THE RELAXATION METHOD OF FINDING THE COMMON POINT OF CONVEX SETS AND ITS APPLICATION TO THE SOLUTION OF PROBLEMS IN CONVEX PROGRAMMING*

L.M. BREGMAN

Leningrad

(*Received* 20 *May* 1966)

IN this paper we consider an iterative method of finding the common point of convex sets. This method can be regarded as a generalization of the methods discussed in [1 - 4]. Apart from problems which can be reduced to finding some point of the intersection of convex sets, the method considered can be applied to the approximate solution of problems in linear and convex programming.

1. The problem of finding the common point of convex sets

Suppose we are given in a linear topological space X some family of closed convex sets A_i, $i \in I$, where I is some set of indices. We shall assume that $R = \bigcap_{i\in I} A_i$ is not empty. It is required to find some point of the intersection of the sets A_i.

Let $S \subset X$ be some convex set such that $S \cap R \neq \Lambda$.

Let us consider the function $D(x, y)$, defined over $S \times S$, and satisfying the following conditions.

I. $D(x, y) \geqslant 0$, $D(x, y) = 0$ if and only if $x = y$.

* *Zh. vȳchisl. Mat. mat. Fiz.* 7, 3, 620 - 631, 1967.

II. For any $y \in S$ and $i \in T$ a point $x = P_i y \in A_i \cap S$ exists such that

$$D(x, y) = \min_{z \in A_i \cap S} D(z, x).$$

This point x will be called the D-projection of the point y onto the set A_i.

III. For each $i \in I$ and $y \in S$ the function $G(z) = D(z, y) - D(z, P_i y)$ is convex over $A_i \cap S$.

IV. A derivative $D_x'(x, y)$ of the function $D(x, y)$ exists when $x = y$, while $D_x'(y, y) = 0$ (i.e. $\lim_{t \to 0} [D(y + tz, y) / t] \rightleftarrows 0$ for all $z \in Y$).

V. For each $z \in R \cap S$ and for every real number L the set $T = \{x \in S | D(z, x) \leqslant L\}$ is compact.

VI. If $D(x^n, y^n) \to 0$, $y^n \to y^* \in \bar{S}$ ($\bar{S}$ is the closure of the set S) and the set of elements of the series $\{x^n\}$ is compact, we have $x^n \to y^*$.

Consider the following iterative process:

(1) take an arbitrary point $x^0 \in S$;

(2) if the point $x^n \in S$ is known, we select in some way the index $i_n(x^n) \in I$ and we find the point x^{n+1} which is the D-projection of the point x^n onto the set $A_{i_n(x^n)}$.

The sequence $\{x^n\}$ obtained as a result of this process will be called the relaxation sequence.

The sequence of indices, chosen during each of the iterations $\{i_0(x^0), i_1(x^1), \dots\}$, will be called, following [5], the control of the relaxation.

Below we shall consider some relaxation controls, under which the relaxation sequence $\{x^n\}$ converges to some point $x^* \in R$.

We have the following lemmas.

Lemma 1

Let $z \in A_i \cap S$. Then for any $y \in S$ the inequality $D(P_i y, y) \leqslant$

$D(z, y) - D(z, P_iy)$ is valid.

Proof. According to condition III, for all $\lambda \in [0, 1]$ we have

$$D(\lambda z+(1-\lambda)P_iy, y)-D(\lambda z+(1-\lambda)P_iy, P_iy)\leqslant$$
$$\leqslant \lambda(D(z, y)-D(z, P_iy))+(1-\lambda)D(P_iy, y).$$

Whence when $\lambda > 0$ we obtain

$$D(z, y)-D(z, P_iy)-D(P_iy, y)\geqslant$$
$$\geqslant \frac{D(\lambda z+(1-\lambda)P_iy, y)-D(P_iy, y)}{\lambda}-\frac{D(\lambda z+(1-\lambda)P_iy, P_iy)}{\lambda}, \tag{1.1}$$

Since $\lambda z+(1-\lambda)P_iy\in A_i\cap S$, the first term on the right-hand side of (1.1) is non-negative (in view of condition II), and the second term tends to zero when $\lambda \to 0$ (in view of condition IV). Hence $D(z,y) - D(z, P_iy) - D(P_iy, y)\geqslant 0$.

