

ON THE SOLUTION OF CONVEX PROGRAMMING PROBLEMS WITH LINEAR CONSTRAINTS BY THE METHOD OF SUCCESSIVE IMPROVEMENT OF A FEASIBLE VECTOR

The basic problem.** Given vectors

1963

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196301.09030>

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

Abstract

Full Text

MATHEMATICS

V. A. BULAVSKII, G. Sh. RUBINSHTEIN

ON THE SOLUTION OF CONVEX PROGRAMMING PROBLEMS WITH LINEAR CONSTRAINTS BY THE METHOD OF SUCCESSIVE IMPROVEMENT OF A FEASIBLE VECTOR

(Presented by Academician S. L. Sobolev on 10 XII 1962)

1. This work is devoted to the generalization of one of the effective techniques for solving linear programming problems—the method of successive improvement of a feasible vector ⁽¹⁾, pp. 315-326—to the case of convex programming problems with linear constraints ⁽²⁾.

The basic problem. Given vectors

$$\alpha^i = (a_{i1}, a_{i2}, \dots, a_{in}), \quad i \in I = \{1, 2, \dots, m\}, \quad \beta = (b_1, b_2, \dots, b_n)$$

and a concave function $f(x)$, twice continuously differentiable on the vectors

$$x = (x_1, x_2, \dots, x_m), \quad x_i \geq 0 \quad (i \in I), \quad (1)$$

i.e., such that for any vectors (1) and $l = (l_1, l_2, \dots, l_m)$

$$\sum_{i,j \in I} f_{ij}(x) l_i l_j \leq 0, \quad (2)$$

where $f_{ij}(x) = \partial^2 f(x) / \partial x_i \partial x_j$. It is required to find a vector (1) from the conditions:

$$1^\circ. \sum_{i \in I} x_i \alpha^i = \beta.$$

2°. The function $f(x)$ attains a maximum.

A vector (1) satisfying 1° is called **feasible**, and the sought vector is called **optimal**.

2. As is known (2), for the optimality of a feasible vector (1) it is necessary and sufficient that there exist a vector

$$y = (y_1, y_2, \dots, y_n) \quad (3)$$

such that

$$(\alpha^i, y) = \sum_{j=1}^n a_{ij} y_j \geq f_i(x), \quad i \in I, \quad (4)$$

where $f_i(x) = \partial f(x) / \partial x_i$, and moreover

$$(\alpha^i, y) = f_i(x) \quad \text{for } i \in I(x) = \{i : x_i > 0\}. \quad (5)$$

For simplicity in presenting the method, we shall assume the following to hold:

- 1) For any feasible vector (1), among the vectors $\alpha^i (i \in I(x))$ there are n linearly independent ones.

If this condition is violated in the course of solving the problem, then, as in the case of linear programming, so-called degeneracy situations may arise; they are overcome here in exactly the same way as in linear programming.

The justification of the algorithm essentially rests on the following proposition concerning the function $f(x)$:

- 2) If, for a given vector l , equality is attained in relation (2) for some feasible vector (1), then for any other feasible vector (1) the equality sign also holds in inequality (2).

Remark. In addition to linear and quadratic functions, condition 2) is satisfied, for example, by the concave function

$$f(x) = \sum_{k=1}^p \exp \left[\sum_{i \in I} c_{ki} x_i + d_k \right],$$

where c_{ki}, d_k are arbitrary real numbers.

If condition 2) is fulfilled, then for the existence of an optimal vector it is necessary and sufficient that there exist a feasible vector and that there be no feasible ray (consisting of feasible elements)

$$\{x + tl; t \in [0, +\infty)\}, \quad (6)$$

on which the function $f(x)$ strictly increases.

3. The process of constructing an optimal vector consists of several parts.

I. By means of techniques developed in linear programming, we either verify that in the problem under consideration there is no feasible vector, and consequently no optimum exists (the process terminates), or else find a feasible vector (1) containing exactly n components different from zero.

II. Suppose there is a feasible vector (1) satisfying the conditions*:

A. The corresponding system (5) is consistent.

B. For every nontrivial solution

$$l = (l_1, l_2, \dots, l_m) \quad (7)$$

of the homogeneous system $\sum_{i \in I} l_i a^i = 0$, $l_i = 0$ ($i \notin I(x)$), strict inequality holds in relation (2).

By virtue of 1), system (5) has the unique solution (3). We find this solution. If inequalities (4) hold for it, then vector (1) is optimal (the process terminates). Otherwise we pass to the next (principal) stage.

III. Suppose there is a feasible vector (1) such that:

a) $\{i_1, i_2, \dots, i_{n+r}\} \subset I(x) \subset \{i_1, i_2, \dots, i_{n+r}, i_{n+r+1}\}$, where $r \geq 0$, and the vectors

$$a^{i_1}, a^{i_2}, \dots, a^{i_n} \quad (8)$$

are linearly independent;

b) the system $(a^{i_k}, y) = f_{i_k}(x)$, $k = 1, 2, \dots, n+r$, is consistent and for its solution (3) (unique by virtue of a)) the inequality

$$(a^{i_{n+r+1}}, y) < f_{i_{n+r+1}}(x);$$

holds;

c) for every nontrivial solution (7) of the system

$$\sum_{i \in I} l_i a^i = 0, \quad l_i = 0 \quad (i \notin \{i_1, i_2, \dots, i_{n+r}\})$$

strict inequality holds in relation (2).

* The feasible vector (1), constructed in I, obviously satisfies these conditions.

