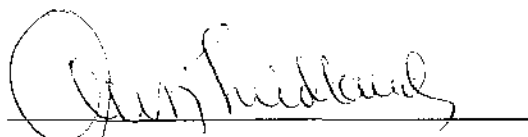


IDENTIFICAÇÃO DAS RESTRIÇÕES ATIVAS PARA UM ALGORITMO DE REGIÃO DE CONFIANÇA EM DOMÍNIOS ARBITRÁRIOS

Este exemplar corresponde à redação final da
tese devidamente corrigida e defendida pelo
Sr. Sandro Dimy Barbosa Bitar e aprovada
pela Comissão Julgadora.

Campinas, 02 de dezembro de 1994



Prof^a Dr^a Ana Friedlander

Orientadora

Dissertação apresentada ao Instituto de Ma-
temática, Estatística e Ciência da Com-
putação, UNICAMP, como requisito parcial
para a obtenção do título de Mestre em Ma-
temática Aplicada.



**IDENTIFICAÇÃO DAS RESTRIÇÕES ATIVAS
PARA UM ALGORITMO DE REGIÃO DE
CONFIANÇA EM DOMÍNIOS ARBITRÁRIOS**

SANDRO DIMY BARBOSA BITAR *n/5-6*

ORIENTADORA: PROF^a DR^a ANA FRIEDLANDER *f*

DMA – IMECC – UNICAMP

*Dedico este trabalho aos meus pais, Floriano
e Lais Bitar, e a minha eterna companheira,
Simone Bitar.*

AGRADECIMENTOS

À CAPES pelo apoio financeiro.

À Prof^ª Ana Friedlander, pela excelente orientação.

Aos professores do IMECC, pelo apoio moral e incentivo ao trabalho.

Aos amigos do IMECC, que de alguma forma contribuíram com este trabalho.

À Sra. Fátima, pela atenção e trabalho de digitação.

Aqueles que dividiram comigo os momentos mais difíceis e alegres dessa jornada: Meus Pais queridos e minha amada esposa, Simone.

Ao meu irmão, pelo crédito atribuído a minha pessoa.

À todos aqueles que torceram pelo sucesso da minha formação.

RESUMO

Neste trabalho, demonstramos os teoremas de identificação de restrições ativas para um algoritmo de região de confiança apresentado em [13] para resolver o problema

$$\begin{array}{ll} \min & f(x) \\ \text{sujeita a} & x \in \Omega, \end{array}$$

onde $\Omega \subset \mathbb{R}^n$ é um subconjunto convexo, não-vazio, definido por restrições de desigualdade.

Assumimos hipóteses de não degenerescência primal e dual, e utilizamos a teoria de cones combinada com alguns resultados da análise convexa, para garantir a identificação das restrições ótimas num número finito de iterações.

ÍNDICE

Capítulo 1 – Introdução	01
Capítulo 2 – Elementos de Convexidade	04
2.1 Definições	04
2.2 Lemas Básicos	11
Capítulo 3 – Identificação de Restrições Ativas	18
3.1 Introdução	18
3.2 O Algoritmo RCARB	18
3.3 Resultados Preliminares	20
3.4 Identificação das Restrições Ativas no Algoritmo RCARB	28
Capítulo 4 – Conclusões e Perspectivas	34
Referências	36

CAPÍTULO 1

INTRODUÇÃO

Neste trabalho, abordamos o problema de identificação de restrições ativas para o problema de programação não-linear restrita

$$\begin{array}{ll} \min & f(x) \\ \text{sujeita a} & x \in \Omega, \end{array} \quad (1.1.1)$$

onde $\Omega \subset R^n$ é convexo, fechado e não-vazio definido por $\Omega = \{x \in R^n \mid c_i(x) \leq 0, i = 1, \dots, m\}$, $f: R^n \rightarrow R$ e $c_i: R^n \rightarrow R$ para $i = 1, \dots, m$, são continuamente diferenciáveis sobre um conjunto aberto B que contém Ω .

Consideremos a i -ésima restrição do conjunto Ω . Pensando no conjunto dos pontos $x \in R^n$ tais que se $c_i(x) \leq 0$, poderíamos ter a seguinte representação geométrica:

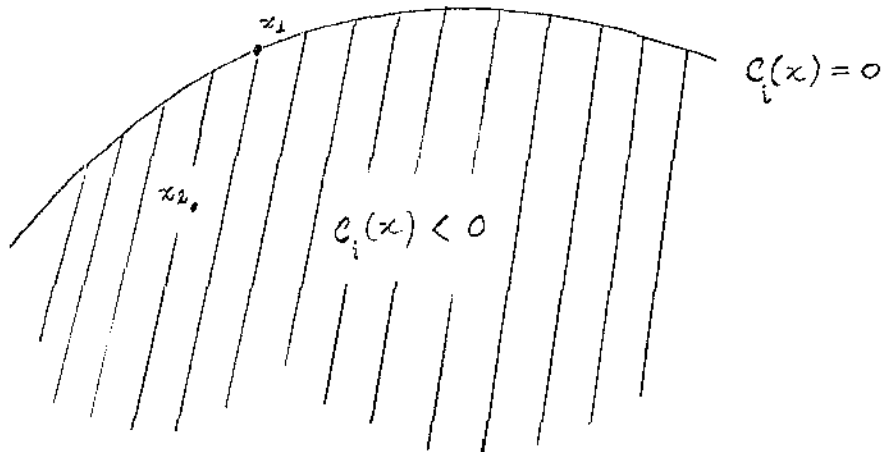


FIGURA 1

Aqui figuram os pontos que estão sobre a curva de nível zero e os pontos que possuem um valor funcional menor que zero.

Notemos que $c_i(x_1) = 0$. Dizemos então, que a restrição i está ativa em x_1 . Por outro lado $c_i(x_2) < 0$, e neste caso dizemos que a restrição i não está ativa em x_2 .

Se $x \in \Omega$ e existem restrições ativas em x , definimos $A(x) \subseteq \{1, 2, \dots, m\}$ o conjunto de índices correspondentes às restrições ativas em x . Se um algoritmo gera uma sequência de pontos $\{x_k\}$ que converge a uma solução x^* de (1.1.1) dizemos que o algoritmo possui a propriedade de identificação das restrições ativas na solução ótima se existe um índice k_0 tal que para todo $k \geq k_0$ $A(x_k) = A(x^*)$ (ver Figura 2).

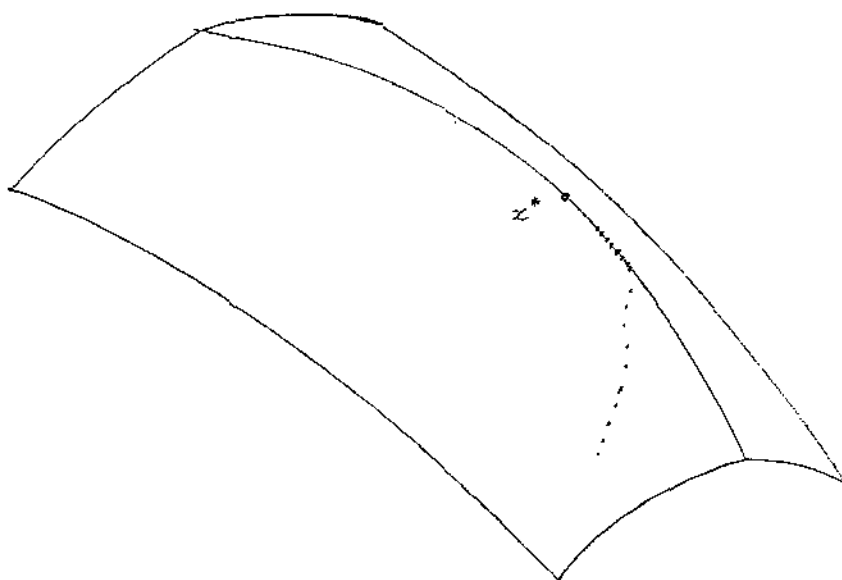


FIGURA 2

Restrições ativas em x^* ; sequência $\{x_k\}$ de vetores do R^n convergindo para x^* .
Para k suficientemente grande $A(x_k) = A(x^*)$.

A identificação das restrições ativas na solução é de grande valia já que neste caso, um algoritmo destinado a resolver o problema (1.1.1), reduz-se a um método de minimização irrestrita através de técnicas de otimização podendo então aplicar os teoremas de convergência local.

Vários autores provam, sob hipóteses de não-degenerescência dual, que os seus algoritmos identificam as restrições ativas na solução em um número finito de iterações (ver

[2, 5, 6]).

Neste trabalho provamos resultados de identificação para o algoritmo apresentado em [13] onde os autores provam resultados de convergência global com Ω arbitrário para um algoritmo de região de confiança que denominam RCARB (região de confiança em domínios arbitrários).

Já foi provada a identificação quando Ω é definido por restrições de canalização (ver [10]).

No algoritmo RCARB, é definido um subproblema “*fácil*” e a solução deste subproblema auxiliar, que chamamos “*ponto auxiliar*”, desempenha o papel do clássico ponto de Cauchy, usado em muitos métodos de região de confiança.

Obtemos resultados de identificação, para o ponto auxiliar, sob as hipóteses de regularidade e não-degenerescência de Dunn (ver [2], [8]). Assumimos a não-degenerescência dual que é um caso particular da não-degenerescência de Dunn [8].

Neste contexto fizemos uso de muitos resultados da análise convexa com destaque à teoria de cones convexos.

É interessante salientar que apesar de considerarmos Ω convexo, todos os resultados apresentados ao longo do trabalho, são igualmente válidas se pensarmos em convexidade local em torno da solução, o que vem justificar o título deste trabalho.

No capítulo 2, apresentamos alguns elementos de análise convexa necessários para o desenvolvimento deste trabalho.

No capítulo 3, desenvolvemos o essencial deste trabalho. Descrevemos o algoritmo RCARB e através de uma série de lemas e teoremas chegamos ao resultado de identificação desejado.

No capítulo 4 apresentamos algumas conclusões e perspectivas.

CAPÍTULO 2

ELEMENTOS DE CONVEXIDADE

Neste capítulo, apresentamos as definições e lemas básicos necessários para o desenvolvimento deste trabalho.

Alguns lemas aqui apresentados, são clássicos da análise convexa. Chamamos, contudo, a atenção para o resultado do lema 2.10 que, além de extremamente importante para os nossos fins, não consta nas bibliografias por nós consultadas.