Lemma 2

For any relaxation control we have the following:

(1) The set of elements of the relaxation sequence $\{x^n\}$ is compact;

(2) For any $z \in S$ there exists $\lim_{n\to\infty} D(z, x^n)$;

(3) $D(x^{n+1}, x^n) \to 0$ when $n \to \infty$.

Proof. We take $z \in R\cap S$. According to Lemma 1,

$$D(x^{n+1}x^n)\leqslant D(z, x^n)-D(z, x^{n+1}). \tag{1.2}$$

Since $D(x^{n+1}, x^n)\geqslant 0$, we have $D(z, x^{n+1})\leqslant D(z, x^n)$. Consequently, there exists $D(z, x^n)$, which together with (1.2) gives $D(x^{n+1}, x^n) \to 0$.

Since the set of elements of the relaxation sequence $\{x^n\}$ is contained in the set $T=\{x\in S|D(z,x)\leqslant D(z,x^0)\}$, which, according to condition V, is compact, therefore statement (1) of the Lemma is also true.

Let us now consider some relaxation controls for which the limiting point of the sequence $\{x^n\}$ belongs to the intersection of the sets A_i.

Theorem 1

Let $I = \{1, 2, \ldots, m\}$ and let the indices be chosen in cyclic order, i.e. $i_0(x^0) = 1$, $i_1(x^1) = 2, \ldots, i_{m-1}(x^{m-1}) = m$, $i_m(x^m) = 1$, and so on. Then any limiting point x^* of the relaxational sequence $\{x^n\}$ is a common point of the sets A_i.

Proof. Let x^* be the limiting point of the sequence $\{x^n\}$ and $x^{n_k} \to x^*$. We separate out from the sequence $\{x^{n_k}\}$ a subsequence which is wholly contained in one of the sets A_i, for example in A_1. We shall assume that $\{x^{n_k}\} \subset A_1$. We separate out from the sequences $\{x^{n_k+i-1}\}$ those which are convergent. It can be assumed that the sequences $\{x^{n_k+i-1}\}$ themselves are convergent. Let

$$x^{n_k} \to x^* = x_1^*,$$
$$x^{n_k+1} \to x_2^*,$$
$$\cdots\cdots$$
$$x^{n_k+m-1} \to x_m^*.$$

Since $\{x^{n_k+i-1}\} \subset A_i$, we have $x_i^* \in A_i$.

According the Lemma 2 we have $D(x^{n_k+1}, x^{n_k}) \to 0$. According to condition VI, $\lim x^{n_k+1} = \lim x^{n_k} = x_1^* = x_2^*$. Consequently, $x^* \in A_2$. It can be shown analogously that $x^* \in A_3, x^* \in A_4$, and so on. Consequently

$$x^* \in \bigcap_{i \in I} A_i.$$

Theorem 2

Suppose that for each $y \in S$ there exists

$$\max_{i \in I} \min_{x \in A_i} D(x, y).$$

For $i_n(x^n)$ we shall choose that index which realizes

$$\max_{i \in I} \min_{x \in A_i} D(x, x^n).$$

Then any limiting point of the relaxation sequence $\{x^n\}$ is a common point of the sets A_i.

Proof. Let $x^{n_k} \to x^*$. We denote the D-projection of the point x^{n_k} onto the set A_i by $y_i^{n_k}$. Then

$$D(y_i^{n_k}, x^{n_k}) \leqslant \max_{j \in I} D(y_j^{n_k}, x^{n_k}) = D(x^{n_k+1}, x^{n_k}).$$

Since according to Lemma 2 we have $D(x^{n_k+1}, x^{n_k}) \to 0$, therefore also

$$D(y_i^{n_k}, x^{n_k}) \to 0. \tag{1.3}$$

According to Lemma 1, for any $z \in R$ we have

$$D(z, y_i^{n_k}) \leqslant D(z, x^{n_k}) \leqslant D(z, x^0).$$

Consequently, according to condition V, the set $\{y_i^{n_k}\}$ is compact which, together with (1.3) on the basis of condition VI, gives $y_i^{n_k} \to x^*$ for all $i \in I$.

Since $y_i^{n_k} \in A_i$, we have $x^* \in \bigcap_{i \in I} A_i$.

Note 1. In many cases the relaxation sequence $\{x_n\}$ has a unique limiting point $x^* \in R$. This happens, for example, if one of the following conditions is satisfied:

(1) The set S is closed, and for any $z_1, z_2 \in R \cap S$ the function $H(y) = D(z_1, y) - D(z_2, y)$ is continuous over S;

(2) the function $D(x, y)$ is defined when $x \in \bar{S}$, and if $y^n \to y^* \in S$, then $D(y^*, y^n) \to 0$.

