Instituto Alberto Luiz Coimbra de
Pós-Graduação e Pesquisa de Engenharia
COPPE UFRJ

MÉTODO DE PONTO PROXIMAL PARA MINIMIZAÇÃO MULTIOBJETIVO QUASE-CONVEXA

Hellena Christina Fernandes Apolinário

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Paulo Roberto Oliveira

Rio de Janeiro
Julho de 2014

MÉTODO DE PONTO PROXIMAL PARA MINIMIZAÇÃO MULTIOBJETIVO QUASE-CONVEXA

Hellena Christina Fernandes Apolinário

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

____

Prof. Paulo Roberto Oliveira, D.Ing.

____

Prof. Erik Alex Papa Quiroz, D.Sc.

____

Prof. Ernesto Prado Lopes, D.Sc.

____

Prof. Jurandir de Oliveira Lopes, D.Sc.

____

Prof. Luis Román Lucambio Pérez, D.Sc.

____

Prof. Orizon Pereira Ferreira, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
JULHO DE 2014

Apolinário, Hellena Christina Fernandes

Método de ponto proximal para minimização multiobjetivo quase-convexa/Hellena Christina Fernandes Apolinário. – Rio de Janeiro: UFRJ/COPPE, 2014.

XI, 55 p.: il.; 29,7 cm

Orientador: Paulo Roberto Oliveira

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2014.

Referências Bibliográficas: p. 52 – 55.

1. Método de ponto proximal. 2. Minimização multiobjetivo. 3. Funções quase-convexas. 4. Fejér convergência. 5. Subdiferencial de Clarke. 6. Ponto Pareto-Clarke crítico. I. Oliveira, Paulo Roberto. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Ao meu pai,
Luiz Apolinário (in memoriam) ,
pelos seus ensinamentos e por
me incentivar a gostar da
matemática.
Aos meus filhos, Gabriela e
Rafael, pela paciência e por
estarem sempre junto de mim
nesta jornada.
À minha mãe, Elvira, pelo
carinho e ternura dedicados a
mim e aos meus filhos.

Agradecimentos

Agradeço a Deus e a Nossa Senhora pelo dom da vida e pelos momentos em que fui amparada nos Seus braços.

Aos meus filhos, Gabriela e Rafael, pelas mensagens de conforto e pela paciência de aguentar a minha ausência, mesmo que momentânea.

À minha mãe, Elvira, por assumir a responsabilidade dos meus filhos na minha ausência e pelo colo dado nos momentos difíceis.

Aos meus irmãos, Ricardo e Aninha, que sempre me apoiaram nesta caminhada, pelo amor e carinho.

Às minhas tias, Laude e Dindinha, pelo carinho, pelo apoio aos meus estudos desde quando eu estudava no Colégio Arquidiocesano Pio XII e por acreditarem junto comigo neste sonho.

Ao Marco Aurélio, pelo apoio tecnológico e pelas palavras de incentivo nos momentos de desânimo.

A toda minha família, Polly, Morena, Beth, Thierry, Yan, Carlinhos, Túlio, Tiago, Segundo, Maria Raquel, Sandra, Tarcísio, Ailton, Teca,...etc,... , pela torcida favorável.

Ao meu orientador, Paulo Roberto Oliveira, pela orientação, por acreditar no meu potencial e pela confiança.

Ao Erik Alex Papa Quiroz, pela co-orientação, pelas ideias de melhoria para a nossa tese, pela paciência, dedicação e sobretudo, pela disponibilidade em me ajudar nos momentos complicados desta jornada.

Aos demais professores da banca, Ernesto Prado Lopes, Jurandir de Oliveira Lopes, Luis Román Lucambio Pérez e Orizon Pereira Ferreira, pela leitura e contribuições referentes à tese.

Aos amigos professores que estiveram presente comigo, Rogério Azevedo Rocha, Sandra Regina e Warley Gramacho, pelas palavras de conforto e incentivo, pelo companherismo, pela ajuda com o Latex e pelo simples fato de ouvir as minhas ansiedades durante esta caminhada.

Aos demais colegas do Colegiado do Curso de Bacharelado em Ciência da Computação da UFT pelo considerável apoio, em especial ao prof. Andreas Kneip, pela coordenação operacional do DINTER-UFT/UFRJ.

Ao Prof. Nelson Maculan, que me ensinou com a sua humildade e sabedoria a dar um valor maior à profissão de ser professor e a esta arte de ensinar.

Aos professores do PESC/COPPE pelas contribuições na minha formação acadêmica.

Aos funcionários do PESC, em especial à Josefina Solange Silva Santos, pela dedicação e por estarem sempre prontos a nos ajudar.

À Universidade Federal do Tocantins-UFT pelo apoio intitucional.

À Coordenação de Aperfeiçoamento de Pessoal da Educação Superior (CAPES), pelo apoio financeiro ao Projeto DINTER-UFT/UFRJ.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

MÉTODO DE PONTO PROXIMAL PARA MINIMIZAÇÃO MULTIOBJETIVO QUASE-CONVEXA

Hellena Christina Fernandes Apolinário

Julho/2014

Orientador: Paulo Roberto Oliveira

Programa: Engenharia de Sistemas e Computação

Neste trabalho, propomos dois métodos de escalarização proximal para resolver problemas de minimização multiobjetivo, sem restrições, com funções vetoriais F quase-convexas. Consideramos em um dos métodos, o subdiferencial de Clarke, e no outro, o subdiferencial de Frechét.

Em ambos os métodos, os subproblemas a serem resolvidos em cada iteração foram regularizados pelo quadrado da distância euclidiana. No primeiro método, considerando F localmente Lipschitz e quase-convexa, provamos a convergência fraca e a convergência global da sequência gerada pelo método para pontos Pareto-Clarke críticos.

No segundo método, considerando F semicontínua inferior, estabelecemos a existência das iteradas. Analisamos a convergência deste método sob alguns aspectos: primeiro, se F além de semicontínua inferior, é quase-convexa, mostramos a convergência para soluções Pareto fracas quando a sequência dos parâmetros regularizadores converge para zero e as iterações minimizam a função regularizada. Se F é contínua e quase-convexa e os parâmetros regularizadores são limitados obtemos a convergência em um sentido generalizado. Posteriormente, considerando F continuamente diferenciável e quase-convexa, mostramos a convergência global da sequência gerada pelas versões, exata e inexata, para pontos Pareto críticos.

Estes dois métodos podem ser vistos como uma extensão para o caso não convexo, do método proximal inexato para problemas de minimização convexa multiobjetivo, estudado por Bonnel et al. (SIAM Journal 15, 4, 953-970, 2005).

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

PROXIMAL POINT METHOD FOR QUASICONVEX MULTIOBJECTIVE MINIMIZATION

Hellena Christina Fernandes Apolinário

July/2014

Advisor: Paulo Roberto Oliveira

Department: Systems Engineering and Computer Science

In this paper , we propose two proximal scalarization methods for solving multiobjective minimization problems without constraints with F a quasiconvex vector function. We consider in one of the methods, the Clarke subdifferential, and in the other, the Frechét subdifferential.

In both methods , the subproblems to be solved in each iteration were regularized by the square of the Euclidean distance . In the first method , considering F locally Lipschitz and quasiconvex, we prove the weak convergence and global convergence of the sequence generated by the method to Pareto-Clarke critical points.

In the second method, considering F lower semicontinuous, we have established the existence of iterations. We analyze the convergence of this method under some hypotheses: first, if F beyond lower semicontinuous, is quasi-convex, we show convergence to weak Pareto solutions when the sequence of regularization parameters converges to zero and each iteration minimizes the regularized function. If F is a continuous and quasiconvex mapping, and the regularization parameters are limited we obtain the convergence in a generalized sense. Subsequently, considering F continuously differentiable and quasiconvex, we show the global convergence of the sequence generated by the exact and inexact versions for Pareto critical points.

These methods may be seen as an extension to the nonconvex case , of the inexact proximal method for multiobjective convex minimization problems studied by Bonnel et al. (SIAM Journal 15, 4, 953-970, 2005).

Sumário

Notações

$\mathbb{N}$	Conjunto dos números naturais.
$\mathbb{R}$	Conjunto dos números reais.
$\mathbb{R}^m$	Espaço Euclidiano m-dimensional.
$\mathbb{R}^m_+$	Ortante não-negativo de $\mathbb{R}^m$.
$\mathbb{R}^m_{++}$	Interior de $\mathbb{R}^m_+$.
$\langle x, y\rangle$	Produto interno Euclidiano de x e y em $\mathbb{R}^m$.
$\|.\|$	Norma euclidiana.
δ_C	Função indicadora do conjunto C.
$\mathcal{N}_C(x)$	Cone normal no ponto x em relaçãoo ao conjunto C.
$\hat{\partial} h(x)$	Subdiferencial Fréchet da função h em x.
$\partial h(x)$	Subdiferencial-limite da função h em x.
$\hat{\partial}_\varepsilon h(x)$	ε-Subdiferencial Fréchet da função h em x.
$\partial^o h(x)$	Subdiferencial de Clarke de h em x.
$h^o(x,d)$	Derivada direcional de Clarke de h em x na direção d.

Introdução

Neste trabalho consideramos o seguinte problema de otimização multiobjetivo irrestrito:

$$\min\{F(x) : x \in \mathbb{R}^n\} \tag{1}$$

onde $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma função vetorial quase-convexa no espaço Euclidiano $\mathbb{R}^n$. As funções quase-convexas podem ser caracterizadas em termos da convexidade de seus conjuntos de nível. Esta classe de funções, que contém a classe das funções convexas, tem sido o ponto de partida para muitas pesquisas em convexidade generalizada. Uma motivação para o estudo deste problema é a teoria da demanda do consumidor na economia onde a quaseconvexidade do vetor de funções objetivo é uma condição natural associada à diversificação do consumo, ver Mas Colell et al. [31]. Outra motivação são as extensões de métodos bem conhecidos em otimização convexa para otimização quase-convexa. Mencionamos os seguintes trabalhos:

- Bello Cruz et al. [9] consideraram o método do gradiente projetado para resolver o problema de encontrar Pareto ótimo de uma função multiobjetivo quase-convexa. Provaram a convergência da seqüência gerada pelo algoritmo para um ponto fixo, e, quando as componentes da função multiobjetivo são pseudoconvexas, obtiveram a convergência para uma solução Pareto fraca.

- da Cruz Neto et al. [14] estenderam o método do subgradiente clássico em minimização de valor real para otimização multiobjetivo. Assumindo a quase-convexidade do vetor de funções objetivo, isto é, quaseconvexidade componente a componente, obtiveram a convergência completa da seqüência a uma solução Pareto.

- Papa Quiroz e Oliveira ([34],[35], [37]) estenderam a convergência do método do ponto proximal para problemas de minimização quase-convexa em variedades Riemannianas gerais as quais incluem o espaço euclidiano. Além disso, em [36], os autores estenderam a convergência do método ponto proximal para o ortante não negativo.

- Kiwiel [25]estendeu a convergência do método do subgradiente para resolver problemas de minimização quase-convexa em espaços de Hilbert.

- Brito et al. [10] propuseram um algoritmo proximal interior inspirado por um método proximal logaritimo-quadrático para problemas de minimização quase-convexa com restrições lineares. Para este método, provaram a convergencia global quando os parametros proximais convergem a zero. Esta última hipótese pode ser descartada quando a função é assumida pseudoconvexa.

- Langenberg e Tichatschke [26] estudaram o método proximal quando a função objetivo é quase-convexa, o problema está restrito a um conjunto convexo fechado arbitrário, e a regularização é uma distancia de Bregman. Assumindo que a função é localmente Lipschitz, e usando o subdiferencial de Clarke, os autores demonstraram a convergência global do método para um ponto crítico.

Neste trabalho estamos interessados em estender as propriedades de convergência do método do ponto proximal para resolver o problema multiobjetivo quaseconvexo (1).

O método do ponto proximal, introduzido por Martinet [29], para resolver o problema $\min\{f(x) : x \in \mathbb{R}^n\}$ onde f é uma função escalar, gera uma sequência $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, a partir de um processo iterativo iniciando com um ponto $x^0 \in \mathbb{R}^n$, arbitrário, e $x^{k+1} \in \operatorname{argmin}\{f(x) + \frac{\lambda_k}{2}\|x - x^k\|^2,\ x \in \mathbb{R}^n\}$, onde $\lambda_k > 0, \forall\ k \in \mathbb{N}$, é um parâmetro de regularização. É bem conhecido, ver Guler [19], que se f é convexa e $\{\lambda_k\}$ satisfaz $\sum\limits_{k=1}^{+\infty}(1/\lambda_k) = +\infty$, então $\lim_{k\to\infty} f(x^k) = \inf\{f(x) : x \in \mathbb{R}^n\}$. Além disso, se o conjunto das soluções ótimas é não vazio, obtemos que $\{x^k\}$ converge para uma solução ótima do problema.

Quando F é convexa em (1), Bonnel at al. [8] provaram a convergência do método do ponto proximal para uma solução Pareto fraca do problema (1) em um contexto geral, ver também Villacorta e Oliveira [44] utilizando distâncias proximais, e, Gregório e Oliveira [18] utilizando um método de escalarização proximal logaritmo quadrático.

A dificuldade encontrada de considerar algoritmos de pontos proximais de valor vetorial para resolver problemas de minimização multiobjetivo quase-convexa, está na impossibilidade de transformar os subproblemas a serem resolvidos em cada iteração em um problema ou em uma família de problemas escalares. Na otimização vetorial convexa esta dificuldade é contornada ao aplicar o Teorema 2.10 do Luc [27], pois este resultado garante uma escalarização dos subproblemas, e por conseguinte, pode-se utilizar os resultados já consolidados da otimização escalar.

Desta forma, introduzimos dois métodos de escalarização proximal para resolver o problema de minimização multiobjetivo quaseconvexo (1). No primeiro método utilizamos o subdiferencial de Clarke, e as iterações são as seguintes: dado $x^k \in \mathbb{R}^n$, encontrar $x^{k+1} \in \Omega_k$ tal que,

$$0 \in \partial^o \left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \ . \ - x^k \|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1})$$

onde ∂^o é o subdiferencial de Clarke, $\Omega_k = \left\{ x \in \mathbb{R}^n : F(x) \preceq F(x^k) \right\}$, $\alpha_k > 0$, $\{e_k\} \subset \mathbb{R}^m_{++}$, $\|e_k\| = 1$, $\{z_k\} \subset \mathbb{R}^m_+ \backslash \{0\}$, $\|z_k\| = 1$ e $\mathcal{N}_{\Omega_k}(x^{k+1})$ o cone normal em relação ao conjunto Ω_k em x^{k+1}. Provamos a boa definição da sequência gerada pelo método, e mostramos a convergência fraca e a convergência global para um ponto Pareto-Clarke crítico ao considerarmos funções vetoriais localmente Lipschitz e quase-convexa. Se o vetor de funções objetivo F é convexo, obtemos a convergência para uma solução Pareto fraca do problema. No segundo método utilizamos o subdiferencial de Fréchet, e as iterações são as seguintes: dado $x^k \in \mathbb{R}^n$, encontrar $x^{k+1} \in \Omega_k$ tal que,

$$0 \in \hat{\partial} \left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \ . \ - x^k \|^2 + \delta_{\Omega_k}(.) \right) (x^{k+1}),$$

onde $\hat{\partial}$ é o subdiferencial de Fréchet e δ_{Ω_k} é a função indicadora do conjunto Ω_k (ver Definição (1.1.1) . Provamos a boa definição da sequência gerada pelo método, e mostramos que esta sequência é Fejér convergente, quando a aplicação F é semi-contínua inferior e quase-convexa. Se as iterações são da forma,

$$x^{k+1} \in \arg\min \left\{ \langle F(x), z_k \rangle + \tfrac{\alpha_k}{2} \langle e_k, z_k \rangle \left\| x - x^k \right\|^2 : x \in \Omega_k \right\},$$

obtemos a convergência para solução Pareto fraca do problema quando $\{\alpha_k\}$ converge para zero. Quando a aplicação F é contínua e quase-convexa, mostramos a convergência do método em um sentido generalizado. Se o vetor de funções objetivo F assume a condição de ser continuamente diferenciável, então mostramos a convergência global para um ponto Pareto crítico, e ainda neste caso, apresentamos também uma versão inexata do método, cujas iterações são : dado $x^k \in \mathbb{R}^n$, encontrar $x^{k+1} \in \Omega_k$ tal que,

$$0 \in \hat{\partial}_{\varepsilon_k} \Psi_k(x) + \alpha_k \langle e_k, z_k \rangle \left(x - x^k \right) + \mathcal{N}_{\Omega_k}(x),$$

$$\lim_{k \to +\infty} \varepsilon_k = 0 \qquad .$$

onde $\hat{\partial}_{\varepsilon_k}$ é o ε_k-subdiferencial de Fréchet (ver Definição 1.1.15).
Para esta versão inexata, mostramos a convergência fraca para um ponto Pareto crítico do problema.

Esta tese está dividida em 4 capítulos. No Capítulo 1 apresentamos algumas notações, conceitos e resultados básicos, os quais serão úteis para o desenvolvimento do trabalho. No Capítulo 2 estudamos o método escalar ponto proximal com subdiferencial de Clarke, o qual denotamos por **MEPPC** para resolver problemas de minimização multiobjetivo quando o vetor de funções objetivo é quaseconvexo e localmente Lipschitz. Este capítulo gerou o artigo, [2], que foi submetido para apreciação e possível publicação na revista Journal of Global Optimization. No Capítulo 3 estudamos o método escalar ponto proximal com subdiferencial de Fréchet, o qual denotamos por **MEPPF** para resolver problemas de minimização multiobjetivo quando a função multiobjetivo é quase-convexa e semicontínua inferior, e neste caso provamos que a sequência gerada é Fejér convergente. Se a função multiobjetivo é contínua e quase-convexa, obtemos a convergência do método em um sentido generalizado. Se a função multiobjetivo é continuamente diferenciável, garantimos a convergência do método para pontos Pareto críticos. Apresentamos também uma versão inexata deste método, e provamos a convergência fraca para pontos Pareto críticos. Este capítulo é parte do artigo em andamento, [38] e o caso particular quando F é continuamente diferenciável e o algoritmo é dado por:

$$x^{k+1} \in \arg\min \left\{ \langle F(x), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \left\| x - x^k \right\|^2 : x \in \Omega_k \right\},$$

gerou o artigo [1], publicado pela SOBRAPO no ano de 2013.

Capítulo 1

Preliminares

Neste capítulo apresentamos alguns conceitos e resultados básicos que são de fundamental importância para o desenvolvimento do nosso trabalho. Estes fatos podem ser encontrados, por exemplo, em Hadjisavvas [20], Mordukhovich [33], e Rockafellar e Wets [41].

1.1 Noções básicas

Ao longo deste documento $\mathbb{R}^n$ denota o espaço euclidiano, isto é, um espaço vetorial real com o produto interno canônico $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$ e a norma dada por $||x|| = \sqrt{\langle x, x\rangle}$.

Definição 1.1.1 *Sejam $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$, e D um subconjunto do $\mathbb{R}^n$.*

1. *O domínio efetivo de f, denotado por* dom (f), *é definido por:*

$$\text{dom }(f) = \{x \in \mathbb{R}^n : f(x) < +\infty\}.$$

2. *A função f é própria se* dom $(f) \neq \emptyset$.

3. *A função f é coerciva se $\lim_{\|x\| \to +\infty} f(x) = +\infty$.*

4. *A função f é semicontínua inferior no ponto $\bar{x} \in \mathbb{R}^n$, quando para qualquer sequência $\{x_k\} \subset \mathbb{R}^n$ tal que $\lim_{k \to +\infty} x_k = \bar{x}$ tem-se $f(\bar{x}) \leq \liminf_{k \to +\infty} f(x_k)$.*

5. *O conjunto dos minimizadores da função f denotaremos por* $\text{argmin}\{f(x) : x \in \mathbb{R}^n\}$ *e o valor ótimo do problema* $\min \{f(x) : x \in \mathbb{R}^n\}$, *denotaremos por f^*.*

6. *A função indicadora $\delta_D : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ do conjunto D, é dada por:*

$$\delta_D(x) := \begin{cases} 0, & if\ x \in D \\ +\infty, & if\ x \notin D \end{cases}$$

Temos que δ_D é convexa, se D é convexo; δ_D é semicontínua inferior, se D é fechado, e δ_D é coerciva, se D é limitado.