For fixed linearly independent elements (8), to each $a^{i_{n+s}}$ ($s = 1, 2, \dots, r + 1$) we associate the vector

$$x^s = (x_1^s, x_2^s, \dots, x_m^s), \quad (9)$$

where $x_i^s = 0$ for $i \in \{i_1, i_2, \dots, i_n, i_{n+s}\}$, $x_{i_{n+s}}^s = 1$, and the remaining components are determined by the relation:

$$a^{i_{n+s}} + \sum_{k=1}^n x_{i_k}^s a^{i_k} = 0.$$

Then conditions) and) can be rewritten in the form:

$$') \quad \sum_{i \in I} f_i(x) x_i^s = 0 \quad (s = 1, 2, \dots, r), \quad \sum_{i \in I} f_i(x) x_i^{r+1} > 0;$$

') for any linear combination of the vectors (9)

$$l = \sum_{s=1}^r c_s x^s \neq 0 \quad (10)$$

a strict inequality holds in relation (2), i.e., the homogeneous system

$$\sum_{\sigma=1}^r \left(\sum_{i,j \in I} f_{ij}(x) x_i^s x_j^\sigma \right) c_\sigma = 0 \quad (s = 1, 2, \dots, r)$$

has only the trivial solution.

For fixed ε there exists at most one vector

$$\bar{x} = x + \varepsilon x^{r+1} + \sum_{s=1}^r g_s x^s, \quad (11)$$

satisfying the conditions:

$$\bar{x}_{i_k} \geq 0 \quad (k = 1, 2, \dots, n + r); \quad (12)$$

$$\sum_{i \in I} f_i(\bar{x}) x_i^{r+1} \geq 0; \quad (13)$$

$$\sum_{i \in I} f_i(\bar{x}) x_i^s = 0 \quad (s = 1, 2, \dots, r). \quad (14)$$

Indeed, if for some ε there were two vectors \bar{x} and $\bar{\bar{x}}$ satisfying (12)–(14), then for their difference $l = \bar{\bar{x}} - \bar{x}$, of the form (10), by virtue of (14) we would have

$$\sum_{i \in I} f_i(\bar{\bar{x}})l_i = 0, \quad \sum_{i \in I} f_i(\bar{x})l_i = 0.$$

However, these relations contradict the fact that for some $0 < \theta < 1$

$$\sum_{i \in I} f_i(\bar{\bar{x}})l_i = \sum_{i \in I} f_i(\bar{x})l_i + \sum_{i,j \in I} f_{ij}(\bar{x} + \theta l)l_i l_j$$

and, in view of (1) and assumption 2),

$$\sum_{i,j \in I} f_{ij}(\bar{x} + \theta l)l_i l_j < 0.$$

Consider the set E of nonnegative ε for which there does not exist a vector (11) such that conditions (12)–(14) are satisfied for it, with strict inequalities in (12) and (13). This set is, obviously, closed and, in view of 6'), does not contain $\varepsilon = 0$.

If E is the empty set, then, as can be shown, there exists an admissible ray (6) on which $f(x)$ strictly increases, and therefore in the problem under consideration there is no optimal vector (the process is finished).

If, however, $E \neq \Lambda$, then for the minimal $\varepsilon \in E$ there exists a vector (11) satisfying (12)–(14), and equality is attained in at least one of the inequalities (12) or (13). In this case two possibilities may occur:

- 1) In (13) a strict inequality holds, and among the vectors a'_k to which the strict inequalities (12) correspond there are n linearly independent ones.
- 2) At least one of the conditions of the preceding item is violated.

In the first case, part III can again be applied to the obtained vector (11); in the second, vector (11) satisfies conditions A and B, and therefore part II is applicable to it.

The monotone process described cannot continue indefinitely, since at each subsequent application of part II a new set $I(x) \subset I$ appears, while between two successive applications of II, part III of the algorithm can occur no more than $r + 1$ times.

4. Taking into account 6'), the system of equations (14) with respect to the vector $c = (c_1, c_2, \dots, c_r)$ can be replaced by the following:

$$\sum_{i \in I} \frac{1}{\varepsilon} [f_i(\bar{x}) - f_i(x)] x_i^s = 0 \quad (s = 1, 2, \dots, r), \quad (15)$$

or, approximately (for small ε):

$$\sum_{i,j \in I} f_{ij}(x) \left(x_j^{r+1} + \sum_{\sigma=1}^r c_\sigma x_j^\sigma \right) x_i^s = 0 \quad (s = 1, 2, \dots, r),$$

i.e., the linear system

$$\sum_{\sigma=1}^r \left(\sum_{i,j \in I} f_{ij}(x) x_i^s x_j^\sigma \right) c_\sigma = - \sum_{i,j \in I} f_{ij}(x) x_i^s x_j^{r+1} \quad (s = 1, 2, \dots, r). \quad (16)$$

This makes it possible, at stage III, to find vector (11) approximately, by giving small increments to the quantity ε .

In the case of convex quadratic programming, systems (15) and (16), obviously, coincide. Therefore, for this case an exact finite algorithm is obtained, in which at each step only linear systems have to be solved.

Institute of Mathematics with Computing Center
Siberian Branch of the Academy of Sciences of the USSR

Received
7 XII 1962

REFERENCES

1. L. V. Kantorovich, *Economic Calculation of the Best Use of Resources*, Publishing House of the Academy of Sciences of the USSR, 1959.
2. H. W. Kuhn, A. W. Tucker, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, 1951, p. 481.

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

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