

ON ONE METHOD OF REGULARIZATION OF AN EXTREMAL PROBLEM FOR A CONTINUOUS CONVEX FUNCTIONAL

MATHEMATICS

1969

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196901.41919>

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

Abstract

Full Text

UDC 518:62-50

MATHEMATICS

B. M. BUDAK, A. VIGNOLI, Yu. L. GAPONENKO

ON ONE METHOD OF REGULARIZATION OF AN EXTREMAL PROBLEM FOR A CON- TINUOUS CONVEX FUNCTIONAL

(Presented by Academician A. N. Tikhonov, 30 IV 1968)

1°. Statement of the problem. Idea of the method. Let $J(u)$ be a continuous convex functional defined on some bounded, closed convex set $U \subseteq H$; H is a Hilbert space. $J(u)$ attains the minimal value

$$J^* = \inf_{u \in U} J(u)$$

on the set $U^* \subseteq U$ of elements $u^* \in U$, where U is closed, convex, bounded, and nonempty. In the general case U^* may consist of more than one element. In this connection, a sequence $\{u_n\}$ minimizing $J(u)$ on U may, generally speaking, fail to converge strongly in the norm of H .

The first regularization method for constructing a strongly convergent minimizing sequence in such problems was proposed by A. N. Tikhonov ^(1, 2). In works ^(3, 4) a direct development of this method is given. In the present note, regularization is carried out by constructing a sequence of sets, in each of which its own minimizing sequence converges strongly to its limiting point; the sequence of these limiting points converges strongly to the solution of the problem u_{\min}^* —the element of U^* with least norm.

2°. Basic definitions and lemmas.

Lemma 1. *If the continuous convex functional $J(u)$ attains its least value in the closed convex bounded set U_α at the unique point u_α^* , and this point lies on the boundary of U_α and*

$$\|u_\alpha^*\| = \sup_{u \in U_\alpha} \|u\| = \alpha,$$

then an arbitrary sequence minimizing $J(u)$ on U_α converges strongly to u_α^ .*

Denote by $\mathcal{S}[0, \alpha]$ the set of elements $u \in H$ satisfying the condition $\|u\| \leq \alpha$, i.e., the ball with center at zero and radius α in the Hilbert space H . Without loss

of generality, it may be assumed that $u = 0$ is contained in the set U on which the functional $J(u)$, $u \in U$, under consideration is defined. This can always be achieved by a “parallel translation.” Let

$$\sup_{u \in U} \|u\| < R < +\infty.$$

Then U lies strictly inside the ball $S[0, R]$. Denote by U_α the convex set

$$U_\alpha = U \cap S[0, \alpha]$$

for any α , $0 \leq \alpha \leq R$. Put

$$f(\alpha) = \inf_{u \in U_\alpha} J(u), \quad 0 \leq \alpha \leq R.$$

Lemma 2. $f(\alpha)$ is a continuous monotonically decreasing (nonincreasing) function of α on the segment $0 \leq \alpha \leq R$.

Lemma 3. On the set U_α , for $0 \leq \alpha \leq \alpha^*$, the functional $J(u)$ attains the minimum

$$\inf_{u \in U_\alpha} J(u) = f(\alpha)$$

at the unique point $u_\alpha^* \in U_\alpha$, and u_α^* belongs to that part of the boundary of U_α which lies on the surface of the ball $S[0, \alpha]$; $\alpha^* = \|u_{\min}^*\|$.

3°. Regularization algorithm. We now consider a regularization algorithm for the problem that leads to a sequence minimizing $J(u)$ on U and converging strongly to u_{\min}^* . The proposed algorithm is essentially based on the established properties of the function $f(\alpha)$. We shall assume

that we have a method that makes it possible to solve the extremal problem for $J(u)$ on U_α , $0 \leq \alpha \leq R$, with arbitrarily high accuracy in the functional, i.e., we shall assume that for any $\varepsilon > 0$ our method makes it possible to construct an element $u_{\varepsilon, \alpha}$ such that:

$$1) \quad u_{\varepsilon, \alpha} \in U_\alpha; \quad 2) \quad \left| J(u_{\varepsilon, \alpha}) - \inf_{u \in U_\alpha} J(u) \right| = |J(u_{\varepsilon, \alpha}) - f(\alpha)| < \varepsilon.$$

We shall consider one such method in §4.

