

# Método proximal vetorial com regularização quase-distância em programação multiobjetivo

**Rogério Azevedo Rocha**

Universidade Federal do Tocantins  
Curso de Graduação em Ciência da Computação, ALC NO 14 (109 Norte) AV.NS.15 S/N  
CEP 77001-090, Palmas, Brazil  
rogerioar@cos.ufrj.br

**Paulo Roberto Oliveira**

Universidade Federal do Rio de Janeiro  
Departamento de Engenharia de Sistemas e Computação, Caixa Postal 68511  
CEP 21945-970, Rio de Janeiro, Brazil  
poliveir@cos.ufrj.br

**Ronaldo Gregório**

Universidade Federal Rural do Rio de Janeiro  
Departamento de Tecnologia e Linguagens, Rua Capitão Chaves, Nr. 60, Centro, Nova Iguaçu  
CEP 26221-010, Rio de Janeiro, Brazil  
rgregor@ufrj.br

**Michael Souza**

Universidade Federal do Ceará  
Departamento de Estatística e Matemática Aplicada, Campus do Pici,  
CEP 60455-760, Fortaleza, Brazil  
michael@ufc.br

## Resumo

Desenvolvemos um algoritmo proximal vetorial com regularização quase-distância para o problema de otimização multiobjetivo de encontrar soluções Pareto fraco para aplicações convexas de  $\mathbb{R}^n$  em  $\mathbb{R}^m$  com respeito a ordem lexicográfica gerada pelo cone  $\mathbb{R}_+^m$ . Apresentamos uma versão exata e uma versão inexata de nosso algoritmo. Desenvolvemos ainda um método de escalarização que foi fundamental na demonstração de que os pontos de acumulação de qualquer sequência gerada pela versão exata são soluções Pareto fraco.

**Palavras Chave:** Programação multiobjetivo, método proximal, método de escalarização.

**Área principal:** Programação matemática.

## Abstract

We developed a vector proximal algorithm with regularization given by a quasi-distance function to the multiobjective optimization problem of finding weak Pareto solutions to convex applications from  $\mathbb{R}^n$  to  $\mathbb{R}^m$  with respect to the lexicographical order generated by the cone  $\mathbb{R}_+^m$ . We present an exact and an inexact version of our algorithm. We also developed a scalarization method which ensures that all accumulation points of any sequence generated by the exact version are weak Pareto solutions.

**Keywords:** Multiobjective programming, proximal method, scalarization methods.

**Main area:** Mathematical programming.

# 1 Introdução

Bonnell et al. (2005) consideraram o problema de otimização vetorial de encontrar pontos fracamente eficientes para aplicações  $F$ , de um espaço de Hilbert  $X$  para um espaço de Banach  $Y$ , com respeito à ordem parcial  $\preceq_C$  em  $Y$  gerada por um cone  $C \subset Y$ , fechado, convexo, pontudo e com interior não vazio, dada por  $y_1 \preceq_C y_2 \Leftrightarrow y_2 - y_1 \in C$ , com sua relação associada  $\prec_C$  dada por  $y_1 \prec_C y_2 \Leftrightarrow y_2 - y_1 \in \text{int}(C)$ . Um ponto  $a \in X$  é dito ser fracamente eficiente se não existe  $x \in X$  tal que  $F(x) \prec_C F(a)$  (Para mais detalhes sobre problemas de otimização vetorial, conferir Luc [7]). Para este problema, Bonnell et al. (2005) desenvolveram uma extensão do método do ponto proximal com regularização quadrática para otimização convexa com valor escalar. Além do trabalho de Bonnell et al. (2005), destacamos, como métodos de ponto proximal em otimização vetorial, os trabalhos de Ceng e Yao (2007), Gregório e Oliveira (2011) e Villacorta e Oliveira (2011).

Quando  $X = \mathbb{R}^n$ ,  $Y = \mathbb{R}^m$  e  $C$  o octante não negativo de  $\mathbb{R}^m$ , denotado por  $\mathbb{R}_+^m$ , temos uma importante subclasse de problemas de otimização vetorial, conhecida como otimização multiobjetivo (Para mais detalhes, conferir seção 2.3 e Miettinen (1999)).

Nosso objetivo é desenvolver métodos de ponto proximal com regularização quase-distância para encontrar soluções Pareto fraco para o problema de otimização multiobjetivo irrestrito

$$(P) \quad \text{MINIMIZE}\{F(x) \mid x \in \mathbb{R}^n\}, \quad (1)$$

onde  $F = (F_1, \dots, F_m)^T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é uma função convexa, isto é,  $\forall x, y \in \mathbb{R}^n$  e  $\lambda \in (0, 1)$ ,  $F_i(\lambda x + (1 - \lambda)y) \leq \lambda F_i(x) + (1 - \lambda)F_i(y)$ ;  $i = 1, \dots, m$ . Além das hipóteses de convexidade, vamos exigir que uma das funções objetivo seja coerciva, isto é, que existe  $r \in \{1, \dots, m\}$  tal que  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$ .

Vale ressaltar a contribuição de Moreno et al. (2011) pelo desenvolvimento de um método proximal com a regularização dada por uma quase-distância, aplicado a funções mono-objetivo não-convexas e não-diferenciáveis e satisfazendo a propriedade de Kurdyka-Lojasiewicz.

Na seção 2, apresentamos conceitos e resultados sobre a quase-distância, a teoria do subdiferencial e a teoria da programação multiobjetivo. Na seção 3, apresentamos um resultado de escalarização e a versão exata do nosso método onde provamos a existência das iterações e a convergência da sequência gerada. Na seção 4, apresentamos a versão inexata do nosso método. E, finalmente, na seção 5, apresentamos os resultados de alguns experimentos numéricos.

## 2 Preliminares

### 2.1 Quase-distância

**Definição 2.1.1** *Seja  $X$  um conjunto. Uma função  $q : X \times X \rightarrow \mathbb{R}_+$  é uma quase-distância em  $X$  se, para todos  $x, y, z \in X$ , **a**)  $q(x, y) = q(y, x) = 0 \iff x = y$  e **b**)  $q(x, z) \leq q(x, y) + q(y, z)$ .*

**Proposição 2.1.1** *Seja  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  uma quase-distância tal que*

$$\alpha \|x - y\| \leq q(x, y) \leq \beta \|x - y\|, \quad \forall x, y \in \mathbb{R}^n. \quad (2)$$

*Então, para cada  $\bar{z} \in \mathbb{R}^n$ , as funções  $q(\bar{z}, \cdot)$  e  $q(\cdot, \bar{z})$  são lipschitzianas em  $\mathbb{R}^n$ , as funções  $q^2(\bar{z}, \cdot)$  e  $q^2(\cdot, \bar{z})$  são localmente lipschitzianas em  $\mathbb{R}^n$  e as funções  $q(\bar{z}, \cdot)$ ,  $q(\cdot, \bar{z})$ ,  $q^2(\bar{z}, \cdot)$  e  $q^2(\cdot, \bar{z})$  são coercivas.*

**Prova:** Proposições 3.6 e 3.7 e observação 5 de Moreno et al. (2011) ■

Em Moreno et al. (2011), encontramos o seguinte exemplo de uma função quase-distância em  $\mathbb{R}^n$ , convexa em cada um dos seus argumentos e que satisfaz a hipótese da proposição 2.1.1.

**Exemplo 2.1.1** Para cada  $i = 1, \dots, n$ , considere  $c_i^-, c_i^+ > 0$  e  $q_i : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$  definida por

$$q_i(x_i, y_i) = \begin{cases} c_i^+(y_i - x_i) & \text{if } y_i - x_i > 0 \\ c_i^-(x_i - y_i) & \text{if } y_i - x_i \leq 0. \end{cases}$$

Portanto  $q(x, y) = \sum_{i=1}^n q_i(x_i, y_i)$  é uma quase-distância sobre  $\mathbb{R}^n$ .

## 2.2 Teoria do subdiferencial

Como a quase-distância não necessariamente é uma função convexa nem diferenciável em nenhum de seus argumentos, os subproblemas do nosso algoritmo (exato) não necessariamente serão convexos nem diferenciáveis. Portanto, para obtermos as condições de otimalidade, necessárias ao desenvolvimento do nosso trabalho, faz-se necessário o estudo de algum tipo de subdiferencial generalizado.

