RAIRO Operations Research
RAIRO Oper. Res. **36** (2002) 73-94
DOI: 10.1051/ro:2002006

CONVERGENCE OF PROX-REGULARIZATION METHODS FOR GENERALIZED FRACTIONAL PROGRAMMING

Ahmed Roubi[1]

Communicated by Jean-Pierre Crouzeix

Abstract. We analyze the convergence of the prox-regularization algorithms introduced in [1], to solve generalized fractional programs, without assuming that the optimal solutions set of the considered problem is nonempty, and since the objective functions are variable with respect to the iterations in the auxiliary problems generated by Dinkelbach-type algorithms DT1 and DT2, we consider that the regularizing parameter is also variable. On the other hand we study the convergence when the iterates are only η_k-minimizers of the auxiliary problems. This situation is more general than the one considered in [1]. We also give some results concerning the rate of convergence of these algorithms, and show that it is linear and some times superlinear for some classes of functions. Illustrations by numerical examples are given in [1].

Keywords: Generalized fractional programs, Dinkelbach-type algorithms, proximal point algorithm, rate of convergence.

1. Introduction

The prox-regularization algorithms for generalized fractional programming studied in [1] are two methods based on Dinkelbach-type algorithms introduced in [2,3].

Received March, 2001.

[1] Faculté des Sciences et Techniques, Département de Mathématiques et Informatique, Route de Casablanca, BP. 577, Settat, Morocco; e-mail: roubi@uh1.ac.ma

These algorithms are for solving problems of the form

$$(P) \qquad \inf_{x\in X}\left\{\max_{i\in I}\{f_i(x)/g_i(x)\}\right\}$$

where $I = \{1,\dots,m\}$, $m \geq 1$, and X a nonempty, closed subset of $\mathbb{R}^n$, and the functions f_i and g_i are defined on X, continuous and satisfy $g_i(x) > 0$ for all $x \in X$ and $i \in I$. Next we will use the notation

$$f(x) := \max_{i\in I}\{f_i(x)/g_i(x)\}\cdot$$

The Dinkelbach-type algorithms DT1 and DT2 [2, 3] generalize Dinkelbach algorithm [4] to the case $m > 1$. In these algorithms, at each iteration an auxiliary problem that has in some situations simpler structure than the original problem is solved. The algorithms DT1 and DT2 are based on the same principle but DT2 is generally faster than DT1 (see [2, 3, 5]).

The algorithms DTR1 and DTR2 introduced in [1] combine the last algorithms with the proximal point algorithm [6–9]. These algorithms are useful in some situations and particularly when the auxiliary problems generated by DT1 and DT2 has no solutions. On the other hand, DTR1 and DTR2 extend the proximal point algorithm to a class of nonconvex problems, but under the assumption that the optimal solutions set of (P) is nonempty and with a constant regularizing parameter.

Since regularization is useful in the case of ill-conditioned problems (see for example [10]), and since the objective functions in DT1 and DT2 are variable with respect to the iterations, it is natural to consider a variable regularizing parameter in DTR1 and DTR2. In this paper we analyze the convergence of these algorithms with a variable parameter of regularization, and without the assumption that optimal solutions set is nonempty used in [1].

On the other hand, by refining Lemma 3.5 given in [1], we establish the convergence of these algorithms under other conditions on the approximate solutions of the intermediate problems, and that are weaker in some cases than those used in [1]. We will also analyze the rate of convergence of these algorithms and show that it is linear and superlinear in some cases.

2. Convergence and rate of convergence of algorithm DTR1

We will denote by $\bar\lambda$ the optimal value of (P), and by X^* its optimal solutions set and we will assume in all what follows that:

1) there exists $\gamma > 0$, such that for all $x \in X$ and $i \in I$, $0 < g_i(x) \leq \gamma$;
2) X is convex, and for all $\lambda \geq \bar\lambda$, and $i \in I$, the functions $f_i - \lambda g_i$ are convex.

The last hypothesis is fulfilled for example when the functions f_i and $-g_i$ are convex and $\bar\lambda \geq 0$, or when the functions f_i are convex and the functions g_i are affine.

For $x \in X$ and $\lambda \in \mathbb{R}$, we define the function

$$J(\lambda, x) = \max_{i \in I}\{f_i(x) - \lambda g_i(x)\}\cdot$$

Also, for $\alpha \in \mathbb{R}$, $\lambda \in \mathbb{R}$ and $x \in X$, we define the function

$$G_\alpha(\lambda, x) = \inf_{y \in X}\{J(\lambda, y) + \alpha\|y - x\|^2\}\cdot$$

In the following we describe DTR1; the algorithm DTR2 will be described later.

Algorithm 2.1. *Let $\{\eta_k\}$ be a given sequence of nonnegative reals.*

0. *Choose $x^0 \in X$ and set $\lambda_0 = f(x^0)$.*
1. *At Step k we have x^k and λ_k. Then, find $x^{k+1} \in X$ satisfying*

$$\psi(\lambda_k, x^k, \eta_k) \geq J(\lambda_k, x^{k+1}) + \alpha\|x^{k+1} - x^k\|^2,$$

set $\lambda_{k+1} = f(x^{k+1})$, $k \leftarrow k+1$ and go back to 1.

It is shown in [1], Theorem 3.1, with

$$\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_\alpha(\lambda_k, x^k) + \eta_k\} \quad \text{and} \quad \sum_{k=0}^{\infty}\sqrt{\eta_k} < +\infty,$$

that if the optimal solutions set of (P) is nonempty, then the sequence $\{\lambda_k\}$ converges to $\bar{\lambda}$ and $\{x^k\}$ converges towards a solution of (P). When the function $J(\bar{\lambda}, .)$ is strongly convex, it is also shown that the rate of convergence is linear.

Next, we will consider Algorithm 2.1 with a regularizing parameter α_k at each iteration k, and with the following choices of ψ:

(i) $\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$;

where G_k denotes G_{α_k}, so that $x^{k+1} \in X$ satisfies

$$\psi(\lambda_k, x^k, \eta_k) \geq J(\lambda_k, x^{k+1}) + \alpha_k\|x^{k+1} - x^k\|^2. \tag{1}$$

Notice that only the choice (i) with $\alpha_k = \alpha$, which corresponds to the algorithm DTR1 is considered in [1], and that in (ii), x^{k+1} is an η_k-minimizer of $J(\lambda_k, .)$ $+\alpha_k\|. - x^k\|^2$.

In what follows we will show that the sequence $\{\lambda_k\}$ converges towards $\bar{\lambda}$ even if the optimal solutions set of (P) is empty. The convergence is established, with (i) or (ii), under the usual assumption, $\sum_{k=0}^{\infty} 1/\alpha_k = +\infty$, made for the proximal algorithm (see for example [8]). Then we will show that the rate of convergence is linear for other classes of functions which include the strongly convex case (studied in [1]) and the polyhedral case. Similar results as for the proximal point algorithm are given in Proposition 2.2 and Proposition 2.3.

Now let us consider the following interesting lemma that refine Lemma 3.5 given in [1]. This lemma may be used in different situations and under different forms in this paper, in particular with $\mu_k = 0$.

Lemma 2.1. *Let $\{\mu_k\}$ and $\{\beta_k\}$ be sequences of nonnegative reals such that*

$$\sum_{k\geq 0} \mu_k < +\infty, \quad \sum_{k\geq 0} \beta_k < +\infty,$$

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. For all $k \in \mathbb{N}$ and $p = k+2, k+3, \dots$, we can show that

$$u_p \leq \prod_{i=k}^{p-1}(1+\mu_i)u_k + \sum_{i=k}^{p-2} \prod_{j=i+1}^{p-1} (1+\mu_j)\beta_i + \beta_{p-1}.$$

Since $1+\mu_j \geq 1$, for all $j \in \mathbb{N}$, then $\prod_{j=i+1}^{p-1}(1+\mu_j) \leq \prod_{j=k}^{p-1}(1+\mu_j)$ for all $i = k, k+1, \dots$, and thus we get

$$u_p \leq \prod_{i=k}^{p-1}(1+\mu_i)u_k + \prod_{j=k}^{p-1}(1+\mu_j) \sum_{i=k}^{p-1} \beta_i.$$

It follows that

$$\limsup_{p\to\infty} u_p \leq \prod_{i=k}^{\infty}(1+\mu_i)u_k + \prod_{j=k}^{\infty}(1+\mu_j) \sum_{i=k}^{\infty} \beta_i.$$

Since $\sum_{k\geq 0} \beta_k < \infty$ and $\sum_{k\geq 0} \mu_k < \infty$, then we have $\prod_{j=k}^{\infty}(1+\mu_j) \to 1$ and $\sum_{i=k}^{\infty} \beta_i \to 0$ as $k \to \infty$, which implies that

$$\limsup_{p\to\infty} u_p \leq \liminf_{k\to\infty} u_k,$$

and that the sequence $\{u_k\}$ converges to some $u \in \mathbb{R} \cup \{-\infty\}$. □

Lemma 2.2 is derived from Lemma 2.1; and Lemmas 2.3, 2.4 are based on results in [1], but we reformulate here some proofs since the parameter α is variable in our case. However, these proofs remain close to the ones given in [1]. To facilitate reading of the two papers, we will often use notations used in [1].

