Applied Numerical Analysis

CURTIS F. GERALD

second edition

second edition

Applied Numerical Analysis

CURTIS F. GERALD

California Polytechnic State University
San Luis Obispo

Addison-Wesley Publishing Company

 Library of Congress Catalog Card No. 77-79469.

ISBN 0-201-02696-1
BCDEFGHIJK-HA-798

Preface

This second edition of *Applied Numerical Analysis* continues the emphasis that the first edition placed on applications. The range of topics, much broader than in most elementary texts, has been retained, so the student is introduced to a wide selection of numerical procedures, and several newer methods have been added. The level of the material is still directed to the sophomore and junior student in engineering, science, mathematics, and computer science. An informal tone and dependence more on demonstrations than on rigorous proofs makes it accessible to a student who has had only the usual calculus background and an introduction to ordinary differential equations. This also makes the book a valuable guide and reference to the practicing engineer or applied mathematician.

The book has been strengthened in many ways. The already extensive problem sets have been expanded, and each chapter now includes applied problems and projects, in addition to standard exercises for practice in using the methods. These applied problems are placed in a separate section at the end of each chapter, while the exercises are keyed to the chapter sections, in order to ease the instructor's burden in making assignments. Answers to a few selected exercises are included in the text; in the exercise sections for each chapter we have identified such exercises by a small pointer (triangle) in the margin.

Significant new material has been added and many of the chapters have been extensively rewritten to improve the clarity of exposition and to provide additional illustrative examples. The order of chapters has been modified; most significantly, the chapter on solving linear systems has been moved up to become Chapter Two. At the same time, topics from linear algebra are more extensively covered so that the treatment is now relatively independent of any prior exposure in this area. The discussion of eigenvalues and eigenvectors is greatly expanded; however, for motivational reasons it continues to be placed in a later chapter with boundary-value problems, and so it is closer to the chapters on partial-differential equations, where the concepts are used in discussing stability. Other topics have been regrouped to improve the logical flow of ideas. Back reference to earlier topics frequently occurs when interrelated subjects are considered.

The strong interdependence between computers and numerical analysis is recognized by completely rewritten computer programs. FORTRAN is still the language of choice. The programs are grouped at the conclusion of each chapter for ease of reference and also because the student should normally gain some insight and understanding of the methods by personally performing the computations before he calls on the computer to grind out numbers for him. Most of the computer programs are in the form of subroutines, which the instructor may wish to incorporate in the FORTRAN library for ready access by student-written programs. The programs are written so as to be readily understandable by the student, and some concessions of efficiency are made in the interests of clarity. In the first chapter, an example of the program-development process is presented in order to help the students learn how they should write programs. An informal structured language is used to describe many of the algorithms. Since FORTRAN is not a structured language, the programs may serve as examples of how one makes the transition from algorithm to computer program.

A section on errors, with emphasis on computer arithmetic and the round-off problem, is included. Rather than place this separately as an introduction, it is incorporated within Chapter One at a time when the student is ready to appreciate the importance of the topic. Throughout the book the subject of errors is continually stressed in connection with the effectiveness and efficiency of the methods. Alternative methods are contrasted and compared from the standpoint of the computational effort required for a desired accuracy of result.

There is more than enough material for a full year's course but, by a judicious selection of topics, the book can serve shorter courses. Because the book is readable by itself, many students will find it a helpful reference to topics and procedures during their professional careers, even though they have not specifically covered the material in a class.

I wish to thank several colleagues who have suggested improvements in the book: Paul W. Davis (Worcester Polytechnic Institute), Richard Franke (Naval Postgraduate School), Stanley Preiser (Polytechnic Institute of New York), and Stanley L. Spiegel (University of Lowell).

