"TÉCNICAS DE OTIMIZAÇÃO NÃO-DIFERENCIAVÉL E APLICAÇÕES"

Este éxemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. MARKO ANTONIO ROJAS MEDAR e aprovado pela Comissão Julgadora.

Campinas, 30 de setembro de 1988.

José Luiz Boldrini

Prof. Dr. JOSÉ LUIZ BOLDRINI
Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para obtenção do título de Mestre em MATEMÁTICA APLICADA.

Agredecimentos

Ao meu orientador e amigo, Prof. José Luiz Boldrini, pela sua valiosa contribuição para que este trabalho fosse realizado.

Aos Profs. Rodney Carlos Bassanezi, José Antonio Scaramucci e José Mário Martinez, pelo Interesse mostrado por este trabalho, e por sua gentiliza ao aceitar fazer parte da banca examinadora.

Aos Profs. Victor Hugo Salinas (Universidad de La Serena Chile), Rafael Correa e Servet Martínez (Univerdidad de Chile), pelo incentivo ao meu ingresso na Pós-Graduação.

À Sra. Maria Aparecida, por sua paciência e dedicação ao "traduzir" e datilograr o manuscrito desta tese.

A meus pais e irmãos pelo seu apoio incondicional.

Agradecemos a toda a comunidade do Departamento de Matemática Aplicada - professores, colegas e pessoal administrativo- pelo ambiente cordial e de trabalho, o qual desenvolví meus estudos e a presente tese.

Devo também agradecer o apoio econômico prestado pelo CNPq, CAPES e FAPESP.

Finalmente agradeço a minha esposa Mitza pela paciência e compreensão durante o desenvolvimento deste trabalho.

Campinas, 30 de setembro de 1988.

ÍNDICE

Introdução

O presente trabalho apresenta alguns conceitos do Análise não-linear e não-diferenciável, bem como algumas de suas aplicações.

O primeiro capítulo está dedicado ao estudo da noção de Multiaplicação, isto é, uma relação que a cada x de um certo conjunto X associa um subconjunto $\Gamma(x)$ de outro conjunto Y. Esta ferramenta matemática é muito boa para modelar sistemas (econômicos, físicos, etc.) onde há várias possibilidade de comportamento.

Começaremos primeiramente dando as diferentes noções de continuidade e semicontinuidades, e suas equivalências, estudamos também a composição de elas. Logo passamos a explicar o problema de inclusão diferencial, isto é, relações que generalizam o conceito de equações diferencias e são do tipo $\dot{x}(t) \in \Gamma(x(t))$, onde é uma multiaplicação. Apresentamos as duas técnicas básicas para abordar o problema anterior, que são: seleções e pontos fixos. Para mostrar alguns dos resultados de ponto fixo, faremos uso do célebre princípio variacional de Ekeland. Como término do capítulo, apresentamos duas aplicações à otimização: algoritmos e ótimos de Pareto.

No segundo capítulo extendemos o cálculo diferencial clássico para as funções localmente lipschitzianas, apresentando a noção de Gradiente Generalizado, devida a F. H. Clarke, e suas propriedades (soma, produto, regras da cadeia, Teorema do valor médio de Lebourg, etc), e tembém conceitos geométricos a ela associados (cones tangente, normal, etc). A seguir passamos às aplicações à otimização: caracterização dos pontos ótimos de problemas de otimização não diferenciáveis (isto é, onde a função a otimizar e as restrições são não-diferenciáveis), apresentamos o Teorema de Kuhn - Tucker generalizado, como também resultados relativos à programação geométrica.

O terceiro capítulo esta dedicado ao uso das técnicas dos capítulos anteriores a problemas do cálculo das variações e controle ótimo não regulares, faz-se um estudo detalhado das condições suficientes de otimalidade devidas a V. Zeidan, e apresentamos uma aplicação ao problema do "regular linear com diodo".

Os requesitos para lêr o presente trabalho se reduzem a um curso básico de topologia, por exemplo: L. Schwartz "Analyse" Hermann, Paris (1970), e conhecimento elementares de análise convexa [60].

Capítulo I

Multiaplicações, inclusões Diferenciais e Aplicações.

Introdução

A teoria das multiaplicações é uma área da matemática que tem se desenvolvido intensamente nos últimos anos e é uma ferramenta essencial do análise funcional não linear.

Uma multiaplicação, associa aos pontos de algum conjunto X um subconjunto de outro conjunto Y, na forma clássica, elas eram vistas como uma função ordinária de X em P(Y), onde P(Y) é o conjunto de partes de Y (dotado de alguma estrutura), de modo que não havia necessidade de uma teoria independente para as multiaplicações. Porém, o argumento anterior não é satisfatório. na investigação de problemas específicos, os quais nascem naturalmente nesta teoria.

A ampla gama de aplicações, e os distintos métodos de resolução destas aplicações, nos leva à formação da teoria mais desenvolvida das multiaplicações. Elas tem aplicações em topologia, análise, teoria de controle, economia, algorítmos, teoria de jogos, equaçoes diferenciais com discontinuidades, álgebras de Von Neumann, geometria, otimização, programação matemática,etc.

Nesta seção fornecemos algumas definições e questões importantes para a teoría que desenvolveremos no capítulo II, como também, para algumas aplicações neste mesmo capítulo.

Adotaremos as seguintes notações que serão considerados em todo o trabalho.

Seja Z um espaço topológico arbitrário. Definimos as seguintes classes de subconjuntos de Z:

$$Cl(Z) = \{ A \subset Z \mid A \text{ é fechado e não vazio} \}$$

$$C(Z) = \{ A \subset Z \mid A \text{ é compacto e não vazio} \}$$

Se Z tem ademais estrutura vetorial, denotamos

$$CC(Z) = \{ A \subset Z \mid A \text{ é fechado, convexo e não vazio} \}$$

$$K(Z) = \{ A \subset Z \mid A \text{ é convexo e não vazio} \}$$

No espaço métrico (Z,d) para $A \in Cl(Z)$ e $r > 0$ seja

$$K(A,r) = \{ y \in Z \mid d(y,A) < r \}$$

onde $d(y,A) = \inf \{ d(x,y) \mid x \in A \}$.

Chamaremos a $K(A;r)$ a bola aberta com centro A e raio r, no caso que $A = \{x_0\}$, usaremos a notação usual $B(x_0;r)$.

1.1 Multiaplicações

a) Definição e exemplos de multiaplicações

O seguinte conceito será muito utilizado no seguinte trabalho.

Definição 1. Sejam X e Y dois conjuntos não vazios. Se a cada elemento $x \in X$ associamos um subconjunto $\Gamma(x)$ de Y (possivelmente vazio), dizemos que a correspondença $x \to \Gamma(x)$ é uma <u>multiaplicação</u> (ou operador multívoco) <u>de X em Y</u>. Denotaremos a tal multiaplicação por $\Gamma: X \rightrightarrows Y$. Δ

<u>Exemplos</u>

1) Naturalmente uma função $f: X \to Y$ é um caso especial de multiaplicação, pois, basta definir $\Gamma(x) = \{f(x)\}$.

2) Associemos com qualquer função convexa f (isto é, uma função $f: X \to \mathbb{R}$ tal que: $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$ $\forall \lambda \in [0,1]$, $\forall$ $x,y \in X$) semicontínua inferior de um espaço vetorial localmente convexo Hausdorff X a $\mathbb{R} \cup \{+\infty\}$ seu <u>subdiferencial</u> $\partial_c f(x)$, definido por

$$\partial_c f(x) = \{p \in X^* \mid f(y) \geq f(x) + \langle p, y-x \rangle \quad \forall y \in X\}$$

(onde X^* é o dual topológico de X e $\langle .,. \rangle$ é o produto dualidade canônico entre X e X^*), o qual é um subconjunto convexo fechado de X^*, o qual pode ser vazio (porém, no caso que f seja continua, $\partial_c f(x) \neq \emptyset$ $\forall x \in X\}$.

Assim definimos uma multiaplicação $\partial_c f(.): X \rightrightarrows X^*$.

O conceito de subdiferencial generaliza a noção de gradiente, no sentido de que, se f tem um gradiente $\nabla f(x) \in X^*$ em $x \in X$, então $\partial_c f(x) = \{\nabla f(x)\}$.

A multiaplicação $x \rightrightarrows \partial_c f(x)$ tem um papel muito importante nas aplicações, veja por exemplo, J.P. Aubin [6] , H.Brézis [17] , Acker e F. Dickstein [2] , J.J.Moreau [60] , R.T. Rockafellar [66] e I Ekeland e R.Teman [39] .

Definição 2

a) Dizemos que a multiaplicação Γ é própria se existe ao menos um ponto $x \in X$, tal que: $\Gamma(x) \neq \emptyset$, isto é, Γ não é a multiaplicação constante $\emptyset$. Neste caso dizemos que o subconjunto

$$\mathrm{Dom}\ (\Gamma) = \{\, x \in X \mid \Gamma\ (x) \neq \emptyset \,\}$$

é o domínio de Γ.

b) O Gráfico de uma multiaplicação , é o subconjunto de X x Y definido por

$$\mathrm{Graf}(\Gamma) = \{(x,y) \in X\mathrm{x}Y \mid y \in \Gamma(x)\}.$$

Notar que o Dom (Γ) é a projeção sobre X do Graf(Γ).

c) A imagem de Γ , é o subconjunto de Y definido por:

$$\mathrm{Im}(\Gamma) = \bigcup_{x \in X} \Gamma(x) = \bigcup_{x \in \mathrm{Dom}(\Gamma)} \Gamma(x)$$

que corresponde à projeção sobre Y do Graf(Γ).

d) Definimos a inversa da multiaplicação Γ, denotada por

Γ^{-1}, a qual é uma multiaplicação de Y em X, por:

$$\Gamma^{-1}(y) = \{ x \in X \mid y \in \Gamma(x) \}$$

o mais geral, se $B \subset Y$, $B \neq \emptyset$, se tem

$$\Gamma^{-1}(B) = \{ x \in X \mid \Gamma(x) \cap B \neq \emptyset \}, \quad \Gamma^{-1}(\emptyset) = \emptyset$$

Notar que $\Gamma^{-1}(B) \subset \mathrm{Dom}(\Gamma)$. Quando Γ é uma aplicação usual sobre X, isto é $\Gamma(x) = \{ f(x) \}$, Γ^{-1} corresponde à inversa usual. Ademais é evidente que a inversa de Γ^{-1} e Γ, o que quer dizer que:

$$y \in \Gamma(x) \iff x \in \Gamma^{-1}(y) \iff (x,y) \in \mathrm{Graf}(\Gamma) \qquad \Delta$$

Exemplos 4. Nas mesmas notações do exemplo 2), temos que:

$$(\partial_c f)^{-1} = \partial_c f^*$$

onde $\partial_c f^*$ é o subdiferencial da <u>função polar ou conjugada de f</u>, isto é, da seguinte função

$$f^*(y) = \mathrm{Sup}\ \{ \langle x,y \rangle - f(x) \mid x \in X \} \quad \text{com} \quad y \in X^*$$

b) <u>Tipos de Continuidade e equivalências</u>

No seguinte X e Y denotaram dois espaços topológicos e Γ uma multiaplicação de X e Y.

Passemos agora ao estudo de continuidade de multiaplicações, o que será feito estendendo os conceitos usuais. Advertimos, entretanto, que duas definições são possíveis, equivalentes no caso de funções, não se estendem de forma equivalente ao caso de multiaplicações. Após as definições daremos um exemplo dessa situação.

<u>Definição 3</u>. Seja $\Gamma : X \rightrightarrows Y$ uma multiaplicação .

a) Γ diz-se <u>semicontínua superior (s.c.s.)</u> em um ponto $\bar{x} \in X$ se : para todo W aberto com $\Gamma(\bar{x}) \subset W$, existe uma vizinhança V de $\bar{x}$ tal que: $x \in V \Rightarrow \Gamma(x) \subset W$.

b) Γ diz-se <u>semicontínua inferior (s.c.i.)</u> em um ponto $\bar{x} \in X$ se: para todo W aberto com $\Gamma(\bar{x}) \cap W \neq \emptyset$, existe uma vizinhança V de $\bar{x}$ tal que: $x \in V \Rightarrow \Gamma(x) \cap W \neq \emptyset$.

c) Γ diz-se <u>contínua em um ponto $\bar{x} \in X$</u> se é s.c.s. e s.c.i . em $\bar{x}$.

d) Diz-se que Γ é contínua, s.c.i. ou s.c.s. se possui a propriedade respectiva $\forall\, x \in X$. Δ

<u>Observações</u>

(i) Os nomes anteriores s.c.s. e s.c.i. são um pouco infelizes porque eles <u>não</u> corresponden a generalizações imediatas dos conceitos usuais, de mesmo nome, correspondentes às funções.

(ii) Como dissemos anteriormente, se a multiaplicação se reduz a uma função, as definições anteriores coincidem mas no caso geral elas correspondem a conceitos diferentes.

Considere os exemplos Γ, $\Omega: \mathbb{R} \rightrightarrows \mathbb{R}$ definidas por $\Gamma(0)=[-1,1]$ e $\Gamma(x)=\{0\}$ para $x \neq 0$ e $\Omega(0)=\{0\}$ e $\Omega(x)=[-1,1]$ para $x \neq 0$. Então, Γ é s.c.s. em x=0 mas não é s.c.i em x=0 enquantoque Ω é s.c.i. em x=0 mas não é s.c.s. em x=0.

Uma caracterização imediata da s.c.i é:

Proposição 1. $\Gamma: X \rightrightarrows Y$ é s.c.i. em $\bar{x}$ se e somente se dada qualquer rede ("net") x_α convergindo a $\bar{x}$ e qualquer $\bar{y} \in \Gamma(\bar{x})$, existe uma rede $y_\alpha \in \Gamma(x_\alpha)$ que converge a $\bar{y}$. Δ

Nota. Se X e Y forem métricos, a caracterização dada na proposição anterior é verdadeira com sequências usuais (enumeráveis).

Observação As definições das anteriormente também podem ser dadas pensando-se na multiaplicação como uma função $\Gamma: X \to P(Y)$ e colocando-se topologias convenientes em $P(Y)$ (evidentemente diferentes para s.c.s. e s.c.i.) Se X e Y forem métricos e restringimos o estudo a multiaplicações $\Gamma: X \to C(Y)$ (ou $Cl(Y)$) então pode-se introduzir a (pseudo)-métrica de Hausdorff em $C(Y)$ ($Cl(Y)$) e obter os conceitos de continuidade anteriores em função desta (pseudo)-métrica.

53

Para mais detalhes veja E. Klein e A.C. Thompson [53] .

Outro conceito de continuidade muito útil para os propósitos da otimização é o seguinte:

<u>Definição 4</u>. Seja Γ $X \rightrightarrows Y$ uma multiaplicação. Γ diz-se <u>fechada</u> em $\bar{x} \in X$, se para toda rede x_α convergente a $\bar{x}$, e para toda rede de y_α convergente a $\bar{y} \in Y$ com $y_\alpha \in \Gamma(x_\alpha)$ $\forall_\alpha$, tem-se $\bar{y} \in \Gamma(\bar{x})$.

Dizemos que Γ é fechada, se ela é fechada $\forall x \in X$. Δ

<u>Observações</u>

i) Se Γ é fechada em $\bar{x}$, então $\Gamma(\bar{x})$ é fechado em Y.

ii) Se Γ é uma aplicação usual, tem-se que o fato de ser Γ fechada no ponto $\bar{x}$ e equivalente à continuidade (no sentido de funções) no ponto $\bar{x}$.

iii) No caso de espaços métricos basta usar sequências usuais (numeráveis).

<u>Exemplo 5</u>. Se f é uma função convexa s.c.i. de X em $\mathbb{R}$, onde X é um espaço vetorial localmente convexo Hausdorff, o subdiferencial é uma aplicação fechada. Com efeito se $x_\alpha \to x$ e $x_\alpha^* \to x^*$ (para a topologia fraca) e $x_\alpha^* \in \partial_c f(x_\alpha)$, temos que:

$$< x_\alpha^* , v - x_\alpha > + f(x_\alpha) \leq f(v) \quad \forall \ v \in X$$

onde $< .,. >$ é o produto dualidade canônico entre X e X^*. Portanto, na expressão anterior, tomando o limite em α,

$$< x^*, v-x > + \liminf_{\alpha} f(x_\alpha) \leq f(v) \qquad \forall\, v \in X$$

e pela s.c.i. de f(.), temos que:

$$f(x) \leq \liminf_{\alpha} f(x_\alpha)$$

Assim $< x^*, v-x > + f(x) \leq f(v) \quad \forall\, v \in X \iff x^* \in \partial_c f(x)$.

Temos a seguinte caracterização da definição 4.

<u>Proposição 2</u>. $\Gamma: X \rightrightarrows Y$ é fechada em $\bar{x} \in X$ se e somente se, para todo $\bar{y} \notin \Gamma(\bar{x})$, existem vizinhanças V e W de $\bar{x}$ e $\bar{y}$ respectivamente, tais que:

$$x \in V \Rightarrow \Gamma(x) \cap W = \emptyset .$$

<u>Prova</u>. Direta da definição. Δ

Também verifica-se imediatamente o seguinte:

<u>Proposição 3</u>. $\Gamma: X \rightrightarrows Y$ é fechada se e somente se $Graf(\Gamma)$ é um conjunto fechado em X x Y. Δ

<u>Definição 5</u>. Seja $\Gamma: X \rightrightarrows Y$ uma multiaplicação. Dizemos que é localmente compacta em $\bar{x} \in X$ (ou que tem a propriedade de compacidade local), se existe uma vizinhança V de $\bar{x}$ tal que $\bigcup_{x \in V} \Gamma(x)$ está contido num compacto. Δ

Recordemos que um espaço topológico Y é dito regular se todo ponto possui um sistema fundamental de vizinhanças fechadas. Temos:

Teorema 1. Seja $\Gamma: X \rightrightarrows Y$ uma multiaplicação com $\Gamma(x) \neq \emptyset \quad \forall\ x \in X$.

Temos:

a) Suponhamos Y regular. Se Γ é s.c.s. em $\bar{x}$ e $\Gamma(\bar{x})$ é fechado em Y, então Γ é fechada em $\bar{x}$

b) Se Γ é fechada em $\bar{x}$ e é localmente compacta em $\bar{x}$, então Γ é s.c.s. em $\bar{x}$.

Prova.

a) Seja $y \notin \Gamma(\bar{x})$, com $\Gamma(\bar{x})$ e fechado, temos pela regularidade de Y que existem abertos W e V, tais que $y \in V$, $\Gamma(\bar{x}) \subset W$ com $W \cap V = \emptyset$.

Por outro lado como Γ é s.c.s. em $\bar{x}$, existe uma vizinhança U de $\bar{x}$ tal que: $x' \in U \Rightarrow \Gamma(x') \subset W$, de onde $x' \in U \Rightarrow \Gamma(x') \cap V = \emptyset$, de onde o resultado em virtude da proposição 2.

b) Suponhamos que Γ não seja s.c.s. em $\bar{x}$. Então pela definição da s.c.s. existe um aberto W tal que $\Gamma(\bar{x}) \subset W$, uma rede $\{x_\alpha | \alpha \in D\}$ convergente a $\bar{x}$ e uma rede $\{y_\alpha | \alpha \in D\}$ tais que:

$$y_\alpha \in \Gamma(x_\alpha), \quad y_\alpha \notin W \quad \forall\ \alpha$$

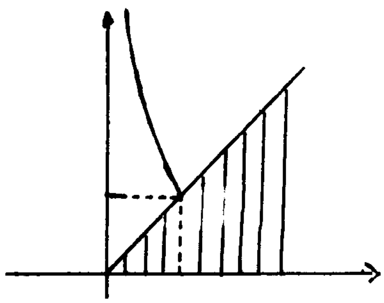
Como Γ é localmente compacta em $\bar{x}$, sem perda de generalidade podemos supor que a rede $\{y_\alpha | \alpha \in D\}$ converge a um ponto y, como Γ é fechada em $\bar{x}$, deduz-se que, $y \in \Gamma(\bar{x})$, mas isto é uma contradição com o fato que $\Gamma(\bar{x}) \subset W$ e $y_\alpha \notin W$. Δ

Observação

i) A hipótese " Γ localmente compacta não pode ser omitida

Considerar por exemplo $X=Y=\mathbb{R}_+$. Definimos Γ por

$$\Gamma(x) = \begin{cases} [\ 0,x\ ] \cup \{1|x\} & \text{se} \quad x > 0 \\ \{0\} & \text{se} \quad x = 0 \end{cases}$$



Então f tem valores compactos, o gráfico de Γ é fechado, mas Γ não é s.c.s. em x=0.

2) Se Γ é fechado em $\bar{x}$ e Y compacto, se tem que Γ é s.c.s. em $\bar{x}$, isto deduz-se trivialmente de b).

Vejamos uma aplicação do teorema anterior:

Exemplo 6. Sob as mesmas notações do exemplo 2, supondo agora f continua (o qual garante que $\partial_c f(x) \neq \emptyset \ \forall\, x \in X$, veja J.P.Laurent [54]) temos que o subdiferencial é uma multiaplicação s.c.s. (x* dotado da topologia fraca*), de fato, em virtude do teorema anterior nos basta provar que $\partial_c f(.) : X \rightrightarrows X^*$ é localmente compacta (para a topologia fraca*), já que no exemplo 5, provamos que ela era fechada. Como é conhecido uma função f convexa contínua é localmente lipschitziana (isto é, dado $x \in X$, existe uma vizinhança V de x e $k > 0$ tal que $\forall\, z,y \in X$, $|f(z)-f(y)| \leq kp(z-y)$, onde p(.) é uma seminorma contínua). Seja $z \in V$ e $x^* \in \partial_c f(z)$, temos: $f(y)-f(z) \geq \langle x^*, y-z \rangle$, logo $\langle x^*, y-z \rangle \leq kp(y-z)$ para todo $y \in V$. Seja $w \in X$, então $w = \lambda(y-z)$ para algum $\lambda \geq 0$ com $y,z \in V$. Assim,

$$\langle x^*, w \rangle = \langle x^*, \lambda(y-z) \rangle \leq \lambda kp(y-z) = k\, p(w)$$

Assim, temos $|\langle x^*, \bar{w} \rangle| \leq 1$ com $\bar{w} = w/kp(w)$.

Pelo teorema de Banach-Alaoglu (Veja W.Rudin [70]), temos a compacidade local.

c) Operações sobre as multiaplicações

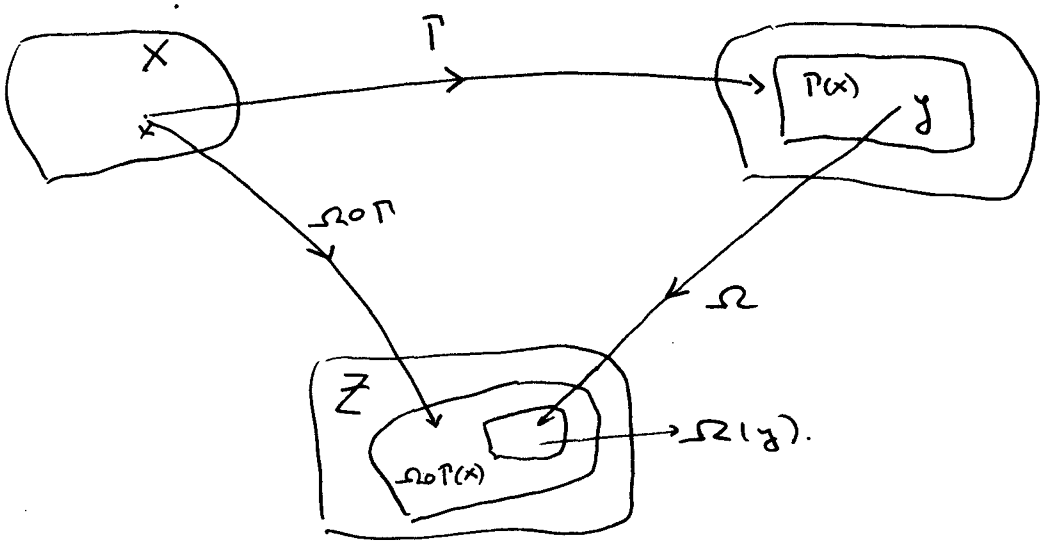
Nesta seção sòmente estudaremos a "composição" entre multiaplicações que necessitaremos na aplicação aos algoritmos.

No seguinte X,Y,Z, denotaram, subconjuntos não vazios dos espaços topológicos T,V,W respectivamente.

Definição 6. Sejam $\Gamma : X \rightrightarrows Y$ e $\Omega : Y \rightrightarrows Z$ duas multiaplicações. Definimos a composição de Γ e Ω, denotada $\Omega o \Gamma : X \rightrightarrows Z$, a qual é uma nova multiaplicação que associa a $x \in X$ o seguinte subconjunto de Z.

$$(\Omega \text{ o } \Gamma)(x) = \bigcup_{y \in \Gamma(x)} \Omega(y). \qquad \Delta$$

Essa definição é ilustrada pela seguinte figura:



Proposição 4. sob as notações anteriores. Suponhamos que

(i) Γ seja fechada em x, e que Ω fechada sobre $\Gamma(x)$;

ii) Y é um conjunto fechado e que toda rede (y_α) tal que: $y_\alpha \in \Gamma(x_\alpha)$ com $x_\alpha \to x$ admite uma subrede convergente (esta condição é satisfeita automaticamente se Y for compacto).

Então, $\Omega \circ \Gamma$ é fechada em x.

Prova

Seja $x_\alpha \to x$ e $z_\alpha \to z$ com $z_\alpha \in (\Omega \circ \Gamma)(x_\alpha)$. Devemos mostrar que $z \in (\Omega \circ \Gamma)(x)$. Selecionemos $y_\alpha \in \Gamma(x_\alpha)$ tal que $z_\alpha \in \Omega(y_\alpha)$ e, conforme a hipótese, sejam y e y_β tal que $y_\beta \to y$. Como Γ é fechada em x, segue que $y \in \Gamma(x)$.

Analogamente, como $y_\beta \to y$, $z_\beta \to z$ e, sendo Ω fechada em y, segue que $z \in \Omega(y) \subset (\Omega \circ \Gamma)(x)$. Δ

Um corolário imediato é:

Corolário 1. Seja $f: X \to Y$ uma aplicação (no sentido usual) e $\Omega : Y \rightrightarrows Z$ uma multiaplicação, então se f for contínua no ponto x e se Ω é fechada em $f(x)$, então $\Omega$of é uma multiaplicação fechada em x.

Prova

É imediatamente da proposição anterior, posto que se f é contínua em x, e se $x_\alpha \to x$, a rede $y_\alpha = f(x_\alpha)$ converge a $y = f(x)$. Δ

Comentário. Muitos outros resultados sobre multiaplicações são disponíveis. De particular interesse são aqueles que garantem semicontinuidade ou fechadura de multiaplicações construídas de forma especial. Por exemplo, pelo uso das operações usuais (soma, produto, limites, união, interseção , etc), pode-se consultar sobre estes tópicos C. Berge [12], C. Castaing e M. Valadier [20], E. Klein e A.C. Thompson [53] e as referencias citadas nos mesmos.

1.2 Inclusoes Diferenciais

As Inclusões Diferenciais forem primeiramente estudadas por Marchand (1934) e Zaremba (1936) sob o nome equações contingentes (ou paratingentes).

Uma inclusão diferencial é definida como a relação

$$\dot{x}(t) \in \Gamma(t,x(t)) \;,\; x(0)=x_0 \qquad \text{(ID)}.$$

onde $\dot{x}(t)$ denota a derivada da função desconhecida $x=x(t)$. Com respeito a t, e Γ é uma multiaplicação dada, a qual associa a cada par $(t,x(t))$ (em algúm domínio U) o conjunto $\Gamma(t,x)$, $x(t)$ deve pertencer a um espaço vetorial (usualmente $\mathbb{R}^n$) e $t \in \mathbb{R}$ (usualmente $t \geq 0$)

Se para cada $(t,x) \in U$ o conjunto $\Gamma(t,x)$ consiste de um único ponto, então a inclusão Diferencial (ID) é uma equação diferencial usual.

Uma equação diferencial específica em cada ponto de U uma direção que deve ser tangente à solução (mais precisamente, ao gráfico da solução, é o chamado campo vetorial associado), porém, a inclusão diferencial especifica em cada ponto de U um conjunto de direções, e a solução pode ser tangente a qualquer destas. È claro que o problema (ID) é muito geral e complexo, nós só descreveremos nesta seção resultados sobre a existência de soluções da (ID), obviamente, sob algumas restrições sobre Γ (s.c.i. ou s.c.s.) e sobre suas imagens (convexidade, compacidade).

a. Definição e exemplos

Por simplicidade de exposição trataremos apenas do caso da inclusão diferencial "autônoma", descrita a seguir.

Definição 7. Consideremos uma multiaplicação $\Gamma: X \rightrightarrows X$, onde X é um espaço de Banach. Uma função absolutamente contínua $x: [t_0,T) \to X$ é chamada uma solução do problema de inclusão diferencial se para quase todo $t \in [t_0,T)$ tem-se:

$$\dot{x}(t) \in \Gamma(x(t)) \quad , \quad x(t_0)=x_0 \in X. \qquad \Delta$$

Vejamos alguns exemplos:

Exemplos

7) Provavelmente este é o exemplo mais conhecido, pois du-

rante os anos 60 e 70 foi muito pesquisado e deu um novo impulso à teoria de equações diferenciais não lineares. Os tipos de inclusões a que estamos nos referindo são da forma:

$$\dot{x}(t) \in -A(x(t)),\ x(0)=x_0$$

onde A é um operador multívoco maximal monótono.

Notar que o subdiferencial da análise convexa é um caso particular destes operadores.

Boas referencias sobre estes tipos de inclusões diferenciais são H. Brézis [17] e L. Barbu [11].

8) (Desigualdades Variacionais). Seja V: $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ uma função contínua e consideremos a desigualdade diferencial:

$$V(x,\dot{x}) \leq 0 \quad , \quad x(0)=x_0.$$

O problema anterior pode ser formulado em termos de um problema de inclusão diferencial com

$$\Gamma(x) = \{ v \in \mathbb{R}^n \ / \ V(x,v) \leq 0, \quad v=\dot{x} \}.$$

9) (Controle) Consideremos o sistema controlavel

$$\dot{x} = f(x,u) \qquad \text{(c)}$$

onde x é o vetor de fase em $\mathbb{R}^n$, $\dot{x}$ é o vetor velocidade, e $u=u(t)$ é o controle sujeito à restrição geométrica: $u(t) \in U$. Com U um conjunto de $\mathbb{R}^n$. Consideremos o conjunto de todas as velocidades admissíveis do sistema no ponto x no espaço de fase $\mathbb{R}^n$. Este conjunto $f(x,U)$ consiste de todos os vetores $f(x,u)$, onde u é um ponto arbitrário de U. Se agora $x(t)$ é uma trajetória do sistema controlável (c) com controle admissível $u(t)$, então

$$\dot{x}(t) \in f(x(t),U) \quad \text{para quase todo } t.$$

Isto nos leva ao conceito de inclusão diferencial

$$\dot{x} \in f(x,U) \qquad (*)$$

sob hipóteses fracas o sistema controlável (c) com as restrições geométrica: $u(t) \in U$ é equivalente à inclusão diferencial (*), i.e, para qualquer solução $x(t)$ de (*) existe um controle admissível $U(t) \in U$ tal que é trajetória de (c) com este controle 14
$u(t)$ (Lema de Filippov, veja por exemplo J.L. Boldrini [14]).

Existem duas técnicas básicas para provar a existência de soluções do problema de inclusão diferencial que são: Técnica de seleções e a técnica de ponto fixo, que é o que apresentaremos nas duas próximas seções.

b) <u>A técnica de Seleções</u>

Suponhamos que exista uma função $f:[t_o,T] \times \mathbb{R}^n \to \mathbb{R}^n$ suficientemente regular, contínua, por exemplo, com a propriedade de que para todo $(t,x) \in [t_o,T) \times \mathbb{R}^n$ tem-se $f(t,x) \in \Gamma(t,x)$. Então, qualquer solução da equação diferencial $\dot{x}=f(t,x)$, que no caso de continuidade de $g(.,.)$, é garantida existir por um teorema de Caratheodory, será necessariamente solução de (ID).

A argumentação anterior introduz a necessidade do estudo da seguinte questão: dada uma multiaplicação Ω: E $\rightrightarrows$ F, onde E e F são espaços com alguma estrutura (topológicas, de medida, etc), em que condições existe uma função φ: E $\to$ f tal que para todo $x \in E$ tem-se $\varphi(x) \in \Omega(x)$ e $\varphi(.)$ tem certas propriedades de regularidade (mensuralidade, continuidade, etc.)? .
Tal $\varphi(.)$, se existir, é chamada uma <u>seleção</u> (mensurável, contínua, etc) da multiaplicação Ω. A resposta à questão colocada fornece imediatamente resultados sobre existência de soluções de inclusões diferenciais. Por isto estudaremos um pouco esta questão.

<u>Seleções Contínuas</u>

<u>Definição 8</u>. Sejam X e Y espaços topológicos e $\Gamma: X \rightrightarrows Y$ uma multiaplicação. <u>Uma seleção contínua de Γ</u> é uma função contínua

$f: X \to Y$ tal que $f(x) \in \Gamma(x) \quad \forall x \in X$. Δ

Evidentemente, para garantir a existência de seleções contínuas são necessárias hipóteses de continuidade sobre a multiaplicação. Infelizmente a s.c.s. da multiaplicação não é suficiente para garantir a existência de seleção contínua, como pode ser facilmente visto pelo exemplo seguinte: a multiaplicação s.c.s. $\Gamma: \mathbb{R} \rightrightarrows \mathbb{R}$ dada por $\Gamma(x) = \{0\}$ se $x < 0$, $\Gamma(0) = [0,1]$, $\Gamma(x) = \{1\}$ se $x > 0$ não admite seleção contínua.

Na tentativa de se provar tal existência, pode-se adaptar duas posturas. A primeira consiste em dar uma regra que a cada $x \in X$ seleciona um ponto com uma determinada propriedade de $\Gamma(x)$ e a seguir mostrar que a função assim obtida é contínua (sob certas condições). Com esta postura, obtem-se vários teoremas. Teorema da Seleção Minimal, Teorema de Seleção de Chebishev, Teorema da Seleção Baricêntrica, cada um deles com suas vantagens e desvantagens.

Para que o leitor tenha uma idéia do tipo de resultado obtido, enunciaremos o:

Teorema 2 (da seleção Minimal) Seja X um espaço métrico, Y um espaço de Hilbert e $\Gamma: X \rightrightarrows Y$ com valores convexos fechados ($\Gamma: X \to CC(Y)$ e contínua. Então a função $m: X \to Y$ dada por $m(x) = \Pi_{\Gamma(x)}(0)$, onde $\Pi_{\Gamma(x)}(0)$ indica a projeção de $0 \in Y$ sobre o conjunto convexo fechado $\Gamma(x) \subset Y$, isto é, o ponto de $\Gamma(x)$ de menor norma, é uma seleção contínua de Γ. Δ

A segunda postura a ser adotada é a de obter seleções de caráter local (isto é, definidas em uma vizinhança de cada ponto) e construir uma seleção contínua global (isto é, definida para todo X) utilizando uma partição da unidade.

Um teorema clássico com esta postura é o seguinte:

Teorema 3 (de seleção de Michael) Sejam X um espaço métrico, Y um espaço de Banach e $\Gamma: X \rightrightarrows Y$ s.c.i. com valores convexos e fechados. Então, existe uma seleção continua de Γ Δ

Pelo resultado anterior vemos que as multiaplicações s.c.i. (com valores em CC(Y)) se comportam melhor do que as s.c.s. com relação à existência de seleções contínuas. Para estas últimas, entretanto, há teoremas de seleções aproximados que em muitas situações são suficientes.

