Instructor's Solution Manual
for
Numerical Methods: Using MATLAB

John H. Mathews and Kurtis D. Fink

August 2002

Contents

1 Preliminaries **5**
- 1.1 Review of Calculus 5
- 1.2 Binary Numbers 7
- 1.3 Error Analysis 11

2 The Solution of Nonlinear Equations $f(x) = 0$ **15**
- 2.1 Iteration for Solving $x = g(x)$ 15
- 2.2 Bracketing Methods for Locating a Root 18
- 2.3 Initial Approximation and Convergence Criteria 20
- 2.4 Newton-Raphson and Secant Methods 21
- 2.5 Aitken's Process and Steffensen's and Muller's Methods 27

3 The Solution of Linear Systems $\mathbf{AX} = \mathbf{B}$ **33**
- 3.1 Introduction to Vectors and Matrices 33
- 3.2 Properties of Vectors and Matrices 34
- 3.3 Upper-Triangular Linear Systems 37
- 3.4 Gaussian Elimination and Pivoting 38
- 3.5 Triangular Factorization 39
- 3.6 Iterative Methods for Linear Systems 40
- 3.7 Iteration for Nonlinear Systems 44

4 Interpolation and Polynomial Approximation **51**
- 4.1 Taylor Series and Calculation of Functions 51
- 4.2 Introduction to Interpolation 57
- 4.3 Lagrange Approximation 59
- 4.4 Newton Polynomials 62
- 4.5 Chebyshev Polynomials (Optional) 65
- 4.6 Padé Approximations 68

5 Curve Fitting **73**
- 5.1 Least-Squares Line 73
- 5.2 Curve Fitting 77
- 5.3 Interpolation by Spline Functions 81

5.4 Fourier Series and Trigonometric Polynomials 87
5.5 Bézier Curves 89

6 Numerical Differentiation **93**
6.1 Approximating the Derivative 93
6.2 Numerical Differentiation Formulas 97

7 Numerical Integration **101**
7.1 Introduction to Quadrature 101
7.2 Composite Trapezoidal and Simpson's Rule 105
7.3 Recursive Rules and Romberg Integration 114
7.4 Adaptive Quadrature 118
7.5 Gauss-Legendre Integration 118

8 Numerical Optimization **121**
8.1 Minimization of a function 121
8.2 Nelder-Mead and Powell's Methods 125
8.3 Gradient and Newton's Methods 130

9 Solution of Differential Equations **135**
9.1 Introduction to Differential Equations 135
9.2 Euler's Method 138
9.3 Heun's Method 140
9.4 Taylor Series Method 142
9.5 Runge-Kutta Methods 144
9.6 Predictor-Corrector Methods 147
9.7 Systems of Differential Equations 147
9.8 Boundary Value Problems 149
9.9 Finite-Difference Method 151

10 Solution of Partial Differential Equations **155**
10.1 Hyperbolic Equations 155
10.2 Parabolic Equations 157
10.3 Elliptic Equations 160

11 Eigenvalues and Eigenvectors **163**
11.1 Homogeneous Systems: The Eigenvalue Problem 163
11.2 Power Method 168
11.3 Jacobi's Method 169
11.4 Eigenvalues for Symmetric Matrices 170

Chapter 1

Preliminaries

1.1 Review of Calculus

1. (a) $L = \lim_{n\to\infty} \frac{4n+1}{2n+1} = 2$
   $\lim_{n\to\infty} \epsilon_n = \lim_{n\to\infty}(2 - \frac{4n+1}{2n+1}) = 2 = \frac{4}{2} = 0$

   (b) $\lim_{n\to\infty} \frac{2n^2+6n-1}{4n^2+2n+1} = \frac{2}{4} = \frac{1}{2}$
   $\lim_{n\to\infty} \epsilon_n = (\frac{1}{2} - \frac{2n^2+6n-1}{4n^2+2n+1}) = \frac{1}{2} - \frac{1}{2} = 0$

2. (a) $\lim_{n\to\infty} sin(x_n) = sin(\lim_{n\to\infty} x_n) = sin(2)$

   (b) $\lim_{n\to\infty} ln(x_n^2) = ln(\lim_{n\to\infty} x_n^2) = ln(4)$

3. (a) Since f is continuous on $[-1, 0]$; solve

$$\begin{aligned} -x^2 + 2x + 3 &= 2 \\ x^2 - 2x - 1 &= 0 \\ x &= \tfrac{2\pm\sqrt{2^2-4(1)(-1)}}{2} \\ c &= 1 - \sqrt{2} \in [-1, 0] \end{aligned}$$

   (b) Since f is continuous on $[6, 8]$; solve

$$\begin{aligned} \sqrt{x^2 - 5x - 2} &= 3 \\ x^2 - 5x - 11 &= 0 \\ x &= \tfrac{5\pm\sqrt{5^2-4(1)(-11)}}{2} \\ c &= \tfrac{5+\sqrt{69}}{2} \in [6, 8] \end{aligned}$$

4. (a) $f'(x) = 2x - 3 = 0$, thus the critical points are $c = \pm 1$. Thus
   $min\{f(-1), f(1), f(2)\} = min\{5, -1, -1\} = -1$ and
   $max\{f(-1), f(1), f(2)\} = max\{5, -1, -1\} = 5$

   (b) $f'(x) = -2\cos(x)\sin(x) - \cos(x) = -\cos(x)(2\sin(x) + 1) = 0$, thus the critical points are $c = \pi, 7\pi/6, 11\pi/6$. Thus

$min\{f(0), f(\pi), f(7\pi/6), f(11\pi/6), f(2\pi)\} = min\{1, 1, 5/4, 5/4, 1\} =$ 1 and
$max\{f(0), f(\pi), f(7\pi/6), f(11\pi/6), f(2\pi)\} = max\{1, 1, 5/4, 5/4, 1\} =$ 5/4

