RAIRO-Oper. Res. 51 (2017) 985–1004
DOI: 10.1051/ro/2017004

RAIRO Operations Research
www.rairo-ro.org

CONVERGENCE OF A PROXIMAL ALGORITHM FOR SOLVING THE DUAL OF A GENERALIZED FRACTIONAL PROGRAM*

Mostafa El Haffari[1] and Ahmed Roubi[1]

Abstract. We propose to use the proximal point algorithm to regularize a "dual" problem of generalized fractional programs (GFP). The proposed technique leads to a new dual algorithm that generates a sequence which converges from below to the minimal value of the considered problem. At each step, the proposed algorithm solves approximately an auxiliary problem with a unique dual solution whose every cluster point gives a solution to the dual problem. In the exact minimization case, the sequence of dual solutions converges to an optimal dual solution. For a class of functions, including the linear case, the convergence of the dual values is at least linear.

Mathematics Subject Classification. 90C30, 90C32, 49K35, 49M29, 49M37.

Received December 9, 2015. Accepted December 22, 2016.

1. Introduction

In this paper, we will be interested to generalized fractional programs of the form

$$(P) \qquad \lambda_* = \inf_{x \in X} \max_{i \in I} \left\{ \frac{f_i(x)}{g_i(x)} \right\}$$

where $I = \{1, \ldots, m\}$, $m \geq 1$, and X a non empty subset of $\mathbb{R}^n$. The functions f_i and g_i are defined on an open subset K containing X, continuous and satisfy $g_i(x) > 0$ for all $x \in X$ and $i \in I$.

In the literature, several algorithms were considered for solving generalized fractional programs ([1–4, 7–10, 13, 16, 18, 19, 21]).

The Dinkelbach-type algorithms proposed in [9, 10] generalize Dinkelbach algorithm [11] to the multi-ratios case. In these algorithms, the problem is reduced to a sequence of auxiliary problems.

Later, it was proposed in [13] a method based on the proximal point algorithm to surmount the difficulties that can occur when the feasible set is unbounded or when the fractional program does not have a unique solution.

In the same way, but by using the concept of bundle methods (see [6, 15] for example), the authors proposed in [21] new algorithms that use the bundle methods for solving generalized fractional programs. In this approach, the auxiliary problems appearing in Dinkelbach-type algorithms are replaced by quadratic programs.

Keywords. Multi-ratio fractional programs, Dinkelbach-type algorithms, Lagrange duality, proximal point algorithm.

* *The authors thanks the two anonymous referees.*

[1] Faculté des Sciences et Techniques, Settat, Morocco. roubia@hotmail.com

Dual algorithms for solving GFP were introduced in [2,3], and are based on duality approach.

The purpose of this paper is to use the proximal point algorithm to regularize a "dual" problem for generalized fractional programs (GFP) proposed in [2]. At each step, the proposed algorithm solves approximately an auxiliary problem with a unique pair of primal-dual solutions whose every cluster point gives a solution to the primal and dual problems. The proposed technique leads to a new primal-dual algorithm that generates a sequence which converges from below to the minimal value of the considered problem. For a class of problems, including linear fractional programs, we establish that this algorithm converges at least linearly.

2. Dinkelbach-type algorithms and the proximal regularization methods

Before introducing and analyzing the dual approaches, we first recall the primal methods based on Dinkelbach procedure ([11]) and their prox-regularized versions.

Dinkelbach-type algorithm introduced in [7,9,10] starts from an initial point $x_0 \in X$ and generates a sequence $\{\lambda_k\}$ by solving the parametric auxiliary problems

$$\inf_{x\in X} \max_{i\in I} \{f_i(x) - \lambda_k g_i(x)\} \quad \text{where} \quad \lambda_k = \max_{i\in I} \frac{f_i(x_k)}{g_i(x_k)}\cdot \tag{2.1}$$

The sequence $\{\lambda_k\}$ is decreasing and converges at least linearly to the minimal value of (P).

To improve the speed of convergence, an other version of the last method was introduced in [10]. The speed of convergence may be superlinear without significant additional computational efforts. This version is based on the same principle but solves at each iteration the following auxiliary problem

$$\inf_{x\in X} \max_{i\in I} \left\{\frac{f_i(x) - \lambda_k g_i(x)}{g_i(x_k)}\right\} \quad \text{where} \quad \lambda_k = \max_{i\in I} \frac{f_i(x_k)}{g_i(x_k)}\cdot \tag{2.2}$$

The idea of the proximal point algorithm (see for example [14]) was used in [13] to regularize the auxiliary parametric programs generated by each one of the previous Dinkelbach-type methods. This technique is useful when the parametric problems in Dinkelbach-type algorithms does not have unique solution or if the feasible set is unbounded. In the prox-regularization algorithms proposed in [1,13,19] the generated subproblems have unique solutions.

To describe the approach given in [13], we define for $\alpha > 0$, and for all $(\lambda, x, z) \in \mathbb{R} \times X \times X$ the functions

$$J(x,\lambda) = \inf_{y\in X} \max_{i\in I} \{f_i(y) - \lambda g_i(y) + \alpha\|x-y\|^2\},$$

$$J_z(x, \lambda) = \inf_{y\in X} \max_{i\in I} \left\{\frac{f_i(y) - \lambda g_i(y)}{g_i(z)} + \alpha\|x-y\|^2\right\},$$

where $\|\ .\ \|$ stands for the euclidean norm on $\mathbb{R}^n$.

For a given sequence $\{\eta_k\}$ of nonnegative numbers such that $\sum_{k=0}^{\infty} \sqrt{\eta_k} < \infty$, the prox-regularization algorithm computes at iteration k a point $x_{k+1} \in X$ such that

$$\max_{i\in I} \{f_i(x_{k+1}) - \lambda_k g_i(x_{k+1})\} + \alpha\|x_{k+1} - x_k\|^2 \le \min\{0, J(x_k, \lambda_k) + \eta_k\},$$

where

$$\lambda_k = \max_{i\in I} \frac{f_i(x_k)}{g_i(x_k)}\cdot$$

The sequence $\{\lambda_k\}$ decreasingly converges to the minimal value of (P).

As for the scaled version derived from the Dikelbach-type algorithm, a modified version of the previous algorithm was also given in [13]. In this method one has to find $x_{k+1} \in X$ such that

$$\max_{i\in I} \left\{\frac{f_i(x_{k+1}) - \lambda_k g_i(x_{k+1})}{g_i(x_k)}\right\} + \alpha\|x_{k+1} - x_k\|^2 \le \min\{0, J_{x_k}(x_k, \lambda_k) + \eta_k\},$$

where

$$\lambda_k = \max_{i\in I} \frac{f_i(x_k)}{g_i(x_k)}\cdot$$

With this minor modification, the resulting method appears to be more efficient than the previous one.

3. Dual algorithms

The algorithms we have presented above are primal ones. In [2] Barros et *al.* proposed a partial dual formulation of (P) and then developed algorithms based on this description.

Next we describe these methods. For this, let

$$f(x) = \big(f_1(x), \dots, f_m(x)\big)^\top \quad \text{and} \quad g(x) = \big(g_1(x), \dots, g_m(x)\big)^\top.$$

From the equality

$$\max_{1\le i\le m} \left\{\frac{f_i(x)}{g_i(x)}\right\} = \max_{y\in\Sigma} \left\{\frac{y^\top f(x)}{y^\top g(x)}\right\}$$

where

$$\Sigma = \left\{ y\in\mathbb{R}^m \;\middle|\; \sum_{i=1}^m y_i = 1,\, y_i \ge 0,\, i=1,\dots,m \right\},$$

the authors consider in [2] the following problem

$$(\mathcal{R}) \qquad \max_{y\in\Sigma} c(y)$$

where the function $c\ :\ \Sigma \longrightarrow \mathbb{R}$ is given by

$$c(y) = \min_{x\in X} \left\{\frac{y^\top f(x)}{y^\top g(x)}\right\}\cdot$$

For all $(\lambda, y) \in \mathbb{R}\times\mathbb{R}^m$, define the function

$$\phi(\lambda, y) = \min_{x\in X}\{y^\top(f(x) - \lambda g(x))\}.$$

The dual algorithm solves at each iteration a parametric problem of the form

$$\max_{y\in\Sigma} \phi(\lambda, y). \tag{3.1}$$

Below we summarize this algorithm.

Algorithm 3.1.