**Definição 2.2.1** Seja  $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$  uma função própria e semicontínua inferior e seja  $x \in \mathbb{R}^n$ .

1. O subdiferencial Fréchet de  $h$  em  $x \in \mathbb{R}^n$ ,  $\hat{\partial}h(x)$ , é dado por:

$$\hat{\partial}h(x) := \begin{cases} \left\{ x^* \in \mathbb{R}^n : \liminf_{y \neq x, y \rightarrow x} \frac{h(y) - h(x) - \langle x^*, y - x \rangle}{\|x - y\|} \geq 0 \right\}, & \text{if } x \in \text{dom}(h) \\ \emptyset, & \text{if } x \notin \text{dom}(h) \end{cases}$$

2. O subdiferencial-limite de  $h$  em  $x \in \mathbb{R}^n$ ,  $\partial h(x)$ , é dado por:

$$\partial h(x) := \left\{ x^* \in \mathbb{R}^n : \exists x_n \rightarrow x, \quad h(x_n) \rightarrow h(x), \quad x_n^* \in \hat{\partial}h(x_n) \rightarrow x^* \right\}$$

**Observação 2.2.1** Nas condições da definição acima temos: **a)**  $\hat{\partial}h(x) \subset \partial h(x)$  e **b)** Se  $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma função convexa e própria e  $\bar{x} \in \text{dom}(h)$  então  $\hat{\partial}h(\bar{x}) = \partial h(\bar{x}) = \{y \in \mathbb{R}^n : h(z) \geq h(\bar{x}) + \langle y, z - \bar{x} \rangle \forall z \in \mathbb{R}^n\}$ .

**Proposição 2.2.1 (Condição de Otimalidade)** Se uma função própria  $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  possui um mínimo local em  $\bar{x} \in \text{dom}(h)$ , então  $0 \in \hat{\partial}h(\bar{x})$  e  $0 \in \partial h(\bar{x})$ .

**Prova:** Teorema 10.1 de Rockafellar e Wets (1998). ■

**Observação 2.2.2** Seja  $C \subset \mathbb{R}^n$ . Se uma função própria  $h : C \rightarrow \mathbb{R} \cup \{\infty\}$  possui um mínimo local em  $\bar{x} \in C$ , então  $0 \in \hat{\partial}(h + \delta_C)(\bar{x})$ ,  $0 \in \partial(h + \delta_C)(\bar{x})$ , onde  $\delta_C$  é a função indicadora do conjunto  $C$ , definida por  $\delta_C(x) = 0$  se  $x \in C$  e  $\delta_C(x) = \infty$  se  $x \notin C$ .

**Proposição 2.2.2** Se  $f_1$  é localmente lipschitziana em  $\bar{x}$ ,  $f_2$  é semicontínua inferior e própria com  $f_2(\bar{x})$  finito, então:  $\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x})$ .

**Prova:** Exercício 10.10 de Rockafellar e Wets (1998). ■

**Proposição 2.2.3** Sejam  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, 2$ , funções localmente lipschitzianas em  $\bar{x}$ . Se  $f_i \geq 0, i = 1, 2$ , então se tem uma regra do produto na forma de igualdade:

$$\partial(f_1 \cdot f_2)(\bar{x}) = \partial(f_2(\bar{x})f_1 + f_1(\bar{x})f_2)(\bar{x}).$$

**Prova:** Teorema 7.1 de Mordukhovich e Shao (1996). ■

## 2.3 Programação multiobjetivo

**Notação:** Sejam  $x, y \in \mathbb{R}^m$ , então

$$x \leq y \iff x_i \leq y_i \quad \forall i = 1, \dots, m \quad \text{e} \quad x \ll y \iff x_i < y_i \quad \forall i = 1, \dots, m.$$

Considere o problema (P) e uma aplicação  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

**Definição 2.3.1** Dizemos que  $a \in \mathbb{R}^n$  é uma **solução Pareto local** para o problema (P) se existe uma bola aberta  $B_\delta(a) \subset \mathbb{R}^n$ , com  $\delta > 0$ , tal que não existe  $x \in B_\delta(a)$  satisfazendo  $F(x) \leq F(a)$  e  $F_j(x) < F_j(a)$  para pelo menos um índice  $j \in \{1, \dots, m\}$ .

**Definição 2.3.2** Dizemos que  $a \in \mathbb{R}^n$  é uma **solução Pareto local fraco** para o problema (P) se existe uma bola aberta  $B_\delta(a) \subset \mathbb{R}^n$ , com  $\delta > 0$ , tal que não existe  $x \in B_\delta(a)$  satisfazendo  $F(x) \ll F(a)$ .

Vamos denotar por  $\text{argmin}\{F(x) | x \in \mathbb{R}^n\}$  e  $\text{argmin}_w\{F(x) | x \in \mathbb{R}^n\}$  o conjunto das soluções Pareto local e o conjunto das soluções Pareto local fraco para o problema (P), respectivamente. É fácil ver que  $\text{argmin}\{F(x) | x \in \mathbb{R}^n\} \subset \text{argmin}_w\{F(x) | x \in \mathbb{R}^n\}$ .

Se um problema de otimização multiobjetivo é convexo, i.e., se a função objetivo  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é uma função convexa, então toda solução Pareto local (fraco) é também uma solução Pareto global (fraco) (Teorema 2.2.3 de Miettinen (1999)).

**Definição 2.3.3** Uma função de valor real  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é dita ser uma **representação escalar estrita** de uma aplicação  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  quando dados  $x, y \in \mathbb{R}^n$

$$F(x) \leq F(y) \implies f(x) \leq f(y) \quad \text{e} \quad F(x) \ll F(y) \implies f(x) < f(y).$$

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

$$F(x) \ll F(y) \implies f(x) < f(y).$$

É óbvio que toda representação escalar estrita é também uma representação escalar fraca. O próximo resultado estabelece uma importante relação entre os conjuntos  $\text{argmin}\{f(x) | x \in \mathbb{R}^n\}$  e  $\text{argmin}_w\{F(x) | x \in \mathbb{R}^n\}$ . Sua demonstração segue imediatamente da definição 2.3.3.

**Proposição 2.3.1** Seja  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  uma representação escalar fraca de uma aplicação  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $\text{argmin}\{f(x) | x \in \mathbb{R}^n\}$  o conjunto dos minimizadores local de  $f$ . Temos a inclusão:  $\text{argmin}\{f(x) | x \in \mathbb{R}^n\} \subset \text{argmin}_w\{F(x) | x \in \mathbb{R}^n\}$ .

**Proposição 2.3.2** Se  $S \subseteq \mathbb{R}^n$  é um conjunto convexo e  $G : S \rightarrow \mathbb{R}^m$  é uma aplicação convexa, então

$$\text{argmin}_w\{G(x) | x \in S\} = \bigcup_{\bar{z} \in \mathbb{R}_+^m \setminus \{0\}} \text{argmin}\{\langle G(x), \bar{z} \rangle | x \in S\}.$$

**Prova:** Teorema 2.10 de Luc (1989). ■

**Observação 2.3.1** Conforme destacado em Huang e Yang (2004), dada uma aplicação  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  considere o problema de otimização multiobjetivo irrestrito

$$\text{MINIMIZE}\{\exp(F(x)) \mid x \in \mathbb{R}^n\}, \tag{3}$$

onde  $\exp(F(x)) = (\exp(F_1(x)), \dots, \exp(F_m(x)))$ . Então, os conjuntos das soluções Pareto fraco dos problemas (P) e (3) são iguais. Portanto, sem perda de generalidades, podemos supor que para todo  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\langle F(x), \bar{z} \rangle > 0$ ,  $\forall x \in \mathbb{R}^n$ . Observamos ainda, que se o problema (P) for um problema convexo então o problema (3) também será convexo. Além disto, se a função multiobjetivo do problema (P) tiver uma das funções objetivo coerciva então o mesmo acontece com o problema (3).

