G. I. Marchuk

Methods of Numerical Mathematics

Second Edition

Translated by Arthur A. Brown

G. I. Marchuk

Methods of Numerical Mathematics

Second Edition

Translated by Arthur A. Brown

With 28 Illustrations

Springer-Verlag
New York Heidelberg Berlin

G. I. Marchuk
Presidium of the Academy of Sciences
of the USSR
Leninprospekt 14
Moscow
USSR

Translator:
Arthur A. Brown
10709 Weymouth Street
Garrett Park, MD 20896
USA

Managing Editor:
A. V. Balakrishnan
Systems Science Department
University of California
Los Angeles, CA 90024
USA

AMS Subject Classification (1980): 65-XX

Library of Congress Cataloging in Publication Data
Marchuk, G. I. (Guriĭ Ivanovich), 1925–
Methods of numerical mathematics.
(Applications of mathematics; 2)
Translation of Metody vychislitel'noĭ matematiki.
Includes bibliographies and index.
1. Numerical analysis. I. Title. II. Series.
QA297.M3413 1981 519.4 81-16634
AACR2

The original Russian edition *Metody Vychislitel'noi Matematiki* was published in 1973 by Nauka, Novosibirsk.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90614-2 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90614-2 Springer-Verlag Berlin Heidelberg New York

Preface to the First Edition

The present volume is an adaptation of a series of lectures on numerical mathematics which the author has been giving to students of mathematics at the Novosibirsk State University during the span of several years. In dealing with problems of applied and numerical mathematics the author sought to focus his attention on those complicated problems of mathematical physics which, in the course of their solution, can be reduced to simpler and theoretically better developed problems allowing effective algorithmic realization on modern computers.

It is usually these kinds of problems that a young practicing scientist runs into after finishing his university studies. Therefore this book is primarily intended for the benefit of those encountering truly complicated problems of mathematical physics for the first time, who may seek help regarding rational approaches to their solution.

In writing this book the author has also tried to take into account the needs of scientists and engineers who already have a solid background in practical problems but who lack a systematic knowledge in areas of numerical mathematics and its more general theoretical framework.

Consequently, the author has selected a form of exposition which, in his opinion, helps to attract the attention of a wide range of researchers to problems of numerical mathematics. This style has required certain concessions in the exposition, thus allowing concentration only on basic ideas and approaches. As for the details (sometimes important) and the possible generalizations (such as minimal smoothness requirements, constraints on the input data, etc.), they are obvious to the specialist and present useful exercises for a beginner.

Chapter 10 is an expanded version of the paper given by the author at the International Congress of Mathematicians in Nice (1970). This chapter gives some idea of the material considered in the previous chapters, and of various methods and problems of numerical mathematics that are of fundamental importance but have not found their way into this volume.

In the process of preparation for publication this book has undergone considerable changes in response to advice and comments obtained by the author from his colleagues and associates. Those whose help is gratefully acknowledged include M. M. Lavrentiev, V. I. Lebedev, I. Marek, M. K. Fage, and N. N. Yanenko. They have made a number of constructive comments regarding the exposition of individual chapters, especially the first and fifth. The changes in the second chapter, which are due to Yu. A. Kuznetsov, are so profound that the nature of his contribution in this part is essentially that of coauthorship. The author has also enjoyed valuable advice and comments from V. T. Vasil'ev, V. P. Il'in, A. N. Konovalov, V. P. Kochergin, V. V. Penenko, V. V. Smelov, U. M. Sultangazin, and others. G. S. Rivin did considerable work in editing the manuscript. To all these, as well as M. S. Yudin who took part in preparing the book for publication, the author expresses his deep gratitude.

Preface to the Second English Edition

The second edition is a re-written version of the first, remedying various ambiguities and typographical errors, and including new material that in the author's opinion extends the scope of the methods covered therein. This edition includes a new chapter dealing with optimization theory, which is today an indispensible part of the development of mathematical models and of methods for implementing them.

Part of the material in this edition was published in 1973, in a monograph having the same title. The present text is a manual differing essentially from the monograph in that it contains a number of new ideas and algorithms of methodological and practical interest. In particular it includes:

(1) new optimization algorithms based on variational methods;
(2) problems of automating the numerical processes by use of the so-called method of "fictive" domains;
(3) consideration of iterative algorithms for the splitting process for non-commutative operators;
(4) the method of incomplete factorization, and other topics.

