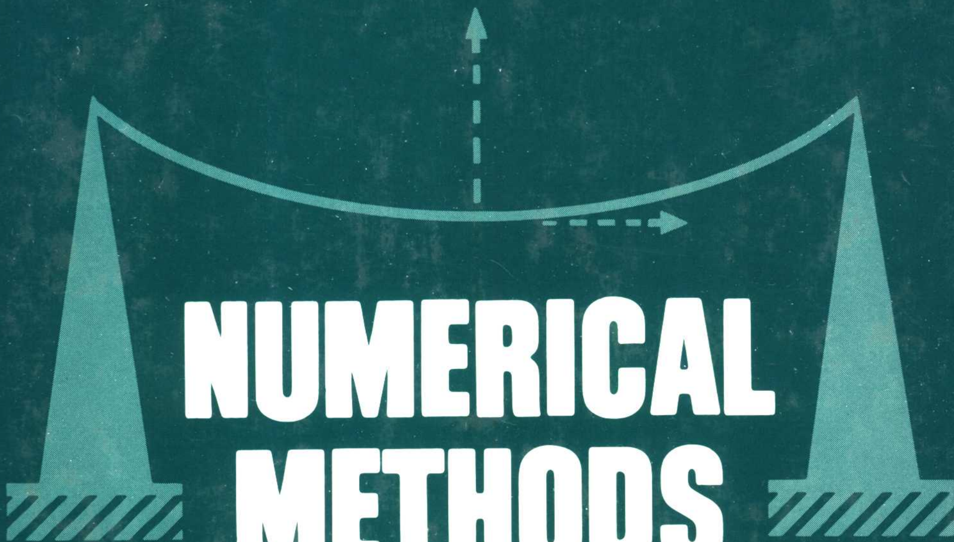
CARL E. PEARSON



NUMERICAL METHODS IN ENGINEERING AND SCIENCE

NUMERICAL METHODS IN ENGINEERING AND SCIENCE

Carl E. Pearson

University of Washington

VAN NOSTRAND REINHOLD COMPANY

New York

Library of Congress Catalog Card Number: 85-22516
ISBN: 0-442-27344-4

Manufactured in the United States of America.

Published by Van Nostrand Reinhold Company Inc.
115 Fifth Avenue
New York, New York 10003

Van Nostrand Reinhold Company Limited
Molly Millars Lane
Wokingham, Berkshire RG11 2PY, England

Van Nostrand Reinhold
480 La Trobe Street
Melbourne, Victoria 3000, Australia

Macmillan of Canada
Division of Canada Publishing Corporation
164 Commander Boulevard
Agincourt, Ontario MIS 3C7, Canada

15 14 13 12 11 10 9 8 7 6 5 4 3 2

Library of Congress Cataloging-in-Publication Data
Pearson, Carl E.
Numerical methods in engineering and science.
Includes bibliographies and index.
1. Engineering mathematics. 2. Science—Mathematics.
3. Numerical analysis. I. Title.
TA335.P43 1986 620'.0042 85-22516
ISBN 0-442-27344-4

Preface

A course in numerical analysis has become accepted as an important ingredient in the undergraduate education of engineers and scientists. *Numerical Methods in Engineering and Science* reflects my experience in teaching such a course for several years. Related work in industry and research has influenced my choice of content and method of presentation.

Most students at the undergraduate level will have had, at the very least, an introductory course in ordinary differential equations. Tutorial appendixes on complex variables, determinants, partial differentiation, and Taylor expansions are included at the end of this book. Other background material is developed as needed. For example, Chapter 2 (linear equations) begins with an outline of matrix algebra; this will represent review material for some students, but it will spare others the necessity of consulting references. Proofs for almost all results of importance are given, in what I hope is palatable form. Overall, the book should be reasonably self-contained.

A number of illustrative computer programs are provided. The language chosen is FORTRAN (ANSI 77) because of the wide availability of service and application programs in that language. I am aware of course that some readers, for good reasons, will prefer other languages; however, since most languages are sufficiently similar to FORTRAN there should be little difficulty in translating from one language to another as required.

