RAIRO-Oper. Res. 58 (2024) 3119–3142
https://doi.org/10.1051/ro/2024097

RAIRO Operations Research
www.rairo-ro.org

SUCCESSIVE UPPER APPROXIMATION METHODS FOR GENERALIZED FRACTIONAL PROGRAMS

Karima Boufi[1], Abdessamad Fadil[2] and Ahmed Roubi[2,*]

Abstract. The majorization approximation procedure consists in replacing the resolution of a nonlinear optimization problem by solving a sequence of simpler ones, whose objective and constraint functions upper estimate those of the original problem. For generalized fractional programming, *i.e.*, constrained minimization programs whose objective functions are maximums of finite ratios of functions, we propose an adapted scheme that simultaneously upper approximates parametric functions formed by the objective and constraint functions. For directionally convex functions, that is, functions whose directional derivatives are convex with respect to directions, we will establish that every cluster point of the generated sequence satisfies Karush–Kuhn–Tucker type conditions expressed in terms of directional derivatives. The proposed procedure unifies several existing methods and gives rise to new ones. Numerical problems are solved to test the efficiency of our methods, and comparisons with different approaches are given.

Mathematics Subject Classification. 90C32, 90C26, 90C25, 90C55, 90C46.

Received September 20, 2023. Accepted April 26, 2024.

1. Introduction

This paper deals with minimax or generalized fractional programming (GFP) problems, which are defined as minimization problems having an objective function expressed as the maximum of finite ratios of functions.

Problems of such type arise in management applications of goal programming, in mathematical economics and numerical analysis, and in telecommunications, information theory and computer science. More applications of single and multi-ratio programming can be found in [4–6, 32, 41, 47].

There is a rich literature that deals with the subject, and a multitude of methods destined to solve such problems. The first approaches for solving GFP are based on the parametric approach which gave rise to several efficient methods such as [2, 3, 13, 23, 25–27, 33–35, 37, 44, 45, 48, 55]. The parametric approach was first introduced in [28] for single fractional programs, see [46] for a bibliography. Several other results and algorithms for solving GFP were developed using duality, but are also based on the parametric approach, see *e.g.*, [9, 10, 12, 15–20, 24, 29, 30, 36, 38]. Interested readers can consult [48–54] for a detailed bibliography.

Keywords. Fractional programming, optimality conditions, successive majorizations methods, successive quadratic approximation, gradient method.

[1] Hassan First University of Settat, Institut Supérieur des Sciences de la Santé & Laboratoire MISI, Settat, Morocco.

[2] Hassan First University of Settat, Faculté des Sciences et Techniques, Laboratoire MISI, Settat, Morocco.

*Coressponding author: `roubia@hotmail.com`

To position our work in relation to previous works, we return to the parametric approach and recall that algorithms for solving GFP are generally based on globally solving intermediate subproblems. But for almost all of these methods, some convexity assumptions are required, and unfortunately, apart from the situation when the parametric subproblems are convex, these approaches may fail since it is assumed that one can compute their global minimum, at least approximately. Our intention is then to design a method based on the sequential upper approximation procedure idea, which consists in replacing the resolution of the original problem by solving a sequence of simpler or more regular ones, whose objective and constraint functions upper approximate those of that problem. We will use this technique to upper approximate the objective functions of the parametric subproblems appearing in [44] to have more tractable ones. By doing so, we aim two objectives: to cover a large class of problems to be solved, and to unify several methods under a generic algorithm.

In the setting of general minimization problems, the technique of sequential upper approximation dates back several years, see *e.g.*, [40]. Since then, it has undergone several extensions and developments. To give a brief overview of this method, we go through some references related to the subject. In [11, 40], the authors treated problems with convex objective function and nonconvex constraints, whereas in [43], the objective function may be nonconvex, but the constraint set is assumed to be convex. The authors deal in [7] with the case where both the objective and the constraints are approximated.

We recall that the objective function of each parametric problem in the method of centers given in [44] is formed from the objective and constraint functions of the original problem. We propose in what follows a model that over estimates this function, in such away to simultaneously approximate by a single function the objective and the constraint functions. We focus on directionally convex functions, *i.e.*, functions that are directionally differentiable whose directional derivatives are convex with respect to directions. This is a quite large class of functions that contains continuously differentiable functions, convex functions, Lipschitz functions regular in the sense of Clarke, and the sum of convex and continuously differentiable or Lipschitz regular functions, etc.

We firstly establish Karush–Kuhn–Tucker type optimality conditions expressed in terms of directional derivatives of the objective and constraint functions. Sufficient optimality conditions are also given for a class of nonconvex problems. We will then establish that every cluster point of the sequence generated by our procedure satisfies these conditions.

By means of several examples, we establish that several earlier methods inscribe in our generic scheme.

An application of the proposed model to constrained programs with continuously differentiable functions that have Lipschitz gradients gives rise to a sequential quadratic method. This seems to be the first application of a sequential quadratic method to GFP. With additional convexity assumptions, we show that the sequence generated by our algorithm converges to an optimal solution of the GFP.

By passing through a sort of dual of the previous method, we obtain a gradient type method, in which the descent direction is obtained from the gradients of the objective and the constraint functions. The cost to obtain this direction is the resolution of a simple quadratic program.

The paper is organized as follows. We introduce in Section 2 our problem and some preliminary notions. Section 3 is devoted to necessary and sufficient optimality conditions for GFP. In Section 4 we define the notion of upper bounding approximating functions and problems, followed by concrete examples and preliminary properties and results. We introduce and describe our successive majorizations algorithm in Section 5, and analyze its convergence. We detail in Section 6 a primal sequential quadratic method and its dual, we will call the constrained gradient algorithm. We end by some numerical tests in Section 7 and a conclusion in Section 8.

2. Preliminaries

We are interested in this work in solving minimax fractional problems of the form

$$\begin{aligned} \bar{\lambda} = \inf_{x\in C}\ & \lambda(x) := \max_{i\in I} \frac{f_i(x)}{g_i(x)} \\ \text{s.t.}\ & h_j(x) \le 0, \quad j \in J := \{1, \cdots, p\} \end{aligned} \tag{P}$$

where the functions f_i, g_i and h_j, for $i \in I$ and $j \in J$ are real valued functions defined on $\mathbb{R}^n$, directionally differentiable at points and directions that will be specified in the forthcoming development, and $C \subset \mathbb{R}^n$ is a nonempty, closed and convex set of $\mathbb{R}^n$. We suppose in addition that for all $i \in I$, the function g_i is positive on X, the constraint's set of (P). For simplicity, we will pose

$$h(x) = \max_{j \in J} h_j(x)$$

so that $X := \{x \in C \mid h(x) \leq 0\}$.

As mentioned in the beginning, we will present in what follows a general algorithm based on the sequential upper approximation procedure idea to solve problem (P). We will begin our analysis by defining specific upper approximating functions and related approximating sub-problems for minimax fractional problems. For this, instead of approximating directly and only the objective function of the initial problem, we will first pass through parametric functions, which usually appear in the frameworks of fractional programming, to which we append the constraint functions, to be all upper approximated by the same approximating function. To help understanding the different steps of our general algorithm, we will give some preliminary results related to such approximations, before describing it. To complete convergence results, we will focus on optimality conditions, or stationarity, expressed in terms of directional derivatives. The points found by our algorithm actually verify these conditions. For some classes of problems, these conditions even characterize optimal solutions.

To solve problem (P), Roubi considered in [44], for $\mu \in \mathbb{R}$, the parametric problem

$$\inf_{x \in C} \max_{(i,j) \in I \times J} \{f_i(x) - \mu g_i(x), h_j(x)\}. \tag{P_μ}$$

This problem is generally simpler than (P). We recall that under mild assumptions, when $\mu = \bar{\lambda}$, problems (P) and (P_μ) have the same optimal solutions set, see *e.g.*, [44]. The problem (P_μ) is the key ingredient in the extended method of centers developed in [44]. A brief description of this method is as follows: from an arbitrary $x^0 \in X$ a sequence of points x^k is generated by solving a sequence of subproblems (P_{λ_k}) where $\lambda_k = \lambda(x^k)$ and x^{k+1} is an optimal solution of (P_{λ_k}). It is worth noting that a crucial step in this method is to find a global optimal solution to (P_{λ_k}). Apart from the convex case, it is difficult to find such a global optimal solution to (P_{λ_k}). Since the problems (P_{λ_k}) may be nonconvex for the case of our problem (P), we will focus in the following analysis on how to escape this difficulty.

For all $x \in \mathbb{R}^n$, we define the following functions, parameterized by $\mu \in \mathbb{R}$,

$$\mathcal{F}(x, \mu) = \max_{(i,j) \in I \times J} \{f_i(x) - \mu g_i(x), h_j(x)\}. \tag{1}$$

This is precisely the objective function of the problem (P_μ). As mentioned above, finding a global minimum on C of this function, for $\mu \in \mathbb{R}$, may be difficult or even impossible. For this, we approximate the latter, locally by an upper bounding function, to get an approximate optimal solution of (P_μ) by the minimization of the last function on C. Our approach upper approximates simultaneously the objective and the constraint functions.

For the next developments, we will need some preliminary notions and results. We start by recalling that the function f is directionally differentiable at the point $\bar{x} \in \mathbb{R}^n$, in the direction $d \in \mathbb{R}^n$, if the one-sided directional derivative defined by

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

exists. We also recall the following result that we will use later. This result concerns the directional derivative of the following function

$$\varphi(x) = \max_{i \in I} \varphi_i(x),$$

where the functions φ_i are directionally differentiable at $\bar{x}$, in a direction $d \in \mathbb{R}^n$. Then φ is also directionally differentiable at $\bar{x}$ in the direction d, and we have

$$\varphi'(\bar{x}; d) = \max_{i \in I(\bar{x})} \varphi_i'(\bar{x}; d), \tag{2}$$

where $I(\bar{x}) := \{i \in I \mid \varphi_i(\bar{x}) = \varphi(\bar{x})\}$.

To specify the kind of points that will be found by our procedure, we give in the next section various necessary and sufficient optimality conditions for problem (P), expressed in terms of directional derivatives.

3. Optimality conditions for GFP

In the following, we will use the notation

$$\Sigma = \Big\{(\alpha, \beta) \in \mathbb{R}^m \times \mathbb{R}^p \;\Big|\; \textstyle\sum_{i\in I} \alpha_i + \sum_{j\in J} \beta_j = 1,\ \alpha_i,\ \beta_j \geq 0 \quad \forall (i, j) \in I \times J\Big\}.$$

In the following we gives an expression of the directional derivative of $\mathcal{F}(\cdot, \mu)$.

Proposition 1. *Let $\mu \in \mathbb{R}$ and $x, d \in \mathbb{R}^n$. The directional derivative of $\mathcal{F}(\cdot, \mu)$ at x in the direction d is given by*

$$\mathcal{F}'((x, \mu); d) = \max_{(\alpha,\beta)\in\Sigma(x)} \left\{ \sum_{i\in I} \alpha_i (f_i'(x; d) - \mu g_i'(x; d)) + \sum_{j\in J} \beta_j h_j'(x; d) \right\}, \tag{3}$$

where $\Sigma(x) := \{(\alpha, \beta) \in \Sigma \mid \sum_{i\in I} \alpha_i(f_i(x) - \mu g_i(x)) + \sum_{j\in J} \beta_j h_j(x) = \mathcal{F}(x, \mu)\}$.

