A Logarithmic-Quadratic Proximal Point Scalarization Method for Multiobjective Programming

Ronaldo Gregório* and Paulo Roberto Oliveira†
COPPE-Sistemas, Universidade Federal do Rio de Janeiro
Caixa Postal 68511, CEP 21945-970 - Rio de Janeiro, RJ, Brazil.

August 11, 2008

Abstract

We present a proximal point method to solve multiobjective problems based on the scalarization for maps. We build a family of a convex scalar strict representation of a convex map F with respect to the lexicographic order on R^m and we add a variant of the logarithm-quadratic regularization of Auslender, where the unconstrained variables in the domain of F are introduced on the quadratic term and the constrained variables employed in the scalarization we put on the logarithmic term. We show that the central trajectory of the scalarized problem is bounded and converges to a weak Pareto of the multiobjective problem .

Keywords: proximal point algorithm, scalar representations, multiobjective programming.

1 Introduction

We consider the class of problems that minimize a set of objective functions denominated *Multiobjective Programming.* Many important applications in the literature have this structure and this class is related with *decision-making* problems. There is a more general class of problems that contains

*Research of this author was partially supported by CAPES, Brazil. e-mail: rgregor@cos.ufrj.br
†e-mail: poliveir@cos.ufrj.br

Multiobjetive Programming known by Vector Optimization, see for example The Luc[7]. On the other side, the methods developed for this class of problems can be classified in two types, first we have the scalarization methods based on the inner product $\langle F(x), z\rangle$, for any $z \in R^m_+$. Second, it appears the extensions of nonlinear algorithms to the Vector Optimization.

Proximal point method is a particular kind of algorithms that has been extended to Vector Optimization. The first in this line was the Multiobjective Proximal Bundle method presented by Kiwiel in 1990, see Miettinen[6]. Göpfert, [4], also discussed a proximal point method for the scalar representation $\langle F(x), z\rangle$ with a regularization based on Bregman functions on finite dimensional spaces. After Bonnel at all, [2], presented a proximal algorithm with a quadratic regularization in the vector form.

Alternatively, we present a proximal method for abstract scalar strict representation with a variant of the logarithmic-quadratic function of Auslender, [1], as regularization. We introduce the variables of the domain of F on the quadratic term and the scalarization variables are introduced on the logarithm term. Our method is developed for the specific Multiobjective case.

First, we introduce in sections 2, 3 and 4 some concepts and results of the general theory of Vector Optimization that can be found in Luc, [7]. After, we present our method in section 5 where we show the existence of solutions in each iteration and the convergence of the method. We finish with a consideration on problems with inequality constraints on section 6 and some conclusions in the following section.

2 Preliminaries

Consider the unconstrained multiobjective problem

$$\min \ F(x) \qquad (1)$$
$$x \in R^n,$$

where F is a map from R^n to R^m. The nonnegative orthant of R^m (R^m_+) defines the partial order relation $\leq$ ($<$) given by $y, \bar{y} \in R^m$, $y \leq \bar{y}$ iff $y_i \leq \bar{y}_i, i = 1, \cdots, m$ ($y < \bar{y}$ iff $y_i \leq \bar{y}_i, \ \ i = 1, \cdots, m$ and $y_l < \bar{y}_l$, for any $1 \leq l \leq m$). We also have $y \ll \bar{y}$ to design that $y_i < \bar{y}_i$, for all $i = 1, \cdots, m$. It is easy to see that $\leq$ satisfies the axioms of partial order relation on R^m. In a more general closed convex pointed cone K on R^m we can build a partial order relation $\leq_K$ assuming that $y \leq_K \bar{y}$ iff $\bar{y} - y \in K$ ($y <_K \bar{y}$ iff $\bar{y} - y \in \text{int}(K)$).

We say that $a \in R^n$ is a *local Pareto* or *local solution* to the problem 1 iff there exists a disc $B_\delta(a) \subset R^n$, with $\delta > 0$, such that does not exist $x \in B_\delta(a)$ satisfying $F(x) < F(a)$. In the same way, $a \in R^n$ is said a *weak local Pareto* iff there exist a disc $B_\delta(a) \subset R^n$, with $\delta > 0$, such that does not exist $x \in B_\delta(a)$ satisfying $F(x) \ll F(a)$. We will denote by $\text{argmin}\{F(x)/x \in R^n\}$ and $\text{argmin}_w\{F(x)/x \in R^n\}$ the local Pareto and the local weak Pareto set to the problem 1. It is easy to see that $\text{argmin}\{F(x)/x \in R^n\} \subset \text{argmin}_w\{F(x)/x \in R^n\}$.