In fact, let condition (1) be satisfied and let

$$x^{n_k} \longrightarrow x^* \in R, \quad x^{n_l} \longrightarrow x^{**} \in R.$$

According to Lemma 2, there exists

$$\lim H(x^n) = \lim(D(x^*, x^n) - D(x^{**}, x^n)).$$

For the subsequences x^{n_k}

$$\lim H(x^{n_k}) = -D(x^{**}, x^*) \leqslant 0.$$

For the subsequences x^{n_l}

$$\lim H(x^{n_l}) = D(x^*, x^{**}) \geqslant 0.$$

Consequently $D(x^*, x^{**}) = D(x^{**}, x^*) = 0$ and according to condition (1),

$x^* = x^{**}$.

Suppose condition (2) is satisfied and again $x^{n_k} \to x^* \in R$, $x^{n_l} \to x^{**} \in R$. Then

$$0 = \lim D(x^*, x^{n_k}) = \lim D(x^*, x^{n_l}) = \lim D(x^*, x^n).$$

Hence, according to condition VI, it follows that $x^* = x^{**}$.

Let us now consider some examples of functions satisfying the conditions I - VI.

1. Let X be a real Hilbert space, $S = X$ and $D(x, y) = (x - y, x - y)$.

The function $D(x, y)$, obviously satisfies condition I.

Condition II is satisfied, since the D-projection onto a convex set is in this case the same as an ordinary projection.

The function defined in condition III

$$G(z) = D(z, y) - D(z, P_i y) = (z - y, z - y) - \\ - (z - P_i y, z - P_i y) = 2(z, P_i y - y) + (y, y) - (P_i y, P_i y)$$

is linear, and consequently this condition is also satisfied.

Furthermore,

$$D_x'(y, y) = \lim_{t \to 0} \frac{(y + tz - y, y + tz - y)}{t} = \lim_{t \to 0} t(z, z) = 0.$$

Consequently condition IV is satisfied.

The set $T = \{y \in S \mid (x - y, x - y) \leqslant L\}$ occurring in condition V is not compact, but is bounded and it will therefore be compact if we assume that a weak topology is introduced in X.

Condition VI is satisfied if by convergence we mean weak convergence. In fact, let $(x^n - y^n, x^n - y^n) \to 0$, $y^n \dashrightarrow y^*$, and let the set of elements of the sequence $\{x^n\}$ be weakly compact. Let $x^{n_k} \dashrightarrow x^*$. Then for every $u \in X$ we have

$$|(u, x^{n_k} - y^{n_k})| \leqslant \|u\| \, \|x^{n_k} - y^{n_k}\| \to 0.$$

Consequently, $\lim(u, x^{n_k}) = \lim(u, y^{n_k}) = (u, y^*)$, and this means that

$x^* = y^*$.

Condition (1) is also satisfied since the function

$$H(y) = (z_1 - y, z_1 - y) - (z_2 - y, z_2 - y) = (z_1, z_1) - (z_2, z_2) + 2(z_2 - z_1, y)$$

is linear and therefore continuous in the weak topology.

Hence it follows that for certain relaxation controls (for example those satisfying the conditions of Theorems 1 or 2) the relaxation sequence $\{x^n\}$ will be weakly convergent to some element $x^* \in \bigcap_{i \in I} A_i$

This has been proved earlier in [4].

2. Let $f(x)$ be a strictly convex differentiable function given over the convex set $S \subset E^p$, and let $g(x)$ be its gradient at the point s.

Let us consider the function

$$D(x, y) = f(x) - f(y) - (g(y), x - y). \tag{1.4}$$

We shall show that $D(x, y)$ satisfies conditions I - IV.

Indeed, condition I represents one of the properties of convex functions, which is that the graph of a convex function lies only on one side of a tangential plane (see [6]).

Condition II is satisfied, since for each $y \in S$, $\min_{x \in S} D(x, y) = 0$ exists and consequently $\min_{x \in S \cap A_i} D(x, y)$ exists for every closed convex set A.

The function $G(z)$ occurring in condition III is convex, since

$$G(z) = -f(y) + f(P_i y) - (g(y), y) + (g(P_i y), P_i y) - \\ - (g(y) - g(P_i y), z).$$

Furthermore, $D_x'(y, y) = g(y) - g(y) = 0$, so condition IV is satisfied.

