

Proximal-gradient algorithms for fractional programming

Radu Ioan Bot^{*†} Ernő Robert Csetnek[‡]

December 8, 2016

Abstract. In this paper we propose two proximal gradient algorithms for fractional programming problems in real Hilbert spaces, where the numerator is a proper, convex and lower semicontinuous function and the denominator is a smooth function, either concave or convex. In the iterative schemes, we perform a proximal step with respect to the nonsmooth numerator and a gradient step with respect to the smooth denominator. The algorithm in case of a concave denominator has the particularity that it generates sequences which approach both the (global) optimal solutions set and the optimal objective value of the underlying fractional programming problem. In case of a convex denominator the numerical scheme approaches the set of critical points of the objective function, provided the latter satisfies the Kurdyka-Łojasiewicz property.

Key Words. fractional programming, forward-backward algorithm, convergence rate, convex subdifferential, limiting subdifferential, Kurdyka-Łojasiewicz property

AMS subject classification. 65K05, 90C25, 90C32

1 Introduction and preliminaries

Consider the fractional programming problem

$$\bar{\theta} := \inf_{x \in S} \frac{f(x)}{g(x)}, \quad (1)$$

where S is a nonempty subset of a real Hilbert space \mathcal{H} , the function f is nonnegative and the function g is positive on S . One of the classical methods to handle (1) is Dinkelbach's procedure (see [13, 14]) which relates it to the following optimization problem

$$\inf_{x \in S} \{f(x) - \bar{\theta}g(x)\}. \quad (2)$$

If (1) has an optimal solution $\bar{x} \in S$, then this is also an optimal solution to (2) and the optimal objective value of the latter is equal to zero. Vice-versa, if (2) has $\bar{x} \in S$ as an optimal solution and its optimal objective value is equal to zero, then \bar{x} is an optimal

^{*}University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: radu.bot@univie.ac.at.

[†]Invited Associate Professor, Babeş-Bolyai University, Faculty of Mathematics and Computer Sciences, str. Mihail Kogălniceanu 1, 400084 Cluj-Napoca, Romania

[‡]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: ernoe.robert.csetnek@univie.ac.at. Research supported by FWF (Austrian Science Fund), Lise Meitner Programme, project M 1682-N25.

solution to (1), too. This shows that finding an optimal solution to (1) can be approached by algorithms which solve (2). However, one drawback of this procedure is that this can be done in the very restrictive case when the optimal objective value of (1) is known.

One can find in the literature (see [13, 14, 16, 17, 23]) an iterative scheme which, in the attempt to overcome this drawback in finite-dimensional spaces, requires the solving in each iteration $k \geq 0$ of the optimization problem

$$\inf_{x \in S} \{f(x) - \theta_k g(x)\}, \quad (3)$$

while θ_k is updated by $\theta_{k+1} := \frac{f(x^k)}{g(x^k)}$, where x^k is an optimal solution of (3). However, solving in each iteration an optimization problem of type (3) can be as expensive and difficult as solving the fractional programming problem (1).

The aim of this note is to propose an alternative to this approach. Namely, we formulate two iterative schemes for solving (1), where $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is proper, convex and lower semicontinuous and $g : \mathcal{H} \rightarrow \mathbb{R}$ is differentiable with Lipschitz continuous gradient and either concave or convex. Instead of solving in each iteration (3), the proposed iterative methods perform a gradient step with respect to g and a proximal step with respect to f . In this way, the functions f and g are processed separately in each iteration. A further advantage of the algorithm investigated in case g is concave comes from the fact that it generates sequences that concomitantly approach the set of optimal solutions and the optimal objective value of (1). The second numerical scheme, proposed in case g is convex, has the particularity that it approaches the set of critical points of the objective function of (1), provided the latter satisfies the Kurdyka-Łojasiewicz property.

For the notations used in this paper we refer the reader to [4, 9, 15, 24]. Let \mathcal{H} be a real Hilbert space with *inner product* $\langle \cdot, \cdot \rangle$ and associated *norm* $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. The symbols \rightharpoonup and \rightarrow denote weak and strong convergence, respectively.

For a function $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is the extended real line, we denote by $\text{dom } f = \{x \in \mathcal{H} : f(x) < +\infty\}$ its *effective domain* and say that f is *proper* if $\text{dom } f \neq \emptyset$ and $f(x) \neq -\infty$ for all $x \in \mathcal{H}$. The *subdifferential* of f at $x \in \mathcal{H}$, with $f(x) \in \mathbb{R}$, is the set $\partial f(x) := \{v \in \mathcal{H} : f(y) \geq f(x) + \langle v, y - x \rangle \forall y \in \mathcal{H}\}$. We take by convention $\partial f(x) := \emptyset$, if $f(x) \in \{\pm\infty\}$. Let $S \subseteq \mathcal{H}$ be a nonempty set. The *indicator function* of S , $\delta_S : \mathcal{H} \rightarrow \overline{\mathbb{R}}$, is the function which takes the value 0 on S and $+\infty$ otherwise.

An efficient tool for proving weak convergence of a sequence in Hilbert spaces (without a priori knowledge of its limit) is the Opial Lemma, which we recall in the following.

Lemma 1 (*Opial*) *Let C be a nonempty set of \mathcal{H} and $(x_k)_{k \geq 0}$ be a sequence in \mathcal{H} such that the following two conditions hold:*

- (a) *for every $x \in C$, $\lim_{k \rightarrow +\infty} \|x_k - x\|$ exists;*
- (b) *every weak sequential cluster point of $(x_k)_{k \geq 0}$ is in C ;*

Then $(x_k)_{k \geq 0}$ converges weakly to an element in C .

When proving the first part of the Opial Lemma one usually tries to show that for every $x \in C$ the sequence $(\|x_k - x\|)_{k \geq 0}$ fulfills a Fejér-type inequality. In this sense the following result is very useful.

Lemma 2 Let $(a_k)_{k \geq 0}$, $(b_k)_{k \geq 0}$ and $(\varepsilon_k)_{k \geq 0}$ be real sequences. Assume that $(a_k)_{k \geq 0}$ is bounded from below, $(b_k)_{k \geq 0}$ is nonnegative, $(\varepsilon_k)_{k \geq 0} \in \ell^1$ and $a_{k+1} - a_k + b_k \leq \varepsilon_k$ for every $k \geq 0$. Then $(a_k)_{k \geq 0}$ is convergent and $(b_k)_{k \geq 0} \in \ell^1$.

The following summability result will be useful in Subsection 2.2.