5. (a) $f'(x) = 4x^3 - 8x = 4x(x^2 - 2) = 0$, thus $c = 0, \pm\sqrt{2} \in [-2, 2]$

   (b)

$$\begin{aligned} f'(x) &= \cos(x) + 2\cos(2x) \\ &= \cos(x) + 2(2\cos^2(x) - 1) \\ &= 4\cos^2(x) + \cos(x) - 2 \\ &= 0 \end{aligned}$$

$$\begin{aligned} x &= (-1 \pm \sqrt{33})/8 \\ c &= \cos^{-1}((-1 \pm \sqrt{33})/8), 2\pi - \cos^{-1}((-1 \pm \sqrt{33})/8) \end{aligned}$$

6. (a) $f'(x) = \frac{1}{2\sqrt{x}}$ and $\frac{f(4)-f(0)}{4-0} = \frac{1}{2}$. Solving $\frac{1}{2\sqrt{x}} = \frac{1}{2}$ yields $c = 1$.

   (b) $f'(x) = (x^2 + 2x)/(x + 1)^2$ and $\frac{f(1)-f(0)}{1-0} = \frac{1}{2}$. Solving $f'(x) = (x^2 + 2x)/(x + 1)^2 = \frac{1}{2}$ yields $c = -1 + \sqrt{2}$

7. The given function satisfies the hypotheses of the Generalized Rolle's Theorem. Since $f(0) = f(1) = f(3) = 0$, there exista a $c \in (0, 3)$ such that $f''(c) = 0$. Solve $6c - 8 = 0$ to find $c = 4/3$.

8. (a) $\int_0^2 xe^x dx = xe^x - e^x|_0^2 = e^2 + 1$

   (b) $\int_{-1}^1 \frac{3x}{x^2+1} dx = \frac{3}{2}\ln(x^2 + 1)|_{-1}^1 = 0$ (The integrand is an odd function)

9. (a) $\frac{d}{dx}\int_0^x t^2\cos(t)dt = x^2\cos(x)$

   (b) $\frac{d}{dx}\int_1^{x^3} e^{t^2} dt = e^{(x^3)^2}(3x^2) = 3x^2e^{x^6}$

10. (a) $\frac{1}{4-(-3)}\int_{-3}^4 6x^2 dx = \frac{2}{7}x^3|_{-3}^4 = 52$. Solving $6x^2 = 52$ yields $c = \pm\sqrt{26/3} \in [-3, 4]$.

    (b) $\frac{2}{3\pi}\int_0^{3\pi/2} x\cos(x)dx = \frac{2}{3\pi}(x\sin(x) + \cos(x))_0^{3\pi/2} = -(1 + \frac{2}{3\pi})$. Use a calculator to approximate the solution(s): $x\cos(x) = -(1 + \frac{2}{3\pi})$; $c \approx 2.16506, 4.43558 \in [0, 3\pi/2]$.

11. (a) $\frac{1}{1-\frac{1}{2}} = 2$

    (b) $\frac{1}{1-\frac{2}{3}} = 3$

    (c) $\sum_{n=1}^\infty \frac{3}{n(n+1)} = 3\sum_{n=1}^\infty(\frac{1}{n} - \frac{1}{n+1}) = 3\lim_{k\to\infty}\sum_{n=1}^k(\frac{1}{n} - \frac{1}{n+1}) = 3$

    (d) $\sum_{k=1}^\infty \frac{1}{4k^2-1} = \frac{1}{2}\sum_{k=1}^\infty(\frac{1}{2k-1} - \frac{1}{2k+1})$
    $= \frac{1}{2}\lim_{n\to\infty}\sum_{k=1}^n(\frac{1}{2k-1} - \frac{1}{2k+1}) = \frac{1}{2}$

12. (a) $-\frac{5}{128}(x-1)^4 + \frac{1}{16}(x-1)^3 - \frac{1}{8}(x-1)^2 + \frac{1}{2}(x-1) + 1$
    (b) $4x^2 + 3x + 1$
    (c) $\frac{1}{24}x^4 - \frac{1}{2}x^2 + 1$

13. The Taylor polynomial of degree $n = 4$ expanded about $x_0 = 0$ for $f(x) = \sin(x)$ is $P(x)$.

14. (a) $P(3) = -24$
    (b) $P(-1) = 20$

15. The average area is given by; $\frac{1}{3-1}\int_1^3 \pi r^2 dr = \frac{\pi}{2}(\frac{r^3}{3})_1^3 = \frac{13\pi}{3}$.

16. Any polynomial $P(x)$ satisfies the hypotheses of Rolle's Theorem on the interval $[a, b]$. Thus $P'(x)$ has at least $n-1$ real roots in the interval $[a, b]$, $P''(x)$ has at least $n-2$ real roots in the interval $[a, b], \ldots,$ and $P^{(n-1)}$ has at least $n - (n-1) = 1$ real root in the interval $[a, b]$.

17. If f, f' and f'' are defined on the interval $[a, b]$, then f is continuous on the interval $[a, b]$ and f is differentiable on the interval (a, b). By Theorem 1.6 (Mean Value Theorem) there exists numbers $c_1 \in (a, c)$ and $c_2 \in (c, b)$ such that:

$$f'(c_1) = \frac{f(c) - f(a)}{c - a} \text{ and } f'(c_2) = \frac{f(b) - f(c)}{b - c}$$

. But, since $f(a) = f(b) = 0$ it follows that $f'(c_1) = f(c)/(c-a)$ and $f'(c_2) = f(c)/(c-b)$. Given that f' and f'' are defined in the interval $[a, b]$, it follows that f' also satisfies the hypotheses of Theorem 1.6. Thus there exists a number $d \in (a, b)$ such that:

$$f''(d) = \frac{f'(c_2) - f'(c_1)}{c_2 - c_1} = \frac{\frac{f(c)}{c-b} - \frac{f(c)}{c-a}}{c_2 - c_1} = \frac{f(c)(b-a)}{(c_2 - c_1)(c-b)(c-a)} < 0,$$

since $f(c) > 0$.

