

**CARACTERIZAÇÃO FUNCIONAL DE PONTOS ÓTIMOS
PARA PROBLEMAS DE OTIMIZAÇÃO,
REGULARES E NÃO-REGULARES**

Ana Lúcia Carlos Reis da Silva

**DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA,
ESTATÍSTICA E CIÊNCIA DA COMPUTAÇÃO
DA
UNIVERSIDADE ESTADUAL DE CAMPINAS
PARA
OBTENÇÃO DO GRAU DE MESTRE
EM
MATEMÁTICA APLICADA**

**Área de Concentração: Análise Aplicada
Orientador: Prof. Dr. Marko Antonio Rojas Medar**

Durante a elaboração deste trabalho a autora recebeu apoio financeiro da CAPES

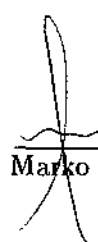
– Campinas, julho de 1996 –

“Caracterização Funcional de Pontos Ótimos para Problemas de Otimização Regulares e Não-Regulares”

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pela Sra. Ana Lúcia Carlos Reis da Silva e aprovada pela comissão julgadora.

Campinas, 16 de junho de 1996.

Prof. Dr.


Marko Antônio Rojas Medeiros
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.

UNIDADE	BC
N.º CHAMADA:	77 Unicamp
Si	38c
V.	E.
TOMBO BC/	29.538
PROC.	281.197
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	08/02/97
N.º CPD	

CM-00096672-8

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Silva, Ana Lúcia Carlos Reis da

Si38c Caracterização funcional de pontos ótimos para problemas de
otimização, regulares e não-regulares / Ana Lúcia Carlos Reis da
Silva -- Campinas, [S.P. :s.n.], 1996.

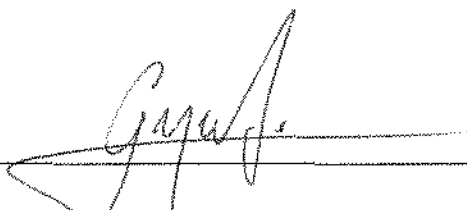
Orientador : Marko Antonio Rojas Medar

Dissertação (mestrado) - Universidade Estadual de Campinas,
Instituto de Matemática, Estatística e Ciência da Computação.

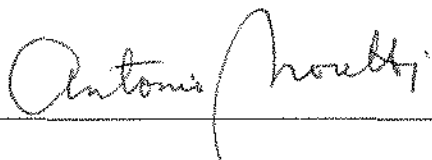
1. Teoria de controle. 2. Otimização matemática. 3. Banach,
Espaços de. 4. Equações de Lagrange. I. Rojas Medar, Marko
Antonio. II. Universidade Estadual de Campinas. Instituto de
Matemática, Estatística e Ciência da Computação. III. Título.

Tese de Mestrado defendida e aprovada em 16 de Julho de 1996

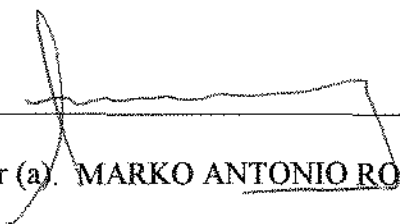
pela Banca Examinadora composta pelos Profs. Drs.

A handwritten signature in dark ink, appearing to read 'Gerald', written over a horizontal line.

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

A handwritten signature in dark ink, appearing to read 'Antonio Carlos Moretti', written over a horizontal line.

Prof (a). Dr (a). ANTONIO CARLOS MORETTI

A handwritten signature in dark ink, appearing to read 'Marko Antonio Rojas Medar', written over a horizontal line.

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

*Aos meus queridos pais,
Paulo e Nilza*

Agradecimentos

Primeiramente, aos meus pais e irmãos, cujo amor e dedicação me fizeram chegar até aqui.

Aos colegas do IMECC, dentre os quais destaco Pedrinho, Luciane, Gisele, Eliane e Patrícia, pois deles obtive um enorme apoio em momentos difíceis pelos quais passei durante o curso.

Às colegas Cristina Vaz e Cristina Wolff, e ao professor Petrônio Pulino, pela imensa ajuda que me prestaram desde a minha viagem para o curso de verão no IMECC, até o momento.

Aos amigos Lúcia, Ayumi, Fábio e Jair, pela ajuda na fase de digitação.

Ao meu orientador, professor Marko, que verdadeiramente teve que exercer uma enorme paciência, e demonstrou grande amizade, em meus diversos momentos de desânimo e irritação.

Finalmente, ao meu esposo Juaci, pelas preciosas horas perdidas na digitação deste trabalho e pelo incentivo dado desde o início do curso até o final.

Sumário

1	Introdução	1
2	Caso regular, formalismo de Dubovitskii-Milyutin	3
2.1	A equação de Euler-Lagrange	4
2.2	Direções de descida	9
2.3	Direções factíveis	12
2.4	Direções tangentes e o teorema de Lyusternik	15
2.5	Cálculo dos cones duais	21
2.6	Multiplicadores de Lagrange e o teorema de Kuhn-Tucker	33
3	Extensão para o caso não regular	40
3.1	Uma generalização do teorema de Lyusternik	41
3.2	Os multiplicadores de Lagrange	47
3.3	Considerações finais	64

Apêndice A	65
Apêndice B	68
Apêndice C	72
Bibliografia	75

Capítulo 1

Introdução

Neste trabalho fazemos uma exposição de problemas de otimização sob a visão da Análise Funcional. Na verdade o problema básico é

$$\begin{cases} \min & f(x) \\ \text{s.a} & F(x) = 0 \\ & f_i(x) \leq 0, \quad i = 1, \dots, m \end{cases}$$

onde f e f_i , $i = 1, \dots, m$, são funcionais sobre o espaço de Banach X e $F : U \subseteq X \longrightarrow Y$ é uma função do aberto $U \subseteq X$ no espaço de Banach Y .

Utilizando-se da Análise Funcional, Dubovitskii e Milyutin estabeleceram a regra de multiplicadores de Lagrange que, por ser tratada em espaços de Banach quaisquer, refere-se tanto a problemas de programação matemática, quanto a problemas relacionados com a teoria de controle ótimo. No primeiro capítulo, tratamos deste assunto, apresentando alguns conceitos mais comumente encontrados na teoria de controle.

A teoria desenvolvida por Dubovitskii e Milyutin, contudo, torna-se ineficiente nos casos em que a função F é não-regular em pontos suspeitos de serem extremos no problema (P) , pois nestes pontos o multiplicador λ_0 associado à função objetivo pode ser nulo, impossibilitando-nos de relacionar a função a ser minimizada com as restrições do problema.

Tendo, então, essa visão analítico-funcional do problema, Avakov [13] conseguiu estender a regra de multiplicadores para problemas em que se exige menos do que a regularidade da função F em pontos suspeitos de serem de mínimo, conseguindo assim $\lambda_0 \neq 0$. Este é o assunto do nosso segundo capítulo, no qual consideraremos apenas o problema de minimização com restrições de igualdade.

Nas demonstrações feitas nos dois capítulos são utilizados alguns resultados de análise funcional que, para que não se tornem obscuras as explicações, são indicados por números entre colchetes, indicando sua colocação em um dos três apêndices, que foram feitos apenas para referência, pois suas demonstrações não são feitas, mas damos as referências bibliográficas, onde podem ser encontradas, entre colchetes.

Capítulo 2

Caso regular, formalismo de Dubovitskii-Milyutin

Abordaremos aqui a teoria desenvolvida por Dubovitskii e Milyutin para a determinação de condições necessárias de otimalidade para o problema

$$\left\{ \begin{array}{ll} \min & f(x) \\ \text{s.a} & F(x) = 0 \\ & f_i(x) \leq 0, \quad i = 1, \dots, n. \end{array} \right.$$

Neste sentido, damos algumas definições preliminares, seguidas de alguns resultados sobre os cones de direções de descida de f , direções factíveis e de direções tangentes. Logo depois, apresentamos o teorema de Dubovitskii-Milyutin, que nada mais é do que a regra dos multiplicadores de Lagrange em uma versão funcional. Para que possamos utilizá-la na prática, seguimos com o desenvolvimento dos cálculos dos cones e de seus duais. Concluimos com a aplicação dessa teoria a alguns problemas de programação matemática.

2.1 A equação de Euler-Lagrange

Nesta seção, $F(x)$ será um funcional (não necessariamente linear) cujo domínio de definição será uma vizinhança U de um ponto x_o num espaço vetorial normado E . Serão também considerados dois tipos de conjuntos que restringirão a variável x , a saber, Q_i , $i = 1, \dots, n$, que serão conjuntos de interior não-vazio, e Q_{n+1} , que será um conjunto sem pontos interiores.

Em geral, Q_i , $i = 1, \dots, n$, serão determinados por restrições de desigualdade, e Q_{n+1} por um sistema de restrições de igualdade.

Considerando $Q = \bigcap_{i=1}^{n+1} Q_i$, nosso problema será encontrar $x_o \in Q$ tal que

$$F(x_o) = \min_{Q \cap U} F(x).$$

Seguiremos agora com um conjunto de definições que serão a base de todo este capítulo.

Definições 2.1.1 :

- (i) Um vetor h é dito uma **direção de descida** do funcional $F(x)$ no ponto x_o , se existe uma vizinhança U de h , um número estritamente negativo $\alpha = \alpha(F, x_o, h)$ e $\varepsilon_o > 0$ tais que, quaisquer que sejam ε ($0 < \varepsilon < \varepsilon_o$) e $\bar{h} \in U$

$$F(x_o + \varepsilon \bar{h}) \leq F(x_o) + \varepsilon \alpha. \quad (2.1)$$

- (ii) $F(x)$ se diz um **funcional regular de descida** se suas direções de descida no ponto x_o formam um conjunto convexo.
- (iii) Sendo Q dado por restrições de desigualdade, dizemos que o vetor h é uma **direção factível** para Q no ponto x_o , se existe uma vizinhança U do vetor h e $\varepsilon_o > 0$ tal que, para quaisquer ε ($0 < \varepsilon < \varepsilon_o$) e $\bar{h} \in U$, os vetores $x_o + \varepsilon \bar{h} \in Q$ (ver figura 1.1).

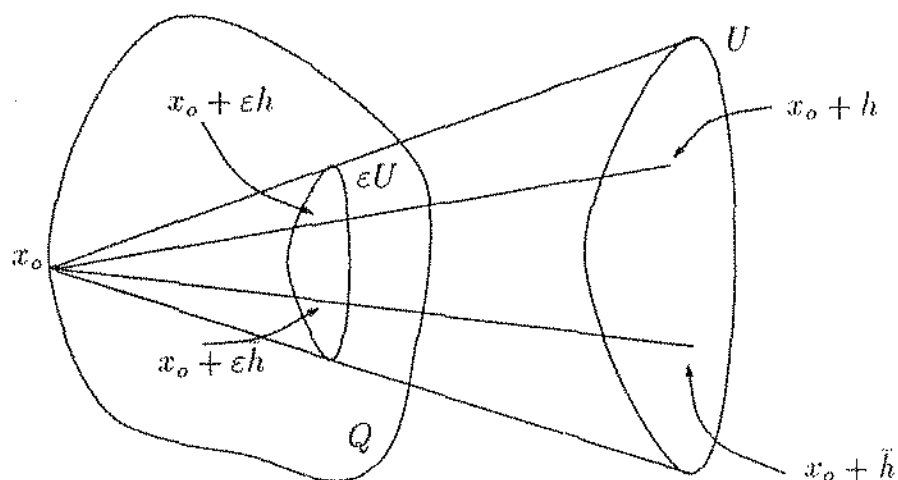


Figura 2.1:

- (iv) Uma restrição de desigualdade Q é **regular** no ponto x_o se o conjunto de direções factíveis para Q em x_o é convexo.

Quando trabalhamos com restrições de igualdades, por ser Q um conjunto sem pontos interiores, não podemos utilizar o conjunto de direções factíveis, como dado acima. Assim, torna-se necessário uma ligeira modificação em tal conceito para este tipo de restrições.

- (v) Um vetor h se diz um **vetor tangente unilateral** ou, simplesmente, **direção tangente** a Q no ponto x_o se, para qualquer ε ($0 < \varepsilon < \varepsilon_o$), existe um ponto $x(\varepsilon) \in Q$ tal que, se colocarmos $x(\varepsilon) = x_o + \varepsilon h + r(\varepsilon)$ então o vetor $r(\varepsilon) \in E$ é tal que, para qualquer vizinhança U de zero, $\frac{1}{\varepsilon} r(\varepsilon) \in U$ para todo $\varepsilon > 0$ suficientemente pequeno, ou, equivalentemente,

$$\|r(\varepsilon)\| = o(\varepsilon).$$

(ver figura 1.2)

- (vi) Uma restrição da igualdade Q é **regular** no ponto x_o se o conjunto das direções tangentes a Q em x_o é convexo.

Observações 2.1.1 :

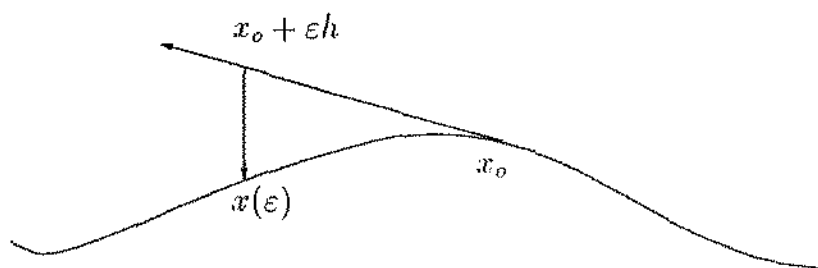


Figura 2.2:

- (i) As direções de descida geram um cone aberto K com vértice O , isto é, $\lambda K = K$ para todo $\lambda > 0$.

De fato, se h é uma direção de descida, então, λh , $\lambda > 0$, também o é (basta trocar U por λU , ϵ_0 por $\frac{\epsilon_0}{\lambda}$ e α por $\lambda\alpha$), e K é um cone com vértice em O . Além disso, K é aberto, pois se $h \in K$ e $\bar{h} \in U$, então, $\bar{h} \in K$ (basta tomar o mesmo U , $\alpha \epsilon_0$ na desigualdade (1.1)).

- (ii) As direções factíveis geram um cone aberto K com vértice em O .

De fato, se $h \in K$, trocando U por λU e ϵ_0 por $\frac{\epsilon_0}{\lambda}$, tem-se que $\lambda h \in K$, para $\lambda > 0$, e todo $\bar{h} \in U$ também estará em K .

- (iii) As direções tangentes também geram um cone K com vértice em O .

Definição 2.1.1 Se K é um cone em E , seu cone dual K^* é o conjunto dado por $K^* = \{f \in E^* : f(x) \geq 0, \forall x \in K\} \subset E^*$

O uso da notação K^* para o cone dual não parece próprio, pois se $K \equiv E$, então $K^* = \{0\} \neq E^*$. Contudo, isto é usual na literatura, não devendo, portanto, causar confusão.

Lema 2.1.1 Sejam $K_1, \dots, K_n, K_{n+1}$ cones convexos com vértices em 0 , onde $K_1, \dots, K_n$, são abertos, e suponha que $\bigcap_{i=1}^{n+1} K_i = \emptyset$. Então existem funcionais lineares $f_i \in K_i^*$, não simultaneamente nulos, tais que

$$f_1 + \dots + f_{n+1} = 0.$$

Demonstração:

Suponha $\bigcap_{i=1}^n K_i = K \neq \emptyset$, $K_{n+1} \cap K = \emptyset$. Como K é aberto, pelo teorema de separação de Hahn-Banach, existe $f \in E^*$ ($f \neq 0$) tal que

$$\begin{aligned} f(x) &\geq 0, & x \in K \\ f(x) &\leq 0, & x \in K_{n+1}. \end{aligned}$$

É claro que

$$K^* = \left(\bigcap_{i=1}^n K_i \right)^* = \sum_{i=1}^n K_i^*$$

e portanto,

$$f = \sum_{i=1}^n f_i, \quad \text{com } f_i \in K_i^* \quad i = 1, \dots, n.$$

Tomando $f_{n+1} = -f$, temos $f_{n+1} \in K_{n+1}^*$.

Supondo agora que $\bigcap_{i=1}^n K_i = \emptyset$, existe algum s ($1 \leq s \leq n$) tal que

$$\bigcap_{i=1}^{s+1} K_i = \emptyset \quad \text{e} \quad K = \bigcap_{i=1}^n K_i \neq \emptyset.$$

Usando o que foi provado acima (trocando n por s) obtemos $f_i \in K_i^*$, $i = 1, \dots, s+1$, $f_{s+1} \neq 0$ tais que

$$f_i + \dots + f_{s+1} = 0.$$

Fazendo $f_{s+2} = \dots = f_{n+1} = 0$, obtemos os referidos funcionais.

□

A recíproca deste lema também é verdadeira e só não a demonstramos aqui por não ser de nosso interesse imediato. (ver [*]).

Tendo em mãos tais fatos podemos apresentar o seguinte

Teorema 2.1.1 (Dubovitskii-Milyutin) Suponha que o funcional $F(x)$ assume um mínimo local sobre $Q = \bigcap_{i=1}^{n+1} Q_i$ no ponto $x_o \in Q$. Suponha ainda que $F(x)$ é regular de descida em x_o , com direções de descida K_o , as restrições de desigualdade Q_i , $i = 1, \dots, n$, são regulares em x_o , com direções factíveis K_i , a restrição de igualdade Q_{n+1} é também regular em x_o , com direções tangentes K_{n+1} . Então, existem funcionais lineares contínuos f_i , $i = 0, 1, \dots, n+1$, não simultaneamente nulos, tais que $f_i \in K_i^*$, $i = 0, \dots, n+1$, os quais satisfazem a equação de Euler-Lagrange:

$$f_0 + f_1 + \dots + f_n + f_{n+1} = 0.$$

Demonstração:

Basta mostrar que se $F(x)$ tem um mínimo em x_o , então

$$\bigcap_{i=0}^{n+1} K_i \neq \emptyset.$$

(nenhuma direção de descida do funcional $F(x)$ pode ser factível para todas as restrições).

Depois é só usar o lema acima.

Suponha que isto é falso, isto é, existe uma vizinhança U de um vetor h tal que, quando $0 \leq \varepsilon \leq \varepsilon_o$, qualquer vetor $x_o + \varepsilon \bar{h}$, $\bar{h} \in U$, está em $\bigcap_{i=1}^n Q_i$ e satisfaz

$$F(x_o + \varepsilon \bar{h}) \leq F(x_o) + \varepsilon \alpha.$$

Sejam o vetor $x(\varepsilon) = x_o + \varepsilon h + r(\varepsilon) \in Q_{n+1}$ como na definição de direção tangente, e ε_1 tais que

$$\frac{1}{\varepsilon} r(\varepsilon) \in U - h \quad \text{ou} \quad \bar{h}(\varepsilon) = h + \frac{1}{\varepsilon} r(\varepsilon) \in U,$$

para $0 \leq \varepsilon \leq \varepsilon_1$.

