Computational Optimization and Applications, 1, 307–326 (1992)

(28)

Convergence Criteria for Generalized Gradient Methods of Solving Locally Lipschitz Feasibility Problems

DAN BUTNARIU*
Department of Mathematics and Computer Science, Haifa University, 31905 Haifa, Israel.

ABRAHAM MEHREZ
The Faculty of Engineering Sciences, Ben-Gurion University of the Negev, 84105 Beer-Sheva, Israel.

Received December 23, 1991, Revised June 23, 1992

Abstract. We prove the convergence of a class of iterative algorithms for solving locally Lipschitz feasibility problems, that is, finite systems of inequalities $f_i(x) \leq 0$, $(i \in I)$, where each f_i is a locally Lipschitz functional on $\mathbb{R}^n$. We also obtain a new convergence criterion for the so-called block-iterative projection methods of finding common points of finite families of convex closed subsets of $\mathbb{R}^n$ as defined by Aharoni and Censor ([3]).

Keywords: Locally Lipschitz feasibility problem, convex feasibility problem, (Clarke) generalized derivative, (Clarke) generalized gradient, regularity point, projection method, subgradient method, generalized gradient method.

1. Introduction

A *feasibility problem* is a problem of computing solutions of a system of inequalities

$$f_i(x) \leq 0, \qquad (i \in I), \tag{1}$$

where I is a finite set and, for each $i \in I$, f_i is a continuous real functional on $\mathbb{R}^n$. The feasibility problem (1) is called *locally Lipschitz* (respectively, *convex*) if *all* functionals f_i, $(i \in I)$, are locally Lipschitz (respectively, convex).

Feasibility problems are of interest in fields like optimization (see [5], [7], [26], [32], [45] and [46]), computerized tomography (cf. [13], [14], [18] and [47]), game theory (cf. [9], [22] and [23]), computer-aided design (cf. [36]), location theory ([31]), etc.

Feasibility problems encountered in practice are frequently locally Lipschitz. For instance, *convex* feasibility problems appearing in optimization and computerized tomography (see [5], [11], [14] and [47]), the *convex-concave* feasibility problem appearing in the "pixel by pixel model" of radioactive emission ([13]),

*The work of Dan Butnariu was done while visiting the Department of Mathematics of the University of Texas at Arlington.

the problem of determining the starting point of the optimization procedure devised in [28], and the (one- or two-way) transport-efficient emplacement of distribution and storage centers are all locally Lipschitz feasibility problems of practical interest (see Section 5.5 below).

In this note we study the convergence of a class of "generalized gradient methods" for solving locally Lipschitz feasibility problems given in the form (1). A sequence $\{x^k | \kappa \in I\!N\}$ in $\mathbb{R}^n$ is called *generated by the generalized gradient method* (for short, *GGM-generated*) if it is recursively defined as follows: Choose an *initial point* $x^0 \in \mathbb{R}^n$ and, for each nonnegative integer κ, choose

$$x^{k+1} \in x^k - \lambda_k \cdot \Sigma_{i \in I} w_k(i) \cdot \partial f_i(x^k), \tag{2}$$

where $\partial f_i(\cdot)$ denotes the Clarke-generalized gradient multifunction of f_i ($\partial f_i(x^k)$ is necessarily nonempty since the functionals f_i are locally Lipschitz (cf. [16, Proposition 2.1.2]), $\lambda_k \in \mathbb{R}_+$ is called *relaxation parameter* and $w_k : I \longrightarrow \mathbb{R}_+$ is a *weight function* (i.e., it satisfies $\Sigma_{i \in I} w_k(i) = 1$). The question is whether and in which conditions GGM–generated sequences converge to solutions of the locally Lipschitz feasibility problem from which they are derived.

Theorem 1, given in Section 2, shows that, in some circumstances, GGM–*properly generated* sequences (i.e., GGM–generated sequences such that, for each $\kappa \in I\!N$, $w_k(i) = 0$, whenever $f_i(x^k) \neq \max_{i \in I} f_i(x^k)$) converge to solutions of the feasibility problem (1) provided that all functionals f_i are locally Lipschitz and their *envelope* $f(x) := \max_{i \in I} f_i(x)$ is globally Lipschitz, that is, the Clarke-generalized gradient multifunction ∂f is uniformly bounded. Recall that the multifunction ∂f is called *uniformly bounded* if there exists an $M > 0$ such that $\|y\| \leq M$ for all $y \in \partial f(x)$ and for each $x \in \mathbb{R}^n$. A number M satisfying this condition is usually called a *uniform bound* of the multifunction ∂f.

In general, any feasibility problem given in the form (1) and involving continuous functionals f_i can be equivalently rewritten as an *intersection problem in normal form*, i.e., as a system of inequalities

$$d_i(x) \leq 0, (i \in I),$$

where, for each $i \in I$, $d_i(x)$ denotes the distance of x to the closed set

$$C_i := \{x \in \mathbb{R}^n | f_i(x) \leq 0\}. \tag{3}$$

Intersection problems in normal form are *globally* Lipschitz feasibility problems, and, thus, the envelope of the functionals d_i involved in them is globally Lipschitz too. Therefore, theoretically speaking, Theorem 1 provides tools for solving large classes of intersection problems in normal form and, hence, large classes of feasibility problems involving functionals that are not even locally Lipschitz but only continuous. However, keeping in mind that computing the distances $d_i(x)$ may sometimes be quite difficult and that the algorithms whose convergence is guaranteed by Theorem 1 may require such computations quite often, it is more practical to apply Theorem 1 and the *synchronized maximal-functions reduction*

(SMFR) algorithms (see Section 2) whose convergence follows from it, directly to the original problem (1) whenever this is possible.

The uniform boundedness condition for the generalized gradient multifunction ∂f which is involved in Theorem 1 via the globally Lipschitz continuity requirement *is not* a necessary condition for the convergence of GGM-properly generated sequences to solutions of the feasibility problem (1). Theorem 2, which is a new convergence criterion for a class of "block-iterative projection methods" (in the sense given to this term in [3]), shows that, in some cases, GGM-properly generated sequences approximate solutions of the feasibility problems from which they are derived even if the generalized gradient multifunction ∂f is not uniformly bounded (and, hence, f is not globally Lipschitz).

Recall that "projection methods" are algorithms of computing points in the intersection of a finite family $\{Q_i | i \in I\}$ of convex closed subsets of $\mathbb{R}^n$. A *projection method* is a GGM-generated sequence applied to a feasibility problem in the form (1) in which the functionals f_i are defined by $f_i(x) = d^2_{Q_i}(x)$, with $d^2_{Q_i}(x)$ the square of the distance of x to Q_i, $(i \in I)$. In this case, the functionals f_i are convex and differentiable, and their gradients at the point $x \in \mathbb{R}^n$ are determined (cf. [4, p. 24]) by the projections $P_{Q_i}(x)$ of the point x on the convex closed sets Q_i via the equations:

$$\nabla d^2_{Q_i}(x) = 2 \cdot (x - P_{Q_i}(x)). \tag{4}$$

In Theorem 2 we take advantage of these particularities of the functionals $d^2_{Q_i}$ in order to show that a significant class of projection methods that are neither "simultaneous" (see [33]) nor necessarily "sequential" (see [3]) converge to solutions of the given intersection problem.