1.2 Binary Numbers

1. Answers will depend on specific platform.

2. (a) 21 (b) 56 (c) 254 (d) 519

3. (a) 0.75 (b) 0.65625 (c) 0.6640625 (d) 0.85546875

4. (a) 1.4140625 (b) 3.1416015625

5. (a) $\sqrt{2} - 1.4140625 = 0.00015109\ldots$
   (b) $\pi - 3.1416015625 = -0.000008908\ldots$

6. (a) $23 = 10111_{two}$

$$\begin{aligned} 23 &= 2(11)+1 & b_0 = 1 \\ 11 &= 2(5)+1 & b_1 = 1 \\ 5 &= 2(2)+1 & b_2 = 0 \\ 2 &= 2(1)+0 & b_3 = 0 \\ 1 &= 2(0)+1 & b_4 = 1 \end{aligned}$$

(b) $87 = 1010111_{two}$

$$\begin{aligned} 87 &= 2(43)+1 & b_0 = 1 \\ 43 &= 2(21)+1 & b_1 = 1 \\ 21 &= 2(10)+1 & b_2 = 1 \\ 10 &= 2(5)+0 & b_3 = 0 \\ 5 &= 2(2)+1 & b_4 = 1 \\ 2 &= 2(1)+0 & b_5 = 0 \\ 1 &= 2(0)+1 & b_6 = 1 \end{aligned}$$

(c) $378 = 101111010_{two}$

(d) $2388 = 100101010100_{two}$

7. (a) 0.0111_{two} (b) 0.1101_{two} (c) 0.10111_{two} (d) 0.1001011_{two}

8. (a) $0.0\overline{0011}_{two}$

(b) $\frac{1}{3} = 0.\overline{d_1 d_2}_{two} = 0.\overline{01}_{two}$

$$\begin{array}{lll} 2R = \frac{2}{3} & d_1 = 0 = INT(\frac{2}{3}) & F_1 = \frac{2}{3} = FRAC(\frac{2}{3}) \\ 2F_1 = \frac{4}{3} & d_2 = 1 = INT(\frac{4}{3}) & F_2 = \frac{1}{3} = FRAC(\frac{4}{3}) \\ 2F_2 = \frac{2}{3} & d_3 = 0 = INT(\frac{2}{3}) & F_3 = \frac{2}{3} = FRAC(\frac{2}{3}) \\ 2F_3 = \frac{4}{3} & d_4 = 1 = INT(\frac{4}{3}) & F_4 = \frac{1}{3} = FRAC(\frac{4}{3}) \\ \vdots & \vdots & \vdots \end{array}$$

(c) $\frac{1}{7} = 0.\overline{d_1 d_2 d_3}_{two} = 0.\overline{001}_{two}$

$$\begin{array}{lll} 2R = \frac{2}{7} & d_1 = 0 = INT(\frac{2}{7}) & F_1 = \frac{2}{7} = FRAC(\frac{2}{7}) \\ 2F_1 = \frac{4}{7} & d_2 = 0 = INT(\frac{4}{7}) & F_2 = \frac{4}{7} = FRAC(\frac{4}{7}) \\ 2F_2 = \frac{8}{7} & d_3 = 1 = INT(\frac{8}{7}) & F_3 = \frac{1}{7} = FRAC(\frac{8}{7}) \\ 2F_3 = \frac{2}{7} & d_4 = 0 = INT(\frac{2}{7}) & F_4 = \frac{2}{7} = FRAC(\frac{2}{7}) \\ \vdots & \vdots & \vdots \end{array}$$

9. (a)

$$\begin{aligned} \tfrac{1}{10} - 0.0001100_{two} &= 0.0\overline{0011}_{two} - 0.0001100_{two} \\ &= 0.00000000\overline{1100}_{two} \\ &= \tfrac{1}{160} \\ &= 0.00625 \end{aligned}$$

(b)

$$\begin{aligned}
\tfrac{1}{7} - 0.0010010_{two} &= 0.\overline{001}_{two} - 0.0010010_{two} \\
&= 0.000000001\overline{001}_{two} \\
&= \tfrac{1}{448} \\
&= 0.0022321428\ldots
\end{aligned}$$

10. In Theorem 1.14 let $c = \frac{1}{8}$ and $r = \frac{1}{8}$, then

$$\frac{1}{8} + \frac{1}{64} + \frac{1}{512} + \cdots = \frac{\frac{1}{8}}{1 - \frac{1}{8}} = \frac{1}{7}$$

11. In Theorem 1.14 let $c = 3/16$ and $r = 1/16$, then

$$\frac{3}{16} + \frac{3}{256} + \frac{3}{4096} + \cdots = \frac{\frac{3}{16}}{1 - \frac{1}{16}} = \frac{1}{5}$$

12. $\frac{1}{2} = \frac{5}{10}$. Assume $\left(\frac{1}{2}\right)^k = \frac{5^k}{10^k}$. Then

$$\begin{aligned}
\left(\frac{1}{2}\right)^{k+1} &= \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right) \\
&= \left(\frac{5^k}{10^k}\right)\left(\frac{5}{10}\right) \\
&= \frac{5^{k+1}}{10^{k+1}}
\end{aligned}$$

Therefore, by the principle of mathematical induction, 2^{-N} can be represented as a decimal number that has N digits.