Definição 1.1.2 *Seja $D \subset \mathbb{R}^n$ um conjunto convexo e $\bar{x} \in D$. O cone normal no ponto $\bar{x}$ em relação ao conjunto D é dado por*

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

Definição 1.1.3 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria e $D \subset dom(f)$ um conjunto aberto. A função f é Lipschitz em D se existe uma constante real $L \geq 0$ tal que, $|f(x) - f(y)| \leq L\|x - y\|$, $\forall x, y \in D$.*
Além disso, f é dita localmente Lipschitz em $\bar{x} \in D \subset dom(f)$ se existe $\varepsilon_{\bar{x}} > 0$ e $L_{\bar{x}} > 0$ tal que $|f(x) - f(y)| \leq L_{\bar{x}}\|x - y\|$, $\forall x, y \in B(\bar{x}, \varepsilon_{\bar{x}})$, onde $B(\bar{x}, \varepsilon_{\bar{x}}) = \{y \in \mathbb{R}^n : \|y - \bar{x}\| < \varepsilon_{\bar{x}}\}$. Finalmente, f é dita localmente Lipschitz em D, se f é localmente Lipschitz para cada $\bar{x} \in D$.

Definição 1.1.4 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Dizemos que $d \in \mathbb{R}^n$ é uma direção de descida de f no ponto $\bar{x} \in \mathbb{R}^n$, se exite $\varepsilon > 0$ tal que*

$$f(\bar{x} + td) < f(\bar{x}),\ \forall\ t \in (0, \varepsilon].$$

O próximo resultado garante que o conjunto dos minimizadores de uma função, sob algumas hipóteses, é não vazio.

Proposição 1.1.1 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior, coerciva e própria. Então o valor f^* é finito e o conjunto* argmin $\{f(x) : x \in \mathbb{R}^n\}$ *é não vazio e compacto.*

Prova: Conferir, Rockafellar e Wets ([41], Teorema 1.9). ■

Um resultado importante que envolve sequências de números não negativos é o seguinte:

Lema 1.1.1 *Sejam $\{w_k\}$, $\{p_k\}$ e $\{q_k\}$ sequências de números reais não-negativos. Se*

$$w_{k+1} \leq (1 + p_k)\, w_k + q_k, \qquad \sum_{i=1}^{\infty} p_k < \infty \ \text{ e } \ \sum_{i=1}^{\infty} q_k < \infty,$$

então a sequência $\{w_k\}$ é convergente.

Prova: Conferir Polyak ([40], Lema 2.2.2). ■

1.1.1 Otimização Multiobjetivo

Nesta subseção apresentamos algumas propriedades e notações em otimização multiobjetivo. Estes fatos básicos podem ser encotrados, por exemplo, em Miettinen [32].

Neste trabalho consideramos o cone $\mathbb{R}^m_+ = \{y \in \mathbb{R}^m : y_i \geq 0, \forall\ i = 1, ..., m\}$, o qual induz uma ordem parcial $\preceq$ em $\mathbb{R}^m$, dada por, para $y, y' \in \mathbb{R}^m$, $y \ \preceq\ y'$ se e somente se $y' \ - \ y \in \mathbb{R}^m_+$, isto significa que $y_i \leq \ y'_i$ para todo $i = 1, 2, ..., m$, com associada relação de ordem estrita $\prec$ induzida por $\mathbb{R}^m_{++} = \{y \in \mathbb{R}^m : y_i > 0, \forall\ i = 1, ..., m\}$, que representa o interior do cone, dada por $y \ \prec\ y'$, se e somente se $y' \ - \ y \in \mathbb{R}^m_{++}$, isto significa que $y_i < \ y'_i$ para todo $i = 1, 2, ..., m$. Esta ordem parcial estabelece uma classe de problemas conhecida na literatura como Otimização Multiobjetivo.

Vamos considerar o problema de otimização multiobjetivo (POM) irrestrito:

$$\min \{G(x) : x \in \mathbb{R}^n\} \tag{1.1}$$

onde $G : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, com $G = (G_1, G_2, ..., G_m)^T$.

Definição 1.1.5 (Miettinen [32], Definição 2.2.1) *Um ponto $x^* \in \mathbb{R}^n$ é uma solução Pareto para o problema* (1.1)*, se não existe $x \in \mathbb{R}^n$ tal que $G_i(x) \leq G_i(x^*)$, para todo $i \in \{1, ..., m\}$ e $G_j(x) < G_j(x^*)$, para algum $j \in \{1, ..., m\}$.*

Definição 1.1.6 (Miettinen [32], Definição 2.5.1) *Um ponto $x^* \in \mathbb{R}^n$ é uma solução Pareto fraca para o problema* (1.1)*, se não existe $x \in \mathbb{R}^n$ tal que $G_i(x) < G_i(x^*)$, para todo $i \in \{1, ..., m\}$.*

Denotaremos por $\text{argmin}\{G(x) : x \in \mathbb{R}^n\}$ e por $\text{argmin}_w \ \{G(x) : x \in \mathbb{R}^n\}$ o conjunto das soluções Pareto e Pareto fraca do problema (1.1), respectivamente. É fácil verificar que $\arg\min\{G(x) : x \in \mathbb{R}^n\} \subset \arg\min_w \ \{G(x) : x \in \mathbb{R}^n\}$.

No que segue assumimos que a função G é continuamente diferenciável. Para $x \in \mathbb{R}^n$, o jacobiano de G em x, denotado por $JG(x)$, é uma matriz de ordem $m \times n$ cujas entradas são definidas por $(JG(x))_{i,j} = \frac{\partial G_i}{\partial x_j}(x)$. E a imagem do jacobiano de G em x denotaremos por,

$$Im\,(JG\,(x)) := \{JG\,(x)\,v = (\langle \nabla G_1(x), v\rangle, \langle \nabla G_2(x), v\rangle, ..., \langle \nabla G_m(x), v\rangle) : v \in \mathbb{R}^n\}.$$

onde ∇G_i, com $i \in \{1, ..., m\}$, é o vetor gradiente de cada função componente de G.

No próximo resultado temos uma condição necessária de otimalidade de primeira ordem para o problema (1.1) de um ponto $x^* \in \mathbb{R}^n$.

Lema 1.1.2 *Seja $G : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, uma aplicação continuamente diferenciável e x^* uma solução Pareto para o problema* (1.1), *então*

$$Im\left(JG\left(x^*\right)\right) \cap \left(-\mathbb{R}^m_{++}\right) = \emptyset \tag{1.2}$$

Prova: Por hipótese x^* é solução Pareto do problema (1.1), então não existe $\bar{x} \in \mathbb{R}^n$ tal que $G_i(\bar{x}) \leq G_i(x^*)$, para todo $i \in \{1, ..., m\}$ e $G_j(\bar{x}) < G_j(x^*)$, para algum indice $j \in \{1, ..., m\}$. Suponha por contradição que $d \in Im\left(JG\left(x^*\right)\right) \cap \left(-\mathbb{R}^m_{++}\right)$, então existe $v \in \mathbb{R}^n$ tal que $\langle \nabla G_i(x^*), v\rangle < 0$, para todo $i \in \{1, ..., m\}$. Como G é continuamente diferenciável,

$$\lim_{t \to 0} \frac{G_i(x^* + tv) - G_i(x^*)}{t} = \langle \nabla G_i(x^*), v\rangle < 0, \ \forall \ i \in \{1, ..., m\}.$$

Isto implica dizer que, v é uma *direção de descida*, para cada função componente G_i, isto é, $\exists \ \ t_0 > 0$, tal que

$$G_i(x^* + tv) < G_i(x^*), \forall \ t \in (0, t_0], \forall \ i \in \{1, ..., m\}.$$

Portanto, v é uma *direção de descida* para G em x^*, (conferir, Fliege e Svaiter [15], Fliege et al. [16], Graña Drummond e Svaiter [17]), isto é, $\exists \ \ t_0 > 0$ tal que

$$G(x^* + tv) \prec G(x^*), \ \forall \ t \in (0, t_0].$$

Isto contradiz o fato que x^* é solução Pareto para o problema (1.1). ∎

Observação 1.1.1 *Observa-se que o Lema*1.1.2 *continua sendo válido quando x^* é solução Pareto fraca para o problema* (1.1). *Desta forma,* 1.2 *é uma condição necessária de otimalidade Pareto.*

Equivalentemente, de (1.2), $\forall \ v \in \mathbb{R}^n$, existe $i_0 = i_0(v) \in \{1, ..., m\}$ tal que

$$\langle \nabla G_{i_0}(x), v\rangle \geq 0.$$

Observe que a condição (1.2) generaliza para otimização multiobjetivo, a clássica condição, "gradiente igual a zero", para o caso de valor real.

Em geral, (1.2) é necessária, mas não é suficiente para otimalidade. Considere o exemplo a seguir.

Exemplo 1.1.1 *$G : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, tal que $G(x_1, x_2) = (x_1^3, x_2^3)$. Observe que:*

$$JG(x_1, x_2) = \begin{bmatrix} 3x_1^2 & 0 \\ 0 & 3x_2^2 \end{bmatrix}$$

e $Im\,(JG\,(x_1,x_2)) = \{(\langle(3x_1^2,0),v\rangle\,,\langle(0,3x_2^2),v\rangle)\,,\forall v\in\mathbb{R}^n\}$.

Observe que para $x^* = (0,0)$ *temos que* $Im\,(JG\,(x^*))\cap\left(-\mathbb{R}^m_{++}\right) = \emptyset$, $\forall v\in\mathbb{R}^n$. *Porém, se considerarmos* $\bar{x} = (-1,-1)$, *obtemos que* $G(\bar{x})\prec G(x^*)$. *Logo* x^* *não é solução Pareto fraca.*

A partir da condição (1.2), temos a definição a seguir.

Definição 1.1.7 *Um ponto* $x^*\in\mathbb{R}^n$ *que satisfaz* (1.2) *é chamado um ponto Pareto crítico.*

1.1.2 Escalarização

Nesta seção, com o objetivo de superar as dificuldades causadas pela não completude da ordem estabelecida no espaço objetivo R^m, apresentamos uma técnica útil na otimização multiobjetivo, a Escalarização. Esta técnica permite substituir o problema de otimização multiobjetivo original em um problema de otimização escalar, ou em uma família de problemas escalares e, assim, podemos usar os resultados existentes na literatura relacionados à otimização escalar.

Definição 1.1.8 *Uma função* $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ *é dita uma* **representação escalar estrita** *de uma função vetorial* $F:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ *quando dados* $x,\bar{x}\in\mathbb{R}^n$:

$$F(x)\preceq F(\bar{x})\Longrightarrow f(x)\leq f(\bar{x})\quad e\quad F(x)\prec F(\bar{x})\Longrightarrow f(x)<f(\bar{x}).$$

Além disso, dizemos que f *é uma* **representação escalar fraca** *de* F *se*

$$F(x)\prec F(\bar{x})\Longrightarrow f(x)<f(\bar{x}).$$

É fácil observar que toda representação escalar estrita também é uma representação escalar fraca. O seguinte resultado propõe uma forma de obter uma representação escalar estrita para funções $F\mathbb{R}^n\longrightarrow\mathbb{R}^m$.

Proposição 1.1.2 *Seja* $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ *uma função.* f *é uma representação escalar estrita de* $F:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ *se, e somente se, existe uma função estritamente crescente* $g:F\,(\mathbb{R}^n)\longrightarrow\mathbb{R}$ *tal que* $f=g\circ F$.

Prova: Conferir Luc ([27], Proposição 2.3). ∎

Exemplo 1.1.2 *Seja a aplicação* $F:\mathbb{R}^n\longrightarrow\mathbb{R}^m$. *A função* $f:\mathbb{R}^n\longrightarrow\mathbb{R}$, *definida por* $f(x)=\langle F(x),z\rangle$, *com* $z\in\mathbb{R}^m_+\backslash\{0\}$, *fixo, é uma representação escalar estrita de* F.

Com efeito, dados x, $\bar{x}\in\mathbb{R}^n$, *se* $F(x)\preceq F(\bar{x})$, *então* $F(\bar{x})-F(x)\in\mathbb{R}^m_+$, *isto é,* $F_i(\bar{x})-F_i(x)\geq 0$ *para todo* $i\in\{1,...,m\}$. *Segue que, para todo* $z\in\mathbb{R}^m_+\backslash\{0\}$,

$$[F_i(\bar{x}) - F_i(x)]z_i \geq 0, \; \forall\, i \in \{1, ..., m\},$$

Logo,

$$\langle F(\bar{x}) - F(x), z\rangle \geq 0, \; \forall\, i \in \{1, ..., m\},$$

ou, equivalentemente,

$$\langle F(x), z\rangle \leq \langle F(\bar{x}), z\rangle.$$

Além disso, se $F(x) \prec F(\bar{x})$*, então* $F(\bar{x}) - F(x) \in \mathbb{R}^m_{++}$*, isto é,* $F_i(\bar{x}) - F_i(x) > 0$ *para todo* $i \in \{1, ..., m\}$*. Assim, para todo* $z \in \mathbb{R}^m_+ \setminus \{0\}$*,*

$$\langle F(\bar{x}) - F(x), z\rangle > 0 \Longrightarrow \langle F(x), z\rangle < \langle F(\bar{x}), z\rangle.$$

Proposição 1.1.3 *Seja* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *uma representação escalar fraca da aplicação* $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *e* $\text{argmin}\,\{f(x) : x \in \mathbb{R}^n\}$ *o conjunto dos minimizadores de* f*. Então temos a inclusão:*

$$\text{arg min}\,\{f(x) : x \in \mathbb{R}^n\} \subseteq \text{arg min}_w\{F(x) : x \in \mathbb{R}^n\}.$$

Prova: Seja $\bar{x} \in \text{arg min}\,\{f(x) : x \in \mathbb{R}^n\}$, onde f é uma representação escalar fraca, arbitrária, de F. Agora suponha que $\bar{x} \notin \text{arg min}_w\{F(x) : x \in \mathbb{R}^n\}$, então existe $\widehat{x} \in \mathbb{R}^n$ tal que $F(\widehat{x}) \prec F(\bar{x})$. Segue que, para todo $z \in \mathbb{R}^m_+ \setminus \{0\}$,

$$\langle F(\widehat{x}), z\rangle < \langle F(\bar{x}), z\rangle. \tag{1.3}$$

Como $\langle F(.), z\rangle$ é uma representação escalar estrita, e portanto, representação escalar fraca de F, e ao considerarmos $f(x) = \langle F(x), z\rangle$, de (1.3), obtemos:

$$f(\widehat{x}) < f(\bar{x}),$$

o que nos leva a uma contradição pois $\bar{x} \in \text{arg min}\,\{f(x) : x \in \mathbb{R}^n\}$. ■

1.1.3 Funções Quase-Convexas

Nesta seção apresentamos o conceito e a caracterização de funções escalar e multiobjetivo: convexas e quase-convexas, e funções multiobjetivo localmente Lipschitz. Esta teoria pode ser encontrada em Bazaraa et al. [4], Luc [27], Mangasarian [28], Rockafellar e Wets [41], e suas referências.

Definição 1.1.9 *Seja* $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ *uma função própria. Então,* f *é convexa, se para todo* $x, y \in \mathbb{R}^n$*, e para todo* $t \in [0, 1]$*,*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Definição 1.1.10 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria. Então, f é quase-convexa se, para todo $x, y \in \mathbb{R}^n$, e para todo $t \in [0,1]$,*

$$f(tx + (1-t)y) \leq \max\{f(x), f(y)\}.$$

Quando esta desigualdade é estrita, isto é, $f(tx + (1-t)y) < \max\{f(x), f(y)\}$ e $t \in (0,1)$, para $x \neq y$, a função f é *estritamente quase-convexa.*

Observe que se f é uma função quase-convexa então $\text{dom}(f)$ é um conjunto convexo. Por outro lado, enquanto uma função convexa pode ser caracterizada pela convexidade do seu epígrafo, uma função quase-convexa pode ser caracterizada pela convexidade dos seus conjuntos de nível:

Teorema 1.1.1 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria. Então fé quase-convexa se, e somente se,*

$$S_\alpha = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$$

é convexo para cada número real α.

Prova: Conferir Bazaraa et al. ([4], Teorema 3.5.2). ∎

Teorema 1.1.2 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ uma função diferenciável em $\mathbb{R}^n$. Então f é quase-convexa se, e somente se, a seguinte afirmação é satisfeita:*

$$\text{se } f(x) \leq f(y) \text{ então } \langle \nabla f(y), x - y \rangle \leq 0.$$

Prova: Conferir Mangasarian ([28], Teorema 4). ∎

Nas definições a seguir são consideradas funções vetoriais, e a noção de convexidade, quase-convexidade e localmente Lipschitz para aquelas aplicações.

Definição 1.1.11 *Seja $F = (F_1, ..., F_m)^T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação, então F é $\mathbb{R}^m_+$-convexa, se e somente se toda função componente de F, $F_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, é convexa.*

Definição 1.1.12 *Seja $F = (F_1, ..., F_m)^T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação, então F é $\mathbb{R}^m_+$ - quase-convexa, se e somente se toda função componente de F, $F_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, é quase-convexa.*

Definição 1.1.13 *Seja $F = (F_1, ..., F_m)^T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação, então F é localmente Lipschitz em $\mathbb{R}^n$, se e somente se toda função componente de F, $F_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, é localmente Lipschitz em $\mathbb{R}^n$.*

1.1.4 Subdiferenciais Fréchet e Limite

A seguir recordamos algumas definições e resultados que tratam do cálculo subdiferencial.

Definição 1.1.14 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria.*

(a) Para cada $x \in \operatorname{dom} f$ o conjunto dos subgradientes regulares, também chamado de subdiferencial de Fréchet de f em x, denotado por $\hat{\partial} f(x)$, é o conjunto dos vetores $v \in \mathbb{R}^n$ tal que

$$f(y) \geq f(x) + \langle v, y - x \rangle + o(\|y - x\|), \text{ onde } \lim_{y \to x} \tfrac{o(\|y-x\|)}{\|y-x\|} = 0.$$

Ou equivalentemente,

$$\hat{\partial} f(x) := \left\{ v \in \mathbb{R}^n : \liminf_{y \neq x,\ y \to x} \frac{f(y) - f(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

Se $x \notin \operatorname{dom}(f)$ então $\hat{\partial} f(x) = \emptyset$.

(b) O conjunto dos subgradientes generalizados, também chamado de subdiferencial limite de f em $x \in \mathbb{R}^n$, denotado por $\partial f(x)$, é definido como segue:

$$\partial f(x) := \left\{ v \in \mathbb{R}^n : \exists\ x_n \to x, \quad f(x_n) \to f(x), \quad v_n \in \hat{\partial} f(x_n) \text{ com } v_n \to v \right\}.$$

Proposição 1.1.4 *Para uma função $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ e um ponto $\bar{x} \in \operatorname{dom}(f)$, os conjuntos $\partial f(\bar{x})$ e $\hat{\partial} f(\bar{x})$ são fechados, com $\hat{\partial} f(\bar{x})$ convexo e $\hat{\partial} f\,(\bar{x}) \subset \partial f\,(\bar{x})$.*

Prova: Conferir, Rockafellar e Wets ([41], Teorema 8.6). ■

Os subdiferenciais Fréchet e limite generalizam o subdiferencial de funções convexas. Segue o resultado.

Proposição 1.1.5 *Sejam $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria, convexa e $x \in dom(f)$. Então*

$$\hat{\partial} f(x) = \partial f(x) = \{ y \in \mathbb{R}^n : f(z) \geq f(x) + \langle y, z - x \rangle,\ \forall\ z \in \mathbb{R}^n \}.$$

Prova: Conferir Rockafellar e Wets ([41], Proposição 8.12). ■

Além disso, quando f é convexa, a *derivada direcional* da f em x na direção d é dada por:

$$f'(x, d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

A Proposição a seguir é conhecida como regra generalizada de Fermat, e trata da condição necessária de otimalidade. Além disto, em um problema de otimização em que a função e as restrições do problema são convexas, então esta condição torna-se também suficiente.

Proposição 1.1.6 *Se uma função própria* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *tem um mínimo local em* $\bar{x} \in dom(f)$, *então* $0 \in \hat{\partial} f\,(\bar{x})$ *e portanto* $0 \in \partial f(\bar{x})$.

Prova: Conferir Rockafellar e Wets ([41], Teorema 10.1). ■

Observação 1.1.2 *Seja* $C \subset \mathbb{R}^n$ *e* $f : C \to \mathbb{R} \cup \{+\infty\}$. *A minimização de* f *sobre o conjunto* C *é um caso especial da condição de otimalidade da regra generalizada de Fermat, na Proposição* (1.1.6). *Isto implica que se* f *possui um mínimo local em* $\bar{x} \in C$, *então* $0 \in \hat{\partial}(f + \delta_C)(\bar{x})$, *onde* δ_C *é a função indicadora do conjunto* C.