### 3 Um algoritmo proximal exato

Nesta seção, vamos propor um algoritmo proximal vetorial com a regularização sendo uma quase-distância. Com as hipóteses que a aplicação  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é convexa tal que existe  $r \in \{1, \dots, m\}$  satisfazendo  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$  e que a quase-distância  $q$  satisfaz a propriedade (2), mostraremos que qualquer sequência gerada pelo algoritmo é limitada e que seus pontos de acumulação são soluções Pareto fraco para o problema de otimização multiobjetivo irrestrito (P).

O método gera uma sequência  $\{x^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$  da seguinte forma:

#### Algoritmo QDV-I

1. Tome  $x^0 \in \mathbb{R}^n$ .
2. Dado  $x^k$ , se  $x^k \in \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$ , então  $x^{k+p} = x^k, \forall p \geq 1$ .
3. Dado  $x^k$ , se  $x^k \notin \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$ , então tome como um próximo iterando qualquer  $x^{k+1}$  tal que

$$x^{k+1} \in \operatorname{argmin}_w \left\{ F(x) + \frac{\alpha_k}{2} q^2(x, x^k)e, x \in \Omega^k \right\}, \quad (4)$$

onde  $0 < m < \alpha_k < M; \forall k = 1, 2, \dots, e = (1, 1, \dots, 1) \in \mathbb{R}^m$  e  $\Omega^k = \{x \in \mathbb{R}^n | F(x) \leq F(x^k)\}$ .

#### 3.1 Existência das iterações

Sejam  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma função e  $\bar{x} \in \mathbb{R}^n$ . Defina  $\bar{\Omega} = \{x \in \mathbb{R}^n | F(x) \leq F(\bar{x})\}$ . Temos o seguinte resultado [Conf. Rocha et al. (2012)].

**Lema 3.1.1** *Seja  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação convexa tal que existe  $r \in \{1, \dots, m\}$  satisfazendo  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$ . Então  $\bar{\Omega}$  é um conjunto convexo e compacto.*

**Proposição 3.1.1 (Boa Definição)** *Sejam  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação convexa tal que existe  $r \in \{1, \dots, m\}$  satisfazendo  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$  e  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  uma aplicação quase-distância satisfazendo (2). Então, para cada  $k \in \mathbb{N}$ , se  $x^k \notin \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$ , existe uma solução  $x^{k+1}$  para o problema (4).*

**Prova:** Seja  $x^0 \in \mathbb{R}^n$  escolhido na etapa de inicialização. Supondo que o algoritmo atingiu a iteração  $k$ , vamos mostrar que um apropriado  $x^{k+1}$  existe. Pelo critério de parada, se  $x^k \in \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$  então  $x^{k+p} = x^k, \forall p \geq 1$ . Caso contrário, tome qualquer  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$  e defina  $\varphi^k : \mathbb{R}^n \rightarrow \mathbb{R}$  tal que

$$\varphi^k(x) = \left\langle F(x) + \frac{\alpha_k}{2} q^2(x, x^k)e, \bar{z} \right\rangle = \langle F(x), \bar{z} \rangle + \frac{\alpha_k}{2} \|\bar{z}\|_1 q^2(x, x^k). \quad (5)$$

A convexidade de  $F$  em  $\mathbb{R}^n$  implica na convexidade de  $\langle F(\cdot), \bar{z} \rangle$  em  $\mathbb{R}^n$  e portanto que  $\langle F(\cdot), \bar{z} \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma aplicação contínua. Pela proposição 2.1.1,  $q^2(\cdot, x^k) : \mathbb{R}^n \rightarrow \mathbb{R}_+$  é uma aplicação contínua. Portanto  $\varphi^k : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma aplicação contínua em  $\mathbb{R}^n$ . Logo, desde que  $\Omega^k, \forall k$  é um conjunto compacto (conf. Lema 3.1.1), o conjunto  $\operatorname{argmin}\{\varphi^k(x), x \in \Omega^k\}$  é não vazio. Seja

$$x^{k+1} \in \operatorname{argmin}\{\varphi^k(x); x \in \Omega^k\}. \quad (6)$$

Como  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\varphi^k : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma representação escalar estrita de  $\tilde{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  dada por  $\tilde{F}(x) = F(x) + \frac{\alpha_k}{2} q^2(x, x^k)e$ . Portanto, como toda representação escalar estrita é também uma representação escalar fraca, por (6) e pela proposição 2.3.1,  $x^{k+1} \in \operatorname{argmin}_w \{F(x) + \frac{\alpha_k}{2} q^2(x, x^k)e, x \in \Omega^k\}$ . ■

**Observação 3.1.1** *Seja  $\{x^k\}$  uma sequência gerada pelo algoritmo QDV-I. Pelo lema 3.1.1,  $\Omega^k, \forall k \in N$  é um conjunto compacto. Portanto, como  $\Omega^{k+1} \subseteq \Omega^k, \forall k \in N$ , temos:  $\Omega = \bigcap_{k=0}^{\infty} \Omega^k \neq \emptyset$ .*

Bonnell et al. (2005) utilizaram a proposição 2.3.2 para escalarizar os subproblemas de seu algoritmo e então provaram a convergência de qualquer sequência gerada. Como a quase-distância não necessariamente é uma função convexa em nenhum de seus argumentos não será possível aplicar esta proposição nos subproblemas de nosso algoritmo QDV-I. Então, com o intuito de solucionar esta situação propomos uma extensão da proposição 2.3.2. Sua demonstração é baseada na demonstração do teorema 2.10 em Luc (1989). Segue o resultado:

**Proposição 3.1.2** *Considere  $S \subseteq \mathbb{R}^n$  um conjunto convexo e  $G : S \rightarrow \mathbb{R}^m$  uma aplicação. Se para todo subconjunto convexo  $T$  de  $S$ ,  $G(T) + \mathbb{R}_+^m$  é um conjunto convexo em  $\mathbb{R}^m$ , então*

$$\operatorname{argmin}_w \{G(x) \mid x \in S\} = \bigcup_{\bar{z} \in \mathbb{R}_+^m \setminus \{0\}} \operatorname{argmin} \{ \langle G(x), \bar{z} \rangle \mid x \in S \}.$$

**Prova:** Suponha que  $x^* \in \bigcup_{\bar{z} \in \mathbb{R}_+^m \setminus \{0\}} \operatorname{argmin} \{ \langle G(x), \bar{z} \rangle \mid x \in S \}$ . Então  $x^* \in \operatorname{argmin} \{ \langle G(x), \bar{z} \rangle \mid x \in S \}$  para algum  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$ . Desde que  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\langle G(\cdot), \bar{z} \rangle$  é uma representação escalar estrita de  $G$  e então pela proposição 2.3.1,  $x^* \in \operatorname{argmin}_w \{G(x) \mid x \in S\}$ . Suponha agora que  $x^* \in \operatorname{argmin}_w \{G(x) \mid x \in S\}$ . Então existe uma bola aberta  $B_\delta(x^*) \subset \mathbb{R}^n$ , com  $\delta > 0$ , tal que não existe  $x \in B_\delta(x^*) \cap S$  satisfazendo  $G(x) \ll G(x^*)$ . Logo

$$(G(B_\delta(x^*) \cap S) + \mathbb{R}_+^m) \cap (G(x^*) - \mathbb{R}_{++}^m) = \emptyset. \quad (7)$$

De fato, suponha por contradição, que existam  $x \in B_\delta(x^*) \cap S$  e  $c \in \mathbb{R}_+^m$  tal que  $G(x) + c \in (G(x^*) - \mathbb{R}_{++}^m)$ . Então,  $G(x) \in [G(x^*) - (\mathbb{R}_{++}^m + c)] \subset [G(x^*) - \mathbb{R}_{++}^m]$ , o que é uma contradição. Como  $B_\delta(x^*) \cap S$  é um conjunto convexo e contido em  $S$ , por hipótese,  $G(B_\delta(x^*) \cap S) + \mathbb{R}_+^m$  é também um conjunto convexo. Portanto, como vale (7), os conjuntos  $A = G(B_\delta(x^*) \cap S) + \mathbb{R}_+^m$  e  $B = G(x^*) - \mathbb{R}_{++}^m$  são convexos e disjuntos. Então, pelo teorema da separação (conf. Rockafellar e Wets (1998), teorema 2.39) existem  $z \in \mathbb{R}^m \setminus \{0\}$  e  $d \in \mathbb{R}$  tal que  $\langle x_1, z \rangle \leq d \leq \langle x_2, z \rangle, \forall x_1 \in B$  e  $x_2 \in A$ , isto é,

$$\langle G(x^*) - c_b, z \rangle \leq \langle G(x) + c_a, z \rangle, \forall x \in B_\delta(x^*) \cap S, c_a \in \mathbb{R}_+^m \text{ e } c_b \in \mathbb{R}_{++}^m. \quad (8)$$