(1) Take $y_0\in\Sigma$, compute $\lambda_0 = c(y_0)$ and let $k=0$.
(2) Determine

$$y_{k+1} \in \operatorname*{argmax}_{y\in\Sigma}\ \phi(\lambda_k, y).$$

(3) If $\phi(\lambda_k, y_{k+1}) = 0$, then y_{k+1} is an optimal solution of $(\mathcal{R})$ and λ_k is the optimal value, and stop.
(4) Compute $\lambda_{k+1} = c(y_{k+1})$, let $k = k+1$ and go to 2.

The scaled version of Algorithm 3.1 follows the same strategy used to derive the procedure (3.1). Before presenting this variant we introduce for $x_k \in X$ the vector-valued functions f^k, g^k given by

$$f_i^k(x) := \frac{f_i(x)}{g_i(x_k)} \text{ and } g_i^k(x) := \frac{g_i(x)}{g_i(x_k)}.$$

We can now define the optimization problem

$$(\mathcal{R}^k) \qquad \max_{y\in\Sigma} c^k(y)$$

where the function $c^k : \Sigma \longrightarrow \mathbb{R}$ is given by

$$c^k(y) = \min_{x\in X} \frac{y^\top f^k(x)}{y^\top g^k(x)}.$$

For all $(\lambda, y) \in \mathbb{R} \times \mathbb{R}^m$, define the function

$$\phi^k(\lambda, y) = \min_{x\in X} \left\{ y^\top \left(f^k(x) - \lambda g^k(x) \right) \right\}.$$

The scaled version of Algorithm 3.1 is described by the following procedure.

Algorithm 3.2.

(1) Take $y_0 \in \Sigma$ and $x^0 \in X$. Compute $\lambda_0 = c^0(y_0)$ and let $k = 0$.
(2) Determine $(x_{k+1}, y_{k+1}) \in X \times \Sigma$ a solution of

$$\max_{y\in\Sigma} \min_{x\in X} \left\{ y^\top \left(f^k(x) - \lambda_k g^k(x) \right) \right\}.$$

(3) If $\phi^k(\lambda_k, y_{k+1}) = 0$, then y_{k+1} is an optimal solution of $(\mathcal{R}^k)$ and λ_k is the optimal value, and stop.
(4) Compute $\lambda_{k+1} = c^{k+1}(y_{k+1})$, let $k = k + 1$ and go to 2.

The convergence of these algorithms is established under some convexity assumptions (see the next section) and the compactness of the constraints set X.

4. The regularization of the dual problem

In this section we propose to prox-regularize the "dual" problem $(\mathcal{R})$ of (P) with the same procedure used in [19] to regularize (P). Observe that $(\mathcal{R})$ is a continuous fractional program and so the technique used in [19] do not directly apply.

In all what follows we will assume that the feasible set X is nonempty and the functions $f_i, g_i : K \longrightarrow \mathbb{R}$, $i \in I$, are continuous and satisfy one of the following assumptions:

(C1) For every $i \in I$, the function $f_i : K \longrightarrow \mathbb{R}$ is convex on X and nonnegative on X and the function g_i: $K \longrightarrow \mathbb{R}$ is positive and concave on X.
(C2) For every $i \in I$, the function $f_i : K \longrightarrow \mathbb{R}$ is convex on X and the function g_i: $K \longrightarrow \mathbb{R}$ is positive and affine on X.

Hereafter, we consider the problem (P) with the notations of Section 3.

Let $\alpha > 0$. For $\lambda \in \mathbb{R}$ and $y \in \Sigma$, we associate to the parametric problem (3.1), the regularized problem

$$(\mathcal{R}(\lambda, y, \alpha)) \qquad \max_{z\in\Sigma} \{\phi(\lambda, z) - \alpha\|z - y\|^2\}.$$

The method we propose replaces in Algorithm 3.1, the step (2) by solving the regularized auxiliary problem $(\mathcal{R}(\lambda_k, y_k, \alpha_k))$. Next, we describe in detail our algorithm.

Algorithm 4.1. Let $\{\alpha_k\}$ and $\{\eta_k\}$ be two chosen or constructed sequences of nonnegative numbers such that $\sum_{k\geq 0}\sqrt{\eta_k/\alpha_k} < \infty$ and $\sum_{k\geq 0}\eta_k < \infty$.

(1) Choose a point $y_0 \in \Sigma$, calculate $c_0 = c(y_0)$ and let $k = 0$.
(2) Find $y_{k+1} \in \Sigma$ such that

$$\phi(c_k, y_{k+1}) - \alpha_k\|y_{k+1} - y_k\|^2 \geq \max_{y\in\Sigma}\{\phi(c_k, y) - \alpha_k\|y - y_k\|^2\} - \eta_k.$$

(3) Calculate $c_{k+1} = c(y_{k+1})$, set $k = k+1$ and go to 2.

Remark 4.2.

(1) In order to make sense to our algorithm, we will assume in all what follows that the following problems

$$c(y) = \inf_{x\in X}\frac{y^\top f(x)}{y^\top g(x)} \quad \text{and} \quad \phi(\lambda, y) = \inf_{x\in X}\{y^\top(f(x) - \lambda g(x))\}$$

have solutions for all $\lambda \in \mathbb{R}$ and $y \in \Sigma$.
(2) Notice that if the hypothesis (C1) is satisfied, then the positivity assumption implies that $c(y) \geq 0$ for all $y \in \Sigma$.

Notice that the important step in Algorithm 4.1 is to solve the minimax problem

$$(\mathcal{R}(c_k, y_k, \alpha_k)) \qquad \max_{y\in\Sigma}\big\{\phi(c_k, y) - \alpha_k\|y - y_k\|^2\big\}.$$

Before analyzing the convergence and the rate of convergence of the algorithm, we will give an equivalent simpler formulation for $(\mathcal{R}(\lambda, y, \alpha))$.

Let

$$\psi(\lambda, y, \alpha) = \max_{z\in\Sigma}\big\{\phi(\lambda, z) - \alpha\|z - y\|^2\big\}.$$

First of all, recall that

$$\phi(\lambda, z) = \min_{x\in X}\big\{z^\top\big(f(x) - \lambda g(x)\big)\big\},$$

and remark that

$$z^\top\big(f(x) - \lambda g(x)\big) - \alpha\|z - y\|^2 = z^\top\big(f(x) - \lambda g(x) + 2\alpha y\big) - \alpha\|z\|^2 - \alpha\|y\|^2.$$

Now, let

$$L^y_{\alpha,\lambda}(x, z) = z^\top\big(f(x) - \lambda g(x) + 2\alpha y\big) - \alpha\|z\|^2.$$

Then, it follows from these notations that

$$\psi(\lambda, y, \alpha) = \max_{z\in\Sigma}\min_{x\in X} L^y_{\alpha,\lambda}(x, z) - \alpha\|y\|^2. \tag{4.1}$$

Let

$$\Psi^y_{\alpha,\lambda}(x) = f(x) - \lambda g(x) + 2\alpha y.$$

Then we have

$$L^y_{\alpha,\lambda}(x, z) = -\alpha\|z\|^2 + z^\top\Psi^y_{\alpha,\lambda}(x). \tag{4.2}$$

Proposition 4.3. *Let $\lambda \in \mathbb{R}$, $\alpha > 0$ and $y \in \mathbb{R}^m$. Assume that either* (C1) *is fulfilled and $\lambda \geq 0$ or* (C2) *is satisfied. Then we have*

$$\psi(\lambda, y, \alpha) = \inf_{\substack{x\in X \\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}} \Big\{ \alpha\|z\|^2 + \mu - \alpha\|y\|^2 \;\Big|\; \Psi^y_{\alpha,\lambda}(x) - \mu e - 2\alpha z \leq 0 \Big\},$$

where e denotes the m-vector whose components are all equal to 1.

Proof. Let $\lambda \in \mathbb{R}$, $\alpha > 0$, $y \in \mathbb{R}^m$ and $x \in X$; and consider the minimization problem

$$\begin{cases} \min \alpha\|u\|^2 - u^\top \Psi^y_{\alpha,\lambda}(x) \\ \mathrm{e}^\top u = 1, \\ u \geq 0. \end{cases}$$

Observe that by assumptions, the functions $f_i - \lambda g_i$ are convex and so is the function $u^\top \Psi^y_{\alpha,\lambda}(.)$ for all $u \in \Sigma$.