Proposição 1.1.7 *Seja* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *uma função própria. Então, as seguintes propriedades são verdadeiras:*

(i) Se f *é diferenciável em* $\bar{x}$ *então* $\hat{\partial} f(\bar{x}) = \{\nabla f(\bar{x})\}$, *e portanto* $\nabla f(\bar{x}) \in \partial f(\bar{x})$;

(ii) Se f *é continuamente diferenciável em uma vizinhança de* x, *então* $\hat{\partial} f(x) = \partial f(x) = \{\nabla f(x)\}$;

(iii) Se $g = f + h$ *com* f *finita em* $\bar{x}$ *e* h *continuamente diferenciável em uma vizinhança de* $\bar{x}$ *então* $\hat{\partial} g(\bar{x}) = \hat{\partial} f(\bar{x}) + \nabla h(\bar{x})$ *e* $\partial g(\bar{x}) = \partial f(\bar{x}) + \nabla h(\bar{x})$.

Prova: Conferir Rockafellar e Wets ([41], Exercício 8.8, pp 304). ■

O subdiferencial limite possui importantes propriedades, e uma delas, de grande importância para o nosso trabalho, trata do subdiferencial da soma. Segue o resultado.

Proposição 1.1.8 *Sejam* $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *funções próprias tais que* f *é localmente Lipschitz em* $\bar{x} \in dom(f) \cap dom(g)$ *e* g *é semicontínua inferior neste ponto. Então,*

$$\partial(f + g)(\bar{x}) \subset \partial f(\bar{x}) + \partial g(\bar{x}).$$

Prova: Conferir Mordukhovich ([33] Theorem 2.33). ■

1.1.5 ε-Subdiferencial

Apresentamos alguns conceitos e resultados importantes sobre ε-subdiferencial. A teoria sobre estes fatos pode ser encontrada em Jofré et al.[23], Rockafellar e Wets [41].

Definição 1.1.15 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função semicontínua inferior, própria e ε um número real dado não negativo. O ε-subdiferencial de Fréchet de f em $x \in \mathrm{dom}(f)$ é definido por:*

$$\hat{\partial}_\varepsilon f(x) := \left\{ x^* \in \mathbb{R}^n : \liminf_{\|h\| \to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\|h\|} \geq -\varepsilon \right\} \tag{1.4}$$

Observação 1.1.3 *Quando $\varepsilon = 0$,* (1.4) *reduz ao subdiferencial de Fréchet,o qual é denotado por $\hat{\partial} f(x)$, de acordo com a* Definição 1.1.14. *Mais precisamente,*

$$\begin{gathered} x^* \in \hat{\partial} f(x), \textit{ se e somente se, para cada } \eta > 0 \textit{ existe } \delta > 0 \textit{ tal que} \\ \langle x^*, y - x\rangle \leq f(y) - f(x) + \eta\|y - x\|, \textit{ para todo } y \in x + \delta \mathrm{B}, \end{gathered}$$

onde B representa a bola unitária fechada em $\mathbb{R}^n$.

Assim $\hat{\partial} f(x) = \hat{\partial}_0 f(x) \subset \hat{\partial}_\varepsilon f(x)$, para toda função $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ própria, semicontínua inferior, para todo $x \in \mathrm{Dom} f$ e $\varepsilon \in \mathbb{R}_+$.

Como observado por Treiman [43]:

$$x^* \in \hat{\partial}_\epsilon f(x) \Leftrightarrow x^* \in \hat{\partial}(f + \epsilon\|. - x\|)(x).$$

Equivalentemente, isto significa dizer que $x^* \in \hat{\partial}_\epsilon f(x)$, se e somente se, para cada $\eta > 0$, existe $\delta > 0$ tal que

$$\langle x^*, y - x\rangle \leq f(y) - f(x) + (\epsilon + \eta)\|y - x\|, \text{ para todo } y \in x + \delta \mathrm{B}.$$

A seguir definiremos um outro tipo de subdiferencial aproximado.

Definição 1.1.16 *O ε-subdiferencial limite de Fréchet de f em $x \in \mathrm{dom}(f)$ é definido por*

$$\partial_\varepsilon f(x) := \limsup_{y \xrightarrow{f} x} \hat{\partial}_\varepsilon f(y) \tag{1.5}$$

onde $\limsup_{y \xrightarrow{f} x} \hat{\partial}_\varepsilon f(y) := \{x^* \in \mathbb{R}^n : \exists\ x_n^* \xrightarrow{f} x, x_n^* \longrightarrow x^* \text{com } x_n^* \in \hat{\partial}_\varepsilon f(x_n) \forall n \in \mathbb{N}\}$

No caso em que f é continuamente diferenciável, o ε-subdiferencial limite de Fréchet assume um forma muito simples, de acordo com a proposição a seguir, a qual pode ser encontrada em Jofré et al.[23].

Proposição 1.1.9 *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função continuamente diferenciável em x com derivada $Df(x) = \nabla f(x)$. Então:*

$$\partial_\varepsilon f(x) = \nabla f(x) + \varepsilon B.$$

Prova: Conferir Jofré et al.([23], Proposição 2.8). ■

1.1.6 Derivada direcional e Subdiferencial de Clarke

Definição 1.1.17 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria e localmente Lipschitz em $x \in \mathrm{dom}(f)$ e $d \in \mathbb{R}^n$. A derivada direcional de Clarke de f em x na direção d, denotada por $f^o(x, d)$, é definida por*

$$f^o(x, d) = \lim_{t \downarrow 0} \sup_{y \to x} \frac{f(y + td) - f(y)}{t}$$

e o subdiferencial de Clarke de f em x, denotado por $\partial^o f(x)$, é definido por

$$\partial^o f(x) = \{w \in \mathbb{R}^n : \langle w, d\rangle \leq f^o(x, d), \forall\ d \in \mathbb{R}^n\}.$$

Observação 1.1.4 *Segue diretamente das definições acima, que para todo $x \in \mathbb{R}^n$, temos que: $\hat{\partial} f(x) \subset \partial f(x) \subset \partial^o f(x)$.(Conferir Bolte et al. [7], (7)).*

Lema 1.1.3 *Sejam $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ funções localmente Lipschitz em $x \in \mathbb{R}^n$. Então, $\forall d \in \mathbb{R}^n$:*

(i) $(f + g)^o (x, d) \leq f^o (x, d) + g^o (x, d)$;

(ii) $(\lambda f)^o (x, d) = \lambda (f^o(x, d))$, $\forall \lambda \geq 0$;

(iii) $f^o (x, \lambda d) = \lambda f^o(x, d)$, $\forall \lambda > 0$.

Prova: A prova é imediata da definição da derivada direcional de Clarke de f em x. ■

Lema 1.1.4 *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ uma função localmente Lipschitz em x e λ um escalar arbitŕario. Então*

$$\partial^o (\lambda f) (x) \subset \lambda \partial^o f(x)$$

Prova: Conferir Clarke ([12], Proposição 2.3.1). ∎

Lema 1.1.5 *Sejam $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ funções próprias, localmente Lipschitz em x, então*

$$\partial^o \left(\sum_{i=1}^{n} f_i \right)(x) \subset \sum_{i=1}^{n} \partial^o f_i(x)$$

Prova: Conferir Clarke ([12], Proposição 2.3.3). ∎

Proposição 1.1.10 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria e localmente Lipschitz em $\mathbb{R}^n$. Então f^o é semicontínua superior, isto é, se $\{(x^k, d^k)\}$ é uma sequência em $\mathbb{R}^n$ tal que $\lim\limits_{k\to+\infty}(x^k, d^k) = (x, d)$, então $\limsup\limits_{k\to+\infty} f^o(x^k, d^k) \leq f^o(x, d)$.*

Prova: Conferir Clarke ([12], Proposição 2.1.1, (b)). ∎

A Proposição que segue será útil na demonstração da Féjer convergência da sequência gerada pelo nosso método.

Proposição 1.1.11 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ uma função localmente Lipschitz em $\mathbb{R}^n$ e quase-convexa. Se $g \in \partial^o f(x)$, tal que $\langle g, \tilde{x} - x\rangle > 0$ então, $f(x) \leq f(\tilde{x})$.*

Prova: Conferir Aussel ([3], Teorema 2.1). ∎

Proposição 1.1.12 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ uma função convexa. Então $\partial^o f(x)$ coincide com o subdiferencial em x para funções convexas, e $f^o(x, d)$ coincide com a derivada direcional $f'(x, d)$ para cada d.*

Prova: Conferir Clarke ([12], Proposição 2.2.7) ∎

1.1.7 Pareto-Clarke crítico

Na Seção 1.1.1 introduzimos a definição de ponto Pareto crítico para aplicações continuamente diferenciáveis. Agora, com a teoria do subdiferencial de Clarke, introduziremos a definição de ponto Pareto-Clarke crítico para aplicações localmente Lipschitz em $\mathbb{R}^n$, que representa uma importante definição em nosso trabalho.

Definição 1.1.18 (Custódio et al. ([13], Definição 4.6)) *Seja $F = (F_1, ..., F_m)^T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação localmente Lipschitz em $\mathbb{R}^n$. Dizemos que $x^* \in \mathbb{R}^n$ é um ponto Pareto-Clarke crítico de F, se para todas as direções $d \in \mathbb{R}^n$, existe $i_0 = i_0(d) \in \{1, ..., m\}$ tal que $F^o_{i_o}(x^*, d) \geq 0$.*

A Definição 1.1.18 diz essencialmente que não existe direção em $\mathbb{R}^n$ que seja direção de descida para todas as funções objetivo (conferir, por exemplo, (Custódio et al. [13]). Se um ponto é Pareto ótimo (local ou global), então é necessariamente um ponto Pareto-Clarke crítico.

Observação 1.1.5 *Segue da definição anterior que se um ponto x não é Pareto-Clarke crítico, então existe uma direção $d \in \mathbb{R}^n$ satisfazendo*

$$F_i^o(x,d) < 0, \forall\ i \in \{1,...,m\}$$

Isto implica que, para cada $i \in \{1,...,m\}$, d é uma direção de descida, para cada função F_i, i.e, existe $\varepsilon > 0$, tal que

$$F_i(x+td) < F_i(x), \forall\ t \in (0,\varepsilon], \forall\ i \in \{1,...,m\}.$$

Este fato é bem conhecido, que d é uma direção de descida para a aplicação multi-objetivo F em x, isto é, $\exists\ \varepsilon > 0$ tal que

$$F(x+td) \prec F(x),\ \forall\ t \in (0,\varepsilon].$$

Proposição 1.1.13 *Seja $\bar{x}$ um ponto Pareto-Clarke crítico de uma aplicação localmente Lipschitz $G : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Se G é $\mathbb{R}^m_+$-convexa, então $\bar{x}$ é uma solução Pareto fraca do problema (1.1).*

Prova: Como $\bar{x}$ é uma ponto Pareto-Clarke crítico de G, então existe $i_0 = i_0(d) \in \{1,...,m\}$ tal que $G^o_{i_o}(\bar{x},d) \geq 0, \forall d$. Agora, visto que G é $\mathbb{R}^m_+-$ convexa então a derivada direcional de Clarke coincide com a derivada direcional para funções convexas, ver Proposição 1.1.12, isto é, podemos escrever:

$$G^o_{i_o}(\bar{x},d) = G'_{i_o}(\bar{x},d) \geq 0, \forall d \in \mathbb{R}^n \tag{1.6}$$

onde $G'_{i_0}(\bar{x},d)$ é a derivada direcional da função convexa G_{i_0} em $\bar{x}$ na direção d. Por outro lado, suponha por contradição que $\bar{x}$ não é solução Pareto fraca do problema (1.1), então existe $x^* \in \mathbb{R}^n$ tal que

$$G(x^*) \prec G(\bar{x}),\ \text{i.e,}\ G_i(x^*) < G_i(\bar{x}), \forall i \in 1,...,m.$$

Portanto existe $\alpha > 0$ tal que $G_{i_0}(x^*) = G_{i_0}(\bar{x}) - \alpha$. Defina $x_\lambda = \lambda x^* + (1-\lambda)\bar{x}$, $\lambda \in (0,1)$. Da $\mathbb{R}^m_+$-convexidade da G temos

$$G_{i_0}(x_\lambda) = G_{i_0}(\lambda x^* + (1-\lambda)\bar{x}) \leq \lambda G_{i_0}(x^*) + (1-\lambda)G_{i_0}(\bar{x}) = -\alpha\lambda + G_{i_0}(\bar{x})$$

Segue que

$$\frac{G_{i_0}(\bar{x}+\lambda(x^*-\bar{x}))-G_{i_0}(\bar{x})}{\lambda} \leq -\alpha < 0, \, \forall \lambda \in (0,1).$$

Tomando $d = x^* - \bar{x} \in \mathbb{R}^n$ e o limite quando λ converge a zero na desigualdade acima, obtemos uma contradição com (1.6). Portanto $\bar{x}$ é uma solução Pareto fraca do problema (1.1). ■

1.1.8 Fejér convergência

Nesta seção apresentamos alguns conceitos e resultados em Fejér convergência, os quais podem ser encontrados, por exemplo, em Schott [42].

Definição 1.1.19 *Uma sequência* $\{y_k\} \subset \mathbb{R}^n$ *é dita Fejér convergente a um conjunto* $U \subseteq \mathbb{R}^n$, $U \neq \emptyset$, *com respeito à norma Euclidiana, se,* $\|y_{k+1} - u\| \leq \|y_k - u\|$, $\forall \ k \in \mathbb{N}$, $\forall \ u \in U$.

O seguinte resultado em Féjer convergência é bem conhecido. E a vantagem é que, ao demonstrarmos que a sequência gerada pelo nosso algoritmo é Féjer convergente, então garantimos a sua limitação, e consequentemente, a existência de pontos de acumulação.

Lema 1.1.6 *Se* $\{y_k\} \subset \mathbb{R}^n$ *é Fejér convergente a um conjunto* $U \neq \emptyset$, *então:*

(i) A sequência $\{y_k\}$ *é limitada.*

(ii) Se um ponto de acumulação y *de* $\{y_k\}$ *pertence ao conjunto* U, *então* $\lim_{k \to +\infty} y_k = y$.

Prova: Conferir Schott ([42], Teorema 2.7). ■

1.2 Modelos Matemáticos de Otimização Quase-convexa Multiobjetivo

Nesta seção apresentamos dois modelos matemáticos, os quais advém da Economia e teoria de localização, e que representam na prática problemas multiobjetivos.

Exemplo 1.2.1 *(Teoria da demanda) Consideramos um número finito* ***n*** *de bens de consumo. Um consumidor é um agente que deve escolher quanto consumir de cada bem. Um conjunto ordenado de números representando as quantidades consumidas de cada bem é chamado de vetor de consumo, e denotado por $x = (x_1, x_2, ..., x_n)$ onde x_i, com $i = 1, 2, ..., n$, representa a quantidade consumida do bem* ***i****. Denotaremos por X o conjunto viável destes vetores, o qual será chamado de conjunto de consumo, e como se trata de quantidades, então $X \subset \mathbb{R}^n_+$. Assim, na clássica abordagem da demanda do consumidor, a análise do comportamento do consumidor inicia-se especificando a relação de preferência, denotada por $\succeq$, sobre o conjunto X. A relação "$x \succeq y$" significa que "x é ao menos tão bom quanto y "ou "y não é preferido a x ". O consumidor escolhe, de acordo com suas preferências, a combinação de bens(vetor de consumo) que lhe proporciona o maior nível de satisfação. Esta relação de preferência $\succeq$ é assumida racional, isto é, é completa, pois o consumidor é capaz de ordenar todas as combinações possíveis de bens e/ou serviços, e transitiva, pois as preferências do consumidor são consistentes, o que significa se o consumidor prefere A a B, e B a C, então ele prefere A a C (conferir definição 3.B.1 de Mas-Colell et al. [31]).*
Uma função $\mu : X \longrightarrow \mathbb{R}$ é dita uma função de utilidade que representa uma relação de preferência em X, se a seguinte condição é satisfeita:

$$x \succeq y \Longleftrightarrow \mu(x) \geq \mu(y) \tag{1.7}$$

A função utilidade é a forma que se tem de representar as preferências entre dois vetores de consumo, o que tiver o valor da função utilidade mais elevada é o preferido. Se tiverem o mesmo valor da função utilidade, então o consumidor é indiferente. Além disto, se tivermos várias relações de preferências(múltiplos critérios), que satisfazem a condição 1.7, então teríamos uma função utilidade para cada uma destas preferências.

Observe que nem sempre existe uma função utilidade para uma preferência. Para isto, basta definir em $X = \mathbb{R}^2$ uma relação lexicográfica, dada por:
para $x, y \in \mathbb{R}^2$, $x \succeq y$ se, e somente se, "$x_1 > y_1$" ou "$x_1 = y_1$ e $x_2 \geq y_2$". Felizmente, uma classe mais geral de relação de preferência pode ser representada por funções utilidades, ver por exemplo Proposição 3.C.1 de Mas-Colell et al. [31].

Se uma relação de preferência $\succeq$ é representada por uma função utilidade μ, *então o problema de maximizar a preferência do consumidor em X é equivalente a resolver o problema de otimização*

$$\text{(P)} \max\{\mu(x) : x \in X\}$$

Agora, considere múltiplos critérios, isto é, considere m relações de preferências $\succeq$. Para cada uma das m preferências, sempre existe uma função utilidade, $\mu_1, \mu_2, ..., \mu_m$, então o problema de maximizar a preferência do consumidor em X é equivalente a resolver o problema de otimização multiobjetivo :

$$\text{(P')} \max\{(\mu_1(x), \mu_2(x), ..., \mu_m(x)) \in \mathbb{R}^m : x \in X\}$$

Por outro lado, um pressuposto natural da economia é que o consumidor tende a diversificar o seu consumo entre todos os produtos, isto é, a preferência $\succeq$ satisfaz a seguinte propriedade da convexidade: X é convexo e se $x \succeq z$ e $y \succeq z$ então $\lambda x + (1-\lambda)y \succeq z$, $\forall \lambda \in [0,1]$.

Se uma relação de preferência, $\succeq$, pode ser representada por uma função utilidade $\mu : X \longrightarrow \mathbb{R}$, então a propriedade da convexidade de $\succeq$ é equivalente a quaseconcavidade da função utilidade μ. De fato, pelo fato de μ representar a preferência $\succeq$, para quaisquer $x, y \in X$ tal que $x \succeq y$ temos que $\mu(x) \geq \mu(y)$. Pela convexidade de $\succeq$, para $0 \leq \alpha \leq 1$ temos $\alpha x + (1-\alpha)y \succeq y$. Assim, $\mu(\alpha x + (1-\alpha)y) \geq \mu(y) = min(\mu(x), \mu(y))$ logo μ é quase-côncava.

Portanto (P') *torna-se um problema de maximização com função multiobjetivo quase-côncava, pois cada função componente é quase-côncava. Considerando $F = (-\mu_1, -\mu_2, ..., -\mu_m)$, obtemos um problema de minimização multiobjetivo quase-convexa , pois cada função componente é quase-convexa.*

Existem várias classe de funções utilidades que são frequentemente utilizadas para gerar funções de demanda. Uma das mais comuns é a função utilidade de Cobb-Douglas, a qual é definida em $\mathbb{R}^2$ por: $\mu(x_1, x_2) = kx_1^{\alpha} x_2^{1-\alpha}$, com $\alpha \in (0,1)$ e $k > 0$. Outra forma comum de utilidade, é a função utilidade CES (Constant Elasticity of Substitution), definida em $\mathbb{R}^2$ por: $\mu(x_1, x_2) = (\lambda_1 x_1^{\rho} + \lambda_2 x_2^{\rho})^{1/\rho}$, onde $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$, e ρ é uma constante.

Para melhor entender o modelo anteriormente descrito, considere uma situação em que o consumidor, o qual é proprietário de uma panificadora, compra quatro tipos de queijos, q_1, q_2, q_3 e q_4, os quais são utilizados para fabricação de pães, biscoitos e outros produtos de sua panificadora. Seja, x_1 a quantidade, em quilogramas (kg), comprada do queijo q_1, x_2 a quantidade, em quilogramas (kg), comprada do queijo q_2, e assim por diante. Assim, o vetor consumo é $x = (x_1, x_2, x_3, x_4)$. O consumidor utiliza-se de alguns critérios para efetuar a compra dos queijos. Vamos considerar os seguintes critérios(preferências):

1. *Preferência 1: Preço*

2. *Preferência 2: Qualidade do queijo*

3. *Preferência 3: Sabor*

4. *Preferência 4: Teor de sódio*

5. *Preferência 5: Teor de gordura e colesterol*

Cada preferência define uma função utilidade μ, assim teremos as funções utilidades μ_1, μ_2, μ_3, μ_4 e μ_5, as quais representam numericamente o julgamento do consumidor. Portanto o problema de maximizar a preferência do consumidor em $X \subset \mathbb{R}^4_+$ é equivalente a resolver o problema de otimização multiobjetivo :

$$\text{(POM)} \max\{(\mu_1(x), \mu_2(x), \mu_3(x), \mu_4(x), \mu_5(x)) \in \mathbb{R}^5 : x \in X\}$$

Tomando $F = (-\mu_1, -\mu_2, -\mu_3, -\mu_4, -\mu_5)$, obtemos um problema de minimização multiobjetivo quase-convexa , pois cada função componente é quase-convexa, no ortante não negativo.

