



METHODS OF NONLINEAR ANALYSIS

Volume 61-I

Richard Bellman

METHODS OF NONLINEAR ANALYSIS

Volume I

This is Volume 61 in

MATHEMATICS IN SCIENCE AND ENGINEERING

A series of monographs and textbooks

Edited by **RICHARD BELLMAN**, *University of Southern California*

A complete list of the books in this series appears at the end of this volume.

METHODS OF NONLINEAR ANALYSIS

Richard Bellman

Departments of Mathematics,
Electrical Engineering, and Medicine
University of Southern California
Los Angeles, California

VOLUME I



ACADEMIC PRESS New York and London

COPYRIGHT © 1970, BY ACADEMIC PRESS, INC.

ALL RIGHTS RESERVED.

NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM,
BY PHOTOSTAT, MICROFILM, RETRIEVAL SYSTEM, OR ANY
OTHER MEANS, WITHOUT WRITTEN PERMISSION FROM
THE PUBLISHERS.

ACADEMIC PRESS, INC.

111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
Berkeley Square House, London W1X 6BA

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 78-91424

PRINTED IN THE UNITED STATES OF AMERICA

To EMIL SELETZ

Surgeon, Sculptor, Humanitarian, and Friend

This page intentionally left blank

PREFACE

The demands of modern science inexorably force the mathematician to explore the nonlinear world. That it is a difficult and often humbling journey with painfully crude maps and rather primitive direction-finders cannot be gainsaid, but in return it can be asserted that it is richly rewarding. The few areas that have been so far examined with any care have been full of surprises and vastly stimulating to the imagination. There is every reason to believe from what so far has been glimpsed that many more surprises lay in store, novel phenomena which will open up undreamt of vistas for mathematics. It is an exciting prospect in an exciting field in an exciting time.

Explicit analytic solutions of nonlinear equations in terms of the familiar, well-tamed functions of analysis are not to be expected, although fortuitous and treasured examples occur here and there. Consequently, if either analytic or computational results are desired, various approximate methods must be applied. By and large, the effective solution of particular problems is an art. However, there do exist a number of powerful procedures for solving nonlinear problems which have been slowly forged and polished over the last one hundred years. As in all intellectual areas, art combined with method is more effective than untutored art. This book is intended as an introduction to the study of certain systematic and sophisticated techniques.

The power and versatility of these methods has been tremendously amplified by the digital computer. Even more, this new tool has motivated a careful reexamination of older methods and thus a creation of new techniques specifically designed to exploit the peculiar properties of the electronic calculator. That a good deal of mathematical ingenuity and experience is required to study significant problems with the aid of a computer hardly needs emphasizing.

This volume may also be regarded as a contribution to a new mathematical theory that is slowly emerging, a theory of closure of operations. Abstractly, the general problem may be described in the following terms. We are given the privilege of using a certain limited number of mathematical operations, such as, for example, the solution of finite systems of linear or nonlinear differential equations subject to initial conditions, or the solution of a finite system of linear algebraic equations. The task is then that of solving a particular equation, such as a partial differential equation, a two-point boundary value problem for

ordinary differential equations, or an integral equation, to a specified degree of accuracy using only these algorithms. The study becomes still more interesting and significant if we impose a constraint on the number of operations of particular types that can be employed, or on the time that may be consumed in the overall calculation. Usually, the computational facilities available automatically impose these constraints.

The two types of operations mentioned above have been carefully singled out for explicit mention since they correspond to the two major capabilities of the digital computer in the field of analysis. That they are not guaranteed capabilities merely adds to the interest of the zest of using computers to obtain numerical results.

We will present a spectrum of methods which can be used for a variety of purposes, ranging from the derivation of a simple exponential or algebraic approximation to a sequence of algorithms of increasing complexity which require a digital computer. At the moment, our orientation, as far as large-scale computing is concerned, is toward a digital computer, which is to say a leaning toward initial value problems. As hybrid computers become more powerful and prevalent, a certain mix of methods involving two-point boundary value problems and initial value problems will occur.