Teorema 4 (de Seleção aproximada I) Seja X um espaço métrico compacto, Y um espaço normado. Seja $\Gamma: X \rightrightarrows K(Y)$ s.c.s.. Então para $\varepsilon > 0$ arbitrário, existe uma função contínua $f: X \to K(\mathrm{Img}(\Gamma), \varepsilon) \cap \mathrm{co}\,\mathrm{Im}(\Gamma)$, (onde coA denota a envolvente convexa de A) que depende de ε, tal que se F e G são os gráficos de f e Γ respectivamente,

$$\sup \{ d(z,G) \mid z \in F \} \leq \varepsilon$$

Prova. Pode ser vista em A Cellina [21]. Δ

Outro resultado na linha do teorema anterior é o seguinte:

Teorema 5 (de Seleção Aproximada II) Sejam X um espaço de Banach, Y um espaço de Hilbert e $\Gamma: X \rightrightarrows Y$ s.c.s. com valores convexos fechados. Então, para todo $\varepsilon > 0$ existe uma função localmente lipschitziana $f_\varepsilon : X \to Y$ tal que $f_\varepsilon(x) \subset \overline{co}\ Im(\Gamma)$ e

$$Graf(f_\varepsilon) \subset Graf(\Gamma) + \varepsilon B$$

onde $B = B_X \times B_Y$, sendo B_X e B_Y as bolas unitárias com centro em 0 de X e Y, respectivamente.

Prova. A prova deste resultado pode ser encontrado em J.P.Aubin e A.Cellina [7], Teorema 1, pág 84. Δ

Comentários

Existem muitos outros resultados relativos a seleções, uma excelente apresentação é dada em [7]. Para seleções mensuráveis, recomendamos o "Survey" de D.H.Wagner [77] e o de A. Ioffe [50].

Vejamos algumas aplicações do exposto a inclusões diferenciais.

Proposição 6. Sejam $\Gamma: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ uma multiaplicação s.c.i. com valores convexos fechados não vazios. Então existe uma solução da inclusão diferencial associada a Γ.

Prova: A prova é extremamente simples: utilizando o Teorema de seleção de Michael, obtemos uma seleção contínua $f: \mathbb{R}^n \to \mathbb{R}^n$ de Γ. Um teorema clássico de Caratheodory garante então a existência de uma solução $x: [t_0, T) \to \mathbb{R}^n$, ($|\tau - t_0|$ suficientemente pequeno) da equação diferencial $\dot{x} = f(x)$, $x(t_0) = x_0 \in \mathbb{R}^n$ (Veja J.Hale [86]). Esta é também uma solução de $\dot{x} \in \Gamma(x)$, $x(t_0) = \dot{x}_0$. Δ

A utilização do Teorema de Caratheodory restringe a resultado anterior a dimensão finita. Para multiaplicações s.c.s. em espaços de Hilbert, para os quais se dispõe da informação de compacidade local da seleção minimal, é possível se utilizar a seguinte linha de raciocínio:

O teorema da seleção Aproximada II fornece uma seleção localmente lipschitz da multiaplicação e isto garante a existência de solução local mesmo em dimensão infinita.

Infelizmente a solução obtida é apenas uma solução aproximada (já que a seleção é aproximada) e tem-se que utilizar um teorema de convergência (Veja [7]) juntamente com a informação sobre a compacidade local da seleção minimal para passar o limite e obter a solução procurada.

Proposição 7 . Seja X um espaço de Hilbert, $x_0 \in X$ e $\Gamma : X \rightrightarrows X$ uma multiaplicação s.c.s. com valores convexos, fechados e não vazios. Suponhamos que a seleção minimal $m: X \to X$ dada por $m(x) = \Pi_{\Gamma(x)}(0)$ é localmente compacta, então existe $T > 0$ e uma função absolutamente contínua $x: [t_0, T) \to X$ que é solução de

$$\dot{x}(t) \in \Gamma(x(t)), \quad x(t_0) = x_0$$

Prova Veja J. P. Aubin e A. Cellina [7], Teorema 3, pág 98. Δ

Observação. O resultado anterior é verdadeiro para espaços de Banach uniformemente convexos.

c) Tecnica do Ponto Fixo

Outra técnica muito utilizada para achar soluções de inclusões diferenciais, é o que utiliza teoremas de ponto fixo. Dada a inclusão diferencial (ID), vamos a introduzir um "Operador integral", §, (na verdade uma multiaplicação que fará o papel correspondente no caso de equações) atuando sobre $C([to, t]) = \{z: [to, T] \to \mathbb{R}^n, \text{ contínua}\}$, isto é, $§: C([to, T]) \rightrightarrows C[to, T])$ dada por

$$§(z(.)) = \{ \varphi \in C([to, T]) \mid \varphi(.) \text{ é absolutamente contínua}$$

$$\text{e } \frac{d\varphi}{dt}(t) = \Gamma(t, z(t)) \text{ para quase todo } t \in [t, T] \} .$$

Então, se $x(t)$ é solução de (ID) devemos ter

$x(.) \in \S(x(.))$

isto é, $x(.)$ deve ser um <u>ponto fixo</u> da multiaplicação $\S$ (a igualdade $x = g(x)$ na definição de ponto fixo de uma função g é substituida pela tertenência $x \in \Gamma(x)$ no caso de uma multiaplicação). Desta forma, somos levados a estudar condições que garantam a existência de pontos fixos e, em particular, um teorema equivalente Clássico de Schauder para operadores compactos atuando sobre convexos, como também generalizações do teorema de ponto fixo de Banach-Picard.

Vejamos primeiramente o seguinte princípio Variacional devido a I. Ekeland [37], o qual nos permitira provar alguns resultados de pontos fixos para multiaplicações, como também alguns resultados do próximo capítulo.

<u>Teorema</u>. (Princípio Variacional de Ekeland [37]). Seja (X,d) um espaço métrico completo, e seja $f:X \to \mathbb{R}\cup\{+\infty\}$ uma função s.c.i., propria (i.e, $f \not\equiv +\infty$) e limitada inferiormente. Seja $\bar{x} \in X$ e $\varepsilon,\lambda > 0$ dados, e suponhamos que

$$f(\bar{x}) \leq \inf\{f(x) \mid x \in X\} + \varepsilon \qquad (*)$$

Então, existe $v \in X$ tal que

(a) $d(\bar{x},v) \leq \lambda$

(b) $f(v) \leq f(\bar{x})$

(c) $f(x) + (\varepsilon \mid \lambda)\, d\,(x,v) > f(v)$, se $x \neq v$.

<u>Prova</u>. Construimos uma sequência encaixada de conjuntos fechados (C_n) da forma seguinte. Seja $x_0 = \bar{x}$, e suponhamos que x_n é dado. Definamos

$$C_n = \{x \in X \mid f(x)+(\varepsilon \mid \lambda)d(x,x_n) \leq f(x_n)\}\ , \tag{1}$$

$$M_n = \inf\{f(x) \mid x \in C_n\} \tag{2}$$

e escolhamos x_{n+1} em C_n com

$$f(X_{n+1}) \leq \frac{1}{2}\,(m_n+f(x_n)). \tag{3}$$

Como $x_{n+1} \in C_n$, a desigualdade triangular mostra que $C_{n+1} \subset C_{n+1}$, e assim $m_n \leq m_{n+1}$. Logo por (3),

$$f(x_{n+1}) - m_{n+1} \leq \frac{1}{2}(f(x_n)+m_n)-m_{n+1} \leq \frac{1}{2}\,(f(x_n)-m_n) \tag{4}$$

Agora, para $x \in C_n$, temos, de (1), (2),

$$d(x,x_n) \leq (\lambda|\varepsilon)\ (f(x_n) - m_n) \leq (\lambda|2^n \varepsilon)\ (f(x_0) - m_0), \tag{5}$$

onde a última desigualdade segue-se de (4). Então (5) mostra que dim $C_n \to 0$. Como X é completo, temos que:

$$\bigcap_{n \in \mathrm{IN}} C_n = \{ v \},\ \text{e que } x_n \to v \quad \text{quando} \quad n \to \infty$$

Agora (a), (b) são consequências diretas de (*) e de fato que $v \in C_0$. Finalmente, suponhamos que

$$f(x)+(\varepsilon|\lambda)d(x,v) \leq f(v). \tag{6}$$

Então, como $v \in C_n$,

$$f(v) + (\varepsilon|\lambda)\ d\ (v,x_n) \leq f(x_n). \tag{7}$$

Agora (6), (7) e a desigualdade triangular mostram que $x \in C_n$ para cada $n \in \mathrm{IN}$. Segue-se que $x=v$, e (c) é estabelecido. Δ

Alguns corolários imediatos são:

<u>Corolário 1</u>. Suponhamos que X seja um espaço métrico completo e que $f\colon X \to \mathrm{IR}_+ \cup \{+\infty\}$, seja própria, positiva e s.c.i. Consideremos $x_0 \in \mathrm{Dom}(f)$ e $\varepsilon > 0$. Existe $\bar{x} \in E$ tal que

(a) $f(\bar{x}) + \varepsilon d(x_0,\bar{x}) \leq f(x_0)$

(b) $\forall\ x \neq \bar{x},\quad f(\bar{x}) < f(x) + \varepsilon d(x,\bar{x})$. Δ

Corolário 2. Sob as hipótese do corolário anterior. Sejam ε , $\lambda > 0$ e um ponto x_0 tal que $f(x_0) \leq \inf f(x) + \varepsilon\lambda$. Existe então $\bar{x} \in X$ tal que

a) $f(\bar{x}) \leq f(x_0)$

b) $d(x_0,\bar{x}) \leq \lambda$

c) $\forall\ x \in X,\quad f(\bar{x}) \leq f(x) + \varepsilon d(x,\bar{x})$. Δ

Observações

1) É suficiente que C_0 seja completo mais que X, como ocorre se Dom(f) é um subconjunto completo de um conjunto incompleto X.

2) O princípio Variacional Ekeland assegura que v é minimal para a ordem parcial implícito em (1) e minora $\bar{x}$.

3) Uma consequência um pouco surpreendente do princípio va- viacional de Ekeland, foi obtida por F. Sullivan, a qual diz ,

que o princípio variacional de Ekeland caracteriza a completitude de um espaço métrico [75].

4) Muitos resultados podem-se derivar do Teorema de Ekeland, por exemplo, em equações diferenciais parciais recomendamos Chang [22] , Figueiredo e Solimini [41] e Szulkin [76], na área de Geometria dos espaços de Banach, pode-se consultar Giles [85].

Para um "survey" recomendamos o excelente artigo de I. Ekeland [38], e a monografia de D.J. de Figueiredo [40], como também o livro de J.P. Aubin e I. Ekeland [8].

Outras referências são Penot [62], Clarke [29], e Borwein [16].

<u>Definição 9</u>. Seja $\Gamma: X \rightrightarrows X$ uma multiaplicação, dizemos que $\bar{x} \in X$ é um <u>ponto fixo de</u> Γ se: $\bar{x} \in \Gamma(\bar{x})$. Δ

<u>Teorema 5</u>. (Caristi). Suponhamos que X seja um espaço métrico completo, e que $\Gamma: X \rightrightarrows X$ uma multiaplicação com imagens não vazias. Além disso que existe uma função $f: X \to \mathbb{R}_+ \cup \{+\infty\}$ semicontínua inferior e não identicamente igual a $+\infty$ tal que

$$\forall\, x \in X \quad \exists\, y \in \Gamma(x) : \quad f(y)+d(x,y) \leq f(x) \qquad (1)$$

Então Γ tem um ponto fixo.

Se f satisfaz a seguinte relação

$$\forall x \in X, \ \forall y \in \Gamma(x) : \quad f(y)+d(x,y) \leq f(x) \tag{2}$$

então existe $\bar{x} \in X$ tal que $\Gamma(\bar{x}) = \{\bar{x}\}$.

Prova. A idéia é fazer uso do Corolário 1 do Princípio Variacional de Ekeland, seja $\bar{x}$ verificando a condição (b) de tal corolário com $\varepsilon < 1$ e $\bar{y} \in \Gamma(\bar{x})$ verificando $f(\bar{y}) \leq d(\bar{x},\bar{y}) \leq f(\bar{x})$. Se $\bar{y}$ é diferente de $\bar{x}$, da desigualdade (b) do corolário mencionado com $x=\bar{y}$ implica que $d(\bar{x},\bar{y}) \leq \varepsilon d(\bar{x},\bar{y})$ o qual é impossível pois $\varepsilon < 1$. Logo $\bar{x}=\bar{y}$. Assim temos que se a condição (1) é satisfeita existe ao menos um ponto fixo e se a condição (2) é satisfeita todos $\bar{y} \in \Gamma(\bar{x})$ são iguais a $\bar{x}$, isto é $\Gamma(\bar{x}) = \{\bar{x}\}$. Δ

Observações

1) Uma função satisfazendo a condição (1) é chamada uma débil entrópia para Γ, e a multiaplicação Γ diz-se débilmente dissipativa.

2) Uma função satisfazendo a condição (2) é chamada uma entropia para Γ, e a multiaplicação Γ diz-se dissipativa.

3) Os nomes anteriores, são devidos ao fato que a multiapli-

cação se associa com um sistema dinâmico, para resultados nesta direção recomendamos o artigo de J.P. Aubin e J. Siegel [9].

Vamos agora mostrar um resultado na qual f não é necessariamente semicontínua inferior; mas, a multiaplicação Γ deve ter gráfico fechado.

<u>Teorema 6</u>. Seja X um espaço métrico completo. Consideremos uma multiaplicação $\Gamma: X \rightrightarrows X$ com gráfico fechado. Se existe uma função f: $X \to \mathbb{R}_+ \cup \{+\infty\}$ não identicamente igual a $+\infty$. Verificando a condição (1) do teorema anterior, a multiaplicação Γ tem ponto fixo

<u>Prova</u>. Tomemos um $x_0 \in \mathrm{Dom}(f)$; vamos construir por recursividade uma sequência de elementos $x_n \in X$ tais que, graças à condição (1) do teorema 5, se tenha:

$$x_{n+1} \in \Gamma(x_n), \quad d(x_{n+1}, x_n) \leq f(x_n) - f(x_{n+1}). \qquad (*)$$

O qual implica que a sequência de números positivos $f(x_n)$ é descrescente: ela converge a um número α . Somando as desigualdades (*) de n=p a n=q-1, a desigualdade triangular implica que

$$d(x_p, x_q) \leq \sum_{n=p}^{q-1} d(x_{n+1}, x_n) \leq f(x_p) - f(xq)$$

Como o membro da direita tende a $\alpha - \alpha = 0$ quando p e q tendem a infinito, temos que a sequência (x_n) é Cauchy, como X é completo, existe $\bar{x} \in X$ tal que $x_n \to \bar{x}$ quando $n \to \infty$. Por outro lado como $(x_n, x_{n+1}) \in Graf(\Gamma)$, e $(x_n, x_{n+1}) \to (\bar{x}, \bar{x})$ quando $n \to \infty$, temos do fato que $Graf(\Gamma)$ é fechado, que $\bar{x}$ é um ponto fixo de Γ. Δ

Agora vamos mostrar o teorema de ponto fixo de Kakutani (1961), o qual é o analogo a teorema do ponto fixo de Schauder [47], para o caso de multiaplicações Primeiramente damos uma definição e um Lema.

<u>Definição 10</u>.Dizemos que um espaço métrico X tem a propriedade de ponto fixo, se toda função contínua de X em X tem um ponto fixo (no sentido usual). Δ

<u>Exemplos</u> 10) $X = \bar{B}(0,1) = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ onde $\|.\|$ é qualquer norma em $\mathbb{R}^n$, tem a propriedade de ponto fixo, graças ao teorema de Brouwer [47].

11) Seja K um convexo compacto de um espaço de Banach X, então K tem a propriedade de ponto fixo, em virtude do teorema de Schauder.

<u>Lema 1</u>. Seja X um espaço métrico compacto que tem a propriedade de ponto fixo. Seja $\Gamma: X \rightrightarrows X$ uma multiaplicação fechada. Suponhamos também que Γ tem uma ε - seleção contínua para todo $\varepsilon > 0$, então Γ tem ponto fixo.

Prova. Denotemos $F = Graf(f)$ e $G = Graf(\Gamma)$. Como Γ tem um ε-seleção contínua para todo $\varepsilon > 0$, consideremos uma sequência $\varepsilon_n \to 0$, logo para cada n, existe $f_n: X \to X$ continua, como X tem a propriedade de ponto fixo, cada f_n tem ponto fixo em X, denotemos do y_n. Como X é compacto, existe uma subsequência (y_m) de (y_n) tal que $y_m \to y_0 \in X$ quando $n \to \infty$. Vamos mostrar que $y_0 \in \Gamma(y_0)$. Temos que:

$$d((y_0,y_0),G \leq d((y_0,y_0),(y_m,y_m))+ d((y_m, f(y_m)),G)$$
$$\leq d((y_0,y_0),(y_m,y_m))+ \sup \{ d(z,G) / z \in F_m \}$$

onde $F_m = Graf(f_m)$. Como $y_m \to y_0$ $(m \to \infty)$, temos que $(y_m,y_m) \to (y_0, y_0)$ $(m \to \infty)$ e pela hipóteses $\sup \{d(z,G) \mid z \in F_m\}$ $\leq \varepsilon_m$, de onde $\sup \{d(z,G) \mid z \in F_m\} \to 0$ quando $m \to \infty$ já que $\varepsilon_m \to 0$. Logo, $d((y_0,y_0,G)=0$, e como G é um conjunto fechado (pois Γ é uma multiaplicação fechada), se tem $(y_0,y_0) \in G$, ou $y_0 \in \Gamma(y_0)$. Δ

Teorema 7. (Kakutani). Seja K um conjunto compacto e convexo de um espaço de Banach X e $\Gamma: K \rightrightarrows K$ com valores convexos fechados e não vazios (i.e., $\Gamma: K \rightrightarrows CC(x)$) s.c.s..

Então Γ tem ponto fixo.

Prova. (A. Cellina [21], 1969). Pelo teorema de seleção aproximada I, existe uma função $f: K \to K$ continua tal que

$\sup \{d(x,G) \mid x \in F\} < \varepsilon \quad \forall\, \varepsilon > 0$, onde F e G são como no lema anterior. Pelo teorema de Schauder, K tem a propriedade de ponto fixo. Pelo lema anterior Γ tem ponto fixo em K. Δ

Fazendo uso do Teorema anterior, pode-se provar o seguinte resultado:

<u>Proposição 8</u>. Sejam X um espaço de Banach, $x_0 \in X$ e $\Gamma : X \rightrightarrows X$ é s.c.s., com valores convexos, fechados e não vazios e tal que exista $M > 0$ tal que para todo $x \in X$, $\|\Gamma\| = \sup\{\|z\| \mid z \in \Gamma(x)\} \leq M$. Então existe uma solução da inclusão diferencial

$$\dot{x}(t) \in \Gamma(x(t)), \quad x(t_0) = x_0. \qquad \Delta$$

<u>Comentário</u> .

1) Existem uma extensa literatura relativa a teoremas de ponto fixo para multiaplicações, como também uma extensa bibliografia com aplicações, com mais detalhes, recomendamos J. P.Aubin e A. Cellina [7], J.P.Aubin [6] , I.Kaneko [52] R.Wegrzyk [78] Z.Dzedzey [36] , e as referencias citadas alí.

2) Muitos outros resultados sobre inclusões diferenciais existem. Por exemplo, para resultados com multiaplicações sem valores convexos e associados a operadores maximais monotônicos, consulte [87] , [8]; para soluções clássicas de inclusões diferenciais [84] .

Os resultados de existência de soluções que apresentarmos são locais; como em equações diferenciais, resultados sobre existência global podem ser obtidos impondo condições de crescimento sobre a multiaplicação.

1.3 Aplicações

a) Otimização Multiobjetivo: ótimos de Pareto.

Nesta seção abordaremos o problema canônico de programação multiobjetivo, e veremos que a existência de um ótimo de Pareto é equivalente a achar um ponto fixo de uma certa multiaplicação.

Começaremos com uma breve revisão de cones, para logo formalizar o problema há ser estudado.

No seguinte X e Y denotaram dois espaços de Banach.

Definição 11. Um subconjunto $K \subset Y$ é dito um cone se para todo $y \in K$ e para todo $\lambda \geq 0$, temos $\lambda y \in K$. Δ

Definição 12. Um cone convexo K, é um cone, o qual é convexo , isto é:

$$\forall\, y_1, y_2 \in K \text{ e } \forall \lambda_1, \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 = 1,\ \lambda_1 y_1 + \lambda_2 y_2 \in K,$$

ou, equivalentemente, $K + K \subset K$, o que nos diz:

$\forall\, y_1, y_2 \in K$ e $\forall \alpha, \beta \geq 0$: $\alpha y_1 + \beta y_2 \in K$. Δ

Um cone K diz-se <u>ponteado</u> se: $K \cap -K = \{0\}$.

Para um cone ponteado K escreveremos

$$y_1 \leq_k y_2 \iff y_2 - y_1 \in K \quad \text{e}$$

$$y_1 <_k y_2 \iff y_2 - y_1 \in K \mid \{0\}\,.$$

<u>Exemplos de Cones</u>.

12) Um subespaço sempre é um cone convexo.

13) Os semiespaços (abertos ou fechados) limitados por um hiperplano que contem a origem.

14) Seja $B \subset Y$ não vazio, o seguinte conjunto é um cone:

$$\{\, y \in Y \mid y = \lim \lambda_n (y_n - y_0) \text{ com } y_0 \in \bar{B},\ \lambda_n \geq 0 \text{ e } y_n \in B,\ y_n \to y_0 \,\}$$

e é chamado <u>cone de deslocamentos aderentes</u> a B em y_0 e é denotado por $D(B;y_0)$ observemos que $D(B, y_0)$ é um cone fechado.

15) Seja A um subconjunto de Y . O subconjunto seguinte:

$$\text{CONE}(A) = \bigcup_{\lambda \geq 0} \lambda \, \text{co}(A)$$

é o mais pequeno cone convexo que contém a A.

O cone anterior chama-se cone convexo engendrado por A.

A aderência do cone anterior, chama-se cone convexo engendrado por A.

Propriedades dos cones convexos

1) Toda interseção de cones convexos é um cone convexo (assim todo subconjunto de Y definido por um número qualquer de desigualdades lineares homogêneas).

2) A imagem de um cone convexo por uma aplicação linear é um cone convexo.

3) O produto Cartesiano $K_1 \times K_2 \times \ldots \times K_n$ é um cone convexo se e somente se $K_1, K_2, \ldots, K_n$ são cones convexos.

4) Para que um cone convexo seja ponteado é suficiente que $\mathrm{ext}(K) = \{0\}$, onde $\mathrm{ext}(K)$ é o conjunto de pontos extremos de K, isto é, $\{ x \in K, \nexists\ x_1, x_2 \in K, x_1 \neq x_2 \text{ e } x = (x_1+x_2)/2 \}$.

Definição 13. Seja B um subconjunto não vazio de Y, $y_o \in B$ é chamado um elemento maximal de B, denotado $y_0 \in \max(B;K)$ se: $\nexists\ y \in B$: $y_0 <_k y$.

Exemplos

16) Seja $Y = \mathbb{R}^2$ e $B = B_1 \cup B_2$, onde

$$B_1 = \{(\alpha_1, \alpha_2) \mid 0 \leq \alpha_1 \leq 1, \quad 0 \leq \alpha_2 \leq 1\}$$

$$B_2 = \{(\alpha_1, \alpha_2) \mid (\alpha_1 - 1)^2 + \alpha_2^2 \leq 1, \; \alpha_1 \geq 1, \quad \alpha_2 \geq 0\}$$

$$K = \{(\alpha_1, \alpha_2) \mid \alpha_1 \geq \alpha_2, \; \alpha_2 \geq 0\}$$

Temos que: $\max(B;K) = \{(\alpha_1,\alpha_2) \mid (\alpha_1 - 1)^2 + \alpha_2^2 = 1, \; \alpha_1 \geq 0\}$

17) Consideremos $Y = \mathbb{R}^2$ e $B = B_1 \cup B_2$, onde

$$B_1 = \{(\alpha_1,\alpha_2) \mid \alpha_1^2 + \alpha_2^2 \leq 1, \quad \alpha_1 \geq 0, \; \alpha_2 \geq 0\}$$

$$B_2 = \{(\alpha_1,\alpha_2) \mid \alpha_1 \geq 1, \quad \alpha_2 \geq 0\}$$

$$K = \{(\alpha_1,\alpha_2) \mid \alpha_1 \geq 0, \quad \alpha_2 \geq 0\}$$

Temos neste caso: $\max(B;K) = \emptyset$

Sejam $f: X \to Y$ uma função $K \subset Y$ um cone ponteado, consideremos o problema

$$\text{Maximizar} \quad \{f(x) \mid x \in A\} \qquad \text{P)}$$

onde $A \subset X$ e não vazio , i.e., determinar todos os $x_0 \in A$ para os quais

$$f(x_0) \in \max (f(A); K)$$

tais pontos x_0, são chamados pontos maximais de P).

Relação do problema anterior com ótimos de Pareto.

Recordemos primeiramente a definição do ótimo de Pareto.

Definição 14 . Seja X um espaço de Banach, e $g_i : X \to \mathbb{R}$ $i=1,\dots,n$ funções dadas. Dizemos que $x_0 \in X$ é um ótimo de Pareto se:

$$\nexists \; x \in X \text{ tal que } \forall \; i=1,\dots,m \quad g_i(x) > g_i(x_0). \qquad \Delta$$

Agora vejamos a relação entre ponto maximal é ótimos de Pareto.

Tomando $A = X$, $Y = \mathbb{R}^m$ e $K = \mathbb{R}^m_+$ e

$f = (g_1,\dots,g_m)$ no problema P) anterior, temos que $x_0 \in X$ é um ponto maximal se

$$f(x_0) \in \max(f(X);\ \mathbb{R}^m_+) \qquad (*)$$

ou equivalentemente, x_0 é um ponto maximal se não existe $y \in f(X)$ para o qual $f(x_0) <_{\mathbb{R}_+^m} y$

Já que $y \in f(X)$, temos que existe $x \in X$ tal que :

$y = (g_1(x),\ldots,g_m(x))$.

Assim, temos que (*) é equivalente a

$\nexists\ x \in X$, tal que $\forall\ i=1,\ldots,m \quad g_i(x) > g_i(x_0)$ isto é, x_0 é um ótimo Pareto.

O problema de achar os pontos de Pareto, recive, usualmente o nome de <u>programa multiobjetivo</u>.

Para mais detalhes de ótimos de Pareto, bem como sua aplicação aos problemas de Economia, recomendamos Moulin e Fogelman [61], Aubin [3].

Agora veremos um teorema devido a Corley [35], que relaciona o problema P) com um problema de ponto fixo para multiaplicações.

<u>Teorema 8</u>. (H.W. Corley) Seja $\Gamma : Y \rightrightarrows Y$ uma multiaplicação definida por

$$\Gamma(y) = \{f(x) / \ x \in A, \ f(x) \in K+y\}.$$

Então x_0 é um ponto maximal de P) se e somente se

$$\Gamma(f(x_0)) = \{f(x_0)\} \ . \qquad (*)$$

Prova. Suponhamos que x_0 é um ponto maximal. Temos então que, $f(x_0) \in \Gamma(f(x_0))$. Se existir $f(x_1) \neq f(x_0)$ tal que $f(x_1) \in \Gamma(f(x_0))$, temos que x_1 é factível para P) e satisfaz

$$0 \neq f(x_1) - f(x_0) \in K$$

o que contradiz as hipóteses

Inversamente, suponhamos que (*) se satisfaz. Então

$$0 \notin f(x_1) - f(x_0) \in K$$

não pode ser satisfeito para qualquer factível a P), de modo que x_0 é um ponto maximal. Δ

O seguinte exemplo mostra como o teorema pode ser aplicado diretamente.

Exemplo 18. Sejam α_i, β_i números reais. Consideremos o seguinte problema

$$\text{Maximizar } (3\alpha_1 + \alpha_2 \ , \ \alpha_1 - 2\,\alpha_2)$$

sujeito a $\qquad \alpha_1^2 + \alpha_2^2 \leq 1$

$$\alpha_1, \ \alpha_2 \geq 0$$

Colocando $y = (\beta_1, \beta_2)$ e $x = (\alpha_1, \alpha_2)$, buscamos um ponto fixo de

$$\Gamma(\beta_1,\beta_2) = \{(3\alpha_1+\alpha_2,\ \alpha_1-2\alpha_2) \mid \alpha_1^2+\alpha_2^2 \leq 1;\ \alpha_1,\alpha_2 \geq 0;$$

$$3\alpha_1+\alpha_2 \geq \beta_1\ ,\ \alpha_1-2\alpha_2 \geq \beta_2\}$$

no sentido do teorema 8. É evidente de considerações gráficas que os pontos (α_1, α_2) que são pontos fixos são aqueles que satisfazem

$$\alpha_1^2 + \alpha_2^2 = 1$$

$$3\alpha_1 + \alpha_2 = \beta_1,$$

onde β_1 é qualquer escalar produzindo não negatividade dos α_i.

O conjunto de pontos maximais é

$$\{(\alpha_1,\alpha_2) \mid \alpha_1=(6\beta_1+\sqrt{40-4\beta_1^2})/20,\ \alpha_2=2\beta_1-3\sqrt{40-4\beta_1^2}/20;\ 1\leq \beta_1 \leq \sqrt{10}\}.$$

Vamos estabelecer teoremas que garantem pontos fixos de multiaplicações do tipo: $\Gamma(y) = \{y\}$ existe. Estes resultados estão então relacionados ao problema P) via Teorema 8.

Recordemos primeiramente algumas definições.

<u>Definição 15</u>. Sejam $B \subset Y$ não vazio e $\Gamma: B \rightrightarrows B$ tal que $\Gamma(y) \subset B$ é fechado em Y para todo $y \in B$. Dizemos que B é <u>Γ-semicompacto</u> se toda cobertura aberta de complementos da forma.

$$\{\Gamma^c(y_\alpha) \mid y_\alpha \in B,\ \alpha \in \Lambda\}$$

onde Λ é um conjunto de índices arbitrários, tem uma sub - cobertura finita. Δ

<u>Nota</u>. Claramente se B é compacto, então B é Γ-semicompacto.

<u>Definição 16</u>. Seja $\Gamma: B \rightrightarrows B$ uma multiaplicação. Dizemos que Γ é:

a) <u>reflexiva sobre B</u> se : $y \in \Gamma(y) \quad \forall\ y \in B$.

b) <u>Antisimétrica sobre</u> B se: $y_1 = y_2$ quando $y_2 \in \Gamma(y_1)$ e $y_3 \in \Gamma(y_2)$.

c) <u>Transitiva sobre B</u> se: $y_3 \in \Gamma(y_1)$ quando $y_2 \in \Gamma(y_1)$ e $y_3 \in \Gamma(y_2)$. Δ

Temos o seguinte Teorema.

<u>Teorema 9</u>. Seja $\Gamma: B \rightrightarrows B$ uma multiaplicação tal que $\Gamma(y) \subset B$ é fechado em Y para todo $y \in B$, reflexiva, antisimétrica e transi-

tiva sobre B, e seja B Γ-semicompacto. Então existe $y \in B$ tal que $\Gamma(y) = \{y\}$.

<u>Prova</u>. Notemos que Γ induz uma ordem parcial sobre B definido por

$$y_1 \leq y_2 \quad \text{se} \quad y_2 \in \Gamma(y_1).$$

Mostraremos que B é inductivamente ordenado para logo aplicar o Lema de Zorn. Suponhamos que B não o seja. Então existe um conjunto totalmente ordenado $T = \{ y_\alpha \mid \alpha \in \Lambda \}$ em B o qual não tem cota superior em B. Assim $\bigcap_{\alpha\in\Lambda}\Gamma(y_\alpha) = \emptyset$: senão qualquer elemento desta interseção seria uma cota superior de T em B. Logo para qualquer $y \in B$ existe $y_\alpha \in T$ tal que $y \notin \Gamma(y_\alpha)$ de modo que $\{ \Gamma^c(y_\alpha) \mid \alpha \in \Lambda\}$ forma uma cobertura aberta de B.

A Γ-semicompacto de B agora implica que B tem uma subcobertura finita denotada por

$$\{ \Gamma^c(y_i) \mid i=1,\ldots,n \}$$

onde $y_1 \leq y_2 \leq \ldots \leq y_n$ sem perda de generalidade.

Mas então, das hipóteses de transitividade

$$\Gamma(y_i) \subset \Gamma(y_{i-1}) \quad , \; i=2,\ldots,n.$$

Segue-se que $B \subset \Gamma^c(y_n)$ e assim $\Gamma(y_n) \subset B^c$, o qual é uma contradição já que $y_n \in \Gamma(y_n)$.

Assim B é inductivamente ordenado e tem um elemento maximal pelo Lema de Zorn, com a propriedade

$$z \in \Gamma(y) \implies z = y.$$

Em outras palavras, existe $y \in B$ para o qual $\{y\} = \Gamma(y)$. Δ

Um corolário imediato é o seguinte resultado de existência para o problema P).

Corolário 2 . Se $f: X \to Y$ é contínua, A e não vazio e compacto , e K um cone convexo ponteado em Y, então existe um ponto maximal para P).

Prova. O conjunto $B = f(A)$ é compacto, já que f é contínua, e $\Gamma(y) = B \cap [K+y]$ satisfaz as hipóteses do teorema 9, logo existe $x_0 \in A$ para o qual (*) do teorema 8 de satisfaz, e a conclusão se obtem do teorema 8. Δ

Observação. O seguinte exemplo mostra que o teorema pode ser falso se suprimir-se a hipoteses B é compacto, por B fechado e limitado em um espaço de Banach infinito dimensional.

Seja $Y = c_0$, o espaço de Banach das sequências $y = \langle \alpha_n \rangle$ de números reais convergentes a zero com a norma $\|y\| = \sup_n |\alpha_n|$. O conjunto K de todas as sequências não negativas em c_0 é um cone convexo fechado ponteado , como pode verificar-se facilmente . Seja $B = \{y \mid \|y\| \leq 1\}$ o qual é um conjunto fechado e limitado e $\Gamma(y) = B \cap [K+y]$ para $y \in B$. Fixemos $y = \langle \alpha_n \rangle$ em B, notar que existe $k \in \mathbb{N}$ tal que $|\alpha_k| < 1$. Seja y_k a sequência a qual é zero exceto no k-ésimo lugar, onde vale $1-|\alpha_k|$.

Segue-se que $y+y_k \in B \cap [K+y]$. Já que y é arbitrário, $\{y\} \neq \Gamma(y) \forall y \in B$, isto é, $\max(B;K)=\emptyset$.

b) <u>Aplicação</u> : Convergência de Algoritmos.

A maioria dos métodos de otimização são interativos, i.e, que a partir de um ponto inicial dado x_0, se gera uma sequência potencialmente infinita $x_1, x_2, \ldots, x_k, \ldots$, que se espera que convirja ao ótimo buscado.

Um algoritmo de resolução é um procedimento que permite a partir do ponto inicial dado x_0, engendrar a sequência (x_k). Um algoritmo está então perfeitamente definido por uma função f dada, que a x_k associa $x_{k+1} = f(x_k)$. Isto permite confundir o algoritmo com a função associada a este. Logo o estudo da convergência de um algoritmo se remite ao estudo das propriedades da função f.

Consideremos por exemplo o seguinte problema

$$\text{Min } x - 10$$
$$\text{s.a. } x \geq 0.$$

onde $x \in \mathbb{R}$, o ponto ótimo é obviamente zero.

O melhor algoritmo possível para o problema anterior usa a função f(x)=0; o problema e assim, resolvido num só passo.

Um outro algoritmo, que é do mesmo tipo, usa a função f(x)=x/2. Começando no ponto x=2, a sequência gerada seria

$$\{2,1,1/2,1/4,1/8,\ldots\}.$$

Para esse algoritmo, devido a $x_{k+1} = x_k/2$, em cada interação, o próprio ponto é 1/2, a distância que resta até chegar a zero, e o algoritmo converge, porque o limite da sequência $\{x_k\}$ é o ponto ótimo

$$\lim_{k\to\infty} x_k = 0$$

Um modelo Geral de Algoritmos

O modelo anterior de algoritmo de resolução permite o estudo de algoritmos particulares e não está bem adaptado ao estudo de uma classe de algoritmos. Já que o objetivo geral é estudar globalmente o comportamento de uma classe de algoritmos ao invés de se estudar todos os algoritmos de uma classe, os quais diferem na "prática", sejam $A_1, \ldots A_p$ um certo número de algoritmos pertencentes a uma mesma família. Em princípio poder-se-ia estudar globalmente seus comportamentos, não obstante, é melhor considerar globalmente a família como uma aplicação que a um ponto x_k associa os p pontos:

$A_1(x_k)$, $A_2(x_k)$, ..., $A_p(x_k)$. ou, analogamente como a aplicação : $x_k \rightarrow \{A_1(x_k), \ldots, A_p(x_k)\}$

Isto nos leva em forma muito natural a um modelo geral no qual os algoritmos (ou classes de algoritmos) são representados pelas multiaplicações .