Em particular, fazendo  $c_a = (0, \dots, 0)$ , obtemos  $\langle G(x^*), z \rangle \leq \langle G(x), z \rangle + \langle c_b, z \rangle, \forall x \in B_\delta(x^*) \cap S$  e  $c_b \in \mathbb{R}_{++}^m$  e então, fazendo  $c_b \rightarrow 0$ , obtemos  $\langle G(x^*), z \rangle \leq \langle G(x), z \rangle, \forall x \in B_\delta(x^*) \cap S$ , i.e.,  $x^* \in \operatorname{argmin} \{ \langle G(x), z \rangle \mid x \in S \}$ . Para finalizar, mostraremos que  $z \in \mathbb{R}_+^m$ . Tomando  $x = x^*$  em (8), obtemos  $\langle G(x^*) - c_b, z \rangle \leq \langle G(x^*) + c_a, z \rangle, \forall c_a \in \mathbb{R}_+^m$  e  $c_b \in \mathbb{R}_{++}^m$ , i.e.,  $\langle c_a + c_b, z \rangle \geq 0 \forall c_a \in \mathbb{R}_+^m$  e  $c_b \in \mathbb{R}_{++}^m$ . Então, fazendo  $c_a = 0$ , temos  $\langle c_b, z \rangle \geq 0 \forall c_b \in \mathbb{R}_{++}^m$ . Logo,  $z \in \mathbb{R}_+^m$ . ■

**Observação 3.1.2** *Observe que se  $T \subseteq S$  for um conjunto convexo e  $G : S \rightarrow \mathbb{R}^m$  for uma aplicação convexa então  $G(T) + \mathbb{R}_+^m$  será um conjunto convexo. Portanto a proposição 3.1.2 estende a proposição 2.3.2.*

## 3.2 Convergência

**Lema 3.2.1** *Se  $X \subset \mathbb{R}^n$  é um conjunto conexo, então a imagem de toda função real contínua  $f : X \rightarrow \mathbb{R}$  é um intervalo.*

**Prova:** Corolário 7 de Lima (2004). ■

**Lema 3.2.2** *Sejam  $y \in \mathbb{R}^n$ ,  $S \subseteq \mathbb{R}^n$  um subconjunto convexo e  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow R_+$  uma aplicação quase-distância satisfazendo a propriedade (2). Então o conjunto  $q^2(S, y)e$  é um segmento de reta em  $\mathbb{R}^m$  onde  $e = (1, \dots, 1) \in \mathbb{R}^m$ . Em particular,  $q^2(S, y)e$  é um conjunto convexo em  $\mathbb{R}^m$ .*

**Prova:** Pela proposição 2.1.1,  $q^2(\cdot, y) : \mathbb{R}^n \rightarrow R_+$  é uma aplicação contínua. Logo, desde que  $S$  um conjunto convexo, pelo lema 3.2.1,  $q^2(S, y)$  é um intervalo em  $R$ , que será denotado por  $I$ . Então, temos

$$\begin{aligned} q^2(S, y)e &= \{q^2(u, y)e \mid u \in S\} = \{(q^2(u, y), \dots, q^2(u, y)) \mid u \in S\} \\ &= \{(x, x, \dots, x) \mid x \in I\} = \{x(1, 1, \dots, 1) \mid x \in I\}. \end{aligned}$$

Portanto, o conjunto  $q^2(S, y)e$  é um segmento de reta em  $\mathbb{R}^m$ . ■

Seja  $\{x^k\}_{k \in N}$  uma sequência gerada pelo algoritmo **QDV-I**. Para cada  $k \in N$  fixado, considere  $S^k \subseteq \Omega^k$  um subconjunto convexo. Pelo lema 3.2.2,  $q^2(S^k, x^k)e$  é um conjunto convexo. Logo  $\frac{\alpha_k}{2} q^2(S^k, x^k)e$  também será um conjunto convexo. Como  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é convexa,  $F(S^k) + \mathbb{R}_+^m$  é um conjunto convexo. Portanto, para todo  $S^k \subseteq \Omega^k$  convexo,  $F(S^k) + \frac{\alpha_k}{2} q^2(S^k, x^k)e + \mathbb{R}_+^m$  é um conjunto convexo, e então, pela proposição 3.1.2:  $\forall k \in N$ ,

$$\operatorname{argmin}_w \{F(x) + \frac{\alpha_k}{2} q^2(x, x^k)e \mid x \in \Omega^k\} = \bigcup_{\bar{z} \in \mathbb{R}_+^m \setminus \{0\}} \operatorname{argmin}\{\langle F(x), \bar{z} \rangle + \frac{\alpha_k}{2} q^2(x, x^k) \|\bar{z}\|_1 \mid x \in \Omega^k\}. \quad (9)$$

Suponha agora que o critério de parada nunca se aplica e seja  $\{x^k\}_{k \in N} \subset \mathbb{R}^n$  uma sequência gerada pelo algoritmo **QDV-I**, isto é, tal que

$$x^{k+1} \in \operatorname{argmin}_w \left\{ F(x) + \frac{\alpha_k}{2} q^2(x, x^k)e, \quad x \in \Omega^k \right\}.$$

Portanto, por (9), existe  $\{z^k\}_{k \in N} \subset \mathbb{R}_+^m \setminus \{0\}$  tal que

$$x^{k+1} \in \operatorname{argmin}\{\langle F(x), z^k \rangle + \frac{\alpha_k}{2} q^2(x, x^k) \|z^k\|_1, \quad x \in \Omega^k\}. \quad (10)$$

Além disto, como as soluções de (10) não se alteram quando multiplicamos  $z^k$  por um número real positivo, sem perda de generalidades, podemos supor que  $\|z^k\|_1 = 1$ .

**Proposição 3.2.1** *Sejam  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação convexa tal que existe  $r \in \{1, \dots, m\}$  satisfazendo  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$  e  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow R_+$  uma aplicação quase-distância satisfazendo (2). Se  $\{x^k\}_{k \in N}$  é uma sequência gerada pelo algoritmo **QDV-I** então  $\{x^k\}_{k \in N}$  satisfaz (10) com  $\|z^k\|_1 = 1$  e*

(a) **(Caracterização)** *Seja  $k \in N$  fixado. Se  $\Psi_k(x) = \langle F(x), z^k \rangle$ ;  $x \in \mathbb{R}^n$ , existem  $u^{k+1} \in \partial \Psi_k(x^{k+1})$ ,  $\zeta^{k+1} \in \partial(q(\cdot, x^k))(x^{k+1})$  e  $v^{k+1} \in N_{\Omega^k}(x^{k+1})$  tal que*

$$u^{k+1} = -\alpha_k q(x^{k+1}, x^k) \zeta^{k+1} - v^{k+1}. \quad (11)$$

(b) (i)  $\{x^k\}_{k \in N}$  é limitada; (ii)  $\forall z \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\{\langle F(x^k), z \rangle\}_{k \in N}$  é não crescente e convergente; (iii)  $\lim_{k \rightarrow \infty} \|F(x^{k+1}) - F(x^k)\| = 0$ ; (iv)  $\lim_{k \rightarrow \infty} q(x^{k+1}, x^k) = 0$ ; (v)  $\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0$ .

