Soviet-era science, translated into English

A RELAXATION METHOD FOR FINDING A COMMON POINT OF CONVEX SETS AND ITS APPLICATION TO OPTIMIZATION PROBLEMS

MATHEMATICS

1966

SovietRxiv

View the original and related papers at https://sovietrxiv.org/items/ru-196601.22781

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

Abstract

Full Text

UDC 519.25/26+519.3:330.115

MATHEMATICS

L. M. BREGMAN

A RELAXATION METHOD FOR FINDING A COMMON POINT OF CONVEX SETS AND ITS APPLICATION TO OPTIMIZATION PROBLEMS

(Presented by Academician L. V. Kantorovich on 26 II 1966)

Let closed convex sets A_i $(i \in I)$ be given in a linear topological space X. Suppose that $R = \bigcap_{i \in I} A_i$ is nonempty.

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

Consider a function $D(x, y)$, defined on $S \times S$ and possessing the following properties:

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

(II). For each $i \in I$ and $y \in S$ there exists a point $x = P_i y \in A_i \cap S$ such that $D(x, y) \leq D(z, y)$ for all $z \in A_i \cap S$. We shall call this point x 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 on $A_i \cap S$.

(IV). There exists the derivative $D'_x(x, y)$ of the function $D(x, y)$ at $x = y$, and moreover

$$D'_x(y, y) = 0 \quad \left(\text{i.e. } \lim_{t \to 0} \frac{D(y + tz, y)}{t} = 0 \text{ for all } z \in X\right).$$

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

(VI). If $D(x_n, y_n) \to 0$, $y_n \to y^* \in \bar{S}$, and the set of elements of the sequence $\{x_n\}$ is compact, then $x_n \to y^*$.

Lemma 1. Let $z \in A_i \cap S$. Then for any $y \in S$ the inequality holds

$$D(P_i y, y) \leq D(z, y) - D(z, P_i y).$$

Proof. According to condition (III), we have

$$D(\lambda z + (1-\lambda) P_i y, y) - D(\lambda z + (1-\lambda) P_i y, P_i y) \leq$$

$$\leq \lambda \big(D(z, y) - D(z, P_i y)\big) + (1-\lambda) D(P_i y, y)$$

for all $\lambda \in [0, 1]$. Hence for $\lambda > 0$ we obtain

$$D(z, y) - D(z, P_i y) - D(P_i y, y) \geq$$

$$\geq \frac{D(\lambda z + (1-\lambda) P_i y, y) - D(P_i y, y)}{\lambda} - \frac{D(\lambda z + (1-\lambda) P_i y, P_i y)}{\lambda} \geq$$

$$\geq -\frac{D(\lambda z + (1-\lambda) P_i y, P_i y)}{\lambda}.$$

Passing to the limit in this inequality as $\lambda \to 0$, we obtain

$$D(z, y) - D(z, P_i y) - D(P_i y, y) \geq 0.$$

The lemma is proved.

Consider the following iterative process: take an arbitrary point $x_0 \in S$, then choose $i_1(x_0) \in I$ and find a point x_1—the D-projection of the point x_0 onto the set $A_{i_1(x_0)}$; then choose $i_2(x_1) \in I$ and find a point x_2—the D-projection of the point x_1 onto the set $A_{i_2(x_1)}$, and so on. The sequence $\{x_n\}$ obtained as a result of this process will be called a **relaxation sequence**, and the set of functions $\{i_1(x), i_2(x), ...\}$ will be called a **relaxation control** (cf. (1)).

Lemma 2. For any relaxation control:

1. The set of elements of the relaxation sequence $\{x_n\}$ is compact.
2. $D(x_{n+1}, x_n) \to 0$ as $n \to \infty$.
3. For any $z \in R \cap S$ there exists $\lim\limits_{n \to \infty} D(z, x_n)$.

Proof. Take $z \in R \cap S$. By Lemma 1,

$$D(x_{n+1}, x_n) \leq D(z, x_n) - D(z, x_{n+1}). \tag{1}$$

Since $D(x_{n+1}, x_n) \geq 0$, it follows that $D(z, x_{n+1}) \leq D(z, x_n)$, and consequently $\lim D(z, x_n)$ exists; together with (1) this gives $D(x_{n+1}, x_n) \to 0$. Since the set of elements of the relaxation sequence is contained in the set

$$T = \{x \mid D(z, x) \leq D(z, x_0)\},$$

and, according to (V), the set T is compact, assertion 1 is true.

Let us now consider some relaxation controls.

Theorem 1. Let $I = \{1, 2, \ldots, m\}$, and let the indices be chosen in cyclic order, i.e.

$$i_1(x) = 1, \quad i_2(x) = 2, \ldots, \quad i_m(x) = m, \quad i_{m+1}(x) = 1, \ldots$$

Then any limit point x^* of the relaxation sequence $\{x_n\}$ is a common point of the sets A_i.

Proof. By Lemma 2, from the sequences $\{x_{nm+i}\} \subset A_i$ one can extract convergent subsequences. Let $x_{n_k m+i} \to x_i^* \in A_i$. According to Lemma 2,

$$D(x_{n_k m+2}, x_{n_k m+1}) \to 0.$$

By property (VI), $x_{n_k m+1} \to x_2^*$. Therefore, $x_1^* = x_2^*$. Analogously one can show that

$$x_2^* = x_3^*, \quad x_3^* = x_4^*, \ldots, \quad x_{m-1}^* = x_m^*.$$

Consequently, the limit point

$$x^* = x_1^* = \cdots = x_m^* \in R.$$

Theorem 2. Suppose that for each $y \in S$ there exists

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

and let $i_n(x) = i(x)$. Then any limit point x^* of the relaxation sequence $\{x_n\}$ is a common point of the sets A_i.

