M. A. Krasnosel'skii et al.

Approximate solution of operator equations

M. A. Krasnosel'skii, G. M. Vainikko, P. P. Zabreiko,
Ya. B. Rutitskii, V. Ya. Stetsenko

APPROXIMATE SOLUTION OF OPERATOR EQUATIONS

translated by

D. LOUVISH

Israel Program for Scientific Translations

WOLTERS-NOORDHOFF PUBLISHING GRONINGEN

Preface

One of the most important chapters in modern functional analysis is the theory of approximate methods for solution of various mathematical problems. Besides providing considerably simplified approaches to numerical methods. the ideas of functional analysis have also given rise to essentially new computation schemes in problems of linear algebra, differential and integral equations, nonlinear analysis, and so on.

The general theory of approximate methods includes many known fundamental results. We refer to the classical work of Kantorovich; the investigations of projection methods by Bogolyubov, Krylov, Keldysh and Petrov, much furthered by Mikhlin and Pol'skii; Tikhonov's methods for approximate solution of ill-posed problems; the general theory of difference schemes; and so on.

During the past decade, the Voronezh seminar on functional analysis has systematically discussed various questions related to numerical methods; several advanced courses have been held at Voronezh University on the application of functional analysis to numerical mathematics. Some of this research is summarized in the present monograph. The authors' aim has not been to give an exhaustive account, even of the principal known results.

The book consists of five chapters.

In the first chapter we study iterative processes: conditions for convergence, estimates of convergence rate, effect of round-off errors, etc. Much attention is paid to the convergence of iterative processes under conditions incompatible with the contracting mapping principle (the theory of concave operators, the role of uniformly convex norms, and so on). The second chapter studies linear problems: methods for approximate solution of linear equations, estimates for the spectral radius of a linear operator, approximate determination of eigenvalues, etc. The theory of semiordered spaces plays an important role. The third chapter considers equations with smooth nonlinear operators, employing ideas close to those of Kantorovich. Considerable attention is paid to the situations arising in approximate methods which utilize various simplified formulas. Topological methods are proposed for *a posteriori* error estimates. Much of Chapters 1 to 3 borrows from the above-mentioned advanced courses, which were given alternately by Krasnosel'skii and Rutitskii; a few sections in the first and second chapters were written by Krasnosel'skii and Stetsenko.

Chapter 4 is devoted to a systematic theory of projection methods (method of least squares, the methods of Galerkin, Galerkin–Petrov, et al.) as applied to the approximate solution of linear and nonlinear equations, and approximate determination of eigenvalues. Most of this chapter was written by G. M. Vainikko.

The fifth and last chapter considers approximate methods in a difficult field of nonlinear analysis—the theory of branching of small solutions. The authors have seen fit to present a short account of the basic theory of formal power series. This chapter was written by Zabreiko and Krasnosel'skii.

The book includes a large number of exercises, ranging from the simple to the very difficult.

G. A. Bezsmertnykh, N. N. Gudovich, A. Yu. Levin, E. A. Lifshits, V. B. Melamed and A. I. Perov offered valuable remarks and advice in discussing various parts of the book. Several important remarks were made by L.V. Kantorovich and G.P. Akilov after reading the manuscript. The authors are deeply indebted to all those mentioned.

The authors

Contents

Chapter 2

Linear equations

Chapter 3

Equations with smooth operators

Chapter 4

Projection methods

1

Successive approximations

§ 1. Existence of the fixed point of a contraction operator

1.1. *Contraction operators.* Let A be an operator defined on a set $\mathfrak{M}$ in a Banach space E, satisfying a Lipschitz condition

$$\|Ax - Ay\| \leqq q\,\|x - y\| \qquad (x, y \in \mathfrak{M}). \tag{1.1}$$

If $q < 1$, A is called a *contraction operator* (or *contracting operator*).

Theorem 1.1. (*Contracting mapping principle*). *Let $\mathfrak{M}$ be a closed set and assume that the contraction operator A maps $\mathfrak{M}$ into itself: $A\mathfrak{M} \subset \mathfrak{M}$. Then A has a unique fixed point x^* in $\mathfrak{M}$; in other words, the equation*

$$x = Ax \tag{1.2}$$

has a unique solution x^* in $\mathfrak{M}$.