13. (a)

$$\begin{array}{cclcl}
\frac{1}{3} & \approx & 0.1011_{two} \times 2^{-1} & = & 0.1011 \times 2^{-1} \\
\frac{1}{5} & \approx & 0.1101_{two} \times 2^{-2} & = & 0.01101 \times 2^{-1} \\
\hline
\frac{8}{15} & & & & 0.100011_{two} \times 2^{0}
\end{array}$$

$$\begin{array}{cclcl}
\frac{8}{15} & \approx & 0.1001_{two} \times 2^{0} & = & 0.1001_{two} \times 2^{0} \\
\frac{1}{6} & \approx & 0.1011_{two} \times 2^{-2} & = & 0.001011_{two} \times 2^{0} \\
\hline
\frac{7}{10} & & & & 0.101111_{two} \times 2^{0}
\end{array}$$

Thus $(\frac{1}{3} + \frac{1}{5}) + \frac{1}{6} \approx 0.1100_{two}$

(b)

$$\begin{array}{cclcl}
\frac{1}{10} & \approx & 0.1101_{two} \times 2^{-3} & = & 0.001101 \times 2^{-1} \\
\frac{1}{3} & \approx & 0.1011_{two} \times 2^{-1} & = & 0.101100 \times 2^{-1} \\
\hline
\frac{13}{30} & & & & 0.111001_{two} \times 2^{-1}
\end{array}$$

$$\begin{array}{rclcl}
\frac{13}{30} & \approx & 0.1110_{two} \times 2^{-1} & = & 0.011100_{two} \times 2^{0} \\
\frac{1}{5} & \approx & 0.1101_{two} \times 2^{-2} & = & 0.001101_{two} \times 2^{0} \\
\hline
\frac{19}{30} & & & & 0.101001_{two} \times 2^{0}
\end{array}$$

Thus $(\frac{1}{10} + \frac{1}{3}) + \frac{1}{5} \approx 0.1010_{two}$

(c)

$$\begin{array}{rclcl}
\frac{3}{17} & \approx & 0.1011_{two} \times 2^{-2} & = & 0.01011 \times 2^{-1} \\
\frac{1}{9} & \approx & 0.1110_{two} \times 2^{-3} & = & 0.001110 \times 2^{-1} \\
\hline
\frac{44}{153} & & & & 0.100100_{two} \times 2^{-1}
\end{array}$$

$$\begin{array}{rclcl}
\frac{44}{153} & \approx & 0.1001_{two} \times 2^{-1} & = & 0.1001_{two} \times 2^{-1} \\
\frac{1}{7} & \approx & 0.1001_{two} \times 2^{-2} & = & 0.01001_{two} \times 2^{-1} \\
\hline
\frac{461}{1071} & & & & 0.11011_{two} \times 2^{-1}
\end{array}$$

Thus $(\frac{3}{17} + \frac{1}{9}) + \frac{1}{7} \approx 0.1110_{two} \times 2^{-1}$

(d)

$$\begin{array}{rclcl}
\frac{7}{10} & \approx & 0.1011_{two} \times 2^{0} & = & 0.1011000 \times 2^{0} \\
\frac{1}{9} & \approx & 0.1110_{two} \times 2^{-3} & = & 0.0001110 \times 2^{0} \\
\hline
\frac{73}{90} & & & & 0.1100110_{two} \times 2^{0}
\end{array}$$

$$\begin{array}{rclcl}
\frac{73}{90} & \approx & 0.1101_{two} \times 2^{0} & = & 0.110100_{two} \times 2^{0} \\
\frac{1}{7} & \approx & 0.1011_{two} \times 2^{-2} & = & 0.001001_{two} \times 2^{0} \\
\hline
\frac{601}{630} & & & & 0.111101_{two} \times 2^{0}
\end{array}$$

Thus $(\frac{7}{10} + \frac{1}{9}) + \frac{1}{7} \approx 0.1111_{two}$

14. (a) $10 = 101_{three}$
   (b) $23 = 212_{three}$
   (c) $421 = 120121_{three}$
   (d) $1784 = 211002_{three}$

15. (a) $\frac{1}{3} = 0.1_{three}$
   (b) $\frac{1}{2} = 0.\overline{1}_{three}$
   (c) $\frac{1}{10} = 0.\overline{0022}_{three}$
   (d) $\frac{11}{27} = 0.102_{three}$

16. (a) (a) $10 = 20_{five}$
   (b) (b) $35 = 120_{five}$
   (c) (c) $721 = 1034_{five}$
   (d) (d) $734 = 10414_{five}$

17. (a) $\frac{1}{3} = 0.\overline{13}_{five}$
   (b) $\frac{1}{2} = 0.\overline{2}_{five}$
   (c) $\frac{1}{10} = 0.0\overline{2}_{five}$
   (d) $\frac{154}{625} = 0.1104_{five}$

1.3 Error Analysis

1. (a) $x - \hat{x} = 0.00008182$, $\frac{x-\hat{x}}{x} = 0.0000300998\ldots$, 4- significant digits

   (b) $y - \hat{y} = 350$, $\frac{y-\hat{y}}{y} = 0.0355871\ldots$, 2-significant digits

   (c) $z = \hat{z} = 0.000008$, $\frac{z-\hat{z}}{z} = 0.117647$, 0-significant digits

2.

$$\begin{aligned}
\int_0^{1/4} e^{x^2} dx &\approx \int_0^{1/4} (1 + x^2 + \tfrac{x^4}{3} + \tfrac{x^6}{3!}) dx \\
&= (x + \tfrac{x^3}{3} + \tfrac{x^5}{5(2!)} + \tfrac{x^7}{7(3!)})_{x=0}^{x=1/4} \\
&= \tfrac{1}{4} + \tfrac{1}{192} + \tfrac{1}{10240} + \tfrac{1}{688128} \\
&= \tfrac{292807}{1146880} \approx 0.2553074428 = \hat{p}
\end{aligned}$$

3. (a) $p_1 + p_2 = 1.414 + 0.09125 = 1.505$
   $p_1 p_2 = (2.1414)(0.09125) = 0.1290$

   (b) $p_1 + p_2 = 31.415 + 0.027182 = 31.442$
   $p_1 p_2 = (31.415)(0.27182) = 0.85392$

4. (a) $\frac{0.70711385222 - 0.70710678110}{0.00001} = \frac{0.00000707103}{0.00001} = 0.707103$ The error involves loss of significance.

