BIT 30 (1990), 289–300

RELAXED OUTER PROJECTIONS, WEIGHTED AVERAGES AND CONVEX FEASIBILITY

SJUR D. FLÅM[1] and JOCHEM ZOWE

Dep. of Economics,
Univ. of Bergen,
N-5008 Bergen, Norway

Dep. of Mathematics,
Univ. of Bayreuth,
D-8580, BRD

Abstract.

A new algorithmic scheme is proposed for finding a common point of finitely many closed convex sets. The scheme uses weighted averages (convex combinations) of relaxed projections onto approximating halfspaces. By varying the weights we generalize Cimmino's and Auslender's methods as well as more recent versions developed by Iusem & De Pierro and Aharoni & Censor. Our approach offers great computational flexibility and encompasses a wide variety of known algorithms as special instances. Also, since it is "block-iterative", it lends itself to parallel processing.

CR categories: F.2.2; I.3.5.

Key words: Outer Projections, Relaxations, Convex Feasibility Problem, Row-action Methods, Block-iterations.

1. Introduction.

A *convex feasibility problem* can be cast in the form

(CF): Find, if possible, at least one $x \in \bigcap_{i=1}^{m} C_i$.

Here each C_i, a closed convex subset of $\mathbb{R}^n$, reflects some prescribed constraint. Since problem (CF) is essentially geometrical, our discussion will be coached in corresponding terms [13]. It should not be forgotten however, that (CF) has an equivalent functional counterpart involving *convex inequalities*

(CI): Find, if possible, at least one x such that

$$f_i(x) \le y_i, \quad i = 1, \ldots, m,$$

[1] The research of the first author was partially supported by Ruhrgas under a NAVF grant.

Received February 1989. Revised August 1989.

where each f_i: $\mathbb{R}^n \to \mathbb{R}$ is quasi-convex and lower semicontinuous. For the equivalence of (CF) and (CI) note that, given f_i and y_i, we can define $C_i = \{x\colon f_i(x) \le y_i\}$; and conversely, given C_i, we could fix $y_i = 0$, and let

$$f_i(x) := d(x, C_i) := \inf\{\|x - c\|\colon c \in C_i\}$$

be the *distance (function)* to C_i.

We remark that (CI) emerges in many situations of great practical importance, and is often naturally called an *inversion problem.* Then, for $i = 1, \ldots, m$, y_i is typically a *known* output (measurement) of a controlled experiment, whose design is fixed via the function f_i, and x is an *unknown* "model" which, in the sense of (CI), must conform with observed data. Censor [8] gives an account of such problems in *image reconstruction* or *radiation therapy treatment planning.*

Not surprisingly, *linear* instances of problem (CI) have a long history starting with Kaczmarz [16], Cimmino [9], and continued with the work of Agmon [1], Motzkin and Schoenberg [18] on socalled *relaxation* methods. Later the treatment of *nonlinear* instances of (CI) has benefitted from developments in nonsmooth analysis, see [7], [10] and [23]. Now, the state of the art seems rather advanced, as already made clear in the review [5]. Nonetheless, the field continues to be very active with new impetus coming from techniques in decomposition [20], duality [24] and parallel processing [3], [17], [14], [15].

The purpose of this note is to provide an algorithm that subsumes a large variety of existing, well known iterative methods as special instances. We thus add to the flexibility of those methods and offer a unifying framework.

For motivation it is appropriate to review briefly some of the established approaches.

Auslender [4] has generalized the classical method of Cimmino [9] employing weighted averages of simultaneous projections. Specifically, he uses weights $\lambda_i > 0$ (actually, $\lambda_i = 1/m$), $\sum_{i=1}^m \lambda_i = 1$, and constructs the algorithmic map A as follows:

$$x^{k+1} := Ax^k := \sum_{i=1}^{m} \lambda_i P_i x^k, \quad x^0 \text{ arbitrary}, \tag{1.1}$$

where the orthogonal *projection* $P_i x^k$ is the point in C_i which is closest to x^k.

