Reports of the Academy of Sciences of the USSR

1964

SovietRxiv

Abstract

Full Text

Reports of the Academy of Sciences of the USSR

1964. Vol. 159, No. 3

MATHEMATICS

S. I. Zukhovitskii, M. E. Primak

AN ALGORITHM FOR SOLVING THE PROBLEM OF CHEBYSHEV APPROXIMATION IN HILBERT SPACE

(Presented by Academician N. N. Bogolyubov, 28 V 1964)

1. By the problem of Chebyshev approximation in a Hilbert space we shall, following ($^{1-3}$), understand the following.

Let an operator-function $A(q)$ be defined on a compact set Q, which for each $q \in Q$ is a closed linear operator acting from the Hilbert space H_1 into the Hilbert space H_2, with a domain of definition D common to all $A(q)$ and dense in H_1; moreover, for each fixed $x \in D$ the function $A(q)x$, with values in H_2, is continuous on Q. Let, further, $f(q)$ be a function continuous on Q with values in H_2. One seeks a vector $x^* \in D$ such that

$$\max_{q\in Q} \|A(q)x^* - f(q)\| = \inf_{x\in D} \max_{q\in Q} \|A(q)x - f(q)\|.$$

It is clear that, in order to construct a numerical solution of the problem, one may consider it not on the whole compact set Q, but on an ε-net $\{q_1, \dots, q_n\}$ of this compact set, for sufficiently small $\varepsilon > 0$. Then, denoting

$$A(q_i) = A_i, \qquad f(q_i) = f_i \qquad (i = 1, \dots, n),$$

we arrive at the problem of Chebyshev approximation in a Hilbert space of the system

$$A_i x - f_i \qquad (i = 1, \dots, n), \tag{1}$$

i.e., at the problem of finding such a vector $x^* \in D$ (the Chebyshev vector of the system (1)) for which

$$\max_i \|A_i x^* - f_i\| = \inf_{x\in D} \max_i \|A_i x - f_i\|. \tag{2}$$

If we denote by R the subspace of those vectors $x \in D$ for which $A_i x = 0$ for all $i = 1, \dots, n$, and by S the orthogonal complement to R in H_1, then for every vector $x \in D$ we shall have the representation

$$x = x_R + x_S \qquad (x_R \in R;\ x_S \in S),$$

$$\inf_{x\in D} \max_i \|A_i x - f_i\| = \inf_{x_S \in D\cap S} \max_i \|A_i x_S - f_i\|,$$

so that one may assume $R = 0,\ S = H_1$.

As is known (^{3}), in order that for each system of vectors $f_1, \dots, f_n$ from H_2 there exist a vector $x^* \in D$ (a Chebyshev vector) for which (2) holds, it is necessary and sufficient that the operators $A_1, \dots, A_n$ satisfy the condition

$$\max_i \|A_i x\| \geq m\|x\| \quad \text{for all } x \in D, \tag{3}$$

where $m > 0$ is a constant.

In the case when $n = 1$, the operator A is bounded and the equation $Ax = f$ has an exact solution, a descent algorithm for finding this solution was constructed in (^{4})*. Other algorithms for solving a similar problem are given or may be obtained from the methods described in ($^{5-7}$).

* There the case is also considered where the operator A is unbounded but satisfies certain special conditions.

In the proposed paper we give an algorithm for solving problem (1)–(2) for the case when the operators A_i $(i = 1, \dots, n)$ are bounded.

2. As the initial approximation take an arbitrary vector $x^{(0)} \in H_1$ and an arbitrary sufficiently small $\delta_1 > 0$. Let

$$\max_i \|A_i x^{(0)} - f_i\|^2 = \|A_{i_1} x^{(0)} - f_{i_1}\|^2;$$

$$\|A_{i_\nu} x^{(0)} - f_{i_\nu}\|^2 > \|A_{i_1} x^{(0)} - f_{i_1}\|^2 - \delta_1 \quad (\nu = 1, \dots, \nu_1),$$

$$\|A_i x^{(0)} - f_i\|^2 \leq \|A_{i_1} x^{(0)} - f_{i_1}\|^2 - \delta_1 \quad (i \neq i_1, \dots, i_{\nu_1}).$$

The set of indices $\{i_1, i_2, \dots, i_{\nu_1}\}$ for which the deviations $\|A_i x^{(0)} - f_i\|^2$ differ from the maximal ones by less than δ_1, i.e., are "almost maximal," will be denoted by $I(x^{(0)}, \delta_1)$.