According to Lemma 3, on the set U_α , for $0 \leq \alpha \leq \alpha^*$, the functional $J(u)$ attains its least value at the unique point u_α^* , which lies on the boundary of U_α and on the boundary of the ball $[0, \alpha]$, i.e. $\|u_\alpha^*\| = \alpha$.

According to Lemma 2, every sequence minimizing $J(u)$ on U_α for $0 \leq \alpha \leq \alpha^*$ converges strongly to this boundary element $u_\alpha^* \in U_\alpha$, at which $J(u_\alpha^*) = f(\alpha) = \inf_{u \in U_\alpha} J(u)$. We define the regularization algorithm as follows.

Let us choose a numerical sequence $\{\varepsilon_n\}$, $n = 1, 2, \dots$; $\varepsilon_n > 0$; $\varepsilon_n \rightarrow 0$ as $n \rightarrow +\infty$. Introduce the notation:

$$\Delta_n(\alpha) = [J(u_{\varepsilon_n, \alpha}) - \varepsilon_n, J(u_{\varepsilon_n, \alpha}) + \varepsilon_n].$$

We have: $f(\alpha) \in \Delta_n(\alpha)$, and the length of the interval $\Delta_n(\alpha)$ tends to zero as $n \rightarrow +\infty$.

0. The zero step consists in putting $u_0 = 0$, $\alpha_0 = 0$. We shall call the zero step successful.

I. The first step. Two extremal problems for $J(u)$ on the sets $U_{R/2}$ and U_R are solved with accuracy up to ε_1 .

As a result of solving these two problems, two cases are possible:

$$\text{a) } \Delta_1(R/2) \cap \Delta_1(R) \neq \emptyset; \quad \text{b) } \Delta_1(R/2) \cap \Delta_1(R) = \emptyset.$$

In case a), we shall call step I unsuccessful and put $u_1 = u_0 = 0$, $\alpha_1 = \alpha_0 = 0$. In case b), we shall call step I successful and put $u_1 = u_{\varepsilon_1, R/2}$, $\alpha_1 = R/2$.

II. The second step. There are two possible cases:

- 1) The first step was unsuccessful. Then two extremal problems for $J(u)$ on the sets $U_{R/4}$ and U_R are solved with accuracy up to ε_2 . As a result of solving these two problems, two cases are possible:

$$\text{a) } \Delta_2(R/4) \cap \Delta_2(R) \neq \emptyset; \quad \text{b) } \Delta_2(R/4) \cap \Delta_2(R) = \emptyset.$$

In case a), we shall call step II unsuccessful and put: $u_2 = u_1$, $\alpha_2 = \alpha_1$. In case b), we shall call step II successful and put $u_2 = u_{\varepsilon_2, R/4}$, $\alpha_2 = R/4$.

- 2) The first step was successful. Then two extremal problems for $J(u)$ on $U_{3/4R}$ and U_R are solved with accuracy up to ε_2 . As a result of solving these two problems, two cases are possible:

$$\text{a) } \Delta_2(3/4R) \cap \Delta_2(R) \neq \emptyset; \quad \text{b) } \Delta_2(3/4R) \cap \Delta_2(R) = \emptyset.$$

In case a), we shall call step II unsuccessful and put $u_2 = u_1$, $\alpha_2 = \alpha_1$. In case b), we shall call step II successful and put $u_2 = u_{\varepsilon_2, 3/4R}$, $\alpha_2 = 3/4R$.

The n -th step. Two extremal problems for $J(u)$ on the sets $U_{\alpha_{n-1} + (R - \alpha_{n-1})/2^{k+1}}$, U_R are solved with accuracy up to ε_n , where k is the number of consecutive unsuccessful steps immediately preceding the n -th step.