This method can be seen as an instance of *successive projections;* namely, consider the closed convex subset $\boldsymbol{C} := C_1 x \ldots . x C_m$ of the m-fold product space $(\mathbb{R}^n)^m$, and endow the latter with inner product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \lambda_1 \langle x_1, y_1 \rangle + \ldots . + \lambda_m \langle x_m, y_m \rangle$$

where $\boldsymbol{x} = (x_i)$, $\boldsymbol{y} = (y_i) \in (\mathbb{R}^n)^m$. In this larger space (1.1) can be rephrased on the equivalent form:

$$\boldsymbol{x}^{k+1} := P_{\boldsymbol{D}} P_{\boldsymbol{C}} \boldsymbol{x}^k, \quad \boldsymbol{x}^0 \text{ arbitrary}, \tag{1.1*}$$

with P_D, P_C denoting the orthogonal projections onto the *diagonal* $\boldsymbol{D} := \{\boldsymbol{d} = (d,\ldots,d) \mid d \in \mathbb{R}^m\}$ and the set $\boldsymbol{C}$, respectively. For details see Pierra [20] who proved that the sequence in (1.1), or equivalently (1.1*), converges to some point $x \in \cap_i C_i$, provided, of course, this intersection is nonempty.

Moreover, if int $\cap_i C_i \neq \emptyset$, then $\{x^k\}$ converges with a geometric rate.

A closely related method has been explored by De Pierro and Iusem [21]. They let

$$x^{k+1} = x^k + \alpha_k(Ax^k - x^k), \quad x^0 \text{ arbitrary}, \tag{1.2}$$

where α_k are *relaxation parameters* contained in a compact subset of (0, 2), and A is the same operator as in (1.1). Actually, they deal only with the case when each C_i is a halfspace, but, as will be demonstrated below, (1.2) converges for all closed convex sets C_i, $i = 1,\ldots,m$, having nonempty intersection.

To see the relation between (1.1) and (1.2) more clearly, let P be an orthogonal projection, and denote by

$$R(\ ,\alpha) := I + \alpha(P - I),$$

an associated *relaxed projection* with parameter $\alpha \in (0, 2)$. Then (1.1), respectively (1.2), can be rewritten on the form

$$x^{k+1} = \sum_{i=1}^{m} \lambda_i R_i(x^k, 1), \quad \text{and} \tag{1.1$'$}$$

$$x^{k+1} = \sum_{i=1}^{m} \lambda_i R_i(x^k, \alpha_k), \tag{1.2$'$}$$

x^0 being arbitrary in both cases. Evidently, algorithm (1.2′) is more general, yielding (1.1′) when $\alpha_k = 1$ for all k.

We conclude our review by mentioning a method of Aharoni et al. [2] designed as follows: Suppose that in (1.2′), at iteration k, all weight is assigned to *one* set $C_{i(k)}$, i.e.,

$$\lambda_j = \delta_{i(k),j} = \begin{cases} 1 & \text{if } j = i(k), \\ 0 & \text{otherwise.} \end{cases}$$

Here $\{i(k),\ k = 0, 1, \ldots\}$ is a so-called *control* sequence in $\{1,\ldots,m\}$ governing which set is brought into consideration at stage k. Also, suppose that instead of projecting onto $C_{i(k)}$ itself, we contend, if $x^k \notin C_{i(k)}$, with projecting onto a hyperplane separating $C_{i(k)}$ from x^k. Thus we arrive at the method

$$x^{k+1} = R_{(k)}(x^k, \alpha_k), \quad x^0 \text{ arbitrary}, \tag{1.3}$$

where $\{\alpha_k\}$ again is a sequence of relaxation parameters contained in a compact subset of (0, 2) and $R_{(k)}$ denotes the relaxed projection onto a half-space which contains the currently regarded set $C_{i(k)}$. It is required, if $x^k \notin C_{i(k)}$, that the halfspace in question lies at a distance at least

$$\beta_{ik} d(x^k, C_{i(k)}), \quad \beta_{ik} \in (0, 1], \tag{1.4}$$

from x^k. Also, every set must be picked up *repetitively*, i.e.,

$$k = j \quad \text{infinitely often for every } j = 1, \ldots, m. \tag{1.5}$$

Clearly, choosing $\beta_{ik} = 1$ for all i, k in (1.4), and letting the weights vary according to the rule $\lambda_j = \delta_{i(k),j}$ we recover (1.2′).

As mentioned we shall forge one single scheme that encompasses (1.1), (1.2) and (1.3) as particular implementations. It should, by now, not come as a surprise that we intend to use weights that may change from step to step. Such a strategy has also been investigated recently by Aharoni and Censor [3]. Their scheme, kindly brought to our attention during the revision of this paper, is, in one respect, not quite as general as ours: We contend with relaxed projections onto approximating halfspaces instead of using the sets $C_1, \ldots, C_m$ themselves. This is a substantial advantage if the projection onto some C_i is hard to execute.