In general, the word "solution" must be defined operationally in terms of various technological tools available for obtaining numerical results. The arrival of the digital computer has already drastically changed the significance of this term "solution," and there will be further radical changes over the next twenty-five years.

The majority of the methods we present here can be applied to the study of partial differential equations and to the still more complex functional equations that the determined engineer and intrepid physicist are forced to face. The applications within this broader context are naturally of greater significance than those that can be made using ordinary differential equations. Despite this, we have deliberately refrained from any excursion in force into the area of partial differential equations. In this volume, the first of two, we have discussed only ordinary differential equations. However, since any complete separation between ordinary and partial differential equations is unnatural, we have broken this self-imposed vow in the second volume. This is particularly the case in the treatment of dynamic programming and invariant imbedding.

Since we are primarily concerned with introducing the reader to a variety of fundamental methods, we feel that there is considerable pedagogical force to keeping the setting as familiar as possible while new ideas are being introduced. Once acquainted with the concepts, the reader can readily apply them to all types of functional equations with a small amount of additional background. References will be found throughout to their utilization in the theory of partial differential equations.

Another strong argument for using ordinary differential equations as a proving ground is that it is relatively easy to provide a number of numerical examples in this area to illustrate different methods. In addition, a large number of interesting analytic results are available as exercises. These have principally been taken from research papers.

Having briefly described our overall aims, let us examine the structure of the book. The first three chapters contain some of the fundamental results and methods that will serve throughout both as foundation and ancillary tools. Chapter 1 discusses first- and second-order linear differential equations, subject to initial and boundary value problems, with some attention to the Riccati differential equation and a detailed study of the behavior of the physical solutions of nonlinear equations of the form

$$u' = \frac{p(u, t)}{q(u, t)}, \quad (1)$$

where p and q are polynomials in their arguments. This last represents a brief account of extensive work by Borel and Hardy, very important and useful results which are still not as well known as they should be. Some related results for the Emden–Fowler (or Fermi–Thomas) equation will be presented at the end of Chapter 4 as applications of stability theory.

Throughout we have tried to preserve some sort of a careful balance between general methods and particular problems. We have constantly kept in mind the famous dictum of Hurwitz, “It is easier to generalize than particularize.”

In Chapter 2, we present a brief account of basic results in algebraic aspects of matrix analysis that will be employed throughout the remainder of the book. The principal contents are the reduction of quadratic forms to canonical forms and associated variational problems and the Perron theorem for positive matrices. Chapter 3 discusses the use of matrices in the study of systems of linear differential equations with both constant and variable coefficients. It is impossible to study multidimensional problems in any meaningful fashion without matrix theory.

Chapter 4 contains some basic results concerning stability theory which we will employ in subsequent chapters to validate certain methods of approximation. Following the lines initiated by Poincaré and Lyapunov, we wish to compare the solutions of

$$T(u) = 0 \quad (2)$$

with those of

$$T(u) = N(u), \quad (3)$$

where $N(u)$ is “small” in some sense. The most important case is that where T is a linear operator with the property that $T(u) = 0$ possesses a convenient

solution. Closely connected with this question is the problem of estimating the difference between the solution of (2) and a function w satisfying the inequality

$$\|T(w)\| \leq \epsilon, \quad (4)$$

where $\|\cdots\|$ denotes some appropriate norm.

With these “back-up” results available, we can turn to our principal goal, the study of certain powerful methods of analytic and computational approximation. In Chapter 5, we present the Bubnov–Galerkin method, and in Chapters 7 and 8 that of Rayleigh–Ritz. Although in certain special, but important, cases the methods overlap, they are quite different conceptually and extend in different ways.

The guiding ideas are simple, as are all fundamental mathematical ideas. As always, effort and ingenuity enter in making these procedures work in particular cases.

