Elementary Numerical Analysis Problems

Bent E. Petersen

Fall 2003 – 20031203

Contents

Note $\log x$ means the *natural logarithm* of x.

This list of problems will be augmented during the quarter. Many of the problems are from old lists of sample problems, old tests or old assignments. Thus there may be some duplication, or at least, near duplication.

If you notice any duplicate problems please let me know so I can remove the duplicates.

Additional problems will be added, and duplicates removed, in such a way that the (automatically generated) problem numbers will not change (I hope).

---

1 Binary representation of numbers

Problem 1.1: For this problem assume a machine using a normalized binary floating point representation for real numbers. Assume that the mantissa occupies 8 bits (a silly design), there is no packing trickery, chopping is used and enough bits are available for the exponent that we wont have to worry about overflow nor underflow. (A) What bit pattern would be stored in the mantissa to represent the decimal fraction $\frac{1}{7}$? (B) What is the exact decimal value of the floating point representation of $\frac{1}{7}$?

Problem 1.2: For this problem assume a machine using a normalized binary floating point representation for real numbers. Assume that the mantissa occupies 10 bits (a silly design), there is no packing trickery, chopping is used and enough bits are available for the exponent that we will not have to worry about overflow nor underflow (in this problem). What bit pattern would be stored in the mantissa to represent the decimal fraction $\frac{3}{13}$?

Problem 1.3: Use the rule of 1's to convert $0.000111\overline{000111}$ (binary) to a decimal fraction.

Problem 1.4: Use the rule of 1's to convert $0.10011\overline{0011}$ (binary) to a decimal fraction.

Problem 1.5: For this problem assume a machine using a normalized binary floating point representation for real numbers. Assume that the mantissa occupies 12 bits (not realistic!), there is no packing trickery, chopping is used and enough bits are available for the exponent that we will not have to worry about overflow nor underflow. **(A).** What bit pattern would be stored in the mantissa to represent the decimal fraction $\frac{4}{15}$? **(B).** What is the exact decimal value of normalized binary floating point representation of $\frac{4}{15}$? **(C).** What is the roundoff error?

Problem 1.6: For this problem assume a machine using a normalized binary floating point representation for real numbers. Assume that the mantissa occupies 10 bits (not realistic!), there is no packing trickery, chopping is used and enough bits are available for the exponent that we will not have to worry about overflow nor underflow. **(A).** What bit pattern would be stored in the mantissa to represent the decimal fraction $\frac{67}{160}$? **(B).** What is the exact decimal value of normalized binary floating point representation of $\frac{67}{160}$? **(C).** What is the roundoff error?

A real number x_a is said to have m significant decimal digits as an approximation to the real number $x_t \neq 0$ provided that

$$\left| \frac{x_t - x_a}{x_t} \right| \leq 5 \times 10^{-m-1}.$$

The best value of m is clearly given by

$$m = \text{floor}\left(-\log_{10} \left(2 \left| \frac{x_t - x_a}{x_t} \right| \right) \right).$$

Here floor( ) indicates the *greatest integer function*.

There are also various digit-counting schemes for determining the number of significant digits. In exceptional cases various methods yield slightly different results.

Problem 1.7: The fractions $\frac{22}{7}$, $\frac{333}{106}$ and $\frac{355}{113}$ are well-known approximations to π. Use the formula above to find the the number of significant decimal digits in each. Does your understanding of significant digits agree with your calculated result?

Problem 1.8: Convert $\frac{1}{3}$ into a normalized base 2 floating point value with an 8 bit mantissa. Use chopping. What is the error?

Problem 1.9: Let $a > 0$ be an integer. Show a requires

$$q = 1 + \text{floor}\left(\frac{\log a}{\log 2}\right)$$

bits for its binary representation. The base of the logarithm is immaterial, but base 2 leads to a nicer expression. The composition of floor and $\log_2$ is called the *integer logarithm base 2* and is denoted by ilog_2. Thus

$$q = 1 + \mathrm{ilog}_2(a).$$

Similarly the number of decimal digits required is

$$p = 1 + \mathrm{ilog}_{10}(a).$$

Deduce that for large a we have

$$\frac{p}{q} \sim \log_{10} 2.$$

Problem 1.10: How many binary bits in 1,048,543?

Problem 1.11: In the binary representation division by 2 amounts to shifting the binary point to the left. Thus the remainder obtained when dividing an integer a by 2 is the least significant binary digit in a. By applying the same process to the integer quotient of a divided by 2 we produce the next binary digit. Use this process to write 9,345 in binary.

Problem 1.12: In the binary representation multiplication by 2 amounts to shifting the binary point to the right. Thus the integer part obtained when multiplying a fraction a by 2 is the most significant binary digit in a. If we then toss the integer part and multiply the resulting fraction by 2 we obtain the next binary digit in a. Use this process to obtain the binary expansion of $\frac{1}{7}$.

Problem 1.13: Find the first 12 binary digits in π.

2 Polynomials

Problem 2.1: Write a brief description of HORNER's method (synthetic division). Be explicit. Use code to describe the algorithm if you prefer code to equations.

Problem 2.2: The polynomial $p(x) = x^4 + 4x^3 + 2x^2 - 5x - 2$ has roots -2 and 1. Find the remaining roots. (This is an algebra problem – not a numerical analysis problem).

3 TAYLOR polynomials

Problem 3.1: Suppose we approximate e^{2x} on the interval $[-1,1]$ by the the TAYLOR polynomial

$$p(x) = 1 + 2x + 2x^2 + \frac{4}{3}x^3.$$

Use the TAYLOR remainder to give a good upper bound for the truncation error that we make at $x = -\frac{1}{2}$ and at $x = \frac{1}{2}$.

Problem 3.2: Suppose we approximate $f(x) = \log(x+1)$ on the interval $(-1,1)$ by the the TAYLOR polynomial

$$p(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3.$$

Use the TAYLOR remainder to give a good upper bound for the absolute value of the *truncation* error, $|f(x) - p(x)|$, that we make at $x = -\frac{1}{2}$ and at $x = \frac{1}{2}$. Note 10^{97} is a correct upper bound, but it is not a *good* one.

