



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO EM MATEMÁTICA

**Método do Ponto Proximal para Funções Localmente
Lipschitz em Otimização Multiobjetivo**

Edimilson Lopes Dias Júnior

Teresina - 2018

Edimilson Lopes Dias Júnior

Dissertação de Mestrado:

**Método do Ponto Proximal para Funções Localmente Lipschitz
em Otimização Multiobjetivo**

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:

Prof. Dr. João Xavier da Cruz Neto

Co-Orientador:

Prof. Dr. Jurandir de Oliveira Lopes

Teresina - 2018

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

D541m Dias Júnior, Edimilson Lopes.

Método do ponto proximal como aplicação das condições de otimalidade para um problema em otimização multiobjetivo / Edimilson Lopes Dias Júnior. – Teresina, 2018.

58f.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Matemática, 2018.

Orientador: Prof. Dr. João Xavier da Cruz Neto.

Co-orientador: Prof. Dr. Jurandir de Oliveira Lopes.

1. Otimização Matemática. 2. Método do Ponto Proximal. 3. Otimização Multiobjeto. 4. Funções Lipschitz.
I. Título.

CDD 519.3



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Método do Ponto Proximal para Funções Localmente Lipschitz em Otimização Multiobjetivo

EDIMILSON LOPES DIAS JUNIOR

Esta Dissertação foi submetida como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**, outorgado pela Universidade Federal do Piauí.

A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

Dissertação aprovada em 03 de Agosto de 2018.

Banca Examinadora:


Prof. Dr. João Xavier da Cruz Neto - Presidente


Prof. Dr. João Carlos de Oliveira Souza - membro interno

Prof. Dr. Luis Román Lucambio Pérez - membro externo

*Dedico esse trabalho aos meus pais e a minha família,
em especial minha vó Maria José de Jesus Dias(In
memoriam)*

Agradecimentos

Findados dois anos de trabalho, venho aqui agradecer a todas as pessoas que contribuíram de alguma forma para a realização desse sonho. Agradeço primeiramente a Deus por me dar força e coragem para lutar pelos meus sonhos, aos meus pais e familiares pelo amor, apoio e por sempre acreditarem em mim. A minha namorada Evelyn Larissa por seu amor e carinho para comigo durante todo esse tempo (Te amo). Aos amigos do departamento, que tive a honra e o prazer de conhecer: Arilson, Rafaelber, Ronaldo, Rafael (Gordines), José Edilson, Leonardo, Mauro (Professor), Marcos Paulo (M.P.), Luan, Luciano (em memória), Ronnyê, Cícero Nadiel, Kelvin, Juliana, Raul, Ray, Tiago e Hércules. Quero dirigir meus agradecimentos aos professores do departamento de matemática e programa de pós-graduação da UFPI com tive o privilégio de estudar: Antônio Wilson, Barnabé, Carlos Humberto, Franciane, Gleison, José Francisco, Paulo Alexandre e Mário Gomes. Gostaria de fazer um agradecimento especial aos professores João Xavier(orientador), Jurandir(co-Orientador) pela orientação e ajuda nos momentos de dificuldade que ocorreram no decorrer dessa dissertação. Queria também agradecer a Yuri Rafel, que teve papel fundamental na realização desse trabalho. Finalmente agradeço aos professores João Carlos e Luis Román por aceitarem participar de minha banca, e a CAPES e ao CNPq pelo apoio financeiro.

“A educação tem raízes amargas, porém
seus frutos são doces.”

Aristóteles.

Resumo

Nessa dissertação estudamos o problema de otimização multiobjetivo para encontrar os pontos Pareto-críticos de uma função vetorial. Nosso trabalho se baseia em Bento et al. (SIAM J. Optim., (2018),pp 1104-1120), que por sua vez faz uma extensão do Método do Ponto Proximal considerado por Bonnel et al. (SIAM J. Optim., (2005),pp. 953-970) no sentido de que substituímos funções convexas por localmente Lipschitz. Destaca-se como o grande diferencial desse trabalho a não utilização de técnicas de escalarização para resolver o referido problema, para isso utilizamos as condições de otimalidade feitas por Minami (J. Optim. Theory Appl., 41 (1983), pp.451-461) em substituição as condições de primeira ordem para problemas de escalarização.

Palavras-Chave: Método do Ponto Proximal, Otimização Multiobjetivo, Funções Localmente Lipschitz.

Abstract

In this work, we study the multiobjective optimization problem of finding Pareto critical points of a vector-valued function. Our work is based on Bento et al. (SIAM J. Optim., (2018),pp 1104-1120), which in turn is an extension of the Proximal Point Method considered by Bonnafant-Bertrand et al. (SIAM J. Optim., (2005),pp. 953-970) is proposed in the sense of considering locally Lipschitz functions instead of convex functions. It is worth to mention that this work does not use any scalarization technique to solve the problem. Therefore, we use an optimality condition proposed by Minami (J. Optim. Theory Appl., 41(1983), pp.451-461) instead of the first order optimality condition of the scalarized problem.

Keywords: Proximal Point Method, Multiobjective Optimization, Locally Lipschitz Function.

Sumário

Resumo	iv
Abstract	v
1 Preliminares	3
1.1 O espaço \mathbb{R}^n	3
1.2 Elementos de Análise Convexa	7
2 Otimização Multiobjetivo	14
2.1 Ordenação Parcial	14
2.2 Otimização Vetorial	16
3 Método do Ponto Proximal Multiobjetivo	32
3.1 O Algoritmo	32
4 Análise de Convergência	40
4.1 Caso Localmente Lipschitz	40
4.2 Caso Quase Convexo	42
5 Considerações Finais	46
Referências Bibliográficas	47

Introdução

Em otimização de um modo geral, estamos sempre buscando resolver o problema de encontrar quando possível o ponto que minimiza uma dada função f . Tal tarefa pode se tornar bastante dificultosa ou até mesmo impossível uma vez que está intimamente relacionada com as características de f . Em se tratando de otimização vetorial, o problema anteriormente mencionado ganha novas peculiaridades, uma vez que minimizar uma função $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ teoricamente seria minimizar cada uma das m -funções escalares $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$. Devido a dificuldade de resolver um problema dessa natureza Vilfredo Pareto (1848-1923) desenvolveu a ideia de ponto Pareto, segundo a qual \mathbf{x}^* é um ponto Pareto para uma dada função $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ quando não existir $\mathbf{y} \in \mathbb{R}^n$ de modo que $F(\mathbf{y}) \preceq F(\mathbf{x}^*)$, ou seja, em pelo menos uma das funções coordenadas de F vale que $f_i(\mathbf{y}) \geq f_i(\mathbf{x}^*)$ $i \in \{1, \dots, m\}$. Assim, conclui-se que estamos em busca não de um ponto que minimize todas as funções coordenadas de F , mas que seja um ponto de equilíbrio em relação as funções coordenadas. Essa dissertação tem como referência o trabalho de Bento et al. em [2], onde é feita uma extensão do trabalho de Bonnel et al. em [4], diferenciando-se desse em alguns aspectos. No trabalho desenvolvido em [4], a função objetivo é assumidamente convexa e utiliza-se a técnica de escalarização, onde para encontrar o ponto ótimo Pareto de F , encontramos o minimizador para a função escalar $\psi(F(\cdot))$, para alguma função escalar ψ . Esse minimizador possui um correspondente no problema vetorial, onde tal propriedade se verifica desde que F seja convexa. Outros trabalhos que usaram da técnica utilizada em [4] podem ser encontrados em [5, 17]. Por sua vez, o trabalho realizado em [2] não utiliza técnicas de escalarização, fazendo uso das condições de otimalidade propostas por Minami em [16] em substituição as condições de primeira ordem para problemas de escalarização.

Nesse contexto, o trabalho se divide em cinco capítulos, onde no Capítulo 1 apresentamos inicialmente alguns elementos básicos referentes à Análise no espaço \mathbb{R}^n e Análise

Convexa necessários para a compreensão da teoria a ser desenvolvida no transcorrer dos capítulos seguintes. No Capítulo 2 fazemos um breve comentário acerca da ordenação, de um modo especial em \mathbb{R}_+^m , estendemos ainda os conceitos já mencionados anteriormente para funções escalares, agora para funções vetoriais, bem como alguns resultados importantes, entre eles destacamos o resultado de [16] que trata da condição necessária para que um ponto seja um ponto ótimo Pareto para uma função vetorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ em um problema de Otimização Multiobjetivo. No Capítulo 3 introduzimos algumas definições necessárias ao desenvolvimento da teoria e que são utilizadas no transcorrer do capítulo. Além disso, definimos o Algoritmo do Ponto Proximal Multiobjetivo e provamos sua boa definição, bem como alguns resultados. No Capítulo 4 fazemos uma análise da convergência do Algoritmo acerca das hipóteses de F ser localmente Lipschitz e posteriormente quase-convexa. Finalmente no Capítulo 5 fazemos uma análise do trabalho desenvolvido no transcorrer da dissertação e também uma menção a futuros trabalhos a serem desenvolvidos.

Capítulo 1

Preliminares

No decorrer desse capítulo, enunciaremos uma série de definições e resultados que darão suporte a teoria e as demonstrações que serão desenvolvidas posteriormente. Iniciamos falando um pouco sobre o espaço \mathbb{R}^n , bem como alguns tópicos de sua Topologia. Em seguida abordamos alguns resultados de Análise Convexa, os quais serão de fundamental importância durante esse trabalho. As demonstrações desse capítulo serão apenas referenciadas, uma vez que são resultados considerados clássicos e com o qual o leitor certamente está familiarizado.

1.1 O espaço \mathbb{R}^n

Nesta seção faremos um breve comentário sobre o espaço \mathbb{R}^n , e alguns resultados a ele pertinentes. Esperamos que o leitor já tenha uma certa familiaridade com alguns conceitos relacionados a análise, de modo a tornar a leitura dessa seção mais confortável e natural. Os resultados abordados nesta seção são baseados em [13], o qual pode ser consultado pelo leitor caso necessário.

Nesse trabalho utilizaremos os vetores do espaço \mathbb{R}^n como vetor(coluna), de modo que $\mathbf{x}^\top = (x_1, \dots, x_n)$, com cada $x_i \in \mathbb{R}$.

Definição 1. *Um produto interno em \mathbb{R}^n é uma aplicação que faz corresponder a cada par $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ um número real, indicado por $\langle \mathbf{x}, \mathbf{y} \rangle$ de modo que, para quaisquer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$ se tenham :*

$$i) \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle;$$

$$ii) \langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle ;$$

$$iii) \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha \mathbf{y} \rangle ;$$

$$iv) \mathbf{x} \neq 0 \Rightarrow \langle \mathbf{x}, \mathbf{x} \rangle > 0.$$

Um exemplo de produto interno em \mathbb{R}^n é o produto interno canônico, que é dado por:

$$\mathbf{x}^\top \mathbf{y} := \langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \dots + x_n y_n,$$

$$\text{onde } \mathbf{x}^\top = (x_1, \dots, x_n) \text{ e } \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

A menos que seja mencionado o contrário, esse será o produto interno que faremos uso no decorrer dessa dissertação.

Através do produto interno definido anteriormente, podemos definir a norma euclidiana de um vetor $\mathbf{x} \in \mathbb{R}^n$, como: $\|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_n^2}$. O número $\|\mathbf{x}\|$ chama-se a norma euclidiana do vetor \mathbf{x} , e essa goza das seguintes propriedades:

- a) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (desigualdade triangular);
- b) $\|\alpha \mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\| \forall \mathbf{x} \in \mathbb{R}^n \text{ e } \forall \alpha \in \mathbb{R}$;
- c) $\mathbf{x} \neq 0 \Rightarrow \|\mathbf{x}\| > 0 \forall \mathbf{x} \in \mathbb{R}^n$.

Uma vez definida, a norma euclidiana nos trás alguns conceitos a ela relacionados, tais como bola aberta, bola fechada, esfera e distância. Os quais formalizaremos a seguir:

Definição 2. *Seja $\mathbf{a} \in \mathbb{R}^n$ e $r > 0$ um número real, definimos:*

- a) **Bola Aberta:** *Conjunto dos pontos $\mathbf{x} \in \mathbb{R}^n$ tais que $\|\mathbf{x} - \mathbf{a}\| < r$;*
- b) **Bola Fechada:** *Conjunto dos pontos $\mathbf{x} \in \mathbb{R}^n$ tais que $\|\mathbf{x} - \mathbf{a}\| \leq r$;*
- c) **Esfera:** *Conjunto dos pontos $\mathbf{x} \in \mathbb{R}^n$ tais que $\|\mathbf{x} - \mathbf{a}\| = r$.*

Usaremos as notações $B_r(\mathbf{a})$, $B_r[\mathbf{a}]$ e $S_r(\mathbf{a})$ para representar respectivamente a Bola Aberta, Bola Fechada e Esfera de centro \mathbf{a} e raio r .

Definição 3. *Denotamos a função distância $d : \mathbb{R}^n \rightarrow \mathbb{R}$ de um ponto $\mathbf{x} \in \mathbb{R}^n$ para um certo conjunto $C \subset \mathbb{R}^n$ como sendo*

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

A seguir falamos um pouco a respeito de seqüências no espaço \mathbb{R}^n , bem como alguns resultados a elas relacionados.

Definição 4. Uma seqüência em \mathbb{R}^n é uma aplicação $\chi : \mathbb{N} \rightarrow \mathbb{R}^n$, onde o valor que essa aplicação assume no número k é indicado como χ^k e chama-se o k -ésimo termo da seqüência. Usaremos as notações $\{\chi^k\}$ ou $\{\chi^k\}_{k \in \mathbb{N}}$ para indicar uma seqüência cujo o k -ésimo termo é $\chi^k \in \mathbb{R}^n$.

Exemplo 1. $\{\chi^k\} \in \mathbb{R}^2$ onde, $(\chi^k)^\top = \left((-1)^{k+1} + \frac{1}{k+1}, \frac{1}{k^3} \right)$

Definição 5. Definimos uma subsequência de $\{\chi^k\}$ como sendo uma restrição da seqüência a um subconjunto infinito $\mathbb{N}^1 = \{k_1 < k_2 < \dots < k_j < \dots\} \subset \mathbb{N}$ e usaremos as seguintes notações em sua representação: $\{\chi^k\}_{k \in \mathbb{N}^1}$ ou ainda $\{\chi^{k_i}\}_{i \in \mathbb{N}}$.

Exemplo 2. Do Exemplo 1 temos que se olharmos para os termos cujo índice é par, obtemos a subsequência $(\chi^{2k})^\top = \left(-1 + \frac{1}{2k+1}, \frac{1}{(2k)^3} \right)$. Analogamente, ao tomarmos os elementos cujo o índice é ímpar obtemos a subsequência $(\chi^{2k+1})^\top = \left(1 + \frac{1}{2(k+1)}, \frac{1}{(k+1)^3} \right)$.

Definição 6. Dizemos que uma seqüência $\{\chi^k\}$ é limitada quando existe uma constante real $M > 0$ tal que $\|\chi^k\| \leq M \forall k \in \mathbb{N}$.

Exemplo 3. Dada a seqüência $\{\chi^k\} \in \mathbb{R}^2$ definida por $(\chi^k)^\top = \left((-1)^k, \frac{1}{k} \right)$. Como $\|\chi^k\| \leq 2$, conclui-se que tal seqüência é limitada.

Definição 7. Diz-se que o ponto $\mathbf{a} \in \mathbb{R}^n$ é limite da seqüência de pontos $\{\chi^k\} \in \mathbb{R}^n$ quando, para todo $\varepsilon > 0$ dado, existe $k_0 \in \mathbb{N}$ tal que $k > k_0 \Rightarrow \|\chi^k - \mathbf{a}\| < \varepsilon$. Neste caso, diz-se também que $\{\chi^k\}$ converge para \mathbf{a} ou tende para \mathbf{a} e escreve-se $\lim_{k \rightarrow \infty} \chi^k = \mathbf{a}$ ou simplesmente $\chi^k \rightarrow \mathbf{a}$.

Quando o limite acima mencionado existe, dizemos que a seqüência $\{\chi^k\}$ é convergente. Caso contrário, dizemos que a seqüência é divergente.

Definição 8. Dizemos que o ponto \mathbf{a} é um valor de aderência para a seqüência $\{\chi^k\}$ se existe alguma subsequência de $\{\chi^k\}$ que converge para \mathbf{a} .

Observação 1. A seqüência dada no Exemplo 3 não é convergente, uma vez que mesmo sendo limitada, duas de suas subsequências possuem limites distintos e conseqüentemente dois valores de aderência. A saber $\{(1, 0)^\top (-1, 0)^\top\}$.