Let us begin with a description of the Bubnov–Galerkin method. Let $T(u) = 0$ be the equation whose solution is desired. This is equivalent to minimizing the scalar quantity $\|T(u)\|$, for any norm, over the class of admissible functions. We now introduce a closure technique by restricting u to some smaller class of functions, for example, one defined by a finite set of parameters. The most important example of this is that where the restricted class is defined by

$$u = \sum_{k=1}^N a_k u_k, \quad (5)$$

where the u_k are fixed functions and the a_k are parameters. The infinite-dimensional problem of minimizing $\|T(u)\|$ is then replaced by the approximating finite-dimensional problem of minimizing the function

$$g(a_1, a_2, \dots, a_N) = \left\| \left(\sum_{k=1}^N a_k u_k \right) \right\| \quad (6)$$

with respect to the a_k .

This problem may be attacked by any of a number of techniques developed in optimization theory over the last twenty years: search techniques, gradient methods, Newton–Raphson, nonlinear programming, expansion methods, and so on. Let us note that with such methods in mind, we have deliberately refrained from any automatic use of the usual quadratic norm in the foregoing description. In the text, however, succumbing to the lure of analytic simplicity, we have considered principally quadratic functionals. More general nonlinear functionals give us an opportunity to discuss the Newton–Raphson–Kantorovich method and the use of the Lagrange expansion.

Closely associated with the Galerkin method are the methods of mean-square

approximation and differential approximation. The first is discussed at the end of Chapter 5, the second in Chapter 6.

The technique of mean-square approximation may be described in the following terms. Let $T(u) = 0$, as usual, be the original equation, and let

$$S(v, a) = 0 \quad (7)$$

be another equation, depending on the vector parameter a , which is analytically or computationally more tractable than the original equation. Thus, for example, the equation in (7) may be linear with the original equation nonlinear, or it may be a nonlinear differential equation subject to an initial value condition with the original condition linear and subject to multipoint boundary conditions. Alternatively, the original equation may contain stochastic elements, while (7) is deterministic, or conversely. The existence of analog, digital, and hybrid computers, as well as the availability of many powerful analytic theories has considerably altered the concept of "tractable." A great deal of flexibility now exists. Many different types of mathematical models are available to treat various kinds of physical processes. We have avoided stochastic processes in this volume since a good deal of effort is required to make various useful methods rigorous.

We wish to determine the parameter a so that

$$\|T(u) - S(u, a)\| \quad (8)$$

is small, where u is the solution of $T(u) = 0$ and some convenient norm is employed. Presumably, this ensures that v , the solution of (7), is close to u . This is a stability question.

Interesting complications arise from the fact that u itself is unknown. There are various "bootstrap" methods that can be employed to circumvent this annoying "little detail." Here we make brief contact with "self-consistent" methods, of such importance in modern physics. A major analytic problem is that of choosing the operator $S(v, a)$ in such a way that the function v preserves certain desirable properties of u . Very little is known in this area.

The method of differential approximation is the following. Let $R(u, b)$ be a family of operators depending on a finite-dimensional vector b , and let b be chosen so that

$$\|R(u, b)\| \quad (9)$$

is minimized where u is now given implicitly as the solution of $T(u) = 0$. We then use the solution of

$$R(v, b) = 0 \quad (10)$$

as an approximation to u . Once again, any discussion of the validity of this approach requires stability considerations.

A case of particular importance is that where

$$R(u, b) = \sum_{k=1}^N b_k R_k(u), \quad (11)$$

with the $R_k(u)$ differential operators. Observe that the aim of this procedure is once again closure. We want to solve complex functional equations using only the algorithms required to solve the more familiar and placid differential equations.

We now turn to an entirely different type of artifice. The Rayleigh–Ritz method hinges upon the observation that many equations of the form $T(u) = 0$ may be viewed as the Euler equation of an associated functional $J(u)$. By this we mean that a solution of $T(u) = 0$ is a stationary point for $J(u)$ as u varies over an appropriate space. Let us suppose that we are looking for a minimum value. The question of determining the minimum of $J(u)$ over the original infinite-dimensional space is then replaced by the finite-dimensional problem of minimizing $J(u)$ over a finite-dimensional subspace, each element of which is characterized by a finite number of parameters, say

$$u = \sum_{k=1} b_k u_k. \quad (12)$$

Here the u_k are carefully chosen functions.