Now, the Lagrangian associated to this problem is defined on $\mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^m$ by

$$\mathcal{L}^y_{\alpha,\lambda}(u, \mu, \nu) = \alpha\|u\|^2 - u^\top \Psi^y_{\alpha,\lambda}(x) + \mu(\mathrm{e}^\top u - 1) - \nu^\top u,$$

and its dual problem is

$$\begin{cases} \sup \inf_{u\in\mathbb{R}^m} \mathcal{L}^y_{\alpha,\lambda}(u, \mu, \nu) \\ \mu \in \mathbb{R}, \\ \nu \geq 0. \end{cases}$$

Let z be a critical point of $\mathcal{L}^y_{\alpha,\lambda}(.\,, \mu, \nu)$. Then

$$\nabla_u \mathcal{L}^y_{\alpha,\lambda}(z, \mu, \nu) = 0$$

which implies that

$$2\alpha z - \Psi^y_{\alpha,\lambda}(x) + \mu e - \nu = 0$$

and

$$z = \frac{1}{2\alpha}(\Psi^y_{\alpha,\lambda}(x) - \mu e + \nu).$$

Then

$$\inf_{u\in\mathbb{R}^m} \mathcal{L}^y_{\alpha,\lambda}(u, \mu, \nu) = \alpha\|z\|^2 - z^\top \Psi^y_{\alpha,\lambda}(x) + \mu(\mathrm{e}^\top z - 1) - \nu^\top z.$$

Thus

$$\sup_{\mu\in\mathbb{R},\nu\geq 0} \inf_{u\in\mathbb{R}^m} \mathcal{L}^y_{\alpha,\lambda}(u, \mu, \nu) = \sup_{z\in\mathbb{R}^m,\mu\in\mathbb{R},\, \nu\geq 0} \left\{ \alpha\|z\|^2 - z^\top \Psi^y_{\alpha,\lambda}(x) + \mu(\mathrm{e}^\top z - 1) - \nu^\top z \;\middle|\; z = \frac{1}{2\alpha}(\Psi^y_{\alpha,\lambda}(x) - \mu e + \nu)) \right\}. \tag{4.3}$$

Replacing $\Psi^y_{\alpha,\lambda}(x)$ by $2\alpha z + \mu e - \nu$ we get

$$\alpha\|z\|^2 - z^\top \Psi^y_{\alpha,\lambda}(x) + \mu(\mathrm{e}^\top z - 1) - \nu^\top z = -\alpha\|z\|^2 - \mu.$$

So,

$$\begin{aligned} &\sup_{\substack{z\in\mathbb{R}^m \\ \mu\in\mathbb{R},\, \nu\geq 0}} \left\{ \alpha\|z\|^2 - z^\top \Psi^y_{\alpha,\lambda}(x) + \mu(\mathrm{e}^\top z - 1) - \nu^\top z \;\middle|\; z = \frac{1}{2\alpha}(\Psi^y_{\alpha,\lambda}(x) - \mu e + \nu)) \right\} \\ &\quad = \sup_{\substack{z\in\mathbb{R}^m \\ \mu\in\mathbb{R},\, \nu\geq 0}} \big\{ -\alpha\|z\|^2 - \mu \mid \Psi^y_{\alpha,\lambda}(x) - \mu e - 2\alpha z + \nu = 0 \big\}. \end{aligned} \tag{4.4}$$

By (4.3) and (4.4), we get

$$\sup_{\mu\in\mathbb{R},\nu\geq 0}\inf_{u\in\mathbb{R}^m}\mathcal{L}^y_{\alpha,\lambda}(u,\mu,\nu) = \sup_{\substack{z\in\mathbb{R}^m\\ \mu\in\mathbb{R},\,\nu\geq 0}}\big\{-\alpha\|z\|^2-\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z+\nu=0\big\}$$

$$= \sup_{z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}\big\{-\alpha\|z\|^2-\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\} \tag{4.5}$$

$$= -\inf_{z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}\big\{\alpha\|z\|^2+\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\}. \tag{4.6}$$

Using ([17], Cor. 28.2.2 and Thm. 28.4), we get

$$\sup_{\substack{\mu\in\mathbb{R}\\ \nu\geq 0}}\inf_{u\in\mathbb{R}^m}\mathcal{L}^y_{\alpha,\lambda}(u,\mu,\nu) = \min_{\substack{e^\top u=1\\ u\geq 0}}\big\{\alpha\|u\|^2-u^\top\Psi^y_{\alpha,\lambda}(x)\big\}$$

$$= \min_{u\in\Sigma}\big\{\alpha\|u\|^2-u^\top\Psi^y_{\alpha,\lambda}(x)\big\}. \tag{4.7}$$

Then using (4.6) and (4.7), we obtain

$$\min_{u\in\Sigma}\big\{\alpha\|u\|^2-u^\top\Psi^y_{\alpha,\lambda}(x)\big\} = -\inf_{z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}\big\{\alpha\|z\|^2+\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\}.$$

Therefore,

$$\max_{u\in\Sigma}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)\big\} = \inf_{z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}\big\{\alpha\|z\|^2+\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\}. \tag{4.8}$$

Since the function $x\longmapsto -\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)$ is convex for all $u\in\Sigma$, and the function $u\longmapsto -\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)$ is concave for all $x\in X$; and the sets X and Σ are convex and Σ is compact, then Sion's theorem ([20], Thm. 3.4 and Cor. 3.3) implies that

$$\min_{x\in X}\max_{u\in\Sigma}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)\big\} = \max_{u\in\Sigma}\min_{x\in X}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)\big\}.$$

It follows from (4.1), (4.2) and (4.8) that

$$\psi(\lambda,y,\alpha)+\alpha\|y\|^2 = \min_{x\in X}\max_{u\in\Sigma}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)\big\}$$

$$= \min_{x\in X}\inf_{z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}\big\{\alpha\|z\|^2+\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\}$$

$$= \inf_{\substack{x\in X\\ z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}}\big\{\alpha\|z\|^2+\mu \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\}.$$

Therefore

$$\psi(\lambda,y,\alpha) = \inf_{\substack{x\in X\\ z\in\mathbb{R}^m,\,\mu\in\mathbb{R}}}\big\{\alpha\|z\|^2+\mu-\alpha\|y\|^2 \mid \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\big\}. \tag{4.9}$$

□

Now we present an equivalent ordinary convex problem to $(\mathcal{R}(\lambda,y,\alpha))$.

Proposition 4.4. *Let $\lambda\in\mathbb{R}$, $\alpha>0$ and $y\in\mathbb{R}^m$. Assume that either C1 is fulfilled and $\lambda\geq 0$ or C2 is satisfied. Let $(\bar{x},\bar{\mu},\bar{z})\in X\times\mathbb{R}\times\mathbb{R}^m$ be an optimal solution of the problem*

$$(\mathcal{R}^{eq}(\lambda,y,\alpha)) \qquad \inf_{\substack{x\in X\\ \mu\in\mathbb{R},\,z\in\mathbb{R}^m}}\Big\{\alpha\|z\|^2+\mu \;\Big|\; \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\Big\},$$

Then $\bar{z}$ is the solution of $(\mathcal{R}(\lambda,y,\alpha))$.

Proof. Let $\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)$ denotes the Lagrangian associated to $(\mathcal{R}^{\rm eq}(\lambda,y,\alpha))$. Then

$$\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)=\alpha\|z\|^2+\mu+u^\top(\Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z).$$

We start first by finding the dual of $(\mathcal{R}^{\rm eq}(\lambda,y,\alpha))$. So, for all $u\geq 0$ we have

$$\begin{aligned}\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)&=\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\Big\{\alpha\|z\|^2+\mu+u^\top(\Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z)\Big\}\\&=\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\Big\{\alpha\|z\|^2-2\alpha u^\top z+\mu(1-u^\top e)+u^\top\Psi^y_{\alpha,\lambda}(x)\Big\}.\end{aligned}$$

Observe that for all $u\in\mathbb{R}^m$, the function

$$z\longmapsto\alpha\|z\|^2-2\alpha u^\top z+\mu(1-u^\top e)+u^\top\Psi^y_{\alpha,\lambda}(x)$$

achieves its infimum at the unique point $z=u$ for all $(x,\mu)\in X\times\mathbb{R}$. It follows that

$$\begin{aligned}\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)&=\inf_{x\in X,\,\mu\in\mathbb{R}}\Big\{-\alpha\|u\|^2+\mu(1-u^\top e)+u^\top\Psi^y_{\alpha,\lambda}(x)\Big\}\\&=\begin{cases}\min_{x\in X}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)\big\} & \text{if}\quad u^\top e=1\\ -\infty & \text{if}\quad u^\top e\neq 1.\end{cases}\end{aligned}$$

Notice that from the assumption on $\phi(\lambda,y)$ in Remark 4.2 we have $\psi(\lambda,y,\alpha)>-\infty$, and that from (4.9) we have $\vartheta(\mathcal{R}^{\rm eq}(\lambda,y,\alpha))=\psi(\lambda,y,\alpha)+\alpha\|y\|^2$, where $\vartheta(\mathcal{R}^{\rm eq}(\lambda,y,\alpha))$ denotes the infimal value of $(\mathcal{R}^{\rm eq}(\lambda,y,\alpha))$. So,

$$\sup_{u\geq0}\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)=\sup_{\substack{{\rm e}^\top u=1\\ u\geq 0}}\inf_{x\in X}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)\big\}. \tag{4.10}$$