Um fato a ser observado é que uma sequência $\{\mathbf{x}^k\}$ em \mathbb{R}^n equivale a n seqüências de números reais. Assim para cada $k \in \mathbb{N}$ podemos escrever $(\mathbf{x}^k)^\top = (x_{k1}, \dots, x_{kn})$ em que cada seqüência x_{ki} é uma seqüência de números reais. Usando desse raciocínio podemos enunciar o seguinte teorema.

Teorema 1. *Uma seqüência $\{\mathbf{x}^k\} \in \mathbb{R}^n$ converge para um ponto $\mathbf{a}^\top = (a_1, \dots, a_n)$ se, e somente se, para cada $i = 1, \dots, n$ tem-se $\lim_{k \rightarrow \infty} x_{ki} = a_i$, ou seja, cada coordenada de \mathbf{x}^k converge para a coordenada correspondente de \mathbf{a} .*

Demonstração. Para a demonstração ver [13] pág. 16. □

Teorema 2. *(Bolzano-Weierstrass) Toda seqüência limitada em \mathbb{R}^n possui uma subseqüência convergente.*

Demonstração. Para a demonstração ver [13] pág. 17. □

Definição 9. *Um ponto $\mathbf{a} \in \mathbb{R}^n$ é um ponto de acumulação para o conjunto $X \subset \mathbb{R}^n$ se existe uma seqüência de pontos $\{\mathbf{x}^k\} \in X$ com $\lim_{k \rightarrow \infty} \mathbf{x}^k = \mathbf{a}$ e $\mathbf{x}^k \neq \mathbf{a} \forall k \in \mathbb{N}$.*

Uma vez abordados os principais conceitos relacionados a seqüências em \mathbb{R}^n , apresentamos a definição de continuidade para uma função definida em um subconjunto do espaço \mathbb{R}^n . Em seguida traremos um resultado que relaciona continuidade e seqüências.

Definição 10. *Seja $X \subset \mathbb{R}^n$ um conjunto não-vazio, dizemos que a função $f : X \rightarrow \mathbb{R}^m$ é contínua no ponto \mathbf{a} , quando para qualquer $\varepsilon > 0$ dado, se pode obter $\delta > 0$ de modo que todo ponto de $\mathbf{x} \in X$ cuja distância para o ponto \mathbf{a} seja menor que δ seja transformado por f num ponto $f(\mathbf{x})$ cuja distância para o ponto $f(\mathbf{a})$ é menor que ε .*

Simbolicamente: Dado $\varepsilon > 0$, $\exists \delta > 0$ tal que $\|\mathbf{x} - \mathbf{a}\|_{\mathbb{R}^n} < \delta \Rightarrow \|f(\mathbf{x}) - f(\mathbf{a})\|_{\mathbb{R}^m} < \varepsilon$.

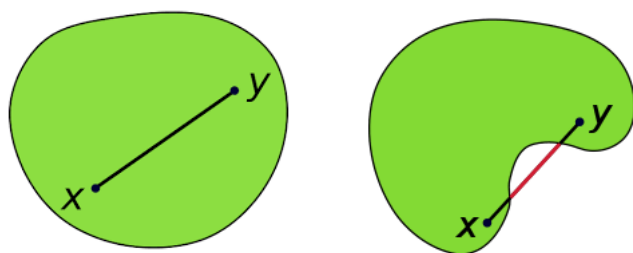
Teorema 3. *Uma aplicação $f : X \rightarrow \mathbb{R}^m$ com $X \subset \mathbb{R}^n$ é contínua no ponto $\mathbf{a} \in X$ se, e somente se, para cada seqüência de pontos $\{\mathbf{x}^k\} \in X$ com $\lim_{k \rightarrow \infty} \mathbf{x}^k = \mathbf{a}$, tem-se $\lim_{k \rightarrow \infty} f(\mathbf{x}^k) = f(\mathbf{a})$.*

Demonstração. Para a demonstração ver [13] pág. 27. □

1.2 Elementos de Análise Convexa

Nesta seção abordaremos algumas definições e resultados relacionados à área de Análise Convexa que faremos uso no decorrer desse trabalho. Grande parte dos resultados aqui mencionados são baseados em [12] e em alguns artigos, cujas referências poderão ser encontradas nas Referências Bibliográficas.

Definição 11. Um conjunto $D \subset \mathbb{R}^n$ é chamado convexo se para quaisquer $x, y \in D$ e $\alpha \in [0, 1]$, tem-se que $\alpha x + (1 - \alpha)y \in D$.



Nas imagens acima, temos respectivamente um conjunto convexo e um não-convexo. O ponto $\alpha x + (1 - \alpha)y$, onde $\alpha \in [0, 1]$ é uma combinação convexa de x, y (com parâmetro α). O espaço \mathbb{R}^n , e um conjunto que contém um único ponto são trivialmente convexos.

Exemplo 4. A bola $B_r(\mathbf{a})$ em \mathbb{R}^n é um conjunto convexo.

Com efeito, seja $B = B_r(\mathbf{a})$ a bola aberta de centro \mathbf{a} e raio $r > 0$. Se $x, y \in B$ então $\|x - \mathbf{a}\| < r$ e $\|y - \mathbf{a}\| < r$. Para qualquer $t \in [0, 1]$, temos:

$$\begin{aligned} \|(1-t)x + ty - \mathbf{a}\| &= \|(1-t)x + ty - t\mathbf{a} + t\mathbf{a} - \mathbf{a}\| \\ &= \|(1-t)(x - \mathbf{a}) + t(y - \mathbf{a})\| \\ &\leq (1-t)\|x - \mathbf{a}\| + t\|y - \mathbf{a}\| < r. \end{aligned}$$

Exemplo 5. A esfera $S_r(0)$ não é um conjunto convexo.

Com efeito, denotando $S = S_r(0)$ devemos provar que para quaisquer $x, y \in S$, com $x \neq y$ o ponto $(1-t)x + ty, \notin S$ com $t \in (0, 1)$. Suponha por absurdo que $(1-t)x + ty \in S$,

daí tem-se que

$$\begin{aligned}
 \|(1-t)x + ty\| = r &\Leftrightarrow \|(1-t)x + ty\|^2 = r^2 \\
 &\Leftrightarrow \langle (1-t)x + ty, (1-t)x + ty \rangle = r^2 \\
 &\Leftrightarrow (1-t)^2\|x\|^2 + 2t(1-t)\langle x, y \rangle + t^2\|y\|^2 = r^2 \\
 &\Leftrightarrow (1-2t+t^2)r^2 + 2t(1-t)\langle x, y \rangle + t^2r^2 = r^2 \\
 &\Leftrightarrow r^2 - 2tr^2 + t^2r^2 + 2t(1-t)\langle x, y \rangle + t^2r^2 = r^2 \\
 &\Leftrightarrow -2tr^2 + 2t^2r^2 + 2t(1-t)\langle x, y \rangle = 0 \\
 &\Leftrightarrow -tr^2 + t^2r^2 + t(1-t)\langle x, y \rangle = 0 \\
 &\Leftrightarrow -r^2t(1-t) + t(1-t)\langle x, y \rangle = 0 \\
 &\Leftrightarrow t(1-t)(\langle x, y \rangle - r^2) = 0.
 \end{aligned}$$

Uma vez que $t(1-t) \neq 0$, tem-se que $\langle x, y \rangle - r^2 = 0$ e conseqüentemente $\langle x, y \rangle = r^2$. Tal igualdade ocorre somente se $y = x$, e isso contradiz nossa hipótese inicial.

A seguir apresentamos algumas propriedades a cerca de conjuntos convexos. As provas de tais propriedades serão apenas referenciadas.

Proposição 1. *Sejam $D_i \subset \mathbb{R}^n$, $i \in I$, conjuntos convexos, onde I é um conjunto qualquer não necessariamente finito. Então a interseção $D = \bigcap_{i \in I} D_i$ também é um conjunto convexo.*

Demonstração. Para uma demonstração ver [12] pág. 82. □

Proposição 2. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Então $\text{cl}(D)$ e $\text{int}(D)$, onde $\text{cl}(D)$ e $\text{int}(D)$ representam respectivamente o fecho e o interior topológico de D , são conjuntos convexos.*

Demonstração. Para uma demonstração ver [12] pág. 83. □

Definição 12. *Se $D \subset \mathbb{R}^n$ é um conjunto convexo, diz-se que a função $f : D \rightarrow \mathbb{R}$ é convexa em D quando para quaisquer $x, y \in D$ e $\alpha \in [0, 1]$, tem-se que*

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y).$$

A função diz-se estritamente convexa quando a desigualdade acima é estrita para todos $x \neq y$ e $\alpha \in (0, 1)$.

Exemplo 6. $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \|\mathbf{x}\|^2$ é convexa.

De fato, queremos mostrar para quaisquer $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e qualquer $t \in [0, 1]$ tem-se

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}),$$

ou seja,

$$\|t\mathbf{x} + (1-t)\mathbf{y}\|^2 \leq t\|\mathbf{x}\|^2 + (1-t)\|\mathbf{y}\|^2.$$

Para tal, ao desenvolvermos o quadrado da soma temos :

$$\|t\mathbf{x} + (1-t)\mathbf{y}\|^2 = t^2\|\mathbf{x}\|^2 + 2t(1-t)\langle \mathbf{x}, \mathbf{y} \rangle + (1-t)^2\|\mathbf{y}\|^2.$$

Lembremos que $2\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, e usando esse fato na equação anterior, temos

$$\begin{aligned} t^2\|\mathbf{x}\|^2 + 2t(1-t)\langle \mathbf{x}, \mathbf{y} \rangle + (1-t)^2\|\mathbf{y}\|^2 &\leq t^2\|\mathbf{x}\|^2 + t(1-t)(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) + (1-t)^2\|\mathbf{y}\|^2 \\ &= t^2\|\mathbf{x}\|^2 + t\|\mathbf{x}\|^2 + t\|\mathbf{y}\|^2 - t^2\|\mathbf{x}\|^2 - t^2\|\mathbf{y}\|^2 + \|\mathbf{y}\|^2 - 2t\|\mathbf{y}\|^2 + t^2\|\mathbf{y}\|^2. \end{aligned}$$

Fazendo as devidas simplificações e cancelamentos, temos que

$$\|t\mathbf{x} + (1-t)\mathbf{y}\|^2 \leq t\|\mathbf{x}\|^2 + (1-t)\|\mathbf{y}\|^2.$$

Definição 13. O epígrafo da função $f : D \rightarrow \mathbb{R}$ é o conjunto

$$E_f = \{(\mathbf{x}, c) \in D \mid f(\mathbf{x}) \leq c\}.$$

A relação entre a convexidade de f e seu epígrafo é dada pelo seguinte teorema.

Teorema 4. Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : D \rightarrow \mathbb{R}$ é convexa em D se, e somente se, o epígrafo de f é um conjunto convexo em $\mathbb{R}^n \times \mathbb{R}$.

Demonstração. Para uma demonstração ver [12] pág. 78. □

Definição 14. Dizemos que $\bar{v} \in [-\infty, +\infty)$ definido por

$$\bar{v} = \inf_{\mathbf{x} \in D} f(\mathbf{x}),$$

é o valor ótimo do problema de minimizar f restrita ao conjunto D .

Definição 15. Seja $D \subset \mathbb{R}^n$ um conjunto não-vazio e $f : D \rightarrow \mathbb{R}$. Dizemos que um ponto $\bar{\mathbf{x}} \in D$ é:

a) **Minimizador Global** de f em D se

$$f(\bar{x}) \leq f(x) \quad \forall x \in D;$$

b) **Minimizador Local** de f em D se existe uma vizinhança V de \bar{x} tal que

$$f(\bar{x}) \leq f(x) \quad \forall x \in V \cap D.$$

Observação 2. Se \bar{x} é minimizador global (respectivamente local) para a função f em D , então $f(\bar{x})$ é chamado **valor ótimo global (respectivamente local)**.

Perceba que o problema de minimizar uma função pode ter várias soluções (globais), isto é, vários minimizadores (globais). Porém cada problema apresenta apenas um único valor ótimo.

Dizemos que

$$\min f(x) \text{ sujeito a } x \in D$$

é um problema de minimização convexo quando $D \subset \mathbb{R}^n$ é um conjunto convexo e a função $f : D \rightarrow \mathbb{R}$ é convexa em D .

A seguir apresentamos um teorema de grande importância em análise convexa, uma vez que relaciona mínimos globais com a convexidade de uma função.

Teorema 5. Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $f : D \rightarrow \mathbb{R}$ uma função convexa em D . Então todo minimizador local de f em D é um minimizador global. Além disso, o conjunto dos minimizadores é convexo. Se f é estritamente convexa, não pode haver mais de um minimizador.

Demonstração. Para uma demonstração ver [12] pág. 79. □

Definição 16. Dados $x^i \in \mathbb{R}^n$, $\alpha_i \in [0, 1]$, $i = 1, \dots, p$, tais que $\sum_{i=1}^p \alpha_i = 1$, o ponto $\sum_{i=1}^p \alpha_i x^i$ chama-se combinação convexa de pontos $x^i \in \mathbb{R}^n$ com parâmetros α_i , $i = 1, \dots, p$.

A seguir apresentamos um resultado que nos dá uma nova caracterização para conjuntos convexos.

Teorema 6. Um conjunto $D \subset \mathbb{R}^n$ é convexo se, e somente se, para quaisquer $p \in \mathbb{N}$, $x^i \in D$ e $\alpha_i \in [0, 1]$, $i = 1, \dots, p$, tais que $\sum_{i=1}^p \alpha_i = 1$, a combinação convexa $\sum_{i=1}^p \alpha_i x^i \in D$.

Demonstração. Para uma demonstração ver [12] pág. 85. □

Definição 17. *Seja $D \subset \mathbb{R}^n$ um conjunto qualquer. O fecho convexo de D , denotado por $\text{conv}D$, é o menor conjunto convexo em \mathbb{R}^n que contém D (ou, equivalentemente, a interseção de todos os conjuntos convexos em \mathbb{R}^n que contêm D).*

Observação 3. *Note que $\text{conv}D$ é um conjunto convexo para qualquer $D \subset \mathbb{R}^n$, e se, D é convexo então $\text{conv}D = D$.*

Proposição 3. *Seja $D \subset \mathbb{R}^n$ um conjunto qualquer. O fecho convexo de D é o conjunto de todas as combinações convexas de pontos de D .*

Demonstração. Para uma demonstração ver [12] pág. 91. □

Definição 18. *Um conjunto $K \subset \mathbb{R}^n$ chama-se cone quando*

$$d \in K \Rightarrow td \in K \quad \forall t \in \mathbb{R}_+.$$

Perceba que se $K \neq \emptyset$, necessariamente $0 \in K$. Como exemplo de cone podemos citar o espaço

$$\mathbb{R}_+^m := \{x \in \mathbb{R}^m \mid x_i \geq 0, i = \{1, 2, \dots, m\}\}.$$

Definição 19. *O Cone dual de um cone $K \subset \mathbb{R}^n$ é definido por*

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

A diante apresentamos os Cones Normal e Tangente, mas para apresentar tal definição, necessitamos antes de alguns elementos, os quais introduziremos a seguir.

Definição 20. *Uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, é localmente Lipschitz em um ponto $x \in D$ se existe uma vizinhança V deste ponto e um número real $L > 0$ tal que:*

$$|f(y) - f(x)| \leq L\|y - x\|, \quad \forall x, y \in V.$$

Uma função f é localmente Lipschitz, quando o for em todos os pontos de seu domínio.