3 Scalar representation

Scalarization is an importante concept in vector optimization. It plays a fundamental role to develop methods to solve that class of problems and it is also employed as a tool to get the convergence of others algorithms as for example the proximal point method presented by Göpfert, [4], and Bonnel et all, [2].

Definition 1 *Given $x, \bar{x} \in R^n$, a real valued function $f : R^n \longrightarrow R$ is said a strict scalar representation of a map $F : R^n \longrightarrow R^m$ if $F(x) \leq F(\bar{x})$ implies $f(x) \leq f(\bar{x})$ and $F(x) \ll F(\bar{x})$ implies $f(x) < f(\bar{x})$. If we only have that $F(x) \ll F(\bar{x})$ implies $f(x) < f(\bar{x})$, we say that f is a weak scalar representation of F.*

It is obvious that all strict scalar representations are weak scalar representations.

Proposition 1 *Let f be a weak scalar representation of F and* $\text{Argmin}\,\{f(x)/x \in R^n\}$ *the local minimizer set of f. We have the inclusion*

$$\text{argmin}\,\{f(x)/x \in R^n\} \subset \text{argmin}_w\{F(x)/x \in R^n\}.$$

Proof. The Proposition follows immediately from the Definition 1. ■

4 Scalar representation and convexity

Convexity represents an important concept to develop efficient algorithms in Multiobjective Programming. We say that F is a convex map if, and only if

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y), \tag{2}$$

for all $x, y \in R^n$. In this case, the Problem 1 is said convex. If the inequality 2 is strict, we say that F is strictly convex. The inequality 2 still implies that F is a convex map if, and only if each component $F_i : R^n \longrightarrow R, i = 1, \cdots, m$, is a convex function. The importance of the convexity in Multiobjective Programming is due to the fact that every local (weak) Pareto is also global (weak) Pareto for unconstrained or constrained problems. This result is discussed in Theorem 2.2.3 in Miettinen, [6].

If F is convex, we can think on a scalar representation f that be also convex. The Proposition 2.3 in Luc, [7], gives us a good idea how to get scalar representations through the composition of F with an increasing scalar-valued functions on $F(R^n)$. We can add the convexity in the scalar representation of F admitting that f is convex.

The inner product $\langle F(x), z \rangle$, for any $z \in R^m_+ \setminus \{0\}$, is an example of convex scalar strict representation of F, with respect to x. Bonnel et all, [2], employ this representation to get the convergence of the classical proximal point algorithm extended to vector optimization. Göpfert, [4], also presents a scalar proximal point algorithm with Bregman distances in the same lines for vector optimization on spaces of finite dimension.

In this paper, we work with a function $f : R^n \times R^m_{++} \longrightarrow R$ that satisfies the following properties

(P1) $f(x, z) \geq 0$ for all $(x, z) \in R^n \times R^m_{++}$;

(P2) f is convex;

(P3) f is a scalar weak representation of F, with respect to x, i.e.,

$$F(x) \ll F(y) \text{ implies } f(x, z) < f(y, z)$$

for all $x, y \in R^n$ and $z \in R^m_{++}$;

(P4) f is differentiable, with respect to z and

$$\tfrac{\partial}{\partial z} f(x, z) = h(x, z),$$

where h is a continuous map from $R^n \times R^m$ to R^m_+.

The set of functions satisfying those properties is not empty. As an example it is easy to show that $f(x, z) = \displaystyle\sum_{i=1}^{m} \exp\left(z_i + F_i(x)\right)$ satisfies (P1) to (P4).

Our goal is find $x^* \in \text{argmin} f(x, \bar{z})$, for any $\bar{z} \in R^m_{++}$. Therefore, by P3 and Proposition 1 we conclude that x^* is weak Pareto for the unconstrained multiobjective problem.