Problem 3.3: Let $p(x)$ be the TAYLOR–MACLAURIN polynomial of degree 5 for the function $f(x) = \sin(x) + \exp(x)$ **(A)** Compute $p(x)$. **(B)** Suppose we use $p(x)$ to estimate $f(x)$ on the interval $[0,1]$. Use TAYLOR's remainder to bound the error in $p(1)$ and in $p(\frac{1}{2})$. Use the estimate $e \leq 3$ in computing your error bounds.

Problem 3.4: We know we can approximate $\exp(x)$ by the TAYLOR polynomial

$$\sum_{k=0}^{n} \frac{x^k}{k!}.$$

If we take $x = 1$ we obtain a way to approximate EULER's number $e = \exp(1)$. How large m<u>u</u>st we take n in order to compute e with an error no bigger than 8.0×10^{-6}.

Problem 3.5: We approximate a certain function $f(x)$ by an nth degree TAYLOR polynomial about the origin. If the absolute value of the kth derivative of $f(x)$ on the interval $[0,1]$ is bounded by 4^k for each k, how large whould we take n in order to estimate $f(0.5)$ with an error no larger that 0.2 in absolute value?

Problem 3.6: Suppose $f(0) = 0$, $f'(0) = 2$, $f''(0) = -\frac{1}{3}$ and $|f'''(x)| \leq 0.024$ for $0 \leq x \leq 2$. Estimate $f(1)$ by using a TAYLOR polynomial of degree 2. Compute a good bound for the absolute error.

Problem 3.7: Suppose $f(0) = 1$, $f^{(1)}(0) = -2$, $f^{(2)}(0) = 1$, $f^{(3)}(0) = -1$ and $|f^{(4)}(x)| \leq 0.384$ for $|x| \leq 1$. **(A)** Estimate $f(0.5)$ by using a TAYLOR polynomial of degree ≤ 3. **(B)** Compute a good bound for the absolute error.

Problem 3.8: Suppose $f(1) = 1$, $f^{(1)}(1) = -2$, $f^{(2)}(1) = 1$, $f^{(3)}(1) = -1$ and $f^{(4)}(1) = 24$. Estimate $f(1.5)$ by using a TAYLOR polynomial of degree ≤ 4.

Problem 3.9: Find the TAYLOR polynomial with center at 0 and degree 4 for the function f given by

$$f(x) = \mathrm{e}^{\cos\left(\frac{x}{2}\right) - 1 + \frac{1}{8}x^2}.$$

Given that $\left| f^{(5)}(x) \right| \le 0.005$ for $|x| \le 0.5$ estimate the error in approximating $f(0.5)$ by means of its TAYLOR polynomial of degree 4. Compare your estimate with the actual error.

Problem 3.10: Suppose $f(0) = 1$, $f'(0) = -2$, $f''(0) = -\frac{1}{5}$, $f'''(0) = \frac{2}{3}$ and $\left| f^{(4)}(x) \right| \le 0.016$ for $0 \le x \le 2$. Estimate $f(1)$ by using a TAYLOR polynomial of degree 3. Compute a good bound for the absolute error.

Problem 3.11: Suppose $f(1) = 1$, $f'(1) = 1$, $f''(1) = -2$, $f'''(1) = 4$. Estimate $f(1.25)$ by using a TAYLOR polynomial of degree 3.

Problem 3.12: We approximate a certain function $f(x)$ by an n^{th} degree TAYLOR polynomial about the origin. If the absolute valuse of the k^{th} derivative of $f(x)$ on the interval $[-1, 1]$ is bounded by 2^k for each k, how large should we take n in order to estimate $f(0.25)$ with an error no larger than 0.00002?

Problem 3.13: We know

$$x - \frac{x^3}{6} \le \sin(x) \le x - \frac{x^3}{6} + \frac{x^5}{120}, \qquad \text{for } x > 0.$$

Use this inequality (and only this inequality) to obtain an estimate of

$$\int_0^1 \frac{\sin(x)}{x}\, dx$$

with an error no bigger than 0.0008. Be sure to explain what you are doing and to back up your claim that the error is no bigger than 0.0008. (**Hint:** Find an interval containing the value of the integral and use the midpoint as your estimate.)

Problem 3.14: **Part A:** Let $f(x) = \mathrm{e}^{\sin x}$. The TAYLOR polynomial of degree 6 for $f(x)$ is

$$1 + x + \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{15}x^5 - \frac{1}{240}x^6.$$

Given that $\left| f^{(7)}(x) \right| \le 120$ for $|x| \le \frac{1}{2}$ estimate the error in approximating $f(x)$ by the TAYLOR polynomial on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

Part B: Compute the value of the TAYLOR polynomial at $x = \frac{1}{2}$.

Part C: Use your calculator to compute $f\left(\frac{1}{2}\right)$ and compare this value with the estimate obtained in part B. How does the actual error compare with your estimate of the error obtained in part A?

Problem 3.15: Compute the TAYLOR polynomial of degree 4 with center at 1 for $\sqrt{x}$. Evaluate the TAYLOR polynomial at $x = 1.2$ and compare your result with $\sqrt{1.2}$ (as given by a calculator). Compare the actual error with the error estimate given by the TAYLOR remainder.

Problem 3.16: Substitute $t = -s^2$ in

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \frac{t^4}{1-t}$$

and integrate with respect to s to obtain

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \int_0^x \frac{s^8}{1+s^2}\, \mathrm{d}s.$$

Now use $\tan\left(\frac{\pi}{8}\right) = \sqrt{2} - 1$ to estimate π. Note that the absolute error in your estimate of π is bounded by

$$\left| 8 \int_0^{\sqrt{2}-1} \frac{s^8}{1+s^2}\, \mathrm{d}s \right| \le \left| 8 \int_0^{\sqrt{2}-1} s^8\, \mathrm{d}s \right| \le \frac{8}{9}(\sqrt{2}-1)^9 \le .0003190556$$

How does this estimate compare with the actual error?

Problem 3.17: Estimate π as in the previous problem but this time use

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + t^4 + t^5 + t^6 + \frac{t^7}{1-t}.$$

Again compare the estimated error and the actual error.