2.1 DEFINIÇÕES

Definição 2.1: Dizemos que um conjunto $M \subset \mathbb{R}^n$ é convexo se, para todo $x, y \in M, \lambda \in [0, 1]$

$$\lambda x + (1 - \lambda)y \in M.$$

Definição 2.2: Dizemos que $M \subset \mathbb{R}^n$ é um conjunto afim se, para todo $x, y \in M, \lambda \in \mathbb{R}$

$$\lambda x + (1 - \lambda)y \in M.$$

Definição 2.3: Seja $S \subset \mathbb{R}^n$ e consideremos a interseção de todos os conjuntos afins M tais que $M \supset S$. Essa interseção é obviamente um conjunto afim e a denotamos por

$$\text{aff}(S).$$

Definição 2.4: Seja $C \subset \mathbb{R}^n$ um conjunto convexo. Definimos o interior relativo de C , que denotaremos por $r_i(C)$ como o interior de C relativo à $\text{aff}(C)$

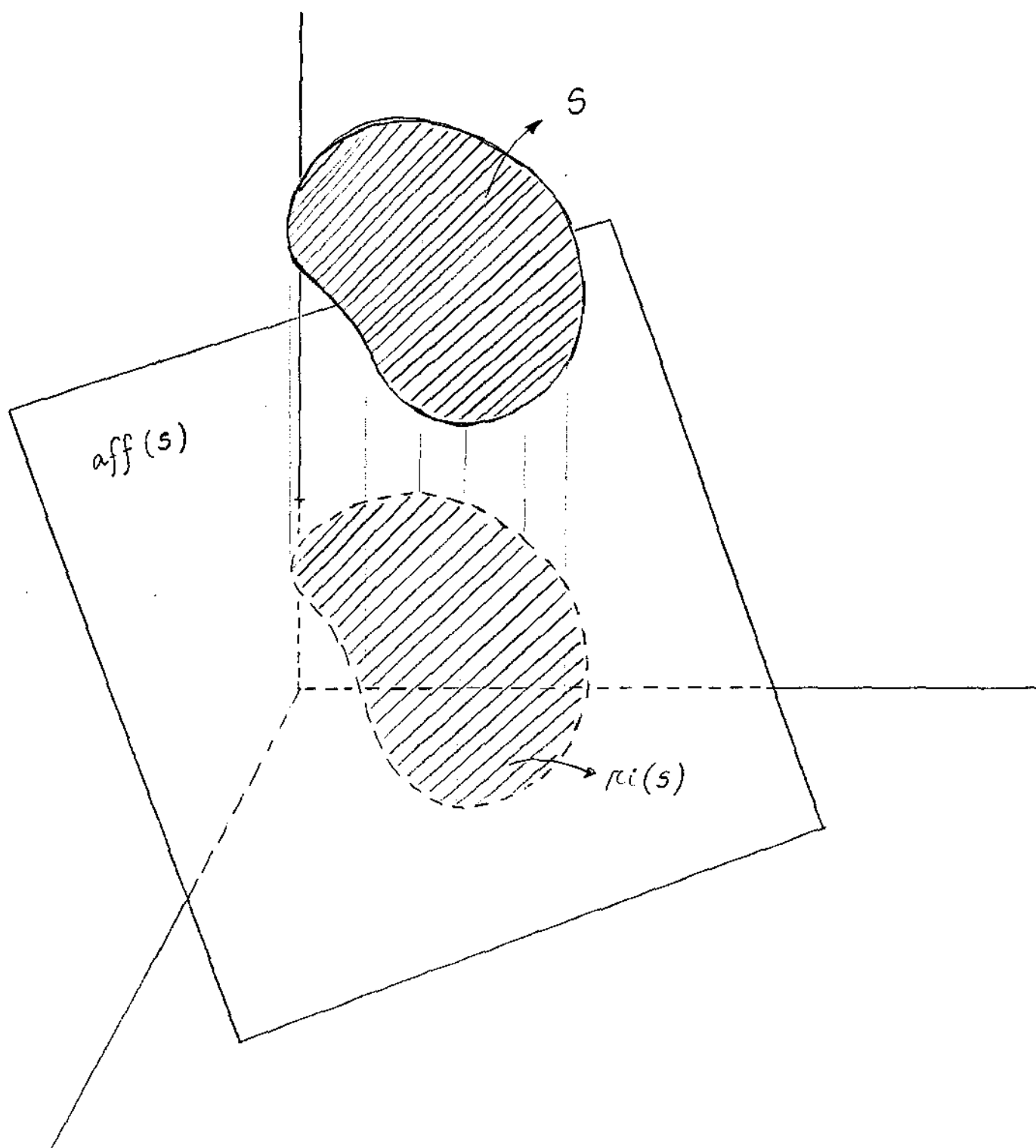


FIGURA 3

Aqui temos $S \subset R^n$ em destaque, interior relativo de S e $aff(S)$.

Definição 2.5: Um conjunto $K \subset \mathbb{R}^n$ é um cone, se para todo $x \in K, \lambda \geq 0, \lambda x \in K$. Se K é convexo, então dizemos que K é um cone convexo.

Definição 2.6: Seja $K \subset \mathbb{R}^n$ um cone. Definimos o cone polar K^0 de K como

$$K^0 = \{y \in \mathbb{R}^n \mid \langle y, v \rangle \leq 0, \forall v \in K\}.$$

Definição 2.7: Seja $C \subset \mathbb{R}^n$ um conjunto convexo e $x \in C$. O cone normal no ponto x é definido por

$$N(x) = \{u \in \mathbb{R}^n \mid \langle u, y - x \rangle \leq 0, y \in C\}.$$

Definição 2.8: Seja $C \subset \mathbb{R}^n$ um conjunto convexo e $x \in C$. O cone tangente $T(x)$ no ponto x é definido como o cone polar do cone normal no ponto x , ou seja

$$T(x) = N(x)^0.$$

Podemos também de forma equivalente definir $T(x)$ como o fecho do cone das direções viáveis no ponto x .

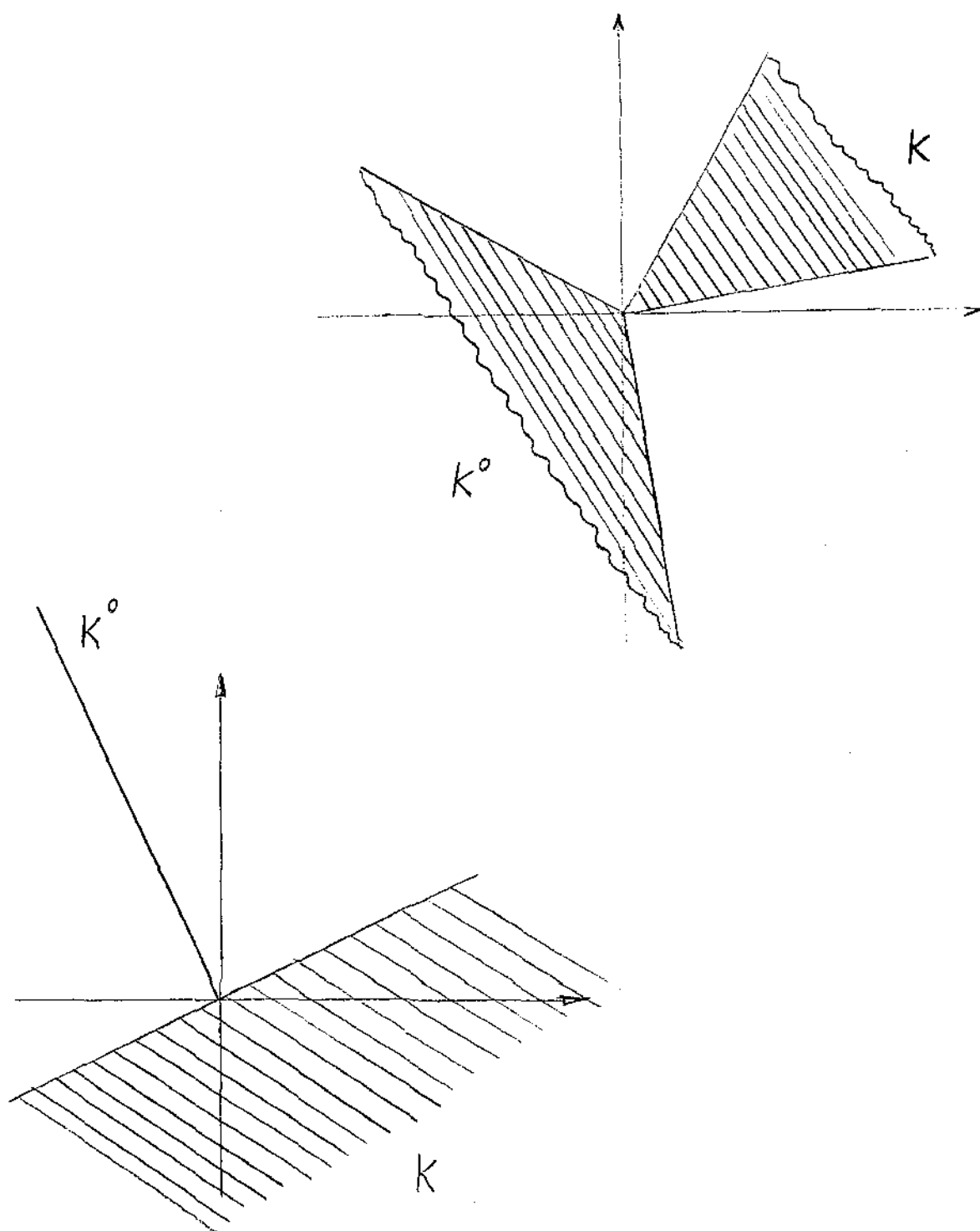


FIGURA 4.a

Cones convexos K e respectivos cones polares.