That portion of the book dealing with the interpolation of functions by the use of splines has been extended, and in this edition forms a self-contained chapter. Also, a separate chapter has been devoted to the notions connected with Richardson's extrapolation methods for obtaining a higher order of approximation in the solution of problems. The portions of the book dealing with variational-difference methods contain a number of new ideas, e.g., the representation of continuous functions by piecewise-discontinuous bases, and the construction of bases that take into account the singularities of solutions, as well as other new notions. The chapter dealing with the solution of inverse problems has been expanded by the inclusion of new results in perturbation theory for the solution of nonlinear problems of mathematical physics and for the analysis of the sensitivity of mathematical models with respect to variations in the initial data. There are also other expansions of the text.

This new material should provide for a better understanding of the methods of numerical mathematics for the solution of complex problems in applied mathematics.

The author expresses his deep gratitude to Yu. A. Kuznetsov, whose contribution to the preparation of the book cannot be over-estimated. For the contributions of V. V. Smelov, V. P. Il'in, V. V. Shaidurov, V. I. Agoshkov, V. A. Vasilenko, A. M. Matsokin, V. A. Bulavskii, A. L. Buchheim, Yu. S. Zav'yalov, V. A. Kuzin, G. S. Rivin, V. A. Tsetsokho, and also V. I. Drobyshevich, V. P. Dymnikov, and V. V. Penenko, the author is also deeply grateful.

The manuscript was read by N. S. Bakhvalov, V. I. Lebedev, and M. K. Fage, who made many valuable comments which helped to improve the book. That portion of the text concerned with variational inequalities was written on the basis of material graciously made available to the author by the French mathematicians J. L. Lions and R. Glowinski. To all those mentioned, the author expresses his warmest thanks.

In studying the book, the reader is recommended to make use of the exercises contained in *Problems in Numerical Mathematics*, by Drobyshevich, Dymnikov, and Rivin, Moscow, Nauka, 1980, which corresponds to the expository material in this text.

The English translation has been reviewed by the author, and fully corresponds to the Russian original version. The author appreciates very much the cooperation of Springer-Verlag.

G. I. Marchuk

Contents

Introduction

Modern electronic computers have put into the hands of research workers an effective means for using mathematical models of complex problems in science and technology. In consequence, quantitative methods of research have spread into practically all fields of human endeavor, and mathematical models have become a tool of knowledge.

The role of mathematical models is far from being exhausted in studying natural laws. Their significance is constantly being increased by the natural tendency toward optimization of technical processes and technological systems for planning experiments. In the process of research, and in the desire to develop a detailed representation of the processes under study, we are driven to the construction of ever more complex mathematical models, which require refined and generally applicable mathematical methods. Mathematical models are implemented on an electronic computer by the methods of numerical mathematics, which are continually being perfected, in keeping with developments in computing technology.

Every reduction of the problems of mathematical physics or of technology usually comes down, in the end, to a set of algebraic equations having some definite structure. Therefore, the subject of numerical mathematics is, as a rule, connected with methods of reducing a problem to a system of algebraic equations and subsequently solving them.

The construction of a set of algebraic equations corresponding to a problem with continuously varying arguments relies, in general, on *a priori* information arising from the original problem. We may know, for instance, that the solution must belong to a given class of functions characterized by given smoothness properties, or by properties of the operator associated with the problem, or by properties of the boundary conditions, etc. Such information in many cases has a decisive influence on the choice of the numerical method to be used for the solution of the corresponding algebraic equations. As a rule, the properties of the algebraic analog of the original problem must reflect our *a priori* information on the original constraints. This refers primarily to the operator of the problem and to the preservation of its properties during the reduction of the problem from one of continuous arguments to a discrete version.

Clearly, such a principle may be taken as a basic assumption in many problems. At the same time, we must note that the inheritance of the properties of operators during reduction opens up possibilities for the use of well-developed methods of functional analysis, which usually give us a simple and universal way of studying the effectiveness of the algorithms of numerical mathematics.