5 Logarithmic-quadratic proximal point scalarization method

Let F be a convex map, $\beta, \mu > 0$, $y \in R^n$ and $u \in R^m_{++}$. Consider the Logarithmic-quadratic scalarization problem

$$\min \varphi(x,z) = f(x,z) + \beta\langle \frac{z}{u} - \log \frac{z}{u} - e, e\rangle + \frac{\mu}{2}\|x-y\|^2 \qquad (3)$$

$$x \in R^n, z \in R^m_{++},$$

where $f(x,z)$ verifies the properties (P1) to (P4), z/u is the vector whose $i-th$ component is given by z_i/u_i and e is the vector with all components equal to 1.

The objective function $\varphi(x,z)$, defined in 3, is a function build through the addition of a variant of the logarithmic-quadratic function of Auslender[1], where the logarithmic-quadratic regularization presented here can be seen as a sum of two strictly convex separable functions, a logarithmic term with respect to the variable z and a quadratic term with respect to the variable x.

The next lemma shows that the problem 3 has at least one solution.

Lemma 1 *Let $F : R^n \longrightarrow R^m$ be a convex map. The function $\varphi : R^n \times R^m_{++} \longrightarrow R$, given by*

$$\varphi(x,z) = f(x,z) + \beta\langle \frac{z}{u} - \log \frac{z}{u} - e, e\rangle + \frac{\mu}{2}\|x-y\|^2,$$

is $1-$coercive.

Proof. We show first that $\varphi(x,z) \geq 0$, for all $(x,z) \in R^n \times R^m_{++}$. Indeed, taking $t = z_i/u_i$, we have the function $g : R_{++} \longrightarrow R$, defined by

$$g(t) = t - \log t - 1,$$

is strictly convex ($g''(t) = \frac{1}{t^2} > 0$), with its minimum $g(1) = 0$. We conclude that $g(t) \geq 0$. This implies that each parcel of $z/u - \log\,(z/u) - e$ is nonnegative.

Now, Define $\|(x,z)\| = \|x\| + \|z\|$. Suppose that $\|(x,z)\| \longrightarrow +\infty$. This implies $x_l \longrightarrow \infty$, for some $1 \leq l \leq n$ or $z_q \longrightarrow +\infty$, for some $1 \leq q \leq m$. In the first case, we have that

$f(x,z) + \beta\langle \frac{z}{u} - \log \frac{z}{u} - e, e\rangle + \frac{\mu}{2}\|x-y\|^2 \geq \frac{\mu}{2}\|x-y\|^2 \geq \frac{\mu}{2}(x_l - y_l)^2$.
Therefore

$$\lim_{\|(x,z)\|\longrightarrow+\infty}\varphi(x,z)\geq\lim_{x_l\longrightarrow+\infty}\frac{\mu}{2}\left(x_l-y_l\right)^2=+\infty.$$

This proves the lemma in that case.

The second case is proved by a similar argument. ■

Given a initial point $(x^0, z^0) \in R^n \times R^m_{++}$, the logarithmic-quadratic proximal point scalarization method generates sequences $\left\{x^k\right\}_{k\in N} \subset R^n$ and $\left\{z^k\right\}_{k\in N} \subset R^m_{++}$ defined by

$$\left(x^{k+1}, z^{k+1}\right) = \text{argmin}_{x\in R^n, z\in R^m_{++}}\ \varphi^k(x,z),$$

where $\varphi^k(x,z) = f(x,z) + \beta^k\langle \frac{z}{z^k} - \log\frac{z}{z^k} - e, e\rangle + \frac{\mu^k}{2}\|x - x^k\|^2$.

Lemma 2 (Well-posedness) *There is only one solution $\left(x^{k+1}, z^{k+1}\right)$ to the logarithmic-quadratic proximal scalarization problem characterized by*

$$\mu^k(x^k - x^{k+1}) \in \partial_{x^{k+1}} f\left(x^{k+1}, z^{k+1}\right) \tag{4}$$

and

$$\frac{1}{z_i^{k+1}} - \frac{1}{z_i^k} = \frac{h_i\left(x^{k+1}, z^{k+1}\right)}{\beta^k}, \quad i = 1, \cdots m. \tag{5}$$

Proof. The existence and the uniqueness of solution are guaranteed by Lemma 1 and by the strict convexity of φ^k, respectively. Take the optimality conditions for the exponential logarithmic-quadratic proximal scalarization problem, and the Lemma follows. ■

Note that each function h_i is not necessary separable, i. e., we do not need to have $h_i(x,z) = \xi(x) + \eta(z)$ as we will see in the following.