Problem 3.18: Show $n! \ge 2^{n-1}$ if $n \ge 2$. Then use

$$\mathrm{e} = \sum_{n=0}^{\infty} \frac{1}{n!}$$

to deduce $\mathrm{e} \le 3$.

Problem 3.19: Recall

$$\mathrm{e}^x = \sum_{k=0}^{n} \frac{1}{k!} x^k + \frac{1}{(n+1)!}\, \mathrm{e}^{\xi} x^{n+1}$$

for some ξ between 0 and x. If we use the polynomial part here (the Taylor polynomial of degree n) to estimate $\mathrm{e}^{\,x}$ how large must we take n in order to ensure the absolute truncation error is bounded by $10^{\,-9}$ on the interval $[-1, 1]$. How large must we take n in order to ensure the same error bound, but just on the interval $[-\frac{1}{2}, \frac{1}{2}]$.

Problem 3.20: If we have a full TAYLOR series we can use it to estimate the error in a Taylor polynomial. This method is sometimes more convenient the estimating the TAYLOR remainder directly (and avoids the mysterius ξ.) For example, if we integrate $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$ we obtain

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{1}{k} x^k, \quad |\, x\, | < 1.$$

Thus

$$-\log(1-x) = \sum_{k=1}^{n} \frac{1}{k} x^k \;+\; R_n(x) \quad \text{where} \quad R_n(x) = \sum_{k=n+1}^{\infty} \frac{1}{k} x^k.$$

An easy estimate yields

$$|\, R_n(x) \,| \le \frac{1}{n+1} \sum_{k=n+1}^{\infty} |\, x \,|^k \le \frac{1}{n+1} \frac{|\, x \,|^{n+1}}{1 - |\, x \,|}.$$

Taking $x = \frac{1}{2}$ we obtain

$$\log 2 \cong \sum_{k=1}^{n} \frac{1}{k 2^k}.$$

with an absolute error bounded by $\dfrac{1}{(n+1)2^n}$.

How large should we take n to guarantee an estimate of $\log 2$ with an absolute error no larger than 10^{-12}?

Problem 3.21: Replacing x by $-x$ in the previous problem yields

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \, x^k, \quad |\, x \,| < 1.$$

Thus

$$\log\left(\frac{1+x}{1-x}\right) \;=\; -\log(1-x) + \log(1+x) \;=\; 2 \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}, \quad |\, x \,| < 1.$$

Taking $x = \frac{1}{3}$ we obtain

$$\log 2 = \frac{2}{3} \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{1}{9}\right)^k.$$

Estimate the absolute error $|\, R_n \,|$ in

$$\log 2 = \frac{2}{3} \sum_{k=0}^{n} \frac{1}{2k+1} \left(\frac{1}{9}\right)^k \;+\; R_n$$

where

$$R_n = \frac{2}{3} \sum_{k=n+1}^{\infty} \frac{1}{2k+1} \left(\frac{1}{9}\right)^k.$$

How large should we take n to guarantee the estimate

$$\log 2 \cong \frac{2}{3} \sum_{k=0}^{n} \frac{1}{2k+1} \left(\frac{1}{9}\right)^k$$

has an absolute error no larger than 10^{-12}?

Problem 3.22: The Taylor polynomial of degree 6 with center at the origin for the function

$$f(x) = \mathrm{e}^{x+\sin(x)}$$

is given by

$$p(x) = 1 + 2x + 2x^2 + \frac{7}{6}x^3 + \frac{1}{3}x^4 - \frac{7}{120}x^5 - \frac{37}{360}x^6.$$

One can show

$$\frac{1}{7!}\left| f^{(7)}(x) \right| \leq 0.095 \quad \text{for} \quad 0 \leq x \leq 0.6.$$

Use this fact to estimate the error in approximating $f(0.5)$ by $p(0.5)$. Simplify. (The actual error is actually quite a bit smaller.) **Ans:** 0.0007421875000

Problem 3.23: For a certain function f we have the TAYLOR polynomial of degree 7 at the origin is given by

$$p(x) = \frac{2}{9}x^2 - \frac{1}{27}x^3 + \frac{10}{243}x^4.$$

Suppose we know

$$-2 \leq f^{(5)}(x) \leq 0$$

for $-1 \leq x \leq 1$. Use the TAYLOR POLYNOMIAL to estimate $f(1/2)$. Then find a bound for the error that we make in approximating $f(1/2)$ by $p(1/2)$.

Problem 3.24: Let $f(x) = \log(1+x)$ and note that

$$f^{(k)}(x) = (-1)^{k-1}\frac{(k-1)!}{(1+x)^k}, \quad k \geq 1.$$

Thus the TAYLOR polynomial of degree 8 with center at the origin is

$$p(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \frac{x^7}{7} - \frac{x^8}{8}.$$

Find an upper bound for the absolute error in $p(x)$ as an approximation to $\log(1+x)$ for $0 \leq x \leq \frac{1}{2}$ by using the TAYLOR remainder or facts about alternating series.

Use the TAYLOR remainder to obtain an upper bound for the absolute error in $p(x)$ for $-\frac{1}{2} \leq x \leq 0$.

4 Numeric differentiation

FIRST DERIVATIVE

$$
\begin{aligned}
f'(a) &= \frac{f(a+h)-f(a)}{h} + \bigcirc(h) \\
f'(a) &= \frac{f(a+h)-f(a-h)}{2h} + \bigcirc(h^2) \\
f'(a) &= \frac{-f(a+2h)+4f(a+h)-3f(a)}{2h} + \bigcirc(h^2) \\
f'(a) &= \frac{-f(a+2h)+8f(a+h)-8f(a-h)+f(a-2h)}{12h} + \bigcirc(h^4)
\end{aligned}
$$

SECOND DERIVATIVE

$$
f''(a) = \frac{f(a+h)-2f(a)+f(a-h)}{h^2} + \bigcirc(h^2)
$$

THIRD DERIVATIVE

$$
f'''(a) = \frac{f(a+2h)-2f(a+h)+2f(a-h)-f(a-2h)}{2h^3} + \bigcirc(h^2)
$$

Problem 4.1: Given $f(1.8) = 2.21,\ f(1.9) = 2.19,\ f(2.0) = 2.17$ and $f(2.1) = 2.18$ estimate $f'(2.0)$ and $f''(2.0)$ by using second order estimates.