Lemma 3 Let $(a_k)_{k \geq 0}$ and $(\varepsilon_k)_{k \geq 0}$ be nonnegative real sequences, such that $\sum_{k \geq 0} \varepsilon_k < +\infty$ and $a_{k+1} \leq a \cdot a_k + \varepsilon_k$ for every $k \geq 0$, where $a \in \mathbb{R}$, $a < 1$. Then $\sum_{k \geq 0} a_k < +\infty$.

Finally, the descent lemma which we recall next is a helpful tool in the convergence analysis of the algorithms proposed in this manuscript.

Lemma 4 (see [21, Lemma 1.2.3]) Let $g : \mathcal{H} \rightarrow \mathbb{R}$ be (Fréchet) differentiable with L -Lipschitz continuous gradient. Then

$$g(y) \leq g(x) + \langle \nabla g(x), y - x \rangle + \frac{L}{2} \|y - x\|^2 \quad \forall x, y \in \mathcal{H}.$$

2 Two proximal-gradient algorithms

In this section we propose two proximal-gradient algorithms for solving (1) and investigate their convergence properties. We treat the situations when g is either a convex or a concave function separately.

2.1 Concave denominator

The problem that we investigate throughout this subsection has the following formulation.

Problem 5 We are interested in solving the fractional programming problem

$$\bar{\theta} := \inf_{x \in S} \frac{f(x)}{g(x)} \quad (4)$$

where \mathcal{H} is a real Hilbert space, S is a nonempty, convex and closed subset of \mathcal{H} , $C := \{\bar{x} \in S : \bar{\theta} = \frac{f(\bar{x})}{g(\bar{x})}\} \neq \emptyset$ and the following conditions hold:

- (H_f) $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is proper, convex, lower semicontinuous such that $\text{dom } f \cap S \neq \emptyset$ and $f(x) \geq 0 \quad \forall x \in S$;
- (H_g) $g : \mathcal{H} \rightarrow \mathbb{R}$ is concave, (Fréchet) differentiable with L -Lipschitz continuous gradient, and there exists $M > 0$ such that $0 < g(x) \leq M \quad \forall x \in S$.

To this aim we propose the following algorithm.

Algorithm 6

Initialization: Choose $x^0 \in S \cap \text{dom } f$ and set $\theta_1 := \frac{f(x^0)}{g(x^0)}$;
For $k \geq 1$: **If** $\theta_k = 0$, **then** $x^{k-1} \in C$: **STOP**;
 $\eta_k := \frac{1}{2L\theta_k}$;
 $x^k := \operatorname{argmin}_{x \in S} \left[f(x) + \frac{1}{2\eta_k} \left\| x - (x^{k-1} + \theta_k \eta_k \nabla g(x^{k-1})) \right\|^2 \right]$;
 $\theta_{k+1} := \frac{f(x^k)}{g(x^k)}$.

We are now in position to present the convergence statement of this algorithm. To this end we assume that the algorithm does not stop after finitely many iterations.

Theorem 7 *In the setting of Problem 5, consider the sequences generated by Algorithm 6. The following statements hold:*

(i) *The sequence $(\theta_k)_{k \geq 1}$ is nonincreasing and $\lim_{k \rightarrow +\infty} \theta_k = \bar{\theta}$. Moreover,*

$$0 \leq \theta_{k+1} - \bar{\theta} \leq \frac{\theta_1(M + L\|\bar{x} - x^0\|^2)}{kg(\bar{x})} \quad \forall \bar{x} \in C \quad \forall k \geq 1. \quad (5)$$

(ii) *Additionally, assume that $\inf_{x \in S} \frac{f(x)}{g(x)} > 0$. Then the sequence $(x^k)_{k \geq 0}$ converges weakly to an element in C .*

Proof. According to the first order optimality conditions we have

$$0 \in \partial(f + \delta_S)(x^k) + \frac{1}{\eta_k} \left(x^k - x^{k-1} - \theta_k \eta_k \nabla g(x^{k-1}) \right) \quad \forall k \geq 1. \quad (6)$$

A direct consequence of the definition of the convex subdifferential is the inequality

$$\begin{aligned} f(x) - f(x^k) &\geq \left\langle \frac{1}{\eta_k} (x^{k-1} - x^k) + \theta_k \nabla g(x^{k-1}), x - x^k \right\rangle \\ &= \frac{1}{2\eta_k} \left(\|x^{k-1} - x^k\|^2 + \|x - x^k\|^2 - \|x - x^{k-1}\|^2 \right) \\ &\quad + \theta_k \langle \nabla g(x^{k-1}), x - x^k \rangle \quad \forall x \in S \quad \forall k \geq 1. \end{aligned} \quad (7)$$

Invoking the concavity of g and using that $\theta_k \geq 0$, we have

$$-\theta_k g(x) + \theta_k g(x^{k-1}) \geq \theta_k \langle \nabla g(x^{k-1}), x^{k-1} - x \rangle \quad \forall x \in S \quad \forall k \geq 1. \quad (8)$$

Combining (7) and (8), we obtain

$$\begin{aligned} f(x) - \theta_k g(x) &\geq \frac{1}{2\eta_k} \left(\|x^{k-1} - x^k\|^2 + \|x - x^k\|^2 - \|x - x^{k-1}\|^2 \right) \\ &\quad + f(x^k) - \theta_k g(x^{k-1}) + \theta_k \langle \nabla g(x^{k-1}), x^{k-1} - x^k \rangle \quad \forall x \in S \quad \forall k \geq 1. \end{aligned} \quad (9)$$

Lemma 4 applied to the function $-g$ yields the inequality

$$-\theta_k g(x^{k-1}) + \theta_k \langle \nabla g(x^{k-1}), x^{k-1} - x^k \rangle \geq -\theta_k g(x^k) - \frac{L\theta_k}{2} \|x^k - x^{k-1}\|^2 \quad \forall k \geq 1,$$

hence from (9) we derive

$$\begin{aligned} f(x) - \theta_k g(x) &\geq \frac{1}{2\eta_k} \left(\|x^{k-1} - x^k\|^2 + \|x - x^k\|^2 - \|x - x^{k-1}\|^2 \right) \\ &\quad + f(x^k) - \theta_k g(x^k) - \frac{L\theta_k}{2} \|x^k - x^{k-1}\|^2 \quad \forall x \in S \quad \forall k \geq 1. \end{aligned}$$

Taking into account the relation $f(x^k) = \theta_{k+1}g(x^k)$ and the way η_k is defined, we obtain for every $x \in S$ and $k \geq 1$ the inequality

$$f(x) - \theta_k g(x) \geq (\theta_{k+1} - \theta_k)g(x^k) + \frac{L\theta_k}{2} \|x^k - x^{k-1}\|^2 + L\theta_k \|x - x^k\|^2 - L\theta_k \|x - x^{k-1}\|^2. \quad (10)$$

(i) Taking $x := x^{k-1}$ in (10) we get

$$(\theta_{k+1} - \theta_k)g(x^k) + \frac{3L\theta_k}{2}\|x^k - x^{k-1}\|^2 \leq 0 \quad \forall k \geq 1. \quad (11)$$

This further implies that $(\theta_k)_{k \geq 1}$ is a nonincreasing sequence, hence convergent, since it is bounded from below by 0.