**Prova:** (a) Como  $x^{k+1}$  satisfaz (10) com  $\|z^k\|_1 = 1$ , pela observação 2.2.2 temos

$$0 \in \partial \left( \langle F(\cdot), z^k \rangle + \frac{\alpha_k}{2} q^2(\cdot, x^k) + \delta_{\Omega^k} \right) (x^{k+1}). \quad (12)$$

A convexidade de  $F$  em  $\mathbb{R}^n$ , implica que  $\langle F(\cdot), z^k \rangle$  é contínua em  $x^{k+1}$ ; pela proposição 2.1.1,  $\frac{\alpha_k}{2} q^2(\cdot, x^k)$  é localmente lipschitziana em  $x^{k+1}$ ; a convexidade de  $\Omega^k$  implica na convexidade de

$\delta_{\Omega^k}$  e portanto que  $\delta_{\Omega^k}$  é localmente lipschitziana, então, usando a proposição 2.2.2 em (12), obtemos

$$0 \in \partial(\langle F(\cdot), z^k \rangle)(x^{k+1}) + \partial\left(\frac{\alpha_k}{2} q^2(\cdot, x^k)\right)(x^{k+1}) + \partial(\delta_{\Omega^k})(x^{k+1}). \quad (13)$$

Como  $\Omega^k$  é convexo e  $x^{k+1} \in \Omega^k$  temos que  $\partial(\delta_{\Omega^k})(x^{k+1}) = N_{\Omega^k}(x^{k+1})$ , onde  $N_{\Omega^k}(x^{k+1})$  denota o cone normal no ponto  $x^{k+1}$  em relação ao conjunto  $\Omega^k$ . Pela proposição 2.1.1,  $q(\cdot, x^k)$  é lipschitziana em  $\mathbb{R}^n$ . Portanto, tomando  $f_1 = f_2 = q$  na proposição 2.2.3, temos  $\partial\left(\frac{\alpha_k}{2} q^2(\cdot, x^k)\right)(x^{k+1}) = \alpha_k q(x^{k+1}, x^k) \partial(q(\cdot, x^k))(x^{k+1})$  e então de (13),

$$0 \in \partial(\langle F(\cdot), z^k \rangle)(x^{k+1}) + \alpha_k q(x^{k+1}, x^k) \partial(q(\cdot, x^k))(x^{k+1}) + N_{\Omega^k}(x^{k+1}).$$

Logo, existem  $u^{k+1} \in \partial\Psi_k(x^{k+1})$ ,  $\zeta^{k+1} \in \partial(q(\cdot, x^k))(x^{k+1})$  e  $v^{k+1} \in N_{\Omega^k}(x^{k+1})$  tal que  $u^{k+1} = -\alpha_k q(x^{k+1}, x^k) \zeta^{k+1} - v^{k+1}$ .

(b)(i) Desde que  $\Omega^k \supseteq \Omega^{k+1}$ ,  $k = 0, 1, \dots$ , temos  $x^k \in \Omega^{k-1} \subseteq \Omega^0$ ,  $\forall k \geq 1$ . Como  $\Omega_0$  é limitado (conf. Lema 3.1.1), temos  $\{x^k\}_{k \in \mathbb{N}}$  limitada.

(b)(ii) Como  $x^{k+1} \in \Omega^k$ ,  $F(x^{k+1}) \leq F(x^k)$ . Logo, desde que  $z \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\langle F(x^{k+1}), z \rangle \leq \langle F(x^k), z \rangle \forall k \in \mathbb{N}$ , i.e.,  $\{\langle F(x^k), z \rangle\}_{k \in \mathbb{N}}$  é não crescente. Pela observação 2.3.1,  $\{\langle F(x^k), z \rangle\}_{k \in \mathbb{N}}$  é limitada inferiormente e portanto convergente.

(b)(iii) Seja  $\bar{z} \in \mathbb{R}_+^m$  fixado. Por (b) (ii),  $\{\langle F(x^k), \bar{z} \rangle\}_{k \in \mathbb{N}}$  é convergente. Logo,

$$\lim_{k \rightarrow \infty} \langle F(x^k) - F(x^{k+1}), \bar{z} \rangle = \lim_{k \rightarrow \infty} \sum_{i=1}^m (F_i(x^k) - F_i(x^{k+1})) \bar{z}_i = 0. \quad (14)$$

Como  $x^{k+1} \in \Omega^k$  temos  $F(x^{k+1}) \leq F(x^k)$ , isto é,  $F_i(x^{k+1}) \leq F_i(x^k)$ ,  $\forall i = 1, \dots, m$ . Logo  $(F_i(x^k) - F_i(x^{k+1})) \bar{z}_i \geq 0$ ,  $\forall i = 1, \dots, m$ . Então, de (14),  $\lim_{k \rightarrow \infty} (F_i(x^k) - F_i(x^{k+1})) \bar{z}_i = 0$ ,  $\forall i = 1, \dots, m$ . Desde que  $\bar{z}_i > 0$ ,  $i = 1, \dots, m$  temos  $\lim_{k \rightarrow \infty} (F_i(x^k) - F_i(x^{k+1})) = 0$ ,  $\forall i = 1, \dots, m$  e então,  $\lim_{k \rightarrow \infty} (F(x^k) - F(x^{k+1})) = 0 \in \mathbb{R}^m$ . Portanto,  $\lim_{k \rightarrow \infty} \|F(x^{k+1}) - F(x^k)\| = 0$ .

(b)(iv) Como  $x^k \in \Omega^k$ ,  $\forall k$ , por (10),  $\varphi^k(x^{k+1}) \leq \varphi^k(x^k)$  onde  $\varphi^k(x) = \langle F(x), z^k \rangle + \frac{\alpha_k}{2} q^2(x, x^k)$ ;  $z^k \in \mathbb{R}_+^m$  com  $\|z^k\| = 1$ . Então, desde que  $q(x^k, x^k) = 0$  temos  $\langle F(x^{k+1}), z^k \rangle + \frac{\alpha_k}{2} q^2(x^{k+1}, x^k) \leq \langle F(x^k), z^k \rangle$ . Logo, como  $0 < m < \alpha_k$  e  $\|z^k\| = 1$  temos

$$0 \leq q^2(x^{k+1}, x^k) \leq \frac{2}{m} \langle F(x^k) - F(x^{k+1}), z^k \rangle \leq \frac{2}{m} \|F(x^k) - F(x^{k+1})\|.$$

Então, por (b) (iii),  $\lim_{k \rightarrow \infty} q^2(x^{k+1}, x^k) = 0$ . Como  $q(x, y) \geq 0$ ,  $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$  temos

$$\lim_{k \rightarrow \infty} q(x^{k+1}, x^k) = 0.$$

(b)(v) Por (2),  $0 \leq \alpha \|x^{k+1} - x^k\| \leq q(x^{k+1}, x^k)$ ,  $\forall k \in \mathbb{N}$ . Portanto, por (b) (iv)  $\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0$ . ■

**Proposição 3.2.2** *Sejam  $u, v \in \mathbb{R}^n$  fixados e  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  uma aplicação quase-distância satisfazendo (2). Então, existe  $M > 0$ , que independe de  $u$  e  $v$ , tal que  $\|\bar{x}\| \leq M$  para todo  $\bar{x} \in \partial(q(\cdot, v))(u)$ .*

**Prova:** Lema 2.2.1 de Moreno (2011). ■

**Teorema 3.2.1 (Convergência)** *Sejam  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação convexa tal que existe  $r \in \{1, \dots, m\}$  satisfazendo  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$  e  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  uma quase-distância satisfazendo (2). Se  $\{x^k\}_{k \in \mathbb{N}}$  é uma sequência gerada pelo algoritmo **QDV-I**, então  $\{x^k\}_{k \in \mathbb{N}}$  é limitada e todo ponto de acumulação desta sequência é uma solução Pareto fraco para o problema de otimização multiobjetivo irrestrito (P).*