Problem 4.2: (Deleted)

Problem 4.3: The function f has the values shown in the table below.

x	0	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000
$f(x)$	2.802	2.873	2.911	3.109	3.002	2.999	2.981	2.786	2.544

(A) Estimate $f'(0.875)$ and $f''(0.875)$ by using second order symmetric estimates. **(B)** Estimate $f'(1.000)$ by using a second order estimate.

Problem 4.4: The function f has the values shown in the table below.

x	0	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000
$f(x)$	2.802	2.873	2.911	3.109	3.002	2.999	2.981	2.786	2.544

Estimate $f'(0.500)$ and $f''(0.750)$ by using second order symmetric estimates with the smallest possible step size.

Problem 4.5: The function f has the values shown in the table below.

x	0	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000
$f(x)$	2.802	2.873	2.911	3.109	3.002	2.999	2.981	2.786	2.544

What estimate of $f'(0.500)$ do you get from the fourth order symmetric method

$$f'(a) \approx \frac{f(a-2h) - 8f(a-h) + 8f(a+h) - f(a+2h)}{12h}$$

with step size 0.125.

Problem 4.6: The function f has the values shown in the table below.

x	0	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000
$f(x)$	2.802	2.873	2.911	3.109	3.002	2.999	2.981	2.786	2.544

Estimate $f'(0.250)$ and $f''(0.250)$ by using second order symmetric estimates with the smallest possible step size.

Problem 4.7: Find α, β and γ such that

$$f'(a) \approx \frac{\alpha f(a+3h) + \beta f(a+h) + \gamma f(a)}{h}$$

is a second order method.

Problem 4.8: Find constants A, B and C such that

$$\frac{Af(a) + Bf(a+2h) + Cf(a+3h)}{h} = f'(a) + \mathcal{O}(h^2)$$

(for any suitably differentiable function f). Note you can use TAYLOR polynomials or l'HÔPITAL's rule for this problem.

Problem 4.9: Recall the central symmetric finite difference estimate of order 2 for the first derivative,

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} + \mathcal{O}\left(h^2\right).$$

The $\mathcal{O}\left(h^2\right)$ term is bounded in absolute value by Kh^2 for some constant $K > 0$, for all small h. If $|\, f(x)\,| \leq M$ in a neighborhood of a then the roundoff error in the fraction is bounded in absolute value by $\frac{u2M}{2h}$ where u is the unit round. Thus the worst case absolute error is given by

$$g(h) = \frac{uM}{h} + Kh^2.$$

Find the optimal value of the step size h, that is, the value which minimizes the error estimate $g(h)$.

Do not take this estimate too seriously. Increasing our estimate of K will overemphasize our estimate of the truncation error and so will decrease our calculated optimal h. On the other hand increasing our estimate of M will overemphasize the roundoff error and so will increase our calculated optimal h. To obtain a reasonable estimate of the optimal value of h we need to know the best values of M and K rather accurately.

5 Fixed point iterations

Problem 5.1: Let

$$g(x) = 1 + \frac{1}{x}.$$

Show that g has a unique positive fixed point α. Let $x_0 = 2$. Does the iteration $x_{n+1} = g(x_n)$ appear to converge to α.

Problem 5.2: If $G(x) = 9 + 5\sqrt{x}$ then G maps the interval $I = [9, 49]$ into itself and $|\, G'(x) \,| \leq \frac{5}{6}$ for each $x \in I$. **Part (A).** How do you know that G has a unique fixed point in $[9, 49]$? **Part (B).** If $x_0 = 49$ and we define

$$x_{n+1} = G(x_n)$$

can you guarantee that the iterates x_n converge to the fixed point? **Part (C).** [Algebra] Find the exact fixed point.

Problem 5.3: Plot $y = \cos(\sin(x))$ and $y = x$. Estimate the point where the graphs intersect as an estimate of a fixed point for $\cos(\sin(x))$ and deduce also that $\cos(\sin(x))$ has a unique fixed point. Do a few fixed point iterations to improve your fixed point estimate. **Remark** the fixed point is about $0.768169156\ldots$.

Problem 5.4: Plot $y = \cos(\cos(x))$ and $y = x$. Estimate the point where the graphs intersect as an estimate of a fixed point for $\cos(\sin(x))$ and deduce also that $\cos(\sin(x))$ has a unique fixed point. Do a few fixed point iterations to improve your fixed point estimate. **Remark** the fixed point is about $0.739085133\ldots$.

Problem 5.5: The function f given by $f(x) = x^{x-\sin(x)}$ has a fixed point at 1 and another fixed point near 2. Graphically estimate the fixed point near 2. What happens if you run the fixed point iteration with initial guesses 2.0 and 1.9? Estimate the fixed point by applying the mid-point method to approximate a root of $f(x) - x$. **Ans:** $1.93456321\ldots$.

6 Root finding. Binary search or midpoint method

Problem 6.1: Show analytically that the polynomial

$$p(x) = x^3 - 6x^2 + 9x - 5$$

has a root in the interval $[4, 5]$. Suppose we bisect the interval. (A) Which half of the interval can we guarantee contains a root? (B) If we use the midpoint of the subinterval containing the root to estimate the root give an upper bound for the error.

Problem 6.2: The polynomial

$$p(x) = x^3 - 3x + 1$$

has a root in the interval $[0, 1]$. Use the *bisection method* (also called *binary search method*) to locate the root with an error no larger than $\frac{1}{16}$. (Your work must show clearly that you used the *bisection method*.)

Problem 6.3: Let $p(x) = 2x^3 - 3x^2 + 6x + 5$. The polynomial $p(x)$ has a root in the interval $[-1, 0]$. **(A)** Suppose we bisect the interval. Which half of the interval can we guarantee contains a root. Why? **(B)** Bisect the subinterval known to contain a root. Which sub-sub-interval can we now guarantee contains the root? **(C)** If we use the midpoint of this last interval as an estimate of the root find an bound for the error that we make.