San Luis Obispo
November 1977

C. F. G.

Contents

one

Solution of Nonlinear Equations

two

Solving Sets of Equations

three

Interpolating Polynomials

four

Numerical Differentiation and Numerical Integration

five

Numerical Solution of Ordinary Differential Equations

six

Boundary-Value Problems and Characteristic-Value Problems

seven

Numerical Solution of Elliptic Partial-Differential Equations

eight

Parabolic Partial-Differential Equations

nine

Hyperbolic Partial-Differential Equations

ten

Curve-Fitting, Splines, and Approximation of Functions

Appendixes

Chapter One

Solution of Nonlinear Equations

1.1 THE LADDER IN THE MINE

It is not uncommon, in applied mathematics, to have to solve a nonlinear equation. If you worked for a mining company the following might be a typical problem:

> *There are two intersecting mine shafts that meet at an angle of 123°, as shown in Fig. 1.1. The straight shaft has a width of 7 feet, while the entrance shaft is 9 feet wide. What is the longest ladder that can negotiate the turn? You can neglect the thickness of the ladder members, and assume it is not tipped as it is maneuvered around the corner. Your solution should provide for the general case in which the angle, A, is a variable, as well as the widths of the shafts.*

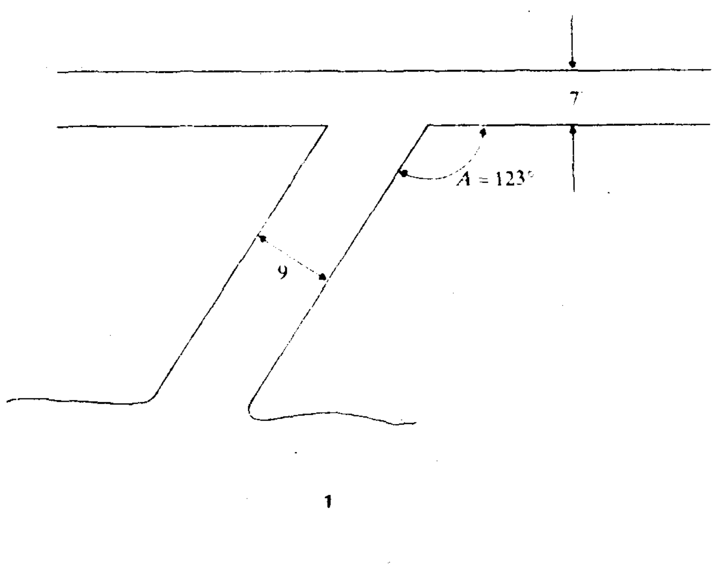

Figure 1.1

As the analysis below shows, to solve this problem we must solve a transcendental equation for the value of C:

$$\frac{9\cos(\pi-123°-C)}{\sin^2(\pi-123°-C)}-\frac{7\cos C}{\sin^2 C}=0,$$

and then substitute C into

$$\ell=\frac{9}{\sin(\pi-123°-C)}+\frac{7}{\sin C}.$$

Finding the solution to an algebraic or transcendental equation, as we must do here, is the topic of this first chapter.

Here is one way to analyze our ladder problem. Visualize the ladder in successive locations as we carry it around the corner; there will be a critical position in which the two ends of the ladder touch the walls while a point along the ladder touches the corner where the two shafts intersect. (See Fig. 1.2.) Let C be the angle between the ladder and the wall when in this critical position.

Consider a series of lines drawn in this critical position—their lengths vary with the angle C, and the following relations hold (angles are expressed in radian measure):

$$\ell_1=\frac{w_2}{\sin B}; \qquad \ell_2=\frac{w_1}{\sin C};$$

$$B=\pi-A-C;$$

$$\ell=\ell_1+\ell_2=\frac{w_2}{\sin(\pi-A-C)}+\frac{w_1}{\sin C}.$$

The maximum length of ladder that can negotiate the turn is the minimum of ℓ as a function of angle C. We hence set $d\ell/dC=0$.

$$\frac{d\ell}{dC}=\frac{w_2\cos(\pi-A-C)}{\sin^2(\pi-A-C)}-\frac{w_1\cos C}{\sin^2 C}=0.$$

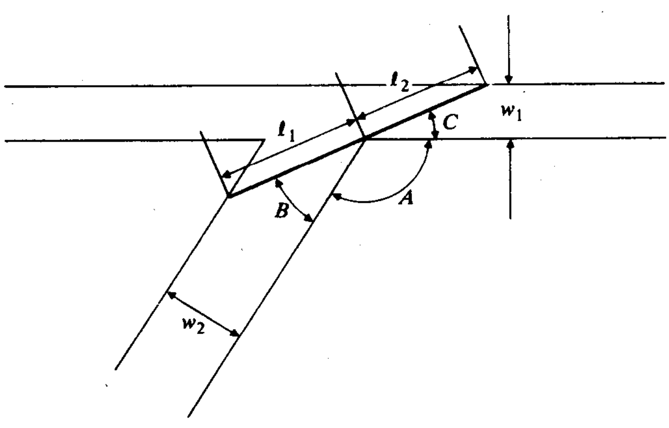

Figure 1.2

We can solve the problem if we can find the value of C that satisfies this equation. With the critical angle determined, the ladder length is given by:

