



Soviet-era science, translated into English

L. M. Bregman

1965

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196501.42016>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

L. M. Bregman

FINDING A COMMON POINT OF CONVEX SETS BY THE METHOD OF SUCCESSIVE PROJECTION

(Presented by Academician L. V. Kantorovich, 7 XII 1964)

In the present article an iterative method is considered for finding a common point of convex sets. This method can be applied to problems of optimal programming and to some others.

Let closed convex sets A_i , $i \in I$, where I is some set of indices, be given in a real Hilbert space H with distance ρ . Let $R = \bigcap_{i \in I} A_i$ be nonempty. It is required to find some point $x \in R$. Consider the following iterative process: take an arbitrary point $x_0 \in H$, then choose $i(x_0) \in I$, and in the set $A_{i(x_0)}$ find the point x_1 nearest to x_0 ; then in the same way choose $i(x_1) \in I$, and in the set $A_{i(x_1)}$ find the point x_2 nearest to x_1 , and so on.

It is natural to call this method the **method of successive projection**. Various variants of this method are possible, differing from one another in the rule for choosing the index $i(x)$. We shall consider some ways of choosing $i(x)$ and prove that in these cases the sequence $\{x_n\}$ will converge to some point $x^* \in R$. We shall need the following lemma:

Lemma. Let A be a closed convex set in H ; $x \in H$; y the point of the set A nearest to x . Then for any point $z \in A$,

$$(x - y, y - z) \geq 0.$$

The proof of the lemma is carried out in the same way as the proof of the theorem on a supporting hyperplane in (1).

Corollary 1. $\|y - z\| \leq \|x - z\|$.

Indeed, from the lemma it follows that

$$\|y - z\|^2 \leq (x - z, y - z).$$

Hence,

$$\|y - z\|^2 \leq \|x - z\| \|y - z\|$$

and

$$\|y - z\| \leq \|x - z\|.$$

Corollary 2.

$$\|y - x\|^2 \leq \|x - z\|^2 - \|y - z\|^2;$$

$$\|y - x\|^2 = \|x - z\|^2 - \|y - z\|^2 - 2(x - y, y - z) \leq \|x - z\|^2 - \|y - z\|^2.$$

A sequence $\{x_n\} \subset H$ converges weakly to $x \in H$ ($x_n \rightarrow x$) if $(u, x_n) \rightarrow (u, x)$ for every $u \in H$.

1. Let I be a finite set, $I = \{1, 2, \dots, p\}$. We shall take the indices in cyclic order, i.e.,

$$i(x_0) = 1, \quad i(x_1) = 2, \dots, \quad i(x_{p-1}) = p,$$

$$i(x_p) = 1, \quad i(x_{p+1}) = 2,$$

and so on. Accordingly the sequence $\{x_n\}$ is constructed as follows: x_0 is an arbitrary point; x_1 is the point of the set A_1 nearest to x_0 ; x_2 is the point of the set A_2 nearest to x_1 , ..., x_{p+1} is the point of the set A_1 nearest to x_p , and so on.

Theorem 1. x_n converges weakly to $x^* \in R$.

Proof. Take an arbitrary point $z \in R$. On the basis of Corollary 1 the set $\{x_n\}$ is bounded and, consequently, weakly compact. Let y_k^i be the notation for x_{kp+i} ($k = 0, 1, \dots$; $i = 0, 1, \dots, p - 1$). From the sequence $\{y_k^1\}$ choose a subsequence $\{y_{k_\nu}^1\}$ converging weakly to some point x^* . Since $y_{k_\nu}^1 \in A_1$ and the set A_1 is closed and convex, $x^* \in A_1$ (2). On the basis of Corollary 2,

$$\|y_{k_\nu}^1 - y_{k_\nu}^2\|^2 \leq \|y_{k_\nu}^1 - z\|^2 - \|y_{k_\nu}^2 - z\|^2.$$

Since there exists

$$\lim \|x_n - z\|,$$

it follows that

$$\|y_{k_\nu}^2 - y_{k_\nu}^1\| \rightarrow 0$$

and for every $u \in H$,

$$(u, y_{k_\nu}^1 - y_{k_\nu}^2) \rightarrow 0.$$

Consequently,

$$\lim(u, y_{k_\nu}^2) = \lim(u, y_{k_\nu}^1) = (u, x^*).$$

Hence $y_{k_\nu}^2 \rightarrow x^*$, and, since $y_{k_\nu}^2 \in A_2$, we have $x^* \in A_2$. Similarly one can show that $x^* \in A_3, \dots, A_p$. Consequently, $x^* \in R$.