Proof. It is straightforward to verify that we can write

$$\mathcal{F}(x, \mu) = \max_{(\alpha,\beta)\in\Sigma} \left\{ \sum_{i\in I} \alpha_i (f_i(x) - \mu g_i(x)) + \sum_{j\in J} \beta_j h_j(x) \right\}.$$

It is clear that the maximum in the last expression of $\mathcal{F}(x, \mu)$ is achieved at all $(\alpha, \beta) \in \Sigma$ such that $\sum_{i\in I} \alpha_i(f_i(x) - \mu g_i(x)) + \sum_{j\in J} \beta_j h_j(x) = \mathcal{F}(x, \mu)$, that is at all $(\alpha, \beta) \in \Sigma(x)$. Thus, by using Theorem 3.2 of [42], we get

$$\mathcal{F}'((x, \mu); d) = \max_{(\alpha,\beta)\in\Sigma(x)} \left\{ \sum_{i\in I} \alpha_i (f_i'(x; d) - \mu g_i'(x; d)) + \sum_{j\in J} \beta_j h_j'(x; d) \right\},$$

where $\Sigma(x)$ is as defined in the proposition. □

The next theorem gives necessary optimality conditions of Karush–Kuhn–Tucker type expressed in terms of directional derivatives of the objective and constraint functions, that any optimal solution of the minimax fractional problem (P) must satisfy. This result is based on the convexity of the directional derivatives, at the desired point $\bar{x}$, with respect to directions, that is, for a function f, say, one has $f'(\bar{x}; \alpha u + (1 - \alpha)v)) \leq \alpha f'(\bar{x}; u) + (1 - \alpha) f'(\bar{x}; v)$ for all $u, v \in \mathbb{R}^n$ and $\alpha \in (0, 1)$. This class of functions contains convex functions and continuously differentiable functions. The sum of a convex and a continuously differentiable function, which is an example of nonconvex functions that have convex directional derivatives, also belongs to this class. Locally Lipschitz functions that are regular in the sense of Clarke is another example of such functions.

Theorem 1. *Let $\bar{x} \in X$ and $\bar{\lambda} = \lambda(\bar{x})$. Assume that the functions f_i, g_i and h_j are directionally differentiable at $\bar{x}$ and that $f_i'(\bar{x}; \cdot) - \bar{\lambda} g_i'(\bar{x}; \cdot)$ and $h_j'(\bar{x}; \cdot)$ for $(i, j) \in I \times J$, are convex as functions of the directions. If $\bar{x}$ is an optimal solution of (P), then there exists $(\bar{\alpha}, \bar{\beta}) \in \Sigma$ such that*

$$\sum_{i\in I} \bar{\alpha}_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C,$$

with $\bar{\alpha}_i[f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $(i, j) \in I \times J$.

Proof. Let $\bar{x} \in X$ be an optimal solution of (P), and let $\bar{\lambda} = \lambda(\bar{x})$. We have already seen that $\mathcal{F}(x, \bar{\lambda}) \geq 0$ for all $x \in C$, and that $\mathcal{F}(\bar{x}, \bar{\lambda}) = 0$. Thus, with the convexity of C we obtain

$$\mathcal{F}'\big((z, \bar{\lambda}); x - \bar{x}\big) \,|_{z=\bar{x}} \geq 0 \quad \text{for all } x \in C. \tag{4}$$

By using (3) with $x = \bar{x}$, $\mu = \bar{\lambda}$, and $d = x - \bar{x}$ we get

$$\mathcal{F}'\big((z, \bar{\lambda}); x - \bar{x}\big)|_{z=\bar{x}} = \max_{(\alpha,\beta)\in\Sigma(\bar{x})} \left\{ \sum_{i\in I} \alpha_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \beta_j h_j'(\bar{x}; x - \bar{x}) \right\}, \tag{5}$$

where

$$\Sigma(\bar{x}) := \left\{ (\alpha, \beta) \in \Sigma \;\middle|\; \sum_{i\in I} \alpha_i \big(f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})\big) + \sum_{j\in J} \beta_j h_j(\bar{x}) = 0 \right\}$$

since we have $\mathcal{F}(\bar{x}, \bar{\lambda}) = 0$. Therefore, with (4) and (5) we arrive to

$$\inf_{x\in C} \max_{(\alpha,\beta)\in\Sigma(\bar{x})} \left\{ \sum_{i\in I} \alpha_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \beta_j h_j'(\bar{x}; x - \bar{x}) \right\} \geq 0. \tag{6}$$

Note that the function

$$(x, (\alpha, \beta)) \mapsto \sum_{i\in I} \alpha_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \beta_j h_j'(\bar{x}; x - \bar{x})$$

is linear with respect to $(\alpha, \beta) \in \Sigma(\bar{x})$, with $\Sigma(\bar{x})$ being compact, and by our assumption, convex with respect to x. Then, by Theorem 2 of [31] we can interchange the "inf" and "sup" in (6) to get

$$\max_{(\alpha,\beta)\in\Sigma(\bar{x})} \inf_{x\in C} \left\{ \sum_{i\in I} \alpha_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \beta_j h_j'(\bar{x}; x - \bar{x}) \right\} \geq 0.$$

From the compactness of the set $\Sigma(\bar{x})$ and the upper semicontinuity of the function

$$(\alpha, \beta) \mapsto \inf_{x\in C} \left\{ \sum_{i\in I} \alpha_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \beta_j h_j'(\bar{x}; x - \bar{x}) \right\}$$

we conclude that the latter achieves its maximum on $\Sigma(\bar{x})$. Thus, there exists $(\bar{\alpha}, \bar{\beta}) \in \Sigma(\bar{x})$ such that

$$\inf_{x\in C} \left\{ \sum_{i\in I} \bar{\alpha}_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \right\} \geq 0.$$

On the other hand, $\sum_{i\in I} \bar{\alpha}_i [f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] + \sum_{j\in J} \bar{\beta}_j h_j(\bar{x}) = 0$ since $(\bar{\alpha}, \bar{\beta}) \in \Sigma(\bar{x})$. Taking into account that $f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x}) \leq 0$ and $h_j(\bar{x}) \leq 0$ for all $(i, j) \in I \times J$, we conclude that $\bar{\alpha}_i [f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $(i, j) \in I \times J$. Finally, we have

$$\sum_{i\in I} \bar{\alpha}_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C, \tag{7}$$

with $\bar{\alpha}_i [f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $(i, j) \in I \times J$. ☐

Remark 1. Let for $i \in I$, $r_i(x) = f_i(x)/g_i(x)$. If the functions f_i and g_i are directionally differentiable at $\bar{x}$ and that $f_i'(\bar{x};\cdot) - \bar{\lambda} g_i'(\bar{x};\cdot)$, for those $i \in I(\bar{x}) := \{i \in I \mid \bar{\lambda} = f_i(\bar{x})/g_i(\bar{x})\}$, are convex as functions of the directions, then the functions r_i are directionally differentiable at $\bar{x}$ and $r_i'(\bar{x};\cdot)$ are convex as functions of the directions, and *vice versa*. This is because for $i \in I(\bar{x})$, we have

$$\begin{aligned} f_i'(\bar{x};\cdot) - \bar{\lambda} g_i'(\bar{x};\cdot) &= f_i'(\bar{x};\cdot) - \frac{f_i(\bar{x})}{g_i(\bar{x})} g_i'(\bar{x};\cdot) \\ &= g_i(\bar{x}) \frac{g_i(\bar{x}) f_i'(\bar{x};\cdot) - f_i(\bar{x}) g_i'(\bar{x};\cdot)}{g_i(\bar{x})^2} \\ &= g_i(\bar{x}) r_i'(\bar{x};\cdot). \end{aligned}$$

In what follows we will express the conditions of the last theorem in terms of the directional derivatives of the ratios r_i.

Corollary 1. *Let $\bar{x} \in X$ and $\bar{\lambda} = \lambda(\bar{x})$. Assume that the functions f_i, g_i and h_j are directionally differentiable at $\bar{x}$ and that $f_i'(\bar{x};\cdot) - \bar{\lambda} g_i'(\bar{x};\cdot)$ and $h_j'(\bar{x};\cdot)$ for $(i,j) \in I \times J$, are convex as functions of the directions. If $\bar{x}$ is an optimal solution of (P), then there exists $(\bar{\mu}, \bar{\nu}) \in \Sigma$ such that*

$$\sum_{i\in I} \bar{\mu}_i r_i'(\bar{x}; x - \bar{x}) + \sum_{j \in J} \bar{\nu}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C, \tag{8}$$

with $\bar{\mu}_i[r_i(\bar{x}) - \bar{\lambda}] = 0$ and $\bar{\nu}_j h_j(\bar{x}) = 0$ for all $(i,j) \in I \times J$.

Proof. From Theorem 1, there exists $(\bar{\alpha}, \bar{\beta}) \in \Sigma$ such that

$$\sum_{i\in I} \bar{\alpha}_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C, \tag{9}$$

with $\bar{\alpha}_i[f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $(i,j) \in I \times J$. Note firstly that the condition $\bar{\alpha}_i[f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ entails that $\bar{\alpha}_i = 0$ whenever $f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x}) < 0$, or equivalently when $r_i(\bar{x}) < \bar{\lambda} := \lambda(\bar{x})$, which implies that $\bar{\alpha}_i[r_i(\bar{x}) - \bar{\lambda}] = 0$. Therefore, we can write (9) as

$$\sum_{i\in I} \bar{\alpha}_i \left(f_i'(\bar{x}; x - \bar{x}) - \frac{f_i(\bar{x})}{g_i(\bar{x})} g_i'(\bar{x}; x - \bar{x}) \right) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C.$$

Then we immediately have, for all $x \in C$,

$$\sum_{i\in I} \bar{\alpha}_i g_i(\bar{x}) \left(\frac{g_i(\bar{x}) f_i'(\bar{x}; x - \bar{x}) - f_i(\bar{x}) g_i'(\bar{x}; x - \bar{x})}{g_i^2(\bar{x})} \right) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0.$$

This nothing but the following formula

$$\sum_{i\in I} \bar{\alpha}_i g_i(\bar{x}) \left(\frac{f_i}{g_i} \right)' (\bar{x}; x - \bar{x}) + \sum_{j\in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C. \tag{10}$$

Let for $(i,j) \in I \times J$,

$$\bar{\mu}_i = \frac{\bar{\alpha}_i g_i(\bar{x})}{\sum_{i\in I} \bar{\alpha}_i g_i(\bar{x}) + \sum_{j\in J} \bar{\beta}_j} \quad \text{and} \quad \bar{\nu}_j = \frac{\bar{\beta}_j}{\sum_{i\in I} \bar{\alpha}_i g_i(\bar{x}) + \sum_{j\in J} \bar{\beta}_j}.$$

Clearly, $(\bar{\mu}, \bar{\nu}) \in \Sigma$, and (10) becomes

$$\sum_{i\in I} \bar{\mu}_i r_i'(\bar{x}; x-\bar{x}) + \sum_{j\in J} \bar{\nu}_j h_j'(\bar{x}; x-\bar{x}) \geq 0 \quad \text{for all } x \in C. \tag{11}$$

Since $\bar{\alpha}_i[f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$, then it can easily seen that $\bar{\mu}_i[r_i(\bar{x}) - \bar{\lambda}] = 0$ and $\bar{\nu}_j h_j(\bar{x}) = 0$ for all $(i,j) \in I \times J$. □

Now we can express the previous optimality conditions in terms of the directional derivative of λ. For this we need the following Mangasarian–Fromovitz type condition to hold, see *e.g.*, [39].

Assumption 1. *For $\bar{x} \in C$ such that $h(\bar{x}) = 0$, there exist $\hat{x} \in C$ such that $h'(\bar{x}; \hat{x} - \bar{x}) < 0$.*

Corollary 2. *Let $\bar{x} \in X$ and $\bar{\lambda} = \lambda(\bar{x})$. Assume that the functions f_i, g_i and h_j are directionally differentiable at $\bar{x}$ and that $f_i'(\bar{x}; \cdot) - \bar{\lambda} g_i'(\bar{x}; \cdot)$ and $h_j'(\bar{x}; \cdot)$ for $(i,j) \in I \times J$, are convex as functions of the directions. Suppose in addition that Assumption 1 holds. Then, if $\bar{x}$ is an optimal solution of (P), there exist $\bar{\xi}_j \geq 0$ for $j \in J$, such that*

$$\lambda'(\bar{x}; x-\bar{x}) + \sum_{j\in J} \bar{\xi}_j h_j'(\bar{x}; x-\bar{x}) \geq 0 \quad \text{for all } x \in C, \tag{12}$$

with $\bar{\xi}_j h_j(\bar{x}) = 0$ for all $j \in J$.