I have not tried to include an extensive collection of library-type programs. At this stage, it seems to me that it is important for the student to acquire facility in writing actual programs (and this is called for, in the text and in exercises). This kind of programming experience should help solidify the understanding of numerical techniques, as well as provide perspective for the eventual use of subroutine libraries.

This book contains more material than would normally be included in an introductory undergraduate course. I feel, however, that it is useful to cover the various topics with some degree of completeness so that the book may serve the student as a subsequent reference and source of ideas. Illustrative examples are given throughout the text. It is important for the student to work problems, and a fairly large number of problems that illustrate (and in some cases, extend) the material of the text will be found at the end of each chapter. In a course based on this book, the instructor may want to supplement these problems with some of the usual kind of drill-type exercises.

I am grateful to many past and present associates who have had an influence on *Numerical Methods in Engineering and Science.* Several colleagues have been

kind enough to critically read the manuscript, in various stages of preparation, and to offer suggestions concerning appropriate material or treatment. I want to express particular appreciation to Professors David Benney (Massachusetts Institute of Technology), Graham Carey (University of Texas), Walter Christiansen (University of Washington), Robin Esch (Boston University), Robert MacCormack (Stanford University), and Chris Newbery (University of Kentucky). The responsibility for inaccuracies or for inelegances of exposition remains, of course, mine alone.

It is a pleasure to thank Kathy Hamilton for her painstaking efforts to make everything legible and for her patience in dealing with many changes. In the process, she has become something of a numerical analyst herself.

CARL E. PEARSON

Contents

CHAPTER

1

NONLINEAR EQUATIONS

The purpose of this chapter is to discuss a number of methods applicable to the solution of a single nonlinear equation, usually algebraic or transcendental in character. Sets of equations are considered in Chapter 2 (linear case) and Chapter 3 (nonlinear case).

1.1 A SAMPLE PROBLEM

Suppose that electrical cables are to be strung between a series of towers. To design the towers one has to know the tension in the cables, and this depends on the ground clearance desired. Consider such a cable, as shown in Figure 1.1. It is a standard exercise in differential equations texts to show that the cable takes the shape of a catenary. If ρ denotes the linear density of the cable, T the horizontal component of the cable tension, and g the acceleration of gravity, then in terms of horizontal distance s, the height w of the cable is given by

$$\begin{aligned} w &= \frac{1}{\alpha}\,[\cosh \alpha s - 1] \\ &= \frac{1}{\alpha}\left[\frac{e^{\alpha s} + e^{-\alpha s}}{2} - 1\right], \end{aligned} \tag{1.1}$$

where $\alpha = \rho g/T$. Here the origin of the (s, w) coordinate system has been made to coincide with the lowest point on the cable, as shown in Figure 1.1.

If l denotes the half-span, then at $s = l$ the height of the cable above its lowest point is given by

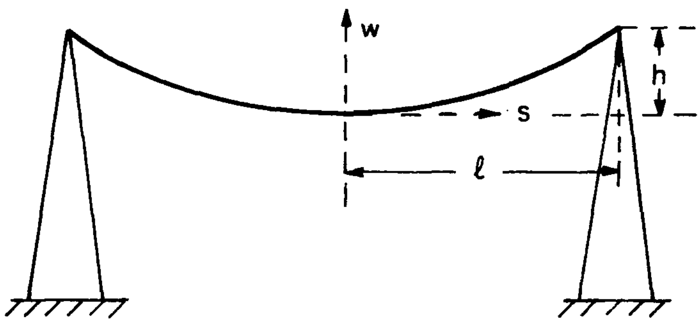


Figure 1.1. Cable problem.

$$h = \frac{1}{\alpha} [\cosh \alpha l - 1] . \qquad (1.2)$$

The problem is then to determine α (and so T) if the sag h is specified. It is often useful to introduce dimensionless variables, and we do so here by the definitions

$$x = \alpha l = \frac{\rho g l}{T} , \qquad \lambda = \frac{h}{l} .$$

The quantities x and λ are dimensionless. Equation (1.2) becomes

$$\lambda x = \cosh x - 1 . \qquad (1.3)$$

This equation is now to be solved for x, where λ is specified. This problem will serve as an example to which several solution methods will be applied. The problem is, of course, a rather simple one, but it has the advantage that we can concentrate on solution methods without encumbrances of algebraic complexity.