Então, quando $0 \leq \varepsilon < \min\{\varepsilon_o, \varepsilon_1\}$, os vetores $x(\varepsilon) = x_o + \varepsilon \bar{h}(\varepsilon)$ estão, por um lado, em $\bigcap_{i=1}^n Q_i$ e, por outro lado, em Q_{n+1} . Assim, eles satisfazem todas as restrições. Acontece que

eles também satisfazem a desigualdade

$$F(x(\varepsilon)) = F(x_o + \varepsilon \bar{h}(\varepsilon)) \leq F(x_o) + \varepsilon \alpha < F(x_o)$$

o que contradiz o fato de que x_o é um ponto de mínimo. Logo,

$$\bigcap_{i=0}^{n+1} K_i = \emptyset.$$

Além disso, $K_0, K_1, \dots, K_n$, são cones convexos abertos, e K_{n+1} é um cone convexo.

Obtivemos, assim, as hipóteses necessárias para a utilização do lema anterior e consequente conclusão do teorema.

□

Do teorema de Dubovitskii-Milyutin vemos que, para obtermos uma caracterização analítica (ou funcional) das condições geométricas de otimalidade obtidas no lema, devemos conseguir calcular os cones associados às diferentes restrições do problema dado e, posteriormente, devemos saber calcular os cones duais associados. Este será o assunto das seções seguintes.

2.2 Direções de descida

Nesta seção serão apresentados alguns teoremas que servirão para o cálculo das direções de descida de funcionais. Em toda ela K denotará o conjunto das direções de descida para o funcional $F(x)$ no ponto x_o , $F'(x_o; h)$ indicará a derivada de $F(x)$ no ponto x_o na direção h e $F'(x_o)$ a derivada segundo Fréchet de $F(x)$ no ponto x_o .

Teorema 2.2.1 *Se $h \in K$ e $F'(x_o; h)$ existe, então,*

$$F'(x_o; h) < 0.$$

Demonstração:

Temos, da definição de direção de descida, que existe $\varepsilon_o > 0$ e $\alpha < 0$ tais que

$$F(x_o + \varepsilon h) \leq F(x_o) + \varepsilon \alpha$$

mas,

$$F'(x_o; h) = \lim_{\varepsilon \downarrow 0} \frac{F(x_o + \varepsilon h) - F(x_o)}{\varepsilon} \leq \lim_{\varepsilon \downarrow 0} \frac{F(x_o) + \varepsilon \alpha - F(x_o)}{\varepsilon} = \alpha < 0.$$

□

Teorema 2.2.2 *Seja E um espaço de Banach e suponha que $F(x)$ satisfaz uma condição de Lipschitz numa vizinhança de x_o (isto é, existe $\varepsilon_o > 0$ e $\beta > 0$ tais que $|F(x_1) - F(x_2)| \leq \beta \|x_1 - x_2\|$ sempre que $\|x_1 - x_2\| \leq \varepsilon_o$, $\|x_2 - x_o\| \leq \varepsilon_o$) e que existe h tal que $F'(x_o; h) < 0$. Então, $h \in K$.*

Demonstração:

Sejam $\delta > 0$ e $F'(x_o; h) = -\delta < 0$. Da definição de derivada direcional temos:

$$F(x_o + \varepsilon h) \leq F(x_o) - \frac{\varepsilon \delta}{2}$$

para $0 < \varepsilon < \varepsilon_1$.

Se $\bar{h}$ é um vetor arbitrário tal que $\|h - \bar{h}\| \leq \frac{\delta}{4\beta}$ e $0 < \varepsilon < \min\{\varepsilon_o, \varepsilon_1\}$, então

$$\begin{aligned} F(x_o + \varepsilon \bar{h}) &\leq F(x_o + \varepsilon h) + \beta \|\varepsilon \bar{h} - \varepsilon h\| \\ &\leq F(x_o) - \frac{\varepsilon \delta}{2} + \frac{\varepsilon \delta}{4} \\ &= F(x_o) - \frac{\varepsilon \delta}{4} \end{aligned}$$

e a condição para que h seja uma direção de descida é satisfeita tomando $\alpha = -\frac{\delta}{4}$, isto é, $h \in K$.

□

Teorema 2.2.3 *Seja E um espaço de Banach. Suponha que $F(x)$ satisfaz uma condição de Lipschitz numa vizinhança do ponto $x_o \in E$, e que é direcionalmente diferenciável em x_o , em qualquer direção. Se $F'(x_o; h)$ é convexo como função de h , então $F(x)$ é regular de descida em x_o , e*

$$K = \{h : F'(x_o; h) < 0\}.$$

Demonstração:

Dos teoremas 2.2.1 e 2.2.2 concluímos que

$$K = \{h : F'(x_o; h) < 0\}.$$

Como $F'(x_o; h)$ é convexa em h , segue das propriedades de funções convexas que o conjunto K acima é convexo.

□

Temos ainda dois casos especiais em que podemos determinar as direções de descida usando derivadas ou derivadas direcionais:

- i) Se $F(x)$ é um funcional convexo contínuo num espaço normado E , então $F(x)$ é regular de descida em qualquer ponto e

$$K = \{h : F'(x_o; h) < 0\}.$$

- ii) Se $F(x)$ é Fréchet-diferenciável, ou, simplesmente, diferenciável em x_o , então $F(x)$ é regular de descida em x_o e

$$K = \{h : \langle F'(x_o), h \rangle < 0\}.$$

Vamos agora a alguns exemplos, nos quais calculamos a derivada direcional e aplicamos os teoremas já vistos.

Exemplo 2.2.1 Seja $F(x) = \langle f, x \rangle$, com $f \in E^*$. Então, $F(x)$ é Fréchet-diferenciável, $F'(x) = f$ e

$$K = \{h : \langle f, h \rangle < 0\}.$$

Exemplo 2.2.2 Seja o funcional integral

$$F(x) = \int_0^T \Phi(x(t), t) dt$$

sobre $E = (C[0, T])^n$, $x = x(t)$, com $\Phi(x, t)$ contínua em x, t , Fréchet-diferenciável em relação a x e com derivada parcial $\Phi_x(x, t)$ contínua em x, t . Então, $F(x)$ é Fréchet-diferenciável, e

$$\langle F'(x_o), h \rangle = \int_0^T \langle \phi_x(x_o(t), t), h(t) \rangle dt$$

$$K = \{h \in C : \int_0^T \langle \Phi_x(x_o(t), t), h(t) \rangle dt < 0\}.$$

Exemplo 2.2.3 Seja $F(x) = \max_{1 \leq i \leq n} F_i(x)$, onde $F_i(x)$, $i = 1, \dots, n$, são funções num espaço normado, diferenciáveis em x_o na direção h . Então, $F(x)$ é diferenciável em x_o na direção h , e

$$F'(x_o; h) = \max_{i \in I} F'_i(x_o; h)$$

onde $I = \{i : F_i(x_o) = F(x_o)\}$.

Se cada função $F_i(x)$ é convexa e contínua, ou diferenciável, então $F(x)$ é regular de descida em x_o e

$$K = \{h : F'_i(x_o, h) < 0, i \in I\}.$$

2.3 Direções factíveis

Nosso objetivo agora é determinar o cone de direções factíveis no ponto x_o para o conjunto Q . Tal cone será denotado por K_b .

O caso que nos interessa é aquele em que x_o é ponto de fronteira de Q , já que, se $x_o \in \text{int } Q$, segue que $K_b \equiv E$.

Além disso, o conjunto Q deve ter pontos interiores, pois, caso contrário, $K_b = \emptyset$. Trataremos o caso onde Q é definido por funcional, isto é,

$$Q = \{x : F(x) \leq F(x_o)\}. \quad (2.2)$$

Pelo que já foi dito, se $F(x)$ é contínuo podemos estudar o caso

$$Q = \{x : F(x) \leq \lambda, \quad \lambda \neq F(x_o)\}.$$

Aqui, o cone de direções de descida em x_o do funcional $F(x)$ será denotado por K_y .

Lema 2.3.1 $K_y \subseteq K_b$.

Demonstração:

Se $h \in K_y$, temos

$$F(x_o + \varepsilon \bar{h}) \leq F(x_o) + \varepsilon \alpha < F(x_o),$$

para todo $\bar{h}$ em uma vizinhança U do vetor h e todo ε tal que $0 < \varepsilon < \varepsilon_o$. Além disso, para tais $\bar{h}$ e ε , temos $x_o + \varepsilon \bar{h} \in Q$.

Portanto, h é também uma direção factível.

□

Teorema 2.3.1 *Suponha que $F(x)$ é Fréchet-diferenciável em x_o em qualquer direção. Se existe $\bar{h}$ tal que $F'(x_o; \bar{h}) < 0$, então*

$$K_b = \{h : F'(x_o; h) < 0\}.$$

Demonstração:

Suponhamos $h \in K_b$. Então $x_o + \varepsilon h \in Q$ para $0 < \varepsilon < \varepsilon_o$, isto é,

$$F(x_o + \varepsilon h) \leq F(x_o).$$

Temos então que $F'(x_o, h) \leq 0$. Como K_b é aberto, existe uma vizinhança U do vetor h tal que $U \subseteq K_b$. Seja $\gamma > 0$ tal que $h_\gamma = h + \gamma(h - \tilde{h}) \in U$. Então,

$$F'(x_o; h_\gamma) \leq 0$$

e segue-se da convexidade de $F'(x_o; h)$ em h que

$$\begin{aligned} F'(x_o; h) &= F'\left(x_o; \frac{1}{1+\gamma} h_\gamma + \frac{\gamma}{1+\gamma} \tilde{h}\right) \\ &\leq \frac{1}{1+\gamma} F'(x_o; h_\gamma) + \frac{\gamma}{1+\gamma} F'(x_o; \tilde{h}) \\ &\leq \frac{\gamma}{1+\gamma} F'(x_o; \tilde{h}). \end{aligned}$$

□

Corolário 2.3.1 *Suponha que qualquer das seguintes condições se satisfaz:*

- i) *E é um espaço de Banach, $F(x)$ satisfaz uma condição de Lipschitz numa vizinhança do ponto x_o e é diferenciável no ponto x_o em qualquer direção, $F'(x_o, h)$ é convexo em h , e existe h tal que $F'(x_o; \tilde{h}) < 0$;*
- ii) *$F(x)$ é um funcional convexo contínuo e existe $\tilde{x}$ tal que $F(\tilde{x}) < F(x_o)$;*
- iii) *E é um espaço de Banach, $F(x)$ é Fréchet-diferenciável em x_o , $F'(x_o) \neq 0$.*

Então

$$K_b = K_y = \{h : F'(x_o; h) < 0\}.$$

□

Finalmente, se Q é um conjunto convexo, o seu cone de direcções factíveis no ponto x_o é dado por

$$K_b = \{\lambda(\text{int } Q - x_o) : \lambda > 0\},$$

isto é,

$$K_b = \{h : h = \lambda(x - x_o), x \in \text{int } Q, \lambda > 0\}.$$

E isto é obtido simplesmente usando a definição.

2.4 Direcções tangentes e o teorema de Lyusternik

Nesta secção demonstraremos o teorema de Lyusternik, que é de grande importância para a determinação das direcções tangentes, utilizando o teorema das funções implícitas numa versão em que não necessitamos que a função seja de classe C^1 , mas apenas estritamente diferenciável.

Definição 2.4.1 *Sejam X e Y espaços de Banach. A aplicação $P : X \longrightarrow Y$ é dita estritamente diferenciável em x_o se existe uma aplicação $\Lambda \in \mathcal{L}(X, Y)$ tal que, para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, quaisquer que sejam $x_1, x_2 \in B(x_o, \delta)$ tem-se*

$$\|P(x_1) - P(x_2) - \Lambda(x_1 - x_2)\| \leq \varepsilon \|x_1 - x_2\| \quad (2.3)$$

Observações 2.4.1 :

- (i) Se P é estritamente diferenciável, então P é Fréchet-diferenciável. Basta fazer $x_1 = x_o$ e $x_2 = x_o + h$ em (1.3), e considerar que $\Lambda = P'(x_o)$.
- (ii) Se P é estritamente diferenciável em x_o , então existe uma vizinhança de x_o onde P é contínua.

De fato, para quaisquer $x_1, x_2 \in B(x_o, \delta)$ tem-se

$$\|P(x_1) - P(x_2)\| - \|\Lambda(x_1 - x_2)\| \leq \|P(x_1) - P(x_2) - \Lambda(x_1 - x_2)\| \leq \varepsilon \|x_1 - x_2\|.$$

Portanto,

$$\begin{aligned}\|P(x_1) - P(x_2)\| &\leq \varepsilon\|x_1 - x_2\| + \|\Lambda(x_1 - x_2)\| \\ &\leq \varepsilon\|x_1 - x_2\| + M\|x_1 - x_2\| \\ &= (\varepsilon + M)\|x_1 - x_2\|.\end{aligned}$$

Logo, P é Lipschitz e portanto contínua em $B(x_0, \delta)$.

(iii) P pode ser Fréchet-diferenciável e não ser estritamente diferenciável, como no seguinte exemplo: $P : \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$P(x) = \begin{cases} x^2, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}$$

em $x = 0$.

De fato, em qualquer vizinhança de $x = 0$, fazemos

$$\Lambda(h) = 0 \quad \text{para todo } h$$

$$\alpha(h) = \begin{cases} |h|, & \text{se } h \in \mathbb{Q} \\ 0, & \text{se } h \notin \mathbb{Q} \end{cases}$$

e obtemos

$$P(0 + h) = P(0) + \Lambda(h) + \alpha(h)\|h\|$$

onde $\Lambda \in \mathcal{L}(\mathbb{R}, \mathbb{R})$ e

$$\lim_{\|h\| \rightarrow 0} \|\alpha(h)\| = \|\alpha(0)\| = 0.$$

Logo, $P(x)$ é Fréchet-diferenciável em $x = 0$. Contudo não é estritamente diferenciável, pois se o fosse, pelo que foi visto no item anterior, ela seria contínua em uma vizinhança de $x = 0$, mas $P(x)$ só é contínua no ponto $x = 0$.

(iv) Se P é de classe C^1 , então P é estritamente diferenciável.

Teorema 2.4.1 (das funções implícitas) *Sejam X um espaço topológico, Y e Z espaços de Banach, W uma vizinhança do ponto (x_o, y_o) em $X \times Y$, $\psi : W \longrightarrow Z$ e $\psi(x_o, y_o) = z_o$. Suponhamos que:*

- i) $x \longmapsto \psi(x, y_o)$ é contínua em x_o ;
- ii) Existe $\Lambda \in \mathcal{L}(Y, Z)$ tal que, para todo $\varepsilon > 0$, existe $\delta > 0$ e uma vizinhança V de x_o possuindo a seguinte propriedade: a condição $x \in V$ e $y', y'' \in B(y_o, \delta)$ implica a desigualdade
$$\|\psi(x, y') - \psi(x, y'') - \Lambda(y' - y'')\| < \varepsilon \|y' - y''\|;$$
- iii) $\Lambda Y = Z$.

Então, existe um número $k > 0$, uma vizinhança U do ponto (x_o, y_o) em $X \times Y$ e uma aplicação $\varphi : U \longrightarrow Y$ tais que

- a) $\phi(x, \varphi(x, z)) = z$;
- b) $\|\varphi(x, z) - y_o\| \leq k \|\phi(x, z_o) - z\|$.

□

A demonstração deste teorema pode ser vista em [1].

Definição 2.4.2 *Seja $M \subset E$. O elemento $h \in E$ é dito vetor tangente a M no ponto $x_o \in M$ se existe um número $\varepsilon > 0$ e uma aplicação $r : [-\varepsilon, \varepsilon] \longrightarrow E$ tais que*

- i) $x_o + \alpha h + r(\alpha) \in M$,
- ii) $r(\alpha) = o(\alpha)$ para $\alpha \rightarrow 0$.

O elemento $h \in M$ é dito vetor tangente unilateral (ou direção tangente) a M no ponto $x_o \in M$ se existem $\varepsilon > 0$ e uma aplicação $r(0, \varepsilon) \longrightarrow E$ tais que i) se verifica e ii) se verifica para $\alpha \downarrow 0$.

O conjunto de todos os vetores tangentes a M no ponto x_o é denotado por $T_{x_o}M$, e o conjunto de todos os vetores tangentes unilaterais é denotado por $T_{x_o}^+M$.

Da definição segue

$$T_{x_o}M \subseteq T_{x_o}^+M \quad \text{e}$$

$$T_{x_o}M = T_{x_o}^+M \cap (-T_{x_o}^+M).$$

Quando o conjunto $T_{x_o}M$ é um subespaço de E ele é chamado de **espaço tangente a M no ponto x_o** .

Exemplo 2.4.1 Se $E = Y \times \mathbb{R}$, onde Y é um espaço normado e $M = \{(y, z) : z = \|y\|\}$, então

$$T_{(0,0)}M = \{0\} \quad T_{(0,0)}^+M = M.$$

Teorema 2.4.2 (Lyusternik) Sejam X, Z espaços de Banach, U uma vizinhança do ponto $x_o \in X$, $P : U \rightarrow Z$ tal que $P(x_o) = 0$. Se P é estritamente diferenciável no ponto x_o e $P'(x_o)X = Z$, então o conjunto $M = \{x : P(x) = 0\}$ possui no ponto x_o um espaço tangente

$$T_{x_o}M = T_{x_o}^+M = \ker P'(x_o) = \{h : P'(x_o)h = 0\}$$

Demonstração:

Primeiramente, mostraremos, em (a), que $T_{x_o}^+M \subset \ker P'(x_o)$, e depois, em (b), que $\ker P'(x_o) \subset T_{x_o}^+M$.

(a) Seja $h \in T_{x_o}^+M$. Então,

$$P(x_o + \alpha h + r(\alpha)) = 0$$

pois, pela definição de vetor tangente unilateral, temos que $x_o + \alpha h + r(\alpha) \in M$.