Exemplo 1.2.2 *(Teoria da localização) Problemas de localização estão relacionados com a determinação da localização para uma ou várias instalações, considerando um determinado conjunto de pontos de demanda, com os quais as interações devem ser estabelecidas. Estes termos não fazem parte de uma terminologia padrão, as vezes são substituídos por: clientes, instalações existentes, usuários ou comércios.*

Considere o problema de localização de uma instalação, e para cada $i = 1, ..., m$, considere o conjunto D_i de p pontos de demanda distintos, dados por $D_i = \{d^i_1, d^i_2, ..., d^i_p\} \subset \mathbb{R}^n$, $n \geq 2$. Necessitamos encontrar um local ideal(ótimo) $x \in \mathbb{R}^n$ para uma instalação de tal modo que este local minimize alguma função real envolvendo a distância entre este novo local e cada um dos conjuntos de pontos de demanda. Para cada $i = 1, ..., m$, se C^i_j, $j = 1, ..., p$ são conjuntos compactos convexos com $0 \in \operatorname{int}(C^i_j)$, onde $\operatorname{int}(C^i_j)$, denota o interior de C^i_j, então, para cada $i = 1, ..., m$, definimos a distância entre x e d^i_j por $\gamma_{C^i_j}(x - d^i_j)$, com $\gamma_{C^i_j}$ o indicador ou funcional de Minkowsky do conjunto C^i_j, isto é, $\gamma_{C^i_j}(x) = \inf\{t > 0 : x \in tC^i_j\}$. Observe que se C^i_j é a bola unitária em $\mathbb{R}^n$, então $\gamma_{C^i_j}(x)$ é a distância Euclidiana de x até 0.

Para introduzir o modelo, considere, para cada $i = 1, ..., m$, a função $\gamma_i : \mathbb{R}^n \longrightarrow \mathbb{R}^p_+$, dada por $\gamma_i(x) = (\gamma_{C^i_j}(x - d^i_1), ..., \gamma_{C^i_p}(x - d^i_p))$. Suponha que as funções $f_j : \mathbb{R}^p_+ \longrightarrow \mathbb{R}_+$, com $j = 1, ..., p$ são não-decrescentes em $\mathbb{R}^p_+$, isto é, se $x, y \in \mathbb{R}^p_+$, satisfazendo para cada $j = 1, ..., p$, $x_j \leq y_j$, então $f_j(x) \leq f_j(y)$.

O modelo de localização é dado por

$$min\{(\phi_1(x), \phi_2(x), ..., \phi_m(x)) : x \in \mathbb{R}^n\},$$

onde, para cada $i = 1, ..., m$, $\phi_i(x) = max_{1 \leq j \leq p} f_j(\gamma_i(x))$.

Se as funções $f_j : \mathbb{R}^p_+ \longrightarrow \mathbb{R}^+$ *são quase-convexas em* $\mathbb{R}^p_+$, *e pelo fato da função* $\gamma_i(x)$ *ser uma função vetorial convexa (conferir Hiriart-Urruty e Lemaréchal ([21], Teorema 1.2.5, Definição 1.1.1), então para* $i = 1, ..., m$, *cada função* $\phi_i(.)$ *é quase-convexa em* $\mathbb{R}^n$.

Capítulo 2

Método proximal com subdiferencial de Clarke

Estamos interessados em resolver o problema de otimização multiobjetivo (POM) irrestrito:

$$\min\{F(x) : x \in \mathbb{R}^n\} \tag{2.1}$$

onde $F : \mathbb{R}^n \to \mathbb{R}^m$ satisfaz as seguintes hipóteses:

$(\mathbf{H_1})$ F é uma aplicação localmente Lipschitz em $\mathbb{R}^n$;

$(\mathbf{H_2})$ F é $\mathbb{R}^m_+$-quase-convexa.

2.0.1 O algoritmo

Nesta seção propomos um método de ponto proximal escalar com regularização quadrática utilizando o subdiferencial de Clarke, denotado por **MEPPC**, para resolver o problema (2.1). Este método gera uma sequência pela seguinte recursão:

Algoritmo MEPPC

Inicialização: Escolha um ponto inicial arbitrário

$$x^0 \in \mathbb{R}^n \tag{2.2}$$

Passo principal: Dado x^k, encontrar x^{k+1} tal que

$$0 \in \partial^o \left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \parallel . \; - x^k \|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}) \tag{2.3}$$

onde ∂^o é o subdiferencial de Clarke, $\Omega_k = \left\{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\right\}$, $\alpha_k > 0$, $\{e_k\} \subset \mathbb{R}^m_{++}$, $\|e_k\| = 1$, $\{z_k\} \subset \mathbb{R}^m_+ \backslash \{0\}$ e $\|z_k\| = 1$.

Critério de parada: Se $x^{k+1} = x^k$ ou x^{k+1} é um ponto Pareto-Clarke crítico, então pare. Caso contrário, faça $k \leftarrow k+1$ e retorne ao Passo principal.

Observação 2.0.1 *Se F é $\mathbb{R}^m_+-$convexa, então o passo principal* (2.3) *é equivalente a:*

$$x^{k+1} = argmin\left\{\langle F(x), z_k\rangle + \frac{\alpha_k}{2}\langle e_k, z_k\rangle \left\|x - x^k\right\|^2 : x \in \Omega_k\right\} \tag{2.4}$$

Esta iteração foi estudada por Bonnel et al. [8]. Assim, podemos dizer que, em um certo sentido, nossa iteração é uma extensão para o caso não convexo do trabalho citado em [8] . Por outro lado, quando F é $\mathbb{R}^m_+-$quase-convexa, a função regularizada $F_k(x) = \langle F(x), z_k\rangle + \frac{\alpha_k}{2}\langle e_k, z_k\rangle \left\|x - x^k\right\|^2$ não necessariamente é quase-convexa e, dessa forma, (2.4) é um problema de otimização global, isto é, um problema de buscar a melhor solução, dentro do conjunto das soluções viáveis, e isto no sentido computacional não é tão eficiente. Esta é a razão pela qual consideramos a iteração mais fraca (2.3).

Ressaltamos que o conjunto $\Omega_k = \left\{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\right\}$, para cada $k \in \mathbb{N}$ é um conjunto fechado, pelo fato da F satisfazer a hipótese $(\mathbf{H_1})$, e é um conjunto convexo, pelo fato da F satisfazer a hipótese $(\mathbf{H_2})$.

2.0.2 Existência das iteradas

Nesta seção, garantimos que a sequência gerada pelo algoritmo **MEPPC** está bem definida.

Teorema 2.0.1 *Seja $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação satisfazendo $(\mathbf{H_1})$, $(\mathbf{H_2})$ e $0 \prec F$. Então a sequência $\left\{x^k\right\}$, gerada pelo algoritmo* **MEPPC**, *dado por* (2.2) *e* (2.3), *está bem definida.*

Prova: Seja $x^0 \in \mathbb{R}^n$ um ponto arbitrário dado no passo da inicialização. Dado x^k, defina $\varphi_k(x) = \langle F(x), z_k\rangle + \frac{\alpha_k}{2}\langle e_k, z_k\rangle \left\|x - x^k\right\|^2 + \delta_{\Omega_k}(x)$, onde $\delta_{\Omega_k}(.)$ é a função indicadora de Ω_k. Então temos que $\min\{\varphi_k(x) : x \in \mathbb{R}^n\}$ é equivalente a $\min\{\langle F(x), z_k\rangle + \frac{\alpha_k}{2}\langle e_k, z_k\rangle \left\|x - x^k\right\|^2 : x \in \Omega_k\}$. Desde que $0 \prec F$, para cada $z \in \mathbb{R}^m_+ \setminus \{0\}$, a função $\langle F(.), z\rangle$ é limitada inferiormente. Então, pela limitação inferior e continuidade da função $\langle F(.), z\rangle$, e também, pela continuidade e coercividade da função norma, da Proposição 1.1.1, existe $x^{k+1} \in \Omega_k$, não necessariamente único devido a não convexidade da F, o qual é um mínimo global de $\varphi_k(.)$. Assim, a Proposição 1.1.6, Observação 1.1.2, garante que x^{k+1} satisfaz

$$0 \in \hat{\partial}\left(\langle F(.), z_k\rangle + \frac{\alpha_k}{2}\langle e_k, z_k\rangle \| \;.\; - x^k\|^2 + \delta_{\Omega_k}(.)\right)(x^{k+1}).$$

Da Proposição 1.1.4 e Proposição 1.1.8 temos que

$$0 \in \partial \Big(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \ . \ - x^k \|^2 \Big) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}), \tag{2.5}$$

e pela Observação 1.1.4, a iteração (2.3) é obtida de (2.5). ■

Observação 2.0.2 (Huang e Yang [22]) *Sem perda de generalidade, podemos sempre assumir que a aplicação $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfaz $0 \prec F$. De fato, considere o seguinte problema de otimização multiobjetivo*

$$(P') \quad \min \big\{ (e^{F(x)} : x \in \mathbb{R}^n \big\}.$$

Agora observe que se $\bar{x} \in \mathbb{R}^n$ é um ponto Pareto-Clarke crítico para o problema (2.1)*, então para todas as direções $d \in \mathbb{R}^n$, existe $i_0 = i_0(d) \in \{1, ..., m\}$ tal que $F^o_{i_o}(\bar{x}, d) \geq 0$, isto é:*

$$F^o_{i_0}(\bar{x}, d) = \limsup_{t \downarrow 0 \quad y \to \bar{x}} \frac{F_{i_0}(y + td) - F_{i_0}(y)}{t} \geq 0$$

Segue que para $\varepsilon > 0$ arbitrário, $F_{i_0}(y) \leq F_{i_0}(y + td)$, $\forall \ y \in B(\bar{x}, \varepsilon)$ e $t \in (0, \varepsilon]$. Pelo fato da função $e^{F_{i_0}(.)}$ ser crescente, temos que

$$[e^{F_{i_0}(.)}]^o(\bar{x}, d) = \limsup_{t \downarrow 0 \quad y \to \bar{x}} \frac{e^{F_{i_0}(y+td)} - e^{F_{i_0}(y)}}{t} \geq 0$$

Portanto, os problemas (2.1) *e (P') possuem o mesmo conjunto de pontos Pareto-Clarke críticos. Além disso, se F é $\mathbb{R}^m_+$ - quase-convexa e localmente Lipschitz em $\mathbb{R}^n$, então $e^{F(x)}$ também é $\mathbb{R}^m_+$ - quase-convexa e localmente Lipschitz em $\mathbb{R}^n$. Portanto, ao longo deste capítulo, a partir de agora, assumiremos que $0 \prec F$.*

Observação 2.0.3 *Estamos interessados na convergência assintótica do método. Assim assumiremos ao longo deste capítulo que em cada iteração, x^k não é um ponto Pareto-Clarke crítico e $x^{k+1} \neq x^k$. Isto implica, da Observação 1.1.5, que o interior de Ω_{k+1}, denotado por Ω^0_{k+1}, não é vazio.*

Quando a condição, $x^{k+1} \neq x^k$, não é satisfeita, isto é, se existe k_0 tal que $x^{k_0+1} = x^{k_0}$ então provamos que este ponto é Pareto Clarke crítico da F (conferir Proposição (2.0.5)*).*

2.0.3 Convergência Fraca

Nesta subseção provamos, sob a suposição de que as iterações consecutivas converge para zero, que qualquer ponto de acumulação é um ponto Pareto-Clarke crítico do problema (2.1).

Proposição 2.0.1 *Seja* $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação satisfazendo* $(\mathbf{H_1})$ *e* $(\mathbf{H_2})$. *Se* $0 < \alpha_k < \tilde{\alpha}$, *com* $\tilde{\alpha} > 0$, *e a sequência* $\{x^k\}$ *gerada pelo algoritmo* **MEPPC**, (2.2) *e* (2.3), *satisfaz*

$$\lim_{k\to+\infty} ||x^{k+1} - x^k|| = 0, \tag{2.6}$$

e tem um ponto de acumulação, então este ponto é Pareto-Clarke crítico do problema (2.1).

Prova: Por hipótese, existe uma subsequência convergente $\left\{x^{k_j}\right\}$ de $\left\{x^k\right\}$ cujo limite é algum $\widehat{x} \in \mathbb{R}^n$. Desde que F é localmente Lipschitz em $\mathbb{R}^n$, então a função $\langle F(.), z\rangle$ é também localmente Lipschitz em $\mathbb{R}^n$ e portanto, contínua para todo $z \in \mathbb{R}^m$, em particular, para todo $z \in \mathbb{R}^m_+\backslash \{0\}$, e $\lim_{j\to+\infty} \left\langle F(x^{k_j}), z\right\rangle = \langle F(\widehat{x}), z\rangle$. Por outro lado, como $x^{k+1} \in \Omega_k$, temos que $F(x^{k+1}) \preceq F(x^k)$ e visto que $z \in \mathbb{R}^m_+\backslash \{0\}$, concluímos que a sequência $\left\{\left\langle F(x^k), z\right\rangle\right\}$ é convergente para $\langle F(\widehat{x}), z\rangle$, pois é uma sequência não-crescente que admite uma subsequência convergente para $\langle F(\widehat{x}), z\rangle$. Assim $\lim_{k\to+\infty} \left\langle F(x^k), z\right\rangle = \langle F(\widehat{x}), z\rangle = inf_{k\in\mathbb{N}} \left\{\left\langle F(x^k), z\right\rangle\right\} \leq \left\langle F(x^k), z\right\rangle$. Portanto $\left\langle F(x^k) - F(\widehat{x}), z\right\rangle \geq 0, \forall\ k \in \mathbb{N}, \forall\ z \in \mathbb{R}^m_+\backslash \{0\}$. Concluímos que $F(x^k) - F(\widehat{x}) \in \mathbb{R}^m_+$, isto é, $F(\widehat{x}) \preceq F(x^k), \forall\ k \in \mathbb{N}$. Isto implica que $\widehat{x} \in \Omega_k$.
Supondo, para chegarmos a uma contradição, que $\widehat{x}$ não é ponto Pareto-Clarke crítico em $\mathbb{R}^n$, então existe uma direção $d \in \mathbb{R}^n$ tal que

$$F_i^o(\widehat{x}, d) < 0, \forall\ i \in \{1, ..., m\} \tag{2.7}$$

Logo d é uma direção de descida para a função multiobjetivo F em $\widehat{x}$, assim, $\exists\ \varepsilon > 0$ tal que

$$F(\widehat{x} + \lambda d) \prec F(\widehat{x}),\ \forall\ \lambda \in (0, \varepsilon].$$

Portanto $\widehat{x} + \lambda d \in \Omega_k$.
Por outro lado, como $\left\{x^k\right\}$ é gerada pelo algoritmo **MEPPC**, temos, do Teorema 2.0.1, de (2.3), do Lema 1.1.5 e do Lema 1.1.4, que existe $\beta_k(x^k - x^{k+1}) - v_k \in \partial^o\left(\langle F(.), z_k\rangle\right)(x^{k+1})$, com $v_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ e $\beta_k = \alpha_k \langle e_k, z_k\rangle > 0$, tal que

$$\beta_k\langle x^k - x^{k+1}, p\rangle - \langle v_k, p\rangle \leq \langle F(.), z_k\rangle^o(x^{k+1}, p), \forall p \in \mathbb{R}^n. \tag{2.8}$$

Considere $p = (\widehat{x} + \lambda d) - x^{k+1}$ e, como $v_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$, de (2.8) temos

$$\beta_k\langle x^k - x^{k+1}, \widehat{x} + \lambda d - x^{k+1}\rangle \leq \langle F(.), z_k\rangle^o(x^{k+1}, \widehat{x} + \lambda d - x^{k+1}) \tag{2.9}$$

Agora considere um sequência $\{z_k\} \subset \mathbb{R}^m_+\backslash \{0\}$ com $\|z_k\| = 1$. Como $\{z_k\}$ é limitada, então existe uma subsequência, $\left\{z^{k_j}\right\}$ tal que $\lim_{j\to+\infty} z^{k_j} = \bar{z}$, com $\bar{z} \in \mathbb{R}^m_+\backslash \{0\}$. De

(2.9), temos:

$$\beta_{k_j}\langle x^{k_j} - x^{k_j+1}, \widehat{x} + \lambda d - x^{k_j+1}\rangle \leq \langle F(.), z_{k_j}\rangle^o(x^{k_j+1}, \widehat{x} + \lambda d - x^{k_j+1})$$

Do Lema 1.1.3, (i) e (ii), temos:

$$\beta_{k_j}\langle x^{k_j} - x^{k_j+1}, \widehat{x} + \lambda d - x^{k_j+1}\rangle \leq \sum_{i=1}^{m} z_{k_j}^i F_i^o(x^{k_j+1}, \widehat{x} + \lambda d - x^{k_j+1}),$$

onde $z_{k_j}^i$ são as componentes do vetor z_{k_j}. Então utilizando o Lema 1.1.3, (iii), obtemos:

$$\beta_{k_j}\langle x^{k_j} - x^{k_j+1}, \widehat{x} + \lambda d - x^{k_j+1}\rangle \leq \sum_{i=1}^{m} F_i^o\left(x^{k_j+1}, z_{k_j}^i(\widehat{x} + \lambda d - x^{k_j+1})\right),$$

Tomando o limite superior na desigualdade acima, utilizando a condição (2.6), a Proposição 1.1.10 e como $\lambda > 0$, concluímos que

$$0 \leq F_1^o(\widehat{x}, d)\bar{z}_1 + ... + F_m^o(\widehat{x}, d)\bar{z}_m \tag{2.10}$$

Sem perda de generalidade, considere o conjunto $J = \{i \in I : \bar{z}_i > 0\}$, onde $I = \{1, ..., m\}$. Logo, de (2.10), existe $i_0 \in J$ tal que $F_{i_0}^o(\widehat{x}, d)\bar{z}_{i_0} \geq 0$, portanto $F_{i_0}^o(\widehat{x}, d) \geq 0$, contradizendo (2.7). ■

2.0.4 Convergência Global

Nesta subseção consideramos a seguinte hipótese em relação à aplicação F e um ponto inicial x^0 :

($\mathbf{H_3}$) O conjunto $\left(F(x^0) - \mathbb{R}^m_+\right) \cap F(\mathbb{R}^n)$ é $\mathbb{R}^m_+$ - completo, isto significa que para toda sequência $\{a_k\} \subset \mathbb{R}^n$, com $a_0 = x^0$, tal que $F(a_{k+1}) \preceq F(a_k)$, existe $a \in \mathbb{R}^n$ tal que $F(a) \preceq F(a_k)$, $\forall\ k \in \mathbb{N}$.

Observação 2.0.4 *A hipótese* ($\mathbf{H_3}$) *é citada em vários trabalhos que envolvem o método ponto proximal para funções convexas, ver Bonnel et al. [8], Ceng e Yao [11] e, Villacorta e Oliveira [44].*

Como a sequência $\left\{x^k\right\}$ gerada pelo algoritmo **MEPPC** satisfaz a hipótese ($\mathbf{H_3}$), e das hipóteses ($\mathbf{H_1}$) e ($\mathbf{H_2}$), então o conjunto

$$E = \left\{x \in \mathbb{R}^n : F(x) \preceq F\left(x^k\right) \quad \forall\ k \in \mathbb{N}\right\}$$

é não vazio, fechado e convexo.

Proposição 2.0.2 (*Fejér convergência*)
Se as hipóteses $(\mathbf{H_1})$, $(\mathbf{H_2})$ *e* $(\mathbf{H_3})$ *são satisfeitas, então a sequência* $\{x^k\}$ *gerada pelo algoritmo* **MEPPC**, (2.2) *e* (2.3), *é Fejér convergente ao conjunto* E.