Referring back to the successive projection method (1.1*) of Pierra (op. cit.) one could say that our algorithm approximates, at each step, the product set **C** from the outside, and changes, via the weights, the inner product and possibly also the "dimension" of the product space.

The paper is organized as follows. Section 2 introduces notations and states the algorithm. Convergence is established in Section 3. Section 4 elaborates on an *interior points algorithm* due to Aharoni et al. [2]. Section 5 concludes with some remarks.

2. The algorithm

It is convenient to introduce some notation. For a given closed convex set $C \subset \mathbb{R}^n$, and a point x outside C, let $H(x)$ be the set of all hyperplanes H in $\mathbb{R}^n$ which separates C from x.

We follow Aharoni et al. [2] and define for each number $\alpha \in [0, 2]$, a parametrized correspondence $R(., \alpha)$ from the entire space $\mathbb{R}^n$ to subsets of $\mathbb{R}^n$ as follows:

$$R(x, \alpha) = \begin{cases} \{x\} & \text{when } x \in C, \\ \{x + \alpha[P_H(x) - x]: \ H \in H(x)\} & \text{otherwise.} \end{cases} \tag{2.1}$$

Here P_H denotes the orthogonal projection onto the hyperplane H. In short, any element $y \in R(x, \alpha)$, different from x, is the result of relaxed projection, with parameter α, of x onto a hyperplane which separates C from x.

In the sequel we shall write $R_i(x, \alpha)$ and $d_i(x)$ to indicate that the set in question is C_i. With this notation the *algorithm* may now be stated as follows: For arbitrary initial x^0, choose iteratively.

(2.2) $$x^{k+1} \in \sum_{i=1}^{m} \lambda_{ik} R_i(x^k, \alpha_{ik}).$$

The right hand side of (2.2) is a weighted Minkowski sum of $R_i(x^k, \alpha_{ik})$ using weights λ_{ik}.

Clearly, if some set C_i drops out of attention at stage k, this corresponds to $\lambda_{ik} = 0$. As will be seen, this makes no harm as long as C_i is not permanently ignored from some stage onwards. In fact, every set has to be considered from time to time. To formalize this concern we need the following

ASSUMPTION (On weights). At stage $k \geq 0$, *weights* $\lambda_{ik} \geq 0$, $i = 1, \ldots, m$, $\sum_{i=1}^{m} \lambda_{ik} = 1$, are assigned to the different sets subject to the following restrictions: *Positive weights must exceed a threshold*, i.e. for some $\lambda > 0$, we have that

(2.3) $$\lambda_{ik} > 0 \quad \text{implies} \quad \lambda_{ik} \geq \lambda.$$

Also, *every set is repetitively considered* in the sense that

(2.4) $$\lambda_{ik} > 0 \quad \text{infinitely often for} \quad i = 1, \ldots, m.$$

Clearly, (2.3-4) are satisfied in (1.1-2). They are also satisfied in (1.3) provided (1.5) holds. In particular, following Aharoni et al. [2], one could use *cyclic control*:

$$i(k) = k (\operatorname{mod} m) + 1,$$

or, more generally, *almost cyclical* control requiring that for some integer period $p \geq m$, we should have

$$\{1, \ldots, m\} \subset \{i(k), \ldots, i(k + p - 1)\} \quad \text{for all} \quad k \geq 0.$$

As an alternative to (2.4) we could *always consider some worst violation*, meaning that $\lambda_{ik} > 0$, for at least one index i such that $d_i(x^k) > 0$, is currently maximal.

One might also require that *only violated constraints be taken into account*, i.e. that $x^k \in C_i$ implies $\lambda_{ik} = 0$, but we shall not do so. In any event, we regard conditions (2.3-4) as purely a matter of implementation and rather free choice.

That last remark goes equally well for the following innocuous restrictions on relaxations, saying that we should avoid reflexions and invariably seek to advance at least somewhat towards a closest point.