Conditions VI and (1) are satisfied if some auxiliary assumptions are made with respect to the function $f(x)$. For example, they are satisfied if the set S is closed, and the function $f(x)$ is continously differentiable.

Condition V is not a consequence of the convexity of $f(x)$, and for this reason we shall only consider functions for which condition V is satisfied.

An example of a function $f(x)$, for which the function $D(x, y)$, constructed according to formula (1.4), satisfies conditions I - VI and (1), is given by a positive definite quadratic form given over the whole space E^p. The corresponding function $D(x, y) = (x - y, C(x, y))$, as is easy to see, satisfies conditions I - VI and (1).

Another example of such a function will be

$$f(x)=\sum_{j=1}^{p} x_j \ln x_j,$$

given over the set $S=\{x\in E^p | x>0\}$. Its corresponding function $D(x, y)$ has the form

$$D(x,y)=\sum_{j=1}^{p} (y_j-x_j+x_j(\ln x_j-\ln y_j)). \tag{1.5}$$

According to the previous considerations, function (1.5) satisfies conditions I - VI. It is easy to see that $D(x, y)$ also satisfies condition V.

Let us verify that function (1.5) satisfies condition VI. Let $D(x^n, y^n) \to 0$ and $y^n \to y^* = (y_1^*, y_2^*, \ldots, y_p^*)$. If $y_j^* = 0$, then also $x_j^n \to 0$, since otherwise $D(x^n, y^n) \to \infty$. If $y_j^* > 0$, then $x_j^n \to y_j^*$ in view of the continuous nature of the function $D(x, y)$ when $y_j > 0$.

As can be seen from (1.5), the function $D(x, y)$ satisfies condition (2).

Consequently a relaxation sequence $\{x^n\}$ constructed by means of the function (1.5) will for any system of closed convex sets A_i converge to the point $x^*\in\bigcap_{i\in I} A_i$ with an appropriate selection of the relaxation control.

In concluding this section we remark that for the functions $f_1(x)$ and $f_2(x)$, the difference between which is a linear function, the functions

$$D_1(x,y)=f_1(x)-f_1(y)-(g_1(y),x-y)$$

and

$$D_2(x,y)=f_2(x)-f_2(y)-(g_2(y),x-y)$$

are identical.

2. Solving some problems of convex programming

If $R=\bigcap_{i\in I} A_i$ does not consist of a single point, the limit of the relaxation sequence will depend on the choice of the initial approximation and of the relaxation control. Therefore by a suitable choice of the initial approximation and of the relaxation control it can be ensured that the limiting point x^* will have certain specified properties, for example, that it will minimize a certain function over R. We shall make use of these considerations to solve certain problems in convex programming.

Let $f(x)$ be a strictly convex function which is continuously differentiable over the convex set $S \subset E^p$, and let it be continuous over $\bar{S}$. Consider the following problem:

to minimize

$$f(x) \tag{2.1}$$

subject to the conditions

$$Ax=b, \tag{2.2}$$

$$x\in\bar{S}. \tag{2.3}$$

Here $A = \| a_{ij} \|$ is a matrix with m rows and p columns, $x \in E^p$, $b\in E^m$.

Let us denote the i-th row of the matrix A by A_i. We assume that all $A_i \neq 0$.

Let R be the set of permissible vectors of problem (2.1) - (2.3), i.e. $R=\{x\in E^p | Ax=b, x\in\bar{S}\}$. We shall assume that R is not empty.

We assume that the function $D(x, y)$ constructed according to formula (1.4) satisfies conditions I - VI and condition (2).

We note that in this case condition (2) is a consequence of condition (1). Indeed, if $f(x)$ is continuously differentiable over the closed set S, $D(x, y)$ is continuous, and condition (2) is satisfied.

Let us denote by Z the set of those $x \in S$ for which there exists such $u \in E^m$, that $g(x) = uA$ (($g(x)$ is the gradient of the function $f(x)$). Let $\bar{Z}$ be the closure of the set Z.

Lemma 3

If $y^* \in R \cap \bar{Z}$, y^* is a solution of the problem (2.1) - (2.3).

Proof. Since $y^* \in R$, therefore

$$f(y^*) \geqslant \inf_{x \in R} f(x).$$

Therefore there exists $x^* \in R$ such that

$$f(y^*) - f(x^*) = a \geqslant 0. \tag{2.4}$$

In order to prove this lemma it is sufficient to show that $a = 0$. Since $y^* \in \bar{Z}$, we can find a sequence $\{y^n\}$ such that $y^n \in Z$ and $y^n \to y^*$.