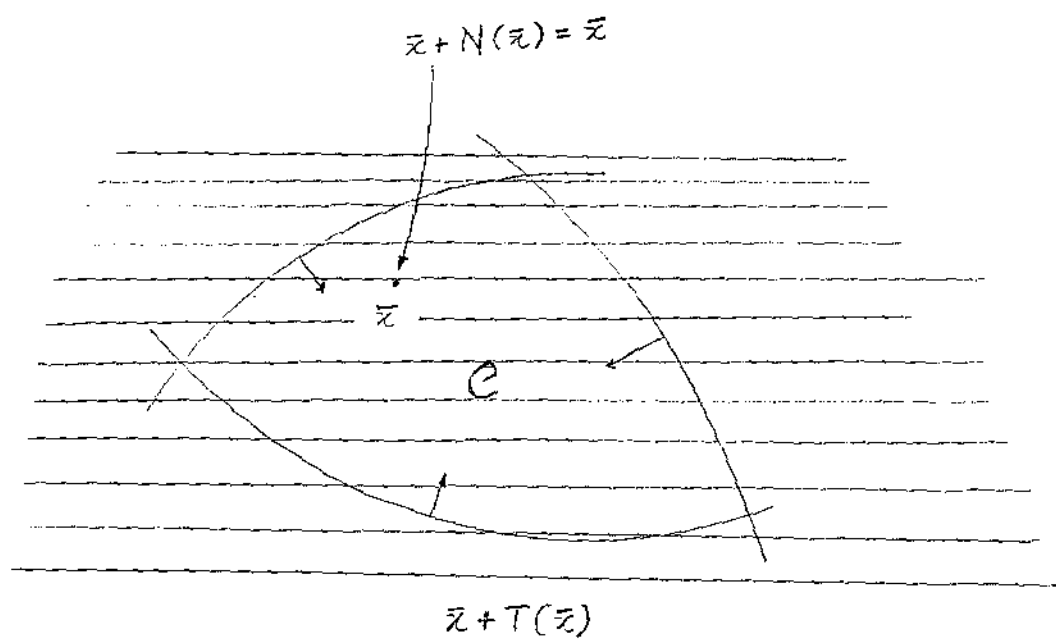
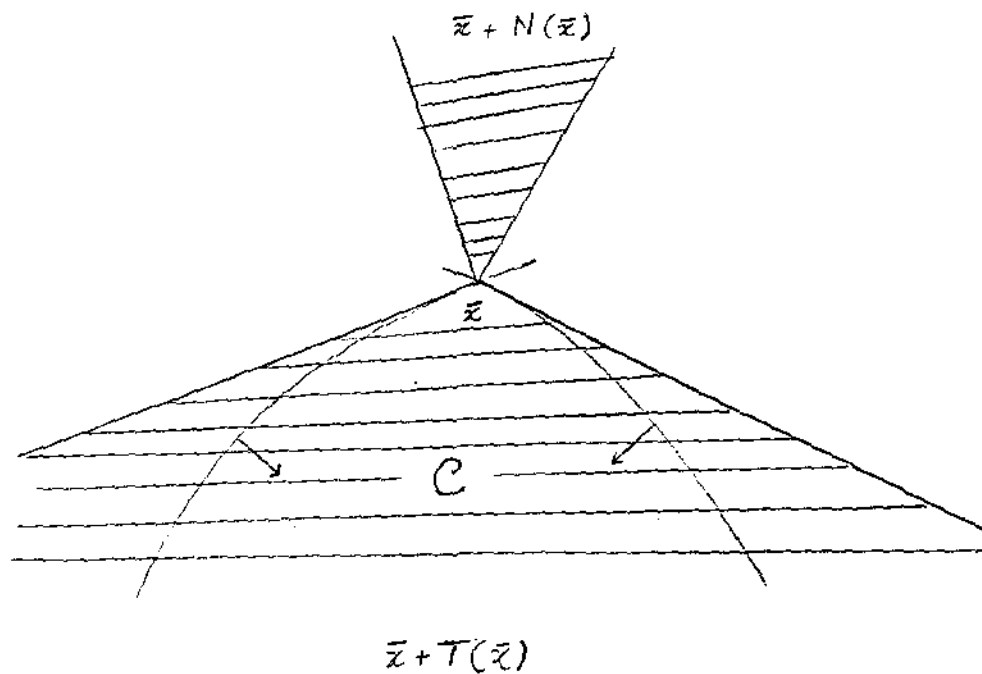


FIGURA 4.b

Cones normal e tangente do ponto $\bar{x}$.

Definição 2.9: Seja H um espaço de Hilbert e $\|\cdot\|$ a respectiva norma. Se $\Omega \subset H$ é um conjunto não-vazio, convexo e fechado, a projeção sobre Ω é definida como a aplicação

$$P : H \longrightarrow \Omega, \text{ onde } P(x) = \operatorname{argmin}\{\|z - x\| \mid z \in \Omega\}.$$

Referências para as definições e estudo do assunto, são: [1, 2, 16].

Usamos neste trabalho a decomposição de Moreau

$$x = P_{T(y)}(x) + P_{N(y)}(x),$$

que é válida para todo $x \in R^n$ e $y \in \Omega$ convexo (ver [15]), onde $P_{T(y)}(x)$ e $P_{N(y)}(x)$ são as projeções de x no cone tangente e normal no ponto y respectivamente.

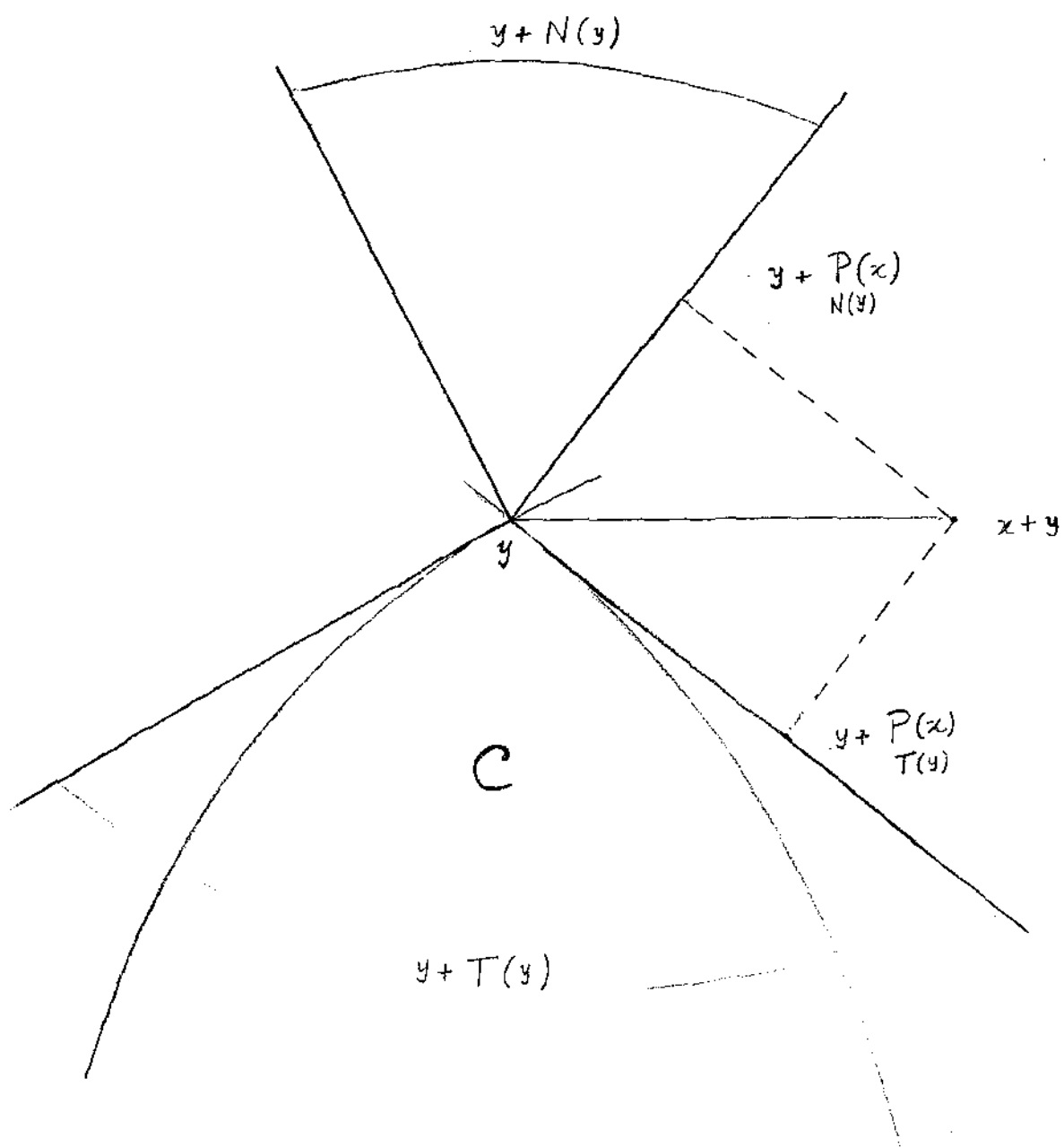


FIGURA 5

Decomposição de Moreau

2.2 LEMAS BÁSICOS

Lema 2.1: Seja H um espaço de Hilbert e $C \subset H$ um subconjunto convexo fechado.

Dados $x \in H$ e $P_C(x) = x_0$ a projeção de x sobre C , então

$$\langle x - x_0, y - x_0 \rangle \leq 0 \text{ para todo } y \in C.$$

Demonstração: Por definição $\|x - P_C(x)\| \leq \|x - z\|$ para todo $z \in C$. Suponhamos que existe $\bar{z} \in C$ tal que

$$\langle x - x_0, \bar{z} - x_0 \rangle > 0. \quad (2.2.1)$$

Definindo $z(t) = x_0 + t(\bar{z} - x_0)$, então para $t \in (0, 1)$,

$$z(t) \in C.$$

Por outro lado $\|z(t) - x\|^2 = \|x_0 - x + t(\bar{z} - x_0)\|^2 = \|x_0 - x\|^2 - 2t\langle x - x_0, \bar{z} - x_0 \rangle + t^2\|\bar{z} - x_0\|^2$.

Agora por (2.2.1), para t suficientemente pequeno, $\|z(t) - x\| < \|x_0 - x\|$ contradizendo que

$$x_0 = P_C(x).$$

□

Lema 2.2: Seja H um espaço de Hilbert. Se C é um subconjunto convexo fechado de H e $x_1, x_2 \in H$ então

$$\|P_C(x_1) - P_C(x_2)\| \leq \|x_1 - x_2\|.$$

Demonstração: Pelo Lema 2.1

$$\langle x_1 - P_C(x_1), P_C(x_2) - P_C(x_1) \rangle \leq 0 \quad (2.2.2)$$

e

$$\langle x_2 - P_C(x_2), P_C(x_1) - P_C(x_2) \rangle \leq 0. \quad (2.2.3)$$

Somando (2.2.2) e (2.2.3) obtemos

$$\langle P_C(x_1) - P_C(x_2), P_C(x_1) - P_C(x_2) \rangle \leq \langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle.$$

e assim,

$$\|P_C(x_1) - P_C(x_2)\|^2 \leq \langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle.$$

Agora

$$\langle P_C(x_1) - P_C(x_2), x_1 - x_2 \rangle \leq \|P_C(x_1) - P_C(x_2)\| \|x_1 - x_2\|,$$

portanto

$$\|P_C(x_1) - P_C(x_2)\| \leq \|x_1 - x_2\|.$$

□

Lema 2.3: Seja $C \subset R^n$ um conjunto convexo, $x \in r_i(C)$ e $y \in \overline{C}$ (fecho de C). Então $(1 - \lambda)x + \lambda y \in r_i(C)$ para todo $0 \leq \lambda < 1$.

Demonstração: Como $x \in r_i(C)$ então existe $\varepsilon > 0$ tal que

$$B(x, \varepsilon) \cap \text{aff}(C) \subset C.$$

Definimos $\varepsilon_\lambda = \frac{\varepsilon}{(1 + \lambda)(1 - \lambda)^{-1}}$, para $0 \leq \lambda < 1$. Notemos que $\varepsilon_\lambda > 0$.