Remember that,

$$\begin{aligned}\psi(\lambda,y,\alpha)&=\min_{x\in X}\max_{u\in\Sigma}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)-\alpha\|y\|^2\big\}\\&=\max_{u\in\Sigma}\min_{x\in X}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)-\alpha\|y\|^2\big\}.\end{aligned}$$

So, by virtue of (4.10) we have

$$\begin{aligned}\psi(\lambda,y,\alpha)&=\sup_{u\geq0}\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)-\alpha\|y\|^2\\&=\max_{u\in\Sigma}\min_{x\in X}\big\{-\alpha\|u\|^2+u^\top\Psi^y_{\alpha,\lambda}(x)-\alpha\|y\|^2\big\}.\end{aligned}$$

It follows that if $\bar u$ is an optimal solution of the problem

$$\sup_{u\geq0}\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u) \tag{4.11}$$

then $\bar u$ is an optimal solution of $(\mathcal{R}(\lambda,y,\alpha))$; and *vice versa*, if $\bar u$ is an optimal solution of $(\mathcal{R}(\lambda,y,\alpha))$, then $\bar u$ is an optimal solution of (4.11).

On the other hand, by using the convexity assumptions and ([17], Thms. 28.2 and 28.4) we get

$$\sup_{u\geq0}\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u)=\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}}\Big\{\alpha\|z\|^2+\mu\ \Big|\ \Psi^y_{\alpha,\lambda}(x)-\mu e-2\alpha z\leq 0\Big\},$$

and if $\bar{u}$ is an optimal solution of $(\mathcal{R}(\lambda, y, \alpha))$ and $(\bar{x}, \bar{\mu}, \bar{z})$ is an optimal solution of $(\mathcal{R}^{\mathrm{eq}}(\lambda, y, \alpha))$, then

$$\begin{aligned}
\sup_{u\geq 0} \inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}} \mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,u) &= \inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}} \mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,\bar{u})\\
&= \sup_{u\geq 0} \mathscr{L}^y_{\alpha,\lambda}(\bar{x},\bar{\mu},\bar{z},u)\\
&= \mathscr{L}^y_{\alpha,\lambda}(\bar{x},\bar{\mu},\bar{z},\bar{u}).
\end{aligned}$$

Since the function

$$z \longmapsto \alpha\|z\|^2 - 2\alpha\bar{u}^\top z + \mu(1-\bar{u}^\top e) + \bar{u}^\top \Psi^y_{\alpha,\lambda}(x)$$

achieves its infimum at the unique point $z=\bar{u}$ for all $(x,\mu)\in X\times\mathbb{R}$, then it follows that

$$\begin{aligned}
\inf_{\substack{x\in X\\ \mu\in\mathbb{R},\, z\in\mathbb{R}^m}} \mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,\bar{u}) &= \inf_{x\in X,\,\mu\in\mathbb{R}} \inf_{z\in\mathbb{R}^m} \mathscr{L}^y_{\alpha,\lambda}(x,\mu,z,\bar{u})\\
&= \inf_{x\in X,\,\mu\in\mathbb{R}} \mathscr{L}^y_{\alpha,\lambda}(x,\mu,\bar{u},\bar{u})\\
&= \mathscr{L}^y_{\alpha,\lambda}(\bar{x},\bar{\mu},\bar{z},\bar{u})
\end{aligned}$$

and $\bar{z}=\bar{u}$.

Finally, $\bar{u}=\bar{z}$ is the solution of $(\mathcal{R}(\lambda, y, \alpha))$. □

5. Convergence and rate of convergence of Algorithm 4.1

To prove the convergence and give the rate of convergence of our algorithm, we begin by showing some intermediate results.

Later, we will use the following notations:

$$\delta = \inf_{x\in X}\min_{i\in I} g_i(x) \quad\text{and}\quad \Delta = \sup_{x\in X}\max_{i\in I} g_i(x).$$

We assume that $\delta > 0$ and $\Delta < \infty$.

5.1. Convergence of Algorithm 4.1

Lemma 5.1. *Let $y\in\Sigma$ and $\lambda\leq\mu$. Then*

(1) $\phi(\lambda,y)\geq\phi(\mu,y)+(\mu-\lambda)\delta$,
(2) $\phi(\mu,y)\geq\phi(\lambda,y)-(\mu-\lambda)\Delta$.

Proof. For all $y\in\Sigma$ and $\lambda\leq\mu$ we have

(1)

$$\begin{aligned}
\phi(\lambda,y) &= \min_{x\in X}\{y^\top(f(x)-\lambda g(x))\}\\
&= \min_{x\in X}\{y^\top(f(x)-\mu g(x)+(\mu-\lambda)g(x))\}
\end{aligned}$$

Thus,

$$\begin{aligned}
\phi(\lambda,y) &\geq \min_{x\in X}\{y^\top(f(x)-\mu g(x))\} + (\mu-\lambda)\min_{x\in X}\{y^\top g(x)\}\\
&= \phi(\mu,y) + (\mu-\lambda)\min_{x\in X}\{y^\top g(x)\}\\
&\geq \phi(\mu,y) + (\mu-\lambda)\delta.
\end{aligned}$$

(2)

$$\begin{aligned}\phi(\mu, y) &= \min_{x\in X}\{y^\top(f(x) - \mu g(x))\}\\ &= \min_{x\in X}\{y^\top(f(x) - \lambda g(x) - (\mu-\lambda)g(x))\}.\end{aligned}$$

Therefore,

$$\begin{aligned}\phi(\mu, y) &\geq \min_{x\in X}\{y^\top(f(x) - \lambda g(x))\} - (\mu-\lambda)\max_{x\in X}\{y^\top g(x)\}\\ &= \phi(\lambda, y) - (\mu-\lambda)\max_{x\in X}\{y^\top g(x)\}\\ &\geq \phi(\lambda, y) - (\mu-\lambda)\varDelta.\end{aligned}$$

□

Lemma 5.2. *Let λ_* denotes the infimal value of (P). Then we have*

(1) *for all $y \in \varSigma$, $\phi(c(y), y) = 0$,*
(2) *if the problem (P) has an optimal solution, then $\phi(\lambda_*, y_*) = 0$ and $c(y_*) = \lambda_*$ for all $y_* \in \operatorname{argmax}_{y\in\varSigma}\phi(\lambda_*, y)$.*

Proof.

(1) For all $y \in \varSigma$ let $x_y \in X$ be such that

$$\begin{aligned}c(y) &:= \min_{x\in X}\frac{y^\top f(x)}{y^\top g(x)}\\ &= \frac{y^\top f(x_y)}{y^\top g(x_y)}.\end{aligned}$$

Then $y^\top(f(x_y) - c(y)g(x_y)) = 0$. Since

$$\phi(c(y), y) = \min_{x\in X} y^\top\big(f(x) - c(y)g(x)\big),$$

then $\phi(c(y), y) \leq 0$.

On the other hand, from the definition of $c(y)$ we have

$$c(y) \leq \frac{y^\top f(x)}{y^\top g(x)} \quad \text{for all } x\in X,$$

which implies that $y^\top(f(x) - c(y)g(x)) \geq 0$ for all $x \in X$. So, $\phi(c(y), y) \geq 0$. Finally $\phi(c(y), y) = 0$.

(2) Let $y_* \in \operatorname{argmax}_{y\in\varSigma}\phi(\lambda_*, y)$. Then

$$\begin{aligned}\phi(\lambda_*, y_*) &= \max_{y\in\varSigma}\min_{x\in X} y^\top\big(f(x) - \lambda_* g(x)\big)\\ &= \min_{x\in X}\max_{y\in\varSigma} y^\top\big(f(x) - \lambda_* g(x)\big)\\ &= \min_{x\in X}\max_{1\leq i\leq m}\big\{f_i(x) - \lambda_* g_i(x)\big\}.\end{aligned}$$

It is then easy to see that $\phi(\lambda_*, y_*) = 0$ when (P) has an optimal solution. Now, let $x^* \in X$ be such that

$$\begin{aligned}\phi(\lambda_*, y_*) &= \min_{x\in X} y_*^\top\big(f(x) - \lambda_* g(x)\big)\\ &= y_*^\top\big(f(x^*) - \lambda_* g(x^*)\big).\end{aligned} \tag{5.1}$$

Then $y_*^\top(f(x^*) - \lambda_* g(x^*)) = 0$ which implies that

$$\lambda_* = \frac{y_*^\top f(x^*)}{y_*^\top g(x^*)} \geq c(y^*). \tag{5.2}$$

On the other hand, since $\phi(\lambda_*, y_*) = 0$ then from (5.1) we get

$$y_*^\top \big(f(x) - \lambda_* g(x)\big) \geq 0 \quad \text{for all } x \in X.$$

Therefore

$$\frac{y_*^\top f(x)}{y_*^\top g(x)} \geq \lambda_* \quad \text{for all } x \in X.$$

This means that $c(y_*) \geq \lambda_*$. The equality follows from (5.2). □

Remark 5.3. Let $x_y \in X$ be such that

$$\begin{aligned} c(y) &:= \min_{x\in X} \frac{y^\top f(x)}{y^\top g(x)} \\ &= \frac{y^\top f(x_y)}{y^\top g(x_y)}. \end{aligned}$$

Then even if $c(y) = \lambda_*$, x_y is not necessarily a solution of (P). This is shown by the next example.