Proof. Consider the following nonnegative functional on $\mathfrak{M}$:

$$\Phi(x) = \|x - Ax\|. \tag{1.3}$$

Let x_n be a minimizing sequence for $\Phi(x)$:

$$\lim_{n\to\infty} \Phi(x_n) = \inf_{x\in\mathfrak{M}} \Phi(x) = \alpha.$$

Obviously,

$$\alpha \leqq \Phi(Ax_n) = \|Ax_n - A^2x_n\| \leqq q\|x_n - Ax_n\| = q\Phi(x_n),$$

and hence $\alpha \leqq q\alpha$. It follows that $\alpha = 0$.

Since

$$\|x_n - x_m\| \leqq \|x_n - Ax_n\| + \|Ax_n - Ax_m\| + \|Ax_m - x_m\| \leqq$$
$$\leqq \Phi(x_n) + q\|x_n - x_m\| + \Phi(x_m),$$

it follows that

$$\|x_n - x_m\| \leqq \frac{\Phi(x_n) + \Phi(x_m)}{1 - q} \to 0 \qquad (n, m \to \infty),$$

i.e., x_n is a Cauchy sequence. Its limit x^* belongs to $\mathfrak{M}$ ($\mathfrak{M}$ is closed!). The functional $\Phi(x)$ is continuous, since by (1.1)

$$|\Phi(x) - \Phi(y)| = |\,\|x - Ax\| - \|y - Ay\|\,| \leqq$$

$$\leqq \|x - y\| + \|Ax - Ay\| \leqq (1 + q)\|x - y\|.$$

Therefore

$$\Phi(x^*) = \lim_{n \to \infty} \Phi(x_n) = 0,$$

and this means that x^* is a solution of equation (1.2).

Suppose that y^* is another solution of (1.2) in $\mathfrak{M}$. Then

$$\|x^* - y^*\| = \|Ax^* - Ay^*\| \leqq q\|x^* - y^*\|$$

and so $\|x^* - y^*\| = 0$, i.e., $x^* = y^*$. ∎

The most commonly studied contraction operators are defined on the entire space E or on some ball T. In the latter case, it is convenient to replace Theorem 1.1 by the following special case:

Theorem 1.2. *Let A be a contraction operator on a closed ball $T = \{x : \|x - x_0\| \leqq r\}$ in a Banach space E, and let*

$$\|Ax_0 - x_0\| \leqq (1 - q)r. \tag{1.4}$$

Then A has a unique fixed point in T.

To prove this theorem it suffices to verify the inclusion $AT \subset T$, which is obvious: if $\|x - x_0\| \leqq r$, then (1.1) and (1.4) imply the inequality

$$\|Ax - x_0\| \leqq \|Ax - Ax_0\| + \|Ax_0 - x_0\| \leqq q\|x - x_0\| + (1 - q)r \leqq r.$$

Exercise 1.1. Let the function $K(t, s; u)$ be jointly continuous in $t, s \in [a, b]$, $-\infty < u < \infty$, and assume that

$$|K(t, s; u) - K(t, s; v)| \leqq M(\rho)\,|u - v| \quad (|u|, |v| \leqq \rho).$$

Show that the operator

$$Ax(t) = \int_a^b K[t, s; x(s)]\, ds$$

is defined on the space C of continuous functions on $[a, b]$ and satisfies a Lipschitz condition (1.1) on every ball of C. Under what conditions is it a contraction operator on the ball $\|x\| \leqq \rho_0$?

Exercise 1.2. Under the assumptions of Theorem 1.1, show that the sequence $Ax_n = x_{n+1}$ $(n = 0,1,2,\ldots)$ minimizes the functional (1.3) for arbitrary $x_0 \in \mathfrak{M}$.

1.2. *Use of an equivalent norm.* Suppose that, apart from the basic norm $\|\cdot\|$, the space E has another norm, which we shall denote by $\|\cdot\|_*$. Recall that the norms $\|\cdot\|$ and $\|\cdot\|_*$ are said to be *equivalent* if there exist positive constants m and M such that

$$m\,\|x\| \leqq \|x\|_* \leqq M\,\|x\| \qquad (x \in E).$$

When the basic norm is replaced by an equivalent one, convergent sequences of elements in the space remain convergent, closed (open) sets remain closed (open), etc.