Suponhamos agora que $r \in B(z_\lambda, \varepsilon_\lambda) \cap \text{aff}(C)$ onde

$$z_\lambda = (1 - \lambda)x + \lambda y. \quad (2.2.4)$$

Afirmamos que $r \in C$. Com efeito, como $y \in \overline{C}$, então existe $x_y \in C$ tal que $\|y - x_y\| < \varepsilon_\lambda$ e fazendo $v_y = \frac{y - x_y}{\varepsilon_\lambda}$ temos $y = x_y + \varepsilon_\lambda v_y$ e

$$\|v_y\| < 1. \quad (2.2.5)$$

Definimos agora

$$v_r = \frac{r - z_\lambda}{\varepsilon_\lambda}. \quad (2.2.6)$$

Por (2.2.4), (2.2.5) e (2.2.6),

$$r = (1 - \lambda)x + \lambda y + \varepsilon_\lambda v_r = (1 - \lambda)x + \lambda(x_y + \varepsilon_\lambda v_y) + \varepsilon_\lambda v_r,$$

portanto

$$r = (1 - \lambda) \left[x + \frac{\lambda}{1 + \lambda} \varepsilon v_y + \frac{1}{1 + \lambda} \varepsilon v_r \right] + \lambda x_y.$$

Agora notemos que $\left\| \frac{\lambda}{1 + \lambda} \varepsilon v_y + \frac{1}{1 + \lambda} \varepsilon v_r \right\| < \varepsilon$, considerando (2.2.5), (2.2.6) e o fato de que $r \in B(z_\lambda, \varepsilon_\lambda)$. Por outro lado $\frac{1}{1 - \lambda} r - \frac{\lambda}{1 - \lambda} x_y \in \text{aff}(C)$ já que r e $x_y \in \text{aff}(C)$. Logo, $x + \frac{\lambda}{1 + \lambda} \varepsilon v_y + \frac{1}{1 + \lambda} \varepsilon v_r \in B(x, \varepsilon) \cap \text{aff}(C) \subset C$. Portanto $r \in (1 - \lambda)C + \lambda C = C$.

□

Lema 2.4: Seja C um conjunto convexo não-vazio de $\mathbb{R}^n$. Então $z \in r_i(C)$, se e somente se, para todo $x \in C$, existe um $\theta > 1$ tal que $(1 - \theta)x + \theta z \in C$.

Demonstração: Se $z \in r_i(C)$ então existe $\varepsilon > 0$ tal que $B(z, \varepsilon) \cap \text{aff}(C) \subset C$. Se $x = z$, então é óbvio que $(1 - \theta)x + \theta z \in C$ para todo $\theta \in \mathbb{R}$ e em particular, $\theta > 1$. Se $x \neq z$ tomemos θ tal que $1 < \theta < \frac{\varepsilon}{\|z - x\|} + 1$. Para este θ escolhido vale então $(\theta - 1)\|z - x\| < \varepsilon$ implicando que

$$\|x + (z - x)\theta - z\| < \varepsilon.$$

Portanto $x + (z - x)\theta \in \text{aff}(C) \cap B(z, \varepsilon) \subset C$ e concluímos que $x + (z - x)\theta \in C$.

Suponhamos agora que $(1 - \theta)x + \theta z \in C$ com $\theta > 1$.

Fazendo $y = (1 - \theta)x + \theta z$ segue que $\frac{1}{\theta}y - \left(\frac{1 - \theta}{\theta}\right)x = z$.

Suponhamos que a escolha de x foi feita com $x \in r_i(C)$ (Observação: Se $C \neq \emptyset$ então $r_i(C) \neq \emptyset$. Ver [16]).

Tomando $\lambda = \theta^{-1}$ no Lema 2.3, concluímos que $z \in r_i(C)$.

□

Lema 2.5: Seja C um conjunto convexo e $f : R^n \rightarrow R, f \in C^1(A), C \subset A$ aberto. Para $x \in C$, seja $P_{T(x)}(-\nabla f(x))$ a projeção do vetor $-\nabla f(x)$ sobre o cone $T(x)$. Então

$$\min\{\langle \nabla f(x), v \rangle \mid v \in T(x), \|v\| \leq 1\} = -\|P_{T(x)}(-\nabla f(x))\|.$$

Demonstração: Inicialmente observamos que, se $u \in N(x)$ e $d \in T(x)$ então $\langle u, d \rangle \leq 0$.

Com efeito $\langle u, y - x \rangle \leq 0$ para todo $y \in C$. Por outro lado, se $\{d_k\}$ é uma sequência de direções factíveis no ponto x tal que $d_k \rightarrow d$, então $\langle d_k, u \rangle \leq 0$ e por continuidade do produto interno $\langle d, u \rangle \leq 0$.

Usando a decomposição de Moreau e o fato de que $v \in T(x)$ temos

$$\langle -\nabla f(x), v \rangle = \langle P_{T(x)}(-\nabla f(x)), v \rangle + \langle P_{N(x)}(-\nabla f(x)), v \rangle, \quad (2.2.7)$$

e pelo observado acima $\langle -\nabla f(x), v \rangle \leq \langle P_{T(x)}(-\nabla f(x)), v \rangle$ portanto $\langle -\nabla f(x), v \rangle \leq \|P_{T(x)}(-\nabla f(x))\|$, já que $\|v\| \leq 1$.

Se $P_{T(x)}(-\nabla f(x)) = 0$, então o resultado é imediato já que neste caso teríamos $\langle \nabla f(x), v \rangle \geq 0$, e como $0 \in T(x)$

$$\min\{\langle \nabla f(x), v \rangle \mid v \in T(x), \|v\| \leq 1\} = 0 = -\|P_{T(x)}(-\nabla f(x))\|.$$

Suponhamos que $P_{T(x)}(-\nabla f(x)) \neq 0$. Neste caso definimos $\bar{v} = \frac{P_{T(x)}(-\nabla f(x))}{\|P_{T(x)}(-\nabla f(x))\|}$.

Temos $\bar{v} \in T(x)$ e $\|\bar{v}\| = 1$.

Agora, por (2.2.7)

$$\langle -\nabla f(x), \bar{v} \rangle = \|P_{T(x)}(-\nabla f(x))\|.$$

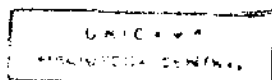
Usando o fato de que $\min f(x) = -\max [-f(x)]$, segue que

$$\min\{\langle \nabla f(x), v \rangle \mid v \in T(x), \|v\| \leq 1\} = -\|P_{T(x)}(-\nabla f(x))\|.$$

□

Lema 2.6: Seja A uma matriz $n \times n$, inversível. Existe $a > 0$ tal que $\|Ax\| \geq a\|x\|$ para todo $x \in R^n$.

Demonstração: Se $a = \frac{1}{\|A^{-1}\|}$, então para todo $x \in R^n, x = A^{-1}Ax$ logo $\|x\| \leq \|A^{-1}\|\|Ax\|$ e $a\|x\| \leq \|Ax\|$. □



Lema 2.7: (*Lema de Farkas*)

Seja $A \in R^{m \times n}$ e $c \in R^n$. Então um e somente um dos dois sistemas abaixo tem solução.

$Ax \leq 0, c^T x > 0$ para algum $x \in R^n$.

$A^T y = c, y \geq 0$ para algum $y \in R^m$.

Demonstração: Ver [1]. □

Lema 2.8: Consideremos o conjunto $V = \{v_1, v_2, \dots, v_m\}, v_i \in R^n$. Suponhamos que $T = \{u \in R^n \mid u^T v_i \leq 0, i = 1, \dots, m\}$, então

$$T^0 = \{x \in R^n \mid x = \sum_{i=1}^m \rho_i v_i, \rho_i \geq 0\}.$$

Demonstração: Seja A a matriz cujas linhas são formadas pelos vetores $v_i, i = 1, \dots, m$. Seja $c \in T^0$, então o sistema

$$\begin{aligned} Ad &\leq 0 \\ c^T d &> 0, d \in R^n \end{aligned} \tag{2.2.8}$$

não tem solução.

Com efeito se $Ad \leq 0$ então $d \in T^0$. Como $c \in T^0$, segue que $c^T d \leq 0$. Portanto o sistema (2.2.8) não tem solução. Pelo Lema de Farkas existe $y \in R^m, y \geq 0$ tal que $A^T y = c$, ou seja, $c = \sum_{i=1}^m y_i v_i$ com $y_i \geq 0$. Como “ c ” é genérico, vale o lema. □

Lema 2.9: Seja Ω um conjunto convexo e fechado, $x \in \Omega$ e $v \in N(x)$. Consideremos $P_\Omega(\cdot)$ a projeção do R^n sobre o conjunto Ω . Então

$$P_\Omega(x + v) = x.$$

Demonstração: Pelo Lema 2.1

$$\langle (x + v) - P_\Omega(x + v), x - P_\Omega(x + v) \rangle \leq 0.$$

Portanto

$$\begin{aligned} \langle v, x - P_{\Omega}(x + v) \rangle &\leq -\|x - P_{\Omega}(x + v)\|^2 \\ \text{e} \\ \|x - P_{\Omega}(x + v)\|^2 &\leq \langle v, P_{\Omega}(x + v) - x \rangle. \end{aligned}$$

Agora, se $x \neq P_{\Omega}(x + v)$ temos $\|x - P_{\Omega}(x + v)\| > 0$ e assim $\langle v, P_{\Omega}(x + v) - x \rangle > 0$ o que é um absurdo já que $v \in N(x)$. Portanto,

$$P_{\Omega}(x + v) = x.$$

□

Lema 2.10: Seja $\Omega \subset \mathbb{R}^n$ um conjunto convexo fechado e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continuamente diferenciável sobre $A \subset \mathbb{R}^n$, conjunto aberto que contém Ω . Se $x \in \Omega$, $\rho > 0$ e $P_{\Omega}(x - \rho \nabla f(x)) = y$ é a projeção do vetor $x - \rho \nabla f(x)$ sobre o conjunto Ω , então

$$\|P_{T(y)}(-\nabla f(y))\| \leq \|\nabla f(y) - \nabla f(x)\| + \frac{1}{\rho}\|x - y\|.$$

Demonstração: Pelo Lema 2.1,

$$\langle y - (x - \rho \nabla f(x)), y - z \rangle \leq 0 \quad (2.2.9)$$

para todo $z \in \Omega$.