In what follows we will denote for $\lambda \in \mathbb{R}$ and $x \in X$,

$$I(\lambda, x) = \{i \in I \mid \lambda = f_i(x)/g_i(x)\}$$

and

$$K(\lambda, x) = \{i \in I \mid J(\lambda, x) = f_i(x) - \lambda g_i(x)\}\cdot$$

Next we will set

$$\bar{\lambda} = \inf_{x \in X} f(x), \quad X^* = \operatorname*{argmin}_{x \in X} f(x), \quad \nu = \inf_{x \in X} \min_{i \in I} g_i(x) \quad \text{and} \quad \tau = \inf_{x \in X^*} \min_{i \in I} g_i(x)$$

2.1. Convergence of algorithm DTR1

For the proof of the main results, we will use Lemmas 2.2–2.5 below.

Lemma 2.2. (a) *The sequence $\{\lambda_k\}$ converges to some $\hat{\lambda} \in \mathbb{R} \cup \{-\infty\}$ if one of the two following cases is realized:*
(i) $\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, $\sum_{k \geq 0} \eta_k < +\infty$ *and* $\nu > 0$.
(b) *If $\bar{\lambda} > -\infty$, then $J(\lambda_k, x^{k+1}) \to 0$ and $\alpha_k \|x^{k+1} - x^k\| \to 0$ as $k \to \infty$.*

Proof. (a) From (1) we have

$$\psi(\lambda_k, x^k, \eta_k) \geq J(\lambda_k, x^{k+1}) + \alpha_k \|x^{k+1} - x^k\|^2 \geq J(\lambda_k, x^{k+1}).$$

On the other hand,

$$J(\lambda_k, x^{k+1}) \geq f_i(x^{k+1}) - \lambda_k g_i(x^{k+1})$$

for all $i \in I$. For $i \in I(\lambda_{k+1}, x^{k+1})$, we have $f_i(x^{k+1}) = \lambda_{k+1} g_i(x^{k+1})$ and then

$$\psi(\lambda_k, x^k, \eta_k) \geq J(\lambda_k, x^{k+1}) \geq g_i(x^{k+1})(\lambda_{k+1} - \lambda_k). \tag{2}$$

In the case (i) we have $\psi(\lambda_k, x^k, \eta_k) \leq 0$, which implies that $\lambda_{k+1} \leq \lambda_k$ for all $k \in \mathbb{N}$, and the conclusion follows.

Now with (ii) we have $\psi(\lambda_k, x^k, \eta_k) \leq \eta_k$ since $G_k(\lambda_k, x^k) \leq 0$. On the other hand let $\operatorname{sg}(\lambda) = 1$ if $\lambda > 0$ and $\operatorname{sg}(\lambda) = -1$ if $\lambda \leq 0$ and let

$$\delta_k = \max\{\nu \operatorname{sg}(\lambda_{k+1} - \lambda_k), -\gamma \operatorname{sg}(\lambda_{k+1} - \lambda_k)\}\cdot$$

Notice that if $\lambda_{k+1} - \lambda_k \leq 0$, then $\delta_k = \gamma$, and $\delta_k = \nu$ if $\lambda_{k+1} - \lambda_k > 0$, implying that

$$\eta_k \geq J(\lambda_k, x^{k+1}) \geq \delta_k(\lambda_{k+1} - \lambda_k). \tag{3}$$

Then it follows that

$$\eta_k / \delta_k + \lambda_k \geq \lambda_{k+1}. \tag{4}$$

Remarking that $\delta_k \geq \nu$, the result follows by using Lemma 2.1 with $u_k = \lambda_k$, $\mu_k = 0$ and $\beta_k = \eta_k/\delta_k$.

(b) If $\bar{\lambda} > -\infty$, then the sequence $\{\lambda_k\}$ converges to some $\hat{\lambda} \geq \bar{\lambda}$. Following (3), $J(\lambda_k, x^{k+1}) \to 0$ as $k \to \infty$, and since

$$\eta_k \geq J(\lambda_k, x^{k+1}) + \alpha_k \|x^{k+1} - x^k\|^2 \geq \delta_k(\lambda_{k+1} - \lambda_k), \tag{5}$$

then $\alpha_k \|x^{k+1} - x^k\| \to 0$ as $k \to \infty$. □

Lemma 2.3. *For all $k \in \mathbb{N}$, let $\bar{x}^{k+1} = \operatorname{argmin}_{x \in X}\{J(\lambda_k, x) + \alpha_k \|x - x^k\|^2\}$ and x^{k+1} be defined by (1). Then we have*

$$\eta_k/\alpha_k \geq \|\bar{x}^{k+1} - x^{k+1}\|^2.$$

Proof. From the definition of $\bar{x}^{k+1}$ we have

$$J(\lambda_k, x) \geq J(\lambda_k, \bar{x}^{k+1}) + 2\alpha_k \left\langle x^k - \bar{x}^{k+1}, x - \bar{x}^{k+1} \right\rangle,$$

for all $x \in X$ (see for example [11], Prop. 2.2, p. 37). Using the equality

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

we obtain

$$J(\lambda_k, x) \geq J(\lambda_k, \bar{x}^{k+1}) + \alpha_k \Big(\|\bar{x}^{k+1} - x^k\|^2 + \|x - \bar{x}^{k+1}\|^2 - \|x - x^k\|^2 \Big). \tag{6}$$

On the other hand, relation (1) implies that

$$J(\lambda_k, x) + \alpha_k \|x - x^k\|^2 + \eta_k \geq J(\lambda_k, x^{k+1}) + \alpha_k \|x^{k+1} - x^k\|^2,$$

for all $x \in X$. For $x = \bar{x}^{k+1}$, we get

$$J(\lambda_k, \bar{x}^{k+1}) \geq J(\lambda_k, x^{k+1}) + \alpha_k \|x^{k+1} - x^k\|^2 - \alpha_k \|\bar{x}^{k+1} - x^k\|^2 - \eta_k.$$

Considering relation (6) we obtain

$$J(\lambda_k, x) \geq J(\lambda_k, x^{k+1}) + \alpha_k \Big(\|x^{k+1} - x^k\|^2 + \|x - \bar{x}^{k+1}\|^2 - \|x - x^k\|^2 \Big) - \eta_k, \tag{7}$$

for all $x \in X$. Thus, remplacing x by x^{k+1} in the last inequality we obtain

$$\eta_k/\alpha_k \geq \|x^{k+1} - \bar{x}^{k+1}\|^2.$$

□

Lemma 2.4. *For all $k = 1, 2, \dots,$ and for all $x \in \mathbb{R}^n$, we have*

$$\|x - x^k\|^2 \leq \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|x - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1}.$$

Proof. For every $k = 1, 2, \dots,$ and for every $x \in \mathbb{R}^n$ we have

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

Since $\left\langle x - \bar{x}^k, \bar{x}^k - x^k\right\rangle \leq \|x - \bar{x}^k\|\|\bar{x}^k - x^k\|$ and $\|\bar{x}^k - x^k\| \leq \sqrt{\eta_{k-1}/\alpha_{k-1}}$ (from Lem. 2.3), then we obtain

$$\|x - x^k\|^2 \leq \|x - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\|x - \bar{x}^k\| + \eta_{k-1}/\alpha_{k-1}.$$

Remarking that $\|x - \bar{x}^k\| \leq 1 + \|x - \bar{x}^k\|^2$ we deduce the result. □

Lemma 2.5. *If $\lambda \geq \bar{\lambda}$ is such that $J(\lambda, x) \geq 0$ for all $x \in X$, then $\lambda = \bar{\lambda}$.*

Proof. For $x \in X$, we have $J(\lambda, x) \geq 0$, if and only if, there exists $i \in I$ such that $f_i(x) - \lambda g_i(x) \geq 0$. It follows that $f_i(x)/g_i(x) \geq \lambda$. But $f(x) \geq f_i(x)/g_i(x)$ and thus, $f(x) \geq \lambda$. This is true for all $x \in X$ and so, $\bar{\lambda} \geq \lambda$. Therefore, equality holds. □

Theorem 2.1. *Suppose that*

$$\sum_{k\geq 0} 1/\alpha_k = +\infty \text{ and that } \sum_{k\geq 0} \sqrt{\eta_k/\alpha_k} < +\infty,$$

and consider Algorithm 2.1 with one of the two following choices:

(i) $\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, $\sum_{k\geq 0} \eta_k < +\infty$ *and* $\nu > 0$.