For every n we can find $u^n \in E^m$ such that $g(y^n) = u^n A$. Hence it follows that $(g(y^n), v) = 0$ for all v, for which $Av = 0$. We put $v = y^* - x^*$. Then for all n we have

$$(g(y^n), y^* - x^*) = 0. \tag{2.5}$$

Taking into account (2.5) and (1.4) we have

$$\begin{aligned} a = f(y^*) - f(x^*) = (g(y^n), y^* - y^n) + D(y^*, y^n) - \\ - (g(y^n), x^* - y^n) - D(x^*, y^n) = D(y^*, y^n) - D(x^*, y^n). \end{aligned} \tag{2.6}$$

From (2.6) it follows that $a \leqslant D(y^*, y^n)$.

Hence, using condition (2) and (2.4), we obtain $a = 0$.

Theorem 3

Let the D-projection of any point x belonging to the interior of the set S onto the set

$$A_i = \left\{ x \in E^p \;\middle|\; \sum_{j=1}^{p} a_{ij} x_j = b_i \right\}$$

also belong to the interior of S. Suppose we select such a relaxation control for which $x^n \to x^* \in R$ (for example, of the kind as in Theorems 1 - 3). Then if the initial approximation is $x_0 \in Z \cap \text{int } S$ (int S denotes the interior of the set S), x^* is a solution of the problem (2.1) - (2.3).

Proof. Let x^{n+1} be a D-projection of the point x^n onto the set $A_{i\cdot}$. Then

$$g(x^{n+1}) = g(x^n) + \lambda A_{i\cdot}, \tag{2.7}$$

$$(A_{i\cdot}, x^{n+1}) = b_i. \tag{2.8}$$

Hence it can be seen that if $x^n \in Z$, we have $x^{n+1} \in Z$ also. Consequently, $x^* \in Z$ and, according to Lemma 3, x^* is a solution of the problem (2.1) - (2.3).

Note 2. Both x^{n+1} and λ are uniquely determined from the conditions (2.7) and (2.8).

Indeed, let there be $y, z \in E^p$ and numbers λ and μ such that

$$g(y) = g(x^n) + \lambda A_{i\cdot}, \quad g(z) = g(x^n) + \mu A_{i\cdot},$$
$$(A_{i\cdot}, y) = b_i, \quad (A_{i\cdot}, z) = b_i \text{ and } y \neq z.$$

Then

$$f(y) - f(z) > (g(z), y-z) = (g(x^n), y-z) + \mu(A_{i\cdot}, y-z) = (g(x^n), y-z), \tag{2.9}$$

$$f(z) - f(y) > (g(y), z-y) = (g(x^n), z-y) + \lambda(A_{i\cdot}, z-y) = (g(x^n), z-y). \tag{2.10}$$

Adding equations (2.9) and (2.10) we obtain $0 > 0$. Consequently $y = z$. This means that $\lambda = \mu$ also.

Note 3. If $f(x)$ has a global minimum inside S, then as an initial approximation x^0 we can take a point at which this minimum is achieved, since $g(x^0) = 0$, $u^0 = 0$ and consequently $x^0 \in Z$.

Let us now consider a problem in which the restrictions are given in the form of inequalities:

Minimize

$$f(x) \tag{2.11}$$

under the conditions

$$Ax \geqslant b, \tag{2.12}$$

$$x \in \bar{S}. \tag{2.13}$$

Let the function $f(x)$ have the same properties as earlier, and let $R = \{x \in E^p | Ax \geqslant b, x \in \bar{S}\} \neq \Lambda$.

Let $Z_0 = \{x \in S|$, there exists $u = (u_1, u_2, \ldots, u_m)$ such that $u \geqslant 0$.

To solve the problem (2.11) - (2.13) the relaxational method requires some modifications.

We shall assume that the conditions of Theorem 3 are satisfied.

Let us consider the following method of solving the problem (2.11) - (2.13).

1. We shall assume that a cyclic relaxation control is adopted. We shall denote by i_n the index selected on the n-th iteration.

2. For the initial approximation we choose the point $x^0 \in Z_0 \cap$ int S. The vector $u^0 = (u_1{}^0, u_2{}^0, \ldots, u_m{}^0)$ is such that $g(x^0) = u^0 A$.

