra o Teorema anterior, conhecer o conjunto $\partial f(x)$ é equivalente a conhecer a função $f^o(x;v)$. Este é um exemplo de um fato mais geral: conjuntos convexos e fechados são caracterizados pelos seus funcionais suporte. Lembremos que o *funcional suporte* de um subconjunto não vazio C de E é a função $\sigma_C : E^* \longrightarrow R \cup \{+\infty\}$ dada por

$$\sigma_C(\zeta) := \sup\{\langle \zeta, x\rangle : x \in C\}.$$

Se Σ é um subconjunto de E^*, seu funcional suporte é definido em E^{**}. Se identificarmos E com um subconjunto de E^{**}, então, para cada $x \in E$, temos que

$$\sigma_\Sigma(x) := \sup\{\langle \zeta, x\rangle : \zeta \in \Sigma\}.$$

Teorema 1.5. *Sejam C, D subconjuntos convexos fechados e não vazios de E; Σ, Δ subconjuntos convexos w^*-fechados e não vazios de E^*; $K \subset E$ um cone convexo e fechado; K^o seu cone polar e $\mu, \lambda \geq 0$. Então*

(i) $C \subset D$ se e somente se $\sigma_C(\zeta) \leq \sigma_D(\zeta)\ \ \forall \zeta \in E^*$;
(ii) $\Sigma \subset \Delta$ se e somente se $\sigma_\Sigma(x) \leq \sigma_\Delta(x)\ \ \forall x \in E$;
(iii) $\mu\sigma_C(v) + \lambda\sigma_D(v) = \sigma_{\{\mu C + \lambda D\}}(v)$;
(iv) $\mu\sigma_\Sigma(v) + \lambda\sigma_\Delta(v) = \sigma_{\{\mu\Sigma + \lambda\Delta\}}(v)$.

(v) $\sigma_K(v) = \psi_{K^o}(v)$ onde ψ_{K^o} é a função indicatriz de K^o.

O seguinte teorema mostra que o gradiente de Clarke é uma boa generalização dos casos diferenciável e convexo:

Teorema 1.6. *Seja $f : \Omega \to \Re$, localmente Lipschitz. Então,*
(i) Se f é convexa, então $\partial f(x)$ coincide com o subdiferencial de f no sentido da análise convexa;
(ii) Se f é continuamente diferenciável então $\partial f(x) = \{f'(x)\}$, onde $f'(x)$ é a derivada usual de f em x.

Definição 1.7. *Dizemos que f é regular (Clarke regular) em $x \in \Omega$ se*
(i) para todo $v \in E$, a derivada direcional $f'(x;v)$ existe;
(ii) para todo $v \in E, f'(x;v) = f^o(x;v)$.
Se f é regular para todo $x \in \Omega$ então f é regular em Ω.

Seja C um subconjunto não vazio de Ω, e consideremos sua função distância $d_C(.) : E \to \Re$ definida por

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

A função d_C não é diferenciável mas é globalmente Lipschitz.

Definição 1.8. *Seja $x \in C$. Um vetor $v \in E$ é dito tangente a C em x se $d_C^o(x;v) = 0$. O conjunto de todos os tangentes a C em x é um cone convexo e fechado em E denotado por $T_C(x)$ chamado o cone tangente de Clarke a C em x. Definimos o cone normal a C em x por polaridade com $T_C(x)$:*

$$N_C(x) = \{\xi \in E^* : \langle \xi, v \rangle \leq 0, \ \ \forall v \in T_C(x)\}.$$

$N_C(x)$ é um cone convexo e w^*-fechado em E^*.

No caso em que E é um espaço finito dimensional, digamos, $E = \Re^n$, podemos dar uma boa caracterização de $\partial f(x)$. De fato o Teorema de Rademacher nos diz que se uma função $f : \Omega \to \Re$ é Lipschitz no aberto

$\Omega \subset \Re^n$ então f é diferenciável em quase todo ponto de Ω. Chamemos Ω_f o conjunto dos pontos onde f falha em ser diferenciável.

Teorema 1.9. *Seja $f : \Omega \to \Re$ Lipschitz próxima de x e suponha que S é uma conjunto de medida de nula no $\Re^n$. Então*

$$\partial f(x) = co\{\lim \nabla f(x_i) : x_i \to x, x_i \notin S, x_i \notin \Omega_f\}.$$

A caracterização acima é muito útil no cálculo de $\partial f(x)$. Por exemplo, consideremos a função $f : \Re \to \Re$ dada por $f(x) = |x|$. f é claramente uma função localmente Lipschitz e só não é diferenciável no ponto zero. Calculemos $\partial f(0)$. Para qualquer sequência de números positivos x_i convergindo a zero, $\lim \nabla f(x_i) = 1$. Analogamente, para qualquer sequência de números negativos x_i convergindo a zero, $\lim \nabla f(x_i) = -1$. Portanto $\partial f(0) = co\{-1, 1\} = [-1, 1]$. Neste caso $f'(0, v) = f^o(0, v) = |v|$ e f é Clarke regular em $x = 0$. Agora, se tomarmos $f(x) = -|x|$, um cálculo análogo ao anterior mostra que $\partial f(0) = [-1, 1]$ e portanto $f^o(0, v) = |v|$. No entanto é fácil verificar que $f'(0, v) = -|v|$ e temos então um exemplo de uma função que não é Clarke regular.

Teorema 1.10. *Suponha que $f : \Omega \to \Re$ é uma função localmente Lipschitz e que f atinge um mínimo sobre C em $\overline{x}$. Então*

$$0 \in \partial f(\overline{x}) + N_C(\overline{x}). \tag{1.1}$$

Definição 1.11. *Dizemos que $\overline{x} \in C$ é um ponto estacionário generalizado de f sobre C se vale (1.1).*

1.2.3 Invexidade

Consideremos o seguinte problema de programação matemática:

$$\left.\begin{array}{lc} \text{Minimizar} & f_0(x) \\ \text{sujeito a} & f_i(x) \leq 0, \quad i \in I = \{1, \ldots, m\}, \end{array}\right\} (PH)$$

onde $f_i : \Re^n \longrightarrow \Re$ $i \in I \cup \{0\}$, são funções diferenciáveis.

Impondo-se algum tipo de regularidade nas restrições do problema acima, uma condição necessária para que $\bar{x} \in \Re^n$ seja um mínimo local de (PH) é que existam $\lambda_i \in \Re, i \in I$, tais que

$$\lambda_i \geq 0, \tag{1.2}$$

$$\lambda_i f_i(\bar{x}) = 0, \tag{1.3}$$

$$\nabla f_0(\bar{x}) + \sum_{i \in I} \lambda_i \nabla f_i(\bar{x}) = 0. \tag{1.4}$$

A regra de multiplicadores acima nos permite encontrar apenas candidatos a solução do problema (PH). A situação ideal seria aquela em que tivessemos alguma hipótese que garantisse a suficiência das condições (1.2)-(1.4) numa forma global, isto é, se $\bar{x}$ satisfaz (1.2)-(1.4) então $\bar{x}$ é uma solução global de (PH). Um resultado bastante conhecido e que num certo sentido responde a esta questão é o seguinte: se todas as funções envolvidas são convexas e se um ponto $\bar{x}$, satisfaz (1.2)-(1.4) então $\bar{x}$ é solução ótima global de (PH). Uma crítica natural a esse resultado é que esta hipótese de convexidade é, em boa parte dos casos, muito restritiva. É natural então a tentativa de introduzir conceitos que em algum sentido generalizem a noção usual de convexidade e que mantenham algumas de suas boas propriedades.

Em 1981 Hanson [28] introduziu uma nova classe de funções. Ele considerou funções diferenciáveis para as quais exista $\eta(x, y) \in \Re^n$ tal que

$$f(y) - f(x) \geq \langle \nabla f(x), \eta(y, x) \rangle, \quad \forall x, y \in \Re^n. \tag{1.5}$$

É óbvio que funções convexas diferenciáveis satisfazem (1.5) com $\eta(y, x) = y - x$. Hanson mostrou que, se no problema (PH), $f_0, f_i, \in I$ satisfazem (1.5), com a mesma função η comum a todas, então as condições (1.2)-(1.4) garantem otimalidade global.

Neste mesmo trabalho Hanson demonstra que essa classe de funções é "maior" do que a classe das funções convexas diferenciáveis, através do seguinte exemplo de programa não convexo mas que satisfaz (1.5) para uma mesma função η:

$$\begin{array}{ll} \text{Minimizar} & x_1 - \sin x_2 \\ \\ \text{sujeito a} & \sin x_1 - 4\sin x_2 \leq 0, \\ & 2\sin x_1 + 7\sin x_2 + x_1 - 6 \leq 0, \\ & 2x_1 + 2x_2 - 3 \leq 0, \\ & 4x_1^2 + 4x_2^2 - 9 \leq 0, \\ & -\sin x_1 \leq 0, \\ & -\sin x_2 \leq 0, \\ & (x_1, x_2) \in \Re^2. \end{array}$$

Notemos a natureza não convexa do conjunto de restrições e da função objetivo. No entanto todas elas satisfazem (1.5) com $\eta(y,x)$ dada por

$$\eta(y,x) = (\frac{\sin y_1 - \sin x_1}{\cos x_1}, \frac{\sin y_2 - \sin x_2}{\cos x_2}),$$

onde $x = (x_1, x_2)$ e $y = (y_1, y_2)$.

Em [19] Craven observou que se $F : \Re^n \to \Re$ é uma função convexa e $\phi : \Re^n \to \Re^n$ é uma aplicação diferenciável com derivada inversível então, tomando $f = F \circ \phi$ e $x, y \in \Re^n$ temos que

$$\begin{aligned} f(y) - f(x) = & \; F(\phi(y)) - F(\phi(x)) \\ \geq & \; \langle \nabla F(\phi(x)), \phi(y) - \phi(x) \rangle \\ = & \; \langle \nabla f(x), J\phi(x)^{-1}[\phi(y) - \phi(x)] \rangle \\ = & \; \langle \nabla f(x), \eta(y,x) \rangle, \end{aligned}$$

onde $\eta(y,x) = J\phi(x)^{-1}[\phi(y) - \phi(x)]$ e $J\phi(x)$ é o Jacobiano de ϕ em x. A conta acima mostra o que resta da convexidade de F após seu domínio ter sido distorcido por ϕ. Essa propriedade que permanece (invariante convexo) é justamente a propriedade (1.5). Baseado nisso Craven chamou tal propriedade de *"invexity"* e denominou as funções que satisfazem (1.5) de *"invex functions"* numa alusão a invariante convexo. Numa "livre tradução" para o português, nós utilizaremos as denominações *invexidade* e *funções invexas*, respectivamente.

Os trabalhos de Hanson e Craven foram muito importantes na teoria de otimização global e inspiraram uma grande quantidade de trabalhos subsequentes. Em meados da década de oitenta temos os trabalhos de Craven e Glover [20] e Ben-Israel e Mond [3] que são referências importantes para quem quer se iniciar no estudo do conceito de invexidade no caso de funções diferenciáveis. Um resultado importante que aparece nestes trabalhos é a seguinte caracterização de funções invexas diferenciáveis: Uma função é invexa se e somente se todo ponto estacionário é mínimo global. Em [21], Craven estendeu o conceito de invexidade para funções Lipschitz e em [49], Reiland introduz o conceito de invexidade para funções localmente Lipschitz. Em [47] Phuong, Sach e Yen, introduzem um novo conceito de invexidade para funções localmente Lipschitz definidas num espaço finito dimensional e que generaliza todos os anteriores. A seguinte definição é a extensão natural da definição de Phuong, Sach e Yen para o caso infinito dimensional:

Definição 1.12. *Seja $C \subset \Omega$ um conjunto não vazio. Uma função localmente Lipschitz $f : \Omega \to \Re$ é dita invexa em $x \in C$ se, para todo $y \in C$, existe $\eta(y,x) \in T_C(x)$, tal que,*

$$f(y) - f(x) \geq f^o(x; \eta(y,x)).$$

Dizemos que f é invexa em C se f é invexa em todo ponto $x \in C$.

Toda função localmente Lipschitz convexa é invexa. De fato, suponha que $C \subset \Omega$ é um conjunto convexo e que $f : \Omega \to \Re$ é uma função localmente Lipschitz e convexa em C. Como f é convexa em C então f é Clarke regular em C, isto é, para todo $x \in C$ e $v \in \Re^n$, existe a derivada direcional $f'(x; v)$ e $f^o(x, v) = f'(x; v)$. Portanto, para todo $y \in C$,

$$f(y) - f(x) \geq f'(x; y - x) = f^o(x; y - x).$$

Desde que C é um conjunto convexo então, para todo $y \in C, \eta(y,x) = y - x \in T_C(x)$. Logo, concluímos que f é invexa em C. Portanto, toda função convexa é invexa. No entanto, nem toda função invexa é convexa. De fato, consideremos a função $f : \Re \to \Re$ dada por

$$f(x) = \begin{cases} |x| & \text{se} \quad x \leq 1, \\ \\ \frac{x+1}{2} & \text{se} \quad x > 1. \end{cases}$$

Esta função é claramente localmente Lipschitz. Usando o Teorema 1.9 é fácil calcular $\partial f(x)$:

$$\partial f(x) = \begin{cases} -1 & \text{se} & x < 0, \\ [-1, 1] & \text{se} & x = 0, \\ 1 & \text{se} & 0 < x < 1, \\ [1/2, 1] & \text{se} & x = 1, \\ 1/2 & \text{se} & 1 < x. \end{cases}$$

É fácil verificar que esta função não é convexa. No entanto, f é invexa com $\eta(x, y)$ dada por:

$$\eta(x, y) = \begin{cases} \frac{f(y) - f(x)}{\nabla f(x)} & \text{se} \quad x \neq 1 \text{ ou } x \neq 0 \\ \\ 0 & \text{se} \quad x = 0 \\ \\ f(y) - f(x) & \text{se} \quad x = 1 \text{ e } |x| \geq 1 \\ \\ 2(f(y) - f(x)) & \text{se} \quad x = 1 \text{ e } |x| < 1. \end{cases}$$

Vamos agora obter uma caracterização de funções invexas. Phuong, Sach e Yen [47] obtiveram tal caracterização para funções localmente Lipschitz invexas no caso em que E é finito dimensional. Este resultado é uma extensão para o caso não diferenciável da caracterização obtida em [3] para o caso diferenciável. Tal resultado pode ser estendido para um espaço de Banach qualquer como veremos a seguir.

Teorema 1.13. *Seja C um subconjunto não vazio de Ω. Uma função localmente Lipschtz $f : \Omega \to \Re$ é invexa em C se e somente se todo ponto estacionário de f sobre C é um mínimo global.*

Prova. Suponhamos que f é invexa em C. Se $x \in C$ é um ponto estacionário de f sobre C então $0 \in \partial f(x) + N_C(x)$. Logo, existe $\xi \in \partial f(x)$ e $\mu \in$

$N_C(x)$ tal que $0 = \xi + \mu$. Então, $0 = \langle \xi, v \rangle + \langle \mu, v \rangle$ para todo $v \in E$. Em particular, para $v \in T_C(x)$, $0 \leq -\langle \mu, v \rangle = \langle \xi, v \rangle$. Portanto, para todo $y \in C$, $f(y) - f(x) \geq f^o(x, \eta(y,x)) \geq \langle \xi, \eta(y,x) \rangle \geq 0$, isto é, x é um mínimo global de f em C. Suponha agora que todo ponto estacionário é mínimo global. Para provar que f é invex em C, tomemos $x \in C$. Se $0 \in \partial f(x) + N_C(x)$ então $f(y) \geq f(x)$ qualquer que seja $y \in C$. Basta tomar $\eta(y,x) = 0$. Suponha agora que $0 \notin \partial f(x) + N_C(x)$. Então existe $v \in T_C(x)$ tal que $f^o(x,v) < 0$, pois do contrário se $f^o(x,v) \geq 0$ para todo $v \in T_C(x)$ então $f^o(x,v) + \psi_{T_C(x)}(v) \geq 0$ para todo $v \in E$. Mas isso é equivalente a $\sigma_{\{\partial f(x) + N_C(x)\}}(v) \geq 0$ para todo $v \in E$ o qual implica que $0 \in \partial f(x) + N_C(x)$ (veja o Teorema 1.5), contrariando o fato que $0 \notin \partial f(x) + N_C(x)$. Portanto, existe $v \in T_C(x)$ tal que $f^o(x,v) < 0$. Dado $y \in C$ escolhamos $t > 0$ tal $f(y) - f(x) \geq t f^o(x,v) = f^o(x, tv)$. Basta agora, tomar $\eta(y,x) = tv$. □

Usando o Teorema acima é fácil mostrar que a classe das funções invexas é estritamente maior do que as convexas sem a necessidade de exibir uma função η. De fato, consideremos novamente a função $f : \Re \to \Re$ dada por

$$f(x) = \begin{cases} |x| & \text{se} \quad x \leq 1, \\ \frac{x+1}{2} & \text{se} \quad x > 1. \end{cases}$$

O único ponto estacionário de f é $x = 0$ pois $0 \in [-1, 1] = \partial f(0)$. É evidente que $x = 0$ é um mínimo global de f. Portanto, pelo teorema 1.13, f é invexa.

Para um o leitor interessado em outros conceitos de invexidade para funções não diferenciáveis sugerimos [27].

Um outro tipo de convexidade generalizada é o conceito de função pré-invexa introduzido por Ben-Israel e Mond em [3]. Uma função $f : \Omega \to \Re$ é pré-invexa em $C \subset \Omega$ se existe $\eta(x,y)$ tal que

$$f(y + \lambda \eta(x,y)) \leq \lambda f(x) + (1 - \lambda) f(y) \quad \text{para todo } x, y \in C \text{ e } \lambda \in [0,1].$$

Na definição acima está implícito que C satisfaz a seguinte propriedade de conexidade: $y + \lambda \eta(x,y) \in C$ para todo $x, y \in C$ e $\lambda \in [0,1]$. Conjuntos satisfazendo essa propriedade são chamados conjuntos invexos. É óbvio que se f é pré-invexa e diferenciável então f é invexa no sentido de Hanson. É interessante notar que esta propriedade não vale para funções apenas localmente Lipschitz. Por exemplo, $f : \Re \to \Re$ dada por $f(x) = -|x|$ é localmente Lipschitz e pré-invexa com $\eta(x, y)$ dado por

$$\eta(x,y) = \begin{cases} x - y & \text{se} \quad x \leq 0 \text{ e } y \leq 0, \\ x - y & \text{se} \quad x \geq 0 \text{ e } y \geq 0, \\ y - x & \quad\quad \text{caso contrário.} \end{cases}$$

No entanto f não é invexa pois seu único ponto estacionário $x = 0$ não é mínimo global. Notemos também que toda função convexa é pré-invexa e todo conjunto convexo é invexo com $\eta(x, y) = x - y$. A recíproca não é verdadeira como mostra o exemplo acima. Para outras informações sobre funções pré-invexas sugerimos [59].