Prova: Do Teorema 2.0.1, de (2.3), do Lema 1.1.5 e do Lema 1.1.4, obtemos que existe $g_i^k \in \partial^o F_i(x^{k+1}), i = 1, ..., m$ tal que

$$0 \in \sum_{i=1}^{m} z_k^i g_i^k + \alpha_k \langle e_k, z_k \rangle (x^{k+1} - x^k) + \mathcal{N}_{\Omega_k}(x^{k+1})$$

onde z_k^i são as componentes de z_k. Portanto existem vetores $g_i^k \in \partial^o F_i(x^{k+1}), i = 1, ..., m$, e $v_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ tais que

$$\sum_{i=1}^{m} z_k^i g_i^k = \beta_k (x^k - x^{k+1}) - v_k \tag{2.11}$$

onde $\beta_k = \alpha_k \langle e_k, z_k \rangle$, $\forall\ k \in \mathbb{N}$. Note que $\beta_k > 0$, pois $\alpha_k > 0$, e_k pertence a $\mathbb{R}^m_{++}$, e z_k pertence a $\mathbb{R}^m_{+} \backslash \{0\}$. De (2.11), temos

$$x^k - x^{k+1} = \frac{1}{\beta_k} \left(\sum_{i=1}^{m} z_k^i g_i^k + v_k \right) \tag{2.12}$$

Agora tome $x^* \in E$, então pela definição do conjunto E, $x^* \in \Omega_{k+1}$ para todo k, e da Observação 2.0.3, existe $\{x^l\} \in \Omega^0_{k+1}$ tal que $x^l \to x^*$. Observe que, $\forall\ x \in \mathbb{R}^n$:

$$\left\| x^k - x \right\|^2 = \left\| x^k - x^{k+1} \right\|^2 + \left\| x^{k+1} - x \right\|^2 + 2 \left\langle x^k - x^{k+1}, x^{k+1} - x \right\rangle. \tag{2.13}$$

Agora, combinando (2.13), com $x = x^l$, e (2.12), temos:

$$\left\| x^k - x^l \right\|^2 = \left\| x^k - x^{k+1} \right\|^2 + \left\| x^{k+1} - x^l \right\|^2 + \frac{2}{\beta_k} \left(\sum_{i=1}^{m} z_k^i \left\langle g_i^k,\ x^{k+1} - x^l \right\rangle + \left\langle v_k\ ,\ x^{k+1} - x^l \right\rangle \right) \tag{2.14}$$

Como $F(x^l) \prec F(x^{k+1})$, então $F_i(x^l) < F_i(x^{k+1})$, Além disso, $g_i^k \in \partial^o F_i(x^{k+1})$ e F_i é quase-convexa, utilizando a Proposição 1.1.11, temos que

$$\left\langle g_i^k,\ x^{k+1} - x^l \right\rangle \geq 0, \forall i = 1, ..., m. \tag{2.15}$$

Agora, como $v_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$, (2.14) e a desigualdade (2.15), implicam, tomando $l \to \infty$

$$0 \leq \left\| x^{k+1} - x^k \right\|^2 \leq \left\| x^k - x^* \right\|^2 - \left\| x^{k+1} - x^* \right\|^2, \forall\ k \in \mathbb{N}. \tag{2.16}$$

Portanto

$$\left\| x^{k+1} - x^* \right\| \leq \left\| x^k - x^* \right\| \tag{2.17}$$

■

Proposição 2.0.3 *Se as hipóteses* $(\mathbf{H_1})$, $(\mathbf{H_2})$ *e* $(\mathbf{H_3})$, *são satisfeitas, então a sequência* $\left\{x^k\right\}$ *gerada pelo algoritmo* **MEPPC**, (2.2) *e* (2.3), *satisfaz*

$$\lim_{k \to +\infty} \left\| x^{k+1} - x^k \right\| = 0.$$

Prova: Segue de (2.17) que $\forall x^* \in E$, $\left\{\left\| x^k - x^* \right\|\right\}$ é uma sequência não crescente de termos não negativos, portanto convergente. Assim, o lado direito de (2.16) converge a 0 quando $k \to +\infty$, e o resultado é obtido. ■

Proposição 2.0.4 *Se as hipóteses* $(\mathbf{H_1})$, $(\mathbf{H_2})$ *e* $(\mathbf{H_3})$ *são satisfeitas, então a sequência* $\left\{x^k\right\}$ *gerada pelo algoritmo* **MEPPC** *converge para algum ponto do conjunto* E.

Prova: Da Proposição 2.0.2 e do Lema 1.1.6, (i), $\left\{x^k\right\}$ é limitada, então existe uma subsequência $\left\{x^{k_j}\right\}$ tal que $\lim\limits_{j \to +\infty} x^{k_j} = \widehat{x}$. Desde que F é localmente Lipschitz em $\mathbb{R}^n$, então a função $\langle F(.), z \rangle$ é também localmente Lipschitz em $\mathbb{R}^n$ portanto contínua para todo $z \in \mathbb{R}^m$, em particular, para todo $z \in \mathbb{R}^m_+ \backslash \{0\}$, e $\lim\limits_{j \to +\infty} \left\langle F(x^{k_j}), z \right\rangle = \langle F(\widehat{x}), z \rangle$. Por outro lado, como $x^{k+1} \in \Omega_k$, temos que $F(x^{k+1}) \preceq F(x^k)$. Visto que $z \in \mathbb{R}^m_+ \backslash \{0\}$, concluímos que $\left\langle F(x^{k+1}), z \right\rangle \leq \left\langle F(x^k), z \right\rangle$. Além disso, da Observação 2.0.2, podemos assumir que a função $\langle F(.), z \rangle$ é limitada inferiormente para cada $z \in \mathbb{R}^m_+ \backslash \{0\}$. Então a sequência $\left\{\left\langle F(x^k), z \right\rangle\right\}$ é não crescente e limitada inferiormente, portanto convergente. Assim, $\lim\limits_{k \to +\infty} \langle F(x^k), z \rangle = \langle F(\widehat{x}), z \rangle = inf_{k \in \mathbb{N}} \left\{\left\langle F(x^k), z \right\rangle\right\} \leq \left\langle F(x^k), z \right\rangle$. Logo $\left\langle F(x^k) - F(\widehat{x}), z \right\rangle \geq 0, \forall\ k \in \mathbb{N}, \forall\ z \in \mathbb{R}^m_+ \backslash \{0\}$. Concluímos que $F(x^k) - F(\widehat{x}) \in \mathbb{R}^m_+$, isto é, $F(\widehat{x}) \preceq F(x^k), \forall\ k \in \mathbb{N}$. Portanto $\widehat{x} \in E$, então pelo Lema 1.1.6, (ii), obtemos o resultado. ■

Finalmente provaremos que a sequência das iterações converge para um ponto Pareto-Clarke crítico, quando a sequência dos parâmetros de regularização $\{\alpha_k\}$ é limitada.

Teorema 2.0.2 *Considere* $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação satisfazendo as hipóteses* $(\mathbf{H_1})$, $(\mathbf{H_2})$ *e* $(\mathbf{H_3})$. *Se* $0 < \alpha_k < \tilde{\alpha}$, *então a sequência* $\{x^k\}$ *gerada pelo algoritmo* **MEPPC**, (2.2) *e* (2.3), *converge para um ponto Pareto-Clarke crítico do problema* (2.1).

Prova:

Do Teorema 2.0.1, de (2.3), do Lema 1.1.5 e do Lema 1.1.4, implica que existem $\beta_k(x^k - x^{k+1}) - v_k \in \partial^o\left(\langle F(.), z_k\rangle\right)(x^{k+1})$, com $v_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ e $\beta_k = \alpha_k \langle e_k, z_k\rangle > 0$, tais que

$$\beta_k\langle x^k - x^{k+1}, p\rangle - \langle v_k, p\rangle \leq \langle F(.), z_k\rangle^o(x^{k+1}, p), \forall p \in \mathbb{R}^n \tag{2.18}$$

Seja $x^* \in E$ fixo, e considere $p = x^* - x^{k+1}$. Como $v_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$, de (2.18), temos

$$\beta_k\langle x^k - x^{k+1}, x^* - x^{k+1}\rangle \leq \langle F(.), z_k\rangle^o(x^{k+1}, x^* - x^{k+1}) \tag{2.19}$$

Seja $\lim\limits_{k\to+\infty} x^k = \bar{x}$ e considere uma sequência $\{z_k\} \subset \mathbb{R}^m_+ \setminus \{0\}$ com $\|z_k\| = 1$. Como $\{z_k\}$ é limitada, então existe uma subsequência, $\left\{z^{k_j}\right\}$ tal que $\lim\limits_{j\to+\infty} z^{k_j} = \bar{z}$, com $\bar{z} \in \mathbb{R}^m_+ \setminus \{0\}$. De (2.19), temos que:

$$\beta_{k_j}\langle x^{k_j} - x^{k_j+1}, x^* - x^{k_j+1}\rangle \leq \langle F(.), z_{k_j}\rangle^o(x^{k_j+1}, x^* - x^{k_j+1})$$

Pelo Lema 1.1.3, (i) e (ii), temos:

$$\beta_{k_j}\langle x^{k_j} - x^{k_j+1}, x^* - x^{k_j+1}\rangle \leq \sum_{i=1}^{m} z^i_{k_j} F^o_i(x^{k_j+1}, x^* - x^{k_j+1}),$$

onde $z^i_{k_j}$ são as componentes do vetor z_{k_j}. Então utilizando o Lema 1.1.3, (iii), obtemos:

$$\beta_{k_j}\langle x^{k_j} - x^{k_j+1}, x^* - x^{k_j+1}\rangle \leq \sum_{i=1}^{m} F^o_i\left(x^{k_j+1}, z^i_{k_j}(x^* - x^{k_j+1})\right),$$

Tomando o limite superior na desigualdade acima, pela Proposição 2.0.3 e pela Proposição 1.1.10, obtemos que

$$0 \leq F^o_1(\bar{x}, x^* - \bar{x})\bar{z}_1 + ... + F^o_m(\bar{x}, x^* - \bar{x})\bar{z}_m, \forall x^* \in E \tag{2.20}$$

Sem perda de generalidade, considere o conjunto $J = \{i \in I : \bar{z}_i > 0\}$, onde $I = \{1, ..., m\}$. Portanto, de (2.20), existe $i_0 \in J$ tal que $F^o_{i_0}(\bar{x}, x^* - \bar{x})\bar{z}_{i_0} \geq 0$. Portanto

$$F^o_{i_0}(\bar{x}, x^* - \bar{x}) \geq 0, \ \forall\ x^* \in E. \tag{2.21}$$

Agora provaremos que $\bar{x}$ é ponto Pareto-Clarke crítico em $\mathbb{R}^n$.

Suponha por contradição que $\bar{x}$ não é ponto Pareto-Clarke crítico em $\mathbb{R}^n$, então existe um direção $d \in \mathbb{R}^n$ tal que

$$F^o_i(\bar{x}, d) < 0, \forall\ i \in \{1, ..., m\} \tag{2.22}$$

Assim, da Observação 1.1.5, d é uma direção de descida para a função multiobjetivo F em $\bar{x}$, logo, $\exists\ \varepsilon > 0$ tal que

$$F(\bar{x} + \lambda d) \prec F(\bar{x}),\ \forall\ \lambda \in (0, \varepsilon]. \tag{2.23}$$

Como $\bar{x} \in E$, então de (2.23) concluímos que $\bar{x}+\lambda d \in E$. Portanto, de (2.21), com $x^* = \bar{x}+\lambda d$, e do Lema 1.1.3, (iii), temos que: $F^o_{i_0}(\bar{x}\ ,\ \bar{x} + \lambda d - \bar{x}) = F^o_{i_0}(\bar{x}\ ,\ \lambda d) = \lambda F^o_{i_0}(\bar{x}\ ,\ d) \geq 0$. Visto que $\lambda > 0$, concluímos que $F^o_{i_0}(\bar{x}\ ,\ d) \geq 0$, o que implica em um contradição com (2.22). Portanto, $\bar{x}$ é ponto Pareto-Clarke crítico do problema (2.1). ∎

Teorema 2.0.3 *Se $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é $\mathbb{R}^m_+$-convexa e $\bar{x}$ o ponto de convergência da sequência gerada pelo algoritmo* **MEPPC**, *dado por* (2.2) *e* (2.3), *então $\bar{x}$ é solução Pareto fraca do problema* (2.1).

Prova: A demonstração é imediata da Proposição 1.1.13. ∎

Proposição 2.0.5 **(Critério de parada do MEPPC)** *Seja $\{x^k\}$ uma sequência gerada pelo algoritmo* **MEPPC**. *Se, para algum inteiro k_0 , $x^{k_0+1} = x^{k_0}$, então x^{k_0} é ponto Pareto-Clarke crítico do Problema* (2.1).

Prova: Suponha que o critério de parada é verificado na k_0-ésima iteração. Visto que x^{k_0+1} satisfaz (2.3) e $x^{k_0+1} = x^{k_0}$, existe $-v_{k_0} \in \partial^o\left(\langle F(.), z_{k_0}\rangle\right)(x^{k_0})$ com $\left\langle v_{k_0}\ ,\ x - x^{k_0}\right\rangle \leq 0, \forall x \in \Omega_{k_0}$ tal que

$$-\langle v_{k_0}, d_{k_0}\rangle \leq \langle F(.), z_{k_0}\rangle^o(x^{k_0}, d_{k_0}), \forall d_{k_0} \in \mathbb{R}^n$$

Seja $x^* \in E$ fixo, e considere $d_{k_0} = x^* - x^{k_0}$. Portanto, da última desigualdade, obtemos

$$\langle F(.), z_{k_0}\rangle^o(x^{k_0}, x^* - x^{k_0}) \geq 0. \tag{2.24}$$

Suponha por contradição que x^{k_0} não é ponto Pareto-Clarke crítico, então existe uma direção $d \in \mathbb{R}^n$ tal que

$$F^o_i(x^{k_0}, d) < 0, \forall\ i \in \{1, ..., m\}. \tag{2.25}$$

Segue que existe $\varepsilon > 0$ tal que

$$F(x^{k_0} + \lambda d) \prec F(x^{k_0}),\ \forall\ \lambda \in (0, \varepsilon]. \tag{2.26}$$

De (2.26), $x^{k_0} + \lambda d \in E$ e substituindo em (2.24), obtemos $\langle F(.), z_{k_0}\rangle^o(x^{k_0}, \lambda d) \geq 0$. Visto que $\lambda > 0$, obtemos que $\langle F(.), z_{k_0}\rangle^o(x^{k_0}, d) \geq 0$. De forma análoga provada na Proposição 2.0.1 e no Teorema 2.0.2, concluímos que existe $i_0 \in J =$

$\{i \in I : z_{k_0}^i > 0\}$, onde $I = \{1, ..., m\}$, tal que $F_{i_0}^o(x^{k_0}, d) \geq 0$, contradizendo (2.25).
■

Capítulo 3

Método proximal com subdiferencial de Fréchet

Estamos interessados em resolver o problema de otimização multiobjetivo (POM) irrestrito:

$$\min\{F(x) : x \in \mathbb{R}^n\} \tag{3.1}$$

onde $F : \mathbb{R}^n \to \mathbb{R}^m$ satisfaz a seguinte hipótese:

$(\mathbf{C_1})$ F é uma aplicação semicontínua inferior(sc.i) em $\mathbb{R}^n$.

3.1 Algoritmo exato

Nesta seção propomos um método de ponto proximal escalar com regularização quadrática utilizando o subdiferencial de Fréchet, denotado por **MEPPF**, para resolver o problema (3.1). Este método gera uma sequência pela seguinte recursão:

Algoritmo MEPPF

Inicialização: Escolha um ponto inicial arbitrário

$$x^0 \in \mathbb{R}^n \tag{3.2}$$

Passo principal: Dado x^k encontrar x^{k+1} tal que

$$0 \in \hat{\partial}\left(\langle F(.), z_k\rangle + \frac{\alpha_k}{2}\langle e_k, z_k\rangle \|\ .\ - x^k\|^2 + \delta_{\Omega_k}(.)\right)(x^{k+1}) \tag{3.3}$$

onde $\hat{\partial}$ é o subdiferencial de Fréchet, $\Omega_k = \left\{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\right\}$, $\alpha_k > 0$, $\{e_k\} \subset \mathbb{R}^m_{++}$, $\|e_k\| = 1$, $\{z_k\} \subset \mathbb{R}^m_+ \backslash \{0\}$ e $\|z_k\| = 1$.

Critério de parada: Se $x^{k+1} = x^k$ ou x^{k+1} é um ponto Pareto crítico, então pare. Caso contrário, faça $k \leftarrow k+1$ e retorne ao Passo principal.

Podemos supor, sem perda de generalidade, que a aplicação F é limitada inferiormente. Segue a observação.

Observação 3.1.1 *De forma análoga considerada na Observação* 2.0.2*, sem perda de generalidade, podemos sempre assumir que a aplicação $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfaz $0 \prec F$. Caso contrário, considere o seguinte problema de otimização multiobjetivo*

$$(P') \quad \min\left\{e^{F(x)} : x \in \mathbb{R}^n\right\}.$$

Agora observe que se $\bar{x} \in \mathbb{R}^n$ é uma solução Pareto fraca para o problema (3.1)*, então não existe $x \in \mathbb{R}^n$ tal que $F_i(x) < F_i(\bar{x})$, para todo $i \in \{1, ..., m\}$. Como, para cada $i \in \{1, ..., m\}$, a função exponencial $e^{F_i(x)}$ é monótona crescente, então podemos concluir que não existe $x \in \mathbb{R}^n$ tal que $e^{F_i(x)} < e^{F_i(\bar{x})}$, para todo $i \in \{1, ..., m\}$. E portanto, os problemas* (3.1) *e* (P') *possuem o mesmo conjunto de soluções Pareto fracas. Além disto, de forma análoga concluída na Observação* 2.0.2*, utilizando agora o conceito de ponto Pareto crítico, concluímos que estes problemas possuem o mesmo conjunto de pontos Pareto críticos. E, se F é $\mathbb{R}^m_+$ - quase-convexa e continuamente diferenciável em $\mathbb{R}^n$, então $e^{F(x)}$ também é $\mathbb{R}^m_+$ - quase-convexa e continuamente diferenciável em $\mathbb{R}^n$,.*

Portanto, ao longo deste capítulo, a partir de agora, assumiremos que $0 \prec F$.

No Teorema a seguir garantimos que a sequência gerada pelo algoritmo **MEPPF** está bem definida.

Teorema 3.1.1 *(***Existência das iteradas)**
Se $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma aplicação satisfazendo $(\mathbf{C_1})$*, então a sequência $\{x^k\}$, gerada pelo algoritmo* **MEPPF***, dado por* (3.2) *e* (3.3)*, está bem definida.*

Prova: Seja $x^0 \in \mathbb{R}^n$ um ponto arbitrário dado no passo da inicialização. Dado x^k, defina $\varphi_k(x) = \langle F(x), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \left\|x - x^k\right\|^2 + \delta_{\Omega_k}(x)$, onde $\delta_{\Omega_k}(.)$ é a função indicadora de Ω_k. Então temos que $\min\{\varphi_k(x) : x \in \mathbb{R}^n\}$ é equivalente a $\min\{\langle F(x), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \left\|x - x^k\right\|^2 : x \in \Omega_k\}$. Visto que φ_k é sc.i pois Ω_k é fechado, desde que $0 \prec F$, e pela coercividade da função $\|.\|^2$, obtemos que φ_k é sc.i e coerciva. Logo da Proposição 1.1.1, existe $x^{k+1} \in \Omega_k$, não necessariamente único devido à não convexidade da F, o qual é um mínimo global de $\varphi_k(.)$, assim, a Proposição 1.1.6, Observação 1.1.2, garante que x^{k+1} satisfaz

$$0 \in \hat{\partial}\left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \, . \, - x^k\|^2 + \delta_{\Omega_k}(.)\right)(x^{k+1}).$$

■

3.1.1 Fejér convergência

Para esta subseção suponha as seguintes hipóteses:

($\mathbf{C_2}$) F é uma aplicação $\mathbb{R}^m_+$ - quase-convexa.

($\mathbf{C_3}$) O conjunto $\left(F(x^0) - \mathbb{R}^m_+\right) \cap F(\mathbb{R}^n)$ é $\mathbb{R}^m_+$ - completo, isto significa que para toda sequência $\{a_k\} \subset \mathbb{R}^n$, com $a_0 = x^0$, tal que $F(a_{k+1}) \preceq F(a_k)$, existe $a \in \mathbb{R}^n$ tal que $F(a) \preceq F(a_k)$, $\forall\ k \in \mathbb{N}$.