Then the sequence $\{\lambda_k\}$ converges to the optimal value $\bar{\lambda}$ of (P).

Proof. Following Lemma 2.2(a), the sequence $\{\lambda_k\}$ converges to $\hat{\lambda} \in \mathbb{R} \cup \{-\infty\}$. Then, $\lambda_k \geq \hat{\lambda} \geq \bar{\lambda}$ since $x^k \in X$. If $\hat{\lambda} = -\infty$, then $\hat{\lambda} = \bar{\lambda}$. So, suppose that $\hat{\lambda} > -\infty$. From (7), for all $k \in \mathbb{N}$, and for all $x \in X$ we have

$$J(\lambda_k, x) - J(\lambda_k, x^{k+1}) + \alpha_k\|x - x^k\|^2 + \eta_k \geq \alpha_k\|x - \bar{x}^{k+1}\|^2.$$

Using Lemma 2.4 in this inequality we obtain

$$\begin{aligned}\|x - \bar{x}^{k+1}\|^2 \leq &\left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|x - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ &+\eta_k/\alpha_k + \left(J(\lambda_k, x) - J(\lambda_k, x^{k+1})\right)/\alpha_k. \end{aligned} \quad (8)$$

Let us show that for all $x \in X$,

$$\lim_{k\to\infty} \Big(J(\lambda_k, x) - J(\lambda_k, x^{k+1})\Big) \geq 0. \tag{9}$$

For this, assume the contrary. So, there exists $\varepsilon > 0$, $\tilde{x} \in X$ and l such that

$$J(\lambda_k, x) - J(\lambda_k, x^{k+1}) < -\varepsilon$$

for all $k \geq l$. Thus, substituting $\tilde{x}$ to x in (8) yields

$$\begin{aligned}\|\tilde{x} - \bar{x}^{k+1}\|^2 \leq \Big(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\Big)\|\tilde{x} - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} \\ +\eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k - \varepsilon/\alpha_k. \end{aligned}\tag{10}$$

From the assumptions of the theorem,

$$\sum_{k\geq 1} \sqrt{\eta_{k-1}/\alpha_{k-1}} < +\infty \text{ and } \sum_{k\geq 1} \Big(\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k\Big) < +\infty, \tag{11}$$

and thus Lemma 2.1 implies that the sequence $\{\bar{x}^k\}$ converges. Consequently,

$$\sum_{k\geq 1} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\tilde{x} - \bar{x}^k\|^2 < +\infty. \tag{12}$$

Summing in (10) over $k = l, \dots, n$, we obtain

$$\begin{aligned}\|\tilde{x} - \bar{x}^{n+1}\|^2 - \|\tilde{x} - x^l\|^2 \leq 2\sum_{k=l}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\tilde{x} - \bar{x}^k\|^2 + 2\sum_{k=l}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}} \\ + \sum_{k=l}^{n} \Big(\eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k\Big) - \varepsilon \sum_{k=l}^{n} 1/\alpha_k. \end{aligned}\tag{13}$$

But since $\sum_{k=l}^{n} 1/\alpha_k \to +\infty$ as $n \to \infty$, the inequality (13) cannot hold. It follows that (9) must hold. Consequently, $J(\hat{\lambda}, x) \geq 0$ for all $x \in X$. Since $\hat{\lambda} \geq \bar{\lambda}$, Lemma 2.5 implies that $\hat{\lambda} = \bar{\lambda}$. □

Proposition 2.1. *Assume that the sequence $\{\alpha_k\}$ is such that $\alpha_k \geq \bar{\alpha} > 0$ for all $k \in \mathbb{N}$, that the assumptions of Theorem 2.1 are fulfilled, and that the optimal solutions set of (P) is nonempty. If Algorithm 2.1 is considered in one of the following situations:*

(i) $\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;

(ii) $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, $\sum_{k\geq 0} \eta_k < +\infty$, $\nu > 0$, $\gamma \leq \nu + \tau$, *and* $\alpha_{k+1} \geq \alpha_k$ *for all* $k \in \mathbb{N}$,

then the sequence $\{x^k\}$ converges to some solution of (P). □

Proof. Observe that under these assumptions the sequence $\{\lambda_k\}$ converges to $\bar{\lambda}$. Assume that (i) occurs. For $\bar{x} \in X^*$, $J(\lambda_k, \bar{x}) \leq 0$. On the other hand, since $\psi(\lambda_k, x^k, \eta_k) \leq 0$ then from (2), $\lambda_{k+1} - \lambda_k \leq 0$ and thus from (3) we have $J(\lambda_k, x^{k+1}) \geq \gamma(\lambda_{k+1} - \lambda_k)$. Using this and (8) with $x = \bar{x}$, we obtain

$$\begin{aligned}\|\bar{x} - \bar{x}^{k+1}\|^2 \leq{} & \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x} - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ & + \eta_k/\alpha_k + (\lambda_k - \lambda_{k+1})\gamma/\bar{\alpha}.\end{aligned}$$

Following Lemma 2.1, the sequence $\{\bar{x}^k\}$ converges. Let $\tilde{x} \in X$ be its limit. Since f is continuous, then $\hat{\lambda} = f(\tilde{x})$ and from Theorem 2.1, $\tilde{x}$ is a solution of (P). From Lemma 2.3 and the fact that $\eta_k/\alpha_k \to 0$ as $k \to \infty$, we deduce that $\{x^k\}$ also converges to $\tilde{x}$.

Suppose now that (ii) occurs. For $\bar{x} \in X^*$ we have $J(\lambda_k, \bar{x}) \leq \tau(\bar{\lambda} - \lambda_k)$, and on the other hand

$$J(\lambda_k, x^{k+1}) \geq \delta_k(\lambda_{k+1} - \lambda_k) \geq \nu(\lambda_{k+1} - \bar{\lambda}) + \gamma(\bar{\lambda} - \lambda_k).$$

Considering (8) with these inequalities we get

$$\begin{aligned}\|\bar{x} - \bar{x}^{k+1}\|^2 \leq{} & \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x} - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ & + \eta_k/\alpha_k + (\gamma - \tau)(\lambda_k - \bar{\lambda})/\alpha_k - \nu(\lambda_{k+1} - \bar{\lambda})/\alpha_k \\ \leq{} & \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x} - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ & + \eta_k/\alpha_k + \nu(\lambda_k - \bar{\lambda})/\alpha_k - \nu(\lambda_{k+1} - \bar{\lambda})/\alpha_{k+1} \\ & + \nu(1/\alpha_{k+1} - 1/\alpha_k)(\lambda_{k+1} - \bar{\lambda}).\end{aligned}$$

Notice that we used in the last inequality the assumption $\gamma \leq \nu + \tau$. Then from the assumption that $\alpha_{k+1} \geq \alpha_k$, it follows that

$$\begin{aligned}\|\bar{x} - \bar{x}^{k+1}\|^2 + \nu(\lambda_{k+1} - \bar{\lambda})/\alpha_{k+1} \leq{} & \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x} - \bar{x}^k\|^2 \\ & + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ & + \eta_k/\alpha_k + \nu(\lambda_k - \bar{\lambda})/\alpha_k \\ \leq{} & \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\left(\|\bar{x} - \bar{x}^k\|^2 + \nu(\lambda_k - \bar{\lambda})\right) \\ & + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k.\end{aligned}$$

Applying Lemma 2.1 with $u_k = \|\bar{x} - \bar{x}^k\|^2 + \nu(\lambda_k - \bar{\lambda})$, $\mu_k = 2\sqrt{\eta_{k-1}/\alpha_{k-1}}$ and $\beta_k = 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k$, we deduce that the sequence $\{\bar{x}^k\}$ converges. The rest is as in the previous case. □

2.2. Rate of convergence of algorithm DTR1

The rate of convergence of DTR1 was analyzed in [1] when the function $J(\bar{\lambda}, .)$ is strongly convex. Next, we will see that the results about the rate of convergence remain still valid for other classes of functions which include the strongly convex case and the polyhedral case.

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

$$(H)\quad \exists\delta > 0, \exists\kappa > 0 \quad \text{such that} \quad J(\bar{\lambda}, x) \geq \kappa d(x, X^*)^2 \text{ for all } x \in B(X^*, \delta) \cap X,$$

where $B(X^*, \delta) = \bigcup_{\bar{x}\in X^*} B(\bar{x}, \delta)$, $B(x, \delta) = \{z \in \mathbb{R}^n \mid \|z - x\| < \delta\}$ and $d(x, X^*) = \inf_{\bar{x}\in X^*} \|x - \bar{x}\|$.