We now turn to a brief overview of the book in order to point out the weightings and the new ideas presented in it.

Chapter 1 is devoted to general questions in the theory of difference schemes. Along with the classical concepts such as approximation, countable stability, and convergence of solutions of difference equations, we present some important results connected with the general properties of basic and adjoint problems. These will be used in many later chapters. Section 1.1.2 contains contemporary algorithms for computing the bounds of the non-negative spectrum of matrices, and is of special interest. It is well known that the upper bound of the spectrum is found by a well-developed iterative process, and the implementation usually gives no trouble. The smallest eigenvalue—the lower bound of the spectrum—is usually difficult to compute.

The simplest method, theoretically, for finding the smallest eigenvalue is by estimating the maximal eigenvalue of the inverse operator and is of little use algorithmically. We present another approach, based on shifting the spectrum, which allows us to find the smallest eigenvalue rather easily. We have dwelt on this topic at some length because many numerical algorithms, especially those connected with the optimization of iterative processes, rely essentially on *a priori* information about the bounds of the spectrum.

In Chapter 2 we consider methods for constructing difference schemes, and we focus our attention on two approaches: the method of integral relationships and variational methods. Each of these approaches has advantages and weaknesses. We note only that they are not independent and under certain conditions lead to identical difference schemes approximating the original differential problems.

Nevertheless, it must be noted that the variational approach is to be preferred in many cases since it preserves the definiteness of the initial operators in the passage to the difference scheme. It is important to observe that this happens automatically in a wide class of problems.

We limit ourselves to the consideration of three methods of constructing difference schemes by the variational approach: namely, the methods of Ritz, Galërkin, and least-squares. These, of course, do not exhaust the great multiplicity of variational approaches, but they do provide an acquaintance with the general principles of the construction of difference schemes, which can be easily extended to other cases.

A few words on the method of finite elements. We may characterize it as a convenient way to construct difference schemes using the variational approach. At its methodological roots it is closely connected with Fourier analysis; instead of a basis of continuous functions (e.g., trigonometric functions, Legendre, or Hermite polynomials, etc.) we deal with polynomials which vanish outside a comparatively small region in the space of their arguments. These functions have been called finite elements.

The application of variational methods to the construction of difference schemes is not accidental. In fact, it follows from theory that a variational functional which adequately reflects definite laws of mechanics, mathematical physics, dynamics, etc., attains its extremal value on the solution of the problem that interests us. Therefore, if we are given a variational functional and a definite class of functions on which we are to find the minimum of the functional, the rest of the task consists of an algorithmic search for the function yielding an extremal of the functional.

If we restrict the class of admissible functions by imposing additional constraints, the minimizing function may be not a solution of the original problem but merely an approximation to the exact solution.

As the means for numerical technology becomes more powerful in the future, the role of variational methods for constructing solutions of problems of mathematical physics will continue to grow. Goal-directed methods of enumerating trial functions belonging to a wide class are beginning to appear, providing an effective means for finding extremal solutions. Thus, the use of variational methods for finding solutions to problems is ever more closely linked to the question of optimal organization of the algorithm for obtaining a solution to these problems with a given precision, i.e., to the theory of optimization.

Together with the classically formulated problems for the solution of tasks in science and technology, it is often necessary to deal with nonclassically formulated problems, for instance, those with constraints. Of course, the simplest of the constrained problems are classical, as in the case of boundary conditions for differential problems.

More complex problems with constraints demand a more complex mathematical apparatus for their solution. For instance, if we are required to find the deflection of a membrane under the action of various forces, while its position is constrained from above and below by given functions of the coordinates, the customary classical approach is powerless. However, if we set up a correspondence between this problem and some variational functional, and seek the minimum of the functional over the class of functions that each satisfy the given constraints, the minimizing function will be the solution to our problem.

A wide collection of studies in this direction has been carried out by French mathematicians. They have considered the so-called variational inequalities, which are specially adapted to the solution of problems with constraints. These questions are discussed in Section 2.8.

In Chapter 3 we deal with the interpolation of net functions. The interpolation problem arises whenever we must extend a function defined on the net to a continuous function over the whole region. Here we are concerned with the task of extending an approximate solution to the whole region, given its values on the vertices of the net, and with the task of reducing experimental data given on a discrete set of points.