Now, we present the convergence results for our algorithm.

Theorem 1 (Convergence) *Let $F : R^n \longrightarrow R^m$ be a convex map. If $\left\{\mu^k\right\}_{k\in N}$ and $\left\{\beta^k\right\}_{k\in N}$ are bounded sequences of real positive numbers and β^k satisfies $\sum_{k=0}^{\infty}\frac{1}{\beta^k} < +\infty$ then the sequence $\{(x^k, z^k)\}_{k\in N}$ generated by the logarithmic-quadratic proximal point scalarization method is bounded and $x^* = \lim_{n\longrightarrow+\infty} x^k$ is a weak Pareto point to the unconstrained multiobjective problem.*

Proof. The relation 4 and subgradient inequality applied to the convex function $f(x, z^{k+1})$ imply

$$f(x, z^{k+1}) \geq f(x^{k+1}, z^{k+1}) + \mu^k \langle x^k - x^{k+1}, x - x^{k+1} \rangle,$$

that can be rewritten as

$$\mu^k \langle x^k - x^{k+1}, x^{k+1} - x \rangle \geq f(x^{k+1}, z^{k+1}) - f(x, z^{k+1}),$$

for all $x \in R^n$. On the other hand, we can write

$$\|x^k - x\|^2 = \|x^k - x^{k+1} + x^{k+1} - x\|^2 =$$

$$\|x^k - x^{k+1}\|^2 + \|x^{k+1} - x\|^2 + 2\langle x^k - x^{k+1}, x^{k+1} - x \rangle,$$

for all $x \in R^n$. Now, putting together this inequality and the precedent, we get

$$\|x^k - x\|^2 \geq \|x^k - x^{k+1}\|^2 + \|x^{k+1} - x\|^2 + \frac{2}{\mu^k} \left(f(x^{k+1}, z^{k+1}) - f(x, z^{k+1}) \right),$$

for all $x \in R^n$. Since f satisfies P3, we have that $f(x^{k+1}, z^{k+1}) - f(\bar{x}, z^{k+1}) \geq 0$, for all $\bar{x} \in \text{Argmin}_w \ \{F(x)/x \in R^n\}$.This implies

$$\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2.$$

for all $\bar{x} \in \text{Argmin}_w \ \{F(x)/x \in R^n\}$.

We conclude that $\{x^k\}_{k \in N}$ is a Fejér convergent sequence to argmin_w $\{F(x)/x \in R^n\}$ and therefore, bounded. This is a general result for Fejér convergent sequence on complete metric spaces. See, Iusem[5].On the other side, since $h_i(x, z) \geq 0$, the condition 5 implies $\left\{z_i^k\right\}_{k \in N}$ is a nonincreasing monotone sequence, for all $i = 1, \cdots, m$. Therefore, $\{z^k\}_{k \in N}$ is convergent (it is easy to see that each sequence $\{z_i^k\}_{k \in N}$ $i = 1, \cdots, m$ is monotone nonincreasing and bounded). Now, we show that $\lim_{k \longrightarrow +\infty} z^k \neq 0$. Let $\{x^{k_j}\}_{j \in N}$ be a subsequence convergent of $\{x^k\}_{k \in N}$ and x^* its cluster point. We have, associated to the subsequence $\{x^{k_j}\}_{j \in N}$, the subsequence $\{z^{k_j}\}_{j \in N}$. By the equality 5 we have

$$\frac{1}{z_i^{k_{j+1}}} - \frac{1}{z_i^{k_j}} = \frac{h_i\left(x^{k_{j+1}}, z^{k_{j+1}}\right)}{\beta^{k_j}}.$$

Let $\{y^j\}_{j \in N}$, $\{w^j\}_{j \in N}$ and $\{\gamma^j\}_{j \in N}$ be sequences defined by $y^j = x^{k_j}$, $w^j = z^{k_j}$ and $\gamma^j = \beta^{k_j}$, respectively. This implies, after summing, from 0 to $l - 1$, both sides of the above equality, that

$$\frac{1}{w_i^l} - \frac{1}{w_i^0} = \sum_{j=0}^{l-1} \frac{h_i\left(y^{j+1}, w^{j+1}\right)}{\gamma^j}.$$