1.3 Sobre as contribuições da tese

1.3.1 Introdução

Nosso objetivo nessa seção é adiantar ao leitor as principais contribuições dessa tese à otimização não diferenciável. Faremos uma análise capítulo por capítulo. Esperamos que esse procedimento facilite a leitura da tese.

1.3.2 Sobre o capítulo 2

Consideremos a seguinte versão do problema (PG) na forma de um problema de programação matemática usual:

$$\left.\begin{array}{lc} \text{Minimizar} & f_0(x) \\ \\ \text{sujeito a} & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ \\ & x \in C, \end{array}\right\} \quad (P)$$

onde Ω é um aberto do $\Re^n$, $f_0, f_i : \Omega \to \Re$, $i = 1, \ldots, m$, são funções Lipschitz em Ω, e $C \subset \Omega$ um conjunto fechado e não vazio;

No capítulo 2 obteremos, basicamente, condições necessárias e suficientes de otimalidade para o problema (P) na forma de uma regra de multiplicadores de Karush-Kuhn-Tucker. Grosso modo, tal regra pode ser sumarizada da seguinte maneira: se as funções envolvidas são funções invexas e as restrições do problema satisfazem uma condição do tipo Slater então, as condições de Karush-Kuhn-Tucker

$$0 \in \partial(f_0 + \sum_{i \in I} \bar{\lambda}_i f_i)(\bar{x}) + N_C(\bar{x}),$$

$$\bar{\lambda}_i \geq 0,$$

$$\bar{\lambda}_i f_i(\bar{x}) = 0,$$

são necessárias e suficientes para a otimalidade de $\bar{x}$ em (P). Além disso estabeleceremos resultados de dualidade. De fato consideremos o seguinte problema dual relacionado a (P):

$$\left.\begin{array}{ll} \text{Maximizar} & f_0(w) + \sum_{i \in I} \lambda_i f_i(w) \\ \text{sujeito a} & (w, \lambda) \in C \times \Re^m_+, \\ & 0 \in \partial(f_0 + \sum_{i \in I} \lambda_i f_i)(w) + N_C(w). \end{array}\right\} (D)$$

Sob hipóteses de invexidade e uma condição de regularidade do tipo Slater, provaremos que os valores deste problema dual e o problema (P) coincidem.

Notemos que existem vários desses resultados em problemas de programação matemática não diferenciável (veja por exemplo [35], [49]). No entanto, não conhecemos nenhum resultado na literatura que se aplique diretamente ao problema (P), aonde agregamos uma restrição abstrata $x \in C$, sendo C um conjunto não necessariamente convexo. De qualquer maneira, os resultados acima não são a principal contribuição desse capítulo e sim resultados auxiliares que nos permitirão obter nosso resultado principal, o Teorema de Alternativa Invexo. As conecções entre teoremas de alternativa e problemas de otimização são bem conhecidas ([23], [26], [39]). Tais teoremas servem, por exemplo, para garantir a existência de multiplicadores para problemas de programação matemática. Um protótipo de teorema de alternativa pode ser estabelecido como se segue: sejam $f_1, \cdots, f_p$ funções definidas num conjunto $C \subset \Re^n$, e consideremos os dois seguintes sistemas:

1. $\exists x \in C$ tal que $f_i(x) < 0, i = 1, \cdots, p$;
2. $\exists \lambda_i \geq 0,\ i = 1, \cdots, p$, não todos nulos, tal que $\sum_{i=1}^{p} \lambda_i f_i(x) \geq 0\ \forall x \in C$.

Então, das duas uma, ou o sistema 1 tem solução ou o sistema 2 tem solução, mas nunca ambos simultaneamente.

Existem várias versões desse tipo de teorema onde as funções envolvidas pertencem a várias classes de funções. Quando as f_i são lineares temos o Teorema de Gordan. Se as funções são convexas o resultado é o Teorema de Gordan Generalizado (veja [23] e [39]). Teoremas de alternativa similares foram estabelecidos para outras classes de funções tendo alguma propriedade de convexidade generalizada (veja [35], [30] e [59]). Nosso objetivo nesse ponto foi obter uma versão do teorema de alternativa acima no caso em que as funções f_i são Lipschitz e invexas e o conjunto C é um fechado não necessariamente convexo. O único resultado relacionado é aquele obtido por Craven em [23]. No entanto no seu Teorema de alternativa ele exige que o conjunto de restrições abstratas C seja, além de fechado, convexo. Também, as técnicas usadas por Craven na obtenção de seu resultado são diversas das usadas por nós. Nós baseamos a demonstração de nosso teorema num artigo de Geoffrion [26], em que ele usa resultados de dualidade convexa para provar o Teorema de Gordan Generalizado (caso convexo).

Como aplicação de nosso Teorema de Alternativa Invexo, caracterizaremos as soluções do seguinte problema de otimização vetorial, o qual é uma versão especializada do problema inicial (PG):

$$\left.\begin{array}{ll} \text{Minimizar} & f(x) = (f_1(x), \ldots, f_m(x)) \\ \\ \text{sujeito a} & x \in C, \end{array}\right\} (PV)$$

onde $C \subset \Omega$ é um compacto e $f_i : \Omega \to \Re\ i = 1, \ldots, m$, são funções Lipschitz e invexas em C.

Os principais resultados deste capítulo também podem ser encontrados em [10], [11].

1.3.3 Sobre o capítulo 3

Em [1] Thibault e Abdoni obtiveram condições necessárias de otimalidade para (PG) na forma de uma regra de multiplicadores do tipo Fritz-John. Nesse trabalho os autores supõem que as aplicações satisfazem certa propriedade do tipo Lipschitz e obtém seus resultados sem nenhuma hipótese de diferenciabilidade fazendo uso dos gradientes generalizados de Clarke [16] e Ioffe [32]. O trabalho de Thibault foi a fonte de inspiração para obtermos os resultados desse capítulo. De fato, obteremos aqui, condições necessárias e suficientes de otimalidade para (PG), na forma de uma regra de multiplicadores do tipo Karush-Kuhn-Tucker. Além do mais, estabeleceremos um problema dual relacionado a (PG) e provaremos resultados de dualidade fraca e forte entre eles. Todos esses resultados serão obtidos sem nenhuma hipótese de convexidade mas sim fazendo uso de um conceito de invexidade entre espaços de Banach adaptado do conceito usual de invexidade (veja definição 1.10). Resultados relacionados, mas que ou envolvem outras classes de funções ou são menos gerais que o nosso, podem ser encontrados em [17], [36], [44].

Os principais resultados deste capítulo também podem ser encontrados em [12].

1.3.4 Sobre o capítulo 4

Consideremos o seguinte problema de programação matemática com tempo contínuo:

$$\left.\begin{array}{c}\text{Minimizar } \ \phi(x) = \int_0^T f(t, x(t))dt, \\ \\ \text{sujeito a } \quad g_i(t, x(t)) \leq 0 \ \text{ a.e. in } [0, T], \\ i \in I = \{1, \ldots, m\}, \quad x \in X,\end{array}\right\} \ (PNC)$$

onde X é um subconjunto aberto, convexo, e não vazio do espaço de Banach $L_\infty^n[0, T]$ de todas as funções vetoriais n-dimensionais Lebesgue mensuráveis, que são essencialmente limitadas, definidas no intervalo compacto $[0, T] \subset I\!R$, com a norma $\|\cdot\|_\infty$ definida por

$$\|x\|_\infty = \max_{1 \leq j \leq n} \operatorname{ess\,sup}\{|x_j(t)|, \ 0 \leq t \leq T\},$$

onde, para cada $t \in [0, T]$, $x_j(t)$ é a $j-$ésima componente de $x(t) \in I\!R^n$, ϕ é uma função a valores reais, definida em X, $g(t, x(t)) = \gamma(x)(t)$ e $f(t, x(t)) = \Gamma(x)(t)$, onde γ é uma aplicação de X no espaço normado $\Lambda_1^m[0, T]$ de todas as funções vetoriais m-dimensionais, essencialmente limitadas, Lebesgue mensuráveis definidas em $[0, T]$, com a norma $\|\cdot\|_1$ defina por

$$\|y\|_1 = \max_{1 \leq j \leq m} \int_0^T |y_j(t)| dt,$$

e Γ é uma aplicação de X no espaço normado $\Lambda_1^1[0, T]$.

É fácil verificar que o problema acima pode ser colocado na forma do problema (PG). De fato basta colocar $E = L_\infty^n[0, T]$, $F = \Re$, $G = \Lambda_1^m[0, T]$, $f(x) = \phi(x)$, $g(x) = \gamma(x)$, e $K = \{y \in \Lambda_1^m[0, T] : y(t) \geq 0 \text{ qtp em } [0, T]\}$. Geralmente, na obtenção de condições necessárias de otimalidade para (PG), uma hipótese crucial é que o cone associado K seja convexo, fechado e com interior não vazio (veja [1], [23]). Infelizmente, no caso do cone K relacionado ao problema com tempo contínuo acima, esta hipótese não é satisfeita. Zalmai, numa série de artigos [63], [64], [65], [66], conseguiu solucionar esta questão, fazendo uso de um teorema de alternativa com tempo contínuo estabelecido por ele mesmo. Zalmai obteve, além de condições necessárias, resultados em dualidade e condições suficientes de primeira e segunda ordem, tudo isso num contexto diferenciável. Mais especificamente, em [66], Zalmai obteve condições suficientes de otimalidade de primeira ordem para (PNC), supondo que as funções envolvidas são continuamente diferenciáveis e fazendo uso de várias noções de convexidade generalizada tais como pseudo-convexidade e quasiconvexidade. Ele também obteve condições suficientes de otimalidade de segunda ordem, supondo que as funções são duas vezes continuamente diferenciáveis e fazendo uso da noção usual de Hessiano. Nosso ganho aqui foi estender os resultados de Zalmai em [66], da seguinte maneira: obtivemos condições suficientes de otimalidade de primeira ordem supondo apenas que as funções são Lipschitz e invexas. Também obtivemos condições suficientes de otimalidade de segunda ordem, supondo somente que as funções em (PNC) são de classe $C^{1,1}$, isto é, são funções diferenciáveis com gradiente Lipschitz.

Os principais resultados deste capítulo também podem ser encontrados em [52], [53].

1.3.5 Sobre o capítulo 5

No capítulo 5 faremos uma discussão a respeito de várias possibilidades de pesquisas futuras.

1.3.6 Sobre algumas observações

Nossos resultados originais (teoremas, corolários, proposições) estão basicamente nos capítulos 2, 3 e 4. Nestes capítulos quando nos referirmos a algum resultado que não seja nosso, colocaremos uma referência ao autor. Caso contrário não colocaremos nada. Por exemplo, o seguinte Teorema é um resultado usados por nós, devido a Clarke:

Teorema 2.1. (Clarke) *Suponha que $\overline{x}$ resolve (P). Então, para k suficientemente grande, existem $\lambda_0, \lambda_i \in \Re$, $i = 1, \ldots, m$, não todos nulos tais que*

$$0 \in \partial(\lambda_0 f_0 + \sum_{i=1}^{m} \lambda_i f_i + k d_C)(\overline{x}),$$

$$\lambda_i \geq 0, \; i = 1, \ldots, m,$$

$$\lambda_i f_i(\overline{x}) = 0, \; i = 1, \ldots, m.$$

Por outro lado, na proposição abaixo, nossa autoria está implícita, pois não existe nenhuma referência na frente da proposição:

Proposição 4.5. *Seja $\bar{x} \in \mathbb{F}$. Se existe $\bar{\lambda} \in L_\infty^m[0,T]$ tal que $(\bar{x}, \bar{\lambda})$ satisfaz (4.18)-(4.20), e se a função Lagrangeana $L(x;\bar{\lambda})$ é invexa em $\bar{x}$ (com respeito a $\mathbb{F}$), então $\bar{x}$ é uma solução ótima global de (PNC).*

Capítulo 2

Programação matemática

2.1 Introdução

Seja Ω um aberto não vazio do $\Re^n$, $f_0, f_i : \Omega \to \Re$, $i = 1, \ldots, m$, funções localmente Lipschitz em Ω, e $C \subset \Omega$ um conjunto fechado e não vazio. Consideremos o seguinte problema de programação matemática:

$$\left.\begin{array}{lc} \text{Minimizar} & f_0(x) \\ \\ \text{sujeito a} & f_i(x) \leq 0, \quad i \in I = \{1, \ldots, m\}, \\ \\ & x \in C. \end{array}\right\} \quad (P)$$

O conjunto dos pontos factíveis para o problema (P), (o qual suporemos não vazio) será dado por

$$\mathbb{F} := \{x \in C : f_i(x) \leq 0, i \in I\}.$$

Dado $\overline{x} \in \mathbb{F}$, seja $I(\overline{x})$ o conjunto dos índices correspondentes às restrições ativas de (P) em $\overline{x}$, isto é,

$$I(\overline{x}) = \{i \in I \; / \; f_i(\overline{x}) = 0\}.$$

Em ([16] pp. 227-230), Clarke obteve condições necessárias de otimalidade para um problema de programação matemática não diferenciável, na forma de uma regra de multiplicadores do tipo Fritz-John. O problema (P) acima é um caso particular daquele estudado por Clarke (o problema de Clarke continha restrições de igualdade). Logo, podemos estabelecer o seguinte teorema:

Teorema 2.1. **(Clarke)** *Suponha que $\bar{x}$ resolve (P). Então, para k suficientemente grande, existem $\lambda_0, \lambda_i \in \Re$, $i = 1, \ldots, m$, não todos nulos tais que*

$$0 \in \partial(\lambda_0 f_0 + \sum_{i=1}^{m} \lambda_i f_i + k d_C)(\bar{x}), \tag{2.1}$$

$$\lambda_i \geq 0, \ i = 1, \ldots, m, \tag{2.2}$$

$$\lambda_i f_i(\bar{x}) = 0, \ i = 1, \ldots, m. \tag{2.3}$$

Observemos que no Teorema acima o multiplicador associado à função objetivo pode ser nulo e neste caso o problema é dito *anormal*. Sem nenhuma hipótese suplementar não é possível garantir que $\lambda_0 \neq 0$. Como é usual, necessitamos impor alguma condição de regularidade sobre as restrições do problema para garantir que o multiplicador λ_0 seja não nulo. Neste caso o problema é dito *normal*. No caso diferenciável as condições de regularidade mais conhecidas talvez sejam as condições de Mangasarian-Fromovitz e de Slater [39]. Tais condições podem ser facilmente generalizadas para o caso não diferenciável. Impondo algumas das hipóteses citadas anteriormente, o problema (P) torna-se normal e é possível obter uma regra de multiplicadores do tipo Karush-Kuhn-Tucker, aonde podemos tomar $\lambda_0 = 1$. No nosso caso vamos obter um Teorema do tipo Karush-Kuhn-Tucker (o Teorema 2.2) supondo que no problema (P) as funções envolvidas são invexas, para a mesma função η, e que as restrições satisfazem uma condição do tipo Slater:

$$\text{existe } \hat{x} \in C \text{ tal que } f_i(\hat{x}) < 0, \quad i \in I(\bar{x}). \tag{2.4}$$

O motivo de tal escolha particular é bem conveniente (a priori poderíamos escolher qualquer condição de regularidade disponível na literatura como por exemplo as condições citadas acima): na seção 2.3, na demonstração do teorema de alternativa invexo (Teorema 2.8), vamos fazer uma aplicação do Teorema 2.2 na resolução de um problema de otimização invexo (problema $(\overline{P})$), o qual satisfaz a condição de Slater.

2.2 Condições de otimalidade e dualidade

Teorema 2.2. *(Condições necessárias de otimalidade) Seja $\overline{x}$ um ponto ótimo para (P) e suponha que $f_i, i \in I$ são invexas em $\overline{x}$, para o mesmo η. Se as restrições de (P) satisfazem a condição de Slater, então existem $\overline{\lambda} = (\overline{\lambda}_1, \ldots, \overline{\lambda}_m)$ satisfazendo as condições de Karush-Kuhn-Tucker*

$$0 \in \partial(f_0 + \sum_{i \in I} \bar{\lambda}_i f_i)(\overline{x}) + N_C(\overline{x}), \tag{2.5}$$

$$\bar{\lambda}_i \geq 0, \tag{2.6}$$

$$\bar{\lambda}_i f_i(\overline{x}) = 0, \tag{2.7}$$

Prova. Se $\overline{x}$ é uma solução de (P) então, pelo Teorema 2.1, existem $\lambda_0 \geq 0, \lambda_i \in \Re, i \in I$, não todos nulos tal que

$$0 \in \partial(\lambda_0 f_0 + \sum_{i \in I} \lambda_i f_i)(\overline{x}) + N_C(\overline{x}), \tag{2.8}$$

$$\lambda_i \geq 0, \tag{2.9}$$

$$\lambda_i f_i(\overline{x}) = 0. \tag{2.10}$$

Basta então provar que $\lambda_0 > 0$. Suponha que $\lambda_0 = 0$. Desde que existe $\hat{x} \in C$ tal que $f_i(\hat{x}) < 0, i \in I(\overline{x})$, e os λ_i não são todos nulos então segue-se de (2.9)-(2.10) que

$$\sum_{i \in I} \lambda_i f_i(\hat{x}) - \sum_{i \in I} \lambda_i f_i(\overline{x}) < 0. \tag{2.11}$$

Como $f_i, i \in I$, são funções localmente Lipschitz e invexas em $\overline{x}$, para uma função η comum, então $\sum_{i\in I} \lambda_i f_i$ é um função localmente Lipschitz e invexa em $\overline{x}$. Logo, existe $\eta(\widehat{x}, \overline{x}) \in T_C(\overline{x})$ tal que

$$\sum_{i\in I} \lambda_i f_i(\widehat{x}) - \sum_{i\in I} \lambda_i f_i(\overline{x}) \geq (\sum_{i\in I} \lambda_i f_i)^o(\overline{x}; \eta(\widehat{x}, \overline{x})). \tag{2.12}$$

Portanto, (2.11) junto com (2.12) implicam que

$$\text{existe } \eta(\widehat{x}, \overline{x}) \in T_C(\overline{x}) \text{ tal que } (\sum_{i\in I} \lambda_i f_i)^o(\overline{x}; \eta(\widehat{x}, \overline{x})) < 0. \tag{2.13}$$

Mas, (2.8) implica que

$$(\sum_{i\in I} \lambda_i f_i)^o(\overline{x}; \eta) \geq 0, \text{ for all } \eta \in T_C(\overline{x}),$$

o que contradiz (2.13). Portanto $\lambda_0 > 0$ e o Teorema segue fazendo $\overline{\lambda}_i = \lambda_i/\lambda_0$, $i \in I$. □