A ambiguidade aparenta nessa definição de um algoritmo não significa que o carater de um algoritmo seja aleatório. Em implementações reais, os algoritmos não são definidos de uma maneira ambígua.

De fato, um determinado programa de computador, que é executado duas vezes do mesmo ponto inicial, vai gerar duas cópias da mesma sequência. Em outras palavras, na prática, algoritmos são funções.

A utilidade da definição mais geral é que ela permite analizar, de uma só vez, a convergência de uma família infinita de algoritmos parecidos.

Então dois programas de computador, desenhados da mesma idéia básica, são talvez, diferentes em alguns pequenos detalhes e, talvez, não produziram resultados idênticos quando iniciados do mesmo ponto inicial.

Os dois programas podem, porém, ser considerados como implantações da mesma multiaplicação.

Noção de Convergência Global

Definição 17. Dizemos que um algoritmo descrito por uma multiaplicação é globalmente convergente, se qualquer que seja o ponto de partida x_0 escolhido, a sequência (x_k) gerada por $x_{k+1} \in A(x_k)$ ou uma subsequência converge a um ponto que satisfaz alguma condição necessaria de otimalidade. Δ

Comentários. A condição necessária que intervem na definição anterior é geralmente a estacionaridade (isto é, $\nabla f(x)=0$ no caso diferenciável, $0 \in \partial_c f(x)$ no caso convexo), no caso sem restrições, e a condição de Kuhn-Tucker no caso de otimização com restrições.

Naturalmente a convergência global não implica que a sequência convirja a um ótimo global. Impor que um algoritmo convirja globalmente a um ótimo global é muito forte.

Um Teorema de Convergência Global

Consideremos um problema de otimização sobre um espaço X e seja Ω o conjunto pontos que satisgazem uma certa condição de otimalidade necessária.

Suponhamos que para resolver um problema de otimização se utiliza um algoritmo representado por $A: X \rightrightarrows X$.

Mostraremos abaixo em que condições o algoritmo A tem a propriedade de convergência global (i.e., a convergência da sequência, ou a uma subsequência gerada por A, a um elemento de Ω, a partir de qualquer ponto x_0, chamado usualmente ponto de partida).

<u>Definição 18</u>. Diz-se que $z: X \to \mathbb{R}$ é uma função de descenso (relativo ao algoritmo A), se ela é contínua e possui as seguintes propriedades:

i) $x \notin \Omega \implies z(y) < z(x) \quad \forall\, y \in A(x)$

ii) $x \in \Omega \implies z(y) \leq z(x) \quad \forall\, y \in A(x)$ Δ

<u>Teorema 10</u> . (Zangwill, 1969). Seja um problema de otimização sobre X e Ω o conjunto de pontos que satisfazem uma certa condição necessária de otimalidade. Seja $A: X \rightrightarrows X$ um algoritmo e consideremos uma sequência $\{x_k\}$ engendrada por A, i.e., $x_{k+1} \in A(x_k)$ se

H_1) $\{x_k\}$ esta contido num compacto $C \subset X$,

H_2) existe uma função de descenso para Γ ,

H_3) A fechado sobre $X \setminus \Omega$ e para todo $x \in X \setminus \Omega$,$A(x) \neq \emptyset$.

Então o limite de qualquer subsequência convergente de $\{x_k\}$ pertence a Ω (Tais limites existem pela hipótese H_1)) .

<u>Prova</u>. Como $\{x_k\}$ esta contida num compacto, pode-se extrair uma subsequência convergente. Seja $\{x_m / m \in M\}$ uma tal subsequência e seja x seu limite ($x \in C$). Da continuidade de z (função de descenso para A), se tem $z(x_m) \to z(x)$ quando $m \to \infty$. Mostremos que se tem também $z(x_k) \to z(x)$ quando $k \to \infty$

$\forall\ \varepsilon > 0$ (dado), existe, m_ε tal que:

$$\forall m \geq m_\varepsilon \quad (m \in M): \quad z(x_m) - z(x) < \varepsilon .$$

Por outro lado para todo $k \geq m_\varepsilon$ ($k \in \mathbb{N}$) tem-se:

$$z(x_k) - z(x) = z(x_k) - x(x_{m_\varepsilon}) + z(x_{m_\varepsilon}) - z(x).$$

Já que z é monôtonica, $z(x_k) - z(x_{m_\varepsilon}) \leq 0$, logo

$$z(x_k) - z(x) \leq z(x_{m_\varepsilon}) - z(x) < \varepsilon \qquad \forall\ k \geq m_\varepsilon$$

Nos resta provar que $x \in \Omega$.

Para ela consideremos a sequência $\{x_{m+1} / m \in M\}$. Todos os elementos desta sequência estão contidos no compacto C , logo pode-se extrair uma subsequência desta sequência convergente.

$$\{x_{m+1} / m \in M' \subset M\} \quad \text{e:} \quad x_{m+1} \to x' \quad \text{quando} \quad m \to \infty$$

Assim, para $m \to \infty$ e $m \in M'$: $x_m \to x$, $x_{m+1} \to x'$ $x_{m+1} \in A(x_m)$.

Se $x \notin \Omega$, como A é fechada (por H_3)), se deduz que: $x' \in A(x)$.

Já que $z(x_k) \to z(x)$ quando $k \to \infty$ $(k \in \mathbb{N})$, em particular

$z(x_{m+1}) \to z(x') = z(x)$ quando $m \to \infty$, $m \in M'$.

Então, a desigualdade estrita: $z(y) < z(x)$ não se verifica para $y=x' \in A(x)$, o que contradiz o fato que z é uma função de descenso.

Logo, necessariamente $x \in \Omega$. Δ

Observações

1) Este resultado muito geral nos permite estabelecer a convergência global de muitos algoritmos.

2) Existem várias extensões deste resultado como por exemplo os dados por Huard (1979) [49], Adhigama, PolaK e Klessig (1979) [1].

3) Com relação à aplicação prática do teorema de Zamgwill, podemos notar:

a) A função de descenso z se toma em geral como a função a minimizar, logo uma hipótese que em geral se faz é a continuidade de tal função.

b) Dada a função f a minimizar, para satisfazer H_1), se faz em geral a hipótese que o conjunto

$$S_\alpha(f) = \{ x \in X \mid f(x) \leq \alpha = f(x_0)\}$$

Seja limitado, isto assegura que a sequência $\{x_k\}$ esta contida num fechado limitado que em $\mathbb{R}^n$ é um compacto.

A outra forma de satisfazer H_1) é que f seja inf-limitado, isto é, que $||x|| \to +\infty \implies f(x) \to +\infty$

c) Contrariamente às condições H_1) e H_2) que são geralmente satisfeitas, a condição H_3) é de importancia crítica. A não convergência de um algoritmo provem frequentemente do fato que não se satisfaz esta hipótese.

Exemplo 19. (Luemberger 1973): Seja $X = [0,1]$, definamos A por, $A(x) = [0,x[$ se $0 < x \leq 1$ e $\{0\}$ se $x=0$.

Para $\Omega = 0$, a função $z(x)=x$ é uma função de descenso : $y = z(y) < z(x)=x \quad \forall y \in A(x),\ x \neq 0$.

A sequência $x_0 = 1$, $x_{k+1} = x_k - (2^{k+2})^{-1}$ verifica para todo k: $x_{k+1} \in A(x_k)$, mas

$$x_{k+1} = 1 - \sum_{m=0}^{k} (2^{m+2})^{-1} = 1-(4)^{-1} \sum_{m=0}^{k} (2^m)^{-1} \to \frac{1}{2} \notin \Omega$$

A razão pela qual x_k não converge a Ω é que a multiaplicação A não é fechada em X/Ω.

Consideremos alguns algoritmos particulares.

Otimização Unidimencional exata

A iteração típica de numerosos algoritmos de otimização é construida no esquema seguinte:

(a) No ponto em andamento, x, se escolhe uma direção de deslocamento d;

(b) a partir de x, se busca o mínimo da função f na direção d; i.e., se determina $\bar{\alpha}$ tal que:

$$f(x+\bar{\alpha}d) = \min \{f(x+\alpha d) \mid \alpha \geq 0\}$$

O ponto $y = x+\bar{\alpha}d$ e então tomando como o ponto de inicio na interação seguinte.

Note-se que este esquema conduz de forma natural a se considerar o algoritmo como a composição de duas multiaplicações D e U:

1) A aplicação D que associa a x uma direção de deslocamento $d \in D(x)$.

2) A aplicação U que associa à dupla (x,d) um ponto $y \in U(x,d)$,

onde $U(x,d)$ é definido como o conjunto de pontos da forma $x+\alpha d$ onde f atinge seu mínimo:

$$U(x,d)=\{ y \mid y = x+\alpha d \mid \alpha \geq 0 \text{ e } f(y)=f(x+\bar{\alpha} d)\}.$$

O resultado seguinte é de grande utilidade para a demonstração da convergência de numerosos algoritmos de otimização.

<u>Teorema 11</u>. Se a função f é contínua sobre $\mathbb{R}^n$, então a multiaplicação $U(x,d)$ definida por:

$$U(x,d)= \{y \mid y=x+\alpha d (\alpha \geq 0); \quad f(y)= \min f(x+\alpha d \mid \alpha \geq 0\} .$$

é fechada no ponto (x,d) $\forall\, d \neq 0$.

<u>Prova</u>. Consideremos duas sequências $\{x_k\}$ e $\{d_k\}$ tais que $x_k \to x$ e $d_k \to d \neq 0$ quando $k \to \infty$. Suponhamos que $\forall\, k$, $y_k \in U(x_k, d_k)$ e que $y_k \to y$. Mostremos então que $y \in U(x,d)$.

Pela definição de U, para todo k, existe $\alpha_k \geq 0$, tal que: $y_k = x_k + \alpha_k d_k$. Assim:

$$\alpha_k = \frac{\| y_k - x_k \|}{\| d_k \|}$$

e por conseguinte quando $k \to \infty$:

$$\alpha_k \to \bar{\alpha} = \frac{\|y-x\|}{\|d\|}$$

Assim $y = x + \bar{\alpha}d$.

É suficiente agora demosntrar que o mínimo de f, no segmento de reta que inicia em x, na direção d, é atingido no ponto y.

Tem-se $\forall$ k e $\forall \alpha \geq 0$: $f(y_k) \leq f(x_k+\alpha d_k)$, o que prova que:

$$f(y) = \min \{f(x+\alpha d) \mid \alpha \geq 0\}$$

de onde $y \in U(x,d)$. Δ

A condição $d \neq 0$ é absolutamente necessária. Com efeito, se $d = 0$ a noção de otimização de f na direção d não tem sentido e a multiaplicação $U(x,d)$ não está definida. Esta condição não é restritiva, posto que, muitos dos algoritmos, quando a direção de deslocamento é $d = 0$, significa que o ponto x satisfaz uma condição necessária de otimalidade ((estacionaridade no caso sem restrições, Kuhn-Tucker no caso com restrições).

Otimização Unidimensional Aproximada

Como a maioria dos métodos para minimizar a função

$g(\alpha) = f(x+\alpha d)$ ($\alpha \geq 0$) é de natureza iterativa (dicotomia, fibonacci, etc), o mínimo exato da função g(.) não pode ser obtido em forma exata em um número finito de iterações. Na pratica, sempre se acha um mínimo aproximado. É então fundamental verificar se a propriedade de fechamento estabelecida anteriormente para o caso idealizado de uma otimização unidimensional exata, se conserva.

Consideremos por exemplo, os seguintes tipos de aproximação:

<u>Tipo 1</u> Corresponde ao caso onde a busca é interrompida quando o erro relativo sobre $\bar{\alpha}$, o mínimo de $g(\alpha)$, é inferior a uma porcentagem δ fixa, i.e.,

$$\frac{|\alpha - \bar{\alpha}|}{\bar{\alpha}} \leq \delta$$

(por exemplo, nos casos de dicotonia, Fibonacci, tem-se).

<u>Tipo 2</u> Corresponde ao caso onde a busca é interrompida quando se obtem uma aproximação com tolerância ε do ótimo procurado, i.e., quando

$$|g(\alpha_k) - g(\bar{\alpha})| \leq \varepsilon$$

A aproximação do tipo 1 conduz ao estudo da multiaplicação

$$U_\delta(x,d) = \{y \mid y = x+\alpha d(\alpha \geq 0),\ |\alpha - \bar{\alpha}| \leq \delta\bar{\alpha}\}$$

com $\bar{\alpha}$ verificando $g(\bar{\alpha}) = \min\{g(\alpha) \mid \alpha \geq 0\}$, enquanto que, o tipo 2 leva a

$$U_\varepsilon(x,d) = \{y \mid y=x+\alpha d \quad (\alpha \geq 0);\ g(\alpha) \leq g(\bar{\alpha}) + \varepsilon\}$$

onde $g(\bar{\alpha}) = \min\{g(\alpha) \mid \alpha \geq 0\}$.

Podemos enunciar:

<u>Teorema 12</u>. Se a função f é contínua sobre $\mathbb{R}^n$, e inf-limitada (i.e., $f(x) \to +\infty$ para $||x|| \to +\infty$), então $\dot{U}_\delta(.,.)$ é fechada em todo ponto (x,d) tal que $d \neq 0$.

<u>Prova</u>. Podemos considerar $U_\delta(.,.)$ como a composição das seguintes multiaplicações:

a) a multiaplicação que a $(x,d) \rightrightarrows U(x,d) = \{y \mid y = x+\alpha d$, $(\alpha \geq 0)$ e $f(y) = \min\{f(x+\alpha d) \mid \alpha \geq 0\}\}$. (a qual é em geral um intervalo $[\alpha_{min}, \alpha_{max}]$);

b) A multiaplicação V_δ que a um intervalo $[\alpha_{min}, \alpha_{max}]$ com $\alpha_{min} \geq 0$ associa o intervalo "extendido".

$$V_\delta([\alpha_{min}, \alpha_{max}]) = [\alpha_{min}(1-\delta), \alpha_{max}(1+\delta)] .$$

Pelo teorema 11, a multiaplicação $U(.,.)$ é fechada e é elementar se constatar que V_δ é fechada.

Se o conjunto $U(x,d)$ é limitado (este é o caso por exemplo se $f(x) \to +\infty$ para $||x|| \to +\infty$) então se tem que U_δ é fechada em virtude da proposição 4. Δ

É fácil mostrar que se f é contínua, a multiaplicação U_ε é fechada.

Consideremos agora um algoritmo mais complexo.

<u>Método da Máxima descida</u> (Steepest descent) (Cauchy 1847 , Curry 1944)

Este método é para resolver problemas do tipo

$$\text{Min } \{f(x) \mid x \in \mathbb{R}^n\}, \quad \text{com } f \text{ diferenciável.}$$

O algoritmo é o seguinte:

(a) Escolher um ponto inicial x_0; $k=0$

(b) Na iteração k: $d_k = -\nabla f(x_k)$

Pesquisar λ_k tal que:

$$f(x_k+\lambda_k d_k) = \min \{f(x_k+\lambda d_k) \mid \lambda \geq 0\}$$

fazer $x_{k+1} = x_k+\lambda_k d_k$

(c) Testar o erro: se verifica: FIM, se não fazer $k \leftarrow k+1$ e retornar a (b).

Como a convergência não é, em geral, finita, deve-se definir um teste de parada. Alguns critérios são os seguintes:

critério 1: $\text{Max}\ \{ |\frac{\partial f}{\partial x_i}| < \varepsilon / 1 \leq i \leq n \}$ ($\varepsilon > 0$ dado)

Critério 2: $\| \nabla f \|^2 = \sum_{i=1}^{n} (\frac{\partial f}{\partial x_i})^2 < \varepsilon$ ($\varepsilon > 0$ dado).

critério 3: $|f(x_{k+1}) - f(x_k)| < \eta$ ($\eta > 0$ dado).

Temos o seguinte teorema de convergência global para o método da máxima descida.

Teorema 13. Se a função f é continuamente derivável e inf-limitada, então, para todo ponto de partida x_0 (com $\nabla f(x_0) \neq 0$), o método da máxima descida (com otimização unidimensional exata ou aproximada) converge a um ponto estacionário de f (i.e., x^*: $\nabla f(x^*)=0$)

Prova. O algoritmo da máxima descida pode ser representado pela composição de duas multiaplicações D e U:

- a aplicação D, que a x_k associa a direção de deslocamento $d_k = - \nabla f(x_k)$.

\- a multiaplicação U representando o processo de otimização unidimensional e que a (x_k, d_k) associa: $x_{k+1} \in U(x_k, d_k)$.

Se a função f e continuamente diferenciável, a aplicação D é continua (no sentido usual, i.e., de funções); por outro lado, como f é continua, a multiaplicação U é fechada (teorema 11 para a otimização unidimensional exata; teorema 12 para a otimização unidimensional aproximada).

Assim, pelo teorema , U o D é fechada. Temos que da hipótese de crescimento no infinito que todos os x_k estão contidos num fechado limitado.

Observemos que o fato de $\nabla f(x_k) \neq 0$, implica que $f(x_{k+1}) < f(x_k)$, então, tomando f como a função de descenso, e o conjunto Ω como o conjunto de pontos estacionarios de f($\Omega = \{\bar{x} \mid \nabla f(\bar{x}) = 0\}$), se pode aplicar o teorema de Zangwill , de onde a convergência global do método a um ponto estacionário. Δ

Observações

1) Recordemos que a convergência global não significa que se obtem necessariamente um ótimo global de f. Se f é continuamente diferenciável, a unica coisa que podemos afirmar é que se obtem um ponto estacionário $\bar{x}$ de f. Se f é duas vezes continuamente diferenciável em $\bar{x}$, e $\nabla^2 f(\bar{x})$ é definido positivo, então $\bar{x}$ é um mínimo local de f.

Somente em casos bem particulares (f convexa diferenciável, por exemplo) se pode garantir a obtenção de um mínimo global de f.

2) Também existem teoremas para problemas com restrições, por exemplo, pode consultar as seguintes referências Contesse [82] , [83], Huard [49], Luenberger [57] , Minoux [59] e Fritzsche [43] .

Capítulo II.

Análise Não Diferenciável e Aplicações.

Introdução

É devida a Pierre de Fermat (1601-1655) o descobrimento crucial da primeira idéia do cálculo diferencial. A qual está associada ao problema da determinação dos extremos,de funçoes , descrita sem demonstração, em um pequeno tratado "Methodus ad disquerendam Maximam e Minimam" escrito em 1637. Entretanto seus trabalhos em teoria dos números eclipsaram suas contribuições que este homen notavel fez nos outros domínios das matemáticas.

Mas Fermat não conhecia o conceito de derivada o qual foi descoberto mais tarde por Newton (1671) para estudar problemas da Física e por Leibniz na sua publicação sobre o calculo diferencial, em 1684, titulada "Nova Methodus por Maximis et Minimus". A analogia entre o método de Fermat (limitado as funções algébricas) e o de Leibniz é surpreendente. Como é conhecido, esta regra consiste em achar os extremos de uma função f através das soluções da equação $f'(x)=0$.

Esta regra com o passar do tempo foi-se adaptando aos novos problemas que surgiram na teoria de otimização, do cálculo das variacões, e da teoria do controle ótimo. Matemáticos importantes tais como Euler, Lagrange, Jacobi, Poincaré , Hilbert e Frechet estiveram ligados ao desenvolvimento das novas técnicas. A regra de Fermat e Leibniz mantinha-se válida; se a função atinge seu mínimo em x, o gradiente de f deve-se anular nesse ponto.

Na teoria da otimização, teoria de jogos, economia, etc. aparecem frequentemente funções que não são diferenciáveis no sentido usual, posto que as operações de supremo e infímo não conservam as propriedades de diferenciabilidade usuais.
Por exemplo a função $x \to |x|$, é obtida como a envolvente superior das funções $x \to x$ e $x \to -x$, as quais são diferenciáveis, mas $|x|$ não é diferenciável em $x=0$.

Portanto, se desejamos conservar a regra de Fermat, devemos modificar o conceito de gradiente e a generalização deve ser adequada. O exame da função $x \to |x|$ pode-nos servir de guia: como $x \to |x|$ é a envolvente superior das funções $x \to x$ e $x \to -x$ onde as derivadas em 0 são +1 e -1, respectivamente, por que não tomar a envolvente convexa $[-1,+1]$ destas derivadas como candidato?. Evidentemente, é necessário dominar a resistencia provocada pelo carater multívoco desta solução, que não é devido a um fato de familiaridade (e a nossa tendência ao conservatismo). Mas, para nos convencermos do interesse desta ousadia, é necessário ressaltar sobre este exemplo que a regra de Fermat ainda permanece válida, posto que:

$$0 \text{ pertence a } [-1,+1]$$

No quadro de outras teorias, nas equações diferenciais parciais, por exemplo, outras generalizações são certas, como o é a teoria das distribuições devida a L. Schwartz.

Veremos neste capítulo que o conceito que se adapta muito bem aos problemas de otimização e termos afins, é o gradiente generalizado, denotado $\partial f(x)$, devido a F.H. Clarke (1973), e mostramos que a regra de Fermat permanece válida se $\bar{x}$ minimiza f (suposta localmente lipschitziana) então $0 \in \partial f(\bar{x})$

Tal conceito coincide com a gâteaux-derivada se f é de classe C^1, e com a Subdiferencial da análise convexa quando f é convexa e continua. Esta última noção, ocupou muito interesse dos matemáticos na década dos 60 e 70, e foi introduzida por J.J.Moreau e R.T.Rockafellar independente.

Desenvolveremos também o cálculo Generalizado (Regras da cadeia, teorema do valor médio, etc) e apresentaremos sua aplicação aos problemas de otimização não diferenciáveis.

No capítulo seguinte aplicaremos tais resultados a problemas do cálculo das variações e de controle ótimo.

Primeiramente recordemos a definição da função suporte, e o teorema de Hormander (1954), que nos serviram muito na teoria que apresentaremos.

Função Suporte

A seguir X denotará um espaço de Banach, X* seu dual topológico e $\langle .,. \rangle$ o produto dualidade canônico entre X e X*.

Definição. Seja A um subconjunto arbitrário de um espaço de Banach X. Denominaremos função suporte de A à função ,

$\sigma_A: X^* \to \mathbb{R} \cup \{+\infty\}$ definida por

$$\sigma_A(x^*) = \sup \{ \langle x, x^* \rangle \mid x \in A \} . \qquad \Delta$$

Às vezes denotaremos à função suporte de A por $\sigma(A;x^*)$.

Notar que $\sigma_A = \sigma_{\overline{co}(A)}$ onde $\overline{co}$ (A) denota a envolvente convexa fechada de A.

Se $D \subset X^*$, sua função suporte esta definida sobre X^{**}. Se fizermos a identificação canônica de X como subconjunto de X^{**}, temos para $x \in X$,

$$\sigma_D(x) = \sup \{ \langle x^*, x \rangle \mid x^* \in D \} .$$

Os seguintes fatos são conhecidos.

<u>Teorema</u> (L.Hörmander 1954, [48]). Sejam C,K subconjuntos não vazios convexos fechados de X, e sejam D,P subconjuntos não vazios convexos w*-fechados de X*. Então

(a) $C \subset K \iff \sigma(C;x^*) \leq \sigma(K;x^*) \quad \forall\, x^* \in X^*$

(b) $D \subset P \iff \sigma(D;x) \leq \sigma(P;x) \quad \forall\, x \in X$

(c) D é w*-compacto $\iff$ $\sigma(D;.)$ é com valores finitos sobre X.

(d) Dada uma função $\psi: X \to \mathbb{R} \cup \{+\infty\}$, ψ é positivamente homogênea, subaditiva, semicontinua inferior (forte o fraca) e não identicamente igual a $+\infty$ se e somente se existe um sub conjunto não vazio, convexo, w^*-fechado D de X^* tal que $\psi = \sigma_D$. (e tal D é único). Δ

2.1 Recordação do Cálculo Diferencial Clássico.

Nesta seção faremos uma recordação das definições usuais de derivadas, de funções $f: X \to \mathbb{R} \cup \{+\infty\}$. Boas referências para estes tópicos são Cartan [19] , Avez [10] , Flett [42].

Recordemos que o limite seguinte

$$f'(x;d) = \lim_{t \to 0+} \frac{f(x+td) - f(x)}{t}$$

quando existe, é denominado derivada direcional de f em x na direção d e que f é direcionalmente derivável se f'(x;d) existe para todo d.

Dizemos que f é derivável no sentido de Gâteaux em x se f é direcionalmente derivavel e se $d \to f'(x;d) = \langle \nabla f(x), d \rangle$ é linear e continua. Denominamos $\nabla f(x) \in X^*$ o gradiente de f em x.

Dizemos que f é continuamente derivável (ou de classe C^1) se para todo $d \in X$, a função $y \to \langle \nabla f(y), d \rangle$ é continua em x.

Também temos a dizer que f é devivável no sentido de Fréchet em x se

$$\lim_{d \to 0} \left| \frac{f(x+d) - f(x) - \langle \nabla f(x), d \rangle}{\|d\|} \right| = 0$$

e que f é <u>estritamente derivável no sentido de Bourbaki</u> se

$$\lim_{\substack{y \to x \\ d \to 0}} \left| \frac{f(y+d) - f(y) - \langle \nabla f(x), d \rangle}{\|d\|} \right| = 0$$

Ligações entre estas noções, são dadas no teorema seguinte:

<u>Teorema 1</u> C^1 ⇒ estritamente derivável no sentido de Bourbaki ⇒ derivável no sentido de Fréchet ⇒ Gâteaux-derivável. Δ

Recordemos ademais a <u>derivada de Dini direcional</u> de f em x na direção d

$$Df(x)(d) = \limsup_{\theta \to 0} \frac{f(x+\theta d) - f(x)}{\theta}$$

2.2 Derivada de Clarke.

Agora vamos a introduzir uma definição de derivada direcional generalizada devida a F.H. Clarke [24].

<u>Definição 1.</u> Seja $f: X \to \mathbb{R} \cup \{+\infty\}$ cujo domínio é não vazio.

Definimos a <u>derivada direcional de Clarke</u> de f no ponto x na direção d por:

$$f^0(x;d) = \lim_{\substack{y\to x \\ t\to 0^+}} \sup \frac{f(y+td)-f(y)}{t}$$

quando o lim sup é finito. Δ

Dizemos que f é <u>derivável no sentido de Clarke</u> em x se para todo $d \in X$, o limite $f^0(x;d)$ é finito.

<u>Lema 1</u>. As funções $f'(x;.)$, $f^0(x;.)$, $Df(x)(.)$ são positivamente homogêneas, além disso quando estes limites existem , se obtem as desigualdades.

$$f'(x;v) = Df(x)(d) \leq f^0(x;d).$$

<u>Prova</u>. Direta das definições. Δ

Uma propriedade interessante da derivada de Clarke é a seguinte:

<u>Proposição 1</u>. (Cominetti e Correa [31]). Seja $f: X \to \mathbb{R}$ Continua . Então

$$f^0(x;d) = \lim_{\substack{y\to x \\ t\to 0}} \sup \frac{f(y+td)-f(y)}{t}$$

para cada $x,d \in X$.

<u>Prova</u>. Claramente é suficiente mostrar que:

$$q(x,d) = \limsup_{\substack{y \to x \\ t \to 0}} \frac{f(y+td)-f(y)}{t} \leq f^0(x;d)$$

Sejam $\varepsilon > 0$ e V uma vizinhança de x arbitrarios. Também escolhemos $\delta > 0$ e uma vizinhança W de x tal que $\delta > \varepsilon$, $W \subseteq V$ e $W + [0, \delta] d \subseteq V$.

Para cada $y \in W$ e $|t| < \delta$, temos

$$t > 0 \Rightarrow y \in V,\ 0 < t < \varepsilon \Rightarrow \frac{f(y+td)-f(y)}{t} \leq \sup_{\substack{Z \in V \\ 0 < S < \varepsilon}} \frac{f(Z+Sd)-f(Z)}{S}$$

$$t < 0 \Rightarrow y+td \in V,\quad 0 < -t \leq \varepsilon \Rightarrow \frac{f(y+td)-f(y)}{t} = \frac{f(y+td-td)-f(y+td)}{-t}$$

$$\leq \sup_{\substack{Z \in V \\ 0 < S < \varepsilon}} \frac{f(Z+Sd)-f(Z)}{S}$$

Logo,

$$q(x,d) \leq \sup_{\substack{y \in W \\ |t| < \delta}} \frac{f(y+td)-f(y)}{t} \leq \sup_{\substack{Z \in V \\ 0 < S < \varepsilon}} \frac{f(Z+Sd)-f(Z)}{S}$$

de modo que o resultado segue-se tomando o ínfimo quando V percorre a familia de vizinhanças de x e $\varepsilon > 0$. Δ

2.3 Funções que são derivaveis no sentido de Clarke.

Agora vamos exibir classes de funções que são derivaveis no sentido de Clarke. Notemos que uma função f derivavel no sentido de Gâteaux, não é necessariamente derivável no sentido de Clarke, como mostra o seguinte exemplo:

Exemplo. Consideremos a função

$$f(x) = \begin{cases} x^2 \operatorname{sen} \frac{1}{x}, & x \neq 0 \\ 0 & x = 0 \end{cases}$$

temos que

$$f'(x) = \begin{cases} 2x \operatorname{sen} \frac{1}{x} - \cos \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Mas $\quad f^0(0;\, d) = |d|$.

Entretanto, se obtem o resultado seguinte.

Teorema 2. Suponhamos que f seja continuamente derivável em x Então f é derivável no sentido de Clarke e

$$< \nabla f(x),\, d > \; = \; f^0(x;\, d)$$

Prova. Como $f \in C^1$ no ponto x, se tem que para todo $\varepsilon > 0$, existe $\eta > 0$ tal que

$$|< \nabla f(z), d > - < \nabla f(x), d>| \leq \varepsilon \quad \text{quando} \quad \| z-x \| \leq \eta$$

se $\| y \| \leq \eta/2$ e se $0 < t \leq \eta/2 \, \|d\|$, pondo $g(t) = f(y+td)$, temos que g é derivável e

$$g'(t) = \lim_{\theta \to 0} \frac{f(y+td+\theta d) - f(y+td)}{\theta} = < \nabla f(y+td), d >$$

de onde se $0 \leq \eta/2 \, \|d\|$, se obtem:

$$\frac{f(y+\theta d) - f(y)}{\theta} - < \nabla f(x), d > = \frac{g(\theta) - g(0)}{\theta} - < \nabla f(x), d >$$

$$= \frac{1}{\theta} \int_0^\theta (<\nabla f(y+td), d > - < \nabla f(x), d >) \, dt$$

e por conseguinte, posto que $\| y+td -x \| \leq \eta$,

$$\left| \frac{f(y+\theta d) - f(y)}{\theta} - < \nabla f(x), d > \right| \leq \varepsilon$$

quando $\| y-x \| \leq \eta/2$ e $\theta \leq \eta/2 \, \|d\|$, o qual implica que se $\alpha \leq \eta/2$ e $\beta \leq \eta/2 \|d\|$

$$\sup_{\|y-x\| \leq \alpha} \ \sup_{\theta \leq \beta} \ \frac{f(y+\theta d) - f(y)}{\theta} \leq \langle \nabla f(x), d \rangle + \varepsilon \ .$$

Tomando o ínfimo com respeito a α e β , temos

$$f^0(x;d) \leq \langle \nabla f(x), d \rangle + \varepsilon$$

o que termina a prova. Δ

<u>Definição 2</u>. Diremos que uma função $f: X \to \mathbb{R} \cup \{+\infty\}$ de dominio não vazio é <u>localmente lipschitziana em</u> $x_0 \in \text{int}\ (\text{Dom}(f))$ se, existe $\lambda > 0$ e uma vizinhança do ponto x_0 tais que:

$$|f(y) - f(z)| \leq \lambda \ \| y-z \| \qquad \forall \ y, z \in V$$

onde V é uma vizinhança do ponto x_0 e λ depende de x_0. Δ

Dizemos que f é localmente lipschitziana sobre o interior de seu domínio, se ela é localmente lipschitziana em cada ponto $x_0 \in \text{int}\ (\text{Dom}\ (f))$.

Temos o seguinte teorema.

<u>Teorema 3</u>. Toda função localmente lipschitziana $f: X \to \mathbb{R} \cup \{+\infty\}$ é derivável no sentido de Clarke sobre o interior de seu domínio.

Além disso:

a) $\forall\ x \in \text{int}\ (\text{Dom}(f))$, a função $d \to f^0(x;d)$ é positivamente homogênea, convexa e continua.

b) e a função: $(x,d) \in \text{int}\ (\text{Dom}\ (f)) \times X \to f^0(x;d)$ é s.c.s.

<u>Prova</u>. Seja $x \in \text{int}\ (\text{Dom}(f))$. Posto que f é localmente lipschi ziana, existe $\eta > 0$ e $\lambda > 0$ tais que

$$|\ f(y) - f(z)\ | \leq \lambda \|y-z\| \qquad \forall\ y,z \in B(x;\eta\ ).$$

De onde para todo $\alpha \leq \eta/_2$ e $\beta \leq \eta/_2\ \|d\|$

$$-\lambda\ \|\ d\ \| \leq \frac{f(y+\theta d) - f(y)}{\theta} \leq \lambda\ \|d\|$$

quando $y \in B\ (x;\ \alpha\ )$ e $\theta \leq \beta$. De onde

$$-\lambda\ \|\ d\ \| \leq f^0(x;d) = \inf_{\alpha,\beta>0}\ \sup_{\substack{\|y-x\| \leq \alpha \\ 0 < \theta \leq \beta}} \frac{f(y+\theta d)-f(y)}{\theta} \leq \lambda\ \|\ d\ \|$$

é dizer que f é derivável no sentido de Clarke; em particular se obtem a desigualdade:

$$|\ f^0(x;d)\ | \leq \lambda\ \|d\|$$

Já sabemos que $d \to f^0(x;d)$ é positivamente homogênea. Mostraremos que esta função é convexa

$$\frac{f(y+\theta(td+(1-t)w))-f(y)}{\theta} = \frac{(1-t)[f(z+\alpha w)-f(z)]}{\alpha} + \frac{t\,[f(y+\beta d)-f(y)]}{\beta}$$

onde $z = y+\theta td$ converge a x, $\alpha=(1-t)\theta$ e $\beta=t\theta$ convergem a zero, se obtem tomando os limites superiores de os dois membros que:

$$f^0(x;td+(1-t)w) \leq tf^0(x;d)+(1-t)f^0(x;w).$$

Nos resta provar que $\{x,d\} \to f^0(x;d)$ é s.c.s.. Pela definição de $f^0(x,d)$, podemos associar a todo $\varepsilon > 0$ um número α_0 tal que

$$\sup_{\substack{\|z-x\| \leq 2\alpha_0 \\ t \leq \alpha_0}} \frac{f(z+td)-f(z)}{t} \leq f^0(x;d) + \frac{\varepsilon}{2}$$

Se $\|z-y\| \leq \alpha_0$ e $\|y-x\| \leq \alpha_0$, se tem, posto que f é localmente lipschitziana

$$\frac{f(z+tw)-f(z)}{t} \leq \frac{f(z+td)-f(z)}{t} + \lambda\,\|d-w\|$$

Por conseguinte, se $y \in B(x;\alpha_0)$, se $\alpha \leq \alpha_0$ e se $\beta \leq \beta_0$

$$\sup_{\substack{\|z-y\| \leq \alpha \\ t \leq \beta}} \frac{f(z+tw)-f(z)}{t} \leq \sup_{\substack{\|z-x\| \leq 2\alpha_0 \\ t \leq \alpha_0}} \frac{f(z+td)-f(z)}{t} + \lambda\,\|d-w\|$$

$$\leq f^0(x;d) + \frac{\varepsilon}{2} + \lambda \|d-w\|$$

$$\leq f^0(x;d) + \varepsilon$$

quando $\|d-w\| \leq \frac{\varepsilon}{2\lambda}$

Fazendo tender α e β a 0, deduzimos que:

$f^0(y;w) \leq f^0(x;v) + \varepsilon$ quando $\|y-x\| \leq \alpha_0$ e $\|v-w\| \leq \frac{\varepsilon}{2\lambda}$ é dizer $f^0(x;v)$ é s.c.s. em $\{x,v\}$. Δ

Ademais, temos que:

<u>Proposição 2</u>. Seja $f: X \to \mathbb{R} \cup \{+\infty\}$ Uma função convexa não identicamente igual a $+\infty$. As condições seguintes são equivalentes:

a) f é limitada superiormente sobre um conjunto aberto (necessariamente contido no domínio de f).

b) f é localmente lipschitziana sobre o int (Dom(f)).