Example 5.4. Let $n = 1$, $m = 2$, $f_1(x) = 1$, $g_1(x) = x$, $f_2(x) = x$, $g_2(x) = 1$ and $X = [1, 2]$.

The problem (P) is to solve

$$\min_{x\in X} \max\left\{\frac{1}{x}, x\right\}.$$

Then the minimal value is $\lambda_* = 1$ and achieved at $x = 1$.

On the other hand, for $y = (y_1, y_2)^\top \in \Sigma$ we have

$$c(y) = \min_{x\in X} \frac{y_1 + y_2 x}{y_1 x + y_2}. \tag{5.3}$$

For $y_1 = y_2 = 1/2$, we have $c(y) = 1 = \lambda_*$, and

$$\begin{aligned} c(y) &= \min_{x\in X} \frac{y_1 + y_2 x}{y_1 x + y_2} \\ &= \min_{x\in X} \frac{(1+x)/2}{(1+x)/2} \\ &= 1 \end{aligned}$$

and the minimum is attained at all $x \in X$. This shows that a minimum of (5.3) is not necessarily a solution of (P) even when y is a solution of the dual.

Furthermore,

$$\begin{aligned} \phi(\lambda_*, y) &= \min_{x\in X} \{y_1(1-x) + y_2(x-1)\} \\ &= \min_{x\in X} \{(1-x)/2 + (x-1)/2\} \\ &= 0 \end{aligned}$$

where the minimum is reached at every $x \in X$. This also shows that a solution of the last problem is not necessarily a solution of (P) even when y is a solution of the dual.

Lemma 5.5. *Let $\bar{y}_{k+1}$ be the optimal solution of $(\mathcal{R}(c_k, y_k, \alpha_k))$. Then, for all $y \in \Sigma$ we have*

$$\begin{aligned}\phi(c_k, \bar{y}_{k+1}) - \phi(c_k, y) &\geq -2\alpha_k \langle \bar{y}_{k+1} - y_k, y - \bar{y}_{k+1} \rangle \\ &= -\alpha_k \|y - y_k\|^2 + \alpha_k \|y - \bar{y}_{k+1}\|^2 + \alpha_k \|y_k - \bar{y}_{k+1}\|^2.\end{aligned}$$

Proof. Apply for example ([12], Prop. 2.2, p. 37) to the function $-\phi(c_k, .)$. □

Lemma 5.6. *Let $\{\mu_k\}$, $\{\beta_k\}$ be sequences of nonnegative reals such that*

$$\sum_{j=1}^{\infty} \mu_j < \infty, \sum_{j=1}^{\infty} \beta_j < \infty,$$

and let $\{u_k\}$ be a sequence of reals such that

$$u_{k+1} \leq (1 + \mu_k) u_k + \beta_k.$$

Then, the sequence $\{u_k\}$ converges to some $u \in \mathbb{R} \cup \{-\infty\}$.

Proof. See [19]. □

Lemma 5.7. *Assume that $\sum_{k \geq 0} \eta_k < \infty$. Then, the sequence $\{c_k\}$ generated by Algorithm 4.1 converges to some $c_* \in \mathbb{R}$.*

Proof. From the definition of y_{k+1} in Algorithm 4.1, we have

$$\psi(c_k, y_k, \alpha_k) - \eta_k \leq \phi(c_k, y_{k+1}) - \alpha_k \|y_{k+1} - y_k\|^2 \leq \phi(c_k, y_{k+1}).$$

From the definition of $\phi(c_k, y_{k+1})$, we get

$$\phi(c_k, y_{k+1}) \leq y_{k+1}^\top \big(f(x) - c_k g(x)\big) \qquad \forall x \in X.$$

On the other hand, since $c_{k+1} = c(y_{k+1})$, there exists some $\hat{x}_k \in X$ such that

$$c_{k+1} = \frac{y_{k+1}^\top f(\hat{x}_k)}{y_{k+1}^\top g(\hat{x}_k)},$$

which implies that

$$\psi(c_k, y_k, \alpha_k) - \eta_k \leq \phi(c_k, y_{k+1}) \leq y_{k+1}^\top g(\hat{x}_k)(c_{k+1} - c_k).$$

From the definition of $\psi(c_k, y_k, \alpha_k)$, we deduce that $\psi(c_k, y_k, \alpha_k) \geq \phi(c_k, y_k) = 0$, where the last equality follows from Lemma 5.2. Thus,

$$-\eta_k \leq \phi(c_k, y_{k+1}) \leq y_{k+1}^\top g(\hat{x}_k)(c_{k+1} - c_k). \tag{5.4}$$

It follows that

$$y_{k+1}^\top g(\hat{x}_k)(\lambda_* - c_{k+1}) \leq y_{k+1}^\top g(\hat{x}_k)(\lambda_* - c_k) + \eta_k. \tag{5.5}$$

Inequality (5.5) implies that

$$\lambda_* - c_{k+1} \leq \lambda_* - c_k + \frac{\eta_k}{y_{k+1}^\top g(\hat{x}_k)} \leq \lambda_* - c_k + \frac{\eta_k}{\delta}. \tag{5.6}$$

It is easy to see that since

$$c_k := c(y_k) := \min_{x \in X} \left\{ \frac{y_k^\top f(x)}{y_k^\top g(x)} \right\},$$

then $c_k \leq \lambda_*$, so that $\lambda_* - c_k \geq 0$.

Using Lemma 5.6 with $u_k = \lambda_* - c_k \geq 0$, $\mu_k = 0$, $\beta_k = \eta_k/\delta$, we conclude that $\{\lambda_* - c_k\}$ converges in $\mathbb{R}$. This implies that $\{c_k\}$ converges to some $c_* \in \mathbb{R}$. □

Lemma 5.8. *For all $k \in \mathbb{N}$, let $\bar{y}_{k+1}$ be the optimal solution of $(\mathcal{R}(c_k, y_k, \alpha_k))$. Then*

$$\|y_{k+1} - \bar{y}_{k+1}\| \leq \sqrt{\frac{\eta_k}{\alpha_k}}.$$

Proof. From Lemma 5.5 we have for all $y \in \Sigma$,

$$\phi(c_k, \bar{y}_{k+1}) \geq \phi(c_k, y) + \alpha_k \big[-\|y - y_k\|^2 + \|y - \bar{y}_{k+1}\|^2 + \|\bar{y}_{k+1} - y_k\|^2 \big]. \tag{5.7}$$

From the definition of y_{k+1}, we have for all $y \in \Sigma$

$$\phi(c_k, y_{k+1}) - \alpha_k \|y_{k+1} - y_k\|^2 \geq \phi(c_k, y) - \alpha_k \|y - y_k\|^2 - \eta_k.$$

For $y = \bar{y}_{k+1}$ in the last inequality, we obtain

$$\phi(c_k, y_{k+1}) - \alpha_k \|y_{k+1} - y_k\|^2 + \alpha_k \|\bar{y}_{k+1} - y_k\|^2 + \eta_k \geq \phi(c_k, \bar{y}_{k+1}). \tag{5.8}$$

Considering this inequality and (5.7) with $y = y_{k+1}$, we arrive to

$$\eta_k \geq \alpha_k \|y_{k+1} - \bar{y}_{k+1}\|^2. \qquad \square$$

Lemma 5.9. *Let $\bar{y}_k$ be the optimal solution of $(\mathcal{R}(c_{k-1}, y_{k-1}, \alpha_{k-1}))$. Then for all $k \in \mathbb{N}^*$ and for all $y \in \Sigma$, we have*

$$\|y - y_k\|^2 \leq \left(1 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\right) \|y - \bar{y}_k\|^2 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}} + \frac{\eta_{k-1}}{\alpha_{k-1}}.$$