Teorema 2.3. *(Condições suficientes de otimalidade) Dado $\overline{x} \in \mathbb{F}$, suponha que f_0, f_i sejam invexas em $\overline{x}$, com relação a C, para um mesmo η, e que existam $\overline{\lambda}_i \in \Re$, $i \in I$ tal que valem as condições K-K-T (2.5)-(2.7). Então $\overline{x}$ é mínimo global para o problema (P).*

Prova. É óbvio que se $\overline{\lambda}_i \geq 0$ e f_0, f_i são invexas em $\overline{x}$ com respeito a C para o mesmo η, então $f_0 + \sum_{i\in I} \overline{\lambda}_i f_i$ é invexa em $\overline{x}$ com relação a C. Logo, segue-se de (2.5) e do Teorema 1.13 que $\overline{x}$ é mínimo de $f_0 + \sum_{i\in I} \overline{\lambda}_i f_i$ sobre C, isto é,

$$f_0(x) + \sum_{i\in I} \overline{\lambda}_i f_i(x) \geq f_0(\overline{x}) + \sum_{i\in I} \overline{\lambda}_i f_i(\overline{x}), \quad \forall x \in C. \tag{2.14}$$

Então, (2.7) e (2.14) implicam que

$$f_0(x) + \sum_{i\in I} \overline{\lambda}_i f_i(x) \geq f_0(\overline{x}), \quad \forall x \in C. \tag{2.15}$$

Desde que $\overline{\lambda}_i \geq 0$, por (2.15), concluímos que

$$f_0(x) \geq f_0(\overline{x}), \quad \forall x \in C \text{ tal que } f_i(x) \leq 0, i \in I, \tag{2.16}$$

ou seja, $\overline{x}$ é solução do problema (P). □

Obs. Notemos que se no problema (P) as funções envolvidas são invexas em $\overline{x}$ e satisfazem alguma hipótese de regularidade (Slater por exemplo) então, no Teorema acima, tanto faz usar a condição (2.5) quanto as condições $0 \in \partial(f_0 + \sum_{i\in I}\overline{\lambda}_i f_i + kd_C)(\overline{x})$ e $0 \in \partial f_0(\overline{x}) + \sum_{i\in I}\overline{\lambda}_i \partial f_i(\overline{x}) + N_C(\overline{x})$. De fato, chamemos $X = \{x \in E : 0 \in \partial(f_0 + \sum_{i\in I}\overline{\lambda}_i f_i + kd_C)(x)\}$, $Y = \{x \in E : 0 \in \partial(f_0 + \sum_{i\in I}\overline{\lambda}_i f_i)(x) + N_C(x)\}$ e $Z = \{x \in E : 0 \in \partial f_0(x) + \sum_{i\in I}\overline{\lambda}_i \partial f_i(x) + N_C(x)\}$. As inclusões $X \subset Y \subset Z$ são bem conhecidas. Agora, se $\overline{x} \in Z$ então, pelo Teorema 2.3, $\overline{x}$ resolve (P). Mas se $\overline{x}$ resolve (P) então, para k suficientemente grande, $\overline{x} \in X$. Portanto, nessas condições, $X = Y = Z$. Em [49] Reiland obtém condições suficientes de otimalidade para um problema similar a (P) separando em dois casos conforme x pertença a Y ou Z. Nossa observação mostra que isto é desnecessário. Além disso mostramos outra propriedade de funções convexas preservada por funções invexas.

Definição 2.4. *Dizemos que o ponto $(\overline{x}, \overline{\lambda})$, $\overline{\lambda} \geq 0$, é uma solução ponto de sela para (P) se*

$$f_0(\overline{x}) + \sum_{i\in I}\lambda_i f_i(\overline{x}) \leq f_0(\overline{x}) + \sum_{i\in I}\overline{\lambda}_i f_i(\overline{x}) \leq f_0(x) + \sum_{i\in I}\overline{\lambda}_i f_i(x), \tag{2.17}$$

para todo $x \in C$ e $\lambda \geq 0$.

Teorema 2.5. *(Ponto de Sela Escalar) Se $\overline{x}$ é ótimo para (P), as restrições do problema satisfazem alguma condição de regularidade e $f_0 + \sum_{i\in I}\lambda_i f_i$ é invexa em $\overline{x}$ com relação à C para todo $\lambda_i \geq 0$, então existe $\overline{\lambda} \geq 0$ tal que $(\overline{x}, \overline{\lambda})$ é uma solução ponto de sela de (P).*

Prova. Se $\overline{x}$ é ótimo e vale alguma condição de regularidade nas restrições de (P) (A condição de Slater por exemplo) então, pelo Teorema 2.2, existem $\overline{\lambda}_i \in \Re$, $i \in I$ tal que valem as condições K-K-T (2.5)-(2.7). Portanto, (2.5) mais a hipótese de invexidade implicam, pelo Teorema 1.13, que $\overline{x}$ é ponto de mínimo para $f_0 + \sum_{i\in I}\overline{\lambda}_i f_i$ sobre C, isto é, vale a desigualdade no lado direito de (2.17). A desigualdade no lado esquerdo de (2.17) vale por (2.7), $f_i(\overline{x}) \leq 0$ e $\lambda \geq 0$. □

Vamos estabelecer agora resultados de dualidade para (P). Consideremos o seguinte problema dual (do tipo Wolfe) relacionado a (P):

$$\left.\begin{array}{lc} \text{Maximizar} & f_0(w) + \sum_{i \in I} \lambda_i f_i(w) \\ \text{sujeito a} & (w, \lambda) \in C \times \Re^m_+, \\ & 0 \in \partial(f_0 + \sum_{i \in I} \lambda_i f_i)(w) + N_C(w). \end{array}\right\} (D)$$

Teorema 2.6. *(Dualidade fraca). Se f_0 e $f_i, i \in I$, são funções invexas em C, com uma função η comum a todas então:*

$$f_0(x) \geq f_0(w) + \sum_{i \in I} \lambda_i f_i(w),$$

para todo x factível para (P) e (w, λ) factível para (D).

Prova. Seja x factível para (P) e (w, λ) factível para (D). Já que $\lambda \in \Re^m_+$ e $x \in I\!\!F$ então

$$f_0(x) - f_0(w) - \sum_{i \in I} \lambda_i f_i(w) \geq f_0(x) + \sum_{i \in I} \lambda_i f_i(x) - f_0(w) - \sum_{i \in I} \lambda_i f_i(w). \quad (2.18)$$

Como $\lambda \in \Re^m_+$ e f_0, f_i são invexas em C então $f_0 + \sum_{i \in I} \lambda_i f_i$ é invexa em C. Portanto, a condição $0 \in \partial(f_0 + \sum_{i \in I} \lambda_i f_i)(w) + N_C(w)$ implica, pelo Teorema 1.13, que

$$f_0(x) + \sum_{i \in I} \lambda_i f_i(x) \geq f_0(w) + \sum_{i \in I} \lambda_i f_i(w), \text{ for all } x \in I\!\!F. \quad (2.19)$$

O resultado segue facilmente por (2.18) e (2.19). □

Teorema 2.7. *(Dualidade forte). Suponha que f_0 e $f_i, i \in I$ são funções invexas em C, para a mesma função η. Suponha que $\overline{x}$ é uma solução ótima de (P) e que existe $\bar{\lambda} \in \Re^m$ tal que as condições K-K-T (2.5)-(2.7) são satisfeitas em $(\overline{x}, \bar{\lambda})$. Então (D) é maximizado em $(\overline{x}, \bar{\lambda})$ e os valores ótimos de (P) e (D) coincidem.*

Prova. Seja $\overline{x}$ uma solução ótima para (P). Se existe $\bar{\lambda}$ tal que $(\overline{x}, \bar{\lambda})$ satisfaz as condições de K-K-T (2.5)-(2.7) então $(\overline{x}, \bar{\lambda})$ é factível para (D) e, pelo resultado de dualidade fraca, concluímos que

$$f_0(\overline{x}) + \sum_{i \in I} \bar{\lambda}_i f_i(\overline{x}) = f_0(\overline{x}) \geq f_0(x) + \sum_{i \in I} \bar{\lambda}_i f_i(x),$$

para todo (x, λ) factível para (D). Então $(\overline{x}, \bar{\lambda})$ é ótimo para (D) e os valores ótimos de (P) e (D) coincidem. □

2.3 Sobre um teorema de alternativa invexo

Em ([26], p. 34) Geoffrion obteve uma nova demonstração do conhecido Teorema de Gordan generalizado (teorema de alternativa para funções convexas), fazendo uso de resultados de dualidade entre problemas de programação convexa. Seguindo sua técnica usaremos os resultados de dualidade invexa obtidos na seção anterior, para estabelecermos um teorema de alternativa do tipo Gordan para funções invexas e localmente Lipschitz, o qual é nosso principal resultado nesse capítulo.

Teorema 2.8. *Seja C um subconjunto fechado e não vazio de Ω. Suponha que $f_i : \Omega \to \Re, i \in I = \{1, \ldots, m\}$, são localmente Lipschitz e invexas em C, para um mesmo η. Suponha ainda que*

$$f(x) = \max_{i \in I} f_i(x) \text{ atinge um mínimo em } C. \tag{2.20}$$

Então,

$$\text{ou existe } x \in C \text{ tal que } f_i(x) < 0, i \in I, \tag{2.21}$$

ou existem $\bar{\lambda}_i \geq 0, i \in I$, tal que

$$\sum_{i \in I} \bar{\lambda}_i = 1 \quad e \quad \inf_{x \in C} \sum_{i \in I} \bar{\lambda}_i f_i(x) \geq 0. \tag{2.22}$$

Prova. Primeiro vamos provar a parte fácil. Suponha que (2.21) tem uma solução $\tilde{x} \in C$. Então, qualquer que sejam $\lambda_i \geq 0$, não todos nulos,

$$\sum_{i\in I} \lambda_i f_i(\tilde{x}) < 0.$$

Logo, (2.22) não pode ter solução. Notemos que nessa parte da demonstração não utilizamos a hipótese (2.20).

Reciprocamente, suponha que (2.21) não tenha solução e consideremos o seguinte problema auxiliar:

$$\left.\begin{array}{ll} \text{Minimizar} & F_0(x,y) := y \\ \\ \text{sujeito a} & (x,y) \in C \times \Re, \\ \\ & F_i(x,y) := f_i(x) - y \leq 0, \quad i \in I. \end{array}\right\} \quad (\bar{P})$$

Desde que f_i são funções localmente Lipschitz em Ω então concluímos que F_0, F_i são localmente Lipschitz em $\Omega \times \Re$. As derivadas de Clarke de F_0, F_i em $(x,y) \in \Omega \times \Re$ na direção $(u,v) \in \Re^n \times \Re$ são dadas por

$$F_0^o(x,y;u,v) = v,$$

$$F_i^o(x,y;u,v) = f_i^o(x,u) - v, i \in I.$$

Sejam $(x,y),(x',y') \in C \times \Re$. Então, pelas hipóteses de invexidade,

$$\begin{aligned} F_i(x',y') - F_i(x,y) &= f_i(x') - y' - f_i(x) + y \\ &\geq f_i^o(x;\eta(x',x)) - (y'-y) \\ &= F_i^o(x,y;\eta(x',x),y'-y), \end{aligned}$$

$$\begin{aligned} F_0(x',y') - F_0(x,y) &= y' - y \\ &= F_0^o(x,y;\eta(x',x),y'-y). \end{aligned}$$

Desde que $\eta(x',x) \in T_C(x)$ e $y'-y \in T_\Re(y)$ então $(\eta(x',x),y'-y) \in T_C(x) \times T_\Re(y) = T_{C\times\Re}(x,y)$. Portanto, dados $(x,y),(x',y') \in C \times \Re$ existe $\tilde{\eta}(x',y';x,y) = (\eta(x',x),y'-y) \in T_{C\times\Re}(x,y)$ tal que

$$F_i(x',y') - F_i(x,y) \geq F_i^o(x,y;\tilde{\eta}(x',y';x,y)), \quad i \in I \cup \{0\}, \tag{2.23}$$

isto é, $F_0, F_i, i \in I$, são funções invexas em $C \times \Re$, para um $\tilde{\eta}$ comum.

Vamos mostrar que se (2.20) vale então o problema acima tem pelo menos uma solução. Primeiramente notemos que o conjunto $E(f) = \{(x,y) \in C \times \Re : f(x) \leq y\}$ coincide com os pontos factíveis de $(\overline{P})$ ($E(f)$ é chamado o epigrafo de f). Suponhamos que vale (2.20). Então, existe $\tilde{x} \in C$ tal que

$$\tilde{y} = f(\tilde{x}) = \max_{x \in C} f(x).$$

Como $(\tilde{x}, \tilde{y}) \in C \times \Re$ satisfaz $f(\tilde{x}) - \tilde{y} = 0$ então $(\tilde{x}, \tilde{y})$ é factível para $(\overline{P})$. Além disso, qualquer que seja (x, y) factível para $(\overline{P})$,

$$\tilde{y} = f(\tilde{x}) \leq f(x) \leq y,$$

ou seja

$$F_0(\tilde{x}, \tilde{y}) \leq F_0(x, y).$$

Portanto, o problema $(\overline{P})$ tem solução.

Seja $(\overline{x}, \overline{y})$ uma solução ótima de $(\overline{P})$. As restrições do problema $(\overline{P})$ satisfazem a condição de Slater. De fato, se tomarmos $(\widehat{x}, \widehat{y}) = (\overline{x}, \overline{y} + 1)$, então, para $i \in I(\overline{x}, \overline{y})$, $F_i(\widehat{x}, \widehat{y}) = f_i(\widehat{x}) - \widehat{y} = f_i(\overline{x}) - (\overline{y} + 1) = F_i(\overline{x}, \overline{y}) - 1 < 0$.

O problema dual relacionado a $(\bar{P})$ é dado por:

$$\left.\begin{array}{ll} \text{Maximizar} & L(x,y,\lambda) := y + \sum_{i \in I} \lambda_i (f_i(x) - y) \\[1em] \text{sujeito a} & (x, y, \lambda) \in C \times \Re \times \Re^m_+, \\[1em] & 0 \in \partial_{(x,y)} L(x,y,\lambda) + N_{C \times \Re}(x,y). \end{array}\right\} (\bar{D})$$

Desde que as hipótese de invexidade e a condição de Slater são satisfeitas, concluímos, pelo Teorema 2.2, que existe $\bar{\lambda} \in \Re^m$ tal que $(\overline{x}, \overline{y}, \overline{\lambda})$ satisfaz as condições de K-K-T. Portanto, todas as hipótese do Teorema 2.7 (Dualidade forte) estão em vigor para $(\bar{P})$ e segue que o programa dual $(\bar{D})$ tem uma solução ótima $(\bar{x}, \bar{y}, \bar{\lambda})$ e os valores $F_0(\overline{x}, \overline{y})$ e $L(\bar{x}, \bar{y}, \bar{\lambda})$ são iguais.

Se o sistema de desigualdades (2.21) não tem solução em C então segue-se que $F_0(x, y) \geq 0$ para todo (x, y) factível para $(\bar{P})$. De fato, se existisse

$(x,y) \in C \times \Re$ tal que $f_i(x) - y \leq 0, i \in I$, $F_0(x,y) = y < 0$, então existiria $x \in C$ tal que $f_i(x) \leq y < 0$, o que contradiz (2.21). Logo, o programa $(\bar{P})$ tem valor ínfimo $F_0(\overline{x},\overline{y}) = \overline{y} \geq 0$ e pelos resultados de dualidade isto implica que $L(\bar{x},\bar{y},\bar{\lambda}) \geq 0$.

Consideremos a função $\Phi : \Omega \times \Re \to \Re$ dada por $\Phi(x,y) = L(x,y,\bar{\lambda})$. Φ é obviamente uma função localmente Lipschitz invexa em $C \times \Re$ e satisfaz $0 \in \partial\Phi(\bar{x},\bar{y}) + N_{C\times\Re}(\bar{x},\bar{y})$. Portanto, pelo Teorema 1.13, concluímos que $(\bar{x},\bar{y})$ é um mínimo global para Φ em $C \times \Re$, isto é,

$$0 \leq \Phi(\bar{x},\bar{y}) = \inf_{(x,y)\in C\times\Re} \Phi(x,y).$$

Mas isto é o mesmo que

$$\begin{aligned} 0 \leq\ & \inf_{(x,y)\in C\times\Re}\{y + \textstyle\sum_{i\in I}\bar{\lambda}_i(f_i(x) - y)\} \\ =\ & \inf_{x\in C}\{\textstyle\sum_{i\in I}\bar{\lambda}_i f_i(x)\} + \inf_{y\in\Re}\{y(1 - \textstyle\sum_{i\in I}\bar{\lambda}_i)\}. \end{aligned}$$

O segundo termo acima deve ser 0, isto é, $\sum_{i\in I}\bar{\lambda}_i = 1$ deve valer, pois do contrário seu valor seria $-\infty$. Logo, o primeiro termo é não negativo e o teorema está provado. □

Observação 2.9 É óbvio que se C é um conjunto compacto então a condição (2.20) é válida. Outra hipótese que implica (2.20) é a de que pelo menos uma das funções f_i seja inf-limitada, isto é, exista um índice $i \in I$ tal que o conjunto de níveis $L_\alpha f_i = \{x \in C : f_i(x) \leq \alpha\}$ seja compacto para todo $\alpha \in \Re$.

Agora, apresentaremos um teorema do tipo Bohnenblust-Karlin-Shapley ([39], p. 67), o qual pode envolver uma família, possivelmente infinita, de funções invexas localmente Lipschitz.