Proof. From these assumptions we have the conclusions of Corollary 1. On the other hand Assumption 1 entails that in (8), $\bar{\mu}_i \neq 0$ for at least one $i \in I$. Dividing in (8) by $\sum_{i\in I} \bar{\mu}_i$ and letting for $(i,j) \in I \times J$,

$$\bar{\mu}_i' = \frac{\bar{\mu}_i}{\sum_{i\in I} \bar{\mu}_i} \quad \text{and} \quad \bar{\nu}_j' = \frac{\bar{\nu}_j}{\sum_{i\in I} \bar{\mu}_i},$$

we get

$$\sum_{i\in I} \bar{\mu}_i' r_i'(\bar{x}; x-\bar{x}) + \sum_{j\in J} \bar{\nu}_j' h_j'(\bar{x}; x-\bar{x}) \geq 0 \quad \text{for all } x \in C,$$

with $\bar{\mu}_i'[r_i(\bar{x}) - \bar{\lambda}] = 0$ and $\bar{\nu}_j' h_j(\bar{x}) = 0$ for all $(i,j) \in I \times J$. Since $\sum_{i\in I} \bar{\mu}_i' = 1$ we get $\sum_{i\in I} \bar{\mu}_i' r_i(\bar{x}) = \bar{\lambda}$, and we can conclude that $\sum_{i\in I} \bar{\mu}_i' r_i'(\bar{x}; x-\bar{x}) = \lambda'(\bar{x}; x-\bar{x})$, which gives the proof by taking $\bar{\xi}_j = \bar{\nu}_j'$ for $j \in J$. □

Below we discuss sufficiency of the optimality conditions stated in Theorem 1. For this we define the function

$$f(x,\mu) := \max_{i\in I}\{f_i(x) - \mu g_i(x)\} \tag{13}$$

for $x \in \mathbb{R}^n$ and $\mu \in \mathbb{R}$, and we need the following assumptions. The first assumption is

Assumption 2. *The constraint set X is nonempty and satisfies the equality $X = \mathrm{cl}\{x \in \mathrm{int}\, C \mid h(x) < 0\}$, where "cl" and "int" stand respectively for the closure and the interior.*

The second one is the non-degeneracy condition of Mangasarian–Fromovitz type.

Assumption 3. *For $\bar{x} \in X$ such that $h(\bar{x}) = 0$, there exist $u, v \in \mathbb{R}^n$ such that $\lambda'(\bar{x}; u) < 0$ and $h'(\bar{x}; v) < 0$.*

Remark 2. It is straightforward to verify that $\lambda'(\bar{x}; u) < 0$ is equivalent to saying that $f_i'(\bar{x}; u) - \bar{\lambda} g_i'(\bar{x}; u) < 0$ for all $i \in I$ such that $f_i(\bar{x})/g_i(\bar{x}) = \bar{\lambda}$ where $\bar{\lambda} := \lambda(\bar{x})$; which is equivalent to $f'((z,\bar{\lambda}); u)|_{z=\bar{x}} < 0$, where $f'((z,\bar{\lambda}); u)|_{z=\bar{x}}$ is the directional derivative of the function $f(\cdot, \bar{\lambda})$, defined in (13), at $\bar{x}$ in the direction u.

The next theorem is a reciprocal to Theorem 1.

Theorem 2. *Let $\bar{x} \in X$ be such that $h(\bar{x}) = 0$, and let $\bar{\lambda} = \lambda(\bar{x})$. Assume that the functions f_i, g_i and h_j for $(i, j) \in I \times J$, are directionally differentiable at $\bar{x}$, and that the functions λ and h are continuous. Suppose on the other hand that $\bar{x}$ satisfies the KKT conditions of Theorem 1. In addition to Assumptions 2 and 3, if the set $S := \{x \in C \mid \mathcal{F}(x, \bar{\lambda}) \leq 0\}$ is convex then $\bar{x}$ is an optimal solution for (P).*

Proof. Let $\bar{x} \in X$ be such that $h(\bar{x}) = 0$, $\bar{\lambda} = \lambda(\bar{x})$ and let $S := \{x \in C \mid \mathcal{F}(x, \bar{\lambda}) \leq 0\}$ where $\mathcal{F}(x, \bar{\lambda}) = \max\{f(x, \bar{\lambda}), h(x)\}$. Note that $\bar{x} \in S$, and that $\mathcal{F}(x, \bar{\lambda}) \leq 0 = \mathcal{F}(\bar{x}, \bar{\lambda})$ and $h(x) \leq 0 = h(\bar{x})$ for all $x \in S$. So, from the convexity of S we conclude that

$$f'\big((z, \bar{\lambda}); x - \bar{x}\big) \mid_{z=\bar{x}} \leq 0 \text{ and } h'(\bar{x}; x - \bar{x}) \leq 0 \quad \text{for all } x \in S. \tag{14}$$

Recall that by assumption we have

$$\sum_{i \in I} \bar{\alpha}_i \big(f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big) + \sum_{j \in J} \bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) \geq 0 \quad \text{for all } x \in C. \tag{15}$$

Note on the other hand that the conditions $\bar{\alpha}_i[f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$ entail that $\bar{\alpha}_i = 0$ when $f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x}) < 0$, *i.e.*, $f_i(\bar{x})/g_i(\bar{x}) < \bar{\lambda}$, and $\bar{\beta}_j = 0$ when $h_j(\bar{x}) < 0$, which implies, thanks to (2), that for all $(i, j) \in I \times J$, it holds

$$\bar{\alpha}_i \big[f_i'(\bar{x}; x - \bar{x}) - \bar{\lambda} g_i'(\bar{x}; x - \bar{x})\big] = \bar{\alpha}_i f'\big((z, \bar{\lambda}); x - \bar{x}\big) \mid_{z=\bar{x}}$$

and

$$\bar{\beta}_j h_j'(\bar{x}; x - \bar{x}) = \bar{\beta}_j h'(\bar{x}; x - \bar{x}).$$

Therefore, equation (15) gives, in particular for $x \in S$,

$$f'\big((z, \bar{\lambda}); x - \bar{x}\big) \mid_{z=\bar{x}} \sum_{i \in I} \bar{\alpha}_i + h'(\bar{x}; x - \bar{x}) \sum_{j \in J} \bar{\beta}_j \geq 0 \quad \text{for all } x \in S. \tag{16}$$

Thus, equation (16) together with (14) entails that

$$f'\big((z, \bar{\lambda}); x - \bar{x}\big) \mid_{z=\bar{x}} \sum_{i \in I} \bar{\alpha}_i = 0 \text{ and } h'(\bar{x}; x - \bar{x}) \sum_{j \in J} \bar{\beta}_j = 0 \quad \text{for all } x \in S. \tag{17}$$

Assume now that there exists $\hat{x} \in \operatorname{int} C$ such that $\lambda(\hat{x}) < \lambda(\bar{x})$ and $h(\hat{x}) < 0$. A fortiori, $\hat{x} \in S$. The continuity of the functions λ and h implies then that for all $d \in \mathbb{R}^n$, there exists a $t > 0$ such that $\lambda(\hat{x} + td) < \lambda(\bar{x})$ and $h(\hat{x} + td) < 0$. This implies, in particular, that $\hat{x} + td \in S$. So, by virtue of (17), it follows that

$$f'\big((z, \bar{\lambda}); \hat{x} + td - \bar{x}\big) \mid_{z=\bar{x}} \sum_{i \in I} \bar{\alpha}_i = 0 \text{ and } h'(\bar{x}; \hat{x} + td - \bar{x}) \sum_{j \in J} \bar{\beta}_j = 0.$$

Using the sublinearity and the homogeneity of $f'((z, \bar{\lambda}); \cdot) \mid_{z=\bar{x}}$ and $h'(\bar{x}; \cdot)$ we arrive to

$$\big[f'\big((z, \bar{\lambda}); \hat{x} - \bar{x}\big) \mid_{z=\bar{x}} + t f'\big((z, \bar{\lambda}); d\big) \mid_{z=\bar{x}}\big] \sum_{i \in I} \bar{\alpha}_i \geq 0 \tag{18}$$

and

$$(h'(\bar{x}; \hat{x} - \bar{x}) + t h'(\bar{x}; d)) \sum_{j \in J} \bar{\beta}_j \geq 0. \tag{19}$$

Taking into account the equalities in (17), with $x = \hat{x}$, we deduce from (18) and (19) that

$$f'\big((z, \bar{\lambda}); d\big) \mid_{z=\bar{x}} \sum_{i \in I} \bar{\alpha}_i \geq 0 \text{ and } h'(\bar{x}; d) \sum_{j \in J} \bar{\beta}_j \geq 0 \quad \text{for all } d \in \mathbb{R}^n.$$

Since $(\bar{\alpha}, \bar{\beta}) \in \Sigma$, we get $f'((z, \bar{\lambda}); d) \mid_{z=\bar{x}} \geq 0$ for all $d \in \mathbb{R}^n$, or $h'(\bar{x}; d) \geq 0$ for all $d \in \mathbb{R}^n$, which contradicts our assumption (see Rem. 2). Thus, our hypothesis is false, that is, for all $x \in \operatorname{int} C$ such that $h(x) < 0$, we must have $\lambda(x) \geq \lambda(\bar{x})$. Now, if $x \in C$ is such that $h(x) = 0$, then by Assumption 2 there must exist a sequence $\{x_k\} \subset \operatorname{int} C$ converging to x such that $h(x_k) < 0$ for all $k \in \mathbb{N}$. Therefore, from the previous conclusion we must have $\lambda(x_k) \geq \lambda(\bar{x})$. The continuity of the function λ entails that $\lambda(x) \geq \lambda(\bar{x})$. Finally, we showed that $\lambda(x) \geq \lambda(\bar{x})$ for all $x \in X$, and this gives the desired result. □

4. Upper bounding functions for GFP

We define an upper approximating function $\mathcal{U}(\cdot, y)$ of $\mathcal{F}(\cdot, \mu)$, with $\mu = \lambda(y)$, on C at $y \in X$, to be a function defined on $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying

(A1) $\mathcal{U}(y, y) = \mathcal{F}(y, \mu)$ where $\mu = \lambda(y)$ and $y \in X$,
(A2) $\mathcal{U}(x, y) \geq \mathcal{F}(x, \mu)$ when $\mu = \lambda(y)$ for all $(x, y) \in C \times X$,
(A3) $\mathcal{U}'((z, y); x - y) \mid_{z=y} = \mathcal{F}'((z, \mu); x - y) \mid_{z=y}$ with $\mu = \lambda(y)$, for all $(x, y) \in C \times X$, where we denote by $\mathcal{U}'((z, y); x - y) \mid_{z=y}$ and $\mathcal{F}'((z, \mu); x - y) \mid_{z=y}$, respectively, the directional derivatives of $\mathcal{U}(\cdot, y)$ and $\mathcal{F}(\cdot, \mu)$ at y in the direction $x - y$.

Below we list some examples of problems for which we have upper approximating functions $\mathcal{U}$ that satisfies (A1)–(A3).

Example 1. Assume that the functions f_i, g_i for $i \in I$, and h_j for $j \in J$, are directionally differentiable. For any $y \in X$, let $\lambda = \lambda(y)$ and define

$$\mathcal{U}(x, y) := \mathcal{F}(x, \lambda).$$

This function is the objective function of the parametric problems appearing in the method of centers for GFP given in [44].

Example 2. Assume that the functions f_i, g_i for $i \in I$, and h_j for $j \in J$, are directionally differentiable and let $\alpha > 0$. For any $y \in X$, let $\mu = \lambda(y)$ and define

$$\mathcal{U}(x, y) := \mathcal{F}(x, \mu) + \tfrac{\alpha}{2}\|x - y\|^2.$$

This function is used in the prox-method of centers given in [1].

Example 3. Assume that the functions f_i, g_i for $i \in I$, and h_j for $j \in J$, are directionally differentiable and let $\alpha > 0$. For any $y \in X$, let $\mu = \lambda(y)$ and define

$$\mathcal{U}(x, y) := \mathcal{F}(x, \mu) + \alpha \mathcal{D}_\psi(x, y),$$

where $\mathcal{D}_\psi$ is a Bregman function [22]. This function is a generalization of the function given in the previous example, and was introduced in [1].