As a result of solving these two problems, two cases are possible:

- a) $\Delta_n(\alpha_{n-1} + (R - \alpha_{n-1})/2^{k+1}) \cap \Delta_n(R) \neq \emptyset$;
- b) $\Delta_n(\alpha_{n-1} + (R - \alpha_{n-1})/2^{k+1}) \cap \Delta_n(R) = \emptyset$.

In case a), we shall call the n -th step unsuccessful and put $u_n = u_{n-1}$, $\alpha_n = \alpha_{n-1}$. In case b), we shall call the n -th step successful and put:

$$u_n = u_{\varepsilon_n \alpha_n}, \quad \alpha_n = \alpha_{n-1} + (R - \alpha_{n-1})/2^{k+1}.$$

As a result of the indicated algorithm, we obtain two sequences $\{a_n\}$, $\{u_n\}$. Denote by u_{\min}^* an element satisfying the following conditions:

- 1) $u_{\min}^* \in U^*$,
- 2) $\|u_{\min}^*\| = \inf_{u \in U^*} \|u\|$.

Such an element exists and is unique, because the set is closed, convex, bounded, and nonempty. Note that $f(a_n) \geq f(a^*)$, and $J(u_n) - f(a_n) \leq \varepsilon_n$ for all $n = 1, 2, \dots$

Theorem 1. *If $\{a_n\}$ and $\{u_n\}$ are constructed by means of the indicated algorithm, then as $n \rightarrow +\infty$:*

- 1) $a_n \rightarrow a^* - 0$, not decreasing;
- 2) $u_n \rightarrow u_{\min}^*$ strongly in the metric of H .

4°. On the use of the conditional-gradient method in the general algorithm. As a concrete method for solving the extremal problem at each step of the general algorithm, any method may be taken that permits one to solve the extremal problem with prescribed accuracy (with respect to the functional). In particular, if the functional is differentiable and its gradient satisfies certain additional requirements (see, for example, (5, 6)), then the conditional-gradient method can be proposed. This method gives the estimate $J(v) - J^* \leq (I(v), v - \bar{v}^\alpha)$, where

$$J^* = \inf_{v \in U^\alpha} J(v),$$

$(I(v), v - \bar{v}^\alpha)$ is a functional linear (with respect to the difference $v - \bar{v}^\alpha$ standing on the right) representing the gradient of the functional $J(v)$ at the point v ; \bar{v}^α is a solution of the linear extremal problem for $(I(v), w)$ on U_α , $w \in U_\alpha$. It is clear that for the effective application of the conditional-gradient method one must be able to find \bar{v}^α , i.e., be able to solve the extremal problem for the linear functional $(I(v), w)$ with $w \in U_\alpha$. Let us show how \bar{v}^α can be found in the following practically important case:

$$H = L_2^r[0, T]; \quad v(t) = (v^1(t), \dots, v^r(t)); \quad v^i(t) \in L_2[0, T],$$

$$i = 1, \dots, r; \quad U : |v^i(t)| \leq 1 \text{ for all } t \in [0, T] \text{ and all } i = 1, 2, \dots, r.$$

In the present case we have the following problem: find $\bar{v}^\alpha \in U_\alpha$ such that

$$(I(v), \bar{v}^\alpha) = \min_{w \in U_\alpha} (I(v), w),$$

where

$$U_\alpha = U \cap [0, \alpha]; \quad U : |w^i(t)| \leq 1, \quad t \in [0, T], \quad i = 1, 2, \dots, r;$$

$$[0, \alpha] : \|w\|_{L_2}^2 = \int_0^T w^2(t) dt \leq \alpha^2.$$

In view of the linearity, we represent the functional $(I(v), w)$ in the form

$$(I(v), w) = \int_0^T \psi(t)w(t) dt,$$

where $\psi(t)$ is a known function.