Teorema 2.10. *Seja $C \subset \Omega$ um conjunto compacto e $f_i : \Omega \to \Re, i \in M$ uma família (finita ou infinita) de funções invexas localmente Lipschitz em C. Se*

$$f_i(x) \leq 0, i \in M, \text{ não tem solução } x \in C,$$

então, para alguma subfamília finita $f_{i_j}, j = 1, \ldots, k$, existe $p_j \geq 0$ tal que

$$\sum_{j=1}^{k} p_j = 1 \quad e \quad \inf_{x \in C} \sum_{j=1}^{k} p_j f_{i_j}(x) > 0.$$

Prova. Afirmamos que o sistema

$$f_i(x) \leq \varepsilon, \forall i \in M, \forall \varepsilon > 0, \tag{2.24}$$

não tem solução $x \in C$. De fato, se tivesse uma solução $\overline{x}$, então $f_i(\overline{x}) \leq \varepsilon$, para todo $i \in M, \varepsilon > 0$ e isto implica que $f_i(\overline{x}) \leq 0$ para todo $i \in M$ (pois, caso contrário se $f_{i_0}(\overline{x}) > 0$ para algum $i_0 \in M$ então colocando $\varepsilon_0 = 1/2 f_{i_0}(\overline{x}) > 0$, teríamos que existe $i_0 \in M$ e $\varepsilon_0 > 0$ tal que $f_{i_0}(\overline{x}) > \varepsilon_0$, o que nos levaria a uma contradição). Isto, entretanto, contradiz a hipótese do teorema e portanto, o sistema (2.24) não tem solução $x \in C$. Os conjuntos

$$C(i, \varepsilon) = \{x \in C : f_i(x) \leq \varepsilon\}$$

são subconjuntos fechados do compacto C, e sua interseção é vazia. Logo, pelo Teorema da interseção finita (see [39], p. 189), existe um número finito de tais conjuntos tal que sua interseção é vazia. Assim, nós obtemos índices $i_j, j = 1, \ldots, k$ e números reais $\varepsilon_j > 0, j = 1, \ldots, k$, tal que o sistema

$$f_{i_j}(x) - \varepsilon_j \leq 0, \quad j = 1, \ldots, k,$$

não tem solução $x \in C$. Portanto, pelo Teorema 2.8, existem $p_j \geq 0$ tal que

$$\sum_{j=1}^{k} p_j = 1 \quad \text{e} \quad \sum_{j=1}^{k} p_j f_{i_j}(x) \geq \sum_{j=1}^{k} p_j \varepsilon_j > 0,$$

como queríamos demonstrar. □

2.4 Programação multiobjetivo

Nesta seção, daremos uma interessante aplicação dos resultados da seção anterior, caracterizando as soluções de um problema de otimização vetorial.

Antes de estabelecermos o resultado propriamente dito, vamos falar um pouco sobre o assunto bem como introduzir alguns conceitos de solução para problemas de otimização vetorial.

Muitos problemas de tomada de decisão com múltiplos critérios podem ser modelados na forma de um problema de otimização vetorial que consiste, basicamente, em maximizar ou minimizar várias (e as vezes conflitantes) funções objetivo, sujeitas a alguma restrição das variáveis envolvidas no problema. Tais problemas tem outras denominações como, por exemplo, otimização multicritério, programação multiobjetivo, problemas de ótimo de Pareto (para maiores detalhes veja por exemplo [15]). No entanto, na formulação de muitos modelos, aparecem funções não muito bem comportadas, não tendo boas propriedades tais como diferenciabilidade ou convexidade. Neste caso a teoria existente é relativamente pobre para uma boa caracterização das soluções de tais problemas. Seguindo o mesmo caminho da teoria clássica tentaremos relacionar as soluções do programa multiobjetivo com as soluções ótimas de algum programa escalar. Como veremos adiante, o Teorema de alternativa invexo estabelecido no capítulo anterior será essencial na obtenção de tais resultados.

Consideremos o seguinte problema de programação multiobjetivo:

$$\left.\begin{array}{ll}\text{Minimizar} & f(x)=(f_1(x),\ldots,f_m(x))\\[1ex] \text{sujeito a} & x\in C,\end{array}\right\}(PV)$$

onde $C \subset \Omega$ é um subconjunto não vazio e $f : \Omega \to \Re^m$ é uma função vetorial dada.

Em 1896, Pareto [46] introduziu o seguinte conceito de solução para problemas vetoriais. De acordo com ele, um ponto factível para (PV) é uma solução eficiente (ou não dominada, não inferior ou ótima de Pareto) de (PV), se não existe outra solução factível que melhore um objetivo sem degradar pelo menos um dos outros. Tal conceito pode ser formalizado da seguinte maneira:

Definição 2.11. *Dizemos que $\bar{x}$ é uma solução eficiente para (PV) se*

$\overline{x}$ é factível para (PV) e se não existe outro x factível para (PV) tal que $f_i(x) \leq f_i(\overline{x})$, para todo $i = 1, \dots m$, com a desigualdade estrita valendo para pelo menos um índice i.

Em [25] Geoffrion introduziu o conceito de solução propriamente eficiente:

Definição 2.12. *Dizemos que $\overline{x}$ é uma solução propriamente eficiente de (PV) se ela é eficiente e se existe um escalar $M > 0$ tal que, para cada i,*

$$\frac{f_i(\overline{x}) - f_i(x)}{f_j(x) - f_j(\overline{x})} \leq M$$

para algum j tal que $f_j(x) > f_j(\overline{x})$ sempre que x é factível para (PV) e $f_i(x) < f_i(\overline{x})$.

Observação 2.13 O conceito de solução propriamente eficiente de Geoffrion elimina certas soluções eficientes indesejáveis, cujo raio da taxa de ganho de um objetivo em relação à taxa de perda de outro objetivo é ilimitado. De fato consideremos o seguinte programa:

$$\text{Minimizar} \quad f(x) = (f_1(x) = -x^2, f_2(x) = x^3)$$

$$\text{sujeito a} \quad x \in C = \{x \in \Re : x \geq 0\}.$$

O ponto $x = 0$ é eficiente mas não é propriamente eficiente. De fato, o raio

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

é ilimitado quando x tende a zero pela direita.

Como já dissemos na introdução, a estratégia mais comum para caracterizar tais soluções é tentar relacioná-las com as soluções ótimas de um problema escalar. Existem várias maneiras de escalarizar o problema (PV) (veja [15]). Em nosso trabalho nos ateremos a uma escalarização usada por Geoffrion. Para isso consideremos o seguinte problema escalar associado a

(PV):

$$\left.\begin{array}{r}\text{Minimizar } \langle \mu, f(x)\rangle = \sum_{i=1}^{m} \mu_i f_i(x) \\ \\ \text{sujeito a } x \in C,\end{array}\right\} \quad (P(\mu))$$

onde $\mu \in \Lambda^+ = \{\mu \in \Re_+^m : \mu_i > 0,\ i \in I,\ \sum_i^m \mu_i = 1\}$.

Geoffrion [25] estabeleceu o seguinte resultado fundamental:

Teorema 2.14. (Geoffrion) *Sejam $\mu_i > 0$, $i \in I$, fixados. Se $\overline{x}$ é uma solução ótima de $(P(\mu))$, então $\overline{x}$ é uma solução propriamente eficiente de (PV).*

Supondo que f_i são funções convexas e C é um conjunto convexo, Geoffrion também estabeleceu a recíproca do Teorema 2.14. Este resultado é baseado em uma propriedade fundamental de funções convexas ([4], p.62, Teorema fundamental) o que na prática corresponde ao Teorema de Gordan generalizado. Então, trocando este Teorema fundamental pelo nosso Teorema 2.8, também obteremos uma recíproca do Teorema 2.14 numa versão invexa:

Teorema 2.15. *Suponha que C é um conjunto compacto, e que f_i são funções localmente Lipschitz e invexas em C, para o mesmo η. Se $\overline{x}$ é uma solução propriamente eficiente de (VP) então $\overline{x}$ é uma solução ótima de $(P(\mu))$ para algum μ com componentes estritamente positivas.*

Prova. Se $\overline{x}$ é uma solução propriamente eficiente de (PV) então existe $M > 0$ tal que, para cada $i = 1, \ldots, m$, o sistema

$$f_i(x) - f_i(\overline{x}) < 0,$$

$$f_i(x) + M f_j(x) - f_i(\overline{x}) - M f_j(\overline{x}) < 0, \ \ \forall j \neq i,$$

não admite solução $x \in C$. Portanto, pelo Teorema 2.8, para o $i-$ésimo sistema, existem $\lambda_j^i \geq 0$ $(j = 1, \ldots, m)$ com $\sum_{j=1}^{m} \lambda_j^i = 1$, tal que, $\forall x \in C$,

$$\lambda_i^i(f_i(x) - f_i(\overline{x})) + \sum_{j\neq i} \lambda_j^i(f_i(x) + Mf_j(x) - f_i(\overline{x}) - Mf_j(\overline{x})) \geq 0. \qquad (2.25)$$

Como $\sum_{j=1}^{m} \lambda_j^i = 1$ então (2.25) é equivalente a

$$f_i(x) + M\sum_{j\neq i} \lambda_j^i f_j(x) \geq f_i(\overline{x}) + M\sum_{j\neq i} \lambda_j^i f_j(\overline{x}), \quad \forall x \in C.$$

A desigualdade acima vale para cada i, isto é,

$$f_1(x) + M\sum_{j\neq 1} \lambda_j^1 f_j(x) \geq f_1(\overline{x}) + M\sum_{j\neq 1} \lambda_j^1 f_j(\overline{x}), \quad \forall x \in C,$$

$$f_2(x) + M\sum_{j\neq 2} \lambda_j^2 f_j(x) \geq f_2(\overline{x}) + M\sum_{j\neq 2} \lambda_j^2 f_j(\overline{x}), \quad \forall x \in C,$$

$$\vdots$$

$$f_m(x) + M\sum_{j\neq m} \lambda_j^m f_j(x) \geq f_m(\overline{x}) + M\sum_{j\neq m} \lambda_j^m f_j(\overline{x}), \quad \forall x \in C.$$

Somando as desigualdades acima e rearranjando os termos, obtemos

$$\sum_{j=1}^{m}(1 + M\sum_{i\neq j} \lambda_j^i) f_j(x) \geq \sum_{j=1}^{m}(1 + M\sum_{i\neq j} \lambda_j^i) f_j(\overline{x}), \quad \forall x \in C,$$

o que completa a prova. □

De posse do Teorema acima podemos aplicar resultados conhecidos sobre problemas usuais de programação matemática (tanto do ponto de vista teórico quanto computacional) à problemas de programação multiobjetivo.

Em [59], Weir e Mond obtiveram um Teorema de Alternativa pré-invexo e estabeleceram a recíproca do Teorema 2.14 no caso em que f é uma função pré-invexa.

Daremos agora uma outra aplicação dos Teoremas da seção anterior, desta vez na busca de uma solução Ponto de Sela de um problema derivado do problema (PV). Tal problema é explicitado abaixo:

$$\left.\begin{array}{c}\text{Minimizar } f(x)\\ \\ \text{sujeito a } x \in C = \{x \in S : g_j(x) \leq 0, j \in J = \{1, \ldots, k\}\}.\end{array}\right\} \quad (PV')$$

where $C \subset \Omega$ é um conjunto compacto e não vazio (isto é verdade, por exemplo, se S é um compacto não vazio), $g_j : \Omega \to \Re$ são funções localmente Lipschitz.

Consideremos o seguinte Lagrangiano associado a (PV') [59]:

$$L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle e,$$

onde $e = (1, \ldots, 1) \in \Re^m$, $g(x) = (g_1(x), \ldots, g_k(x))$, e $\lambda \in \Re^k$.

Nós adotaremos a seguinte convenção para vetores no $\Re^m$: se $u, v \in \Re^m$ então $u \succeq v$ é equivalente a $u_i \geq v_i$, para todo $i = 1, \ldots, m$ e $u \neq v$; $u \not\succeq v$ é a negação de $u \succeq v$.

O problema do ponto de sela é o problema de encontrar $\overline{x} \in S, \overline{\lambda} \in \Re^k_+$, tal que

$$f(\overline{x}) + \langle \lambda, g(\overline{x}) \rangle e \not\succeq f(\overline{x}) + \langle \overline{\lambda}, g(\overline{x}) \rangle e, \tag{2.26}$$

$$f(\overline{x}) + \langle \overline{\lambda}, g(\overline{x}) \rangle e \not\succeq f(x) + \langle \overline{\lambda}, g(x) \rangle e, \tag{2.27}$$

para todo $x \in S, \lambda \in \Re^k_+$.

Theorem 2.16. *(Ponto de sela vetorial) Seja $\overline{x}$ uma solução propriamente eficiente de (PV'). Suponha que as restrições de (PV') satisfaçam a* *condição de Slater e $f_i, i \in I$ e $g_j, j \in J$ sejam invexas em S para o mesmo η. Então, existe $\overline{\lambda} \in \Re^k_+$ tal que $(\overline{x}, \overline{\lambda})$ é uma solução ponto de sela de (PV')*

Prova. Desde que $\overline{x}$ é uma solução propriamente eficiente de (PV') então, pelo Teorema 2.15, $\overline{x}$ minimiza $\langle \mu, f(x) \rangle$ sujeito a $x \in C, g_i(x) \leq 0$, para algum $\mu \in \Lambda^+$. Como as hipóteses de invexidade e a condição de Slater são satisfeitas então pelo Teorema 2.5 (Ponto de sela escalar), existe $\overline{\lambda} \in \Re^k_+$, tal que $\langle \overline{\lambda}, g(\overline{x}) \rangle = 0$ e

$$\phi(\overline{x}, \lambda) \leq \phi(\overline{x}, \overline{\lambda}) \leq \phi(x, \overline{\lambda}), \tag{2.28}$$

onde $\phi(x, \lambda) = \langle \mu, f(x) + \langle \lambda, g(x) \rangle e \rangle$. Logo, se (2.26) não fosse verdadeira então para algum $i \in I$

$$f_i(\overline{x}) + \langle \lambda, g(\overline{x}) \rangle > f_i(\overline{x}) + \langle \overline{\lambda}, g(\overline{x}) \rangle \tag{2.29}$$

e

$$f_j(\overline{x}) + \langle \lambda, g(\overline{x}) \rangle \geq f_j(\overline{x}) + \langle \overline{\lambda}, g(\overline{x}) \rangle, \text{ for all } j \neq i. \tag{2.30}$$

Multiplicando por $\mu_i, i \in I$, e somando sobre todos os valores de i chegaríamos numa contradição com (2.28). Um argumento similar aplicado a (2.27) também nos levaria a uma contradição. □

Obs.: Com algumas adaptações os resultados obtidos nesse capítulo podem ser estendidos para espaços de Banach não necessariamente finito dimensionais. Por exemplo, se E é um espaço de Banach infinito dimensional, a bola unitária fechada em E não é um conjunto compacto na topologia forte associada à norma (veja [13]). Nesse caso a hipótese de compacidade forte sobre um conjunto de restrições de algum problema (por exemplo o Teorema 2.15) é muito restritiva. Torna-se necessário impor hipóteses de compacidade associadas a alguma topologia mais fraca.

Os principais resultados deste capítulo também podem ser encontrados em [10], [11].

Capítulo 3

Otimização vetorial entre espaços de Banach

3.1 Introdução

Consideremos novamente o problema de otimização entre espaços de Banach (PG):

$$\left.\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & -g(x) \in K \\ & x \in C. \end{array}\right\} \quad (PG)$$

Vamos impor agora, algumas hipóteses adicionais. De fato vamos supor que E, F e G são espaços de Banach; $C \subset E$ é um subconjunto fechado e não vazio de E; $Q \subset F$ e $K \subset G$ são cones convexos fechados e com interior não vazio.

Este problema tem sido muito estudado nos últimos anos sob diferentes pontos de vista. No caso em que F e G são espaços de Banach finito dimensionais e as funções envolvidas são localmente Lipschitz, o problema (PG) foi estudado por [16], [40] (condições necessárias de otimalidade), [41] (condições suficientes de otimalidade e dualidade) e por nós, no final do último capítulo desta tese (condições suficientes de otimalidade com invexidade). O Caso Lipschitz infinito dimensional foi considerado em [1] (condições necessárias) e [17] (condições suficientes e dualidade) . Em [17] o problema foi estudado na ausência da restrição $-g(x) \in K$ e alguma regularidade sobre f e C foi assumida. O estudo do problema (PG) e os resultados deste capítulo foram inspirados principalmente no trabalho de Abdoni e Thibault [1]. Neste artigo eles estudaram um problema de programação matemática multiobjetivo mais geral do que o problema (PG). Eles consideraram também restrições de igualdade, mas com espaço imagem finito dimensional. Sob a hipótese de que as funções envolvidas são fortemente compactamente Lipschitz e sem nenhum outra hipótese de diferenciabilidade eles obtiveram condições necessárias de otimalidade do tipo Fritz-John para seu problema na forma de uma regra de multiplicadores envolvendo as derivadas generalizadas de Ioffe [32] e de Clarke [16].

Nosso objetivo é, sob as mesmas hipóteses básicas de [1], obter condições necessárias de otimalidade do tipo KKT e condições suficientes de otimalidade para o problema (PG), utilizando o conceito de invexidade apresentado no primeiro capítulo, adaptado para aplicações entre espaços de Banach. Resultados relacionados utilizando outros conceitos de convexidade generalizada podem ser encontrados em [36] e [44].

Na próxima seção detalharemos melhor algumas questões relacionadas a (PG). Na última seção obteremos os principais resultados relativos a (PG), isto é, condições de otimalidade e dualidade.

3.2 Preliminares técnicos

Nessa seção vamos supor que A e B são dois espaços de Banach, $H \subset B$ é um cone convexo, fechado e com interior não vazio, satisfazendo $H \cap \{-H\} = \{0\}$. Vamos denotar por H^* o seu cone dual:

$$H^* = \{w^* \in B^* : \langle w^*, y\rangle \geq 0,\ \forall y \in H\}.$$

Uma multifunção $\Gamma : A \to B$ é uma aplicação de A em subconjuntos não vazios de B. Dizemos que Γ é semicontínua superior em $x \in A$ se para todo $\varepsilon > 0$, existe um $\delta > 0$ tal que

$$\Gamma(x') \subset \Gamma(x) + \varepsilon\mathcal{B}_B, \text{ para todo } x' \in x + \delta\mathcal{B}_A.$$

Dizemos que uma aplicação $\phi : A \to B$ é fortemente compactamente Lipschitz em $\bar{x} \in A$ se existe uma multifunção $R : A \to Comp(B)$, onde $Comp(B)$ denota o conjunto de todos os subconjuntos compactos não vazios de B, e uma função $r : A \times A \to \Re_+$ satisfazendo

(i) $\lim_{x \to \bar{x}, d \to 0} r(x, d) = 0$;

(ii) Existe $\alpha > 0$ tal que

$$t^{-1}[\phi(x + td) - \phi(x)] \in R(d) + \|d\| r(x, t)\bar{\mathcal{B}}_B,$$

para todo $x \in \bar{x} + \alpha\bar{\mathcal{B}}_A$ e $t \in]0, \alpha[$ (Aqui, $\bar{\mathcal{B}}_A$ e $\bar{\mathcal{B}}_B$ denotam as bolas unitárias fechadas em torno da origem de A e B respectivamente);

(iii) $R(0) = \{0\}$ e R é semicontínua superior.

Notemos que no caso em que B é finito dimensional então ϕ é fortemente compactamente Lipschitz em $\bar{x}$ se e somente se ϕ é Lipschitz próxima de $\bar{x}$. Se ϕ é fortemente compactamente Lipschitz em $\bar{x}$ então, qualquer que seja $w^* \in H^*$, $w^* \circ \phi(\cdot) = \langle w^*, \phi(\cdot)\rangle$ é uma função Lipschitz próxima de $\bar{x}$. Para maiores detalhes sobre aplicações fortemente compactamente Lipschitz veja [1], [32], [57].