Exemplo 7. *Seja $E \subset \mathbb{R}^n$. A função $f(x) = d(x, E)$ é localmente Lipschitz.*

Com efeito, temos que

$$\inf_{z \in E} \|x - z\| \leq \inf_{z \in E} \{\|x - y\| + \|y - z\|\} = \|x - y\| + \inf_{z \in E} \|y - z\|,$$

ou seja,

$$d(x, E) \leq \|x - y\| + d(y, E).$$

Ou ainda,

$$d(x, E) + d(y, E) \leq \|x - y\|. \quad (1.1)$$

Analogamente se prova que:

$$d(y, E) + d(x, E) \leq \|x - y\|.$$

Ou seja,

$$-\|x - y\| \leq d(x, E) + d(y, E) \quad (1.2)$$

De 1.1 e 1.2 segue que

$$\|d(x, E) - d(y, E)\| \leq \|x - y\|, \forall x, y \in E.$$

Definição 21. *Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função localmente Lipschitz e seja $v \in \mathbb{R}^n$. A derivada direcional de Clarke da função f no ponto x na direção de v , a qual é denotada por $f^\circ(x; v)$ é definida por:*

$$f^\circ(x; v) := \lim_{\delta \downarrow 0} \sup_{\|h\| + t < \delta} \frac{f(x + h + tv) - f(x + h)}{t}, \quad t > 0.$$

Definição 22. *Seja $d : \mathbb{R}^n \rightarrow \mathbb{R}$ a função distância em \mathbb{R}^n . Definimos o Cone Tangente em relação ao conjunto Ω no ponto $x \in \Omega$, o qual denotamos por $T_\Omega(x)$, como:*

$$T_\Omega(x) := \{v \in \mathbb{R}^n : d_\Omega^\circ(x, v) = 0\}$$

Definição 23. *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in D$. O cone normal no ponto \bar{x} em relação ao conjunto D , o qual é denotado por $\mathcal{N}_D(\bar{x})$, é dado por*

$$\mathcal{N}_D(\bar{x}) = \{d \in \mathbb{R}^n | \langle d, x - \bar{x} \rangle \leq 0 \quad \forall x \in D\}.$$

O resultado a seguir mostra que no caso convexo, o dual do Cone Tangente coincide com o Cone Normal definido acima.

Teorema 7. *Sejam $D \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in D$. Então*

$$(\mathbb{T}_D(\bar{x}))^* = \mathcal{N}_D(\bar{x}).$$

Demonstração. Para uma demonstração ver [12] pág. 65. □

Assim, podemos utilizar qualquer uma das duas definições para o Cone Normal, utilizando-se sempre daquela que for mais cômoda.

A seguir apresentaremos um conceito de fundamental importância na análise convexa, o qual faremos uso posteriormente.

Definição 24. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função localmente Lipschitz. Definimos o subdiferencial de Clarke para f no ponto $x \in \mathbb{R}^n$ como $\partial f(x) = \{y \in \mathbb{R}^n \mid \langle y, d \rangle \leq f^\circ(x; d) \forall d \in \mathbb{R}^n\}$.*

Observação 4. *Vale ressaltar que quando f é uma função convexa, o subdiferencial de Clarke coincide com o subdiferencial usual da análise convexa.*

Capítulo 2

Otimização Multiobjetivo

No decorrer desse capítulo trazemos algumas definições e resultados relacionados à funções vetoriais, as quais denotaremos por $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Tais resultados darão suporte para a teoria a ser desenvolvida nos capítulos posteriores. Caso o leitor sinta necessidade poderá consultar O livro texto em [15]. Iniciaremos o presente capítulo falando um pouco sobre como ocorre a ordenação no cone \mathbb{R}_+^n . Em seguida, elencaremos uma série de lemas e proposições a serem usados na demonstração das condições de otimalidade para um problema de otimização multiobjetivo.

2.1 Ordenação Parcial

Recordemos inicialmente que o conjunto \mathbb{R} é totalmente ordenado, o que o diferencia dos demais espaços vetoriais. Uma vez que \mathbb{R} é bem ordenado, temos que dados quaisquer x e $y \in \mathbb{R}$ podemos determinar uma relação de ordem entre os mesmos, utilizando para tal os sinais \leq ou \geq bem como suas representações $<$ ou $>$ que são utilizadas quando x e y são necessariamente distintos. Em espaços onde não existe uma boa ordenação como a apresentada pelo conjunto dos números reais utilizamos a ideia de ordenação parcial, a qual definiremos a seguir.

Definição 25. *Um conjunto A é dito totalmente ordenado com relação de ordem (\preceq), se dados quaisquer elementos x, y e $z \in A$ sempre vale $x \preceq y$ ou $y \preceq x$ e além disso as*

seguintes propriedades são satisfeitas:

$$x \preceq x \quad (\text{reflexividade})$$

$$x \preceq y, y \preceq z \Rightarrow x \preceq z \quad (\text{transitividade})$$

$$x \preceq y, y \preceq x \Rightarrow x = y \quad (\text{antisimetria}).$$

Temos ainda a definição de ordenação parcial, que por sua vez está descrita abaixo.

Definição 26. Dizemos que A é parcialmente ordenado quando valem as propriedades acima, mas A não é totalmente ordenado. Em outras palavras, não é verdade que dados quaisquer $x, y \in A$ é possível estabelecer entre eles uma relação da forma $x \preceq y$ ou $x \succeq y$.

No espaço \mathbb{R}^n a ordenação é dada por:

$$x \preceq y \Rightarrow \{x_i \leq y_i, i = 1, 2, \dots, n\}$$

$$x \prec y \Rightarrow \{x_i < y_i, i = 1, 2, \dots, n\}$$

$$x = y \Rightarrow \{x_i = y_i, i = 1, 2, \dots, n\}.$$

O operador \neq é dado por:

$$x \neq y \Rightarrow \{x_i \neq y_i, \text{ para algum } i = 1, 2, \dots, n\}.$$

Exemplo 8. Considere os vetores $x, y, z \in \mathbb{R}^3$. Temos:

$$x^\top = (1, 3, 5), y^\top = (3, 5, 7), z^\top = (2, 1, 4)$$

$$x \prec y,$$

$$z \prec y.$$

Porém, os vetores z e x não podem ser comparados segundo a ordem utilizada em \mathbb{R}^n .

Dessa forma, temos que o espaço \mathbb{R}^n com a ordem acima é parcialmente ordenado, porém não é totalmente ordenado, visto que no Exemplo 8 os pontos z e x não puderam ser comparados.

Exemplo 9. Em \mathbb{R}_+^m temos que $x \preceq_{\mathbb{R}^m} y \Leftrightarrow y - x \in \mathbb{R}_+^m$, ou ainda, $y_i \geq x_i \forall i \in \{1, \dots, m\}$.

2.2 Otimização Vetorial

Definição 27. Dado um conjunto não-vazio $\Omega \subset \mathbb{R}^n$ e $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função vetorial, onde $F^\top = (f_1, f_2, \dots, f_m)$. Um ponto $x^* \in \Omega$ é chamado um ponto ótimo Pareto (ou ponto Pareto) de F em Ω se não existe $x \in \Omega$ tal que $F(x) \preceq F(x^*)$ e $F(x) \neq F(x^*)$. Dizemos ainda que, um ponto $x^* \in \Omega$ como um ponto ótimo Pareto fraco (ou Pareto fraco) de F em Ω se não existe $x \in \Omega$ tal que $F(x) \prec F(x^*)$.

Observemos da definição acima que para todo $x \in \Omega$ existe pelo menos um índice $j(x) = j \in \{1, \dots, m\}$ tal que $F_j(x) - F_j(x^*) \geq 0$.

Utilizaremos (VP) para denotar o problema de encontrar o ponto ótimo Pareto fraco para a função F , e escreveremos

$$\min_w \{F(x) : x \in \Omega\}. \quad (2.1)$$

Definição 28. O conjunto de todos os pontos Pareto fracos de F em Ω é denotado por $\arg \min_w \{F(x) : x \in \Omega\}$.

Exemplo 10. Seja $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 + 2 \\ y^2 - 2y \end{pmatrix}$, então o ponto $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ é um ótimo Pareto.

De fato, caso contrário existiria $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$;

$$F \begin{pmatrix} x \\ y \end{pmatrix} \prec \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ e } F \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Logo, $x^2 + 2 < 2$ e $y^2 - 2y < -1$. Isto implicaria em

$$x^2 + 2 < 2 \Rightarrow x^2 < 0$$

e

$$y^2 - 2y < -1 \Rightarrow y^2 - 2y + 1 < 0,$$

ou seja, um absurdo.

No que diz respeito a funções vetoriais, recordemos algumas propriedades básicas.

Definição 29. Para uma função vetorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ temos que:

i) F é \mathbb{R}_+^m -convexa se, para cada $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ valem as seguintes propriedades:

$$F((1-t)\mathbf{x} + t\mathbf{y}) \preceq (1-t)F(\mathbf{x}) + tF(\mathbf{y}), \quad t \in [0, 1].$$

ii) F é \mathbb{R}_+^m -quase-convexa se, para cada $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, ocorrer :

$$F((1-t)\mathbf{x} + t\mathbf{y}) \preceq \max\{F(\mathbf{x}), F(\mathbf{y})\} \quad \forall t \in [0, 1],$$

$$\text{onde } \max\{F(\mathbf{x}), F(\mathbf{y})\} = \begin{pmatrix} \max\{f_1(\mathbf{x}), f_1(\mathbf{y})\} \\ \vdots \\ \max\{f_m(\mathbf{x}), f_m(\mathbf{y})\} \end{pmatrix}.$$

Definição 30. Seja $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, onde $F(\mathbf{x}) = \begin{pmatrix} f_1(\cdot) \\ \vdots \\ f_m(\cdot) \end{pmatrix}$. Dizemos que F é localmente Lipschitz com constante L se cada $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ é localmente Lipschitz com constante L_i com $L = \max\{L_1, \dots, L_m\}$.

A partir de agora, a menos que se mencione o contrário, F será uma função localmente Lipschitz.

Definição 31. Seja $\Omega \subset \mathbb{R}^n$ um conjunto fechado e não-vazio. Dizemos que um ponto \mathbf{x} é crítico Pareto-Clarke (ponto Pareto crítico) de F em Ω se, para qualquer $\mathbf{v} \in T_\Omega(\mathbf{x})$, existe pelo menos um $i \in I = \{1, \dots, m\}$ tal que

$$f_i^\circ(\mathbf{x}; \mathbf{v}) \geq 0. \tag{2.2}$$

Observação 5. O conjunto dos pontos Pareto-crítico das funções vetoriais

$$F(\cdot) \text{ e } \mathbf{e}^{F(\cdot)} := \begin{pmatrix} \mathbf{e}^{f_1(\cdot)} \\ \vdots \\ \mathbf{e}^{f_m(\cdot)} \end{pmatrix}$$

são iguais. Com efeito, seja \mathbf{x}_0 um ponto Pareto-crítico para F então para qualquer vetor $\mathbf{v} \in T_\Omega(\mathbf{x})$ existe um índice $i \in \{1, \dots, m\}$ tal que $f_i^\circ(\mathbf{x}_0; \mathbf{v}) \geq 0$. Aplicando a função exponencial temos que $\mathbf{e}^{f_i^\circ(\mathbf{x}_0; \mathbf{v})} \geq 0$ e conseqüentemente \mathbf{x}_0 é um ponto Pareto crítico para $\mathbf{e}^{F(\cdot)}$. Agora chame $\mathbf{G}(\mathbf{x}) = \mathbf{e}^{F(\mathbf{x})}$ e para um ponto \mathbf{x}^* Pareto-crítico para \mathbf{G} temos que $\mathbf{G}_i^\circ(\mathbf{x}^*; \mathbf{v}) \geq 0$ e conseqüentemente $\mathbf{e}^{F(\mathbf{x}^*)} \cdot f_i^\circ(\mathbf{x}^*; \mathbf{v}) \geq 0$ e assim temos também que \mathbf{x}^* é Pareto-crítico para F .

Definição 32. *O subdiferencial de Clarke para a função F em $x \in \mathbb{R}^n$, que será denotado por $\partial F(x)$, é definido como*

$$\partial F(x) := \{U \in \mathbb{R}^{m \times n} \mid U d \preceq F^\circ(x; d), \forall d \in \mathbb{R}^n\},$$

onde $F^\circ(x; d) := \begin{pmatrix} f_1^\circ(x; d) \\ \vdots \\ f_m^\circ(x; d) \end{pmatrix}$

Proposição 4. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função localmente Lipschitz. Então vale $f^\circ(x; v) = \max\{\langle \zeta, v \rangle : \zeta \in \partial f(x)\}$.*

Demonstração. Para uma demonstração ver [7] pág.248-249. □

Observação 6. *Note que se $m = 1$ na definição de Pareto crítico, recuperamos a definição clássica para pontos críticos para funções localmente Lipschitz, que nos diz que x é um ponto crítico para a função f , quando $0 \in \partial f(x)$. Assim, temos que combinando a definição 31 com a Proposição 4 obtem-se uma definição alternativa para um ponto $x \in \mathbb{R}^n$ ser Pareto-crítico: Um ponto $x \in \mathbb{R}^n$ é Pareto-crítico para a função F em Ω se, para qualquer $v \in T_\Omega(x)$, existe $i \in \{1, \dots, m\}$ e $\zeta \in \partial f_i(x)$ tal que $\langle \zeta, v \rangle \geq 0$. Desse modo, temos que se x não é um ponto Pareto-crítico para F em Ω , existe um $v \in T_\Omega(x)$ de modo que*

$$Uv \prec 0, \forall U \in \partial F(x).$$

A seguir apresentamos um resultado que nos dá uma condição necessária para que um ponto seja um ponto Pareto-crítico para uma função vetorial.

Lema 1. *Seja $w \in \mathbb{R}_+^m \setminus \{0\}$ e assuma que Ω é um conjunto não-vazio e fechado. Se $-U^T w \in \mathcal{N}_\Omega(x)$, para algum $U \in \partial F(x)$, então x é um ponto Pareto-crítico para F .*

Demonstração. Seja $x \in \Omega$ tal que $-U^T w \in \mathcal{N}_\Omega(x)$ e suponhamos por contradição que x não é um ponto Pareto-crítico para F . Da Observação 6 temos que existe $v \in \Omega$ tal que

$$Uv \prec 0. \forall U \in \partial F(x)$$

Uma vez que $w \in \mathbb{R}_+^m \setminus \{0\}$, temos que $\langle w, Uv \rangle < 0$. Isto contradiz o fato de que $-U^T w \in \mathcal{N}_\Omega(x)$, uma vez que

$$\begin{aligned} 0 > \langle w, Uv \rangle &= \langle U^T w, v \rangle \\ &= -\langle -U^T w, v \rangle \geq 0. \end{aligned}$$

Essa contradição finaliza a prova. □

A seguir apresentamos uma série de resultados que serão utilizados afim de demonstrar o principal resultado deste capítulo.

Lema 2. *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}$, $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I = \{1, \dots, m\}$ funções localmente Lipschitz, onde I representa um conjunto não-vazio e finito de índices. Então seguem algumas propriedades:*

- i) $\partial f(x)$ é um subconjunto não-vazio, convexo e compacto;
- ii) $\partial(f + g)(x) \subset \partial f(x) + \partial g(x)$;
- iii) $\partial(\lambda f)(x) = \lambda \partial f(x)$, para cada $\lambda \geq 0$;
- iv) Se x é uma ponto de mínimo local para f , então $0 \in \partial f(x)$. ;
- v) A função

$$m(x) = \max\{f_i(x) : i \in I = \{1, \dots, m\}\}$$

é localmente Lipschitz e $\partial m(x)$ está contido no fecho convexo de $\{\partial f_i(x) : i \in M(x)\}$, onde $M(x)$ representa o conjunto de índices em I no qual o máximo de $m(x) = \max\{f_i(x) : i \in I = \{1, \dots, m\}\}$ é atingido;

vi) Se um ponto $x \in \mathbb{R}^n$ minimiza f em um subconjunto não-vazio $C \subset \mathbb{R}^n$; então existe uma vizinhança V de x na qual x minimiza $f + Ld_C$ para qualquer constante de Lipschitz local L de f , onde

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

Demonstração. Para uma demosntração ver [6] □

Sejam $F : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função semi-continua inferiormente, limitada inferiormente e própria. Para cada $\varepsilon > 0$ existe algum ponto $u \in \mathbb{R}^n$ tal que

$$\inf F \leq F(u) \leq \inf F + \varepsilon. \tag{2.3}$$

Dado $\alpha > 0$, definimos em $\mathbb{R}^n \times \mathbb{R}$ a ordenação \prec que estabelece

$$\begin{pmatrix} v^1 \\ a_1 \end{pmatrix} \preceq \begin{pmatrix} v^2 \\ a_2 \end{pmatrix} \Leftrightarrow (a_2 - a_1) + \alpha \|v_1 - v_2\| \leq 0.$$

Lema 3. *Seja S um subconjunto fechado de $\mathbb{R}^n \times \mathbb{R}$ tal que*

$$\exists m \in \mathbb{R} : \begin{pmatrix} v \\ a \end{pmatrix} \in S \Rightarrow a \geq m.$$

Então para a ordem \preceq utilizada em $\mathbb{R}^n \times \mathbb{R}$ existe um elemento maximal $\begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix} \in S$, isto é, $\begin{pmatrix} v \\ a \end{pmatrix} \preceq \begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix}$, $\forall \begin{pmatrix} v \\ a \end{pmatrix} \in S$.