<u>Prova</u>. Veja I. Ekeland e R. Teman [39] página 12. Δ

<u>Lema 2</u>. Suponhamos que $f: X \to \mathbb{R} \cup \{+\infty\}$ seja uma função convexa não identicamente igual a $+\infty$. Tomemos $x_0 \in \text{Dom}(f)$ e $v \in X$. Então

$$f'(x_0;v) = \lim_{t \to 0+} \frac{f(x_0+tv) - f(x_0)}{t}$$

existe em $\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$ e verifica :

$$f(x_0) - f(x_0-v) \leq f'(x_0;v) \leq f(x_0+v) - f(x_0).$$

<u>Prova</u>. A função $t \to \dfrac{f(x_0+tv)- f(x_0)}{t}$ é crecente.

Em efeito, se $t_1 \leq t_2$, então

$$f(x_0 + t_1 v) - f(x_0) = f\left(\frac{t_1}{t_2} (x_0 + t_2 v) + (1- \frac{t_1}{t_2})x_0\right)- f(x_0),$$

e fazendo uso da convexidade de f, temos que:

$$\frac{f(x_0+ t_1 v)- f(x_0)}{t_1} \leq \frac{f(x_0 + t_2 v) - f(x_0)}{t_2}$$

já que ${}^{t_1}/t_2 < 1$, de onde estes quocientes diferenciais tem um limíte em $\overline{\mathbb{R}}$ quando $t \to 0^+$:

$$f'(x_0;v) = \inf_{t>0} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

Tornando $t = 1$, em a igualdade anterior, se tem:

$$f'(x_0; v) \leq f(x_0 + v) - f(x_0).$$

Escrevendo $x_0 = \frac{1}{1+t}(x_0 + tv) + \frac{t}{1+t}(x_0 - v)$ e fazendo uso novamente da convexidade de f, temos que

$$f(x_0) \leq \frac{1}{1+t} \quad f(x_0 + tv) + \frac{t}{1+t} \quad f(x_0 - v).$$

então: $\forall\ t > 0$, $f(x_0) - f(x_0 - v) \leq \dfrac{f(x_0 + tv) - f(x_0)}{t}$ o

qual implica que $f(x_0) - f(x_0 - v) \leq f'(x_0;\ v)$.

Logo se conclui a prova. Δ

<u>Corolário 1</u>. Suponhamos que uma função convexa $f: X \to \mathbb{R} \cup \{+\infty\}$ seja contínua em um ponto x do interior de seu domínio. Então

a) f é derivável no sentido de Clarke e

b) $\forall\ v \in X$, $f^0(x;v) = f'(x;v)$.

<u>Prova</u>. a) A derivabilidade resulta da proposição 2 junto com o teorema 3.

b) De acordo ao Lema 2, temos que $f'(x;\ v)$ existe e é finita pois $f(x) < +\infty$. De acordo ao Lema 1, nos resta provar que:

$$f^0(x;v) \leq f'(x;v) \quad \forall\ v \in X.$$

Posto que a função $(t,y) \to \dfrac{f(y+tu) - f(y)}{t}$ é continua em $(\beta,\ x)$, existe $\alpha > 0$ tal que:

$$\frac{f(y+tu) - f(y)}{t} \leq \frac{f\ (x+\beta u) - f(x)}{\beta} + \varepsilon$$

quando $|t-\beta| \leq \alpha$ e $\|y-x\| \leq \alpha$. Isto implica em particular, de acordo à monotonia da função $t \to \frac{f(y+tu)-f(y)}{t}$ que :

$$\sup_{\|y-x\| \leq \alpha} \ \sup_{0 \leq t \leq \beta+\alpha} \frac{f(y+tu)-f(y)}{t} \leq \frac{f(x+\beta u)-f(x)}{\beta} + \varepsilon$$

Tomando o ínfimo com respeito a β e α , se obtem:

$$f^0(x;u) \leq f'(x;u) + \varepsilon$$

é suficiente agora fazer ε tender a zero. Δ

Pode-se provar a seguinte caracterização de lipschitzianidade local.

Proposição 3. Seja $f: X \to \mathbb{R}\{+\infty\}$ f é localmente lipschitziana em $x_0 \in \operatorname{int}(\operatorname{Dom}(f))$ se e somente se existem vizinhanças V de x_0 e U de 0 tal que $\bigcup_{x\in V} f^0(x;U)$ é um conjunto limitado em $\mathbb{R}$.

Prova. Veja G. Lebourg [55], página 126. Δ

Fazendo uso da caracterização anterior, temos o seguinte resultado.

Proposição 4. Seja $f: X \to \mathbb{R} \cup \{+\infty\}$ localmente lipschitziana em $x_0 \in \operatorname{Int}(\operatorname{Dom}(f))$, então existe uma vizinhança V de x_0 tal que $\{f^0(x;.) \mid x \in V\}$ é uma familia de funcionais sublineares uniformemente equicontinuos.

Prova. Por hipoteses existe uma vizinhança V de x_0 e uma vizinhança U de 0 em X tal que: $\bigcup_{x\in V} f^0(x;U)$ ē limitado em $\mathbb{R}$.

Seja $M = \sup \bigcup_{x\in V} f^0(x;U)$; para todo $\varepsilon > 0$ e $x \in V$, temos que

$$h-k \in \varepsilon M^{-1}U \Rightarrow f^0(x;h) - f^0(x;k) \leq f^0(x;h-k).$$

Assim

$$f^0(x;h) - f^0(x;k) \leq \varepsilon M^{-1} \sup_{x\in A} \bigcup f^0(x;U) = \varepsilon$$

Similarmente; $f^0(x,k) - f^0(x;h) \leq \varepsilon$; logo, temos que $\forall \varepsilon > 0$, existe uma vizinhança $\theta = \varepsilon M^{-1}U$ de 0 em X tal que, para $x \in V$,

$$h-k \in V \Rightarrow | f^0(x;h) - f^0(x;k) | \leq \varepsilon$$

o qual nos da o resultado. Δ

2.4 Propriedades Bāsicas da Derivada de Clarke

Nesta seção desenvolveremos as propriedades de cālculo da derivada de Clarke.

Proposição 5. Sejam $f,g: X \to \mathbb{R} \cup \{+\infty\}$ localmente lipschitzianas e $x_0 \in \text{Int}\,(\text{Dom}(f)) \cap \text{Int}\,(\text{Dom}(g))$. Então

a) $(sf)^0(x;v) = sf^0(x;v) \qquad \forall\ s \in \mathbb{R}$

b) $(f+g)^0(x;v) \leq f^0(x;v)+g^0(x;v)$

Prova.

a) Notar que sf também é localmente lipschitziana em x_0. Quando $s \geq 0$, claramente $(sf)^0 = sf^0$. Assim é suficiente provar que $(-f)^0 = -f^0$. Notemos que

$$\frac{-f(y+tv)-(-f(y))}{t} = \frac{f(z+t(-v))-f(z)}{t}$$

onde $z=y+tv$ converge a x_0 quando $y \to x_0$ e $t \to 0^+$ $(t>0)$. Logo, tomando o lim sup quando y e z convergem a x_0 e $t \to 0^+$, o membro da direita converge a $f^0(x-v)$ e o da esquerda a $(-f^0)(x;v)$.

b) Direta da definição de lim sup. Δ

Corolário 2. Sob as hipóteses da proposição anterior se tem

$$(\alpha f+\beta g)^0(x;v) \leq \alpha f^0(x;v)+\beta g^0(x;v) \qquad \forall\ \alpha,\beta \in \mathbb{R}. \qquad \Delta$$

Para termos uma igualdade no corolário anterior, temos que pedir certa regularidade das funções envolvidas. Para isto, introduzimos a seguinte definição.

Definição 3. Seja $f: X \to \mathbb{R} \cup \{+\infty\}$, f se diz regular (ou Clarke regular) em $x_0 \in \text{Int}(\text{Dom}(f))$ se

a) $\forall\, v,\quad f'(x_0;v)$ existe

b) $\forall\, v,\quad f'(x_0;v) = f^0(x_0,v).$ Δ

A pergunta natural é: Que funções são regulares?, uma resposta parcial é dada pela seguinte proposição.

Proposição 6. Seja $f: X \to \mathbb{R} \cup \{+\infty\}$ localmente lipschitziana em $x_0 \in \mathrm{Int}(\mathrm{Dom}(f))$.

a) Se f é estritamente derivável em x_0 então f é regular em x_0

b) Se f é convexa, então f é regular em x_0.

c) Uma combinação linear (por escalares não negativos) finita de funções regulares em x_0 é regular em x_0. Δ

2.5 Gradiente Generalizado e Propriedades básicas.

Definição 4. Seja $f: X \to \mathbb{R} \cup \{+\infty\}$ uma função derivável no sentido de Clarke no ponto $x \in X$. Denominaremos gradiente generalizado de f em x ao subconjunto ∂f de X^* definido por

$$\partial f(x) = \{\, x^* \in X^* \mid \langle x^*, v \rangle \leq f^0(x,v) \quad \forall\, v \in X \,\}. \qquad \Delta$$

Teorema 4. Seja $f: X \to \mathbb{R} \cup \{+\infty\}$ uma função localmente lipschitiziana, ela possui um gradiente generalizado não vazio $\partial f(x)$ em todo ponto $x \in \mathrm{Int}(\mathrm{Dom}(f))$, o qual é convexo, fechado e W^*- compacto e onde a função suporte $\sigma(\partial f(x),v) = \sup\{\langle x^*,v \rangle \mid x^* \in \partial f(x)\}$ Verifica:

$$\sigma(\partial f(x),v) = f^0(x;v). \qquad (*)$$

Prova. Como $f^0(x;.)$ é convexa positivamente homogênea e continua, ela é a função suporte de um subconjunto convexo fechado de elementos $x^*\in X^*$ tais que $< x^*,v > \leq f^0(x;v)$ para todo $v\in X$, é dizer, do gradiente generalizado: $\partial f(x)$, assim $\partial f(x)$ e não vazio, convexo, fechado e $(*)$ se tem.

Nos resta provar que $\partial f(x)$ é w^*-compacto, como $\sigma(\partial f(x),v) \leq \lambda \, ||v|| = \lambda\sigma(B^*,v)$ (donde B^* é a bola unitária de X^*), tem-se que:

$$\partial f(x) \subset \lambda B^*$$

para concluir faz-se uso do teorema de Banach-Alaoglu-Bourbaki (Veja Brézis [18] , Teorema III 15, pág. 42). $\wedge$

Assim, para o caso localmente lipschitziano, temos que o gradiente generalizado é bem comportado, ademais para este caso ∂f: Int(Dom(f)) $\rightrightarrows X^*$ é uma multiaplicação com imagens não vazias. Então podemos perguntar se ela tem alguma propriedade de continuidade.

Proposição 7. Seja $f: x \to \mathbb{R}\cup\{+\infty\}$, localmente lipschitziana , dotemos a X^* com a topologia W^*-fraca, então a multiplicação ∂f: Int(Dom(f)) $\rightrightarrows C(X^*)$ Verifica:

a) Ela é fechada em todo ponto $x\in$Int(Dom(f)), isto é quaisquer que sejam as sequências $x_n^* \to x^*$ (convergência w^*- fraca) e $x_n \to X$ (convergência forte), com $x_n^* \in \partial f(x_n)$ se tem $x^* \in \partial f(x)$.

b) $$\partial f(x) = \bigcap_{\varepsilon>0} \bigcup_{y\varepsilon B(x,\varepsilon)} \partial f(y)$$

c) ∂f é uma multiplicação s.c.s..

Prova.

a) Sejam $x_n \to x$ e $x_n^* \to x^*$ com $x_n^* \in \partial f(x_n)$, temos qualquer que seja $d \in X$

$$< x^*,d> = \lim_{n\to\infty} < x_n^*,d> \leq \lim_{n\to\infty} \sup f^0(x_n;d)$$
$$\leq f^0(x;d)$$

de onde $x^* \in \partial f(x)$ e pelo qual $x \rightrightarrows \partial f(x)$ é fechada.

b) é imediato a partir de a)

c) Em virtude de a) Nos basta provar a local w* -compacidade de ∂f. seja U uma vizinhança de 0 em X e $h \in U$, temos; para todo $x \in Int(Dom(f))$

$$-f^0(x;-h) \leq <x^*,h> \leq f^0(x;h)$$

logo:

$$|< x^*,h >| \leq f^0(x,h)$$

Se segue facilmente da proposição 4 que todo $x \in \mathrm{Int}(\mathrm{Dom}(f))$ tem uma vizinhança V em $\mathrm{Int}(\mathrm{Dom}(f))$ tal que: $\bigcup_{x \in V} \partial f(x)$ esta contido em um equicontinuo, logo w*-compacto, subconjunto de X*. Δ

A proposição 5 e seu corolário se traduz em termos da gradiente generalizado da seguinte maneira:

Proposição 8. Sejam f e g duas funções localmente lipschitzianas se $x \in \mathrm{Int}(\mathrm{Dom}(f)) \cap \mathrm{Int}(\mathrm{Dom}(g))$, então, se α e $\beta \in \mathbb{R}$,

(a) $\partial(\alpha f+\beta g)(x) \subset \alpha\partial f(x) + \beta\partial g(x)$

se $x \in \mathrm{int}(\mathrm{Dom}(f))$.

(b) $\partial(-f)(x) = -\partial f(x)$. Δ

Temos Também a seguinte proposição.

Proposição 9. Sob as hipóteses da proposição 8, se $\alpha, \beta > 0$ e se f, g são regulares, temos

$$\partial(\alpha f+\beta g)(x) = \alpha\partial f(x)+\beta\partial g(x).$$

Prova Basta notar que a função suporte do conjunto da esquerda é a mesma que a função suporte da soma dos conjuntos da direita. Δ

Observações.

1) Pode-se mostrar o seguinte fato. Consideremos $X=\mathbb{R}^n$.

Teorema 5 (Clarke [24]). Seja f localmente lipschitziana em x e suponhamos que S é qualquer conjunto de medida nula (no sentido de Lebesgue) em $\mathbb{R}^n$. Então

$$\partial f(x)= co\{\lim \nabla f(x_i) \mid x_i \to x, x_i \notin S, x_i \notin Nf\}$$

onde Nf é o conjunto onde f é não derivável, ($Nf \neq \mathbb{R}^n$ em virtude ao teorema de Rademacher [88], que estabelece que uma função localmente lipschitziana em $\mathbb{R}^n$ é diferenciável em quase todo ponto), e ademais

$$f^0(x;v)= \limsup_{y\to x} \{ \nabla f(y).v \mid y \notin S \cup Nf\}$$

Nesta forma o uso de algoritmos fica mais claro, veja por exemplo Polak, [64] , Mifflin [81] , Shor [72] e as referências citadas alí.

2) No caso de dimensão infinita, a fórmula anterior não pode ser generalizada diretamente, porém, o resultado de Rademacher tem uma extensão para os espaços de Banach separáveis, (obviamente com outra noção de conjunto de medida de Lebesgue nula, para os chamados conjuntos Haar-nulos, em $\mathbb{R}^n$ as duas noções coincidem), veja Christensen [23] , com isto, Thibault [89] , extende o resultado para este tipo de espaços, do seguinte modo:

$$\partial f(x)= \overline{co}\{w^*\text{-}\lim \nabla f(y) \mid y \to x, y \notin D\}$$

onde a adherença é tomada com relação à topologia fraca *, e $w^*\text{-}\lim \nabla f(y)$, o limite de gradientes convergentes na w*-topologia.

Porém para, espaços mais gerais não se tem uma generaliza ção muito boa.

Isto levou à seguinte definição:

Definição 5. Seja $f\colon X\to\mathbb{R}$ localmente lipschitiziana. Consideremos um conjunto D denso em X, f diz-se D-representavel em x se $\nabla f(x)$ (Gateaux-derivada) existe em qualquer ponto de D e

$$\partial f(x)=\overline{co}^{\,w^*}\{w^*\text{-}\lim_{y\to x}\nabla f(y)\mid y\in D\}. \qquad \Delta$$

Há vários teoremas relativos à definição anterior, por exemplo consultar Shi Shu-Chung [71], Correa e Jofrê [33], Correa e Thibault [34], Bernal e Rojas [13].

2.6 Relação com o subdiferencial e derivadas clássicas

Proposição 10 Seja $f\colon x\to\mathbb{R}$.

a) se f é convexa e finita em x, então $\partial f(x)=\partial_c f(x)$, onde $\partial_c f(x)=\{x^*\in X^*\mid \langle x^*,d\rangle\leq f'(x;d)\ \forall\ d\in X\}$

e o subdiferencial do análise convexo

b) se $f\in C^1$, então: $\partial f(x)=\{\nabla f(x)\}$

Prova (a) é direta já que $f'(x;v)=f^0(x;v)\ \forall v$.

(b) Também é direta do Teorema 2. Δ

Observações

Podem-se provar fatos mais finos, como por exemplo:

Teorema. (Clarke [29]). Se f é estritamente derivavel no sentido de Bourbaki em x, então f é localmente lipschitziana em x e $\partial f(x)=\{\nabla f(x)\}$ ($\nabla f(x)$ derivada estrita).

Inversamente, se f é localmente lipschitziana em x e $\partial f(x)$ se reduz a um singleton $\{x^*\}$, então f é estritamente diferenciável em x e $x^*=\nabla f(x)$. Δ

Ademais, temos a seguinte proposição.

Proposição 11. Seja f localmente lipschitziana em x tal que tenha uma Gâteaux-derivada (ou estrita, Fréchet, Hadamard) $Df(x)$. Então $Df(x)\in\partial f(x)$.

Prova. Por definição, $f'(x;d)$ existe para cada d e é igual a $\langle Df(x),d\rangle$. Claramente tem-se $f'\leq f^0$, de modo que $f^0(x;d)\geq\langle Df(x),d\rangle$ prove todo $d\in X$, ou equivalentemente $Df(x)\in\partial f(x)$. Δ

Nota. A inclusão pode ser estrita. Considere por exemplo $X=\mathbb{R}$: $f(x) = x^2 \operatorname{Sen}(1/x)$. Esta função é localmente lipschitziana em 0, e $\partial f(0)=[-1,1]$, mas $Df(0)=0$

Temos a seguinte caracterização para ótimos locais:

Proposição 12. Se f tem um máximo ou mínimo local em x então $0\in\partial f(x)$.

<u>Prova</u>. Em virtude da proposição 8 b), basta provar a proposição quando x é um mínimo local. Porém, neste caso é evidente que $\forall d$, $f^0(x;d) \geq 0$. Assim $0 \in \partial f(x)$. Δ

<u>Observação</u>. Podemos constatar que $\partial f(x)=\partial_c f^0(x;0)$, isto é, o gradiente generalizado de f em x é igual ao subdiferencial do análise convexa da função $d \to f^0(x;d)$, em d=0, em efeito, denotemos por $\varphi(d)=f^0(x;d)$, logo, $x^* \in \partial\varphi(0)$ se e somente se:

$$< x^*,d > \leq \varphi(d)- \varphi(0) \qquad \forall\ d$$

$$<=> \quad < x^*,d > \leq f^0(x;d) \qquad \forall\ d$$

$$<=> \quad x^* \in \partial f(x)$$

já que $\varphi(0)=f^0(x;0)=0$.

Este fato é importante, já que muitos resultados conhecidos do análise convexa podem-se transmitir para a teoria do gradiente generalizado via a relação anterior.

2.7 Teorema do valor médio de Lebourg, Regras da Cadeia e Gradientes Generalizados Parciais

a) <u>Teorema do valor Médio de Lebourg</u>.

Agora veremos o teorema do valor médio de Lebourg,o qual é o análogo, para o caso não diferenciável, do teorema do valor médio usual. A prova dada a seguir é devida a R. Cominetti e R.Correa [31], a qual é bastante elementar. Primeiramente recordemos algumas definições.

Definição 6. Seja $f:\mathbb{R}\to\mathbb{R}$, e $c\in\mathbb{R}$: As quatro derivadas de Dini de f em c são definidas como:

$$D^+f(c)=\limsup_{t\to 0^+} t^{-1}[f(c+t)-f(c)]$$

$$D^-f(c)=\limsup_{t\to 0^-} t^{-1}[f(c+t)-f(c)]$$

$$D_+f(c)=\liminf_{t\to 0^+} t^{-1}[f(c+t)-f(c)]$$

$$D_-f(c)=\liminf_{t\to 0^-} t^{-1}[f(c+t)-f(c)] \qquad \Delta$$

Primeiro Teorema do valor médio de Dini.

Se $f:\mathbb{R}\to\mathbb{R}$ é contínua em $[a,b]\subset\mathbb{R}$ com $b>a$, então existe $c\in(a,b)$ tal que

$$D_-f(c)\leq D^-f(c)\leq \frac{f(b)-f(a)}{b-a}\leq D_+f(c)\leq D^+f(c) \qquad (1)$$

ou

$$D_+f(c)\leq D^+f(c)\leq \frac{f(b)-f(a)}{b-a}\leq D_-f(c)\leq D^-f(c) \qquad (2)$$

Prova. Sem perda de generalidade, podemos assumir que $f(a)=f(b)$. Além disso, ou f é constante e (1) e (2) imediatamente verdadei ras, ou então f tem um mínimo ou máximo local em algum ponto $c\in(a,b)$. Se c é um mínimo local concluimos (1) da positividade de $f(c+t)-f(c)$ para todo t perto de 0. Similarmente se c é um máximo local obtemos as desigualdades (2). Δ

Segundo Teorema do valor médio de Dini.

Seja $f: \mathbb{R} \to \mathbb{R}$ contínua sobre $[a,b] \subset \mathbb{R}$ com $b>a$. Para qualquer derivada de Dini Df de f, existe $c\in(a,b)$, tal que:

$$\frac{f(b)-f(a)}{b-a} \leq Df(c) \tag{3}$$

Prova. Como na prova anterior, podemos assumir $f(a)=f(b)$. Faremos a prova para a derivada de Dini D^+f.

Se existe um ponto $d\in(a,b)$, tal que $f(d)\leq f(a)$, então f tem um mínimo num ponto $c\in(a,b)$. O resultado segue-se da positividade de $t^{-1}[f(c+t)-f(c)]$ para todo $t>0$ suficientemente pequeno. Se não, tomamos um máximo global $t_0\in(a,b)$ de f e $t_1\in(a,t_0)$. Se $D^+f(t_1) \geq 0$ tomamos $c=t_1$ para concluir; se não, f tem um mínimo local em algúm ponto $c\in(t_1 t_2) \subset (a,b)$ e ademais $D^+f(c)\geq 0$. Δ

Nota. A desigualdade inversa em (3) é obtida aplicando o mesmo resultado a $-f$.

No seguinte denotamos por $D^+f(x;v),\ldots, D_-f(x;v)$ as correspondentes derivadas de Dini de $h(t)=f(x+tv)$ em $t=0$.

Notação. Dados $x,y\in X$, a notação $[x,y]$ significa o segmento fechado o qual consiste de todos os pontos $tx+(1-t)y$ para $t\in [0,1]$; $]x,y[$ significa o segmento aberto.

Teorema do valor médio de Lebourg.

Seja $f:X\to\mathbb{R}$ localmente lipschitziana e sejam $x,y\in X$. Então existe um ponto $c\in]x,y[$ tal que:

$$-f^0(c;-(y-x)) \leq f(y)-f(x) \leq f^0(c;y-x)$$

ou equivalente,

$$f(y) \in f(x)+ \langle \partial f(c),y-x\rangle$$

Prova. Pela proposição , é fácil ver que toda derivada de Dini Df de f e para cada $c,v\in X$, temos

$$-f^0(c;-v) \leq Df(c;v) \leq f^0(c;v)$$

logo, o resultado segue-se imediatamente do primeiro teorema do valor médio de Dini. Δ

Agora faremos uso do segundo teorema do valor médio de Dini, para obter uma caracterização muito boa da derivada de Clarke.

Teorema 6. Seja $f: X\to\mathbb{R}$ localmente lipschitziana e seja Df qualquer das quatro derivadas de Dini de f. Então

$$f^0(x;v)= \limsup_{y\to x} Df(y;v).$$

<u>Prova</u>. Como foi notado na prova anterior, $Df(y;v) \leq f^0(y;v)$. Logo, tomando límite superior quando y converge a x, a s.c.s. de $f^0(.;v)$ nos da: $\limsup_{y\to x} Df(y;v) \leq f^0(x;v)$.

Para provar a desigualdade contraria observemos que para $y \in X$ arbitrário e $t>0$, o segundo teorema de Dini garante que existe $\alpha(t) \in (0,1)$ tal que:

$$\frac{f(y+tv)-f(y)}{t} \leq Df(y+\alpha(t)tv;v).$$

e tomando o limíte superior quando $y\to x$ e $t\to 0^+$, concluimos que

$$f^0(x;v) \leq \limsup_{\substack{y\to x\\ t\to 0^+}} Df(y+\alpha(t)tv;v) \leq \limsup_{y\to x} Df(y;v) \qquad (*)$$

Δ

<u>Comentários</u>

1) Notemos, ademais, que o teorema do valor médio de Lebourg segue-se também como uma consequência imediata da caracterização do teorema anterior e o primeiro teorema do valor médio de Dini.

2) Existem outros teoremas do valor médio generalizado; para isto podem ser consultados as referências, J.B.Hiriart-Urruty [45], M. Studniarski [74].

3) Da caracterização dada no teorema anterior se seguem derivações simples de muitos resultados em Análise Não-diferên-

ciável, especialmente os referentes a funções regulares. Com efeito, é óbvio de (*) que f é regular em x se e somente se qualquer derivada de Dini $Df(.;d)$ é s.c.s. em x para cada $d \in X$. Mais resultados nesta direção podem ser encontrados em R. Correa e A. Jofré [33], R. Correa e L. Thibault [34].

b) <u>Regras da cadeia para o Gradiente Generalizado</u>.

Em numerosos problemas (Controle ótimo, programação fracional, melhor aproximação, etc) a função objetivo e as funções que definem o conjunto de restrições são funções compostas.

<u>Exemplos</u>

1) A resposta de um sitema físico é uma quantidade $y(t)$ satisfazendo uma relação do tipo $y(t)=\theta(t,a)$ onde θ é uma função conhecida de t e a, t é um parâmetro auxiliar (tempo por exemplo) e a um parâmetro desconhecido de $\mathbb{R}^n$. Para vários valores de t, $t_1, t_2, \ldots, t_m$ $(m >> n)$ tem-se acesso a medidas aproximadas y_i de $y(t_i)$. Quer-se obter $a^* \in A$ (onde A é um conjunto de restrições) tal que minimize

$$f(a)= \Phi(y_1-\theta(t_1;a), \ldots, y_m-\theta(t_m;a))$$

sobre A. Em particular, a escolha de $\Phi(u_1, \ldots, U_m)=\sum_{i=1}^{m} u_i^2$ corresponde à estimação dos mínimos quadrados.

2) Generalização do problema de Fermat-Weber.

Seja $(K_i)_{i=0}^{p}$ uma família de subconjuntos não vazios de X, seja $\{\varphi_i\}_{i=1}^{p}$ uma família de funções de $\mathbb{R}_+$ em $\mathbb{R}$. O problema de otimização corresponde a minimizar $\sum_{i=1}^{p} \varphi_i(d(K_i;x))$ sobre K_0, onde $d(K_i;x)$ é a distância de x ao conjunto K_i.

No problema de Fermat-Weber, os subconjuntos K_i são pontos e $K_0=X$. A consideração de distâncias a subconjuntos e a introdução de um conjunto de restrições são inerentes aos problemas, especialmente para os problemas de localização. Em particular, o conjunto das restrições K_0 pode ter a seguinte estrutura:

$d(F_j, x) \leq d_j \quad \forall\, j=1,\ldots,q \quad d(F_m, x) \geq d_m \ \forall\, m=q+1,\ldots,r$ onde $\{F_l\}$ é uma familia de subconjuntos de X.

Se as funções φ_i são localmente lipschitzianas, a função critério $f= \Sigma\, \varphi_i o d(K_i;.)$ é localmente lipschitziana.

De forma mais geral, podem-se considerar critérios do seguinte tipo: $x \to \varphi(d(K_1;x),\ldots,d(K_p;x))$ com φ localmente lipschitziana.

Vemos assim, em vista dos problemas anteriores, que teoremas tipo regra da cadeia, são necessários.

Regras da Cadeia para o Gradiente Generalizado.

Sejam X,Y,Z espaços de Banach; não faremos distinção dos produtos de dualidade $<.,.>$ entre os diferentes espaços e seus duais topológicos.

Teorema 7. (Regra 1). Sejam $F: Y\to Z$ e $f:Z\to \mathbb{R}$ aplicações. F de classe C^1 e f localmente lipschitziana. Então denotando $DF(x_0) \in \mathcal{L}(Y,Z)$ o operador diferencial de F em x_0, temos:

$$\partial (f\circ F)(x_0) \subset \partial f(F(x_0))\circ DF(x_0).$$

A igualdade é satisfeita se f (ou-f) é regular em $F(x_0)$ ou $DF(x_0)$ é sobrejetora. Δ

Teorema 8. (Regra 2). Sejam $f:X\to \mathbb{R}$ e $\emptyset: \mathbb{R} \to \mathbb{R}$ localmente lipschitzianas, então

$$\partial (\emptyset o f)(x_0) \subset co \quad \{\partial \emptyset(f(x_0)).\partial f(x_0)\}$$

Ademais se $\emptyset \in C^1$, ou se $\emptyset$ (ou-$\emptyset$) é regular em $f(x_0)$ e $f \in C^1$, então:

$$\partial (\emptyset o f)(x_0) = \partial \emptyset(f(x_0)).\partial f(x_0). \qquad \Delta$$

<u>Teorema 9</u> (Regra 3). Sejam $F: X\to\mathbb{R}^m$ e $\varphi:\mathbb{R}^m\to\mathbb{R}$ localmente lipschitzianas. Então

$$\partial(\varphi\circ F)(x_0)\subset\overline{co}\,\{\sum_{i=1}^{m} u_i x_i^* \mid (u_1,\ldots,u_m)\in\partial\varphi(F(x_0)) \quad \text{e}$$

$$(x_1^*,\ldots,x_m^*)\in\prod_{i=1}^{m}\partial f_i(x_0)\}$$

onde $f_1,\ldots,f_m$ são as funções componentes de F, isto é $F=(f_1,\ldots,f_m)^T$.

Além disso, se as funções f_i, $i=1,\ldots,m$ são regulares em x_0, se φ é regular em $F(x_0)$ e se $\partial\varphi(F(x_0))\subset\mathbb{R}^m_+$, a igualdade se satisfaz. Δ

<u>Observação</u>. Notar que a estimativa dada na regra 2 aparece como um caso particular da inclusão dada na regra 3).

Faremos somente a prova da regra 3; as outras são análogas.

<u>Prova</u> (Regra 3). Seja $d\in X$. Consideremos uma sequência $\{x_n\}$ convergente a x_0 e uma sequência $\{\lambda_n\}$ de reais positivos (estritamente) convergente a zero;

denotemos

$$T_n = [\varphi(F(x_n+\lambda_n d)) - \varphi(F(x_n))] \, . \, \lambda_n^{-1}$$

φ: $\mathbb{R}^m \to \mathbb{R}$ é localmente lipschitziana, e pelo teorema do valor médio de Lebourg, existe $F_n \in]F(x_n), F(x_n+\lambda_n d)[$ em $\mathbb{R}^m$ e $\bar{u}_n \in \partial\varphi(F_n)$ tal que

$$T_n = \langle F(x_n+\lambda_n d) - F(x_n), \bar{u}_n \rangle \lambda_n^{-1} \tag{1}$$

$\bar{u}_n = (\bar{u}_n^1, \ldots, \bar{u}_n^m)$.

De acordo com o mesmo teorema, existe

$\bar{x}_n^i \in]x_n, x_n+\lambda_n d[$ e $\bar{y}_n^i \in \partial f_i(\bar{x}_n^i)$ para todo $i=1,\ldots,m$ tal que

$$f_i(x_n+\lambda_n d) - f_i(x_n) = \lambda_n \langle \bar{y}_n^i, d \rangle$$

Brevemente de (1);

$$T_n = \sum_{i=1}^{m} \bar{u}_n^i \langle \bar{y}_n^i, d \rangle = \langle \sum_{i=1}^{m} \bar{u}_n^i \, \bar{y}_n^i, d \rangle \tag{2}$$

As multiaplicações $\partial\varphi$: $\mathbb{R}^m \rightrightarrows \mathbb{R}^m$ e ∂f_i: $X \rightrightarrows X^*$ (X^* dotado da topologia w*-fraca) são s.c.s.. Por outro lado quando $n\to\infty$, $F_n \to F(x_0)$ e $\bar{x}_n^i \to x_0$.

Dado que a sequência $\{\bar{y}_n^i\}_m$ (resp. $\{\bar{u}_n\}_n$) é limitada em X^* (resp. em $\mathbb{R}^m$) podemos supor sem perdida de generalidade de que converge em X^* para a topologia w*-fraca (resp. que converge em $\mathbb{R}^m$) e tem-se:

$$\bar{y}_n^{\,i} \to \bar{y}^i \in \partial f_i(x_0) \qquad i=1,\dots, m \text{ (convergência w*-fraca)}$$

$$\bar{u}_n \to \bar{u} \in \partial\varphi(F(x_0)) \qquad (\text{em } \mathbb{R}^m).$$

Seja $D = \{x^* \in X^* \mid x^* = \sum_{i=1}^{m} u_i x_i^*,\ (u_1,\dots,u_m) \in \partial\varphi(F(x_0)),$

$$(x_1^*,\dots, x_m^*) \in \prod_{i=1}^{m} \partial f_i(x_0)\ \}$$

Notemos que D é um subconjunto w*-compacto de X^*, logo de acordo com (2) e (3), temos:

$$\lim_{n\to\infty} \sup T_n \leq \max \{\langle x^*,d\rangle \mid x^* \in D\}$$

De modo que:

$$(\varphi\circ F)^0(x_0;d) \leq \sigma(D;d) \quad \forall\ d \in X$$

o qual é equivalente a:

$$\partial(\varphi\circ F)(x_0) \subset \overline{co}\ D.$$

b) Consideremos a expressão $T_n^0 = [\varphi(F(x_0+\lambda_n d))-\varphi(F(x_0))]\lambda_n^{-1}$, f_i sendo regular em x_0, temos que para cada i:

$$f_i(x_0+\lambda_n d)-f_i(x_0)= \lambda_n f_i^0\ (x_0;d)+ \lambda_n\ \varepsilon_n^i$$

Com $\lim\limits_{n\to\infty} \varepsilon_n^i = 0.$

Seja ε_n o vetor $(\varepsilon_n^1, \dots, \varepsilon_n^m)^T$, seja $V(x_0;d)$ o vetor $(f_1^0(x_0;d),\dots, f_m^0(x_0;d))^T$. Já que φ é lipschitziana, T_n^0 pode ser escrito como:

$$T_n^0 = \{\varphi(F(x_0)+\lambda_n V(x_0;d)) - \varphi(F(x_0))\}\ \lambda_n^{-1}+o(1)\,||\varepsilon_n||$$

φ é regular em $F(x_0)$, de modo que;

$$\lim_{n\to\infty} T_n^0= \varphi^0(F(x_0);\ V(x_0;d)). \qquad (*)$$

Seja $(u_1,\dots,u_m) \in \partial\varphi(F(x_0))$ e $(x_1^*,\dots x_m^*) \in \prod\limits_{i=1}^m \partial f_i(x_0)$.