Example 4. Assume that the functions f_i and g_i for $i \in I$, are convex, and that h_j for $j \in J$, are directionally differentiable. Suppose on the other hand that the function g_i is continuously differentiable for all $i \in I$. Let, for $i \in I$, the functions

$$g_{i,y}(x) := \begin{cases} g_i(y) + \langle \nabla g_i(y), x - y \rangle & \text{if } \lambda(y) > 0 \\ g_i(x) & \text{if } \lambda(y) \leq 0. \end{cases}$$

Define the approximating function

$$\mathcal{U}(x, y) := \max_{(i,j) \in I \times J} \{f_i(x) - \lambda(y) g_{i,y}(x), h_j(x)\}.$$

If for $j \in J$, the functions h_j are convex, $\mathcal{U}(\cdot, y)$ is then an upper convex approximation of $\mathcal{F}(\cdot, \mu)$ at y.

Example 5. Assume that the functions f_i and $-g_i$ for $i \in I$, are convex, and that h_j for $j \in J$, are directionally differentiable. Suppose on the other hand that the function g_i is continuously differentiable for all $i \in I$. Let, for $i \in I$, the functions

$$g_{i,y}(x) := \begin{cases} g_i(y) + \langle \nabla g_i(y), x - y \rangle & \text{if } \lambda(y) < 0 \\ g_i(x) & \text{if } \lambda(y) \geq 0. \end{cases}$$

Define the approximating function

$$\mathcal{U}(x, y) := \max_{(i,j) \in I \times J} \{ f_i(x) - \lambda(y) g_{i,y}(x), h_j(x) \}.$$

If for $j \in J$, the functions h_j are convex, the $\mathcal{U}(\cdot, y)$ is an upper convex approximation of $\mathcal{F}(\cdot, \mu)$ at y.

Example 6. Assume that the functions f_i, g_i for $i \in I$, and h_j for $j \in J$, are continuously differentiable with L-Lipschitz gradients. Then one can easily prove that $f_i - \lambda(y) g_i$ is L_y-Lipschitz gradient, where $L_y := L(1+|\lambda(y)|)$, and from the descent lemma, see *e.g.*, Proposition A.24 of [14], for any $x, z \in \mathbb{R}^n$, $y \in X$, and all $i \in I$, we have

$$f_i(x) - \lambda(y) g_i(x) \leq f_i(z) - \lambda(y) g_i(z) + \langle \nabla f_i(z) - \lambda(y) \nabla g_i(z), x - z \rangle + \tfrac{L_y}{2} \|x - z\|^2 \tag{20}$$

and, since $L \leq L_y$, for all $j \in J$,

$$h_j(x) \leq h_j(y) + \langle \nabla h_j(y), x - y \rangle + \tfrac{L_y}{2} \|x - y\|^2. \tag{21}$$

Define the function

$$\mathcal{U}(x, y) := \mathcal{L}(x, y) + \tfrac{L_y}{2} \|x - y\|^2,$$

where

$$\mathcal{L}(x, y) := \max_{(i,j) \in I \times J} \{ \langle \nabla f_i(y) - \lambda(y) \nabla g_i(y), x - y \rangle + f_i(y) - \lambda(y) g_i(y), \langle \nabla h_j(y), x - y \rangle + h_j(y) \}.$$

This example will be the subject of further details in Section 6.

Example 7. Let the functions $f_i = f_i^1 - f_i^2$, $g_i = g_i^1 - g_i^2$ and $h_j = h_j^1 - h_j^2$ for $(i, j) \in I \times J$, and assume that f_i^ℓ, g_i^ℓ and h_j^ℓ for $\ell = 1, 2$ and $(i, j) \in I \times J$, are convex. Suppose on the other hand that the functions f_i^2, g_i^ℓ for $\ell = 1, 2$, and h_j^2 for $j \in J$, are continuously differentiable. Define the functions

$$\begin{aligned} f_{i,y}(x) &= f_i^1(x) - \left[f_i^2(y) + \left\langle \nabla f_i^2(y), x - y \right\rangle \right], \\ g_{i,y}(x) &:= \begin{cases} g_i^1(x) - \left[g_i^2(y) + \left\langle \nabla g_i^2(y), x - y \right\rangle \right] & \text{if } \lambda(y) < 0 \\ -g_i^2(x) + \left[g_i^1(y) + \left\langle \nabla g_i^1(y), x - y \right\rangle \right] & \text{if } \lambda(y) \geq 0, \end{cases} \\ h_{j,y}(x) &= h_j^1(x) - \left[h_j^2(y) + \left\langle \nabla h_j^2(y), x - y \right\rangle \right], \end{aligned}$$

and let

$$\mathcal{U}(x, y) = \max_{(i,j) \in I \times J} \{ f_{i,y}(x) - \lambda(y) g_{i,y}(x), h_{j,y}(x) \}.$$

This example was developed in [21].

Now we will investigate consequences and elementary properties induced by the introduction of upper bounding approximation functions, especially assumptions (A1) and (A2). We will see in particular, what happens for λ by minimizing $\mathcal{U}(\cdot, y)$ over C.

Proposition 2. *For all $y \in X$, we have*

(1) $\mathcal{U}(y,y)=0$.
(2) *Let* x_y *be a global minimizer of* $\mathcal{U}(\cdot,y)$ *over* C*. Then,* $h(x_y)\leq 0$ *and* $\lambda(x_y)\leq\lambda(y)$*.*

Proof. (1) The definition of $\lambda(y)$ entails that $f_i(y)/g_i(y)\leq\lambda(y)$ for all $i\in I$, with equality for some index i_0. Since g_i is assumed to be positive on X we conclude that $f_i(y)-\lambda(y)g_i(y)\leq 0$ for all $i\in I$, with equality for $i=i_0$. On the other hand, the fact that $h(y)\leq 0$, *i.e.*, $h_j(y)\leq 0$ for all $j\in J$, gives $\mathcal{F}(y,\lambda(y))=0$. It follows from (A1) that $\mathcal{U}(y,y)=0$ for all $y\in X$.

(2) Since the point x_y is an optimal solution on C of $\mathcal{U}(\cdot,y)$ we get, in particular, $\mathcal{U}(x_y,y)\leq\mathcal{U}(y,y)$. By using Item 1 we obtain $\mathcal{U}(x_y,y)\leq 0$. It follows from (A2) that $\mathcal{F}(x_y,\lambda(y))\leq 0$. The definition of $\mathcal{F}(x_y,\lambda(y))$ entails that $f_i(x_y)-\lambda(y)g_i(x_y)\leq 0$ for all $i\in I$ and $h(x_y)\leq 0$, or equivalently $f_i(x_y)/g_i(x_y)\leq\lambda(y)$ for all $i\in I$, and $h(x_y)\leq 0$. Finally we have $\lambda(x_y)\leq\lambda(y)$ and $x_y\in X$.

□

Remark 3. It is clear from Item 2 that for $y\in X$, the minimization of $\mathcal{U}(\cdot,y)$ permits to get a feasible point that decreases λ.

In the next proposition we see links between minimizing λ on X and minimizing $\mathcal{U}(\cdot,y)$ on C, for particular parameters $y\in X$.

Proposition 3. *Let* $\bar{x}\in X$ *be a global optimal solution of* (P)*, and let the function* $\mathcal{U}$ *satisfies (A1) and (A2). Then* $\bar{x}$ *globally minimizes* $\mathcal{U}(\cdot,\bar{x})$ *over* C*. Conversely, for all global optimal solution* $\bar{x}$ *of* (P)*, every optimal solution* $x_{\bar{x}}$ *of* $\mathcal{U}(\cdot,\bar{x})$ *over* C *also globally solves* (P)*.*

Proof. Let $\bar{x}\in X$ be a global optimal solution of (P), and let $\bar{\lambda}=\lambda(\bar{x})$. By using the condition (A2), with $y=\bar{x}$, we get $\mathcal{U}(x,\bar{x})\geq\mathcal{F}(x,\bar{\lambda})$ for all $x\in C$. Note that $\mathcal{F}(x,\bar{\lambda})\geq 0$ for all $x\in C$, because for $x\in X$ we have $\lambda(x)\geq\bar{\lambda}$, and thus $f_i(x)-\bar{\lambda}g_i(x)\geq 0$ for at least one $i\in I$; and for $x\notin X$ but $x\in C$ we have $h(x)>0$. It follows that $\mathcal{F}(x,\bar{\lambda})\geq 0$ for all $x\in C$, thereby implying that $\mathcal{U}(x,\bar{x})\geq 0$ for all $x\in C$. The fact that $\mathcal{F}(\bar{x},\bar{\lambda})=0$ entails that $\mathcal{U}(\bar{x},\bar{x})=0$ by assumption (A1). Finally $\mathcal{U}(x,\bar{x})\geq\mathcal{U}(\bar{x},\bar{x})$ for all $x\in C$, which means that $\bar{x}$ minimizes $\mathcal{U}(\cdot,\bar{x})$ over C. Conversely, from Proposition 2, Item 2, we have with $y=\bar{x}$ that $\lambda(x_{\bar{x}})\leq\lambda(\bar{x})$ and $h(x_{\bar{x}})\leq 0$. Thus, $\lambda(x_{\bar{x}})=\lambda(\bar{x})$ and $x_{\bar{x}}$ solves problem (P). □

The following result will serve as practical information to detect points satisfying the necessary optimality KKT conditions stated in Theorem 1.

Theorem 3. *Let for* $y\in X$*,* x_y *be a global minimizer of* $\mathcal{U}(\cdot,y)$ *over* C*. Assume that the functions* f_i*,* g_i *and* h_j *are directionally differentiable at* y *and that* $f_i'(y;\cdot)-\lambda(y)g_i'(y;\cdot)$ *and* $h_j'(y;\cdot)$ *for* $(i,j)\in I\times J$*, are convex as functions of the directions. Suppose on the other hand that the upper bounding function* $\mathcal{U}$ *satisfies conditions (A1)–(A3). If* $\mathcal{U}(x_y,y)=0$*, then there exists* $(\alpha^y,\beta^y)\in\Sigma$ *such that*

$$\sum_{i\in I}\alpha_i^y(f_i'(y;x-y)-\lambda(y)g_i'(y;x-y))+\sum_{j\in J}\beta_j^y h_j'(y;x-y)\geq 0\quad\text{for all } x\in C,$$

with $\alpha_i^y[f_i(y)-\lambda(y)g_i(y)]=0$ *and* $\beta_j^y h_j(y)=0$ *for all* $(i,j)\in I\times J$*.*

Proof. Since x_y is a global optimal solution on C of $\mathcal{U}(\cdot,y)$, we have

$$\mathcal{U}(x,y)\geq\mathcal{U}(x_y,y)\quad\text{for all } x\in C. \tag{22}$$

The assumption that $\mathcal{U}(x_y,y)=0$, together with (22) and the fact that $\mathcal{U}(y,y)=0$ gives

$$\mathcal{U}(x,y)\geq 0=\mathcal{U}(y,y)\quad\text{for all } x\in C,$$

that is y is also a global optimal solution on C of $\mathcal{U}(\cdot, y)$. Taking into account that C is convex, it follows that

$$\mathcal{U}'((z,y); x-y) \mid_{z=y} \geq 0 \quad \text{for all } x \in C.$$

From assumption (A3) we conclude that

$$\mathcal{F}'((z,\lambda(y)); x-y) \mid_{z=y} \geq 0 \quad \text{for all } x \in C.$$

This is what we obtained in (4). It suffices to note that in this case the set $\Sigma(y)$ is given by

$$\Sigma(y) := \left\{ (\alpha,\beta) \in \Sigma \;\middle|\; \sum_{i\in I} \alpha_i (f_i(y) - \lambda(y) g_i(y)) + \sum_{j\in J} \beta_j h_j(y) = 0 \right\},$$

the rest of the proof is analogous to that of Theorem 1. □

At this introductory stage of our analysis, in order to construct an iterative procedure based on $\mathcal{U}$ to solve problem (P), the previous results suggest us to approximate the unknown optimal solution $\bar{x}$ by some x^k, at an iteration k, and then to minimize over C the function $\mathcal{U}(\cdot, x^k)$ to obtain another approximation x^{k+1} of $\bar{x}$ and so on, hoping to converge to an optimal solution. The next algorithm summarizes this procedure which we will call successive upper bounding approximation method.

5. Successive upper bounding approximation algorithm for GFP

In this section we introduce the general sequential upper bounding algorithm to solve the problem (P). It is assumed that one has in hand an upper bounding approximation function $\mathcal{U}$ satisfying (A1)–(A3).