Remark 2.1. 1) Assumption (H) is satisfied when the function $J(\bar{\lambda}, .)$ is polyhedral (for example when the functions f_i are polyhedral and the functions g_i are affine) and X is polyhedral. Indeed, following ([12], Th. 3.5 and Cor. 3.6), $J(\bar{\lambda}, .)$ satisfies the property

$$(H')\qquad \exists\kappa > 0, \quad \text{such that} \quad J(\bar{\lambda}, x) \geq \kappa d(x, X^*), \text{ for all } x \in X,$$

since $X^* = \operatorname{argmin}_{x\in X} J(\bar{\lambda}, x)$ and $J(\bar{\lambda}, \bar{x}) = 0$ (see also [13] and some references therein for some characterizations of such functions). Therefore, for $0 < \delta \leq 1$, and $x \in B(X^*, \delta) \cap X$, we have $d(x, X^*) \leq 1$ and thus $d(x, X^*) \geq d(x, X^*)^2$. It follows that $J(\bar{\lambda}, x) \geq \kappa d(x, X^*)^2$ for all $x \in B(X^*, \delta) \cap X$ and the assumption (H) is satisfied.

2) If the function $J(\bar{\lambda}, .)$ is strongly convex, the assumption (H) is also satisfied. Indeed, the strong convexity assumption implies that there exists $\kappa > 0$ such that

$$J(\bar{\lambda}, x) \geq J(\bar{\lambda}, \bar{x}) + \langle \bar{x}^*, x - \bar{x}\rangle + \kappa\|x - \bar{x}\|^2,$$

for all $x, \bar{x} \in X$ and $\bar{x}^* \in \partial J(\bar{\lambda}, \bar{x})$ where $\partial J(\bar{\lambda}, \bar{x})$ is the subdifferential of $J(\bar{\lambda}, .)$ at $\bar{x}$. For $\bar{x} \in X^* = \{\bar{x}\}$, we have $0 \in \partial J(\bar{\lambda}, \bar{x})$ and $J(\bar{\lambda}, \bar{x}) = 0$. It follows that

$$J(\bar{\lambda}, x) \geq \kappa d(x, X^*)^2 = \kappa\|x - \bar{x}\|^2$$

for all $x \in X$, and the assumption (H) is fulfilled.

Theorem 2.2. *Suppose that the optimal solutions set X^* of (P) is nonempty and that the function $J(\bar{\lambda}, .)$ satisfies the assumption (H) above. Assume on the other hand, that $\eta_k/(\lambda_k - \bar{\lambda}) \to 0$ as $k \to \infty$, that $\tau > 0$ (this condition is fulfilled for example when X^* is compact), and that the sequence $\{x^k\}$ converges to a minimizer of (P) (this is the case for example when the conditions of Prop. 2.1 are fulfilled). If Algorithm 2.1 is considered in one of the two following situations:*

(i) $\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$; *$\nu > 0$ and $\gamma < \nu + \tau$;*

then for α_k sufficiently small, the sequence $\{\lambda_k\}$ converges linearly to $\bar{\lambda}$.

Proof. From the definition of x^{k+1} we have for all $x \in X$,

$$J(\lambda_k, x) + \alpha_k \|x - x^k\|^2 + \eta_k \geq J(\lambda_k, x^{k+1}). \tag{14}$$

Since $\{x^k\}$ converges to a solution of (P), then $x^k \in B(X^*, \delta) \cap X$ for k large enough. Let $\tilde{x}^k \in X^*$ be such that $\|x^k - \tilde{x}^k\| = d(x^k, X^*)$. Then we have

$$J(\bar{\lambda}, x^k) \geq \kappa \|x^k - \tilde{x}^k\|^2.$$

For $i \in K(\bar{\lambda}, x^k)$ we have $J(\bar{\lambda}, x^k) = f_i(x^k) - \bar{\lambda} g_i(x^k)$, and then

$$J(\bar{\lambda}, x^k) \leq g_i(x^k)(\lambda_k - \bar{\lambda}) \leq \gamma(\lambda_k - \bar{\lambda}).$$

For $i \in K(\lambda_k, \tilde{x}^k)$ we get $J(\lambda_k, \tilde{x}^k) = f_i(\tilde{x}^k) - \lambda_k g_i(\tilde{x}^k)$, and thus

$$J(\lambda_k, \tilde{x}^k) \leq g_i(\tilde{x}^k)(\bar{\lambda} - \lambda_k) \leq \tau(\bar{\lambda} - \lambda_k).$$

On the other hand, $J(\lambda_k, x^{k+1}) \geq \delta_k(\lambda_{k+1} - \lambda_k)$. Then, with $x = \tilde{x}^k$ in (14) we obtain

$$\tau(\bar{\lambda} - \lambda_k) + \gamma(\lambda_k - \bar{\lambda})\alpha_k/\kappa + \eta_k \geq \delta_k(\lambda_{k+1} - \lambda_k). \tag{15}$$

With the choice (i) in Algorithm 2.1 we have $\lambda_{k+1} - \lambda_k \leq 0$ and $\delta_k = \gamma$. It follows that

$$(1 - \tau/\gamma + \alpha_k/\kappa)(\lambda_k - \bar{\lambda}) + \eta_k/\gamma \geq \lambda_{k+1} - \bar{\lambda}.$$

So, if $\limsup_{k\to\infty} \alpha_k < \kappa\tau/\gamma$, then $\limsup_{k\to\infty} (\lambda_{k+1} - \bar{\lambda})/(\lambda_k - \bar{\lambda}) < 1$.

Now with the choice (ii) we have $\nu > 0$. By writing

$$\delta_k(\lambda_{k+1} - \lambda_k) = \delta_k(\lambda_{k+1} - \bar{\lambda}) + \delta_k(\bar{\lambda} - \lambda_k) \geq \nu(\lambda_{k+1} - \bar{\lambda}) + \gamma(\bar{\lambda} - \lambda_k)$$

and considering (15) we get

$$[\gamma/\nu - \tau/\nu + \alpha_k\gamma/(\kappa\nu)](\lambda_k - \bar{\lambda}) + \eta_k/\nu \geq \lambda_{k+1} - \bar{\lambda}.$$

Thus, if $\limsup_{k\to\infty} \alpha_k < \kappa(\nu + \tau - \gamma)/\gamma$, then $\limsup_{k\to\infty}(\lambda_{k+1} - \bar{\lambda})/(\lambda_k - \bar{\lambda}) < 1$. □

Proposition 2.2. *For all $n \in \mathbb{N}$, let $\sigma_n = \sum_{k=0}^{n} 1/\alpha_k$. If the optimal solutions set X^* of (P) is nonempty and $\psi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$, then for all $\bar{x} \in X^*$ we have the following estimate*

$$\tau(\lambda_{n+1} - \bar{\lambda}) \leq \sigma_n^{-1}\Big[\sum_{k=1}^{n}\big(\eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k + 2(1 + \|\bar{x} - \bar{x}^k\|^2)\sqrt{\eta_{k-1}/\alpha_{k-1}}\big)$$
$$+ \|\bar{x} - \bar{x}^1\|^2 + \tau(\lambda_0 - \bar{\lambda})/\alpha_0 + (\gamma - \tau)\sum_{k=1}^{n}(\lambda_k - \lambda_{k+1})/\alpha_k\Big].$$

Proof. With $x = \bar{x} \in X^*$ in (8), we have

$$\|\bar{x} - \bar{x}^{k+1}\|^2 \leq \Big(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\Big)\|\bar{x} - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1}$$
$$+\eta_k/\alpha_k + \Big(J(\lambda_k, \bar{x}) - J(\lambda_k, x^{k+1})\Big)/\alpha_k.$$

By summing in this inequality over $k = 1, 2, \dots, n$, we get

$$\begin{aligned}\|\bar{x} - \bar{x}^{n+1}\|^2 - \|\bar{x} - \bar{x}^1\|^2 \leq\ & 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\bar{x} - \bar{x}^k\|^2 + 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}} \\ & + \sum_{k=1}^{n} \eta_{k-1}/\alpha_{k-1} + \sum_{k=1}^{n} \eta_k/\alpha_k \\ & + \sum_{k=1}^{n} [J(\lambda_k, \bar{x}) - J(\lambda_k, x^{k+1})]/\alpha_k. \qquad (16)\end{aligned}$$