Consider now an arbitrary $\bar{x} \in C$ and take $x := \bar{x}$ in (10). We derive

$$\begin{aligned} \frac{L\theta_k}{2}\|x^k - x^{k-1}\|^2 + L\theta_k\|\bar{x} - x^k\|^2 - L\theta_k\|\bar{x} - x^{k-1}\|^2 &\leq (\theta_k - \theta_{k+1})M + f(\bar{x}) - \theta_k g(\bar{x}) \\ &= (\theta_k - \theta_{k+1})M + (\bar{\theta} - \theta_k)g(\bar{x}) \\ &\leq (\theta_k - \theta_{k+1})M \quad \forall k \geq 1. \end{aligned} \quad (12)$$

This yields the inequality

$$\frac{L\theta_k}{2}\|x^k - x^{k-1}\|^2 + L\theta_{k+1}\|\bar{x} - x^k\|^2 - L\theta_k\|\bar{x} - x^{k-1}\|^2 \leq (\theta_k - \theta_{k+1})M \quad \forall k \geq 1. \quad (13)$$

Since $(\theta_k)_{k \geq 1}$ is bounded from below by 0, the sequence on the right-hand side of inequality (13) belongs to ℓ^1 . We derive from Lemma 2 that

$$\sum_{k \geq 1} \theta_k \|x^k - x^{k-1}\|^2 < +\infty \quad (14)$$

and

$$(\theta_{k+1}\|\bar{x} - x^k\|^2)_{k \geq 0} \text{ is convergent.} \quad (15)$$

Coming back to (12) and using $\theta_k \geq \theta_{k+1}$, we obtain

$$(\bar{\theta} - \theta_k)g(\bar{x}) \geq (\theta_{k+1} - \theta_k)M + \frac{L\theta_k}{2}\|x^k - x^{k-1}\|^2 + L\theta_{k+1}\|\bar{x} - x^k\|^2 - L\theta_k\|\bar{x} - x^{k-1}\|^2 \quad \forall k \geq 1. \quad (16)$$

Relying on (14) and (15) and the convergence of the sequence $(\theta_k)_{k \geq 1}$, the right-hand side of the above inequality is a sequence which converges to 0 as $k \rightarrow +\infty$. Invoking also the fact that $(\theta_k)_{k \geq 1}$ is bounded from below by $\bar{\theta}$, we conclude that $\lim_{k \rightarrow +\infty} \theta_k = \bar{\theta}$.

Let us prove now the convergence rate result stated in (5). Let $\bar{x} \in C$ and $n \geq 1$ be arbitrary. From (11) we obtain

$$\sum_{k=1}^n (k-1)(\theta_k - \theta_{k+1})g(x^k) \geq \frac{3}{2}L \sum_{k=1}^n (k-1)\theta_k \|x^k - x^{k-1}\|^2,$$

hence

$$\begin{aligned} \sum_{k=1}^n \left((k-1)\theta_k - k\theta_{k+1} + \theta_{k+1} \right) &= \sum_{k=1}^n (k-1)(\theta_k - \theta_{k+1}) \\ &\geq \frac{3L}{2M} \sum_{k=1}^n (k-1)\theta_k \|x^k - x^{k-1}\|^2. \end{aligned}$$

Noticing the telescoping sum in the left-hand side of the previous inequality, we obtain

$$-n\theta_{n+1} + \sum_{k=1}^n \theta_{k+1} \geq \frac{3L}{2M} \sum_{k=1}^n (k-1)\theta_k \|x^k - x^{k-1}\|^2. \quad (17)$$

Summing up the inequalities in (16) for k from 1 to $n + 1$ we obtain

$$\begin{aligned} & \left((n+1)\bar{\theta} - \sum_{k=1}^{n+1} \theta_k \right) g(\bar{x}) \geq \\ & (\theta_{n+2} - \theta_1)M + \frac{L}{2} \sum_{k=1}^{n+1} \theta_k \|x^k - x^{k-1}\|^2 + L\theta_{n+2}\|\bar{x} - x^{n+2}\|^2 - L\theta_1\|\bar{x} - x^0\|^2, \end{aligned}$$

hence

$$\begin{aligned} & (n+1)\bar{\theta} - \sum_{k=1}^n \theta_{k+1} - \theta_1 \geq \\ & \frac{1}{g(\bar{x})} \left[(\theta_{n+2} - \theta_1)M + \frac{L}{2} \sum_{k=1}^{n+1} \theta_k \|x^k - x^{k-1}\|^2 + L\theta_{n+2}\|\bar{x} - x^{n+2}\|^2 - L\theta_1\|\bar{x} - x^0\|^2 \right]. \end{aligned} \quad (18)$$

Summing up the inequalities (17) and (18) and discarding the nonnegative terms on the right-hand side we derive

$$n(\bar{\theta} - \theta_{n+1}) + \bar{\theta} - \theta_1 \geq -\frac{\theta_1 M}{g(\bar{x})} - \frac{L\theta_1}{g(\bar{x})}\|\bar{x} - x^0\|^2.$$

Noticing that $\theta_1 \geq \bar{\theta}$, the last inequality implies (5) after rearranging the terms.

(ii) For the remaining of the proof we assume that $\inf_{x \in S} \frac{f(x)}{g(x)} > 0$. In this situation, $\lim_{k \rightarrow +\infty} \theta_k = \bar{\theta} > 0$ and from (14) and (15) we derive

$$\lim_{k \rightarrow +\infty} (x^k - x^{k-1}) = 0 \quad (19)$$

and

$$(\|\bar{x} - x^k\|)_{k \geq 1} \text{ is convergent } \forall \bar{x} \in C. \quad (20)$$

Thus the first condition in the Opial Lemma is fulfilled.