Como $\partial\varphi(F(x_0)) \subset \mathbb{R}_+^m$, temos

$$\sum_{i=1}^m u_i < x_i^*,d > \ \leq \sum_{i=1}^m u_i\ f_i^0(x_0;d)$$

$$\leq \max \{ \sum_{i=1}^{m} v_i \; f_i^0(x_0;d) \mid (v_1,\ldots v_m) \in \partial\varphi(F(x_0)) \}$$

$$= \max \{ <v,V(x_0;d)> \mid v=(v_1,\ldots,v_m) \in \partial\varphi(F(x_0)) \}$$

$$= \varphi^0(F(x_0);\; V(x_0;d))$$

por (*) $= \lim\limits_{n\to\infty} T_n^0$

$$= \lim_{n\to\infty} \frac{(\varphi\circ F)(x_0+\lambda_n d)-(\varphi\circ F)(x_0)}{\lambda_n}$$

$= (\varphi\circ F)^0(x_0;d)$ por ser φ e f regulares

logo, $\mathrm{Max}\,\{ < x^*,d> \mid x^* \in D\} \leq (\varphi\circ F)^0(x_0;d)$ $\forall$ d de onde o resultado. Λ

Comentários

1) No livro de Clarke [29], podem ser achados resultados um pouco mais finos das regras anteriores.

2) Na tese de Cominetti [30], encontramos também regras mais ge

rais (isto é, em espaços mais gerais e funções mais gerais).

Agora daremos algumas corolários das regras anteriores.

Corolários da regra de 3

Corolário 1. Se $\varphi \in C^1$, então

$$\partial(\varphi \circ F)(x_0) \subset \sum_{i=1}^{m} \frac{\partial \varphi}{\partial xi} (F(x_0)) \quad \partial f_i(x_0)$$

Com igualdade se as funções f_i são regulares em x_0 e se $\nabla \varphi(F(x_0)) \in \mathbb{R}^m_+$

Prova. Basta aplicar a regra 3. De fato, dado que $\partial f_i(x_0)$ é um convexo resulta que

$\frac{\partial \varphi}{\partial x_i} (F(x_0)) \; \partial f_i(x_0)$ é convexo e também

$$\Sigma \frac{\partial \varphi}{\partial x_i} (F(x_0)) \; \partial f_i(x_0)$$

é convexo, de modo que em na fórmula da regra 3, o símbolo $\overline{co}$ é superfluo. Δ

Observação. Tomando $\varphi(x) = \prod_{i=1}^{m} x_i$ no corolário anterior, temos:

$$\partial\left(\prod_{i=1}^{m} f_i\right)(x_o) \subset \sum_{i=1}^{m} \prod_{j\neq i} f_j(x_o)\,\partial f_i(x_0).$$

Com a igualdade se as funções f_i são regulares em x_0 e se $f_i(x_0) \geq 0$ $\forall i$. Δ

<u>Corolário 2</u>. Se $f_i: X \to \mathbb{R}$, $i=1,\ldots,m$ são localmente lipschitzianas, então

$$\partial\left(\max_{1\leq i\leq m} f_i\right)(x_0) \subset \operatorname{co}\left\{\bigcup_{i\in I(x_0)} \partial f_i(x_0)\right\}$$

onde $I(x_0)=\{i \mid f_i(x_0)= \max\limits_{1\leq i\leq m} f_i(x_0)\}$

Se ademais as f_i são regulares, se tem a igualdade.

<u>Prova</u>. É consequência da regra 3, considerando

$\varphi(x)= \max\limits_{1\leq i\leq n} x_i$, e notando que:

$$\partial\varphi(x)= \operatorname{co}\{e_i \mid i,\ x_i= \max_{1\leq j\leq m} x_j\}$$

onde e_i é o i-ésimo vetor coluna da matriz identidade mxm. Δ

<u>Observação</u>. Se o conjunto de índice I e não finito, pode ser mostrado um resultado análogo ao corolário 2. Como referência futura o citaremos, primeiramente vejamos uma definição.

Seja f_t uma familia de funções sobre X parametrizadas por $t \in T$, onde T é um espaço topológico. Suponhamos que para algum ponto $x \in X$, cada função f_t é localmente lipschitziana em x.

Denotemos por $\partial_{[t]} f_t(x)$ o conjunto ($\overline{co}$ denota a envoltura convexa w*-fechada)

$$\overline{co}\{ x^* \in X^* \mid x^*_n \in \partial f_{t_n}(x_n), \quad x_n \to x, \ t_n \to t,$$

$$t_n \in T, \ x^* \text{ é um ponto w*-adherente de } x^*_n \}.$$

<u>Definição 7</u>. A multiaplicação $(\tau, y) \rightrightarrows \partial f_\tau(y)$ diz-se <u>envolutiva</u> em (t,x) se:

$$\partial_{[T]} f_t(x) = \partial f_t(x)$$

(Note que esta condição certamente é satisfeita se t é um ponto isolado de T, de acordo à proposição 7(a)). Δ

Faremos as seguintes hipóteses:

(i) T é um espaço sequencialmente compacto.

(ii) Para alguma vizinhança U de x, a função $t \to f_t(y)$ é s.c.s.

para cada $y \in U$

(iii) Cada f_t, $t \in T$, é lipschitziana com constante K sobre U, e $\{f_t \mid t \in T\}$ é limitado.

Definamos a função $f: X \to \mathbb{R}$ por

$f(y) = \max \{f_t(y) \mid t \in T\}$

Notemos que as hipoteses anteriores aseguram que f esta definida e é finita sobre U. Também temos que f é lipschitz sobre U (de constante K), já que cada f_t o é.

Denotemos por $M(y)$ o conjunto $\{t \in T \mid f_t(y)=f(y)\}$

É fácil ver que $M(y)$ é não vazio e fechado para cada $y \in U$. Finalmente, para qualquer subconjuntos S de T, $P(S)$ significa a coleção de medidas de probabilidade de Randon suportadas sobre S.

<u>Teorema 10</u>. (Clarke). Além das hipóteses anteriores, suponhamos que :

(iv) X é separável, ou

(iv)' T é metrizável (o qual é verdadeiro em particular se T é separável). Então , temos

$$\partial f(x) \subset \{ \int_T \partial_{[T]} f_t(x) \; \mu(dt) \mid \mu \in P[M(x)] \}. \qquad (1)$$

Ademais, se a multiaplicação $(\tau,y) \rightrightarrows \partial f_\tau(y)$ é envolutiva em

(t,x) para cada $t \in M(x)$, e se f_t é regular em x para cada $t \in M(x)$, então f é regular e a igualdade se satisfaz na expressão (1) (com $\partial_{[T]} f_t(x) = \partial f_t(x)$).

Prova. Veja Clarke [29], Teorema 2.8.2, pág. 86, ou Levin e Ioffe [51], para o caso convexo. Δ

Nota. A interpretação do conjunto do lado direito da expressão (1) é o seguinte, um elemento x^* esta nesse conjunto, se $x^* \in X^*$ e existe uma função correspondente $t \to x_t^* \in \partial_{[T]} f_t(x)$ de T a X^* e um elemento μ de $P[M(x)]$ tal que, para todo $v \in X$, $t \to \langle x^*_t, v \rangle$ é μ-integravel, e

$$\langle x^*, v \rangle = \int_T \langle x_t^*, v \rangle \mu(dt) \qquad \Delta$$

Corolário da Regra 1.

Corolário 1. Sejam x_0, $d \in X$, denotemos por $f_{x_0,d}$ a função definida sobre $\mathbb{R}$ por $f_{x_0,d}(\lambda) = f(x_0+\lambda d)$. Então:

$$\partial f_{x_0,d}(0) \subset \langle \partial f(x_0), d \rangle$$

Com igualdade se f (ou $-f$) é regular em x_0.

Prova. É uma aplicação direta da Regra 1. De fato, neste caso $Y=\mathbb{R}$, $Z=X$ e $F = F_{x_0,d}: \lambda \to x_0+\lambda d$. Δ

<u>Observação</u>. Se $X=\mathbb{R}^n$, pode-se estabelecer:

$$\partial f_{x_0,d}(\lambda) = \langle \partial f(x_0+\lambda d), d \rangle \quad \forall \lambda \quad \text{e q.t.p. } x_0.$$

Com esta relação, o teorema do valor médio de Lebourg pode ser provado de uma forma fácil, para mais detalhes veja F.H.Clarke [29], pág 41.

<u>Corolários da Regra 2</u>

<u>Corolário 1</u>. Sejam $f: X \to \mathbb{R}$, $g:X \to \mathbb{R}$ localmente lipschitzianas. Se $g((x_0) \neq 0$, se tem que:

$$\partial(f \,|g)(x_0) \subset \frac{g(x_0)\ \partial f(x_0) - \partial g(x_0) f(x_0)}{[\,g(x_0)]^2}$$

Com igualdade se f e $-g$ são regulares em x_0, $f(x_0) \geq 0$ e $g(x_0) \geq 0$.

<u>Prova</u>.Aplicando a Regra 2, tem-se se $g(x_0) \neq 0$,

$$\partial(1/g)(x_0) = \frac{-\partial g(x_0)}{[\,g(x_0)\ ]^2}$$

De fato, basta considerar:

$$h(x) = 1/x \qquad x \neq 0$$

e como h(.) é diferenciável, o símbolo "co" é superfluo, já que $h'(gx_0))\partial g(x_0)$ é convexo. Δ

Finalmente, aplicando a observação seguinte ao corolário 1 da Regra 3 se obtem o resultado.

Corolário 2 . Se $f \in C^1$ na regra 2 e se $f(x_0)=0$, então,

$$\partial|f|\,(x_0) = [-\nabla f(x_0),\ \nabla f(x_0)]\ .$$

Prova. De fato, basta aplicar a Regra 2 com $\emptyset(x)=|x|$. Δ

Observação. 1) De forma geral, tem-se apenas a inclusão

$$\partial|f|(x_0) \subset \text{co}\ \{\partial f(x_0) \cup - f(x_0)\}.$$

Se $f(x_0)=0$; a igualdade em geral não se satisfaz, Veja J.B. Hiriart-Urruty [46] , mesmo que as funções sejam regulares em x_0.

2) Existe outro caso onde a operação de convexificação "co" não é necessária na fórmula da Regra 2; é quando $\emptyset$ é monótona em uma vizinhança de $f(x_0)$. De fato, se $\emptyset$: $\mathbb{R} \to \mathbb{R}$ é crescente (resp. decrescente) em uma vizinhança de u_0, tem - se $\partial\emptyset(u_0) \subset \mathbb{R}_+$ (resp. $\mathbb{R}_-$) (este resultado permanece válido se $\emptyset$: $\mathbb{R} \to \overline{\mathbb{R}}$ não é necessáriamente localmente lipschitziana.

3) A Regra 3 aplicada com m=1 nos da outro caso de igualda-

de no contexto da Regra 2, isto é:

Se f é regular em x_0, se $\emptyset$ é regular em $f(x_0)$ e se $\partial\emptyset(f(x_0)) \subset \mathbb{R}_+$, então

$$\partial(\emptyset \circ f)(x_0) = \partial\emptyset(f(x_0))\partial f(x_0).$$

<u>Nota</u>. Devido à natureza local da noção de gradiente generalizado, as propriedades requeridas sobre as funções (propriedade de lipschitz, C^1,...) necessitam sòmente ser assumidas em umas vizinhanças dos pontos considerados.

As regras anteriores (regra 2, por exemplo) podem ser aplicados para derivar condições de otimalidade (necessárias e |ou suficientes) para programação fraccional, isto é, problemas onde a função objetiva tem a forma

$$x \to f(x) = [\prod_{i=1}^{m} f_i(x)]\,[\prod_{j=1}^{p} g_j(x)]^{-1}$$

Para resultados nesta direção pode-se consultar J.M. Borwein [15], para o caso de funções quasi-diferenciáveis.

c) <u>Gradientes Generalizados Parciais</u>

Sejam X_1, X_2 dois espaços de Banach e $f: X_1 \times X_2 \to \mathbb{R}$ localmente lipschitziana em (x_1, x_2). Denotemos por $\partial_1 f(x_1,x_2)$ o <u>gradiente generalizado</u> (parcial) de $f(.,x_2)$ em x_1 e por

$\partial_2 f(x_1,x_2)$ o de $f(x_1,.)$ em x_2. A notação $f_1^0(x_1,x_2;v)$ representará a derivada generalizada de Clarke em x_1 na direção $v \in X_1$ da função $f(.,x_2)$, analogamente $f_2^0(x_1,x_2;v)$.

A questão natural é saber qual é a relação entre $\partial f(x_1,x_2)$ e $\partial_1 f(x_1,x_2) \times \partial_2 f(x_1,x_2)$.

Em geral, não há nenhuma relação de inclusão, como mostra o seguinte exemplo:

Consideremos $f: \mathbb{R}^2 \to \mathbb{R}$ dada por:

$f(x,y) = \max [\ \min [\ x,-y\ ], y-x\ ]$,

Calculemos $\partial f(0,0)$. Definamos

$C_1 = \{(x,y) \mid y \leq 2x \quad e \quad y \leq -x\}$

$C_2 = \{(x,y) \mid y \leq x/2 \quad e \quad y \geq -x\}$

$C_3 = \{(x,y) \mid y \geq 2x \quad ou \quad y \geq x/2\}$.

Então $\mathbb{R}^2 = C_1 \cup C_2 \cup C_3$, e temos:

$$f(x,y) = \begin{cases} x & \text{se} \quad (x,y) \in C_1 \\ -y & \text{se} \quad (x,y) \in C_2 \\ y-x & \text{se} \quad (x,y) \in C_3 \end{cases}$$

Notar que a fronteira destes tres conjuntos C_1, C_2, C_3, forman um conjunto S de medida nula, e que se $(x,y) \notin S$, então f é diferenciável e $\nabla f(x,y)$ é um dos pontos (1,0), (0,-1) ou (-1,1), logo pela caracterização do gradiente generalizado para o caso $X=\mathbb{R}^n$, temos que

$$\partial f(0,0) = \overline{co}\ \{(1,0),\ (0,-1),\ (-1,1)\}$$

Ademais $f(0,y) = \max[0,y]$, de modo que $\partial_y f(0,0) = [0,1]$, similarmente, $\partial_x f(0,0) = [-1,0]$, logo:

$$\partial_x f(0,0) \times \partial_y f(0,0) \not\subset \partial f(0,0) \not\subset \partial_x f(0,0) \times \partial_y f(0,0).$$

Mas no caso regular, tem-se a seguinte relação:

<u>Proposição 13</u>. Se f é regular em $x=(x_1, x_2)$, então

$$\partial f(x_1, x_2) \subset \partial_1 f(x_1, x_2) \times \partial_2 f(x_1, x_2).$$

<u>Prova</u>. Seja $z=(z_1, z_2) \in \partial f(x_1, x_2)$. Devemos provar que $z_1 \in \partial_1 f(x_1, x_2)$, o qual é equivalente a provar que:

$$\langle z_1, v\rangle \leq f_1^0(x_1, x_2; v) \qquad \forall\, v \in X_1$$

Mas, como f é regular, temos que:

$$f_1^0(x_1, x_2; v) = f_1'(x_1, x_2; v)$$

$$= f'((x_1, x_2) \; ; \; (v,0))$$

$$= f^0((x_1, x_2) \; ; \; (v,0))$$

e, pela definição do gradiente generalizado, temos que:

$$\langle (z_1, z_2), (v,0) \rangle = \langle z_1, v \rangle \leq f^0((x_1, x_2); (v,0))$$

$$= f_1^0 \; (x_1, x_2; v) \qquad \Delta$$

Para obtermos uma relação no caso não regular, definimos a <u>projeção</u> $\Pi_1 \, \partial f(x_1, x_2)$ como o conjunto

$$\{ z_1 \in X_1^* \mid \text{para algum } z_2 \in X_2^*, (z_1, z_2) \in \partial f(x_1, x_2)\} \text{ e}$$

analogamente $\Pi_2 \, \partial f(x_1, x_2)$.

<u>Proposição 14</u>. $\partial_1 f(x_1, x_2) \subset \Pi_1 \, \partial f(x_1, x_2)$.

<u>Prova</u> . Fixemos x_2, e seja f_1 a função $f_1(x)= f(x,x_2)$ definida sobre X_1. Seja $F: X_1 \to X_1 x \, X_2$ definida por $F(x)= (x,x_2)$, notemos que $DF(x)$ é dada por $\langle DF(x), v \rangle = (v,0)$ já que $f_1 = f o F$ aplicando a regra da cadeia 3 com $x_0 = x_1$ tem-se o resultado.

<u>Corolário 3</u>. $\partial_1 \, f(x_1, x_2) x \; \partial_2 f(x_1,x_2) \subset \Pi_1 \; \partial f(x_1,x_2) x \; \Pi_2 \partial f(x_1,x_2)$. Δ

2.8 Conceitos Geométricos Associado: Cones Tangente e Normal

Seja K um subconjunto de vazio de X. Designaremos por $d_K(.)$ ou $d(K;.)$ a função "distância" a K definida por:

$$d_K(y) = \inf \{ \| x-y \| \mid y \in K \}$$

Ela é evidentemente lipschitziana com constante de Lipschitz $\lambda=1$: $|d_K(y)-d_K(z)| \leq \| y-z \|$

Consequentemente, ela é derivavel no sentido de Clarke.

Definição 8. Seja $x \in K$. Um vetor $v \in \dot{X}$ é tangente a K em x se $d_K^0(x;v)=0$. O conjunto de todos os tangentes a K em x é denotado $T_K(x)$ ou $T(K;x)$.

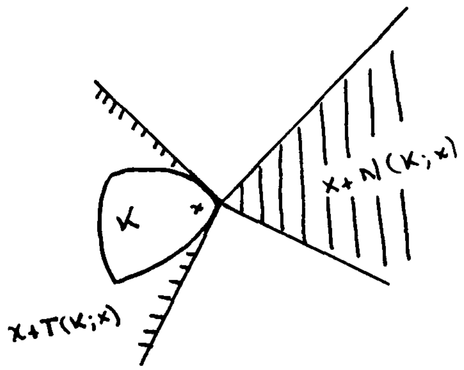
Observações.

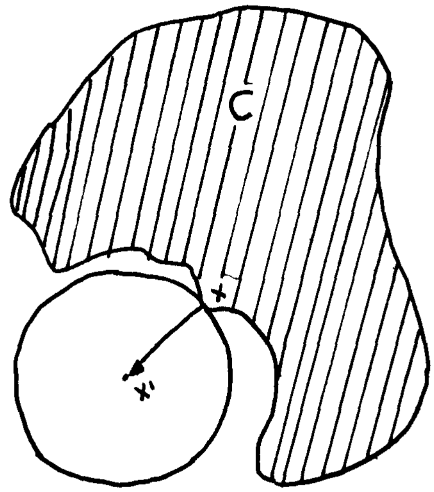
1) Na definição anterior somente involucra a natureza local de K em x.

2) Como $d_K^0(x;.)$ é convexa e continua, $T(K;x)$ é um cone fechado. (Em particular $0 \in T(K,x)$).

Definição 9. Definimos o cone normal a K em x por

$$N_K(x) = N(K;x) = \{x^* \in X^* \mid \langle x^*, v \rangle \leq 0 \ \forall\, v \in T(K;x)\}$$





o vetor $\frac{x'-x}{|x'-x|} \in N(C;x)$.

Temos a seguinte caracterização alternativa de $N(K;x)$ em termos de gradientes generalizados:

<u>Proposição 15</u>.

$$N(C;x) = \overline{\{ \bigcup_{\lambda \geq 0} \lambda \partial d_C(x) \}}$$

onde a barra denota a w*-adherença.

<u>Prova</u>. Pela definição de $T(C;x)$, temos que

$$v \in T(C;x) \iff \langle x^*, v \rangle \leq 0 \quad \forall x^* \in \partial d_C(x).$$

Com o qual segue-se que o cone polar a $T(C;x)$ é o cone w*-fechado gerado por $\partial d_C(x)$, que é o que diz a proposição. Δ

Proposição 16. Seja f lipschitziana com constante L sobre o conjunto $S \in X$. Seja $x \in K \subset S$ e suponhamos que f atinge um mínimo sobre K em x. Então para qualquer $\hat{L} \geq L$ a função $g(y)=f(y)+\hat{L}d(K;y)$ atinge um mínimo sobre S em x. Se $\hat{L} > L$ e K é fechado, então qualquer outro ponto que minimize g sobre K deve estar também em K.

Prova. Para provar a primeira conclusão, o faremos por contradição. Então existe um ponto $y \in S$ e $\varepsilon>0$ tal que $f(y)+\hat{L}d(K;x) < f(x)-\hat{L}\varepsilon$. Seja p um ponto em K tal que $\|y-p\| \leq d(K;x)+\varepsilon$.

Então:

$$f(p) \leq f(y)+\hat{L}\|y-p\| \leq f(y)+\hat{L}(d(K;x)+\varepsilon) < f(x),$$

o qual contradiz o fato que x minimiza f sobre K.

Agora seja $\hat{L} > L$, e seja y também um mínimo de g sobre K. Então

$$f(y)+\hat{L}d(K;y)=f(x) \leq f(y)+(L+\hat{L})d(K;y)|2$$

(pela primeira conclusão aplicada a $(L+\hat{L})|2$), o qual implica que $d(K;y)=0$; e assim $y \in K$. Δ

Corolário 4. Suponhamos que f é localmente lipschitziana em x e que atinge um mínimo sobre K em x. Então $0 \in \partial f(x)+N(K;x)$ ou equivalentemente $N(K;x) \cap -\partial f(x) \neq \emptyset$.

<u>Prova</u>. Seja S uma vizinhança de x sobre a qual f é lipschitziana (de constante L). Podemos supor $K \subset S$ (já que K e $K \cap S$ tem os mesmos cones normais em x), de modo que x minimiza $f(y)+ Ld(K;y)$ localmente. Assim,

$$0 \in \partial(f+Ld_k)(x) \subset \partial f(x)+ L\,\partial d_k(x).$$

Agora o resultado segue-se da proposição 15. △

Quando K é convexo, se tem um conceito bem conhecido de vetor normal: <u>$x^* \in X^*$ diz-se normal a K em x (no sentido da análise convexa)</u> se $< x^*, x-k > \geq \quad \forall\, k \in K$.

<u>Lema 3</u>. Se K é convexo a função $d_K(.)$ é convexa.

<u>Prova</u>. Em efeito sejam $x,y \in X$ e $\lambda \in (0,1)$ dados.

Para qualquer $\varepsilon>0$, tomamos c_x, c_y em K tal que

$\| c_x-x \| \leq d(K;x)+ \varepsilon$, $\| c_y-y \| \leq d(K;y)+ \varepsilon$, e seja $c \in K$ dado por $c=\lambda c_x+(1-\lambda)\, c_y$. Então

$$\begin{aligned} d(K;\lambda x+(1-\lambda)y) &\leq \| c-\lambda x-(1-\lambda)y\| \\ &\leq \lambda \| c_x-x \| +(1-\lambda)\| c_y-y\| \\ &\leq \lambda\, d(K;\lambda) + (1-\lambda)\, d(K;y) + \varepsilon \end{aligned}$$

fazendo agora ε tender o zero, tem-se o resultado. △

<u>Proposição 17</u>. Se K é convexo, $N(K;x)$ coincide com o cone de normais no sentido da análise convexa.

<u>Prova</u>. Seja x^* normal a K em x no sentido da análise convexa. Então o ponto $k = x$, minimiza $f(k) = < x^*, x-k >$ sobre K, de modo que:

$$0 \in - x^* + N(K;x)$$

o qual mostra que $x^* \in N(K;x)$.

Para completar a prova, é suficiente, em vista da proposição 15, provar que todo elemento $x^* \in \partial d_k(x)$ é normal a K em x no sentido de análise convexa (já que o conjunto de tais normais é um cone convexo w^*-fechado).

Como $d_K(.)$ é convexa (pelo lema anterior), temos que $\partial d_K(x) = \partial_c d_K(x)$, logo:

$$d_K(y) - d_K(x) \geq < x^*, y-x > \quad \forall\, y \in X$$

o qual implica que:

$$< x^*, k-x > \leq 0 \qquad \forall\, k \in K$$

com o qual termina-se a prova. Δ

<u>Corolário 5</u>. Se K é convexo, então

$$v \in T(K;x) \iff d_k^0(x;v) = d_k'(x;v) = 0.$$

<u>Prova</u>. Basta notar que neste caso $d_K(.)$ é convexa, logo ela é regular. Δ

Vamos mostrar agora que o conceito de vetor tangente definido anteriormente é independente da norma (e assim da função distância) usada sobre X.
Este fato é muito bom no sentido que podemos fazer uso, em certos problemas, de distâncias nas quais seja mais fácil calcular tangentes (o normais).

<u>Teorema 11</u>. Um elemento v de X é tangente e K em x se e somente se, para toda seguência x_n em K convergente a x e toda sequência t_n em $(0,+\infty)$ convergente a zero, existe uma sequência v_n em X convergente a v tal que:

$$X_n+t_n v_n \in K \qquad \forall\, n.$$

<u>Prova</u>. Veja, Clarke [29], teorema 2.4.5., pág 54. Λ

<u>Proposição 18</u>. Um elemento $x^* \in X^*$ pertence a $\partial f(x)$ e se somente se $(x^*,-1) \in N(epi(f);(x,f(x))$.

<u>Prova</u> Veja F.H. Clarke corolário, pág 61. Λ

<u>Proposição 19</u>. Seja $x \in K$. Então

$$\partial \psi_K (x) = N(K;x).$$

onde $\psi_K(.)$ é a função indicatriz de K, isto é:

$$\psi_K(x) = 0 \text{ se } x \in k, \ \psi_K(x) = +\infty \quad \text{se} \quad x \notin K.$$

Prova. Temos que $x^* \in \partial\psi_K(x)$ se e somente se $(x^*,-1) \in N(\text{epi}(\psi_K);(x,0))$. Mas claramente epi $(\psi_K) = K \times [0,\infty)$ o que é equivalente, pelo Teorema 11, a

$$X^* \in N(K;x), \quad -1 \in N([0,\infty);0).$$

Mas, $-1 \in N([0,\infty);0)$ sempre é verdadeiro (proposição 16); logo a fórmula segue-se. Δ

Para futuras referências, mencionemos também o seguinte resultado de F.H. Clarke.

Proposição 20. Seja f localmente lipschitziana em x. Então o epígrafo de $f^0(x;.)$ é $T(\text{epi}(f);(x,f(x))$; isto é

$$(v,r) \in T\ (\text{epi}(f),(x,f(x)) <=> r \geq f^0(x;v).$$

Prova. Veja F.H. Clarke [29], Teorema 2.4.9 a, pág 59. Δ

2.9 Aplicações à Otimização: Condições necessárias e suficientes de otimalidade.

Consideremos primeiramente o problema de otimização em sua forma geral. Seja A uma parte não vazia de X e $f:X \to \mathbb{R} \cup \{+\infty\}$.

Consideremos o seguinte problema:

P) Minimizar $f(x)$

s.a. $x \in A$.

<u>Definição 10</u>. $x_0 \in A$ diz-se um <u>mínimo local</u> de f sobre A, se f é finita em x_0 e se existe uma vizinhança V de x_0 tal que : $f(x) \geq f(x_0) \quad \forall\, x \in A \cap V$. Δ

Recordemos (Veja capítulo I) que $D(A;x_0)$ denota o cone de deslocamentos aderentes a A em x_0, isto é:

$$D(A;x_0)=\{d \in X \mid d = \lim_{n\to\infty} \lambda_n(x_n-x_0),\ \text{com } (x_n) \subset A,\quad x_n \to x_0\ ,$$

$$(\lambda_n) \subset \mathbb{R}_+,\ n\to\infty\ ,\ x_0 \in \bar{A}\ \}.$$

a) <u>Condições Necessárias</u>

Temos os seguintes teoremas.

<u>Teorema 12</u> . Se $x_0 \in A$ é um mínimo local de f sobre A e se f é localmente lipschitziana em x_0, então

$$f^0(x_0;d) \geq 0 \qquad \forall\, d \in D(A;x_0). \qquad (*)$$

<u>Prova</u>. Seja $d \in D(A;x_0)$ arbitrário, então temos que $d=\lim d_n$ com $d_n= \lambda_n(x_n-x_0)$ com $\lambda_n \geq 0$, $x_n \to x_0$ e $x_n \in A$.

Para d = 0, o resultado é trivial, pois $f^0(x_0;0)=0$. Assim podemos supor que $\lambda_n > 0 \quad \forall\, n$.

Como x_0 é mínimo local de P), e $x_n \to x_0$, existe $n_0 \in \mathbb{N}$ tal que:

$$f(x_0) \leq f(x_n) \qquad \forall\, n \geq n_0$$

Denotando $\alpha_n = \lambda_n^{-1} > 0$, temos que:

$$\frac{f(x_0+\alpha_n d_n) - f(x_0)}{\alpha_n} \geq 0$$

ou

$$0 \leq \frac{1}{\alpha_n}[f(x_0+\alpha_n d_n) - f(x_0+\alpha_n d)] + \frac{1}{\alpha_n}[f(x_0+\alpha_n d) - f(x_0)]$$

Como f é localmente lipschitziana em x_0, deduzimos que:

$$0 \leq L||d_n-d|| + \frac{1}{\alpha_n}[f(x_0+\alpha_n d) - f(x_0)]$$

onde L é a constante de lipschitz local de f, logo, tomando o $\limsup\limits_{n\to\infty}$ e recordando que $d_n \to d$, temos que:

$$f^0(x_0;d) \geq 0$$

e pela arbitrariedade de $d \in D(A;x_0)$, temos a conclusão. Δ

<u>Teorema 13</u>. Seja $x_0 \in A$ e consideremos um cone convexo M com vértice 0 tal que $M \subset D(A;x_0)$.
Se f é uma função localmente lipschitziana em x_0 e se x_0 é um mínimo local de f sobre A, então

$$0 \in \partial f(x_0) + M^0$$

onde $M^0 = \{x^* \in X^* \mid \langle x^*,x \rangle \leq 0 \quad \forall\, x \in M\}$ é o <u>cone polar</u> de M.

<u>Prova</u>. Denotemos por $g(d) = f^0(x_0;d)$, e consideremos o problema: Min $\{ g(d) \mid d \in D(A;x_0)\}$, pelo teorema anterior o ponto ótimo deste problema e d=0.

Por outro lado, como $D(A;x_0)$ é fechado, temos que $\bar{M} \subset D(A;x_0)$ e como M é um cone com vértice 0, $d=0 \in \bar{M}$, logo d=0 também é a solução ótima do problema: min $\{g(d) \mid d \in \bar{M}\}$.
Pelo corolário 4, temos que:

$$0 \in \partial g(0) + N(\bar{M},0) \qquad (1)$$

Mas g(.) é convexa pelo Teorema 3 , logo, $\partial g(0) = \partial_c\, g(0)$, mas pela observação seguinte à proposição 12, temos que: $\partial g(0) = \partial f(x_0)$.

Além disso como M e convexo, $\bar{M}$ também e convexo e $N(\bar{M},0)$ é o cone normal da análise convexa, pela qual: $N(\bar{M},0)=M^0$.
Assim (1) é equivalente a:

$$0 \in \partial f(x_0)+M^0. \qquad \Delta$$

b) <u>Condições Suficientes</u>

Para obter condições suficientes de otimalidade devemos pedir algumas hipóteses extra sobre a função e/ou o conjuntode restrições A. Umas delas são as seguintes

<u>Definição 11</u>. Uma função $f: X \to \mathbb{R}$ localmente lipschitziana diz-

se <u>localmente pseudoconvexa</u> em x_0 se existe uma vizinhança V de x_0 tal que:

$$\forall\, x \in V, \quad f^0(x_0;x-x_0) \geq 0 \quad \Rightarrow f(x) \geq f(x_0). \qquad \Delta$$

Se esta propriedade é satisfeita para todo $x_0 \in X$, dizemos simplesmente que f é <u>localmente pseudoconvexa</u>. Se a relação anterior é satisfeita globalmente (i.e., V=X), f é dita <u>pseudoconvexa</u>.

<u>Definição 12</u>. Dizemos que o conjunto de restrições A verifica a <u>condição (L)</u> se:

existe V, vizinhança de x_0, tal que para todo $x \in V \cap A$, temos $x-x_0 \in D(A;x_0)$, onde $D(C;y_0)$ e o cone de deslocamentos adherentes a C em x_0.

<u>Teorema 14</u>. Sob as hipóteses seguintes:

(a) f é localmente pseudoconvexa em x_0,

(b) A verifica a condição (L).

Se a condição necessária de otimalidade (*) se verifica em x_0, então x_0 é um mínimo local de f sobre A.

<u>Prova</u>. Temos por (*) que $f^0(x_0;d) \geq 0 \;\; \forall d \in D(A;x_0)$, como A verifica a condição (L), temos que existe uma vizinhança V_1 de x_0 tal que $\forall\, x \in V_1 \cap A$, se tem $d=x-x_0 \in D(A;x_0)$, logo $f^0(x_0;x-x_0) \geq 0$

$\forall\ x \in V_1 \cap A$ e fazendo uso da pseudoconvexidade local de f em x_0 concluímos que $f(x) \geq f(x_0)$. Isto é x_0 é um mínimo local de $f(x)$ sobre A. Δ

Teorema 15. Seja $x_0 \in A$ tal que $D(A;x_0)$ seja convexo; uma condição necessária para que x_0 seja um mínimo local de f sobre A é que:

$$\partial f(x_0) \cap -[D(A;x_0)]^0 \neq \emptyset \qquad (**)$$

Ademais, se f é localmente pseudoconvexa em x_0 e se A verifica a propriedade (L) em x_0, então (**) é uma condição suficiente para que x_0 seja um mínimo local de f sobre A.

Prova. Basta considerar $M = D(A;x_0)$ no teorema 13 para concluir (**).

A suficiença é similar à prova do teorema anterior. Δ

c) Kuhn-Tucker generalizado

Há situações que o conjunto de restrições tem uma estrutura do seguinte tipo

$$A = \{ x \mid x \in C,\ \ f_i(x) \leq 0,\ g_j(x)=0,\ i \in I,\ j \in J\}$$

onde $I = \{1,2,...,n\}$, $J = \{1,2,...m\}$ são conjuntos de índices não simultaneamente vazios.

O teorema de Kuhn-Tucker usual (referências [32], [56]) nos permite resolver o seguinte problema:

$$\begin{array}{llll} \text{Minimizar } f_0(x) & & & \\ \text{s.a.} & f_i(x) \leq 0 & i \in I & \\ & g_j(x) = 0 & j \in J & \text{P)} \\ & x \in C & & \end{array}$$

quando as funções f_0, $f_i (i \in I)$, $g_j (j \in J)$ são diferenciáveis . Agora veremos uma generalização do teorema anterior para funções que não são necessáriamente diferenciáveis, mas localmente lipschitziana, devida a F.H.Clarke [25].

Suponhamos que $f_0, f_1, \ldots, f_n : X \to \mathbb{R}$ e $g_1, \ldots, g_m : X \to \mathbb{R}$ são localmente lipschitzianas e C fechado, consideremos o problema P) descrito anteriormente.

Sejam $F: X \to \mathbb{R}^n$ e $G: X \to \mathbb{R}^m$ definidos por $F = [f_1, \ldots, f_n]$ e $G = [g_1, \ldots, g_m]$ respectivamente.

<u>O Lagrangeano</u> é a função $L(x,\lambda,\mu,\gamma,\rho): X \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$ definida por

$$L(x,\lambda,\mu,\gamma,\rho) = \lambda f_0(x) + \langle \mu, F(x) \rangle + \langle \gamma, G(x) \rangle + \rho \, \|(\lambda,\mu,\gamma)\| \, d_C(x).$$

Quando $C=X$, $d_c(x) \equiv 0$, logo o anterior se reduz à expressão usual da Teoría clássica.