By considering the inequalities $J(\lambda_k, \bar{x}) \leq \tau(\bar{\lambda} - \lambda_k)$ and $J(\lambda_k, x^{k+1}) \geq \gamma(\lambda_{k+1} - \lambda_k)$ we get

$$(\gamma - \tau)(\lambda_k - \bar{\lambda}) - \gamma(\lambda_{k+1} - \bar{\lambda}) \geq J(\lambda_k, \bar{x}) - J(\lambda_k, x^{k+1}),$$

and thus by setting $u_k = \lambda_k - \bar{\lambda}$, we obtain

$$\begin{aligned}\|\bar{x} - \bar{x}^{n+1}\|^2 - \|\bar{x} - \bar{x}^1\|^2 \leq\ & 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\bar{x} - \bar{x}^k\|^2 + 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}} \\ & + \sum_{k=1}^{n} \eta_{k-1}/\alpha_{k-1} + \sum_{k=1}^{n} \eta_k/\alpha_k \\ & + (\gamma - \tau)\sum_{k=1}^{n} u_k/\alpha_k - \gamma\sum_{k=1}^{n} u_{k+1}/\alpha_k. \qquad (17)\end{aligned}$$

On the other hand, we have $\lambda_{k+1} - \lambda_k \leq 0$, that is $u_k - u_{k+1} \geq 0$, and thus by writing $\sigma_{k-1} = \sigma_k - 1/\alpha_k$, we get

$$0 \leq \sigma_{k-1}u_k - \sigma_k u_{k+1} + u_{k+1}/\alpha_k.$$

By summing over $k = 1, 2, \dots, n$, we get

$$0 \leq \sum_{k=1}^{n} u_{k+1}/\alpha_k + \sigma_0 u_1 - \sigma_n u_{n+1},$$

where $\sigma_0 = 1/\alpha_0$. Now multiplying this inequality by τ and summing together with (17) yields

$$\begin{aligned}\|\bar{x} - x^{n+1}\|^2 - \|\bar{x} - \bar{x}^1\|^2 \leq\ & 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\bar{x} - \bar{x}^k\|^2 + 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}} \\ & + \sum_{k=1}^{n} \eta_{k-1}/\alpha_{k-1} + \sum_{k=1}^{n} \eta_k/\alpha_k \\ & + (\gamma - \tau)\sum_{k=1}^{n} (u_k - u_{k+1})/\alpha_k + \tau u_1/\alpha_0 - \tau\sigma_n u_{n+1},\end{aligned}$$

which gives the desired result. □

Proposition 2.3. *For all $n \in \mathbb{N}$, let $\sigma_n = \sum_{k=0}^{n} 1/\alpha_k$. If the optimal solutions set X^* of (P) is nonempty and $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, then for all $\bar{x} \in X^*$ we have the following estimate*

$$\begin{aligned}\nu(\lambda_{n+1} - \bar{\lambda}) \leq \sigma_n^{-1}\Big[& \sum_{k=1}^{n} \big(\eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k + 2(1 + \|\bar{x} - \bar{x}^k\|^2) \\ & \times \sqrt{\eta_{k-1}/\alpha_{k-1}} + \sigma_{k-1}\eta_k\big) + \|\bar{x} - \bar{x}^1\|^2 + \nu(\lambda_0 - \bar{\lambda})/\alpha_0 \\ & + (\gamma - \tau)\sum_{k=1}^{n} (\lambda_k - \bar{\lambda})/\alpha_k\Big].\end{aligned}$$

Proof. Since $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$ then we have $J(\lambda_k, x^{k+1}) \geq \delta_k(\lambda_{k+1} - \lambda_k) \geq \nu(\lambda_{k+1} - \bar{\lambda}) - \gamma(\lambda_k - \bar{\lambda})$. By considering this inequality and the inequalities $J(\lambda_k, \bar{x}) \leq \tau(\bar{\lambda} - \lambda_k)$ we get

$$(\gamma - \tau)(\lambda_k - \bar{\lambda}) - \nu(\lambda_{k+1} - \bar{\lambda}) \geq J(\lambda_k, \bar{x}) - J(\lambda_k, x^{k+1}),$$

and thus by setting $u_k = \lambda_k - \bar{\lambda}$ and considering the last inequality in (16), we obtain

$$\begin{aligned}\|\bar{x} - \bar{x}^{n+1}\|^2 - \|\bar{x} - \bar{x}^1\|^2 \leq\ & 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\bar{x} - \bar{x}^k\|^2 + 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}} \\ & + \sum_{k=1}^{n} \eta_{k-1}/\alpha_{k-1} + \sum_{k=1}^{n} \eta_k/\alpha_k \\ & + (\gamma - \tau)\sum_{k=1}^{n} u_k/\alpha_k - \nu\sum_{k=1}^{n} u_{k+1}/\alpha_k. \qquad (18)\end{aligned}$$

Remember that since $\psi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$ we have $\eta_k \geq \delta_k(\lambda_{k+1} - \lambda_k)$, and $\eta_k \geq \nu(\lambda_{k+1} - \lambda_k)$ since $\eta_k \geq 0$ and $\delta_k \geq \nu$. It follows that $\nu u_k - \nu u_{k+1} + \eta_k \geq 0$, and thus by writing $\sigma_{k-1} = \sigma_k - 1/\alpha_k$, we get

$$0 \leq \nu\sigma_{k-1}u_k - \nu\sigma_k u_{k+1} + \nu u_{k+1}/\alpha_k + \sigma_{k-1}\eta_k.$$

By summing over $k = 1, 2, \ldots, n$, we get

$$0 \leq \nu\sigma_0 u_1 - \nu\sigma_n u_{n+1} + \nu\sum_{k=1}^{n} u_{k+1}/\alpha_k + \sum_{k=1}^{n} \sigma_{k-1}\eta_k.$$

By adding this inequality to (18) we obtain

$$\begin{aligned}
\|\bar{x} - x^{n+1}\|^2 - \|\bar{x} - \bar{x}^1\|^2 \leq\ & 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}}\|\bar{x} - \bar{x}^k\|^2 + 2\sum_{k=1}^{n} \sqrt{\eta_{k-1}/\alpha_{k-1}} \\
& + \sum_{k=1}^{n} \eta_{k-1}/\alpha_{k-1} + \sum_{k=1}^{n} \eta_k/\alpha_k + \sum_{k=1}^{n} \sigma_{k-1}\eta_k \\
& + (\gamma - \tau)\sum_{k=1}^{n} u_k/\alpha_k + \nu u_1/\alpha_0 - \nu\sigma_n u_{n+1},
\end{aligned}$$

which gives the desired result. □

Remark 2.2. When $m = 1$ and $g_1 = 1$, that is when f is convex, then $\gamma = \tau = \nu = 1$ and DTR1 coincides with the proximal point algorithm. For this last algorithm, with the conditions of Proposition 2.2, we obtain

$$\begin{aligned}
\lambda_{n+1} - \bar{\lambda} \leq \sigma_n^{-1}\Big[& \sum_{k=1}^{n} \big(\eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k + 2(1 + \|\bar{x} - \bar{x}^k\|^2)\sqrt{\eta_{k-1}/\alpha_{k-1}}\big) \\
& + \|\bar{x} - \bar{x}^1\|^2 + (\lambda_0 - \bar{\lambda})/\alpha_0\Big],
\end{aligned}$$

and with the conditions of Proposition 2.3 we have

$$\begin{aligned}
\lambda_{n+1} - \bar{\lambda} \leq \sigma_n^{-1}\Big[& \sum_{k=1}^{n} \big(\eta_{k-1}/\alpha_{k-1} + \eta_k/\alpha_k + 2(1 + \|\bar{x} - \bar{x}^k\|^2)\sqrt{\eta_{k-1}/\alpha_{k-1}} \\
& + \sigma_{k-1}\eta_k\big) + \|\bar{x} - \bar{x}^1\|^2 + (\lambda_0 - \bar{\lambda})/\alpha_0\Big].
\end{aligned}$$

In particular, when $\eta_k = 0$ we find the estimate

$$\lambda_{n+1} - \bar{\lambda} \leq \sigma_n^{-1}\big[\|\bar{x} - \bar{x}^1\|^2 + (\lambda_0 - \bar{\lambda})/\alpha_0\big].$$

3. Convergence and rate of convergence of algorithm DTR2

The algorithm DTR2 is based on the algorithm DT2 introduced in [3]. In the algorithm DTR2, the function $J(.,.)$ is replaced by the function

$$J(\lambda, x, y) = \max_{i \in I}\{(f_i(x) - \lambda g_i(x))/g_i(y)\},$$

and $G(.,.)$ is replaced by

$$G(\lambda, x) = \inf_{y \in X}\{J(\lambda, y, x) + \alpha\|y - x\|^2\},$$

where α is a given positive real.

Algorithm 3.1. *Let $\{\eta_k\}$ be a given sequence of positive reals.*

0. *Choose $x^0 \in X$ and set $\lambda_0 = f(x^0)$;*
1. *At Step k we have x^k and λ_k. Then, find x^{k+1} satisfying*

$$\varphi(\lambda_k, x^k, \eta_k) \geq J(\lambda_k, x^{k+1}, x^k) + \alpha\|x^{k+1} - x^k\|^2,$$

set $\lambda_{k+1} = f(x^{k+1})$, $k \leftarrow k + 1$ and go back to 1.

