N. I. Polsky

Mathematics

1962

SovietRxiv

Abstract

Full Text

N. I. Polsky

Projection Methods in Applied Mathematics

(Presented by Academician A. Yu. Ishlinsky, 18 IX 1961)

Mathematics

1°. In the present paper we study approximate methods called in ([1]) projection methods. These include the methods of Ritz, Galerkin, moments, least squares, etc. All these methods are often classed as variational, reducing the study of their convergence to the search for the minimum of some functional. However, the convergence of all these methods is, in essence, equivalent to the fulfillment of a certain simple geometric condition, established already in ([2,3]) and there called condition (A).

Below we give a completely elementary proof of the indicated fact and present classes of operators for which condition (A) is fulfilled. At the same time, for a Hilbert space condition (A) is formulated with the aid of the decomposition of subspaces established in ([4]). We note that in all works devoted to the justification of methods of the Ritz–Galerkin type, the matter is ultimately reduced to establishing convergence of the process for an equation of the second kind with a completely continuous operator in one or another space. As is shown below, the geometric condition (A), and hence the convergence of projection methods, holds for a broader class of problems that are not reducible to problems with a completely continuous operator.

2°. We first give the following definition:

Definition 1. A sequence of closed subspaces M_n of a Hilbert space H, together with the sequence of corresponding operators P_n projecting H onto these subspaces, will be called **projection complete** if, for every $f \in H$,

$$|f - P_n f| \to 0 \quad \text{as } n \to \infty.$$

Consider Hilbert spaces H_1, H_2, an arbitrary linear operator L (generally speaking, unbounded), whose domain of definition is dense in H_1 and whose range lies in H_2, and the equation $Lx = f,\ f \in H_2$.

The **projection method** for solving this equation consists in the following. Let R_n be a sequence of such finite-dimensional subspaces of the domain of definition of the operator L that the sequence $L_n = LR_n$ is projection complete in H_2.* Let M_n be another projection-complete sequence of subspaces of H_2, and let $\dim M_n = \dim L_n$. Denote by P_n and Π_n the operators of orthogonal projection onto M_n and L_n, respectively. One seeks an approximate solution $x_n \in R_n$ from the condition

$$P_n L x_n = P_n f.$$

Denote $Lx = y \in H_2,\ Lx_n = y_n \in L_n \subset H_2$. Then the original equation takes the form $y = f$, and the approximate one $P_n y_n = P_n f$. We shall investigate the question of convergence of y_n to y. We shall consider the operator P_n **only** on L_n. It maps L_n into M_n. Obviously, $P_n L_n$ coincides with all of M_n if and only if in L_n there is no element $y_n \neq 0$ that would be mapped by the operator P_n to zero. This condition means that the number

$$\tau_n = \min_{\substack{y_n \in L_n; \\ |y_n|=1}} |P_n y_n| \neq 0.$$

* All projection operators considered below project orthogonally. However, the scheme presented is easily generalized to the case where the original spaces are Banach spaces, and the projection operators of each sequence are subject only to the condition that their norms are bounded in the aggregate. We also note that in applications the subspaces R_n are usually chosen nested, i.e. $R_n \subset R_{n+1}$ $(n = 1, 2, ...)$.

If the condition $\tau_n \neq 0$ is satisfied, then the approximate equation $P_n y_n = P_n f$ is uniquely solvable and $y_n = P_n^{-1} P_n f \in L_n$. In this case $\|P_n^{-1}\| = \tau_n^{-1}$.

For finite-dimensional subspaces L_n and M_n of the same dimension, the aperture θ_n can be defined as

$$\theta_n = \max_{\substack{y_n \in L_n:\ |y_n|=1; \\ z \in M_n}} |y_n - z|.$$

Obviously,

$$\theta_n^2 + \tau_n^2 = 1.$$

Thus the lemma is proved. *For the unique solvability of the approximate equation* $P_n y_n = P_n f$ *for any* $f \in H_2$, *it is necessary and sufficient that the condition* $\tau_n > 0$ *(or* $\theta_n < 1$*) be satisfied.*

Let us now denote $\tau = \underline{\lim}_{n\to\infty} \tau_n;\ \theta = \overline{\lim}_{n\to\infty} \theta_n$. Obviously, $\theta^2 + \tau^2 = 1$.

Theorem 1. *In order that the sequence* $y_n = Lx_n$ *converge strongly to* $y = Lx$ *for every* $f \in H_2$, *it is necessary and sufficient that condition (A) be satisfied:* $\tau > 0$ *(or* $\theta < 1$*). In this case the estimate*

$$|Lx - Lx_n| \leq \left(1 + \frac{1}{\tau_n}\right) |f - \Pi_n f|$$

holds.