Proof. We have

$$\|y - y_k\|^2 = \|y - \bar{y}_k\|^2 + \|\bar{y}_k - y_k\|^2 + 2\langle y - \bar{y}_k, \bar{y}_k - y_k \rangle.$$

The Schwartz inequality implies

$$\|y - y_k\|^2 \leq \|y - \bar{y}_k\|^2 + \|\bar{y}_k - y_k\|^2 + 2\|y - \bar{y}_k\| \|\bar{y}_k - y_k\|.$$

Using Lemma 5.8, we get

$$\|y - y_k\|^2 \leq \|y - \bar{y}_k\|^2 + \frac{\eta_{k-1}}{\alpha_{k-1}} + 2\|y - \bar{y}_k\| \sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}.$$

Remarking that $\|y - \bar{y}_k\| \leq 1 + \|y - \bar{y}_k\|^2$, we get

$$\|y - y_k\|^2 \leq \left(1 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\right) \|y - \bar{y}_k\|^2 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}} + \frac{\eta_{k-1}}{\alpha_{k-1}}. \qquad \square$$

Lemma 5.10. *If $\lambda \leq \lambda_*$ is such that $\phi(\lambda, y) \leq 0$ for all $y \in \Sigma$, then $\lambda = \lambda_*$.*

Proof. For all $y \in \Sigma$, if $\phi(\lambda, y) \leq 0$ then from the assumption in Remark 4.2, there exists $x \in X$ such that $\phi(\lambda, y) = y^\top (f(x) - \lambda g(x)) \leq 0$. Obviously, the last inequality is equivalent to

$$\frac{y^\top f(x)}{y^\top g(x)} \leq \lambda.$$

Thus $c(y) \leq \lambda$ for all $y \in \Sigma$. It follows that

$$\max_{y \in \Sigma} c(y) \leq \lambda. \tag{5.9}$$

Using the convexity assumptions on f_i and g_i, the function $x \longmapsto \dfrac{y^\top f(x)}{y^\top g(x)}$ is quasiconvex and continuous on X for every $y \in \Sigma$ and the function $y \longmapsto \dfrac{y^\top f(x)}{y^\top g(x)}$ is quasiconcave and continuous on Σ for every $x \in X$.

Combining the last results with the convexity of X and Σ and the compactness of Σ, and using Sion's theorem, we get

$$\max_{y\in\Sigma}\min_{x\in X}\frac{y^\top f(x)}{y^\top g(x)} = \min_{x\in X}\max_{y\in\Sigma}\frac{y^\top f(x)}{y^\top g(x)}\cdot$$

On the other hand, we have

$$\min_{x\in X}\max_{y\in\Sigma}\frac{y^\top f(x)}{y^\top g(x)} = \min_{x\in X}\max_{i\in I}\frac{f_i(x)}{g_i(x)} = \lambda_*. \tag{5.10}$$

But

$$\max_{y\in\Sigma}\min_{x\in X}\frac{y^\top f(x)}{y^\top g(x)} = \max_{y\in\Sigma} c(y).$$

So, this equality together with (5.9) and (5.10) give $\lambda_* \le \lambda$. The equality $\lambda_* = \lambda$ follows from the assumption $\lambda \le \lambda_*$. □

Theorem 5.11. *Suppose that* $\sum_{k\ge 0} 1/\alpha_k = +\infty$, *that* $\sum_{k\ge 0}\sqrt{\eta_k/\alpha_k} < +\infty$ *and that* $\sum_{k\ge 0}\eta_k < \infty$. *Then the sequence* $\{c_k\}$ *converges to* λ_*.

Proof. From (5.7) and (5.8), we obtain for all $y \in \Sigma$,

$$\phi(c_k, y_{k+1}) - \phi(c_k, y) + \alpha_k\|y - y_k\|^2 + \eta_k \ge \alpha_k\|y - \bar{y}_{k+1}\|^2. \tag{5.11}$$

Using Lemma 5.9 in this inequality, we obtain for all $y \in \Sigma$,

$$\|y - \bar{y}_{k+1}\|^2 \le \left(1 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\right)\|y - \bar{y}_k\|^2 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}} + \frac{\eta_{k-1}}{\alpha_{k-1}} + \frac{\eta_k}{\alpha_k}$$
$$+\frac{1}{\alpha_k}\big(\phi(c_k, y_{k+1}) - \phi(c_k, y)\big) \tag{5.12}$$

Since $\delta \le y^\top g(x) \le \Delta$ for all $y \in \Sigma$ and $x \in X$, and the sequence $\{c_k\}$ converges, by Lemma 5.7, inequalities (5.4) give,

$$\lim_{k\longrightarrow\infty}\phi(c_k, y_{k+1}) = 0. \tag{5.13}$$

We assert that

$$\lim_{k\longrightarrow\infty}\phi(c_k, y_{k+1}) - \phi(c_k, y) \ge 0 \qquad \text{for all } y \in \Sigma. \tag{5.14}$$

Indeed, assume the contrary. Then there exists $\hat{y} \in \Sigma$, $\epsilon > 0$ and $k_0 \in \mathbb{N}$ such that

$$\phi(c_k, y_{k+1}) - \phi(c_k, \hat{y}) < -\epsilon \qquad \text{for all } k \ge k_0.$$

So, for all $k \ge k_0$, inequality (5.12) with $y = \hat{y}$ gives

$$\|\hat{y} - \bar{y}_{k+1}\|^2 \le \left(1 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\right)\|\hat{y} - \bar{y}_k\|^2 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}} + \frac{\eta_{k-1}}{\alpha_{k-1}} + \frac{\eta_k}{\alpha_k} - \frac{\epsilon}{\alpha_k}\cdot \tag{5.15}$$

Then

$$\|\hat{y} - \bar{y}_{k+1}\|^2 \le \left(1 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\right)\|\hat{y} - \bar{y}_k\|^2 + 2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}} + \frac{\eta_{k-1}}{\alpha_{k-1}} + \frac{\eta_k}{\alpha_k}\cdot \tag{5.16}$$

Using the assumptions of the theorem, we have

$$\sum_{k\geq 1}\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}} < +\infty \quad \text{and} \quad \sum_{k\geq 1}\left(2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}+\frac{\eta_{k-1}}{\alpha_{k-1}}+\frac{\eta_k}{\alpha_k}\right) < +\infty,$$

and from (5.16), Lemma 5.6 entails that $\{\|\hat{y}-\bar{y}_k\|\}$ converges. Consequently,

$$\sum_{k\geq 1}\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\|\hat{y}-\bar{y}_k\|^2 < +\infty.$$

Summing in (5.15) over $k = k_0, \dots, n$, we get

$$\begin{aligned}\|\hat{y}-\bar{y}_{n+1}\|^2 - \|\hat{y}-\bar{y}_{k_0}\|^2 \leq{}& 2\sum_{k=k_0}^{n}\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}\|\hat{y}-\bar{y}_k\|^2 \\ &+ \sum_{k=k_0}^{n}\left(2\sqrt{\frac{\eta_{k-1}}{\alpha_{k-1}}}+\frac{\eta_{k-1}}{\alpha_{k-1}}+\frac{\eta_k}{\alpha_k}\right) - \sum_{k=k_0}^{n}\frac{\epsilon}{\alpha_k}\cdot \end{aligned} \tag{5.17}$$

Evidently, the inequality (5.17) can not hold because of the assumption $\sum_{k\geq 0} 1/\alpha_k = +\infty$. We then conclude that (5.14) holds and we have

$$\phi(c_*, y) \leq 0 \qquad \text{for all } y \in \Sigma.$$

Since $c_* \leq \lambda_*$, Lemma 5.10 implies that $c_* = \lambda_*$. □

Proposition 5.12. *The sequence $\{y_k\}$ is bounded and, if the assumptions of Theorem 5.11 are fulfilled then every accumulation point is a solution of $(\mathcal{R})$.*

Proof. For all $k \in \mathbb{N}$, c_k is defined by

$$c_k = \min_{x\in X}\frac{y_k^\top f(x)}{y_k^\top g(x)}\cdot$$

This means that

$$c_k \leq \frac{y_k^\top f(x)}{y_k^\top g(x)} \quad \text{for all } x \in X.$$

The sequence $\{y_k\}$ is bounded since $\{y_k\} \subset \Sigma$. Let $\hat{y} \in \Sigma$ be an accumulation point of $\{y_k\}$. From Theorem 5.11 the sequence $\{c_k\}$ converges to λ_*, and we have

$$\lambda_* \leq \frac{\hat{y}^\top f(x)}{\hat{y}^\top g(x)} \quad \text{for all } x \in X.$$