$$\ell = \frac{w_2}{\sin(\pi - A - C)} + \frac{w_1}{\sin C}.$$

In this chapter we study methods to find the roots of an equation such as in our ladder-in-the-mine example. Much of algebra is devoted to the "solution of equations." In simple situations, this consists of a rearrangement to exhibit the value of the unknown variable as a simple arithmetic combination of the constants of the equation. For second-degree polynomials, this can be expressed by the familiar *quadratic formula*. For third- and fourth-degree polynomials, formulas exist but are so complex as to be rarely used; for higher-degree equations it has been proved that finding the solution through a formula is impossible. Most transcendental equations (involving trigonometric or exponential functions) are likewise intractable.

Even though it is difficult if not impossible to exhibit the solution of such equations in explicit form, numerical analysis provides a means where a solution may be found, or at least approximated as closely as desired. Many of these numerical procedures follow a scheme that may be thought of as providing a series of successive approximations, each more precise than the previous one, so that enough repetitions of the procedure eventually give an approximation which differs from the true value by less than some arbitrary error tolerance. Numerical procedures are thus seen to resemble the *limit concept* of mathematical analysis.

1.2 METHOD OF HALVING THE INTERVAL

The first numerical procedure that we shall study is that of *interval-halving*.* Consider the cubic

$$f(x) = x^3 + x^2 - 3x - 3 = 0.$$

At $x = 1$, f has the value -4. At $x = 2$, f has the value $+3$. Since the function is continuous, it is obvious that the change in sign of the function between $x = 1$ and $x = 2$ guarantees at least one root on the interval $(1, 2)$. See Fig. 1.3.

Suppose we now evaluate the function at $x = 1.5$ and compare the result to the function values at $x = 1$ and $x = 2$. Since the function changes sign between $x = 1.5$ and $x = 2$, a root lies between these values. We can obviously continue this interval-halving to determine a smaller and smaller interval within which a root must lie. For this example, continuing the process leads eventually to an approximation to the root at $x = \sqrt{3} = 1.7320508075\ldots$. The process is illustrated in Fig. 1.4.

---

* The method, also known as the *Bolzano method*, is of ancient origin. Some authors call it the *bisection method*.