3. (a) If $(A_{i_n \cdot}, x^n) < b_{i_n}$, then for x^{n+1} we take the D-projection of the point x^n on the set $A_{i_n} = \{x \in E^p | (A_{i_n \cdot}, x) = b_{i_n}\}$, i.e. x^{n+1} is determined from the conditions

$$g(x^{n+1}) = g(x^n) + \lambda_n A_{i_n \cdot}, \tag{2.14}$$

$$(A_{i_n \cdot}, x^{n+1}) = b_{i_n}. \tag{2.15}$$

(According to Note 2, x^{n+1} and λ_n are uniquely determined from conditions (2.14) - (2.15).) Furthermore

$$u_{i_n}^{n+1} = u_{i_n}^{n} + \lambda_n, \quad u_i^{n+1} = u_i^n \text{ when } i \neq i_n.$$

(b) If $(A_{i_n \cdot}, x^n) = b_{i_n}$ or $(A_{i_n \cdot}, x^n) > b_{i_n}$, but $u_{i_n} = 0$, we have $x^{n+1} = x^n$ and $u_i{}^{n+1} = u_i{}^n$.

(c) If $(A_{i_n \cdot}, x^n) > b_{i_n}$ and $u_{i_n} > 0$, we determine x^{n+1} from the relationship

$$g(x^{n+1}) = g(x^n) - \mu_n A_{i_n},\qquad(2.16)$$

and u^{n+1} is determined from the formula $u_{i_n}^{n+1} = u_{i_n}^{n} - \mu_n,\ u_i^{n+1} = u_i$ when $i \neq i_n$.

Here

$$\mu_n = \min(\mu_n', \mu_n''),\qquad(2.17)$$

where μ_n' is determined from the conditions

$$g(y) = g(x^n) - \mu_n' A_{i_n},\qquad(2.18)$$

$$(A_{i_n}, y) = b_{i_n},\qquad(2.19)$$

and $\mu_n' = u_{i_n}^n$.

Theorem 4

The sequence $\{x^n\}$ obtained as a result of the process just described converges to the point x^*, which is a solution of the problem (2.11) - (2.13).

Proof. 1. We shall show that $x^n \in Z_0$ for all n. According to (2), we have $x^0 \in Z_0$. Let $x^n \in Z_0$, when $n \leqslant k$. If for the index i_k or (3c) is satisfied then, as can be seen from formula (2.16), $x^{k+1} \in Z_0$ (since $u^{k+1} \geqslant 0$, in view of the fzct that $\mu_k \leqslant u_{i_k}{}^k$).

Let us take the case (3a). Vrom (2.14) we obtain

$$(g(x^{k+1}) - g(x^k),\ x^{k+1} - x^k) = \lambda_k (A_{i_k},\ x^{k+1} - x^k).\qquad(2.20)$$

It is easy to obtain from formula (1.4) that the left-hand side of (2.20) is equal to $D(x^{k+1}, x^k) + D(x^k, x^{k+1})$. Hence, taking into account (2.15), we obtain

$$\lambda_k = \frac{D(x^{k+1}, x^k) + D(x^k, x^{k+1})}{b_{i_k} - (A_{i_k}, x^k)} > 0\qquad(2.21)$$

Consequently, $u^{k+1} \geqslant 0$ and $x^{k+1} \in Z_0$.

2. Let us consider the function $\varphi(x, u) = f(x) - (u, Ax - b)$. We show that

$$\varphi(x^{n+1}, u^{n+1}) \geqslant \varphi(x^n, u^n);\qquad(2.22)$$

$$\varphi(x^{n+1}, u^{n+1}) - \varphi(x^n, u^n) = f(x^{n+1}) - f(x^n) - (u^{n+1}, Ax^{n+1} - b) +$$
$$+ (u^n, Ax^n - b) = D(x^{n+1}, x^n) + (g(x^n), x^{n+1} - x^n) - (g(x^{n+1}), x^{n+1}) +$$
$$+ (g(x^n), x^n) + (u^{n+1} - u^n, b) = D(x^{n+1}, x^n) +$$
$$+ (u^{n+1} - u^n, b) + (g(x^n) - g(x^{n+1}), x^{n+1}).$$

If for the index i_n condition (3a) is valid

$$\varphi(x^{n+1}, u^{n+1}) - \varphi(x^n, u^n) = D(x^{n+1}, x^n) \qquad (2.23)$$

and consequently (2.22) is satisfied.