   (b) $\frac{0.69317218025 - 0.6931478056}{0.00005} = \frac{0.00002499969}{0.00005} = 0.4999938$ The error involves loss of significance.

5. (a) $ln(\frac{x+1}{x})$

   (b) $\frac{1}{\sqrt{x^2+1}+x}$

   (c) $cos(2x)$

   (d) $cos(x/2)$

6. (a)

$$\begin{aligned}
P(2.72) &= (2.72)^3 - 3(2.72)^2 + 3(2.72) - 1 \\
&= 20.12 - 3(7.398) + 8.16 - 1 \\
&= 20.12 - 22.19 + 8.16 - 1 \\
&= 5.09
\end{aligned}$$

$$\begin{aligned}
Q(2.72) &= ((2.72 - 3)(2.72) + 3)(2.72) - 1 \\
&= (-0.2800)(2.72) + 3)(2.72) - 1 \\
&= (-0.7616 + 3)(2.72) - 1 \\
&= (2.2384)(2.72) - 1 \\
&= 6.088 - 1 \\
&= 5.088
\end{aligned}$$

$$\begin{aligned}
R(2.72) &= (2.72 - 1)^3 \\
&= (1.72)^3 \\
&= 5.088
\end{aligned}$$

(b)

$$\begin{aligned}
P(0.975) &= (((0.975)^3 - 3(0.975)^2) + 3(0.975)) - 1 \\
&= ((0.9268 - 3(0.9506)) + 2.925) - 1 \\
&= ((0.9268 - 2.852) + 2.925) + 1 \\
&= (-1.925 + 2.925) - 1 \\
&= 1 - 1 = 0
\end{aligned}$$

$$\begin{aligned}
Q(0.975) &= ((0.975 - 3)(0.975) + 3)(0.975) - 1 \\
&= ((-2.025)(0.975) + 3)(0.975) - 1 \\
&= (-1.9774 + 3)(0.975) - 1 \\
&= (1.026)(0.975) - 1 \\
&= 1 - 1 = 0
\end{aligned}$$

$$\begin{aligned}
R(0.975) &= (0.975 - 1)^3 \\
&= (-0.025)^3 \\
&= -0.00001562
\end{aligned}$$

7. (a) $\frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81} + \frac{1}{243} + \frac{1}{729} \approx 0.498$

   (b) $\frac{1}{729} + \frac{1}{243} + \frac{1}{81} + \frac{1}{27} + \frac{1}{9} + \frac{1}{3} \approx 0.499$

8. (a) The propagation of error is $\epsilon_p + \epsilon_q + \epsilon_r$.

   (b)

$$\frac{p}{q} = \frac{\hat{p} + \epsilon_p}{\hat{q} + \epsilon_q} = \frac{\hat{p}}{\hat{q}} + \frac{\epsilon_p + \frac{\hat{p}}{\hat{q}}\epsilon_q}{\hat{q} + \epsilon_q}$$

   Hence, if $1 < |\hat{q}| < |\hat{p}|$, then there is a possibility of magnification of the original error.

   (c)

$$\begin{aligned}
pqr &= (\hat{p} + \epsilon_p)(\hat{q} + \epsilon_q)(\hat{r} + \epsilon_r) \\
&= \hat{p}\hat{q}\hat{r} + \hat{p}\hat{r}\epsilon_q + \hat{q}\hat{r}\epsilon_p + \hat{p}\hat{q}\epsilon_r + \hat{r}\epsilon_p\epsilon_q + \hat{q}\epsilon_p\epsilon_r + \hat{p}\epsilon_q\epsilon_r + \epsilon_p\epsilon_q\epsilon_r \\
&= \hat{p}\hat{q}\hat{r} + (\hat{p}\hat{r}\epsilon_q + \hat{q}\hat{r}\epsilon_p + \hat{p}\hat{q}\epsilon_r) \\
&\quad + (\hat{r}\epsilon_p\epsilon_q + \hat{q}\epsilon_p\epsilon_r + \hat{p}\epsilon_q\epsilon_r) + \epsilon_p\epsilon_q\epsilon_r
\end{aligned}$$

   Depending on the absolute values of $\hat{p}$, $\hat{q}$, and $\hat{r}$, there is a possibility of magnification of the original errors ϵ_p, ϵ_q, and ϵ_r.

9.

$$\begin{aligned}
\tfrac{1}{1-h} + \cos(h) &= 2 + h + \tfrac{h^2}{2} + h^3 + \boldsymbol{O}(h^4) \\
(\tfrac{1}{1-h})\cos(h) &= 1 + h + \tfrac{h^2}{2} + \tfrac{h^3}{2} + \boldsymbol{O}(h^4)
\end{aligned}$$

10.

$$\begin{aligned} e^h + \sin(h) &= 1 + 2h + \tfrac{h^2}{2} + \boldsymbol{O}(h^4) \\ e^h sin(h) &= h + h^2 + \tfrac{h^3}{3} + \boldsymbol{O}(h^5) \end{aligned}$$

An intermediate computation was

$$(1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \frac{h^4}{4!})(h - \frac{h^3}{3!}) = h + h^2 + \frac{h^3}{3} - \frac{h^5}{24} - \frac{h^6}{36} - \frac{h^7}{144}$$

11.