**Prova:** Como  $\{x^k\}_{k \in \mathbb{N}}$  é uma sequência gerada pelo algoritmo **QDV-I**, existe  $\{z^k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+^m \setminus \{0\}$  tal que  $\{x^k\}_{k \in \mathbb{N}}$  satisfaz a equação (10) com  $\|z^k\|_1 = 1$ . Logo, existe  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$  e  $z^{k_l}$  subsequência de  $\{z^k\}_{k \in \mathbb{N}}$  tal que  $\lim_{l \rightarrow \infty} z^{k_l} = \bar{z}$ . Pela proposição 3.2.1(b)(i),  $\{x^k\}_{k \in \mathbb{N}}$  é uma sequência limitada. Logo existem  $x^* \in \mathbb{R}^n$  e  $\{x^{k_j}\}_{j \in \mathbb{N}}$  subsequência de  $\{x^k\}_{k \in \mathbb{N}}$  tal que  $\lim_{j \rightarrow \infty} x^{k_j} = x^*$ . Fixe  $z \in \mathbb{R}_+^m \setminus \{0\}$ . A convexidade da aplicação  $\Psi = \langle F(\cdot), z \rangle : \mathbb{R}^n \rightarrow R$  em  $\mathbb{R}^n$  implica em sua continuidade. Pela proposição 3.2.1(b)(ii),  $\forall z \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\{\langle F(x^k), z \rangle\}_{k \in \mathbb{N}}$  é não crescente e convergente. Logo,  $\forall z \in \mathbb{R}_+^m \setminus \{0\}$ ,  $\lim_{j \rightarrow \infty} \langle F(x^{k_j}), z \rangle = \langle F(x^*), z \rangle = \inf_{k \in \mathbb{N}} \{\langle F(x^k), z \rangle\}$ . Portanto,

$$\langle F(x^*), z \rangle \leq \langle F(x^k), z \rangle, \forall z \in \mathbb{R}_+^m \setminus \{0\} \text{ e } k \in \mathbb{N}. \quad (15)$$

Pela proposição 3.2.1(a), existem  $\zeta^{k+1} \in \partial(q(\cdot, x^k))(x^{k+1})$  e  $v^{k+1} \in N_{\Omega^k}(x^{k+1})$  tal que  $-\alpha_k q(x^{k+1}, x^k) \zeta^{k+1} - v^{k+1} \in \partial \Psi_k(x^{k+1})$ , onde  $\Psi_k(x) = \langle F(x), z^k \rangle$ ;  $x \in \mathbb{R}^n$ . Portanto, pela desigualdade do subgradiente para função convexa  $\Psi_k$ , temos:  $\forall x \in \mathbb{R}^n$ ,

$$\Psi_{k_l}(x) \geq \Psi_{k_l}(x^{k_l+1}) - \alpha_{k_l} q(x^{k_l+1}, x^{k_l}) < \zeta^{k_l+1}, x - x^{k_l+1} > - < v^{k_l+1}, x - x^{k_l+1} >.$$

Como  $v^{k_l+1} \in N_{\Omega^{k_l}}(x^{k_l+1})$  temos  $- < v^{k_l+1}, x - x^{k_l+1} > \geq 0 \forall x \in \Omega^{k_l}$ . Pela observação 3.1.1,  $\Omega = \bigcap_{k=0}^{\infty} \Omega^k \neq \emptyset$ . Portanto, em particular,

$$\langle F(x), z^{k_l} \rangle \geq \langle F(x^{k_l+1}), z^{k_l} \rangle - \alpha_{k_l} q(x^{k_l+1}, x^{k_l}) < \zeta^{k_l+1}, x - x^{k_l+1} >, \forall x \in \Omega. \quad (16)$$

Por (15),  $\langle F(x^{k_l+1}), z^{k_l} \rangle \geq \langle F(x^*), z^{k_l} \rangle$ . Logo de (16),

$$\langle F(x), z^{k_l} \rangle \geq \langle F(x^*), z^{k_l} \rangle - \alpha_{k_l} q(x^{k_l+1}, x^{k_l}) < \zeta^{k_l+1}, x - x^{k_l+1} >, \forall x \in \Omega. \quad (17)$$

Pela proposição 3.2.2,  $\|\zeta^{k_l+1}\| \leq M_1$ . Logo, como as sequências  $\{x^k\}_{k \in \mathbb{N}}$  e  $\{\alpha_k\}_{k \in \mathbb{N}}$  são limitadas, utilizando a desigualdade de Cauchy-Swartz, concluímos que  $|\alpha_{k_l} q(x^{k_l+1}, x^{k_l}) < \zeta^{k_l+1}, x - x^{k_l+1} >| \leq M_2$ . Então, desde que, pela proposição 3.2.1 (b)(iv),  $\lim_{k \rightarrow \infty} q(x^{k+1}, x^k) = 0$ , concluímos que

$$|\alpha_{k_l} q(x^{k_l+1}, x^{k_l}) < \zeta^{k_l+1}, x - x^{k_l+1} >| \rightarrow 0 \text{ quando } l \rightarrow \infty.$$

Portanto, recordando que  $\lim_{l \rightarrow \infty} z^{k_l} = \bar{z}$ , por (17)

$$\langle F(x), \bar{z} \rangle \geq \langle F(x^*), \bar{z} \rangle, \forall x \in \Omega. \quad (18)$$

Demonstraremos que  $x^* \in \operatorname{argmin}_w \{F(x)/x \in \mathbb{R}^n\}$ . Suponha, por contradição, que exista  $\bar{x} \in \mathbb{R}^n$  tal que

$$F(\bar{x}) \ll F(x^*). \quad (19)$$

Como  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$  temos:

$$\langle F(\bar{x}), \bar{z} \rangle < \langle F(x^*), \bar{z} \rangle. \quad (20)$$

Desde que,  $\Omega^{k+1} \subseteq \Omega^k$ ,  $\forall k \geq 0$  e  $x^{k_j} \in \Omega^{k_j-1}$ ,  $\forall j$  com  $x^{k_j} \rightarrow x^*$ ;  $j \rightarrow \infty$ , temos  $x^* \in \Omega$ , i.e.,  $F(x^*) \leq F(x^k)$ ,  $\forall k \in \mathbb{N}$ . Portanto, de (19),  $F(\bar{x}) \ll F(x^k)$ ,  $\forall k \in \mathbb{N}$ , i.e.,  $\bar{x} \in \Omega$ , a qual contradiz (18) e (20). ■

## 4 Um algoritmo proximal inexacto

Aqui, vamos considerar uma sequência de parâmetros  $\{\beta_k\}_{k \in \mathbb{N}}$  com as mesmas características da sequência de parâmetros  $\{\alpha_k\}_{k \in \mathbb{N}}$  do algoritmo **QDV-I**, i.e., satisfazendo  $0 < m < \beta_k < M$ ,  $\forall k \in \mathbb{N}$ . Além disto, vamos requerer uma sequência  $\{z^k\}_{k \in \mathbb{N}} \subset \mathbb{R}_+^m \setminus \{0\}$  tal que  $\|z^k\|_1 = 1, \forall k \in \mathbb{N}$  e uma sequência  $\{\varepsilon_k\} \subset R_+$  satisfazendo  $\lim_{k \rightarrow \infty} \varepsilon_k = 0$ .

A sequência  $\{x^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$  é gerada da seguinte forma:

### Algoritmo QDV-II

1. Tome  $x^0 \in \mathbb{R}^n$ .
2. Dado  $x^k$ , se  $x^k \in \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$ , então  $x^{k+p} = x^k, \forall p \geq 1$ .
3. Dado  $x^k$ , se  $x^k \notin \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$ , então defina  $\Psi_k : \mathbb{R}^n \rightarrow R$  dada por  $\Psi_k(x) = \langle F(x), z^k \rangle$ . Tome como  $x^{k+1}$  qualquer vetor  $x \in \Omega^k$  tal que existe  $\varepsilon_k \in R_+$  satisfazendo:

$$0 \in \partial_{\varepsilon_k} \Psi_k(x) + \beta_k q(x, x^k) \partial(q(\cdot, x^k))(x) + N_{\Omega^k}(x), \quad (21)$$

$$q^2(x, x^k) \leq \frac{2}{m} \|F(x) - F(x^k)\| \quad \text{e} \quad \lim_{k \rightarrow \infty} \varepsilon_k = 0. \quad (22)$$

onde  $0 < m < \beta_k < M; \forall k \in N, \{z^k\}_{k \in N} \subset \mathbb{R}_+^m \setminus \{0\}; \|z^k\|_1 = 1, \forall k \in N$  e  $\Omega^k = \{x \in \mathbb{R}^n | F(x) \leq F(x^k)\}$ .