Escrevendo (2.2.9) de forma conveniente e usando a desigualdade de Cauchy – Schwarz

$$\rho \langle \nabla f(x), y - z \rangle \leq -\langle y - x, y - z \rangle \leq \|y - x\| \|y - z\|. \quad (2.2.10)$$

Agora, se v é uma direção viável em y tal que $\|v\| \leq 1$, então $z_0 = y + \theta v \in \Omega$ para algum $\theta > 0$. Substituindo z_0 em (2.2.10) obtemos

$$-\langle \nabla f(x), v \rangle \leq \frac{1}{\rho} \|y - x\|. \quad (2.2.11)$$

Por outro lado

$$\langle -\nabla f(y) + \nabla f(x), v \rangle = -\langle \nabla f(y), v \rangle + \langle \nabla f(x), v \rangle,$$

logo

$$-\langle \nabla f(y), v \rangle \leq \|\nabla f(y) - \nabla f(x)\| - \langle \nabla f(x), v \rangle. \quad (2.2.12)$$

Substituindo (2.2.11) em (2.2.12) pelo Lema 2.5 resulta

$$\|P_{T(y)}(-\nabla f(y))\| \leq \|\nabla f(y) - \nabla f(x)\| + \frac{1}{\rho}\|y - x\|.$$

□

CAPÍTULO 3

IDENTIFICAÇÃO DE RESTRIÇÕES ATIVAS

3.1 INTRODUÇÃO

Neste capítulo apresentaremos os resultados de identificação para o problema

$$\begin{array}{ll} \min & f(x) \\ \text{sujeita a} & x \in \Omega, \end{array} \quad (3.1.1)$$

onde $\Omega \subset R^n$ é convexo, fechado e não vazio definido por

$$\Omega = \{x \in R^n \mid h_i(x) \leq 0, i = 1, \dots, m\},$$

$f, h_i \in C^1(A), i = 1, \dots, m, \Omega \subset A$ aberto.

Se $x \in \Omega$ e existem restrições ativas no ponto x , definimos o conjunto de índices $A(x) \subset I = \{1, \dots, m\}$ como

$$A(x) = \{i \in I \mid h_i(x) = 0\}.$$

3.2 O ALGORITMO RCARB

Nesta seção apresentamos o algoritmo RCARB abordado em [13], cuja sigla significa Regiões de Confiança em Conjuntos Arbitrários.

Algoritmo RCARB: Dados $\gamma_1, \gamma_2, \theta, \Delta_{min}, M, \gamma$ tais que $0 < \gamma_1 < \gamma_2 < 1, \theta \in (0, 1], \Delta_{min} > 0, M > 0$ e $\gamma \in (0, 1]$, um ponto inicial $x_0 \in \Omega$, uma matriz B_0 tal que $\|B_0\| \leq M$ e um raio inicial $\Delta^0 \geq \Delta_{min}$, obtidos $x_k \in \Omega, B_k = B_k^T \in R^{n \times n}$ tal que $\|B_k\| \leq M$ e $\Delta^k \geq \Delta_{min}$, os passos para definir Δ_k e x_{k+1} são os seguintes:

PASSO 0: Faça $\Delta \leftarrow \Delta^k$

PASSO 1: (Subproblema Fácil) Calcule $s_k^Q(\Delta)$ a solução global de

$$\begin{aligned} \min \quad & Q_k(s) = \frac{1}{2}M\|s\|^2 + g_k^T s, \\ \text{sujeita a} \quad & x_k + s \in \Omega, \\ & \|s\| \leq \Delta, \end{aligned} \tag{3.2.1}$$

onde $g_k = g(x_k)$. Se $Q_k(s_k^Q) = 0$ pare.

PASSO 2: Calcule $\bar{s}_k(\Delta)$ tal que

$$\begin{aligned} \psi_k(\bar{s}_k(\Delta)) &\leq \gamma Q_k(s_k^Q(\Delta)), \\ x_k + \bar{s}_k(\Delta) &\in \Omega, \\ \|\bar{s}_k(\Delta)\| &\leq \Delta, \end{aligned}$$

onde ψ_k é definida por

$$\psi_k(s) = \frac{1}{2}s^T B_k s + g_k^T s$$

para todo $s \in R^n$.

PASSO 3: Se

$$f(x_k + \bar{s}_k(\Delta)) \leq f(x_k) + \theta \psi_k(\bar{s}_k(\Delta)),$$

então $x_{k+1} = x_k + \bar{s}_k(\Delta)$

$$\Delta_k = \Delta.$$

Escolha $\Delta^{k+1} \geq \Delta_{min}$ e $B_{k+1} \in R^{n \times n}$ simétrica tal que $\|B_{k+1}\| \leq M$.

Senão $\Delta \leftarrow \Delta_{novo}$, onde

$$\Delta_{novo} \in [\gamma_1 \|\bar{s}_k(\Delta)\|, \gamma_2 \Delta],$$

e volte para o Passo 1.

Se $\langle, \rangle$ é um produto interno e $\|\cdot\|$, a correspondente norma, então a solução $s_k^Q(\Delta)$ de (3.2.1) verifica

$$\begin{aligned} s_k^Q(\Delta) = \operatorname{argmin} & \left\| s + \frac{g_k}{M} \right\|, \\ & x_k + s \in \Omega, \\ & \|s\| \leq \Delta. \end{aligned}$$

Com efeito

$$\begin{aligned} \left\| s + \frac{g_k}{M} \right\|^2 &= \left\langle s + \frac{g_k}{M}, s + \frac{g_k}{M} \right\rangle = \\ &= \langle s, s \rangle + 2\left\langle s, \frac{g_k}{M} \right\rangle + \frac{1}{M^2} \|g_k\|^2. \end{aligned}$$

$$\text{Mas } \|s\|^2 + \frac{2}{M} \langle s, g_k \rangle = \frac{2}{M} Q_k(s).$$

Logo, $\left\| s + \frac{g_k}{M} \right\|^2 = \frac{2}{M} Q_k(s) + \frac{1}{M^2} \|g_k\|^2$. Portanto,

$$\begin{aligned} s_k^Q(\Delta) = \operatorname{argmin} & \left\| s + \frac{g_k}{M} \right\|, \\ & x_k + s \in \Omega, \\ & \|s\| \leq \Delta. \end{aligned} \tag{3.2.2}$$

Todos os resultados apresentados a seguir, exploram a propriedade acima observada. Notemos que $s_k^Q(\Delta)$ está bem definido em (3.2.2) já que o conjunto factível é convexo e fechado.

3.3 RESULTADOS PRELIMINARES

Lema 3.1: Consideremos o conjunto Ω definido em (3.1.1). Seja x^* um ponto regular de Ω . Então existe um $\varepsilon > 0$ tal que, se $x \in B(x^*, \varepsilon) \cap \Omega$, o conjunto $\{\nabla h_i(x)\}_{i \in A(x^*)}$ é linearmente independente.

Demonstração: Como x^* é regular segue por definição, que o $\{\nabla h_i(x^*)\}_{i \in A(x^*)}$ é linearmente independente.

Suponhamos agora que para $\varepsilon = \frac{1}{k}$, existe $x_k \in B(x^*, \varepsilon) \cap \Omega$ tal que o conjunto $\{\nabla h_i(x_k)\}_{i \in A(x^*)}$ é linearmente dependente. Assim para cada $k \in \mathbb{N}$, existem $\{c_i^k\}$, não todos nulos, tais que $\sum_{i \in A(x^*)} c_i^k \nabla h_i(x_k) = 0$.

Para cada $k \in \mathbb{N}$, consideremos o vetor c^k cujas componentes são $(c_i^k)_{i \in A(x^*)}$, pelo afirmado acima $\|c^k\| \neq 0$ para todo $k \in \mathbb{N}$. Consideremos a sequência $u_k = \frac{c^k}{\|c^k\|}$. Sendo u_k limitada, existe $K_1 \subset \mathbb{N}$ infinito tal que

$$\lim_{k \in K_1} u_k = u \neq 0. \quad (3.3.1)$$

Assim, se $h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix}$ então por definição de u_k , $J_h^T(x_k)u_k = 0$ para todo $k \in K_1$.

Como $h \in C^1$, segue que

$$\lim_{k \in K_1} J_h^T(x_k)u_k = J_h^T(x^*)u = 0. \quad (3.3.2)$$

Mas (3.3.2) é uma contradição já que $J_h^T(x^*)$ tem posto completo e $u \neq 0$. Portanto, vale o resultado do Lema. $\square$

O Lema a seguir, permite-nos caracterizar o cone tangente a um ponto regular $x \in \Omega$, onde Ω é definido em 3.1.1.

Lema 3.2: Seja Ω o conjunto definido em (3.1.1). Se x^* é um ponto regular de Ω , então o cone tangente em x^* é:

$$T(x^*) = \{v \in \mathbb{R}^n \mid \langle v, \nabla h_j(x^*) \rangle \leq 0, j \in A(x^*)\}.$$

Demonstração: Consideremos o conjunto

$$S = \{v \in \mathbb{R}^n \mid \langle v, \nabla h_j(x^*) \rangle \leq 0, j \in A(x^*)\}.$$

Se $d \in T(x^*)$, então existe uma sequência de direções factíveis $\{d_k\}$, em x^* , tal que $d_k \rightarrow d$. Mas $\nabla h_j^T(x^*)d_k \leq 0$ para todo $j \in A(x^*)$ e pela continuidade do produto interno, $\nabla h_j^T(x^*)d \leq 0$ para todo $j \in A(x^*)$. Logo $d \in S$ e

$$T(x^*) \subset S. \quad (3.3.3)$$

Para mostrar a outra inclusão, consideremos dois casos:

Caso 1: $v = 0, 0 \in T(x^*)$ e a inclusão é trivial.

Caso 2: $v \neq 0$ e $\langle v, \nabla h_j(x^*) \rangle \leq 0$ para todo $j \in A(x^*)$.