ASSUMPTION (On relaxation). *Mirror images are forbidden*, i.e.,

(2.5a) $$\limsup_{k \to \infty} \alpha_{ik} < 2 \quad \text{for} \quad i = 1, \ldots, m,$$

and some positive step must always be made, i.e.,

$$0 < \liminf_{k \to \infty} \alpha_{ik} \quad \text{for} \quad i = 1, \ldots, m. \tag{2.5b}$$

These conditions (2.5a-b) imply no real restriction on the design of the algorithm. By contrast, it is more difficult to ensure at stage k, when projecting onto a halfspace containing, say C_i, that some definite progress towards C_i be made. This concern motivates the

ASSUMPTION (On approximation). When $x^k \notin C_i$, denote by $H_{ik} \in H(x^k)$ the hyperplane selected in (2.1) for use in (2.2), and define the *quality of approximation*

$$\beta_{ik} := d(x^k, H_{ik})/d(x, C_i) \leq 1.$$

We require that

$$\liminf_{k \to \infty} \beta_{ik} > 0 \quad \text{for each} \quad i = 1, \ldots, m. \tag{2.6}$$

3. Convergence.

Throughout this section we assume that (CI) is *consistent*, i.e., $\cap_i C_i \neq \emptyset$. Our main result is now easy to state.

THEOREM 1. (*Global convergence*). *Suppose the assumptions above concerning weights, relaxation, approximation and consistency are in force. Then every sequence* $\{x^k\}$ *generated according to algorithm* (2.2) *converges, for arbitrary initial point* x^0, *to a point solving problem* (*CF*).

It simplifies the main argument to single out two auxiliary results.

LEMMA 1. *Suppose C is a closed convex non-empty subset of some Hilbert space. Let P be the orthogonal projection onto C, and denote by*

$$R = I + \alpha(P - I), \quad \alpha \in [0, 2],$$

an associated relaxation. Then

$$a\,\|x - Px\|^2 \leq \|x - c\|^2 - \|Rx - c\|^2 \quad \textit{for all} \quad c \in C, \tag{3.1}$$

where $a = \alpha^2$ *when* $\alpha \in [0, 1]$, *and* $a = 1 - (1 - \alpha)^2$ *otherwise.*

PROOF. When $\alpha \in [0, 1]$, we have

$$\|x - Rx\|^2 \leq \|x - c\|^2 - \|Rx - c\|^2 \quad \text{for all} \quad c \in C. \tag{3.2}$$

This Pythagoras type assertion is proven in Auslender [4, Thm. V.1.1] when

$\alpha = 1$. It can also be shown to hold for $\alpha \in [0, 1]$. The left hand side of (3.2) equals $\alpha^2 \|x - Px\|^2$, so that (3.1) follows. In general, $Rx - Px = (1 - \alpha)(x - Px)$, so that

$$\begin{aligned}\|Rx - x\|^2 &= \|Rx - Px\|^2 + 2\langle Rx - Px, Px - c\rangle + \|Px - c\|^2 \\ &= (1 - \alpha)^2 \|x - Px\|^2 + 2(1 - \alpha)\langle x - Px, Px - c\rangle + \|Px - c\|^2\end{aligned}$$

where $\langle x - Px, Px - c\rangle \geq 0$ for all $c \in C$. Thus, $\alpha > 1$ implies

$$\|Rx - c\|^2 \leq (1 - \alpha)^2 \|x - Px\|^2 + \|Px - c\|^2 \quad \text{for all } c \in C. \tag{3.3}$$

In (3.2) we set $\alpha = 1$, and add to (3.3). This yields immediately the desired conclusion. ■

LEMMA 2 [2]. *Let C be a closed convex non-empty subset of some Hilbert space. Then for every $\alpha \in [0, 2]$, $c \in C$, and $x \notin C$, we have*

$$\|x' - c\| \leq \|x - c\| (1 - a\beta^2 d^2(x, C)/2 \|x - c\|^2), \tag{3.4}$$

where $x' = x + \alpha(P_H x - x) \in R(x, \alpha)$ for some $H \in H(x)$; a is defined in (3.1) and $\beta = d(x, H)/d(x, C)$.

PROOF. For arbitrary $x \notin C$, and $H \in H(x)$, a hyperplane selected in (2.1), denote by H^- the associated halfspace containing C.