Demonstração. Inicialmente definamos uma sequência $\begin{pmatrix} v^n \\ a_n \end{pmatrix} \in S$, onde pomos $\begin{pmatrix} v^1 \\ a_1 \end{pmatrix} \in S$ como ponto inicial. Em seguida denotemos :

$$S_n = \left\{ \begin{pmatrix} v \\ a \end{pmatrix} \in S; \begin{pmatrix} v \\ a \end{pmatrix} \succeq \begin{pmatrix} v^n \\ a_n \end{pmatrix} \right\} \quad (2.4)$$

e

$$m_n = \inf \left\{ a \in \mathbb{R}; \begin{pmatrix} v \\ a \end{pmatrix} \in S_n \right\} \quad (2.5)$$

Agora perceba que $m_n \geq m$ para cada $n \in \mathbb{N}$, pois m_n é a maior das cotas inferiores de S_n . Em seguida, definimos o ponto $\begin{pmatrix} v^{n+1} \\ a_{n+1} \end{pmatrix} \in S_n$ de modo que

$$a_n - a_{n+1} \geq \frac{1}{2}(a_n - m_n) \geq 0. \quad (2.6)$$

Vejamos que para cada $n \in \mathbb{N}$, temos que S_n é fechado, não-vazio e $S_{n+1} \subset S_n$. De fato, seja $\{z^n\} \subset S_n$ tal que $z^n \rightarrow z$. Mostraremos que $z \in S_n$ e assim concluiremos que o referido conjunto é fechado. Supondo $z^n = \begin{pmatrix} x^n \\ y_n \end{pmatrix}$ de modo que $\begin{pmatrix} x^n \\ y_n \end{pmatrix} \succeq \begin{pmatrix} v^n \\ a_n \end{pmatrix}$, o que por definição nos dá que

$$(y_n - a_n) + \alpha \|x^n - v^n\| \leq 0.$$

Tomando o limite com $n \rightarrow +\infty$ e usando das propriedades de limite juntamente com a continuidade da função norma, segue que

$$(y - a_n) + \alpha \|x - v^n\| \leq 0.$$

Assim, segue que o ponto $z = \begin{pmatrix} x \\ y \end{pmatrix} \in S_n$ e consequentemente S_n é fechado. Como $\begin{pmatrix} v^n \\ a_n \end{pmatrix} \in S$ e também $\begin{pmatrix} v^n \\ a_n \end{pmatrix} \succeq \begin{pmatrix} v^n \\ a_n \end{pmatrix}$, tem-se que $S_n \neq \emptyset$. Resta-nos mostrar que

$S_{n+1} \subset S_n$. Com efeito, dado $w = \begin{pmatrix} p \\ q \end{pmatrix} \in S_{n+1}$ temos que

$$\begin{pmatrix} v^{n+1} \\ a_{n+1} \end{pmatrix} \succeq \begin{pmatrix} p \\ q \end{pmatrix} \text{ e } \begin{pmatrix} v^n \\ a_n \end{pmatrix} \succeq \begin{pmatrix} v^{n+1} \\ a_{n+1} \end{pmatrix}.$$

Por transitividade segue que $\begin{pmatrix} v^n \\ a_n \end{pmatrix} \succeq \begin{pmatrix} p \\ q \end{pmatrix}$ e conseqüentemente $w = \begin{pmatrix} p \\ q \end{pmatrix} \in S_n$.

Temos ainda de (2.6) que

$$\begin{aligned} \frac{1}{2}(a_n - m_n) &\leq a_n - a_{n+1} \\ a_{n+1} &\leq \frac{1}{2}(m_n + a_n). \end{aligned}$$

lembrando que $\{m_n\}$ é não decrescente temos

$$\begin{aligned} a_{n+1} - m_{n+1} &\leq \frac{1}{2}(m_n + a_n) - m_{n+1} \\ &\leq \frac{1}{2}(a_n + m_n) - m_n \\ &= \frac{1}{2}(a_n - m_n) \leq \dots \leq \frac{1}{2^n}(a_1 - m). \end{aligned} \tag{2.7}$$

Perceba que $a_{n+1} - m_{n+1} \geq 0$ e assim passando ao módulo conclui-se que

$$|a_{n+1} - m_{n+1}| \leq \frac{1}{2}|a_n - m_n| \leq \dots \leq \frac{1}{2^n}|a_1 - m|. \tag{2.8}$$

Conseqüentemente para cada $\begin{pmatrix} v \\ a \end{pmatrix} \in S_{n+1}$, temos de (2.4) que

$$\begin{aligned} |a_{n+1} - a| &\leq \frac{1}{2^n}|a_1 - m| \\ \|v_{n+1} - v_n\| &\leq \left(\frac{1}{2^n}\right) \left(\frac{1}{\alpha}\right) |a_1 - m|. \end{aligned} \tag{2.9}$$

Desse modo, temos que o diâmetro de S_n tende para zero quando $n \rightarrow +\infty$. Uma vez que $\mathbb{R}^n \times \mathbb{R}$ é um espaço métrico completo, existe um ponto $\begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix}$ tal que

$$\begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix} = \bigcap_{n=1}^{\infty} S_n$$

Por definição, $\begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix} \succeq \begin{pmatrix} v^n \\ a_n \end{pmatrix}$ para cada $n \in \mathbb{N}$ e em particular para $n = 1$. Se existisse $\begin{pmatrix} v^* \\ a^* \end{pmatrix} \in S$ de modo que $\begin{pmatrix} v^* \\ a^* \end{pmatrix} \succeq \begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix}$. Então teríamos que $\begin{pmatrix} v^* \\ a^* \end{pmatrix} \in \bigcap_{n=1}^{\infty} S_n$ e assim $\begin{pmatrix} v^* \\ a^* \end{pmatrix} = \begin{pmatrix} \bar{v} \\ \bar{a} \end{pmatrix}$. \square

Teorema 8. (Princípio Variacional de Ekeland)

Sejam $C \subset \mathbb{R}^n$ um subconjunto fechado e não-vazio e $F : C \rightarrow \mathbb{R}$ uma função contínua e limitada inferiormente. Se $u \in C$ satisfaz:

$$F(u) \leq \inf_{x \in C} F(x) + \varepsilon, \quad \varepsilon > 0,$$

então existe algum $v \in C$, tal que

$$\|u - v\| \leq \sqrt{\varepsilon},$$

e

$$F(x) + \sqrt{\varepsilon}\|x - v\| > F(v), \forall x \neq v \text{ em } C.$$

Demonstração. Seja S o epígrafo de F , ie,

$$S = \left\{ \begin{pmatrix} v \\ a \end{pmatrix}; v \in C, a \geq F(v) \right\} \tag{2.10}$$

Veja que S é não-vazio, e fechado em $\mathbb{R}^n \times \mathbb{R}$, uma vez que F é semicontínua e limitada inferiormente. Sejam $\varepsilon, \lambda > 0$, pondo $\alpha = \frac{1}{\sqrt{\varepsilon}}$, e aplicando observando que para cada

$\begin{pmatrix} v \\ a \end{pmatrix} \in S$ temos que $a \geq F(v) \geq \inf F$. Fazendo $m = \inf F$ temos que aplicando o Lema

3 ao epígrafo de F , obtemos um elemento maximal $\begin{pmatrix} v \\ a \end{pmatrix} \in S$ satisfazendo

$$\begin{pmatrix} p \\ q \end{pmatrix} \preceq \begin{pmatrix} v \\ a \end{pmatrix}, \forall \begin{pmatrix} p \\ q \end{pmatrix} \in S. \tag{2.11}$$

Como $\begin{pmatrix} v \\ a \end{pmatrix} \in S$, temos também que $\begin{pmatrix} v \\ a \end{pmatrix} \preceq \begin{pmatrix} v \\ F(v) \end{pmatrix}$, pois $(F(v) - a) + \frac{\varepsilon}{\lambda} \leq 0$.

Uma vez que $\begin{pmatrix} v \\ a \end{pmatrix}$ é maximal, temos que $a = F(v)$. Ainda do fato de $\begin{pmatrix} v \\ F(v) \end{pmatrix}$ ser

elemento maximal, temos que se dado

$$\begin{pmatrix} \mathbf{w} \\ \mathbf{b} \end{pmatrix} \in \mathcal{S} \text{ é tal que } \begin{pmatrix} \mathbf{v} \\ F(\mathbf{v}) \end{pmatrix} \prec \begin{pmatrix} \mathbf{w} \\ \mathbf{b} \end{pmatrix} \text{ então teríamos que } (\mathbf{b} - F(\mathbf{v})) + \frac{\varepsilon}{\lambda} \|\mathbf{w} - \mathbf{v}\| > 0,$$

desde que suponhamos $\mathbf{b} \neq F(\mathbf{v})$ e $\mathbf{w} \neq \mathbf{v}$. Agora pondo $\mathbf{b} = F(\mathbf{w})$ obtêm-se a relação $F(\mathbf{w}) > F(\mathbf{v}) - \frac{1}{\sqrt{\varepsilon}} \|\mathbf{w} - \mathbf{v}\|$. Agora perceba que $\begin{pmatrix} \mathbf{u} \\ F(\mathbf{u}) \end{pmatrix} \in \mathcal{S}$, logo de (2.11) temos que

$$(F(\mathbf{v}) - F(\mathbf{u})) + \frac{\varepsilon}{\lambda} \|\mathbf{v} - \mathbf{u}\| \leq 0. \quad (2.12)$$

Veja que $F(\mathbf{v}) \leq F(\mathbf{u})$. Por outro lado, de (2.3), temos que $F(\mathbf{v}) \geq F(\mathbf{u}) - \varepsilon$ e ao substituir na expressão (2.12) obtemos que $\frac{\varepsilon}{\lambda} \|\mathbf{v} - \mathbf{u}\| \leq \varepsilon$. Para concluir o resultado basta tomar $\lambda = \sqrt{\varepsilon}$. \square

Lema 4. *Sejam Ω_1, Ω_2 subconjuntos de \mathbb{R}^n não-vazios e compactos. Temos as seguintes propriedades:*

- i) Para $\lambda_1, \lambda_2 \geq 0$ e $\mathbf{v} \in \mathbb{R}^n$, $\max\{\langle \zeta, \mathbf{v} \rangle : \zeta \in \lambda_1 \Omega_1 + \lambda_2 \Omega_2\} = \lambda_1 \max\{\langle \zeta, \mathbf{v} \rangle : \zeta \in \Omega_1\} + \lambda_2 \max\{\langle \zeta, \mathbf{v} \rangle : \zeta \in \Omega_2\}$;*
- ii) $\Omega_1 \subset \Omega_2$ se, e somente se, para qualquer $\mathbf{v} \in \mathbb{R}^n$ $\max\{\langle \zeta, \mathbf{v} \rangle : \zeta \in \Omega_1\} \leq \max\{\langle \zeta, \mathbf{v} \rangle : \zeta \in \Omega_2\}$.*

Lema 5. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ localmente Lipschitz. Se uma sequência $\{z^r\}_{r=1}^\infty$ converge para x_0 , então*

$$\limsup_{r \rightarrow \infty} f^\circ(z^r; \mathbf{v}) \leq f^\circ(x_0; \mathbf{v}), \quad \text{para cada } \mathbf{v} \in \mathbb{R}^n$$

Demonstração. Primeiramente lembremos que

$$f^\circ(x; \mathbf{v}) := \lim_{\delta \downarrow 0} \sup_{\|\mathbf{h}\| + t < \delta} \frac{f(x + \mathbf{h} + t\mathbf{v}) - f(x + \mathbf{h})}{t}, \quad t > 0.$$

Vejamus que o caso em que $f^\circ(x_0; \mathbf{v}) = \infty$ é óbvio, então é suficiente provarmos o caso em que $f^\circ(x_0; \mathbf{v}) < \infty$. Assim para cada $\mu \in \mathbb{R}$ fixado, de modo que $f^\circ(x_0; \mathbf{v}) < \mu$, temos a existência de um $\delta > 0$ de tal forma que sempre que $\|\mathbf{h}\| + t < \delta$ e $t > 0$ tenhamos

$$\frac{f(x_0 + \mathbf{h} + t\mathbf{v}) - f(x_0 + \mathbf{h})}{t} < \mu.$$

Agora seja $\{z^r\}_{r=1}^\infty \in \mathbb{R}^n$ uma sequência tal que $z^r \rightarrow x_0$. Da definição de convergência

$$\|z^r - x_0\| < \frac{\delta}{2} \quad \forall r > r_0.$$

Para todo $h \in \mathbb{R}^n$, $t > 0$ tal que $\|h\| + t < \frac{\delta}{2}$, ponhamos $z^r - x_0 = h'$ em seguida veja que $\|h + h'\| + t \leq \|h'\| + (\|h\| + t) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$ e por conseguinte

$$\frac{f(z^r + h + tv) - f(z^r + h)}{t} = \frac{f(x_0 + h' + h + tv) - f(x_0 + h' + h)}{t} < \mu,$$

Tomando-se o \limsup na expressão anterior com $\delta \rightarrow 0$ ganhamos que

$$f^\circ(z^r; v) \leq \mu.$$

Queremos provar que

$$\limsup_{r \rightarrow \infty} f^\circ(z^r; v) \leq f^\circ(x_0; v).$$

Suponhamos por absurdo que tal desigualdade não é verdadeira. Assim teríamos que

$$f^\circ(x_0; v) < \limsup_{r \rightarrow \infty} f^\circ(z^r; v),$$

logo podemos encontrar uma constante μ_0 de modo que

$$f^\circ(x_0; v) < \mu_0 < \limsup_{r \rightarrow \infty} f^\circ(z^r; v).$$

Assim procedendo de maneira análoga para a constante μ_0 , conclui-se que

$$f^\circ(z^r; v) \leq \mu_0.$$

Essa contradição conclui a prova do Lema. □

Lema 6. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função localmente Lipschitz. Então seguem algumas propriedades:*

i) $(-f)^\circ(x; v) = f^\circ(x; -v)$, para $x, v \in \mathbb{R}^n$.

ii) $\partial(-f)(x) = -\partial f(x)$, $\forall x \in \mathbb{R}^n$

Demonstração. i) Provemos que $(-f)^\circ(x; v) \leq (f)^\circ(x; -v)$, pois a desigualdade contrária é obtida, trocando f por $-f$ e v por $-v$. Veja que no caso em que $f^\circ(x; -v) = +\infty$ nada há para se provar. Se $f^\circ(x; -v) < +\infty$ existe $\mu \in \mathbb{R}$, de modo que

$$f^\circ(x; -v) < \mu.$$

Por definição, existe $\delta > 0$ de modo que, se

$$\|\mathbf{h}\| + t < \delta, \text{ com } t > 0$$

temos que

$$\frac{f(\mathbf{x} + \mathbf{h} - t\mathbf{v}) - f(\mathbf{x} + \mathbf{h})}{t} < \mu.$$

Agora escolha $\mathbf{h} \in \mathbb{R}^n$ de modo que

$$\|\mathbf{h}\| + t < \frac{\delta}{\|\mathbf{v}\| + 1}.$$

Em seguida defina, $\mathbf{h}' = \mathbf{h} + t\mathbf{v}$ com $\mathbf{v} \neq 0$, daí temos:

$$\begin{aligned} \|\mathbf{h}'\| + t &= \|\mathbf{h} + t\mathbf{v}\| + t \\ &\leq \|\mathbf{h}\| + t\|\mathbf{v}\| + t \\ &= \|\mathbf{h}\| + t(\|\mathbf{v}\| + 1) \\ &\leq \|\mathbf{h}\|(\|\mathbf{v}\| + 1) + t(\|\mathbf{v}\| + 1) \\ &= (\|\mathbf{h}\| + t)(\|\mathbf{v}\| + 1) < \delta. \end{aligned}$$

E assim temos que

$$\begin{aligned} \frac{f(\mathbf{x} + \mathbf{h}' - t\mathbf{v}) - f(\mathbf{x} + \mathbf{h}')}{t} &< \mu \\ \frac{f(\mathbf{x} + \mathbf{h} + t\mathbf{v} - t\mathbf{v}) - f(\mathbf{x} + \mathbf{h} + t\mathbf{v})}{t} &< \mu \\ \frac{-f(\mathbf{x} + \mathbf{h} + t\mathbf{v}) + f(\mathbf{x} + \mathbf{h})}{t} &< \mu. \end{aligned}$$