<u>Teorema 16</u>. (Kuhn-Tucker generalizado) sob as hipoteses feitas anteriormente. Seja $\bar{x}$ que resolve P). Então para todo ρ suficientemente grande existe $\lambda \geq 0$, $\mu \geq 0$, e γ não todos nulos tal que:

(a) $\quad < \mu, F(\bar{x}) > = 0,$

(b) $\quad 0 \in \partial_x L(\bar{x}, \lambda, \mu, \gamma, \rho).$

<u>Notas</u> .

1) A prova mostrará que a conclusão é válida para qualquer $\rho > \hat{\rho}$, onde $\hat{\rho}$ é uma constante de Lipschitz para a função $[f_0, F, G]$ em uma vizinhança de $\bar{x}$.

2) A notação $\mu \geq 0$ significa que cada componente μ_i de μ é não negativa. A condição $< \mu, F(\bar{x}) > = 0$ é algumas vezes chamada "Complementary Slackness". Isto é equivalente a dizer que $\mu_i = 0$ quando $f_i(\bar{x}) < 0$ (i.e., quando a restrição $f_i(\bar{x}) \leq 0$ é (localmente) inativa).

3) Por suposto as conclusões do teorema se satisfazem para mínimos locais, sômente há que trocar C por $C \cap \bar{B}(\varepsilon, \bar{x})$ para qualquer $\varepsilon > 0$.

4) A condição $0 \in \partial_x L$ implica a condição "separada"

$$0 \in \lambda\partial f_0(\bar{x}) + \sum_{i\in I} \mu_i\, \partial f_i(\bar{x}) + \sum_{j\in J} \gamma_j\, \partial g_j(\bar{x}) + \rho\,||(\lambda,\mu,\gamma)||\,\partial d_C(\bar{x})$$

(Como também algumas variantes na qual o último termo é trocado por $N(C;\bar{x})$). Esta condição é em geral mas fraca que a forma dada no teorema (Veja Proposição 8) sem embargo às duas são equivalentes para "dados diferenciáveis (ou convexos)".

5) O vetor (λ,μ,γ) anterior é chamado um multiplicador. Note que se (λ,μ,γ) é um multiplicador (i.e., se satisfaz as conclusões do teorema para algum ρ), então também $(t\lambda,t\mu,\lambda\gamma)$ $\forall\, t>0$. Assim o teorema poderia também estipular $||(\lambda,\mu,\gamma)||=1$.

<u>Prova do Teorema16.</u> Seja H o conjunto

$$\{h=(\lambda,\mu,\gamma) \in \mathbb{R}^{1+n+m} \mid \lambda \geq 0,\ \mu \geq 0,\ ||\lambda,\mu,\gamma)||=1\},$$

Seja $\varepsilon > 0$ dado arbitrário, e definamos $\varphi: X \to \mathbb{R}$ por

$$\varphi(y)=\max_H \{(\lambda,\mu,\gamma).(f_0(y)-f_0(\bar{x})+\varepsilon),F(y),G(y)\}$$

Como f_0, F, G são localmente lipschitzianas em $\bar{x}$, temos que φ é localmente lipschitziana em $\bar{x}$, e $\varphi(\bar{x})=\varepsilon$.

Observemos que φ é positiva sobre C, de fato, se φ não for positiva, isto é, $\varphi(y) \leq 0$, então $f_i(y) \leq 0$, $h_j(y)=0$ (i.e, y é factivel para P)), e $f(y) \leq f(\bar{x})-\varepsilon$, o que uma contradição com o fato de ser $\bar{X}$ otimal.

Além disso $\bar{x}$ satisfaz

$$\varphi(\bar{x}) \leq \inf\{\varphi(y) \mid y \in C\} + \varepsilon,$$

logo pelo pricípio Variacional de Ekeland, existe um ponto $u \in \bar{B}(\sqrt{\varepsilon}, \bar{x})$ tal que:

$$\varphi(y) + \sqrt{\varepsilon}\ \|y-u\| \geq \varphi(u) \qquad \forall\ y \in C.$$

Se $\hat{\rho}$ é a constante de lipschitz da nota 2 anterior, então é fácil ver que qualquer $\rho > \hat{\rho}$ é uma constante de lipschitz local (quando ε é suficiente pequeno) para a função $\varphi(y) + \sqrt{\varepsilon}\ \|y-u\|$ perto do ponto $y=u$.
Pela proposição 16, u também minimiza, sobre alguma vizinhança de u, a função

$$y \to \varphi(y) + \sqrt{\varepsilon}\ \|y-u\| + \rho\ d_C(y)$$

Mas

$$\theta(y) = \varphi(y) + \rho\ d_C(y) = \max_H\{(\lambda,\mu,\gamma).(f(y)-f(\bar{x})+\varepsilon, F(y), G(y)\} + \rho d_C(y)$$

$$= \max_H \{\lambda f(y) - \lambda f(\bar{x}) + \lambda\varepsilon + \langle \mu, F(y)\rangle + \langle \gamma, G(y)\rangle\} + \rho d_C(y)$$

$$= \operatorname{Max}_H \{\lambda f(y) + \langle \mu, F(y)\rangle + \langle \gamma, G(y)\rangle + \rho d_C(y) - \lambda f(\bar{x}) + \lambda\varepsilon\}$$

$$(\text{Def. de } L) = \operatorname{Max}_H \{L(y,\lambda,\mu,\gamma,\rho) - \lambda f(\bar{x}) + \varepsilon\lambda\}$$

Assim, para $\varepsilon > 0$ suficientemente pequeno, tem-se

$$0 \in \partial\theta(u) + \bar{B}_*(0, \sqrt{\varepsilon})$$

onde $\bar{B}_*(0, \sqrt{\varepsilon}) = \{x^* \in X^* \mid \|x^*\| \leq \sqrt{\varepsilon}\}$. (1)

Agora intentamos estimar $\partial\theta(u)$ pelo teorema 10. Para isto, devemos primeiramente mostrar que a multiaplicação

$$(h,y) \rightrightarrows \partial_x L(y,h,\rho) \tag{2}$$

é envolutiva (Veja definição 7). Notemos que para qualquer par h_1, h_2 em H, a função

$$y \to L(y,h_1,\rho) - L(y,h_2,\rho) = (h_1-h_2)\cdot(f_0,F,G)(y)$$

é localmente lipschitziana em x com constante $\rho \| h_1-h_2 \|$; assim,

$$\partial_x L(y,h_1,\rho) \subset \partial_x L(y,h_2,\rho) + \rho \| h_1-h_2 \| \bar{B}_*(0,1),$$

junto com o fato que o gradiente generalizado é uma multiaplicação fechada, temos que a multiaplicação (2) é envolutiva.

Como $\varphi(u)$ é positiva, existe um único ponto h_u em H no qual φ (e assim θ) atinge seu máximo. Agora podemos aplicar o teorema 10, e obtemos de (1):

$$0 \in \partial_x L(u, h_u,\rho) + \bar{B}_*(0,\sqrt{\varepsilon}) \tag{3}$$

Note que se $f_i(u) \leq 0$, então o maximizante h_u tem $\mu_i = 0$ necessariamente.

Se no anterior tomamos uma sequência $\varepsilon_k \downarrow 0$, então a correspondente $u_k \to \bar{x}$, e como a correspondente h_{u_k} é limitada, uma subsequência dela converge a algũm elemento de H. O teorema segue-se agora de (3) junto com o fato que a multiaplicação dada pela relação (2) é envolutiva. Δ

Comentarios. Analogamente ao caso diferenciável, para assegurar que $\lambda \neq 0$, deve-se impor alguma condição, F.H. Clarke introduz a condição de "Calm", para mais detalhes recomendamos ver o artigo de F.H.Clarke [25].

d) Programação Geométrica

Sejam $C, P \subset \mathbb{R}^n$ não vazios, P um cone, e $g: C \to \mathbb{R}$, o problema resultante da programação geométrica, problema P.G., é definido por Peterson [63], do seguinte modo:

Problema P.G.. Sobre o conjunto de soluções factiveis

$$S = C \cap P$$

achar o ĩnfimo

$$\Psi = \inf \{ g(x) \mid x \in S \}.$$

e o conjunto de soluções otimais.

$$S^* \equiv \{ x \in S \mid g(x) = \Psi \}.$$

Observações .

O problema clássico de programação matemática, é quando o cone $P = \mathbb{R}^n$. Em geral, existem várias formas para se expressar um problema de otimização particular na forma do problema P.G. Por exemplo, o potencial da programação geométrica consiste em selecionar um cone não trivial (i.e, $P \neq \emptyset$, $P \neq \mathbb{R}^n$) o qual permita explorar as linearidades presentes no problema.

Denotamos Q o cone dual de P, isto é:

$$Q = \{ y \in \mathbb{R}^n \mid < y,x > \leq 0 \quad \forall x \in P \}.$$

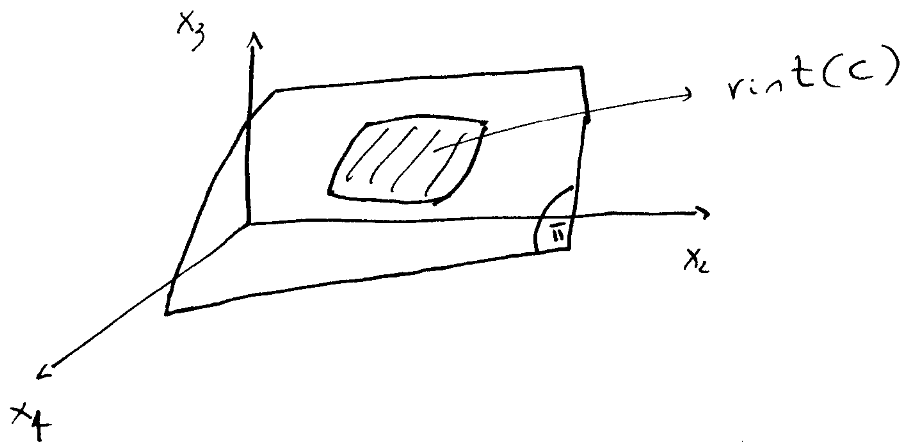
Note que $Q = - P^0$, onde P^0 é o cone polar de P, para um conjunto não vazio $D \subset \mathbb{R}^n$, usamos, a notação:

$$< x,D > = \{ < x,d > \mid d \in D\}.$$

Recordemos os seguintes fatos do Análise Convexa.

Definição 13. O interior relativo, denotado $\mathrm{rint}(C)$ de um conjunto convexo $C \subset \mathbb{R}^n$ é o interior do conjunto C relativamente à variedade afim de menor dimensão que o contém. Δ

A figura abaixo representa em $\mathbb{R}^3$ um conjunto convexo contido em um plano Π.



(rint(C) é então o interior do conjunto C considerado como subconjunto de Π isomorfo a $\mathbb{R}^2$ (sua definição utiliza vizinhanças de dimensão 2).

Os seguintes fatos podem ser provados (Veja a demonstração em Rockafellar [66]).

<u>Teorema 17</u>. Todo subconjunto não vazio de $\mathbb{R}^n$ tem um interior relativo não vazio. Δ

<u>Lema 4</u>. Para qualquer função convexa f,

$$\text{rint}(\text{epi}(f)) = \{(x,\lambda) \mid x \in \text{rint}(\text{Dom}(f) \quad \text{e } f(x) < \lambda < \infty\} \quad \Delta$$

Vamos agora introduzir um conceito de solução devido à Reiland e posteriormente dar um resultado que o relaciona com a solução de P.G.

<u>Definição 14</u>. Um vetor $x \in C$ é uma <u>solução crítica</u> (no sentido de Reiland) para o problema P.G. se g é localmente lipschitziana em x e se as seguintes condições são satisfeitas:

i) $x \in S$,

ii) $0 \in \partial g(x) - Q$ e

iii) $0 \in \langle x, \partial g(x) \rangle$.

Observação.

Se o cone P é um espaço vetorial, então $Q = P^{\perp}$; ademais, a condição $0 \in \langle x, \partial g(x) \rangle$ é redundante já que $0 \in \partial g(x) - P^{\perp}$ implica que existe um $X^* \in \partial g(x) \cap P^{\perp}$, isto é, $\langle x, x^* \rangle = 0$.

Teorema 18. Suponhamos que g é localmente lipschitziana em x

(a) Sepa P convexo. Se x é uma solução otimal para o problema P.G., então x é uma solução crítica no sentido de Reiland para o problema P.G.

(b) Seja g pseudoconvexa em x com respeito a S. Se x é uma solução crítica no sentido de Reiland para o problema P.G. e se x, C e P são tais que $(C \cap P) - x \subset P$, então x é uma solução otimal para o problema P.G.

Prova.

a) A primeira condição para uma solução crítica é satisfeita já que a otimalidade de x implica $x \in C \cap P$.

Ademais, a otimalidade de x e a convexidade de P implicam que:

$$g(x+\lambda v) \geq g(x)$$

para λ suficientemente pequeno e para cada $v \in P$.

Ademais

$$\frac{g(x+\lambda v) - g(x)}{\lambda} \geq 0$$

de onde: $g^0(x;v) \geq 0 \quad \forall\ v \in P$.

Definamos o conjunto D do seguinte modo:

$$D = (P \times \mathbb{R}) \cap T(\mathrm{epi}(g);\ (x,g(x)).$$

Recordemos que $T(\mathrm{epi}(g);\ (x,g(x))$ é o epígrafo da função convexa $v \to g^0(x;v)$ (Veja a proposição 20) D é convexo , pois $T(\mathrm{epi}(g);\ (x,g(x))$ é um cone convexo (observação 2, pág) e $P \times \mathbb{R}$ também é um cone convexo, pois P o é. Além disso,

$$\mathrm{rint}\ (P \times \mathbb{R}) = \mathrm{rint}(P) \times \mathbb{R} \qquad \text{e}$$

$$\mathrm{rint}(T(\mathrm{epi}(g);(x,g(x)) = \{(v,\alpha) \in \mathbb{R}^n \times \mathbb{R} \mid g^0(x,v) < \alpha\}$$

já que $P \subset \mathbb{R}^n$ e $P \neq \emptyset$, temos que;

$$\mathrm{rint}(P) \neq \emptyset$$

Logo, para $v \in \mathrm{rint}(P)$; $g^0(x;v) \leq K\|v\|$, onde K é a constante de lipschitz de g.

Assim, se tem:

$$(\mathrm{rint}(P) \times \mathbb{R} \cap \mathrm{rint}(T(\mathrm{epi}(g);\ (x,g(x))) \neq \emptyset$$

e

$$D^0 \doteq (P \times \mathbb{R})^0 + [\ T(\mathrm{epi}(g);(x,g(x))]^0$$

$$= -\ Q \times \{0\} + N(\mathrm{epi}(g);(x,g(x)))$$

como $g^0(x;v) \geq 0$ para cada $v \in P$, temos que $(0,-1) \in D^0$, o qual implica que

$$0 \in -\ Q + \partial g(x).$$

Logo, a segunda condição para uma solução crítica é satisfeita.

Como $x + \lambda(-x) \in P$ para $0 \leq \lambda \leq 1$, a otimalidade de x implica que:

$$g^0(x;-x) \geq 0$$

e assim,

$$-g^0(x;-x) \leq \ < x,x^* > \ \leq g^0(x,x) \quad \forall\ x^* \in \partial g(x) \quad (*)$$

Agora bem, para $x^* \in \partial g(x)$, definamos a função $h(x^*) = < x,x^* >$

Como g^0 é a função suporte do gradiente generalizado temos que existem x_1^*, $x_2^* \in \partial g(x)$, tais que

$$h(x_1^*) = g^0(x;x) = < x,x_1^* > \qquad \text{e}$$

$$-\ h(x_2^*) = g^0(x;-x) = < -x,x_2^* > ,$$

Se $h(x_1^*)$ ou $h(x_2^*)$ é zero, então a terceira condição para uma solução crítica: $0 \in \langle x, \partial g(x) \rangle$ é estabelecida.

Se $h(x_1{}^*) \neq 0$ e $h(x_2{}^*) \neq 0$, então (*) implica que:

$h(x_1{}^*)h(x_2{}^*) < 0$ e pela continuidade de h, existe $x^* \in \partial g(x)$ tal que:

$$h(x^*) = \langle x, x^* \rangle = 0$$

isto completa a prova de a).

b) A condição $0 \in \partial g(x) - Q$ implica que existe $x^* \in \partial g(x)$ e $q \in -Q$ tal que:

$$0 = x^* + q,$$

logo

$$0 = \langle z, x^* \rangle + \langle z, q \rangle \quad \forall\ z \in \mathbb{R}^n$$

Em particular, a condição $(C \cap P) - x \subset P$ implica:

$$\langle w-x, q \rangle \leq 0 \quad \text{e} \quad \langle w-x, x^* \rangle \geq 0 \qquad \forall\ w \in C \cap P.$$

Ademais:

$$g^0(x; w-x) \geq \langle w-x, x^* \rangle \geq 0 \qquad \forall w \in C \cap P,$$

o qual, pela pseudoconvexidade de g em x com respeito a $S = C \cap P$, temos que:

$$g(w) \geq g(x) \qquad \forall\ x \in C \cap P. \qquad \Delta$$

Observação. Uma condição suficiente natural para que $(C \cap P)-x \subset P$ se verifique é que P seja um subespaço de $\mathbb{R}^n$. Logo nos casos usuais de programação matemática (diferenciável ou não) $(C \cap P)-x \subset P$ é sempre verdadeira.

Aplicação: Estatística.

Problema: Para determinar a melhor L_1- aproximação de um conjunto de pontos observados $(z_1,y_1),\ldots,(z_n,y_n)$ por uma reta $y= \alpha z + \beta$, é necessário resolver o problema

$$\min_{\alpha,\beta} \sum_{i=1}^{n} |\alpha z_i + \beta - y_i| \tag{1}$$

onde $|\alpha z_i+ \beta-y_i|$ é o erro no i-ésimo ponto.

Para colocar o problema anterior no formato do problema P.G., definimos as novas variaveis $x_i= \alpha z_i+\beta$ $i=1,\ldots,n$; (1) é então equivalente a resolver o problema P.G. com $C=\mathbb{R}^n$, $g(x) = \sum_{i=1}^{n} | x_i-y_i |$, e P = espaço coluna de $[z,1]$,onde $z=(z_1,\ldots,z_n)^T$ e 1 é o vetor coluna de n 1. Assim o problema de programação geométrica equivalente a (1) é:

$$\min_{x\in P} \sum_{i=1}^{n} |x_i-y_i|$$

onde P = espaço coluna de $[z,1]$

Apliquemos o teorema anterior. Se x é otimal, então temos que:

$$x \in P, \quad 0 \in \partial g(x) - P^{\perp}$$

isto é, $x \in P$ e existe $x^* \in \partial g(x) \cap P^{\perp}$

O gradiente generalizado, $\partial g(x)$, é o conjunto seguinte.

$$\left\{ x^* = (x_1^*, \ldots, x_n^*) \in \mathbb{R}^n \;\middle|\; x_i^* = \begin{bmatrix} 1 & \text{se } x_i - y_i > 0 \\ -1 & \text{se } x_i - y_i < 0 \\ [-1,1] & \text{se } x_i - y_i = 0 \end{bmatrix} \right\} \qquad (2)$$

Com efeito: $g(x) = (\sum_{j=i}^{n} f_j)(x)$, onde

$f_j(x) = |x_j - y_j|$ e como f_j $(j=1,\ldots,n)$ são convexas, logo Clarke-regular, temos que:

$$\partial g(x) = \sum_{j=1}^{n} \partial f_j(x)$$

Mas, calcular $\partial f_j(x)$ é fácil, posto que:

$$f_j = u \circ hj$$

onde $u(t) = |t|$ $(t \in \mathbb{R})$ $h_j(x) = x_i - y_i$ e podemos aplicar a regra da cadeia 1.

Introdução.

Problemas de controle ótimo são muito frequantes em aplicações às Engenharias e geralmente aparecem sob a forma de um sistema de equações diferenciais no qual uma das variáveis (o controle) deve ser escolhido no sentido de minimizar um certo funcional que depende da solução do sistema e do controle escolhido.

Muito tem sido feito na análise de tais problemas mas geralmente sob a hipótese de regularidade das funções envolvidas. Entretanto, em muitas situações (tais como aquelas em que o sistema estudado muda de estrutura) tais hipóteses de regularidade não são verdadeiras e muita da teoría usual não pode ser usada. Um exemplo de tal situação e apresentado na seção 3.4..

Neste capítulo aplicaremos os conceitos dos capítulos anteriores a este tipo de problemas, estudaremos condições necessárias e suficientes para que os problemas do cálculo das variações e do controle ótimo, como também suas relaçoes.

3.1 Formulação do Problema.

Consideremos os seguintes problemas de otimização:

1) Problema de Lagrange

$$\text{Minimizar} \qquad J(x) = \int_a^b \tilde{L}(t,x(t),\ x(t))dt$$

$$\text{Sujeito a:} \qquad x(a) = A\ , \quad x(b) = B$$

$$g(t\ ,\ x(t),\ \dot{x}(t)) \leq 0.$$

sobre todas a funções absolutamente contínuas $x: [a,b] \to \mathbb{R}^n$. $\tilde{L}: [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $g: [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ são funções dadas, A e B são vetores constantes fixos dados. No caso $g \equiv 0$, é o problema usual do cálculo das variações.

2) Problema de Inclusão Diferencial. Dada uma multiplicação $\Gamma : [a,b] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, e um subconjunto K de $\mathbb{R}^n \times \mathbb{R}^n$. O problema proposto é:

$$\text{Minimizar} \qquad \Phi(x(b))$$

$$\text{sujeito a} \qquad (x(a),\ x(b)) \in K$$

$$\dot{x}(t) \in \Gamma(t,x(t)) \ \text{q.t.p.}$$

sôbre todos as funções absolutamente contínuas $x: [a,b] \to \mathbb{R}^n$, e $\Phi: \mathbb{R}^n \to \mathbb{R}$ é uma função dada.

3) Problema de Controle otimo: Consideremos o seguinte problema.

$$\text{Minimizar} \qquad J(x,u) = \Phi(x(b)) + \int_a^b g(t,x(t),u(t))dt$$

sujeito a :

$$\dot{x}(t) = f(t,x(t),u(t)) \qquad \text{q.t.p.}$$

$$x(a) \in C_0, \quad x(b) \in C_1 \quad u(t) \in U \quad \text{q.t.p.}$$

sobre todas as funções absolutamente contínuas $x:[a,b] \to \mathbb{R}^n$, e $u:[a,b] \to \mathbb{R}^m$ mensuraveis, onde $\Phi: \mathbb{R}^n \to \mathbb{R}$, $g: [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ $f: [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, C_0, $C_1 \subset \mathbb{R}^n$ são dados.

Todos os problemas anteriores podem ser posto sob uma mesma forma, a qual recebe o nome de Problema de Bolza Generalizado

Agora formularemos o problema de Bolza Generalizado e veremos como reescrever os problemas anteriores nestes termos, para logo deduzir algumas condições necessárias e suficientes de otimalidade.

O seguinte é uma condensação dos trabalhos de R.T. Rockofellar, F.H. Clarke e V. Zeidan.

<u>Problema de Bolza Generalizado</u> (a seguir denotado problema (P_B)).

Seja $J(x)=l(x(a),x(b))+\int_a^b L(t, x(t),\dot{x}(t))dt$

O problema de Bolza Generalizado consiste em:

Minimizar $J(x)$

onde x é uma função absolutamente contínua de $[a,b] \to \mathbb{R}^n$ com derivada $\dot{x}$ (em quase todo ponto), chamaremos a tal função x um <u>arco</u> e onde $L: [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to (-\infty, +\infty]$ e $l: \mathbb{R}^n \times \mathbb{R}^n \to (-\infty, +\infty]$ são funções dadas.

<u>Definição 1</u>. O <u>Hamiltoniano</u> do problema (P_B) é:

$$H(t,x,p) = \sup \{\langle p,v\rangle - L(t,x,v) \mid v \in \mathbb{R}^n\} \quad . \Delta \text{ (D)}$$

o primeiro a tratar este problema nesta forma foi R.Rockafellar, sob as hipotesis de que L(t, ., .) e l (.,.) são funções convexas, e, em forma mais geral, F. H. Clarke.

Formulação dos problemas anteriores na forma (P_B).

1') Consideremos o problema 1). Definamos l e L na forma seguinte:

$$L(t,w,v) = \begin{cases} \tilde{L}(t,w,v) & \text{se} \quad g(t,w,v) \leq 0 \\ +\infty & \text{se não} \end{cases}$$

e

$$l(x(a),\, x(b)) = \begin{cases} 0 & \text{se} \quad x(a) = A, \quad x(b) = B \\ +\infty & \text{se não} \end{cases}$$

2') Neste problema, pomos:

$$L(t,w,v) = \begin{cases} 0 & \text{se} \quad v \in \Gamma(t,w) \quad \text{q.t.p.} \\ +\infty & \text{se não} \end{cases}$$

$$l(x(a),x(b)) = \begin{cases} \Phi\ (x(b)) & \text{se} \quad (x(a),x(b)) \in K \\ +\infty & \text{se não} \end{cases}$$

3') Aqui fazemos

$$l(u,v) = \begin{cases} \Phi(v) & \text{se} \quad u \in C_0, \quad v \in C_1 \\ +\infty & \text{se não} \end{cases}$$

$$L(t,x,v) = \begin{cases} \operatorname{Inf}\{g(t,x,u) \mid u \in U, \ v = f(t,x,u)\} \\ +\infty & \text{se não} \end{cases}$$

<u>Definição 2</u>. Dado um arco $\bar{x}$ tal que $J(\bar{x})$ é finita.

Dizemos que $\bar{x}$ é um <u>mínimo local forte</u> para o probléma (P_B), se podemos achar um número positivo ε tal que $\bar{x}$ minimize $J(x)$ sobre todos os arcos satisfazendo

$$\| x(t) - \bar{x}(t) \| < \varepsilon \qquad \forall \ t \in [a,b]$$

se $\varepsilon = +\infty$, dizemos que $\bar{x}$ é um <u>mínimo global</u> para (P_B). Δ

Suponhamos que são dados um arco $\bar{x} : [a,b] \to \mathbb{R}^n$ e um número positivo ε. Definimos

$$V_\varepsilon(\bar{x}) = \{ x \in \mathbb{R}^n \mid \quad \| x - \bar{x}(t) \| < \varepsilon \ \text{ para algúm } t \in [a,b] \}.$$

seja $\pounds$ a coleção de subconjuntos mensuráveis segundo Lebesgue

de $[a,b]$ e $@$ os subconjuntos de Borel de $\mathbb{R}^n \times \mathbb{R}^n$.
Denotaremos por $\mathcal{L} \times @$ a σ álgebra de subconjuntos de $[a,b] \times \mathbb{R}^n \times \mathbb{R}^n$ gerada pelos produtos de conjuntos em $\mathcal{L}$ e $@$.

Suporemos o seguinte:

$$\left.\begin{array}{ll} \text{a)} & L \text{ é } \mathcal{L} \times @ \text{ mensurável} \\ \text{b)} & l \text{ é semicontínua inferior} \\ \text{c)} & \text{para cada } t \in [a,b], \text{ a função } L(t,.,.) \text{ é semicontí-} \\ & \text{nuas inferior.} \end{array}\right\} \quad (I)$$

A seguir apresentaremos um teorema de condição necessária para que o $\bar{x}$ seja um mínimo do problema (P_B), a prova deste teorema é muito longa, e, por isto, não o faremos. Ela pode ser encontrada no livro de Clarke [29] pág. 180, Teorema 4.4.1.

Consideremos o <u>caso autônomo</u> (i.e., L não tem uma dependencia explícita do tempo t). Nossas hipóteses implicaram que $L(y,v)$ é localmente lipschitziana (portanto, finita) como função de (y,v). Mais especificamente, requeremos que o crecimento de L em (y,v) seja no máximo de ordem exponencial, no seguinte sentido: existem constantes k_0, k_1, c_0 tal que para todo $(y,v) \in \mathbb{R}^n$ se tem:

(d) $\quad | \partial L(y,v) | \leq k_0 | L(y,v) | + k_1 |(y,v)| + c_0$

Ademais, suporemos que para todo $y \in \mathbb{R}^n$, tem-se

(e) $\displaystyle \lim_{||v|| \to +\infty} \frac{L(y,v)}{\| v \|} = +\infty$

(neste caso é fácil ver que o Hamiltoniano é finito).

Suponhamos além disso que l tem a forma

(f) $\quad l(S_0, S_1) = h(S_1) + \psi_{C_0}(S_0) + \psi_{C_1}(S_1)$

onde h é localmente lipschitz e C_0, C_1 são fechados.

A (d), (e) e (f) nos referiremos como hipóteses II.

ψ_A (x) denota a função indicatiz do conjunto A, isto é:

ψ_A (x) = 0 se $x \in A$ e $+\infty$ se $x \notin A$.

3.2: Condição Necessária de otimalidade

Teorema 1 . (F.H. Clarke). seja $\bar{x}$ uma solução de (P_B) sob as hipóteses feitas I e II sobre L. Então $\dot{\bar{x}}$ é essencialmente limitada, H é lipschitz sobre subconjuntos limitados, e existe um arco $\bar{p}$ tal que:

$$\begin{bmatrix} -\dot{\bar{p}}(t) \\ \\ \dot{\bar{x}}(t) \end{bmatrix} \in \quad \partial H(\bar{x}(t), \bar{p}(t)) \qquad \text{q.t.p}$$

$$\bar{p}(a) \in N(C_0; x(a)); -\bar{p}(b) \in \partial\, h(x(b)) + N(C_1; x(b))$$

$$H(\bar{x}(t), \bar{p}(t)) = \text{constante.} \qquad \Lambda$$

Quando L não possui as propriedades de crescimentos necessários para H ser finita, é muito conveniente poder-mos expressar as condições necessárias em termos de L.

Como a ilustração, é facilitar a comparação com as condições necessárias clássicas, consideremos o seguinte teorema.

Teorema 1' . Seja $\bar{x}$ que resolva P_B no caso no qual $L(y,v)$ é uma função localmente lipschitziana independente de t, e suponhamos que $\dot{\bar{x}}(t)$ é essencialmente limitada. Então existe um arco $\bar{p}$ tal que:

(1) $\dot{\bar{p}}(t) \in \partial_y L(t), \qquad \bar{p}(t) \in \partial_v L(t)$ q.t.p.

(2) $L(t) - \langle \bar{p}(t), \dot{\bar{x}}(t) \rangle =$ constante q.t.p.

$$(3) \quad L(\bar{x}(t), \dot{\bar{x}}(t) + v) \geq L(\bar{x}(t), \dot{\bar{x}}(t)) + < \bar{p}(t), v > \qquad v,\text{q.t.p.}$$

$$(4) \quad \begin{matrix} \bar{p}(a) \\ \\ \bar{p}(b) \end{matrix} \quad \in \quad \partial l(\bar{x}(a), \bar{x}(b))$$

onde $L(t) \equiv L(\bar{x}(t), \dot{\bar{x}}(t))$.

Prova. Veja ClarKe [26]. Δ

As que são familiar com as condições de primeira ordem do cálculo das variações clássico, reconhecemos em (1) a contraporte da equação de Euler-Lagrange $(d/dt) \nabla_V L = \nabla_Y L$. A primeira condição de Erdmann corresponde à continuidade de $\bar{p}$, e a segunda à equação (2). A condição necessária de Weierstrass está em (3), por outro lado (4) reflexa as condições de transversabilidade (ou fronteira normal).

Para outros resultados relacionados recomendamos ver Clarke [27].

3.3. Condições Suficientes de Otimalidade.

Proposição 1. Assumamos que L é $\mathcal{L} \times \mathcal{B}$ mensurável e que $\bar{x}$ é um arco dado, com $J(\bar{x})$ finito.

Suponhamos que exista um número $\varepsilon > 0$ positivo e uma função $W(t,x)$ definida sobre $[a,b] \times V_\varepsilon(\bar{x})$ tal que, para todos os arcos x satisfazendo

$$\|x(t) - \bar{x}(t)\| < \varepsilon \quad \forall\ t \in [a,b]$$

a função $W(.,x(.))$ é absolutamente continua e (a) $\frac{d}{dt} W(t,x(t)) -$

$$-L(t, x(t), \dot{x}(t)) \le \frac{d}{dt} W(t,\bar{x}(t)) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) \quad \text{q.t.p} ;$$

(b) para todo c,d tal que $\|c\| < \varepsilon$ $\|d\| < \varepsilon$, temos:

$$W(a, \bar{x}(a)+c) - W(a, \bar{x}(a)) + W(b, \bar{x}(b)) - W(b,\bar{x}(b)+d) \le$$
$$l(\bar{x}(a) + c, \bar{x}(b) + d) - l(\bar{x}(a), \bar{x}(b)).$$

Então $J(x)$ está bem definida (possivelmente $+\infty$) para x numa vizinhança de raio ε de $\bar{x}$, e $\bar{x}$ é um mínimo local forte para (P_B). Ademais se $\varepsilon = +\infty$ então $\bar{x}$ é um mínimo global para (P_B).

Prova. seja x(.) um arco satisfazendo

$$\|x(t) - \bar{x}(t)\| < \varepsilon \quad \forall \quad t \in [a,b].$$

Para tais $x(.)$, a condição (a) se satisfez; isto é,

$$L(t,x(t),\dot{x}(t)) \geq \frac{d}{dt} W(t,x(t)) - \frac{d}{dt} W(t,\bar{x}(t))$$

$$+ L(t, \bar{x}(t), \dot{\bar{x}}(t)) \quad \text{q.t.p.}$$

Como o lado direito da desigualdade anterior é integravel , a mensuralidade de L implica que $J(x)$ está bem definida (possivelmente $+\infty$).

De desigualdade anterior, temos:

$$\int_a^b L(t,x(t), \dot{x}(t))dt \geq \int_a^b \frac{d}{dt} W(t,x(t))\, dt - \int_a^b \frac{d}{dt} W(t,\bar{x}(t))dt$$

$$+ \int_a^b L(t,\bar{x}(t),\dot{\bar{x}}(t))dt$$

$$= W(b,x(b)) - W(a,x(a)) - W(b,\bar{x}(b)) +$$

$$W(a,\bar{x}(a)) + \int_a^b L(t,\bar{x}(t),\dot{\bar{x}}(t))dt$$

Como $\|x(a) - \bar{x}(a)\| < \varepsilon$ e $\|x(b) - \bar{x}(b)\| < \varepsilon$ por (b) temos que:

$$\int_a^b L(t,x(t),\ \dot{x}(t))dt \geq l(\bar{x}(a),\ \bar{x}(b)) + \int_a^b L(t,\bar{x}(t),\ \dot{\bar{x}}(t))dt$$

$$- l\ (x(a),x(b)),$$

ou seja, $J(x) \geq J(\bar{x})$. ; isto é $\bar{x}$ é um minimo local forte para (P_B).

Se as condições (a) e (b) se satisfazem para todos os arcos x e para todo c,d, em $\mathbb{R}^n$, neste caso, o argumento anterior implica que:

$J(x) \geq J(\bar{x})$ para todos os arcos x e assim, $\bar{x}$ é um mínimo global para (P_B). Δ.

Observações.(i) Se a função W (.,.) é lipschitziana então W(., x (.)) é absolutamente continua para qualquer arco x.

(ii) No caso que os valores de fronteira são fixos $(x(a) = A,\ x(b) = B)$,então a condição (b) é automaticamente satisfeita para todo c, d em $\mathbb{R}^n$ e para qualquer função W. Notemos que neste caso $l(x_1,\ x_2) = \psi_{\{A\}}\ (x_1) + \psi_{\{B\}}(x_2)$

onde para qualquer conjunto C, $\psi_C(x) = 0$ se $x \in C$ e $+ \infty$ no caso contrario.