The algorithm DTR2 considered in [1] corresponds to the case

$$\varphi(\lambda_k, x^k, \eta_k) = \min\{0, G(\lambda_k, x^k) + \eta_k\}, \quad \text{with} \quad \sum_{k=0}^{\infty}\sqrt{\eta_k} < +\infty.$$

With this choice and when the regularizing parameter is constant, convergence of the sequence $\{\lambda_k\}$ may be established under the following weaker choice of φ:

$$\varphi(\lambda_k, x^k, \eta_k) = G(\lambda_k, x^k) + \eta_k.$$

As was done for Algorithm 2.1, we will consider that the regularizing parameter α is variable with respect to the iterations since the objective function in DT2 is also variable with respect to the iterations; and we will study Algorithm 3.1 under two choices of $\varphi(\lambda_k, x^k, \eta_k)$. We will establish in particular, that the algorithm DTR2 converges without assuming that the optimal solutions s of (P) is nonempty. Also we will establish convergence under other conditions on the approximate solutions of the auxiliary problems. Later, we will show that the rate of convergence is linear when the assumption (H), introduced in Section 2.2, is satisfied. Under some additional conditions we establish superlinear convergence.

For given $\{\alpha_k\}$ and $\{\eta_k\}$ as for Algorithm 2.1, we define the point x^{k+1} as a point in X satisfying

$$\varphi(\lambda_k, x^k, \eta_k) \geq J(\lambda_k, x^{k+1}, x^k) + \alpha_k\|x^{k+1} - x^k\|^2,$$

and consider Algorithm 3.1 with this definition in the two following cases:
(i) $\varphi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\varphi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$,
where

$$G_k(\lambda_k, x^k) = \inf_{x\in X}\{J(\lambda_k, x, x^k) + \alpha_k \|x - x^k\|^2\}\cdot$$

Notice that only case (i) is studied in [1] and that with $\varphi(\lambda_k, x^k, \eta_k)$ as in (ii), x^{k+1} is an η_k-minimizer of $J(\lambda_k, ., x^k) + \alpha_k\|. - x^k\|$.

3.1. Convergence of algorithm DTR2

In the rest of the paper, we assume that our basic assumptions given in Section 2 are fulfilled. In the proofs of convergence of Algorithm 3.1, we will often use the previous results. Also we use the same notations

$$\bar\lambda = \inf_{x\in X} f(x), \quad X^* = \operatorname*{argmin}_{x\in X} f(x), \quad \nu = \inf_{x\in X}\min_{i\in I} g_i(x) \text{ and } \tau = \inf_{x\in X^*}\min_{i\in I} g_i(x),$$

and we will assume in all what follows that $\nu > 0$. Next, we will use the notations

$$I(\lambda, x) = \{i \in I \mid \lambda = f_i(x)/g_i(x)\}$$

and

$$K(\lambda, x, y) = \{i \in I \mid J(\lambda, x, y) = (f_i(x) - \lambda g_i(x))/g_i(y)\}\cdot$$

Lemma 3.1. *If one of the two following cases is realized in Algorithm 3.1*
(i) $\varphi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\varphi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, $\sum_{k\geq 0}\eta_k < +\infty$;
then the sequence $\{\lambda_k\}$ *converges to some* $\hat\lambda \in \mathbb{R}\cup\{-\infty\}$.

Proof. For all $i \in I$ we get

$$\varphi(\lambda_k, x^k, \nu_k) \geq J_k(\lambda_k, x^{k+1}) \geq (f_i(x^{k+1}) - \lambda_k g_i(x^{k+1}))/g_i(x^k).$$

For $i \in I(\lambda_{k+1}, x^{k+1})$ we have $\lambda_{k+1} = f_i(x^{k+1})/g_i(x^{k+1})$ and thus,

$$\varphi(\lambda_k, x^k, \nu_k) \geq J_k(\lambda_k, x^{k+1}) \geq (\lambda_{k+1} - \lambda_k) g_i(x^{k+1})/g_i(x^k). \tag{19}$$

Let $\mathrm{sg}(\lambda) = 1$ if $\lambda > 0$ and $\mathrm{sg}(\lambda) = -1$ if $\lambda \leq 0$ and let

$$\delta'_k = \max\{-\gamma/\nu\,\mathrm{sg}(\lambda_{k+1} - \lambda_k), \nu/\gamma\,\mathrm{sg}(\lambda_{k+1} - \lambda_k)\}\cdot$$

Then $\delta'_k = \gamma/\nu$ if $\lambda_{k+1} - \lambda_k \leq 0$ and $\delta'_k = \nu/\gamma$ if $\lambda_{k+1} - \lambda_k > 0$, which implies that

$$\varphi(\lambda_k, x^k, \nu_k) \geq J(\lambda_k, x^{k+1}) \geq \delta'_k(\lambda_{k+1} - \lambda_k). \tag{20}$$

With (i) we have $\varphi(\lambda_k, x^k, \nu_k) \leq 0$ and it follows that $\lambda_{k+1} \leq \lambda_k$ for all $k \in \mathbb{N}$, and $\{\lambda_k\}$ converges to $\hat{\lambda} \in \mathbb{R} \cup \{-\infty\}$.

With (ii) we have $\varphi(\lambda_k, x^k, \nu_k) \leq \eta_k$ since $G_k(\lambda_k, x^k) \leq 0$ and (20) implies that

$$\eta_k/\delta'_k + \lambda_k \geq \lambda_{k+1}. \tag{21}$$

Notice that $0 < \nu \leq \delta'_k$ and thus the result follows by using Lemma 2.1 with $u_k = \lambda_k$, $\mu_k = 0$ and $\beta_k = \eta_k/\delta'_k$.

Theorem 3.1. *Suppose that*

$$\sum_{k \geq 0} 1/\alpha_k = +\infty \text{ and that } \sum_{k \geq 0} \sqrt{\eta_k/\alpha_k} < +\infty.$$

If one of the two following conditions is satisfied

(i) $\varphi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;

(ii) $\varphi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, $\sum_{k \geq 0} \eta_k < +\infty$;

then the sequence $\{\lambda_k\}$ converges to the optimal value $\bar{\lambda}$ of (P).

Proof. For all $k \in \mathbb{N}$ we set $J(.,., x^k) = J_k(.,.)$. As in the proof of Theorem 2.1, we will have with the function J_k,

$$\begin{aligned} \|x - \bar{x}^{k+1}\|^2 \leq \Big(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\Big)\|x - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ + \eta_k/\alpha_k + \Big(J_k(\lambda_k, x) - J_k(\lambda_k, x^{k+1})\Big)/\alpha_k, \end{aligned} \tag{22}$$

where $\bar{x}^{k+1} = \operatorname{argmin}_{x \in X}\{J_k(\lambda_k, x) + \alpha_k\|x - x^k\|^2\}$. Notice that under the assumptions of the theorem, Lemma 3.1 implies that the sequence $\{\lambda_k\}$ converges to some $\hat{\lambda}$. It remains to show that $\hat{\lambda} = \bar{\lambda}$. If $\hat{\lambda} = -\infty$, then $\hat{\lambda} = \bar{\lambda}$. So, we will assume that $\hat{\lambda} > -\infty$. On the other hand, for $i \in K(\lambda_k, x, x^k)$ we have $J_k(\lambda_k, x) = (f_i(x) - \lambda_k g_i(x))/g_i(x^k)$, and then

$$J_k(\lambda_k, x) \leq (f(x) - \lambda_k) g_i(x)/g_i(x^k).$$

Then, it follows that

$$J_k(\lambda_k, x) \leq (f(x) - \lambda_k)\gamma/\nu \quad \text{if} \quad f(x) - \lambda_k \geq 0, \tag{23}$$

and

$$J_k(\lambda_k, x) \leq (f(x) - \lambda_k)\nu/\gamma \quad \text{if} \quad f(x) - \lambda_k \leq 0. \tag{24}$$

We will show that for all $x \in X$ we have

$$\limsup_{k \to \infty} \Big(J_k(\lambda_k, x) - J_k(\lambda_k, x^{k+1})\Big) \geq 0. \tag{25}$$

Suppose, for contradiction, that there exists $\varepsilon > 0$, $\tilde{x} \in X$ and l such that $J_k(\lambda_k, \tilde{x}) - J_k(\lambda_k, x^{k+1}) < -\varepsilon$ for all $k \geq l$. As in the proof of Theorem 2.1 we show, using (22) with $x = \tilde{x}$, that this is impossible and that (25) is true.