Proof. Let $x_{n_k} \to x^*$. Let $y_{n_k}^i = P_i x_{n_k} \in A_i$. Then

$$D(y_{n_k}^i, x_{n_k}) \leq \max_{i \in I} D(y_{n_k}^i, x_{n_k}) = D(x_{n_k+1}, x_{n_k}).$$

Consequently, $D(y_{n_k}^i, x_{n_k}) \to 0$ for all $i \in I$. Since for any $z \in R \cap S$

$$D(z, y_{n_k}^i) \leq D(z, x_{n_k}) \leq D(z, x_0),$$

then, according to (V), the set $\{y^i_{n_k}\}$ is compact. Consequently, by property (VI), $y^i_{n_k} \to x^*$ for all $i \in I$. Hence it follows that $x^* \in R$.

Remark. In many cases the relaxation sequence $\{x_n\}$ has a unique limit point. This is the case, for example, if the following condition is fulfilled:

(A). 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 on S.

In fact, let x^* and x^{**} be limit points of the relaxation sequence $\{x_n\}$ and let $x^*, x^{**} \in R$. Let $x_{n_k} \to x^*$, $x_{n_l} \to x^{**}$. By Lemma 2, there exists

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

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

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

Since $\lim H(x_{n_k}) = \lim H(x_{n_l}) = \lim H(x_n)$, it follows that $D(x^{**}, x^*) = 0$, and hence $x^{**} = x^*$.

Let us consider some examples of functions satisfying conditions (I)–(VI).

1. Let X be a real Hilbert space. The function

   $$D(x, y) = (x - y, x - y)$$

   satisfies conditions (I)–(VI) and condition (A) for any system of closed convex sets A_i, if convergence is understood as weak convergence, and compactness as weak compactness. In this case the D-projection onto a convex set coincides with the usual projection. The corresponding relaxation sequence $\{x_n\}$ under the conditions of Theorem 1 or 2 will converge weakly to an element $x^* \in R$. In this case the relaxation method coincides with the method of successive projection (2).

2. Let $f(x)$ be a strictly convex twice differentiable function defined on some convex closed set $S \subset E^n$. Let $g(x)$ be the gradient of this function at the point x. Consider the function

   $$D(x, y) = f(x) - f(y) - (g(y), x - y).$$

   If this function satisfies condition (V), then, as is easy to verify, it satisfies conditions (I)–(IV), (VI) and (A) for any system of closed convex sets A_i.

Suppose the following conditions are fulfilled:

1) The absolute minimum of the function $f(x)$ is attained at a point x_0 lying in the interior of S.
2) The sets A_i are hyperplanes, i.e.

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

   and the set I is finite.
3) If y belongs to the interior of S, then the D-projection of the point y onto any set A_i also belongs to the interior of S.

In this case one can show that the corresponding relaxation sequence with initial approximation x_0 converges to the point x^*, which delivers the minimum of the function $f(x)$ under the conditions

$$\sum_{j=1}^{n} a_{ij} x_j = b_i$$

$(i \in I)$.

3. Let

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

   be given on the set

   $$\{x \geq 0,\ y > 0\}.$$

   This function satisfies conditions (I)–(VI), but does not satisfy condition (A). Nevertheless, in this case too the relaxation sequence will have a unique limit point.

Indeed, if $x^* \in R$ is a limit point of the sequence $\{x_n\}$ and $x_{n_k} \to x^*$, then $\lim D(x^*, x_{n_k}) = 0$, and since $\lim D(x^*, x_n)$ exists, by property (VI) $x_n \to x^*$.

If

$$A_i = \left\{ x \;\middle|\; \sum_{j=1}^{n} a_{ij} x_j = b_i \right\}, \qquad \left(\sum_{j=1}^{n} a_{ij}^2 > 0 \right),$$

then x^* maximizes

$$\sum_{j=1}^{n} x_j \ln \frac{p_j}{x_j} \quad (p_j > 0)$$

under the conditions

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad x_j \geqslant 0,$$

if the vector $x_0 = \{p_j/e\}$ is taken as the initial approximation. The D-projection x of the point y onto the set A_i is found from the formulas $x_j = y_j e^{\lambda a_{ij}}$, where λ is the unique root of the equation

$$\sum_{j=1}^{n} a_{ij}y_j e^{\lambda a_{ij}} = b_i.$$

This equation becomes linear with respect to e^{λ} if all a_{ij} are equal to 0 or 1.

The latter occurs, for example, in the following problem: maximize

$$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} \ln \frac{p_{ij}}{x_{ij}} \quad (p_{ij} \geqslant 0)$$

under the conditions

$$\sum_{j=1}^{n} x_{ij} = a_i, \qquad \sum_{i=1}^{m} x_{ij} = b_j;$$

$x_{ij} \geqslant 0$; $x_{ij} = 0$ if $p_{ij} = 0$. Such problems arise in the calculation of passenger flows in cities, and the corresponding relaxation method for this problem coincides with the method of G. V. Sheleikhovsky (see (^{3})).

Leningrad State University
named after A. A. Zhdanov

Received
13 II 1966

REFERENCES

[1] D. K. Faddeev, V. N. Faddeeva, *Computational Methods of Linear Algebra*, Moscow, 1963. [2] L. M. Bregman, DAN, **162**, No. 3, 487 (1965). [3] A. G. Dynkin, E. G. Movchan, in: *Application of Mathematical Methods and Electronic Computers in Urban Planning*, Kiev, 1966.

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

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