A seguir damos um outro resultado de suficiência

Teorema 2. Seja a função L £ x @ mensuravel e seja $\bar{x}$ um arco tal que $J(\bar{x})$ é finita. Suponhamos que exista um número $\varepsilon > 0$ e uma função lipschitziana $W(t,x)$ definida sobre $[a,b] \times V_\varepsilon(\bar{x})$ satisfazendo:

(i) para todo c,d tal que $\|c\| < \varepsilon$ e $\|d\| < \varepsilon$

$$W(a,\bar{x}(a)+c) - W(a,\bar{x}(a)) + W(b,\bar{x}(b)) - W(b,\bar{x}(b)+d)$$

$$\leq l(\bar{x}(a)+c,\bar{x}(b)+d) - l(\bar{x}(a),\bar{x}(b));$$

(2) para $Z(t,x) = \sup\{\alpha + H(t,x,\beta) \mid (\alpha\ \beta) \in \partial W(t,x)\}$

$$Z(t,x) \leq Z(t,\bar{x}(t)) \qquad t \in [a,b] \qquad \text{q.t.p.}$$

$$Z(t,\bar{x}(t)) = \frac{d}{dt} W(t,\bar{x}(t)) - L(t,\bar{x}(t),\dot{\bar{x}}(t)) \quad \text{q.t.p.}$$

Então $J(x)$ está bem definida (possivelmente $+\infty$) para x perto de $\bar{x}$, e $\bar{x}$ é um mínimo local forte para (P_B). Ademais, se $\varepsilon = +\infty$, então $\bar{x}$ é um mínimo global para (P_B).

Observação.

Notar que $W(.,.)$ é lipschitz e $\bar{x}(.)$ é absolutamente continua, então $w(.,\bar{x}(.))$ é absolutamente continua e assim, o uso de $\frac{d}{dt}w(t,\bar{x}(t))$ na condição (2) é justificado

<u>Prova</u>. Basta provar que para qualquer arco com $\|x(t)-x(t)\| < \varepsilon$ $\forall\, t \in [\, a,b\,]$, a função W satisfaz a condição (a) da proposição 2.

Seja x um arco tal que $\|x(t)-\bar{x}(t)\| < \varepsilon$ $\forall\, t \in [a,b]$.

Consideremos o conjunto

$$G= \{t \mid \frac{d}{dt} W(t,x(t)) \text{ existe}\} \cap \{ t \mid \dot{x}(t) \text{ existe} \}$$

G tem medida de Lebesgue (b-a), pois x e W são localmente lipschitzianas. Para qualquer $t \in G$, definamos a função:

$$f_t(\tau)= W(t+\tau,\ x(t)+\tau\dot{x}(t)),$$

onde $t \in [\, 0,\ \alpha]$, para um certo $\alpha > o$. Pela regra da cadeia I capítulo II), temos

$$\partial f_t(0) \subset \partial W(t,x(t))\ (1,\dot{x}(t)).$$

Por outro lado, para qualquer $t \in G$, temos

$$x(t+\tau) = x(t) + \tau\dot{x}(t) + o(\tau)$$

onde $\lim\limits_{\tau\to 0} \dfrac{o(\tau)}{\tau}=0.$

Então, como W é localmente lipschitziana,

$$W(t+\tau,x(t+\tau))-W(t,x(t))$$

$$= W(t+\tau,\ x(t) + \tau\,\dot{x}(t))-W(t,x(t))+0(\tau).$$

Assim, para qualquer $t \in G$

$$\frac{d}{dt} W(t,\ x(t)) = \lim_{\tau \to 0} \frac{f_t(\tau) - f_t(0)}{\tau} \in \partial f_t(0) \quad ,$$

logo,

$$\frac{d}{dt} W(t,\ x(t)) \in \partial W(t,\ x(t))\ (1,\ \dot{x}(t)) \quad \text{q.t.p.}$$

Agora, da definição do Hamiltoniano dado por (D), relação anterior e a condição 2 implicam que, para quase todo $t \in [a,b]$,

$$\frac{d}{dt} W(t,\ x(t)) - L(t,\ x(t),\ \dot{x}(t))$$

$$\leq Z(t,\ x(t)) \leq Z(t,\ \bar{x}(t))$$

$$\frac{d}{dt} W(t,\ \bar{x}(t)) - L(t,\ \bar{x}(t),\ \dot{\bar{x}}(t)).$$

Isto é, a condição (a) da proposição 1 é satisfeita. Δ

<u>Corolário 1</u>. Seja L $\mathcal{L} \times \mathcal{B}$ mensurável e $\bar{x}$ um arco com $J(\bar{x})$ finita. Suponhamos que exista um $\varepsilon > 0$ e uma função C^1

$\bar{p}: [a,b] \to \mathbb{R}^n$ tal que $l(.,.)$ e $-H(t,.,p)$ são convexas respectivamente sobre $V_\varepsilon(\bar{x}(a)) \times V_\varepsilon(\bar{x}(b))$ e $V_\varepsilon(\bar{x})$. Além disso, assumamos:

(a) $L(t, \bar{x}(t), \dot{\bar{x}}(t)+v) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) \geq \langle \bar{p}(t), v \rangle$

para todo $v \in \mathbb{R}^n$ e para quase todo $t \in [a,b]$,

(b) $(-\dot{\bar{p}}(t), \dot{\bar{x}}(t)) \in \partial H(t, \bar{x}(t), \bar{p}(t))$ q.t.p.

(c) $(\bar{p}(a), -\bar{p}(b)) \in \partial l(\bar{x}(a), \bar{x}(b))$.

Então as condições do teorema anterior se satisfazem.

<u>Observação</u>.i) O conjunto $\partial H(t, \bar{x}(t), \bar{p}(t))$ é definido neste caso como o produto dos subgradientes em $\bar{x}(t)$ da função convexa $-H(t,.,p(t))$ com os subgradientes em $\bar{p}(t)$ da função convexa $H(t, \bar{x}(t), .)$. Similarmente, ∂l é o conjunto dos subgradientes da função convexa $l(.,.)$.

ii) Em [90] e [26] F.H. Clarke mostra que uma condição necessária para a otimalidade de $\bar{x}$ é a condição de Weierstrass generalizada, isto é, existe uma função $\xi(t)$ tal que $\bar{p}(t) = \xi(t)$ satisfaz a condição (a). Também, Clarke a mostrado em [27] que a existência de um arco $\bar{p}$ satisfazendo as inclusões Hamiltonianas, i.e., condição (b), e a condição de transversabilidade, i.e., a condição (c) é uma condição necessária para a otimalidade de $\bar{x}$. Assim, o corolario é obtido confirmando as condições necessárias.

Prova. Definamos sobre $[a,b] \times V_\varepsilon(\bar{x})$ a função $W(t,x) = \langle \bar{p}(t), x \rangle$. Temos que $W(.,.)$ é uma função continuamente diferenciável, e assim pelo Teorema 2 do capítulo II,

$$\partial W(t,x) = \nabla W(t,x) = (\langle \dot{\bar{p}}(t), x \rangle, \bar{p}(t)).$$

Assim a função $Z(t,x)$ da condição (2) do teorema 2 é dada por:

$$Z(t,x) = \langle \dot{\bar{p}}(t), x \rangle + H(t,x, \bar{p}(t)).$$

A condição (a) implica que:

$$L(t,\bar{x}(t),w) - L(t,\bar{x}(t), \dot{\bar{x}}(t)) \geq \langle \bar{p}(t), w-\bar{x}(t)) \ \forall \ w$$ o qual implica:

$$\langle \bar{p}(t), \dot{\bar{x}}(t) \rangle \geq \sup \{\langle \bar{p}(t), w \rangle - L(t,x,w) \mid w \in \mathbb{R}^n \}$$

Logo:

$$H(t,\bar{x}(t),\bar{p}(t)) = \langle \bar{p}(t), \dot{\bar{x}}(t) \rangle - L(t,\bar{x}(t),\dot{\bar{x}}(t)) \text{ q.t.p.}$$

o que nos dá:

$$Z(t, \bar{x}(t)) = \frac{d}{dt} W(t, \bar{x}(t)) - L(t,\bar{x}(t), \dot{\bar{x}}(t)) \text{ q.t.p.}$$

Por outro lado, (b) e a convexidade de $-H(t,.,\bar{p}(t))$ implicam que:

$-H(t,x,p) + H(t,\bar{x}(t), \bar{p}(t)) \geq \langle \dot{\bar{p}}(t), x - \bar{x}(t)\rangle - \langle \dot{\bar{x}}(t), p - \bar{p}(t)\rangle$

para todo $x \in V_\varepsilon(\bar{x})$ e para todo p. Em particular para $p = \bar{p}(t)$, assim:

$$-H(t,x,\bar{p}(t)) + H(t,\bar{x}(t),\bar{p}(t)) \geq \langle \dot{\bar{p}}(t), x-\bar{x}(t)\rangle$$

o equivalentemente

$$Z(t, \bar{x}(t)) = \langle \dot{\bar{p}}(t), \bar{x}(t) \rangle + H(t, \bar{x}(t), \bar{p}(t))$$

$$\geq \langle \dot{\bar{p}}(t), x \rangle + H(t,x, \bar{p}(t)) = Z(t, x)$$

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

Analogamente, fazendo uso da condição (c) e a convexidade de $l(.,.)$, temos que:

$$l(\bar{x}(a) + c, \bar{x}(b) + d) - l(\bar{x}(a), \bar{x}(b))$$

$$\geq \langle \bar{p}(a), c \rangle - \langle \bar{p}(b), d \rangle.$$

Além disso, W satisfaz as condições (1) e (2) do teorema 2. Δ

Teorema 3 (V. Zeidan [79]). Assumamos que L é £ x@ mensurável e que $\bar{x}$ é um arco dado com $J(\bar{x})$ finito. Suponhamos que existe um número positivo ε, um arco $\bar{p}$ e uma função absolutamente continua $Q(t)$, com valores no espaço das matrízes reais $n \times n$ simétricas $\forall t$, tal que:

(i) $L(t,\bar{x}(t),\dot{\bar{x}}(t)+v) - L(t,\bar{x}(t),\dot{\bar{x}}(t)) \geq \langle \bar{p}(t),v \rangle$ para quase todo $t \in [a,b]$ e para todo $v \in \mathbb{R}^n$;

(ii) $$H(t,x,\bar{p}(t) - Q(t)(x-\bar{x}(t)) - H(t,\bar{x}(t),\bar{p}(t))$$

$$\leq -\langle \dot{\bar{p}}(t),x-\bar{x}(t)\rangle - \langle \dot{\bar{x}}(t),Q(t)(x-\bar{x}(t))\rangle$$

$$+ \frac{1}{2}\langle x-\bar{x}(t), \dot{Q}(t)(x-\bar{x}(t))\rangle$$

para quase todo $t \in [a,b]$ e para todo $x \in V_\varepsilon(\bar{x})$:

(iii) para todo c,d com $\|c\| < \varepsilon$, $\|d\| < \varepsilon$,

$$l(\bar{x}(a)+c, \bar{x}(b)+d) - l(\bar{x}(a),\bar{x}(b)) \geq$$

$$\langle \bar{p}(a),c\rangle - \langle \bar{p}(b),d\rangle - \frac{1}{2}\langle c,Q(a)c\rangle + \frac{1}{2}\langle d,Q(b)d\rangle.$$

Então $J(x)$ esta bem definido (possivelmente $+\infty$) para x numa vizinhança de raio ε de $\bar{x}$, e $\bar{x}$ é um mínimo local forte para (P_B). Se $\varepsilon=+\infty$ então $\bar{x}$ é um mínimo global para (P_B).

Prova. Para as funções $\bar{x}$, $\bar{p}$ e Q, definimos:

$$W(t,x) = < \bar{p}(t),x > - \frac{1}{2} <x-\bar{x}(t), Q(t)(x-\bar{x}(t))>$$

como função sobre $[a,b] \times V_{\varepsilon}(\bar{x})$.

Vamos mostrar que esta função W satisfaz as condições da proposição 1. Primeiramente observemos que, para qualquer arco x tal que $\| x(t) - \bar{x}(t) \| < \varepsilon$, a função $W(., x(.))$ e absolutamente continua com derivada.

$$\frac{d}{dt} W(t,x(t)) = < \dot{\bar{p}}(t), x(t) > + < \bar{p}(t), \dot{x}(t) >$$

$$- < \dot{x}-\dot{\bar{x}}(t), Q(t)(x-\bar{x}(t)) > - \frac{1}{2} < x(t)-\bar{x}(t), Q(t)(x-\bar{x}(t) > \quad \text{q.t.p.}$$

Usando a condição (i) do teorema e a definição do Hamiltoniano, obtemos

$$H(t,\bar{x}(t),\bar{p}(t)) = < \bar{p}(t), \dot{\bar{x}}(t) > - L(t,\bar{x}(t), \dot{\bar{x}}(t)) \text{ q.t.p.}$$

e

$$H(t,x(t),\bar{p}(t) - Q(t)(x(t)-\bar{x}(t))) \geq$$

$$< \dot{\bar{x}}(t), \bar{p}(t)-Q(t)(x(t)-\bar{x}(t)) > - L(t,x(t), \dot{x}(t) \text{ q.t.p.}$$

das relações anteriores e da condição (ii) do teorema temos que para qualquer arco x com $\| x(t) - \bar{x}(t) \| < \varepsilon \quad \forall\, t \in [a,b]$

$$\frac{d}{dt} W(t,x(t)) - L(t, x(t), \dot{x}(t)) - \frac{d}{dt} W(t,\bar{x}(t)) +$$

$$L(t,\bar{x}(t),\dot{\bar{x}}(t)) \leq H(t,x(t),\bar{p}(t)-Q(t)(x(t)-\bar{x}(t)))$$

$$-H(t,\bar{x}(t),\bar{p}(t)) + < \dot{\bar{p}}(t),x(t)-\bar{x}(t) >+$$

$$< \dot{\bar{x}}(t), Q(t) (x(t)-\bar{x}(t)) > - \frac{1}{2} < x(t)-\bar{x}(t),Q(t)(x(t)-\bar{x}(t)) \leq 0$$

q.t.p.

Assim a condição (a) da proposição 1 é satisfeita.

Agora, usando a definição de W, e a condição (iii) do teorema, obtemos que, para qualquer c,d com $||c|| < \varepsilon$ e $||d|| < \varepsilon$

$$W(a, \bar{x}(a)+c) - W(a,\bar{x}(a)) + W(b,\bar{x}(b)) - W(b,\bar{x}(b)+d)$$

$$= < \bar{p}(a), c > - < \bar{p}(b),d > - \frac{1}{2} <c, Q(a)c > + \frac{1}{2} < d,Q(b)d) >$$

$\leq l(\bar{x}(a)+c,\bar{x}(b)+d) - l(\bar{x}(a),\bar{x}(b))$ o qual termina a prova graças à proposição 1. Δ

<u>Exemplo</u>. Consideremos o seguinte problema

$$\text{Minimizar} \quad J(x) = \int_0^1 \{ \frac{x^2(t)}{4} \quad \text{sen} \quad \frac{1}{x(t)} + \dot{x}^2(t) \} \, dt \qquad P)$$

sujeito a $x(0) = x(1) = 0$.

O Hamiltoniano correspondente a este problema é

$$H(x,p) = \sup \{ pv - \frac{x^2}{4} \quad \text{sen} \, \frac{1}{x} - v^2 \mid v \in \mathbb{R} \} = \frac{p^2}{4} - \frac{x^2}{4} \quad \text{sen} \, \frac{1}{x}$$

Seja o arco $\bar{x} \equiv 0$ em $[0,1]$. Então $\bar{x}$ satisfaz as condições de fronteira. Vamos usar o teorema 3 para mostrar que $\bar{x}$ é um ótimo para P).

Como os valores de fronteira são fixos, temos que a condição (iii) do teorema 4 é satisfeita para qualquer $\bar{p}$ e qualquer Q.

Agora, seja $\bar{p} \equiv 0$. É claro que a condição (i) do mesmo teorema se satisfaz. Neste caso a condição (ii) se reduz a achar uma função absolutamente continua $q(t)$, satisfazendo , para $t \in [0,1]$ q.t.p., e para todo $x \in \mathbb{R}$,

$$q^2 - \text{sen} \, \frac{1}{x} \leq 2 \dot{q} .$$

Definimos $q_0(t) = \tan \frac{t}{2}$, que é solução da equação diferencial $1 + q^2 = 2 \dot{q}$. Assim,

$$q_0^2 (t) - \text{sen} \, \frac{1}{x} \leq q_0^2 (t) + 1 = 2 \dot{q}_0(t)$$

e $q_0(t)$ satisfaz a condição (ii) do teorema,

Portanto $\bar{x} \equiv 0$ é um mínimo global para P).

Nota. O Hamiltoniano H neste exemplo não é côncavo em x e também não diferenciável em $\bar{x} \equiv 0$, pelo qual as técnicas clássicas não são aplicaveis.

Vamos agora especializar alguns dos resultados anteriores para problemas de controle ótimo, isto é,

$$\left.\begin{array}{ll} \text{Minimizar} \quad \Phi(x,u) = l^0(x(b)) + \int_a^b g(t,x(t),u(t))\,dt & \\ \text{sujeito a} & \\ \dot{x}(t) = f(t,x(t),u(t)) \text{ q.t.p.} & (1) \\ x(a) = A & (2) \\ U(t) \in U \text{ q.t.p.} & (3) \end{array}\right\} (P_c)$$

onde $U \subset \mathbb{R}^m$ é fechado, $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ $g: [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$; $l^0 : \mathbb{R}^n \to \mathbb{R}$ são funções dadas; onde $x: [a,b] \to \mathbb{R}^n$ é absolutamente contínua com derivada $\dot{x}$ (q.t.p.) e $u: [a,b] \to \mathbb{R}^m$ é mensuravel

Definição 3. O Hamiltoniano para o problema (P_c) é definido por

$$H(t,x,p) = \sup \{ \langle p, f(t,x,u) \rangle - g(t,x,u) \mid u \in U \}. \qquad \Delta$$

<u>Definição 4</u>. Sejam $x: [a,b] \to \mathbb{R}^n$ absolutamente continua e $u: [a,b] \to \mathbb{R}^m$ mensuravel. O par (x,u) diz-se <u>admissivel</u> para (P_c) se ele satisfaz (1), (2), (3). Δ

<u>Definição 5</u>. Seja $(\bar{x},\bar{u})$ um par admissivel. Dizemos que $(\bar{x}, \bar{u})$ é um <u>mínimo local forte</u> para (P_c) se existe $\varepsilon > 0$ tal que $(\bar{x}, \bar{u})$ minimize $\Phi(x,u)$ sobre todos os pares admissiveis tais que $\| x(t) - \bar{x}(t) \| < \varepsilon \;\; \forall\, t \in [a,b]$.

se $\varepsilon = +\infty$, dizemos que $(\bar{x},\bar{u})$ é um mínimo global para (P_c).

Δ

Faremos as seguintes hipóteses.

Dado um par $(\bar{x},\bar{u})$ e $\varepsilon > o$, f e g são mensuraveis sobre $[a,b] \times V_\varepsilon(\bar{x}) \times U$ e, para cada t, $g(t,.,.)$ é semicontinua inferior e $f(t,.,.)$ é continua sobre $V_\varepsilon(\bar{x}) \times U$ (H)

<u>Condições Suficientes de otimalidade para</u> (P_c)

A idéia básica para achar condições suficientes de otimalidade para o problema de controle (P_c) é formulá-lo como um problema de Bolza Generalizado (P_B) e interpretar os resultados de V. Zeidan dados anteriormente. Existem varias maneiras para escrever (P_C) na forma (P_B). Uma delas é a seguinte:

Colocamos

$$L(t,x,v)=\begin{cases} \inf\ \{ g(t,x,u) \mid f(t,x,u)=v,\ u\in U \} & (E_1) \\ +\ \infty \qquad \text{se não} & \end{cases}$$

$$l(S_0,\ S_1)=\begin{cases} l^0(S_1) \quad \text{se } S_0\in C_0,\ S_1\in C_1 & (E_2) \\ +\ \infty \qquad \text{se não} & \end{cases}$$

(Notar que como é usual o infimo sobre o conjunto vazio é $+\ \infty$). É razoavel esperar que um arco $\bar{x}$ resolva o problema (P_B) associado se e somente se, para algum controle correspondente $\bar{u}$, $(\bar{x},\bar{u})$ resolve (P_C). Sob hipóteses fracas, isto é verdade, como mostra o seguinte teorema de R.T.Rockafellar.

(Omitiremos a prova, a qual esta baseada em argumentos da Teoria de seleções mensuráveis).

O seguinte resultado é uma simplificação de um resultado mais geral de R. Rockafellar.

Teorema 4. (Rockafellar 1975)

suponhamos que os dados do Problema (P_c) satisfazem as seguintes condições :

i) $g(t,x,u)$ é $\mathcal{L} \times \mathcal{B}$ mensuravel e s.c.i em (x,u).

ii) $f(t,x,u)$ é mensuravel em t e continua em (x,u).

iii) U é fechado, e $\{(t,u,(t)) \mid t \in [a,b], u(t) \in U\}$ é $\mathcal{L} \times \mathcal{B}$ mensuravel

iv) C_0, C_1 são fechados; f é semicontinua inferior em x.

v) Para cada t, para todo subconjunto limitado S de $\mathbb{R}^n \times \mathbb{R}^n$, o seguinte conjunto é limitado

$$\{ u \in U \mid \text{para algum } (x,v) \text{ em } S, \ v = f(t,x,u) \}$$

Então os dados, l, L do problema de Bolza associado, (P_B), satisfazem as hipóteses I, e um arco $\bar{x}$ resolve (P_B) se e somente se existe um controle $\bar{u}$ correspondente a $\bar{x}$ tal que $(\bar{x},\bar{u})$ resolve (P_c).

Prova Veja Rockafellar [68], equivalence Theorem, pág 316. Δ

Relação entre o Hamiltoniano de (P_B) e (P_C)

Recordemos que o Hamiltoniano para (P_B) é definido por:

$$H(t,x,p) = \sup\{< p,v > - L(t,x,v) \mid v \in \mathbb{R}^n\}$$

já vimos na seção anterior que o problema (P_C) é "equivalente" a minimizar $l(S_0, S_1) + \int L(t,x,\dot{x})dt$ onde L e l são dados pelas equações (E_1), (E_2).

Substituindo este lagrangeano, temos

$$H(t,x,p) = \sup\{< p, f(t,x,u)> - g(t,x,u) \mid u \in U\}$$

Assim este é o verdadeiro Hamiltoniano para (P_C); notar ademais que no caso clássico, considera-se o seguinte Hamiltoniano.

$$H_C(t,x,p,u) = < p, f(t,x,u)> - g(t,x,u)$$

chamado Hamiltoniano de Pontryagin, e assim temos que:

$$H(t,x,p) = \sup\{H_C(t,x,p,u) \mid u \in U\}$$

Teorema 5 . (V. Zeidan [80]). Seja $(\bar{x},\bar{u})$ um par admissível para (P_C) tal que $\Phi(\bar{x},\bar{u})$ é finita. Assumamos que para algum $\varepsilon > 0$ a hipótese (H) se satifaz e que exista um arco $\bar{p}$ é uma função absolutamente continua $Q(t)$, com valores no espaço das matrizes reais n x n simétricas tal que:

(a) $g(t,x,u) - g(t,\bar{x}(t), \bar{u}(t)) -$

$$< \bar{p}(t)-Q(t)(x-\bar{x}(t)),\ f(t,x,u)-f(t,\bar{x}(t),\bar{u}(t)) >$$

$$\geq \ <\dot{\bar{p}}(t),\ x-\bar{x}(t)> - \frac{1}{2}< x-\bar{x}(t),\dot{Q}(t)(x-\bar{x}(t)) >$$

para quase todo $t \in [a,b]$, para todo $x \in V_\varepsilon(\bar{x})$ e para todo $u \in U$;

(b) $l^0(\bar{x}(b)+d) - l^0(\bar{x}(b)) \geq - <\bar{p}(b),d> + \frac{1}{2}<d, Q(b)\,d>$

para todo d com $||d|| < \varepsilon$. Então $(\bar{x},\bar{u})$ é um mínimo local forte para (P_C). Ademais, se $\varepsilon = +\infty$, $(\bar{x},\bar{u})$ é um mínimo global.

<u>Prova</u>. A idéia da prova é converter o problema (P_C) no problema de Bolza Generalizado (P_B), e assim aplicar o teorema 4.

Definamos as seguintes funções

$$l(x_1\ x_2) = \psi_{\{A\}}(x_1) + l^0(x_2)$$

onde $\psi_{\{A\}}(x) = 0$ se $x = A$ e $+\infty$ no caso contrario, e $\quad L(t,x,v) = \inf \{ K(t,x,v,u) \mid u \in \mathbb{R}^n \}$ (1)

onde $K(t,x,v,u) = g(t,x,u)$ se $u \in U$ e $v = f(t,x,u)$ e $+\infty$ no caso contrário.

Consideremos o problema de Bolza Generalizado, (P_B), associado ao problema de controle ótimo:

$$\text{Minimizar } J(x) = l(x(a), x(b)) + \int_a^b L(t,x(t), \dot{x}(t))\, dt.$$

O problema (P_C) e (P_B) anterior tem o mesmo Hamiltoniano.

A condição (a) do teorema implica que

$$< \bar{p}(t), \; f(t,\bar{x}(t),\bar{u}(t) > - g(t,\bar{x}(t), \bar{u}(t)) \tag{2}$$

$$\geq \; < \bar{p}(t), \; f(t,\bar{x}(t),u > - g(t,\bar{x}(t), u)$$

(basta avaliar (a) em $x = \bar{x}(t)$, para todo $u \in U$, e para quase todo $t \in [a,b]$).

Assim, usando a equação (1) obtemos

$$L(t, \bar{x}(t), \dot{\bar{x}}(t)) = g(t,\bar{x}(t), \bar{u}(t)) \quad \text{q.t.p.} \tag{3}$$

por outro lado, (2) diz que, para qualquer par admissível (x,u)

$$\Phi(\bar{x},\bar{u}) = J(\bar{x}) \leq J(x) \leq \Phi(x,u).$$

Notemos que, do anterior, para mostrar que $(\bar{x},\bar{u})$ é um mínimo local (resp. global) forte para (P_C) é suficiente mostrar que $\bar{x}$ é um minimo local (resp.global) forte para (P_B).

Para provar este último fato faremos uso do Teorema 4 . Verifiquemos então as hipóteses do teorema mencionado.

Devemos testar a $\mathcal{L} \times @$ mensuravilidade de L.
Para isto basta provar que $K(.,.,.,.)$ e $\mathcal{L} \times \tilde{@}$ mensuravel (onde, $\mathcal{L}$ é a coleção de subconjuntos mensuráveis segundo Lebesgue e $\tilde{@}$ os subconjuntos de Borel de $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$, e como é usual $\mathcal{L} \times \tilde{@}$ a σ- álgebra de subconjuntos de $[a,b] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ gerada pelo produto de conjuntos em $\mathcal{L}$ e $\tilde{@}$) e que $K(t,.,.,.)$ é semicontínua inferior, em virtude do teorema 1 de R.T. Rockafellar [67] .

Definamos $G =[a,b] \times V_\varepsilon(\bar{x}) \times U$ e para cada $t \in [a,b]$ e $\alpha \in \mathbb{R}$

$$G_{t,\alpha} = \{ (x,u) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \in V_\varepsilon(\bar{x}) ,u \in U, g(t,x,u) \leq \alpha \}$$

Temos que G é $\mathcal{L} \times @$ mersuravel, pois $V_\varepsilon(\bar{x})$ é aberto e

U é fechado.

Por outro lado pela s.c.i. de $g(t,.,.)$ temos que $G_{t,\alpha}$ é fechado. Os seguintes conjuntos

$a_1)$ $\{ (t,x,v,u) \mid (t,x,u) \in G,\ g(t,x,u) \leq \alpha\ f(t,x,u) - v = 0 \}$

$b_1)$ $\{ (x,v,u) \mid (x,u) \in G_{t,\alpha}\ ,\ f(t,x,u) - v = 0 \}$ $t \in [a,b]$

São £ x $\tilde{@}$ mensuraveis e fechados respectivamente, já que $f(t,.,.)$ é continua sobre $G_{t,\alpha}$ (por hipóteses) e g: $G \to \mathbb{R}$ e f: $G \to \mathbb{R}^n$ são mensuráveis. Logo temos que $K(t,x,v,u)$ é s.c.i. em (x,v,u) e £ e $\tilde{@}$ mensuravel em (t,x,v,u). Além disso, também temos que $J(\bar{x})$ é finita.

Notemos ademais que (2) e (3) implicam que a condição (i) do teorema 4 se satisfaz, e que

$$H(t,\bar{x}(t),\ \bar{p}(t)) = \langle \bar{p}(t),\ f(t,\bar{x}(t),\ \bar{u}(t) \rangle -$$

$$- g(t,\ \bar{x}(t),\ \bar{u}(t)) \text{ q.t.p.}$$

Da definição do Hamiltoniano para (P_C) e a condição (a) do teorema, deduzimos que:

$$H(t,x,\bar{p}(t) - \Omega(t)\ (x-\bar{x}(t))) - H(t,\ \bar{x}(t),\ \bar{p}(t))$$

$$\leq \; - < \dot{\bar{p}}(t),\; x-\bar{x}(t) > + \frac{1}{2} < x-\bar{x}(t),\; \dot{Q}(t)\;(x-\bar{x}(t)) >$$

$$- < \dot{\bar{x}}(t),\; Q(t)\;(x-\bar{x}(t)) >,$$

para $t \in [a,b]$ q.t.p. e $\forall\; x \in V_\varepsilon(\bar{x})$. Logo a condição (ii) do teorema 4 está satisfeita. Só nos resta verificar a condição (iii) do teorema 4 o que é obtido pela condição b) e da definição de l : $l(x_1, x_2) = \psi_{\{A\}}(x_1) + l^0(x_2)$. Logo $\bar{x}$ é um mínimo local forte para (P_c) e, quando $\varepsilon = +\infty$, $\bar{x}$ é um mínimo global. Δ

O seguinte resultado mostra que não é necessário verificar a condição (a) do teorema 5 para todo $u \in U$; é suficiente verifica-la em uma vizinhança de $\bar{u}$, se certas outras condições forem satisfeitas.

<u>Corolário</u>. Seja $(\bar{x}, \bar{u})$ admisível para (P_c) com $\Phi(\bar{x},\bar{u})$ finita.

Assumamos que U é compacto e que, para algúm $\gamma > 0$, f e g são contínuas sobre $[a,b] \times V_\gamma(\bar{x}) \times U$. Suponhamos que exista um arco $\bar{p}$ e uma função absolutamente continua $Q(t)$ com valores no espaço das matrizes reais nxn simétricas tal que a condição b) do teorema 5 se satisfaça e

(i) condição (a) do teorema 5 se satisfaça para $t \in [a,b]$ em quase todo ponto, para todo $x \in V_\gamma(\bar{x})$ e para todo $u \in V_\gamma(\bar{u}) \cap U$

(ii) para todo $t \in [a,b]$ e para todo $u \in U$ com $u \neq \bar{u}(t)$

$$< \bar{p}(t),\ f(t,\bar{x}(t),\ u)) > - \quad g(t,\ \bar{x}(t),\ u)$$

$$< \quad < \bar{p}(t),\ f(t,\bar{x}(t),\ \bar{u}(t)) > - \ g(t,\ \bar{x}(t),\ \bar{u}(t));$$

(iii) para todo $t \in [a,b]$, a função

$$< p,\ f(t,x,.) > - \quad g(t,x,.)$$

é estritamente convexa sobre $V_\gamma(\bar{u})$ para todo $(x,p) \in N_\gamma(\bar{x},\bar{p})$.
Então $(\bar{x},\bar{u})$ é um mínimo local forte para (P_C).

<u>Prova</u>. Pelo teorema 5, é suficiente mostrar que a condição (a) do teorema é satisfeita para todo $u \in U$.
Definamos

$$h(t,x,p,u) = < p,f(t,x,u) > - \ g(t,x,u)$$

Como uma função de $[a,b] \times V_\gamma(\bar{x},\bar{p}) \times U$.

Como U é compacto, e f e g são continuas, então o máximo de H

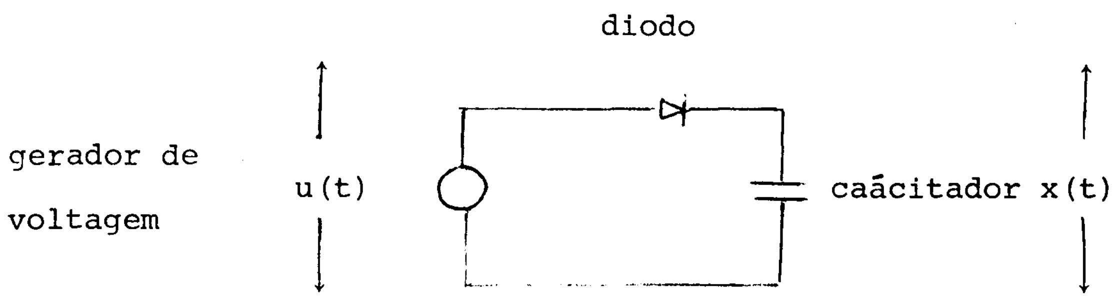
$$H(t,x,p) = \max \ \{ h(t,x,p,u) \mid u \in U \}$$

é sempre atingido. Temos que a condição (ii) e (iii) implicam que existe δ, $0 < \delta \leq \gamma$, tal que para todo $t \in [a,b]$ e para $(x,p) \in V_\delta(\bar{x},\bar{p})$ o máximo de h é atingido em um único valor $u(t,x,p)$ com $u(t,\bar{x}(t),\bar{p}(t)) = \bar{u}(t)$. Ademais a função $u(t,x,p)$ é continua. Assim, podemos encontrar um número positivo $\varepsilon \leq \delta$, tal que, para todo t e para todo $(x,p) \in V_\varepsilon(\bar{x},\bar{p})$, temos $u(t,x,p) \in V_\varepsilon(\bar{u}) \cap U$.

Ademais a condição (i) implica que a condição (a) do teorema é satisfeita para todo $u \in U$. Δ

3.4 Aplicação: Regulador Linear com Diodo.

Consideremos o problema de se obter o ponto de operação ótimo de um circuito regulador linear com diodo: se $x(t)$ é a voltagem, no instante t, a que se encontra submetido o capacitador no circuito representado na figura abaixo,



pode-se mostrar (Veja McClamroch [58]) que a equação diferencial que governa o circuito é:

$$\dot{x}(t) = \begin{cases} \alpha(u(t)-x(t)) & \text{se} \quad x(t) \leq u(t) \\ -\beta(x(t)-u(t)) & \text{se} \quad x(t) \geq u(t) \end{cases}$$

onde α e β são constantes positivas (com $\alpha > \beta$). Quer se escolher $u(\cdot)$ em $[0,T]$ que minimize

$$\frac{1}{2}\int_0^T u^2(t)\,dt$$

e tal que o estado x satisfaça condições de contorno dadas $x(0) = x_0$ e $x(T) = x_T$

O problema de controle ōtimo é

$$\text{Minimizar} \quad \int_0^T \frac{u^2(t)}{2}\, dt$$

sujeito a:

$$(P_c)$$

$$\dot{x}(t) = \max\{\alpha(u(t)-x(t)),\ \beta(u(t)-x(t))\}$$

$$x(0) = x_0 \quad ; \quad x(T) = x_T$$

$$u(t) \in \mathbb{R}$$

Para poder aplicar o teorema 1 (condições necessārias) devemos reescrever nosso problema na forma (P_B).

Como $f(x,u) = \alpha(u-x)$ se $u \leq x$; chamando $\dot{x}=v$, temos que: $u = v/\alpha + x$ se $v \geq 0$, logo

$$L(x,v) = \frac{(x+v/\alpha)^2}{2} \quad \text{se} \quad v \geq 0 \tag{1}$$

Analogamente

$$L(x,v) = \frac{(x + v/\beta)^2}{2} \quad \text{se} \quad v \leq 0 \tag{2}$$

e definimos

$$l(s_0,s_1) = \begin{cases} 0 & \text{se} \quad s_0 = x_0, \quad s_1 = x_T \\ +\infty & \text{no outro caso} \end{cases} \tag{3}$$

já que $C_0 = \{x_0\}$, $C_1 = \{x_T\}$.

Logo o problema (P_B) associado (formalmente) a (P_C) é:

$$\text{Minimizar} \quad \int_0^T L(x,v)dt + l(s_0,s_1) \qquad (P_B)$$

com L e l definidos pelas relações (1), (2) e (3).

Além disso como os dados do problema (P_c) satisfazem trivialmente as hipóteses do teorema 4, temos que os dados do problema (P_B) satisfazem as hipoteses I, e um arco $\bar{x}$ resolve (P_B) se e somente se existe um controle $\bar{u}$ correspondente a $\bar{x}$ tal que $(\bar{x},\bar{u})$ resolve (P_c).