**Teorema 4.0.2** *Sejam  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação convexa tal que existe  $r \in \{1, \dots, m\}$  satisfazendo  $\lim_{\|x\| \rightarrow \infty} F_r(x) = \infty$  e  $q : \mathbb{R}^n \times \mathbb{R}^n \rightarrow R_+$  uma quase-distância satisfazendo (2). Então toda sequência  $\{x^k\}_{k \in N}$  gerada pelo algoritmo QDV-II é limitada e cada ponto de acumulação desta sequência é uma solução Pareto fraco para o problema de otimização multiobjetivo irrestrito (P).*

**Prova:** [Existência das iterações]: Seja  $x^0 \in \mathbb{R}^n$  escolhido na etapa de inicialização. Supondo que o algoritmo atingiu a iteração  $k$ , vamos mostrar que um apropriado  $x^{k+1}$  existe. Pelo critério de parada, se  $x^k \in \operatorname{argmin}_w \{F(x), x \in \mathbb{R}^n\}$  então  $x^{k+p} = x^k, \forall p \geq 1$ . Caso contrário, defina  $\varphi^k : \mathbb{R}^n \rightarrow R$  tal que  $\varphi^k(x) = \Psi_k(x) + \frac{\beta_k}{2} q^2(x, x^k)$ . Então, de forma análoga a demonstração da proposição 3.1.1, conclui-se que o conjunto  $\operatorname{argmin}\{\varphi^k(x)/x \in \Omega^k\}$  é não vazio. Seja  $x^{k+1} \in \operatorname{argmin}\{\varphi^k(x)/x \in \Omega^k\}$ . Logo pela observação 2.2.2,  $0 \in \partial(\Psi_k + \frac{\beta_k}{2} q^2(\cdot, x^k) + \delta_{\Omega^k})(x^{k+1})$ . E então, de forma análoga a demonstração da proposição 3.2.1 a), conclui-se que  $0 \in \partial(\Psi_k)(x^{k+1}) + \beta_k q(x^{k+1}, x^k) \partial(q(\cdot, x^k))(x^{k+1}) + N_{\Omega^k}(x^{k+1})$ . Portanto, desde que  $\partial f(x) = \partial_0 f(x) \subset \partial_\varepsilon f(x)$  para toda função convexa  $f : \mathbb{R}^n \rightarrow R$ , todo  $x \in \mathbb{R}^n$  e todo  $\varepsilon \in R_+, x^{k+1}$  satisfaz (21) para  $\varepsilon_k \equiv 0$ . Desde que  $x^{k+1} \in \operatorname{argmin}\{\varphi^k(x)/x \in \Omega^k\}$  e  $x^k \in \Omega^k; \forall k$ , de forma análoga a demonstração da proposição 3.2.1, item (b)(iv), prova-se que  $q^2(x^{k+1}, x^k) \leq \frac{2}{m} \|F(x^{k+1}) - F(x^k)\|$ . Portanto  $x^{k+1}$  satisfaz (21) e (22).

[Propriedades]: Seja  $\{x^k\}_{k \in N}$  uma sequência gerada de forma aleatória pelo algoritmo QDV-II. Temos:  $\{x^k\}_{k \in N}$  é limitada;  $\forall z \in \mathbb{R}_+^m \setminus \{0\}, \{\langle F(x^k), z \rangle\}_{k \in N}$  é não crescente e convergente;  $\lim_{k \rightarrow \infty} \|F(x^{k+1}) - F(x^k)\| = 0$  e  $\lim_{k \rightarrow \infty} q(x^{k+1}, x^k) = 0$ . As provas destas propriedades são análogas às demonstrações da proposição 3.2.1.

[Convergência]: Seja  $\{x^k\}_{k \in N}$  uma sequência gerada pelo algoritmo QDV-II. Então existem  $x^* \in \mathbb{R}^n$  e  $\{x^{k_j}\}_{j \in N}$  subsequência de  $\{x^k\}_{k \in N}$  tal que  $\lim_{j \rightarrow \infty} x^{k_j} = x^*$ . Como  $\|z^k\|_1 = 1, \forall k \in N$ , existem  $\bar{z} \in \mathbb{R}_+^m \setminus \{0\}$  e  $z^{k_l}$  subsequência de  $\{z^k\}_{k \in N}$  tal que  $\lim_{l \rightarrow \infty} z^{k_l} = \bar{z}$ . Por (21), existem  $\zeta^{k_l+1} \in \partial(q(\cdot, x^{k_l}))(x^{k_l+1})$  e  $v^{k_l+1} \in N_{\Omega^{k_l}}(x^{k_l+1})$  tal que  $-\beta_{k_l} q(x^{k_l+1}, x^{k_l}) \zeta^{k_l+1} - v^{k_l+1} \in \partial_{\varepsilon_{k_l}} \Psi_{k_l}(x^{k_l+1})$ . Logo, pela desigualdade do  $\varepsilon_{k_l}$ -subgradiente para a função convexa  $\Psi_{k_l}$ , temos:  $\forall x \in \mathbb{R}^n, \Psi_{k_l}(x) \geq \Psi_{k_l}(x^{k_l+1}) - \beta_{k_l} q(x^{k_l+1}, x^{k_l}) \langle \zeta^{k_l+1}, x - x^{k_l+1} \rangle - \langle v^{k_l+1}, x - x^{k_l+1} \rangle - \varepsilon_{k_l}$ . Portanto, como  $\lim_{k \rightarrow \infty} \varepsilon_k = 0$ , de forma análoga a demonstração do teorema 3.2.1 conclui-se que  $x^*$  é uma solução Pareto fraco para o problema de otimização multiobjetivo irrestrito (P). ■

## 5 Experimentos numéricos

Em nossos experimentos, analisamos o desempenho do método QDV-I em duas instâncias que foram propostas por Li e Zhang (2009). Consideramos as seguintes funções testes: (Abaixo, considere  $J_1 = \{j : j \text{ mpar e } 2 \leq j \leq n\}$  e  $J_2 = \{j : j \text{ par e } 2 \leq j \leq n\}$ )

**(LZF1):**  $F_1 = (f_{11}(x), f_{12}(x))$ ;  $x \in [0, 1] \times [-1, 1]^{n-1}$  onde  $f_{11}(x) = x_1 + (2/|J_1|) \sum_{j \in J_1} (x_j - \sin(6\pi x_1 + j\pi/n))^2$  e  $f_{12}(x) = 1 - \sqrt{x_1} + (2/|J_2|) \sum_{j \in J_2} (x_j - \sin(6\pi x_1 + j\pi/n))^2$ . O conjunto Pareto solução (PS) de (LZF1) é dado por:  $x_j = \sin(6\pi x_1 + j\pi/n)$ ;  $x_1 \in [0, 1]$ ,  $j = 2, \dots, n$ .  
**(LZF2):**  $F_2 = (f_{21}(x), f_{22}(x))$ ;  $x \in [0, 1] \times [-1, 1]^{n-1}$  onde  $f_{21}(x) = x_1 + (2/|J_1|) \sum_{j \in J_1} (x_j - 0.8x_1 \cos(6\pi x_1 + j\pi/n))^2$  e  $f_{22}(x) = 1 - \sqrt{x_1} + (2/|J_2|) \sum_{j \in J_2} (x_j - 0.8x_1 \sin(6\pi x_1 + j\pi/n))^2$ . O conjunto Pareto solução (PS) de (LZF2) é dado por:  $x_j = 0.8x_1 \cos(6\pi x_1 + j\pi/n)$ ;  $x_1 \in [0, 1]$  se  $j \in J_1$  e  $x_j = 0.8x_1 \sin(6\pi x_1 + j\pi/n)$ ;  $x_1 \in [0, 1]$  se  $j \in J_2$ .

No método **QDV-I**, temos o seguinte problema de otimização restrita a cada iteração  $[\min f(x) = \langle F(x), z^k \rangle + (\alpha_k/2)q^2(x, x^k) \|z^k\|_1$  s.a.  $c_j(x) = F_j(x^k) - F_j(x) \geq 0$ ], onde  $\{z^k\} \subset \mathbb{R}_+^m \setminus \{0\}$  é uma sequência gerada de forma aleatória e satisfazendo  $\|z^k\|_1 = 1$ . A rotina de minimização local para problemas restritos utilizada foi a VF04 disponibilizada em linguagem Fortran pela *Harwell Mathematical Software Library* (<http://www.hsl.rl.ac.uk/>).