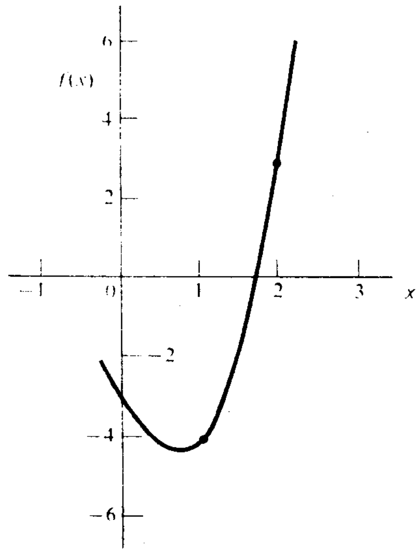

Figure 1.3

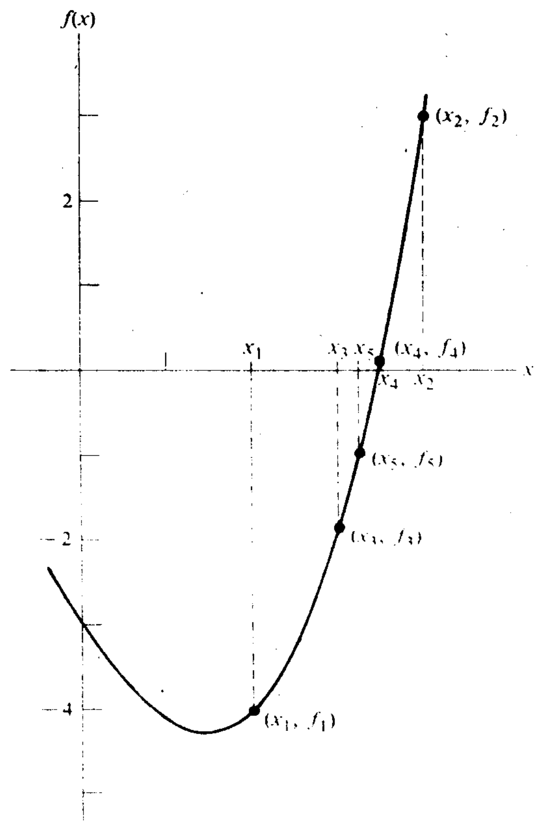

Figure 1.4

While a graphic method, as illustrated in Fig. 1.4, may be suitable if we want only an approximate answer, to obtain more accuracy we need to write a rule to do it mathematically. We should also express our algorithm (the technical name for a systematic procedure) in a way that makes it easy to implement the method with a computer program. We shall adopt a style of expressing algorithms that emphasizes the orderly structure.

Method of halving the interval (Bisection method)

To determine a root of $f(x) = 0$, accurate within a specified tolerance value, given values of x_1 and x_2 such that $f(x_1)$ and $f(x_2)$ are of opposite sign.

```
DO  WHILE ½|x1 − x2| ≥ tolerance value,
    Set x3 = (x1 + x2)/2.
    IF      f(x3) of opposite sign to f(x1):
            Set x2 = x3.
    ELSE    Set x1 = x3.
    ENDIF.
ENDDO.
```

The final value of x_3 approximates the root.

Note. The method may give a false root if $f(x)$ is discontinuous on $[x_1, x_2]$.

Applying the method to $f(x) = x^3 + x^2 - 3x - 3 = 0$, we get the results of Table 1.1. The repetition of our algorithm is called *iteration* and the successive approximations are termed the *iterates.*

The entries in Table 1.1 indicate the necessity of representing values of the argument, x, as well as of the function, $f(x)$, only approximately when we carry a limited number of decimal figures. In floating-point operations on digital computers, there is a similar inaccuracy in our work because computers retain only a limited number of significant digits. Note that this is true in all computations, not just in numerical methods. We shall give attention to such "round-off errors" later. The distinction between *numerical methods* and *numerical analysis* is that the latter term implies the consideration of errors in the procedure used. Certainly the blind use of any calculation method without concern for its accuracy is foolish.

Whether one rounds to the nearest fractional value or chops off the extra digits will make a difference in the effect of the round-off errors. In Table 1.1, the figures have been chopped after five places, which is similar to the action of most digital computers.

In addition to the limitation on accuracy because we retain only a limited number of figures in our work, there is an obvious limitation if we terminate the procedure itself too soon. One important advantage of the interval-halving

Table 1.1 Method of halving the interval for $f(x) = x^3 + x^2 - 3x - 3 = 0$.

Iteration number	x_1	x_2	x_3	$f(x_1)$	$f(x_2)$	$f(x_3)$	Maximum error in x_3
1	1	2	1.5	−4.0	3.0	−1.875	0.5
2	1.5	2	1.75	−1.875	3.0	0.17187	0.25
3	1.5	1.75	1.625	−1.875	0.17187 . . .	−0.94335 . . .	0.125
4	1.625	1.75	1.6875	−0.94335 . . .	0.17187	0.40942	0.0625
5	1.6875	1.75	1.71875	−0.40942 . . .	0.17187 . . .	−0.12478	0.03125
6	1.71875	1.75	1.73437 . . .	−0.12478 . . .	0.17187 . . .	−0.02198	0.015625*
7	1.71875	1.73437 . . .	1.72656 . . .				0.0078125
	⋮	⋮	⋮				
∞			1.73205 . . .			−0.00000 . . .	

* Actual error in x_3 after 5 iterations is 0.01330.

Table 1.2 Halving the interval for $f(x) = e^x - 3x = 0$.