Let us prove that the entire sequence x_n converges weakly to x^* . Suppose there exists a subsequence $x_{n_l} \rightarrow x^{**}$. Arguing as before, we obtain $x^{**} \in R$. By Corollary 1, there exists

$$a = \lim(\|x_n - x^*\|^2 - \|x_n - x^{**}\|^2) = \lim 2(x_n, x^{**} - x^*) + \|x^*\|^2 - \|x^{**}\|^2.$$

For the subsequence $\{y_{k_\nu}^1\}$, $a = -\|x^* - x^{**}\|^2$. For the subsequence $\{x_{n_l}\}$,

$$a = \|x^* - x^{**}\|^2.$$

Consequently, $a = 0$ and $x^* = x^{**}$. Hence $x_n \rightarrow x^*$.

2. Let I be an arbitrary set of indices, and suppose that the following condition is fulfilled:

$$\text{for every } x \in H \text{ there exists } \max_{i \in I} \rho(x, A_i). \quad (1)$$

In this case, as $i(x_n)$ we choose the index for which $\max_{i \in I} \rho(x_n, A_i)$ is attained (if such an index is not unique, we take any one of these indices). As x_{n+1} we take the point of the set $A_{i(x_n)}$ nearest to x_n .

Theorem 2. x_n converges weakly to $x^* \in R$.

Proof. Take an arbitrary point $z \in R$. On the basis of Corollary 1,

$$\rho(x_{n+1}, z) \leq \rho(x_n, z).$$

Consequently, the set $\{x_n\}$ is weakly compact. Choose a subsequence x_{n_k} that converges weakly to some point x^* . Since $\lim \rho(x_n, z)$ exists and, by Corollary 2,

$$(\rho(x_n, x_{n+1}))^2 \leq (\rho(x_n, z))^2 - (\rho(x_{n+1}, z))^2,$$

we have $\rho(x_n, x_{n+1}) \rightarrow 0$. Since $\rho(x_n, A_i) \leq \rho(x_n, x_{n+1})$, it follows that $\rho(x_n, A_i) \rightarrow 0$ for every $i \in I$. Let $\{y_k\} \subset A_i$ and $\rho(x_{n_k}, y_k) \rightarrow 0$. The set $\{y_k\}$ is weakly compact. Suppose

$$y_{k_l} \rightarrow y^*.$$

Then $y^* \in A_i$, and since $\|x_{n_{k_l}} - y_{k_l}\| \rightarrow 0$, we have $x_{n_{k_l}} \rightarrow y^*$, and, consequently, $x^* = y^* \in A_i$. Since this is true for every i , $x^* \in R$. The fact that the entire sequence $\{x_n\}$ converges weakly to x^* is proved in the same way as in Theorem 1.

3. Let the sets A_i be half-spaces, i.e.

$$A_i = \{x \mid (f_i, x) \leq a_i\}, \quad f_i \in H,$$

and let a_i be real numbers. Suppose the following conditions are fulfilled:

$$\text{for every } x \in H \text{ there exists } \max_{i \in I} [(f_i, x) - a_i]; \quad (2)$$

$$\text{there exists } M \text{ such that } \|f_i\| \leq M \text{ for all } i \in I. \quad (3)$$

In this case, as $i(x_n)$ we take the index for which

$$\max_{i \in I} [(f_i, x_n) - a_i]$$

is attained. Accordingly we construct the sequence $\{x_n\}$. If $x_n \notin R$, then it is easy to see that

$$x_{n+1} = x_n + f_{i(x_n)} [a_{i(x_n)} - (f_{i(x_n)}, x_n)] \|f_{i(x_n)}\|^{-2}. \quad (4)$$

Theorem 3. $\{x_n\}$ converges weakly to $x^* \in R$.