For $x \in X$, if there is an infinite set of indices k such that $f(x) \geq \lambda_k$, then $f(x) \geq \hat{\lambda}$. Else, there exists $\bar{k}$ such that $f(x) < \lambda_k$ for all $k = \bar{k}, \bar{k}+1, \ldots$ But from (24) we have

$$(f(x) - \lambda_k)\nu/\gamma - J_k(\lambda_k, x^{k+1}) \geq J_k(\lambda_k, x) - J_k(\lambda_k, x^{k+1}),$$

for all $k = \bar{k}, \bar{k}+1, \ldots$, and then because of (25) and the fact that $J_k(\lambda_k, x^{k+1}) \to 0$ as $k \to \infty$, $f(x) \geq \hat{\lambda}$. So, $f(x) = \hat{\lambda}$. In both cases, $f(x) \geq \hat{\lambda}$ for all $x \in X$ which implies that $\hat{\lambda} = \bar{\lambda}$. □

Proposition 3.1. *Assume that the sequence $\{\alpha_k\}$ is such that $\alpha_k \geq \bar{\alpha}$ for all $k \in \mathbb{N}$, that the assumptions of Theorem 3.1 are fulfilled, and that the optimal solutions set of (P) is nonempty. If Algorithm 3.1 is considered in one of the following situations:*

(i) $\varphi(\lambda_k, x^k, \eta_k) = \min\{0, G_k(\lambda_k, x^k) + \eta_k\}$;
(ii) $\varphi(\lambda_k, x^k, \eta_k) = G_k(\lambda_k, x^k) + \eta_k$, $\sum_{k\geq 0} \eta_k < +\infty$, $\gamma^2 \leq \nu(\nu + \tau)$ *and* $\alpha_{k+1} \geq \alpha_k$ *for all* $k \in \mathbb{N}$;

then the sequence $\{x^k\}$ converges to a solution of (P).

Proof. Consider first the case (i). Then $\varphi(\lambda_k, x^k, \eta_k) \leq 0$ and thus, following (20) we obtain

$$0 \geq J_k(\lambda_k, x^{k+1}) \geq (\lambda_{k+1} - \lambda_k)\gamma/\nu.$$

On the other hand, for $\bar{x} \in X^*$, we have $J_k(\lambda_k, \bar{x}) \leq 0$. With $x = \bar{x}$ in (22) we then obtain

$$\begin{aligned}\|\bar{x} - \bar{x}^{k+1}\|^2 \leq \left(1 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x} - \bar{x}^k\|^2 + 2\sqrt{\eta_{k-1}/\alpha_{k-1}} + \eta_{k-1}/\alpha_{k-1} \\ +\eta_k/\alpha_k + (\gamma/\nu)(\lambda_k - \lambda_{k+1})/\bar{\alpha}. \qquad (26)\end{aligned}$$

The rest is as in the proof of Proposition 2.1 by using Lemma 2.1.

Suppose now that (ii) occurs. For $\bar{x} \in X^*$ we have $J_k(\lambda_k, \bar{x}) \leq (\bar{\lambda} - \lambda_k)\tau/\gamma$, and on the other hand

$$J(\lambda_k, x^{k+1}) \geq \delta_k(\lambda_{k+1} - \lambda_k) \geq (\lambda_{k+1} - \bar{\lambda})\nu/\gamma + (\bar{\lambda} - \lambda_k)\gamma/\nu.$$

Considering (22) with these inequalities we get

$$\begin{aligned}
\|\bar{x}-\bar{x}^{k+1}\|^2 \le & \left(1+2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x}-\bar{x}^k\|^2+2\sqrt{\eta_{k-1}/\alpha_{k-1}}+\eta_{k-1}/\alpha_{k-1}\\
&+\eta_k/\alpha_k+(\gamma/\nu-\tau/\gamma)(\lambda_k-\bar{\lambda})/\alpha_k-(\nu/\gamma)(\lambda_{k+1}-\bar{\lambda})/\alpha_k\\
\le & \left(1+2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x}-\bar{x}^k\|^2+2\sqrt{\eta_{k-1}/\alpha_{k-1}}+\eta_{k-1}/\alpha_{k-1}\\
&+\eta_k/\alpha_k+(\nu/\gamma)(\lambda_k-\bar{\lambda})-(\nu/\gamma)(\lambda_{k+1}-\bar{\lambda})/\alpha_{k+1}\\
&+(\nu/\gamma)(1/\alpha_{k+1}-1/\alpha_k)(\lambda_{k+1}-\bar{\lambda}),
\end{aligned}$$

where we used the assumption $\gamma^2 \le \nu(\nu+\tau)$ in the last inequality. Since $\alpha_{k+1} \ge \alpha_k$ we have

$$\begin{aligned}
\|\bar{x}-\bar{x}^{k+1}\|^2+(\nu/\gamma)(\lambda_{k+1}-\bar{\lambda})/\alpha_{k+1} \le & \left(1+2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\|\bar{x}-\bar{x}^k\|^2\\
&+2\sqrt{\eta_{k-1}/\alpha_{k-1}}+\eta_{k-1}/\alpha_{k-1}+\eta_k/\alpha_k\\
&+(\nu/\gamma)(\lambda_k-\bar{\lambda})\\
\le & \left(1+2\sqrt{\eta_{k-1}/\alpha_{k-1}}\right)\left(\|\bar{x}-\bar{x}^k\|^2\right.\\
&\left.+(\nu/\gamma)(\lambda_k-\bar{\lambda})\right)+2\sqrt{\eta_{k-1}/\alpha_{k-1}}\\
&+\eta_{k-1}/\alpha_{k-1}+\eta_k/\alpha_k.
\end{aligned}$$

By invoking Lemma 2.1 with $u_k=\|\bar{x}-\bar{x}^k\|^2+(\nu/\gamma)(\lambda_k-\bar{\lambda})$, $\mu_k=2\sqrt{\eta_{k-1}/\alpha_{k-1}}$ and $\beta_k=2\sqrt{\eta_{k-1}/\alpha_{k-1}}+\eta_{k-1}/\alpha_{k-1}+\eta_k/\alpha_k$, we deduce that the sequence $\{\bar{x}^k\}$ converges. The rest is as in the previous case. □

3.2. Rate of convergence of algorithm DTR2

Following [1], the algorithm DTR2 is generally much faster than DTR1. This is quite natural since DT2 is generally faster than DT1 [3, 5]. In this section we show that Algorithm 3.1 converges linearly when the assumption (H) is fulfilled and superlinearly under additional assumptions.

Theorem 3.2. *Suppose that the optimal solutions set X^* of (P) is nonempty and that the function $J(\bar{\lambda},.)$ satisfies the condition (H) in Section 2.2. Also suppose that $\eta_k/(\lambda_k-\bar{\lambda}) \to 0$ as $k\to\infty$ and that the sequence $\{x^k\}$ converges to a solution of (P) (see Prop. 3.1). If Algorithm 3.1 is considered in one of the two following situations:*

(i) $\varphi(\lambda_k, x^k, \eta_k)=\min\{0, G_k(\lambda_k, x^k)+\eta_k\}$;

(ii) $\varphi(\lambda_k, x^k, \eta_k)=G_k(\lambda_k, x^k)+\eta_k$, *and* $\gamma^2<\nu(\nu+\tau)$;

then for α_k small enough, the sequence $\{\lambda_k\}$ converges linearly to $\bar{\lambda}$.

Proof. For all $k\in\mathbb{N}$, let $\tilde{x}^k \in X^*$ be such that $d(x^k, X^*)=\|x^k-\tilde{x}^k\|$. We have

$$J_k(\lambda_k,\tilde{x}^k)+\alpha_k\|\tilde{x}^k-x^k\|^2+\eta_k \ge \delta_k'(\lambda_{k+1}-\lambda_k)\cdot \tag{27}$$

For k large enough, we have $x^k \in B(X^*, \delta) \cap X$ and thus $J(\bar\lambda, x^k) \geq \kappa \|x^k - \tilde{x}^k\|^2$. So,

$$\kappa \|\tilde{x}^k - x^k\|^2 \leq J(\bar\lambda, x^k) \leq (\lambda_k - \bar\lambda) \max_{i \in I} g_i(x^k) \tag{28}$$

$$\leq (\lambda_k - \bar\lambda)\gamma. \tag{29}$$

On the other hand,

$$J_k(\lambda_k, \tilde{x}^k) \leq (\bar\lambda - \lambda_k) \min_{i \in I} g_i(\tilde{x}^k)/g_i(x^k) \tag{30}$$

$$\leq (\bar\lambda - \lambda_k)\tau/\gamma. \tag{31}$$

Therefore by using (27–29) and (31), for k large enough, we obtain

$$(\bar\lambda - \lambda_k)\tau/\gamma + (\lambda_k - \bar\lambda)\alpha_k\gamma/\kappa + \eta_k \geq \delta'_k(\lambda_{k+1} - \lambda_k). \tag{32}$$

With (i) we have $\lambda_{k+1} - \lambda_k \leq 0$ and $\delta'_k = \gamma/\nu$ and this yields

$$(1 - \nu\tau/\gamma^2 + \alpha_k\nu/\kappa)(\lambda_k - \bar\lambda) + \nu\eta_k/\gamma \geq \lambda_{k+1} - \bar\lambda.$$

So, if $\limsup_{k\to\infty} \alpha_k < \kappa\tau/\gamma^2$, then $\limsup_{k\to\infty}(\lambda_{k+1} - \bar\lambda)/(\lambda_k - \bar\lambda) < 1$.