Tomando o respectivo limite, vem que $(-f)^\circ(\mathbf{x}; \mathbf{v}) \leq \mu$. Suponha por absurdo que $(-f)^\circ(\mathbf{x}; \mathbf{v}) > (f)^\circ(\mathbf{x}; -\mathbf{v})$. Assim, existe $\mu \in \mathbb{R}$ tal que $(f)^\circ(\mathbf{x}; -\mathbf{v}) < \mu < (-f)^\circ(\mathbf{x}; \mathbf{v})$. O que contradiz o argumento provado anteriormente, e isso finaliza a prova do item i)

ii) Pelo item i) dado $\zeta \in \mathbb{R}^n$ tal que :

$$\begin{aligned} \zeta \in \partial(-f(\mathbf{x})) &\Rightarrow \langle \zeta, \mathbf{v} \rangle \leq -f^\circ(\mathbf{x}, \mathbf{v}) \\ &\Rightarrow \langle \zeta, \mathbf{v} \rangle \leq f^\circ(\mathbf{x}, -\mathbf{v}). \end{aligned}$$

Faça $-\mathbf{v} = \mathbf{w}$, daí

$$\begin{aligned} \langle \zeta, -\mathbf{w} \rangle &\leq f^\circ(\mathbf{x}, \mathbf{w}) \\ \langle -\zeta, \mathbf{w} \rangle &\leq f^\circ(\mathbf{x}, \mathbf{w}) \Rightarrow \zeta \in -\partial f(\mathbf{x}). \end{aligned}$$

Seja agora $\mathbf{p} \in -\partial f(\mathbf{x}) \Leftrightarrow -\mathbf{p} = \mathbf{y} \in \partial f(\mathbf{x})$, daí

$$\begin{aligned}\langle \mathbf{y}, \mathbf{v} \rangle &\leq f^\circ(\mathbf{x}, \mathbf{v}) \\ \langle -\mathbf{p}, \mathbf{v} \rangle &\leq f^\circ(\mathbf{x}, \mathbf{v}) \\ \langle \mathbf{p}, -\mathbf{v} \rangle &\leq f^\circ(\mathbf{x}, \mathbf{v}).\end{aligned}$$

Fazendo $-\mathbf{v} = \mathbf{z}$ temos

$$\begin{aligned}\langle \mathbf{p}, \mathbf{z} \rangle &\leq f^\circ(\mathbf{x}, -\mathbf{z}) \\ \langle \mathbf{p}, \mathbf{z} \rangle &\leq -f^\circ(\mathbf{x}, \mathbf{z}) \Rightarrow \mathbf{p} \in \partial(-f)(\mathbf{x}).\end{aligned}$$

Assim, finalizamos a prova do item ii) e conseqüentemente o referido Lema. \square

Lema 7. *Sejam $C \subset \mathbb{R}^n$ um subconjunto fechado, não-vazio e $f_i, g_j, h_k : \mathbb{R}^n \rightarrow \mathbb{R}$ funções localmente Lipschitz. Pondo $\Omega := \{\mathbf{x} \in C \mid g_j(\mathbf{x}) \leq 0, j \in J; h_k(\mathbf{x}) = 0, k \in K\}$, onde J e K são conjuntos finitos de índices, temos que se \mathbf{x}_0 é Pareto fraco para F restrita a Ω , então para qualquer $\varepsilon > 0$ e $\mathbf{x} \in C$,*

$$F_\varepsilon(\mathbf{x}) = \max\{f_i(\mathbf{x}) - f_i(\mathbf{x}_0) + \varepsilon, g_j(\mathbf{x}), |h_k(\mathbf{x})| : i \in I, j \in J, k \in K\} > 0.$$

Demonstração. Suponhamos por contradição que existem $\varepsilon_0 > 0$ e $\mathbf{x}_1 \in C$, tal que

$$F_{\varepsilon_0}(\mathbf{x}_1) \leq 0.$$

Então, temos

$$\left\{ \begin{array}{l} f_i(\mathbf{x}_1) < f_i(\mathbf{x}_1) + \varepsilon_0 \leq f_i(\mathbf{x}_0), \quad \forall i \in I \\ g_j(\mathbf{x}_1) \leq 0, \quad \forall j \in J, \\ h_k(\mathbf{x}_1) = 0, \quad \forall k \in K. \end{array} \right.$$

Além disso, temos que $\mathbf{x}_1 \in \Omega$ o que contradiz o fato de \mathbf{x}_0 ser Pareto fraco. \square

Agora enunciaremos o teorema que faz uso dos lemas anteriores em sua demonstração.

Teorema 9. *Sejam $C \subset \mathbb{R}^n$ um subconjunto fechado, não-vazio e $f_i, g_j, h_k : \mathbb{R}^n \rightarrow \mathbb{R}$ funções localmente Lipschitz e $\Omega := \{\mathbf{x} \in C \mid g_j(\mathbf{x}) \leq 0, j \in J; h_k(\mathbf{x}) = 0, k \in K\}$, onde J e K são conjuntos finitos de índices. Se $\mathbf{x}_0 \in \Omega$ é um Pareto fraco para F restrita a Ω , então existem números reais $\lambda_i \geq 0, i \in I, \mu_j \geq 0, j \in J, \nu_k \in \mathbb{R}, k \in K, \tau > 0$ tais que;*

- a) $0 \in \sum_{i \in I} \lambda_i \partial f_i(x_0) + \sum_{j \in J} \mu_j \partial g_j(x_0) + \sum_{k \in K} \nu_k \partial h_k(x_0) + \tau \partial d_C(x_0)$;
 b) $\mu_j g_j(x_0) = 0, j \in J$
 c) $\sum_{i \in I} \lambda_i + \sum_{j \in J} \mu_j + \sum_{k \in K} |\nu_k| = 1$.

Demonstração. Seja $x_0 \in \Omega$ um Pareto fraco para F , então pelo Lema 7 temos que para todo $\varepsilon > 0$ e $x \in C$

$$F_\varepsilon(x) > 0.$$

Como $f_i, i \in I, g_j, j \in J$ e $h_k, k \in K$ são funções localmente Lipschitz, temos pelo Lema 2 item v) que

$$F_\varepsilon(x) = \max\{f_i(x) - f_i(x_0) + \varepsilon, g_j(x), |h_k(x)| : i \in I, j \in J, k \in K\} > 0,$$

é localmente Lipschitz. Uma vez que $F_\varepsilon(x) > 0$ temos que 0 é uma cota inferior e assim $\inf_{x \in C} F_\varepsilon(x) \geq 0$. Desde que $x_0 \in \Omega$ temos que

$$\varepsilon = F(x_0) \leq \varepsilon + \inf_{x \in C} F_\varepsilon(x),$$

logo como consequência do Teorema 8 temos que existe $z(\varepsilon) \in C$ tal que

$$\|x_0 - z(\varepsilon)\| \leq \sqrt{\varepsilon}$$

e

$$F_\varepsilon(z(\varepsilon)) < F_\varepsilon(x) + \sqrt{\varepsilon}\|x - z(\varepsilon)\| \quad \forall x \neq z(\varepsilon) \in C.$$

Da equação acima temos que $z(\varepsilon)$ é um mínimo global para a função:

$$x \mapsto F_\varepsilon(x) + \sqrt{\varepsilon}\|x - z(\varepsilon)\|.$$

Seja L uma constante de Lipschitz para cada uma das funções f_i, g_j, h_k em uma vizinhança de x_0 , então L é uma constante de Lipschitz para $F_\varepsilon(x)$ em x_0 . Agora considere $\varepsilon \leq 1$ de modo que uma vizinhança V_ε de $z(\varepsilon)$ esteja contida numa vizinhança de x_0 . Vejamos que $(L+1)$ é uma constante de Lipschitz para a função $F_\varepsilon(x) + \sqrt{\varepsilon}\|x - z(\varepsilon)\|$ restrita a essa vizinhança. De fato, sejam x, y pertencentes a vizinhança V_ε dada anteriormente, então

$$\|F_\varepsilon(x) + \sqrt{\varepsilon}\|x - z(\varepsilon)\| - F_\varepsilon(y) - \sqrt{\varepsilon}\|y - z(\varepsilon)\|\| \leq \|F_\varepsilon(x) - F_\varepsilon(y)\| + \sqrt{\varepsilon} (\|x - z(\varepsilon)\| - \|y - z(\varepsilon)\|)$$

Assim, tem-se que

$$0 \in \sum_{i \in I} \lambda_i(\varepsilon) \partial f_i(z(\varepsilon)) + \sum_{j \in J} \mu_j(\varepsilon) \partial g_j(z(\varepsilon)) + \sum_{k \in K} v_k(\varepsilon) \partial h_k(z(\varepsilon)) + \sqrt{\varepsilon} \partial(\|\cdot\|)(0) + (L+1) \partial d_C(z(\varepsilon))$$

com $\sum_{i \in I} \lambda_i(\varepsilon) + \sum_{j \in J} \mu_j(\varepsilon) + \sum_{k \in K} |v_k(\varepsilon)| = 1$. Agora, pondo $\varepsilon = \frac{1}{n}$ consideramos as seguintes sequências :

$$\begin{cases} \{\lambda_i(\frac{1}{n})\}_{n=1}^{\infty}, & i \in I \\ \{\mu_j(\frac{1}{n})\}_{n=1}^{\infty}, & j \in J \\ \{v_k(\frac{1}{n})\}_{n=1}^{\infty}, & k \in K \end{cases}$$

Substituindo na Expressão anterior temos:

$$\sum_{i \in I} \lambda_i\left(\frac{1}{n}\right) + \sum_{j \in J} \mu_j\left(\frac{1}{n}\right) + \sum_{k \in K} |v_k\left(\frac{1}{n}\right)| = 1. \quad (2.15)$$

Como tais sequências são não-negativas, temos juntamente com a Equação (2.15) que estas sequências são limitadas e conseqüentemente admitem subsequências convergentes, as quais denotaremos por:

$$\begin{cases} \{\lambda_i(\frac{1}{n_r})\}_{r=1}^{\infty}, & i \in I \\ \{\mu_j(\frac{1}{n_r})\}_{r=1}^{\infty}, & j \in J \\ \{v_k(\frac{1}{n_r})\}_{r=1}^{\infty}, & k \in K \end{cases}$$

que ainda podem ser escritas

$$\begin{cases} \lambda_{i,r} = \lambda_i\left(\frac{1}{n_r}\right), & i \in I \\ \mu_{j,r} = \mu_j\left(\frac{1}{n_r}\right), & j \in J \\ v_{k,r} = v_k\left(\frac{1}{n_r}\right), & k \in K \end{cases}$$

Sendo $\lambda_i, i \in I, \mu_j, j \in J, v_k, k \in K$ seus respectivos limites, reescrevendo a Expressão (2.15) , passando a uma subsequência se necessário, e em seguida tomando o limite com $r \rightarrow \infty$ obtemos o seguinte:

$$\sum_{i \in I} \lambda_i + \sum_{j \in J} \mu_j + \sum_{k \in K} \|v_k\| = 1.$$

E assim segue-se o item c). Pondo $\varepsilon = \frac{1}{n_r}$ e $z_r = z\left(\frac{1}{n_r}\right)$. Temos

$$\|x_0 - z_r\| \leq \sqrt{\frac{1}{n_r}}$$

isso nos diz que $z_r \rightarrow x_0$. Assim, substituindo z_r em (2.14) obtemos

$$0 \in \sum_{i \in I} \lambda_{i,r} \partial f_i(z_r) + \sum_{j \in J} \mu_{j,r} \partial g_j(z_r) + \sum_{k \in K} v_{k,r} \partial h_k(z_r) + \sqrt{\frac{1}{n_r}} \partial(\|\cdot\|)(0) + (L+1) \partial d_C(z_r). \quad (2.16)$$

Agora, escrevendo

$$v_{k,r}^+ = \frac{|v_{k,r}| + v_{k,r}}{2} \quad k \in K, r = 1, 2, \dots$$

e

$$v_{k,r}^- = \frac{|v_{k,r}| - v_{k,r}}{2} \quad k \in K, r = 1, 2, \dots,$$

temos que $v_{k,r}^+$ e $v_{k,r}^-$ são convergentes por se tratarem de somas de seqüências (a menos de uma subsequência) convergentes. Além disso, tem-se ainda que

$$v_{k,r}^+ \rightarrow v_k^+ \geq 0$$

e

$$v_{k,r}^- \rightarrow v_k^- \geq 0,$$

onde a não-negatividade dos limites se dá devido a preservação de sinal por parte do limite. Veja ainda que $v_{k,r} = v_{k,r}^+ - v_{k,r}^-$ e que ao limite com $r \rightarrow \infty$ tem-se $v_k = v_k^+ - v_k^-$. Do Lema 6 item ii) obtemos:

$$\sum_{k \in K} v_{k,r} \partial h_k(z_r) = \sum_{k \in K} v_{k,r}^+ \partial h_k(z_r) + \sum_{k \in K} v_{k,r}^- \partial(-h_k)(z_r).$$

Assim, fazendo o produto interno na Expressão (2.14) com um vetor arbitrário $v \in \mathbb{R}^n$, tomando o máximo na expressão resultante, e utilizando-se ainda da Proposição 4 e do Lema 4 obtemos a seguinte expressão:

$$\begin{aligned} \sum_{i \in I} \lambda_{i,r} f_i^\circ(z_r; v) + \sum_{j \in J} \mu_{j,r} g_j^\circ(z_r; v) + \sum_{k \in K} v_{k,r}^+ h_k^\circ(z_r; v) + \\ \sum_{k \in K} v_{k,r}^- (-h)_k^\circ(z_r; v) + (L+1) d_C^\circ(z_r; v) \geq 0. \end{aligned}$$

Tomando na expressão acima o limite com $r \rightarrow \infty$ e usando o Lema 5 temos como resultado.

$$\sum_{i \in I} \lambda_i f_i^\circ(x_0; v) + \sum_{j \in J} \mu_j g_j^\circ(x_0; v) + \sum_{k \in K} v_k^+ h_k^\circ(x_0; v) + \sum_{k \in K} v_k^- (-h)_k^\circ(x_0; v) + (L+1) d_C^\circ(x_0; v) \geq 0.$$

Então novamente pela Proposição 4 e Lema 4, temos

$$0 \in \sum_{i \in I} \lambda_i \partial f_i(x_0) + \sum_{j \in J} \mu_j \partial g_j(x_0) + \sum_{k \in K} v_k^+ \partial h_k(x_0) + \sum_{k \in K} v_k^- \partial(-h)_k(x_0) + (L+1) \partial d_C(x_0).$$

Finalmente pelo Lema 6, item ii) obtemos a relação a) do Teorema.

Agora de modo obter o item b) temos que se para algum $j \in J$ tivermos que $g_j(x_0) = 0$

implica que $\mu_j g_j = 0$. Agora se $g_j(x_0) < 0$, para algum $j \in J$, segue da continuidade de g_j , juntamente com o Lema 7 que para um inteiro r suficientemente grande, vale que

$$g_j(z_r) < 0 < F_{1/n_r}(z_r)$$

e então temos

$$j \in J \setminus M_j(z_r),$$

daí temos que

$$\mu_{j,r} = 0$$

e conseqüentemente tomando o limite em r , tem-se finalmente que $\mu_j = 0$ e assim conclui-se o item b) e finalizamos a demonstração. \square

Consideremos a seguir o seguinte problema

$$\min_w \{F(x) : x \in \Omega\}$$

de encontrar o ponto Pareto-fraco para a função vetorial F restrita ao conjunto

$$\Omega := \{x \in D : g_s(x) \leq 0, s \in J\}$$

onde $D \subset \mathbb{R}^n$ é um conjunto fechado e não-vazio, e $g_s : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função localmente Lipschitz para cada $s \in J := \{1, \dots, p\}$. O próximo resultado apresenta uma condição necessária para que um ponto $x^* \in \Omega$ seja uma solução Pareto-fraco para o problema

$$\min_w \{F(x) : x \in \Omega\}$$

.

Teorema 10. *Seja $D \subset \mathbb{R}^n$ um conjunto fechado e não-vazio. Sejam $\Omega := \{x \in D : g_s(x) \leq 0\}$ e $f_j, g_s : \mathbb{R}^n \rightarrow \mathbb{R}$, $j \in I$ e $s \in J$, funções localmente Lipschitz. Se $x^* \in \Omega$ é solução para $\min_w \{F(x) : x \in \Omega\}$, então existem números reais $u_j \geq 0, v_s \geq 0$, com $j \in I$ e $s \in J$ e um escalar $\tau > 0$ tais que*

$$\sum_{j \in I} u_j \partial f_j(x^*) + \sum_{s \in J} v_s \partial g_s(x^*) + \tau \partial d_D(x^*) = 0$$

e

$$\sum_{j \in I} u_j + \sum_{s \in J} v_s = 1 \text{ e } v_s g_s(x^*) = 0, s \in J$$

Demonstração. A demonstração segue do Teorema 9, onde a função $h_\kappa(x) \equiv 0$ e o ponto $x_0 = x^*$. \square

Capítulo 3

Método do Ponto Proximal Multiobjetivo

Neste capítulo apresentamos o método do ponto proximal para otimização multiobjetivo. Iniciamos esse capítulo com uma série de definições e resultados que são utilizados a fim de provar que o algoritmo está bem definido. Por conseguinte, mostraremos alguns resultados de grande importância no capítulo que se segue .