If, for example, the functional $J(u)$ has the form

$$J(u) = \int_0^T g(x(t), u(t), t) dt + \Phi(x(T)).$$

where $\dot{x}(t) = f(x(t), u(t), t)$ for $0 < t < T$, $x(0) = x_0$, then

$$\psi(t) = g'_u - (f'_u)^* \psi,$$

where the function $\psi(t)$ is found as the solution of the adjoint Cauchy problem

$$\dot{\psi}(t) = -(f'_x)^* \psi + g'_x \quad \text{for } 0 < t < T, \quad \psi(T) = -\Phi'(x(T))$$

(see, for example, (6)). Without loss of generality, one may assume that $\|\psi\| = \alpha$.

Introduce the notation

$$w_1(t) = \text{sign}[-\psi(t)], \quad w_2(t) = [-\psi(t)].$$

Define the element \bar{w}^α as follows: 1) $\bar{w}^\alpha = w_1$, if $\|w_1(t)\| \leq \alpha$; 2) $\bar{w}^\alpha = w_2$, if $|w^i(t)| \leq 1$, $t \in [0, T]$, $i = 1, 2, \dots, r$; if both conditions 1) and 2) are simultaneously fulfilled, then as the element \bar{w}^α one may take either w_1 or w_2 ; 3) if neither of the two stated conditions is fulfilled, then we set

$$\bar{w}^\alpha(t) = [-\psi_\gamma(t)],$$

where $\psi_\gamma(t)$ is determined in the following way. Consider the one-parameter family of functions $\{\psi_\gamma(t)\}$, $\gamma \geq 1$, constructed according to the following rule:

$$\psi_\gamma(t) = \gamma\psi(t).$$

Form the family of “cut-off” functions:

$$\bar{\psi}_\gamma(t) = \begin{cases} \text{sign } \psi_\gamma(t), & \text{for } t \in \mu_1 = \{t : |\psi_\gamma(t)| > 1\}, \\ \psi_\gamma(t), & \text{for } t \in \mu_2 = \{t : |\psi_\gamma(t)| \leq 1\}. \end{cases}$$

Note that $[-\bar{\psi}_\gamma(t)] \rightarrow w_1(t)$ as $\gamma \rightarrow +\infty$.

Consequently, $\|\bar{\psi}_\gamma(t)\| \rightarrow \|w_1(t)\| > \alpha$ as $\gamma \rightarrow +\infty$, and moreover $\|\bar{\psi}_{\gamma_2}\| > \|\bar{\psi}_{\gamma_1}\|$ for $\gamma_2 > \gamma_1$.

On the other hand, $\|\bar{\psi}_{\gamma=1}\| < \|\psi\| = \alpha$. Therefore there exists, and moreover is unique, a γ such that $\|\bar{\psi}_\gamma(t)\| = \alpha$. It is precisely this value of the parameter γ that is used to determine the function \bar{w}^α in the third case.

Theorem 2. *If the function $\bar{w}^\alpha(t)$ is defined in the manner indicated above, then the following holds:*

$$(I(v), \bar{w}^\alpha) = \min_{w \in U_\alpha} (I(v), w).$$

Moscow State University
named after M. V. Lomonosov

Received
26 IV 1968

References

- ¹ A. N. Tikhonov, DAN, 162, No. 4, 763 (1965).
- ² A. N. Tikhonov, Zhurn. vychisl. matem. i matem. fiz., 6, No. 4, 631 (1966).
- ³ E. S. Levitin, B. T. Polyak, DAN, 168, No. 5, 997 (1966).
- ⁴ V. A. Morozov, *Collection of Papers of the Computing Center of Moscow State University*, Computational Methods and Programming, issue 8, 1967, p. 141.
- ⁵ V. F. Demyanov, A. M. Rubinov, Vestn. LGU, issue 19, 5 (1964).

⁶ E. S. Levitin, B. T. Polyak, Zhurn. vychisl. matem. i matem. fiz., 6, No. 5, 787 (1966).

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

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