



Soviet-era science, translated into English

MATHEMATICS

S. I. ZUKHOVITSKII, R. A. POLYAK, M. E. PRIMAK

1963

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196301.84495>

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

Abstract

Full Text

MATHEMATICS

S. I. ZUKHOVITSKII, R. A. POLYAK, M. E. PRIMAK

AN ALGORITHM FOR SOLVING A CONVEX PROGRAMMING PROBLEM

(Presented by Academician A. Yu. Ishlinskii, 26 VI 1963)

1. Let a system of p convex smooth functions be given,

$$f_i(x) \equiv f_i(x_1, \dots, x_n) \quad (i = 1, \dots, p) \quad (1)$$

and a domain D , defined by q inequalities

$$\varphi_j(x) \equiv \varphi_j(x_1, \dots, x_n) \leq 0 \quad (j = 1, \dots, q), \quad (2)$$

where the functions $\varphi_1(x), \dots, \varphi_q(x)$ are also convex and smooth.

The problem of convex Chebyshev approximation of the system (1) under the constraints (2) consists in finding the Chebyshev point of the system (1)–(2), i.e., a point $x^* \in D$ at which

$$\max_i f_i(x^*) = \min_{x \in D} \max_i f_i(x). \quad (3)$$

By introducing an additional variable ξ , we reduce problem (1)–(3) to the following general problem of convex programming:

Minimize the linear form

$$z = \xi \quad (4)$$

subject to the constraints

$$f_i(x) - \xi \leq 0 \quad (i = 1, \dots, p),$$

$$\psi_j(x) \leq 0 \quad (j = 1, \dots, q). \quad (5)$$

In a more convenient notation:

Minimize the linear form

$$z = p_1 x_1 + \dots + p_n x_n \quad (6)$$

subject to the constraints

$$\psi_j(x) \equiv \psi_j(x_1, \dots, x_n) \leq 0 \quad (j = 1, \dots, m), \quad (7)$$

which define a certain domain Ω , where $\psi_1(x), \dots, \psi_m(x)$ are convex, but not strictly convex, smooth functions.

In the present work, for solving problem (6)–(7) we give two algorithms that develop the algorithm constructed in ⁽¹⁾ for solving the problem of convex Chebyshev approximation without constraints. Other algorithms are indicated in ^(2–5).

The algorithms we present are not purely gradient algorithms. In the first of them, at each step two parameters δ_k and η_k are introduced, while in the second there is one parameter δ_k , which turns out to be convenient for computations and eliminates the danger of “jamming,” when successive corrections become arbitrarily small upon approaching a certain point which, however, is not a solution. The second algorithm, briefly described below in Sec. 5, coincides in idea with the algorithm given in ⁽⁵⁾*

* Work ⁽⁵⁾ became known to us after the present article had been prepared for publication.

2. As the initial approximation we take an arbitrary point $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)}) \in \Omega$, which we find by applying, for example, algorithm ⁽¹⁾ to the system of convex inequalities (7). Without loss of generality it may be assumed that $x^{(0)}$ already lies on the boundary of Ω .

Choose arbitrary sufficiently small $\delta_1 > 0$ and $\eta_1 > 0$, and let

$$-\delta_1 < \psi_{j_\nu}(x^{(0)}) \leq 0 \quad (\nu = 1, \dots, \nu_1);$$

$$\psi_j(x^{(0)}) \leq -\delta_1 \quad (j \neq j_1, \dots, j_{\nu_1}).$$

We define the descent direction $\xi^{(1)} \equiv (\xi_1^{(1)}, \dots, \xi_n^{(1)})$ from the following linear programming problem:

We shall minimize the linear form

$$u = p_1 \xi_1 + \dots + p_n \xi_n \quad (8)$$

subject to the constraints

$$\sum_{i=1}^n \frac{\partial \psi_{j_\nu}(x^{(0)})}{\partial x_i} \xi_i \leq -\eta_1 \quad (\nu = 1, \dots, \nu_1);$$

$$|\xi_i| \leq C \quad (i = 1, \dots, n). \quad (9)$$

Denote by u_1 the minimum of u subject to the constraints (9).