From a historical point of view, the interest in GGM-generated sequences as methods of solving feasibility problems goes back to Fourier ([25]) and Cauchy ([10]). Cauchy has shown that, if the functionals f_i are continuously differentiable and also satisfy some other hypothesis, then particular GGM-generated sequences approximate solutions of the given feasibility problem. More recently, a considerable amount of research was done on the so-called *subgradient methods* of solving *convex* feasibility problems either explicitly or implicitly while analyzing specific convex optimization procedures (see [5], [6], [11], [12], [21], [22], [26], [27], [28], [33], [34], [36], [37], [40], [41], [44], [45], [46], and the references therein). Theorem 1, while applied to convex feasibility problems, can be compared with the main results in [21] and [12] (see Section 5.5). The SMFR algorithms whose convergence is guaranteed by Theorem 1 share with the subgradient methods of solving convex feasibility problems discussed in [21] and [12] the property that they converge to a global minimizer of the functional $f^+(x) = \max(0, f(x))$. However, there are significant differences between the subgradient methods discussed in [21] and [12] and the SMFR algorithms.

First, SMFR algorithms approximate solutions of feasibility problems even if some of the functionals f_i are not convex (Section 5.5., (b)). Second, SMFR algorithms are not "cyclic" or "sequential," that is, they do not reduce the values of the functionals f_i in a cyclical or sequential order but, at each stage, they simultaneously act on all functionals f_i that achieve the value of the envelope f at the latest computed x^k and this improves the initial speed of convergence (see Sections 5.1. and 5.4.).

Theorem 2 is a new addition to the series of convergence criteria for projection methods discussed in [1], [2], [3], [5], [6], [7], [8], [11], [15], [17], [18], [19], [20], [24], [25], [27], [29], [30], [33], [34], [35], [37], [38], [39], and [47], and it can be seen as a generalization of similar results obtained in [1] and [37] for systems of linear inequalities. The *synchronized maximal-distance reduction* (SMDR) algorithms whose convergence is guaranteed by Theorem 2 (see Section 2) are not "simultaneous" as those in [33] and may not be "almost simultaneous" (see [8]) either. The SMDR projection methods combine the accelerated convergence properties achieved by the block-iterative simultaneity of the algorithms in [33] and [8] with the advantage of reducing the sometimes computationally costly task of determining projections on convex closed sets (see Section 5.3.) by allowing a wider range of choices for the relaxation parameters and, thus, an improved initial convergence speed (cf. Section 5.1).

2. Statement of results

In what follows we consider the locally Lipschitz feasibility problem (1) under the assumption that its solution set $C := \bigcap_{i\in I} C_i$, where the sets C_i are defined by (3), has nonempty interior.

For each $x \in \mathbb{R}^n$ denote

$$I(x) := \{i \in I | f_i(x) = f(x)\}.$$

A weight function $w : I \longrightarrow \mathbb{R}_+$ is called *proper at the point* $x \in \mathbb{R}^n$ if for any $j \in I \backslash I(x)$ we have $w(j) = 0$. Since, for any $x \in \mathbb{R}^n$, $I(x) \neq \emptyset$, it follows that for any $x \in \mathbb{R}^n$ there exists at least one weight function $w_x : I \longrightarrow \mathbb{R}_+$ that is proper at x, namely, the weight function defined by $w_x(i) = 1/|I(x)|$ if $i \in I(x)$ and $w_x(i) = 0$ otherwise.

A GGM-generated sequence $\{x^k | \kappa \in \mathbb{N}$ is called *GGM–properly generated* if, for each $\kappa \in \mathbb{N}$ the weight function w_k involved in (2) is proper at x^k. Obviously, for locally Lipschitz feasibility problems, GGM–properly generated sequences always exist. Therefore, we can state the following:

THEOREM 1. *Suppose that the locally Lipschitz feasibility problem (1) is such that the envelope f of the family of functionals f_i, $(i \in I)$, is convex and globally Lipschitz. If $\{x^k | \kappa \in \mathbb{N}\}$ is a GGM–properly generated sequence with relaxation parameters λ_k satisfying the condition*

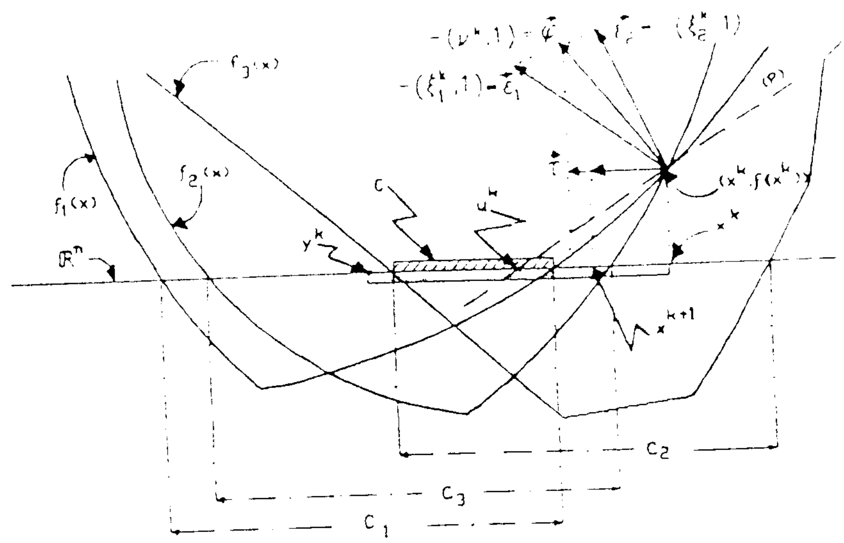

Figure 1. Geometric representation of an iterative step.

$$\max(0, f(x^k)) \leq \lambda_k M^2 \leq 2 \cdot \max(0, f(x^k)), \qquad (\kappa \in \mathbb{N}), \tag{5}$$

for some Lipschitz constant M of the functional f, and if, for each $\kappa \in \mathbb{N}$ and for each $i \in I(x^k)$, the functional f_i is regular at x^k, then the sequence $\{x^k | \kappa \in \mathbb{N}\}$ converges to a solution of the feasibility problem (1).

Theorem 1 is proven in Section 3. Note that, under the assumptions of Theorem 1, the functionals f_i are not required to be globally Lipschitz or convex, even if their envelope f is required to be so. In fact, if each f_i is globally Lipschitz, then so is f, but the converse holds under additional conditions only as, for instance, when each f_i is everywhere regular as follows from [16, Proposition 2.3.12] (everywhere regularity is not implied by the hypothesis of Theorem 1).

The way in which the GGM–properly generated algorithms satisfying (5) proceed at each iterative step is illustrated in Figure 1 where the algorithms are interpreted in $\mathbb{R}^{n+1}$. The main idea is that, at each iterative step κ, either $f(x^k) \leq 0$ and [because of (5)] the algorithm stops at the solution $x^* = x^k$ of the given feasibility problem, or $f(x^k) > 0$ and then a new iterate x^{k+1} is chosen in the interval $[y^k, x^k)$ of the line L passing through $(x^k, 0)$ in the direction $\vec{\tau} = (-\nu^k, 0)$, where $\nu^k := \Sigma_{i \in I} w_k(i) \xi_i^k$ with $\xi_i^k \in \partial f_i(x^k)$ being the generalized gradients chosen at step κ (see (2)). The point y^k is the symmetric of x^k with respect to u^k, where u^k is the intersection of the line L with the hyperplane (P) of normal $\vec{\varphi} = (-\nu^k, -1)$ passing through $(x^k, f(x^k))$.