$$\begin{aligned} \cos(h) + \sin(h) &= 1 + h - \tfrac{h^2}{2} - \tfrac{h^3}{6} + \tfrac{h^4}{24} + \boldsymbol{O}(h^5) \\ \cos(h)\sin(h) &= h - \tfrac{2h^3}{3} + \tfrac{2h^5}{15} + \boldsymbol{O}(h^7) \end{aligned}$$

An intermediate comutation was

$$(1 - \frac{h^2}{2} + \frac{h^4}{24})(h - \frac{h^3}{6} + \frac{h^5}{120}) = h - \frac{2h^3}{3} + \frac{2h^5}{15} - \frac{h^7}{90} + \frac{h^9}{2880}$$

12.

$$\begin{aligned} x_1 &= \tfrac{-b+\sqrt{b^2-4ac}}{2a} \\ &= (\tfrac{-b+\sqrt{b^2-4ac}}{2a})(\tfrac{-b-\sqrt{b^2-4ac}}{-b-\sqrt{b^2-4ac}}) \\ &= \tfrac{b^2-(b^2-4ac)}{2a(-b-\sqrt{b^2-4ac})} \\ &= \tfrac{-2c}{b+\sqrt{b^2-4ac}} \end{aligned}$$

The case for x_2 is handled in a similar manner.

13. (a) $x_1 = -0.001000$, $x_2 = -1000$
    (b) $x_1 = -0.00100$, $x_2 = -10000$
    (c) $x_1 = -0.000010$, $x_2 = -100000$
    (d) $x_1 = -0.000001$, $x_2 = -1000000$

Chapter 2

The Solution of Nonlinear Equations $f(x) = 0$

2.1 Iteration for Solving $x = g(x)$

1. (a) Clearly, $g(x) \in C[0,1]$. Since $g'(x) = -x/2 < 0$ on the interval $[0,1]$, the function $g(x)$ is strictly decreasing on the interval $[0,1]$. If g is strictly decreasing on $[0,1]$, then $g(0) = 1$ and $g(1) = 0$ imply that $g([0,1]) = [0,1] \subseteq [0,1]$. Thus, by Theorem 2.2, the function $g(x)$ has a fixed point on the interval $[0,1]$.

   In addition: $|f'(x)| = |-x/2| = x/2 \leq 1/2 < 1$ on the interval $[0,1]$. Thus, by Theorem 2.2, the function $g(x)$ has a unique fixed point on the interval $[0,1]$.

   (b) Clearly, $g(x) \in C[0,1]$. Since $g'(x) = -\ln(2)2^{-x} < 0$ on the interval $[0,1]$, the function $g(x)$ is strictly decreasing on the interval $[0,1]$. If g is strictly decreasing on $[0,1]$, then $g(0) = 1$ and $g(1) = 1/2$ imply that $g([0,1]) = [1/2,1] \subseteq [0,1]$. Thus, by Theorem 2.2,the function $g(x)$ has a fixed point on the interval $[0,1]$.

   In addition: $|g'(x)| = |-\ln(2)2^{-2}| = \ln(2)2^{-2} \leq \ln(2) < \ln(e) = 1$ on the interval $[0,1]$. Thus, by Theorem 2.2, the function $g(x)$ has an unique fixed point on the interval $[0,1]$.

   (c) Clearly $g(x)$ is continuous on $[0.5, 5.2]$ and $g([0.5, 5.2]) \not\subseteq [0.5, 5.2]$. But, $g([0.5, 2]) \subseteq [0.5, 2]$. Thus, the hypotheses of the first part of Theorem 2.2 are satisfied and g has a fixed point in $[0.5, 2]$. While $(1, 1)$ is the unique fixed point in $[0.5, 2]$, $|f'(1)| = 1 \not< 1$, thus the hypothesee in part (4) of Theorem 2.2 cannot be satisfied.

2. (a)

$$\begin{aligned} g(x) &= x \\ -4 + 4x - \tfrac{1}{2}x^2 &= 0 \\ x^2 - 6x + 8 &= 0 \\ x &= 2, 4 \end{aligned}$$

and

$$\begin{aligned} g(2) &= -4 + 8 - 2 &= 2 \\ g(4) &= -4 + 16 - 8 &= 4 \end{aligned}$$

(b)

$$\begin{aligned} p_0 &= 1.9 \\ p_1 &= 1.795 \\ p_2 &= 1.5689875 \\ p_3 &= 1.04508911 \end{aligned}$$

(c)

$$\begin{aligned} p_0 &= 3.8 \\ p_1 &= 3.98 \\ p_2 &= 3.9998 \\ p_3 &= 3.99999998 \end{aligned}$$

(d) For part (b)

$$\begin{aligned} E_0 &= 0.1 & R_0 &= 0.95 \\ E_1 &= 0.205 & R_1 &= 0.1025 \\ E_2 &= 0.43110125 & R_2 &= 0.21550625 \\ E_3 &= 0.95491089 & R_3 &= 0.477455444 \end{aligned}$$

(e) The sequence in part (b) does not converge to $P = 2$. The sequence in part (c) converges to $P = 4$.

3. (a) $p_1 = \sqrt{13}$, $p_2 = \sqrt{6 + \sqrt{13}}$, converges

   (b) $p_1 = \frac{3}{2}$, $p_2 = \frac{7}{3}$, converges

   (c) $p_1 = 4.083333$, $p_2 = 5.537869$, diverges

   (d) $p_1 = -5.5$, $p_2 = -69.5$, diverges

4. The fixed points are $P = 2$ and $P = -2$. Since $g'(2) = 5$ and $g'(-2) = -3$, fixed- point iteration will not converge to $P = 2$ and $P = -2$, respectively.

5.

$$\begin{aligned} x &= x\cos(x) \\ x(1 - \cos(x)) &= 0 \\ x &= 2n\pi \end{aligned}$$

Thus $g(x)$ has infinitely many fixed points: $P = 2n\pi$, where $n \in Z$. Note:

$$|g'(2n\pi)| \quad = \quad |\cos(2n) - 2n\pi\ in(2n\pi)| \quad = \quad 1.$$

Thus Theorem 2.3 may not be used to find the fixed points of $g(x)$.