It follows that

$$\lambda_* \leq \min_{x\in X}\frac{\hat{y}^\top f(x)}{\hat{y}^\top g(x)} \leq \max_{y\in\Sigma}\min_{x\in X}\frac{y^\top f(x)}{y^\top g(x)} = \lambda_*.$$

Finally

$$\begin{aligned}\min_{x\in X}\frac{\hat{y}^\top f(x)}{\hat{y}^\top g(x)} &= \max_{y\in\Sigma}\min_{x\in X}\frac{y^\top f(x)}{y^\top g(x)} \\ &= \max_{y\in\Sigma} c(y),\end{aligned}$$

and $\hat{y}$ is a solution of $(\mathcal{R})$. □

5.2. Rate of convergence of Algorithm 4.1

Hereafter, we focus on the study of the rate of convergence of Algorithm 4.1. For this, let

$$\Gamma = \Sigma \quad \text{and} \quad \Gamma^* = \operatorname*{argmax}_{y\in\Sigma} \phi(\lambda_*, y).$$

Next, we will denote by (H) the following assumption:

$$(H) \quad \exists \rho > 0\ \exists \kappa > 0 \text{ such that } -\phi(\lambda_*, y) \geq \kappa \operatorname{dist}(y, \Gamma^*)^2 \quad \text{for all } y \in B(\Gamma^*, \rho) \cap \Gamma$$

where

$$B(\Gamma^*, \rho) = \bigcup_{\bar{y}\in\Gamma^*} B(\bar{y}, \rho), \quad B(\bar{y}, \rho) = \{y \in \mathbb{R}^m \mid \|y - \bar{y}\| \leq \rho\}$$

and

$$\operatorname{dist}(y, \Gamma^*) = \inf_{\bar{y}\in\Gamma^*} \|\bar{y} - y\|.$$

Proposition 5.13. *Assume that $f - \lambda_* g$ is a linear map, and X is a polyhedral set. Then assumption (H) is satisfied.*

Proof. Let

$$D = (-f + \lambda_* g)(X).$$

The support function of D is defined, for all $y \in \mathbb{R}^m$ by

$$\begin{aligned}\delta^*(y \mid D) &:= \sup_{x\in X}\{y^\top(-f(x) + \lambda_* g(x))\}\\ &= -\phi(\lambda_*, y).\end{aligned}$$

Theorem 19.3 in [17] implies that D is a polyhedral set. Corollary 19.2.1 in [17] then implies that $\delta^*(y \mid D)$ is polyhedral. Following ([5], Thm. 3.5 and Cor. 3.6), we deduce that

$$\exists \kappa > 0 \text{ such that } -\kappa \operatorname{dist}(y, \Gamma^*) \geq \phi(\lambda_*, y) \quad \text{for all } y \in \Gamma,$$

since for all $y_* \in \Gamma^*$, $\phi(\lambda_*, y_*) = 0$. Then, for $0 < \rho < 1$ and $y \in B(\Gamma^*, \rho) \cap \Gamma$, we have $\operatorname{dist}(y, \Gamma^*) \leq 1$ and thus $\operatorname{dist}(y, \Gamma^*) \geq \operatorname{dist}(y, \Gamma^*)^2$. It follows that

$$-\kappa \operatorname{dist}(y, \Gamma^*)^2 \geq \phi(\lambda_*, y) \quad \text{for all } y \in B(\Gamma^*, \rho) \cap \Gamma.$$

□

Theorem 5.14. *In addition to the hypothesis of Theorem 5.11, assume that the assumption (H) is fulfilled. Assume on the other hand that Algorithm 4.1 is performed in its exact form with $\alpha_k \geq \bar{\alpha} > 0$ for all $k \in \mathbb{N}$. Then the sequence $\{y_k\}$ converges to some solution of $(\mathcal{R})$ and for α_k sufficiently small, the sequence $\{c_k\}$ converges linearly to λ_*.*

Proof. In the case of exact minimization, we have $y_{k+1} = \bar{y}_{k+1}$. Let y_* be an accumulation point of $\{y_k\}$. Then from Proposition 5.12, y_* is an optimal solution of $(\mathcal{R})$. With $y = y_*$ and $\eta_k = 0$ in (5.12) we obtain

$$\|y_* - y_{k+1}\|^2 \leq \|y_* - y_k\|^2 + \frac{1}{\alpha_k}\big(\phi(c_k, y_{k+1}) - \phi(c_k, y_*)\big). \tag{5.18}$$

By considering (5.4) with $\eta_k = 0$, we see that $c_{k+1} - c_k \geq 0$ for all $k \in \mathbb{N}$, and we obtain from Lemma 5.1 and the fact that $\phi(c_{k+1}, y_{k+1}) = 0$ the inequality

$$\phi(c_k, y_{k+1}) \leq \Delta(c_{k+1} - c_k). \tag{5.19}$$

Since from Lemma 5.2 we have $\phi(\lambda_*, y_*) = 0$, then Lemma 5.1 also gives

$$\phi(c_k, y_*) \geq \delta(\lambda_* - c_k). \tag{5.20}$$

Taking into account the fact that by (5.20), $\phi(c_k, y_*) \geq 0$, inequality (5.18) with (5.19) then gives

$$\|y_* - y_{k+1}\|^2 \leq \|y_* - y_k\|^2 + \frac{\Delta}{\bar{\alpha}}(c_{k+1} - c_k).$$

It follows from Lemma 5.6 with $u_k = \|y_* - y_k\|^2$, $\mu_k = 0$ and $\beta_k = \Delta(c_{k+1} - c_k)/\bar{\alpha}$, that the sequence $\{\|y_* - y_k\|\}$ converges. Since it has a subsequence converging to 0, the whole sequence converges to 0, and the sequence $\{y_k\}$ converges to y_*.

Now, since $\{y_k\} \subset \Gamma$ converges to some $y_* \in \Gamma^*$, then for k large, $y_k \in B(\Gamma^*, \rho) \cap \Gamma$. Let $\tilde{y}_k \in \Gamma^*$ be such that

$$\|\tilde{y}_k - y_k\| = \operatorname{dist}(y_k, \Gamma^*).$$

The last equality and the assumption (H) imply that

$$-\kappa\|\tilde{y}_k - y_k\|^2 \geq \phi(\lambda_*, y_k).$$

Since $\phi(c_k, y_k) = 0$, then using Lemma 5.1 and the last inequality we get

$$-\kappa\|\tilde{y}_k - y_k\|^2 \geq \phi(\lambda_*, y_k) \geq \Delta(c_k - \lambda_*). \tag{5.21}$$

Since $\tilde{y}_k \in \Gamma^*$, then $\phi(\lambda_*, \tilde{y}_k) = 0$ and Lemma 5.1 also gives

$$\phi(c_k, \tilde{y}_k) \geq \delta(\lambda_* - c_k). \tag{5.22}$$

On the other hand, the definition of y_{k+1} gives, for all $y \in \Sigma$

$$\phi(c_k, y_{k+1}) - \alpha_k\|y_{k+1} - y_k\|^2 \geq \phi(c_k, y) - \alpha_k\|y - y_k\|^2. \tag{5.23}$$

Then, with $y = \tilde{y}_k$ in (5.23), and by considering (5.21), (5.22) and (5.19), we obtain

$$\Delta(c_{k+1} - \lambda_*) + \Delta(\lambda_* - c_k) \geq \delta(\lambda_* - c_k) + \frac{\alpha_k \Delta}{\kappa}(c_k - \lambda_*).$$

It follows that

$$c_{k+1} - \lambda_* \geq \left(\frac{\alpha_k}{\kappa} - \frac{\delta}{\Delta} + 1\right)(c_k - \lambda_*).$$

Thus,

$$\frac{c_{k+1} - \lambda_*}{c_k - \lambda_*} \leq \frac{\alpha_k}{\kappa} - \frac{\delta}{\Delta} + 1.$$

Therefore, if $\alpha_k \leq \tau < \kappa\delta/\Delta$ then

$$\frac{c_{k+1} - \lambda_*}{c_k - \lambda_*} < 1 - \big(\delta/\Delta - \tau/\kappa\big) < 1. \qquad \square$$

6. Numerical tests

In the following numerical examples, we implemented the algorithms on a personal computer equipped with Matlab R2010A and we use Matlab subroutines linprog and quadprog.