Note that, moving from an iterative step x^k to the next in a GGM generated sequence whose relaxation parameters are satisfying (5) is, essentially, a move toward synchronized reduction of the values of all functionals f_i with $i \in I(x^k)$, while condition (5) determines the possible range of this reduction. Therefore, and in order to distinguish the GGM–generated sequences whose convergence is guaranteed by Theorem 1 from other generalized gradient methods of solving feasibility problems discussed in literature, we call *synchronized maximal-function reduction* (SMFR) *algorithms* associated to the problem (1) any GGM–properly generated sequence $\{x^k | \kappa \in \mathbb{N}\}$ with relaxation parameters chosen according to (5) and consisting of regularity points of all functionals f_i, $(i \in I)$. Now, Theorem 1 can be reinterpreted as saying that, if the feasibility problem (1) is locally Lipschitz, if the envelope f is convex and globally Lipschitz, then SMFR algorithms consisting of regularity points of the functionals f_i converge to a solution of (1). Significant classes of feasibility problems whose solutions can be approximated by SMFR algorithms are pointed out in Section 5.5. below.

A restrictive condition involved in Theorem 1 is the requirement that the envelope f should be globally Lipschitz. Recall that a functional is globally Lipschitz if and only if it is locally Lipschitz and has uniformly bounded generalized gradient multifunction. The local Lipschitz continuity of f is essential in Theorem 1 since it follows from the local Lipschitz continuity of the functions f_i, which guarantees the existence of GGM–generated sequences. By contrast, the uniform boundedness of the generalized gradient multifunction $\partial f(x)$ is not a necessary but a sufficient requirement for ensuring convergence of GGM–properly generated sequences to solutions of the feasibility problem from which they are derived.

The following result shows that, if one considers a "convex intersection problem," i.e., the problem of approximating common points of a family $\{Q_i | i \in I\}$ of closed convex sets that is equivalently represented as a locally Lipschitz feasibility problem (1) with $f_i(x) := d^2_{Q_i}(x)$, $(i \in I)$, then GGM–properly generated sequences may converge to solutions of the intersection problem even if the generalized gradient $\partial f(x)$ may be unbounded. Recall that convex functionals are everywhere regular and locally Lipschitz, and, since each $f_i = d^2_{Q_i}$ is convex, $\partial f(x)$ includes all convex combinations of gradients ∇f_i given by (4) that are not necessarily bounded.

THEOREM 2. *Let Q_i, $(i \in I)$, be a family of closed convex subsets of $\mathbb{R}^n$ and suppose that $Q = \bigcap_{i \in I} Q_i$ has nonempty interior. If $\{x^k | \kappa \in \mathbb{N}\}$ is a GGM–properly generated sequence for the feasibility problem*

$$d^2_{Q_i}(x) \leq 0, \qquad (i \in I) \tag{6}$$

with relaxation parameters $\mu_k \in [\alpha, 1]$, $(\kappa \in \mathbb{N})$, for some $\alpha \in]0, 1]$, then the sequence $\{x^k | \kappa \in \mathbb{N}\}$ converges to a point in Q no matter how the initial point x^0 is chosen.

The proof of Theorem 2 is given in Section 4 and it is essentially based on some results in [8] that are themselves proven using properties of simultaneous projection operators shown in [33]. GGM-properly generated sequences as described in Theorem 2 are recursively defined by

$$x^{k+1} = x^k - 2 \cdot \mu_k \cdot \Sigma_{i \in I} w_k(i) \cdot (x^k - P_{Q_i}(x^k)),$$

where $x^0 \in \mathbb{R}^n$ and, for each $\kappa \in \mathbb{N}$, w_k is proper at x^k (with respect to the functionals $d_{Q_i}^2$) and $\mu_k \in [\alpha, 1]$. In what follows such sequences are called *synchronized maximal-distance reduction methods* (SMDR) because they are built for synchronized reduction of the maximal distances of the actual iterate to the sets Q_i.

3. Proof of Theorem 1

The proof of Theorem 1 is done in several steps. We start with the following:

LEMMA 1. *Let $\{y^k|\kappa \in \mathbb{N}\}$ be a convergent sequence in $\mathbb{R}^n$ and let $\overline{y}$ be its limit. If $\{w_k|\kappa \in \mathbb{N}\}$ is a sequence of weight functions such that, for each $\kappa \in \mathbb{N}$, w_k is proper at y^k, then each cluster point of $\{w_k|\kappa \in \mathbb{N}\}$ in $\mathbb{R}^I$ is a weight function that is proper at $\overline{y}$.*

Proof. Since sequences of weight functions on I are bounded in $\mathbb{R}^I$, it follows that they always have cluster points, and, obviously, cluster points of a sequence of weight functions are weight functions. It remains to show that, if $\{w_k|\kappa \in \mathbb{N}\}$ is a sequence of weight functions such that, for each $\kappa \in \mathbb{N}$, w_k is proper at y^k, and if w_* is a cluster point of $\{w_k|\kappa \in \mathbb{N}\}$, then w_* is proper at $\overline{y}$. To this end, let $i \in I \setminus I(\overline{y})$ and consider a subsequence $\{w_{k_p}|p \in \mathbb{N}\}$ that converges to w_*. Then $f_i(\overline{y}) < f(\overline{y})$. Since the functionals f_i are continuous, there exists a neighborhood U of $\overline{y}$ such that $f_i(y) < f(y)$ for all $y \in U$. Therefore, there exists a positive integer κ^* such that $y^k \in U$ and, hence, such that $f_i(x^k) < f(y^k)$, for all integers $\kappa \geq \kappa^*$. Thus, $i \in I(y^k)$ for all integers $\kappa \geq \kappa^*$. Hence, $w_{k_p}(i) = 0$ whenever $\kappa_p \geq \kappa^*$, since w_{k_p} is proper at y^{k_p}. By consequence, $w_*(i) = \lim_{p \to \infty} w_{k_p}(i) = 0$ and the proof is complete. □

Now, consider an SMFR algorithm $\{x_k|\kappa \in \mathbb{N}\}$, i.e., a GGM-properly generated sequence consisting of regularity points of all functionals f_i, $(i \in I)$, recursively defined [see (2)] by

$$x^{k+1} = x^k - \lambda_k \cdot \nu^k$$

where $x^0 \in \mathbb{R}^n$ and, for each $\kappa \in \mathbb{N}$, λ_k satisfies (5) and

$$\nu^k := \Sigma_{i \in I} w_k(i) \cdot \xi_i^k,$$

with $\xi_i^k \in \partial f_i(x^k)$, $(i \in I)$, and w_k a weight function that is proper at x^k.