Neste caso, sejam $A_1, A_2 \subset A(x^*)$ tais que:

$$\begin{aligned}\langle v, \nabla h_j(x^*) \rangle &< 0 \text{ se } j \in A_1 \\ \langle v, \nabla h_j(x^*) \rangle &= 0 \text{ se } j \in A_2\end{aligned}$$

Pela regularidade de x^* , existe $b > 0$ e $x(t)$, curva passando por x^* tal que $x'(0) = v, x(0) = x^*$ e $h_j(x(t)) = 0 \forall j \in A_2$ e $t \in (0, b)$. Definimos uma sequência $\{t_k\}$ de números reais, do intervalo $(0, b)$, tal que $t_k \rightarrow 0$. Fazendo

$$d_k = \frac{x(t_k) - x(0)}{t_k} \quad (3.3.4)$$

verificamos facilmente que $\lim_{k \rightarrow \infty} d_k = v$. Afirmamos que existe k_0 tal que para $k \geq k_0$ d_k é factível. Para mostrar que a afirmação é verdadeira, verificamos inicialmente que para k suficientemente grande e $k \in A_1, \langle d_k, \nabla h_j(x^*) \rangle < 0$. Com efeito, se $\{d_k\}_{k \in N_1} \subset N$ é uma subsequência de $\{d_k\}_{k \in N}$ satisfazendo $\langle d_k, \nabla h_j(x^*) \rangle \geq 0$ para todo $k \in N_1$ e $j \in A_1$, então

$$\lim_{k \in N_1} \langle d_k, \nabla h_j(x^*) \rangle = \langle v, \nabla h_j(x^*) \rangle \geq 0,$$

que é uma contradição. Usando a fórmula de Taylor, concluímos que existe $k_1 \in \mathbb{N}$ tal que para todo $k \geq k_1$ e $j \in A(x^*), h_j(x(t)) \leq 0$. Suponha agora, que $j \notin A(x^*)$. Como $h_j(x^*) < 0$, então pela continuidade de h_j , existe $\varepsilon > 0$ tal que se $x \in B(x^*, \varepsilon)$, então $h_j(x) < 0$. Pela continuidade da curva $x(t)$, existe $k_2 \in \mathbb{N}$ tal que para todo $k \geq k_2$

$$\|x(t_k) - x(0)\| < \varepsilon.$$

Assim para $k \geq \max\{k_1, k_2\}, h_j(x(t_k)) \leq 0$ para $j = 1, \dots, m$. Como Ω é convexo, segue que d_k é uma direção factível para todo $k \geq \max\{k_1, k_2\}$. Fazendo $k_0 = \max\{k_1, k_2\}$, então $\{d_k\}_{k \in N, k \geq k_0}$ é uma sequência de direções factíveis convergindo para v . Portanto, $v \in T(x^*)$. $\square$

Teorema 3.1: Seja K o cone definido por

$$K = \{v \mid v = \sum_{i=1}^m \lambda_i v_i, \lambda_i \geq 0, v_i \in R^n\}.$$

Suponhamos que $v_0 = \sum \lambda_i v_i$ com $\lambda_i > 0$. Então $v_0 \in r_i(K)$.

Demonstração: Primeiramente notemos que

$$aff(K) = span\{v_i \mid i = 1, \dots, m\}.$$

Definimos I tal que $M = \{v_i \mid i \in I\}$ seja uma base de $aff(K)$. Pela equivalência de normas no R^n , existe $b > 0$, tal que $b\|x\|_\infty \leq \|x\|$. Definimos agora a matriz V cujas colunas são formadas pelos elementos de M . Pelo Lema 2.6, existe $a > 0$ tal que $a\|u\| \leq \|V^T V u\|$ para todo $u \in R^n$. Tomemos agora $\varepsilon = \frac{ab\varepsilon_0}{\|V^T\|}$ onde ε_0 é escolhido de tal sorte que $0 < \varepsilon_0 < \min\{\lambda_i \mid i = 1, \dots, m\}$.

Agora para todo $w = V u$ tal que $\|w\| \leq \varepsilon$ temos

$$a\|u\| \leq \|V^T V u\| \leq \|V^T\| \|w\| \leq \|V^T\| \varepsilon,$$

portanto

$$a\|u\| \leq ab\varepsilon_0 \text{ e } \|u\| \leq b\varepsilon_0.$$

Mas $b\|u\|_\infty \leq \|u\| \leq b\varepsilon_0$ implica $\|u\|_\infty \leq \varepsilon_0$ e pela escolha de ε_0 , segue que $|u_i| < \min\{\lambda_i \mid i = 1, \dots, m\}$. Assim sendo $v_0 + w \in (v_0 + B(0, \varepsilon)) \cap aff(K) \subset K$ e

$$v_0 \in r_i(K).$$

□

Teorema 3.2: Consideremos os conjuntos $G = \{v_1, \dots, v_p\}$ e $C = \{v \in R^n \mid v = \sum_{i=1}^p \lambda_i v_i, \lambda_i \geq 0\}$. Se $0 \in r_i(C)$ então o conjunto G é linearmente dependente.

Demonstração: Tomemos

$$v_0 = \sum_{i=1}^p \lambda_i v_i, \lambda_i > 0 \text{ para todo } i \in \{1, \dots, p\}. \quad (3.3.5)$$

Pelo Lema 2.4, existe $\theta > 1$ tal que $(1 - \theta)v_0 + \theta \cdot 0 \in C$. Assim

$$(1 - \theta)v_0 \in C \text{ implica } (1 - \theta)v_0 = \sum_{i=1}^p r_i v_i, r_i \geq 0. \quad (3.3.6)$$

Substituindo (3.3.5) em (3.3.6) e escrevendo de forma conveniente

$$\sum_{i=1}^p [r_i - (1 - \theta)\lambda_i] v_i = 0,$$

e como $r_i - (1 - \theta)\lambda_i > 0$ segue que o conjunto G é linearmente dependente. $\square$

Corolário 3.2.1: Consideremos o mesmo conjunto G do Teorema 3.2. Se G é linearmente independente então $0 \notin r_i(C)$.

Teorema 3.3: Seja $\Omega \subset \mathbb{R}^n$ definido como em (3.1.1). Consideremos $\{x_k\}$ uma sequência de pontos de Ω convergindo para x^* regular. Suponhamos ainda, que x^* é um ponto estacionário para o problema (3.1.1). Além disso, se $h_i(x^*) = 0$, então o respectivo multiplicador λ_i é positivo ($\lambda_i > 0$). Então, para k suficientemente grande, $M > 0$ e $y_k = P_\Omega\left(x_k - \frac{\nabla f(x_k)}{M}\right)$

$$h_i(y_k) = 0 \text{ para todo } i \in A(x^*).$$

Demonstração: Pelo Lema 2.2, $y_k \rightarrow x^*$. Agora, suponhamos falsa a afirmação do teorema. Isto só pode acontecer de duas maneiras:

1. Existe $k_0 \in \mathbb{N}$ tal que, para todo $k \geq k_0$, y_k está no interior do conjunto Ω (se $\text{int}(\Omega) \neq \emptyset$).

Neste caso é óbvio que $P_{T(y_k)}(-\nabla f(y_k)) = -\nabla f(y_k)$ para $k \geq k_0$. Pelo Lema 2.10

$$\lim_k P_{T(y_k)}(-\nabla f(y_k)) = -\nabla f(x^*) = 0, \quad (3.3.7)$$

e como x^* é estacionário, então $-\nabla f(x^*) = \sum_{i \in A(x^*)} \lambda_i \nabla h_i(x^*)$, com $\lambda_i > 0$ pela hipótese. Pelo Teorema 3.1 $-\nabla f(x^*) \in r_i(N(x^*))$ e por (3.3.7)

$$0 \in r_i(N(x^*)). \quad (3.3.8)$$

Pelos lemas 3.2 e 2.8

$$N(x^*) = \{v \in R^n \mid v = \sum_{i \in A(x^*)} \rho_i \nabla h_i(x^*), \rho_i \geq 0\}.$$

Mas, pelo Corolário 3.2.1, (3.3.8) é uma contradição à hipótese de regularidade de x^* .

2. Existem $\bar{A} \subset A(x^*)$, próprio, e $K_1 \subset \mathbb{N}$ infinito tais que $h_j(y_k) = 0$ para todo $j \in \bar{A}$ e $k \in K_1$, porém, $h_j(y_k) < 0$ se $j \in A(x^*) - \bar{A}$.
Usando a decomposição de Moreau, obtemos

$$-\nabla f(y_k) = P_{N(y_k)}(-\nabla f(y_k)) + P_{T(y_k)}(-\nabla f(y_k)). \quad (3.3.9)$$

Agora, pelo Lema 3.1, para k suficientemente grande o conjunto $\{\nabla h_j(y_k)_{j \in \bar{A}}\}$ é linearmente independente e usando os mesmos argumentos da demonstração do Lema 3.2 e o resultado do Lema 2.8, concluímos que

$$P_{N(y_k)}(-\nabla f(y_k)) = \sum_{j \in \bar{A}} \lambda_j^k \nabla h_j(y_k), \quad (3.3.10)$$

com $\lambda_j^k \geq 0$ para k suficientemente grande.

Afirmamos que a sequência $\lambda^k = (\lambda_j^k)_{j \in \bar{A}}$ é limitada.

Com efeito, se λ^k não é limitada, então existe $K_2 \subset K_1$ tal que $\lim_{k \in K_2} \|\lambda^k\| = \infty$ e

$\lambda^k \neq 0$ para todo $k \in K_2$. Por outro lado como $\frac{\lambda^k}{\|\lambda^k\|}$ é limitada, existe $K_3 \subset K_2$ tal que

$$\lim_{k \in K_3} \frac{\lambda^k}{\|\lambda^k\|} = \lambda \neq 0.$$

Como o primeiro membro e o segundo termo do segundo membro de (3.3.9) convergem, então existe $v^* \in R^n$ tal que

$$\lim_{k \in K_3} P_{N(y_k)}(-\nabla f(y_k)) = v^*. \quad (3.3.11)$$

Definindo para $k \in K_3$ suficientemente grande

$$v_k = P_{N(y_k)}(-\nabla f(y_k)) = \sum_{j \in \bar{A}} \lambda_j^k \nabla h_j(y_k),$$

então

$$\lim_{k \in K_3} \frac{v_k}{\|\lambda^k\|} = 0 = \sum_{j \in \bar{A}} \lambda_j \nabla h_j(x^*) \text{ onde } \lambda_j \text{ é}$$

tal que $\lambda = (\lambda_j)_{j \in \bar{A}}$. Mas isto é uma contradição pois $\{\nabla h_j(x^*)\}_{j \in \bar{A}} \subset \{\nabla h_i(x^*)\}_{i \in A(x^*)}$, e qualquer subconjunto de um conjunto linearmente independente é linearmente independente. Portanto $\{\lambda^k\}$ é limitado.

Extraindo uma subsequência convergente $\{\lambda^k\}_{k \in K_4 \subset K_1}$ e sendo $\bar{\lambda} = (\bar{\lambda}_j)_{j \in \bar{A}}$ tal que $\{\lambda^k\}_{k \in K_4} \rightarrow \bar{\lambda}$, e tomando o limite em (3.3.9) com $k \in K_4$ teremos

$$-\nabla f(x^*) = \sum_{j \in \bar{A}} \bar{\lambda}_j \nabla h_j(x^*) + 0, \bar{\lambda}_j \geq 0. \quad (3.3.12)$$

Agora, se $\bar{\lambda}_j = 0$ para todo $j \in \bar{A}$, então $0 \in r_i(N(x^*))$ contradizendo a regularidade de x^* , pelo resultado do teorema 3.2.