Algorithm 1. Sequential upper bounding algorithm.

0. Choose $x^0 \in X$ and let $k = 0$.
1. Find $x^{k+1} \in C$ such that $\mathcal{U}(x^{k+1}, x^k) \leq \mathcal{U}(x, x^k)$ for all $x \in C$.
2. If $\mathcal{U}(x^{k+1}, x^k) \neq 0$, set $k = k+1$ and return to 1.

Algorithm 1 halts when $\mathcal{U}(x^{k+1}, x^k) = 0$. We explain in the following result this stopping rule.

Proposition 4. *Suppose that at some step k we have $\mathcal{U}(x^{k+1}, x^k) = 0$. Assume on the other hand that the functions f_i, g_i and h_j are directionally differentiable at x^k and that $f_i'(x^k;\cdot) - \lambda_k g_i'(x^k;\cdot)$ and $h_j'(x^k;\cdot)$ for $(i,j) \in I \times J$, are convex as functions of the directions. Finally suppose that the upper bounding function $\mathcal{U}$ satisfies conditions (A1)–(A3). Then, there exists $(\alpha^k, \beta^k) \in \Sigma$ such that*

$$\sum_{i\in I} \alpha_i^k \big(f_i'\big(x^k; x - x^k\big) - \lambda_k g_i'\big(x^k; x - x^k\big)\big) + \sum_{j\in J} \beta_j^k h_j'\big(x^k; x - x^k\big) \geq 0$$

for all $x \in C$, with $\alpha_i^k[f_i(x^k) - \lambda_k g_i(x^k)] = 0$ and $\beta_j^k h_j(x^k) = 0$ for all $(i,j) \in I \times J$.

Proof. This is Theorem 3 with $y = x^k$ and $x_y = x^{k+1}$. □

In the following we will discuss convergence properties of the sequence generated by the previous algorithm, when infinite. Later we will use the notation

$$\Delta := \sup_{x\in X} \max_{i\in I} g_i(x).$$

We also set $\lambda_k = \lambda(x^k)$ for all $k \in \mathbb{N}$.

Proposition 5. (1) *The sequence $\{\lambda_k\}$ is nonincreasing and converges to some $\hat{\lambda} \geq \bar{\lambda}$.*

(2) *If $\Delta < \infty$, the sequences $\{\mathcal{F}(x^{k+1}, \lambda_k)\}$ and $\{\mathcal{U}(x^{k+1}, x^k)\}$ converge to 0. In particular, if $\hat{x}$ is a cluster point of the sequence $\{x^k\}$, and the functions λ and h are continuous then the sequences $\{\mathcal{F}(x^{k+1}, \lambda_k)\}$ and $\{\mathcal{U}(x^{k+1}, x^k)\}$ converge, respectively, to $\mathcal{F}(\hat{x}, \hat{\lambda})$ and $\mathcal{U}(\hat{x}, \hat{x})$.*

Proof. (1) For all $k \in \mathbb{N}$, we have

$$\mathcal{F}\big(x^{k+1}, \lambda_k\big) \leq \mathcal{U}\big(x^{k+1}, x^k\big) \leq \mathcal{U}\big(x^k, x^k\big) = \mathcal{F}\big(x^k, \lambda_k\big),$$

where the first inequality follows from (A2), the second from the definition of x^{k+1} and the equality from (A1). Using the definition of $\mathcal{F}(x^{k+1}, \lambda_k)$ and that $x^k \in X$, we get

$$\max_{(i,j)\in I\times J} \big\{f_i(x^{k+1}) - \lambda_k g_i\big(x^{k+1}\big), h_j\big(x^{k+1}\big)\big\} \leq \mathcal{F}\big(x^{k+1}, \lambda_k\big) \leq \mathcal{U}\big(x^{k+1}, x^k\big) \leq 0. \tag{23}$$

Then $h_j(x^{k+1}) \leq 0$ for all $j \in J$, and $f_i(x^{k+1})/g_i(x^{k+1}) \leq \lambda_k$ for all $i \in I$. It follows that $h(x^k) \leq 0$ and $\lambda_{k+1} \leq \lambda_k$ for all $k \in \mathbb{N}$. Therefore, the fact that $\lambda_k \geq \bar{\lambda}$ for all $k \in \mathbb{N}$ implies that the sequence $\{\lambda_k\}$ converges to some $\hat{\lambda} \geq \bar{\lambda}$.

(2) From (23) we obtain

$$f_i\big(x^{k+1}\big) - \lambda_k g_i\big(x^{k+1}\big) \leq \mathcal{F}\big(x^{k+1}, \lambda_k\big) \leq \mathcal{U}\big(x^{k+1}, x^k\big) \leq 0 \quad \text{for all } i \in I,$$

and in particular, with $i = i_k$ such $\lambda_{k+1} = f_{i_k}(x^{k+1})/g_{i_k}(x^{k+1})$, we get

$$g_{i_k}\big(x^{k+1}\big)\big(f_{i_k}\big(x^{k+1}\big)/g_{i_k}\big(x^{k+1}\big) - \lambda_k\big) \leq \mathcal{F}\big(x^{k+1}, \lambda_k\big) \leq \mathcal{U}\big(x^{k+1}, x^k\big) \leq 0,$$

which gives

$$\Delta(\lambda_{k+1} - \lambda_k) \leq \mathcal{F}\big(x^{k+1}, \lambda_k\big) \leq \mathcal{U}\big(x^{k+1}, x^k\big) \leq 0.$$

This implies that the sequences $\{\mathcal{F}(x^{k+1}, \lambda_k)\}$ and $\{\mathcal{U}(x^{k+1}, x^k)\}$ converge to 0. Let $\hat{x}$ be a cluster point of the sequence $\{x^k\}$. Then by the continuity of λ and h we have $\hat{\lambda} = \lambda(\hat{x})$, $h(\hat{x}) \leq 0$, and consequently $\{\mathcal{F}(x^{k+1}, \lambda_k)\}$ and $\{\mathcal{U}(x^{k+1}, x^k)\}$ converge, respectively, to $0 = \mathcal{F}(\hat{x}, \hat{\lambda})$ and $0 = \mathcal{U}(\hat{x}, \hat{x})$.

□

The following theorem shows that any cluster points of the sequence $\{x^k\}$ satisfies the optimality conditions stated in Theorem 1, and then those of Corollaries 1 and 2. Recall that the sequence $\{\lambda_k\}$ converges to some $\hat{\lambda}$, and consequently, if the function λ is continuous and $\hat{x}$ is any cluster points of the sequence $\{x^k\}$, then $\hat{\lambda} = \lambda(\hat{x})$. Below we assume that λ is continuous.

Theorem 4. *Let $\hat{x}$ be any cluster point of the sequence $\{x^k\}$ and let $\hat{\lambda} = \lambda(\hat{x})$. Assume that the functions f_i, g_i and h_j are directionally differentiable at $\hat{x}$ and that $f_i'(\hat{x}; \cdot) - \hat{\lambda} g_i'(\hat{x}; \cdot)$ and $h_j'(\hat{x}; \cdot)$ for $(i,j) \in I \times J$, are convex as functions of the directions. Suppose on the other hand that the upper bounding function $\mathcal{U}$ satisfies conditions (A1)–(A3), and that $\mathcal{U}(\cdot,\cdot)$, λ and h are continuous. Then, there exists $(\hat{\alpha}, \hat{\beta}) \in \Sigma$ such that*

$$\sum_{i\in I} \hat{\alpha}_i \Big(f_i'(\hat{x}; x - \hat{x}) - \hat{\lambda} g_i'(\hat{x}; x - \hat{x})\Big) + \sum_{j\in J} \hat{\beta}_j h_j'(\hat{x}; x - \hat{x}) \geq 0 \quad \text{for all } x \in C,$$

with $\hat{\alpha}_i[f_i(\hat{x}) - \hat{\lambda} g_i(\hat{x})] = 0$ and $\hat{\beta}_j h_j(\hat{x}) = 0$ for all $(i,j) \in I \times J$.

Proof. By considering a subsequence of $\{x^k\}$ converging to $\hat{x}$, and by using the continuity of h we see from Proposition 5, Item 2 that $\{\mathcal{U}(x^{k+1}, x^k)\}$ converges to $0 = \mathcal{U}(\hat{x}, \hat{x})$. On the other hand, at iteration k, x^{k+1} is a global solution on C of $\mathcal{U}(\cdot, x^k)$, that is,

$$\mathcal{U}\big(x, x^k\big) \geq \mathcal{U}\big(x^{k+1}, x^k\big) \quad \text{for all } x \in C.$$

By using the continuity of $\mathcal{U}$ in the left hand side of the previous inequality, and Proposition 5, Item 2 we get

$$\mathcal{U}(x, \hat{x}) \geq \mathcal{U}(\hat{x}, \hat{x}) \quad \text{for all } x \in C.$$

It follows that $\hat{x}$ is a global optimal solution on C of $\mathcal{U}(\cdot, \hat{x})$. Since $\mathcal{U}(\hat{x}, \hat{x}) = 0$ it suffices to use Theorem 3 with $y = \hat{x}$ and $x_y = \hat{x}$ to conclude. ☐

6. Smooth constrained minimization GFP with Lipschitz gradients

In this section, we will detail Example 6. Recall that we assumed that the functions f_i, g_i and h_j for $(i,j) \in I \times J$, are continuously differentiable with L-Lipschitz gradients. It is easy to see that the function $f_i - \lambda(y) g_i$ is L_y-Lipschitz gradient with $L_y := L(1 + |\lambda(y)|)$, and that from the descent lemma, see *e.g.*, Proposition A.24 of [14], for any $x, z \in \mathbb{R}^n$, $y \in X$, and all $i \in I$, we have

$$f_i(x) - \lambda(y) g_i(x) \leq f_i(z) - \lambda(y) g_i(z) + \langle \nabla f_i(z) - \lambda(y) \nabla g_i(z), x - z \rangle + \tfrac{L_y}{2} \|x - z\|^2 \tag{24}$$

and since $L \leq L_y$ we also have, for all $j \in J$,

$$h_j(x) \leq h_j(y) + \langle \nabla h_j(y), x - y \rangle + \tfrac{L_y}{2} \|x - y\|^2. \tag{25}$$

In the following, we will specify Algorithm 1 for solving problems with data having Lipschitz gradients. Recall that the approximating function $\mathcal{U}$ has the form

$$\mathcal{U}(x, y) := \max_{(i,j) \in I \times J} \Big\{ f_i(y) - \lambda(y) g_i(y) + \langle \nabla f_i(y) - \lambda(y) \nabla g_i(y), x - y \rangle, \\ h_j(y) + \langle \nabla h_j(y), x - y \rangle \Big\} + \tfrac{L_y}{2} \|x - y\|^2.$$

6.1. Primal successive majorization method for GFP

Remark that the problem of finding an optimal solution of the function $\mathcal{U}(., y)$ in step 1 described in Algorithm 1 can be reformulated as a quadratic problem as follow

$$\begin{aligned} \min_{(x,t) \in C \times \mathbb{R}} \quad & t + \tfrac{L_y}{2} \|x - y\|^2 \\ \text{s.t. } & \langle \nabla f_i(y) - \lambda(y) \nabla g_i(y), x \rangle - t \leq \langle \nabla f_i(y) - \lambda(y) \nabla g_i(y), y \rangle - f_i(y) + \lambda(y) g_i(y) \quad \text{for } i \in I, \\ & \langle \nabla h_j(y), x \rangle - t \leq \langle \nabla h_j(y), y \rangle - h_j(y) \quad \text{for } j \in J. \end{aligned} \tag{$P(y)$}$$

By specifying Algorithm 1 to the class of problems fulfilling (24) and (25), we obtain the following primal successive majorization algorithm. In the case $C = \mathbb{R}^n$, it gives a sequential quadratic method for GFP.

Algorithm 2. Primal successive majorization method.