6. $|p_2 - p_1| = |g(p_1) - g(p_0)| = |g'(c_0)(p_1 - p_0)| < K|p_1 - p_2|$
7. $|E_1| = |P - p_1| = |g(P) - g(p_0)| = |g'(c_0)(P - p_0)| > |P - p_0| = |E_0|$
8. (a) By way of contradiction assume there exists k such that $p_{k+1} = g(p_k) \geq p_k$. It follows that:

$$\begin{aligned} -0.0001p_k^2 + p_k &\geq p_k \\ -0.0001p_k^2 &\geq 0 \\ p_k &= 0 \end{aligned}$$

Thus $p_{k-1} = 0$ or $p_{k-1} = 10,000$. Clearly, $p_{k-1} \neq 10,000$, since the maximum value of $g(x)$ is 2500. Thus, if $p_k = 0$, then $p_{k-1}, \ldots, p_1 = 0$. A contradiction to the hypothesis $p_0 = 1$. Therefore, $p_0 > p_1 > \cdots > p_n > p_{n+1} > \cdots$.

(b) By way of contradiction assume there exists k such that $p_j \leq 0$. It follows that:

$$\begin{aligned} g(p_{j-1}) &\leq 0 \\ -0.0001p_{j-1}^2 + p_{j-1} &\leq 0 \\ (-0.0001p_{j-1} + 1)\,p_{j-1} &\leq 0 \end{aligned}$$

From part (a); if $p_{j-1} = 0$, then $p_1 \neq 0$. Thus $p_{j-1} \neq 0$. If $p_{j-1} < 0$, then

$$\begin{aligned} -0.0001p_{j-1} + 1 &\geq 0 \\ p_{j-1} &\geq 10,000, \end{aligned}$$

a contradiction. If $p_{j-1} > 0$, then

$$\begin{aligned} -0.0001p_{j-1} + 1 &\geq 0 \\ p_{j-1} &\leq 10,000, \end{aligned}$$

a contradiction. Therefore, $p_n > 0$ for all n.

(c) $\lim_{n\to\infty} p_n = 0$

9. (a) $g(3) = (0.5)(3) + 1.5 = 3$

(b) $|P - p_n| = |3 - 1.5 - 0.5p_{n-1}| = |1.5 - 0.5p_{n-1}| = \frac{1}{2}|3 - p_{n-1}| = \frac{1}{2}|P - p_{n-1}|$

(c) Using mathematical induction we note that $|P - p_1| = \frac{1}{2}|P - p_0|$ and assume that $|P - p_k| = \frac{1}{2^k}|P - p_0|$. Thus

$$\begin{aligned} |P - p_{k+1}| &= \frac{|P-p_k|}{2} \\ &= \frac{|P-p_0|}{2(2^k)} \\ &= \frac{|P-p_0|}{2^{k+1}} \end{aligned}$$

10. (a) Note: $p_1 = p_0/2$, $p_2 = p_0/2^2, \ldots, p_{k+1} = p_0/2^{k+1}, \ldots$. Thus

$$\frac{|p_{k+1} - p_k|}{|p_{k+1}|} = \frac{|2^{-k-1} - 2^{-k}|}{|2^{-k-1}|} = \frac{2^{-k}(1 - 2^{-1})}{2^{-k}2^{-1}} = 1$$

(b) Clearly, the stopping criteria will (theoretically) never be satisfied.

11. In inequality (11): $|P - p_n| \le K^n|P - p_0|$, where $|g'(x)| \le K < 1$. Therefore, the smaller the value of K the faster fixed-point iteration converges.

2.2 Bracketing Methods for Locating a Root

1.

$$\begin{aligned} I_0 &= (0.11 + 0.12)/2 = 0.115 & A(0.115) &= 254,403 \\ I_1 &= (0.11 + 0.115)/2 = 0.1125 & A(0.1125) &= 246,072 \\ I_2 &= (0.1125 + .125)/2 = 0.11375 & A(0.11375) &= 250,198 \end{aligned}$$

2.

$$\begin{aligned} I_0 &= (0.13 + 0.14)/2 = 0.135 & A(0.135) &= 394,539 \\ I_1 &= (0.135 + 0.14)/2 = 0.1375 & A(0.1375) &= 408,435 \\ I_2 &= (0.135 + 0.1375)/2 = 0.13625 & A(0.13625) &= 401,420 \end{aligned}$$

3. (a) $f(-3) > 0$, $f(0) < 0$, and $f(3) > 0$; thus roots lie in the intervals $[-3, 0]$ and $[0, 3]$.

(b) $f(\pi/4) > 0$ and $f(\pi/2) < 0$; thus a root lies in the interval $[\pi/4, \pi/2]$.

(c) $f(3) < 0$ and $f(5) > 0$; thus a root lies in the interval $[3, 5]$.

(d) $f(3) > 0$, $f(5) < 0$, and $f(7) > 0$; thus roots lie in the intervals $[3, 5]$ and $[5, 7]$.