LEMMA 2. *For each $z \in C$ and for any $\kappa \in \mathbb{N}$ we have*

$$\|x^{k+1} - z\| \leq \|x^k - z\|. \tag{7}$$

Proof. Note that

$$\|x^{k+1} - z\|^2 = \|x^k - z\|^2 + \lambda_k\left[\lambda_k\|\nu^k\|^2 - 2 \cdot \langle \nu^k, x^k - z\rangle\right]. \tag{8}$$

If $\lambda_k = 0$ or $\nu^k = 0$, then (7) clearly holds with equality. Assume that $\lambda_k \neq 0$ and $\nu^k \neq 0$. Then, according to (5), we have $\lambda_k > 0$ and $f(z) \leq 0 < f(x^k)$ for any $z \in C$ .Since f is convex, f is also regular on $\mathbb{R}^n$ (cf. [16, Proposition 2.3.6(b)]) and, therefore, the (Clarke-) generalized derivative of f in the direction $z - x^k$ exists and it is given by

$$f^o(x^k, z - x^k) = \lim_{t \searrow 0} \frac{f(x^k + t(z - x^k)) - f(x^k)}{t}. \tag{9}$$

Also, for any $t \in [0, 1]$ we have

$$f(x^k + t(z - x^k)) - f(x^k) \leq t(f(z) - f(x^k)) \tag{10}$$

because f is convex. From (9) and (10) it follows that

$$f^o(x^k, z - x^k) \leq f(z) - f(x^k) < 0, \qquad (z \in C). \tag{11}$$

Since each f_i is regular at x^k Proposition 2.3.12 in [16] applies and we have

$$\partial f(x^k) = \text{conv}\left\{\bigcup_{i \in I(x^k)} \partial f_i(x^k)\right\}. \tag{12}$$

Hence, for each $\kappa \in \mathbb{N}$, $\nu^k \in \partial f(x^k)$, since the sequence $\{x^k | \kappa \in \mathbb{N}\}$ is properly generated (i.e., w_k is proper at x^k). By consequence, we have

$$\langle \nu^k, z - x^k\rangle \leq f^o(x^k, z - x^k), \qquad (z \in C). \tag{13}$$

Combining (11) and (13) we obtain

$$\langle \nu^k, x^k - z\rangle \geq -f^o(x^k, z - x^k) \geq f(x^k) - f(z) \geq f(x^k) > 0,$$

for any $z \in C$. From that and from (5) and (12) we get

$$2 \cdot \langle \nu^k, x^k - z\rangle \geq 2 \cdot f(x^k) \geq \lambda_k M^2 \geq \lambda_k\|\nu^k\|^2, \qquad (z \in C), \tag{14}$$

because M is a global Lipschitz constant of f and, therefore, a uniform bound of the multifunction ∂f (see [16, Proposition 2.1.2 (a)]). Formula (14) shows that the expression between square brackets in (8) is nonpositive. This completes the proof of the lemma. □

Since $C \neq \emptyset$ and (7) holds, it follows that the sequence $\{\|x^k - z\| \mid \kappa \in \mathbb{N}\}$ is convergent for any $z \in C$ and that the sequence $\{x^k|\kappa \in \mathbb{N}\}$ is bounded. Hence, there exists a convergent subsequence $\{x^{k_p}p \in \mathbb{N}\}$ of the sequence $\{x^k|\kappa \in \mathbb{N}\}$. Let x^* be the limit of this subsequence. The function f is continuous. Hence, the sequence $\{f(x^{k_p})p \in \mathbb{N}\}$ converges to $f(x^*)$. According to (5), the sequence $\{\lambda_{k_p}p \in \mathbb{N}\}$ is bounded because $\{f(x^{k_p}p \in \mathbb{N}\}$ is bounded. The sequence $\{\nu^k|\kappa \in \mathbb{N}\}$ is bounded because of the uniform boundedness of the multifunctions ∂f and because (12) holds for each $\kappa \in \mathbb{N}$. Consequently, using Lemma 1 we deduce that there exists a sequence $\{s_t|t \in \mathbb{N}\}$ of nonnegative integers such that the following limits exist

$$x^* = \lim_{t\to\infty} x^{s_t}, \tag{15}$$

$$\lambda_* = \lim_{t\to\infty} \lambda_{s_t}, \tag{16}$$

$$\nu^* = \lim_{t\to\infty} \nu^{s_t}, \tag{17}$$

$$w_* = \lim_{t\to\infty} w_{s_t}, \tag{18}$$

the weight function w_* is proper at x^* and

$$\nu^* = \Sigma_{i\in I} w_*(i)\xi_i^*, \tag{19}$$

where

$$\xi_i^* = \lim_{t\to\infty} \xi_i^{s_t}, \qquad (i \in I). \tag{20}$$

Since, for each $t \in \mathbb{N}$, we have

$$x^{s_t+1} = x^{s_t} - \lambda_{s_t}\nu^{s_t},$$

it follows that $\{x^{s_t+1}|t \in \mathbb{N}\}$ converges to $x^* - \lambda_* \cdot \nu^*$. On the other hand, for each $z \in C$, $\{\|x^{s_t+1} - z\| \mid t \in \mathbb{N}\}$ and $\{\|x^{s_t} - z\| \mid t \in \mathbb{N}\}$ are subsequences of the same convergent sequence $\{\|x^k - z\| \mid \kappa \in \mathbb{N}\}$ (cf. Lemma 2) and , therefore, they have the same limit, i.e.,

$$\begin{aligned} \|x^* - z\| &= \lim_{t\to\infty} \|x^{s_t} - z\| = \lim_{k\to\infty} \|x^k - z\| \\ &= \lim_{t\to\infty} \|x^{s_t+1} - z\| = \|x^* - \lambda_*\nu^* - z\|, \end{aligned}$$

no matter how z is chosen in C. This implies

$$\|x^* - z\|^2 = \|x^* - z\|^2 + \lambda_*[\lambda_*\|\nu^*\|^2 - 2 \cdot \langle \nu^*, x^* - z\rangle],$$

for any $z \in C$. Hence, for any $z \in C$, we have

$$\lambda_*[\lambda_*\|\nu^*\|^2 - 2 \cdot \langle \nu^*, x^* - z\rangle] = 0. \tag{21}$$

Using this fact, we prove the following

LEMMA 3. *The point x^* defined at (15) belongs to C.*

Proof. Note that, $\nu^* \in \partial f(x^*)$ since, according to (12), we have $\nu^{s_t} \in \partial f(x^{s_t})$, $(t \in \mathbb{N})$, (15) and (18) hold and the subgradient multifunction ∂f is weak closed. Hence, if $\nu^* = 0$, then $0 \in \partial f(x^*)$, that is x^* is a global minimum of f. Therefore, if $\nu^* = 0$, we have $f(x^*) \leq f(z) \leq 0$, $(z \in C)$, and this implies $x^* \in C$.

Assume that $\nu^* \neq 0$. Then, according to (21), we either have $\lambda_* = 0$ or

$$\lambda_* \|\nu^*\|^2 - 2 \cdot \langle \nu^*, x^* - z \rangle = 0, \qquad (z \in C). \tag{22}$$

If (22) holds, then C is a subset of the hyperplane of equation

$$\langle \nu^*, x \rangle = \left(\frac{1}{2}\right) \cdot \left[2 \cdot \langle \nu^*, x^* \rangle - \lambda_* \|\nu^*\|^2\right].$$

This contradicts the assumption that $\mathrm{Int}(C) \neq \emptyset$. Consequently, $\lambda_* = 0$. According to (5), (15) and (17) this implies $f(x^*) \leq 0$, i.e., $x^* \in C$. □

Now, according to Lemmas 1 and 2, the sequence $\{\|x^k - x^*\| \mid \kappa \in \mathbb{N}\}$ converges and it must have the same limit as its subsequence $\{\|x^{s_t} - x^*\| \mid t \in \mathbb{N}\}$, i.e., $\lim_{k\to\infty} \|x^k - x^*\| = 0$. By consequence, the sequence $\{x^k \mid \kappa \in \mathbb{N}\}$ converges to the point x^*, which is a solution of the feasibility problem (1) and this completes the proof of Theorem 1. □