Proof. By virtue of the lemma, beginning with some n, the approximate equations are uniquely solvable. Since $y_n - \Pi_n f \in L_n$, we have

$$|y_n - \Pi_n f| \leq \frac{1}{\tau_n}|P_n(y_n - \Pi_n f)| = \frac{1}{\tau_n}|P_n f - P_n \Pi_n f| \leq \frac{1}{\tau_n}|f - \Pi_n f|.$$

Consequently,

$$|Lx_n - Lx| = |y_n - f| \leq |y_n - \Pi_n f| + |\Pi_n f - f| \leq \left(1 + \frac{1}{\tau_n}\right)|f - \Pi_n f|.$$

Sufficiency is proved. Necessity means that when $\tau = 0$ one can construct an element $f \in H_2$ such that the sequence of approximate solutions y_n grows without bound in norm and, consequently, does not tend to y. This follows from the fact that when $\tau = 0$ there is a subsequence $\tau_{n_k} \to 0$, i.e., $\|P_{n_k}^{-1}\| \to \infty$. The construction of the required element f can be carried out similarly to how it was done in ([5]). The necessity of condition (A), in a somewhat different form, was first shown in ([2,5]), and in the formulation given here in ([6]).

Remark. If there exists a bounded operator L^{-1}, then convergence of x_n to x is also obtained, with the estimate

$$|x - x_n| \leq \|L^{-1}\|\left(1 + \frac{1}{\tau_n}\right)|f - \Pi_n f|.$$

3°. In ([1]) it is shown that from Theorem 1 there follow the theorems on convergence of projection methods established earlier in various works. In particular, from it there immediately follows, in the most general form, the result ([7]) on convergence of the least-squares method for any operator L and for any choice of the sequence R_n (if, of course, the sequence L_n is projection-complete), since in this case $M_n = L_n$ and $\tau_n = 1$.

Methods in which $H_1 = H_2$ and $M_n = R_n$ (the Ritz and Galerkin methods) are very widespread. Of considerable interest is the determination of those operators for which, also here, an arbitrary sequence R_n leads to a convergent process.

Definition 2. We shall call an operator L **proper** if any two projection-complete sequences of subspaces R_n and $L_n = LR_n$ satisfy the condition $\tau > 0$.

Theorem 2. *If -1 does not belong to the spectrum of the completely continuous operator A, then the operator $L = E + A$ is proper (E is the identity operator).*

Proof. Consider an arbitrary element $x \in R_n = M_n$ and $|x| = 1$. By the condition of the theorem, there is a number $\delta > 0$ such that $|Lx| > \delta$, i.e., for every element $y = Lx = x + Ax \in L_n$ one has $|y| > \delta$.

Represent the element Ax in the form $Ax = P_{nAx} + z$, where $z \perp R_n$. Because of the complete continuity of the operator A, for sufficiently large n we shall have $|z| < \varepsilon < \delta$. Moreover,

$$y = x + P_{nAx} + z = P_{ny} + z,$$

and, consequently,

$$|y|^2 = |P_{ny}|^2 + |z|^2 > \delta^2.$$

Hence $|P_{ny}|^2 > \delta^2 - \varepsilon^2$, and in view of $|y| \leq \|L\|$ we obtain the inequality

$$|P_{ny}| \geq \frac{\sqrt{\delta^2 - \varepsilon^2}}{\|L\|}\,|y|.$$

Thus,

$$\tau \geq \frac{\sqrt{\delta^2 - \varepsilon^2}}{\|L\|}.$$

Remark. A stronger fact has been proved here. Namely, only the fact was used that the operator A can be approximated in norm by finite-dimensional operators with accuracy up to some $\varepsilon < \delta$, and not with arbitrary accuracy (the latter characterizes completely continuous operators in separable spaces).

Let us now give one more condition for regularity.

In accordance with (8), the set W of complex numbers assumed by the form (Lx, x) on the unit sphere $|x| = 1$ is called the **set of numerical values** of the operator L. The set W is convex and contains the spectrum of the operator L (8).

Theorem 3. *If the set W of numerical values of the bounded operator L and the zero of the complex plane lie on different sides of some straight line, then the operator L is regular.*

Proof. By the condition of the theorem,

$$|(Lx, x)| \geq \delta > 0$$

for $|x| = 1$. Let $x \in R_n = M_n$, $y = Lx \in L_n$. Then

$$|P_{ny}| = \max_{z \in M_n;\ |z|=1} |(y, z)| \geq |(y, x)| = |(Lx, x)| \geq \delta > 0.$$

Since $|y| \leq \|L\|$, it follows that

$$|P_{ny}| \geq \frac{\delta}{\|L\|}|y|.$$

Thus,

$$\tau \geq \frac{\delta}{\|L\|} > 0.$$

This theorem is a generalization of a well-known fact for a positive-definite operator. Here the operator L may be non-Hermitian. In the case when W is the minimal convex hull of the spectrum, the condition of Theorem 3 means that the spectrum of the operator L can be separated by a straight line from the zero of the complex plane.