Como P é estritamente diferenciável, ela também é Gâteaux-diferenciável, e como $(h + \frac{r(\alpha)}{\alpha}) \in X$, temos

$$\begin{aligned} P(x_o + \alpha h + r(\alpha)) &= P(x_o) + \alpha P'(x_o)h + o(\alpha) \\ &= \alpha P'(x_o)h + o(\alpha). \end{aligned}$$

Logo, $\alpha P'(x_o)h + o(\alpha) = 0$, e passando o limite,

$$P'(x_o)h = 0.$$

Portanto,

$$T_{x_o}^+ M \subset \ker P'(x_o).$$

(b) Agora, vamos aplicar o teorema da função implícita à função

$$\phi(x, y) = P(x + y).$$

Como P é estritamente diferenciável em x_o , ela é contínua em uma vizinhança de x_o , logo $x \mapsto \phi(x, 0) = P(x)$ é contínua em x_o , e

$$\|\phi(x, y') - \phi(x, y'') - P'(x_o)(y' - y'')\| \leq \varepsilon \|y' - y''\|,$$

quando $x \in B(x_o, \delta)$, $y', y'' \in B(0, \delta)$.

Temos também que $P'(x_o)X = Z$.

Assim, as três condições do teorema da função implícita são satisfeitas, Bastando considerar $X = Y$ no teorema.

Logo, existe uma vizinhança $U \subset X$ do ponto x_o tal que $\psi(x, \varphi(x)) = 0$ se, e somente se, $P(x + \varphi(x)) = 0$,

$$\|\varphi(x)\| \leq k \|\phi(x, 0)\| = k \|P(x)\|.$$

Fazendo $r(\alpha) = \varphi(x_o + \alpha h)$, obtemos:

$$P(x_o + \alpha h + r(\alpha)) = P(x_o + \alpha h + \varphi(x_o + \alpha h)) = 0,$$

$$\begin{aligned}
\|r(\alpha)\| = \|\varphi(x_o + \alpha h)\| &\leq k\|P(x_o + \alpha h)\| \\
&= k\|P(x_o + \alpha h) - P(x_o)\| \\
&= k\|P'(x_o)\alpha h + o(\|\alpha h\|)\| \\
&= k\|\alpha P'(x_o)h + o(\alpha)\|,
\end{aligned}$$

e se $h \in \ker P'(x_o)$, então $\|r(\alpha)\| = o(\alpha)$.

Logo, $\ker P'(x_o) \subset T_{x_o}M \subset T_{x_o}^+M$.

□

Observemos que o teorema de Lyusternik é uma ferramenta poderosa para se obter o cone de direções tangentes, o que é feito, basicamente, pelo cálculo da derivada estrita do funcional associado. Porém, não é possível utilizá-lo no caso em que o operador associado não é estritamente diferenciável, ou quando a sua derivada não é sobrejetora. Na prática, a dificuldade se concentra na condição de sobrejetividade do operador derivada. Na terminologia de otimização, esta é uma condição de regularidade. Para o caso não-regular, isto é, quando a derivada não for sobrejetora, o capítulo seguinte terá uma abordagem, devido a Avakov [13], que pode ser vista, em certo sentido, como uma generalização do teorema de Lyusternik.

Vamos agora apresentar um exemplo onde utilizaremos o teorema de Lyusternik para determinar o cone das direções tangentes, que será denotado por K , para o conjunto Q no ponto x_o .

Exemplo 2.4.2 *Seja $x \in \mathbb{R}^m$, $Q = \{x : g_i(x) = 0, i = 1, \dots, n\}$ onde $g_i(x)$ são funções estritamente diferenciáveis no ponto x_o , $g_i(x_o) = 0$, $i = 1, \dots, n$, e os vetores $g'_i(x_o)$ são linearmente independentes. Então,*

$$K = \{h \in \mathbb{R}^n : g'(x_o)h = 0, i = 1, \dots, n\}$$

De fato, se considerarmos $X = \mathbb{R}^n$, $Z = \mathbb{R}^n$ e $P(x) = (g_1(x), \dots, g_n(x))$, temos $P(x_o) = 0$, $P'(x_o)$ é a matriz cuja i -ésima linha é o vetor $g'_i(x_o)$. Como os n vetores $g'_i(x_o)$

são linearmente independentes, $P'(x_o)\mathbb{R}^m = \mathbb{R}^n$. Assim todas as hipóteses do teorema de Lyusternik são satisfeitas, e podemos concluir o resultado acima.

2.5 Cálculo dos cones duais

Logo na primeira seção, no teorema de Dubovitskii-Milyutin, vimos que, sob certas condições de regularidade do funcional $f(x)$ e do conjunto factível Q no ponto x_o , a equação de Euler-Lagrange é satisfeita. Contudo para que possamos utilizá-la precisamos determinar os cones duais, de onde sairão os funcionais f_i , $i = 1, \dots, n$, que a satisfarão. Este será o objetivo desta seção.

Começamos pelo caso mais simples, quando o cone K é um subespaço.

Teorema 2.5.1 *Suponha que o cone K é um subespaço do espaço vetorial normado E . Então*

$$K^* = \{f \in E^* : f(x) = 0, \forall x \in K\}.$$

Demonstração:

Já sabemos que K é um cone e seu cone dual é dado por

$$K^* = \{f \in E^* : f(x) \geq 0, \forall x \in K\}.$$

Suponha que existe $x \in K$ tal que $f(x) > 0$. Então,

$$f(-x) = -f(x) < 0.$$

Mas, $-x \in K$, pois K é subespaço. Logo,

$$f(-x) \geq 0,$$

o que é uma contradição. Portanto,

$$K^* = \{f \in E^* : f(x) = 0, \forall x \in K\}.$$

□

É importante observar que, se $K = E$, então $K^* = \{0\}$ e não E^* .

Teorema 2.5.2 *Sejam $f \in E^*$,*

$$K_1 = \{x : f(x) = 0\}$$

$$K_2 = \{x : f(x) \geq 0\}$$

$$K_3 = \{x : f(x) > 0\}.$$

Então,

$$K_1^* = \{\lambda f : f(x) - \infty < \lambda < \infty\}$$

$$K_2^* = \{\lambda f : 0 \leq \lambda < \infty\}$$

$$K_3^* = \begin{cases} E^*, & \text{se } f = 0 \\ K_2^*, & \text{se } f \neq 0. \end{cases}$$

Demonstração:

Seja $g \in K_1^*$. Como $K_1 = \ker f$ é subespaço de E , pelo teorema 2.5.1 temos que $g(x) = 0$, para todo $x \in K_1$ e, conseqüentemente, $g = \lambda f$, com $\lambda \in \mathbb{R}$.

Como $K_1 \subseteq K_2$, temos $K_2^* \subseteq K_1^*$. Logo, se $g \in K_2^*$, temos

$$g(x) = \lambda f(x) \geq 0, (\forall x \in K_2).$$

Já que $f(x) \geq 0$ sobre K_2 , concluímos que $\lambda \geq 0$.

Finalmente, se $f = 0$, temos $K_3 = \emptyset$ e portanto $K_3^* = E^*$. Sendo $f = 0$, temos $\bar{K}_3 = K_2$. Assim,

$$K_3^* = \bar{K}_3^* = K_2^*.$$

□

Teorema 2.5.3 *Seja $E = E_1 \times E_2$, onde E_1, E_2 são espaços de Banach e $A : E_1 \rightarrow E_2$ uma aplicação linear. Se*

$$K = \{x \in E : x = (x_1, x_2), Ax_1 = x_2\}$$

então

$$K^* = \{f \in E^* : f = (f_1, f_2), f_1 = A^* f_2\}.$$

Demonstração:

Já sabemos que K é um subespaço de E e, assim, pelo teorema 2.5.1, se $f \in K^*$, temos

$$\begin{aligned} 0 = \langle f, x \rangle &= \langle f_1, x_1 \rangle + \langle f_2, x_2 \rangle \\ &= \langle f_1, x_1 \rangle + \langle f_2, Ax_1 \rangle \\ &= \langle f_1, x_1 \rangle + \langle A^* f_2, x_1 \rangle \\ &= \langle f_1 + A^* f_2, x_1 \rangle, (\forall x \in K). \end{aligned}$$

Logo, $f_1 + A^* f_2 = 0$, isto é,

$$K^* \subset \{f \in E^* : f = (f_1, f_2), f_1 = -A^* f_2\}.$$

Agora, se $f = (-A^* f_2, f_2)$, então

$$\begin{aligned} \langle f, x \rangle &= \langle -A^* f_2, x_1 \rangle + \langle f_2, x_2 \rangle \\ &= \langle f_2, Ax_1 \rangle + \langle f_2, x_2 \rangle \\ &= -\langle f_2, Ax_1 \rangle + \langle f_2, x_2 \rangle \\ &= -\langle f_2, Ax_1 - x_2 \rangle \\ &= -\langle f_2, 0 \rangle = 0. \end{aligned}$$

o que prova a inclusão contrária. Portanto,

$$K^* = \{f \in E^* : f = (f_1, f_2), f_1 = -A^* f_2\}.$$

□

Com as mesmas notações do teorema acima, apresentaremos agora um teorema que nos levará a determinar cones duais de cones mais gerais.

Teorema 2.5.4 (Minkowski-Farkas) *Seja K_2 um cone convexo em E_2 com vértice em 0, $K_1 = \{x_1 \in E_1 : Ax_1 \in K_2\}$, e suponha que uma seguintes condições é satisfeita:*

- (i) *Existe $\tilde{x}_1 \in E_1$ tal que $A\tilde{x}_1 \in \text{int } K_2$;*
- (ii) *E_1, E_2 são finito dimensionais e K_2 é o octante positivo de E_2 , isto é,*

$$K_1 = \{x \in E_1 : Ax \geq 0\},$$

onde A é uma $n \times m$ matriz, $m = \dim E_1$ e $n = \dim E_2$.

Então

$$\begin{aligned} K_1^* &= A^* K_2^* \quad (\text{no caso (i)}), \\ K_1^* &= \{A^* y : y \geq 0\} \end{aligned}$$

Demonstração:

Caso (i): Seja

$$\begin{aligned} K &= \{x \in E : x = (x_1, x_2), x_2 \in K_2\} \\ L &= \{x \in E : x = (x_1, x_2), Ax_1 = x_2\}. \end{aligned}$$

Então K é um cone convexo com vértice em 0 e L é um subespaço tal que $\text{int } K \cap L \neq \emptyset$. Logo, pelo teorema [C.0.16],

$$(K \cap L)^* = K^* + L^*.$$

Seja $\tilde{f} = (\tilde{f}_1, 0)$, onde $\tilde{f}_1 \in K_1^*$. Então, $\tilde{f} \in (K \cap L)^*$. De fato, se $x \in K \cap L$, então $x = (x_1, Ax_1)$, $x_1 \in K_1$ e $\tilde{f}(x) = (\tilde{f}_1(x_1), 0) = \tilde{f}_1 \geq 0$. Logo, $\tilde{f} \in K^* + L^*$.

Como

$$\begin{aligned} K^* &= \{f = (f_1, f_2) : f_1 = 0, f_2 \in K_2^*\} \text{ e} \\ L^* &= \{f = (f_1, f_2) : f_1 = -A^*f_2\}, \end{aligned}$$

temos

$$\begin{aligned} (\tilde{f}_1, 0) &= (0, f_2) + (-A^*g_2, g_2) \\ &= (-A^*g_2, f_2 + g_2) \end{aligned}$$

onde $(0, f_2) \in K^*$ e $(-A^*g_2, g_2) \in L^*$.

Logo,

$$\tilde{f}_1 = -A^*g_2 = A^*f_2,$$

isto é, existe $\tilde{f}_2 \in K_2^*$ tal que $\tilde{f}_1 = A^*\tilde{f}_2$, e portanto,

$$K_1^* \subseteq A^*K_2^*.$$

Para provar a inclusão inversa, seja $f_1 = A^*f_2$, $f_2 \in K_2^*$. Então, se $x_1 \in K_1$, temos

$$\langle f_1, x_1 \rangle = \langle A^*f_2, x_1 \rangle = \langle f_2, Ax_1 \rangle \geq 0.$$

Portanto, $f_1 \in K_1^*$.

Caso (ii): Seja $Q_i = \{x; (A_i, x) \geq 0\}$, onde A_i é a i -ésima linha da matriz A . Então,

$$K_1 = \bigcap_{i=0}^n Q_i$$

Assim,

$$Q_i^* = \{A_i y_i : y_i \geq 0\}$$

e

$$\sum_{i=1}^n Q_i^* = \left\{ \sum_{i=1}^n A_i y_i : y_i \geq 0, i = 1, \dots, n \right\}.$$

O segundo membro da igualdade é um conjunto fechado e, portanto, pelo teorema [C.0.13]

$$\left(\bigcap_{i=1}^n Q_i \right)^* = \sum_{i=1}^n Q_i^*$$

o que implica

$$K_1^* = \left\{ \sum_{i=1}^n A_i y_i : y_i \geq 0, i = 1, \dots, n \right\}$$

Na realidade, as condições (i) e (ii) são casos particulares do seguinte fato: o cone $A^* K_2^*$ é w^* -fechado em E_1^* , o que também implica a igualdade $K^* = A^* K_2^*$.

Exemplo 2.5.1 *Sejam $x \in \mathbb{R}^n$, $y_1 \in \mathbb{R}^k$, $y_2 \in \mathbb{R}^{n-k}$, A_1 e A_2 matrizes $k \times n$ e $(n-k) \times n$, respectivamente, A_1^* e A_2^* suas transpostas.*

Se

$$K = \{x : A_1 x \geq 0, A_2 x = 0\},$$

então

$$K^* = A_1^* y_1 + A_2^* y_2, \quad y_1 \geq 0.$$

Exemplo 2.5.2 *Seja*

$$K = \{x \in E : x(t) \geq 0, t \in \mathbb{R}\}$$

onde $E = (C(0, T))^n$ e $R \subseteq [0, T]$ é fechado.

Então, para qualquer $f \in K^*$, existe uma medida vetorial não-negativa $d\mu(t)$ com suporte R tal que

$$f(x) = \int_0^T (x(t), d\mu(x)) = \int_R (x(t), d\mu(t)). \quad (2.4)$$

Sejam $E = (C(0, T))^n$, $R \subseteq [0, T]$ fechado, $h(t) \in C(0, T)$, $h(t) \neq 0$ em R e

$$K = \{x \in C(0, T) : \langle h(t), x(t) \rangle \geq 0, t \in \mathbb{R}\}$$

Então, se $f \in K^*$ existe uma medida escalar não-negativa com suporte em R , tal que

$$\begin{aligned} f(x) &= \int_0^T \langle h(t), x(t) \rangle d\mu(t) \\ &= \int_R \langle h(t), x(t) \rangle d\mu(t). \end{aligned}$$

Para provar isto vamos considerar $E_1 = (C(0, T))^n$, $E_2 = C(0, T)$, $A : E_1 \rightarrow E_2$ uma aplicação linear definida por

$$Ax = \langle x(t), h(t) \rangle,$$

e

$$K_2 = \{y(t) \in C(0, T) : y(t) \geq 0, t \in \mathbb{R}\}$$

Logo, $K_2 = AK$. Se $\tilde{x}(t) = h(t)$, então $A\tilde{x} = |h(x)|^2 > 0$ sobre R , e a condição (i) do teorema de Minkowski-Farkas se satisfaz, isto é, $A\tilde{x} \in \text{int } K_2$.

De acordo com o que vimos no início, dado $f_2 \in K_2^*$ existe uma medida vetorial não-negativa $d\mu(t)$ tal que

$$f_2(x) = \int_0^T \langle x(t), d\mu(t) \rangle = \int_R \langle x(t), d\mu(t) \rangle$$

Usando o teorema 2.5.4, então, qualquer que seja $f \in K^*$, existe $f_2 \in K_2^*$ tal que $f = A^*f_2$, ou seja,

$$\begin{aligned} \langle f, x \rangle &= \langle A^*f_2, x \rangle \\ &= \langle f_2, Ax \rangle \\ &= \int_0^T \langle h(t), x(t) \rangle d\mu(t). \end{aligned}$$

Definição 2.5.1 Um funcional suporte para um conjunto $A \subseteq E$ em $x_o \in A$ é uma forma linear contínua $f \in E^*$ não-nula tal que

$$f(x) \geq f(x_o) \quad (\forall x \in A)$$

Sejam $Q \subseteq E$, $x_o \in Q$, K_b o cone de direções factíveis para Q em x_o , K_k o cone de direções tangentes a Q em x_o e Q^* o cone dos funcionais lineares suportes para Q em x_o , isto é,

$$Q^* = \{f \in E^* : f(x) \geq f(x_o), \forall x \in Q\}$$

Teorema 2.5.5 Se Q é um conjunto convexo fechado, então $K_k^* = Q^*$. Se, além disso, $\text{int } Q \neq \emptyset$, então $K_b^* = Q^*$.

Demonstração:

Mostraremos, primeiramente, que $Q^* \subseteq K_k^*$.

Seja $f \in Q^*$, $h \in K_k$. Como h é uma direção tangente a Q em x_o , dado $0 < \varepsilon < \varepsilon_o$, existe $x(\varepsilon) \in Q$ tal que

$$x(\varepsilon) = x_o + \varepsilon h + r(\varepsilon) \quad \text{e} \quad \frac{1}{\varepsilon} r(\varepsilon) \in U$$

para qualquer vizinhança U de 0 , e ε_o suficientemente pequeno. Pela definição de Q^* temos

$$\langle f, x(\varepsilon) \rangle \geq \langle f, x_o \rangle,$$

isto é, $\langle f, x(\varepsilon) \rangle - \langle f, x_o \rangle \geq 0$, e

$$\begin{aligned} \langle f, x \rangle &= \left\langle f, \frac{x(\varepsilon) - x_o - r(\varepsilon)}{\varepsilon} \right\rangle \\ &= \frac{1}{\varepsilon} \left(\langle f, x(\varepsilon) \rangle - \langle f, x_o \rangle \right) - \frac{1}{\varepsilon} \langle f, r(\varepsilon) \rangle \\ &\geq - \left\langle f, \frac{r(\varepsilon)}{\varepsilon} \right\rangle \end{aligned}$$

Como f é contínua, temos

$$\lim_{\varepsilon \rightarrow 0} \left\langle f, \frac{r(\varepsilon)}{\varepsilon} \right\rangle = \langle f, 0 \rangle = 0.$$

Logo,

$$\langle f, h \rangle \geq 0,$$

e portanto, $f \in K_k^*$, isto é,

$$Q^* \subseteq K_k^*.$$

Mostremos agora que $K^* \subseteq Q$. Seja $f \in K_k^*$. Como Q é convexo, se $x \in Q$ e $h = x - x_o$ temos

$$\begin{aligned} x_o + \varepsilon h &= x_o + \varepsilon(x - x_o) \\ &= (1 - \varepsilon)x_o + \varepsilon x \in Q, \end{aligned}$$

para $0 < \varepsilon < 1$, logo h é uma direção tangente a Q em x_o .