3.1 O Algoritmo

No decorrer desta seção, iremos considerar $D \subset \mathbb{R}^n$ um conjunto fechado e não-vazio e $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que cada função coordenada $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$; $i \in I = \{1, \dots, m\}$, é localmente Lipschitz. Da Observação ?? podemos assumir sem perda de generalidade que $F \succ 0$, onde a ordem adotada aqui será a ordem de \mathbb{R}_+^m .

A seguir, vamos considerar o algoritmo do ponto proximal a fim de encontrar os pontos Pareto críticos de F em D . Seja $\{\lambda_k\}$ uma sequência de números reais estritamente positiva e seja $\{\varepsilon^k\} \subset \mathbb{R}_{++}^m$ uma sequência tal que $\|\varepsilon^k\| = 1, \forall k \geq 0$. O método gera uma sequência $\{x^k\} \subset D$ como segue:

Algoritmo

Inicialização: Escolha $x^0 \in D$.

Regra de Parada: Dado x^k , se x^k é um ponto Pareto crítico, então defina $x^{k+p} =$

$x^k, \forall p \in \mathbb{N}$.

Passo Iterativo: Dado $x^k \in D$ não satisfazendo a regra de parada, escolha a próxima iteração, $x^{k+1} \in D$ de modo que

$$x^{k+1} \in \operatorname{argmin}_w \{F(x) + \frac{\lambda_k}{2} \|x - x^k\|^2 \varepsilon^k : x \in \Omega_k\}, \quad (3.1)$$

onde $\Omega_k = \{x \in D : F(x) \preceq F(x^k)\}$.

Uma pergunta bastante pertinente a se fazer é se realmente é possível tomar x^{k+1} satisfazendo (3.1), isto é, nos questionarmos se o algoritmo está bem definido?

A seguir apresentamos uma série de resultados e definições os quais podem ser consultados em [15].

Lema 8. (*Lema de Zorn*) *Todo conjunto parcialmente ordenado, não-vazio, e no qual todo subconjunto totalmente ordenado tem cota superior, tem elemento maximal.*

Definição 33. *Seja $\{x^k\}$ uma sequência em \mathbb{R}^n . Dizemos que a sequência $\{x^k\}$ é decrescente se $x^\beta \prec x^\alpha$ sempre que $\alpha < \beta$. Onde (\prec) representa a ordem utilizada em \mathbb{R}^n .*

Definição 34. *Dizemos que um dado cone C é Daniell se qualquer sequência decrescente e limitada inferiormente, converge para o ínfimo.*

Exemplo 11. *Um exemplo de cone Daniell é o conjunto \mathbb{R}_+^m , $m \in \mathbb{N}$.*

Definição 35. *Um conjunto $A \subseteq \mathbb{R}^n$ é dito C -completo se A não admite uma cobertura da forma $\{(x^\alpha - cl(C))^c : \alpha \in I\}$ em que $\{x^\alpha\}$ é uma sequência decrescente em A e $I \subset \mathbb{N}$ é um conjunto de índices.*

Definição 36. *Seja C um cone convexo em \mathbb{R}^n , dizemos que C é Correto se vale que $cl(C) + C \setminus (C \cap -C) \subseteq C$.*

Exemplo 12. *Um exemplo de cone Correto é o espaço \mathbb{R}_+^2 , onde tal propriedade pode ser verificada por meio de um esboço geométrico.*

Definição 37. *Seja $x \in \mathbb{R}^n$. O conjunto $A \cap (x - C)$ é chamado uma seção de A em x , e é denotado por A_x .*

Definição 38. *Um conjunto B é dito limitado de ordem completa se qualquer sequência limitada decrescente possui um ínfimo.*

É fácil verificar que \mathbb{R}^n satisfaz a definição acima.

Lema 9. *Se um conjunto $A \subset \mathbb{R}^n$ é fechado e limitado inferiormente, $C \subset \mathbb{R}^n$ é um cone Daniell, então A é C -completo.*

Demonstração. Se a sequência $\{x^\alpha\}$ é decrescente em A , então esta possui um ínfimo para o qual é convergente, uma vez que $A \subset \mathbb{R}^n$ e esse é limitado de ordem completo. Além disso o ponto limite dessa sequência pertence ao conjunto A , pois A é fechado. Veja que tal ponto também pertence a $x^\alpha - \text{cl}(C)$ para algum α . Assim A não admite cobertura da forma

$$\{(x^\alpha - \text{cl}(C))^c; \alpha \in I\}.$$

E assim finda-se a demonstração. □

Observação 7. *A seguinte notação $A|C$ será utilizada para representar que a ordenação utilizada no conjunto A é a mesma utilizada em C .*

Teorema 11. *Assuma que C é um cone convexo Correto e A é um conjunto não-vazio em \mathbb{R}^n . Então $\text{Min}(A|C) \neq \emptyset$ se, somente se A tem uma seção C -completa não-vazia.*

Demonstração. \Rightarrow) Suponhamos que o $\text{Min}(A|C) \neq \emptyset$ e seja $x_0 \in \text{Min}(A|C)$, assim da definição de mínimo temos que não existe $y \in A$ tal que $y \leq_c x_0$. Em seguida tomemos A_{x_0} uma seção em x_0 e pela definição de seção temos:

$$A_{x_0} = A \cap (x_0 - C).$$

Agora seja $z \in A_{x_0}$, então tem-se: $z = x_0 - c_z$, onde $c_z \in C$. Observemos que

$$\begin{aligned} z = x_0 - c_z &\Leftrightarrow c_z = x_0 - z \\ &\Leftrightarrow x_0 - z \in C \\ &\Leftrightarrow x_0 - z \geq_c 0 \\ &\Leftrightarrow z \leq_c x_0. \end{aligned}$$

Uma vez que $x_0 \in \text{Min}(A|C)$, temos que $z = x_0$. Logo conclui-se que $A_{x_0} = \{x_0\}$. Isso nos diz que A_{x_0} apresenta apenas sequências constantes. Consequentemente da definição de seção temos que A_{x_0} é C -completo.

\Leftarrow) Seja A_x uma seção C -completa de A . É suficiente mostrar que $\text{Min}(A_x|C) \neq \emptyset$. De

fato, seja $\mathbf{y} \in \text{Min}(\mathbf{A}_x | \mathbf{C})$ se existisse $\mathbf{z} \in \mathbf{A} \setminus \mathbf{A}_x$ de forma que valesse $\mathbf{z} \leq \mathbf{y}$ teríamos consequentemente que $\mathbf{z} \in \mathbf{A}_x$. Assim, devemos ter $\mathbf{y} \leq \mathbf{z}$ e nesse caso $\mathbf{y} \in \text{Min}(\mathbf{A})$. Assim conclui-se que $\text{Min}(\mathbf{A}_x | \mathbf{C}) \subseteq \text{Min}(\mathbf{A})$.

Definamos o seguinte conjunto

$$\mathbf{P} := \{\{\mathbf{x}^n\} \in \mathbf{A}; \mathbf{x}^n \text{ é decrescente}\}.$$

Como $\mathbf{A} \neq \emptyset$ temos que \mathbf{P} também o é. Agora dadas duas seqüências $\{\mathbf{a}^n\}, \{\mathbf{b}^n\} \in \mathbf{P}$ escrevemos $\mathbf{b}^n \prec \mathbf{a}^n$ para significar que $\{\mathbf{b}^n\}$ é subsequência de $\{\mathbf{a}^n\}$. Note que $(\mathbf{P}; \prec_{\mathbf{P}})$ é um conjunto parcialmente ordenado. Afirmamos que qualquer seqüência em \mathbf{P} possui uma cota superior. De fato, seja $\{\mathbf{a}^\lambda; \lambda \in \Lambda\}$ uma seqüência em \mathbf{P} , denotemos por \mathbf{B}' o conjunto dos subconjuntos finitos \mathbf{B} de Λ munidos da ordem $\prec_{\mathbf{P}}$ e sejam

$$\mathbf{a}^{\mathbf{B}} = \bigcup \{\mathbf{a}^\lambda; \lambda \in \mathbf{B}\}.$$

e

$$\mathbf{a}^0 = \bigcup \{\mathbf{a}^{\mathbf{B}}; \mathbf{B} \in \mathbf{B}'\}.$$

Então \mathbf{a}^0 é um elemento de \mathbf{P} . Além disso $\mathbf{a}^\lambda \prec_{\mathbf{P}} \mathbf{a}^0$, para cada $\lambda \in \Lambda$ isto é, \mathbf{a}^0 é uma cota superior para a seqüência \mathbf{a}^λ . Assim do Lema 8 obtemos que \mathbf{P} possui um elemento maximal, o qual chamaremos de $\mathbf{a}^* = \{\{\mathbf{x}^\alpha\}; \alpha \in \mathbf{I}\} \in \mathbf{P}$. Agora suponhamos por contradição que $\text{Min}(\mathbf{A}_x | \mathbf{C})$ é vazio, provaremos que $\{(\mathbf{x}^\alpha - \text{cl}(\mathbf{C}))^c; \alpha \in \mathbf{I}\}$ forma uma cobertura para \mathbf{A}_x , e assim chegaremos em uma contradição. Seja $\mathbf{y} \in \mathbf{A}_x$ existe $\alpha \in \mathbf{I}$ tal que $(\mathbf{x}^\alpha - \text{cl}(\mathbf{C}))^c$ contém \mathbf{y} . Caso isso não ocorra, $\mathbf{y} \in \mathbf{x}^\alpha - \text{cl}(\mathbf{C})$ para cada $\alpha \in \mathbf{I}$. Como $\text{Min}(\mathbf{A}_x | \mathbf{C})$ é vazio por hipótese, temos que existe algum $\mathbf{z} \in \mathbf{A}_x$ com $\mathbf{z} <_{\mathbf{C}} \mathbf{y}$. Como $\mathbf{y} \in \mathbf{x}^\alpha - \text{cl}(\mathbf{C})$ temos:

$$\begin{aligned} \mathbf{y} = \mathbf{x}^\alpha - \mathbf{c}_y, \mathbf{c}_y \in \text{cl}(\mathbf{C}) &\Leftrightarrow \mathbf{c}_y = \mathbf{x}^\alpha - \mathbf{y} \\ &\Leftrightarrow \mathbf{x}^\alpha - \mathbf{y} \in \text{cl}(\mathbf{C}) \\ &\Leftrightarrow \mathbf{y} \leq_{\text{cl}(\mathbf{C})} \mathbf{x}^\alpha. \end{aligned}$$

Como \mathbf{C} por hipótese é Correto, em particular é fechado, logo da transitividade da ordenação parcial vem que $\mathbf{z} <_{\mathbf{C}} \mathbf{x}^\alpha$ para cada $\alpha \in \mathbf{I}$. Finalmente, adicionando \mathbf{z} a seqüência

$\{\mathbf{a}^*\}$ obtemos uma nova sequência decrescente que é maior (segundo a ordem \prec_P) que a sequência $\{\mathbf{a}^*\}$ e isso contradiz o Lema 8. Logo $(\mathbf{x}^\alpha - \text{cl}(C))^c$ forma uma cobertura para A_x o que novamente gera uma contradição em relação a nossa hipótese inicial. \square

Proposição 5. *O algoritmo está bem definido*

Demonstração. O ponto inicial $\mathbf{x}^0 \in D$ é escolhido no passo de inicialização, em seguida suponhamos que o algoritmo é válido até a iteração k e mostraremos que é possível fazer a iteração $(k + 1)$. Chame $F_k(\mathbf{x}) := F(\mathbf{x}) + \frac{\lambda_k}{2} \|\mathbf{x} - \mathbf{x}^k\|^2 \varepsilon^k$ e veja que $\mathbf{x}^k \in \Omega_k$, uma vez que $F(\mathbf{x}^k) \preceq F(\mathbf{x}^k)$. Além disso, temos também que $F_k(\mathbf{x}^k) = F(\mathbf{x}^k)$, o que implica que $F_k(\Omega_k) \neq \emptyset$. Vejamos ainda que valem as seguintes propriedades:

i) $F_k(\mathbf{x}) \succ 0$;

ii) $F_k(\Omega_k)$ é fechado.

O item i) se justifica do fato de $F_k(\mathbf{x})$ ser soma de funções com essa propriedade. Já para o item ii) observemos inicialmente que a função $F_k(\mathbf{x})$ é coerciva, isto é, vale que $\lim_{\mathbf{x} \rightarrow \infty} \|F_k(\mathbf{x})\| = +\infty$. Seja $\{\mathbf{y}^n\}$ uma sequência em $F_k(\Omega_k)$ tal que $\mathbf{y}^n \rightarrow \mathbf{y}$, mostraremos que $\mathbf{y} \in F_k(\Omega_k)$ e assim obteremos o resultado. Uma vez que $\{\mathbf{y}^n\} \in F_k(\Omega_k)$ temos que existe $\{\mathbf{z}^n\} \in \Omega_k$ tal que $F_k(\mathbf{z}^n) = \mathbf{y}^n$. Veja que se $\{\mathbf{z}^n\}$ é limitada então esta admite uma subsequência convergente, digamos $\mathbf{z}^{n_j} \rightarrow \mathbf{z}^j$. Perceba que $\mathbf{z}^j \in \Omega_k$ uma vez que tal conjunto é fechado e assim, aplicando a função F_k e em seguida utilizando-se da unicidade do limite, temos que $\mathbf{y} \in F_k(\Omega_k)$ e por conseguinte segue que $F_k(\Omega_k)$ é fechado. Resta-nos provar que a sequência $\{\mathbf{z}^n\}$ é limitada. Com efeito, se $\{\mathbf{z}^n\}$ fosse ilimitada, então esta possuiria uma subsequência $\{\mathbf{z}^{n_j}\}$ tal que $\|\mathbf{z}^{n_j}\| \rightarrow +\infty$. Uma vez que F_k é coerciva teríamos por um lado que $\|F_k(\mathbf{z}^{n_j})\| \rightarrow +\infty$, mas também que $\|F_k(\mathbf{z}^{n_j})\| \rightarrow \|\mathbf{y}\|$ pois $\mathbf{y}^{n_j} = F_k(\mathbf{z}^{n_j})$ e $\mathbf{y}^{n_j} \rightarrow \mathbf{y}$ uma vez que $\mathbf{y}^n \rightarrow \mathbf{y}$ gerando assim uma contradição, logo se tem que $\{\mathbf{z}^n\}$ é limitada.

Agora perceba que o cone \mathbb{R}_+^m é Daniell, além disso como $F_k(\Omega_k)$ é fechado e limitado inferiormente. Assim, temos do Lema 9 que $F_k(\Omega_k)$ é \mathbb{R}_+^m -completo. Por conseguinte do Teorema 11 segue que

$$\text{argmin}_w \{F_k(\mathbf{x}) : \mathbf{x} \in \Omega_k\}$$

é não-vazio, e conseqüentemente a boa definição do algoritmo. \square

Teorema 12. *Seja C um subconjunto fechado e não-vazio de \mathbb{R}^m . Se $\mathbf{x} \in C$, então*

$$\partial d_C(\mathbf{x}) \subset B[0, 1] \cap \mathcal{N}_C(\mathbf{x}),$$

onde $B[0, 1]$ denota a bola unitária de \mathbb{R}^m .

Demonstração. Inicialmente lembremos que a função distância é uma função globalmente Lipschitz com constante de Lipschitz igual a 1 (Ver [8] prop.(2.4.1)). Temos ainda de [8] prop.(2.1.2) que se uma função $f : X \rightarrow \mathbb{R}$ é localmente Lipschitz em uma vizinhança de x com constante de Lipschitz L , então vale que $\partial f(x)$ é um subconjunto de \mathbb{R}^n não-vazio, convexo e compacto. Além disso, tem-se que $\|\xi\| \leq L$, para cada $\xi \in \partial f(x)$. Como $L = 1$ para a função distância temos que $\partial d_C(x) \subset B[0, 1]$. Por outro lado temos de [8] prop.(2.4.2) que $\mathcal{N}_C(x) = \text{cl}\{\cup_{\lambda \geq 0} \lambda \partial d_C(x)\}$. Isso nos diz que $\partial d_C(x) \subset \mathcal{N}_C(x)$, o que finaliza a demonstração. \square

De agora em diante, as sequências $\{x^k\}, \{\lambda_k\}, \{e^k\}$, denotarão as sequencias consideradas no Algoritmo. A seguir apresentamos um resultado de grande importância no decorrer do trabalho e será muito utilizado no que diz respeito a análise dos resultados.