Precisamos de uma noção de invexidade para aplicações entre espaços de Banach. Esta noção de invexidade é feita a partir do conceito usual de invexidade introduzido no capítulo anterior: *Seja $\Omega \subset A$ um aberto em A e $S \subset \Omega$ um subconjunto não vazio. Dizemos que uma aplicação $\phi : \Omega \to B$ é invexa em S se a função real $w^* \circ \phi$ é invexa em S no sentido da definição 1.12, para todo $w^* \in H^*$.* Em todo esse capítulo usaremos essa noção generalizada de invexidade.

3.3 Condições de otimalidade e dualidade

Além dos cones Q e K serem convexos, fechados e com interior não vazio, vamos supor que eles satisfazem à seguinte propriedade: $Q \cap \{-Q\} = \{0\}$, $K \cap \{-K\} = \{0\}$. Sob essas hipóteses Q induz uma ordem partial "$\preceq$" em F. De fato, dados $z, z' \in F$, definimos tal ordem parcial como

$$z' \preceq z \text{ se } z - z' \in Q; \tag{3.1}$$

$$z' \prec z \text{ se } z - z' \in \mathrm{int}Q; \tag{3.2}$$

$z' \not\preceq z$ é a negação de (3.1) e $z' \not\prec z$ é a negação de (3.2). Analogamente K induz uma ordem parcial em G. Também, Q^* e K^* denotarão os cones duais de Q e K respectivamente.

De agora em diante vamos supor que $f : E \to F$ e $g : E \to G$ são fortemente compactamente Lipschitz em $x_0 \in E$. Seja $\mathcal{F} = \{x \in C : g(x) \preceq 0\}$ o conjunto dos pontos factíveis para o problema (PG). Dizemos que $x_0 \in \mathcal{F}$ é uma solução ótima de Pareto fraca de (PG) se não existe $x \in \mathcal{F}$ tal que $f(x) \prec f(x_0)$ ou equivalentemente, se para todo $x \neq x_0$ em C temos que $f(x_0) - f(x) \notin \mathrm{int}Q$.

Em [1], Abdoni e Thibault, obtiveram a seguinte regra de multiplicadores do tipo Fritz-John para o problema (PG):

Teorema 3.1. (Abdoni e Thibault) *Se $x_0 \in \mathcal{F}$ é uma solução ótima de Pareto fraca para (PG) então existem $u^* \in Q^*, v^* \in K^*$ não todos nulos tal que, para algum $k > 0$,*

$$0 \in \partial(u^* \circ f + v^* \circ g + kd_C)(x_0),$$

$$\langle v^*, g(x_0)\rangle = 0.$$

Na verdade, Abdoni e Thibault estudaram um problema mais geral do que o problema (P). O problema original estudado por eles continha restrições de igualdade do tipo $h(x) = 0$ onde $h : E \to \Re^k$ e $x \in C$. Além disso eles obtiveram um resultado mais forte do que o Teorema acima pois sua regra de multiplicadores foi baseada no gradiente de Ioffe [31]. O Teorema acima é um corolário dos resultados de Abdoni e Thibault desde que o gradiente de Ioffe está contido no gradiente de Clarke. Notemos também que no Teorema acima u^* pode ser nulo e neste caso as condições de otimalidade seriam de pouco valia na busca da solução do problema. Devemos então impor alguma condição de regularidade sobre as restrições do problema de modo a garantir que $u^* \neq 0$ e obtermos assim um teorema do tipo Karush-Kuhn-Tucker. Novamente utilizaremos uma condição do tipo Slater:

$$\text{Existe } \widehat{x} \in C \text{ tal que } g(\widehat{x}) < 0. \tag{3.3}$$

Teorema 3.2. *Suponha que as restrições do problema (PG) satisfazem a condição de regularidade (3.3) e que f e g são invexas em $x_0 \in \mathcal{F}$ com relação a C para o mesmo η. Então x_0 é uma solução ótima de Pareto fraca para (PG) se e somente se existem $u^* \in Q^*$, $u^* \neq 0$, $v^* \in K^*$ tal que, para algum $k > 0$,*

$$0 \in \partial(u^* \circ f + v^* \circ g + kd_C)(x_0), \tag{3.4}$$

$$\langle v^*, g(x_0)\rangle = 0. \tag{3.5}$$

Prova. Se x_0 é uma solução ótima de Pareto fraca para (PG) então, pelo Teorema 3.1, existem $u^* \in Q^*$, $v^* \in K^*$ não todos nulos tal que valem (3.4) e (3.5). Basta então provar que $u^* \neq 0$. Suponha que $u^* = 0$. Então $v^* \neq 0$ e segue por (3.4) que, qualquer que seja $\eta \in T_C(x_0)$,

$$0 \leq (v^* \circ g)^o(x_0; \eta). \tag{3.6}$$

No entanto, como as restrições do problema satisfazem a condição de regularidade (3.3), então existe $\widehat{x} \in C$ tal que $g(\widehat{x}) < 0$. Como $v^* \neq 0$, segue que

$$\langle v^*, g(\widehat{x})\rangle < 0. \tag{3.7}$$

Agora, a hipótese de invexidade sobre g em x_0 implica que, existe $\eta(\hat{x}, x_0) \in T_C(x_0)$ tal que

$$\langle v^*, g(\hat{x}) - g(x_0)\rangle \geq (v^* \circ g)^o(x_0; \eta(\hat{x}, x_0)).$$

Segue então de (3.5) e (3.7) que existe $\eta(\hat{x}, x_0) \in T_C(x_0)$ tal que

$$(v^* \circ g)^o(x_0; \eta(\hat{x}, x_0)) < 0.$$

Mas isto contraria (3.6). Portanto só pode ser $u^* \neq 0$. Reciprocamente suponha que x_0 não é uma solução ótima de Pareto fraca de (PG). Então, existe $\hat{x} \in \mathcal{F}$ tal que $f(\hat{x}) - f(x_0) < 0$. Portanto, desde que $u^* \neq 0$,

$$\langle u^*, f(\hat{x}) - f(x_0)\rangle < 0. \tag{3.8}$$

Agora, f invexa em x_0 implica que existe $\eta(\hat{x}, x_0) \in T_C(x_0)$ tal que

$$(u^* \circ f)^o(x_0, \eta(\hat{x}, x_0)) \leq \langle u^*, f(\hat{x}) - f(x_0)\rangle.$$

Logo, (3.8) implica que, existe $\eta(\hat{x}, x_0) \in T_C(x_0)$ tal que

$$(u^* \circ f)^o(x_0, \eta(\hat{x}, x_0)) < 0. \tag{3.9}$$

Também, a hipótese de invexidade sobre g em x_0 implica que,

$$(v^* \circ g)^o(x_0, \eta(\hat{x}, x_0)) \leq \langle v^*, g(\hat{x}) - g(x_0)\rangle.$$

Como $\hat{x} \in \mathcal{F}$ segue que $\langle v^*, g(\hat{x})\rangle \leq 0$. Portanto, (3.5) implica que existe $\eta(\hat{x}, x_0) \in T_C(x_0)$ tal que

$$(v^* \circ g)^o(x_0, \eta(\hat{x}, x_0)) \leq 0. \tag{3.10}$$

Então, de (3.9) e (3.10), concluímos que existe $\eta(\hat{x}, x_0) \in T_C(x_0)$ tal que

$$(u^* \circ f)^o(x_0, \eta(\hat{x}, x_0)) + (v^* \circ g)^o(x_0, \eta(\hat{x}, x_0)) < 0. \tag{3.11}$$

No entanto, (3.4) implica que, qualquer que seja $\eta \in T_C(x_0)$,

$$0 \leq (u^* \circ f)^o(x_0; \eta) + (v^* \circ g)^o(x_0, \eta).$$

Mas isto contradiz (3.11). Portanto x_0 é solução ótima de Pareto fraca de (PG). □

Agora, obteremos resultados de dualidade fraca e forte. Para isto consideremos o seguinte problema dual, do tipo Weir-Mond [58], relacionado ao problema (PG):

$$\left.\begin{array}{ll} \text{Maximizar} & f(w) \\ \\ \text{sujeito a} & w \in C,\ u^* \in Q^*, u^* \neq 0,\ v^* \in K^*, \\ \\ & \langle v^*, g(w)\rangle \geq 0,\ 0 \in \partial(u^* \circ f + v^* \circ g + kd_C)(w). \end{array}\right\} (DV)$$

Teorema 3.3. *(Dualidade fraca) Seja x factível para (PG) e (w, u^*, v^*) factíveis para (DV). Suponhamos f e g invexas em w com respeito a C para o mesmo η. Então,*

$$f(x) \not\prec f(w). \tag{3.12}$$

Prova. Suponha que não vale (3.12). Então existe x factível para (PG) e (w, u^*, v^*) factíveis para (DV) tais que $f(x) - f(w) \prec 0$. Portanto, desde que $u^* \neq 0$,

$$\langle u^*, f(x) - f(w)\rangle < 0. \tag{3.13}$$

Agora, f invexa em w implica que, existe $\eta(x, w) \in T_C(w)$ tal que

$$(u^* \circ f)^o(w, \eta(x, w)) \leq \langle u^*, f(x) - f(w)\rangle.$$

Portanto (3.13) implica que, existe $\eta(x, w) \in T_C(w)$ tal que

$$(u^* \circ f)^o(w, \eta(x, w)) < 0. \tag{3.14}$$

Também, a hipótese de invexidade sobre g em w implica que,

$$(v^* \circ g)^o(w, \eta(x, w)) \leq \langle v^*, g(x) - g(w)\rangle.$$

Segue da factibilidade dos pontos que $\langle v^*, g(w)\rangle \geq 0$ e $\langle v^*, g(x)\rangle \leq 0$. Logo, existe $\eta(x,w) \in T_C(w)$ tal que

$$(v^* \circ g)^o(w, \eta(x,w)) \leq 0. \tag{3.15}$$

Então, existe $\eta(x,w) \in T_C(w)$ tal que

$$(u^* \circ f)^o(w, \eta(x,w)) + (v^* \circ g)^o(w, \eta(x,w)) < 0. \tag{3.16}$$

No entanto, $0 \in \partial(u^* \circ f + v^* \circ g + kd_C)(w)$ implica que, qualquer que seja $\eta \in T_C(w)$,

$$0 \leq (u^* \circ f)^o(w; \eta) + (v^* \circ g)^o(w, \eta)$$

Mas isto contradiz (3.16). Portanto vale (3.12).□

Teorema 3.4. *(Dualidade Forte) Suponha que f e g são invexas em w com respeito a C, para o mesmo η, para todo w factível para (DV). Seja x_0 solução ótima de Pareto fraca de (PG) tal que vale a condição de Slater (3.3). Então existem $\bar{u}^*$, $\bar{v}^*$ tal que $\langle \bar{v}^*, g(x_0)\rangle = 0$, $(x_0, \bar{u}^*, \bar{v}^*)$ é uma solução factível para (DV) e o valor das funções objetivo coincidem.*

Prova. Desde que vale a condição de Slater (3.4), segue pelo Teorema 3.2 que existem $\bar{u}^*$, $\bar{v}^*$ tal que $\langle \bar{v}^*, g(x_0)\rangle = 0$ e $(x_0, \bar{u}^*, \bar{v}^*)$ é factível para (DV). Suponha que $(x_0, \bar{u}^*, \bar{v}^*)$ não é uma solução ótima de Pareto fraca de (DV). Então existem (x, u^*, v^*) factíveis para (DV) tal que

$$f(x) \succ f(x_0). \tag{3.17}$$

Então, (3.17) junto com $u^* \neq 0$ e mais a hipótese de invexidade sobre f implicam que existe $\eta(x_0, x) \in T_C(x)$ tal que

$$(u^* \circ f)^o(x, \eta(\_0, x)) \leq \langle u^*, f(x_0) - f(x)\rangle < 0. \tag{3.18}$$

Também, pela factibilidade dos pontos, temos que

$$\langle v^*, g(x_0)\rangle \leq 0, \text{ e } \langle v^*, g(x)\rangle \geq 0. \tag{3.19}$$

Segue então de (3.19) e da hipótese de invexidade sobre g que existe $\eta(x_0, x) \in T_C(x)$ tal que

$$(v^* \circ g)^o(x, \eta(x_0, x)) \leq \langle v^*, g(x_0) - g(x) \rangle \leq 0. \tag{3.20}$$

Portanto, existe $\eta(x_0, x) \in T_C(x)$ tal que

$$(u^* \circ f)^o(x, \eta(x_0, x)) + (v^* \circ g)^o(x, \eta(x_0, x)) < 0. \tag{3.21}$$

Agora, novamente, a factibilidade dos pontos implica que, qualquer que seja $\eta \in T_C(x)$, temos que

$$0 \leq (u^* \circ f)^o(x, \eta) + (v^* \circ g)^o(x, \eta).$$

Mas isto contradiz (3.21). Portanto $(x_0, \bar{u}^*, \bar{v}^*)$ é uma solução ótima de Pareto fraca de (DV). É óbvio que os valores das funções objetivo coincidem nos respectivos pontos de Pareto otimalidade. □

Obs.: Os principais resultados deste capítulo também podem ser encontrados em [12].

Capítulo 4

Programação matemática com tempo contínuo

4.1 Introdução

Consideremos o seguinte problema de programação não linear com tempo contínuo:

$$\left.\begin{array}{c} \text{Minimizar} \;\; \phi(x) = \int_0^T f(t, x(t))dt, \\ \\ \text{sujeito a} \;\; g_i(t, x(t)) \leq 0 \;\; \text{q.t.p. em } [0, T], \\ i \in I = \{1, \ldots, m\}, \quad x \in X. \end{array}\right\} \;\; (PNC)$$

Aqui X é um subconjunto não vazio, aberto e convexo do espaço de Banach $L_\infty^n[0, T]$ de todas as funções vetoriais n-dimensionais Lebesgue mensuráveis que são essencialmente limitadas, definidas no intervalo compacto $[0, T] \subset \Re$, com a norma $\|\cdot\|_\infty$ definida por

$$\|x\|_\infty = \max_{1 \leq j \leq n} \operatorname{ess\,sup}\{|x_j(t)|, \; 0 \leq t \leq T\},$$

onde, para cada $t \in [0, T]$, $x_j(t)$ é a j-ésima componente de $x(t) \in \Re^n$, ϕ é uma função real definida em X, $g(t, x(t)) = \gamma(x)(t)$ e $f(t, x(t)) = \Gamma(x)(t)$, onde γ é uma aplicação de X no espaço normado $\Lambda_1^m[0, T]$, de todas as funções vetoriais m-dimensionais Lebesgue mensuráveis que são essencialmente limitadas, definidas no intervalo compacto $[0, T] \subset \Re$, com a norma $\|\cdot\|_1$ definida por

$$\|y\|_1 = \max_{1 \le j \le m} \int_0^T |y_j(t)| dt,$$

e Γ é uma aplicação de X no espaço normado $\Lambda_1^1[0, T]$.

Nosso objetivo neste capítulo é obter condições suficientes de otimalidade global para (PNC), sem nenhuma hipótese de convexidade sobre as funções envolvidas. Primeiro provaremos a suficiência das condições de Fritz-John e Karush-Kuhn-Tucker, no caso Lipschitz, usando a noção de invexidade. Depois, fazendo uso do conceito de Hessiano generalizado, introduzido por Cominetti e Correa [18], nós provaremos condições suficientes de otimalidade de segunda ordem para (PNC). Resultados relacionados podem ser encontrados em [37], [43], [42], [67]. Entre esses, talvez os melhores resultados sobre condições suficientes de otimalidade para (PNC) foram dados por Zalmai em [66], para funções diferenciáveis. Os outros autores também tratam de problemas diferenciáveis. Em nosso caso, permitimos que as funções sejam somente Lipschitz na segunda variável. Portanto, nossos resultados estendem os resultados anteriores para (PNC).

O Capítulo está organizado da seguinte maneira: Na seção 2, apresentamos condições suficientes de otimalidade para o caso Lipschitz. Na seção 3, discutimos condições suficientes de otimalidade sob a hipótese de Clarke regularidade e sob várias hipóteses de convexidade generalizada. Na seção 4, nós damos condições suficientes de otimalidade de segunda ordem para o caso em que as funções envolvidas são de classe $C^{1,1}$. Finalmente na última seção, formulamos um problema dual relacionado a (PNC) e provamos teoremas de dualidade fraca e forte.

4.2 Condições suficientes de otimalidade: o caso Lipschitz

Nesta seção obtemos condições suficientes de otimalidade global para (PNC) no caso em que as funções envolvidas são localmente Lipschitz sem qualquer hipótese de convexidade. Antes de apresentar os resultados propriamente ditos vamos detalhar algumas questões sobre o problema (PNC). Seja $\mathbb{F}$ o conjunto de todas as soluções factíveis para (PNC) (o qual nós supomos não vazio), isto é,

$$\mathbb{F} = \{x \in X : g_i(t, x(t)) \leq 0 \text{ q.t.p. em } [0,T],\ i \in I\}.$$

Seja V um subconjunto aberto e convexo de $\Re^n$ contendo o conjunto

$$\{x(t) \in \Re^n : x \in X,\ t \in [0,T]\}.$$

Vamos supor que f e g_i, $i \in I$, são funções reais definidas em $[0,T] \times V$. A função $t \to f(t, x(t))$ é assumida ser Lebesgue mensurável e integrável para todo $x \in X$.
Nós assumimos que, dado $a \in V$, existe um $\epsilon > 0$ e um número positivo k tal que, $\forall t \in [0,T]$, e $\forall x_1,\ x_2 \in a + \epsilon B$ (B denota a bola unitária do $\Re^n$) temos que

$$|f(t,x_1) - f(t,x_2)| \leq k\|x_1 - x_2\|.$$

Hipóteses análogas valem para g_i, $i \in I$. Logo, $f(t,\cdot)$ e $g_i(t,\cdot)$, $i \in I$, são localmente Lipschitz em V para cada $t \in [0,T]$.

Sejam $\overline{x} \in X$ e $h \in L_\infty^n[0,T]$ dados. *As derivadas direcionais generalizadas de Clarke com tempo contínuo* de f e g_i's são dadas por

$$f^0(t, \overline{x}(t); h(t)) := \Gamma^0(\overline{x}; h)(t) := \limsup_{\substack{y \to \overline{x} \\ s \to 0^+}} \frac{\Gamma(y + sh)(t) - \Gamma(y)(t)}{s}$$

e

$$g_i^0(t, \overline{x}(t); h(t)) := \gamma_i^0(\overline{x}; h)(t) := \limsup_{\substack{y \to \overline{x} \\ s \to 0^+}} \frac{\gamma_i(y + sh)(t) - \gamma_i(y)(t)}{s}$$

q.t.p. em $[0,T]$.
Segue-se facilmente das hipóteses que

$$t \to f^0(t, \overline{x}(t)); h(t)),$$
$$t \to g_i^0(t, \overline{x}(t)); h(t)),\ i \in I,$$

são Lebesgue mensuráveis e integráveis para todo $\overline{x} \in X$, e $h \in L_\infty^n[0,T]$.