Como $f \in K_k^*$, temos

$$\begin{aligned} 0 \leq \langle f, h \rangle &= \langle f, x - x_o \rangle \\ &= \langle f, x \rangle - \langle f, x_o \rangle \end{aligned}$$

Logo, $\langle f, x \rangle \geq \langle f, x_o \rangle$, isto é, $f \in Q^*$. Portanto,

$$K_k^* \subseteq Q^*,$$

e assim,

$$Q^* = K_k^*.$$

Se $\text{int } Q \neq \emptyset$, então, pela observação feita após o corolário 1.3.1,

$$K_b = \{h : h = \lambda(x - x_o), \lambda \geq 0, x \in \text{int } Q\}.$$

Também, se $f \in K_b^*$ então

$$\langle f, \lambda(x - x_o) \rangle \geq 0,$$

isto é,

$$\lambda \langle f, x \rangle \geq \lambda \langle f, x_o \rangle$$

para todo $\lambda \geq 0$ e todo $x \in \text{int } Q$. Fazendo $\lambda = 1$ obtemos

$$\langle f, x \rangle \geq \langle f, x_o \rangle, \quad (\forall x \in \text{int } Q),$$

e como f é contínua, isto significa que

$$\langle f, x \rangle \geq \langle f, x_o \rangle, \quad (\forall x \in \bar{Q}),$$

pois, como $\text{int } Q \neq \emptyset$ e Q é convexo, temos $\overline{\text{int } Q} = \bar{Q}$. Logo, $f \in Q^*$, de onde se conclui que

$$K_b^* \subseteq K_k^* = Q^*.$$

Por outro lado, temos que $K_b \subseteq K_k$, de onde

$$Q^* = K_k^* \subseteq K_b^*.$$

Portanto,

$$K_b^* = Q^*.$$

□

Apresentamos agora dois exemplos nos quais podemos observar que para calcularmos os cones duais basta determinar os funcionais suportes.

Exemplo 2.5.3 *Seja Q dado por*

$$Q = \{x \in \mathbb{R}^m : \langle a^i, x \rangle \geq b_i, \quad i = 1, \dots, n\}$$

onde $a^i \in \mathbb{R}^m$, $b_i \in \mathbb{R}$, $i = 1, \dots, n$.

Então,

$$K_k^* = Q^* = \left\{ \sum_{i=1}^n \lambda_i a^i : \lambda_i \geq 0, \lambda_i [\langle a^i, x_o \rangle - b_i] = 0 \right\}$$

Demonstração:

Mostremos, primeiramente, que

$$\left\{ \sum_{i=1}^n \lambda_i a^i : \lambda_i \geq 0, \lambda_i [\langle a^i, x_o \rangle - b_i] = 0 \right\} \subseteq Q^*.$$

Seja $f = \sum_{i=1}^n \lambda_i a^i$, $\lambda_i \geq 0$, $\lambda_i [\langle a^i, x_o \rangle - b_i] = 0$. Então, qualquer que seja $x \in Q$,

$$\begin{aligned} \langle f, x \rangle &= \sum_{i=1}^n \lambda_i \langle a^i, x \rangle \\ &\geq \sum_{i=1}^n \lambda_i b_i \\ &= \sum_{i=1}^n \lambda_i \langle a^i, x_o \rangle \\ &= \langle f, x_o \rangle. \end{aligned}$$

ou seja,

$$\langle f, x - x_o \rangle \geq 0.$$

Logo, $f \in Q^*$.

Provemos agora que

$$K_k \supseteq \{h : \langle a^i, h \rangle \geq 0, i \in I\}$$

onde $I = \{i : \langle a^i, x_o \rangle = b_i\}$, ou seja, que

$$K_k^* \subseteq \{h : \langle a^i, h \rangle \geq 0, i \in I\}^*$$

Seja h tal que $\langle a^i, h \rangle \geq 0, i \in I$, e tomemos

$$\varepsilon_o = \min_{i \notin I} \frac{\langle a^i, x_o \rangle - b_i}{\|a^i\| \|h\|} > 0.$$

Como

$$\langle a^i, x_o + \varepsilon h \rangle = \langle a^i, x_o \rangle + \varepsilon \langle a^i, h \rangle \geq \langle a^i, x_o \rangle = b_i, \quad i \in I,$$

e

$$\begin{aligned} \langle a^i, x_o + \varepsilon h \rangle &= \langle a^i, x_o \rangle + \varepsilon \langle a^i, h \rangle \\ &\geq \langle a^i, x_o \rangle - \varepsilon \|a^i\| \|h\| \\ &> b_i, \quad i \notin I, \end{aligned}$$

temos que $x_o + \varepsilon h \in Q$ para $0 < \varepsilon < \varepsilon_o, h \in K_k$. Logo,

$$K_k^* \subseteq \{h : \langle a^i, h \rangle \geq 0, i \in I\}^*$$

Além disso,

$$\{h : \langle a^i, h \rangle \geq 0, i \in I\}^* = \left\{ \sum_{i=1}^n \lambda_i a^i : \lambda_i \geq 0, i \in I \right\}.$$

Portanto,

$$K_k^* \subseteq \left\{ \sum_{i=1}^n \lambda_i a^i : \lambda_i \geq 0, i \in I \right\}.$$

E como, pelo teorema 2.5.5, $K_k^* = Q^*$, temos

$$K_k^* = Q^* = \left\{ \sum_{i=1}^n \lambda_i a^i : \lambda_i \geq 0, \lambda_i [\langle a^i, x_o \rangle - b_i] = 0 \right\}.$$

□

Matricialmente, o que obtivemos neste exemplo foi que, se

$$Q = \{x : Ax \geq b\}$$

então,

$$Q^* = K_k^* = \{A^*y : y \geq 0, y_i [Ax_o - b]_i = 0, i = 1, \dots, n\}.$$

Exemplo 2.5.4 *Seja*

$$Q = \{x \in L_\infty^{(r)}(0, T) : x(t) \in M \forall t \in [0, T]\},$$

onde $M \subset \mathbb{R}^r$.

Devido à complexidade da estrutura do espaço $L_\infty^{(r)}(0, T)$, torna-se difícil a determinação do conjunto dos funcionais suportes a Q em um ponto. Contudo, é possível descrever a forma geral dos funcionais suportes dados por integrais. Mostremos que se

$$f(x) = \int_0^T (a(t), x(t)) dt, \quad a(t) \in L_1^{(r)}(0, T),$$

é um suporte a Q no ponto x_o , então

$$(a(t), x - x_o(t)) \geq 0$$

para todo $x \in M$ e todo $t \in [0, T]$.

Suponha falsa a afirmação acima. Então, existe um subconjunto $R_1 \subset [0, T]$, com $\mu(R_1) \neq 0$ (onde $\mu(R)$ é a medida de Lebesgue de R), tal que, para todo $t \in R_1$, existe $\tilde{x}(t) \in M$ com $(a(t), \tilde{x} - x_o(t)) < 0$. Pelo teorema de Lusin, existem subconjuntos $R_2 \subset [0, T]$, com $\mu(R_2) < \varepsilon/2$, $R_3 \subset [0, T]$, com $\mu(R_3) < \varepsilon/2$, tais que $a(t)$ é contínua em $[0, T] - R_2$ e $x_o(t)$ é contínua em $[0, T] - R_3$. Como

$$\mu(R_2) + \mu(R_3) < \mu(R_1),$$

existe um ponto $t_o \in R_1$ tal que $t_o \notin R_2 \cup R_3$.

Seja $a(t)$ e $x_o(t)$ contínuas em t_o e $(a(t), \tilde{x}(t_o) - x_o(t_o)) = \gamma < 0$, existe $R_4 \subset [0, T]$, com $\mu(R_4) > 0$, tal que

$$(a(t), \tilde{x}(t_o) - x_o(t)) \leq \gamma/2$$

para $t \in R_4$.

Seja a função

$$x_1(t) = \begin{cases} \tilde{x}(t_o) & \text{se } t \in R_4 \\ x(t) & \text{se } t \in [0, T] - R_4 \end{cases}$$

Então, $x_1 \in Q$ e

$$f(x_1) = \int_0^T (a(t), x_o(t)) dt + \int_{R_4} (a(t), \tilde{x}(t_o) - x_o(t)) dt \leq f(x_o) + \frac{\gamma}{2} \mu(R_4).$$

Logo, $f(x_1) < f(x_o)$, o que contradiz o fato de ser f um funcional suporte para Q em x_o .

2.6 Multiplicadores de Lagrange e o teorema de Kuhn-Tucker

Nesta seção aplicaremos a teoria desenvolvida neste capítulo à programação matemática,

o que resultará na já conhecida regra dos multiplicadores de Lagrange e o teorema de Kuhn-Tucker. Faremos isto em dois casos: um com restrições de igualdade e o outro com restrições de desigualdade.

(1) **Extremos condicionados** (com restrições de igualdade).

Nosso problema será

$$(P.1) \begin{cases} \min F_o(x) \\ \text{s.a. } F_i(x) = 0, \quad i = 1, \dots, n, \end{cases}$$

Onde E é um espaço de Banach, $x \in E$, $F_o(x), \dots, F_n(x)$ são funcionais sobre E .

Teorema 2.6.1 *Seja x_o uma solução do problema (P.1), e suponha que os funcionais $F_i(x)$, $i = 1, \dots, n$, são estritamente diferenciáveis em uma vizinhança de x_o . Então, existem escalares $\lambda_o, \lambda_1, \dots, \lambda_n$, não simultaneamente nulos, com $\lambda_o \geq 0$, tais que*

$$\lambda_o F'_o(x_o) + \lambda_1 F'_1(x_o) + \dots + \lambda_n F'_n(x_o) = 0 \quad (2.5)$$

Demonstração:

Se $F'_i(x_o)$, $i = 1, \dots, n$, são linearmente dependentes temos

$$\gamma_1 F'_1(x_o) + \dots + \gamma_n F'_n(x_o) = 0 \quad \text{com} \quad \sum_{i=1}^n \gamma_i^2 \neq 0$$

Tomamos, então, $\lambda_o = 0$ e $\lambda_i = \gamma_i$, $i = 1, \dots, n$, e temos (1.5) satisfeito.

Consideremos, então, o caso em que $F'_i(x_o)$, $i = 1, \dots, n$, são linearmente independentes. Como $Q = \{x : F_i(x) = 0, \quad i = 1, \dots, n\}$ e cada $F'_i(x_o)$, $i = 1, \dots, n$, é estritamente diferenciável, pelo já visto na seção 1.2, temos que o cone de direções de descida para $F_o(x)$ no ponto x_o é dado por

$$K_o = \{h : (F'_o(x_o), h) < 0\}$$

e o seu dual, de acordo com o teorema 2.2.1, por

$$K_o^* = \{\lambda_o F'_o(x_o) : 0 \leq \lambda_o\}$$

desde que $F'_o(x_o) \neq 0$, pois no caso $F'_o(x_o) = 0$ tomamos $\lambda_o = 1$, $\lambda_i = 0$, $i = 1, \dots, n$.

Pelo exemplo 2.4.2, o cone de direções tangentes a Q em x_o é dado por

$$K_1 = \{h : \langle F'_i(x_o), h \rangle = 0, i = 1, \dots, n\},$$

e seu dual, de acordo com o exemplo 2.5.1, por

$$K_1^* = \left\{ \sum_{i=1}^n \lambda_i F'_i(x_o), \lambda_i \geq 0 \right\}.$$

Agora, é só aplicar o teorema de Dubovitskii- Milyutin.

Conforme já observamos na demonstração, se $F'_o(x_o) = 0$ podemos tomar $\lambda_o = 1$, e se $F'_o(x_o) \neq 0$, como $K_o^* = \{\lambda_o F'_o(x_o) : 0 \leq \lambda_o\}$, podemos tomar $\lambda_o \neq 0$, e portanto, $\lambda_o = 1$. Teremos então

$$F'_o(x_o) + \lambda_1 F'_1(x_o) + \dots + \lambda_n F'_n(x_o) = 0.$$

Em particular, quando não existem restrições, temos as condições usuais para um extremo não codicionado, $F'_o(x_o) = 0$.

Os escalares $\lambda_1, \dots, \lambda_n$, são chamados *multiplicadores de Lagrange*.

(2) Problemas de programação linear (restrições de desigualdade)

Nosso problema agora é

$$(P_2) \begin{cases} \min F_o(x) \\ \text{s.a } F_i(x) \leq 0, i = 1, \dots, n, \end{cases}$$

onde E é um espaço de Banach, $x \in E$ e $F_o(x), F_1(x), \dots, F_n(x)$ são funcionais sobre E .

Teorema 2.6.2 Se x_o é uma solução do problema (P.2) e os funcionais $F_i(x)$, $i = 0, \dots, n$, são diferenciáveis em x_o . Então, existem escalares λ_i , $i = 0, \dots, n$, não todos nulos, tais que

$$\lambda_o F'_o(x_o) + \lambda_1 F'_1(x_o) + \dots + \lambda_n F'_n(x_o) = 0 \quad i = 1, \dots, n,$$

com

$$\begin{aligned} \lambda_i &\geq 0, \quad i = 0, \dots, n \\ \lambda_i F_i(x_o) &= 0, \quad i = 1, \dots, n. \end{aligned} \tag{2.6}$$

Demonstração:

Como o que temos agora são restrições de desigualdade, nosso objetivo é determinar, além do cone de direções de descida para $F_o(x)$, os cones de direções factíveis para os conjuntos

$$Q_i = \{x : F_i(x) \leq 0\} \quad i = 1, \dots, n,$$

no ponto x_o , e seus duais. Tais cones serão denotados, respectivamente, por K_i e K_i^* , $i = 1, \dots, n$.

Se $F_i(x_o) < 0$, então $x_o \in \text{int } Q_i$ e, pelo que vimos na seção 1.3, $K_i = E$. No caso em que $F_i(x_o) = 0$ e $F'_i(x_o) \neq 0$, temos $K_i = \{h : (F'_i(x_o), h) < 0\}$, de acordo com o corolário 2.3.1. Logo, do teorema 2.2.1 concluímos que

i) se $F_i(x_o) < 0$, então $K_i^* = 0$,

ii) se $F_i(x_o) = 0$ e $F'_i(x_o) \neq 0$, então

$$K_i^* = \{\lambda_i F'_i(x_o) : 0 \leq \lambda_i\}$$

Quando $F_i(x_o) = 0$ e $F'_i(x_o) = 0$, tomamos $\lambda_i = 1$ e $\lambda_j = 0$ para $j \neq i$, e (1.6) é satisfeito.

O cone K_o das direções de descida para $F_o(x)$ em x_o é

$$K_o = \{h : (F'_o(x_o), h) < 0\}$$

$$K_o^* = \{\lambda_o F'_o(x_o) : 0 \leq \lambda_o\}$$

para $F'_o(x_o) \neq 0$.

Agora basta usar Dubovitskii-Milyutin para concluir o teorema.

O seguinte teorema nos indica duas situações em que podemos assumir $\lambda_o \neq 0$ (problemas normais, ou problemas não-degenerados).

Teorema 2.6.3 (Karush-Kuhn-Tucker) *Sob as hipóteses do teorema anterior, suponha que uma das seguintes condições é satisfeita:*

i) *Os funcionais $F_i(x)$, $i = 1, \dots, n$, são convexos e existe $\tilde{x}$ tal que $F_i(\tilde{x}) < 0$, $i = 1, \dots, n$.*

ii) *As restrições são lineares, isto é,*

$$F_i(x_o) = (a^i, x) - b_i, \quad a^i \in E', \quad b_i \in \mathbb{R}, \quad i = 1, \dots, n.$$

Então, $\lambda_o \neq 0$, $i = 1, \dots, n$, isto é,

$$\begin{aligned} \lambda_i &\geq 0, \quad i = 1, \dots, n, \\ \lambda_i F_i(x_o) &= 0, \quad i = 1, \dots, n \\ F'_o(x_o) + \lambda_1 F'_1(x_o) + \dots + \lambda_n F'_n(x_o) &= 0. \end{aligned} \tag{2.7}$$

Demonstração:

i) Pela prova do teorema de Dubovitskii-Milyutin vemos que uma condição para que $\lambda_o \neq 0$, é que

$$\bigcap_{i=1}^{n+1} K_i \neq \emptyset.$$

Como, neste caso, $F_i(x)$, $i = 1, \dots, n$, são convexos e $F_i(\tilde{x}) < 0$, $i = 1, \dots, n$, temos que $\tilde{x} - x_o$ é uma direção factível. De fato, como cada F_i é contínua e $F_i(\tilde{x}) < 0$, temos que $F_i^{-1}((0, -\infty))$ é aberto não-vazio. Portanto, $\text{int } Q = \{x \in E : F_i(x) < 0, i = 1, \dots, n\}$ é

um aberto não vazio e $\tilde{x} \in \text{int } Q$, isto é, existe uma bola aberta B de centro em $\tilde{x}$ e raio δ tal que $B \subset \text{int } Q$. Como cada F_i é convexo, temos que Q é convexo. Segue de $x_o \in Q$ e $B \subset \text{int } Q$ que para qualquer $y \in B$, o segmento $\lambda x_o + (1 - \lambda)y \subset Q$ ($0 \leq \lambda \leq 1$). Tomando $\varepsilon_o = \|\tilde{x} - x_o\| + \delta$, temos que $h = \tilde{x} - x_o$ é uma direção factível de Q em x_o .

ii) Neste caso, temos

$$Q = \{x : (a^i, x) \leq b_i, i = 1, \dots, n\}$$

e de acordo com o exemplo 1.5.3, o cone dual do cone de direções tangentes a Q em x_o é

$$K_1^* = \left\{ \sum_{i=1}^n \lambda_i a^i : \lambda_i ((a^i, x_o) - b_i) = 0, i = 1, \dots, n \right\}$$

Agora, é só aplicar o teorema de Dubovitskii-Milyutin.

□

O problema de programação linear foi primeiramente investigado por Kuhn e Tucker, entre outros, e as condições (1.7) são chamadas de condições de Karush-Kuhn-Tucker.

Concluindo, vamos juntar os teoremas 2.6.1 e 2.6.2 em um só, o que nos dará uma generalização.

Nosso problema agora será

$$(P.3) \begin{cases} \min F_o(x) \\ F_i(x) \leq 0, i = 1, \dots, k, \\ F_i(x) = 0, i = k + 1, \dots, n, \\ x \in Q \end{cases}$$

onde, além das restrições de igualdade e desigualdade, existe uma restrição adicional não necessariamente determinada por uma função.

Teorema 2.6.4 *Seja x_o uma solução de (P.3) e suponha que os funcionais $F_i(x)$, $i = 1, \dots, n$, são diferenciáveis em uma vizinhança de x_o , Q é convexo e $\text{int } Q \neq \emptyset$. Então,*

existem escalares λ_i , $i = 0, \dots, n$, não todos nulos, $\lambda_i \geq 0$, $i = 0, \dots, k$, $\lambda_i F_i(x_o) = 0$, $i = 1, \dots, k$, tais que

$$f = \sum_{i=0}^n \lambda_i F'_i(x_o)$$

é um funcional suporte para Q no ponto x_o .