Note that

$$d(x, H^-) = d(x, H) = \beta d(x, C).$$

Consequently, setting $C = H^-$, $P = P_{H^-}$, and $x' = Rx$ in (3.1), we obtain

$$\begin{aligned}\|x' - c\|^2 &\leq \|x - c\|^2 - a \|x - P_H x\|^2 \\ &= \|x - c\|^2 - a d^2(x, H) \\ &= \|x - c\|^2 - a\beta^2 d^2(x, C), \\ &< \|x - c\|^2 [1 - a\beta^2 d^2(x, C)/2 \|x - c\|^2]^2.\end{aligned}$$

Now take the square root on both sides to get (3.4). ■

PROOF OF THEOREM 1. The assumption on relaxation (2.5) ensures that we may find $a \in (0, 1)$ such that, apart from finitely many k,

$$a \leq \begin{cases} \alpha_{ik}^2 & \text{if} \quad \alpha_{ik} \in (0, 1] \\ 1 - (1 - \alpha_{ik})^2 & \text{otherwise.} \end{cases}$$

With no loss of generality, let us assume that this inequality holds for *all* k. Similarly, by (2.6) we may find $\beta > 0$, such that

$$\beta_{ik} \geq \beta \quad \text{for all} \quad i \text{ and } k.$$

Then (3.1) implies

(3.5) $$a\beta^2 d_i^2(x^k) \le a\,\|x^k - P_{ik}x^k\|^2 \le \|x^k - c\|^2 - \|R_{ik}x^k - c\|^2$$

for all $c \in C_i$, where P_{ik} denotes the projection onto a halfspace containing C_i, chosen according to the rule (2.1), and R_{ik} is the associated relaxation.

Now multiply (3.5) by λ_{ik}, sum over i, and use the convexity of $\|\cdot - c\|^2$ to obtain

(3.6) $$a\beta^2 \sum_{i=1}^{m} \lambda_{ik} d_i^2(x^k) \le \|x^k - c\|^2 - \|x^{k+1} - c\|^2$$

for all $c \in \cap\{C_i\colon\ \lambda_{ik} > 0\}$.

(3.6) tells that the *sequence* $\{\|x^k - c\|\}$ *is monotone decreasing, thus convergent for every* $c \in \cap_i C_i$. It also tells that $\{x^k\}$ is bounded, hence the sequence has at least one accumulation point x. Suppose, for the sake of the argument, that some such point x belongs to $\cap_i C_i$. Then the monotonicity of $(\|x^k - x\|)$ implies $\|x^k - x\| \to 0$, and the proof would be complete.

Therefore, assume there exists an accumulation point $x \notin \cap_i C_i$, henceforth kept fixed. To derive a contradiction also fix an arbitrary $c \in \cap_i C_i$. We record, as follows from the monotonicity of $(\|x^k - c\|\}$, that

(3.7) $$\|x - c\| \le \|x^k - c\| \quad \text{for all} \quad k \ge 0.$$

Define the index set $I^+ := \{i \mid x \in C_i\}$, and its complement $I^- := \{1, \ldots, m\}\backslash I^+$. Choose a closed ball B centered at x such that $B \cap C_i = \emptyset$ for all $i \in I^-$. For each C_i, $i \in I^-$, we next underestimate the very last term in (3.4) by the number

$$\xi := a\beta^2(\min_{i \in I^-, y \in B} d(y, C_i))/(2 \max_{i \in I^-, y \in B} \|y - c\|^2).$$

Note that $\xi \in (0, 1)$. Also, if $x^k \in B$ for some k, and x^{ik} is selected according to the rule (2.1):

$$x^{ik} \in R_i(x^k, \alpha_{ik}),$$

then by (3.4) and the definition of ξ,

$$\|x^{ik} - c\| \le \begin{cases} \|x^k - c\|\,(1 - \xi) & \text{if} \quad i \in I^-, \text{ and} \\ \|x^k - c\| & \text{otherwise by (3.5).} \end{cases}$$

On such an occasion, when x^k does indeed belong to B, multiply the last inequality by λ_{ik}, and sum over i to obtain

$$\|x^{k+1} - c\| \le \|x^k - c\|\,\{\textstyle\sum_{i \in I^+} \lambda_{ik} + (1 - \xi)\sum_{i \in I^-} \lambda_{ik}\}.$$