Nós também necessitamos de uma noção de invexidade com tempo contínuo que é uma adaptação natural da noção clássica de invexidade. Seja $U \subset \Re^n$

um subconjunto não vazio de $\Re^n$ e $\overline{x} \in X$. Suponha que uma função ψ : $[0,T] \times U \longrightarrow \Re$ é localmente Lipschitz para cada $t \in [0,T]$. A função $\psi(t,.)$ é dita invexa em $\bar{x}(t)$ (com respeito a U) se existe $\eta : U \times U \longrightarrow \Re^n$ tal que a função $t \longrightarrow \eta(x(t),\overline{x}(t))$ está em $L^n_\infty[0,T]$ e

$$\psi(t,x(t)) - \psi(t,\overline{x}(t)) \geq \psi^0(t,\overline{x}(t);\eta(x(t),\overline{x}(t))) \text{ q.t.p. em } [0,T]$$

para todo $x \in X$. Dizemos que ψ é *estritamente invexa* se a desigualdade acima é estrita para $x(t) \neq \overline{x}(t)$ q.t.p. em $[0,T]$.

Teorema 4.1. *Seja $\bar{x} \in \mathbb{F}$. Suponha que $f(t,\cdot)$ é invexa em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0,T]$, e que, para cada $i \in I$, $g_i(t,\cdot)$ é estritamente invexa em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0,T]$, com o mesmo $\eta(x(t),\bar{x}(t))$ para todas as funções. Suponha ainda que existem $\bar{\lambda}_0 \in \mathbb{R}$, $\bar{\lambda} \in L^m_\infty[0,T]$ tal que*

$$0 \leq \int_0^T [\bar{\lambda}_0 f^0(t,\bar{x}(t);h(t)) + \sum_{i=1}^m \bar{\lambda}_i(t) g_i^0(t,\bar{x}(t);h(t))]dt \quad \forall h \in L^n_\infty[0,T], \tag{4.1}$$

$$\bar{\lambda}_0 \geq 0, \bar{\lambda}(t) \geq 0 \quad q.t.p. \ em \ [0,T], \tag{4.2}$$

$$(\bar{\lambda}_0, \bar{\lambda}(t)) \neq 0 \quad q.t.p. \ em \ [0,T], \tag{4.3}$$

$$\bar{\lambda}_i(t) g_i(t,\bar{x}(t)) = 0 \quad q.t.p. \ em \ [0,T], i \in I. \tag{4.4}$$

Então $\bar{x}$ é uma solução ótima global de (PNC).

Prova. Suponha que $\bar{x}$ não é ótima para (PNC). Então, existe $\tilde{x} \in \mathbb{F}$, $\tilde{x} \neq \bar{x}$, tal que

$$\int_0^T f(t,\tilde{x}(t))dt < \int_0^T f(t,\bar{x}(t))dt. \tag{4.5}$$

Desde que $f(t,\cdot)$ é invexa em $\bar{x}(t)$ e, para cada $i \in I$, $g_i(t,\cdot)$ é estritamente invexa em $\bar{x}(t)$ então temos as desigualdades

$$f(t,\tilde{x}(t)) - f(t,\bar{x}(t)) \geq f^o(t,\bar{x}(t);\eta(\tilde{x}(t),\bar{x}(t))) \text{ q.t.p. em } [0,T], \tag{4.6}$$

$$g_i(t,\tilde{x}(t)) - g_i(t,\bar{x}(t)) > g_i^o(t,\bar{x}(t);\eta(\tilde{x}(t),\bar{x}(t))) \text{ q.t.p. em } [0,T], \tag{4.7}$$

$i \in I$, para algum $\eta(\tilde{x}(t),\bar{x}(t))$. Como $\tilde{x} \in \mathbb{F}$ e $\bar{\lambda}_i(t) \geq 0$ q.t.p. em $[0,T]$ para cada $i \in I$, é claro que

$$\bar{\lambda}_i(t) g_i(t,\tilde{x}(t)) \leq 0 \text{ q.t.p. em } [0,T], \ i \in I. \tag{4.8}$$

Agora, de (4.2)-(4.8) segue-se que

$$0 > \int_0^T [\bar{\lambda}_0 f^o(t, \bar{x}(t); \eta(\tilde{x}(t), \bar{x}(t))) + \sum_{i=1}^{m} \bar{\lambda}_i(t) g_i^o(t, \bar{x}(t); \eta(\tilde{x}(t), \bar{x}(t)))]dt,$$

o qual, com $h(t) = \eta(\tilde{x}(t), \bar{x}(t))$, contradiz (1). Portanto, concluímos que $\bar{x}$ é uma solução ótima global de (PNC). □

Obs. Segue-se da prova acima que se, para cada $i \in I$, $g_i(t, \cdot)$ é invexa, e se pelo menos uma destas funções, digamos $g_k(t, \cdot)$, é estritamente invexa em $\bar{x}(t)$ para cada $t \in [0, T]$ tal que o correspondente multiplicador $\bar{\lambda}_k$ é não nulo num subconjunto de $[0, T]$ com medida de Lebesgue positiva, então a afirmação do Teorema permanece válida. Notemos que no teorema anterior, o multiplicador associado à função objetivo pode eventualmente ser nulo. Neste caso as condições (4.1) – (4.4) são chamadas condições do tipo Fritz-John. No próximo teorema o multiplicador associado à função objetivo é tomado igual a um e obtemos um teorema do tipo Karush-Kuhn-Tucker.

Teorema 4.2. *Seja $\bar{x} \in \mathbb{F}$. Suponha que $f(t, \cdot)$, $g_i(t, \cdot)$, $i \in I$, são invexas em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0, T]$, para a mesma função $\eta(x(t), \bar{x}(t))$. Suponha ainda que existe $\bar{\lambda} \in L_\infty^m[0, T]$ tal que*

$$0 \leq \int_0^T [f^0(t, \bar{x}(t); h(t)) + \sum_{i=1}^{m} \bar{\lambda}_i(t) g_i^0(t, \bar{x}(t); h(t))]dt \quad \forall h \in L_\infty^n[0, T], \tag{4.9}$$

$$\bar{\lambda}_i(t) \geq 0 \quad q.t.p \text{ em } [0, T], \ i \in I, \tag{4.10}$$

$$\bar{\lambda}_i(t) g_i(t, \bar{x}(t)) = 0 \quad q.t.p. \text{ em } [0, T], \ i \in I. \tag{4.11}$$

Então $\bar{x}$ é uma solução ótima global de (PNC).

Prova. Seja $x \in \mathbb{F}$. Segue-se de (4.10) e (4.11) que

$$\bar{\lambda}_i(t) g_i(t, x(t)) \leq 0 = \bar{\lambda}_i(t) g_i(t, \bar{x}(t)) \text{ q.t.p. em } [0, T], \ i \in I.$$

Desde que, para cada $i \in I$, $g_i(t, \cdot)$ é invexa em $\bar{x}(t)$ para cada $t \in [0, T]$ e $\bar{\lambda}_i(t) \geq 0$ a.e in $[0, T]$ então nós temos que $\bar{\lambda}_i(t) g_i(t, \cdot)$ também é invexa em $\bar{x}(t)$ para cada $t \in [0, T]$ para a mesma função $\eta(x(t), \bar{x}(t))$. Da invexidade de $\bar{\lambda}_i(t) g_i(t, \cdot)$ nós obtemos que

$$\bar{\lambda}_i(t) g_i^o(t, \bar{x}(t); \eta(x(t), \bar{x}(t))) \leq 0 \text{ q.t.p. em } [0, T], \ i \in I. \tag{4.12}$$

Agora, pondo $h(t) = \eta(x(t), \bar{x}(t))$ em (4.9) nós concluímos que

$$0 \leq \int_0^T [f^o(t, \bar{x}(t); \eta(x(t), \bar{x}(t))) + \sum_{i=1}^m \bar{\lambda}_i(t) g_i^o(t, \bar{x}(t); \eta(x(t). \bar{x}(t)))] dt. \quad (4.13)$$

Combinando (4.12) e (4.13) nós obtemos

$$\int_0^T [f^o(t, \bar{x}(t); \eta(x(t), \bar{x}(t)))] dt \geq 0.$$

A hipótese de invexidade sobre f junto com a última desigualdade implicam que

$$\phi(\bar{x}) \leq \phi(x).$$

Logo, como $x \in \mathbb{F}$ é arbitrário, concluímos que $\bar{x}$ é uma solução ótima global de (PNC). □

4.3 Condições suficientes de otimalidade: o caso Clarke regular

Nesta seção, nós obtemos condições suficientes de otimalidade global para (PNC) sob hipóteses de convexidade generalizada e Clarke regularidade. Os teoremas estabelecidos abaixo generalizam o caso tratado por Zalmai [66], onde os dados de (PNC) são suaves.

Em toda esta seção vamos supor que as funções são Clarke regulares, isto é, as funções possuem derivadas direcionais usuais e as mesmas coincidem com a derivada de Clarke. Vamos introduzir agora alguns conceitos clássicos de convexidade generalizada.

Uma função ψ é dita *pseudoconvexa* em $x_1 \in U$ (com respeito a U) se, para todo $x_2 \in U$,

$$\psi'(x_1; x_2 - x_1) \geq 0 \Longrightarrow \psi(x_2) \geq \psi(x_1).$$

Uma função ψ é dita *quasiconvexa* em $x_1 \in U$ (com respeito a U) se, para todo $x_2 \in U$,

$$\psi(x_2) \leq \psi(x_1) \Longrightarrow \psi'(x_1; x_2 - x_1) \leq 0.$$

Nós definimos a *função Lagrangeano* $L : X \times \Re \times L_\infty^m[0,T] \longrightarrow \Re$ por

$$L(x, \lambda_0, \lambda) := \int_0^T [\lambda_0 f(t, x(t)) + \sum_{i=1}^{m} \lambda_i(t) g_i(t, x(t))] dt.$$

Quando $\lambda_0 \neq 0$, podemos supor que $\lambda_0 = 1$ normalizando os multiplicadores de Lagrange. Neste caso denotamos $L(x, 1, \lambda)$ por $L(x, \lambda)$.

Na sequência $L'_x(\bar{x}, \lambda_0, \lambda; h)$ denota a derivada direcional usual de $L(\cdot, \lambda_0, \lambda)$ em $\bar{x}$ na direção $h \in L_\infty^n[0,T]$ e $\partial_x L(\bar{x}, \lambda_0, \lambda)$ denota o gradiente generalizado de Clarke de $L(\cdot, \lambda_0, \lambda)$. Se f e g_i's são Clarke regulares, então a condição (4.1)é equivalente a $L'_x(\bar{x}, \bar{\lambda}_0, \bar{\lambda}; h) \geq 0$ for all $h \in L_\infty^n[0,T]$ e, portanto, $0 \in \partial_x L(\bar{x}, \bar{\lambda}_0, \bar{\lambda})$. Formalmente, temos os seguintes corolários:

Corolário 4.3. *Seja $\bar{x} \in \mathbb{F}$. Suponha que $f(t, \cdot)$ é invexa em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0,T]$, e que, para cada $i \in I$, $g_i(t, \cdot)$ é estritamente invexa em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0,T]$, com respeito ao mesmo $\eta(x(t), \bar{x}(t))$ para todas as funções. Suponha ainda que existem $\bar{\lambda}_0 \in \mathbb{R}$, $\bar{\lambda} \in L_\infty^m[0,T]$ such that*

$$0 \in \partial_x L(\bar{x}, \bar{\lambda}_0, \bar{\lambda}), \tag{4.14}$$

$$\bar{\lambda}_0 \geq 0, \bar{\lambda}(t) \geq 0 \quad q.t.p. \ em \ [0,T], \tag{4.15}$$

$$(\bar{\lambda}_0, \bar{\lambda}(t)) = (\bar{\lambda}_0, \bar{\lambda}_1(t), \cdots, \bar{\lambda}_m(t)) \neq 0, \quad q.t.p. \ em \ [0,T], \tag{4.16}$$

$$\bar{\lambda}_i(t) g_i(t, \bar{x}(t)) = 0 \quad q.t.p. \ em \ [0,T], \ i \in I. \tag{4.17}$$

Então $\bar{x}$ é uma solução ótima global de (PNC).

Corolário 4.4. *Seja $\bar{x} \in \mathbb{F}$. Suponha que $f(t, \cdot)$, $g_i(t, \cdot), i \in I$, são invexas em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0,T]$, para a mesma função $\eta(x(t), \bar{x}(t))$. Suponha ainda que existe $\bar{\lambda} \in L_\infty^m[0,T]$ tal que*

$$0 \in \partial_x L(\bar{x}, \bar{\lambda}), \tag{4.18}$$

$$\bar{\lambda}_i(t) \geq 0 \quad q.t.p. \ em \ [0,T], \ i \in I, \tag{4.19}$$

$$\bar{\lambda}_i(t) g_i(t, \bar{x}(t)) = 0 \quad q.t.p. \ em \ [0,T], \ i \in I. \tag{4.20}$$

Então $\bar{x}$ é uma solução ótima global para (PNC).

O próxima resultado estabelece um critério de otimalidade global no qual nós supomos apenas que a função Lagrangeana é invexa na primeira variável.

Proposição 4.5. *Seja $\bar{x} \in \mathbb{F}$. Se existe $\bar{\lambda} \in L_\infty^m[0,T]$ tal que $(\bar{x},\bar{\lambda})$ satisfaz (4.18)-(4.20), e se a função Lagrangeana $L(x;\bar{\lambda})$ é invexa em $\bar{x}$ (com respeito a $\mathbb{F}$), então $\bar{x}$ é uma solução ótima global de (PNC).*

Prova. A condição (4.18) implica que $0 \leq L'_x(\bar{x},\bar{\lambda};\eta(x,\bar{x}))\ \ \forall x \in \mathbb{F}$. Da hipótese de invexidade sobre a função Lagrangeana, nós obtemos

$$L(\bar{x},\bar{\lambda}) \leq L(x,\bar{\lambda}) \ \ \forall x \in \mathbb{F}.$$

Esta desigualdade mais (4.19)-(4.20) implicam que $\phi(\bar{x}) \leq \phi(x)\ \ \forall x \in \mathbb{F}$, o qual finaliza a prova. □

A seguir vamos estabelecer dois resultados sobre condições suficientes, do tipo Karush-Kuhn-Tucker, O primeiro é obtido sob as hipóteses de pseudoconvexidade da função objetivo e quasiconvexidade das funções associadas às restrições do problema. O segundo resultado mostra que a otimalidade global é mantida se impusermos a hipótese de quasiconvexidade sobre uma certa função definida em termos das g_i's, ao invés de assumirmos quasiconvexidade diretamente sobre as g_i's.

Proposição 4.6. *Seja $\bar{x} \in \mathbb{F}$. Suponha $\phi(\cdot)$ é pseudoconvexa em $\bar{x}$ (com respeito a $\mathbb{F}$) e que, para cada $i \in I$, $\bar{\lambda}_i(t)g_i(t,\cdot)$ é quasiconvexa em $\bar{x}(t)$ (com respeito a V) para cada $t \in [0,T]$. Se existe $\bar{\lambda} \in L_\infty^m[0,T]$ tal que*

$$0 \in \partial_x L(\bar{x},\bar{\lambda}), \tag{4.21}$$

$$\bar{\lambda}_i(t) \geq 0 \ \ q.t.p. \ em\ [0,T], \ \ i \in I, \tag{4.22}$$

$$\bar{\lambda}_i(t)g_i(t,\bar{x}(t)) = 0 \ \ q.t.p. \ em\ [0,T], i \in I, \tag{4.23}$$

então $\bar{x}$ é uma solução ótima global de (PNC).

Prova. Desde que, para cada $x \in \mathbb{F}$,

$$\bar{\lambda}_i(t)g_i(t,x(t)) \leq 0 = \bar{\lambda}_i(t)g_i(t,\bar{x}(t)) \text{ q.t.p. em } [0,T],\ i \in I,$$

a hipótese de quasiconvexidade implica que

$$\bar{\lambda}_i(t)g_i'(t,\bar{x}(t);x(t)-\bar{x}(t)) \leq 0 \text{ q.t.p. em } [0,T],\ i \in I,$$

e, então,

$$\int_0^T \sum_{i\in I} \bar{\lambda}_i(t)g_i'(t,\bar{x}(t);x(t)-\bar{x}(t))dt \leq 0 \ \ \forall x \in \mathbb{F}. \tag{4.24}$$

Agora, $0 \in \partial_x L(\bar{x},\bar{\lambda})$ implica que, $\forall x \in \mathbb{F}$,

$$0 \leq \int_0^T [f'(t,\bar{x}(t);x(t)-\bar{x}(t)) + \sum_{i\in I} \bar{\lambda}_i(t)g_i'(t,\bar{x}(t);x(t)-\bar{x}(t))]dt. \tag{4.25}$$

De (4.24) e (4.25) concluímos que

$$\int_0^T f'(t,\bar{x}(t);x(t)-\bar{x}(t))dt \geq 0 \ \ \forall x \in \mathbb{F}.$$

Desde que ϕ é pseudoconvexa em $\bar{x}$ a última desiguladade implica

$$\phi(\bar{x}) \leq \phi(x) \ \ \forall x \in \mathbb{F}.$$

Portanto, $\bar{x}$ é uma solução ótima global de (PNC). □

Proposição 4.7. *Seja $\bar{x} \in \mathbb{F}$. Suponha que $\phi(\cdot)$ é pseudoconvexa em $\bar{x}$ (com respeito a $\mathbb{F}$). Se existe $\bar{\lambda} \in L_\infty^m[0,T]$ tal que $(\bar{x},\bar{\lambda})$ satisfaz (4.21)-(4.23), e se a função $G(\cdot;\bar{\lambda}) : L_\infty^n[0,T] \longrightarrow \mathbb{R}$ dada por*

$$G(x,\bar{\lambda}) = \int_0^T \sum_{i=1}^m \bar{\lambda}_i(t)g_i(t,x(t))dt,$$

é quasiconvexa em $\bar{x}$ (com respeito a $\mathbb{F}$), então $\bar{x}$ é uma solução ótima global de (PNC).