Iteration number	x_1	x_2	x_3	$f(x_1)$	$f(x_2)$	$f(x_3)$	Maximum error in x_3
1	1.0	2.0	1.5	−0.028172	1.38906	−0.01831	0.5
2	1.5	2.0	1.75	−0.01831	1.38906	0.50460	0.25
3	1.5	1.75	1.625	−0.01831	0.50460	0.20342	0.125
4	1.5	1.625	1.5625	−0.01831	0.20342	0.08323	0.0625
5	1.5	1.5625	1.53125	−0.01831	0.08323	0.03020	0.03125
6	1.5	1.53125	1.51562	−0.01831	0.03020	0.00539	0.015625*
	.	.	.				
	.	.	.				
	.	.	.				
∞			1.51213 . . .				

* Actual error in x_3 after 5 iterations is −0.01912.

method, beyond its simplicity, is our knowledge of the accuracy of the current approximation to the root. Since a root must lie between the x-values where the function changes sign,† the error in the last approximation can be no more than one-half the last interval of which it is the midpoint, and this interval is known exactly since the original difference, $|x_1 - x_2|$, is halved at each iteration. For other methods, the accuracy determination is much more difficult.

The accuracy of a computed value is usually expressed either as the *absolute error* (true value − approximate value) or as the *relative error* (absolute error divided by true value). The relative error is often the better measure of accuracy for very large or very small values. Sometimes the accuracy is expressed as the number of digits that are correct; in other cases, the number of correct digits after the decimal point is used. When the true value is unknown, it is impossible to express the accuracy with exactness, and approximate accuracy must be specified. Frequently we will put bounds on the size of the error.

The method of halving the interval applies equally well to transcendental equations, as do the other methods of this chapter. Table 1.2 shows the results when we apply the method to $f(x) = e^x - 3x = 0$, which has a root between $x = 1$ and $x = 2$.

The method of interval halving requires that starting values be obtained before the method can begin. This is true of most methods for root finding. Getting these starting values can be done by making a rough graph, by trial calculations, or by writing a search program on a computer or programmable

† Observe that if the function is discontinuous, $f(x)$ may change sign without having a root in the interval. Unknown functions should be examined for continuity before attempting to evaluate their roots.

calculator. Perhaps the best way is through interactive graphics, letting the computer draw curves at the direction of the user, and varying the parameters at the console to find approximate values of roots.

1.3 METHOD OF LINEAR INTERPOLATION

While the interval-halving method is easy and has simple error analysis, it is not very efficient. For most functions, we can improve the rate at which we converge to the root. One such method is the method interpolation.* Suppose we assume that the function is linear over the interval (x_1, x_2), where $f(x_1)$ and $f(x_2)$ are of opposite sign. From the obvious similar triangles in Fig. 1.5 we can write†

$$\frac{x_2 - x_3}{x_2 - x_1} = \frac{f(x_2)}{f(x_2) - f(x_1)},$$

$$x_3 = x_2 - \frac{f(x_2)}{f(x_2) - f(x_1)}(x_2 - x_1).$$

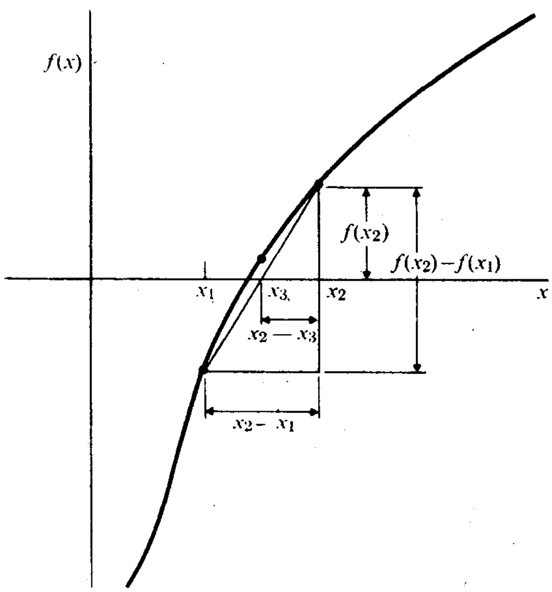

Figure 1.5

---

* This is also known as the method of false position, and by the Latinized version *regula falsi*. It is also a very old method.

† Note that, since $[f(x_2) - f(x_1)]/(x_2 - x_1)$ is the slope of the secant line, which approximates the slope of the function in the neighborhood of the root, the equation can be considered to be $x_3 = x_2 - f(x_2)/(\text{slope of function})$. Compare to Newton's method, in the next section.