If for the index i_n condition (3c) is valid we have

$$\varphi(x^{n+1}, u^{n+1}) - \varphi(x^n, u^n) = D(x^{n+1}, x^n) + \mu_n((A_{i_n \cdot}, x^n) - b_{i_n}).$$

In the same way as formula (2.21), we can obtain

$$\mu_n = -\frac{D(x^{n+1}, x^n) + D(x^n, x^{n+1})}{b_{i_n} - (A_{i_n \cdot}, x^n)} > 0.$$

Since from (2.17) - (2.19) $(A_{i_n \cdot}, x^{n+1}) - b_{i_n} \geqslant 0$, therefore

$$\varphi(x^{n+1}, u^{n+1}) - \varphi(x^n, u^n) \geqslant D(x^{n+1}, x^n), \qquad (2.24)$$

and consequently, (2.22) is satisfied.

3. Let $z \in R$. Then

$$D(z, x^n) = f(z) - f(x^n) - (g(x^n), z - x^n) = \qquad (2.25)$$
$$= f(z) - f(x^n) - (u^n A, z - x^n) \leqslant f(z) - f(x^n) -$$
$$-(u^n, b - Ax^n) = f(z) - \varphi(x^n, u^n) \leqslant f(z) - \varphi(x^0, u^0).$$

Hence, in view of condition V, it follows that the set of elements of the sequence $\{x^n\}$ is compact. In addition, from (2.25) it follows that

$$\varphi(x^n, u^n) \leqslant f(z), \qquad (2.26)$$

This together with (2.22) shows that there exists

$$\lim_{n\to\infty} \varphi(x^n, u^n) \leqslant f(z). \qquad (2.27)$$

4. From (2.23), (2.24) and (2.27) it follows that $D(x^{n+1}, x^n) \to 0$. Therefore, repeating the reasoning of Theorem 1 it can be shown that

any limiting point x^* of the sequence $\{x^n\}$ belongs to the set R. Apart from this, since condition (2) is satisfied, we have $x^n \to x^* \in R$.

5. Let

$$I_1 = \{i \in \{1,2,\dots,m\} \mid (A_{i\cdot}, x^*) > b_i\},$$
$$I_2 = \{i \in \{1,2,\dots,m\} \mid (A_{i\cdot}, x^*) = b_i\}.$$

We choose N so that when $n > N$ for $i \in I_1$, we have $(A_{i\cdot}, x^n) > b_i$. Then when $n > N + m$ for $i \in I_1$, we obtain $u_i^n = 0$. Therefore when $n > N + m$

$$(u^n, Ax^n - b) = \sum_{i \in I_2} u_i^n (A_{i\cdot}, x^n - x^*) = \sum_{i=1}^{m} u_i^n (A_{i\cdot}, x^n - x^*) =$$
$$= (g(x^n), x^n - x^*) = D(x^*, x^n) - f(x^*) + f(x^n).$$

Since condition (2) is satisfied and the function $f(x)$ is continuous over $\bar{S}$, $(n^n, Ax^n - b) \to 0$. Hence $\lim \varphi(x^n, u^n) = \lim f(x^n) - (u^n, Ax^n - b)) = f(x^*)$. Comparing this with (2.27), we obtain that x^* is a solution of the problem (2.11) - (2.13). The theorem is proved.

Note 4. Let us consider a problem which is the dual of (2.11) - (2.13) (see [7]):

to maximize

$$\varphi(x, u) = f(x) - (u, Ax - b) \tag{2.28}$$

subject to the conditions

$$g(x) = uA, \tag{2.29}$$
$$u \geqslant 0, \quad x \in \bar{S}. \tag{2.30}$$

From Theorem 4 we have $\min f(x) = \sup \varphi(x, y)$, where the minimum is taken over the set of vectors x, satisfying conditions (2.12) and (2.13), while the upper bound is taken over the whole set of pairs (x, u), satisfying conditions (2.29) and (2.30).

5. If the sequence $\{u^n\}$ has the limiting point u^*, the pair (x^*, u^*) is a solution of the problem (2.28) - (2.30).

We shall give an example. Suppose we are given the problem:

to minimize

$$\sum_{j=1}^{p} x_j \ln x_j \tag{2.31}$$

subject to the conditions

$$\sum_{j=1}^{p} a_{ij} x_j = b_i \quad (i = 1, 2, \ldots, m), \tag{2.32}$$

$$x \in \overline{S}, \tag{2.33}$$

where $S = \{x \in E^p | x > 0\}$.