If the operator A satisfies a Lipschitz condition with respect to one norm, this is also true for any equivalent norm. Now it often happens that A, though not a contraction, satisfies a Lipschitz condition with respect to the original norm; however, by suitable construction of an equivalent norm, A becomes a contraction. The principal difficulty in investigating concrete equations usually involves constructing this special norm.

As an example, consider a system of ordinary differential equations

$$\frac{d\xi_i}{dt} = f_i(t, \xi_1, \ldots, \xi_n) \qquad (i = 1, \ldots, n),$$

or, in vector notation,

$$\frac{dx}{dt} = f(t, x), \tag{1.5}$$

where $x(t) = \{\xi_1(t), \ldots, \xi_n(t)\}$ is a vector-function with values in n-space R^n. The problem of solving the system (1.5) with the initial condition

$$x(0) = 0 \tag{1.6}$$

is known to be equivalent to that of solving the vector integral equation

$$x(t) = \int_0^t f[s, x(s)]\,ds. \tag{1.7}$$

Assume that $f(t, x)$ is jointly continuous in t, x and satisfies a Lipschitz condition in x:

$$|f(t, x) - f(t, y)| \leqq L|x - y| \qquad (x, y \in R^n;\ 0 \leqq t \leqq T). \tag{1.8}$$

Here and above $|x|$ denotes the length of the vector x in R^n.

Consider the operator A defined by the right member of equation (1.7):

$$Ax(t) = \int_0^t f[s, x(s)]\, ds. \tag{1.9}$$

It is easily seen that, for any τ $(0 < \tau \leqq T)$, this operator is defined in the space C_τ of continuous vector-functions on $[0, \tau]$ with the norm

$$\|x\|_\tau = \max_{0 \leqq t \leqq \tau} |x(t)|. \tag{1.10}$$

It follows from (1.8) that

$$\|Ax - Ay\|_\tau \leqq L\tau \|x - y\|_\tau \qquad (x, y \in C_\tau).$$

Therefore, if

$$L\tau < 1,$$

the operator A satisfies the assumptions of Theorem 1.1. Thus the problem (1.5)–(1.6) has a unique solution $x_*(t)$ on the interval $0 \leqq t < \min\{T, 1/L\}$.

In fact, the problem (1.5)–(1.6) has a solution on the entire interval $[0, T]$. To prove this, we may again use Theorem 1.1, introducing a special norm in the space C_T. Set

$$\|x\|_* = \max_{0 \leqq t \leqq T} e^{-L_1 t} |x(t)| \qquad (x \in C_T), \tag{1.11}$$

where $L_1 > 0$. Clearly,

$$e^{-L_1 T} \|x\|_T \leqq \|x\|_* \leqq \|x\|_T \qquad (x \in C_T),$$

i.e., the norms $\|x\|_*$ and $\|x\|_T$ are equivalent. Now by (1.8)

$$\|Ax - Ay\|_* = \max_{0 \leqq t \leqq T} \left| e^{-L_1 t} \int_0^t \{f[s, x(s)] - f[s, y(s)]\}\, ds \right| \leqq$$

$$\leqq L \max_{0 \leqq t \leqq T} \int_0^t e^{L_1(s-t)} e^{-L_1 s} |x(s) - y(s)|\, ds \leqq$$

$$\leqq L \|x - y\|_* \max_{0 \leqq t \leqq T} \int_0^t e^{L_1(s-t)} ds = \frac{L}{L_1}(1 - e^{-L_1 T}) \|x - y\|_* .$$

Setting $L_1 = L$, we see that the operator (1.9) is a contraction in the norm (1.11), the constant being

$$q = 1 - e^{-L_1 T} < 1 .$$

Thus the problem (1.5)–(1.6) has a unique solution $x_*(t)$, defined for all $t \in [0, T]$.

Another example is the equation

$$x(t) = \int_0^t K(t, s) f[s, x(s)] \, ds + \phi(t) , \tag{1.12}$$

where $x(t)$ is an unknown function, $\phi(t)$ a given continuous function on $[0, T]$, $K(t, s)$ and $f(s, x)$ are jointly continuous for $0 \leqq t, s \leqq T$, $-\infty < x < \infty$ and moreover

$$|f(t, x) - f(t, y)| \leqq L |x - y| \qquad (-\infty < x, y < \infty) . \tag{1.13}$$

The right member of (1.12) obviously defines an operator A on the space C of continuous functions on $[0, T]$. It satisfies the inequality

$$\|Ax - Ay\|_* \leqq \frac{KL}{L_1}(1 - e^{-L_1 T}) \|x - y\|_* \qquad (x, y \in C) ,$$

where $\|\cdot\|_*$ is the norm (1.11) and

$$K = \max_{0 \leqq t, s \leqq T} |K(t, s)| .$$

If $L_1 \geqq KL$, then A is a contraction in the norm (1.11). It follows from Theorem 1.1 that equation (1.12) has a unique solution on $[0, T]$.