Problem 6.4: Show analytically that the polynomial

$$p(x) = 2x^4 - 4x + 1$$

has a root in the interval $[0, 1]$. Suppose we bisect the interval. **(A)** Which half of the interval can we guarantee contains a root? **(B)** If we use the midpoint of the subinterval containing the root to estimate the root, give an upper bound for the error.

Problem 6.5: Show analytically that the polynomial

$$p(x) = x^5 + x + 1$$

has a root in the interval $[-1.0, -0.5]$. Suppose we bisect the interval. **(A)** Which half of the interval can we guarantee contains a root? **(B)** If we use the midpoint of the subinterval containing the root to estimate the root, give an upper bound for the error.

Problem 6.6: Let $f(x) = x + \cos(x)$. **Part (A).** Show that the function f has exactly one root in the interval $[-\pi/2, 0]$. **Part (B).** Use the bisection method to obtain an estimate of the root with an error no larger than $\pi/8$. In part (B) you must give the estimate provided by application of the bisection method and explain it. I have no interest in the more accurate solution provided by your calculator.

7 Root finding. NEWTON's iteration (scalar)

Problem 7.1: The polynomial

$$p(x) = x^5 + x + 1$$

has a root in the interval $[-1, 0]$. Use the midpoint of this interval as an initial guess and apply NEWTON's method once. What is your new approximation to the root?

Problem 7.2: The polynomial

$$p(x) = x^3 - 3x + 1$$

has a root in the interval $[0, 1]$. Use the midpoint of this interval as an initial guess and use *one* iteration of NEWTON's method to obtain a new approximation to the root. (You must obtain *exactly* the value given by *one* iteration of NEWTON's method.)

Problem 7.3: Let $p(x) = x^3 + x^2 + 3x - 4$. Let $x_0 = 1$ be an initial guess to a root. **(A)** Use NEWTON's method (twice) to compute successive approximations $x_1,\ x_2$ to a root. **(B)** Estimate the error in the root estimate x_1. **(C)** Given that $p(x)$ has only one real root, $.866369759\ldots$, find the actual errors in x_1 and x_2. How do they compare with your estimate of the error in x_1?

Problem 7.4: Find two interval of length 1 with integer endpoints, each interval guaranteed to contain a root of $f(x) = \exp(x) - 6x$. Be sure to justify your answer (without finding the roots). Use an initial guess $x_0 = 0$ and apply NEWTON's iteration once to obtain a new estimate x_1 for one of the roots.

Problem 7.5: Let $a > 0$ and let $f(x) = x^3 - a$. Suppose we decide to estimate the cube root $a^{1/3}$ by applying NEWTON's method to estimate the positive root of $f(x)$. For $n \geq 0$ find an expression for the $(n+1)^{st}$ iterate x_{n+1} in terms of x_n. Simplify.

Problem 7.6: One of the roots of the polynomial $p(x) = x^2 - x - 1$ is the golden ratio $(1+\sqrt{5})/2 = 1.6180339\ldots$. Use NEWTON's method with initial "guess" $x_0 = 2$ for the root and compute the iterates x_1, x_2 and x_3 and also the error in each iterate. Is the rate of convergence about what you would expect?

Problem 7.7: Let $p(x) = x^5 - 3x^3 + 4x + 5$. Let $x_0 = -2$ be an initial guess to a root. Use NEWTON's method to compute successive approximations x_1, x_2, x_3, x_4 and x_5. Given that $-1.6277727742136811\!60\ldots$ is the only real root of $p(x)$ find the actual errors in the root estimates. Do they behave as you would expect?

Problem 7.8: Let

$$f(x) = x^7 + x^6 - 4x^5 - 6x^4 + 2x^3 + 8x^2 + 5x + 1$$

and let

$$g(x) = \frac{f(x)}{f'(x)}$$

so g has the same roots as f, but the roots of g are all simple. Use an initial guess $x_0 = 1.1$ and apply NEWTON's method to f and to g. Comment on the rate of convergence (for example, compare the errors) to the root $\frac{\sqrt{5}+1}{2} = 1.6180339887498948482045868343\ldots$ (a double root for f).

Problem 7.9: Let

$$f(x) = x^7 + x^6 - 4x^5 - 6x^4 + 2x^3 + 8x^2 + 5x + 1.$$

The polynomial f has a root of multiplicity 3 at $x = -1$. Take an initail guess of $x_0 = -0.9$ and apply NEWTON's method and the modified NEWTON's method

$$x_{n+1} = x_n - 3\frac{f(x_n)}{f'(x_n)}$$

and compare the errors. In each case comment if the convergence appears to be linear or quadratic.

Problem 7.10: Use NEWTON's method to approximate one of the two roots of

$$f(x) = \log(x) - \frac{1}{9}x^2\log(x) + x.$$

Take for your initial guess $x_0 = 5.0$ and iterate three times to obtain x_3.

8 Root finding. Secant method

Problem 8.1: The polynomial

$$p(x) = x^3 - 3x + 1$$

has a root in the interval $[0, 1]$. Use the points $x_0 = 0.0$ and $x_1 = 0.5$ as initial guesses and use *one* iteration of the secant method to obtain a new approximation to the root. (You must obtain *exactly* the value given by *one* iteration of the secant method.)

Problem 8.2: Let $p(x) = x^5 - 3x^3 + 4x + 5$. Let $x_0 = -2.0$ and $x_1 = -1.8$ be initial guesses to a root. Use the secant method to compute successive approximations x_2, x_3, x_4 and x_5. Given that $-1.6277727742136811160\ldots$ is the only real root of $p(x)$ find the actual errors in the root estimates. Do they behave as you would expect?

9 Root finding. Miscellaneous facts and methods

Problem 9.1: The polynomial $p(x) = x^5 - 7x^4 + 10x^2 - 1000$ has a simple root at $x = 5$. Thus for small $\epsilon > 0$ the polynomial $p(x) + \epsilon x^4$ has a root at $x = \alpha(\epsilon)$ which is continuously differentiable in ϵ and satisfies $\alpha(0) = 5$. Compute the derivative $\alpha'(0)$.