In particular, if $\lambda_{ik} > 0$, for at least one index $i \in I^-$, then the last inequality gives, via (2.3), for such an i,

$$\|x^{k+1} - c\| \le \|x^k - c\|\,(1 - \xi\lambda_{ik}) \le \|x^k - c\|\,(1 - \xi\lambda) < \|x - c\|,$$

provided the radius of B is small enough. However, this contradicts (3.7). Therefore, $\lambda_{ik} = 0$ whenever $i \in I^-$ and $x^k \in B$. This shows, upon invoking (3.6) with x in place of c, that

$$x^k \in B \Rightarrow x^{k+1} \in B,$$

and consequently, since the radius of B can be selected arbitrarily small, the entire sequence $\{x^k\}$ converges to x. By hypothesis $d_i(x) > 0$, for at least one i. Sum (3.6) over k, and use (2.3) to get, for such an i, the contradiction

$$+\infty = a\beta^2 \sum_{k=0}^{m} \lambda_{ik} d_i^2(x^k) \leq \|x^0 - c\|^2. \tag{3.8}$$

Thus, we cannot maintain that an accumulation point x of $\{x^k\}$ falls outside $\cap_i C_i$. This completes the proof. ∎

REMARK. If in algorithm (2.2) some worst violation is always considered (see the remark following the assumption on the weights), then the above proof can be substantially simplified. Indeed, before (3.7) we would note that (2.3) and (3.6) imply:

$$a\beta^2 \lambda \max_i d_i(x^k) \leq \|x^k - c\|^2 - \|x^{k+1} - c\|^2 \quad \text{for all} \quad c \in \cap_i C_i.$$

Consequently, $d_i(x^k) \to 0$ for all $i = 1, \ldots, m$, as $k \to \infty$, and thus *all* accumulation points of $\{x^k\}$ must belong to $\cap_i C_i$. ∎

Clearly, (2.3) and (2.4) imply that

$$\sum_{k=0}^{\infty} \lambda_{ik} = +\infty \quad \text{for} \quad i = 1, \ldots, m. \tag{3.9}$$

We shall see that if $\cap_i C_i$ has non-empty interior, then (3.9) suffices for convergence.

THEOREM 2. *Suppose* int $\cap_i C_i \neq \emptyset$, *that* (3.9) *holds, and that the assumptions on relaxation and approximation are in vigour. Then every sequence* $\{x^k\}$ *generated by* (2.2) *converges, for arbitrary* x^0, *to a point solving problem* (*CF*).

PROOF. For arbitrary $c \in \cap_i C_i$, the monotonicity of $\{\|x_k - c\|\}$, see (3.6), ensures that any two accumulation points x, x' of $\{x^k\}$ satisfy $\|x - c\| = \|x' - c\|$. But int $\cap_i C_i \neq \emptyset$, implies $x = x'$, *i.e.*, $\{x^k\}$ converges to some unique x. If $x \notin C_i$ for some i, then sum (3.6) over k to have the contradiction (3.8) anew. ∎

4. An interior point algorithm.

This section aims at explaining how (2.2) may function in practice. In this endeavour we shall slightly generalize an algorithm of Aharoni et al. [2].

Throughout this section assume that $\operatorname{int} C_i \neq \emptyset$ for $i = 1,\ldots,m$.

This interior point algorithm is executed as follows:

Initialization: Choose $x^0 \in \mathbb{R}^n$ and $c^i \in \operatorname{int} C_i$, $i = 1,\ldots,m$, arbitrary.

$$DO \quad \text{for} \quad k = 0, 1, \ldots \quad \text{until convergence the following}$$

Iterative step: For given x^k, set $x^{k+1} := 0$ and *DO* for $i = 1,\ldots,m$,

1) If $\lambda_{ik} > 0$, go to 2; otherwise continue.
2) If $x^k \notin C_i$, go to 3; otherwise set $x^{ik} := x^k$, and go to 4.
3) Consider the segment $S_{ik} = C_i \cap [c^i, x^k]$. Find a point $z^{ik} \in [c^i, x^k] \backslash \operatorname{int} C_i$ such that

$$\|x^k - z^{ik}\| \geq \delta_{ik} d(x^k, S_{ik}), \tag{4.1}$$

where $\delta_{ik} \in [0, 1]$ is a given parameter. Also find a hyperplane H_{ik} through z_{ik} that separates C_i from x^k. Finally, let $x^{ik} := x^k + \alpha_{ik}(P_{H_{ik}} x^k - x^k)$, and go to 4.
4) Set $x^{k+1} := x^{k+1} + \lambda_{ik} x^{ik}$, and continue.