It is worthwhile to begin with a graphical look at Equation (1.3). Suppose, to be specific, that the designer's choice of cable clearance transforms into a desired value of .1575 for λ. Then we could plot the two curves

$$y = 0.1575x , \qquad y = \cosh x - 1$$

and find that value of x at which the curves intersect. This is done in Figure 1.2 and we obtain $x \cong .32$. A refinement of the graph would give a more accurate value for x, but this process could become cumbersome if several significant figures were required, if the equation to be solved were a complicated one, or if solutions corresponding to several values of λ were required. Consequently, it is appropriate to look for more efficient methods. Nevertheless, a simple preliminary sketch is often useful—it can protect us against a future gross error and it can also provide a starting value for an iterative process.

We remark that it is sometimes useful to interchange the roles of dependent and independent variables. Our interest is in solving Equation (1.3) for x if λ is specified. We could equally well think of Equation (1.3) as determining λ when x

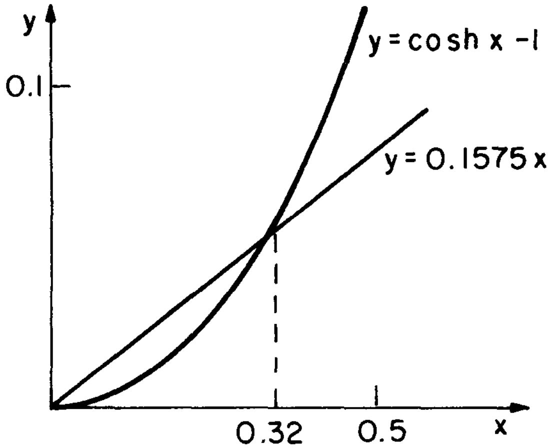


Figure 1.2. Graphical determination of x.

is given. It would be easy to plot λ as a function of x; if a number of values of λ are to be considered, the corresponding values of x could be read off from this graph. Of course, this particular equation happens to be of rather simple form—the parameter λ could easily appear on both sides of a more complicated equation.

1.2 REPEATED BISECTION METHOD

Define

$$y = \cosh x - \lambda x - 1 \; . \qquad \textbf{(1.4)}$$

Then the problem of Section 1.1 requires us to find that value of x, say x_0, for which y vanishes. Again, we take $\lambda = .1575$.

The idea of the bisection method is to start with a pair of values for x, say x_1 and x_2, for which the corresponding values for y (denoted by y_1 and y_2, respectively) are of opposite sign. Then x_0 must lie between x_1 and x_2. We calculate next the midpoint value $x_3 = \frac{1}{2}(x_1 + x_2)$ and determine the corresponding quantity y_3. If y_3 has the same sign as y_1, we deduce that x_0 must lie between the pair x_2 and x_3, whereas if y_3 has the same sign as y_2, then x_0 must lie between x_1 and x_3. Of these two subintervals the one that is known to contain x_0 is then bisected again, and the process continues iteratively.

Figure 1.3 shows the first few steps for the case of Equation (1.4), with $\lambda = .1575$. Guided by the approximate value .32 for x_0, we chose x_1 to be .31 and x_2 to be .33. Calculation shows that $y_1 < 0$ and $y_2 > 0$ (the exact values don't matter much—only the signs), so that x_0 must lie between x_1 and x_2. The midpoint value x_3 is given by $\frac{1}{2}(x_1 + x_2) = .32$, and we find $y_3 > 0$, so that x_0 must lie between x_1 and x_3. The midpoint of these two is $x_4 = \frac{1}{2}(x_1 + x_3) = .315$, for