4. Proof of Theorem 2

Let $\{x^k | \kappa \in \mathbb{N}\}$ be an SMDR algorithm, i.e., a GGM–properly generated sequence associated with the problem (6) with relaxation parameters $\mu_k \in [\alpha, 1]$. Note that $\{x^k | \kappa \in \mathbb{N}\}$ is recursively defined by

$$x^{k+1} = x^k + \lambda_k \cdot \Sigma_{i \in I} w_k(i) \cdot (P_{Q_i}(x^k) - x^k), \tag{23}$$

where $\lambda_k := 2\mu_k \in [2\alpha, 2]$, for each $\kappa \in \mathbb{N}$. Hence, Corollary 4.3 in [8] applies to $\{x^k | \kappa \in \mathbb{N}\}$ and it shows that (7) holds for this sequence for each $\kappa \in \mathbb{N}$ and for any $z \in Q$. Since Q is not empty, this implies that the sequence $\{x^k | \kappa \in \mathbb{N}\}$ is bounded. Hence, it has a convergent subsequence $\{x^{k_p} p \in \mathbb{N}\}$. Let x^* be the limit of this subsequence. Since $\{\lambda_{k_p} p \in \mathbb{N}\}$ is bounded by hypothesis and the sequence $\{x^k | \kappa \in \mathbb{N}\}$ is properly generated, it follows (cf. Lemma 1) that there exists a sequence $\{s_t | t \in \mathbb{N}\}$ of nonnegative integers such that (15), (16), and (18) still hold and such that w_* is proper at x^*. Therefore, the sequence $\{x^{s_t+1} | t \in \mathbb{N}\}$ converges to

$$x^* + \lambda_*^* \Sigma_{i \in I} w_*(i) \cdot (P_{Q_i}(x^*) - x^*).$$

Since, for each $z \in Q$, the sequences $\{\|x^{\kappa} - z\| \mid \kappa \in \mathbb{N}\}$ and $\{\|x^{\kappa+1} - z\| \mid \kappa \in \mathbb{N}\}$ must have the same limit because of (7), it follows that, for any $z \in Q$, we have

$$\|x^{*} + \lambda_{*} \Sigma_{i \in I} w_{*}(i) \cdot (P_{Q_i}(x^{*}) - x^{*}) - z\| = \|x^{*} - z\|.$$

This implies

$$\lambda_{*} \left[\lambda_{*} \|\nu^{*}\|^2 + 2 \cdot \langle \nu^{*}, x^{*} - z \rangle\right] = 0, \qquad (z \in Q). \tag{24}$$

where

$$\nu^{*} = \Sigma_{i \in I} w_{*}(i) \cdot (P_{Q_i}(x^{*}) - x^{*}).$$

By hypothesis, $\lambda_{*} \neq 0$. Therefore, it follows from (24) that either $\nu^{*} = 0$ or Q is a subset of a hyperplane in $\mathbb{R}^n$ of normal ν^{*}. Since Q has nonempty interior, the latter is not possible. Hence, $\nu^{*} = 0$.

Now, denote $J := \{i \in I \mid w_{*}(i) > 0\}$ and let w be the restriction of the weight function w_{*} to J. Obviously, w is still a weight function (on J). Since $\nu^{*} = 0$, it follows that x^{*} is a fixed point of the simultaneous projection operator $\mathbf{P}_w : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined by

$$\mathbf{P}_w(x) = \Sigma_{i \in J} w(i) \cdot P_{Q_i}(x).$$

Note that Proposition 2.2 *(ii)* in [8] applies to $\mathbf{P}_w$ and it shows that $x^{*} \in \bigcap_{i \in J} Q_i$. Hence, $d^2_{Q_i}(x^{*}) = 0$ for each $i \in J$. Observe that $J \subseteq I(x^{*})$ because w_{*} is proper at x^{*}. Hence, if $j \in J$, then

$$d^2_{Q_j}(x^{*}) = \max\{d^2_{Q_i}(x^{*}) \mid i \in I\}.$$

Therefore, we have $d^2_{Q_i}(x^{*}) = 0$, for each $i \in I$, and this completes the proof of Theorem 2. □

5. Comments, applications and examples

In the section we discuss some computational aspects of the application of SMFR or SMDR algorithms. Also, we point out significant classes of feasibility problems that can be solved by them.

5.1. The initial speed of convergence of SMFR and SMDR algorithms

Analyzing the proofs of Theorems 1 and 2 above it can be seen that, if the sequence $\{x^k \mid \kappa \in \mathbb{N}\}$ is an SMFR (or an SMDR) algorithm and if the envelope f is convex [a condition automatically satisfied by the functionals in (6)], then the sequence $\{\|x^k - z\| \mid \kappa \in \mathbb{N}\}$ is nonincreasing whenever $z \in C$ (respectively, $z \in Q$). Therefore, the quantity

$$\Delta_k = \inf_z \{ \|x^k - z\| - \|x^{k+1} - z\| \},$$

where z runs over all solutions of the problem (1) [or (6)], is nonnegative. It is called *initial speed of convergence (at step κ)* since it measures how much closer to the solution set one gets by moving from the actual iterate x^k to the next iterate x^{k+1}. By increasing the initial speed of convergence at each step κ we obtain the desired approximation of x^* in fewer iterative steps. Observe that, if $\{x^k | \kappa \in \mathbb{N}\}$ is an SMFR algorithm, then for each $\kappa \in \mathbb{N}$, we have

$$\|x^k - z\|^2 - \|x^{k+1} - z\|^2 = \lambda_k \cdot \left[2 \cdot \langle x^k, x^k - z\rangle - \lambda_k \cdot \|\nu^k\|^2\right]$$

and, consequently,

$$\Delta_k \geq \lambda_k \cdot \inf_z \frac{2 \cdot \langle \nu^k, x^k - z\rangle - \lambda_k \cdot \|\nu^k\|^2}{2 \cdot \|x^k - z\|}.$$

By (11) and (13), this implies

$$\Delta_k \geq \lambda_k \cdot \inf_z \frac{\langle \nu^k, x^k - z\rangle - f(x^k)}{\|x^k - z\|} \tag{25}$$

provided that x^k is not a solution of the feasibility problem. Note that, according to (11), the right-hand side (25) is nonnegative and it increases whenever λ_k increases. This shows that the initial speed of convergence of SMFR algorithms can be eventually improved by choosing relaxation parameters with values as large as possible [see (5)]. This happens to be true for block-iterative projection methods too and, in particular, for SMDR algorithms as shown in [8].

5.2. The global Lipschitz condition for the envelope f

This is essential for our proof of Theorem 1.

(a) The value of the Lipschitz constant M of the function f is involved in operating SMFR algorithms by determining the range of the possible relaxation parameters at each iterative step [cf. (5)]. It follows from (5) and from 5.1. above that determining the "smallest" Lipschitz constant M of f (that is, the Lipschitz rank of f) may help improve the initial speed of convergence and, thus, the overall behavior of SMFR algorithms.

(b) If all functionals f_i involved in (1) are globally Lipschitz, then we can use SMFR algorithms (provided that all other requirements of Theorem 1 are satisfied) with M being the maximum of the Lipschitz ranks of the functionals f_i. There are significant classes of feasibility problems involving globally Lipschitz functionals whose Lipschitz ranks are known a priori or can be easily determined and for which SMFR algorithms efficiently converge.

Intersection problems in normal form and systems of linear inequalities are typical examples in this sense.