Proposição 6. *Para todo $k \in \mathbb{N}$, existe $A_k \in \mathbb{R}^{m \times n}$, $u^k, v^k \in \mathbb{R}_+^m$, $w^k \in \mathbb{R}^m$ e $\tau_k \in \mathbb{R}_{++}$ tal que*

$$A_k^\top (u^k + v^k) + \lambda_{k-1} \langle e^{k-1}, u^k \rangle (x^k - x^{k-1}) + \tau_k w^k = 0, \quad (3.2)$$

onde

$$w^k \in B[0, 1] \cap \mathcal{N}_D(x^k), \quad \|u^k + v^k\|_1 = 1, \quad A_k^\top = [a_1^k, \dots, a_m^k]^\top \quad \forall k \in \mathbb{N} \quad (3.3)$$

Demonstração. Para a demonstração do referido Teorema, utilizaremos o Teorema 10 aplicado a função $F_{k-1}(x) = F(x) + \frac{\lambda_{k-1}}{2} \|x - x^{k-1}\|^2 e^{k-1}$, em que o ponto x^k é solução Pareto-fraca para o problema

$$\min_w \{F_{k-1}(x) : x \in \Omega_{k-1}\},$$

onde $\Omega_{k-1} = \{x \in D : F(x) \preceq F(x^{k-1})\}$. Denotando $G_{k-1}(x) = F(x) - F(x^{k-1})$, temos que $G_{k-1}(x) \preceq 0, \forall x \in \Omega_{k-1}$. Além disso observemos que como F é localmente Lipschitz, consequentemente suas funções coordenadas também são e com mais razão as funções coordenadas de $G_{k-1}(x)$ a saber

$$(g_{k-1})_j(\cdot) = f_j(\cdot) - f_j(x^{k-1}), \text{ com } j \in I$$

serão funções localmente Lipschitz. Logo, para cada $k \in \mathbb{N}$ fixado, aplicamos o Teorema 10 e assim obtemos $u_j^k, v_j^k \geq 0$ e $\tau_k > 0$ tais que

$$\sum_{j \in I} u_j^k \partial (f_{k-1})_j(x^k) + \sum_{j \in I} v_j^k \partial (g_j)(x^k) + \tau^k \partial d_D(x^k) = 0.$$

Uma vez que $(f_{k-1})_j(\cdot) = f_j(\cdot) + \frac{\lambda_{k-1}}{2} \|\cdot - \mathbf{x}^{k-1}\|^2 \varepsilon_j^{k-1}$, substituindo $(f_{k-1})_j(\cdot)$ e $g_j(\cdot)$ na expressão acima temos:

$$\sum_{j \in I} \mathbf{u}_j^k \partial \left((f_j(\cdot) + \frac{\lambda_{k-1}}{2} \|\cdot - \mathbf{x}^{k-1}\|^2 \varepsilon_j^{k-1}) \right) (\mathbf{x}^k) + \sum_{j \in I} \mathbf{v}_j^k \partial (f_j(\cdot) - f_j(\mathbf{x}^{k-1})) (\mathbf{x}^k) + \tau_k \partial \mathbf{d}_D(\mathbf{x}^k) = 0.$$

Rearrmando a expressão acima obtemos:

$$\sum_{j \in I} \mathbf{u}_j^k \partial f_j(\mathbf{x}^k) + \sum_{j \in I} \mathbf{u}_j^k \frac{\lambda_{k-1}}{2} 2(\mathbf{x}^k - \mathbf{x}^{k-1}) \varepsilon_j^k + \sum_{j \in I} \mathbf{v}_j^k \partial f_j(\mathbf{x}^k) + \tau_k \partial \mathbf{d}_D(\mathbf{x}^k) = 0.$$

Fazendo algumas manipulações algébricas na expressão anterior, tem-se

$$\sum_{j \in I} (\mathbf{u}_j^k + \mathbf{v}_j^k) \partial f_j(\mathbf{x}^k) + \lambda_{k-1} \langle \mathbf{u}^k, \varepsilon^k \rangle (\mathbf{x}^k - \mathbf{x}^{k-1}) + \tau_k \mathbf{w}^k = 0. \quad (3.4)$$

Onde $\sum_{j \in I} \mathbf{u}_j^k + \sum_{j \in I} \mathbf{v}_j^k = 1$ e $\mathbf{w}^k \in \partial \mathbf{d}_D(\mathbf{x}^k)$. Ainda do Teorema 10, temos as seguintes observações:

a) $\partial \mathbf{d}_D(\mathbf{x}^k) \subset \mathbf{B}[0, 1] \cap \mathcal{N}_D \forall k \in \mathbb{N}$.

b) Existem $\mathbf{a}_1^k, \dots, \mathbf{a}_m^k \in \partial f_1(\mathbf{x}^k), \dots, \partial f_m(\mathbf{x}^k)$ respectivamente, que ao serem substituídos na expressão (3.2) nos fornecem a seguinte equação:

$$\sum_{i=1}^m \mathbf{a}_i^k (\mathbf{u}_j^k + \mathbf{v}_j^k) + \lambda_{k-1} \langle \mathbf{u}^k, \varepsilon^k \rangle (\mathbf{x}^k - \mathbf{x}^{k-1}) + \tau^k \mathbf{w}^k = 0. \quad (3.5)$$

E assim findamos a demonstração. □

Observação 8. *Veja que da equação (3.2), (3.3) temos que $\{\mathbf{u}^k\}, \{\mathbf{v}^k\}$ e $\{\mathbf{w}^k\}$ são sequências limitadas. Além disso temos também que ∂f_j é limitado quando restrito a conjuntos compactos ([3] obs. 1). Assim, temos que $\{\mathbf{A}_k\}$ é limitado, uma vez que $\{\mathbf{x}^k\}$ é limitado, e $\mathbf{a}_j^k \in \partial f_j(\mathbf{x}^k)$, $j \in I$. Agora, se $\{\lambda_k\}$ e $\{\mathbf{x}^k\}$ são sequências limitadas, então temos de (3.2) que $\{\tau_k\}$ é limitada.*

Como consequência da proposição anterior, obtemos a seguinte regra de parada para o Algoritmo.

Corolário 1. *Seja $k_0 \in \mathbb{N}$ tal que $\mathbf{u}^{k_0} = 0$. Então \mathbf{x}^{k_0} é um ponto Pareto crítico de F .*

Demonstração. Se existe $k_0 \in \mathbb{N}$ tal que $\mathbf{u}^{k_0} = 0$ então temos da expressão (3.2) que

$$\mathbf{A}_{k_0}^\top \mathbf{v}^{k_0} + \tau_{k_0} \mathbf{w}^{k_0} = 0$$

$$-\mathbf{A}_{k_0}^\top \mathbf{v}^{k_0} = \tau_{k_0} \mathbf{w}^{k_0}$$

Uma vez que $\tau_{k_0} > 0$ e $w^{k_0} \in \mathcal{N}_D(x^{k_0})$ temos que

$$-A_{k_0}^\top v^{k_0} \in \mathcal{N}_D(x^{k_0}),$$

além disso como $\|u_0^k + v_0^k\|_1 = 1$ temos que $v_0^k \in \mathbb{R}_+^m \setminus \{0\}$. Temos ainda que $A_{k_0} \in \partial F(x_0^k)$, e assim o resultado segue usando o Lema 1, onde fazemos $U = A_{k_0}$, $w = v_0^k$ e $x = x_0^k$. \square

Como pode ser visto em [4], podemos trocar a regra de parada para o Algoritmo pela seguinte regra: Após computar o termo x^{k+1} o algoritmo para se $x^{k+1} = x^k$, isto é, definimos $x^{k+p} = x^k \forall p \geq 1$. Justificamos tal fato no corolário seguinte.

Corolário 2. *Se $x^{k+1} = x^k$, então x^k é um ponto crítico Pareto de F .*

Demonstração. Se $x^{k+1} = x^k$, então de (3.2) temos que

$$-A_k^\top v^k = \tau_k w^k,$$

daí o resultado segue diretamente do Lema 1, onde $U = A_k$, $w = u^k + v^k$ e $x = x^k = x^{k+1}$. \square

Observação 9. *Perceba que se o Algoritmo termina após um número finito de iterações, a última iteração produz um ponto Pareto crítico. Isto nos leva a supor que $\{x^k\}$ é uma sequência infinita, e conseqüentemente $u^k \neq 0 \forall k \in \mathbb{N}$, uma vez que caso contrário, o ponto Pareto-crítico é garantido pelos corolários da Proposição 6.*

Capítulo 4

Análise de Convergência

Nesse capítulo fazemos uma análise da convergência do Método do Ponto Proximal para os casos onde F é localmente Lipschitz e posteriormente o caso em que F é Quase-covexa, bem como alguns resultados decorrente dessas hipóteses.

4.1 Caso Localmente Lipschitz

Teorema 13. *Suponha que existam escalares a, b e $c \in \mathbb{R}$ tais que $0 < a \leq \lambda_k \leq b$ e $0 < c \leq \varepsilon_j^k \forall k \in \mathbb{N}$ e $j = 1, \dots, m$. Então, cada ponto de acumulação de $\{\mathbf{x}^k\}$, quando existir, é um ponto Pareto crítico de F .*

Demonstração. Suponhamos que \mathbf{x}^k é solução para o seguinte problema

$$\min_{\mathbf{x}} \left\{ F(\mathbf{x}) + \frac{\lambda_{k-1}}{2} \|\mathbf{x} - \mathbf{x}^{k-1}\|^2 \varepsilon^{k-1} : \mathbf{x} \in \Omega_{k-1} \right\}.$$

Isso nos diz que não existe $\mathbf{x} \in \Omega_{k-1}$ tal que

$$F_{k-1}(\mathbf{x}) < F_{k-1}(\mathbf{x}^k),$$

em que $F_{k-1}(\mathbf{x}) = F(\mathbf{x}) + \frac{\lambda_{k-1}}{2} \|\mathbf{x} - \mathbf{x}^{k-1}\|^2 \varepsilon^{k-1}$. Assim, para $\mathbf{x}^{k-1} \in \Omega_{k-1}$, $\exists i \in \{1, \dots, m\}$ tal que

$$(F_{k-1})_i(\mathbf{x}^{k-1}) \geq (F_{k-1})_i(\mathbf{x}^k)$$

e conseqüentemente,

$$(F_{k-1})_i(\mathbf{x}^{k-1}) - (F_{k-1})_i(\mathbf{x}^k) \geq 0.$$

Agora tomando o máximo na expressão acima obtemos :

$$\max_{1 \leq i \leq m} \left\{ F_i(\mathbf{x}^{k-1}) - \frac{\lambda_{k-1}}{2} \|\mathbf{x}^{k-1} - \mathbf{x}^{k-1}\|^2 \varepsilon_i^{k-1} - \left(F_i(\mathbf{x}^k) + \frac{\lambda_{k-1}}{2} \|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2 \varepsilon_i^{k-1} \right) \right\} \geq 0.$$

E conseqüentemente tem-se

$$\max_{1 \leq i \leq m} \left\{ F_i(\mathbf{x}^{k-1}) - F_i(\mathbf{x}^k) - \frac{\lambda_{k-1}}{2} \|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2 \varepsilon_i^{k-1} \right\} \geq 0.$$

Pondo $j_0(k) = j_0 \in \{1, \dots, m\}$ o índice para o qual o valor máximo da inequação acima é atingido, temos que

$$\frac{\lambda_{k-1}}{2} \|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2 \varepsilon_{j_0}^{k-1} \leq F_{j_0}(\mathbf{x}^{k-1}) - F_{j_0}(\mathbf{x}^k).$$

Daí utilizando-se das hipóteses a cerca de $\{\lambda_k\}, \{\varepsilon^k\}$ temos:

$$\begin{aligned} \frac{\alpha c}{2} \|\mathbf{x}^k - \mathbf{x}^{k-1}\|^2 &\leq F_{j_0}(\mathbf{x}^{k-1}) - F_{j_0}(\mathbf{x}^k) \leq \sqrt{(F_{j_0}(\mathbf{x}^{k-1}) - F_{j_0}(\mathbf{x}^k))^2} \leq \sqrt{\sum_{i=1}^n (F_i(\mathbf{x}^{k-1}) - F_i(\mathbf{x}^k))^2} \\ &= \|F(\mathbf{x}^{k-1}) - F(\mathbf{x}^k)\|. \end{aligned}$$

Como $\{F(\mathbf{x}^k)\}$ é não crescente e $F \succ 0$, temos que o lado direito na desigualdade acima tende para zero, quando $k \rightarrow +\infty$. Conseqüentemente, vem que

$$(\mathbf{x}^k - \mathbf{x}^{k-1}) \rightarrow 0, \quad k \rightarrow \infty.$$

Agora, seja $\hat{\mathbf{x}}$ um ponto de acumulação de $\{\mathbf{x}^k\}$ e seja $\{\mathbf{x}^{k_l}\}$ uma subsequência de $\{\mathbf{x}^k\}$ convergindo para $\hat{\mathbf{x}}$. Aplicando a Proposição 6 para a sequência $\{\mathbf{x}^{k_l}\}$ temos que existem sequências $\{A_{k_l}\} \subset \mathbb{R}^{m \times n}$, $\{\mathbf{u}^{k_l}\}, \{\mathbf{v}^{k_l}\} \subset \mathbb{R}_+^m$, $\{\mathbf{w}^{k_l}\} \subset \mathbb{R}^m$ e $\{\tau_{k_l}\} \subset \mathbb{R}_{++}$ satisfazendo :

$$A_{k_l}^\top (\mathbf{u}^{k_l} + \mathbf{v}^{k_l}) + \lambda_{k_l-1} \langle \varepsilon^{k_l-1}, \mathbf{u}^{k_l} \rangle (\mathbf{x}^{k_l} - \mathbf{x}^{k_l-1}) + \tau_{k_l} \mathbf{w}^{k_l} = 0. \quad (4.1)$$

Agora percebe-se que $\{\lambda_{k_l}\}$ é limitada e uma vez que $\{\mathbf{x}^{k_l}\}$ converge para $\hat{\mathbf{x}}$ temos que a mesma também é limitada. Além disso, da Observação 8 podemos assumir que as sequências $\{A_{k_l}\}, \{\mathbf{u}^{k_l}\}, \{\mathbf{v}^{k_l}\}, \{\mathbf{w}^{k_l}\}$, e $\{\tau_{k_l}\}$ são limitadas. (A menos de uma subsequência) Podemos supor $A_{k_l} \rightarrow \hat{A}$, $\mathbf{u}^{k_l} \rightarrow \hat{\mathbf{u}}$, $\mathbf{v}^{k_l} \rightarrow \hat{\mathbf{v}}$ e $\tau_{k_l} \rightarrow \hat{\tau}$ quando $l \rightarrow \infty$ (Usaremos a mesma notação para os índices, mesmo que seja necessário extrair uma nova subsequência). Uma vez que $\{\lambda_{k_l-1} \langle \varepsilon^{k_l-1}, \mathbf{u}^{k_l} \rangle\}$ é limitado e usando o fato de que $(\mathbf{x}^{k_l} - \mathbf{x}^{k_l-1}) \rightarrow 0$ quando $l \rightarrow \infty$ temos da expressão (4.1)

$$\hat{A}^\top \hat{\mathbf{y}} + \hat{\tau} \hat{\mathbf{w}} = 0, \quad (4.2)$$

onde $\hat{\mathbf{y}} := (\hat{\mathbf{u}} + \hat{\mathbf{v}}) \in \mathbb{R}_+^m \setminus \{0\}$, $\hat{A} \in \partial F(\hat{\mathbf{x}})$ e $\hat{\mathbf{w}} \in \mathcal{N}_D(\hat{\mathbf{x}})$, pois $\partial F(\cdot)$ e $\mathcal{N}_D(\cdot)$ são fechados. Finalmente de (4.2) obtemos

$$-A_{k_l}^\top \hat{\mathbf{y}} \in \mathcal{N}_D(\hat{\mathbf{x}}).$$

Agora novamente utilizando do Lema 1, temos que $\hat{\mathbf{x}}$ é um ponto Pareto crítico de F . \square

4.2 Caso Quase Convexo

Nessa seção, consideramos o Algoritmo com a hipótese adicional de que a função $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é \mathbb{R}_+^m -quase convexa. D é um conjunto convexo e com a hipótese de \mathbb{R}_+^m -completude para o conjunto $(F(x^0) - \mathbb{R}_+^m) \cap F(D)$:

(\mathcal{H}) : para cada sequência $\{\mathbf{a}^k\} \subset D$, com $\mathbf{a}^0 = x^0$, tal que $F(\mathbf{a}^{k+1}) \preceq F(\mathbf{a}^k)$, $\forall k \in \mathbb{N}$, então existe $\mathbf{a} \in D$ tal que

$$F(\mathbf{a}) \preceq F(\mathbf{a}^k), \forall k \in \mathbb{N}.$$

Definição 39. Uma sequência $\{\mathbf{y}^k\}$ é dita Féjer convergente(Féjer monótona) para um conjunto não-vazio $\mathcal{U} \subset \mathbb{R}^n$ se, para todo $k \in \mathbb{N}$

$$\|\mathbf{y}^{k+1} - \mathbf{y}\| \leq \|\mathbf{y}^k - \mathbf{y}\|, \forall \mathbf{y} \in \mathcal{U}.$$

Proposição 7. Seja $\mathcal{U} \subset \mathbb{R}^n$ um conjunto não-vazio e $\{\mathbf{y}^k\}$ uma sequência Féjer convergente em \mathcal{U} . Então $\{\mathbf{y}^k\}$ é limitada. Além disso, se \mathbf{y} é um ponto de acumulação de $\{\mathbf{y}^k\}$ que pertence a \mathcal{U} , então $\{\mathbf{y}^k\}$ converge para \mathbf{y} .