Proposição 3.1.1 *Seja $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação que satisfaz as hipóteses* ($\mathbf{C_1}$) *e* ($\mathbf{C_2}$). *Se $g \in \hat{\partial}\left(\langle F(.), z\rangle + \delta_\Omega\right)(x)$, com $z \in \mathbb{R}^m_+ \backslash \{0\}$, e $F(y) \preceq F(x)$, com $y \in \Omega$, sendo $\Omega \subset \mathbb{R}^n$ um conjunto fechado e convexo, então $\langle g, y - x\rangle \leq 0$.*

Prova: Seja $t \in (0, 1]$, pela $\mathbb{R}^m_+$-quase-convexidade da F e pela hipótese que $F(y) \preceq F(x)$, temos que: $F_i(ty + (1-t)x) \leq \max\{F_i(x), F_i(y)\} = F_i(x)$, $\forall\ i \in \{1, ...m\}$. Segue-se daí que para cada $z \in \mathbb{R}^m_+ \backslash \{0\}$, temos

$$\langle F(ty + (1-t)x), z\rangle \leq \langle F(x), z\rangle \tag{3.4}$$

Seja $g \in \hat{\partial}\left(\langle F(.), z\rangle + \delta_\Omega\right)(x)$, obtemos

$$\langle F(ty + (1-t)x), z\rangle + \delta_\Omega(ty + (1-t)x) \geq \langle F(x), z\rangle + \delta_\Omega(x) + t\,\langle g, y - x\rangle + o(t\,\|y - x\|) \tag{3.5}$$

De (3.4) e (3.5), concluímos

$$t\,\langle g, y - x\rangle + o(t\,\|y - x\|) \leq 0 \tag{3.6}$$

Por outro lado, temos que $\lim\limits_{t\to 0} \frac{o(t\|y-x\|)}{t\|y-x\|} = 0$. Assim, $\lim\limits_{t\to 0} \frac{o(t\|y-x\|)}{t} = \lim\limits_{t\to 0} \frac{o(t\|y-x\|)}{t\|y-x\|}\,\|y - x\| = 0$. Portanto, dividindo (3.6) por t, e considerando o limite quando $t \to 0$, obtemos o resultado desejado. ■

Como a sequência $\left\{x^k\right\}$ gerada pelo algoritmo **MEPPF** satisfaz a hipótese ($\mathbf{C_3}$) então o conjunto

$$E = \left\{x \in \mathbb{R}^n : F\left(x\right) \preceq F\left(x^k\right) \quad \forall\ k \in \mathbb{N}\right\}$$

é não vazio.

Proposição 3.1.2 (*Fejér convergência*)
Suponha que as hipóteses ($\mathbf{C_1}$), ($\mathbf{C_2}$) *e* ($\mathbf{C_3}$) *sejam satisfeitas. Então a sequência $\left\{x^k\right\}_{k\in\mathbb{N}}$ gerada pelo algoritmo* **MEPPF** *é Fejér convergente em E.*

Prova:
Observe que $\forall\ x \in \mathbb{R}^n$:

$$\left\|x^k - x\right\|^2 = \left\|x^k - x^{k+1}\right\|^2 + \left\|x^{k+1} - x\right\|^2 + 2\left\langle x^k - x^{k+1}, x^{k+1} - x\right\rangle. \tag{3.7}$$

Do Teorema 3.1.1, de (3.3) e da Proposição 1.1.7, (iii), existe $g_k \in \hat{\partial}\left(\langle F(.), z_k\rangle + \delta_{\Omega_k}\right)(x^{k+1})$ tal que:

$$g_k = \beta_k(x^k - x^{k+1}) \tag{3.8}$$

onde $\beta_k = \alpha_k \langle e_k, z_k\rangle > 0$, $\forall\ k \in \mathbb{N}$. De (3.8) segue que

$$x^k - x^{k+1} = \frac{1}{\beta_k} g_k \tag{3.9}$$

Agora tome $x^* \in E$ fixo. Pela definição de E, $x^* \in \Omega_k$ para todo $k \in \mathbb{N}$. Combinando (3.7) com $x = x^*$ e (3.9), obtemos:

$$\left\|x^k - x^*\right\|^2 = \left\|x^k - x^{k+1}\right\|^2 + \left\|x^{k+1} - x^*\right\|^2 + \frac{2}{\beta_k}\left\langle g_k,\ x^{k+1} - x^*\right\rangle \geq \left\|x^k - x^{k+1}\right\|^2 + \left\|x^{k+1} - x^*\right\|^2 \tag{3.10}$$

A última desigualdade segue pela Proposição 3.1.1.
De 3.10, implica, em particular, que

$$0 \leq \left\|x^{k+1} - x^k\right\|^2 \leq \left\|x^k - x^*\right\|^2 - \left\|x^{k+1} - x^*\right\|^2, \forall\ x^* \in E \text{ e } \forall\ k \in \mathbb{N} \tag{3.11}$$

Portanto

$$\left\|x^{k+1} - x^*\right\| \leq \left\|x^k - x^*\right\| \tag{3.12}$$

■

Proposição 3.1.3 *Se as hipóteses* $(\mathbf{C_1})$, $(\mathbf{C_2})$ *e* $(\mathbf{C_3})$ *são satisfeitas, então a sequência* $\left\{x^k\right\}$ *gerada pelo algoritmo* **MEPPF**, (3.2) *e* (3.3)*, satisfaz*

$$\lim_{k \to +\infty} \left\|x^{k+1} - x^k\right\| = 0.$$

Prova: Segue de (3.12), que $\forall x^* \in E$, $\left\{\left\|x^k - x^*\right\|\right\}$ é uma sequência não crescente de termos não negativos, portanto convergente. Assim, o lado direito de (3.11) converge a 0 quando $k \to +\infty$, e o resultado é obtido. ■

3.1.2 Análise de convergência - caso não diferenciável

Nesta subseção analisaremos a convergência do método.

Proposição 3.1.4 (**Convergência para algum ponto de E**)
Se as hipóteses $(\mathbf{C_1})$*,* $(\mathbf{C_2})$ *e* $(\mathbf{C_3})$ *são satisfeitas, então a sequência* $\{x^k\}$ *gerada pelo algoritmo* **MEPPF** *converge para algum ponto do conjunto* E*.*

Prova: Da Proposição 3.1.1 e do Lema 1.1.6, (i), $\{x^k\}$ é limitada, então existe uma subsequência $\{x^{k_j}\}$ tal que $\lim_{j\to+\infty} x^{k_j} = \widehat{x}$. Desde que F é semicontínua inferior em $\mathbb{R}^n$, então a função $\langle F(.), z\rangle$ é também semicontínua inferior em $\mathbb{R}^n$ para todo $z \in \mathbb{R}^m$, em particular, para todo $z \in \mathbb{R}^m_+ \setminus \{0\}$, e $\langle F(\widehat{x}), z\rangle \leq \liminf_{j\to+\infty} \langle F(x^{k_j}), z\rangle$. Por outro lado, como $x^{k+1} \in \Omega_k$, temos que $F(x^{k+1}) \preceq F(x^k)$, e visto que $z \in \mathbb{R}^m_+ \setminus \{0\}$, concluímos que $\langle F(x^{k+1}), z\rangle \leq \langle F(x^k), z\rangle$. Além disso, da Observação 3.1.1, podemos assumir que a função $\langle F(.), z\rangle$ é limitada inferiormente, para cada $z \in \mathbb{R}^m_+ \setminus \{0\}$. Então a sequência $\{\langle F(x^k), z\rangle\}$ é não crescente e limitada inferiormente, portanto convergente. Assim,

$$\langle F(\widehat{x}), z\rangle \leq \liminf_{j\to+\infty} \langle F(x^{k_j}), z\rangle = \lim_{j\to+\infty} \langle F(x^{k_j}), z\rangle = inf_{k\in\mathbb{N}} \{\langle F(x^k), z\rangle\} \leq \langle F(x^k), z\rangle.$$

Logo $\langle F(x^k) - F(\widehat{x}), z\rangle \geq 0, \forall\ k \in \mathbb{N}, \forall\ z \in \mathbb{R}^m_+ \setminus \{0\}$. Concluímos que $F(x^k) - F(\widehat{x}) \in \mathbb{R}^m_+$, isto é, $F(\widehat{x}) \preceq F(x^k), \forall\ k \in \mathbb{N}$. Portanto $\widehat{x} \in E$, então pelo Lema 1.1.6, (ii), obtemos o resultado. ■

<u>Convergência para solução Pareto fraca</u>

Teorema 3.1.2 *Considere* $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação contínua satisfazendo as hipóteses* $(\mathbf{C_2})$ *e* $(\mathbf{C_3})$*. Se* $\lim_{k\to+\infty} \alpha_k = 0$ *e as iterações são dadas na forma*

$$x^{k+1} \in \arg\min \left\{ \langle F(x), z_k\rangle + \frac{\alpha_k}{2} \langle e_k, z_k\rangle \left\| x - x^k \right\|^2 : x \in \Omega_k \right\}, \tag{3.13}$$

então a sequência $\{x^k\}$ *gerada por* (3.2) *e* (3.13) *converge para uma solução Pareto fraca do problema* (3.1)*.*

Prova: Seja $x^{k+1} \in \arg\min \left\{ \langle F(x), z_k\rangle + \frac{\alpha_k}{2} \langle e_k, z_k\rangle \left\| x - x^k \right\|^2 : x \in \Omega_k \right\}$, isto implica que $\forall\ x \in \Omega_k$,

$$\left\langle F(x^{k+1}), z_k \right\rangle + \frac{\alpha_k}{2} \langle e_k, z_k\rangle \left\| x^{k+1} - x^k \right\|^2 \leq \langle F(x), z_k\rangle + \frac{\alpha_k}{2} \langle e_k, z_k\rangle \left\| x - x^k \right\|^2. \tag{3.14}$$

Visto que a sequência $\{x^k\}_{k\in\mathbb{N}}$ converge em E, então existe $x^* \in E$ tal que $\lim_{k\to+\infty} x^k = x^*$. Além disto considere uma sequência $\{z_k\} \subset \mathbb{R}^m_+ \setminus \{0\}$ com $\|z_k\| = 1$.

Como $\{z_k\}$ é limitada, então existe uma subsequência,$\{z_{k_l}\}_{l\in\mathbb{N}}$ tal que $\lim\limits_{l\to+\infty} z_{k_l} = \bar{z}$, com $\bar{z} \in \mathbb{R}^m_+ \setminus \{0\}$. Daí, a desigualdade em (3.14), $\forall\ x \in E$ torna-se:

$$\left\langle F(x^{k_l+1}), z_{k_l} \right\rangle + \frac{\alpha_{k_l}}{2} \langle e_{k_l}, z_{k_l} \rangle \left\| x^{k_l+1} - x^{k_l} \right\|^2 \leq \langle F(x), z_{k_l} \rangle + \frac{\alpha_{k_l}}{2} \langle e_{k_l}, z_{k_l} \rangle \left\| x - x^{k_l} \right\|^2. \tag{3.15}$$

Pela Proposição 3.1.3, $\lim\limits_{l\to+\infty} \left\| x^{k_l+1} - x^{k_l} \right\| = 0$, pelo fato das sequências $\left\{x^k\right\}$, $\{e_k\}$ e $\{z_k\}$ serem limitadas, e a sequência $\{\alpha_k\}$ convergir a 0, utilizando a desigualdade de Cauchy-Schwartz, concluímos que :

$$|\tfrac{\alpha_{k_l}}{2} \langle e_{k_l}, z_{k_l} \rangle \left\| x^{k_{l+1}} - x^{k_l} \right\|^2 | \to 0 \text{ e } |\tfrac{\alpha_{k_l}}{2} \langle e_{k_l}, z_{k_l} \rangle \left\| x - x^{k_l} \right\|^2 | \to 0 \text{ quando } l \to +\infty$$

Portanto, pela continuidade da F, tomando o limite em (3.15), obtemos que

$$\langle F(x^*), \overline{z} \rangle \leq \langle F(x), \overline{z} \rangle, \forall\ x \in E \tag{3.16}$$

Logo $x^* \in \arg\min \{\langle F(x), \overline{z} \rangle : x \in E\}$. Como $\langle F(.), \overline{z} \rangle$, com $\bar{z} \in \mathbb{R}^m_+ \setminus \{0\}$, é uma representação escalar estrita da F, portanto uma representação escalar fraca, então pela Proposição 1.1.3 temos que $x^* \in \arg\min_w \{F(x) : x \in E\}$.

Provaremos a seguir que $x^* \in \arg\min_w \{F(x) : x \in \mathbb{R}^n\}$. Suponha por contradição que $x^* \notin \arg\min_w \{F(x) : x \in \mathbb{R}^n\}$, então $\exists\ \widetilde{x} \in \mathbb{R}^n$ tal que

$$F(\widetilde{x}) \prec F(x^*) \tag{3.17}$$

Logo, para $\bar{z} \in \mathbb{R}^m_+ \setminus \{0\}$, segue-se

$$\langle F(\widetilde{x}), \bar{z} \rangle < \langle F(x^*), \bar{z} \rangle \tag{3.18}$$

Por outro lado, pela Proposição 3.1.4, temos que $x^* \in E$, logo de (3.17) temos que $F(\widetilde{x}) \prec F(x^k)\ \forall\ k \in \mathbb{N}$, isto é, $\widetilde{x} \in E$. Portanto de (3.16) e (3.18) chegamos a uma contradição. ■

Finalmente faremos a análise do método no sentido generalizado.

Teorema 3.1.3 *(***Análise de convergência generalizada***)*
Considere $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação contínua satisfazendo as hipóteses* $(\mathbf{C_2})$, $(\mathbf{C_3})$. *Se* $0 < \alpha_k < \tilde{\alpha}$, *então a sequência* $\{x^k\}$ *gerada pelo algoritmo* **MEPPF**, (3.2) *e* (3.3), *converge para um ponto* $\bar{x}$ *tal que :*

$$\lim_{j\to+\infty}[\langle F(x^{k_j+1}), z_{k_j}\rangle + \delta_{\Omega_{k_j}}(x^{k_j+1})] = \langle F(\bar{x}), \bar{z}\rangle \ e \lim_{j\to+\infty} g_{k_j} = 0,$$

onde $g_k \in \hat{\partial}\left(\langle F(.), z_k\rangle + \delta_{\Omega_k}\right)(x^{k+1})$, *e* $\bar{z} \in \mathbb{R}^m_+ \backslash \{0\}$ *é o ponto de convergência de* $\{z_{k_j}\}$ *.*

Prova: Do Teorema 3.1.1, de (3.3) e da Proposição 1.1.7, (iii), existe $g_k \in \hat{\partial}\left(\langle F(.), z_k\rangle + \delta_{\Omega_k}\right)(x^{k+1})$ tal que:

$$g_k = \beta_k(x^k - x^{k+1}) \tag{3.19}$$

onde $\beta_k = \alpha_k \langle e_k, z_k\rangle > 0$, $\forall\ k \in \mathbb{N}$. Desde que $0 < \alpha_k < \tilde{\alpha}$, então

$$|\beta_k| \leq \tilde{\alpha}\|e_k\|\|z_k\| = \tilde{\alpha}. \tag{3.20}$$

Da Proposição 3.1.3, $\lim_{k\to+\infty} \left\|x^{k+1} - x^k\right\| = 0$, e de 3.20, temos:

$$0 \leq \|g_k\| = |\beta_k| \left\|x^{k+1} - x^k\right\| \leq \tilde{\alpha} \left\|x^{k+1} - x^k\right\|$$

E portanto, $\lim_{k\to+\infty} g_k = 0$.

Temos que :

- Existe $\bar{x} \in E$ tal que $\lim_{k\to+\infty} x^k = \bar{x}$.
- F é uma aplicação contínua, então, para cada $i = 1, ..., m$, $\lim_{k\to+\infty} F_i(x^k) = F_i(\bar{x})$.
- $\{z_k\}$ é limitada, logo existe $\bar{z} \in \mathbb{R}^m_+ \backslash \{0\}$ tal que $\lim_{j\to+\infty} z_{k_j} = \bar{z}$.

Assim existem sequências $\{x^{k_j}\}$, $\{\langle F(x^{k_j}), z_{k_j}\rangle)\}$ e $\{g_{k_j}\}$, com $g_{k_j} \in \hat{\partial}\left(\left\langle F(.), z_{k_j}\right\rangle + \delta_{\Omega_{k_j}}\right)(x^{k_j+1})$ tais que:

- $\lim_{j\to+\infty} x^{k_j+1} = \bar{x}$.
- $\lim_{j\to+\infty}[\langle F(x^{k_j+1}), z_{k_j}\rangle + \delta_{\Omega_{k_j}}(x^{k_j+1})] = \lim_{j\to+\infty}\langle F(x^{k_j+1}), z_{k_j}\rangle = \langle F(\bar{x}), \bar{z}\rangle$.
- $\lim_{j\to+\infty} g^{k_j} = 0$.

■

3.1.3 Análise de convergência- caso continuamente diferenciável

Observe que se a função F é continuamente diferenciável, então os algoritmos (**MEPPC**) e (**MEPPF**) são os mesmos e por isso, usando os resultados do capítulo anterior, imediatamente obtemos a convergência da sequência dada pelo algoritmo **MEPPF** para um ponto Pareto Crítico. Embora esta observação, por questões didáticas e de completude, obteremos a seguir este resultado a partir do algoritmo (**MEPPF**).

A partir de agora assume-se a seguinte hipótese:

($\mathbf{C_4}$) F é uma aplicação continuamente diferenciável em $\mathbb{R}^n$.

Neste caso, com esta hipótese, os resultados desta subseção, são um caso particular do que foi feito no Capítulo 2.
Desta forma, mostramos que a sequência gerada pelo algoritmo **MEPPF** converge para pontos Pareto críticos, sob a hipótese de que a sequência dos parâmetros de regularização é limitada.

O próximo resultado caracteriza em E uma função vetorial quase-convexa e diferenciável.

Proposição 3.1.5 *Seja $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação diferenciável, que satisfaz as hipóteses* ($\mathbf{C_2}$), ($\mathbf{C_3}$), *e $x \in E$. Então $\left\langle \nabla F_i(x^k), x - x^k \right\rangle \leq 0$, $\forall\ k \in \mathbb{N}$ e $\forall\ i \in \{1, ..., m\}$.*

Prova: Como F é $\mathbb{R}^m_+$-quase-convexa, então cada função componente, F_i, $i = 1, ..., m$, é quase-convexa. Então o resultado segue pela caracterização de funções escalares diferenciáveis quase-convexas (conferir Mangasarian [28],p.134). ■

Convergência para ponto Pareto Crítico

Teorema 3.1.4 *Considere $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação satisfazendo as hipóteses* ($\mathbf{C_2}$), ($\mathbf{C_3}$) *e* ($\mathbf{C_4}$). *Se $0 < \alpha_k < \tilde{\alpha}$, então a sequência $\{x^k\}$ gerada pelo algoritmo* **MEPPF**, (3.2) *e* (3.3) *converge para um ponto Pareto crítico do problema* (3.1).