Consider now a subsequence $(x^{k_l})_{l \geq 0}$ of $(x^k)_{k \geq 0}$ that weakly converges to \bar{x} as $l \rightarrow +\infty$. From (6) we deduce

$$\frac{1}{\eta_{k_l}} (x^{k_l-1} - x^{k_l}) + \theta_{k_l} \nabla g(x^{k_l-1}) \in \partial(f + \delta_S)(x^{k_l}),$$

hence

$$\frac{1}{\eta_{k_l}} (x^{k_l-1} - x^{k_l}) + \theta_{k_l} \nabla g(x^{k_l-1}) - \bar{\theta} \nabla g(x^{k_l}) \in \partial(f + \delta_S - \bar{\theta}g)(x^{k_l}) \quad \forall l \geq 1, \quad (21)$$

due to the concavity of g and $\bar{\theta} > 0$. Since for every $l \geq 1$ we have

$$\begin{aligned} \|\theta_{k_l} \nabla g(x^{k_l-1}) - \bar{\theta} \nabla g(x^{k_l})\| & \leq |\theta_{k_l} - \bar{\theta}| \|\nabla g(x^{k_l-1})\| + \bar{\theta} \|\nabla g(x^{k_l-1}) - \nabla g(x^{k_l})\| \\ & \leq L|\theta_{k_l} - \bar{\theta}| \|x^{k_l-1} - x^0\| + |\theta_{k_l} - \bar{\theta}| \|\nabla g(x^0)\| + \bar{\theta} L \|x^{k_l-1} - x^{k_l}\|, \end{aligned}$$

from (i), (19) and the fact that $(x^{k_l})_{l \geq 0}$ is bounded, we conclude that

$$\frac{1}{\eta_{k_l}} (x^{k_l-1} - x^{k_l}) + \theta_{k_l} \nabla g(x^{k_l-1}) - \bar{\theta} \nabla g(x^{k_l}) \rightarrow 0 \text{ as } l \rightarrow +\infty.$$

Since $(x^{k_l})_{l \geq 0}$ weakly converges to \bar{x} as $l \rightarrow +\infty$, from (21) and the fact that the graph of the convex subdifferential of a proper, convex and lower semicontinuous function is sequentially closed with respect to the weak-norm topology (see [4, Proposition 20.33]), we derive that

$$0 \in \partial(f + \delta_S - \bar{\theta}g)(\bar{x}),$$

hence $\bar{x} \in \text{dom } f \cap S$. The definition of the convex subdifferential yields the inequality

$$f(y) - \bar{\theta}g(y) \geq f(\bar{x}) - \bar{\theta}g(\bar{x}) \quad \forall y \in S.$$

From here, by choosing $y \in C$, we get

$$0 \geq f(\bar{x}) - \bar{\theta}g(\bar{x}),$$

hence

$$\bar{\theta} \geq \frac{f(\bar{x})}{g(\bar{x})}. \quad (22)$$

Relation (22) implies now that $\bar{x} \in C$. Thus the second condition in the Opial Lemma is also fulfilled. The conclusion follows now from Lemma 1. \blacksquare

2.2 Convex denominator

In this subsection we consider the case when g is a convex function.

Problem 8 *We are interested in solving the fractional programming problem*

$$\inf_{x \in S} \frac{f(x)}{g(x)} \quad (23)$$

where \mathcal{H} is a real Hilbert space, S is a nonempty, convex and closed subset of \mathcal{H} , and the following conditions hold:

(H_f) $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ is proper, convex, lower semicontinuous such that $\text{dom } f \cap S \neq \emptyset$ and $f(x) \geq 0 \quad \forall x \in S$;

(\tilde{H}_g) $g : \mathcal{H} \rightarrow \mathbb{R}$ is convex, continuously (Fréchet) differentiable and there exists $M > 0$ such that $0 < g(x) \leq M \quad \forall x \in S$.

The algorithm we propose in this context has the following formulation.

Algorithm 9

Initialization: Choose $x^0 \in S \cap \text{dom } f$ and set $\theta_1 := \frac{f(x^0)}{g(x^0)}$;

For $k \geq 1$: **Choose** $\eta_k > 0$;

$$x^k := \operatorname{argmin}_{x \in S} \left[f(x) + \frac{1}{2\eta_k} \left\| x - (x^{k-1} + \theta_k \eta_k \nabla g(x^{k-1})) \right\|^2 \right];$$

$$\theta_{k+1} := \frac{f(x^k)}{g(x^k)}.$$

The proof of the first result in this subsection reveals the fact that when g is convex one cannot expect convergence of the whole sequence $(x^k)_{k \geq 0}$. Furthermore, if this is the case, then the limit is not necessarily an optimal solution of (23), but a critical point of

the objective function $\frac{f+\delta_S}{g}$ in the sense of the limiting subdifferential. In order to explain this notion, we need some prerequisites of nonsmooth analysis.

For the following generalized subdifferential notions and their basic properties we refer to [20, 22]. Let $h : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function. If $x \in \text{dom } h$, we consider the *Fréchet (viscosity) subdifferential* of h at x as being the set

$$\hat{\partial}h(x) := \left\{ v \in \mathcal{H} : \liminf_{y \rightarrow x} \frac{h(y) - h(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}.$$

For $x \notin \text{dom } h$ we set $\hat{\partial}h(x) := \emptyset$. The *limiting (Mordukhovich) subdifferential* is defined at $x \in \text{dom } h$ by

$$\partial_L h(x) := \{v \in \mathcal{H} : \exists x_k \rightarrow x, h(x_k) \rightarrow h(x) \text{ and } \exists v_k \in \hat{\partial}h(x_k), v_k \rightarrow v \text{ as } k \rightarrow +\infty\},$$

while for $x \notin \text{dom } h$, one takes $\partial_L h(x) := \emptyset$. Therefore $\hat{\partial}h(x) \subseteq \partial_L h(x)$ for each $x \in \mathcal{H}$.

When h is continuously differentiable around $x \in \mathcal{H}$ we have $\partial_L h(x) = \{\nabla h(x)\}$. Notice that in case h is convex, these two subdifferential notions coincide with the *convex subdifferential*, thus $\hat{\partial}h(x) = \partial_L h(x) = \{v \in \mathcal{H} : h(y) \geq h(x) + \langle v, y - x \rangle \ \forall y \in \mathcal{H}\}$ for all $x \in \mathcal{H}$.