□

Capítulo 3

Extensão para o caso não regular

Sejam X e Y espaços de Banach, $f : U_x \subseteq X \rightarrow \mathbb{R}$ e $F : X \rightarrow Y$, onde U_x é um aberto de X . Neste capítulo trataremos do problema

$$(P) \begin{cases} \min f(x) \\ \text{s.a } F(x) = 0 \end{cases}$$

De acordo com o princípio de Lagrange, as condições necessárias para um mínimo neste problema são as mesmas para um mínimo não-condicionado da função lagrangeana

$$\mathcal{L}(x, \lambda_o, y^*) = \lambda_o f(x) + \langle y^*, F(x) \rangle.$$

Existem casos em que o princípio de Lagrange não nos ajuda a resolver o problema (P) , pois nos pontos suspeitos de serem mínimos a função F não é regular e poderemos ter $\lambda_o = 0$.

Para exemplificar o que foi dito acima, consideremos o conjunto

$$Q = \{x \in \mathbb{R}^4; f_1(x) = 0, f_2(x) = 0\}$$

onde

$$f_1(x) = x_2^2 - 2x_1x_3 - x_4^4$$

$$f_2(x) = x_1^2 + 2x_1x_3 - 8x_3^2 - x_3^4$$

e denotemos o conjunto dos pontos regulares de Q (isto é, os pontos de Q onde os gradientes das funções f_1 e f_2 são l.i.) por R e o conjunto dos pontos irregulares de Q por IR . Então,

$$IR = \{x \in Q; x_1 = x_3 = 0, x_2 = \pm x_4^2\}.$$

Neste caso, as condições necessárias de Lagrange serão satisfeitas, com $\lambda_0 = \lambda_1 = 0$, e $\lambda_2 = 1$, por todos os elementos de IR , independentemente da função a ser minimizada sobre Q . Contudo, isto não nos interessa pois não poderemos relacionar o funcional a ser minimizado com as restrições do problema. Isto se deve ao fato de F não ser regular em IR .

Desta forma, torna-se necessário um novo desenvolvimento da regra de multiplicadores de Lagrange, que garanta também $\lambda_0 \neq 0$ para hipóteses mais fracas que a regularidade da função F . Este será o nosso objetivo a partir de agora.

3.1 Uma generalização do teorema de Lyusternik

Nesta seção consideraremos F uma função duplamente Fréchet-diferenciável no ponto $x_* \in X$. Ao longo de todo este capítulo faremos uso de uma função linear que definiremos agora.

Fixado um elemento $h \in X$ arbitrário, a função $G(x_*, h)$ é definida por

$$G(x_*, h) : X \longrightarrow \text{Im } F'(x_*) \times Y / \text{Im } F'(x_*),$$

$$G(x_*, h)[x] = (F'(x_*)[x], \pi \circ F''(x_*)[h, x]),$$

onde $\text{Im } F'(x_*)$ é a imagem de X pela função $F'(x_*)$, $Y / \text{Im } F'(x_*)$ é o espaço-quociente de Y em relação a $\text{Im } F'(x_*)$ e $\pi : Y \longrightarrow Y / \text{Im } F'(x_*)$ é a função projeção canônica.

Definimos também os conjuntos

$$\begin{aligned}
Z &= Y \times Y / \text{Im } F'(x_*) \\
Z_1 &= \text{Im } F'(x_*) \times Y / \text{Im } F'(x_*) \\
H(x) &= \{h \in X : F'(x)h = 0, F''(x)[h, h] \in \text{Im } F'(x)\}, \\
H_o(x) &= \{h \in H(x) : \text{Im } G(x, h) = Z_1\} \\
M(x_*) &= \{x \in X : F(x) = F(x_*)\}.
\end{aligned} \tag{3.1}$$

Teorema 3.1.1 *Sejam X e Y espaços de Banach, U uma vizinhança do ponto $x_* \in X$ e F uma função de U em Y duplamente Fréchet-diferenciável no ponto x_* tal que $\text{Im } F'(x_*)$ é um subespaço fechado de Y . Suponha que existe um elemento $h \in X$, $\|h\| = 1$, tal que $\text{Im } G(x_*, h) = Z_1$. Então, existem números $\varepsilon = \varepsilon(h)$, $k = k(h)$, e uma vizinhança $U' \subseteq U$ do ponto x_* tais que, quaisquer que sejam $x \in U'$ que satisfaça*

$$\left\| \frac{x - x_*}{\|x - x_*\|} - h \right\| \leq \varepsilon$$

e elementos arbitrários $y \in \text{Im } F'(x_)$, existe uma função $x \mapsto r(x)$ de U' em X tal que*

$$F(x + r(x)) = F(x_*), \tag{3.2}$$

$$\|r(x)\| \leq K[\|F(x) - F(x_*) - y\| \|x - x_*\|^{-1} + \|y\|]. \tag{3.3}$$

Demonstração:

Como, por hipótese, $\text{Im } F'(x_*)$ é fechado e Y é um espaço de Banach, pelo teorema do espaço quociente $Y / \text{Im } F'(x_*)$ é um espaço de Banach com a norma

$$\|\pi(y)\|_{Y/\text{Im } F'(x_*)} = \min_{\tilde{y} \in F'(x_*)} \|y - \tilde{y}\|.$$

Portanto, o espaço Z_1 é também um espaço de Banach com a norma

$$\|z\|_{Z_1} = \|(y, \pi(y))\| = \max\{\|y\|_Y, \|\pi(y)\|_{Y/\text{Im } F'(x_*)}\}$$

Assim, como $G(x_*, h)$ é uma aplicação linear contínua de X em Z_1 que é um epimorfismo, existe uma função $M : Z_1 \longrightarrow X$ e um número $C = C(h)$ tais que

$$G(x_*, h) \circ M = I_{Z_1}, \tag{3.4}$$

$$\|M(z)\| \leq C\|z\|. \tag{3.5}$$

Além disso, visto que F é duplamente Fréchet-diferenciável em x_* , existe um número $\delta > 0$ tal que, para quaisquer $x', x'' \in X$, $\|x' - x_*\| \leq \delta$, $\|x'' - x_*\| \leq \delta$, tem-se $B(x_*, \delta) = \{x \in X : \|x - x_*\| \leq \delta\} \subseteq U$ e

$$\begin{aligned} \|F(x') - F(x'') - F(x_*)[x' - x'']\| &= F''(x_*)[z, x' - x''] \\ &= \frac{1}{2} F''(x_*)[x' + x'' - 2(z + x_*), x' - x'']\| \\ &\leq (\|x' - x_*\| + \|x'' - x_*\|)\|x' - x''\|, \end{aligned} \quad (3.6)$$

onde z é um elemento arbitrário de X .

Consideremos agora as constantes

$$\varepsilon = (8\|F'(x_*)\|C)^{-1}, \quad (3.7)$$

$$K = \max\{C^2\|F'(x_*)\|16, 4C\}, \quad (3.8)$$

$$k = \min\left\{\frac{\delta}{2}, (2[1 + \|F'(x_*)\|C])^{-1}\right\}, \quad (3.9)$$

e sejam $x \in U' = B(x_*, k)$ tal que

$$\left\| \frac{x - x_*}{\|x - x_*\|} - h \right\| \leq \varepsilon$$

e um elemento $y \in \text{Im } F'(x_*)$. Tomemos $z = \|x - x_*\|h$, e

$$\Delta(x) = \|F(x) - F(x_*) - y\|\|x - x_*\|^{-1} + \|y\|.$$

Quando $\|x - x_*\| \leq K\Delta(x)$, fazendo $r(x) = x_* - x$ obtemos (2.2) e (2.3) imediatamente. Vamos então considerar

$$\|x - x_*\| \geq K\Delta(x). \quad (3.10)$$

Usaremos o método de aproximações sucessivas de Newton para construir $r(x)$. Façamos então $x_0 = x$, $y_0 = y$ e

$$x_n = x_{n-1} - M \left(y_{n-1}, \pi \left(\frac{F(x_{n-1}) - F(x_*) - y_{n-1}}{\|x_{n-1} - x_*\|} \right) \right), \quad n = 1, 2, \dots \quad (3.11)$$

$$y_n = F(x_n) - F(x_*) - y_{n-1} + F'(x_*)[z, x_n - x_{n-1}], \quad n = 1, 2, \dots \quad (3.12)$$

Aplicando a função $G(x_*, h)$ em ambos os membros de (2.11) e considerando (2.4), obtemos

$$G(x_*, h)[x_n] = G(x_*, h)[x_{n-1}] - \left(y_{n-1}, \pi \left(\frac{F(x_{n-1}) - F(x_*) - y_{n-1}}{\|x - x_*\|} \right) \right),$$

isto é,

$$G(x_*, h)[x_n - x_{n-1}] = \left(y_{n-1}, \pi \left(\frac{F(x_{n-1}) - F(x_*) - y_{n-1}}{\|x - x_*\|} \right) \right).$$

Como

$$G(x_*, h)[x_n - x_{n-1}] = (F'(x_*)[x_n - x_{n-1}], \pi \circ F''(x_*)[h, x_n - x_{n-1}]),$$

temos

$$\begin{aligned} F'(x_*)[x_n - x_{n-1}] &= -y_{n-1} \quad \text{e} \\ \pi(F''(x_*)[h, x_n - x_{n-1}]) &= \frac{-1}{\|x - x_*\|} \pi(F(x_{n-1}) - F(x_*) - y_{n-1}), \end{aligned} \quad (3.13)$$

isto é,

$$\pi(F''(x_*)[z, x_n - x_{n-1}]) = -\pi(F(x_{n-1}) - F(x_*) - y_{n-1}). \quad (3.14)$$

Vamos agora provar, por indução, a relação

$$F(x_n) - F(x_*) = \alpha_n + y_n, \quad (3.15)$$

$$\|\alpha_n\| \leq 2^{-n} \Delta(x) \|x - x_*\|, \quad (3.16)$$

$$y_n \in \text{Im } F'(x_*), \|y_n\| \leq 2^{-n} \Delta(x), \quad (3.17)$$

$$x_n \in B(x_*, \delta), \|x_n - x_{n-1}\| \leq 2^{1-n} C \Delta(x). \quad (3.18)$$

Tomemos $n = 1$ e utilizemos a igualdade (2.14):

$$\pi(F''(x_*)[z, x_1 - x]) = -\pi(F(x) - F(x_*) - y)$$

ou seja,

$$\pi(F''(x_*)[z, x_1 - x] + F(x) - F(x_*) - y) = 0.$$

Isto significa que $(F''(x_*)[z, x_1 - x] + F(x) - F(x_*) - y) \in \text{Im } F'(x_*)$, ou seja, existe $y_1 \in \text{Im } F'(x_*)$ tal que

$$y_1 = F''(x_*)[z, x_1 - x] + F(x) - F(x_*) - y \quad (3.19)$$

De (2.13) concluimos que

$$F'(x_*)[x_1 - x] = -y, \quad (3.20)$$

e combinando (2.19) e (2.20) obtemos

$$F(x_1) - F(x_*) = \alpha_1 + y_1,$$

onde

$$\alpha_1 = F(x_1) - F(x) - F'(x_*)[x_1 - x] - F''(x_*)[z, x_1 - x].$$

Usando (2.11) e (2.5), com $n = 1$, teremos

$$\begin{aligned} \|x_1 - x\| &= \left\| M \left(y, \pi \left(\frac{F(x) - F(x_*) - y}{\|x - x_*\|} \right) \right) \right\| \\ &\leq C \left\| \left(y, \pi \left(\frac{F(x) - F(x_*) - y}{\|x - x_*\|} \right) \right) \right\| \\ &\leq C \left(\|y\| + \left\| \pi \left(\frac{F(x) - F(x_*) - y}{\|x - x_*\|} \right) \right\| \right) \end{aligned}$$

Como

$$\|\pi(y)\| = \min_{\hat{y} \in \text{Im } F'(x_*)} \|\hat{y} - y\|,$$

e $\hat{y} = 0 \in \text{Im } F'(x_*)$, temos

$$\|\pi(y)\| \leq \|y\|,$$

logo,

$$\|x_1 - x\| \leq C(\|y\| + \|F(x) - F(x_*) - y\|\|x - x_*\|^{-1}) = C\Delta(x) \quad (3.21)$$

Usando (2.7), (2.8), (2.10) e (2.21) obtemos

$$\begin{aligned} \|x_1 - x\| &\leq C\Delta(x) \\ &\leq CK^{-1}\|x - x_*\| \\ &= C(\max\{(4C)^2\|F''(x_*)\|, 4C\})^{-1}\|x - x_*\| \\ &\leq \frac{C}{4C}\|x - x_*\|, \end{aligned}$$

isto é,

$$\|x_1 - x\| \leq \frac{1}{4}\|x - x_*\|.$$

Agora, usando (2.6)-(2.10) e (2.19) temos

$$\|\alpha_1\| \leq \frac{1}{2}\Delta(x)\|x - x_*\|,$$

$$\|y_1\| \leq \frac{1}{2}\Delta(x),$$

e as relações (2.15)-(2.18) ficam provadas para $n = 1$.

Suponhamos agora que as relações (2.15)-(2.18) são válidas para os primeiros m elementos $x_1, \dots, x_m$. Então, quando $n = m + 1$, para x_{m+1} , por (2.12)-(2.14), $y_{m+1} \in \text{Im } F'(x_*)$ e

$$F(x_{m+1}) - F(x_*) = \alpha_{m+1} + y_{m+1},$$

onde $\alpha_{m+1} = F(x_{m+1}) - F(x_m) - F'(x_*)[x_{m+1} - x_m] - F''(x_*)[z, x_{m+1} - x_m]$. De (2.11), usando (2.15)-(2.18), temos

$$\|x_{m+1} - x_m\| \leq 2^{-m}C\Delta(x),$$

$$\|x_{m+1} - x_*\| < 2k.$$

Logo, $x_{m+1} \in B(x_*, \delta)$ e podemos aplicar (2.6). Assim, por (2.7)-(2.10), (2.15)-(2.18) e (2.20) temos que

$$\|\alpha_{m+1}\| \leq 2^{-(m+1)}\Delta(x)\|x - x_*\|,$$

$$\|y_{m+1}\| \leq 2^{-(m+1)}\Delta(x).$$

E isto completa a indução e prova (2.15)-(2.18). Como X é um espaço de Banach, segue de (2.18) que existe

$$g(x) = \lim_{n \rightarrow \infty} x_n.$$

Tomando, então, o limite em (2.15) com $n \rightarrow \infty$, concluímos que $F(g(x)) = F(x_*)$ considerando (2.16)-(2.17) e a continuidade da função F em $B(x_*, \delta)$. Seja $r(x) = g(x) - x$. Então (2.2)-(2.3) segue da última igualdade e de (2.18).

□

Denotemos por $TM(x_*)$ o cone de direções tangentes a $M(x_*)$ no ponto x_* . Do Teorema 3.1.1, quando $h = 0$, obtemos

Corolário 3.1.1 *Sejam X e Y espaços de Banach, U uma vizinhança do ponto $x_* \in X$ e F uma função de U em Y , duplamente Fréchet-diferenciável em x_* , tal que $\text{Im } F'(x_*)$ é um subespaço fechado de Y . Então,*

$$H_o(x_*) \cup \{0\} \subseteq TM(x_*) \subseteq H(x_*).$$

Se $H(x_) \setminus \{0\} = H_o(x_*) \setminus \{0\}$, então*

$$TM(x_*) = H(x_*).$$

Observação 3.1.1 :

É claro que o Teorema de Lyusternik segue diretamente deste corolário, pois se F é regular em x_ , então*

$$H_o(x_*) = H(x_*) = \ker F''(x_*).$$

Portanto, podemos ver o corolário 2.1.1 como uma “generalização” do teorema de Lyusternik para o caso de operadores não regulares, porém o preço que pagamos é a exigência da dupla diferenciabilidade.

Existe uma outra “generalização” que só impõe Fréchet-diferenciabilidade, porém é mantida a hipótese de regularidade de F , isto é, $\text{Im } F'(x_) = Y$ (sobrejetividade da $F'(x_*)$) (veja [*)).*

É um problema ainda em aberto uma verdadeira generalização de Lyusternik, onde sejam enfraquecidas simultaneamente as hipóteses de diferenciabilidade estrita de F em x_ e a sobrejetividade de $F'(x_*)$.*

3.2 Os multiplicadores de Lagrange

Vincularemos ao problema (P) o seu lagrangeano generalizado

$$\hat{\mathcal{L}}(x, \lambda_o, y_1^*, y_2^*, h) = \lambda_o f(x) + \langle y_1^*, F(x) \rangle + \langle y_2^*, F'(x)[h] \rangle,$$

onde $\lambda_o \in \mathbb{R}$, $y_1^* \in Y^*$, $y_2^* \in Y^*$, $h \in H(x)$.

Teorema 3.2.1 *Seja f um funcional duplamente Fréchet-diferenciável e F uma função três vezes Fréchet-diferenciável em x_* , com $F(x_*) = 0$. Se x_* é um ponto de mínimo local no problema (P) e $\text{Im } F'(x_*)$ é um subespaço fechado em Y , então, para todo $h \in H(x_*)$ tal que $\text{Im } G(x_*, h)$ é um subespaço fechado em Z , existem multiplicadores não simultaneamente nulos $\lambda_o(h), y_2^*(h)$, e um multiplicador $y_1^*(h)$ tais que*

$$\begin{aligned}\hat{\mathcal{L}}_x(x_*, \lambda_o(h), y_1^*(h), y_2^*(h), h) &= \lambda_o(h)f'(x_*) + F''(x_*)y_1^*(h) \\ &+ ((F')'(x_*)[h])^* y_2^*(h) \\ &= 0,\end{aligned}\tag{3.22}$$

$$y_2^*(h) \in (\text{Im } F'(x_*))^\perp.\tag{3.23}$$

Para qualquer $h \in H_o(x_*)$ fixo o multiplicador $\lambda_o(h)$ correspondente é não nulo (podemos assumir $\lambda_o(h) = 1$), $y_2^*(h)$ é único, todo multiplicador de Lagrange $y_1^*(h)$ para tal $h \in H_o(x_*)$ pertence à mesma classe de equivalência com respeito a $(\text{Im } F'(x_*))^\perp$ e, além disso,

$$\hat{\mathcal{L}}_{xx}(x_*, 1, y_1^*(h), y_2^*(h), h)[h, h] \geq 0.\tag{3.24}$$

Se ainda $H_o(x_*) \neq \emptyset$ e

$$\bigcap_{h \in H_o(x_*)} \{\text{Im } F''(x_*) + \hat{\text{Im}}((F')'(x_*)[h])^*\} = \text{Im } F''(x_*),\tag{3.25}$$

onde $\hat{\text{Im}}((F')'(x_*)[h])^*$ é a imagem do subespaço $(\text{Im } F'(x_*))^\perp$ pela função

$$y^* \mapsto ((F')'(x_*)[h])^* y^*,$$

então, $y_1^*(h) = y^*$ para todo $h \in H_o(x_*)$ e

$$\mathcal{L}_x(x_*, 1, y^*) = 0\tag{3.26}$$

$$\mathcal{L}_{xx}(x_*, 1, y^*)[h, h] \geq 0,\tag{3.27}$$

onde $\mathcal{L}(x, 1, y^*) = f(x) + \langle y^*, F(x) \rangle$ é o lagrangeano do problema (P).