Prova: Considere $\left\{x^k\right\}$ uma sequência gerada pelo algoritmo **MEPPF**. Visto que esta sequência é Fejér convergente em E, com uma subsequência convergente para algum ponto de E, então existe $\widehat{x} \in E$ tal que $\lim_{k \to +\infty} x^k = \widehat{x}$. Do Teorema 3.1.1 e de (3.3), temos que

$$0 \in \hat{\partial} \left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \ . \ - x^k \|^2 + \delta_{\Omega_k}(.) \right) (x^{k+1})$$

A função $\left(\langle F(.), z_k \rangle + \dfrac{\alpha_k}{2} \langle e_k, z_k \rangle \| \ . \ - x^k \|^2 \right)$ é continuamente diferenciável em $\mathbb{R}^n$, logo, em uma vizinhança de x^{k+1}. Além disto, visto que $x_{k+1} \in \Omega_k$ e Ω_k é fechado, então $\delta_{\Omega_k}(x^{k+1}) = 0$, e portanto finita. Assim, pela Proposição 1.1.7, item (iii), temos que, em uma vizinhança de x^{k+1}

$$0 \in \nabla \left(\langle F(.), z_k \rangle \right) (x^{k+1}) + \alpha_k \langle e_k, z_k \rangle \left(x^{k+1} - x^k \right) + \mathcal{N}_{\Omega_k}(x^{k+1})$$

onde $\mathcal{N}_{\Omega_k}(x^{k+1})$ é o cone normal no ponto x^{k+1} em relação ao conjunto Ω_k.
Assim, existe $\nu_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ tal que:

$$0 = \sum_{i=1}^{m} \nabla F_i(x^{k+1})(z_k)_i + \alpha_k \langle e_k, z_k \rangle \left(x^{k+1} - x^k \right) + \nu_k \tag{3.21}$$

Visto que $\nu_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ então

$$\left\langle \nu_k \ , \ x - x^{k+1} \right\rangle \leq \ 0, \ \forall \ x \in \Omega_k \tag{3.22}$$

Tome $\bar{x} \in E$. Por definição $E \subset \Omega_k, \forall \ k \in \mathbb{N}$, logo $\bar{x} \in \Omega_k \ , \forall \ k \in \mathbb{N}$. Combinando (3.22) com $x = \bar{x}$ e (3.21), temos que:

$$\left\langle \sum_{i=1}^{m} \nabla F_i(x^{k+1})(z_k)_i, \bar{x} - x^{k+1} \right\rangle + \alpha_k \langle e_k, z_k \rangle \left\langle x^{k+1} - x^k, \bar{x} - x^{k+1} \right\rangle \geq 0 \tag{3.23}$$

Além disto, considere uma sequência $\{z_k\} \subset \mathbb{R}^m_+ \backslash \{0\}$ com $\|z_k\| = 1$. Visto que a sequência $\{z_k\}$ é limitada, então existe uma subquência $\left\{ z_{k_j} \right\}_{j \in \mathbb{N}}$, tal que $\lim\limits_{j \to +\infty} z_{k_j} = \bar{z}$, com $\bar{z} \in \mathbb{R}^m_+ \backslash \{0\}$. Assim, a desigualdade em (3.23) torna-se

$$\left\langle \sum_{i=1}^{m} \nabla F_i(x^{k_j+1})(z_{k_j})_i, \bar{x} - x^{k_j+1} \right\rangle + \alpha_{k_j} \left\langle e_{k_j}, z_{k_j} \right\rangle \left\langle x^{k_j+1} - x^{k_j}, \bar{x} - x^{k_j+1} \right\rangle \geq 0 \tag{3.24}$$

Pela Proposição 3.1.3, $\lim\limits_{k \to +\infty} \left\| x^{k+1} - x^k \right\| = 0$. As sequências $\left\{ x^k \right\}$ e $\{e_k\}$ são limitadas. E também a sequência $\{\alpha_k\}$ é limitada. Utilizando a desigualdade de Cauchy Schwartz, obtemos: $\left| \alpha_{k_j} \left\langle e_{k_j}, z_{k_j} \right\rangle \left\langle x^{k_j+1} - x^{k_j}, \bar{x} - x^{k_j+1} \right\rangle \right| \to 0$ quando $j \to +\infty$, e por F ser continuamente diferenciável, a desigualdade em (3.24), para qualquer $\bar{x} \in E$, torna-se:

$$\left\langle \sum_{i=1}^{m} \nabla F_i(\widehat{x}) \bar{z}_i \ , \ \bar{x} - \widehat{x} \right\rangle \geq 0 \Rightarrow \sum_{i=1}^{m} \bar{z}_i \left\langle \nabla F_i(\widehat{x}) \ , \ \bar{x} - \widehat{x} \right\rangle \geq 0 \tag{3.25}$$

Pela quase-convexidade de cada função componente, F_i, para todo $i \in \{1, ..., m\}$, temos que $\langle \nabla F_i(\widehat{x}) \ , \ \bar{x} - \widehat{x} \rangle \leq 0$. E pelo fato de $\bar{z} \in \mathbb{R}^m_+ \backslash \{0\}$, de (3.25), temos que

$$\sum_{i=1}^{m} \bar{z}_i \left\langle \nabla F_i(\widehat{x}) \ , \ \bar{x} - \widehat{x} \right\rangle = 0 \tag{3.26}$$

Sem perda de generalidade, considere o conjunto $J = \{i \in I : \bar{z}_i > 0\}$, onde $I = \{1, ..., m\}$. Portanto, de (3.26), para todo $\bar{x} \in E$ temos

$$\langle \nabla F_i(\widehat{x}) \ , \ \bar{x} - \widehat{x} \rangle = 0, \ \forall \ \ i \in J. \tag{3.27}$$

Mostraremos agora que $\widehat{x}$ é ponto Pareto crítico em $\mathbb{R}^n$.
Suponha por contradição que $\widehat{x}$ não é um ponto Pareto crítico em $\mathbb{R}^n$, então existe uma direção $v \in \mathbb{R}^n$ tal que : $JF(\widehat{x})v \in -\mathbb{R}^m_{++}$, isto é,

$$\langle \nabla F_i(\widehat{x}), v \rangle < 0, \forall \ i \in \{1, ..., m\}\,. \tag{3.28}$$

Logo v é uma direção de descida para a função multiobjetivo F em $\widehat{x}$, e portanto, $\exists \ \epsilon > 0$ tal que

$$F(\widehat{x} + \lambda v) \prec F(\widehat{x}), \ \forall \ \lambda \in (0, \epsilon]. \tag{3.29}$$

Como $\widehat{x} \ \in \ E$, então por (3.29) concluímos que $\widehat{x} + \lambda v \in E$. Assim, de (3.27) com $\bar{x} = \widehat{x} + \lambda v$, obtemos: $\langle \nabla F_i(\widehat{x}) \ , \ \widehat{x} + \lambda v - \widehat{x} \rangle = \langle \nabla F_i(\widehat{x}) \ , \ \lambda v \rangle = \lambda \langle \nabla F_i(\widehat{x}) \ , \ v \rangle = 0$.

Como $\lambda > 0$, concluímos que $\langle \nabla F_i(\widehat{x}) \ , \ v \rangle = 0$, para todo $i \in J$, contradizendo (3.28). Logo $\widehat{x}$ é ponto Pareto crítico para o Problema (3.1). ■

3.2 Um algoritmo proximal inexato

Nesta seção apresentamos uma versão inexata do algoritmo **MEPPF**, o qual denotamos por **MEPPFI**.

3.2.1 O algoritmo MEPPFI

Seja $F : \mathbb{R}^n \to \mathbb{R}^m$ uma aplicação que satisfaz às hipóteses $\mathbf{C_2}$ e $\mathbf{C_4}$, e considere três sequências: a sequência de termos reais positivos dos parâmetros proximais $\{\alpha_k\}$, a sequência $\{e_k\} \subset \mathbb{R}^m_{++}$ tal que $\|e_k\| = 1$, e a sequência $\{z_k\} \subset \mathbb{R}^m_{+} \backslash \{0\}$ tal que $\|z_k\| = 1$.

Este método gera uma sequência $\left\{x^k\right\}_{k \in \mathbb{N}}$ pela seguinte recursão:

Inicialização: Escolha um ponto inicial arbitrário

$$x^0 \in \mathbb{R}^n \tag{3.30}$$

Passo principal: Dado x^k, defina a função $\Psi_k : \mathbb{R}^n \to \mathbb{R}$ tal que $\Psi_k(x) = \langle F(x), z_k \rangle$, e considere $\Omega_k = \left\{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\right\}$. Tome como x^{k+1} qualquer vetor $x \in \Omega_k$ tal que existe $\varepsilon_k \geq 0$, satisfazendo:

$$0 \in \hat{\partial}_{\epsilon_k} \Psi_k(x) + \alpha_k \langle e_k, z_k \rangle \left(x - x^k\right) + \mathcal{N}_{\Omega_k}(x), \tag{3.31}$$

$$\lim_{k \to +\infty} \varepsilon_k = 0 \qquad . \tag{3.32}$$

3.2.2 Existência das Iteradas

Nesta subseção, garantimos que a sequência gerada pelo algoritmo **MEPPFI** está bem definida.

Proposição 3.2.1 (*Existência das iterações*) *Seja* $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação satisfazendo* ($\mathbf{C_4}$). *Então a sequência* $\left\{x^k\right\}$, *gerada pelo algoritmo* **MEPPFI**, *dado por* (3.31) *e* (3.32), *está bem definida.*

Prova: Seja $x^0 \in \mathbb{R}^n$ um ponto escolhido inicialmente. Dado x^k, mostraremos que existe x^{k+1} satisfazendo a condição (3.31). Para isto defina a função $\varphi_k(x) = \langle F(x), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \left\|x - x^k\right\|^2 + \delta_{\Omega_k}(x) = \Psi_k(x) + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \left\|x - x^k\right\|^2$. De forma análoga demonstrada no Teorema 3.1.1 existe $x^{k+1} \in \Omega_k$, não necessariamente único devido à não convexidade da F, o qual é um mínimo global de $\varphi_k(.)$, assim, a Proposição 1.1.6, Observação 1.1.2, garante que x^{k+1} satisfaz

$$0 \in \hat{\partial} \left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \;.\; - x^k \|^2 + \delta_{\Omega_k}(.) \right) (x^{k+1}).$$

Da Proposição 1.1.7 (*iii*), conclui-se que

$$0 \in \hat{\partial} \Psi_k(x^{k+1}) + \alpha_k \langle e_k, z_k \rangle \left(x^{k+1} - x^k\right) + \mathcal{N}_{\Omega_k}(x^{k+1}).$$

Pela Observação 1.1.3, x^{k+1} satisfaz (3.31) para $\varepsilon_k = 0$, o qual satisfaz trivialmente (3.32). ■

A seguir mostraremos que a sequência gerada pelo algoritmo **MEPPFI** é limitada. Para isto, são necessárias hipóteses adicionais sobre as sequências $\{\varepsilon_k\}$, $\{\nu_k\}$, onde esta última é a sequência dos vetores que pertencem ao cone normal em relação ao conjunto Ω_k, e a distância euclidiana.

Hipótese ($\mathbf{B_1}$): $\displaystyle\sum_{k=1}^{\infty} \delta_k < \infty$, onde $\delta_k = \max \left\{ \frac{\varepsilon_k}{\beta_k}, \frac{\|v_k\|}{\beta_k} \right\}$.

Observação 3.2.1 *Considere* $h_1(x) = \sum_{i=1}^{n} x_i^2 = \|x\|^2 = \langle x, x\rangle$. *Logo* $\nabla h_1(x) = 2x$. *A distância de Bregman definida por* h_1, D_{h_1} *é:*

$$\begin{aligned} D_{h_1}(x,y) &= h_1(x) - h_1(y) - \langle \nabla h_1(y), x - y\rangle \\ &= \|x\|^2 - \|y\|^2 - 2\langle y, x - y\rangle \\ &= \|x\|^2 - 2\langle y, x\rangle + \|y\|^2 = \|x - y\|^2. \end{aligned}$$

Por Kaplan e Tichatschke [24], para cada $x \in \mathbb{R}^n$, existem constantes $\gamma(x) > 0$ e $c(x)$ tal que:

$$\|x - z\|^2 + c(x) \geq \gamma(x)\|x - z\|, \; z \in \mathbb{R}^n$$

Proposição 3.2.2 *Seja* $\{x^k\}$ *uma sequência gerada pelo algoritmo* **MEPPFI**. *Se as hipóteses* $(\mathbf{C_2})$, $(\mathbf{C_3})$, $(\mathbf{C_4})$ *e* $(\mathbf{B_1})$ *são satisfeitas, então para cada* $\hat{x} \in E$, $\{\|\hat{x} - x^k\|^2\}$ *é uma sequência convergente.*

Prova: Seja $\{x^k\}_{k \in N}$ uma sequência gerada pelo algoritmo **MEPPFI**,. Pela hipótese $(\mathbf{C_3})$ temos que o conjunto E é não vazio. Por (3.31), existem $g_k \in \hat{\partial}_{\varepsilon_k} \Psi_k(x^{k+1})$ e $\nu_k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ tais que

$$0 = g_k + \beta_k \left(x^{k+1} - x^k\right) + \nu_k, \text{ onde } \beta_k = \alpha_k \langle e_k, z_k\rangle > 0.$$

Segue que para todo $x \in \mathbb{R}^n$ tem-se:

$$\langle -g_k, x - x^{k+1}\rangle + \beta_k \langle x^k - x^{k+1}, x - x^{k+1}\rangle = \langle \nu_k, x - x^{k+1}\rangle \leq \|\nu_k\|\|x - x^{k+1}\|$$

E assim,

$$\langle x^k - x^{k+1}, x - x^{k+1}\rangle \leq \frac{1}{\beta_k}\left(\langle g_k, x - x^{k+1}\rangle + \|\nu_k\|\|x - x^{k+1}\|\right). \tag{3.33}$$

Observe que, $\forall \; x \in \mathbb{R}^n$:

$$\begin{aligned} \left\|x - x^k\right\|^2 &= \left\|x - x^{k+1}\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\left\langle x - x^{k+1}, x^{k+1} - x^k\right\rangle \\ \left\|x - x^k\right\|^2 &- \left\|x - x^{k+1}\right\|^2 \geq \; 2\left\langle x^{k+1} - x^k, x - x^{k+1}\right\rangle \\ \left\|x - x^{k+1}\right\|^2 &- \left\|x - x^k\right\|^2 \leq \; 2\left\langle x^k - x^{k+1}, x - x^{k+1}\right\rangle \end{aligned} \tag{3.34}$$

Substituindo (3.33), em (3.34), obtemos:

$$\left\|x - x^{k+1}\right\|^2 - \left\|x - x^k\right\|^2 \leq \frac{2}{\beta_k}\left(\langle g_k, x - x^{k+1}\rangle + \|\nu_k\|\|x - x^{k+1}\|\right). \tag{3.35}$$

Por outro lado, sendo $\Psi_k(x) = \langle F(x), z_k \rangle$, onde $F : \mathbb{R}^n \to \mathbb{R}^m$ é continuamente diferenciável, então $\Psi_k : \mathbb{R}^n \to \mathbb{R}^m$ é continuamente diferenciável com derivada dada por $\nabla\Psi_k$. Pela Proposição 1.1.9 obtemos:

$$\partial_{\varepsilon_k}\Psi_k(x) = \nabla\Psi_k(x) + \varepsilon_k B, \tag{3.36}$$

onde B é a bola unitária fechada em $\mathbb{R}^n$, centrada em zero.

Além disso, $\hat{\partial}_{\varepsilon_k}\Psi_k(x) \subset \partial_{\varepsilon_k}\Psi_k(x)$, (conferir (2.12)em Jofré et al. [23]). Como $g_k \in \hat{\partial}_{\epsilon_k}\Psi_k(x^{k+1})$, temos que $g_k \in \partial_{\epsilon_k}\Psi_k(x^{k+1})$, então $g_k = \nabla\Psi_k(x^{k+1}) + \varepsilon_k h_k$, com $\|h_k\| \leq 1$.

Seja agora $\hat{x} \in$ E fixo. Assim

$$\begin{aligned}
\langle g_k, \hat{x} - x^{k+1} \rangle &= \left\langle \nabla\Psi_k(x^{k+1}) + \varepsilon_k h_k \ , \ \hat{x} - x^{k+1} \right\rangle \\
&= \sum_{i=1}^{m} \left\langle \nabla F_i(x^{k+1}) \ , \ \hat{x} - x^{k+1} \right\rangle (z_k)_i + \varepsilon_k \left\langle h_k \ , \ \hat{x} - x^{k+1} \right\rangle
\end{aligned} \tag{3.37}$$

Pela Proposição 3.1.5 concluímos que (3.37) torna-se:

$$\langle g_k, \hat{x} - x^{k+1} \rangle \leq \varepsilon_k \left\langle h_k \ , \ \hat{x} - x^{k+1} \right\rangle \leq \varepsilon_k \|\hat{x} - x^{k+1}\| \tag{3.38}$$

Agora, pela Observação 3.2.1 com $x = \hat{x}$ e $z = x^{k+1}$, segue que

$$\|\hat{x} - x^{k+1}\| \leq \frac{1}{\gamma(\hat{x})} \left(\|\hat{x} - x^{k+1}\|^2 + c(\hat{x}) \right) \tag{3.39}$$

Considere $x = \hat{x}$ em (3.35), utilize (3.38), (3.39) e a hipótese $(\mathbf{B_1})$ para obtermos

$$\begin{aligned}
\left\|\hat{x} - x^{k+1}\right\|^2 - \left\|\hat{x} - x^{k}\right\|^2 &\leq \frac{2}{\beta_k} \left(\varepsilon_k + \|\nu_k\|\right) \|\hat{x} - x^{k+1}\| \\
&\leq \frac{2}{\gamma(\hat{x})} \frac{\varepsilon_k}{\beta_k} \left\|\hat{x} - x^{k+1}\right\|^2 + \frac{2c(\hat{x})}{\gamma\,(\hat{x})} \frac{\|\nu_k\|}{\beta_k} \\
&\leq \frac{2\delta_k}{\gamma(\hat{x})} \left\|\hat{x} - x^{k+1}\right\|^2 + \frac{2c(\hat{x})}{\gamma\,(\hat{x})} \delta_k.
\end{aligned} \tag{3.40}$$

Logo

$$\left\|\hat{x} - x^{k+1}\right\|^2 \leq \left(1 - \frac{2\delta_k}{\gamma(\hat{x})}\right)^{-1} \left[\left\|\hat{x} - x^{k}\right\|^2 + \frac{2c(\hat{x})}{\gamma\,(\hat{x})} \delta_k \right]. \tag{3.41}$$

A hipótese ($\mathbf{B_1}$) garante que

$$\frac{\delta_k}{\gamma(\hat{x})} < \frac{1}{4}, \quad k > k_0,$$

para k_0 suficientemente grande. Assim,

$$1 \leq \left(1 - \frac{2\delta_k}{\gamma(\hat{x})}\right)^{-1} \leq \left(1 + \frac{4\delta_k}{\gamma(\hat{x})}\right) < 2, \quad \text{para } \ k \geq k_0,$$

o que combinando com (3.41), resulta em

$$\left\|\hat{x} - x^{k+1}\right\|^2 \leq \left(1 + \frac{4\delta_k}{\gamma(\hat{x})}\right) \left\|\hat{x} - x^k\right\|^2 + \frac{4c(\hat{x})}{\gamma\,(\hat{x})}\delta_k. \tag{3.42}$$

■

Desde que $\displaystyle\sum_{i=1}^{\infty} \delta_k < \infty$, aplicando o Lema 1.1.1 à desigualdade (3.42) , obtém-se a convergência de $\{\left\|\hat{x} - x^k\right\|^2\}$, para cada $\hat{x} \in E$.

Observe que a convergência da sequência $\{\left\|\hat{x} - x^k\right\|^2\}$ implica na convergência desta sequência $\{\left\|\hat{x} - x^k\right\|\}$, e portanto na sua limitação, isto é, existe $M \in \mathbb{R}_+$, tal que

$$\left\|\hat{x} - x^k\right\| \leq M, \quad \forall\ k \in \mathbb{N}. \tag{3.43}$$

Desde que

$$\|x^k\| \leq \|x^k - \hat{x}\| + \|\hat{x}\|,$$

concluímos que $\{x^k\}$ é limitada, e por sua vez garantimos que o conjunto de pontos de acumulação desta sequência é não vazio.

3.2.3 Análise de convergência do MEPPFI

Proposição 3.2.3 *(***Convergência para algum ponto de E)**
Se as hipóteses ($\mathbf{C_2}$)*,* ($\mathbf{C_3}$)*,* ($\mathbf{C_4}$) *e* ($\mathbf{B_1}$) *são satisfeitas, então a sequência* $\left\{x^k\right\}$ *gerada pelo algoritmo* **MEPPFI** *converge para algum ponto do conjunto* E*.*

Prova: Da Proposição 3.2.2 concluímos que a sequência $\left\{x^k\right\}$ é limitada, então existe uma subsequência $\left\{x^{k_j}\right\}$ tal que $\lim\limits_{j\to+\infty} x^{k_j} = \widehat{x}$. Desde que F é contínua em $\mathbb{R}^n$, então a função $\langle F(.), z\rangle$ é também contínua em $\mathbb{R}^n$ para todo $z \in \mathbb{R}^m$, em particular, para todo $z \in \mathbb{R}^m_+ \backslash \{0\}$, e $\langle F(\widehat{x}), z\rangle = \lim\limits_{j\to+\infty} \left\langle F(x^{k_j}), z\right\rangle$. Por outro lado, como $x^{k+1} \in \Omega_k$, temos que $F(x^{k+1}) \preceq F(x^k)$, e visto que $z \in \mathbb{R}^m_+ \backslash \{0\}$, concluímos que $\left\langle F(x^{k+1}), z\right\rangle \leq \left\langle F(x^k), z\right\rangle$. Além disso, da Observação 3.1.1, podemos assumir que a função $\langle F(.), z\rangle$ é limitada inferiormente, para cada $z \in \mathbb{R}^m_+ \backslash \{0\}$. Então a sequência $\left\{\left\langle F(x^k), z\right\rangle\right\}$ é não crescente e limitada inferiormente, portanto convergente. Assim,

$$\langle F(\widehat{x}), z\rangle = \lim_{j\to+\infty}\left\langle F(x^{k_j}), z\right\rangle = \lim_{j\to+\infty}\left\langle F(x^k), z\right\rangle = inf_{k\in\mathbb{N}}\left\{\left\langle F(x^k), z\right\rangle\right\} \le \left\langle F(x^k), z\right\rangle.$$

Logo $\left\langle F(x^k) - F(\widehat{x}), z\right\rangle \ge 0, \forall\ k \in \mathbb{N}, \forall\ z \in \mathbb{R}^m_+ \backslash \{0\}$. Concluímos que $F(x^k) - F(\widehat{x}) \in \mathbb{R}^m_+$, isto é, $F(\widehat{x}) \preceq F(x^k), \forall\ k \in \mathbb{N}$. Portanto $\widehat{x} \in E$. Agora, pela Proposição 3.2.2 temos que a sequência $\{\left\|\widehat{x} - x^k\right\|\}$ é convergente, e visto que $\lim_{k\to+\infty}\left\|x^{k_j} - \hat{x}\right\| = 0$, concluímos que $\lim_{k\to+\infty}\left\|x^k - \widehat{x}\right\| = 0$, isto é, $\lim_{k\to+\infty} x^k = \widehat{x}$. ■

Teorema 3.2.1 *Suponha que as hipóteses* $(\mathbf{C_2})$, $(\mathbf{C_3})$, $(\mathbf{C_4})$ *e* $(\mathbf{B_1})$ *são satisfeitas. Se* $0 < \alpha_k < \widetilde{\alpha}$, *então a sequência* $\{x^k\}$ *gerada pelo algoritmo* **MEPPFI**, (3.30), (3.31) *e* (3.32), *converge para um ponto Pareto crítico do problema* (3.1).