The Fermat rule reads in this nonsmooth setting: if $x \in \mathcal{H}$ is a local minimizer of h , then $0 \in \partial_L h(x)$. An element $x \in \text{dom } h$ fulfilling this inclusion relation is called *critical point* of the function h . The set of all critical points of h is denoted by $\text{crit}(h)$.

The convergence of Algorithm (9) is stated in the following theorem.

Theorem 10 *In the setting of Problem 8, consider the sequences generated by Algorithm 9 such that the additional condition $\liminf_{k \rightarrow +\infty} \eta_k > 0$ is satisfied. The following statements hold:*

(i) *The sequence $(\theta_k)_{k \geq 1}$ is nonincreasing, hence convergent. Moreover,*

$$\sum_{k \geq 1} \frac{1}{\eta_k} \|x^k - x^{k-1}\|^2 < +\infty.$$

(ii) *For every (strong) limit point \bar{x} of $(x^k)_{k \geq 0}$, it holds $\bar{x} \in \text{dom } f \cap S$ and $\lim_{k \rightarrow +\infty} \theta_k = \frac{f(\bar{x})}{g(\bar{x})}$. If we additionally have that $\bar{x} \in \text{int}(\text{dom } f \cap S)$, then $0 \in \partial_L \left(\frac{f+\delta_S}{g} \right) (\bar{x})$.*

Proof. As already seen in the proof of Theorem 7, we have

$$\frac{1}{\eta_k} (x^{k-1} - x^k) + \theta_k \nabla g(x^{k-1}) \in \partial(f + \delta_S)(x^k) \ \forall k \geq 1, \quad (24)$$

and

$$\begin{aligned} f(x) - f(x^k) &\geq \frac{1}{2\eta_k} \left(\|x^{k-1} - x^k\|^2 + \|x - x^k\|^2 - \|x - x^{k-1}\|^2 \right) \\ &\quad + \theta_k \langle \nabla g(x^{k-1}), x - x^k \rangle \ \forall x \in S \ \forall k \geq 1. \end{aligned} \quad (25)$$

By choosing $x := x^{k-1}$ in (25) we obtain

$$f(x^{k-1}) - f(x^k) \geq \frac{1}{\eta_k} \|x^{k-1} - x^k\|^2 + \theta_k \langle \nabla g(x^{k-1}), x^{k-1} - x^k \rangle \ \forall k \geq 1.$$

Further, by combining this with

$$\theta_k \left(g(x^k) - g(x^{k-1}) \right) \geq \theta_k \langle \nabla g(x^{k-1}), x^k - x^{k-1} \rangle,$$

we obtain

$$\begin{aligned} 0 = f(x^{k-1}) - \theta_k g(x^{k-1}) &\geq f(x^k) - \theta_k g(x^k) + \frac{1}{\eta_k} \|x^{k-1} - x^k\|^2 \\ &= (\theta_{k+1} - \theta_k) g(x^k) + \frac{1}{\eta_k} \|x^{k-1} - x^k\|^2 \quad \forall k \geq 1. \end{aligned} \quad (26)$$

(i) From (26) we obtain that $(\theta_k)_{k \geq 1}$ is nonincreasing, hence convergent, since it is bounded from below by 0. Moreover, from (26) we obtain

$$\frac{1}{\eta_k} \|x^{k-1} - x^k\|^2 \leq (\theta_k - \theta_{k+1}) M \quad \forall k \geq 1,$$

hence $\sum_{k \geq 1} \frac{1}{\eta_k} \|x^k - x^{k-1}\|^2 < +\infty$.

(ii) Without losing the generality, we assume that $x^k \rightarrow \bar{x}$ as $k \rightarrow +\infty$. Since S is closed, we have $\bar{x} \in S$. By choosing $x := \bar{x}$ in (25), we obtain

$$\begin{aligned} f(\bar{x}) - f(x^k) &\geq \frac{1}{2\eta_k} \left(\|x^{k-1} - x^k\|^2 + \|\bar{x} - x^k\|^2 - \|\bar{x} - x^{k-1}\|^2 \right) \\ &\quad + \theta_k \langle \nabla g(x^{k-1}), \bar{x} - x^k \rangle \quad \forall k \geq 1. \end{aligned}$$

By using (i), one can see that the right-hand side of the above inequality converges to 0 as $k \rightarrow +\infty$. Hence, $\limsup_{k \rightarrow +\infty} f(x^k) \leq f(\bar{x})$. Since f is lower semicontinuous, the reverse inequality is also true, thus

$$\lim_{k \rightarrow +\infty} f(x^k) = f(\bar{x}).$$

Furthermore, due to the continuity of g , we have

$$\lim_{k \rightarrow +\infty} g(x^k) = g(\bar{x}).$$

Let us denote by θ the limit of the sequence $(\theta_k)_{k \geq 1}$. Passing to the limit as $k \rightarrow +\infty$ in the relation which defines θ_{k+1} in Algorithm 9, we obtain

$$\theta = \frac{f(\bar{x})}{g(\bar{x})}. \quad (27)$$

By using again the closedness property of the graph of the convex subdifferential, from (24) and (i) we obtain

$$\theta \nabla g(\bar{x}) \in \partial(f + \delta_S)(\bar{x}), \quad (28)$$

hence $\bar{x} \in \text{dom } f \cap S$.

Assume now that $\bar{x} \in \text{int}(\text{dom } f \cap S)$. In this situation $f + \delta_S$ is Lipschitz continuous around \bar{x} (see [4, Theorem 8.29]). From (27) and (28) we obtain

$$\begin{aligned} 0 &\in -\frac{(f + \delta_S)(\bar{x})}{g^2(\bar{x})} \nabla g(\bar{x}) + \frac{1}{g(\bar{x})} \partial(f + \delta_S)(\bar{x}) \\ &= (f + \delta_S)(\bar{x}) \nabla \left(\frac{1}{g} \right) (\bar{x}) + \partial \left(\frac{1}{g(\bar{x})} (f + \delta_S) \right) (\bar{x}) \\ &= (f + \delta_S)(\bar{x}) \nabla \left(\frac{1}{g} \right) (\bar{x}) + \partial_L \left(\frac{1}{g(\bar{x})} (f + \delta_S) \right) (\bar{x}) \\ &= \partial_L \left(\frac{1}{g} \cdot (f + \delta_S) \right) (\bar{x}), \end{aligned}$$

where the last equality makes use of [20, Corollary 1.111(i)]. ■

Remark 11 (a) The main ingredient in the proof of the second statement of the above theorem is the rule for the limiting subdifferential of the product (or quotient) of locally Lipschitz continuous functions. We notice that similar rules are valid also for the Clarke subdifferential (see [12, Exercise 10.21]).