5.3. *The regularity condition*

The regularity condition involved in Theorem 1 practically requires the SMFR algorithms to operate in such a way that at each step κ the point x^k is chosen among the regularity points of all functionals f_i involved in (1).

(a) If all f_is are strictly differentiable or convex, then they are necessarily regular at each point in $\mathbb{R}^n$ (cf. [16, Proposition 2.3.6]). Convex functionals on $\mathbb{R}^n$ are also locally Lipschitz (cf. [42]) at each point $x \in \mathbb{R}^n$. Hence, if (1) is a convex feasibility problem involving functionals whose subgradient multifunctions are uniformly bounded, then SMFR algorithms can be applied. This is the case of convex intersection problems in normal form.

(b) The functionals f_i of locally Lipschitz feasibility problems may not be convex or everywhere regular, but they may still have a convex and globally Lipschitz envelope. In such cases, the resolution of (1) via SMFR algorithms can be easily done whenever there are known methods of identifying irregularity points of the functionals f_i, since we can take advantage of the relative freedom of choosing the initial point x^0, the relaxation parameters, and the weight function of the SMFR algorithm in order to ensure that, at each step κ, the actual iterate x^k is a regularity point of all functionals f_i (recall that the set of regularity points of a locally Lipschitz functional is "negligible"). This is the case of the emplacement problems for distribution and storage centers (see 5.5.).

5.4. *Choosing the weight functions in SMFR and in SMDR algorithms*

From the proofs of Theorems 1 and 2 it follows that SMFR and SMDR algorithms are methods of globally minimizing the functional $f^+ = \max(0, f)$, where f is the envelope of the family of functionals f_i (respectively, $d^2_{Q_i}$), $(i \in I)$. The fact that the weight function w_k, at each step κ, has to be proper at x^k is a basic feature of these algorithms. Whenever there are no other requirements to impound upon the choice of the weight functions (supplementary requirements may appear when using SMFR algorithms with functionals f_i that are not everywhere regular as noted in 5.3.), they can be chosen in a way that may improve the computational efficiency of the algorithm. It was noted in Section 1 that SMDR algorithms are block-iterative projection methods and it was observed in [4], [8], and [33] that the freedom of choosing weight functions in such algorithms is a useful tool for block-parallel processing large problems.

The same is true for SMFR algorithms, too. However, there is yet another way of exploiting the relative freedom of choosing weight functions in SMFR and SMDR algorithms. Determining generalized gradients ξ_i^k or projections $P_{Q_i}(x^k)$ can sometimes be a computationally costly process. Therefore, one may be tempted to reduce the need of computing gradients or projections to a minimum. Obviously, one can choose all w_ks as functions with a single nonzero value, in which case the algorithm is called *sequential*. This automatically reduces the need of computing gradients or projections at each step. However, sequential SMFR and SMDR algorithms usually have lower initial speed of convergence than their nonsequential counterparts (this is the effect on sequential algorithms of moving from the actual iterate to the next in directions in which all f_is but one may not decrease or may decrease slowly). Consequently, when using sequential SMFR or SMDR algorithms, one has to perform a sensibly increased (by comparison with nonsequential algorithms) number of steps for reaching a reasonably good approximation of the solution. This considerably increases of the overall number of gradients and/or projections needed in the process. It seems (according to our computational experience) that a better way of achieving the goal of reducing the overall number of gradients and/or projections needed in SMFR/SMDR procedures consists of choosing the weight functions w_k in a way that prevents as many functionals f_i as possible from achieving the value $f(x^{k+1})$ at the next iterative step [i.e., by keeping the cardinality of $I(x^{k+1})$ as low as possible] while still choosing weight functions having positive values for all $i \in I(x^k)$. How this can be practically done strongly depends on the nature of the functionals involved in the given feasibility problem.

5.5. *Classes of problems that can be efficiently solved by SMFR and/or SMDR algorithms*

There are significant classes of feasibility problems that satisfy the requirements of Theorems 1 and/or 2 and can be efficiently solved by SMFR and/or SMDR algorithms.

(a) *Convex feasibility problems* involving functionals f_i whose envelope is globally Lipschitz can be solved via SMFR algorithms. In the particular case when each f_i is globally Lipschitz, such problems can also be solved by (sequential) methods presented in [12] and [21]. SMFR algorithms present the advantage of allowing the choice of relaxation parameters and weight functions in a manner that ensures improved initial convergence speed (cf. 5.1.) and better use of the computational facilities by block-parallel processing (cf. 5.4.).

(b) If the feasibility problem (1) is convex but the envelope f has not uniformly bounded subgradient multifunction (i.e., f is not globally Lipschitz), then the problem can be equivalently rewritten as a convex intersection problem (6) with $Q_i = C_i$, $(i \in I)$, where the sets C_i are defined at (3). In such a case

SMDR algorithms apply provided that $\mathrm{Int}(\bigcap_{i \in I} C_i) \neq \emptyset$. SMDR algorithms are projection methods that share with the methods of the same nature devised in [3], [8], and [33] the property that they strictly Fejer-monotonically (with respect to Q) converge to their limits, which are common fixed points of the operators P_{Q_i}. A phenomenon observed in projection methods is that, while moving from an iterative step to another, the distance of the actual iterate x^k to the set Q may decrease slowly because, for some sets Q_i, the distance $\|P_{Q_i}(x^k) - x^k\|$ practically does not decrease. This cannot happen in SMDR algorithms since, for each $\kappa \in \mathbb{N}$ and for each $j \in I$ we have $\|P_{Q_j}(x^k) - x^k\| \leq \max\{d_i(x^k) | i \in I\}$, implying that, at each iterative step κ, the reduction of all distances $\|P_{Q_i}(x^k) - x^k\|$ is synchronized. In this respect SMDR algorithms are similar to the methods of solving systems of linear inequalities studied in [1] and [37].

(c) *Nonconvex feasibility problems.* The class of nonconvex locally Lipschitz feasibility problems that can be solved by SMFR algorithms is somewhat restricted by the requirement of Theorem 1 that the envelope f of the family of functionals f_i should be convex. In a further paper we shall show that this requirement can be weakened at relatively low cost. In practice, there are significant problems for which this requirement (as well as the other conditions of Theorem 1) is usually satisfied. Among them there are problems of transport efficient emplacement of distribution/storage centers and some convex-concave feasibility problems.