To improve the approximation we shall move (descend) from the point $x^{(0)}$ in some direction $g^{(1)} \in H_1$, i.e., we shall increase $t > 0$ in the expression $x^{(0)} + tg^{(1)}$, choosing the descent direction $g^{(1)}$ so that all the "almost maximal" deviations decrease:

$$\|A_i x^{(0)} + tg^{(1)} - f_i\|^2, \qquad i \in I(x^{(0)}, \delta_1),$$

i.e., so that the derivatives be negative,

$$\frac{1}{2}\frac{d}{dt}\|A_i(x^{(0)} + tg^{(1)}) - f_i\|^2\Big|_{t=0} = (A_i g^{(1)},\, A_i x^{(0)} - f_i) =$$

$$= (g^{(1)},\, A_i^*(A_i x^{(0)} - f_i)) = (g^{(1)},\, g_i(x^{(0)})), \qquad i \in I(x^{(0)}, \delta_1),$$

and so that this decrease (descent) be as steep as possible. Taking into account the negativity of the derivatives, this means that the normalized direction $g^{(1)}$ must satisfy the relation

$$\max_{i \in I(x^{(0)}, \delta_1)} (g^{(1)}, g_i(x^{(0)})) = \min_{g \in H_1} \max_{i \in I(x^{(0)}, \delta_1)} (g, g_i(x^{(0)})).$$

Since the orthogonal complement to the subspace spanned by the vectors $g_i(x^{(0)})$, $i \in I(x^{(0)}, \delta_1)$, evidently plays no role in the problem under consideration, it may be assumed that

$$g = \sum_{j \in I(x^{(0)}, \delta_1)} \xi_j g_j(x^{(0)}).$$

Then the problem of finding the direction of steepest descent from $x^{(0)}$ is reduced to finding

$$\min_{\xi} \max_{i \in I(x^{(0)}, \delta_1)} \sum_{j \in I(x^{(0)}, \delta_1)} (g_j, g_i)\xi_j = \min_{\xi} \max_{i \in I(x^{(0)}, \delta_1)} \sum_{j \in I(x^{(0)}, \delta_1)} a_{ij}\xi_j \tag{4}$$

under the normalization, for example, $|\xi_j| \le 1,\ j \in I(x^{(0)}, \delta_1)$.

The last problem is the usual linear programming problem: minimize

$$u = \xi_{n+1} \tag{5}$$

subject to the constraints

$$\sum_{j\in I(x^{(0)},\delta_1)} a_{ij}\xi_j \le \xi_{n+1}, \qquad i\in I(x^{(0)},\delta_1), \qquad |\xi_j|\le 1, \quad j\in I(x^{(0)},\delta_1). \quad (6)$$

With the normalization

$$\|g\|^2 = \left\| \sum_{j\in I(x^{(0)},\delta_1)} \xi_j g_j(x^{(0)}) \right\|^2 \le 1,$$

the determination of the descent direction is reduced to solving the problem of minimizing a linear form

$$u = \xi_{n+1} \qquad (5\text{'})$$

under the constraints

$$\sum_{j\in I(x^{(0)},\delta_1)} a_{ij}\xi_j \le \xi_{n+1}, \qquad i\in I(x^{(0)},\delta_1), \qquad \left\| \sum_{j\in I(x^{(0)},\delta_1)} \xi_j g_j(x^{(0)}) \right\|^2 \le 1. \quad (6\text{'})$$

Denote by u_1 the minimum u under the constraints $(6')$, and suppose that $u_1 < -\delta_1$. For definiteness choose one of the normalizations, for example the second one.

Having thus constructed a descent direction from $x^{(0)}$, we proceed to determine the approximation step t, i.e., to find the boundary of the admissible increase of t in the formula $x = x^{(0)} + tg^{(1)}$. To this end, let us first observe that problem (1)–(2) is equivalent to the problem of minimizing the function $w = \xi$ under the constraints

$$\|A_i x - f_i\|^2 - \xi \le 0 \qquad (i = 1, \dots, n). \qquad (7)$$

In the descent direction $(g^{(1)}; \eta_1)$ from the point $(x^{(0)}; \mu_1^2)$ (where $\mu_1 = \|A_{i_1} x^{(0)} - f_{i_1}\|$) we use the vector $g^{(1)}$ found above, and determine η_1 from the condition that, along the direction $(g^{(1)}; \eta_1)$, the maximum of the derivatives of the functions $w = \xi$, $\|A_i x - f_i\|^2 - \xi$, $i \in I(x^{(0)}, \delta_1)$, at the point $(x^{(0)}; \mu_1^2)$ be negative and as small as possible, i.e., that the relation

$$\max\left\{\eta_1, \max_{i\in I(x^{(0)},\delta_1)} \left[2(A_i g^{(1)}, A_i x^{(0)} - f_i) - \eta_1\right]\right\} =$$

$$= \min_{\eta} \max \left\{ \eta, \max_{i \in I(x^{(0)}, \delta_1)} \left[2(A_i g^{(1)}, A_i x^{(0)} - f_i) - \eta \right] \right\} =$$

$$= \min_{\eta} \max\{\eta, 2u_1 - \eta\}.$$

It is clear that $\eta_1 = u_1$. To determine the approximation step, we increase t in the expression $(x^{(0)} + tg^{(1)};\ \mu_1^2 + u_1 t) = (x; \xi)$ until we reach the boundary of the domain defined by the inequalities (7), i.e., the approximation step t_1 is the least positive root of the equations

$$\|A_i(x^{(0)} + tg^{(1)}) - f_i\|^2 - \mu_1^2 - u_1 t = 0 \qquad (i = 1, \dots, n).$$