Exercise 1.3. Show that in the arguments of subsection 1.2 the norm (1.11) may be replaced by the norm

$$\|x\|_{**} = \max_{0 \leqq t \leqq T} e^{-L_2 t^2} |x(t)|$$

(for what values of L_2?).

Exercise 1.4. Prove that the linear Volterra integral equation

$$x(t) = \int_0^t K(t, s) x(s) \, ds + \phi(t)$$

with continuous kernel $K(t, s)$ (and continuous $\phi(t)$) has a unique summable solution.

1.3. *Relative uniqueness of the solution.* As a rule, concrete problems lead to equations not associated with any well-defined function space. The same equations may be considered with an operator defined in different subsets $\mathfrak{M}$ of different spaces E. Suppose that the uniqueness of the solution of an equation has somehow been proved (say, using Theorem 1.1) in a subset $\mathfrak{M}$ of a space E. Naturally, this does not imply that there are no other solutions in the space E, or that there are no solutions outside E.

As an example, consider the equation

$$x(t) = \int_0^1 x^2(s)\,ds\,. \tag{1.14}$$

The operator A defined by the right member of this equation satisfies the assumptions of Theorem 1.1 in any ball $\|x\| \leqq \rho\,(\rho < \frac{1}{2})$ in the space C of continuous functions on $[0, 1]$. Thus the unique solution of equation (1.14) in the ball $\|x\| < \frac{1}{2}$ is the trivial solution. However, there is another continuous solution, $x(t) \equiv 1$.

Instead of equation (1.14), we could have considered the scalar equation $x = x^2$, which satisfies the assumptions of Theorem 1.1 on any interval $[-\rho, \rho]\,(\rho < \frac{1}{2})$, but it has two solutions, $x = 0$ and $x = 1$. The scalar equation $x = -x^3$ has a unique solution (the trivial solution) in the space of real numbers, and three solutions in the space of complex numbers.

Now consider the homogeneous Volterra integral equation

$$x(t) = \int_0^t K(t, s)\,x(s)\,ds \tag{1.15}$$

with kernel

$$K(t, s) = \begin{cases} se^{1/t^2 - 1}, & \text{if} \quad 0 \leqq s \leqq te^{1 - 1/t^2}, \\ t, & \text{if} \quad te^{1 - 1/t^2} \leqq s \leqq 1\,. \end{cases}$$

The kernel $K(t, s)$ is continuous on the square $0 \leqq t, s \leqq 1$ (verify!). Equation (1.15) therefore has no nontrivial summable solutions (see Exercise 1.4). However, the equation has a nonsummable solution*:

$$x(t) = \begin{cases} 0, & \text{if} \quad t = 0\,, \\ 1/t\,, & \text{if} \quad 0 < t \leqq 1\,. \end{cases}$$

This example is due to Urysohn.

1.4. *Spectral radius of a linear operator.* If A is a linear operator, it is a contraction if and only if its norm is smaller than 1. This raises the problem of constructing equivalent norms with respect to which a given linear operator A has as small a norm as possible.

It is proved in the theory of linear operators (see, e.g., Kantorovich and Akilov [1], p. 153) that the limit

$$\rho_0 = \lim_{n \to \infty} \sqrt[n]{\|A^n\|} \tag{1.16}$$

exists and is finite. The number ρ_0 is called the *spectral radius* of the bounded linear operator A. Clearly,

$$\rho_0 \leqq \|A\|.$$

The spectral radius ρ_0 is characterized by the fact that the inequality

$$|\lambda| > \rho_0$$

implies that the operator $(A - \lambda I)^{-1}$ exists and is bounded. In particular, if A is a compact linear operator, then $\rho_0 = |\lambda_0|$, where λ_0 is an eigenvalue of A with maximum absolute value. Here one must consider both real and complex eigenvalues of the operator A; we recall that a number $\lambda = \sigma + i\tau$ is an *eigenvalue* of an operator A defined in a real Banach space E if there exist $x, y \in E$ such that

$$Ax = \sigma x - \tau y, \qquad Ay = \tau x + \sigma y \qquad (\|x\| + \|y\| > 0).$$

Exercise 1.5. Show that the spectral radius of a linear operator is invariant with respect to equivalent norms.