The new problem, that of minimizing the expression

$$h(b_1, b_2, \dots, b_N) = J\left(\sum_{k=1}^N b_k u_k\right), \quad (13)$$

can now be approached in a number of ways. In many cases of significance, $J(u)$ is a quadratic functional, the minimization of which leads to linear equations for the b_k . There are, however, numerous difficulties associated with the solution of large systems of linear algebraic equations, which means that the real difficulties often begin at this point.

Observe that in both of the principal methods described above, monotone approximation is obtained immediately upon increasing the dimension of the finite space of functions over which the variation is allowed. Thus, if we set

$$\Delta_N = \min_{\{a_i\}} \left\| T\left(\sum_{k=1}^N a_k u_k\right) \right\|, \quad (14)$$

it is clear that

$$\Delta_1 \geq \Delta_2 \geq \dots \geq \Delta_N \geq \min_u \|T(u)\|. \quad (15)$$

Similarly, if we set

$$d_N = \min_{\{b_i\}} J \left(\sum_{k=1}^N b_k u_k \right), \quad (16)$$

we have

$$d_1 \geq d_2 \geq \cdots \geq d_N \geq \min_u J(u). \quad (17)$$

Several fundamental questions immediately present themselves. The first is that of determining when

$$\begin{aligned} \lim_{N \rightarrow \infty} \Delta_N &= \min_u \|T(u)\|, \\ \lim_{N \rightarrow \infty} d_N &= \min_u J(u). \end{aligned} \quad (18)$$

The second is that of determining when the function $u^{(N)}$ which yields Δ_N converges to the function u which yields $\min_u \|T(u)\|$, with the corresponding problem for the Rayleigh–Ritz functional.

Under reasonable conditions on $T(u)$, $J(u)$, and the spaces over which u varies, these are not difficult to answer. Far more difficult and important are the associated stability problems of estimating $\|u - u^{(N)}\|$ in terms of $\Delta_\infty - \Delta_N$, or $d_\infty - d_N$, and in determining $\|u - u^{(N)}\|$ as a function of N .

These are essential matters when the effective determination of $u^{(N)}$ is of importance. A few numerical examples, together with references to extensive work in this area will be given.

In view of the considerable effort required to treat the finite-dimensional variational problems when N is large, there is considerable motivation for finding ways of obtaining useful estimates for small N . In a sense, the major problem is the converse. It is one of determining the smallest value of N which yields an acceptable approximation. Questions of acceleration of convergence and extrapolation arise in this connection, with techniques that go back to Euler and Kronecker. We shall touch briefly on these matters.

In Chapter 8, we show how a linear equation containing a parameter can be considered to be the Euler equations associated with the minimization of a functional subject to a global constraint. Once again, Rayleigh–Ritz methods can be employed to obtain approximate results.

Many of the problem areas discussed in this volume can be further illuminated, or considered by alternative techniques, using the methods of the second volume. There we consider duality as a technique for providing upper and lower bounds, Caplygin’s method, and differential inequalities, quasilinearization, dynamic programming, invariant imbedding, the theory of iteration, and truncation techniques. The work was divided into two volumes to prevent the single,

massive volume that is so forbidding and discouraging to the newcomer into a field.

Let us encourage the reader with the flat statement that very little is known about nonlinear analysis and that it is not obvious that major breakthroughs will be made in the near future, or ever. Hundreds and thousands of fascinating and significant problems abound, each of which may require a new theory for its elucidation.

I have been fortunate in having three friends read through the book and help considerably with all aspects of preparation of the manuscript: David Collins, Thomas J. Higgins, and Art Lew. I wish to express my appreciation for their help, and to Jeanette Blood and Rebecca Karush for typing the manuscript.