Since $\{y^j\}_{j\in N}$, $\{w^j\}_{j\in N}$ are convergent sequences and h is continuous, the sequence $\{h_i(y^{j+1}, w^{j+1})\}_{j\in N}$ is also convergent and therefore, bounded. Let M be some upper bound. We have that

$$\tfrac{1}{w_i^l} - \tfrac{1}{w_i^0} = \sum_{j=0}^{l-1} \frac{h_i\left(y^{j+1}, w^{j+1}\right)}{\gamma^j} \leq M \sum_{j=0}^{l-1} \frac{1}{\gamma^j}.$$

Assume, by absurd, that $z_i^k \longrightarrow 0$. This implies that $z_i^{k_j} \longrightarrow 0$ or, with above notation, $w_i^j \longrightarrow 0$. As a consequence, we have

$$+\infty = \lim_{l\longrightarrow+\infty} (\frac{1}{w_i^l} - \frac{1}{w_i^0}) \leq M \lim_{l\longrightarrow+\infty} \sum_{j=0}^{l-1} \frac{1}{\gamma^j} < +\infty.$$

This is a contradiction.

Now, we must prove that all cluster points of $\{x^k\}_{k\in N}$ belong to $\mathrm{Argmin}_w\{F(x)/x \in R^n\}$. Notice that

$$f\left(x, z^{k_{j+1}}\right) \geq f\left(x^{k_{j+1}}, z^{k_{j+1}}\right) + \mu^{k_j}\langle x^{k_j} - x^{k_{j+1}}, x - x^{k_{j+1}}\rangle,$$

for all $x \in R^n$. Taking the limit in the inequality above we have

$$f\left(x, z^*\right) \geq f\left(x^*, z^*\right),$$

for all $x \in R^n$. Therefore, $x^* \in \mathrm{Argmin}\left\{f\left(x, z^*\right)/x \in R^n\right\}$.

To finish this proof, remember that $z_i^* > 0$ for all $i = 1, \cdots, m$. We have, by P3, that $f\left(x, z^*\right)$ is a scalar weak representation of F. Invoke the Proposition 1 and conclude that $x^* \in \mathrm{Argmin}_w\{F(x)/x \in R^n\}$. Again by the result of convergence of Fejér sequences (see Iusem[5]), we have that $x^k \longrightarrow x^*$. ■

6 Multiobjective problem with inequality constraints

We consider the addition of the inequality constrained $G(x) \leq 0$ on the Multiobjective problem 1, where G is a convex map from R^n to R^p. The KKT necessary and sufficient conditions for the Logarithmic-quadratic proximal point scalarization problem implies that there exists a Lagrangian multiplier $y^{k+1} \in R_+^p$ such that

$$\mu^k(x^k - x^{k+1}) \in \partial_{x^{k+1}}(f(x^{k+1}, z^{k+1}) + \langle G(x^{k+1}), y^{k+1}\rangle), \qquad (6)$$

$$y_j^{k+1} G_j(x^{k+1}) = 0, \quad j = 1, \cdots, p, \qquad (7)$$

and the equation 5 is assured.

Theorem 2 *Suppose that* $\left(x^{k+1}, z^{k+1}, y^{k+1}\right)$ *satisfies 6, 7 and 5 for every* $k \in N$*. If* G *is a convex map from* R^n *to* R^p *and the hypothesis of the Theorem 1 are assured then the sequence* $\left\{x^k\right\}_{k\in N}$ *converges to a weak Pareto of the constrained multiobjective problem*

$$\min F(x)$$

$$\text{s.t. } G(x) \leq 0$$

$$x \in R^n$$

Proof. The relation 6 implies

$$f(x, z^{k+1}) + \langle G(x), y^{k+1}\rangle \geq$$
$$f(x^{k+1}, z^{k+1}) + \langle G(x^{k+1}), y^{k+1}\rangle + \mu^k\langle x^k - x^{k+1}, x - x^{k+1}\rangle,$$

for all $x \in R^n$. Substituting the equation 7 in the inequality above, we have

$$f(x, z^{k+1}) + \langle G(x), y^{k+1}\rangle \geq f(x^{k+1}, z^{k+1}) + \mu^k\langle x^k - x^{k+1}, x - x^{k+1}\rangle.$$