Prova: Considere $\left\{x^k\right\}$ uma sequência gerada pelo algoritmo **MEPPFI**. Pela Proposição 3.2.3 existe $\widehat{x} \in E$ tal que $\lim_{j\to+\infty} x^k = \widehat{x}$. Além disto, como a sequência $\left\{z^k\right\}$ é limitada, então $\exists\ \left\{z^{k_j}\right\}_{j\in\mathbb{N}}$ tal que $\lim_{j\to+\infty} z^{k_j} = \bar{z}$, com $\bar{z} \in \mathbb{R}^m_+\backslash\{0\}$. Por (3.31) existe $g_{k_j} \in \hat{\partial}_{\varepsilon_{k_j}}\Psi_{k_j}(x^{k_j+1})$, tal que $g_{k_j} = \nabla\Psi_{k_j}(x^{k_j+1}) + \varepsilon_{k_j}h_{k_j}$ com $\|h_{k_j}\| \le 1$, e $\nu_{k_j} \in \mathcal{N}_{\Omega_{k_j}}(x^{k_j+1})$, tal que:

$$0 = \sum_{i=1}^{m} \nabla F_i(x^{k_j+1})(z_{k_j})_i + \varepsilon_{k_j}h_{k_j} + \alpha_{k_j}\left\langle e_{k_j}, z_{k_j}\right\rangle\left(x^{k_j+1} - x^{k_j}\right) + \nu_{k_j} \tag{3.44}$$

Visto que $\nu_{k_j} \in \mathcal{N}_{\Omega_{k_j}}(x^{k_j+1})$ então,

$$\left\langle \nu_{k_j}\ ,\ x - x^{k_j+1}\right\rangle \le\ 0,\ \forall\ x \in \Omega_{k_j} \tag{3.45}$$

Tome $\bar{x} \in E$. Por definição $E \subset \Omega_k, \forall\ k \in \mathbb{N}$, logo $\bar{x} \in \Omega_{k_j}$. Combinando (3.45) com $x = \bar{x}$ e (3.44), temos que:

$$\begin{aligned} 0 &\le \left\langle \sum_{i=1}^{m} \nabla F_i(x^{k_j+1})(z_{k_j})_i\ ,\ \bar{x} - x^{k_j+1}\right\rangle + \varepsilon_{k_j}\left\langle h_{k_j}\ ,\ \bar{x} - x^{k_j+1}\right\rangle + \\ &+ \alpha_{k_j}\left\langle e_{k_j}, z_{k_j}\right\rangle\left\langle x^{k_j+1} - x^{k_j}\ ,\ \bar{x} - x^{k_j+1}\right\rangle \\ &\le \left\langle \sum_{i=1}^{m} \nabla F_i(x^{k_j+1})(z_{k_j})_i\ ,\ \bar{x} - x^{k_j+1}\right\rangle + \varepsilon_{k_j}M + \tilde{\alpha}\left\langle x^{k_j+1} - x^{k_j}\ ,\ \bar{x} - x^{k_j+1}\right\rangle \end{aligned} \tag{3.46}$$

Observe que, $\forall\ x \in \mathbb{R}^n$:

$$\begin{aligned} \left\|x - x^k\right\|^2 &= \left\|x - x^{k+1}\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\left\langle x - x^{k+1}, x^{k+1} - x^k\right\rangle \\ \left\|x^{k+1} - x^k\right\|^2 &= \left\|x - x^k\right\|^2 - \left\|x - x^{k+1}\right\|^2 + 2\left\langle x^k - x^{k+1}, x - x^{k+1}\right\rangle \end{aligned} \tag{3.47}$$

Agora de (3.33) com $x = \bar{x} \in E$, e (3.38), obtemos

$$\left\langle x^k - x^{k+1}, \bar{x} - x^{k+1}\right\rangle \le \left\|\bar{x} - x^{k+1}\right\|\left(\frac{\varepsilon_k}{\beta_k} + \frac{\|\nu_k\|}{\beta_k}\right) \le 2M\delta_k$$

Assim, de (3.47), com $x = \bar{x}$ obtemos

$$0 \leq \left\|x^{k+1} - x^k\right\|^2 \leq \left\|\bar{x} - x^k\right\|^2 - \left\|\bar{x} - x^{k+1}\right\|^2 + 4M\delta_k \tag{3.48}$$

Visto que a sequência $\{\left\|\bar{x} - x^k\right\|\}$ é convergente e que $\sum_{i=1}^{\infty} \delta_k < \infty$, de (3.48) concluímos que $\lim_{k \to +\infty} \left\|x^{k+1} - x^k\right\| = 0$. Além disto, como

$$0 \leq \left\|x^{k_j+1} - \bar{x}\right\| \leq \|x^{k_j+1} - x^{k_j}\| + \|x^{k_j} - \bar{x}\|, \tag{3.49}$$

concluímos que a sequência $\{\left\|\bar{x} - x^{k_j+1}\right\|\}$ é limitada.

Assim, voltando a (3.46), visto que $\lim_{k \to +\infty} \varepsilon_k = 0$, $\lim_{j \to +\infty} x^k = \widehat{x}$ e $\lim_{j \to +\infty} z^{k_j} = \bar{z}$, tomando o limite em (3.46) quando $j \to +\infty$, obtemos

$$\sum_{i=1}^{m} \bar{z}_i \left\langle \nabla F_i(\widehat{x}) \ , \ \bar{x} - \widehat{x} \right\rangle \geq 0 \tag{3.50}$$

Portanto, de forma análoga como demonstrado no Teorema 3.1.4, a partir de (3.25) conclui-se que $\widehat{x}$ é ponto Pareto crítico para o problema de otimização multiobjetivo irrestrito (3.1). ■

Capítulo 4

Conclusão

Este capítulo refere-se às considerações finais, e uma projeção de trabalhos futuros que dá prosseguimento a esta pesquisa.

4.0.4 Considerações finais

Neste trabalho propõe-se estudar o problema

$$\min\{F(x) : x \in \mathbb{R}^n\}$$

onde $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma função vetorial quase-convexa, não necessariamente diferenciável, no espaço Euclidiano $\mathbb{R}^n$.

Propõe-se dois métodos de ponto proximal de valor escalar para este problema, um usando o subdiferencial de Clarke e outro usando o subdiferencial de Frechét, isto é justificado pois no caso não convexo e não diferenciável não existe na atualidade um conceito de subdiferencial apropriado útil para garantir convergência dos métodos proximais. Depois analisa-se a convergência das sequências geradas por estes métodos para pontos Pareto críticos e pontos Pareto fracos. Em geral os resultados obtidos mantém as propiedades de convergência global do caso convexo para pontos Pareto críticos do problema, e obtemos os mesmos resultados conhecidos na literatura, no caso convexo.

Consideramos esta pesquisa como uma primeira abordagem para resolver esta classe de problemas usando métodos proximais e nosso algoritmo como base para a construção de outros métodos mais eficientes que considerem erros computacionais e aproximação da função objetivo como os métodos de feixes.

O método de ponto proximal **MEPPC**, apresentado no capítulo 2, tem como propósito resolver problemas de otimização multiobjetivo quando a função é localmente Lipschitz em todo o espaço e obtém convergência fraca a um ponto Pareto crítico para funções objetivo não convexas e convergência forte (convergência global) para funções quase-convexas satisfazendo à hipótese de completude H_3. Esta hipótese é bem natural e utilizada também para a convergência do método proximal no caso convexo. Uma modificação importante do algoritmo, para uma futura implementação computacional, seria garantir que a sequência gerada pelo método pertença ao interior de Ω_k, para que em cada iteração se obtenha simplesmente encontrar $x^{k+1} \in int(\Omega_k)$ tal que:

$$0 \in \partial^o \left(\langle F(.), z_k \rangle + \frac{\alpha_k}{2} \langle e_k, z_k \rangle \| \ . \ - x^k \|^2 \right) (x^{k+1})$$

Isto reduziria muito o custo computacional em cada iteração. Uma outra coisa que pode ser pesquisada é sobre que condições o ponto de convergência do método **MEPPC** é uma solução Pareto ou Pareto fraca. Neste capítulo não apresentamos uma versão inexata, pois, segundo nossa pesquisa, a teoria de ϵ subdifferencial de Clarke não foi desenvolvida.

O método de ponto proximal **MEPPF**, apresentado no capítulo 3, tem os mesmos objetivos do capítulo anterior, mas trabalha para uma classe maior de funções, isto é, para funções arbitrárias não necessariamente localmente Lipschitzianas. Sobre as hipóteses $C1$, $C2$ e $C3$, (funções s.c.i, quase-convexas e que satisfazem a hipótese de completude) se obtém a convergência global do método para um conjunto que contém o conjunto de soluções Pareto fracos. Se além disso a função F é contínua, obtemos a convergência para uma ponto crítico em um sentido generalizado. Para mostrar sua aplicabilidade apresentamos uma versão exata e inexata para o caso quando F é continuamente diferenciável. Prova-se no Terorema 3.2.1 a convergência fraca para um ponto Pareto crítico.

A extensão dos métodos proximais do caso convexo para o caso quase-convexo não é muito simples como aparentemente se possa perceber. A maior dificuldade enfrentada, foi quando se considera funções quase-convexas, pois perde-se algumas propriedades em relação às funções convexas, por exemplo, o uso do subdiferencial de Fenchel e suas propriedades como a desigualdade do subgradiente no sentido da análise convexa, e a perda da continuidade. Por isso de alguma forma impomos esta condição. Além disso existem resultados que só podem ser utilizados quando a função é convexa, por exemplo, Teorema 2.10 do Luc [27]. Assim para contornar isso consideramos algoritmos de valor escalar, e a definição do cone Normal. Além disto, algumas definições, como por exemplo, ponto Pareto crítico para o problema $\min\{F(X) : x \in \mathbb{R}^n\}$, no caso em que F não é diferenciável, utilizando o subdiferencial de Fréchet, pois no caso em que utilizamos o subdiferencial de Clarke, e F é localmente Lipschitz, foi definido o que significa um ponto Pareto-Clarke crítico do problema. Por este motivo, a convergência no Teorema 3.1.3, foi estabelecida em um sentido generalizado.

Os algoritmos propostos nesta tese foram baseados no método proximal inexato para problemas de minimização convexa multiobjetivo estudado por Bonnel et al. [8]. E observamos que o termo $\langle e_k, z_k \rangle$ não desempenha um papel importante para a convergência dos métodos propostos, e portanto podemos omití-lo.

E por fim, observamos que no artigo publicado recentemente do Bento et al. [6] os autores apresentam um algoritmo de ponto proximal para otimização multiobjetivo quase-convexa assumindo um processo iterativo que utiliza uma função de escalarização variável $f : \mathbb{R}^m \longrightarrow \mathbb{R}$, dada por

$$f(y) = \inf\{t \in \mathbb{R} : te \in y + \mathbb{R}^m_+\}, \quad e = (1, ..., 1) \in \mathbb{R}^m$$

a qual satisfaz algumas propriedades. Enquanto os algoritmos propostos nesta tese utiliza-se a escalarização dada pelo produto interno.

4.0.5 Trabalhos futuros

A seguir destacamos algumas possibilidades para continuidade desta pesquisa:

- No caso irrestrito substituir a regularização quadrática por uma regularização de métrica variável para que em cada iteração tenhamos um subproblema mais accesível do ponto de vista computacional.

- Estender a abrangência dos algoritmos propostos para otimização multiobjetivo com restrições usando outro tipo de regularização, como por exemplo: distância de Bregmann, $\varphi-$divergência, distância proximal ou quase-distância;

- Propor uma versão inexata para o algoritmo **MEPPC**;

- Implementação de **MEPPC** e **MEPPF**;

- Estender a abrangência dos algoritmos propostos para aplicações $F : M \to \mathbb{R}^m$ onde M é uma variedade Riemanniana conexa n-dimensional;

- Estabelecer um algoritmo para resolver o problema restrito:

$$\min\{F(x) : x \in S\},$$

onde X é um espaço de Hilbert real, Y um espaço de Banach real, $S \subset X$ e F uma aplicação quase-convexa de X em $Y \cup \{+\infty\}$. Este tipo de problema tem aplicação em Controle Ótimo.

Referências Bibliográficas

[1] **Apolinário, H.C.F., Villacorta, K.D.V. e Oliveira, P.R.**: Um método proximal para problemas de otimização multiobjetivo quase-convexa. SOBRAPO RN/BRASIL. XLV, 2721-2732 (2013).

[2] **Apolinario H.C.F., Papa Quiroz, E.A., e Oliveira, P.R.**: A Scalarization Proximal Point Method for Quasiconvex Multiobjective Minimization, Optimization Online, $http : //www.optimization - online.org/DB_FILE/2014/03/4261.pdf, http : //arxiv.org/abs/1403.0150$.

[3] **Aussel, D.**: Subdifferential properties of quasiconvex and pseudoconvex functions: unified approach. J. Optim. Theory Appl. 97, 1, 29-45, (1998).

[4] **Bazaraa, M.S., Sherali, H.D. e Shetty, C.M.**: Nonlinear Programming: Theory and Algorithms. 3 ed.,John Wiley and Sons, Inc., New York, (2006).

[5] **Bello Cruz, J.Y., Lucambio Pérez, L.R. e Melo, J.G.**: Convergence of the projected gradient method for quasiconvex multiobjective optimization. Nonlinear Analysis. 74, 5268–5273 (2011).

[6] **Bento, G.C., Cruz Neto, J. X. e Soubeyran, A.**: A Proximal Point-Type Method for Multicriteria Optimization. Set-Valued and Variational Analysis, Theory and Applications. 22, DOI 10.1007/s11228-014-0279-2, Springer, (2014).

[7] **Bolte, J., Daniilidis, A. e Lewis, A.**: Clarke subgradients of stratifiable functions. SIAM Journal on Optimization. 18,556-572 (2007).

[8] **Bonnel, H., Iusem, A.N. e Svaiter, B.F.**: Proximal methods in vector optimization. SIAM Journal on Optimization. 15, 953-970 (2005).

[9] **Bello Cruz, J.Y., Lucambio Pérez, L.R. and Melo, J.G.**: Convergence of the projected gradient method for quasiconvex multiobjective optimization. Nonlinear Analysis. 74, 5268–5273 (2011)

[10] **Brito, A.S, da Cruz Neto, J.X, Lopes, J.O e Oliveira, P.R**: Interior proximal algorithm for quasiconvex programming and variational inequalities with linear constraints. J. Optim Theory Appl. 154, 217-234, (2012).

[11] **Ceng, L. e Yao, J.**: Approximate proximal methods in vector optimization. European Journal of Operational Research. 183, 1-19(2007).

[12] **Clarke, H.F.**: Optimization and nonsmooth analysis. Classics in applied mathematic, SIAM, New York (1983).

[13] **Custodio, A.L., Madeira, J.F.A., Vaz, A.I.F e Vicente, L.N.**: Direct Multisearch for multiobjective optimization. SIAM Journal on Optimization. 21, 1109-1140, (2011).

[14] **Da Cruz Neto, J.X., Da Silva, G.J.P., Ferreira, O.P. e Lopes, J.O.**: A subgradient method for multiobjective optimization. Computational Optimization and Applications. 54 (3), 461-472,(2013).

[15] **Fliege, J. e Svaiter, B.F.**: Steepest descent methods for multicriteria optimization. Mathematical Methods of Operations Research. 51, 479-494,(2000).

[16] **Fliege, J., Granã Drummond, L.M. e Svaiter, B.F.**: Newton's Method for Multiobjective Optimization. SIAM Journal on Optimization. 20, 602-626, (2009).

[17] **Granã Drummond, L.M. e Svaiter, B.F.**: A steepest descent method for vector optimization. Journal of Computational and Applied. Mathematics. 175, 395-414, (2005).

[18] **Gregório, R. e Oliveira, P.R.** : A Logarithmic-quadratic proximal point scalarization method for multiobjective programming. Journal of Global Optimization. 49, 361-378 (2010).

[19] **Güler, O.**: New proximal point proximal algorithms for convex minimization. SIAM Journal Control and Optimization. 2, 649–664 (1992).

[20] **Hadjisavvas, N., Komlosi, S. e Shaible, S.**: Handbook of Generalized Convexity and Generalized Monotonicity. Nonconvex Optimization and its Applications 76, Springer-Verlag, New York, (2005).

[21] **Hiriart-Urruty, J.-B. e Lemaréchal, C.**: Convex Analysis and Minimization Algorithms I. Grundlehren der mathematischen Wissenschaften 305, Springer-Verlag Berlin Heidelberg New York, (1993).

[22] **Huang, X.X. e Yang, X.Q.**: Duality for multiobjective optimization via nonlinear Lagrangian functions. Journal of Optimization Theory and Applications. 120, 111-12(2004).

[23] **Jofré, A., Luc, D.T. e Therá, Michel**: $\varepsilon-$Subdifferential and $\varepsilon-$Monotonicity. Nonlinear Analysis, Theory, Methods and Applications. 33, 71-90(1998).

[24] **Kaplan, A. e Tichatschke**: On inexact generalized proximal methods with a weakened error tolerance criterion. Optimization.53, 3–17 (2004). MR 2040631 (2004k:90095)

[25] **Kiwiel, K.C.**: Convergence and efficience of subgradient methods for quasiconvex minimization. Math Program. A 90, 1-25 (2001)

[26] **Langenberg, N. e Tichatschke, R.**: Interior proximal methods for quasiconvex optimization. J Glob Optim. 52, 641-661 (2012).

[27] **Luc, D.T.**: Theory of vector optimization, Lecture Notes in Economics and Mathematical Systems. Springer, Berlin, (1989).

[28] **Mangasarian, O.L.**: Nonlinear Programming. McGraw-Hill, New York, (1969).

[29] **Martinet, B.**: Regularization d'inequations variationelles par approximations sucessives. Révue Française d'informatique et Recherche Opérationelle. 4, 54-159 (1970).

[30] **Martinez-Legaz, J.E.**: Quasiconvex duality theory by generalized conjugation methods. Optimization: A Journal of Mathematical Programming and Operations Research, 19, 603–652 (1988).

[31] **Mas-Colell, A., Whinston, M.D. e Green, J.R.**: Microeconomic theory. Oxford University Press, New York, NY, USA, (1995).

[32] **Miettinen, K.M.**: Nonlinear multiobjective optimization. Kluwer Academic Publishers, Boston, (1999).

[33] **Mordukhovich, B.S.**: Variational analysis and generalized differentiation I: Basic theory. Grundlehren Series[Fundamental Principles of Mathematical Sciences], vol. 330, Springer-Verlag, Berlin, (2006).

[34] **Papa Quiroz, E.A. e Oliveira, P.R.**: Proximal point methods for quasiconvex and convex functions with Bregman distances on Hadamard manifolds. Journal of Convex Analysis. 16, 1, 49-69 (2009).

[35] **Papa Quiroz, E.A. e Oliveira, P.R.**: Full Convergence of the proximal point method for quasiconvex functions on Hadamard manifolds. ESAIM: COCV, 18, 483-500, (2011).

[36] **Papa Quiroz, E.A. e Oliveira, P.R.**: An extension of proximal methods for quasiconvex minimization on the nonnegative orthant. European Journal of Operational Research. 216, 26–32 (2012).

[37] **Papa Quiroz, E.A. e Oliveira, P.R.**: Proximal point method for minimizing quasiconvex locally Lipschitz functions on Hadamard manifolds. Nonlinnear Analysis. 75, 5924-5932, (2012).

[38] **Papa Quiroz, E.A., Apolinario H.C.F., e Oliveira, P.R.** A Proximal Method for Quasiconvex Multiobjective Minimization, Working paper, Abril 2014.

[39] **Penot, J.-P. and Volle, M.**: Dualité de Fenchel et quasiconvexité. C.R. Acad. Sci. Paris. Serie I 304, 371–374 (1987).

[40] **Polyak, B. T.**: Introduction to optimization. Translations Series in Mathematics and Engineering. Optimization Software Inc. Publications Division. Translated from the Russian, With a foreword by Dimitri P. Bertsekas. New York, (1987), MR 1099605 (92b:49001).

[41] **Rockafellar, R.T. e Wets, R.J-B.**: Variational Analysis. Springer, Berlin, (1998).

[42] **Schott, D.**: Basic properties of Fejer monotone sequences. Rostocker Mathematische Kolloquium. 49, 57-74, (1995).

[43] **Treiman, J.S.**: Clarke's gradients and $\varepsilon-$subgradients in Bancach Spaces. Trans. Amer. Math. Soc. . 294, 65-78, (1984).

[44] **Villacorta, K.D.V. e Oliveira, P.R.**: An interior proximal method in vector optimization. European Journal of Operational Research. 214, 485-492, (2011).