RICHARD BELLMAN

Los Angeles, 1969

CONTENTS

Preface

vii

Chapter 1. First- and Second-order Differential Equations

1.1. Introduction	1
1.2. The First-order Linear Differential Equation	2
1.3. Fundamental Inequality	3
1.4. Second-order Linear Differential Equations	5
1.5. Inhomogeneous Equation	7
1.6. Lagrange Variation of Parameters	8
1.7. Two-point Boundary Value Problem	10
1.8. Connection with Calculus of Variations	11
1.9. Green's Functions	12
1.10. Riccati Equation	14
1.11. The Cauchy-Schwarz Inequality	16
1.12. Perturbation and Stability Theory	18
1.13. A Counter-example	20
1.14. $\int_{-\infty}^{\infty} f(t) dt < \infty$	21
1.15. $\int_{-\infty}^{\infty} f'(t) dt < \infty$	22
1.16. Asymptotic Behavior	23
1.17. The Equation $u'' - (1 + f(t))u = 0$	24
1.18. More Refined Asymptotic Behavior	26
1.19. $\int_{-\infty}^{\infty} f^2 dt < \infty$	27
1.20. The Second Solution	29
1.21. The Liouville Transformation	30
1.22. Elimination of Middle Term	31
1.23. The WKB Approximation	33
1.24. The One-dimensional Schrödinger Equation	33
1.25. $u'' + (1 + f(t))u = 0$; Asymptotic Behavior	33
1.26. Asymptotic Series	35
1.27. The Equation $u' = p(u, t)/q(u, t)$	37
1.28. Monotonicity of Rational Functions of u and t	38
1.29. Asymptotic Behavior of Solutions of $u' = p(u, t)/q(u, t)$	39
Miscellaneous Exercises	42
Bibliography and Comments	51

Chapter 2. Matrix Theory

2.1. Introduction	54
2.2. Determinantal Solution	55

2.3. Elimination	58
2.4. Ill-conditioned Systems	59
2.5. The Importance of Notation	60
2.6. Vector Notation	60
2.7. Norm of a Vector	61
2.8. Vector Inner Product	61
2.9. Matrix Notation	63
2.10. Noncommutativity	64
2.11. The Adjoint, or Transpose, Matrix	65
2.12. The Inverse Matrix	65
2.13. Matrix Norm	67
2.14. Relative Invariants	68
2.15. Constrained Minimization	71
2.16. Symmetric Matrices	72
2.17. Quadratic Forms	74
2.18. Multiple Characteristic Roots	75
2.19. Maximization and Minimization of Quadratic Forms	76
2.20. Min-Max Characterization of the λ_k	77
2.21. Positive Definite Matrices	79
2.22. Determinantal Criteria	81
2.23. Representation for A^{-1}	82
2.24. Canonical Representation for Arbitrary A	82
2.25. Perturbation of Characteristic Frequencies	84
2.26. Separation and Reduction of Dimensionality	85
2.27. Ill-conditioned Matrices and Tychonov Regularization	86
2.28. Self-consistent Approach	88
2.29. Positive Matrices	88
2.30. Variational Characterization of $\lambda(A)$	89
2.31. Proof of Minimum Property	91
2.32. Equivalent Definition of $\lambda(A)$	92
Miscellaneous Exercises	94
Bibliography and Comments	101

Chapter 3. Matrices and Linear Differential Equations

3.1. Introduction	104
3.2. Vector-Matrix Calculus	104
3.3. Existence and Uniqueness of Solution	105
3.4. The Matrix Exponential	107
3.5. Commutators	108
3.6. Inhomogeneous Equation	110
3.7. The Euler Solution	111
3.8. Stability of Solution	113
3.9. Linear Differential Equation with Variable Coefficients	114
3.10. Linear Inhomogeneous Equation	116
3.11. Adjoint Equation	118
3.12. The Equation $X' = AX + XB$	118
3.13. Periodic Matrices: the Floquet Representation	120
3.14. Calculus of Variations	121
3.15. Two-point Boundary Condition	122