0. Choose $x^0 \in X$ and let $k = 0$.
1. Set $\lambda_k = \lambda(x^k)$, and $L_k = L(1 + |\lambda_k|)$, then find $(x^{k+1}, t_{k+1}) \in C \times \mathbb{R}$ an optimal solution to $(P(x^k))$,

$$\begin{aligned}\min_{(x,t)\in C\times\mathbb{R}} \; & t + \tfrac{L_k}{2}\|x - x^k\|^2 \\ \text{s.t. } & \langle \nabla f_i(x^k) - \lambda_k \nabla g_i(x^k), x\rangle - t \le \langle \nabla f_i(x^k) - \lambda_k \nabla g_i(x^k), x^k\rangle - f_i(x^k) + \lambda_k g_i(x^k) \quad \text{for } i \in I, \\ & \langle \nabla h_j(x^k), x\rangle - t \le \langle \nabla h_j(x^k), x^k\rangle - h_j(x^k) \quad \text{for } j \in J. \end{aligned} \qquad (P(x^k))$$

2. If $t_{k+1} + \frac{L_k}{2}\|x^{k+1} - x^k\|^2 \neq 0$, set $k = k + 1$ and return to 1.

Since the functions f_i, g_i and h_j are continuously differentiable, they are directionally differentiable and $f_i'(x; d) = \langle \nabla f_i(\bar{x}), d\rangle$, $g_i'(x; d) = \langle \nabla g_i(\bar{x}), d\rangle$ and $h_j'(x; d) = \langle \nabla h_j(\bar{x}), d\rangle$ which are, together with $f_i'(x; d) - \lambda(y) g_i'(x; d)$, convex with respect to directions. Thus, Theorems 1 and 4 apply, and by direct application of Theorem 4 we get the following result concerning the sequence $\{x^k\}$ generated by Algorithm 2.

Theorem 5. *Assume that the functions f_i, g_i and h_j for $(i, j) \in I \times J$, are continuously differentiable functions having L-Lipschitz gradients. Then, for every cluster point $\bar{x}$ of the sequence $\{x^k\}$ there exists $(\bar{\alpha}, \bar{\beta}) \in \Sigma$ such that*

$$\sum_{i\in I} \bar{\alpha}_i \langle \nabla f_i(\bar{x}) - \bar{\lambda}\nabla g_i(\bar{x}), x - \bar{x}\rangle + \sum_{j\in J} \bar{\beta}_j \langle \nabla h_j(\bar{x}), x - \bar{x}\rangle \ge 0 \quad \text{for all } x \in C \tag{26}$$

with $\bar{\alpha}_i[f_i(\bar{x}) - \bar{\lambda} g_i(\bar{x})] = 0$ and $\bar{\beta}_j h_j(\bar{x}) = 0$ for all $i \in I$ and $j \in J$.

Remark 4. Inequality (26) means that

$$\sum_{i\in I} \bar{\alpha}_i \big(\nabla f_i(\bar{x}) - \bar{\lambda}\nabla g_i(\bar{x})\big) + \sum_{j\in J} \bar{\beta}_j \nabla h_j(\bar{x}) \in -N_C(\bar{x}),$$

where $N_C(\bar{x})$ is the normal cone to C at $\bar{x}$. In the case $C = \mathbb{R}^n$, it reduces to

$$\sum_{i\in I} \bar{\alpha}_i \big(\nabla f_i(\bar{x}) - \bar{\lambda}\nabla g_i(\bar{x})\big) + \sum_{j\in J} \bar{\beta}_j \nabla h_j(\bar{x}) = 0.$$

Remark 5. With the assumptions of Theorem 2, these conditions are also sufficient for $\bar{x} \in X$ to be an optimal solution for (P), provided that the set $\{x \in X \mid \lambda(x) \le \bar{\lambda}\}$ is convex.

6.2. Primal-dual constrained gradient method for GFP

In the case where $C = \mathbb{R}^n$, we will propose in this section a dual form for Algorithm 2. We will then obtain a gradient method for constrained minimization fractional programs, in which the descent direction is obtained from the gradients of the objective and the constraint functions, by solving a simple quadratic program.

To obtain a dual problem for $(P(y))$, note that we can write $\mathcal{U}(x, y)$ as

$$\begin{aligned}\mathcal{U}(x, y) &= \max_{(\alpha,\beta)\in\Sigma} \Bigg\{ \sum_{i\in I} \alpha_i [f_i(y) - \lambda(y) g_i(y) + \langle \nabla f_i(y) - \lambda(y) g_i(y), x - y\rangle] \\ &\qquad + \sum_{j\in J} \beta_j [h_j(y) + \langle \nabla h_j(y), x - y\rangle] \Bigg\} + \tfrac{L_y}{2}\|x - y\|^2 \\ &= \max_{(\alpha,\beta)\in\Sigma} \Bigg\{ \Bigg\langle \sum_{i\in I} \alpha_i [\nabla f_i(y) - \lambda(y) \nabla g_i(y)] + \sum_{j\in J} \beta_j \nabla h_j(y), x - y \Bigg\rangle \end{aligned}$$

$$+ \sum_{i\in I} \alpha_i[f_i(y) - \lambda(y)g_i(y)] + \sum_{j\in J} \beta_j h_j(y) \Bigg\} + \tfrac{L_y}{2}\|x-y\|^2.$$

It follows that,

$$\inf_{x\in\mathbb{R}^n} \mathcal{U}(x,y) = \inf_{x\in\mathbb{R}^n} \max_{(\alpha,\beta)\in\Sigma} \Bigg\{ \Bigg\langle \sum_{i\in I} \alpha_i[\nabla f_i(y) - \lambda(y)\nabla g_i(y)] + \sum_{j\in J} \beta_j \nabla h_j(y), x-y \Bigg\rangle$$
$$+ \sum_{i\in I} \alpha_i[f_i(y) - \lambda(y)g_i(y)] + \sum_{j\in J} \beta_j h_j(y) + \tfrac{L_y}{2}\|x-y\|^2 \Bigg\}. \tag{27}$$

The function defined on $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ by

$$(x,(\alpha,\beta)) \mapsto \Bigg\{ \Bigg\langle \sum_{i\in I} \alpha_i[\nabla f_i(y) - \lambda(y)\nabla g_i(y)] + \sum_{j\in J} \beta_j \nabla h_j(y), x-y \Bigg\rangle$$
$$+ \sum_{i\in I} \alpha_i[f_i(y) - \lambda(y)g_i(y)] + \sum_{j\in J} \beta_j h_j(y) \Bigg\} + \tfrac{L_y}{2}\|x-y\|^2$$

is linear with respect to $(\alpha,\beta)\in\Sigma$, with Σ being compact, and convex with respect to x. Then, by Theorem 2 of [31] we can interchange the "inf" and "sup" in (27) to get

$$\inf_{x\in\mathbb{R}^n} \mathcal{U}(x,y) = \max_{(\alpha,\beta)\in\Sigma} \inf_{x\in\mathbb{R}^n} \Bigg\{ \Bigg\langle \sum_{i\in I} \alpha_i[\nabla f_i(y) - \lambda(y)\nabla g_i(y)] + \sum_{j\in J} \beta_j \nabla h_j(y), x-y \Bigg\rangle$$
$$+ \sum_{i\in I} \alpha_i[f_i(y) - \lambda(y)g_i(y)] + \sum_{j\in J} \beta_j h_j(y) + \tfrac{L_y}{2}\|x-y\|^2 \Bigg\}.$$

The minimum in the right side is achieved at

$$x = y - \frac{1}{L_y}\left(\sum_{i\in I} \alpha_i[\nabla f_i(y) - \lambda(y)\nabla g_i(y)] + \sum_{j\in J} \beta_j \nabla h_j(y) \right).$$

It follows that

$$\inf_{x\in\mathbb{R}^n} \mathcal{U}(x,y) = \max_{(\alpha,\beta)\in\Sigma} \Bigg\{ -\tfrac{1}{2L_y} \left\| \sum_{i\in I} \alpha_i[\nabla f_i(y) - \lambda(y)\nabla g_i(y)] + \sum_{j\in J} \beta_j \nabla h_j(y) \right\|^2$$
$$+ \sum_{i\in I} \alpha_i[f_i(y) - \lambda(y)g_i(y)] + \sum_{j\in J} \beta_j h_j(y) \Bigg\}.$$

This gives rise to the following primal-dual constrained gradient method.

Algorithm 3. Primal-dual constrained gradient method.

0. Choose $x^0 \in X$ and let $k = 0$.
1. Set $\lambda_k = \lambda(x^k)$, and $L_k = L(1 + |\lambda_k|)$, then find $(\alpha^k, \beta^k) \in \Sigma$ an optimal solution to the quadratic program

$$\theta(x^k) = \min_{(\alpha,\beta)\in\Sigma} \Big\{ \tfrac{1}{2L_k} \Big\| \sum_{i\in I} \alpha_i \Big[\nabla f_i\big(x^k\big) - \lambda_k \nabla g_i\big(x^k\big) \Big] + \sum_{j\in J} \beta_j \nabla h_j\big(x^k\big) \Big\|^2 - \sum_{i\in I} \alpha_i \Big[f_i\big(x^k\big) - \lambda_k g_i\big(x^k\big) \Big] - \sum_{j\in J} \beta_j h_j\big(x^k\big) \Big\}.$$

2. Set

$$x^{k+1} = x^k - \frac{1}{L_k} \Bigg(\sum_{i\in I} \alpha_i^k \Big[\nabla f_i\big(x^k\big) - \lambda_k \nabla g_i\big(x^k\big) \Big] + \sum_{j\in J} \beta_j^k \nabla h_j\big(x^k\big) \Bigg).$$

3. If $\theta(x^k) \neq 0$, set $k = k+1$ and return to 1.

Remark 6. Note that if $x^k \in X$ and $\theta(x^k) = 0$ then x^k satisfies the KKT optimality conditions

(1) $\sum_{i\in I} \alpha_i^k [\nabla f_i(x^k) - \lambda_k \nabla g_i(x^k)] + \sum_{j\in J} \beta_j^k \nabla h_j(x^k) = 0$,
(2) $\alpha_i^k [f_i(x^k) - \lambda_k g_i(x^k)] = 0$ for all $i \in I$, and $\beta_j^k h_j(x^k) = 0$ for all $j \in J$.

By remembering that $\theta(x^k) = \mathcal{U}(x^{k+1}, x^k)$ we get from Proposition 5 the following result.

Proposition 6. *The sequence $\{\theta(x^k)\}$ converges to 0 as k tends to ∞.*

6.3. Convex case

We treat in this section the behavior of Algorithm 3 under a convexity assumption we state below.

Assumption 4. *For all $(i,j) \in I \times J$ and $\lambda \geq \bar{\lambda}$, the functions $f_i - \lambda g_i$ and h_j are convex.*

Notice that the function $f_i - \lambda g_i$ is convex for all $\lambda \geq \bar{\lambda}$, for instance, if f_i is convex, g_i is concave and $\bar{\lambda} \geq 0$ if g_i is not affine.

Let

$$\mathcal{L}(x,y) := \max_{(i,j)\in I\times J} \{ \langle \nabla f_i(y) - \lambda(y) \nabla g_i(y), x - y \rangle + f_i(y) - \lambda(y) g_i(y), \langle \nabla h_j(y), x - y \rangle + h_j(y) \}.$$

Note that with this notation we have

$$\mathcal{U}(x,y) = \mathcal{L}(x,y) + \tfrac{L_y}{2} \|x - y\|^2.$$

We immediately have the following result.

Lemma 1. *With Assumption 4, for all $(x,y) \in \mathbb{R}^n \times X$ we have*

$$\mathcal{F}(x, \lambda(y)) + \tfrac{L_y}{2} \|x - y\|^2 \geq \mathcal{U}(x,y) \geq \mathcal{F}(x, \lambda(y)),$$

or equivalently,

$$\mathcal{F}(x, \lambda(y)) \geq \mathcal{L}(x,y) \geq \mathcal{F}(x, \lambda(y)) - \tfrac{L_y}{2} \|x - y\|^2. \tag{28}$$

Proof. The first inequality follows from the subgradient inequalities, the second follows from (24) and (25). □

In order to prove convergence of the sequence $\{x^k\}$, let us first establish the following results.

Proposition 7. *With Assumptions 2–4, the conditions of Theorem 5 are sufficient for $\bar{x}$ to be an optimal solution for (P).*

Proof. Simply note that with Assumption 4 the set $S := \{x \in C \mid \mathcal{F}(x, \bar{\lambda}) \leq 0\}$, stated in Theorem 2 is convex. □

Remark 7. Note that since h is convex, Assumption 2 is fulfilled whenever there is $x \in \mathbb{R}^n$ such that $h(x) < 0$.