Problem 9.2: The polynomial $p(x) = x^5 - 7x^4 + 10x^2 + 1000$ has a simple root at $x = 5$. Thus for small $\epsilon > 0$ the polynomial $p(x) + \epsilon x^4$ has a root at $x = \alpha(\epsilon)$ which is continuously differentiable in ϵ and satisfies $\alpha(0) = 5$. Compute the derivative $\alpha'(0)$.

Problem 9.3: Let $f(x) = \exp(x) - x^2$. Prove that f has precisely one real root. Explain why the root must be in the interval $(-1, 0)$.

Problem 9.4: The polynomial $p(x) = x^3 + x^2 - x - 1$ has a simple root at $x = 1$. Thus for small $\epsilon > 0$ the polynomial $p(x) + \epsilon x^2$ has a root at $x = \alpha(\epsilon)$ which is continuously differentiable in ϵ and satisfies $\alpha(0) = 1$. Compute the derivative $\alpha'(0)$.

10 Vectors. Inner products, orthogonality, norms

Problem 10.1: Compute the angle (in radians) between the vectors

$$[1, 2, -3, 2, 3] \quad \text{and} \quad [2, 4, 5, -1, 2].$$

Problem 10.2: A student playing around with some vectors $\vec{a}_k$, $k = 1, 2, 3, 4$, was amused to find

$$\vec{a}_1 + 2\vec{a}_2 + 3\vec{a}_3 + 4\vec{a}_4 = 4\vec{a}_1 + 3\vec{a}_2 + 2\vec{a}_3 + \vec{a}_4.$$

Are the vectors $\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4$ linearly independent? Justify your answer.

Problem 10.3: Let

$$\vec{v}_1 = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ -1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} -1 \\ 2 \\ 3 \\ 1 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 7 \\ 6 \\ 1 \\ 5 \end{bmatrix},$$

Part A. Write $\vec{v}_4$ as a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3$ (explicitly). **Part B.** Show that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are linearly independent.

Problem 10.4: Let $\vec{v}_1, \cdots, \vec{v}_n$ be vectors in $\mathbb{R}^m$. Let $\vec{w}$ be a vector in $\mathbb{R}^m$ and suppose $\vec{w} \notin \text{span}\{\vec{v}_1, \cdots, \vec{v}_n\}$. (The symbol $\notin$ means "*is not an element of*"). **Part A.** If $\alpha\vec{w} + \beta_1\vec{v}_1 + \cdots + \beta_n\vec{v}_n = \vec{0}$ give a careful explanation of why it follows that $\alpha = 0$. **Part B.** Now show if the set $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is linearly independent, then the set $\{\vec{v}_1, \cdots, \vec{v}_n, \vec{w}\}$ is linearly independent.

Problem 10.5: Consider the vectors $\vec{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -1 \\ -4 \\ 1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 2 \\ 3 \\ a \end{bmatrix}$, and $\vec{v} = \begin{bmatrix} 3 \\ 2 \\ 9 \end{bmatrix}$. For what values of a is $\vec{v}$ in the span of $\vec{v}_1, \vec{v}_2, \vec{v}_3$?

Problem 10.6: Consider the vectors $\vec{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -1 \\ -4 \\ 1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 2 \\ 3 \\ a \end{bmatrix}$, and $\vec{v} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. For what values of a is $\vec{v}$ in the span of $\vec{v}_1, \vec{v}_2, \vec{v}_3$?

11 Matrices. Operations, determinants

Problem 11.1: Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 4 & 5 & 7 \end{bmatrix}.$$

Compute the 1-norm of A where, if B is $m \times n$ then the 1-*norm* of B is given by

$$\|B\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} | \, b_{ij} \, | \, .$$

Problem 11.2: Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 4 & 5 & 7 \end{bmatrix}.$$

Compute the ∞-norm of A where, if B is $m \times n$ then the ∞-*norm* of B is given by

$$\|B\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} | \, b_{ij} \, | \, .$$

Problem 11.3: Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 4 & 5 & 7 \end{bmatrix}.$$

Compute the Frobenius norm of A where, if B is $m \times n$ then the *Frobenius norm* $\|B\|_F$ of B is given by

$$\|B\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} | \, b_{ij} \, |^2}.$$

Problem 11.4: Let C be the matrix

$$C = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

Compute the L^2 operator norm $\|C\|_2$ of the matrix C where the L^2 *operator norm* $\|B\|_2$ of an $m \times n$ matrix B (also called the *spectral norm*) is given by

$$\|B\|_2 = \max_{\|\vec{x}\|_2 = 1} \|B\vec{x}\|_2 \, .$$

Hint: You might try Lagrange multipliers, or just substitute $x_1 = \cos(t)$ and $x_2 = \sin(t)$ in $\|C\vec{x}\|_2^2$.

Problem 11.5: Let C be the matrix

$$C = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

Compute the 1-norm, the ∞-norm and the Frobenius norm of C.

Problem 11.6: Let C be the matrix

$$C = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

Find a monic polynomial p of degree 2 such that $p(C) = 0$, that is, find constants a and b such that $C^2 + a\,C + b\,I = 0$.

Problem 11.7: Let C be the matrix

$$C = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

Note

$$z^5 = (z^2 - 4z + 4)(z^3 + 4z^2 + 12z + 32) + 16(5z - 8).$$

Use this fact to compute C^5 without computing any powers of C. (You will need to do the previous problem first.)

Problem 11.8: Let C be the matrix

$$C = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}.$$

Find a monic polynomial q of degree 2 such that $q(C^T C) = 0$. Find the roots of q. Compare the square root of the largest root of q with $\|C\|_2$. Make a conjecture.

Problem 11.9: If $A = (a_{ij})$ is an $n \times n$ matrix the *trace* or *spur* $\mathrm{tr}(A)$ of A is defined by

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Let

$$A = \begin{bmatrix} 1 & 2 \\ 4 & -5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}.$$

Compute AB, BA, $\mathrm{tr}(AB)$ and $\mathrm{tr}(BA)$. Do you care to make a conjecture?

Problem 11.10: Let A be a 2×2 matrix such that

$$A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Compute the product

$$A \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Problem 11.11: Let A be a 2×2 matrix such that

$$A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Find A.