THEOREM 2. *Suppose the assumptions concerning weights, relaxation and consistency hold. Also suppose that the parameters δ_{ik} in* (4.1) *satisfy*

$$\liminf_{k \to \infty} \delta_{ik} > 0 \quad \textit{for} \quad i = 1,\ldots,m. \tag{4.2}$$

Then the interior point algorithm converges, for arbitrary x^0, to a solution of (CF).

PROOF. It suffices to show that the assumption (2.6) on approximation is satisfied. If $x^k \notin C_i$, then, by elementary geometry,

$$\|x^k - P_{H_{ik}} x^k\| / \|x^k - z^{ik}\| = \|c^i - P_{H_{ik}} c^i\| / \|c^i - z^{ik}\|.$$

Therefore, by (4.1)

$$\beta_{ik} = d(x^k, H_{ik})/d(x^k, C_i) \geq d(x^k, H_{ik})/d(x^k, S_{ik})) \geq \|x^k - P_{H_{ik}} x^k\| \, \delta_{ik} / \|x^k - z^{ik}\|$$
$$= \|c^i - P_{H_{ik}} c^i\| \, \delta_{ik} / \|c^i - z^{ik}\| \geq d(c^i, bdC_i)\delta_{ik} / \|c_i - z^{ik}\|$$

where bdC_i denotes the boundary of C_i. Now the desired inequality (2.6) follows from (4.2), $\min_i d(c^i, bdC_i) > 0$, and the fact that $\max_i \sup_k \|c^i - z^{ik}\| > 0$, is bounded. ■

This algorithm can often be made more amenable to implement. For that purpose suppose $C_i = \{x \mid f_i(x) \leq y_i\}$, with f_i: $\mathbb{R}^n \to \mathbb{R}$ convex and $f_i(c^i) < y_i$ for some c^i, $i = 1,\ldots,m$. Then the *interior point algorithm* stated here above comes naturally in a more concrete version. Namely, step 3 goes as follows:

3′) When $f_i(x^k) > y_i$, and $\lambda_{ik} > 0$, let $c^{ik} \in [c^i, x^k]$ solve the equation $f_i(c^{ik}) = y_i$, and find $z^{ik} \in [c^{ik}, x^k]$ such that

$$f_i(z^{ik}) \geq y_i, \text{ and } \|x^k - z^{ik}\| \geq \delta_{ik} \, \|x^k - c^{ik}\|,$$

where $\delta_{ik} \in [0, 1]$ is a prescribed parameter. Choose a *subgradient* $g^{ik} \in \partial f_i(z^{ik})$, ∂f_i denoting here the *subdifferential* of f_i, see [23], and let

$$x^{ik} = x^k + \alpha_{ik} \frac{\langle g^{ik}, z^{ik} - x^k \rangle}{\|g^{ik}\|^2} g^{ik}.$$

5. Concluding remarks.

Problem (1.1) can also be solved by well established methods of non-differentiable optimization. Granted $\cap_i C_i \neq \emptyset$, one chooses a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\arg\min f = \bigcap_{i=1}^{m} C_i, \tag{5.1}$$

and proceed to minimize f by, e.g., subgradient methods of Shor, Ermoliev and Polyak, see Shor [23]. Consult also Oettli [19]. Such methods have several advantages.

First, the Slater condition int $\bigcap_{i=1}^{m} C_i \neq \emptyset$, implies *finite* convergence. This is also manifest in a method of De Pierro and Iusem [22] akin to Polyak's variant of the subgradient algorithm.

Second, subgradient methods identify inconsistency fairly quickly. Third, since the optimal value $\inf(f)$ is often known in (5.1), one may rather solve a program:

$$\min f_0(x) \text{ s.t. } f(x) \leq \inf(f),$$

where the convex criterion f_0 is chosen to identify solutions enjoying more desirable properties than just feasibility.

By contrast, algorithm (2.2) may suffer from slow convergence even if int $\bigcap_{i=1}^{m} C_i \neq \emptyset$. This phenomenon has been studied by Goffin [11, 12]. Roughly speaking, problem (1.1) is poorly conditioned for (2.2) if $\bigcap_{i=1}^{m} C_i$ is "flat", i.e. if some tangent cone of this set has small solid angles. Also, inconsistency may be harder to detect, and no preference governs the choice of an x solving (1.1).