Exercise 1.6. Let

$$Ax(t) = \int_a^t K(t, s)\, x(s)\, ds$$

be a Volterra integral operator with continuous kernel $K(t, s)$ $(a \leqq t, s \leqq b)$. Show that, as an operator on C, A has spectral radius zero.

Hint. First prove the inequality

$$\|A^n\| \leqq \frac{K^n (b - a)^n}{n!},$$

where $K = \max\limits_{a \leqq t, s \leqq b} |K(t, s)|$.

Exercise 1.7. Let $\beta_n (n = 1, 2, \ldots)$ be a given monotone decreasing numerical sequence converging to zero. Construct a linear operator A such that $\sqrt[n]{\|A^n\|} = \beta_n$ $(n = 1, 2, \ldots)$.

We shall now construct an equivalent norm in the space E such that the norm of the linear operator A is arbitrarily close to its spectral radius. Let $\varepsilon > 0$ be given, and determine n such that

$$\sqrt[n]{\|A^n\|} \leqq \rho_0 + \varepsilon.$$

Now set

$$\|x\|_* = (\rho_0 + \varepsilon)^{n-1} \|x\| + (\rho_0 + \varepsilon)^{n-2} \|Ax\| + \ldots + \|A^{n-1}x\|. \quad (1.17)$$

Clearly,

$$(\rho_0 + \varepsilon)^{n-1} \|x\| \leqq \|x\|_* \leqq$$
$$\leqq [(\rho_0 + \varepsilon)^{n-1} + (\rho_0 + \varepsilon)^{n-2} \|A\| + \ldots + \|A^{n-1}\|] \|x\|,$$

i.e., the norms $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent. A simple calculation shows that

$$\|A\|_* = \sup_{\|x\|_* \leqq 1} \|Ax\|_* \leqq \rho_0 + \varepsilon.$$

Since $\rho_0 \leqq \|A\|_*$ in any norm, we have

$$\rho_0 \leqq \|A\|_* \leqq \rho_0 + \varepsilon. \blacksquare$$

It follows in particular that for any $\varepsilon > 0$ one can construct an equivalent norm in C such that the norm of the linear Volterra integral operator (whose spectral radius is zero; see Exercise 1.6) is smaller than ε.

The norm of a selfadjoint linear operator A defined on a Hilbert space is equal to its spectral radius. In the general case, there are linear operators for which there is no equivalent norm such that $\|A\|_* = \rho_0$. An example is the nonzero Volterra integral operator.

Exercise 1.8. Let A be a compact linear operator on a space E. Assume that the invariant subspace of any eigenvalue of A whose absolute value is equal to the spectral radius consists solely of eigenvectors. Construct an equivalent norm in E for which $\|A\|_* = \rho_0$.

Exercise 1.9. Let A be a compact linear operator defined on a space E. Assume that at least one eigenvalue λ, equal in absolute value to the spectral radius ρ_0, has not only eigenvectors but also generalized eigenvectors.* Show that the inequality $\|A\|_* > \rho_0$ holds for any equivalent norm in E.

1.5. *Operators which commute with contraction operators.* The following simple observation is often useful: Let B be an operator which maps a closed set $\mathfrak{M}$ in a Banach space E into itself, and A an operator which satisfies the conditions of the contracting mapping principle on $\mathfrak{M}$; if B commutes with A ($AB = BA$), then the fixed point of A is also a fixed point of B.

Indeed, if $Ax^* = x^*$, then

$$ABx^* = BAx^* = Bx^*,$$

that is, Bx^* is a fixed point of A, and therefore $Bx^* = x^*$.

* g_0 is a *generalized eigenvector* of a linear operator A, corresponding to the eigenvalue λ_0, if $Ag_0 \neq \lambda_0 g_0$, but $(A - \lambda_0 I)^n g_0 = 0$ for some $n > 1$.