Suponhamos que existe $j \in \bar{A}$ tal que $\bar{\lambda}_j > 0$. Pela hipótese de estacionariedade de x^* ,

$$-\nabla f(x^*) = \sum_{i \in A(x^*)} \lambda_i \nabla h_i(x^*). \quad (3.3.13)$$

Comparando (3.3.12) e (3.3.13) chegamos novamente a uma contradição à regularidade de x^* , já que a inclusão $\bar{A} \subset A(x^*)$ é estrita. $\square$

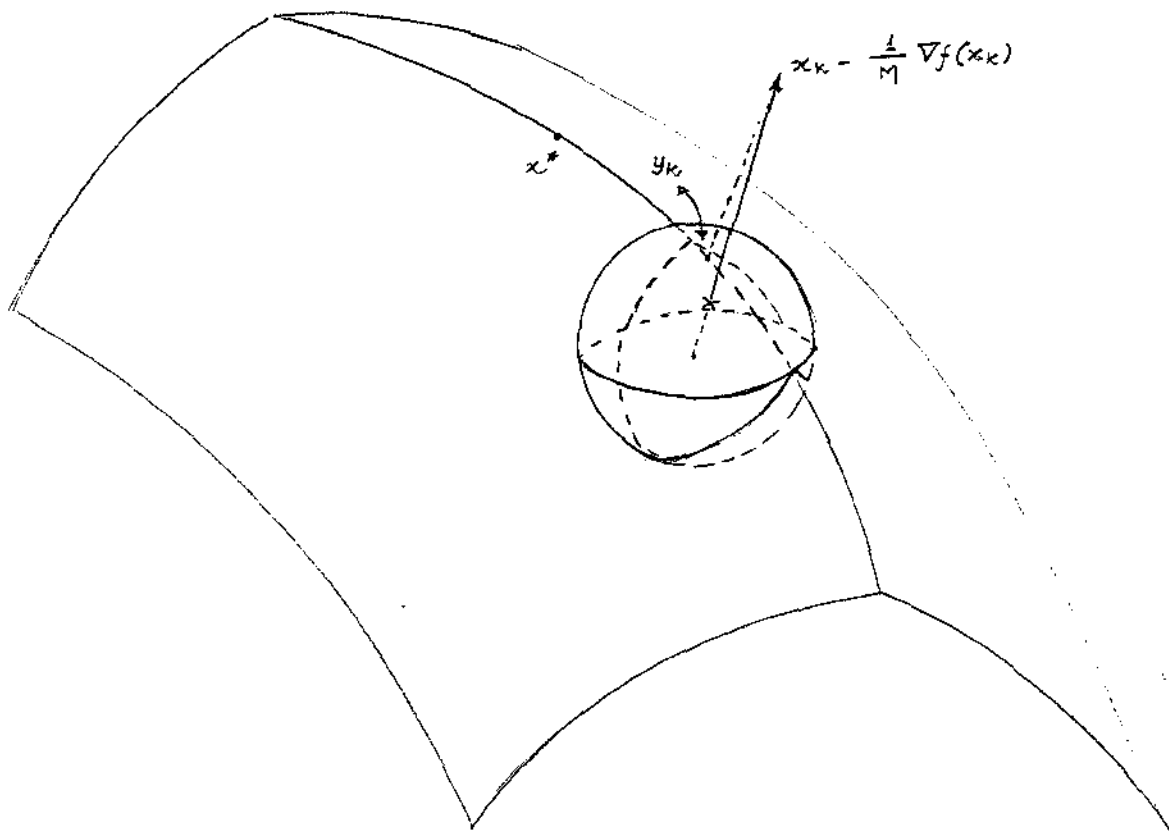


FIGURA 6

Para k suficientemente grande, $h_i(y_k) = 0$ com $i \in A(x^*)$ e

$$y_k = P_{\Omega} \left(x_k - \frac{1}{M} \nabla f(x_k) \right).$$

3.4 IDENTIFICAÇÃO DAS RESTRIÇÕES ATIVAS NO ALGORITMO RCARB

Nesta seção, provamos os teoremas de identificação para o algoritmo RCARB. O lema seguinte é uma consequência da propriedade observada em (3.2.2).

Lema 3.3: Seja $x_k^A = x_k + s_k^Q(\Delta)$ o ponto auxiliar apartir da solução do subproblema (3.2.1).
Então,

$$\|P_{T(x_k^A)}(-\nabla f(x_k^A))\| \leq \|\nabla f(x_k^A) - \nabla f(x_k)\| + M\|x_k^A - x_k\|.$$

Demonstração: Análoga à do Lema 2.10, considerando Ω como sendo a região definida em 3.1.1. $\square$

O próximo teorema destaca uma propriedade do ponto auxiliar que é fundamental para os teoremas de identificação.

Teorema 3.4: Consideremos o problema (3.1.1). Seja $\{x_k\}$ uma sequência infinita gerada pelo algoritmo *RCARB*, $K_1 \subset \mathbb{N}$ um conjunto infinito de índices tal que $\lim_{k \in K_1} x_k = x^*$. Suponhamos que x^* é regular e que se $h_i(x^*) = 0$, então o multiplicador correspondente λ_i é positivo ($\lambda_i > 0$).
Então existe $k_1 \in K_1$ tal que

$$h_i(x_k^A) = 0 \text{ para todo } k \in K_1, k \geq k_1.$$

Demonstração: Como $\{x_k\}$ é uma sequência infinita, então, pelos resultados de convergência global apresentados em [13], x^* é um ponto estacionário do problema (3.1.1). Assim,

$$-\nabla f(x^*) = \sum_{i \in A(x^*)} \lambda_i \nabla h_i(x^*), \quad (3.3.14)$$

com $\lambda_i > 0$, pela hipótese.

Lembrando que $y_k = P_\Omega(x_k - \frac{1}{M}\nabla f(x_k))$, consideraremos dois casos:

Caso 1: Existe $k_3 \in K_1$ tal que $\|y_k - x_k\| \leq \Delta_k$ para todo $k \geq k_3, k \in K_1$.

Caso 2: Existe $K_2 \subset K_1$ infinito tal que $\|y_k - x_k\| > \Delta_k$ para todo $k \in K_2$.

Suponhamos que acontece o Caso 1,

Pelo Teorema 3.3 existe $k_0 \in K_1$ tal que $h_i(y_k) = 0$ para todo $k \geq k_0, k \in K_1$. Como $\|y_k - x_k\| \leq \Delta_k$ para todo, $k \geq k_3, k \in K_1$, então $y_k = x_k^A$ para todo $k \geq \max\{k_0, k_3\}$. Logo, $h_i(x_k^A) = 0$ para todo $k \geq \max\{k_0, k_3\}$.

Suponhamos que acontece o Caso 2,

Pelo Lema 2.2 e a continuidade da aplicação ∇f segue que

$$y_k \rightarrow x^*.$$

Ora, $\|y_k - x_k\| > \Delta_k \geq \|x_k^A - x_k\|$ para todo $k \in K_2$, então

$$\lim_{k \in K_2} \|x_k^A - x_k\| = 0. \quad (3.3.15)$$

Pelo Lema (3.3)

$$\lim_{k \in K_2} P_{T(x_k^A)}(-\nabla f(x_k^A)) = 0. \quad (3.3.16)$$

Suponhamos agora que não existe $k_1 \in K_1$ com a propriedade do enunciado. Então temos duas possibilidades a considerar:

1. Existe $K_3 \subset K_2$ infinito tal que $x_k^A \in \text{int}(\Omega)$ para todo $k \in K_3$.
2. Existe $K_4 \subset K_2$ tal que o $A(x_k^A) \subset A(x^*)$ e $A(x^*) \neq A(x_k^A)$ qualquer que seja x_k^A com $k \in K_4$.

Na primeira possibilidade é imediato que

$$P_{T(x_k^A)}(-\nabla f(x_k^A)) = -\nabla f(x_k^A),$$

e por (3.3.16)

$$\lim_{k \in K_2} P_{T(x_k^A)}(-\nabla f(x_k^A)) = -\nabla f(x^*) = 0,$$

logo por (3.3.14) e o Teorema 3.1, $0 \in r_i(N(x^*))$. Mas isto é uma contradição considerando que x^* é regular e o resultado do Corolário 3.2.1 .

Na segunda possibilidade, como o número de restrições é finito existe $K_5 \subset K_4$ infinito e $\tilde{A} \subset A(x^*)$, próprio tal que:

$$A(x_k^A) = \tilde{A} \text{ para todo } k \in K_5.$$

Agora usando os mesmos argumentos apresentados na demonstração do Teorema 3.3 (Caso 2), concluiremos que o conjunto $\{\nabla h_i(x^*)\}_{i \in A(x^*)}$ é linearmente dependente, o que é uma contradição já que x^* é regular.

Portanto, a segunda possibilidade não pode acontecer se as restrições que são ativas na solução x^* , não são identificadas num número finito de iterações. Esgotadas as duas possibilidades, concluímos que existe $k_0 \in \mathbb{N}$ tal que para $k \geq k_0, k \in K_1$

$$h_i(x_k^A) = 0 \text{ para } i \in A(x^*).$$

□

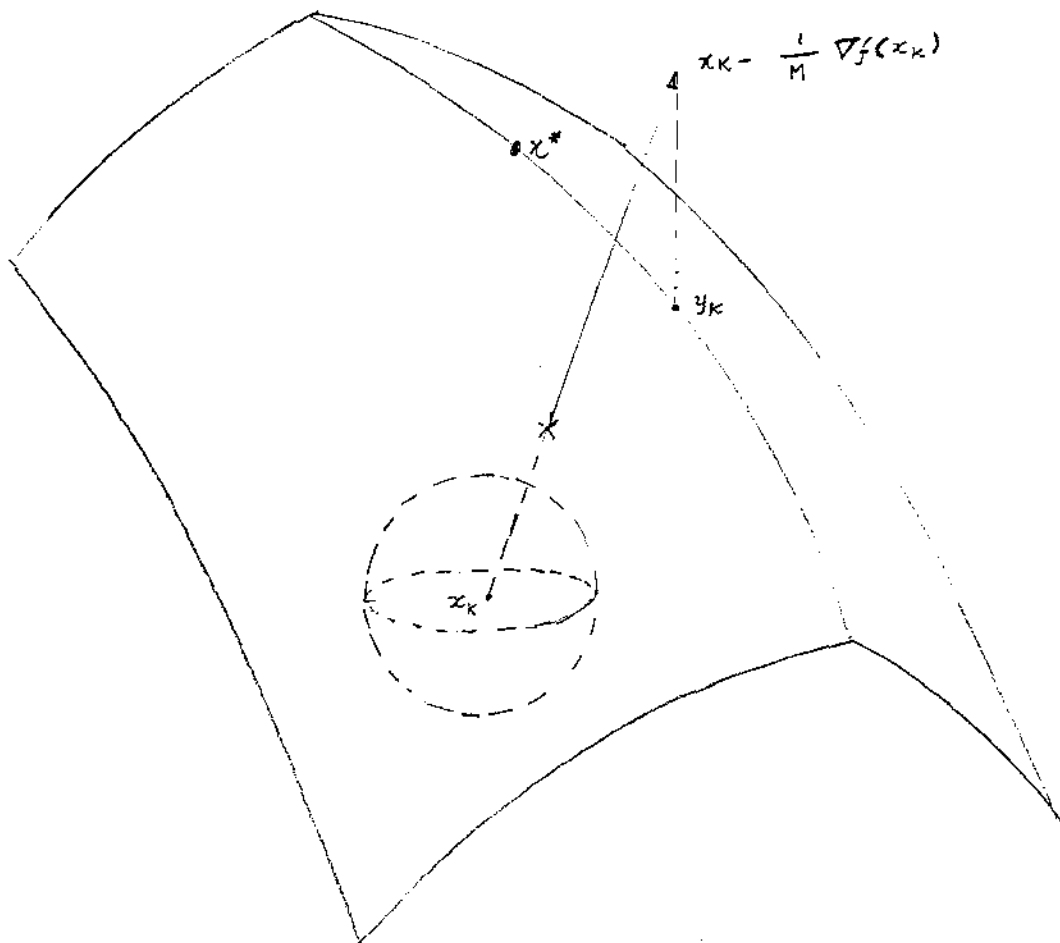


FIGURA 7

Para k suficientemente grande $y_k = x_k^A$.