3.16. Green's Functions	123
3.17. The Matrix Riccati Equation	123
3.18. Kronecker Products and Sums	124
3.19. $AX + XB = C$	125
3.20. Random Difference Systems	127
Miscellaneous Exercises	127
Bibliography and Comments	131

Chapter 4. Stability Theory and Related Questions

4.1. Introduction	134
4.2. Dini-Hukuhara Theorem—I	135
4.3. Dini-Hukuhara Theorem—II	138
4.4. Inverse Theorems of Perron	140
4.5. Existence and Uniqueness of Solution	140
4.6. Poincaré–Lyapunov Stability Theory	142
4.7. Proof of Theorem	143
4.8. Asymptotic Behavior	146
4.9. The Function $\varphi(c)$	148
4.10. More Refined Asymptotic Behavior	149
4.11. Analysis of Method of Successive Approximations	150
4.12. Fixed-point Methods	152
4.13. Time-dependent Equations over Finite Intervals	152
4.14. Alternative Norm	155
4.15. Perturbation Techniques	156
4.16. Second Method of Lyapunov	157
4.17. Solution of Linear Systems	157
4.18. Origins of Two-point Boundary Value Problems	158
4.19. Stability Theorem for Two-point Boundary Value Problem	159
4.20. Asymptotic Behavior	160
4.21. Numerical Aspects of Linear Two-point Boundary Value Problems	161
4.22. Difference Methods	163
4.23. Difference Equations	165
4.24. Proof of Stability	165
4.25. Analysis of Stability Proof	166
4.26. The General Concept of Stability	168
4.27. Irregular Stability Problems	168
4.28. The Emden–Fowler–Fermi–Thomas Equation	170
Miscellaneous Exercises	171
Bibliography and Comments	182

Chapter 5. The Bubnov–Galerkin Method

5.1. Introduction	187
5.2. Example of the Bubnov–Galerkin Method	188
5.3. Validity of Method	189
5.4. Discussion	190
5.5. The General Approach	190
5.6. Two Nonlinear Differential Equations	192

5.7. The Nonlinear Spring	193
5.8. Alternate Average	196
5.9. Straightforward Perturbation	196
5.10. A "Tucking-in" Technique	198
5.11. The Van der Pol Equation	198
5.12. Two-point Boundary Value Problems	200
5.13. The Linear Equation $L(u) = g$	200
5.14. Method of Moments	202
5.15. Nonlinear Case	202
5.16. Newton-Raphson Method	204
5.17. Multidimensional Newton-Raphson	207
5.18. Choice of Initial Approximation	208
5.19. Nonlinear Extrapolation and Acceleration of Convergence	210
5.20. Alternatives to Newton-Raphson	211
5.21. Lagrange Expansion	212
5.22. Method of Moments Applied to Partial Differential Equations	214
Miscellaneous Exercises	215
Bibliography and Comments	222

Chapter 6. Differential Approximation

6.1. Introduction	225
6.2. Differential Approximation	225
6.3. Linear Differential Operators	226
6.4. Computational Aspects—I	226
6.5. Computational Aspects—II	227
6.6. Degree of Approximation	228
6.7. Orthogonal Polynomials	229
6.8. Improving the Approximation	231
6.9. Extension of Classical Approximation Theory	231
6.10. Riccati Approximation	232
6.11. Transcendentally-transcendent Functions	233
6.12. Application to Renewal Equation	233
6.13. An Example	236
6.14. Differential-Difference Equations	238
6.15. An Example	239
6.16. Functional-Differential Equations	240
6.17. Reduction of Storage in Successive Approximations	242
6.18. Approximation by Exponentials	242
6.19. Mean-square Approximation	242
6.20. Validity of the Method	243
6.21. A Bootstrap Method	244
6.22. The Nonlinear Spring	244
6.23. The Van der Pol Equation	246
6.24. Self-consistent Techniques	248
6.25. The Riccati Equation	248
6.26. Higher-order Approximation	250
6.27. Mean-square Approximation—Periodic Solutions	251
Miscellaneous Exercises	253
Bibliography and Comments	255