The interpolation problem is a fundamental part of a system for automating construction project work, where the graphic presentation of information is at the very heart of the problem. The interpolation problem is not new, and classical methods are fully explained in the mathematical literature. A new direction in interpolation theory has been exploited in the last few decades; this is the use of the so-called spline interpolations to which Chapter 3 is essentially devoted.

Spline interpolations offer the best means of smooth completion of net functions on given classes of functions. The optimality of the spline is connected with its special extremal property. Since spline approximation is being used ever more widely in all areas of science and technology, it is necessary in our opinion that the reader should become acquainted with it.

Chapter 4 is essentially given over to iterative methods of solving linear algebraic equations. We discuss the general approaches to the solution of algebraic systems, and specific methods as well, in connection with peculiarities of the approximation of problems of mathematical physics by the use of difference and variational-difference methods. Although the literature on iterative methods is extensive and contains descriptions of many effective methods, we have considered not only the classical processes but also, and basically focused attention on, iterative methods optimizable by quadratic functionals. This constitutes our general approach to questions of optimization, for the development of numerical algorithms and for their implementation.

In the case of specific problems arising from the particular form of the matrices that arise in the numerical solution of problems in mathematical physics, we turn to methods of splitting matrices into the simplest within the general scheme of the iterative process. The splitting method is a natural development of the alternating direction method, playing an exceptional role in the numerical solution of problems in mathematical physics. It has numerous modifications and generalizations, some of which make use of variational principles.

Special attention should be paid to the direct methods of solving finite-difference equations, as discussed at the end of Chapter 4. These are primarily the fast Fourier transform and the method of cyclic reduction. Their application is relatively recent, and they are becoming increasingly popular.

Chapter 5 is devoted to methods of solving nonstationary problems. These essentially use the idea of splitting complex operators into simpler ones. We analyze not only methods well established in practice, such as the stabilization and predictor–corrector methods, but also, in some detail, the method of component-by-component splitting, which is more effective, in our opinion. The ideas are discussed in Sections 5.3.3 and 5.4.

The component-by-component method permits, at each time step, the reduction of a complex problem in mathematical physics to a sequence of very simple one-component problems. As a result, we arrive at an effective algorithm for implementation on an electronic computer which is absolutely stable and yields a second-order approximation in both time and space. It is applied to a wide class of nonstationary problems in mathematical physics.

In Chapter 6 we consider methods for increasing the precision of approximate solutions, developed by Richardson and Runge. A refinement of an approximate solution can be obtained in different ways, generally by using a higher-order approximation to the differential or integral equation in question. Richardson proposed to use a difference approximation of a comparatively low order of accuracy, but to apply it to a sequence of nets. Thus, if the initial difference equation corresponded to an approximation on a net with mesh length h, the next would correspond to a mesh length $h/2$, and so on. As a result, we arrive at difference equations defined on a sequence of nets. It turns out that, if a number of constraints are imposed on the operators and the initial values of the problem, a linear combination of the approximate solutions on the sequence of nets yields a solution with a higher order of precision than the initial solutions.

Richardson's extrapolation method, first proposed for the solution of ordinary differential equations, was successfully applied to boundary-value problems for equations of elliptic and parabolic type. Naturally, various singularities arise, and these are noted in the implementation schemes. It must be emphasized that Richardson's method can be applied to the solution of problems with a small parameter or for conditionally well-posed problems by using the method of regularization. In this case the Richardson method is based on the solution of problems with distinct parameters converging in the limit to some value. Thus the extrapolation method permits us to bring into numerical mathematics new ideas which successfully use various optimization algorithms for the solution of problems.

We must also point out the special place that has been set aside for this method for the solution of problems by variational-difference methods. In fact, we normally have two alternatives: either obtain a solution by difference equations with a very small step on the basis of rather coarse difference approximations, or use a scheme with a higher order of approximation and larger steps in the difference schemes.

The first method is simple but demands a larger volume of computation; the second is logically more difficult but demands fewer arithmetic operations. Therefore, neither is effective for problems of mathematical physics if highly precise results are required. The notion therefore arose of using the simplest