4°. As noted above, in all works devoted to the study of convergence of projection methods, Theorem 2 is essentially used. In doing so, the following procedure is applied. One considers the uniquely solvable equation

$$Lx = f$$

(the operator L may be unbounded). Then a new space H_0 is introduced in such a way that the approximate equations

$$P_{nLx}n = P_{nf}$$

coincide with approximate equations in H_0 for some other initial equation of the form

$$x + Ax = g,$$

where the operator A is already completely continuous in H_0. Such a device is used in $(9, 10)$ and in other works. The indicated procedure is most simply carried out if $L = S + T$, where S is a positive-definite operator with domain of definition dense in H. In this case the new scalar product is introduced by the formula

$$[u, v] = (Su, v).$$

If, in particular, S is bounded, then the new norm is equivalent to the old one, and the following is easily obtained.

Theorem 4. *If zero does not belong to the spectrum of the operator $L = S + T$, where S is a bounded positive-definite operator and T is a completely continuous operator, then the operator L is regular.*

This theorem is known. We note only that, apparently, it remains valid if, instead of positive definiteness of the operator S, one requires that its set of numerical values be separable by a straight line from zero of the complex plane. This fact, however, has not been proved. Its proof would make it possible to unite Theorems 3 and 4.

$5°$. The introduction of a new space makes it possible to study the convergence of projection methods for unbounded operators. In this case one often obtains not only the convergence of x_n to x, but also stronger results. If, for example, $L = S + T$, where S is a positive-definite operator, then under the condition of complete continuity of the operator $\sqrt{S^{-1}T\sqrt{S^{-1}}}$ in the original space one proves the convergence of $\sqrt{Sx_n}$ to $\sqrt{Sx}$. This is obtained under the condition that $R_n = M_n$.

If, moreover, the choice of the subspaces M_n is arranged in a suitable way, then, in accordance with Theorem 1, one can obtain the convergence of Sx_n to Sx or of Lx_n to Lx.

Let M be such a linear (generally speaking, unbounded) operator that the intersection of the domains of definition of the operators L and M is dense in H. Below, the subspaces R_n are chosen in this intersection. Let now $M_n = MR_n$.

Definition 3. We shall say that the operators L and M form a **regular pair** $[L, M]$ if, for any choice of subspaces R_n ensuring the projection completeness of the sequences $L_n = LR_n$ and $M_n = MR_n$, condition (A), i.e. $\tau > 0$, is satisfied.

Obviously, the pair $[L, L]$ is regular. This also guarantees the convergence of the least-squares method. Regularity of the operator L in the sense of Definition 2 is equivalent to regularity of the pair $[L, E]$ in the sense of Definition 3. Theorems 2, 3, 4 give conditions for regularity of the pair $[L, E]$. In (11) it is shown that the pair $[E + A, E + B]$ is regular if the operators A and B are completely continuous and -1 is not their eigenvalue.

It is easy to show that the pair $[L, M]$ is regular if the operators L and M form an acute angle. Indeed, by definition (12), the operators L and M form an acute angle if their domains of definition coincide, are dense in the space, and the inequality $|(Lx, Mx)| \geq \sigma |Lx|\,|Mx|$ holds for some $\sigma > 0$. It is clear that $\sigma < 1$. Further, let $y \in L_n$, $z \in M_n$; $|y| = |z| = 1$ and let $y = Lx$, $x \in R_n$. Then

$$|P_n y| = \max_z |(y, z)| = \max_z \left| \left(\frac{Lx}{|Lx|}, z \right) \right| \geq \left| \left(\frac{Lx}{|Lx|}, \frac{Mx}{|Mx|} \right) \right| \geq \sigma.$$

Consequently, $\tau_n = \min_y |P_n y| \geq \sigma$ and $\tau = \lim_{n\to\infty} \tau_n \geq \sigma$. Let us also note that, in accordance with 2°, the gap between L_n and M_n, as well as between the ranges of the operators L and M, does not exceed $\sqrt{1 - \sigma^2}$.

Institute of Heat Power Engineering
Academy of Sciences of the Ukrainian SSR

Received
13 IX 1961

REFERENCES

1. N. I. Pol' skii, DAN, **111**, No. 6 (1956).

2. N. I. Pol' skii, *On the convergence of approximate methods of Galerkin type*, Dissertation. Kiev University, 1950.

3. N. I. Pol' skii, DAN, **86**, No. 3 (1952).

4. M. G. Krein, M. A. Krasnosel' skii, UMN, **2**, issue 3 (19) (1947).

5. N. I. Pol' skii, Scientific Notes of the Zhitomir Pedagogical Institute, **3** (1956).

6. M. A. Krasnosel' skii, DAN, **78**, No. 6 (1950).

7. S. G. Mikhlin, Scientific Notes of Leningrad University, No. 111 (1949).

8. M. H. Stone, *Linear Transformations in Hilbert Space and their Applications to Analysis*, N. Y., 1932.

9. N. I. Pol' skii, Reports of the Academy of Sciences of the Ukrainian SSR, No. 6 (1949).

10. S. G. Mikhlin, DAN, **61**, No. 2 (1948).

11. N. I. Pol' skii, Ukrainian Mathematical Journal, **7**, No. 1 (1955).

12. P. E. Sobolevskii, DAN, **116**, No. 5 (1957).

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