Lemma 2. *Let $\bar{x}$ be any optimal solution of (P). With Assumption 4, for all $k \in \mathbb{N}$ we have*

$$\lambda(x^k) + \tfrac{L}{2\Delta}\left\|x^k - \bar{x}\right\|^2 \geq \lambda(x^{k+1}) + \tfrac{L}{2\Delta}\left\|x^{k+1} - \bar{x}\right\|^2.$$

Proof. Since x^{k+1} is the minimum of the convex $\mathcal{U}(\cdot, x^k)$ over the convex set C, we get $0 \in \partial\mathcal{U}(x^{k+1}, x^k) + N_C(x^{k+1})$, where $N_C(x^{k+1})$ is the normal cone to C at x^{k+1}. This implies that

$$L_k\left(x^k - x^{k+1}\right) \in \left[\partial\mathcal{L}\left(x^{k+1}, x^k\right) + N_C\left(x^{k+1}\right)\right]. \tag{29}$$

So, there exists $\eta_k \in N_C(x^{k+1})$ such that

$$L_k\left(x^k - x^{k+1}\right) - \eta_k \in \partial\mathcal{L}\left(x^{k+1}, x^k\right). \tag{30}$$

Therefore, there exists $(\alpha^k, \beta^k) \in \Sigma$ such that

$$L_k\left(x^k - x^{k+1}\right) - \eta_k = \sum_{i\in I}\alpha_i^k\left[\nabla f_i\left(x^k\right) - \lambda_k\nabla g_i\left(x^k\right)\right] + \sum_{j\in J}\beta_j^k\nabla h_j\left(x^k\right) \tag{31}$$

and

$$\left\langle\sum_{i\in I}\alpha_i^k\left[\nabla f_i\left(x^k\right) - \lambda_k\nabla g_i\left(x^k\right)\right] + \sum_{j\in J}\beta_j^k\nabla h_j\left(x^k\right), x^{k+1} - x^k\right\rangle + \sum_{j\in J}\beta_j^k h_j\left(x^k\right) = \mathcal{L}\left(x^{k+1}, x^k\right).$$

From (30) and the subgradient inequality we have for all $x \in \mathbb{R}^n$,

$$\begin{aligned}\mathcal{L}\left(x, x^k\right) &\geq \mathcal{L}\left(x^{k+1}, x^k\right) + \left\langle L_k\left(x^k - x^{k+1}\right) - \eta_k, x - x^{k+1}\right\rangle \\ &= \mathcal{L}\left(x^{k+1}, x^k\right) + L_k\left\langle x^k - x^{k+1}, x - x^{k+1}\right\rangle - \left\langle\eta_k, x - x^{k+1}\right\rangle.\end{aligned}$$

But for all $x \in C$ we have $\langle\eta_k, x - x^{k+1}\rangle \leq 0$ since $\eta_k \in N_C(x^{k+1})$. It follows that

$$\mathcal{L}\left(x, x^k\right) \geq \mathcal{L}\left(x^{k+1}, x^k\right) + L_k\left\langle x^k - x^{k+1}, x - x^{k+1}\right\rangle \quad \text{for all } x \in C.$$

By writing

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

we get, for all $x \in C$,

$$\mathcal{L}\left(x, x^k\right) \geq \mathcal{L}\left(x^{k+1}, x^k\right) + \tfrac{L_k}{2}\left[\left\|x^{k+1} - x\right\|^2 - \left\|x^k - x\right\|^2 + \left\|x^{k+1} - x^k\right\|^2\right].$$

By using the second inequality in (28) we obtain

$$\begin{aligned}\mathcal{L}\left(x, x^k\right) &\geq \mathcal{F}\left(x^{k+1}, \lambda_k\right) - \tfrac{L_k}{2}\left\|x^{k+1} - x^k\right\|^2 + \tfrac{L_k}{2}\left[\left\|x^{k+1} - x\right\|^2 - \left\|x^k - x\right\|^2 + \left\|x^{k+1} - x^k\right\|^2\right] \\ &= \mathcal{F}\left(x^{k+1}, \lambda_k\right) + \tfrac{L_k}{2}\left[\left\|x^{k+1} - x\right\|^2 - \left\|x^k - x\right\|^2\right] \quad \text{for all } x \in C.\end{aligned}$$

By considering the first inequality in (28) we obtain from the last inequality,

$$\mathcal{F}(x, \lambda_k) \geq \mathcal{F}\big(x^{k+1}, \lambda_k\big) + \tfrac{L_k}{2}\Big[\big\|x^{k+1} - x\big\|^2 - \big\|x^k - x\big\|^2\Big] \quad \text{for all } x \in C. \tag{32}$$

So, with $x = \bar{x}$, an optimal solution of (P), we have $\mathcal{F}(\bar{x}, \lambda_k) \leq 0$, and the previous inequality gives

$$0 \geq \mathcal{F}\big(x^{k+1}, \lambda_k\big) + \tfrac{L_k}{2}\Big[\big\|x^{k+1} - \bar{x}\big\|^2 - \big\|x^k - \bar{x}\big\|^2\Big].$$

Taking into account that $\mathcal{F}(x^{k+1}, \lambda_k) \geq f_i(x^{k+1}) - \lambda_k g_i(x^{k+1})$ for all $i \in I$, we get

$$0 \geq f_i\big(x^{k+1}\big) - \lambda_k g_i\big(x^{k+1}\big) + \tfrac{L_k}{2}\Big[\big\|x^{k+1} - \bar{x}\big\|^2 - \big\|x^k - \bar{x}\big\|^2\Big] \quad \text{for all } i \in I.$$

It follows that for all $i \in I$,

$$0 \geq g_i\big(x^{k+1}\big)\big(f_i\big(x^{k+1}\big)/g_i\big(x^{k+1}\big) - \lambda_k\big) + \tfrac{L_k}{2}\Big[\big\|x^{k+1} - \bar{x}\big\|^2 - \big\|x^k - \bar{x}\big\|^2\Big].$$

For each $i \in I$ such that $f_i(x^{k+1})/g_i(x^{k+1}) = \lambda_{k+1}$ we get

$$0 \geq g_i\big(x^{k+1}\big)(\lambda_{k+1} - \lambda_k) + \tfrac{L_k}{2}\Big[\big\|x^{k+1} - \bar{x}\big\|^2 - \big\|x^k - \bar{x}\big\|^2\Big],$$

which implies that

$$0 \geq \tfrac{g_i\left(x^{k+1}\right)}{L_k}(\lambda_{k+1} - \lambda_k) + \tfrac{1}{2}\Big[\big\|x^{k+1} - \bar{x}\big\|^2 - \big\|x^k - \bar{x}\big\|^2\Big].$$

Since $0 < L \leq L_k$ and $0 < g_i(x^{k+1}) \leq \Delta$, the fact that $\lambda_{k+1} - \lambda_k \leq 0$ entails that

$$0 \geq \tfrac{\Delta}{L}(\lambda_{k+1} - \lambda_k) + \tfrac{1}{2}\Big[\big\|x^{k+1} - \bar{x}\big\|^2 - \big\|x^k - \bar{x}\big\|^2\Big].$$

Thus,

$$\lambda_k + \tfrac{L}{2\Delta}\big\|x^k - \bar{x}\big\|^2 \geq \lambda_{k+1} + \tfrac{L}{2\Delta}\big\|x^{k+1} - \bar{x}\big\|^2,$$

which is the desired result. □

Theorem 6. *Assume that Assumption 4 and the Mangasarian–Fromovitz condition are fulfilled. Then the sequence $\{x^k\}$ converges to an optimal solution of (P).*

Proof. Lemma 2 tell us that the sequence $\{\|x^k - \bar{x}\|\}$ is convergent, and hence the sequence $\{x^k\}$ is bounded. From Proposition 7 every point $\hat{x}$ of its cluster points is a solution for (P). Reconsidering Lemma 2 with $\hat{x}$ we conclude that $\{\|x^k - \hat{x}\|\}$ converges to zero. Thus, the sequence $\{x^k\}$ converges to $\hat{x}$ which is an optimal solution to (P). □

7. Numerical tests

In this section we test the algorithms presented in this paper on some numerical examples. Recall that our problem has the following form

$$\begin{aligned} \bar{\lambda} = \inf_{x \in C}\ \lambda(x) &:= \max_{i \in I} \frac{f_i(x)}{g_i(x)} \\ \text{s.t. } h_j(x) &\leq 0, \quad j \in J := \{1, \cdots, p\}. \end{aligned}$$

We focus in our tests on Algorithms 2 and 3 described in Section 6, and compare them with the DC method of centers of [21], and the DC-Dinkelbach algorithm considered in [33]. We refer to these algorithms, respectively, by PCGM, DCGM, RCVX and DCC.

We consider a case with $C = \mathbb{R}^n$ and f_i, g_i and h_j are the convex functions considered in [8]. This problem is a particular case of fractional programs with ratios of DC functions. We recall that the data considered in [8] have the following forms:

TABLE 1. Comparison of Algorithms PCGM, DCGM, RCVX and DCC on 10 problems.

Problem	Algorithm	# Iter	T (s)	λ_∞	h_∞
1	PCGM	406	2.49	0.8658	−1.5498e−06
	DCGM	406	1.15	0.8658	−1.5505e−06
	RCVX	31	7.10	0.8654	1.3211e−10
	DCC	32	0.94	0.8654	−8.0562e−07
2	PCGM	388	2.26	0.8152	−1.4446e−06
	DCGM	388	1.11	0.8152	−1.5032e−06
	RCVX	30	7.83	0.8148	3.8748e−07
	DCC	29	0.74	0.8148	−1.0271e−06
3	PCGM	323	2.65	0.7954	−1.5461e−06
	DCGM	323	1.20	0.7954	−1.5429e−06
	RCVX	25	6.86	0.7951	3.9898e−07
	DCC	25	0.72	0.7951	−8.0123e−07
4	PCGM	357	2.01	0.8384	−1.5538e−06
	DCGM	357	1.03	0.8384	−1.5616e−06
	RCVX	28	6.91	0.8381	8.5648e−07
	DCC	29	0.88	0.8381	−2.3104e−07
5	PCGM	389	3.19	0.8614	−1.5299e−06
	DCGM	389	1.56	0.8614	−1.5315e−06
	RCVX	31	7.56	0.8611	7.8052e−07
	DCC	31	0.97	0.8611	−7.7846e−07
6	PCGM	263	1.70	0.7499	−1.5207e−06
	DCGM	263	0.76	0.7499	−1.5608e−06
	RCVX	21	5.39	0.7497	4.0421e−07
	DCC	21	0.67	0.7497	−9.5083e−07
7	PCGM	342	1.20	0.8161	−1.5308e−06
	DCGM	342	0.97	0.8161	−1.5310e−06
	RCVX	26	5.87	0.8158	3.7858e−10
	DCC	27	0.73	0.8158	−8.2695e−07
8	PCGM	428	2.41	0.8541	−1.5687e−06
	DCGM	427	1.21	0.8541	−1.6310e−06
	RCVX	33	9.95	0.8535	9.0192e−07
	DCC	33	1.06	0.8535	−5.3863e−07
9	PCGM	481	3.33	0.8573	−1.4908e−06
	DCGM	480	1.56	0.8573	−1.5099e−06
	RCVX	36	8.50	0.8568	8.9968e−07
	DCC	36	1.19	0.8568	−8.8252e−07
10	PCGM	339	1.99	0.8662	−1.5540e−06
	DCGM	340	1.32	0.8662	−1.5411e−06
	RCVX	27	6.74	0.8658	9.8625e−07
	DCC	27	1.09	0.8658	−7.6033e−07
Averages		Av.Iter	Av.T(s)		
	PCGM	371.6	2.40		
	DCGM	371.5	1.19		
	RCVX	28.8	7.27		
	DCC	29.0	0.89		

- $f_i(x) = \|x - b_i\|^2$ and $g_i(x) = \|x - c_i\|^2$ with $b_i = \frac{v_i}{2\|v_i\|}$, where v_i is a $n \times 1$ random vector with components in $[0, 1]$. We construct c_i in the same way as b_i. The gradient Lipschitz constants for f_i and g_i are equal to 2, for all $i \in I$.
- $h_j(x) = x^\top Q_j x + 2q_j^\top x + a_j$, for all $j \in J$. By some transformations detailed in [8], these functions take the following form

$$h_j(x) = \left\|R_j x + R_j^{-1} w_j\right\|^2 - d_j^2,$$

  where R_j represents the Cholesky decomposition of Q_j. The entries of the problems, as generated in [8], are $Q_j = Y_j D_j Y_j$ where Y_j are random Householder orthogonal matrices given by $Y_j = I - 2\frac{\omega_j \omega_j^\top}{\|\omega_j\|^2}$ with the components of the vectors ω_j randomly chosen from the interval $(-1, 1)$. The elements of the diagonal matrices D_j are drawn randomly from the set $\{10^{\frac{j-1}{n-1}} \mid j = 1, \cdots, n\}$. The Lipschitz constants of the gradients of the functions h_j are equal to 10.