(i) *The problem of transport-efficient emplacement of a two-way distribution and/or storage center.* There are m different commodities denoted $1, 2, \ldots, m$ that should be stored and/or distributed in/from a storage and distribution center whose emplacement should be decided. For each commodity i, there are $p(i)$ given points $y^{i,j} \in \mathbb{R}^n, (1 \leq j \leq p(i))$ to/from which this commodity should be usually transported from/to the center. The price per unit of distance of transporting the commodity i from/to the center to/from the point $y^{i,j}$ is $c_{i,j}$. Some of the prices $c_{i,j}$ may be positive while others are negative. (The positive prices are those paid by the owner of the center as operational expenses. Negative prices are those paid by customers as parts of the commodity distribution cost.) The question is to find a point x where the center should be placed such that the transportation cost of each commodity i will not exceed the affordability limit $c_{i,0}$. This is a feasibility problem that can be represented in the form (1) with the functionals $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by [1]

$$f_i(x) = \Sigma_{j=1}^{p(i)} c_{i,j} \cdot \|x - y^{i,j}\| - c_{i,0}. \tag{26}$$

[1] Note that the typical "one-way" transport-efficient distribution center emplacement problem is that involving (one-way) transportation from the distribution center to consumers placed in the locations $y^{i,j}$. In such a problem all prices $c_{i,j}$ are positive and, consequently, the functions f_i are convex and globally Lipschitz. Hence, (one-way) transport-efficient distribution center emplacement problems can always be solved by SMFR algorithms as shown at (a) above.

for each $i \in I := \{1, 2, \ldots, m\}$. It is reasonable to assume that the envelope f of these functionals is convex [i.e., that placing the center in the point $(a + b)/2$ leads to a maximal transportation cost at most equal to the average of the maximal total transportation costs from a and b] even if some prices $c_{i,j}$ are negative. The functionals f_i defined at (26) may not be convex, but they are globally Lipschitz of ranks $K_i = p(i) \cdot \max_{1 \leq j \leq p(i)} |c_{i,j}|$. Hence, the multifunctions ∂f_i are uniformly bounded (cf. 5.2.). The points at which the functionals f_i may be nonregular are among the $y^{i,j}$'s. Hence, one can apply SMFR algorithms for solving such problems by choosing the initial point x^0, the relaxation parameters, and the weight functions such that none of the points x^k is among the points $y^{i,j}$ (cf. 5.3.). An example of such a problem is given at 5.6. below.

(ii) *Convex-concave feasibility problems* are feasibility problems of form (1), where the set I can be partitioned in two subsets $I = I_1 \cup I_2$ such that the functional f_i is convex when $i \in I_1$ and concave otherwise. Such problems are locally Lipschitz. A typical example is the "pixel by pixel model of emission computerized tomography" presented in [13]. In this particular case the concave functionals are everywhere regular (being continuously differentiable). However, their gradients are unbounded. In another paper we shall show that, in spite of this difficulty, the problem can be solved by adapted SMFR algorithms.

5.6. *A numerical example*

In order to show how SMFR algorithms work while solving a nonconvex nonregular feasibility problem, consider the following particular case of a transport-efficient emplacement problem (see 5.5.). Consider the functionals $f_i : \mathbb{R}^1 \longrightarrow \mathbb{R}$ are defined by $f_1(x) = 6 \cdot |x - 2| - 12$, $f_2(x) = |x - 1| - 2 \cdot |x + 1|$ and $f_3(x) = 2 \cdot |x + 3| - |x - 5| - 10$. The graphs of these functionals are plotted in Figure 2. It is clear that the envelope $f = \max(f_1, f_2, f_3)$ is convex (its plot is the curve marked by "+" in Figure 2). As noted at (a) above, the functionals f_i considered here are globally Lipschitz of ranks 6, 4, and 4, respectively. Thus, we can solve the feasibility problem $f_i(x) \leq 0$, $(i = 1, 2, 3)$, via SMFR algorithms with $M = 6$, while taking care that the choices of weight functions and of relaxation parameters be made in such a way that none of the points x^k to be one of the irregularity points of the functionals f_i (in our case, f_2 is not regular at $x = -1$ and f_3 is not regular at $x = 5$). It is clear that, in this case, $C = [0, 3]$.

We repeatedly solve this problem by always taking $x^0 = 50$ and $w_k(i) = 1/|I(x^k))|$ if $i \in I(x^k)$ and $w_k(i) = 0$ otherwise, while choosing $\lambda_k = \alpha \cdot \max(0, f(x^k))/M^2$ with $\alpha \in \{1, 1.2, 1.4, 1.6, 1.8, 2\}$ as permitted by Theorem 1.

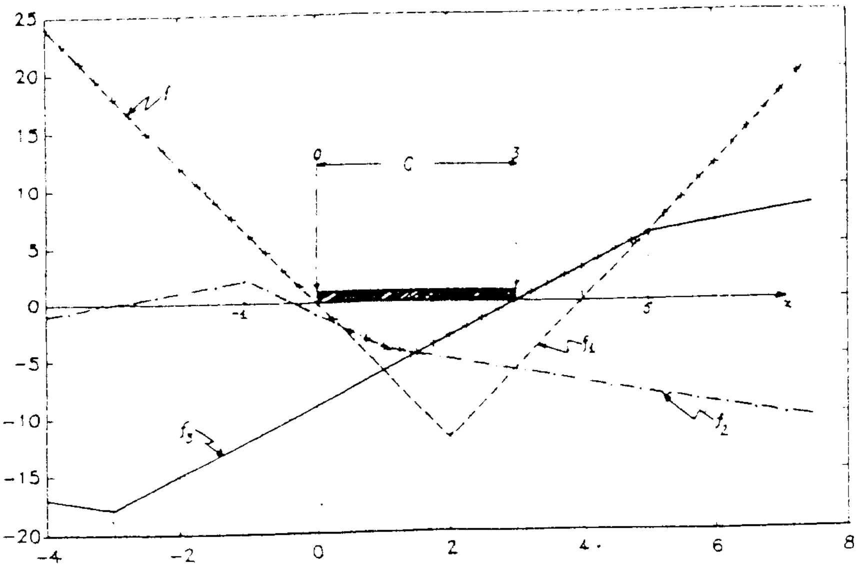

Figure 2. Plot of the functions f_1, f_2, f_3 and f.

We stop the algorithm when it reaches a solution of the problem (i.e., when $\lambda_k = 0$) or the distance of the actual iterate x^k to C is less than 10^{-5}.

The results are summarized in the following table.

Table 1.

$\alpha =$	1	1.2	1.4	1.6	1.8	2
No. of steps	40	2	27	4	6	12
Last iterate	3.00	1.04	3.00	1.25	1.30	2.00

Observe that when $\alpha \in \{1, 1.4, 2\}$, then the number of steps in which the SMFR algorithms reaches a point $x^* \in C$ with the desired error decreases while α (hence, λ_k) increases, and this is consistent with (I) above. In the other cases, i.e., when $\alpha \in \{1.2, 1.6, 1.8\}$, the SMFR algorithms reach interior points of C in finitely many steps and, therefore, they stop even sooner than expected.

Acknowledgments

The authors are grateful to Professor Yair Censor for many helpful discussions concerning the content of this paper. Also, the authors gratefully acknowledge the constructive critical comments of an anonymous referee, which led to improvements of an earlier version of this work.