Demonstração:

Seja x_* um ponto de mínimo local do problema (P). Como $\text{Im } F'(x_*)$ é fechado em Y , os espaços $Z = Y \times Y / \text{Im } F'(x_*)$ e $Z_1 = \text{Im } F'(x_*) \times Y / \text{Im } F'(x_*)$ são espaços de Banach.

Vamos dividir a demonstração em várias etapas.

i) Fixado $h \in H(x_*)$, consideremos o operador conjugado $G^*(x_*, h) : Z^* \longrightarrow X^*$, e provemos a identidade

$$\text{Im } G^*(x_*, h) = \text{Im } F'^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h])^*. \quad (3.28)$$

Para cada $v^* \in (Y/\text{Im } F'(x_*))^*$ seja $\tau(v^*) = v^*\pi$. Pelo Teorema [C.0.9], a função τ é um isomorfismo isométrico do espaço $(Y/\text{Im } F'(x_*))^*$ em $(\text{Im } F'(x_*))^\perp$.

Se $x_* \in \text{Im } G^*(x_*, h)$, ou seja, existe $z^* \in Z^*$ tal que $G^*(x_*, h)z^* = x_*$, então existem $y_1^*(h) \in Y^*$, $v^*(h) \in (Y/\text{Im } F'(x_*))^*$ e $y_2^*(h) \in (\text{Im } F'(x_*))^\perp$ tais que, para todo $x \in X$ tem-se

$$\begin{aligned} \langle G^*(x_*, h)z^*, x \rangle &= \langle z^*, G(x_*, h)x \rangle \\ &= \langle z^*, (F'(x_*), \pi((F')'(x_*)[h]))x \rangle \\ &= \langle y_1^*(h), F'(x_*)x \rangle + \langle v^*(h), \pi(F''(x_*)[h, x]) \rangle \\ &= \langle y_1^*(h), F'(x_*)x \rangle + \langle \tau v^*(h), F''(x_*)[h, x] \rangle \\ &= \langle y_1^*(h), F'(x_*)x \rangle + \langle y_2^*(h), F''(x_*)[h, x] \rangle \\ &= \langle F'^*(x_*)y_1^*(h) + ((F')'(x_*)[h])^*y_2^*(h), x \rangle. \end{aligned}$$

Portanto, $\text{Im } G^*(x_*, h) \subseteq \text{Im } F'^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h])^*$.

É idêntica a demonstração da inclusão contrária, concluindo assim a verificação da identidade desejada.

ii) Ainda considerando $h \in H(x_*)$ fixo, provemos, primeiramente, (2.22)-(2.23) para o caso em que $G(x_*, h)X \neq Z_1$. Como $\text{Im } G(x_*, h)$ é um subespaço fechado próprio de Z_1 , de acordo com o teorema de Hahn-Banach em sua segunda forma geométrica [B.0.8] o anulador do subespaço $\text{Im } G(x_*, h)$ contém um elemento não nulo $z^*(h) \in Z_1^*$. Assim, para todo $x \in X$ temos

$$0 = \langle z^*(h), G(x_*, h)x \rangle = \langle \tilde{y}_1^*(h), F'(x_*)x \rangle + \langle v^*(h), \pi(F''(x_*)[h, x]) \rangle, \quad (3.29)$$

onde $\tilde{y}_1^* \in (\text{Im } F'(x_*))^*$, $v^*(h) \in (Y/\text{Im } F'(x_*))^*$.

Usando agora o teorema de Hahn-Banach em sua forma analítica deduzimos que existe $y_1^*(h) \in Y^*$ tal que

$$\langle \hat{y}_1^*(h), F'(x_*)x \rangle = \langle y_1^*(h), F'(x_*)x \rangle \quad (\forall x \in X). \quad (3.30)$$

Combinando (2.29) e (2.30) obtemos

$$\begin{aligned} 0 &= \langle y_1^*(h), F'(x_*)x \rangle + \langle v^*(h), \pi(F''(x_*)[h, x]) \rangle \\ &= \langle y_1^*(h), F'(x_*)x \rangle + \langle \tau v^*(h), F''(x_*)[h, x] \rangle \\ &= \langle y_1^*(h), F'(x_*)x \rangle + \langle y_2^*(h), F''(x_*)[h, x] \rangle \\ &= \langle F'^*(x_*)y_1^*(h) + ((F'')'(x_*)[h])^*y_2^*(h), x \rangle \quad (\forall x \in X), \end{aligned}$$

isto é,

$$F'^*(x_*)y_1^*(h) + ((F'')'(x_*)[h])^*y_2^*(h) = 0,$$

onde $y_2^*(h) \neq 0$, pois, do contrário, se $0 = y_2^*(h) = \tau v^*(h)$, teríamos $v^*(h) = 0$ (pois τ é um isomorfismo), e por ser $0 \neq z^*(h) = \hat{y}_1^*(h) + v^*(h)$ teríamos $\hat{y}_1^*(h) \neq 0$ e

$$0 = \langle \hat{y}_1^*(h), F'(x_*)x \rangle \quad (\forall x \in X)$$

o que é uma contradição.

iii) Suponhamos agora que existe $h \in H(x_*)$ tal que $G(x_*, h) = Z_1$, ou seja, $H_o(x_*) \neq \emptyset$. Fixemos $h \in H_o(x_*)$ para que todas as condições do Teorema 3.1.1 sejam satisfeitas no ponto x_* . Então, existe $\alpha(t)$ tal que

$$\begin{cases} \alpha(t) \geq 0, \quad \forall t \in \mathbb{R} \\ \alpha(t) = o(t). \end{cases}$$

Além disso, para todo $h_o \in \ker G(x_*, h)$ existem $\delta > 0$ e $r(t) \in X$ tais que

$$F(x_* + th + \alpha(t)h_o + r(t)) = 0 \quad (\forall t \in [-\delta, \delta])$$

$$\lim_{t \rightarrow 0} \frac{\|r(t)\|}{\alpha(t)} = 0.$$

Sendo x_* um ponto de mínimo local do problema (P), a função

$$g(t) = f(x_* + th + \alpha(t)h_1 + r(t))$$

assume seu mínimo local no ponto $t = 0$ e, portanto,

$$g'(0) = \langle f'(x_*), h \rangle = 0.$$

Temos, então,

$$0 \leq f(x_* + th + \alpha(t)h_o + r(t)) - f(x_*) = \alpha(t)\langle f'(x_*), h_o \rangle + o(\alpha(t))$$

Dividindo esta desigualdade por $\alpha(t)$ e fazendo $t \rightarrow 0$

$$\lim_{t \rightarrow 0} \frac{f(x_* + th + \alpha(t)h_o + r(t)) - f(x_*)}{\alpha(t)} = \langle f'(x_*), h_o \rangle + \lim_{t \rightarrow 0} \frac{o(\alpha(t))}{\alpha(t)},$$

logo,

$$0 \leq \lim_{t \rightarrow 0} \frac{g(t) - g(0)}{\alpha(t)} = \langle f'(x_*), h_o \rangle,$$

logo,

$$\langle f'(x_*), h_o \rangle = 0 \quad (\forall h_o \in \ker G(x_*, h)).$$

Portanto,

$$f'(x_*) \in (\ker G(x_*, h))^\perp,$$

e como $\text{Im } G(x_*, h)$ é um subespaço fechado de Z , temos (ver teorema [C.0.10])

$$(\ker G(x_*, h))^\perp = \text{Im } G^*(x_*, h),$$

logo,

$$f'(x_*) \in \text{Im } G^*(x_*, h).$$

Tendo em vista isto e (2.28), temos que existem $y_1^*(h) \in Y^*$, $y_2^*(h) \in (\text{Im } F'(x_*))^\perp$ tais que

$$f'(x_*) = -F''^*(x_*)y_1^*(h) - ((F')'(x_*)[h])^*y_2^*(h),$$

ou

$$f'(x_*) + F''^*(x_*)y_1^*(h) + ((F')'(x_*)[h])^*y_2^*(h) = 0. \quad (3.31)$$

Assim, qualquer que seja $h \in H_o(x_*)$, (2.22) e (2.23) são satisfeitas com $\lambda_o(h) = 1$.

iv) Provemos agora que

$$y_2^*(h) \text{ é único } \iff \tilde{y}_1^*(h) - y_1^*(h) \in (\text{Im } F'(x_*))^\perp$$

qualquer que seja $\tilde{y}_1^*(h)$ que satisfaça (2.31).

Se $y_2^*(h)$ é único, de (2.31) obtemos

$$\langle \tilde{y}_1^*(h) - y_1^*(h), F'(x_*)x \rangle = 0 \quad (\forall x \in X),$$

logo

$$\tilde{y}_1^*(h) - y_1^*(h) \in (\text{Im } F'(x_*))^\perp.$$

Agora, se existe $\tilde{y}_2^*(h) \neq y_2^*(h)$ que satisfaz (2.31), então, já que $\tilde{y}_2^*(h) - y_2^*(h) \neq 0$, existe $y_2 \in Y$ tal que

$$\langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle \neq 0,$$

e isto implica

$$\langle \tilde{y}_1^*(h) - y_1^*(h), F'(x_*)\bar{x} \rangle \neq 0,$$

logo, $\tilde{y}_1^*(h) - y_1^*(h) \notin (\text{Im } F'(x_*))^\perp$, o que nos leva ao resultado desejado.

v) Mostremos agora a unicidade. Suponhamos que existem $\tilde{y}_1^*(h) \in Y^*$ e $\tilde{y}_2^*(h) \in (\text{Im } F'(x_*))^\perp$ tais que

$$\tilde{y}_1^*(h) - y_1^*(h) \notin (\text{Im } F'(x_*))^\perp,$$

$$\tilde{y}_2^*(h) \neq y_2^*(h) \quad \text{e}$$

$$f'(x_*) + F'^*(x_*)\tilde{y}_1^*(h) + ((F')'(x_*)[h])^*\tilde{y}_2^*(h) = 0.$$

Subtraindo (2.31) da equação acima, obtemos

$$F'^*(x_*)(\tilde{y}_1^*(h) - y_1^*(h)) + ((F')'(x_*)[h])^*(\tilde{y}_2^*(h) - y_2^*(h)) = 0,$$

isto é,

$$\langle \tilde{y}_1^*(h) - y_1^*(h), F'(x_*)x \rangle + \langle \tilde{y}_2^*(h) - y_2^*(h), F''(x_*)[h, x] \rangle = 0 \quad (\forall x \in X) \quad (3.32)$$

Por outro lado, existem $y_1 \in \text{Im } F'(x_*)$ e $y_2 \in Y$ tais que

$$\langle \tilde{y}_1^*(h) - y_1^*(h), y_1 \rangle + \langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle \neq 0.$$

De fato, supondo que quaisquer que sejam $y_1 \in \text{Im } F'(x_*)$ e $y_2 \in Y$ tenhamos

$$\langle \tilde{y}_1^*(h) - y_1^*(h), y_1 \rangle = -\langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle,$$

podemos tomar $-y_2 \in Y$ e teremos

$$\langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle = \langle \tilde{y}_1^*(h) - y_1^*(h), y_1 \rangle = -\langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle$$

o que significa

$$\tilde{y}_2^*(h) - y_2^*(h) = 0 \quad \text{e}$$

$$\tilde{y}_1^*(h) - y_1^*(h) \in (\text{Im } F'(x_*))^\perp$$

o que é um absurdo, pois tomamos $\tilde{y}_1^*(h)$ e $\tilde{y}_2^*(h)$ de modo que $\tilde{y}_1^*(h) - y_1^*(h) \notin (\text{Im } F'(x_*))^\perp$ e $\tilde{y}_2^*(h) \neq y_2^*(h)$.

Considerando, então, tais elementos $y_1 \in \text{Im } F'(x_*)$ e $y_2 \in Y$ satisfazendo

$$\langle \tilde{y}_1^*(h) - y_1^*(h), y_1 \rangle + \langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle \neq 0,$$

tomemos os elementos $\tilde{x} \in X$ e $y \in \text{Im } F'(x_*)$ tais que

$$F'(x_*)\tilde{x} = y_1 \quad \text{e}$$

$$F''(x_*)[h, \tilde{x}] = y_2 + y.$$

Então,

$$\begin{aligned} 0 &\neq \langle \tilde{y}_1^*(h) - y_1^*(h), y_1 \rangle + \langle \tilde{y}_2^*(h) - y_2^*(h), y_2 \rangle \\ &= \langle \tilde{y}_1^*(h) - y_1^*(h), F'(x_*)\tilde{x} \rangle + \langle \tilde{y}_2^*(h) - y_2^*(h), F''(x_*)[h, \tilde{x}] \rangle \\ &\quad - \langle \tilde{y}_2^*(h) - y_2^*(h), y \rangle, \end{aligned}$$

o que, junto com (2.28), nos dá

$$\langle \tilde{y}_2^*(h) - y_2^*(h), y \rangle \neq 0,$$

o que é um absurdo, pois $\tilde{y}_2^*(h) - y_2^*(h) \in (\text{Im } F'(x_*))^\perp$.

Como a hipótese que nos levou à contradição foi a de que existem $\tilde{y}_1^*(h) \in Y^*$ e $\tilde{y}_2^*(h) \in (\text{Im } F'(x_*))^\perp$ satisfazendo (2.27) e tais que $\tilde{y}_1^*(h) - y_1^*(h) \notin (\text{Im } F'(x_*))^\perp$ e

$\tilde{y}_2^*(h) \neq y_2^*(h)$, concluímos que, satisfazendo (2.27) e para cada $h \in H_o(x_*)$ fixo, $y_2^*(h) \in (\text{Im } F'(x_*))^\perp$ é único e todo multiplicador de Lagrange $y_1^*(h)$ pertence à mesma classe de equivalência com respeito à $(\text{Im } F'(x_*))^\perp$.

Com os passos (i) a (v) provamos as afirmações relativas a (2.22) e (2.23).

vi) Demonstraremos agora a desigualdade (2.24). De acordo com o teorema 3.1.1, para todo $h_o \in \ker G(x_*, h)$, podemos construir a variação

$$x(t, h, h_o) = x_* + th + \alpha(t)h_o + r(t)$$

$$F(x(t, h, h_o)) = F(x_*) = 0$$

(isto é, uma variação de x no conjunto factível). Tomemos então $h_o = 0$. Como x_* é um ponto de mínimo local para (P),

$$\begin{aligned} 0 &\leq f(x(t, h, 0)) - f(x_*) \\ &= f(x(t, h, 0)) - f(x_*) + \langle y_1^*(h), F(x(t, h, 0)) \rangle + t^{-1} \langle y_2^*(h), F(x(t, h, 0)) \rangle \\ &= t \langle f'(x_*), h \rangle + \langle f'(x_*), r(t) \rangle + \langle y_1^*(h), F(x_*) \rangle \\ &+ t \langle y_1^*(h), F'(x_*)h \rangle + \langle y_1^*(h), F'(x_*)r(t) \rangle + \\ &+ t^{-1} \langle y_2^*(h), F(x_*) \rangle + \langle y_2^*(h), F'(x_*)h \rangle + t^{-1} \langle y_2^*(h), F'(x_*)r(t) \rangle \\ &+ t \langle y_2^*(h), F''(x_*)[h, h] \rangle + \langle y_2^*(h), F''(x_*)[h, r(t)] \rangle \\ &+ t^2 f''(x_*)[h, h] + t^2 \langle y_1^*(h), F'''(x_*)[h, h] \rangle + \\ &+ t^2 \langle y_2^*(h), F''(x_*)[h, h, h] \rangle + o(t^2). \end{aligned}$$

Já que $y_2^*(h) \in (\text{Im } F'(x_*))^\perp$ e $F''(x_*)[h, h] \in \text{Im } F'(x_*)$, por (2.27) temos

$$0 \leq t^2 \hat{\mathcal{L}}''(x_*, 1, y_1^*(h), y_2^*(h), h)[h, h] + o(t^2)$$

ou seja,

$$0 \leq \hat{\mathcal{L}}''(x_*, 1, y_1^*(h), y_2^*(h), h)[h, h] + \frac{o(t^2)}{t^2},$$

e isto implica

$$0 \leq \hat{\mathcal{L}}''(x_*, 1, y_1^*(h), y_2^*(h), h)[h, h] + \lim_{t \rightarrow 0} \frac{o(t^2)}{t^2},$$

logo,

$$\hat{\mathcal{L}}_{xx}(x_*, 1, y_1^*(h), y_2^*(h), h)[h, h] \geq 0.$$

vii) Suponhamos agora que (2.25) é satisfeito. Pela demonstração feita em iii), para todo $h \in H_o(x_*)$ temos

$$f'(x_*) \in \text{Im } G^*(x_*, h) = \text{Im } F'^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h])^*,$$

ou seja,

$$f'(x_*) \in \text{Im } F'^*(x_*) = \bigcap_{h \in H_o(x_*)} \{\text{Im } F'^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h])^*\},$$

Logo, existe $y^* \in Y^*$ tal que

$$\mathcal{L}_x(x_*, 1, y^*) = f'(x_*) + F'^*(x_*)y^* = 0.$$

e fica provado (2.26).

Com desenvolvimento idêntico ao que foi feito para provar (2.2) se demonstra que

$$\mathcal{L}_{xx}(x_*, 1, y^*)[h, h] \geq 0.$$

□

É imediato perceber que se F' é uma função regular no ponto x_* temos $G(x_*, h)X = Z = Y$, $y_2^*(h) \in (\text{Im } F'(x_*))^\perp = 0$ para todo $h \in H_o(x_*) = H(x_*) = \ker F'(x_*)$, e portanto $y_1^*(h) = y^* \in Y^*$ e podemos tomar $\lambda_o(h) = 1$ (pois não se pode ter $\lambda_o(h)$ e $y_2^*(h)$ simultaneamente nulos), as condições necessárias de primeira ordem contidas em (2.22) e (2.23) se transformam nas condições necessárias de Lagrange para um mínimo, e (2.24) se transforma na condição de segunda ordem.