Prova. Para cada $x \in \mathbb{F}$, as condições (4.22) e (4.23) implicam que $G(x,\bar{\lambda}) \leq 0 = G(\bar{x},\bar{\lambda})$. Em vista da quasiconvexidade de G em $\bar{x}$, deduzimos que

$$G_x'(\bar{x},\bar{\lambda};x-\bar{x}) \leq 0 \ \ \forall x \in \mathbb{F},$$

isto é,

$$\int_0^T \sum_{i=1}^m \bar{\lambda}_i(t) g_i'(t, \bar{x}(t); x(t) - \bar{x}(t)) dt \leq 0 \ \ \forall x \in \mathbb{F}.$$

O resto da prova segue-se usando a mesma argumentação usada na prova da proposição anterior. □

4.4 Condições suficientes de otimalidade de segunda ordem

Considerando um problema similar ao problema (PNC) mas com dados duas vezes continuamente diferenciáveis, Zalmai [66] obteve condições suficientes de otimalidade em termos do Hessiano da função Lagrangeana associada a (PNC). Nosso objetivo nesta seção é estender os resultados obtidos por Zalmai para o caso das funções $C^{1,1}$, isto é funções diferenciáveis com derivada localmente Lipschitz. Para isto faremos uso da noção de derivada generalizada de segunda ordem devida a Cominetti e Correa [18]. Entretanto, os resultados apresentados aqui permanecem válidos se, ao invés de usarmos a noção de Hessiano generalizado acima, usarmos qualquer outra noção disponível na literatura, mas com os mesmos aspectos. Para diferentes noções de Hessianos generalizados veja, por exemplo,[14], [45], [61], [62]. As conecções entre as várias derivadas de segunda ordem são discutidas em [45].

Antes de estabelecermos nosso resultado principal vamos lembrar alguns resultados auxiliares do trabalho de Cominetti e Correa [18]. Sejam Z um espaço de Banach, $\psi : Z \longrightarrow \mathbb{R}$ uma função e $x \in Z$ um ponto dado. A *derivada direcional de segunda ordem generalizada* de ψ em x na direção $(u, v) \in Z \times Z$ é dada por

$$\psi^{oo}(x; u, v) = \limsup_{\substack{y \to x \\ s,t \to 0}} \frac{\psi(y + su + tv) - \psi(y + su) - \psi(y + tv) + \psi(y)}{st}.$$

ψ^{oo} pode, eventualmente, assumir os valores $+\infty$ e $-\infty$. O *Hessiano generalizado* de ψ em x é definido como sendo a multifunção $\partial^2\psi(x) : Z \longrightarrow Z^*$ dada por

$$\partial^2\psi(x)(u) = \{x^* \in Z^* : \langle x^*, v\rangle \leq \psi^{oo}(x; u, v) \ \forall v \in Z\}.$$

Teorema 4.8. ([18])

(a) *A função* $(u,v) \to \psi^{oo}(x;u,v)$ *é simétrica* ($\psi^{oo}(x,u,v) = \psi^{oo}(x,v,u)$), *e bisublinear (sublinear em cada variável separadamente);*

(b) *a aplicação* $x \to \psi^{oo}(x;u,v)$ *é semicontínua superior (s.c.s.) em* x *para todo* $(u,v) \in Z \times Z$ *e a multifunção* $x \to \partial^2\psi(x)(u)$ *é fechada em* x *para cada* $u \in Z$ *fixado;*

(c) $\partial^2\psi(x)(u)$ *é convexo e* w^**-fechado;*

(d) $\psi^{oo}(x;-u,v) = \psi^{oo}(x;u,-v) = (-\psi)^{oo}(x;u,v)$.

Uma função ψ é dita de classe $C^{1,1}$ em Z se, para cada $x \in Z, \psi$ é Gâteaux diferenciável em x com a derivada localmente Lipschitz.

Teorema 4.9. ([18]) *Se* ψ *é de classe* $C^{1,1}$ *em* $x \in Z$, *então*

(a) $\partial^2\psi(x)(u)$ *é não vazio;*

(b) $\partial^2\psi(x)(u)$ *é* w^**-compacto e* $\psi^{oo}(x;u,v) = \max_{x^* \in \partial^2\psi(x)(u)}\langle x^*, v\rangle$*;*

(c) $(x,u) \to \partial^2\psi(x)(u)$ *é semicontínua superior.*

Lema 4.10. ([18]) *Suponha que* ψ *é de classe* $C^{1,1}$ *no segmento fechado* $[x,y] \subset Z$. *Então existe* ξ *no segmento aberto* $]x,y[$ *tal que*

$$\psi(y) \in \psi(x) + \langle \nabla\psi(x), y-x\rangle + \frac{1}{2}\langle \partial^2\psi(\xi)(y-x), y-x\rangle.$$

Dizemos que $\partial^2\psi(x)$ é *definido positivo* (d.p.) se

$$-\psi^{oo}(x,u,-u) \geq 0 \;\; \forall u \in Z.$$

Se a desigualdade acima é estrita para $u \neq 0$, então dizemos que $\partial^2\psi(x)$ é *estritamente definido positivo* (e.d.p.).

No restante da seção vamos supor que as funções f e g_i, $i \in I$, em (PNC) são Gâteaux diferenciáveis com respeito à segunda variável e suas derivadas parciais, denotadas por $\nabla f(t,x)$ e $\nabla g_i(t,x)$, $i \in I$, respectivamente, são localmente Lipschitz. Na verdade, vamos supor que existe uma função positiva $k \in L_1[0,T]$, tal que

$$\|\nabla f(t,y) - \nabla f(t,z)\| \leq k(t)\|y - z\| \quad \text{q.t.p. em } [0,T],$$

$$\|\nabla g_i(t,y) - \nabla g_i(t,z)\| \leq k(t)\|y - z\| \quad \text{q.t.p. em } [0,T], \ i \in I,$$

para todo y, z numa vizinhança x. Sob estas hipóteses podemos facilmente provar o seguinte lema:

Lema 4.11. *Seja $\lambda \in L^m_\infty[0,T]$ dado. A função Lagrangeana $L(\cdot,\lambda)$: $X \longrightarrow \Re$ é de classe $C^{1,1}$.*

Prova. Vamos denotar por $\nabla_x L(\bar{x},\lambda)$ a derivada Gâteaux de $L(\cdot,\lambda)$ em $\bar{x} \in X$. Seja $h \in L^n_\infty[0,T]$ não nulo. É fácil verificar que

$$\langle \nabla_x L(\bar{x},\lambda), h\rangle = \int_0^T [\nabla f(t,\bar{x}(t)) + \sum_{i\in I} \lambda_i(t)\nabla g_i(t,\bar{x}(t))]h(t)dt.$$

Então, dados y, z numa vizinhança de $\bar{x}$, segue-se das hipóteses Lipschitz sobre $f, g_i, i \in I$ que

$$|\langle \nabla_x L(y,\lambda) - \nabla_x L(z,\lambda), h\rangle| \leq M\|y - z\|_\infty \|h\|_\infty,$$

onde $M = [1 + \sum_{i\in I} \|\lambda_i\|_\infty] \int_0^T k(t)dt$. Portanto,

$$\|\nabla_x L(y,\lambda) - \nabla_x L(z,\lambda)\|_* \leq M\|y - z\|_\infty,$$

como queríamos demonstrar. □

Vamos denotar por $\partial^2_x L(\bar{x},\lambda)$ o Hessiano generalizado de $L(\cdot,\lambda)$ em $\bar{x}$. Agora estamos em posição de estabelecer e provar o principal resultado desta seção.

Teorema 4.12. *Seja $(\bar{x},\lambda) \in \mathbb{F} \times L^m_\infty[0,T]$. Suponha que $(\bar{x},\lambda)$ satisfaz as condições de otimalidade de primeira ordem*

(i) $\nabla_x L(\bar{x},\lambda) = 0$;

(ii) $\lambda_i(t) \geq 0$ *q.t.p. em* $[0, T]$, $i \in I$;

(iii) $\lambda_i(t) g_i(t, \bar{x}(t)) = 0$ *q.t.p. em* $[0, T]$, $i \in I$.

Suponha também que o Hessiano generalizado $\partial_x^2 L(\bar{x}, \lambda)$ *é e.d.p. Então* $\bar{x}$ *é um mínimo local estrito de (PNC). Se, além disso,* $L(\cdot, \lambda)$ *é quasiconvexa em* $\bar{x}$ *(com respeito a* F*), então* $\bar{x}$ *é um mínimo global estrito de (PNC).*

Prova. Suponha que $\bar{x}$ não é um mínimo local estrito de (PNC). Então $\bar{x}$ não é um mínimo local estrito da função Lagrangeana $L(\cdot, \lambda)$ em $\mathbb{F}$. Logo, existe uma sequência $(x_n) \subseteq \mathbb{F}, x_n \to \bar{x}, x_n \neq \bar{x}$, tal que $L(x_n, \lambda) \leq L(\bar{x}, \lambda)$. Seja

$$u_n = \frac{x_n - \bar{x}}{\|x_n - \bar{x}\|} \to u.$$

Segue-se do Lema 4.11 que $L(\cdot, \lambda)$ é de classe $C^{1,1}$. Logo, pelo Lema 4.10, para cada n, existe $\xi_n \in]\bar{x}, x_n[$ tal que

$$L(x_n, \lambda) - L(\bar{x}, \lambda) \in \langle \nabla_x L(\bar{x}, \lambda), x_n - \bar{x} \rangle + \frac{1}{2} \langle \partial_x^2 L(\xi_n, \lambda)(x_n - \bar{x}), x_n - \bar{x} \rangle.$$

Desde que a hipótese (i) é válida, a inclusão acima é equivalente a

$$a_n = \frac{2(L(x_n, \lambda) - L(\bar{x}, \lambda))}{\|x_n - \bar{x}\|^2} \in \langle \partial_x^2 L(\xi_n, \lambda) u_n, u_n \rangle.$$

Nós construímos sequências $\xi_n \to \bar{x}$, $u_n \to u$ e $x_n^* \in \partial_x^2 L(\xi_n, \lambda)(u_n)$, tal que $a_n = \langle x_n^*, u_n \rangle \leq 0$. Pelo Teorema 4.9, nós podemos supor (tomando subsequências se necessário) que (x_n^*) converge a algum $x^* \in \partial_x^2 L(\bar{x}, \lambda)(u)$. Segue-se que $\langle x^*, u \rangle = \lim \langle x_n^*, u_n \rangle \leq 0$, o que implica que $-L^{oo}(\bar{x}, \lambda; u, -u) \leq 0$. Isto contradiz o fato de que $\partial_x^2 L(\bar{x}, \lambda)$ é e.d.p. . Portanto, $\bar{x}$ é um mínimo local estrito de $L(x, \lambda)$ em $\mathbb{F}$, o que, sob as hipóteses (ii) e (iii), realmente implicam que $\bar{x}$ é um mínimo local estrito de (PNC). A hipótese de quasiconvexidade sobre a função Lagrangeana implica a otimalidade global estrita para (PNC). □

4.5 Dualidade com tempo contínuo

Nesta seção vamos formular e provar alguns teoremas que estabelecem uma relação de dualidade entre (PNC) e um outro problema de otimização, o chamado problema dual de (PNC). Tal problema dual pode ser estabelecido da seguinte maneira:

$$\left.\begin{array}{ll} \text{Maximizar} & L(y,\lambda) := \int_0^T [f(t,y(t)) + \sum_{i=1}^{m} \lambda_i(t) g_i(t,y(t))] dt. \\ \text{sujeito a} & (y,\lambda) \in X \times L_\infty^m[0,T], \\ & \lambda_i(t) \geq 0 \text{ q.t.p. em } [0,T], \quad 0 \in \partial_y L(y,\lambda). \end{array}\right\} (DNC)$$

Seja $I\!D$ o conjunto dos pontos factíveis para (DNC). O próximo resultado estabelece uma relação de dualidade fraca entre (PNC) e (DNC):

Teorema 4.13. *(Dualidade fraca). Suponha que $f, g_i, i \in I$, são invexas em V, com uma função η comum a todas. Então, quaisquer que sejam $x \in I\!F$ e $(y,\lambda) \in I\!D$,*

$$\phi(x) \geq L(y,\lambda).$$

Prova. Sejam $x \in I\!F$ e $(y,\lambda) \in I\!D$. É fácil verificar pela factibilidade dos pontos que

$$\phi(x) \geq \phi(x) + \int_0^T [\sum_{i=1}^{m} \lambda_i(t) g_i(t,x(t))] d = L(x,\lambda).$$

Isto implica que

$$\phi(x) - L(y,\lambda) \geq L(x,\lambda) - L(y,\lambda). \tag{4.26}$$

A hipótese de invexidade sobre $f, g_i, i \in I$, implica que $L(\cdot,\lambda)$ é invexa em X. Logo, existe $\eta(x,y) \in L_\infty^n[0,T]$ tal que

$$L(x,\lambda) - L(y,\lambda) \geq L^o(y;\eta(x,y)). \tag{4.27}$$

A condição $0 \in \partial_y L(y, \lambda)$ implica, em particular, que

$$L^o(y; \eta(x, y)) \geq 0. \tag{4.28}$$

É fácil verificar que (4.26)-(4.28) implicam o resultado do teorema. □

Com algumas hipóteses a mais sobre o teorema anterior conseguimos obter o seguinte teorema que estabelece uma relação de dualidade forte entre (PNC) e (DNC):

Teorema 4.14. *(Dualidade forte). Suponha que $f, g_i, i \in I$ são invexas em V, para a mesma função η. Se $\overline{x}$ é uma solução ótima de (PNC) e se existe $\overline{\lambda} \in L_\infty^m[0, T]$ satisfazendo as condições de K-K-T (4.18)-(4.20), então, (DNC) é maximizado em $(\overline{x}, \bar{\lambda})$ e os valores otimos de (PNC) e (DNC) coincidem.*

Prova. Se $\overline{x}$ é uma solução ótima para (PNC) e se existe $\bar{\lambda} \in L_\infty^m[0, T]$ satisfazendo (4.18)-(4.19) então $(\overline{x}, \bar{\lambda})$ é factível para (DNC). Pelo Teorema anterior (dualidade fraca), mais (4.20), concluímos que

$$L(\overline{x}, \overline{\lambda}) = \phi(\overline{x}) \geq L(x, \lambda), \ \forall \ (x, \lambda) \in I\!D.$$

Portanto, $(\overline{x}, \bar{\lambda})$ é uma solução ótima para (DNC) e os valores ótimos de (PNC) e (DNC) coincidem. □

Obs.: Os principais resultados deste capítulo também podem ser encontrados em [52], [53].

Capítulo 5

Possibilidades Futuras

A seguir listaremos algumas possibilidades de trabalhos futuros, problemas em aberto e mesmo algumas divagações, dando ao leitor uma pequena perspectiva de pesquisa na promissora área de otimização não diferenciável.

5.1 Sobre um algoritmo para resolver problemas de otimização não diferenciáveis

Seja $\mathcal{P}(\Re^n)$ o conjunto das partes do $\Re^n$. Consideremos a seguinte multifunção $G : \Re^n \to \mathcal{P}(\Re^n)$, que a cada ponto $x \in \Re^n$ associa um subconjunto $G(x)$ do $\Re^n$. Em [38] Konnov obteve um algoritmo para encontrar pontos estacionários de G, isto é, encontrar $x \in \Re^n$ tal que $0 \in G(x)$.

No mesmo artigo, Konnov obteve várias aplicações deste algoritmo à problemas de otimização. Em particular ele considerou o seguinte problema:

$$\left.\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & x \in \Omega, \end{array}\right\} \quad (PK)$$

onde $\Omega \subset \Re^n$ é um subconjunto não vazio e $f : \Omega \to \Re$ é uma função localmente Lipschitz e quasiconvexa. Konnov mostrou que se $\Omega = \Re^n$ e f é *Clarke* *regular* para todo $x \in \Re^n$, então seu algoritmo pode ser aplicado ao problema (PK), tomando $G(x) = \partial f(x)$, onde $\partial f(x)$ é o gradiente generalizado de Clarke da função f no ponto x. Lembremos que f é dita Clarke regular no $\Re^n$ se, qualquer que sejam $x, v \in \Re^n$, existe a derivada direcional usual $f'(x; v)$ e a mesma coincide com a derivada generalizada de Clarke $f^o(x; v)$. Como mostra o exemplo abaixo, esta hipótese de Clarke regularidade é muito restritiva.

Exemplo Consideremos a função $f : \Re \to \Re$ dada por

$$f(x) = \begin{cases} -x-2 & \text{se} \quad x \leq -1, \\ x & \text{se} \quad -1 < x < 0, \\ 0 & \text{se} \quad 0 \leq x. \end{cases}$$

f é claramente localmente Lipschitz e quasiconvexa. No entanto f não é Clarke regular. De fato, é fácil verificar que $f'(0, v) = f(v)$. Como f não é convexa então a função $v \to f'(0, v)$ não é convexa e portanto não pode coincidir com a derivada generalizada de Clarke $f^o(0, v)$ a qual é sempre convexa na variável v.

Fazendo uso da teoria de gradientes generalizados de Ioffe [31], conseguimos estender os resultados de Konnov para o caso em que f é apenas localmente Lipschitz e quasiconvexa sem qualquer hipótese adicional de regularidade (veja [51]). Um resultado em aberto é aplicar o algoritmo de Konnov no caso em que (PK) tenha efetivamente um conjunto de restrições, pois Konnov trata apenas do problema (PK) sem restrições tomando $\Omega = \Re^n$.

5.2 Controle ótimo impulsivo

Outra linha de pesquisa de nosso interesse é o estudo de equações diferenciais controladas do tipo

$$dx(t) = f(t, x(t), u(t))dt + g(t, x(t))\mu(dt),$$

onde aparecem, além do controle convencional $u(t)$, um controle chamado *impulsional* , representado por uma medida positiva μ. A motivação do estudo de tais equações origina-se de problemas de controle do gasto de combustível em veículos espaciais. O controle $\mu(\cdot)$ é uma idealização do controle convencional e modela a situação na qual um veículo pode assumir altas velocidades num curto período de tempo. Um problema interessante a ser estudado é o seguinte:

$$\left.\begin{array}{ll} \text{Minimizar} & h(x(0), x(1)) \\ \\ \text{sujeito a} & dx(t) = f(t, x(t), u(t))dt + g(t, x(t))\mu(dt), \\ \\ & u(t) \in U_t, \text{qtp em } [0,1], \\ \\ & \psi(t, x(t)) \leq 0, \forall t \in [0,1], \\ \\ & \mu \geq 0, (x(0), x(1)) \in C. \end{array}\right\} \quad (CP)$$

Aqui $h : \Re^n \times \Re^n \to \Re$, $f : \Re^n \times \Re^n \to \Re^n$ e $g : [0,1] \times \Re^n \to \Re^n$ são funções dadas. U é um subconjunto de Borel de $[0,1] \times \Re^n$, (U_t denota a seção $\{x : (t,x) \in U\}$) e C é um subconjunto de $\Re^n \times \Re^n$. A função $\psi : [0,1] \times \Re^n \to \Re$ é uma restrição de desigualdade no problema (CP).

Um problema interessante é obter condições necessárias e suficientes de otimalidade na forma de um princípio do máximo de Pontriagyn para o problema (CP).