Assume now that (ii) is satisfied. Then from (32) we get

$$(\bar\lambda - \lambda_k)\tau/\gamma + (\lambda_k - \bar\lambda)\alpha_k\gamma/\kappa + \eta_k \geq (\lambda_{k+1} - \bar\lambda)\nu/\gamma + (\bar\lambda - \lambda_k)\gamma/\nu,$$

and this implies that

$$(\gamma/\nu)(\gamma/\nu - \tau/\gamma + \alpha_k\gamma/\kappa)(\lambda_k - \bar\lambda) + \gamma\eta_k/\nu \geq \lambda_{k+1} - \bar\lambda.$$

It follows that if $\limsup_{k\to\infty} \alpha_k < \kappa[\nu(\nu+\tau) - \gamma^2]/(\gamma^2\nu)$, then $\limsup_{k\to\infty}(\lambda_{k+1} - \bar\lambda)/(\lambda_k - \bar\lambda) < 1$. □

Theorem 3.3. *With the assumptions of Theorem 3.2, the sequence* $\{\lambda_k\}$ *converges superlinearly to* $\bar\lambda$ *if* $\alpha_k \to 0$ *as* $k \to \infty$.

Proof. Since the sequence $\{x^k\}$ converges by assumption, then the sequence $\{\tilde{x}^k\}$ defined in the beginning of the proof of Theorem 3.2 also converges to the same limit because of the assumption (H). By using (19) we obtain

$$J_k(\lambda_k, x^{k+1}) \geq (\lambda_{k+1} - \bar\lambda) \min_{i \in I} g_i(x^{k+1})/g_i(x^k) - (\lambda_k - \bar\lambda) \max_{i \in I} g_i(x^{k+1})/g_i(x^k).$$

It follows that

$$J_k(\lambda_k, \tilde{x}^k) + \alpha_k\|\tilde{x}^k - x^k\|^2 + \eta_k \geq (\lambda_{k+1} - \bar\lambda) \min_{i \in I} \frac{g_i(x^{k+1})}{g_i(x^k)} - (\lambda_k - \bar\lambda) \max_{i \in I} \frac{g_i(x^{k+1})}{g_i(x^k)}.$$

Then, using this inequality and the inequalities (29) and (30), we obtain

$$\left[\max_{i\in I}\frac{g_i(x^{k+1})}{g_i(x^k)}-\min_{i\in I}\frac{g_i(\tilde{x}^k)}{g_i(x^k)}+\frac{\gamma\alpha_k}{\kappa}\right](\lambda_k-\bar{\lambda})+\eta_k\geq(\lambda_{k+1}-\bar{\lambda})\min_{i\in I}\frac{g_i(x^{k+1})}{g_i(x^k)}\cdot$$

Since the sequences $\{x^k\}$ and $\{\tilde{x}^k\}$ converge to the same limit, then

$$\max_{i\in I} g_i(x^{k+1})/g_i(x^k)\to 1,\quad \min_{i\in I} g_i(\tilde{x}^k)/g_i(x^k)\to 1 \text{ and } \min_{i\in I} g_i(x^{k+1})/g_i(x^k)\to 1$$

as $k\to\infty$; and thus $(\lambda_{k+1}-\bar{\lambda})/(\lambda_k-\bar{\lambda})\to 0$ as $k\to\infty$. □

Remark 3.1. If the hypothesis (H') is satisfied and $X^*=\{\bar{x}\}$, then $J(\bar{\lambda},x^k)\geq\kappa\|x^k-\bar{x}\|$. So the sequence $\{x^k\}$ converges to $\bar{x}$ since $0\leq J(\bar{\lambda},x^k)\leq\gamma(\lambda_k-\bar{\lambda})$ and $\lambda_k\to\bar{\lambda}$ as $k\to\infty$.

With similar arguments as those used in the proofs of Proposition 2.2 and Proposition 2.3, we obtain the following results.

Proposition 3.2. *For all $n\in\mathbb{N}$, let $\sigma_n=\sum_{k=0}^n 1/\alpha_k$. If the optimal solutions set X^* of (P) is nonempty and $\psi(\lambda_k,x^k,\eta_k)=\min\{0,G_k(\lambda_k,x^k)+\eta_k\}$, then for all $\bar{x}\in X^*$ we have the following estimate*

$$\begin{aligned}\tau(\lambda_{n+1}-\bar{\lambda})/\gamma\leq\sigma_n^{-1}\Bigg[\sum_{k=1}^n\big(&\eta_{k-1}/\alpha_{k-1}+\eta_k/\alpha_k\\&+2(1+\|\bar{x}-\bar{x}^k\|^2)\sqrt{\eta_{k-1}/\alpha_{k-1}}\big)\\&+\|\bar{x}-\bar{x}^1\|^2+(\tau/\gamma)(\lambda_0-\bar{\lambda})/\alpha_0\\&+(\gamma/\nu-\tau/\gamma)\sum_{k=1}^n(\lambda_k-\lambda_{k+1})/\alpha_k\Bigg].\end{aligned}$$

Proposition 3.3. *For all $n\in\mathbb{N}$, let $\sigma_n=\sum_{k=0}^n 1/\alpha_k$. If the optimal solutions set X^* of (P) is nonempty and $\psi(\lambda_k,x^k,\eta_k)=G_k(\lambda_k,x^k)+\eta_k$, then for all $\bar{x}\in X^*$ we have the following estimate*

$$\begin{aligned}\nu(\lambda_{n+1}-\bar{\lambda})/\gamma\leq\sigma_n^{-1}\Bigg[\sum_{k=1}^n\big(&\eta_{k-1}/\alpha_{k-1}+\eta_k/\alpha_k+2(1+\|\bar{x}-\bar{x}^k\|^2)\\&\times\sqrt{\eta_{k-1}/\alpha_{k-1}}+\sigma_{k-1}\eta_k\big)+\|\bar{x}-\bar{x}^1\|^2\\&+(\nu/\gamma)(\lambda_0-\bar{\lambda})/\alpha_0+(\gamma/\nu-\tau/\gamma)\sum_{k=1}^n(\lambda_k-\bar{\lambda})/\alpha_k\Bigg].\end{aligned}$$

References

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

[2] J.-P. Crouzeix, J.A. Ferland and S. Schaible, An Algorithm for Generalized Fractional Programs. *J. Optim. Theory Appl.* **47** (1985) 35-49.
[3] J.-P. Crouzeix, J.A. Ferland and S. Schaible, Note on an Algorithm for Generalized Fractional Programs. *J. Optim. Theory Appl.* **50** (1986) 183-187.
[4] W. Dinkelbach, On Nonlinear Fractional Programming. *Management Sci.* **13** (1967) 492-498.
[5] A. Roubi, Method of Centers for Generalized Fractional Programming. *J. Optim. Theory Appl.* **107** (2000) 123-143.
[6] B. Martinet, Régularisation d'Inéquations Variationnelles par Approximation Successives. *RAIRO: Oper. Res.* **4** (1970) 154-158.
[7] R.T. Rockafellar, Monotone Operators and the Proximal Point Algorithm. *SIAM J. Control Optim.* **14** (1976) 877-898.
[8] O. Güler, On the Convergence of the Proximal Point Algorithm for Convex Minimization. *SIAM J. Control Optim.* **29** (1991) 403-419.
[9] A. Kaplan and R. Tichatschke, *Stable Methods for Ill-Posed Variational Problems.* Akademic Verlag, Berlin, Germany (1994).
[10] C. Lemaréchal and C. Sagastizábal, Practical Aspects of the Moreau–Yosida Regularization: Theoretical Preliminaries. *SIAM J. Optim.* **7** (1997) 367-385.
[11] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels.* Gauthier-Villars, Paris, Bruxelles, Montréal (1974).
[12] J.V. Burke and M.C. Ferris, Weak Sharp Minima in Mathematical Programming. *SIAM J. Control Optim.* **31** (1993) 1340-1359.
[13] O. Cornejo, A. Jourani and C. Zalinescu, Conditioning and Upper-Lipschitz Inverse Subdifferentials in Nonsmooth Optimization Problems. *J. Optim. Theory Appl.* **95** (1997) 127-148.