References

[1] S. Agmon, "The relaxation method for linear inequalities," *Can. J. of Math.*, vol. 6, pp. 382–392, 1954.
[2] R. Aharoni, A. Berman, and Y. Censor, "An interior point algorithm for the convex feasibility problem," *Advances in Appl. Math.*, vol. 4, pp. 479–489, 1983.
[3] R. Aharoni and Y. Censor, "Block-iterative projection methods for parallel computation of solution to convex feasibility problems," *Linear Algebra And Its Applications*, vol. 120, pp. 165–175, 1988.
[4] J.-P. Aubin and A. Cellina, *Differential Inclusions*. Springer Verlag: Berlin, Heidelberg, New York, Tokyo, 1984.
[5] A. Auslender, *Optimization – Méthods Numériques*. Masson: Paris, New York, Barcelona, Milan, 1976.
[6] M. Avriel, "Methods for solving signomial and reverse convex programming problems," in: Optimization and Design, M. Avriel, M. J. Rijkaert and D.J. Wilde, Eds., pp. 307–320, *Prentice Hall: Englewood Cliffs*, 1973.
[7] L.M. Bregman, "The relaxation method of finding the common point of convex sets and its applications to the solution of problems in convex programming," *USSR Comput. Math. and Math. Physics*, vol. 7, pp. 200–217, 1967.
[8] D. Butnariu and Y. Censor, "On the behavior of a block-iterative projection method for solving convex feasibility problems," *Int. J. Comput. Math*, vol. 34, pp. 79–94, 1990.
[9] D. Butnariu and Y. Censor, "On a class of bargaining schemes for points in the core of a n-person cooperative game," *Tech. Report No. 280*, Dept. of Math., Univ. of Texas at Arlington, 1990.
[10] A. Cauchy, "Méthode générale pour la résolution des systémes d'equations simultanées," *Comptes Rendues Hebdomadaires de l'Academie des Sciences*, Paris, vol. XXV: 16, pp. 536–538, 1847.
[11] Y. Censor, "Row-action methods for huge and sparse systems and their applications," *SIAM Review*, vol. 23, pp. 444–466, 1981.
[12] Y. Censor and A. Lent, "Cyclic subgradient projections," *Math. Programming*, vol. 24, pp. 233–235, 1982.
[13] Y. Censor, D.E. Gustafson and A. Lent, "A new approach of the emission computerized tomography problem: Simultaneous calculation of attenuation and activity coefficients," *IEEE Trans. on Nuclear Sci.*, vol. NS-26, pp. 2775–2779, 1979.
[14] Y. Censor, "Parallel applications of block iterative methods in medical imaging and radiation therapy," *Math. Programming*, vol. 42, pp. 307–325, 1988.
[15] G. Cimmino, "Calcolo approsimato per le soluzioni di sistemi di equazioni lineari," *La Riserca Scientifica, Roma*, vol. XVI, IX: 2, pp. 326–333, 1938.
[16] F.H. Clarke, *Optimization and Nonsmooth Analysis*. John Wiley and Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1983.
[17] A.R. De Pierro and A.N. Iusem, "A simultaneous projection method for linear inequalities," *Linear Algebra And Its Applications*, vol. 64, pp. 243–253, 1985.
[18] A.R. De Pierro and A.N. Iusem, "A parallel projection method of finding common points of a family of convex sets," *Pesquisa Operacional*, vol. 5, pp. 1–20, 1985.
[19] F. Deutch, "Applications of von Neumann's alternating projection algorithm," in: *Methods of Operation Research*, P. Kendrov, ed., Sofia, pp. 44–51, 1983.
[20] F. Deutch, "Rate of convergence of the method of alternating projections," in: *Parametric Optimization And Approximations*, P. Brosowski and F. Deutch, eds., Birkhauser, Bassel, pp.

96–107, 1985.

[21] I.I. Eremin, "On systems of inequalities with convex functions in the left side," *Am. Math. Soc. Translations*, vol. 2: 88, pp. 67–83, 1970.

[22] S.D. Flåm, "Solving non-cooperative games by continuous subgradient projection methods," Working Paper No. 1489, *Dept. of Economics, Univ. of Bergen, Norway*, 1989.

[23] S.D. Flåm, "Approaching core solutions by mean of continuous bargaining," Working Paper No. 0690, *Dept. of Economics, Univ. of Bergen, Norway*, 1990.

[24] S.D. Flåm and J. Zowe, "Relaxed outer projections, weighted averages and convex feasibility," *BIT*, vol. 30, pp. 289–300, 1990.

[25] J. J.-B. Fourier, "Solution d'une question particulière du calcul des inégalitées," *Histoire de l'Academie des Sciences*, Paris, pp. 48–49, 1824.

[26] J.L. Goffin, "On convergence rates of subgradient optimization methods," *Math. Programming*, vol. 13, pp. 329–347, 1977.

[27] J.L. Goffin, "On the non-polynomiality of the relaxation method of systems of linear inequalities," *Math. Programming*, vol. 22, pp. 93–103, 1982.

[28] A.A. Goldstein, "Optimization of Lipschitz continuous functions," *Math. Programming*, vol. 13, pp. 14–22, 1977.

[29] C.G. Gubin, B.T. Polyak, and E. V. Raik, "The method of projection for finding the common point of convex sets," *USSR Comput. Math. and Math. Physics*, vol. 7, pp. 1–24, 1964.

[30] I. Halperin, "The product of projection operators," *Acta Sci. Math.*, vol. 23, pp. 94–99, 1962.

[31] P. Hansen, D. Peters, D. Richards, and J.-F. Thissa, "The minisum and the minimax location problems revisited," *Oper. Res.*, vol. 33, pp. 1251–1265, 1985.

[32] R.J. Hillestad, "Optimization problems subject to a budget constraint with economics of scale," *Oper. Res.*, vol. 23, pp. 1091–1098, 1975.

[33] A.N. Iusem and A.R. De Pierro, "Convergence results for an accelerated Cimmino algorithm," *Numerische Mathematik*, vol. 49, pp. 347–368, 1986.

[34] S. Kaczmarz, "Angenaherte auflösung von systemen linearer gleichungen," *Bull. Acad. Polon. Sci. Lett.*, vol. A35, pp. 355–357, 1937.

[35] S. Kayalar and H. Wienert, "Error bounds for the method of alternating projections," *Math. of Control, Signals and Systems*, vol. 1, pp. 43–59, 1988.

[36] L.S. Lasdon and J. Plummer, "Optimal design of efficient acoustic antenna arrays," *Math. Programming*, vol. 39, pp. 131–155, 1987.

[37] T.S. Motzkin and I.J. Scheonberg, "The relaxation method for linear inequalities," *Can. J. of Math.*, vol. 6, pp. 393–404, 1954.

[38] H. Nakano, "*Spectral Theory in Hilbert Space*," Japanese Society For Promotion of Sci., Tokyo, Japan, 1953.

[39] J. von Neumann, "On rings of operators. Reduction theory," *Ann. of Math.*, vol. 50, pp. 401–485, 1949.

[40] B.T. Polyak, "Minimization of nonsmooth functionals," *USSR Comput. Math. and Math. Physics*, vol. 9, pp. 509–521, 1969.

[41] B.T. Polyak, "A general method for solving extremum problems," *Doklady Akademia Nauk SSSR*, vol. 174, pp. 33–36, 1967.

[42] A.W. Roberts and D.E. Varberg, "Another proof that convex functions are locally Lipschitz," *Am. Math. Monthly*, vol. 81, pp. 1014–1016, 1974.

[43] R.T. Rockafellar, *Convex Analysis*, Princeton Univ Press, Princeton, NJ, 1970.

[44] N.Z. Shor, "An application of the method of gradient descent to the solution of the network transportation problem," in: *Materialy Naucnovo Seminara po Theoret i Priklad. Voprosam Kibernet. i Issled. Operacii, Nucnui po Kibernet.* Acad. Nauk Ukrain. SSSR, vyp. 1, Kiev, pp. 9–17, 1962.

[45] N.Z. Shor, *Minimization Methods for Nondifferentiable Functions*, Springer Verlag, Berlin, Heidelberg, New York, Tokyo, 1985.

[46] S. Wright, "An inexact algorithm for composite nondiffrentiable optimization," *Math. Programming*, vol. 44, pp. 221–234, 1989.

[47] D. Youla, "Generalized image restoration by the method of alternating orthogonal projections," *IEEE Trans. on Circuits and Systems*, vol. CAS-25: 9, pp. 694–702, 1978.