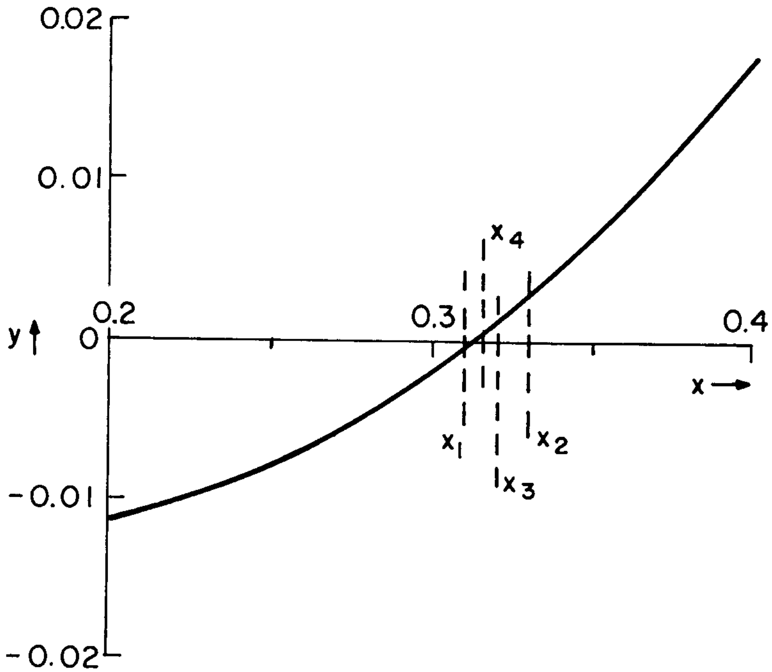


Figure 1.3. Repeated bisection.

which $y_4 > 0$. Then $x_5 = .3125$ (with $y_5 > 0$, so that x_0 must lie between .31 and .3125), and we can continue indefinitely, halving the size of the interval containing x_0 at each step.

Proceeding as far as x_8, we find that x_0 must lie somewhere in the interval (.3124, .3125). More generally, if the original interval length is denoted by δ (in our case, $\delta = x_2 - x_1 = .02$), then after n bisections the interval containing the solution point will have length $\delta/2^n$. Although we will shortly look at more efficient methods, the bisection method has some advantages. At each step, only functional evaluations (and in fact only one new one) are necessary; we do not have to calculate derivatives. Also, convergence is guaranteed, since we have a sequence of intervals, of decreasing size, within which x_0 must lie.

If y is a more intricate function than that described by Equation (1.4), the initial interval may contain several zeros of that function. In that event only one of those zeros will usually be found by the bisection method.

1.3 SECANT METHODS

Let $y = f(x)$, where $f(x)$ is some given function, and let it be required to find that value of x, say x_0, at which y vanishes. A plot of $y = f(x)$ might look something like that shown in Figure 1.4. In the secant method one chooses two x-values, x_1 and x_2, and calculates the corresponding y_1 and y_2 values. This gives a pair of points (x_1, y_1) and (x_2, y_2) lying on the curve. The line joining these points (the *secant*) is drawn, and its intersection point x_3 with the x-axis is calculated. Figure 1.4 suggests that if x_1 and x_2 are reasonably close to the desired root x_0, then x_3 should be an even better approximation to x_0. We can now iterate the process, using the pair (x_2, y_2), (x_3, y_3) for the next step, and so on.

In the case of Equation (1.4), with $\lambda = .1575$ as before, let us take $x_1 = .34$, $x_2 = .33$. Then the straight line through (.33, .002971) and (.34, .004809) is given by

$$y = .002971 + \left(\frac{.004809 - .002971}{.34 - .33}\right)(x - .33)\ .$$

The point x_3 at which this line cuts the x-axis is that value of x for which $y = 0$; we find $x_3 = .3138$. One more step (starting with x_2 and x_3) yields $x_4 = .3125$. Further steps change this value only slightly, so that we can take x_4 as an approximation to x_0, correct to about four figures.

The above method is termed the *secant method*, and the example shows that it can be very effective. Unfortunately, pathological situations, such as the presence of extrema (see Fig. 1.5), can arise in which the method may not converge.