Since $y^{k+1} \in R^p_+$, if we consider just the points that are feasible for the constrained problem, we obtain

$$f(x, z^{k+1}) \geq f(x^{k+1}, z^{k+1}) + \mu^k\langle x^k - x^{k+1}, x - x^{k+1}\rangle,$$

for all $x \in R^n$ such that $G(x) \leq 0$. At this point, the arguments are analogous to the employed on the proof of the Theorem 1, i.e., first, we conclude that $\left\{x^k\right\}_{k\in N}$ is Fejér convergent to $\text{Argmin}_w\left\{F(x)/G(x) \leq 0, x \in R^n\right\}$ and, by the equation 5, $0 \ll \lim\limits_{k\longrightarrow+\infty} z^k = z^*$. To finish, if we work with a subsequence, we have that

$$f\left(x, z^{k_{j+1}}\right) \geq f\left(x^{k_{j+1}}, z^{k_{j+1}}\right) + \mu^{k_j}\langle x^{k_j} - x^{k_{j+1}}, x - x^{k_{j+1}}\rangle,$$

for all $x \in R^n$ such that $G(x) \leq 0$ and

$$y^{k_{j+1}} G\left(x^{k_{j+1}}\right) = 0.$$

Taking the limit in the relations above, we conclude that $0 \in \partial_{x^*} f\left(x^*, z^*\right)$ and $y^*_j G_j\left(x^*\right) = 0, \quad j = 1, \cdots, p$ that is a KKT condition to the problem $\min\left\{f\left(x, z^*\right)/G(x) \leq 0, x \in R^n\right\}$ at x^*. Since f is a weak representation of F with respect to x, we conclude that

$$x^* \in \text{Argmin}_w\left\{F(x)/G(x) \leq 0, x \in R^n\right\}.$$

■

7 Conclusions

In this paper we present a new family of scalarization methods for the Multiobjective Programming. Others algorithms, as Göpfert[4] and Bonnel et all[2], work only with the hypothesis of the existence of some $z \in R^m_+ \setminus \{0\}$ such that $\langle F(x), z\rangle$ is a scalar representation of F. Here, we show that the family of scalar weak representation $f_z(x) = f(x, z)$ converges to a escalar weak representation $f_{z^*}(x) = f(x, z^*)$ with $z^* \in \mathrm{int}(R^m_+)$.

We get the convergence result for this method to a weak Pareto of the unconstrained and constrained Multiobjective cases. But, we do not make any reference about what algorithm must be employed to compute $\left(x^{k+1}, z^{k+1}\right)$ in the unconstrained case or $\left(x^{k+1}, z^{k+1}, y^{k+1}\right)$ in the constrained one. Since φ^k is strictly convex for all $k \in N$, if F is differentiable, we can apply the Newton´s method that has presented a good choice for those problems.

We also present an example of function that satisfies the four properties P1, P2, P3 and P4, but we believe that is not the unique example. We are investigating others functions with the same characteristics, and an implementation of the method.

References

[1] Auslender, A., Teboulle, M., Ben-Tiba, S., *A logarithmic-quadratic proximal method for variational inequalities*, Computational Optimization Applications, v.12, n.1-3, 1999.

[2] Bonnel, H., Iusem, A.N., Svaiter, B.F., *Proximal methods in vector optimization*, SIAM Journal on Optimization, v. 15, n.4, pp. 953-970, 2005.

[3] Ferreira, O.P., Oliveira, P.R., "Proximal point algorithm on Riemannian manifolds", *Optimization*, v. 51, n. 2, pp. 257-270, 2002.

[4] Gopfert, A., Riahi H., Tammer, and Zalinescu, C., *Variational methods in partially ordered spaces*, Springer, New York, 2003.

[5] Iusem, A. N., *An iterative algorithm for the variational inequality problem*, Computational and Applied Mathematics, v. 13, 103-114, 1994.

[6] Miettinen, K.M., *Nonlinear Multiobjective Optimization.* Boston, Kluwer, 1999.

[7] The Luc, D., *Theory of vector optimization*, lecture Notes in Economics and Mathematical Systems, 319, Berlin, Springer, 1989.