No caso não-regular, isto é, quando $F'(x_*)X \neq Y$, tomando $h = 0$, se $\lambda_o(0)$ ou $y_1^*(0)$ for não nulo, então as condições necessárias de Lagrange seguem de (2.22), enquanto que, se $\lambda_o(0) = y_1^*(0) = 0$, então $y_2^*(0) \neq 0$ e as condições necessárias de Lagrange seguem de (2.23).

Mas é importante observar que, no caso não-regular, quando $H(x_*) \neq 0$, não só as condições necessárias de Lagrange são satisfeitas mas toda uma família de condições dependendo do parâmetro $h \in H_o(x_*)$ com $\lambda_o(h) = 1$.

Além disso, ainda pelo teorema 3.2.1, quando (2.25) é satisfeita, tomando $\lambda_0(h) = 1$, com hipóteses mais fracas que a regularidade de F no ponto x_* , as condições necessárias de Lagrange também são satisfeitas.

Observações 3.2.1 :

Quando $Y = L \oplus M$, onde $L = \text{Im } F'(x_*)$ é fechado e M é um outro subespaço fechado de Y , existe uma projeção contínua P tal que

$$\text{Im } P = M \quad \text{e} \quad \ker P = L.$$

Desta forma, podemos substituir o espaço quociente $Y/\text{Im } F'(x_*)$ por M e a projeção π por P , e no lugar da função $G(x_*, h)$, nos teoremas 3.1.1 e 3.2.1, utilizarmos a função

$$G_1(x_*, h) = (F'(x_*), P((F')'(x_*)[h])) : X \longrightarrow Y \times M.$$

Utilizando esta idéia, podemos considerar dois casos especiais do problema (P), a saber:

$$P.a) \quad F = (f_1, \dots, f_m) : X \longrightarrow \mathbb{R}^m$$

$$P.b) \quad F = (f_1, \dots, f_m, \mathcal{F}) : X \longrightarrow \mathbb{R}^m \times Y_1$$

onde f_i , $i = 1, \dots, m$ são funcionais em X , Y_1 é um espaço de Banach e $\mathcal{F} : X \longrightarrow Y_1$ é uma função regular em x_* .

Sob tais condições, os subespaços $F'(x_*)X$, $G(x_*, h)X$, $G_1(x_*, h)X$ são fechados nos respectivos espaços, pois se um subespaço de um espaço de Banach é finito-dimensional ou possui codimensão finita, ele é fechado.

Em relação ao caso P.a) obtemos um corolário do teorema 3.2.1, mas antes precisamos introduzir um conjunto.

Vamos considerar que os primeiros k vetores $f'_i(x)$, $i = 1, \dots, k$, $1 \leq k \leq m$, formam uma base da envolvente linear dos vetores $f'_i(x)$, $i = 1, \dots, m$, isto é, o menor espaço vetorial que contém os vetores $f'_i(x)$, $i = 1, \dots, m$. Se $f'_i(x) = 0$, $i = 1, \dots, m$, tomamos $k = 0$. Então, existem escalares α_{ij} tais que

$$f'_j(x) = \sum_{i=1}^k \alpha_{ij} f'_i(x), \quad j = k+1, \dots, m.$$

Consideremos, então, o seguinte conjunto

$$\begin{aligned} \hat{H}_o(x) &= \{h \in H(x) : f'_i(x), i = 1, \dots, k, \text{ e} \\ & (f'_j)'(x)[h] - \sum_{i=1}^k \alpha_{ij} (f'_i)'(x)[h], j = k+1, \dots, m, \text{ são l.i.}\} \end{aligned}$$

Corolário 3.2.1 *Sejam f, f_i , $i = 1, \dots, m$, funcionais duplamente Fréchet-diferenciáveis no ponto x_* e tais que $f_i(x_*) = 0$, $i = 1, \dots, m$. Se x_* é um ponto de mínimo local do problema (P), com $F = (f_1, \dots, f_m) : X \rightarrow \mathbb{R}^m$, então, para todo $h \in H(x_*)$, existem escalares $\lambda_o(h), \mu_1(h), \dots, \mu_m(h)$, não simultaneamente nulos, e também escalares $\lambda_1(h), \dots, \lambda_m(h)$ tais que*

$$\lambda_o(h) f'(x_*) + \sum_{i=1}^m \lambda_i(h) f'_i(x_*) + \sum_{i=1}^m \mu_i(h) (f'_i)'(x_*)[h] = 0 \quad (3.33)$$

$$\bar{\mu}(h) = (\mu_1(h), \dots, \mu_m(h)) \in (\text{Im } F'(x_*))^\perp. \quad (3.34)$$

Se $\hat{H}_o(x_*) \neq \emptyset$, então, para todo $h \in \hat{H}_o(x_*)$, a equação (2.33) é satisfeita com $\lambda_o(h) = 1$ e, além disso, existem $h_j \in \hat{H}_o(x_*)$, $j = 1, \dots, p$, $p \leq m - \dim \text{Im } F'(x_*) + 1$, tais que

$$\begin{aligned} \bigcap_{h \in \hat{H}_o(x_*)} \{\text{Im } F''^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h])^*\} &= \bigcap_{1 \leq j \leq p} \{\text{Im } F''^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h_j])^*\}, \\ \bigcap_{1 \leq j \leq s} \{\text{Im } F''^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h_j])^*\} &= \bigcap_{1 \leq j \leq s-1} \{\text{Im } F''^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h_j])^*\}, \\ p > 2, s &= 2, \dots, p. \end{aligned}$$

Finalmente, se

$$\bigcap_{1 \leq j \leq p} \{\text{Im } F''^*(x_*) + \hat{\text{Im}}((F')'(x_*)[h_j])^*\} \subset \text{Im } F''^*(x_*)$$

(por exemplo, quando $p = m - \dim \operatorname{Im} F'(x_*) + 1$), podemos assumir $\lambda_i(h) = \lambda_i$, $i = 1, \dots, m$, para todo $h \in \hat{H}_o(x_*)$ e, além disso,

$$\mathcal{L}_x(x_*, 1, \bar{\lambda}) = 0,$$

onde

$$\mathcal{L}(x, 1, \bar{\lambda}) = f(x) + \sum_{i=1}^m \lambda_i f_i(x)$$

é o lagrangeano do problema (P.a).

Demonstração:

Para demonstrar este corolário, usaremos os seguintes fatos:

- a) $H_o(x_*) = \hat{H}_o(x_*)$;
- b) $\dim \operatorname{Im} G(x_*, h) = m$, $\forall h \in \hat{H}_o(x_*)$;
- c) $\dim \operatorname{Im} G(x_*, h) = \dim \operatorname{Im} G^*(x_*, h)$;
- d) $\dim \operatorname{Im} G^*(x_*, h) = \dim \{ \operatorname{Im} F'^*(x_*) + \operatorname{Im} (F')'(x_*)[h] \}^*$;
- e) Se L e $\{L_h\}$, $h \in \hat{H}_o(x_*)$, são, respectivamente, um subespaço e uma família de subespaços do espaço X e $\dim(L + L_h) = m$ para todo $h \in \hat{H}_o(x_*)$, então existem $h_j \in \hat{H}_o(x_*)$, $j = 1, \dots, p$, $p \leq m - \dim L + 1$, tais que

$$\begin{aligned} \bigcap_{h \in \hat{H}_o(x_*)} (L + L_h) &= \bigcap_{h_j \in \hat{H}_o(x_*)} (L + L_{h_j}), \\ \bigcap_{1 \leq j \leq s} (L + L_{h_j}) &= \bigcap_{1 \leq j \leq s-1} (L + L_{h_j}), \quad p \geq 2, \quad s = 2, \dots, p. \end{aligned}$$

Todas as conclusões do corolário saem imediatamente do teorema 3.2.1 associado às afirmações a) - e).

Vamos agora considerar o caso em que $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ e $H(x_*) \setminus \{0\} = H_o(x_*) \setminus \{0\}$. O teorema a seguir fornece-nos condições suficientes para que x_* seja um ponto de mínimo local para o problema (P).

Teorema 3.2.2 *Sejam $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, x_* um ponto factível para o problema (P.a), e suponha que o funcional f e a função $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ são duplamente Fréchet-diferenciáveis no ponto x_* . Se para algum multiplicador de Lagrange $\bar{\lambda} = (\lambda_1, \dots, \lambda_n)$ as relações*

$$\mathcal{L}_x(x_*, 1, \bar{\lambda}) = 0, \quad (3.35)$$

$$\mathcal{L}_{xx}(x_*, 1, \bar{\lambda})[h, h] > 0, \quad h \in H(x_*) \setminus \{0\} \quad (3.36)$$

são satisfeitas, então, x_ é um ponto de mínimo local para o problema (P.a).*

Demonstração:

Primeiramente, observemos que toda sequência do tipo $\{h_k / \|h_k\|\}_{k \in \mathbb{N}}$ tal que $F(x_* + h_k) = 0$ e $\lim_{k \rightarrow \infty} h_k = 0$, possui um subsequência $\{h_{k_j} : \|h_{k_j}\|\}$ que converge a um elemento $h \in H(x_*)$.

Suponhamos, por absurdo, que existe uma tal sequência, de modo que $f(x_*) > f(x_* + h_k)$, consideremos o desenvolvimento de $\mathcal{L}(x_* + h_{k_j}, 1, \bar{\lambda})$ em torno de x_* . Considerando (2.37), obtemos

$$\mathcal{L}(x_* + h_{k_j}, 1, \bar{\lambda}) = \mathcal{L}(x_*, 1, \bar{\lambda}) + \frac{1}{2} \|h_{k_j}\|^2 \mathcal{L}''(x_*, 1, \bar{\lambda})[h_{k_j}, h_{k_j}] + o(\|h_{k_j}\|).$$

Isto significa

$$f(x_* + h_{k_j}) = f(x_*) + \frac{1}{2} \|h_{k_j}\|^2 \mathcal{L}''(x_*, 1, \bar{\lambda})[h_{k_j}, h_{k_j}] + o(\|h_{k_j}\|).$$

Como, por hipótese, $f(x_* + h_{k_j}) < f(x_*)$, temos que

$$\frac{1}{2} \|h_{k_j}\|^2 \mathcal{L}''(x_*, 1, \bar{\lambda})[h_{k_j}, h_{k_j}] + o(\|h_{k_j}\|) < 0.$$

Logo,

$$\frac{1}{2} \|h_{k_j}\|^2 \mathcal{L}''(x_*, 1, \bar{\lambda})[h, h] < 0,$$

o que contradiz (2.38). Portanto, x_* é, de fato, um ponto de mínimo para o problema (P.a).

A fim de aplicar a teoria desenvolvida neste capítulo, voltamos ao exemplo citado no início do mesmo.

Consideraremos dois problemas para ilustrar as condições necessárias para um mínimo conti das no teorema 3.2.1 e seu corolário.

Problema 1

$$\min_{x \in M} g_1(x) = (-x_1 + 2x_3 - x_2^2 + 5x_4^2)$$

Problema 2

$$\min_{x \in M} g_2(x) = (-2x_1 + 2x_2 - 2x_1x_2 + 7x_3^2 + 0.8x_4^5)$$

Vamos analisar os dois juntamente. Suponhamos que as condições necessárias de Lagrange nos fornecem os conjuntos N_i de pontos suspeitos de serem soluções do i -ésimo problema, $i = 1, 2$. Como vimos no início do capítulo, todos os pontos do conjunto $IR = \{x \in M : x_1 = x_3 = 0, \text{ e } x_2 = \pm x_4^2\}$, dos pontos onde $F = (f_1, f_2)$ é não-regular, são suspeitos de serem extremantes. Denotando por R_i o conjunto dos pontos regulares de M que são suspeitos de resolver o i -ésimo problema, temos $N_i = R_i \cup IR$.

Para aplicar o corolário 3.2.1, vamos determinar os conjuntos envolvidos para cada uma das seguintes situações: na primeira, consideremos apenas os pontos de irregularidade que são não nulos, isto é, os pontos do conjunto $IR' = \{x : x_2 \neq 0\}$; na segunda, consideremos $x = 0 \in IR$.

Para $IR' = \{x \in IR : x_2 \neq 0\}$:

Temos, para todo $x \in IR'$,

$$\ker F'(x) = \{h \in \mathbb{R}^4 : x_2 h_2 = 2x_4^3 h_4\},$$

$$\text{Im } F'(x) = \{y \in E^2 : y_2 = 0\},$$

$$H(x) = \{h \in \mathbb{R}^4 : h_2 = 2x_4^3 h_4 / x_2, h_1^2 + 2h_1 h_3 - 8h_3^2 = 0\},$$

$$H_o(x) = \{h \in H(x) : h_1 \neq 0\}$$

$$(\text{Im } F'(x))^\perp = \{y = (y_1, y_2) \in E^2 : y_1 = 0\}$$

Consideremos os vetores $h^1 = (2, 0, 1, 0)$ e $h^2 = (-4, 0, 1, 0)$. Então. h^1 e h^2 pertencem a $H_o(x)$. Além disso,

$$\bigcap_{h \in H_o(x)} \{\text{Im } F'^*(x) + \hat{\text{Im}}((F')'(x)[h])^*\} = \bigcap_{j=1,2} \{\text{Im } F'^*(x) + \hat{\text{Im}}((F')'(x)[h])^*\} \quad (3.37)$$

$$= \text{Im } F'^*(x_*). \quad (3.38)$$

Assim, pelo teorema 3.2.1, se os pontos de mínimo local pertencem a IR' , as condições necessárias de Lagrange são satisfeitas para $\lambda_o = 1$. Tomemos

$$\mathcal{L}_i(x, 1, \bar{\lambda}) = g_i(x) + \lambda_1 f_1(x) + \lambda_2 f_2(x)$$

a função lagrangeana do i -ésimo problema. Então,

$$\mathcal{L}_{1x}(x, 1, \bar{\lambda}) = 0$$

$$\mathcal{L}_{2x}(x, 1, \bar{\lambda}) = 0$$

se, e só se,

$$g'_1(x) + \lambda_1 f'_1(x) + \lambda_2 f'_2(x) = 0$$

$$g'_2(x) + \lambda_1 f'_1(x) + \lambda_2 f'_2(x) = 0$$

Após fazer as contas vemos que a primeira equação não tem solução em IR' , e portanto nenhum ponto de IR' poderá ser candidato a solução do problema (1).

Fazendo as contas com a segunda equação, a referente ao segundo problema, encontramos como solução $\bar{x} = (0, -1, 0, 1) \in IR'$ com $\lambda_1 = 1$, e nenhum outro ponto de IR' satisfaz tal equação. Assim, podemos considerar em IR' apenas os pontos $\bar{x}$ acima como candidatos a extremantes do problema (2).

Para $x = 0 \in IR$:

Neste caso, temos

$$\begin{aligned}
\ker F'(0) &= \mathbb{R}^4 \\
\operatorname{Im} F'(0) &= \{0\} \\
\operatorname{Im}(f'(0))^\perp &= E^2 \\
H(0) &= \{h \in \mathbb{R}^4 : h_2^2 - 2h_1h_3 = 0, h_1^2 + 2h_1h_3 - 8h_3^2 = 0\} \\
&= \{h \in \mathbb{R}^4 : h_1 = 2h_3, h_2 = \pm 2h_3\} \\
H_o(0) &= \{h \in \mathbb{R}^4 : h_1 = 2h_3, h_2 = \pm 2h_3, h_3 \neq 0\}
\end{aligned}$$

Como $\operatorname{Im} F'(0) = \{0\}$ e $(\operatorname{Im} F'(0))^\perp = E^2$, temos que $\operatorname{Im} F''(0) = \{0\}$ e $\hat{\operatorname{Im}}((F')'(0)[h])^* = \operatorname{Im}((F')'(0)[h])^*$. Logo,

$$\operatorname{Im} F''(0) + \hat{\operatorname{Im}}((F')'(x)[h])^* = \operatorname{Im}((F')'(x))$$

Fazendo os cálculos para se determinar uma base para $\bigcap_{h \in H_o(0)} \operatorname{Im}((F')'(0)[h])^*$, como no caso anterior, concluímos que, se $h^1 = (2, 2, 1, 0)$ e $h^2 = (2, -2, 1, 0)$, então

$$\operatorname{Im}((F')'(0)[h^1])^* \bigcap \operatorname{Im}((F')'(0)[h^2])^* = \bigcap_{h \in H_o(0)} \operatorname{Im}((F')'(0)[h])^*$$

$$\operatorname{Im}((F')'(0)[h^1])^* \bigcap \operatorname{Im}((F')'(0)[h^2])^* \neq \operatorname{Im}((F')'(0)[h^1])^*$$

Portanto, pelo corolário 3.2.1, se $x = 0$ é um ponto de mínimo para o problema (1), existem $\mu(h^1) = (\mu_1(h^1), \mu_2(h^1))$ e $\mu(h^2) = (\mu_1(h^2), \mu_2(h^2))$ em E^2 tais que

$$g'_1(0) + \mu_1(h^1)f''_1(0)h^1 + \mu_2(h^1)f''_2(0)h^1 = 0 \quad (3.39)$$

$$g'_1(0) + \mu_1(h^2)f''_1(0)h^2 + \mu_2(h^2)f''_2(0)h^2 = 0. \quad (3.40)$$

Resolvendo as equações acima, encontramos $\mu(h^1) = \mu(h^2) = (0, 1/6)$. Assim, o ponto $x = 0$ satisfaz a família de condições necessárias de primeira ordem com $\lambda_o(h) = 1$ para $h \in H_o(0)$.

Fazendo o mesmo para o problema (2), encontramos para $h^1 = (2, 2, 1, 0)$ a solução

$$\mu(h^1) = (\mu_1(h^1), \mu_2(h^1)) = (-1/2, 1/6),$$

mas para $h^2 = (2, -2, 1, 0)$, o sistema (2.42) é inconsistente.

Como o sistema deveria ser satisfeito para todo $h \in H_o(0)$ (caso $x = 0$ fosse extremante local para o problema 2), temos que $x = 0$ não é solução para o problema (2).

Concluimos então que entre os pontos do conjunto IR , somente $x = 0$ é suspeito de ser ponto de mínimo local para o problema (1), e somente o ponto $x = (0, -1, 0, 1)$ é suspeito de ser ponto de mínimo para o problema (2). Assim, podemos, ao invés de considerar os conjuntos N_i , $i = 1, 2$, iniciais, tomar os subconjuntos $\tilde{N}_i \subset N_i$ dados por

$$\begin{aligned}\tilde{N}_1 &= R_1 \cup \{0\} \\ \tilde{N}_2 &= R_2 \cup \{(0, -1, 0, 1)\}.\end{aligned}$$

Usando o teorema 2.2.2 concluimos, após fazer os cálculos, que $x = 0$ é um ponto de mínimo local para o problema (1) e as condições necessárias de Lagrange são satisfeitas apenas para $\lambda_o = 0$, enquanto que $x = (0, -1, 0, 1)$ é um ponto de mínimo local para o problema (2).