A related method, with guaranteed convergence, is the *rule of false position*— or *regula falsi*, as it was termed by seventeenth-century numerical analysts. In this

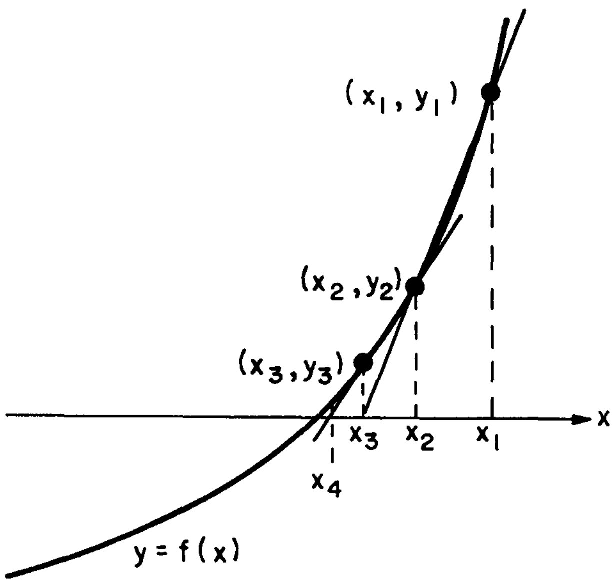


Figure 1.4. Secant method.

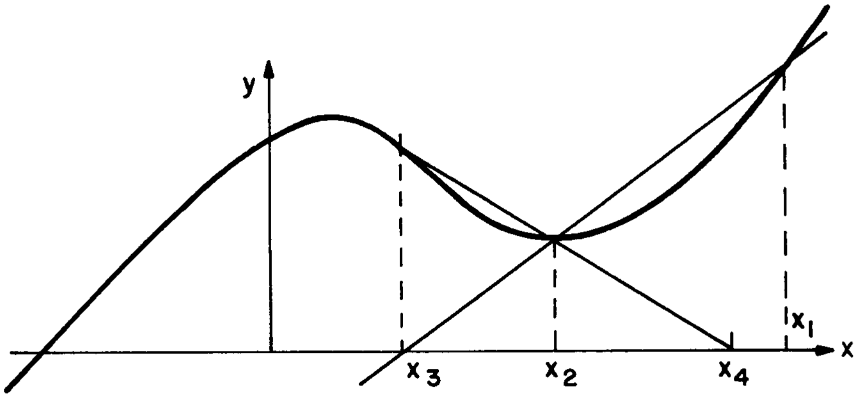


Figure 1.5. Possible nonconvergence of secant method.

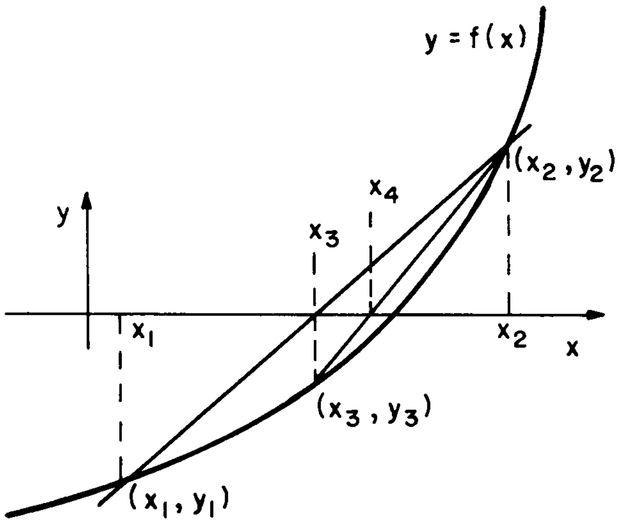


Figure 1.6. Rule of false position.

method the two initial x-values, x_1 and x_2, are chosen to lie on opposite sides of the root x_0—that is, y_1 and y_2 must have opposite signs. The two points (x_1, y_1) and (x_2, y_2) are again connected by a straight line, which intersects the x-axis at x_3. The next iteration step starts with that one of the two possible pairs—the pair (x_1, y_1) and (x_3, y_3), or the pair (x_2, y_2) and (x_3, y_3)—for which the two y-values have opposite signs. The method now continues in the same way. (See Fig. 1.6). If we take $x_1 = .31$, $x_2 = .33$ in our standard example, we find $x_3 = .3123$, $x_4 = .3124$, $x_5 = .3124$, ...