Finalizamos as iterações quando a distância relativa entre as iterações sucessivas foi inferior a  $10^{-5}$ , isto é, quando  $\|x^k - x^{k+1}\|_\infty / (1 + \|x^k\|_\infty) \leq 10^{-5}$ . A sequência de parâmetros foi gerada seguindo a regra  $\alpha_k = k/(1+k)$ . Fixamos os parâmetros  $c^+ = (1, \dots, 1)$  e  $c^- = (0.5, \dots, 0.5)$  da função quase-distância definida no *Exemplo 2.1.1* da seção 2.1. O erro cometido considerado foi  $\|x^* - x_k^*\|_1$  onde  $x^*$  é a solução exata mais próxima da solução aproximada  $x_k^*$  fornecida pelo método.

Na tabela abaixo,  $M(k), m(k), \bar{m}(k), E_M, E_m$  e  $\bar{E}$  são respectivamente, em 300 experimentos, o maior nr. de iterações, o menor nr. de iterações, a média do nr. de iterações, o maior erro, o menor erro e a média dos erros cometidos. Em todos os experimentos consideramos as instâncias envolvendo 30 variáveis, ou seja,  $n = 30$ , as iterações iniciais são escolhidas de forma aleatória e o número máximo de iterações considerado é 1000. Na figura 1 abaixo, vemos as projeções sobre o espaço  $x_1 - x_2 - x_3$  das soluções obtidas e as curvas Pareto relacionadas às instâncias LZF1 e LZF2. Vemos que todas as soluções estão sobre as respectivas curvas.

Instância	$M(k)$	$m(k)$	$\bar{m}(k)$	$E_M$	$E_m$	$\bar{E}$
<b>LZF1</b>	41	2	9,61	1,76E-04	9,05E-08	1,84E-05
<b>LZF2</b>	7	2	2,54	1,96E-06	1,38E-09	2,71E-07

**Tabela 1:** Número de iterações e o erro cometido em 300 experimentos.

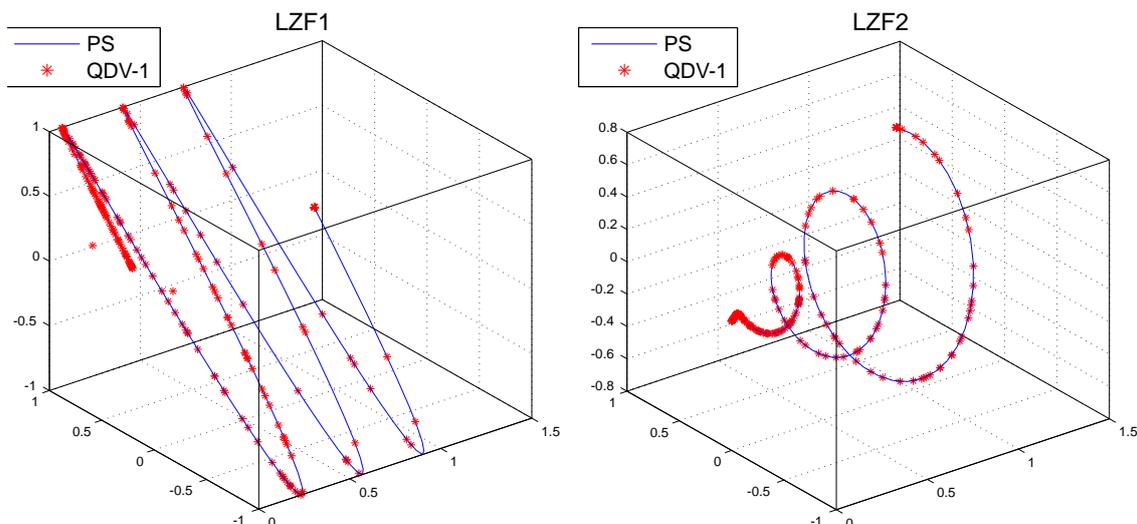


Figura 1: Projeção sobre o espaço  $x_1 - x_2 - x_3$  das 300 soluções obtidas utilizando o método QDV-I nas instâncias LZF1 e LZF2.

## 6 Conclusões finais

Neste trabalho, apresentamos um método proximal com a regularização envolvendo uma função quase-distância para determinar soluções Pareto fraco para o problema de otimização multi-objetivo irrestrito (P) onde o vetor de objetivos  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  é uma aplicação convexa e possuindo uma das funções objetivo coerciva.

Bonnell et al. (2005) desenvolveram em otimização vetorial, uma versão exata e uma versão inexata, de um método proximal vetorial com regularização quadrática. Baseados neste trabalho, desenvolvemos uma versão exata (QDV-I) e uma inexata (QDV-II) de um método proximal vetorial com regularização dada por uma função quase-distância.

Uma vez que a distância Euclidiana é uma função convexa em cada um de seus argumentos, Bonnell et al. utilizaram a proposição 2.3.2 para escalarizar os subproblemas de seu algoritmo e então provaram a convergência de qualquer sequência gerada pelo método. Como a quase-distância não é necessariamente uma função convexa em nenhum de seus argumentos, não foi possível aplicar a proposição 2.3.2 nos subproblemas de nosso algoritmo. Contornamos esta dificuldade desenvolvendo outro resultado de escalarização (Proposição 3.1.2) com menos exigência que a proposição 2.3.2. Desta forma e com argumentos diferentes, provamos a convergência de qualquer sequência gerada pelos nossos algoritmos QDV-I e QDV-II.

## Referências

- [1] **Bonnell, H. , Iusem, A.N. and Svaiter, B.F.** (2005), Proximal methods in vector optimization, *SIAM Journal on Optimization*, v. 15, n. 4, pp. 953-970.
- [2] **Ceng, L. and Yao, J.** (2007) Approximate proximal methods in vector optimization, *European Journal of Operational Research*, 183, pp. 1-19.
- [3] **Gregório, R. and Oliveira, P.R.** (2011), A Logarithmic-quadratic proximal point scalarization method for multiobjective programming, *Journal of Global Optimization*, v. 49, pp. 281-291.
- [4] **Huang, X.X. and Yang, X.Q.** (2004), Duality for multiobjective optimization via nonlinear Lagrangian functions, *J. of Optim. Theory and Applications*, 120 (1), pp. 111-127.
- [5] **Li, H. and Zhang, Q.** (2009), Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II, *IEEE Transactions on Evolutionary Computation*, v. 13, n. 2, pp. 284-302.
- [6] **Lima, E.L.**, *Curso de Análise - Volume 2*, Coleção projeto Euclides, Rio de Janeiro, 2004.
- [7] **Luc, T.D.**, *Theory of vector optimization*, Lecture Notes in Economics and Mathematical Systems, 319, Springer, Berlin, 1989.
- [8] **Miettinen, K.M. ,** *Nonlinear multiobjective optimization*. Kluwer, Boston, 1999.
- [9] **Mordukhovich, B.S. and Shao, Y.** (1996), Nonsmooth Sequential Analysis in Asplund Spaces, *Transactions of the American Mathematical Society* 348(4), pp. 1235-1280.
- [10] **Moreno, F.G.**, *Análise de métodos do tipo proximal com regularização quase-distância*, Tese de doutorado, UFRJ/COPPE/Programa de engenharia de sistemas e computação, 2011
- [11] **Moreno, F.G. , Oliveira, P.R. and Soubeyran, A.** (2011), A proximal Algorithm with Quasi Distance. Application to Habit's Formation, *Optimization: A Journal of Mathematical Programming and Operations Research*, v. 1, pp. 1-21.
- [12] **Rocha, R.A., Oliveira, P.R., and Gregório, R.** (2012), Logarithmic Quasi-distance Proximal Point Scalarization Method for Multi-Objective Programming, *submitted*.
- [13] **Rockafellar, R.T. and Wets, R.J-B.**, *Variational Analysis*, Springer, Berlin, 1998.
- [14] **Villacorta, K.D.V and Oliveira, P. R.** (2011), An Interior Proximal Method in Vector Optimization, *European Journal of Operational Research*, v. 214, pp. 485-492.