(b) Whenever $\frac{f+\delta_S}{g}$ is a convex function, we obtain in the hypotheses of the above theorem that \bar{x} is a global optimal solution of (23) and $\lim_{k \rightarrow +\infty} \theta_k = \frac{f(\bar{x})}{g(\bar{x})} = \inf_{x \in S} \frac{f(x)}{g(x)}$.

In the remaining of this subsection we address the question whether one can guarantee the convergence of the whole sequence $(x^k)_{k \geq 0}$ generated in Algorithm 9. We will see that this is ensured whenever the objective function of (23) satisfies the *Kurdyka-Łojasiewicz property*. To this end we recall some notations and definitions related to the latter.

For the remaining of this section we suppose that \mathcal{H} is finite dimensional. For $\eta \in (0, +\infty]$, we denote by Θ_η the class of concave and continuous functions $\varphi : [0, \eta) \rightarrow [0, +\infty)$ such that $\varphi(0) = 0$, φ is continuously differentiable on $(0, \eta)$, continuous at 0 and $\varphi'(s) > 0$ for all $s \in (0, \eta)$. In the following definition (see [2, 8]) we use also the *distance function* to a set, defined for $A \subseteq \mathcal{H}$ as $\text{dist}(x, A) = \inf_{y \in A} \|x - y\|$ for all $x \in \mathcal{H}$.

Definition 1 (*Kurdyka-Łojasiewicz property*) Let $h : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ be a proper and lower semicontinuous function. We say that h satisfies the *Kurdyka-Łojasiewicz (KL) property* at $\bar{x} \in \text{dom } \partial_L h = \{x \in \mathcal{H} : \partial_L h(x) \neq \emptyset\}$ if there exists $\eta \in (0, +\infty]$, a neighborhood U of \bar{x} and a function $\varphi \in \Theta_\eta$ such that for all x in the intersection

$$U \cap \{x \in \mathcal{H} : h(\bar{x}) < h(x) < h(\bar{x}) + \eta\}$$

the following inequality holds

$$\varphi'(h(x) - h(\bar{x})) \text{dist}(0, \partial_L h(x)) \geq 1.$$

If h satisfies the KL property at each point in $\text{dom } \partial h$, then h is called a *KL function*.

The origins of this notion go back to the pioneering work of Łojasiewicz [19], where it is proved that for a real-analytic function $h : \mathcal{H} \rightarrow \mathbb{R}$ and a critical point $\bar{x} \in \mathcal{H}$ (that is $\nabla h(\bar{x}) = 0$), there exists $\theta \in [1/2, 1)$ such that the function $|h - h(\bar{x})|^\theta \|\nabla h\|^{-1}$ is bounded around \bar{x} . This corresponds to the situation when $\varphi(s) = Cs^{1-\theta}$, where $C > 0$. The result of Łojasiewicz allows the interpretation of the KL property as a re-parametrization of the function values in order to avoid flatness around the critical points. Kurdyka [18] extended this property to differentiable functions definable in an o-minimal structure. Further extensions to the nonsmooth setting can be found in [2, 5–7].

One of the remarkable properties of KL functions is their ubiquity in applications, according to [8]. To this class of functions belong semi-algebraic, real sub-analytic, semi-convex, uniformly convex and convex functions satisfying a growth condition. We refer the reader to [1–3, 5–8] and the references therein for more details regarding KL functions and illustrating examples.

An important role in our convergence analysis will be played by the following uniformized KL property given in [8, Lemma 6].

Lemma 12 *Let $\Omega \subseteq \mathcal{H}$ be a compact and connected set and let $h : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ be a proper and lower semicontinuous function. Assume that h is constant on Ω and h satisfies the KL property at each point of Ω . Then there exist $\varepsilon, \eta > 0$ and $\varphi \in \Theta_\eta$ such that for all $\bar{x} \in \Omega$ and for all x in the intersection*

$$\{x \in \mathcal{H} : \text{dist}(x, \Omega) < \varepsilon\} \cap \{x \in \mathcal{H} : h(\bar{x}) < h(x) < h(\bar{x}) + \eta\} \quad (29)$$

the following inequality holds

$$\varphi'(h(x) - h(\bar{x})) \text{dist}(0, \partial_L h(x)) \geq 1. \quad (30)$$

The techniques used below are well-known in the community dealing with algorithms for optimization problems involving functions with the Kurdyka-Łojasiewicz property (see [3, 8, 10, 11]). We show that this approach can be used also for fractional programming problems.

In the following we denote by $\omega((x^k)_{k \geq 0})$ the set of cluster points of the sequence $(x^k)_{k \geq 0}$. The first statement in the next result is a direct consequence of Theorem 10, while the other statements can be proved similar to [8, Lemma 5], where it is noticed that (b) and (c) are generic for sequences satisfying the relation $\lim_{k \rightarrow +\infty} (x^k - x^{k-1}) = 0$.

Lemma 13 *In the setting of Problem 8, let \mathcal{H} be finite dimensional and consider the sequences generated by Algorithm 9 such that the additional condition*

$$0 < \liminf_{k \rightarrow +\infty} \eta_k \leq \limsup_{k \rightarrow +\infty} \eta_k < +\infty$$

is satisfied. Assume that $(x^k)_{k \geq 0}$ is bounded. The following statements hold:

- (a) $\omega((x^k)_{k \geq 0}) \cap \text{int}(\text{dom } f \cap S) \subseteq \text{crit} \left(\frac{f + \delta_S}{g} \right)$;
- (b) $\lim_{k \rightarrow \infty} \text{dist} \left(x^k, \omega((x^k)_{k \geq 0}) \right) = 0$;
- (c) $\omega((x^k)_{k \geq 0})$ is nonempty, compact and connected;
- (d) $\frac{f + \delta_S}{g}$ is finite and constant on $\omega((x^k)_{k \geq 0})$.

Remark 14 Suppose that $\frac{f + \delta_S}{g}$ is coercive, that is

$$\lim_{\|u\| \rightarrow +\infty} \left(\frac{f + \delta_S}{g} \right)(u) = +\infty.$$

Then the sequence $(x^k)_{k \geq 0}$ generated by Algorithm 9 is bounded. Indeed, this follows from the fact that $(\theta_k)_{k \geq 1}$ is nonincreasing and the lower level sets of $\frac{f + \delta_S}{g}$ are bounded.

We give now the main result concerning the convergence of the whole sequence $(x^k)_{k \geq 0}$.