Proof. If $x_n \notin R$, then

$$(f_i, x_n) - a_i \leq (f_{i(x_n)}, x_n) - a_{i(x_n)} = \|f_{i(x_n)}\| \|x_n - x_{n+1}\| \leq M (\|x_n - z\|^2 - \|x_{n+1} - z\|^2)^{1/2}$$

for every $z \in R$. Since the set $\{x_n\}$ is weakly compact, there exists a subsequence $x_{n_k} \rightarrow x^*$. Since

$$(f_i, x_n) - a_i \leq M (\|x_n - z\|^2 - \|x_{n+1} - z\|^2)^{1/2} \rightarrow 0,$$

we have

$$\lim (f_i, x_{n_k}) - a_i = (f_i, x^*) - a_i \leq 0$$

for every $i \in I$. Consequently, $x^* \in R$. The fact that the entire sequence $\{x_n\}$ converges weakly to x^* is proved in the same way as in Theorem 1.

Remark 1. If H is a Euclidean space, then weak convergence in it coincides with convergence in norm, and, consequently, in all theorems weak convergence can be replaced by ordinary convergence.

Remark 2. Let p sets of indices I_1, I_2, \dots, I_p be given;

$$R_j = \bigcap_{i \in I_j} A_i, \quad R = \bigcap_{j=1}^p R_j$$

is nonempty; all A_i are closed convex sets. Suppose, further, that for each set of indices I_j either condition (1), or conditions (2) and (3), are satisfied. Then one can show that if we choose the index $i(x_n)$ successively from the sets $I_1, I_2, \dots, I_p, I_1, I_2, \dots$, as was done in cases 2 or 3, and accordingly construct the sequence $\{x_n\}$, then $\{x_n\}$ will converge weakly to $x^* \in R$.

Let us consider some applications of these theorems.

I. The linear programming problem. Suppose it is required to **maximize**

$$\sum_{j=1}^n c_j x_j$$

under the conditions $x_j \geq 0$, $\sum_{j=1}^n a_{ij} x_j \leq b_j$ ($i = 1, 2, \dots, m$).

By the duality theorem (3), the vector $x = (x_1, x_2, \dots, x_n)$ is a solution of the problem if and only if there exists a vector $u = (u_1, u_2, \dots, u_m)$ such that the following conditions are satisfied:

$$\begin{aligned} \sum_{j=1}^n a_{ij} x_j \leq b_i \quad (i = 1, 2, \dots, m); \quad x_j \geq 0 \quad (j = 1, 2, \dots, n); \\ \sum_{i=1}^m a_{ij} u_i \geq c_j, \quad u_i \geq 0; \quad \sum_{j=1}^n c_j x_j \geq \sum_{i=1}^m u_i b_i. \end{aligned} \quad (5)$$

Thus, the linear programming problem is reduced to a system of linear inequalities. The system (5) can be solved by the method of successive projection, taking arbitrary initial vectors x_0 and u_0 and obtaining successive approximations by formulas of type (4). This method is convenient for solving large-scale problems with a sparse coefficient matrix.

A similar method was considered by V. A. Bulavskii in (4). Its difference consists in the fact that, instead of changing the vector u , the vector x is shifted along the gradient (the vector $x'_n = x_n + \sigma c$ is considered, and the next iteration begins with the vector x'_n). If σ is a sufficiently small positive number, then $x^* = \lim x_n$ is close to the solution of the problem under consideration. In this way one can obtain a sequence converging to the solution of the dual problem. V. A. Bulavskii's method is more convenient in that it is not necessary to store the vector u ; however, the choice of σ presents considerable difficulties.

II. The quadratic programming problem. Consider the following problem: **find a vector x minimizing** $(p, x) + (x, C, x)$ **under the conditions** $Ax = b$, $x \geq 0$, where p and x are n -dimensional vectors; b is an

m -dimensional vector; A is a matrix of size $m \times n$; C is a nonnegative definite quadratic matrix of order n .