We tested our algorithms on this problem with $n = 50$, $m = 20$ and $p = 30$, and we used the stopping criterion $\mid \mathcal{U}(x^{k+1}, x^k) \mid \leq 1.e - 6$. In our tests, we solved 10 problems and reported the results in Table 1. In this table, # Iter represents the number of iteration, T(s) is the execution time in second, λ_∞ and h_∞ are the images of the obtained solution by the objective and constraint functions. On the other hand, Av.Iter and Av.T(s) represent, respectively, the average of the number of iterations and the total execution time for solving these 10 problems.

8. Conclusion

8.1. Summary of main results

In this paper, we developed several results for GFP, where the objective and constraint functions defining the optimization problem are directionally differentiable with convex directional derivatives, as functions of directions. We gave optimality conditions expressed in terms of such directional derivatives, and fashioned a generic and unifying algorithm based on the notion of successive upper approximations methods. We have listed several examples for which we can effectively construct upper approximation functions. Then, a particular attention has been made for problems with functions that are continuously differentiable with Lipschitz gradients, for which we developed two versions of algorithms: a successive quadratic approximation, and a constrained gradient type method. Finally, we tested these algorithms and compared their efficiency with that of other algorithms.

8.2. Comments on numerical tests

By analysing the results reported in Table 1, we see that:

- all algorithms find the same optimal value;
- algorithm DCGM wins PCGM in term of total execution time, but the two algorithms converge with the same number of iterations;
- the DCC algorithm outperform RCVX in terms of overall execution time, while they require the same number of iterations to converge.

We can conclude from these observations that the DCC algorithm wins, followed by DCGM algorithm, then PCGM, and finally RCVX. This seems to be due to the nature of the subproblems solved in each algorithm.

Acknowledgements

The authors are grateful to the two anonymous referees for their corrections and helpful comments.

References

[1] A. Addou and A. Roubi, Proximal-type methods with generalized Bregman functions and applications to generalized fractional programming. *Optimization* **59** (2010) 1085–1105.

[2] S. Addoune, M. El Haffari and A. Roubi, A proximal point algorithm for generalized fractional programs. *Optimization* **66** (2017) 1495–1517.

[3] S. Addoune, K. Boufi and A. Roubi, Proximal bundle algorithms for nonlinearly constrained convex minimax fractional programs. *J. Optim. Theory Appl.* **179** (2018) 212–239.

[4] A. Aubry, A. De Maio and M.M. Naghsh, Optimizing radar waveform and doppler filter bank via generalized fractional programming. *IEEE J. Sel. Topics Signal Process* **9** (2015) 1387–1399.

[5] A. Aubry, V. Carotenuto and A. De Maio, New results on generalized fractional programming problems with Toeplitz quadratics. *IEEE Signal Process. Lett.* **23** (2016) 848–852.

[6] A. Aubry, A. De Maio, Y. Huang and M. Piezzo, Robust design of radar doppler filters. *IEEE Trans. Signal Process.* **64** (2016) 5848–5860.

[7] A. Aubry, A. De Maio, A. Zappone, M. Razaviyayn and Z.-Q. Luo, A new sequential optimization procedure and its applications to resource allocation for wireless systems. *IEEE Trans. Signal Process.* **66** (2018) 6518–6533.

[8] A. Auslender, R. Shefi and M. Teboulle, A moving balls approximation method for a class of smooth constrained minimization problems. *SIAM J. Optim.* **20** (2010) 3232–3259.

[9] A.I. Barros, J.B.G. Frenk, S. Schaible and S. Zhang, A new algorithm for generalized fractional programs. *Math. Program.* **72** (1996) 147–175.

[10] A.I. Barros, J.B.G. Frenk, S. Schaible and S. Zhang, Using duality to solve generalized fractional programming problems. *J. Glob. Optim.* **8** (1996) 139–170.

[11] A. Beck, A. Ben-Tal and L. Tetruashvili, A sequential parametric convex approximation method with applications to non-convex truss topology design problems. *J. Glob. Optim.* **47** (2010) 29–51.

[12] C.R. Bector, S. Chandra and M.K. Bector, Generalized fractional programming duality: a parametric approach. *J. Optim. Theory Appl.* **60** (1989) 243–260.

[13] J.C. Bernard and J.A. Ferland, Convergence of interval-type algorithms for generalized fractional programming. *Math. Program.* **43** (1989) 349–363.

[14] D.P. Bertsekas, Nonlinear Programming, 2nd edition. Athena Scientific, Belmont, MA (1999).

[15] H. Boualam and A. Roubi, Dual algorithms based on the proximal bundle method for solving convex minimax fractional programs. *J. Ind. Manag. Optim.* **15** (2019) 1897–1920.

[16] H. Boualam and A. Roubi, Proximal bundle methods based on approximate subgradients for solving Lagrangian duals of minimax fractional programs. *J. Global Optim.* **74** (2019) 255–284.

[17] H. Boualam and A. Roubi, Augmented Lagrangian dual for nonconvex minimax fractional programs and proximal bundle algorithms for its resolution. *J. Ind. Manag. Optim.* **19** (2023) 3610–3636.

[18] K. Boufi and A. Roubi, Dual method of centers for solving generalized fractional programs. *J. Glob. Optim.* **69** (2017) 387–426.

[19] K. Boufi and A. Roubi, Duality results and dual bundle methods based on the dual method of centers for minimax fractional programs. *SIAM J. Optim.* **29** (2019) 1578–1602.

[20] K. Boufi and A. Roubi, Prox-regularization of the dual method of centers for generalized fractional programs. *Optim. Methods Softw.* **34** (2019) 515–545.

[21] K. Boufi, M. El Haffari and A. Roubi, Optimality conditions and a method of centers for minimax fractional programs with difference of convex functions. *J. Optim. Theory App.* **187** (2020) 105–132.

[22] L.M. Bregman, The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming. *U.S.S.R. Comput. Math. Math. Phys.* **7** (1967) 200–217.

[23] J.-P. Crouzeix and J.A. Ferland, Algorithms for generalized fractional programming. *Math. Program.* **52** (1991) 191–207.

[24] J.-P. Crouzeix, J.A. Ferland and S. Schaible, Duality in generalized linear fractional programming. *Math. Program.* **27** (1983) 342–354.

[25] J.-P. Crouzeix, J.A. Ferland and S. Schaible, An algorithm for generalized fractional programs. *J. Optim. Theory App.* **47** (1985) 35–49.

[26] J.-P. Crouzeix, J.A. Ferland and S. Schaible, A note on an algorithm for generalized fractional programs. *J. Optim. Theory Appl.* **50** (1986) 183–187.

[27] J.-P. Crouzeix, J.A. Ferland and H.V. Nguyen, Revisiting Dinkelbach-type algorithms for generalized fractional programs. *Opsearch* **45** (2008) 97–110.

[28] W. Dinkelbach, On nonlinear fractional programming. *Manage. Sci.* **13** (1967) 492–498.

[29] M. El Haffari and A. Roubi, Convergence of a proximal algorithm for solving the dual of a generalized fractional program. *RAIRO-Oper. Res.* **51** (2017) 985–1004.

[30] M. El Haffari and A. Roubi, Prox-dual regularization algorithm for generalized fractional programs. *J. Ind. Manag. Optim.* **13** (2017) 1991–2013.

[31] K. Fan, Minimax theorems. *Proc. Natl. Acad. Sci. USA* **39** (1953) 42–47.

[32] J.B.G. Frenk and S. Schaible, Fractional programming. ERIM Report Series, Reference No. ERS-2004-074-LIS (2004).

[33] A. Ghazi and A. Roubi, A DC approach for minimax fractional optimization programs with ratios of convex functions. *Optim. Methods Softw.* **37** (2022) 639–657.

[34] A. Ghazi and A. Roubi, Optimality conditions and DC-Dinkelbach-type algorithm for generalized fractional programs with ratios of difference of convex functions. *Optim. Lett.* **15** (2021) 2351–2375.

[35] M. Gugat, Prox-regularization methods for generalized fractional programming. *J. Optim. Theory Appl.* **99** (1998) 691–722.

[36] R. Jagannathan and S. Schaible, Duality in generalized fractional programming via Farkas' lemma. *J. Optim. Theory App.* **41** (1983) 417–424.

[37] A. Jayswal, I.M. Stancu-Minasian and A.M. Stancu, Multiobjective fractional programming problems involving semilocally type-I univex functions. *Southeast Asian Bull. Math.* **38** (2014) 225–241.

[38] J.-Y. Lin, H.-J. Chen and R.-L. Sheu, Augmented Lagrange primal-dual approach for generalized fractional programming problems. *J. Ind. Manag. Optim.* **4** (2013) 723–741.

[39] O.L. Mangasarian and S. Fromovitz, The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *J. Math. Anal. Appl.* **17** (1967) 37–47.

[40] B.R. Marks and G.P. Wright, A general inner approximation algorithm for nonconvex mathematical programs. *Oper. Res.* **26** (1978) 681–683.

[41] A. Nagih and G. Plateau, Problèmes fractionnaires: tour d'horizon sur les applications et méthodes de résolution. *RAIRO-Oper. Res.* **33** (1999) 383–419.

[42] B.N. Pshenichnyi, Necessary Conditions for an Extremum. Marcel Dekker Inc., New York (1971).

[43] M. Razaviyayn, M. Hong and Z.-Q. Luo, A unified convergence analysis of block successive minimization methods for nonsmooth optimization. *SIAM J. Optim.* **23** (2013) 1126–1153.

[44] A. Roubi, Method of centers for generalized fractional programming. *J. Optim. Theory Appl.* **107** (2000) 123–143.

[45] A. Roubi, Convergence of prox-regularization methods for generalized fractional programming. *RAIRO Oper. Res.* **36** (2002) 73–94.

[46] S. Schaible, Bibliography in fractional programming. *Z. Oper. Res.* **26** (1982) 211–241.

[47] S. Schaible, Fractional programming, in Handbook Global Optimization, edited by R. Horst and P.M. Pardalos. Kluwer, Dordrecht (1995) 495–608.

[48] A.M. Stancu, Mathematical Programming with Type-I Functions. Matrix Rom, Bucharest (2013).

[49] I.M. Stancu-Minasian, Fractional Programming. Theory, Methods and Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands (1997).

[50] I.M. Stancu-Minasian, A sixth bibliography of fractional programming. *Optimization* **55** (2006) 405–428.

[51] I.M. Stancu-Minasian, A seventh bibliography of fractional programming. *Adv. Model. Optim.* **15** (2013) 309–386.

[52] I.M. Stancu-Minasian, An eighth bibliography of fractional programming. *Optimization* **66** (2017) 439–470.

[53] I.M. Stancu-Minasian, A ninth bibliography of fractional programming. *Optimization* **68** (2019) 2125–2169.

[54] I.M. Stancu-Minasian and N. Teodorescu, Programarea fracţionară cu mai multe funcţii-obiectiv. Matrix Rom, Bucureşti (2011).

[55] J.-J. Strodiot, J.-P. Crouzeix, J.A. Ferland and V.H. Nguyen, An inexact proximal point method for solving generalized fractional programs. *J. Glob. Optim.* **42** (2008) 121–138.