Theorem 15 *In the setting of Problem 8, let \mathcal{H} be finite dimensional, ∇g be L -Lipschitz continuous, and consider the sequences generated by Algorithm 9 under the additional conditions $\liminf_{k \rightarrow +\infty} \eta_k > 0$, $\eta_1 \theta_1 < \frac{1}{L}$ and $(\eta_k)_{k \geq 1}$ nonincreasing. Assume that $\frac{f + \delta_S}{g}$ is a KL function. Moreover, suppose that $(x^k)_{k \geq 0}$ is bounded and there exists $k_0 \geq 0$ such that $x^k \in \text{int}(\text{dom } f \cap S)$ for all $k \geq k_0$. Then the following statements are true:*

(a) $\sum_{k \geq 0} \|x^{k+1} - x^k\| < +\infty$;

(b) there exists $x_\infty \in \text{dom } f \cap S$ such that $\lim_{k \rightarrow +\infty} x^k = x_\infty$. If additionally $x_\infty \in \text{int}(\text{dom } f \cap S)$, then $0 \in \partial_L \left(\frac{f + \delta_S}{g} \right) (x_\infty)$.

Proof. (a) Consider the sequences generated by Algorithm 9. According to Lemma 13 we can choose an element $\bar{x} \in \omega((x^k)_{k \geq 0})$. By Theorem 10(ii), we have $\bar{x} \in \text{dom } f \cap S$ and $\lim_{k \rightarrow +\infty} \theta_k = \frac{f(\bar{x})}{g(\bar{x})}$. We separately treat the following two cases.

I. There exists $\bar{k} \geq 1$ such that $\theta_{\bar{k}} = \frac{f(\bar{x})}{g(\bar{x})}$. Since $(\theta_k)_{k \geq 1}$ is nonincreasing, we have $\theta_k = \frac{f(\bar{x})}{g(\bar{x})}$ for every $k \geq \bar{k}$. By using (26), we deduce that the sequence $(x^k)_{k \geq \bar{k}}$ is constant. From here the conclusion follows automatically.

II. For all $k \geq 1$ it holds $\theta_k > \frac{f(\bar{x})}{g(\bar{x})}$. Take $\Omega := \omega((x^k)_{k \geq 0})$.

In virtue of Lemma 13(c) and (d) and Lemma 12, the KL property of $\frac{f + \delta_S}{g}$ leads to the existence of positive numbers ε and η and a concave function $\varphi \in \Phi_\eta$ such that for all

$$x \in \{u \in \mathcal{H} : \text{dist}(u, \Omega) < \varepsilon\} \cap \left\{ u \in \text{dom } f \cap S : \frac{f(\bar{x})}{g(\bar{x})} < \frac{f(u)}{g(u)} < \frac{f(\bar{x})}{g(\bar{x})} + \eta \right\} \quad (31)$$

one has

$$\varphi' \left(\frac{f(x)}{g(x)} - \frac{f(\bar{x})}{g(\bar{x})} \right) \text{dist} \left(0, \partial_L \left(\frac{f + \delta_S}{g} \right) (x) \right) \geq 1. \quad (32)$$

Let $k_1 \geq 0$ be such that $\theta_k < \frac{f(\bar{x})}{g(\bar{x})} + \eta$ for all $k \geq k_1$. According to Lemma 13(b), there exists $k_2 \geq 0$ such that $\text{dist}(x^k, \Omega) < \varepsilon$ for all $k \geq k_2$.

Hence the sequence $(x^k)_{k \geq \bar{k}}$, where $\bar{k} = \max\{k_1, k_2\}$, belongs to the intersection (31). So we have (see (32))

$$\varphi' \left(\frac{f(x^k)}{g(x^k)} - \frac{f(\bar{x})}{g(\bar{x})} \right) \|x^*\| \geq 1 \quad \forall x^* \in \partial_L \left(\frac{f + \delta_S}{g} \right) (x^k) \quad \forall k \geq \bar{k}. \quad (33)$$

Since φ is concave, it holds for all $x^* \in \partial_L \left(\frac{f + \delta_S}{g} \right) (x^k)$ and for all $k \geq \bar{k}$

$$\begin{aligned} \varepsilon_k &:= \varphi \left(\frac{f(x^k)}{g(x^k)} - \frac{f(\bar{x})}{g(\bar{x})} \right) - \varphi \left(\frac{f(x^{k+1})}{g(x^{k+1})} - \frac{f(\bar{x})}{g(\bar{x})} \right) \\ &\geq \varphi' \left(\frac{f(x^k)}{g(x^k)} - \frac{f(\bar{x})}{g(\bar{x})} \right) \cdot \left(\frac{f(x^k)}{g(x^k)} - \frac{f(x^{k+1})}{g(x^{k+1})} \right) \\ &= \varphi' \left(\frac{f(x^k)}{g(x^k)} - \frac{f(\bar{x})}{g(\bar{x})} \right) \cdot (\theta_{k+1} - \theta_{k+2}) \\ &\geq \frac{1}{\|x^*\|} \cdot (\theta_{k+1} - \theta_{k+2}) \\ &\geq \frac{1}{\|x^*\|} \cdot \frac{1}{\eta_{k+1} g(x^{k+1})} \|x^{k+1} - x^k\|^2, \end{aligned} \quad (34)$$

where the last inequality follows from (26).

Further, by using (24) and [20, Corollary 1.111(i)], we have that for every $k \geq k_0$

$$x_k^* := -\frac{\nabla g(x^k)}{g^2(x^k)} (f + \delta_S)(x^k) + \frac{1}{g(x^k)} \left[\frac{1}{\eta_k} (x^{k-1} - x^k) + \theta_k \nabla g(x^{k-1}) \right] \in \partial_L \left(\frac{f + \delta_S}{g} \right) (x^k).$$

Furthermore, notice that

$$x_k^* = \frac{1}{g(x^k)} \left[-\theta_{k+1} \nabla g(x^k) + \theta_k \nabla g(x^{k-1}) + \frac{1}{\eta_k} (x^{k-1} - x^k) \right]$$

and relying on the Lipschitz continuity of the gradient we derive

$$\|x_k^*\| \leq \frac{1}{g(x^k)} \left[\left(\frac{1}{\eta_k} + \theta_k L \right) \|x^k - x^{k-1}\| + (\theta_k - \theta_{k+1}) \|\nabla g(x^k)\| \right].$$