Para impedir que as restrições corretas que foram identificadas pelo ponto auxiliar, sejam abandonadas no Passo 3 do algoritmo RCARB, faremos uma hipótese adicional no próximo teorema. Essa hipótese não compromete os resultados de convergência global.

Teorema 3.5: Consideremos as mesmas hipóteses do Teorema 3.4. Além disso, suponhamos que se $h_i(x_k^A) = 0$, então $h_i(x_{k+1}) = 0$. Então existe $k_0 \in \mathbb{N}$ tal que para todo $k \geq k_0$

$$A(x_k) \subset A(x_{k+1}).$$

Demonstração: Suponhamos por contradição que existe $K_5 \subset K_1$ infinito tal que

$$A(x_k) \not\subset A(x_{k+1}) \text{ para todo } k \in K_5. \quad (3.3.17)$$

Como o número de restrições é finito, existe $i \in \{1, \dots, m\}$ e $K_6 \subset K_5$ infinito tal que

$$i \in A(x_k) \text{ mas } i \notin A(x_{k+1}) \text{ para todo } k \in K_6.$$

Pela continuidade de h_i , se $i \in A(x_k)$, então $i \in A(x^*)$. Pelo Teorema 3.4 $h_i(x_k^A) = 0$ para k suficientemente grande e pela hipótese adicional, $h_i(x_{k+1}) = 0$ implicando que $i \in A(x_{k+1})$. Mas isto contradiz (3.3.17). Segue então o enunciado do teorema. $\square$

Teorema 3.6: (*Identificação das Restrições Ativas*)

Consideremos o problema (3.1.1). Suponhamos que $\{x_k\}$ seja uma sequência infinita gerada pelo algoritmo RCARB tal que $\lim_{k \in K_1 \subset \mathbb{N}} x_k = x^*$. Suponhamos que x^* é regular e que se $h_i(x^*) = 0$, então o correspondente multiplicador λ_i é positivo ($\lambda_i > 0$). Além disso se $h_i(x_k^A) = 0$, então $h_i(x_{k+1}) = 0$. Então existe $k_0 \in \mathbb{N}$ tal que

$$A(x_k) = A(x^*) \text{ para todo } k \geq k_0.$$

Demonstração: Como o número de restrições é finito, pelo Teorema 3.5 existe $k_2 \in \mathbb{N}$ tal que para $k \geq k_2$

$$A(x_k) = A. \quad (3.3.18)$$

Mas (3.3.18) implica que $A \subset A(x^*)$. Resta mostrar que $A(x^*) \subset A$. Suponhamos que existe $i \in A(x^*)$ tal que

$$h_i(x_k) < 0 \text{ para todo } k \geq k_2. \quad (3.3.19)$$

Pelo Teorema 3.4, para k suficientemente grande, $h_i(x_k^A) = 0$ implicando que $h_i(x_{k+1}) = 0$.

Mas isto contradiz (3.3.19). Portanto,

$$A = A(x^*).$$

□

CAPÍTULO 4

CONCLUSÕES E PERSPECTIVAS

Os caminhos para os nossos resultados assemelham-se aos obtidos em [6], para o mesmo contexto de identificação. Contudo, percebemos que a definição do “*ponto auxiliar*” combinado com a idéia da projeção, mostrado em (3.2.2), levou-nos aos objetivos, com alguma “*economia*” de teoria, explorando apenas os resultados da análise convexa.

A flexibilidade na escolha da função modelo, no algoritmo apresentado em Conn, Gould, Sartenaer e Toint (1993) [6], exige hipóteses adicionais que garantam estacionaridade para um ponto limite tais como convergência assintótica entre os gradientes das funções modelos e gradiente da função objetivo, algumas propriedades sobre as curvaturas dos modelos e outros. Acreditamos que essa diversificação não compromete as propriedades do ponto de Cauchy generalizado (CGP), mas exige dos seus autores, para demonstração dos resultados de identificação, um desenvolvimento teórico capaz de “*compensar*” as dificuldades apresentadas pela definição do ponto. Uma interessante observação, ainda sobre esse artigo, é a necessidade do limite

$$\lim_{\substack{k \in K \\ k \rightarrow \infty}} P_{T(y_k)}(-g_k) = 0$$

para alguma sequência $\{y_k\}$ tal que $y_k \rightarrow x^*$, $k \in K$, $y_k \in A(x^*)$ para k suficientemente grande. O algoritmo RCARB, através do subproblema (3.2.2) permite uma definição natural para y_k que além de satisfazer a necessidade acima, coincide com o ponto auxiliar depois de um número finito de iterações.

Pelo resultado do teorema 3.6 concluímos que o conjunto dos pontos de acumulação de uma sequência gerada pelo algoritmo RCARB que gozam da não-degenerescência primal e dual, possuem o mesmo conjunto de restrições ativas.

As hipóteses de regularidade e não-degenerescência de Dunn, são exigidas em muitos trabalhos (ver [2, 4, 5, 6, 8]). A intenção parece ser universal, ou seja, explorar a teoria de cones como um meio de chegar aos resultados. Não sabemos, ainda, se estas condições são

necessárias, mas em alguns trabalhos como por exemplo, os de Burke e Moré (1988) [2], a ausência de uma delas é substituída por uma restrição de qualificação que garanta os mesmos efeitos daquela hipótese além de garantir que um ponto estacionário é um ponto de Kuhn-Tucker.

Pelos resultados obtidos, fica claro que não precisamos supor a convexidade do conjunto factível, mas sim convexidade local na solução.

Pretendemos dar prosseguimento ao nosso trabalho buscando resultados de identificação das restrições ótimas, no algoritmo RCARB, para regiões não-convexas.

REFERÊNCIAS

- [1] Bazarra, M.S., Shetty, C.M. – “*Nonlinear Programming, Theory and Algorithms*”, John Wiley & Sons, 1979.
- [2] Burke, J.V. and Moré, J.J. – “*On the identification of active constraints*”, SIAM J. Numer. Anal., **25** (1988), pp. 1197–1211.
- [3] Burke, J.V., Moré, J.J., Toraldo, G. – “*Convergence properties of trust region methods for linear and convex constraints*”, Mathematical Programming, **47** (1990), pp. 305–336.
- [4] Burke, J. – “*On the identification of active constraints II: on the nonconvex case*”. SIAM J. Numer. Anal., **27** (1990), pp. 1081–1102.
- [5] Conn, A.R., Gould, N.I.M. and Toint, Ph.L. – “*Global convergence of a class of trust region algorithms for optimization with simple bounds*”, SIAM J. Numer. Anal., **25** (1988), pp. 433–460. Correction, Same Journal, **26** (1989), pp. 764–767.
- [6] Conn, A.R., Gould, N.I.M., Sartenaer, A. and Toint, Ph.L. – “*Global convergence of a class of trust region algorithms for optimization using inexact projections on convex constraints*”, SIAM J. Optimization, **25** (1993), pp. 164–221.
- [7] Dennis, J.E. and Schnabel, R.B. – “*Numerical Methods for Unconstrained Optimization and Nonlinear Equations*”, New Jersey, Prentice-Hall, Englewood Cliffs, 1983.
- [8] Dunn, J.C. – “*On the convergence of projected gradient processes to singular critical points*”, J. Optim. Theory Appl., **56**, pp. 203–216.

- [9] Fletcher, R. – “*Practical Methods of Optimization*”, 2 ed., New York, John Wiley & Sons, 1987.
- [10] Friedlander, A., Martínez, J.M. & Santos, S.A. – “*A new trust region algorithm for bound constrained minimization*”. Relatório de Pesquisa RP 19/92, IMECC – UNICAMP, Campinas, 1992. Por aparecer em Journal of Applied Mathematics and Optimization.
- [11] Friedlander, A. & Martínez, J.M. – “*On the maximization of a concave quadratic function with box constraints*”. SIAM Journal on Optimization, vol. 4, nº 2, (1994) pp. 331–339.
- [12] Luenberger, D.G. – “*Linear and Nonlinear Programming*”, 2 ed., California, Addison-Wesley Publishing Company, 1984.
- [13] Martínez, J.M. & Santos, S.A. – “*A trust region strategy for minimization on arbitrary domains*”. Relatório de Pesquisa RP 63/91, IMECC – UNICAMP, Campinas, 1992. Por aparecer em Mathematical Programming.
- [14] Moré, J.J. – “*Trust regions and projected gradients*”, in M. Iri and K. Yajima, eds. Systems Modelling and Optimization: Proceedings of the 13th IFIP Conference on System Modelling and Optimization, Tokyo, Japan, August 31 – September 4, 1987, Lecture Notes in Control Information Sciences, vol. 113 (Springer, Berlin, 1988) pp. 1–13.
- [15] Moreau, J.J. – “*Décomposition orthogonale d’un espace hilbertien selon deux cônes mutuellement polaires*”. Comptes-Rendus Académie des Sciences **255** (1962), pp. 238–240.
- [16] Rockafellar, R.T. – “*Convex Analysis*”, Princeton University Press, Princeton, NJ, 1970.
- [17] Santos, S.A. – “*Métodos de região de confiança em conjuntos arbitrários e minimização em bolas*”, Dissertação de mestrado, Universidade Estadual de Campinas – UNICAMP, Campinas, 1991.

- [18] Santos, S.A. – “*Regiões de confiança em programação Matemática*”, Dissertação de doutorado, Universidade Estadual de Campinas – UNICAMP, Campinas, 1994.