We consider generalized linear fractional programs of the form

$$\inf_{x\in X}\left\{\max_{1\le i\le m}\frac{A_i x + a_i}{B_i x + b_i}\right\}$$

where

$$X = \{x \in \mathbb{R}^n \mid Cx \le \xi,\ x \ge 0\},$$

$A_i^\top, B_i^\top \in \mathbb{R}^n$ and $a_i, b_i \in \mathbb{R}$; C an $p\times n$ matrix and $\xi \in \mathbb{R}^p$. We notice by A and B (resp. a and b) the matrices (resp. vectors) whose rows are the A_i's and B_i's respectively (resp. whose components are a_i and b_i respectively).

The data A_i, B_i, a_i, b_i, C and ξ are generated as follows:

- each element of the vector A_i is uniformly drawn from $[-15, 45]$. Similarly a_i is uniformly drawn from $[-30, 0]$,
- each element of the vector B_i is uniformly drawn from $[0, 10]$. Similarly b_i is drawn uniformly from $[1, 5]$,
- the elements of the matrix C are uniformly distributed within $[0, 10]$. Similarly the elements of the vector ξ are uniformly distributed within $[0, 1]$.

The stopping criterion for Algorithm 4.1 is to reach the accuracy

$$y_{k+1}^\top\left[(A - c_k B)x_{k+1} + a - c_k b\right] \le 10^{-8},$$

where (x_{k+1}, y_{k+1}) is obtained from a solution of $(\mathcal{R}^{eq}(c_k, y_k, \alpha_k))$ in Proposition 4.4.

Observe that if we set $f_i(x) = A_i x + a_i$ and $g_i(x) = B_i x + b_i$ for $i = 1, \dots, m$, then

$$y^\top\left[(A - cB)x + a - cb\right] = y^\top(f(x) - cg(x)).$$

It follows that if

$$y_{k+1}^\top\left[(A - c_k B)x_{k+1} + a - c_k b\right] \le 10^{-8},$$

then

$$y_{k+1}^\top(f(x_{k+1}) - c_k g(x_{k+1})) \le 10^{-8}.$$

This implies that

$$\phi(c_k, y_{k+1}) \le 10^{-8}.$$

We use the previous stopping criterion since (5.13) and (5.14) imply that $\phi(c_*, y) \le 0$ for all $y \in \Sigma$ which implies, from Lemma 5.10, that $c_* = \lambda_*$, wher $c_* = \lim_{k\to\infty} c_k$.

For algorithm [2], we use the stopping criterion $\phi(\lambda_k, y_{k+1}) \le 10^{-8}$ where y_{k+1} is as defined in step (2) of Algorithm 3.1.

During these numerical tests, the two Algorithms will be tested for different sizes ($n = 20$, $m = 10$, $p = 5$), ($n = 50$, $m = 30$, $p = 20$), ($n = 100$, $m = 50$, $p = 30$), where n is the number of variables, m is the number of ratios and p is the number of constraints (without the positivity constraints).

In these tests, we analyze the behavior of Algorithm 4.1 with respect to the regularizing parameter α on sets of five problems, and in the same time, we test the efficiency of the two algorithms. The results are reported in the Tables 1–3.

TABLE 1. The number of iterations and times with $n = 20$, $m = 10$, $p = 5$.

Problems		α					
		10	1	10^{-1}	10^{-2}	10^{-3}	Alg [2]
1	It	176	178	178	178	178	178
	T(s)	10.32	10.62	10.54	10.44	10.47	4.16
2	It	278	66	35	32	31	31
	T(s)	16.54	3.90	2.10	1.91	1.89	0.66
3	It	127	127	127	127	127	127
	T(s)	7.68	7.65	7.59	7.63	7.59	3.11
4	It	68	58	62	62	62	62
	T(s)	4.05	3.44	3.66	3.65	3.66	1.40
5	It	202	79	69	74	75	75
	T(s)	12.03	4.68	4.03	4.28	4.38	1.69

TABLE 2. The number of iterations and times with $n = 50$, $m = 30$, $p = 20$.

Problems		α					
		10	1	10^{-1}	10^{-2}	10^{-3}	Alg [2]
1	It	111	116	117	117	118	118
	T(s)	42.13	43.41	43.92	44.71	43.86	5.53
2	It	123	101	101	102	102	102
	T(s)	46.78	37.75	39.11	39.21	38.43	4.65
3	It	90	91	91	91	91	91
	T(s)	33.26	34.51	33.67	33.24	33.21	4.56
4	It	53	46	46	47	47	47
	T(s)	19.69	17.37	17.50	17.70	17.99	2.12
5	It	76	76	76	76	76	76
	T(s)	27.52	28.60	27.81	27.46	28.00	3.49

TABLE 3. The number of iterations and times with $n = 100$, $m = 50$, $p = 30$.

Problems		α					
		10	1	10^{-1}	10^{-2}	10^{-3}	Alg [2]
1	It	54	53	54	57	56	57
	T(s)	83.11	81.72	83.05	87.37	85.85	3.91
2	It	39	17	17	17	17	17
	T(s)	59.89	25.74	26.32	27.23	27.62	1.31
3	It	56	22	21	20	20	20
	T(s)	85.73	33.55	32.38	31.15	31.07	1.49
4	It	39	11	10	9	9	9
	T(s)	60.48	17.08	15.61	13.92	13.90	0.83
5	It	27	14	13	13	13	13
	T(s)	42.44	22.06	19.89	19.93	20.32	1.00

7. Conclusion and perspectives

As we can observe from these results, the number of iterations decreases when the regularization parameter α becomes small.

On the other hand, the first algorithm requires more time than algorithm [2], in favor of auxiliary problems with unique dual solutions. This is expected because our algorithm treats simultaneously primal and dual variables. But at least for these sets of test problems, both algorithms solve the problems with the same number of iterations when the regularization parameter is small.

Our future research, following this work, is to improve the performance of this algorithm. For example, in step (1) of Algorithm 4.1 a fractional problem has to be solved, we would like to escape this step by using the fact that the parametric problem solves a dual-primal problem which generates a primal and dual sequences. On the other hand, the information that the algorithm generates a sequence of primal solutions $\{x_k\}$ is not used at all.

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] 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.

[3] 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.

[4] J.C. Bernard and J.A. Ferland, Convergence of Interval-Type Algorithms for Generalized Fractional Programming. *Math. Program.* **43** (1989) 349–363.

[5] M.C. Burke, J.V. and Ferris, Weak Sharp Minima in Mathematical Programming. *SIAM J. Control Optim.* **31** (1993) 1340–1359.

[6] R. Correa and C. Lemaréchal, Convergence of Some Algorithms for Convex Minimization. *Math. Program.* **62** (1993) 261–275.

[7] J.P. Crouzeix and J.A. Ferland, Algorithms for Generalized Fractional Programming. *Math. Program.* **52** (1991) 191–207.

[8] J.P. Crouzeix, J.A. Ferland and S. Schaible, Duality in Generalized Linear Fractional Programming. *Math. Program.* **27** (1983) 342–354.

[9] J.P. Crouzeix, J.A. Ferland and S. Schaible, An Algorithm for Generalized Fractional Programs. *J. Optim. Theory Appl.* **47** (1985) 35–49.

[10] 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.

[11] W. Dinkelbach, On Nonlinear Fractional Programming. *Manag. Sci.* **13** (1967) 492–498.

[12] I. Ekeland and R. Temam, Analyse Convexe et Problèmes Variationnels. Gauthier-Villars, Paris, Bruxelles, Montréal (1974).

[13] M. Gugat, Prox-Regularization Methods for Generalized Fractional Programming. *J. Optim. Theory Appl.* **99** (1998) 691–722.

[14] O. Güler, On the Convergence of the Proximal Point Algorithm for Convex Minimization. *SIAM J. Control Optim.* **29** (1991) 403–419.

[15] J-B. Hiriart-Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms II. Springer-Verlag (1993).

[16] J.-Y. Lin, H.-J. Chen and R.-L. Sheu, Augmented Lagrange Primal-Dual Approach for Generalized Fractional Programming Problems. *Ind. Manag. Optim.* **4** (2013) 723–741.

[17] R.T. Rockafellar, Convex Analysis. Princeton University Press, Princeton, N.J. (1971).

[18] A. Roubi, Method of Centers for Generalized Fractional Programming. *J. Optim. Theory Appl.* **107** (2000) 123–143.

[19] A. Roubi, Convergence of Prox-Regularization Methods for Generalized Fractional Programming. *RAIRO: OR* **36** (2002) 73–94.

[20] M. Sion, On General Minimax Theorems. *Pacific J. Math.* **8** (1958) 171–176.

[21] J.J. Strodiot, J.P. Crouzeix, V.H. Nguyen and J.A. Ferland, An Inexact Proximal Point Method for Solving Generalized Fractional Programs. *J. Glob. Optim.* **42** (2008) 121–138.