Problem 11.12: Let A be a 2×2 matrix such that

$$A \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}.$$

Compute the product

$$A \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}.$$

Problem 11.13: Let A be a 2×2 matrix such that

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Find A.

Problem 11.14: Let σ be the permutation which re-arranges $(1, 2, \cdots, n-1, n)$ as $(n, n-1, \cdots, 2, 1)$, call it the *reversor*. For what values of n is the reversor σ even? For what values is the reversor odd?

Problem 11.15: Solve for x

$$\det \begin{bmatrix} x & 2 & x \\ 1 & 1 & 2 \\ 3 & 2 & -1 \end{bmatrix} = 0.$$

Problem 11.16: For a square matrix the notions of upper and lower triangular are defined relative to the main (or *sinister*) diagonal. But there is another diagonal, the *dexter* diagonal. Compute the determinant of a matrix which is "triangular" relative to the dexter diagonal. As a check of your work compute in different ways

$$\det \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Problem 11.17: Let A be an $n \times n$ matrix. Then the series

$$\mathrm{e}^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k,$$

where $A^0 = I$, converges. Let

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Compute e^{tA}.

Problem 11.18: Let

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Let a, α and b, β be scalars. Compute, and express in terms of I and J the matrix

$$(aI + bJ)(\alpha I + \beta J).$$

Also compute $\det(aI + bJ)$. Look familiar?

Problem 11.19: Let $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ and let B be a 2×2 matrix. If

$$AB = B + I$$

where I is the 2×2 identity matrix compute B.

Problem 11.20: Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. First compute $(A^3 + 2 * A)^2$. Then compute

$$A\left((A^3 + 2A)^2 + 17I\right)^{1,423,846,679}$$

where I is the 2×2 identity matrix.

12 Positive definite matrices

(Problems will be added later.)

13 GAUSS reduction. GAUSS–JORDAN reduction

Problem 13.1: Find the complete row reduced row echelon GAUSS–JORDAN canonical form of the matrix A (henceforth abbreviated rref(A)) where

$$A = \begin{bmatrix} 1 & 2 & 1 & 4 & 5 \\ 5 & 4 & 0 & 3 & 1 \\ 0 & 0 & 0 & 5 & 1 \end{bmatrix}$$

Problem 13.2: If

$$B = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \\ 0 & 2 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

then find $\mathrm{rref}(B)$.

Problem 13.3: Consider a system of linear equations $A\vec{x} = \vec{b}$. After an immense number of elementary row operations we reduce the *augmented* matrix $[A, \vec{b}]$ to the row reduced echelon form,

$$\mathrm{rref}\Big([A, \vec{b}]\Big) = \begin{bmatrix} 0 & 1 & 0 & 0 & 3 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Call the variables in the system of linear equations $x_1, x_2, \cdots$ **Part A.** Which columns are pivotal? **Part B.** Which columns are free? **Part C.** Write the general solution of the system in vector parametric form.

Let A be an $m \times n$ matrix. Obviously elementary row operations do not change $\mathrm{row}(A)$, though they may change $\mathrm{col}(A)$. However, since $\dim \mathrm{row}(A) = \dim \mathrm{col}(A)$ we see that elementary row operations do not change $\dim \mathrm{col}(A)$.

Problem 13.4: Consider the vectors

$$\vec{v}_1 = \begin{bmatrix} 8 \\ -8 \\ 21 \\ -11 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ -4 \\ 3 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 6 \\ -7 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} -1 \\ 14 \\ 12 \\ -23 \end{bmatrix}.$$

Find the dimension of the subspace of $\mathbb{R}^4$ spanned by these vectors.

Problem 13.5: **Part A.** Find the row reduced echelon canonical form $\mathrm{rref}(A)$ of the matrix

$$A = \begin{bmatrix} 2 & 1 & 7 & 5 & 4 \\ 1 & 1 & 5 & 3 & 3 \\ -2 & 2 & 2 & -2 & 3 \end{bmatrix}.$$

Part B. Identify the pivotal columns. **Part C.** What is the rank of A? **Part D.** Are the row of A linearly independent? **Part E.** Are the columns of A linearly independent?

14 LU factorization of matrices

Problem 14.1: A certain 3×3 matrix A has an LU decomposition with

$$L := \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{2}{3} & -\frac{1}{3} & 1 \end{bmatrix} \quad U := \begin{bmatrix} 3 & 2 & 5 \\ 0 & -2 & 1 \\ 0 & 0 & -5 \end{bmatrix}.$$

Solve the system $Ax = b$ where $b = [5, 15, 10]^T$. **Answer:** $[9, -6, -2]^T$

Problem 14.2: The 4×4 Hilbert matrix A given by

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}$$

has an LU factorization $A = LU$ where

$$U = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ 0 & \frac{1}{12} & \frac{1}{12} & \frac{3}{40} \\ 0 & 0 & \frac{1}{180} & \frac{1}{120} \\ 0 & 0 & 0 & \frac{1}{2800} \end{bmatrix}.$$

Compute L (recall it is normalized).

15 CHOLESKY factorization of matrices

Problem 15.1: Compute the CHOLESKY factorization $A = C^T C$, where C is upper triangular, of the matrix

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}.$$

Problem 15.2: Compute the CHOLESKY factorization $A = C^T C$, where C is upper triangular, of the matrix

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}.$$

Problem 15.3: Find the CHOLESKY factorization $A = C\,C^T$ (where C is lower triangular) of the matrix

$$\begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix}.$$

Answer:

$$C = \begin{bmatrix} \sqrt{2} & 0 \\ -\sqrt{2} & 1 \end{bmatrix}.$$

16 Iterative methods for fixed points for nonlinear systems. Simultaneous and successive displacements

If we have a fixed-point problem for a system of nonlinear equations, for example,

$$\begin{aligned} x_1 &= g_1(x_1, x_2, x_3) \\ x_2 &= g_2(x_1, x_2, x_3) \\ x_3 &= g_3(x_1, x_2, x_3) \end{aligned}$$

then we can try a fixed-point iteration with *simultaneous displacements*

$$\begin{aligned} x_1^{(n+1)} &= g_1(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}) \\ x_2^{(n+1)} &= g_2(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}) \\ x_3^{(n+1)} &= g_3(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}) \end{aligned}$$

or a fixed-point iteration with *successive displacements*

$$\begin{aligned} x_1^{(n+1)} &= g_1(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}) \\ x_2^{(n+1)} &= g_2(x_1^{(n+1)}, x_2^{(n)}, x_3^{(n)}) \\ x_3^{(n+1)} &= g_3(x_1^{(n+1)}, x_2^{(n+1)}, x_3^{(n)}). \end{aligned}$$

It should be clear how these ideas extend to systems of more equations.