Demonstração. Seja $\{\mathbf{y}^k\}$ uma sequência Féjer convergente, então por definição temos que

$$\|\mathbf{y}^{k+1} - \mathbf{y}\| \leq \|\mathbf{y}^k - \mathbf{y}\|, \forall \mathbf{y} \in \mathcal{U}.$$

Temos também que $\|\mathbf{y}^k - \mathbf{y}\| \leq \|\mathbf{y}^{k-1} - \mathbf{y}\|$. Prosseguindo com o raciocínio temos:

$$\|\mathbf{y}^{k+1} - \mathbf{y}\| \leq \|\mathbf{y}^k - \mathbf{y}\| \leq \|\mathbf{y}^{k-1} - \mathbf{y}\| \leq \dots \leq \|\mathbf{y}^0 - \mathbf{y}\|.$$

Assim temos que $\|\mathbf{y}^k - \mathbf{y}\| \leq \|\mathbf{y}^0 - \mathbf{y}\|$, $\forall k \in \mathbb{N}$. Isso nos diz que $\{\mathbf{y}^k\}$ está contida numa bola de centro \mathbf{y} e raio $\|\mathbf{y}^0 - \mathbf{y}\|$, e conseqüentemente tem-se que $\{\mathbf{y}^k\}$ é limitada. Agora seja \mathbf{y} um ponto de acumulação para $\{\mathbf{y}^k\}$ e seja $\{\mathbf{y}^{k_j}\}$ uma subsequência de $\{\mathbf{y}^k\}$ que converge para \mathbf{y} . Da definição de Féjer convergência temos que a sequência real $\|\mathbf{y}^k - \mathbf{y}\|$ é não-crescente. Como a subsequência $\|\mathbf{y}^{k_j} - \mathbf{y}\| \rightarrow 0$ temos pelo corolário do Teorema 4 de [14] página 111, que

$$\|\mathbf{y}^k - \mathbf{y}\| \rightarrow 0.$$

Isso significa que $\lim_{k \rightarrow \infty} \mathbf{y}^k = \mathbf{y}$. □

A seguir apresentamos um lema que será usado como ferramenta na demonstração do teorema posterior. Para mais detalhes ver [1].

Lema 10. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função semi contínua inferiormente e própria. Então são equivalentes:*

i) f é quase convexa;

ii) $\exists \mathbf{x}^* \in \partial f(\mathbf{x})$ tal que $\langle \mathbf{x}^*, \mathbf{y} - \mathbf{x} \rangle > 0 \Rightarrow f(\mathbf{z}) \leq f(\mathbf{y}), \forall \mathbf{z} \in [\mathbf{x}, \mathbf{y}]$.

Demonstração. Para uma demonstração ver [1] pág. 33-34 □

Teorema 14. *A sequência $\{\mathbf{x}^k\}$ converge para um ponto Pareto crítico de F .*

Demonstração. Inicialmente comecemos definindo o seguinte conjunto

$$E = \{\mathbf{x} \in D : F(\mathbf{x}) \preceq F(\mathbf{x}^k), \forall k \in \mathbb{N}\}$$

Uma vez que estamos supondo que a hipótese \mathcal{H} é válida temos que o conjunto E é não-vazio. Assim, tomamos um ponto arbitrário $\mathbf{x}^* \in E$, o qual por definição nos dá que

$$F(\mathbf{x}^*) \preceq F(\mathbf{x}^k).$$

Isso nos diz que \mathbf{x}^* também pertence a Ω_k para todo $k \in \mathbb{N}$. Em seguida denotamos por $\gamma_{k+1} = \lambda_k \langle \boldsymbol{\varepsilon}^k, \mathbf{u}^{k+1} \rangle$. Perceba que $\gamma_{k+1} > 0$, uma vez que $\lambda_k > 0, \boldsymbol{\varepsilon}^k \in \mathbb{R}_{++}^m$ e $\mathbf{u}^k \in \mathbb{R}_+^m \setminus \{0\}$ para todo $k \in \mathbb{N}$. Em seguida vejamos que :

$$\begin{aligned} \|\mathbf{x}^k - \mathbf{x}^*\|^2 &= \|(\mathbf{x}^k - \mathbf{x}^{k+1}) + (\mathbf{x}^{k+1} - \mathbf{x}^*)\|^2 \\ &= \langle (\mathbf{x}^k - \mathbf{x}^{k+1}) + (\mathbf{x}^{k+1} - \mathbf{x}^*), (\mathbf{x}^k - \mathbf{x}^{k+1}) + (\mathbf{x}^{k+1} - \mathbf{x}^*) \rangle \\ &= \|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2 + 2\langle \mathbf{x}^k - \mathbf{x}^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle + \|(\mathbf{x}^{k+1} - \mathbf{x}^*)\|^2 \end{aligned} \quad (4.3)$$

Da expressão (3.2) temos que

$$\mathbf{x}^k - \mathbf{x}^{k+1} = \frac{1}{\lambda_k \langle \boldsymbol{\varepsilon}^k, \mathbf{u}^{k+1} \rangle} (\mathbf{A}_{k+1}^\top (\mathbf{u}^{k+1} + \mathbf{v}^{k+1})) + \boldsymbol{\tau}_{k+1} \mathbf{w}^{k+1}$$

E conseqüentemente substituindo $(\mathbf{x}^k - \mathbf{x}^{k+1})$ na expressão anterior obtemos:

$$\begin{aligned} \|\mathbf{x}^k - \mathbf{x}^*\|^2 &= \|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2 + \|(\mathbf{x}^{k+1} - \mathbf{x}^*)\|^2 + \frac{2}{\gamma_{k+1}} \langle (\mathbf{A}_{k+1}^\top (\mathbf{u}^{k+1} + \mathbf{v}^{k+1})) \\ &+ \boldsymbol{\tau}_{k+1} \mathbf{w}^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle \\ &= \|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2 + \|(\mathbf{x}^{k+1} - \mathbf{x}^*)\|^2 \\ &+ \frac{2}{\gamma_{k+1}} \sum_{i=1}^m (\mathbf{u}_i^{k+1} + \mathbf{v}_i^{k+1}) \langle \mathbf{a}_i^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle + \boldsymbol{\tau}_{k+1} \langle \mathbf{w}^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle \end{aligned} \quad (4.4)$$

onde $\mathbf{a}_i^{k+1} \in \partial f_i(\mathbf{x}^{k+1}) \forall k \in \mathbb{N}$ e $i = 1, \dots, m$. Por outro lado, uma vez que F é \mathbb{R}^m -quase convexa temos do Lema 10 que $\langle \mathbf{a}_i^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle \geq 0$. Agora como $\gamma_k > 0, \forall k$, obtemos:

$$\frac{2}{\gamma_{k+1}} \sum_{i=1}^m (\mathbf{u}_i^{k+1} + \mathbf{v}_i^{k+1}) \langle \mathbf{a}_i^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle \geq 0. \quad (4.5)$$

Além disso, uma vez que $\mathbf{w}^{k+1} \in \mathcal{N}_D(\mathbf{x}^{k+1})$ e $\tau_k > 0$ segue que

$$\tau_{k+1} \langle \mathbf{w}^{k+1}, \mathbf{x}^{k+1} - \mathbf{x}^* \rangle \geq 0. \quad (4.6)$$

Usando as equações (4.5), (4.6) em (4.4) temos

$$\|\mathbf{x}^k - \mathbf{x}^*\|^2 \geq \|\mathbf{x}^k - \mathbf{x}^{k+1}\|^2 + \|(\mathbf{x}^{k+1} - \mathbf{x}^*)\|^2$$

e conseqüentemente

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \leq \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \|(\mathbf{x}^{k+1} - \mathbf{x}^*)\|^2, \forall k \in \mathbb{N}$$

e assim finalmente vem que

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}^k - \mathbf{x}^*\|^2, \forall \mathbf{x}^* \in E$$

Da expressão acima conclui-se que $\{\mathbf{x}^k\}$ é uma seqüência Féjer convergente em E . Agora mostraremos que o conjunto dos pontos de acumulação de $\{\mathbf{x}^k\}$ está contido em E . Com efeito, uma vez que $\{\mathbf{x}^k\}$ é Féjer convergente, temos da Proposição 7 que a mesma é limitada. Agora seja $\hat{\mathbf{x}}$ um ponto de acumulação de $\{\mathbf{x}^k\}$. Lembremos que da definição do Algoritmo que para $m > k$ temos

$$F(\mathbf{x}^m) \preceq F(\mathbf{x}^k).$$

Usando da continuidade de F e tomando o limite quando $m \rightarrow \infty$ temos que $F(\hat{\mathbf{x}}) \preceq F(\mathbf{x}^k), \forall k$ e assim temos que $\hat{\mathbf{x}} \in E$. Da Proposição 7 temos que a própria seqüência $\{\mathbf{x}^k\}$ converge para $\hat{\mathbf{x}}$. Como $\{\mathbf{x}^k\}$ converge para $\hat{\mathbf{x}}$. Usando a desigualdade triangular temos:

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq \|\mathbf{x}^{k+1} - \hat{\mathbf{x}}\| + \|\hat{\mathbf{x}} - \mathbf{x}^k\|, \forall k \in \mathbb{N}. \quad (4.7)$$

Usando que $\mathbf{x}^k \rightarrow \hat{\mathbf{x}}$, temos que

$$\lim_{k \rightarrow \infty} \|\mathbf{x}^{k+1} - \mathbf{x}^k\| = 0$$

Nos resta mostrar que \hat{x} é um ponto Pareto crítico para a função F . De fato, se \hat{x} é um ponto de acumulação para $\{x^k\}$, temos que $\hat{x} \in E$, e além disso da Proposição 7 temos que

$$x^k \rightarrow \hat{x}.$$

Uma vez que $\hat{x} \in E$, temos da Proposição 6 que existem $A_k \in \mathbb{R}^{m \times n}$, $u^k, v^k \in \mathbb{R}_+^m$, $w^k \in \mathbb{R}^m$ e $\tau_k \in \mathbb{R}_{++}$ tal que

$$A_k^T(u^k + v^k) + \lambda_{k-1} \langle \varepsilon^{k-1}, u^k \rangle (x^k - x^{k-1}) + \tau_k w^k = 0, \quad (4.8)$$

onde $w^k \in B[0, 1] \cap \mathcal{N}_D(\hat{x})$ e $\|u^k + v^k\|_1 = 1$, $\forall k \in \mathbb{N}$.

Tendo em mente que as sequências $\{A_k\}, \{u^k\}, \{v^k\}, \{w^k\}, \{\tau_k\}$ são limitadas e (a menos de uma subsequência) são convergentes para respectivamente $\hat{A}, \hat{u}, \hat{v}, \hat{w}$ e $\hat{\tau}$. Além disso, uma vez que a expressão $(\lambda_{k-1} \langle \varepsilon^{k-1}, u^k \rangle)$ é limitada, temos que tomando o limite com $k \rightarrow \infty$ obtemos de (4.8) que

$$\hat{A}^T \hat{y} + \hat{\tau} \hat{w} = 0, \text{ com } \hat{y} = \hat{u} + \hat{v}. \quad (4.9)$$

Assim finalmente do Lema 1, temos que \hat{x} é um ponto Pareto-crítico para F em Ω . \square

Capítulo 5

Considerações Finais

Nesse trabalho fizemos um estudo do Método do Ponto Proximal, bem como uma análise da convergência da sequência gerada no algoritmo munida das hipóteses de F ser localmente Lipschitz e posteriormente ser quase-convexa, onde nesse último caso foi adicionada a hipótese de completude \mathcal{H} , a qual pode ser removida desde que se suponha de antemão que $\{x^k\}$ possui pelo menos um ponto de acumulação. Em se tratando de trabalhos que futuramente poderão ser desenvolvidos, destaca-se a versão inexata do **Algoritmo** como a utilizada por exemplo em [5, 9] utilizando-se das mesmas técnicas de convergência abordadas nesse trabalho.

Referências Bibliográficas

- [1] Aussel, D., Subdifferential Properties of Quasiconvex and Pseudoconvex Functions: Unified Approach, *Journal of Mathematical Analysis and Applications*, Vol. 97, No. I, pp. 29-45, April 1998.
- [2] Bento, G.C., Cruz Neto, J.X., Lopez, G., Soubeyran, A., Souza, J.C.O., The proximal point method for locally Lipschitz functions in multiobjective optimization with application to the compromise problem.
- [3] Bolte, J., Daniilidis, A., Lewis, A., Shiotani, M., Clarke critical values of subanalytic Lipschitz continuous functions. *Ann. Polon. Math.*, 87(2005), pp13-25.
- [4] Bonnel, H., Iusem, A.N., Svaiter, B.F., Proximal methods in vector optimization. *SIAM J. Optim.*, 15 (2005), pp. 953-970.
- [5] Ceng, L.C., Yao, J.C., Approximate proximal methods in vector optimization. *Eur. J. Oper. Res.*, 183 (2007), pp. 1-19.
- [6] Clarke, F. H., A New Approach to Lagrange Multipliers, *Mathematics of and Applications*, Vol. 47, pp. 324-353, 1974.
- [7] Clarke, F. H., Generalized gradients and applications. *Trans. Amer. Math. Soc.*, 205 (1975), pp. 247-262.
- [8] Clarke, F. H., Optimization and nonsmooth analysis. *Classics in Applied Mathematics* 5, SIAM, Philadelphia, 1990.
- [9] Das. I., Dennis. J.E., Normal-boundary intersection: a new method for generating the pareto surface in nonlinear multicriteria optimization problems, *SIAM J. Optim.*, 8 (1998), pp. 631-657.

-
- [10] Ekeland, I., On the Variational Principle, *Journal of Mathematical Analysis and Applications*, Vol. 47, pp.324-353, 1974.
- [11] Huang, X.X., Yang, X.Q., Duality for multiobjective optimization via nonlinear Lagrangian functions. *J. Optim. Theory Appl.*, 120 (2004), pp. 111-12.
- [12] Izmailov, A., Solodov, M. *Otimização, Volume 1*, Rio de Janeiro: IMPA, (2009).
- [13] Lima, E.L. *Curso de Análise volume 2(décima primeira edição)*, Rio de Janeiro: IMPA,(2015).
- [14] Lima, E.L. *Curso de Análise volume 1(décima quarta edição)*, Rio de Janeiro: IMPA,(2013).
- [15] Luc, D.T., *Theory of Vector Optimization, Lecture Notes in Econom. and Math. Syst.* Springer-Verlag, New York, (1989).
- [16] Minami, M., Weak Pareto-Optimal Necessary Conditions in a Nondifferentiable Multiobjective Program on a Banach Space., *J. Optim. Theory Appl.*, 41 (1983), pp. 451-461.
- [17] Villacorta, K.D.V.,Oliveira, P.R., An interior proximal method in vector optimization. *Eur. J. Oper. Res.*, 214(2011), pp. 485-492.
- [18] www.cienciaestrategia.com.br/teoriadosjogos/capitulo.asp?cap=m6