Então, agora estamos em condições para fazer uso do teorema 1. Vejamos primeiramente se L e $\partial L(x,v)$ satisfazem as hipoteses II, isto é, se:

i) $\quad ||\partial L(x,v)|| \leq k_0 \; |L(x,v)| + k_1 ||(x,v)|| + C_0 \qquad \forall\; x,v \in \mathbb{R}.$

ii) $\quad \lim\limits_{|v|\to\infty} \dfrac{L(x,v)}{|v|} = +\infty$

A condição ii) é claramente satisfeita. Provemos i). Calculemos primeiramente $\partial L(x,v)$. Para $v \neq 0$, $L(x,v)$ é diferenciavel, de onde

$$\partial L(x,v) = \{\nabla L(x,v)\} \;,$$

e para v=0, temos que:

$$\partial L(x,v) = co \{\lim \nabla L(y,v) \mid (y,v) \to (x,0)\} .$$

Assim:

$$\partial L(x,v) = \begin{cases} (x + v/\alpha,\ (x + v/\alpha)/\alpha) & \text{para } (x,v) \text{ com } v > 0 \\ (x + v/\beta,\ (x + v/\beta)/\beta) & \text{para } (x,v) \text{ com } v < 0 \\ co\{(x,x/\alpha),(x,x/\beta)\} & \text{para } (x,0). \end{cases}$$

Para $v > 0$, i), traduz-se em provar que:

$$\|\nabla L(x,v)\| \leq k_0 \mid L(x,v) \mid + k_1 \|(x,v)\| + C_0,$$

ou;

$$\sqrt{(x+v/\alpha)^2 + (x+v/\alpha)^2/\alpha^2} \leq k_0 (x+v/\alpha)^2/2 + k_1 \|(x,v)\| + C_0.$$

Mas o lado esquerdo da expressão anterior é igual a:

$$|(x+v/\alpha)| \sqrt{1+1/\alpha^2} = |(x+v/\alpha)| \sqrt{(\alpha^2+1)}/\alpha$$

$$\leq (x+v/\alpha)^2 \sqrt{\alpha^2+1}/\alpha$$

e portanto basta tomar $k_0^1 = 2\sqrt{\alpha^2+1}/\alpha$ e $C_0 \geq 0$, $k_1 \geq 0$ arbitrários .

Para $v < 0$, com um raciocínio análogo, chegamos a que $k_0^2 = 2\sqrt{\beta^2+1}/\beta$ e $C_0 \geq 0$, $k_1 \geq 0$ arbitrários.

Agora, vejamos o caso $v=0$. Notemos que se $(z,w) \in \partial L(x,v)$, então:

$$(z,w) = \lambda(x,x/\alpha)+(1-\lambda)(x,x/\beta), \qquad 0 \leq \lambda \leq 1$$

o qual implica que:

$$z = \lambda x+(1-\lambda)x \quad => \quad z = x$$

$$w = \lambda x/\alpha + (1-\lambda)x/\beta => w = x(\lambda/\alpha + (1-\lambda)/\beta).$$

Assim a condição i) neste caso é equivalente a:

$$\|(x,x(\lambda/\alpha + (1-\lambda)/\beta)\| \leq k_0\, x^2/2 + k_1\|(x,0)\| + C_0.$$

Mas o lado esquerdo da expressão anterior é equivalente a:

$$\sqrt{x^2+x^2(\lambda/\alpha+(1-\lambda)/\beta)^2} = |x|\sqrt{1+(\lambda/\alpha+(1-\lambda)/\beta)^2}$$

$$\leq x^2\sqrt{1+(\lambda/\alpha+(1-\lambda)/\beta)^2}$$

basta então tomar $k_0^3 = 2\sqrt{1+(\lambda/\alpha+(1-\lambda)/\beta)^2}$ e $k_1 \geq 0$,

$C_0 \geq 0$ arbitrários.

Logo para termos (i), tomamos

$$k_0 = \max\{k_0^1, k_0^2, k_0^3\}$$

e $k_1 \geq 0$, $C_0 \geq 0$ arbitrários.

Assim, temos satisfeitas as hipoteses II.

Devemos agora calcular o Hamiltoniano. Temos que

$$H(x,p) = \sup\{p.v - L(x,v) \mid v \in \mathbb{R}\}.$$

Seja $\varphi(v) = pv-L(x,v)$. Recordemos que uma condição necessária para que v maximize $\varphi(.)$ é que o $\in \partial\varphi(v)$.

Neste caso, $\partial\varphi(v) = p-\partial_v L(x,v)$, já que $p\,v \in C^1$

Logo a condição $0 \in \partial\varphi(v)$ é equivalente a: $p \in \partial_v L(x,v)$.

Calculemos $\partial_v L(x,v)$, como $L(x,.)$ é deferenciável nos casos $v > 0$ e $v < 0$, temos que:

$$\partial_v L(x,v) = \{\nabla_v L(x,v)\}\ .$$

Em virtude do teorema 5(Cap.II),no ponto $v = 0$, temos que:

$$\partial_v L(x,0) = \mathrm{co}\{\lim_{y\to 0} \nabla_v L(x,y) \mid y \neq 0\}.$$

Assim

$$\partial_v L(x,v) = \begin{cases} (x+v/\alpha)/\alpha & \text{se} \quad v > 0 \\ (x+v/\beta)/\beta & \text{se} \quad v < 0 \\ \{tx : \alpha^{-1} \leq t \leq \beta^{-1}\} & \text{se} \quad v = 0, \end{cases}$$

Logo, para $v > 0$, temos que

$$p = (x + v/\alpha)/\alpha \iff v = (\alpha p - x)\alpha ,$$

Como $\alpha > 0$, temos que $\alpha p > x$.

Analogamente para o caso $v < 0$, temos que

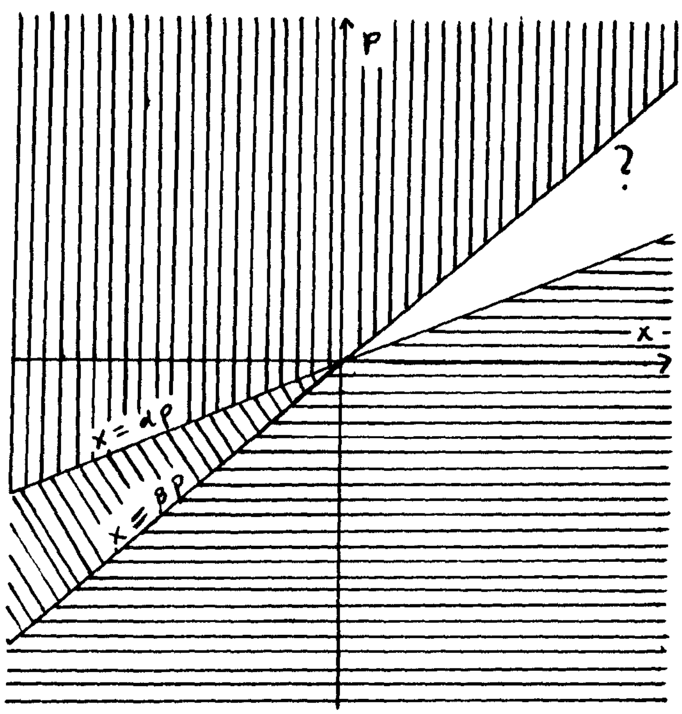
$v = (\beta p - x)\beta$ se $\beta p < x$.

Para o caso $v = 0$, temos que:

$$p \in \{tx: \alpha^{-1} \leq t \leq \beta^{-1}\} \iff p \in \{w: \alpha^{-1} \leq (w/x) \leq \beta^{-1}\}$$

logo $v = 0$ se $\alpha p \leq x \leq \beta p$

Temos o seguinte gráfico.



Na região marcada com raias verticais , o máximo de $H(x,p)$ é atingido em $v = (\alpha p-x)\alpha$.

Na região marcada com raiz horizontais, o máximo é atingido $v = (\beta p-x)\beta$.

Na outra região com raias , o máximo e atingido em $v=0$.

Na região com sinal de interrogação, há dois candidatos para o ponto de máximo: $v_1= (\alpha p-x)\alpha$ e $v_2= (\beta p-x)\beta$.

Temos que comparar os Hamiltonianos avaliados nestes pontos e descobrir para que valores de x, qual deles é maior que o outro, para isto, definimos

$$\Psi(x) = H(x,v_1) - H(x,v_2)$$

e perguntamos: para que valores de x, temos $\Psi(x) \geq 0$ e $\Psi(x) \leq 0$.

Vejamos o caso $\Psi(x) \geq 0$. Temos que:

$$\Psi(x) = \alpha p\ (\alpha p/2 - x) - \beta p(\beta p/2 - x)$$

$$= (p^2/2)\ (\alpha^2 - \beta^2) - (\alpha-\beta)px.$$

Assim, $\Psi(x) \geq 0 \iff (p^2/2)\ (\alpha^2 - \beta^2) \geq (\alpha - \beta)\ px$

$$\iff (p/2)\ (\alpha + \beta) \geq x$$

pois nesta região $p > 0$, e pelas hipóteses $\alpha > \beta > 0$.

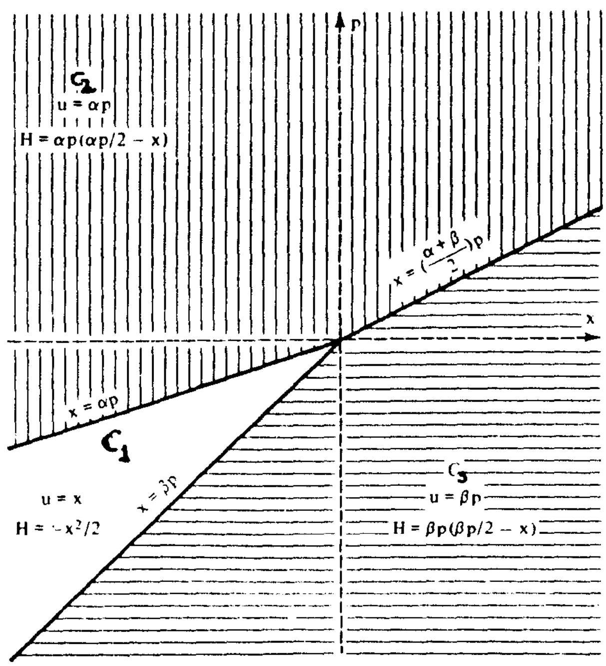
Assim para $0 \leq x \leq (p/2)(\alpha + \beta)$, temos que o máximo é atingido em $v_1 = (\alpha p - x)\alpha$, pelo qual o Hamiltoniano para estes x é:

$$H(x,p) = \alpha p(\alpha p/2 - x).$$

Analogamente para $0 \leq (p/2)(\alpha + \beta) \leq x$, temos que o Hamiltoniano para estes x é:

$$H(x,p) = \beta p\ (\beta p/2 - x)$$

Assim, temos o seguinte gráfico



Assim

$$H(x,p) = \begin{cases} \alpha p(\alpha p/2-x) & \text{em } C_2 \\ -x^2/2 & \text{em } C_1 \\ \beta p(\beta p/2-x) & \text{em } C_3 \end{cases}$$

Para obter um candidato para a otimalidade examinaremos as condições necessárias do teorema 1. Seja $\bar{x}$ ótimo, então existe um arco $\bar{p}$ tal que

$$(-\dot{\bar{p}}(t), \dot{\bar{x}}(t)) \in \partial H(\bar{x}(t), \bar{p}(t)) \quad \text{q.t.p.} \qquad \text{e}$$

$$H(\bar{x}(t), \bar{p}(t)) = \text{constante.}$$

Calculemos o gradiente generalizado de $H(.,.)$.

Em $\text{int}(C_1)$ temos que $\partial H(x,p) = \{(-x,0)\}$.

Em $\text{int}(C_2)$ temos que $\partial H(x,p) = \{(-\alpha p, \alpha(\alpha p-x)\}$.

Em $\text{int}(C_3)$ temos que $\partial H(x,p) = \{(-\beta p, \beta(\beta p-x)\}$.

Na fronteira de C_1 e C_2, temos que:

$$\partial H(x,p) = \text{co}\{\lim_{\substack{y\to x \\ q\to p}} \nabla H(y,p)\} \qquad (x,p) \in \text{Fr}(C_1) \cap \text{Fr}(C_2).$$

Nesta região $x = \alpha p$, de onde:

$$\partial H(x,p) = \text{co}\,\{\lim_{q\to p} \nabla H(\alpha q,q)\}$$

$$= (-\alpha p, 0) \qquad \text{para} \quad (x,p) \in Fr(C_1) \cap Fr(C_2).$$

Em forma análoga

$$\partial H(x,p) = (-\beta p, o) \qquad \text{para} \quad (x,p) \in Fr(C_1) \cap Fr(C_3).$$

Na $Fr(C_2) \cap Fr(C_3)$, temos que $(x,p) \in Fr(C_2) \cap Fr(C_3)$ então $x = (\alpha + \beta)(p/2)$, logo:

$$\partial H(x,p) = \text{co}\,\{\lim_{q\to p} \nabla H((\alpha+\beta)(q/2), q)\}.$$

Para os $q \in int(C_2)$, temos que:

$$\lim_{q\to p} \nabla H((\alpha+\beta)(q/2), q) = \lim_{q\to p} (-\alpha q, \alpha(\alpha q - (\alpha+\beta)(q/2)))$$

$$= (-\alpha p, \alpha p(\alpha - \beta)/2),$$

e para os $q \in int\ (C_3)$, temos que:

$$\lim_{q\to p} \nabla H((\alpha+\beta)(q/2), q) = \lim_{q\to p} (-\beta q, \beta(\beta q - (\alpha+\beta)(q/2)))$$

$$= (-\beta p,\ \beta p(\beta-\alpha)/2).$$

Logo:

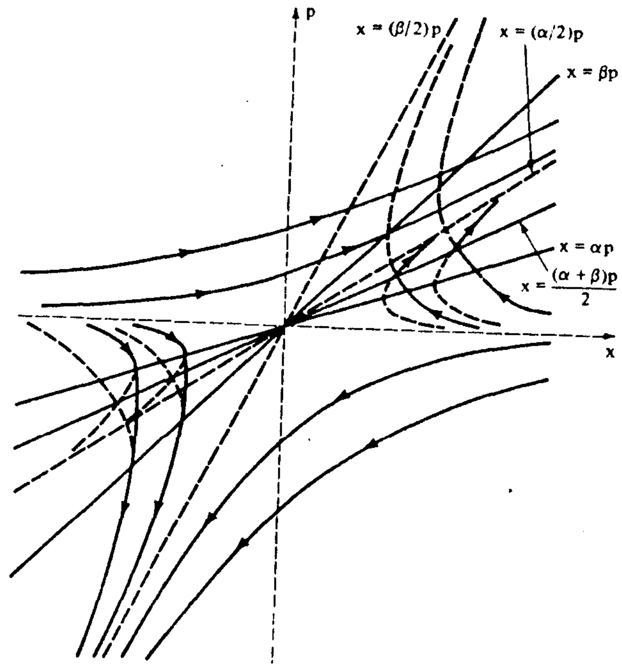
$$\partial H(x,p) = co\{(-\alpha p,\ \alpha p(\alpha-\beta)/2),\ (-\beta p,\ \beta p(\beta-\alpha)/2)\}.$$

Em resumo:

$$\partial H(x,p) \begin{cases} (-x,0) & \text{em } int(C_1) \\ (-\alpha p,\ \alpha(\alpha p-x)) & \text{em } int(C_1) \\ (-\beta p,\ \beta(\beta p-x)) & \text{em } int(C_3) \\ (-\alpha p,\ o) & \text{em } Fr(C_1)\cap Fr(C_2) \\ (-\beta p,\ 0) & \text{em } Fr(C_1)\cap Fr(C_3) \\ Co\{(-\alpha p,\ \alpha p(\alpha-\beta)/2),(-\beta p,\ \beta p)(\beta-\alpha)/2)) & \text{em} \\ Fr(C_2)\cap Fr(C_3)) & \end{cases}$$

Notemos que $(0,0) \notin \partial H(x,p)$ para qualquer $(x,p) \neq (0,0)$ sobre as fronteiras das regiões C_1, C_2, C_3. Isto implica que as trajetórias $(\bar{x}(t),\ \bar{p}(t))$ da inclusão Hamiltoniana não podem "estacionar" ao atingir uma fronteira de duas regiões (C_1, C_2, C_3).

Da natureza de H, as condições mecessárias nos dão, no interior das regiões, equações diferenciais ordinárias cujas soluções são curvas de nível da Hamiltoniana H.



Assim, temos um gráfico qualitativo de $\bar{x}$ em todos os casos. Por exemplo, se $x_0 < 0$ e $x_T < 0$, existem tres segmentos para a solução: um período de tempo durante o qual $\dot{\bar{x}}(t) = \alpha(\alpha\bar{p}(t) - \bar{x}(t)) > 0$, $\dot{\bar{p}}(t) = \alpha\bar{p}(t) < 0$ (na região C_2), seguido por um período $\dot{\bar{x}}(t) = 0$, $\dot{\bar{p}}(t) < 0$ (na região C_1), e finalmente um período em C_3 durante o qual:

$$\dot{\bar{x}}(t) = \beta(\beta\bar{p}(t) - \bar{x}(t)) < 0, \quad \dot{\bar{p}}(t) = \beta\bar{p}(t) < 0.$$

Uma solução numérica é obtida escolhendo p(0) de modo que $\bar{x}$'T) = x_T, resulte, i.e., um problema de valor de fronteira para um sistema de equações diferenciais ordinárias.

Agora, para recuperar o controle $\bar{u}$, associado a tais trajetórias, voltamos à definição dos v otimais.

Na região C_1 temos: $-\bar{x}^2(t)/2 = \gamma_1$, onde γ_1 é uma constante , e o v que da o máximo no Hamiltoniano é: v = 0, de onde:

$$\boxed{\bar{x}(t) = \bar{u}(t)} \qquad \text{na região } C_1.$$

Na região C_2, temos: $\alpha p(\alpha p/2-x) = \gamma_2$, onde γ_2 é uma constante, assim:

$\alpha^2 p^2/2 - \alpha px = \gamma_2 \iff x = \alpha p/2 - \gamma_2/\alpha p$, o v que da o máximo no Hamiltoniano é

$$v = (\alpha p - x)\alpha = (\alpha p - \alpha p/2 + \gamma_2/\alpha p)\alpha$$

$$= (\alpha p/2 + \gamma_2/\alpha p)\alpha \ .$$

Agora,

$$\bar{u}(t) = v/\alpha + x = \alpha p/2 + \gamma_2/\alpha p + \alpha p/2 - \gamma_2/\alpha p$$

i.e.,

$$\boxed{\bar{u}(t) = \alpha\bar{p}(t)} \qquad \text{na região } C_2$$

Análogamente na região C_3, temos que o $\bar{u}(t)$ corespondente para $\bar{x}(t)$ é:

$$\boxed{\bar{u}(t) = \beta\bar{p}(t)} \qquad \text{na região } C_3$$

Agora faremos uso do teorema 5, para provar que o candidato $(\bar{x},\bar{u})$ identificado anteriormente é de fato otimal.

Já que os valores inicial e final de x são fixos, temos que $\Omega \equiv 0$ satisfaz a condição (b) do teorema 5. Só nos resta verificar a condição (a) do teorema 5 para $\Omega \equiv 0$ e $\varepsilon \equiv +\infty$. Isto se traduz em provar a desigualdade:

$$\left.\begin{array}{l} \bar{p}(t)\max\{\alpha(u-x),\beta(u-x)\}-u^2/2 \leq H(\bar{x}(t),\bar{p}(t))-\dot{\bar{p}}(t)\ (x-\bar{x}(t)) \\ \\ \forall\ t\ \in\ [\ 0,t]\quad q.t.p.,\ \forall\ u \in \mathbb{R},\ \forall\ x \in \mathbb{R} \end{array}\right\}\quad (*)$$

o que será feito por regiões.

<u>Região C_2.</u>

Nesta região (*) se reduz a provar:

$\bar{p}(t)\max\ \{\alpha(u-x),\beta(u-x)\}\ -u^2/2\ \leq\ \alpha\bar{p}(t)(\alpha\bar{p}(t)/2\ -\ x)\}$ (1)

já que $\dot{\bar{p}} = \alpha\bar{p}$, $H(x,p) = \alpha p(\alpha p/2-x)$.

Consideremos o caso $u \geq x$; então existe $\delta \geq o$ tal que $u = x+\ \delta$ e (1) é equivalente a:

$x^2+2\delta x+\delta^2+\alpha\bar{p}(t)(\alpha\bar{p}(t)-2x-2\delta) \geq 0.$

Denotando $\varphi(x)= x^2+2\delta x+\ \delta^2+\alpha\bar{p}(t)(\alpha\bar{p}(t)-2x-2\delta\ )$, basta provar que $\varphi(x) \geq 0$.

Busquemos o mínimo de $\varphi(.)$,

$\varphi'(x)\ =\ 2x\ +\ 2\delta\ -\ 2\alpha\bar{p}(t)\ =\ 0$

$\Rightarrow\ \hat{x}\ =\ \alpha\bar{p}(t)\ -\ \delta$

e $\varphi(\hat{x})\ =\ 2\ >\ 0$, logo $\hat{x}\ =\ \alpha\bar{p}(t)\ -\ \delta$ é o ponto de mínimo de $\varphi(.)$, e como $\varphi(\hat{x})\ =\ 0$, temos que $\varphi(x) \geq 0\quad \forall\ x\ \leq\ u$.

Agora consideremos o caso $u\ \leq\ 0$; existe $\delta\ \geq\ 0$ tal que $u = x-\delta$. Logo (1) é equivalente a provar que:

$x^2-2x\delta\ +\ \delta^2+2\beta\delta\bar{p}(t)\ +\alpha\bar{p}(t)(\alpha\bar{p}(t)-2x)\ \geq\ 0.$

Seja $\varphi(x)=x^2-2x\delta+\delta^2+2\beta\delta\bar{p}(t)+\alpha\bar{p}(t)(\alpha\bar{p}(t)-2x)$.

Busquemos o mínimo de $\varphi(.)$.

$\varphi'(x) = 2x-2\delta-2\alpha\bar{p}(t)=0$

$\Rightarrow \hat{x} = \delta+ \alpha\bar{p}(t)$.

Como $\varphi''(x) = 2 > 0$, $\hat{x}$ é mínimo.

Ademais

$\varphi(\hat{x}) = \alpha\bar{p}(t)\ \delta\ (\beta-\alpha) \quad 0$

pois $\bar{p}(t) \leq 0$ e $\beta - \alpha < 0$

O qual prova (*) na região C_2.

<u>Região C_3.</u>

Nesta região (*) se reduz a provar:

$\bar{p}(t)\ \max\ \{\alpha(u-x), \beta(u-x)\} - u^2/2 \leq \beta\bar{p}(t)(\beta\bar{p}(t)/2\ -x)$

já que $\dot{\bar{p}} = \beta\bar{p}$ e $H(x,p) = \beta p(\beta p/2-x)$.

O raciocínio seguinte é análogo que no caso de região C_2.

<u>Região C_1.</u>

Nesta região (*) se reduz a provar:

$$-x^2/2 \quad \bar{x}(t)^2/2 - \bar{x}(t) \quad (x-\bar{x}(t)).$$

já que $u = x$, $\dot{\bar{p}} = \bar{x}$ e $H(x,p) = X^2/2$, ou equivalentemente:

$$3\ \bar{x}(t)^2 - 2\bar{x}(t)\ x + x^2 \geq 0. \qquad (2)$$

Seja $\varphi(x) = x^2 - 2\bar{x}(t) + 3\bar{x}(t)^2$.

$\varphi'(x) = 2x - 2\bar{x}(t) = 0$

$\Rightarrow \quad \hat{x} = \bar{x}(t)$,

e como $\varphi''(x) = 2 > 0$, $\hat{x}$ é ponto de mínimo.

Ademais

$\varphi(\hat{x}) = 2\bar{x}(t)^2 \geq 0$,

temos assim que (2) é satisfeito.

Logo todas as hipóteses do teorema 5 estão satisfeitas, e segue-se que $(\bar{x},\bar{u})$ nos dá em todos os casos uma solução global ao problema.

Bibliografia

[1] Adhigana, S.T., Polak, E., Klessig, R.(1979). A comparative study of several general convergence conditions for algorithms modeled by point-to-set-maps, in: Point- to - set Maps and Mathematical Programming, Mathematical Programming Study 10, (P. Huard ed), North-Holland: 172-190.

[2] Acker, F., Dic Kstein F. (1983) Uma introdução à análise convexa. 14º Colóquio Brasileiro de Matemática IMPA.

[3] Aubin, J.P.(1978). Analyse Fonctionelle nonlinéaire et applications à l'équilibre économique. Ann. Sci. Math. Quebec 2: 5-47.

[4] Aubin, J.P. (1978). Gradients Généralisés de Clarke. Ann. Sci. Math. Quebec 2: 197-252.

[5] Aubin, J.P. (1979). Conês tangents à um sous-ensemble convexo fermé. Ann. Sci. Math. Quebec 3: 63-80.

[6] Aubin, J.P. (1979). Mathematical Methods of Game and Economic Theory. North-Holland, Amsterdan.

[7] Aubin, J.P., Cellina, A. (1984). Differential inclusions Springer-Verlag, Berlin

[8] Aubin, J.P., Ekeland I.(1984). Applied Nonlinear Analysis John-Wiley. New York.

[9] Aubin, J.P., Siegel J. (1980). Fixed points and Stationary points of dissipative Multivalued maps. Proc. Am. Math. Soc. 78: 391-398.

[10] Avez, A.(1983). Calcul Differentiel. Masson, Paris.

[11] Barbu, V. (1976). Nonlinear Semi-Groups and Differential Equations in Banach Spaces. Noordhoff, Leiden, Netherlands.

[12] Berge C. (1959). Espaces Topologiques et Fonctions Multivoques. Dunod, Paris.

[13] Bernal, A., Rojas, M.(1986). Optimización no Diferenciable y Aplicaciones. Tesis Dpto de Matemâticas U. de La Serena. La Serena (Chile).

[14] Boldrini, J.L. (1988). Introdução as Inclusões Diferenciaveis. In Summer School on Diff. Eq. and Dinamical System. Minicourses 1: 43-78.UNICAMP, Campinas.

[15] Borwein, J.M. (1976). Fractional programming without differentiability. Math Programming 11: 283-290.

[16] Borwein, J.M. (1986). Stability and regular points of inequality systems. 48(1): 9-52 .

[17] Brézis, H. (1973). Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert . North-Holland, Amsterdam.

[18] Brézis, H. (1983). Analyse Fonctionelle et Applications. Masson, Paris.

[19] Cartan, H. (1977). Cours de Calcul Différentiel Hermann, Paris.

[20] Castaing, C., Valadier, M. (1977). Convex Analysis and Mensurable Multifunctions. Lectures Notes in Mathematics . Vol 580. Springer-Verlag.

[21] Cellina, A.(1969). Approximation of Set Valued Functions and Fixed Point Theorens. Ann. Math. Pura Appl., 82: 17-24.

[22] Chang, K-C (1981). Variational Methods for non-differentiable functionals and their applications to partial differential equations. J.Math.Anal. and Appl. 80: 102-129.

[23] Christensen, J.P.(1974). Topology and Borel Structure . Math. Studies Nº 10, North-Holland Amsterdam.

[24] Clarke, F.H. (1975). Generalized gradients and applications. Trans. Am. Math. Soc., 205: 247-262.

[25] Clarke, F.H. (1976). A new. approach to Lagrange Multipliers, Math. Oper. Res., 1: 165-174.

[26] Clarke, F.H.(1976). The generalized problem of Bolza , SIAM J. Control opt. 14: 682-699.

[27] Clarke, F.H. (1977). Extremal arcs and extended Hamiltonian systems. Trans. Am. Math. Soc., 231: 349-367.

[28] Clarke, F.H. (1981). Generalized gradients of Lipschitz funtionals, Adv. Math. 40: 52-67.

[29] Clarke, F.H.(1983). Optimization and Nonsmooth Analysis, Wiley, Toronto.

[30] Cominetti, R. (1986). Contribuciones al análisis no-diferenciable: derivadas generalizadas de primer y segundo ordem. Tesis Dpto de Matemáticas e Ciencias de la Computación. Universidad de Chile. Santiago.

[31] Cominetti, R. , Correa, R. (1986). Prepinter.

[32] Contesse L. (1980). Introducción a la optimización con restricciones. Apuntes Universidad de Chile, Santiago.

[33] Correa R., Jofré A. (1986). Prepinter.

[34] Correa R., Thibault. L. (1986). Prepinter

[35] Corley, H.W.(1986). Some Hybrid fixed point theorems related to optimization. J.Math. Anal. Appl. 120:528-532.

[36] Dzedzej, Z. (1985). Fixed-point index theory for a class of nonacyclic multivalued maps. Dissertationes Mathematicae. Vol. CCLIII.

[37] Ekeland, I. (1974) On the variational Principle. J. Math. Anal. Appl. 47: 324-353.

[38] Ekeland, I. (1979). Nonconvex minimization problems Bull. (N-S.) Am. Math. Soc. vol 1, Nº 3: 443-474.

[39] Ekeland, I., Teman, R. (1974). Analyse Convexe et problémes variationnels, Dunod-Gauthier Villars, Paris.

[40] Figueiredo, D.G. (1987). Lectures on the Ekeland variational principle with applications and detours. Monografia, UNICAMP, Campinas.

[41] Figueiredo, D.G., Solimini, S. (1984). A variational approach to superlinear elliptic problems. Comm. in Partial Diff. Eq., 9(7): 699-717.

[42] Flett, T.M. (1980). Differential Analysis. Cambridge University Press, Cambridge.

[43] Fritzsche, H. (1978). Programação não linear análise e métodos. Editora da Universidade de São Paulo, Editora Edgard Blücher Ltda.

[44] Hiriart-Urruty, J.-B. (1979). Tangent cones, generalized gradients and Mathematical programming in Banach Spaces. Math. Oper. Res. 4: 79-97

[45] Hiriart-Urruty, J.-B. (1980). Mean Value theorems in nonsmooth analysis. Numer. Funct. Anal. Optim. 2: 1-30.

[46] Hiriart-Urruty, J.-B. (1981). A better insight into the generalized gradient of the absolute value of a function Appl. Anal. 12: 239-249.

[47] Hönig, C.S. (1976). Aplicações da topologia à análise . "projeto Euclides". IMPA.

[48] Hörmander, L. (1954). Sur la fonction d'appui des ensembles convexes dans um espace localmente convexe. Ark. Math 3: 181-186.

[49] Huard, P. (1979). Extensions of Zangwill's theorem, in : Point-to-set Maps, Mathematical programming Study 10 (P.Huard ed.), North Holland: 98-103.

[50] Ioffe, A.D. (1978). Survey of measurable selection theorems: Russian literatura supplement. SIAM J. Cont.opt. 16:728-733.

[51] Ioffe, A.D., Levin, V.L. (1972). Subdifferentials of convex functions. Trans Moscow. Math Soc. 26:1-72

[52] Kaneko (1984). A Banach type fixed point theorem for multivalued mappings . Kobe J. Math. 1: 163-165.

[53] Klein, E., Thompson, A.C. (1984). Theory of Correspondences. Wiley, Toronto.

[54] Laurent, J.P. (1972). Aproximation et optimizatiom . Herman-Paris.

[55] Lebourg, G. (1979). Generic. differentiability of lipschitzian functions. Trans. Am. Math. Soc. 256: 125-144.

[56] Luemberger, D. (1969) Optimization by Vector Space Methods. J. Wiley, New York.

[57] Luemberger, D. (1973). Introduction to linear and non linear programming, Addison-Wiley.

[58] McClamroch, N.H. (1980). State Models of Dinamical Systems Springer-Verlag, New-York.

[59] Minoux, M. (1983). Programmation mathemãtique, thēorie et algorithmes. Dunod, Paris.

[60] Moreau, J.J.(1966). Fonctionelles Convexes Seminaire Leray, Collēge de France.

[61] Moulin, H., Fogelman F. (1979). La convexite dans les mathēmatiques de la dēcision. Hermann,Paris

[62] Penot, J.P. (1986). The Drop theorem, the Petal theorem and Ekeland's variational principle. Nonlinear Analysis, Theory, Math, Appl. 10(9): 813-822.

[63] Peterson, E.L. (1978). Optimality conditions in generalized Geometric programming. JOTA 26: 3-13.

[64] Polak, E. (1987). On the mathematical foundations of nondifferentiable optimization in engineering desig. SIAM Review 29(1) : 21-89.

[65] Reiland, T.W. (1986). A geometric approach to nonsmooth optimization with sample applications. Nonlinear Analysis, Theory, Meth Appl. 11(10): 1169-1184.

[66] Rockafellar, R.T. (1970). Convex Analysis. Princeton University Press, Princeton N.J.

[67] Rockafellar, R.T. (1971). Convex integral functionals and duality in contributions to Nonlinear Functional Analysis. Academic Press, New York: 215-236.

[68] Rockafellar, R.T. (1975). Existence theorems for general control problems of Bolza and Lagrange. Advances in Math. 15: 312-333.

[69] Rockafellar, R.T. (1984). Network Flows and Monotropic Optimization. Wiley. New York.

[70] Rudin, W. (1979). Analisis Funcional. Ed. revertē , Barcelona.

[71] Shi, Shu-Chung (1980). Remarques sur le gradient généralisē. CRAS 291: 443-446.

[72] Shor, N. Z. (1985). Minimization Methods for Non-Differentiable Functions. Springer-Verlog. Berlīn.

[73] Sotomayor, J. (1979). Lições de equações diferenciais ordinārias. "Projeto Euclides", IMPA.

[74] Studniarski, M. (1985). Meam Value Theorems and sufficient optimality canditions for nonsmooth functions. J. Math. Anal. Appl. 11: 313-326.

[75] Sullivan, F. (1981). A characterization of complete metric spaces. Proc. Am Math. Soc. 83(2): 345-346.

[76] Szulkin, A. (1986). Minimax principles for lower semicontinuous functions and applications to nonlinear boundary Value problens. Ann. Inst. Henri Poincaré, Analyse non linéaire 3(2): 77-109.

[77] Wagner, D.M. (1977). Survey of measurable selection theorems, SIAM J. Control Opt. 15: 859-903.

[78] Wegrzyk, R. (1982). Fixed-point theorems for multivalued functions and their applications to funcional equations. Dissertationes Mathematicae, Vol. CCI.

[79] Zeidan, V. (1984). A modified Hamilton-Jacobi opproach in the generalized problem of Bolza App. Math Opt. 11: 97-109.

[80] Zeidan, V. (1984). First and second ordre Conditions for

optimal Control and the calculus of variations. Appl. Math. Opt. 11: 209-226.

[81] Mifflin, R. (1977). Semismooth and Semiconvex functions in optimization, SIAM J.Control Optm. 15: 959-972.

[82] Contesse, L. (1981). On the continuity of polar point - to - set maps. Publication A.N.O. Université de Lille, France.

[83] Contesse (1982). On the continuity of optimal value functions and of optimal solution sets. Publication A.N.O. Université de Lille, France.

[84] Filippov, A.F. (1967). Classical solutions of differential equations with multivalued right-hard side. SIAM J. Control Optim. 5: 609-621.

[85] Giles, J.R. (1982). Convex analysis with application in the differentiation of convex functions. Pitman Adv. Program nº 58. London.

[86] Hale, J. (1969). Ordinary Differential Equations J. Wiley.

[87] Papargeorgion, N.S. (1987). On multivalued evolution equations and differential equations in Banach spaces Commentari Mathematics, Univ. Sancti Pauli. 36(1): 21-39.

[88] Stein, E.M.(1970). Singular Integrals and Differentiability Ptoperties of Functions. Princeton Math. Sẽries, nº 30, Princeton Univ. Press.

[89] Thibault, L. (1982) On generalized differentials and subdifferential of Lipschitz Vector-Value functions . Nonlinear Analysis 6: 1037 - 1053.

[90] Clarke, F.H. (1975) Admissible relaxation in variational and control problems. J. Math. Anal. Appl. 51: 557 - 576.