As the new approximation we take $x^{(1)} = x^{(0)} + t_1 g^{(1)}$. We regard the obtained point $x^{(1)}$ as the initial one, set $\delta_2 = \delta_1$, determine a descent direction from the point $x^{(1)}$, and so on, until for some $x^{(k)}$ and the corresponding δ_{k+1} the corresponding problem of type (5′)–(6′) leads to $u_{k+1} \geq -\delta_{k+1}$. Then we set $\delta_{k+2} = \delta_{k+1}/2$, and in the case $u_{k+1} < 0$ find t_{k+1} and $x^{(k+1)}$ and continue the computations, taking $x^{(k+1)}$ as the initial point and δ_{k+2} as the parameter value. In the case $u_{k+1} = 0$, in determining the descent direction we leave among the constraints (6′) only those which correspond to strictly maximal deviations. If it again turns out that $\min u = u_{k+1} = 0$, then the point $x^{(k)}$ is a solution of problem (1)–(2). If, however, $u_{k+1} < 0$, then we find t_{k+1} and $x^{(k+1)}$ and continue the computations, taking $x^{(k+1)}$ as the initial point.

3. We outline a proof of convergence of the algorithm described. First note that $\delta_k \to \delta = 0$, since otherwise, for $\delta > 0$, we would have $u_k < -\delta$ beginning with some k, which is impossible, since then it would turn out that, on the one hand, $t_k \to 0$, and on the other hand, $t_k > t_0 > 0$. We now verify that

$$\inf_x \max_{1 \leq i \leq n} \|A_i x - f_i\| = \lim_{k \to \infty} \max_{1 \leq i \leq n} \|A_i x^{(k)} - f_i\| = \mu^*.$$

Indeed, let i_k be the number of the step at which the parameter δ changes. From the bounded sequence $\{x^{(i_k - 1)}\}$ one can extract a weakly convergent ···

a convergent subsequence. We may assume that $x^{(i_k-1)} \xrightarrow{\text{weakly}} \tilde{x}$ and that, at the steps i_k, the descent direction is described by one and the same set I of deviation indices. The set of indices

$$I_0 = \{i \mid \|A_i \tilde{x} - f_i\| = \mu^*\} \cap I$$

is nonempty.

Suppose that the vector $\tilde{x}$ is not a solution of problem (1)–(2), so that from the point $\tilde{x}$ there exists a descent direction $\tilde{g}$. Then, as $i_k \to \infty$ and for sufficiently small $\varepsilon > 0$, we obtain

$$(A_i(\tilde{x} - x^{(i_k-1)} + \varepsilon\tilde{g}),\, A_i x^{(i_k-1)} - f_i) \to \varepsilon(A_i\tilde{g},\, A_i x^{(0)} - f_i) \le 0, \qquad i \in I_0,$$

$$\begin{aligned}(A_i(\tilde{x} - x^{(i_k-1)} + \varepsilon\tilde{g}),\, A_i x^{(i_k-1)} - f_i) \to \|A_i\tilde{x} - f_i\|^2 - (\mu^*)^2 \\ + \varepsilon(A_i\tilde{g},\, A_i\tilde{x} - f_i) < 0, \qquad i \in I \setminus I_0,\end{aligned}$$

i.e., for sufficiently large k, the direction $\xi^{(k)} = \tilde{x} - x^{(i_k-1)} + \varepsilon\tilde{g}$ is a descent direction from the point $x^{(i_k-1)}$. Therefore a number $\alpha < 0$ will be found such that $u_{i_k} < \alpha$. But this is impossible, since from $\delta_{i_k} \to 0$ it follows that $u_{i_k} \to 0$.

Kyiv State Pedagogical Institute
named after A. M. Gorky

Ukrainian Road-Transport
Scientific Research Institute

Received
16 IV 1964

References

1. S. I. Zukhovitskii, "Some questions in the theory of Chebyshev approximations," Dissertation, Kiev, 1950.
2. S. I. Zukhovitskii, *Matem. sborn.*, 37 (79), 1, 3 (1955).
3. S. I. Zukhovitskii, G. I. Eskin, *Izv. AN SSSR*, 24, 93 (1960).
4. L. V. Kantorovich, *Uspekhi Mat. Nauk*, 3, issue 6, 89 (1948).
5. V. K. Ivanov, *Dokl. Akad. Nauk*, 142, No. 5, 997 (1962).
6. V. K. Ivanov, *Dokl. Akad. Nauk*, 145, No. 2, 269 (1962).
7. M. M. Lavrent' ev, *Dokl. Akad. Nauk*, 127, No. 1, 31 (1959).

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

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