3. Suppose that the direction $\xi^{(1)}$ of steepest descent from the point $x^{(0)}$ has already been determined, and suppose moreover that $u_1 < -\delta_1$. Then motion in this direction, i.e. increasing t in the formula $x = x^{(0)} + t\xi^{(1)}$, is possible only until we reach the value t_1 , equal to the smallest of the positive roots of the equations

$$\psi_j(x^{(0)} + t\xi^{(1)}) = 0 \quad (j = 1, \dots, m). \quad (10)$$

As the new approximation we take the point $x^{(1)} = x^{(0)} + t_1\xi^{(1)}$; we regard it as the initial point, set $\delta_2 = \delta_1$ and $\eta_2 = \eta_1$, determine the descent direction $\xi^{(2)}$ from the point $x^{(1)}$, and continue the process until, for some $x^{(k-1)}$ and corresponding δ_k and η_k , the corresponding linear programming problem (8)–(9) leads to a minimum u_k of the function (8) such that $u_k \geq -\delta_k$. Suppose in this case, for example,

$$\psi_{j_\mu}(x^{(k-1)}) = 0 \quad (\mu = 1, \dots, \mu_1); \quad \psi_j(x^{(k-1)}) < 0 \quad (j \neq j_1, \dots, j_{\mu_1}).$$

Then, if $u_k < 0$, we set $\delta_{k+1} = \delta_k/2$, $\eta_{k+1} = \eta_k/2$, and continue the process.

In the case where $u_k \geq 0$, we minimize the function (8) subject to the constraints

$$\sum_{i=1}^n \frac{\partial \psi_{j_\mu}(x^{(k-1)})}{\partial x_i} \xi_i \leq 0 \quad (\mu = 1, \dots, \mu_1); \quad |\xi_i| \leq C \quad (i = 1, \dots, n),$$

and if $\min u = u'_k = 0$, then $x^{(k-1)}$ is a solution of problem (6)–(7) and the process is finished; but if $u'_k < 0$, then we change the values of the parameters δ and η . We set $\delta_{k+1} = \delta_k/2$. To determine η_{k+1} , we maximize the function $v = \xi$ subject to the constraints

$$\sum_{i=1}^n \frac{\partial \psi_{j_\mu}(x^{(k-1)})}{\partial x_i} \xi_i \leq -\xi \quad (\mu = 1, \dots, \mu_1);$$

$$\sum_{i=1}^n p_i \xi_i - \frac{u'_k}{2} \leq 0; \quad |\xi_i| \leq C \quad (i = 1, \dots, n).$$

and set

$$\eta_{k+1} = \min \left\{ \frac{\eta_k}{2}, \max v \right\}.$$

We continue the process, taking δ_{k+1} instead of δ_k and η_{k+1} instead of η_k .

4. Let us dwell on the question of convergence of the process described. Suppose that a sequence of approximations $\{x^{(k)}\}$ has been obtained. The corresponding sequence $\{z^{(k)}\}$ of values of the form (6) is decreasing. Let \tilde{x} be a limit point of the sequence $\{x^{(k)}\}$, and let the subsequence $\{x^{(k_i)}\}$ converge to \tilde{x} . We shall show that \tilde{x} is a solution of problem (6)–(7).

First note that if $\lim_{k \rightarrow \infty} z^{(k)} \neq -\infty$, then $\lim_{k \rightarrow \infty} \delta_k = 0$, and consequently $\lim_{k \rightarrow \infty} \eta_k = 0$. Indeed, if we had $\lim_{k \rightarrow \infty} \delta_k = \delta > 0$, then for all sufficiently large k we would have $|u_k| > \delta$, but then necessarily $z^{(k)} \rightarrow -\infty$.