As indicated in Figure 1.6, the successive iteration points x_i will eventually all lie on the same side of the root x_0. Variants of the method, which provide sets of x_i-values approaching x_0 from both above and below, are possible; see Problem 1.7.

The subroutine FALSE is based on the rule of false position. The input and output parameters are described in comment statements. (Note that the desired tolerance, TOL, should not be chosen so small that the number of significant figures carried by the computer becomes crucial.) A second "driver" program follows the subroutine. This program uses FALSE to again solve for x in $\cosh(x) - .1575x - 1 = 0$. The output is found to be

.31244E+00   −.11921E−0.5   2

which checks the previous answer and also the subroutine program itself.

```
C  SUBROUTINE FALSE(N,XL,XR,XC,FC,I,TOL) USES METHOD OF FALSE
C  POSITION TO ITERATE TOWARDS A ZERO OF A FUNCTION FF(X).  THE
C  ROUTINE WILL ITERATE N TIMES UNLESS FF(X) BECOMES LESS THAN
C  TOL IN ABSOLUTE VALUE.  INPUT VALUES OF X FOR WHICH FF(X) HAS
C  OPPOSITE SIGNS ARE REQUIRED.
C
C  INPUT:
C        XL,XR = VALUES OF X FOR WHICH FF(X) HAS OPPOSITE SIGNS,
C           AND BETWEEN WHICH A ZERO OF FF(X) MUST LIE
C            N = MAXIMUM NUMBER OF ITERATIONS
C          TOL = TOLERANCE.  ROUTINE ITERATES N TIMES, UNLESS
C           ABSOLUTE VALUE OF FF BECOMES LESS THAN TOL
C
C  OUTPUT:
C           XC = LAST ITERATION POINT, AND BEST APPROXIMATION
C                    TO VALUE OF X SUCH THAT FF(X)=0
C           FC = VALUE OF FF(XC)
C            I = NUMBER OF ITERATIONS ACTUALLY PERFORMED
C
C  FUNCTION CALLED:  FF(X)
C
      SUBROUTINE FALSE(N,XL,XR,XC,FC,I,TOL)
      I=0
      FL=FF(XL)
      FR=FF(XR)
3     I=I+1
      XC=(XL*FR-XR*FL)/(FR-FL)
      FC=FF(XC)
      IF(ABS(FC).LE.TOL.OR.I.GE.N) RETURN
C  FOR NEXT ITERATION, CHOOSE TWO POINTS BRACKETING ZERO OF FF
      IF(FL*FC.LT.0.) THEN
      XR=XC
      FR=FC
      ELSE
      XL=XC
      FL=FC
      END IF
      GO TO 3
      END
```

```
C   THE FOLLOWING PROGRAM USES <FALSE> TO APPROXIMATE A ZERO
C   OF THE FUNCTION    COSH(X)-.1575*X-1
C
      PROGRAM TEST
      N=5
      XL=.31
      XR=.33
      TOL=1.E-5
      CALL FALSE(N,XL,XR,XC,FC,I,TOL)
      WRITE(*,100) XC,FC,I
100   FORMAT(2E12.5,I4)
      END
      FUNCTION FF(X)
      FF=COSH(X)-.1575*X-1.
      END
```

1.4 NEWTON'S METHOD

In the secant method of Figure 1.4, a line was drawn through two points (x_1, y_1) and (x_2, y_2) of a curve $y = f(x)$, and the intersection of this line with the x-axis was determined. If the point x_2 is made to approach x_1, then in the limit the secant becomes the tangent to the curve at the point (x_1, y_1), and this leads to *Newton's method* (sometimes called the Newton–Raphson method).