We shall assume that $\sum_{j=1}^{p} a_{ij}^2 > 0$ for all i. As has been shown in Section 1, the function $D(x, y)$ in this case satisfies conditions I - VI and (2). Apart from this, the function (2.31) is continuously differentiable over S and is continuous over $\overline{S}$. Let $y \in S$, and x be the D-projection of the point y onto the set

$$A_i = \left\{ x \,\middle|\, \sum_{j=1}^{p} a_{ij} x_j = b_i \right\}.$$

Then it follows from formulae (2.17) and (2.8) that

$$x_j = y_j \exp(\lambda a_{ij}), \tag{2.34}$$

where λ is the unique root of the equation

$$\sum_{j=1}^{p} a_{ij} y_j \exp(\lambda a_{ij}) = b_i. \tag{2.35}$$

As can be seen from (2.34), if $y \in S$, then also $x \in S$. Consequently, with a suitable relaxation control the conditions of Theorem 3 are satisfied. Therefore the relaxation sequence will converge to the point $\{x^*\}$, that is, to the solution of the problem (2.31) - (2.33), if the point of absolute minimum of the function (2.31), i.e. $x_j^0 = e^{-1}$, is chosen as the initial approximation.

We remark that equation (2.35) is converted into a linear relationship in terms of e^λ if all the coefficients a_{ij} are equal to 0 or 1. This happens in particular for problems with transport restrictions, i.e. restrictions of the kind

$$\sum_{j=1}^{p_1} x_{ij} = a_i, \qquad \sum_{i=1}^{p_2} x_{ij} = b_j.$$

The relaxation method for problems with this kind of restriction is identical with the method of Sheleikhovskii [8].

Finally we remark that the relaxation method can be used for the approximate solution of problems in linear programming.

Suppose we are given the linear programming problem:

to minimize

$$\sum_{j=1}^{p} c_jx_j \tag{2.36}$$

subject to the conditions

$$\sum_{j=1}^{p} a_{ij}x_j = b_i \qquad (i = 1, 2, \ldots, m), \tag{2.37}$$

$$x_j \geqslant 0. \tag{2.38}$$

Instead of this problem we shall solve the problem of minimizing the function

$$\mathop{}_{j=1} c_jx_j + \varepsilon f(x) \tag{2.39}$$

subject to the conditions (2.37) and (2.38). Here $\varepsilon > 0$, and $f(x)$ is the same as earlier.

If the set of vectors x, satisfying conditions (2.37) and (2.38) is finite and ε is sufficiently small, the solution of the problem (2.39), (2.37) - (2.38) will be an approximate solution of the problem (2.36) - (2.38).

If for $f(x)$ we take the function $\sum_{j=1}^{p} x_j \ln x_j$, we can solve the problem of minimizing the function (2.39) with only the restriction (2.37), since the restrictions $x_j \geqslant 0$ will be automatically satisfied.

Translated by G. Kiss

REFERENCES

1. AGMON, S. The relaxation method for linear inequalities. *Can. J. Math.* 6, 3, 382 - 392, 1954.

2. MOTZKIN, T.S. and SCHOENBERG, I.I. The relaxation method for linear inequalities. *Can. J. Math.* 6, 3, 393 - 404, 1954.

3. EREMIN, I.I. A generalization of the Motzkin - Agmon relaxational method. *Usp. mat. Nauk* 20, 183 - 187, 1965.

4. BREGMAN, L.M. Finding the common point of convex sets by the method of successive projections. *Dokl. Akad. Nauk SSSR* 162, 487 - 490, 1965.

5. FADDEEV, D.K. and FADDEEVA, V.N. *Computational methods of linear algebra* (Vychislitel'nye metody lineinoi algebry), Fizmatgiz, Moscow, 1963.

6. KARLIN, S. Mathematical methods in the theory of games, programming and economics (in Russian translation), Mir, Moscow, 1964.

7. DORN, W.D. A duality theorem for convex programms. *IBM J. Res. Dev.* 4, 4, 407 - 413, 1960.

8. BREGMAN, L.M. Proof of the convergence of Sheliekhovskii's method for problems with transport restrictions. *Zh. vȳchisl. Mat. mat. Fiz.* 7, 1, 147 - 156, 1967.