Now suppose that \tilde{x} is not a solution of problem (6)–(7). Then the solution of the corresponding problem (8)–(9) for determining the direction of descent from the point \tilde{x} with $\delta = 0$ and $\eta = 0$ will give $\min u = \tilde{u} < 0$. Let

$$\psi_{j_\lambda}(\tilde{x}) = 0 \quad (\lambda = 1, \dots, \lambda_1); \quad \psi_j(\tilde{x}) < 0 \quad (j \neq j_1, \dots, j_{\lambda_1}).$$

By virtue of the rule for computing η , we obtain $\tilde{\eta} > 0$. Owing to the continuous differentiability of the functions $\psi_1(x), \dots, \psi_m(x)$ and to the tendency of δ_{k_i} and η_{k_i} to zero, we obtain that $u_{k_i} < \tilde{u}/4 < 0$ for all sufficiently large k_i , but this contradicts the fact that $|u_{k_i}| \leq \delta_{k_i} \rightarrow 0$.

5. The second algorithm given below differs from the preceding one in the choice of the direction of descent. Now, in order to simplify the process, we abandon the attempt at each step to decrease the function (6) maximally and to choose the parameter η in the best possible way. Instead, we choose the direction of descent by solving both problems, as it were, in an averaged manner. After choosing δ_1 , we take as the direction of descent $\xi^{(1)}$ the solution of the linear-programming problem consisting in minimizing the form

$$w = \xi$$

under the constraints

$$\sum_{i=1}^n \frac{\partial \psi_{j_\nu}(\tilde{x}^{(0)})}{\partial x_i} \xi_i \leq \xi \quad (\nu = 1, \dots, \nu_1);$$

$$\sum_{i=1}^n p_i^1 \xi_i \leq \xi; \quad |\xi_i| \leq C \quad (i = 1, \dots, n).$$

If $\min w = \xi_1 < -\delta_1$, then we move in the direction $\xi^{(1)}$, as in item 3, obtain the point $x^{(1)}$, which we regard as the initial one, set $\delta_2 = \delta_1$, and so on. If, however, at some step we arrive at $\min w = \xi_k \geq -\delta_k$, then we change the value of the parameter δ_k .

In the case $\xi_k < 0$, set $\delta_{k+1} = \delta_k/2$ and continue the process with the initial $x^{(k+1)}$ and δ_{k+1} .

In the case $\xi_k = 0$ and

$$\psi_{j_l}(x^{(k-1)}) = 0 \quad (l = 1, \dots, l_1); \quad \psi_j(x^{(k-1)}) < 0 \quad (j \neq j_1, \dots, j_{l_1})$$

we minimize $w' = \xi$ subject to the constraints

$$\sum_{i=1}^n \frac{\partial \psi_{j_l}(x^{(k-1)})}{\partial x_i} \xi_i \leq \xi \quad (l = 1, \dots, l_1);$$

$$\sum_{i=1}^n p_i \xi_i \leq \xi; \quad |\xi_i| \leq C \quad (i = 1, \dots, n),$$

and if $\min w' < 0$, then again we set $\delta_{k+1} = \delta_k/2$ and continue the process; while if $\min w' = 0$, then $x^{(k-1)}$ is the solution of the problem.

Kyiv State Pedagogical
Institute named after A. M. Gorky

Ukrainian Road-Transport
Scientific Research Institute

Received
26 VI 1963

CITED LITERATURE

- ¹ S. I. Zukhovitskii, R. A. Polyak, M. E. Primak, *DAN*, **151**, No. 1 (1963).
- ² J. B. Rosen, *J. Soc. Industr. and Appl. Math.*, **8**, No. 1, 181 (1960); **9**, No. 4, 514 (1961).
- ³ K. J. Arrow, L. Hurwicz, H. Uzawa, *Studies in Linear and Nonlinear Programming*, IL, 1962.
- ⁴ V. V. Ivanov, *DAN*, **143**, No. 4, 775 (1962).
- ⁵ G. Zoutendijk, *Methods of Feasible Directions*, IL, 1963.

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

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