These drawbacks should not however, make one overlook the computational efficiency of algorithm (2.2). One attractive feature of (2.2) is that only one or few constraints are considered at a time. In fact, this ability is the basic and great advantage of socalled "row-action" methods [6]. Admittedly, such methods may loose against other methods on small problems (CF), but they are attractive and compete very well for efficient solution of large scale, sparse instances. An additional, important advantage is that different constraints may be handled by different parallel processors. As brought out in (3.5) such processors can, at stage k, perform several relaxed projections before they hand their results over to a center that aggregates by means of weights λ_{ik}, as in (2.2).

Acknowledgements.

We are grateful to Y. Censor and an anonymous referee for helpful comments.

REFERENCES

1. S. Agmon, *The relaxation method for linear inequalities*, Cand. J. Math. 6, 382–392 (1954).
2. R. Aharoni, A. Berman, Y. Censor, *An interior points algorithm for the convex feasibility problem*, Adv. in Appl. Math. 4, 479–489 (1983).
3. R. Aharoni and Y. Censor, *Block-iterative projection methods for parallel computation of solutions to convex feasibility problems, to appear in* Linear Algebra and its Applications (1989).
4. A. Auslender, *Optimisation, Methodes Numériques*, Mason, Paris (1976).
5. Y. Censor, *Row-action methods for huge and sparse systems and their applications*, SIAM Rev. 23, 444–446, (1981).
6. Y. Censor, *Iterative methods for the convex feasibility problem*, Annals of Discrete Math., 20, 83–91, (1984).
7. Y. Censor and A. Lent, *Cyclic subgradient projections*, Math. Programming 24, 233–235 (1982).
8. Y. Censor, *Parallel application of block-iterative methods in medical imaging and radiation therapy*, Mathematical Programming 42, 307–325 (1988).
9. G. Cimmino, *Calcolo approsssimato per le soluzioni di sisteme di equazioni lineari*, La Ricerca Scientifica XVI, 1, 326–333 (1938).
10. V. F. Dem'yanov and L. V. Vasil'ev, *Nondifferentiable Optimization*, Springer Verlag, Berlin (1985).
11. J. L. Goffin, *The relaxation method for solving linear systems of inequalities*, Math. Oper. Res. 5, 388–414 (1980).
12. J. L. Goffin, *Acceleration in the relaxation method for linear inequalities and subgradient optimization*, *in* Progress in Nondifferentiable Optimization, IIASA Proceeding Series, Pergamon, Oxford (1978).
13. L. G. Gubin, B. T. Polyak and E. V. Raik, *The method of projections for finding the common point of convex sets*, USSR Comp. Math. and Math. Phys. 7, 1–24 (1967).
14. A. N. Iusem and A. R. De Pierro, *Convergence results for an accelerated nonlinear Cimmino algorithm*, Numer. Math. 49, 367–378 (1986).
15. A. N. Iusem and A. R. De Pierro, *A simultaneous iterative method for computing projections on ployhedra*, SIAM J. Control and Optimization, 25, 231–243 (1987).
16. S. Kaczmarz, *Angenäherte Auflösung von Systemen linearer Gleichungen*, Bull. Acad. Polon. Sci. Lett. A, 35, 355–357 (1937).
17. J. M. Martinez and R. J. B. De Sampaio, *Parallel and sequential Kaczmarz methods for solving underdetermined nonlinear equations*, J. Comput. Appl. Math. 5, 311–321 (1986).
18. T. S. Motzkin and I. J. Schoenberg, *The relaxation method for linear inequalities*, Canad. J. Math. 6, 393–404 (1954).
19. W. Oettli, *Symmetric duality and convergent subgradient methods for discrete, linear, constrained approximation problems with arbitrary norms appearing in the objective function and in the constraints*, J. Approx. Theory, 14, 43–50 (1975).
20. G. Pierra, *Decomposition through formalization in a product space*, Mathematical Programming 28, 96–115 (1984).
21. A. R. De Pierro and A. N. Iusem, *A simultaneous projections method for linear inequalities*, Lin. Alg. Appl. 64, 243–253 (1985).
22. A. R. De Pierro and A. N. Iusem, *A finitely convergent rowaction method for the convex feasibility problem*, Appl. Math. Optim. 17, 225–235 (1988).
23. N. Z. Shor, *Minimization Methods for Nondifferentiable Functions*, Springer-Verlag, Berlin (1983).
24. J. E. Spingarn, *A primal-dual projection method for solving systems of linear inequalities*, Lin. Alg. Applic. 65, 45–62 (1985).