We use a prime to denote a derivative: so if $y = f(x)$, then $dy/dx = f'(x)$. Then the slope of the tangent line at x_1 is $f'(x_1)$, and the equation of this line is

$$y = f(x_1) + f'(x_1)(x - x_1) .$$

Let x_2 denote the intersection point of this line with the x-axis (see Fig. 1.7). Then

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} . \tag{1.5}$$

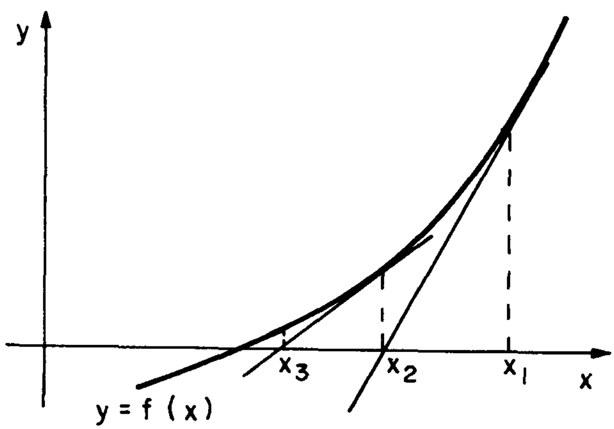


Figure 1.7. Newton's method.

The process can now be repeated, with

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

and so on. In general,

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} . \qquad \textbf{(1.6)}$$

For the previous example, with $f(x) = \cosh(x) - .1575x - 1$, Eq. (1.6) gives

$$x_{i+1} = x_i - \frac{\cosh x_i - .1575x_i - 1}{\sinh x_i - .1575} .$$

Starting at $x_1 = .30$, we obtain the sequence $x_2 = .31300$, $x_3 = .31245$, and at this point the sequence has essentially converged.

It is easy to derive Equation (1.6) analytically by considering the change in $f(x)$ resulting from a small change in x; also, this approach will carry over to the several-variable case of Chapter 3. Suppose x_i has been determined, and let x be any point near x_i. From the definition of a derivative, we know that if we define $\epsilon_i(x)$ by

$$\epsilon_i(x) = f'(x_i) - \frac{f(x) - f(x_i)}{x - x_i} . \qquad \textbf{(1.7)}$$

then $\epsilon_i(x) \to 0$ as $x \to x_i$. This equation may be written

$$f(x) - f(x_i) = f'(x_i) \cdot (x - x_i) - \epsilon_i(x) \cdot (x - x_i) . \qquad \textbf{(1.8)}$$

Since $\epsilon_i(x) \to 0$ as $x \to x_i$, it follows that a first approximation to $f(x) - f(x_i)$, if x is close to x_i, is given by the term $f'(x_i) \cdot (x - x_i)$. Within this approximation, we choose x, denoted now by x_{i+1}, so as to make $f(x)$ vanish; Equation (1.8) becomes

$$0 - f(x_i) \cong f'(x_i) \cdot (x_{i+1} - x_i) ,$$

and this leads again to Equation (1.6).

Newton's method converges very rapidly, once the iteration points are close enough to the root. To show this, we use Taylor's theorem, the derivation of which is sketched in Appendix A. Suppose x_0 is a root of $f(x) = 0$, and suppose the ith iteration point, x_i, is close enough to x_0 that the higher-order terms in Taylor's formula

$$f(x_i) = f(x_0) + f'(x_0) \cdot (x_i - x_0) + \tfrac{1}{2} f''(x_0) \cdot (x_i - x_0)^2 + \text{higher-order terms} \qquad \textbf{(1.9)}$$

are negligible. Then, since $f(x_0) = 0$, Equation (1.6) becomes

$$x_{i+1} \cong x_i - \frac{f'(x_0) \cdot (x_i - x_0) + \frac{1}{2} f''(x_0) \cdot (x_i - x_0)^2}{f'(x_0) + f''(x_0) \cdot (x_i - x_0)} ,$$