3.3 Considerações finais

I) É importante observar que podemos generalizar de modo natural os teoremas 3.1.1 e 3.2.1, considerando as derivadas da função F até a k -ésima ordem, inclusive, com $k > 2$. A função $G(x_*, h) : X \longrightarrow W_1 \times, \dots, \times W_{k-1}$ terá a forma

$$G(x_*, h) = (F'(x_*), \pi_1 \circ F''(x_*)[h], \pi_2 \circ \pi_1 \circ F'''(x_*)[h, h], \dots, \pi_{k-1} \circ \dots \circ \pi_1 \circ F^{(k)}(x_*)[h, \dots, h])$$

com

$$\begin{aligned} W_1 &= Y / \text{Im } F'(x_*), \\ W_2 &= W_{i-1} / \text{Im } \pi_{i-1} \circ \dots \circ \pi_1 \circ F^{(i)}(x_*)[h, \dots, h], \\ i &= 2, 3, \dots, k, \end{aligned}$$

e a prova será semelhante à que foi feita para o teorema 3.1.1 e 3.2.1.

II) Em [14], Avakov ampliou estes resultados para o problema P) acrescido de restrições de desigualdade, sem acréscimo de hipóteses sobre as funções envolvidas e sem maiores complicações.

Em [15] encontramos uma outra generalização do teorema de Lyusternik - onde é considerada apenas a Gâteaux-diferenciabilidade da função F , contudo, é mantida a regularidade da mesma - e esta é utilizada para uma generalização do formalismo de Dubovitskii-Milyutin. Em [7] esta generalização é utilizada para se obter o princípio do máximo de Pontryagin, para problema de controle ótimo, sob a hipótese da Gâteaux-diferenciabilidade.

Em [9], a generalização vista aqui, devido a Avakov, é utilizada para se obter o princípio do máximo de Pontryagin na forma degenerada.

Apêndice A

Teorema A.0.1 *Se F é um subespaço vetorial fechado de um espaço de Banach E , então o espaço vetorial F também é de Banach.*

□

Espaços quocientes e espaços produtos

Seja F um subespaço vetorial de E . Sobre E , a relação $x \sim y$ se $x - y \in F$ ou $x \in (y + F)$ define uma relação de equivalência. Denotaremos por $\dot{x}$ a classe de equivalência que contém o elemento x , e por E/F o conjunto das classes de equivalência (*quociente* de E pelo subespaço F).

Definindo

$$\text{i) } \dot{x} + \dot{y} = \overline{\dot{x} + \dot{y}}, \quad x \in \dot{x} \text{ e } y \in \dot{y}$$

$$\text{ii) } \lambda \dot{x} = \overline{\lambda x}, \quad x \in \dot{x},$$

é fácil verificar que damos a E/F uma estrutura de espaço vetorial e que, se E é um espaço normado, cada classe $\dot{x}$ é um fechado se, e só se, F é fechado.

Geometricamente o que temos é

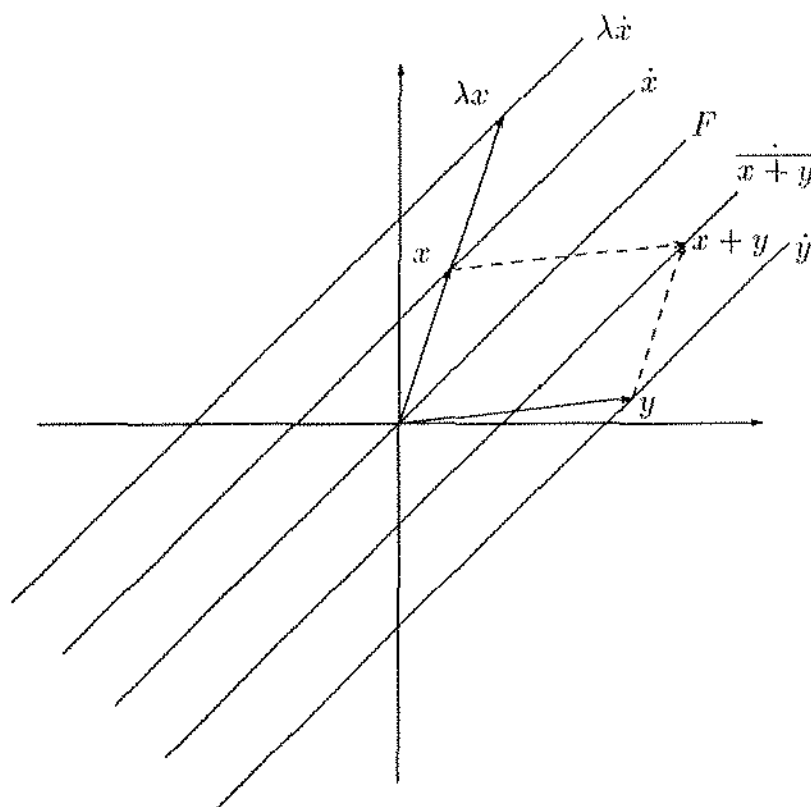


Figura A.1:

Existe uma transformação linear natural $\pi : E \longrightarrow E/F$ definida por $\pi(x) = \dot{x} = x + F$, que é denominada *aplicação quociente* ou *projeção* de E sobre E/F . É imediato que $\ker \pi = F$.

Quando F é fechado, podemos definir sobre E/F uma estrutura de espaço vetorial normado.

Teorema A.0.2 *Se E é um espaço vetorial normado e F é subespaço vetorial de E a expressão*

$$\|\dot{x}\| = \inf_{z \in \dot{x}} \|z\| = \inf_{y \in F, z \in \dot{x}} \|y - z\|$$

define uma norma em E/F se, e só se, F é fechado.

□

Para espaços de Banach, temos

Teorema A.0.3 *O quociente de um espaço de Banach E por um subespaço fechado F é um espaço de Banach.*

□

A demonstração dos teoremas acima pode ser encontrada em [4].

Agora, sejam E_1 e E_2 dois espaços vetoriais. Se definirmos, sobre $E_1 \times E_2$,

$$\alpha x + \beta y = (\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2),$$

onde $x = (x_1, x_2)$, $y = (y_1, y_2)$, $x_i, y_i \in E_i$, $i = 1, 2$ e $\alpha, \beta \in \mathbb{R}$, damos a $E_1 \times E_2$ uma estrutura de espaço vetorial, e se E_1 e E_2 são normados, pode-se definir sobre $E_1 \times E_2$ uma norma por

$$\|(x_1, x_2)\| = \max\{\|x_1\|_{E_1}, \|x_2\|_{E_2}\}.$$

Se os espaços E_1 e E_2 são Banach, então $E_1 \times E_2$, dotado desta norma, é também de Banach.

Apêndice B

Teoremas de Hahn-Banach

Forma analítica

Definição B.0.1 *Uma forma sublinear sobre um espaço vetorial real E é uma aplicação $p : E \longrightarrow \mathbb{R}$ satisfazendo:*

- i) $p(\lambda v) = \lambda p(v), \forall \lambda \geq 0, \forall v \in E,$
- ii) $p(u + v) \leq p(u) + p(v), \forall u, v \in E$

Dizemos que uma forma linear f é *majorada* por p se

$$f(v) \leq p(v) \quad \forall v \in E.$$

Teorema B.0.4 *(Hahn-Banach, forma analítica) Sejam E um espaço vetorial real, p uma forma sublinear sobre E , M um subespaço vetorial de E . Suponha que f é majorada por p sobre M . Então, existe g forma linear sobre E , majorada por p , que prolonga f , isto é, $f(v) = g(v)$ para todo $v \in M$.*

□

Corolário B.0.1 *Seja M um subespaço vetorial do espaço normado E . Seja $m' \in M^*$. Então, existe $v' \in E^*$ tal que*

$$\begin{aligned} v'(v) &= m'(v), \quad \forall v \in M \\ \|v'\|_{E^*} &= \|m'\|_{M^*}. \end{aligned}$$

□

A demonstração do teorema e seu corolário podem ser encontradas em [4].

Forma geométrica

No que segue, E denotará um espaço vetorial normado.

Definições B.0.1 :

(1) *Um hiperplano (afim) é um conjunto da forma*

$$H = \{x \in E ; f(x) = \alpha\} \equiv f^{-1}(\alpha)$$

onde f é uma forma linear sobre E , $f \neq 0$, e $\alpha \in \mathbb{R}$. Dizemos que H é um hiperplano de equação $[f = \alpha]$.

(2) *Sejam $A \subseteq E$ e $B \subseteq E$. Dizemos que o hiperplano H , de equação $[f = \alpha]$, separa A e B no sentido amplo (ver figura B.1), se temos*

$$\begin{aligned} f(x) &\leq \alpha, \quad \forall x \in A \quad e \\ f(x) &\geq \alpha, \quad \forall x \in B. \end{aligned}$$

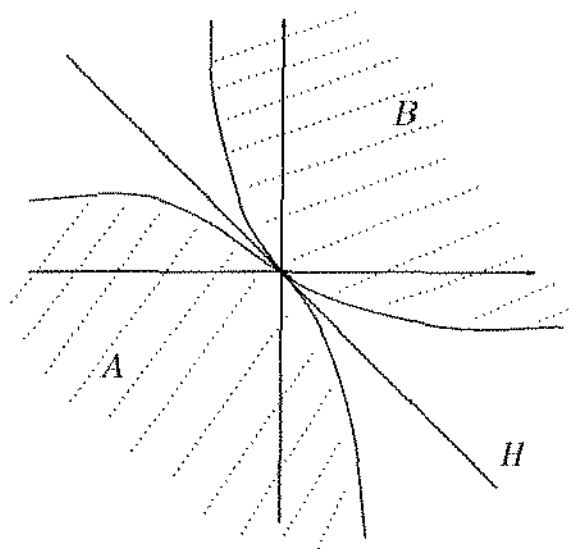


Figura B.1:

Dizemos que H separa A e B no sentido estrito (ver figura B.2) se existe $\varepsilon > 0$ tal que

$$f(x) \leq \alpha - \varepsilon \quad \forall x \in A \quad e$$

$$f(x) \geq \alpha + \varepsilon \quad \forall x \in B.$$

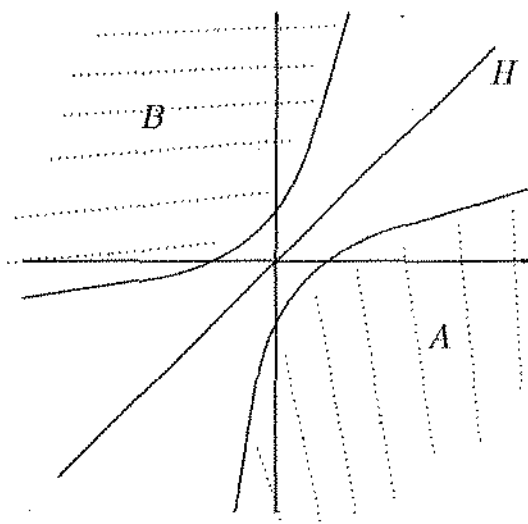


Figura B.2:

Teorema B.0.5 (*Hahn-Banach, primeira forma geométrica*) Sejam $A \subseteq E$, $B \subseteq E$ dois convexos não vazios e disjuntos. Suponha que A é aberto. Então, existe um hiperplano fechado que separa A e B no sentido amplo.

□

Teorema B.0.6 (*Hahn-Banach, segunda forma geométrica*) Sejam $A \subseteq E$ e $B \subseteq E$ dois convexos não-vazios e disjuntos. Suponha que A é fechado e B é compacto. Então, existe um hiperplano fechado que separa A e B no sentido estrito.

□

A demonstração de ambos os teoremas pode ser encontrada em [4].

Apêndice C

Mais alguns resultados de análise funcional

Teorema C.0.7 *Se F é um subespaço vetorial fechado do espaço normado E , então $(E/F)^*$ é canonicamente isomorfo a $F^\perp$, isto é, se π é a projeção de E sobre E/F , a sua transposta*

$$\pi^* : (E/F)^* \longrightarrow E^*$$

leva $(E/F)^$ isomorficamente sobre $F^\perp$.*

□

Para a demonstração deste teorema, ver [*].

Teorema C.0.8 *Seja $f : X \longrightarrow Y$ uma aplicação linear fechada. As seguintes propriedades são equivalentes.*

- i) $\text{Im } f$ é fechada;
- ii) $\text{Im } f^*$ é fechada;

$$iii) \operatorname{Im} f = (\ker f^*)^\perp$$

$$iv) \operatorname{Im} f^* = (\ker f)^\perp$$

□

Demonstração encontrada em [4].

Teorema C.0.9 (Krein) *Seja K um cone convexo, com vértice em 0, que contém pontos interiores, L um subespaço tal que $(\operatorname{int} K) \cap L \neq \emptyset$. Se $\bar{f}(x)$ é uma forma linear sobre L tal que $\bar{f}(x) \geq 0$ sobre $K \cap L$, então, existe uma forma linear contínua $f(x)$ sobre E tal que*

$$f(x) = \bar{f}(x) \quad \forall x \in L$$

$$f(x) \geq 0 \quad \forall x \in K.$$

□

Teorema C.0.10

$$\left(\bigcup_{i \in I} K_i \right)^* = \bigcap_{i \in I} K_i^*$$

onde I é um conjunto arbitrário de índices e K_i ($i \in I$) são cones convexos abertos.

□

Teorema C.0.11 *Sejam K_1 e K_2 cones convexos abertos. Se $K_1 \subseteq K_2$, então, $K_1^* \supseteq K_2^*$.*

□

Teorema C.0.12 *Sejam K_α , $\alpha \in I$, cones convexos abertos, I um conjunto arbitrário de índices. Então,*

$$\left(\bigcap_{\alpha \in I} K_\alpha \right)^* \supseteq \sum_{\alpha \in I} K_\alpha^*,$$

onde $\sum_{\alpha \in I} K_{\alpha}^*$ denota o conjunto de todas as somas finitas $f_{\alpha_1} + \dots + f_{\alpha_n}$, $f_{\alpha_i} \in K_{\alpha_i}^*$, $\alpha_i \in I$

□

Teorema C.0.13 *Se os cones K_i são fracamente fechados e convexos, enquanto $\sum_{i \in I} K_i^*$ é w^* -fechado, então*

$$\left(\bigcap_{i \in I} K_i \right)^* = \sum_{i \in I} K_i^*.$$

□

Teorema C.0.14 *Seja K um cone com vértice 0, e L um subespaço tal que $(\text{int } K) \cap L \neq \emptyset$. Então,*

$$(K \cap L)^* = K^* + L^*.$$

□

Teorema C.0.15 *Se $K_1, \dots, K_n$ são cone convexos abertos, com $\bigcap_{i=1}^n K_i \neq \emptyset$, então*

$$\left(\bigcap_{i=1}^n K_i \right)^* = \sum_{i=1}^n K_i^*.$$

□

As demonstrações dos teoremas C.0.9-C.0.15 podem ser encontradas em [3].

Bibliografia

- [1] Alexéev V., Tikhomirov V., Fomine S., COMMANDE OPTIMALE. Moscou, Mir, 1982.
- [2] A. N. Kolmogorov, S. V. Fomin ELEMENTOS DA TEORIA DAS FUNÇÕES E DE ANÁLISE FUNCIONAL, Editora Mir.
- [3] I. V. Girsanov, LECTURES ON MATHEMATICAL THEORY OF EXTREMUM PROBLEMS / LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS, 67, Springer-Verlag, 1972.
- [4] Haïm Brezis, ANALYSE FONCTIONNELLE, Théorie et applications, Collection Mathématiques appliquées pour la maîtrise, Masson, 1987.
- [5] W. Kotarski, FURTHER GENERALIZATION OF THE DUBOVICKI-MILUTIN THEOREM, Journal of Optimization Theory and Applications, vol. 54, 3, 565-573, SEPTEMBER 1987.
- [6] Urszula Ledzewicz - Kowalewska, ON SOME SPECIFICATION OF THE DUBOVITSKII-MILYUTIN METHOD, Nonlinear Analysis, Theory, Methods & Applications, vol.10, 12 pp 1367-1371, 1986.
- [7] Urszula Ledzewicz - Kowalewska, A NECESSARY CONDITION FOR THE EXTREMAL PROBLEMS UNDER GÂTEAUX DIFFERENTIABILITY, Journal of Mathematical Analysis and Applications 134, pp 158-169, 1988.

- [8] Urszula Ledzewicz - Kowalewska, APPLICATION OF SOME SPECIFICATION OF THE DUBOVITSKII - MILYUTIN METHOD TO PROBLEMS OF OPTIMAL CONTROL. *Nonlinear Analysis, Theory, Methods & Applications*, vol. 12, 2, pp 101-108, 1988.
- [9] Urszula Ledzewicz, EXTENSION OF THE LOCAL MAXIMUM PRINCIPLE TO ABNORMAL OPTIMAL CONTROL PROBLEMS, *Journal of Optimization Theory and Applications* vol 77, 3, 1993.
- [10] Urszula Ledzewicz - Kowalewska, EULER - LAGRANGE EQUATION IN THE CASE NONREGULAR EQUALITY CONSTRAINTS, *Journal of Optimization Theory and Application* vol 71, 3, 1991.
- [11] A. F. Izmailov, OPTIMALITY CONDITIONS FOR DEGENERATE EXTREMUM PROBLEMS WITH INEQUALITY - TYPE CONSTRAINTS. *Comp. Maths. Phys.*, vol 34, 6, pp 723-736, 1994.
- [12] F. Fardella, ON THE IMAGE OF A CONSTRAINED EXTREMUM PROBLEM AND SOME APPLICATIONS TO THE EXISTENCE OF A MINIMUM. *Journal of Optimization Theory and Applications*, vol 60, 1, 1989.
- [13] E. R. Avakov, EXTREMUM CONDITION FOR SMOOTH PROBLEMS WITH EQUALITY - TYPE CONSTRAINTS. *U.S.S.R. Comput. Maths. Math. Phys.*, vol 25, 3, pp 24-32, 1985.
- [14] E. R. Avakov, NECESSARY EXTREMUM CONDITION FOR SMOOTH ANORMAL PROBLEMS WITH EQUALITY - AND INEQUALITY - TYPE CONSTRAINTS.
- [15] M. Altman, AN APPLICATION OF THE METHOD OF CONTRACTOR DIRECTION TO NONLINEAR PROGRAMMING. *Numer. Funct. Anal. Optim.* 1 (6) pp 647-663, 1979.