5.3 Inclusões diferenciais fuzzy

Dada um multifunção $F : [0,T] \times \Re^n \to \mathcal{P}(\Re^n)$, isto é, uma função que a cada $(t,x) \in [0,T] \times \Re^n$ associa o conjunto $F(t,x) \subset \Re^n$, consideremos o problema de encontrar uma função absolutamente contínua $x : [0,T] \to \Re^n$, satisfazendo

$$\dot{x} \in F(t,x), \tag{5.1}$$

onde $\dot{x}$ denota a derivada em relação a t da função x.

A inclusão (5.1) é chamada *inclusão diferencial* e pode ser entendida como uma generalização da equação diferencial clássica

$$\dot{x} = f(t, x),$$

no sentido de que equações diferenciais são inclusões diferenciais onde a multifunção $F(\cdot, \cdot)$ é univalente, isto é, $F(t, x)$ é um conjunto unitário para cada (t, x). Logo, muitas das questões pertinentes a equações diferenciais, como por exemplo, existência e unicidade de soluções, dependência contínua com relação a parâmetros, bifurcação, estabilidade e comportamento assintótico, etc, também são relevantes para o estudo das inclusões diferenciáveis. Apesar da importância do estudo das equações diferenciais tanto do ponto de vista teórico quanto das aplicações, em muitos casos, tais equações parecem ser muito restritivas para descrever certos sistemas de evolução controlados, aparecendo várias dificuldades tais como falta de determinismo das variáveis ou parâmetros envolvidos, desconhecimento das leis que governam o controle para os possíveis estados do sistema, desconhecimento do meio ambiente futuro do sistema, etc. Tais dificuldades são conhecidas de forma genérica como a "fuzziness" do sistema.

É possível traduzir tais problemáticas em termos matemáticos através das chamadas inclusões diferenciáveis. Por exemplo, consideremos o sistema controlado de equações diferenciais

$$\dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U(t), \tag{5.2}$$

onde $x : [0, T] \to \Re^n$ é a função que descreve o estado do sistema, $\dot{x}$ denota a derivada de x com respeito ao tempo, $u : [0, T] \to \Re^k$ é a função controle e $U \subset \Re^k$ é alguma restrição geométrica (física ou econômica, por exemplo) do sistema, no sentido de que $u(t) \in U, \ \forall t \in [0, T]$. Definamos

$$\Gamma(t, x) = f(t, x, U) = \{v \in \Re^n / \text{ existe } w \in U \text{ tal que } v = f(t, x, w)\}.$$

Sob certas hipóteses o sistema controlado (5.2) é equivalente à inclusão diferencial

$$\dot{x}(t) \in \Gamma(t, x(t)). \tag{5.3}$$

Na passagem da equação diferencial para a inclusão diferencial temos uma certa perda, pois de uma forma geral, é mais difícil trabalhar com problemas multívocos, isto é, problemas em que para cada $(t, x(t))$, $\Gamma(t, x(t))$ é um conjunto. Por outro lado, o uso de inclusões diferenciais permite flexibilizar o problema proposto numa tentativa de evitar o confronto direto com a "fuzziness" do sistema.

Um problema em aberto nesta linha de pesquisa é obter um princípio do máximo para o problema de tempo mínimo associado ao sistema (5.2), no caso onde o controle é uma variável fuzzy, via a inclusão diferencial (5.3). Este problema consiste basicamente em encontrar uma solução $x(t)$ da inclusão diferencial (5.3) que parta do estado $x(t_1)$ e chegue no estado $x(t_2)$ de modo a minimizar o tempo percorrido $|t_2 - t_1|$.

Para uma iniciação nos assuntos teoria Fuzzy, "fuzziness" e inclusões diferenciais Fuzzy, sugerimos [24], [55].

5.4 Invexidade e aplicações

Como o leitor deve ter notado, o conceito de invexidade é muito potente e com grandes perspectivas de aplicações. Um possível resultado, o qual já temos alguma coisa em mente, é a extensão do teorema de alternativa invexo 2.8, para o caso com tempo contínuo, isto é, onde as funções envolvidas no teorema de alternativa são do tipo $f_i(t, x(t))$ (veja [56]). Também podemos tentar aplicar o conceito de invexidade (nas suas diversas versões) em problemas de controle ótimo, otimização Fuzzy e outros temas aonde o conceito de convexidade pertinente pode ser substituido por algum tipo de invexidade.

Em termos práticos é importante mostrar sob que condições, um problema de programação matemática é invexo com uma função η comum. Em [29] Hanson e Rueda dão condições suficientes para a existência da função η num programa invexo com funções duas vezes diferenciáveis, usando técnicas de programação linear. Além disso, sob certas hipóteses, eles constroem a função η tornando assim seu método mais eficiente do que outros disponíveis na literatura (veja [19]). Um problema em aberto é obter resultados similares aos de Hanson e Rueda mas diminuindo a diferenciabilidade das funções envolvidas.

Um projeto mais ambicioso é tentar contruir uma teoria de análise invexa, nos moldes da análise convexa clássica (veja [50]). Isto parece um desafio muito grande pois ainda não estamos certos de que os conceitos de conjunto invexo, função pré-invexa e função invexa, na forma como são dados, são os mais apropriados para se construir tal teoria. Como seria um subdiferencial invexo? Sob que condições e tipos de funções tal subdiferencial seria não vazio? Ele teria um cálculo robusto? Em [60] Yang e Craven introduzem um "subdiferencial" para funções invexas e de certa forma respondem a algumas das questões acima. No entanto o artigo deles contém erros e ainda deixam muitas perguntas em aberto.

5.5 Problemas anormais

Os problemas ditos anormais ocorrem quando o multiplicador associado à função objetivo é nulo. Nesse caso as regras de multiplicadores usuais (regra de Lagrange, princípio do máximo de Pontryagin) são satisfeitas trivialmente, sendo de pouca valia na busca da solução ótima.

No contexto diferenciável Avakov [2] (para o caso de restrições de igualdade) e Izmailov [33] (para o caso de restrições de desigualdade) conseguem estender as regras de multiplicadores usuais para alguns casos anormais. A perda nesta extensão é a exigência de mais diferenciabilidade das funções envolvidas. Uma questão em aberto é a seguinte: Será possível tal extensão para o caso não diferenciável?

5.6 Técnicas de deformação em problemas de otimização

Uma técnica interessante para se obter condições suficientes de otimalidade em programação não linear são os chamados métodos de deformação. O método consiste, grosso modo, em "deformar" um problema de otimização em um outro problema mais simples. Entre outras hipóteses esta deformação deveria preservar pontos ótimos. Bobylev [5], [6] estudou vários problemas de deformação. Um problema interessante (e até agora não estudado) seria adaptar o método de deformação para o problema (PNC) (problema com tempo contínuo).

Bibliografia

[1] ABDOUNI, B. EL e THIBAULT, L. (1992) Lagrange multipliers for Pareto nonsmooth programming problems in Banach spaces. *Optimization*. 26, 277-285.

[2] AVAKOV, E. R. (1985) Extremum conditions for smooth problems with equality-type constraints. *Comput. Maths. Math. Phys.*. 25, 24-32.

[3] BEN-ISRAEL, A. e MOND, B. (1986) What is invexity?. *J. Austral. Math. Soc. Ser. B*. 28, 1-9.

[4] BERGE, C. e GHOUILA-HOURI, A. (1965) *Programming, Games and Transportation Networks*. New York, John Wiley.

[5] BOBYLEV, N. A. (1989) Deformation Method of Investigation of Nonlinear Programming Problems I. *Avtomatika i Telemekhanika*. 7, 82-90.

[6] BOBYLEV, N. A. (1989) Deformation Method of Investigation of Nonlinear Programming Problems II. *Avtomatika i Telemekhanika*. 8, 24-33.

[7] BRANDÃO, A. J. V., ROJAS-MEDAR, M. A. e SILVA, G. N. (1996) A Simple Proof of The Multiplier Rule for Nonsmooth Optimization. *Anais do 43o. Seminário Brasileiro de Análise (SBA)*. São Paulo-SP.

[8] BRANDÃO, A. J. V., ROJAS-MEDAR, M. A. e SILVA, G. N. (1996) Programação Matemática não Diferenciável com Tempo Contínuo. *Anais do XIX Congresso Nacional de Matemática Aplicada e Computacional (CNMAC)*. Goiânia-GO.

[9] BRANDÃO, A. J. V., ROJAS-MEDAR, M. A. e SILVA, G. N. (1996) Condiciones Necesarias y Suficientes de Optimalidad en Programación

No-Diferenciable con Tiempo Continuo: Caso Convexo. *Anais do 44o. Seminário Brasileiro de Análise (SBA)*. Ribeirão Preto-SP.

[10] BRANDÃO, A. J. V. e ROJAS-MEDAR, M. A. (1997) Nonsmooth Nonconvex Alternative Theorem and Applications. *Atas do VII Congreso de Matemática Capricórnio COMCA'97.* Antofagasta, Chile.

[11] BRANDÃO, A. J. V., ROJAS-MEDAR, M. A. e SILVA, G. N. (1997) Invex Nonsmooth Alternative Theorem and Applications. Submetido à publicação.

[12] BRANDÃO, A. J. V., ROJAS-MEDAR, M. A. e SILVA, G. N. (1997) Optimality conditions for Pareto nonsmooth nonconvex programming in Banach spaces. Submetido à publicação.

[13] BRÉZIS, H. (1983) *Analyse Fonctionnelle.* Paris, Masson.

[14] CHAN, W. L., HUANG, L. R. e NG, K. F. (1994) On generalized second-order derivatives and Taylor expansions in nonsmooth optimization. *SIAM J. Control Optim.*. 32, 591-611.

[15] CHANKONG, V. e HAIMES, Y.Y. (1983) *Multiojective Decision Making Theory and Methodology Series vol 8.* North-Holland.

[16] CLARKE, F. H. (1990) *Optimization and Nonsmooth Analysis.* Classics in Applied Mathematics 5, SIAM.

[17] COLADAS, L., LI, Z. e WANG, S. (1994) Optimality conditions for multiobjective and nonsmooth minimisation in abstract spaces. *Bull. Austral. Math. Soc.*. 50, 205-218.

[18] COMINETTI, R. e CORREA, R. (1990) A generalized second-order derivative in nonsmooth optimization. *SIAM J. Control Optim.*. 28, 789-809.

[19] CRAVEN, B. D. (1981) Invex functions and constrained local minima. *Bull. Austral. Math. Soc.* 24, 357-366.

[20] CRAVEN, B. D. e GLOVER, M. (1985) Invex functions and duality. *J. Austral. Math. Soc. Ser. A.* 39, 1-20.

[21] CRAVEN, B. D. (1986) Nondifferentiable Optimization by Smooth Approximations. *Optimization*. 17, 3-17.

[22] CRAVEN, B. D. e YANG, X. Q. (1991) A nonsmooth version of alternative theorem. *Utilitas*. 40, 117-128.

[23] CRAVEN, B. D. (1995) *Control and Optimization*. Chapman & Hall.

[24] DIAMOND, P. e KLOEDEN, P. (1994) *Metric Spaces of Fuzzy Sets*. World Scientific.

[25] GEOFFRION, A. M. (1968) Proper efficient and the theory of vector maximization. *J. Math. Anal. Appl.*. 22, 618-630.

[26] GEOFFRION, A. M. (1971) Duality in nonlinear programming: A simplified applications-oriented development. *SIAM rev.*. 13, 1-38.

[27] GIORGI, G. e GUERRAGGIO, A. (1996) Various types of nonsmooth invex functions. *J. Inf. Optim. Sci.* . 17, 137-150.

[28] HANSON, M. A. (1981) On sufficiency of the Kuhn-Tucker conditions. *J. Math. Anal. Appl.*. 80, 545-550.

[29] HANSON, M. A., RUEDA, N. (1989) A sufficient condition for invexity. *J. Math. Anal. Appl.*. 138, 193-198.

[30] HAYASHI, M. e KOMIYA, H. (1980). Perfect Duality and Convexlike Programs. *J. Optim. Theory and Appl.*. 38, 179-189.

[31] IOFFE, A. D. (1984) Approximate Subdifferentials and Applications I: The Finite Dimensional Theory. *Trans. Amer. Math. Soc.*. 281, 389-416.

[32] IOFFE, A. D. (1986) Approximate subdifferentials and aplications II: functions on locally convex spaces. *Mathematika*. 33, 111-128.

[33] IZMAILOV, A. F. (1994) Optimality conditions for degenerate extremum problems with inequality-type constraints. *Comput. Maths. Math. Phys.*. 34 723-736.

[34] JEYAKUMAR, V. (1985) Convexlike Alternative Theorems and Mathematical Programming. *Optimization*. 16, 643-652.

[35] JEYAKUMAR, V. (1987) On optimality conditions in nonsmooth inequality constrained minimization. *Num. Funct. Anal. and Optimiz.*. 9, 535-546.

[36] JEYAKUMAR, V. e ZAFFARONI, A. (1996) Asymptotic conditions for weak and proper optimality in infinite dimensional convex vector optimization. *Numer. Funct. Anal. and Optimiz.*. 17, 323-343.

[37] KIM, D. S., LEE, G. M., PARK, J. Y. e SON, K. H. (1993) Control problems with generalized invexity. *Math. Japonica*. 38, 263-269.

[38] KONNOV, I. V. (1996) A General Approach to Finding Stationary Points and the Solution of Related Problems. *Comp. Maths. Math. Phys.*. 36, 585-593.

[39] MANGASARIAN, O. L. (1994) *Nonlinear Programming*. Classics in Applied Mathematics 10, SIAM.

[40] MINAMI, M. (1983) Weak Pareto-Optimal necessary conditions in a nondifferential multiobjective program on a Banach space. *J. Optim. Theory Appl.*. 41, 451-461.

[41] MISHRA, S. K. e MUKHERJEE, R. N. (1996) On generalised convex multiobjective nonsmooth programming. *J. Austral. Math Soc. Ser B*. 38, 140-148.

[42] MOND, B. e SMART, I. (1988) Duality and sufficiency in control problems with invexity. *J. Math. Anal. Appl.*. 136, 325-333.

[43] MOND, B. e HUSSAIN, I. (1989) Sufficient optimality criteria and duality for variational problems with generalized invexity. *J. Austral. Math. Soc. Ser. B*. 31, 108-121.

[44] MUKHERJEE, R. N. e MISHRA, S. K. (1996) Multiobjective programming with semilocally convex functions. *Journal of Math. Anal. and Applications*. 199, 409-424.

[45] PÁLES, Z. e ZEIDAN, V. (1996) Generalized Hessian for $C^{1,1}$ functions in infinite dimensional normed spaces. *Mathematical Programming*. 74, 59-78.

[46] PARETO, V. (1869) *Cours d'Economic Politique.* Lausanne: Rouge.

[47] PHUONG, T. D., SACH, P. H. Sach e YEN, N. D. (1995) Strict lower semicontinuity of the level sets and invexity of a locally Lipschitz function. *JOTA.* 87, 579-594.

[48] PINI, R. (1991) Invexity and Generalized Convexity. *Optimization.* 22, 513-525.

[49] REILAND, T. W. (1990) Nonsmooth invexity. *Bull. Austral. Math. Soc.*. 42, 437-446.

[50] ROCKAFELLAR, R. T. (1969) *Convex Analysis.* Princeton University Press, Princeton, N.J..

[51] ROJAS-MEDAR, M. A., BRANDÃO, A. J. V. Brandão e SILVA, G. N. (1997) Encontrando Pontos Estacionários via Gradientes Generalizados. *Anais do XX Congresso Nacional de Matemática Aplicada e Computacional (CNMAC).* Gramado-RS.

[52] ROJAS-MEDAR, M. A., BRANDÃO, A. J. V. e SILVA, G. N. (1997) Sufficient conditions of optimality for nonsmooth continuous-time optimization problems. *16o. International Symposium on Mathematical Programming, EPFL.* Laussane, Suiça.

[53] ROJAS-MEDAR, M. A., BRANDÃO, A. J. V. e SILVA, G. N. (1998) Nonsmooth continuous-time optimization problems: sufficient conditions. Aceito para publicação em Journal of Mathematical Analysis and Applications.

[54] ROJAS-MEDAR, M. A., BRANDÃO, A. J. V. e SILVA, G. N. (1997) Nonsmooth continuous-time optimization problems: necessary conditions and nonconvex duality. Submetido à publicação.

[55] ROMÁN-FLORES, H. e ROJAS-MEDAR, M. A. (1997) Introducion al Analisis Fuzzy. *46o. Seminário Brasileiro de Análise.* Niterói-RJ.

[56] SILVA, G. N., ROJAS-MEDAR, M. A. e BRANDÃO, A. J. V. (1997) Invex Nonsmooth and a continuous-time version of Gordan's Theorem. *Anais do XX Congresso Nacional de Matemática Aplicada e Computacional (CNMAC).* Gramado-RS.

[57] THIBAULT, L. (1980) Subdifferentials of compactly Lipschitzian vector valued functions. *Annali Math. Pura Appl.*. 125, 157-192.

[58] WEIR, T. e MOND, B. (1987) Duality for generalized convex programming without a constraint qualification. *Utilitas Mathematica*. 31, 233-242.

[59] WEIR, T. e MOND, B. (1988) Pre-invex functions in multiple objective optimization. *J. Math. Anal. Appl.*. 136, 29-38.

[60] YANG, X. Q. e CRAVEN, B. D. Craven (1992) Necessary optimality conditions with a modified subdifferential. *Optimization*. 22, 387-400.

[61] YANG, X. Q. (1994) Generalized second-order derivatives and optimality conditions. *Nonlinear Analysis: Theory, Methods & Applications*. 23, 767-784.

[62] YANG, X. Q. e JEYAKUMAR, V. (1992) Genereralized second order derivatives and optimization with $C^{1,1}$ functions. *Optimization*. 26, 165-185.

[63] ZALMAI, G. J. (1985) Optimality conditions and Lagrangian duality in continuous-time nonlinear programming. *J. Math. Anal. Appl.*. 109, 426-452.

[64] ZALMAI, G. J. (1985) Continuous-time Generalization of Gordan's transposition theorem. *J. Math. Anal. Appl.*. 110, 130-140.

[65] ZALMAI, G. J. (1985) The Fritz John and Kuhn-Tucker optimality conditions in continuous-time nonlinear programming. *J. Math. Anal. Appl.*. 110, 503-518.

[66] ZALMAI, G. J. (1985) Sufficient optimality conditions in continuous-time nonlinear programming. *J. Math. Anal. Appl.*. 111, 130-147.

[67] ZALMAI, G. J. Zalmai (1990) Generalized sufficiency criteria in continuous-time programming with application to a class of variational-type inequalities. *J. Math. Anal. Appl.*. 153, 331-355.