It can be shown (cf. (5)) that the vector x is a solution of the problem II under consideration if and only if there exists an m -dimensional vector u such that the following conditions are satisfied:

$$Ax = b; \quad x \geq 0; \quad 2xC - uA \geq -p; \quad (6)$$

$$2(xC, x) + (p, x) - (u, b) \leq 0. \quad (7)$$

The conditions (6) are linear. Each of them determines a certain half-space in $(m + n)$ -dimensional Euclidean space. The point in the half-space nearest to a given point can be found by formula (4). Condition (7) determines a certain convex set in $(m + n)$ -dimensional Euclidean space. This set can be represented as an intersection of supporting half-spaces. Hence it follows that the vector x and u satisfy condition (7) if and only if

$$((4yC + p), (x - y)) - (u - v, b) \leq 0 \quad (8)$$

for all y and v such that

$$2(yC, y) + (p, y) - (v, b) = 0. \quad (9)$$

Thus, condition (7) is replaced by the linear constraints (8), and, for any x and u , there exists

$$\max[(4yC + p, x - y) - (u - v, b)]$$

over all y and v satisfying condition (9). This maximum is attained at $y = x$ and $(v, b) = 2(xC, x) + (p, x)$. Consequently, we are in the conditions of Theorem 3, and therefore we can solve the problem II under consideration by the method of successive projection. In doing so, if for the point (x_n, u_n) we seek the nearest point in the set defined by one of the constraints (8), then, choosing it as in case 3 and applying formula (4), we obtain

$$x_{n+1} = x_n + (p + 4x_{nC}) \frac{(b, u_n) - (p, x_n) - 2(x_{nC}, x_n)}{\|p + 4x_{nC}\|^2 + \|b\|^2},$$

$$u_{n+1} = u_n - b \frac{(b, u_n) - (p, x_n) - 2(x_{nC}, x_n)}{\|p + 4x_{nC}\|^2 + \|b\|^2}.$$

III. Infinite programs. Let $H = L^2_T$; A be a continuous linear operator from L^2 to L^2 . Consider the problem: find $\sup(a, x)$ subject to $Ax \leq b$, $x \geq 0$.

As shown in (6), if there exists a function $u \in L^2$ such that $u \geq 0$ and $A^*u > a$ (A^* is the operator adjoint to A), then there exists a function $u \in L^2$ such that the conditions are satisfied:

$$Ax \leq b; \quad (10)$$

$$x \geq 0; \quad (11)$$

$$A^*u \geq a; \quad (12)$$

$$u \geq 0; \quad (13)$$

$$(a, x) + \varepsilon \geq (b, u) \quad \text{for any } \varepsilon > 0. \quad (14)$$

Moreover, if x^* and u^* satisfy conditions (10)–(14), then (a, x^*) differs from $\sup(a, x)$ by no more than ε . We replace condition (10) by the conditions: for all $r \in L^2$ such that $r \geq 0$ and $\|r\| = 1$, $(r, Ax) \leq (r, b)$. For any $x \in L^2$ there exists

$$\max_{r \geq 0, \|r\|=1} (r, Ax - b).$$

It is attained at

$$r = \max(0, Ax - b) \cdot \|\max(0, Ax - b)\|^{-1}.$$

We replace conditions (11)–(13) analogously. Then we are in the conditions of Theorem 3 (or Remark 2), and, consequently, the system (10)–(14) can be solved by the method of successive projection. The solution of this system will give us an approximate solution of the problem III under consideration.

Leningrad State University
named after A. A. Zhdanov

Received
30 XI 1964

REFERENCES

1. C. Wajda, in: *Linear Inequalities and Related Questions*, IL, 1959.
2. N. Dunford, J. Schwartz, *Linear Operators*, 1, IL, 1962.

3. A. J. Goldman, A. W. Tucker, in: *Linear Inequalities and Related Questions*, II, 1959.
4. V. A. Bulavskii, Iterative method for solving a linear programming problem, Abstract of a report, Novosibirsk, 1962.
5. J. B. Dennis, *Mathematical Programming and Electrical Networks*, II, 1961.
6. R. J. Duffin, in: *Linear Inequalities and Related Questions*, II, 1959.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.