Altogether, from (35) we obtain for every $k \geq \max\{\bar{k}, k_0\}$

$$\varepsilon_k \geq \frac{g(x^k)}{g(x^{k+1})} \cdot \frac{\frac{1}{\eta_{k+1}} \|x^{k+1} - x^k\|^2}{\left(\frac{1}{\eta_k} + \theta_k L \right) \|x^k - x^{k-1}\| + (\theta_k - \theta_{k+1}) \|\nabla g(x^k)\|}$$

and from here

$$\begin{aligned} \|x^{k+1} - x^k\| &\leq \sqrt{\eta_{k+1} \left[\left(\frac{1}{\eta_k} + \theta_k L \right) \|x^k - x^{k-1}\| + (\theta_k - \theta_{k+1}) \|\nabla g(x^k)\| \right] \frac{g(x^{k+1})}{g(x^k)} \varepsilon_k} \\ &\leq \frac{\eta_{k+1}}{2} \left[\left(\frac{1}{\eta_k} + \theta_k L \right) \|x^k - x^{k-1}\| + (\theta_k - \theta_{k+1}) \|\nabla g(x^k)\| \right] + \frac{g(x^{k+1})}{2g(x^k)} \varepsilon_k. \end{aligned} \quad (36)$$

Further, we observe that

$$\frac{\eta_{k+1}}{2} \left(\frac{1}{\eta_k} + \theta_k L \right) \leq \frac{\eta_k}{2} \left(\frac{1}{\eta_k} + \theta_k L \right) = \frac{1}{2} + \frac{\eta_k \theta_k L}{2} \leq \frac{1}{2} + \frac{\eta_1 \theta_1 L}{2} \quad \forall k \geq 1.$$

Moreover, $(\nabla g(x^k))_{k \geq 0}$ is bounded and $\limsup_{k \rightarrow +\infty} \frac{g(x^{k+1})}{g(x^k)} < +\infty$, due to $g(x^{k+1}) \leq M$ ($M > 0$) and $\liminf_{k \rightarrow +\infty} g(x^k) > 0$, which follows from the continuity of g , the fact that $(x^k)_{k \geq 0}$ is bounded and Theorem 10(ii). Thus there exist some positive constants $C_1, C_2 > 0$ and $k' \geq 0$ such that

$$\|x^{k+1} - x^k\| \leq \left(\frac{1}{2} + \frac{\eta_1 \theta_1 L}{2} \right) \|x^k - x^{k-1}\| + C_1 (\theta_k - \theta_{k+1}) + C_2 \varepsilon_k \quad \forall k \geq k'.$$

The conclusion follows from Lemma 3 by noticing that $(\theta_k)_{k \geq 1}$ and φ are bounded from below.

(b) It follows from (a) that $(x^k)_{k \geq 0}$ is a Cauchy sequence, hence it is convergent. The conclusion follows from Theorem 10. \blacksquare

3 Future work

We point out some open questions to be followed in the future related to the solving of the fractional programming problem under investigation:

1. Is it possible to evaluate in each iteration the functions f and δ_S separately, which would actually mean that the set S is addressed in the algorithm by means of its projection operator?
2. How to incorporate in Algorithm 6 some extrapolation terms in the sense of Nesterov in order to improve its speed of convergence?
3. Can one consider also other situations, for instance when f is smooth and g is nonsmooth, or even the more general case where both functions are nonsmooth?

References

- [1] H. Attouch, J. Bolte, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*, Mathematical Programming 116(1-2) Series B, 5–16, 2009
- [2] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Lojasiewicz inequality*, Mathematics of Operations Research 35(2), 438–457, 2010
- [3] H. Attouch, J. Bolte, B.F. Svaiter, *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*, Mathematical Programming 137(1-2) Series A, 91–129, 2013
- [4] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011
- [5] J. Bolte, A. Daniilidis, A. Lewis, *The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems*, SIAM Journal on Optimization 17(4), 1205–1223, 2006
- [6] J. Bolte, A. Daniilidis, A. Lewis, M. Shota, *Clarke subgradients of stratifiable functions*, SIAM Journal on Optimization 18(2), 556–572, 2007
- [7] J. Bolte, A. Daniilidis, O. Ley, L. Mazet, *Characterizations of Lojasiewicz inequalities: subgradient flows, talweg, convexity*, Transactions of the American Mathematical Society 362(6), 3319–3363, 2010
- [8] J. Bolte, S. Sabach, M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Mathematical Programming Series A (146)(1–2), 459–494, 2014
- [9] J.M. Borwein, J.D. Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, 2010
- [10] R.I. Boţ, E.R. Csetnek, S. László, *An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions*, EURO Journal on Computational Optimization 4(1), 3–25, 2016
- [11] E. Chouzenoux, J.-C. Pesquet, A. Repetti, *Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function*, Journal of Optimization Theory and its Applications 162(1), 107–132, 2014
- [12] F. Clarke, *Functional analysis, Calculus of Variations and Optimal Control*, Graduate Texts in Mathematics 264, Springer, London, 2013
- [13] J.-P. Crouzeix, J.A. Ferland, S. Schaible, *An algorithm for generalized fractional programs*, Journal of Optimization Theory and Applications 47(1), 35–49, 1985
- [14] W. Dinkelbach, *On nonlinear fractional programming*, Management Science 13, 492–498, 1967

- [15] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, North-Holland Publishing Company, Amsterdam, 1976
- [16] T. Ibaraki, *Solving mathematical programming problems with fractional objective functions*, in S. Schaible and W.T. Ziemba (ed.): *Generalized Concavity in Optimization and Economics*, Academic Press, New York-London, 441–472, 1981
- [17] T. Ibaraki, *Parametric approaches to fractional programs*, *Mathematical Programming* 26(3), 345–362, 1983
- [18] K. Kurdyka, *On gradients of functions definable in o-minimal structures*, *Annales de l'institut Fourier (Grenoble)* 48(3), 769–783, 1998
- [19] S. Łojasiewicz, *Une propriété topologique des sous-ensembles analytiques réels*, *Les Équations aux Dérivées Partielles*, Éditions du Centre National de la Recherche Scientifique Paris, 87–89, 1963
- [20] B. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer-Verlag, Berlin, 2006
- [21] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Kluwer Academic Publishers, Dordrecht, 2004
- [22] R.T. Rockafellar, R.J.-B. Wets, *Variational Analysis*, *Fundamental Principles of Mathematical Sciences* 317, Springer-Verlag, Berlin, 1998
- [23] S. Schaible, *Fractional programming II. On Dinkelbach's algorithm*, *Management Science* 22(8), 868–873, (1975/76)
- [24] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002