4. $[-2.4, -1.6], [-2.0, -1.6], [-2.0, -1.8], [-1.9, -1.8], [-1.85, -1.80]$

5. $[0.8, 1.6], [1.2, 1.6], [1.2, 1.4], [1.2, 1.3], [1.25, 1.30]$

6. $[3.2, 4.0], [3.6, 4.0], [3.6, 3.8], [3.6, 3.7], [3.65, 3.70]$

7. $[6.0, 6.8], [6.4, 6.8], [6.4, 6.6], [6.5, 6.5], [6.40, 6.45]$

8. (a) Starting with $a_0 < b_0$, then either $a_1 = a_0$ and $b_1 = \frac{a_0+b_0}{2}$, or $a_1 = \frac{a_0+b_0}{2}$ and $b_1 = b_0$. In either case we have $a_0 \le a_1 < b_1 \le b_0$. Now assume that the result is true for $n = 1, 2, \ldots, k$; in particular $a_0 \le a_1 \le \cdots \le a_k < b_k \le \cdots \le b_1 \le b_0$. Then either $a_{k+1} = a_k$ and $b_{k+1} = \frac{a_k+b_k}{2}$, or $a_{k+1} = \frac{a_k+b_k}{2}$ and $b_{k+1} = b_k$. In either case we have $a_k \le a_{k+1} < b_{k+1} \le b_k$. Hence $a_0 \le a_1 \le \cdots \le a_k \le a_{k+1} < b_{k+1} \le b_k \le \cdots \le b_1 \le b_0$. Thus by mathematical induction we have proven that $a_0 \le a_1 \le \cdots \le a_n < b_n \le \cdots \le b_1 \le b_0$ for all n.

   (b) From part (a) either $a_1 = a_0$, $b_1 = \frac{a_0+b_0}{2}$, and $b_1 = a_1 = \frac{b_0-a_0}{2}$ or $a_1 = \frac{a_0+b_0}{2}$, $b_1 = b_0$, and $b_1 - a_1 = \frac{b_0-a_0}{2}$. Now assume that the result is true for $n = 1, 2, \ldots, k$, in particular $b_k - a_k = \frac{b_0 - a_0}{2^k}$. Then either $a_{k+1} = a_k$, $b_{k+1} = \frac{a_k+b_k}{2}$, and $b_{k+1} - a_{k+1} = \frac{b_k-a_k}{2} = \frac{b_0-a_0}{2^{k+1}}$ or $a_{k+1} = \frac{a_k+b_k}{2}$, $b_{k+1} = b_k$, and $b_{k+1} - a_{k+1} = \frac{b_k-a_k}{2} = \frac{b_0-a_0}{2^{k+1}}$. Thus by mathematical induction we have proven that $b_n - a_n = \frac{b_0-a_0}{2^n}$ for all n.

   (c) Using part (c) it follows that the sequence $\{a_n\}$ is non-decreasing and bounded above by b_0, hence it is a convergent sequence and we write $\lim_{n\to\infty} a_n = L_1$. Similarly, the sequence $\{b_n\}$ is non-increasing and bounded below by a_0, hence it is a convergent sequence and we write $\lim_{n\to\infty} b_n = L_2$

   To show that the two limits are equal we observe that

$$\begin{aligned} L_2 &= \lim_{n\to\infty} b_n \\ &= \lim_{n\to\infty}(a_n + (b_n - a_n)) \\ &= \lim_{n\to\infty} a_n + \lim_{n\to\infty}(b_n - a_n) \\ &= L_1 + \lim_{n\to\infty} \frac{b_0-a_0}{2^n} \\ &= L_1 + 0 = L_1 \end{aligned}$$

   Since $a_n \le c_n \le b_n$ the squeeze principle for limits implies that

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n \lim_{n\to\infty} b_n$$

9. (a) The function does not change sign on the interval $[3, 7]$.

   (b) $\lim_{n\to\infty} a_n = 2 = \lim_{n\to\infty} b_n$, but $f(x)$ is undefined at 2.

10. (a) It will converge to the zero at $x = \pi$.

    (b) $\lim_{n\to\infty} a_n = \pi/2 = \lim_{n\to\infty} b_n$, but $f(x)$ is undefined at $\pi/2$.

11. Solve:

$$\begin{aligned} \frac{7-2}{2^N} &< 5 \times 10^{-9} \\ \ln(5) - N\ln(2) &< \ln(5 \times 10^{-9}) \\ N &> \frac{]ln(5) - \ln(5 \times 10^{-9})}{\ln(2)} \\ N &> 29.89735 \end{aligned}$$

Thus $N = 30$.

12.

$$\begin{aligned} c_n &= b_n - \frac{f(b_n)(b_n - a_n)}{f(b_n) - f(a_n)} \\ &= \frac{b_n(f(b_n) - f(a_n)) - f(b_n)(b_n - a_n)}{f(b_n) - f(a_n)} \\ &= \frac{-b_n f(a_n) + a_n f(b_n)}{f(b_n) - f(a_n)} \\ &= \frac{a_n f(b_n) - b_n f(a_n)}{f(b_n) - f(a_n)} \end{aligned}$$

13.

$$\begin{aligned} \frac{|b-a|}{2^{N+1}} &< \delta \\ \ln\left(\frac{|b-a|}{2^{N+1}}\right) &< \ln(\delta), \end{aligned}$$

since ln is a strictly increasing function. Thus

$$\begin{aligned} \ln(b-a) - (N+1)\ln(2) &< \ln(\delta) \\ \frac{\ln(b-a) - \ln(\delta)}{\ln(2)} &< N+1 \\ N &> \frac{\ln(b-a) - \ln(\delta)}{\ln(2)} - 1 \end{aligned}$$

Therefore, the smallest value of N is

$$N = int\left(\frac{\ln(b-a) - \ln(\delta)}{\ln(2)}\right)$$

14. The bisection method can't converge to $x = 2$, unless $c_n = 2$ for some $n \geq 1$.

15. We refer the reader to "Which Root Does the Bisection Algorithm Find?" by George Corliss, Mathematical Modeling: Classroom Notes in Applied Mathematics, Murray Klankin Ed., SIAM, 1987.

2.3 Initial Approximation and Convergence Criteria

1. Approximate root location -0.7. Computed root -0.7034674225.
2. Approximate root location 0.7. Computed root 0.7390851332.
3. Approximate root locations -1.0 and 0.6. Computed roots -1.002966954 and 0.6348668712.