There are some sufficent criteria for convergence analogous to the contraction hypothesis in the scalar case.

Problem 16.1: The fixed point problem

$$\begin{cases} x = f(x, y) \\ y = g(x, y) \end{cases}$$

where

$$f(x,y) = \frac{x^2+3y^3+2y^2+4y+1}{15}$$
$$g(x,y) = \frac{3x-2x^2+4y-2y^2+3}{15}$$

has the solution $(\overline{x},\overline{y}) = (0.1645436899, 0.2966770274)$. (There is another solution, $(-2.337069, -2.441771)$, but it does not concern us here.) Use the initial point $(x_0, y_0) = (0,0)$ and compute 8 fixed point iterates (x_n, y_n) by the method of simultaneous displacements (JACOBI iteration). At each step compute also the error $\left((\overline{x}-x_n)^2 + (\overline{y}-y_n)^2\right)^{1/2}$. Do you appear to have convergence?

Problem 16.2: Repeat the previous problem but this time use the method of successive displacements (SEIDEL iteration). Compare the errors to the ones observed in the previous problem.

Problem 16.3: Consider the fixed-point problem

$$x = \frac{3x^3-6x^2+24x+xy-53}{6x^2-12x-1+y}$$
$$y = \frac{3yx^2-24y+3x^2+24-30x}{6x^2-12x-1+y}$$

Start at the initial point $[-2,1]$ and compute 3 iterations using the method of simultaneous displacements. Your iterates should converge towards $[-2.92173131345952223, 1.60954160414971480]$. Compute the 2-norm of the error in each iterate. (Use a spreadsheet or other appropriate tool.)

Problem 16.4: Redo the previous problem, but this time use the method of successive displacements. Comments?

Problem 16.5: The system of equations

$$x = x^2 - xy + y^2$$
$$y = x^2 + xy + y^2$$

has two solutions, $[0,0]$ and $[0.2733011742, 0.6027847152]$. Experiment with the fixed point iteration to see what happens. Convert the equations into a root finding problem and experiment with Newton's method. (This problem is a bit vague.)

Problem 16.6: The system of equations

$$x = \frac{1}{3}x^2 - \frac{2}{3}y + \frac{1}{3}$$
$$y = -\frac{1}{15}x^2 + \frac{1}{5}y^2 - \frac{4}{5}y - \frac{1}{3}.$$

has two solutions. One of them is $[.570592062563225385, -.193100442914760222]$.

Use the initial point $[0.6, -0.2]$ and the implied fixed point iteration with the method of simultaneous displacements with 5 iterations to approximate a solution. Compute the 2-norm of the actual error in each iteration. Comment.

Problem 16.7: The system of equations

$$\begin{aligned} x &= \frac{1}{3}x^2 - \frac{2}{3}y + \frac{1}{3} \\ y &= -\frac{1}{15}x^2 + \frac{1}{5}y^2 - \frac{4}{5}y - \frac{1}{3}. \end{aligned}$$

has two solutions. One of them is $[.570592062563225385, -.193100442914760222]$.

Use the initial point $[0.6, -0.2]$ and the implied fixed point iteration with the method of successive displacements with 5 iterations to approximate a solution. Compute the 2-norm of the actual error in each iteration. Comment. Compare the errors here with the errors in the previous problem.

17 GAUSS–JACOBI and GAUSS–SEIDEL iterative methods for linear systems

For systems of linear equations $A\vec{x} = \vec{b}$ where A is a sparse matrix iterative methods are often used. If A is a very large matrix, storage space constraints may force the use of iterative methods.

Consider the system of linear equations $A\vec{x} = \vec{b}$ where the coefficient matrix A is a $n \times n$ matrix. Iterative methods generally consist of decomposing the matrix A as $A = Q + R$ where Q is nonsingular and is the coefficient matrix of a system of linear equations which may be efficiently solved, for example a triangular or a diagonal matrix. The corresponding fixed point iteration is then

$$Q\vec{x}^{(n+1)} = \vec{b} - R\vec{x}^{(n)}.$$

If we take Q to be the diagonal part of A we obtain the GAUSS–JACOBI iteration (which converges if A is strictly diagonally (row) dominant). If we take Q to be the lower triangular part of A we obtain the GAUSS–SEIDEL iteration (which converges for example if A is symmetric positive definite with positive diagonal elements).

The GAUSS–SEIDEL iteration is equivalent to doing the GAUSS–JACOBI iteration of A as described above, except we use the updated value of each component as it becomes available, rather than waiting until all the new components of the new iterate have been calculated. Thus the GAUSS–SEIDEL method is sometimes known as the *method of successive displacements* whereas the GAUSS–JACOBI method is known as the *method of simultaneous displacements*, though these names really have wider application as we saw in the previous section.

Intuition suggests that GAUSS–SEIDEL should be the better method, but that is not always the case.

Explicitly, the naming convention appears to be:

1. fixed point problems for systems, linear or nonlinear
   (a) method of simultaneous displacements = JACOBI iteration
   (b) method of successive displacements = SEIDEL iteration
2. systems of linear equations
   (a) GAUSS–JACOBI = convert to fixed point problem as above and apply JACOBI iteration
   (b) GAUSS–SEIDEL = convert to fixed point problem as above and apply SEIDEL iteration

Problem 17.1: Consider the linear system $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} -4 & -1 & 2 \\ 2 & 5 & 2 \\ 2 & 2 & -5 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$$

Starting with an initial "guess" $[0,0,0]^T$ compute the first 5 GAUSS–JACOBI iterates.

Compute also the exact solution (any way you like) and use it to find the 2-norm of the error in each iterate. Do the GAUSS–JACOBI iterates appear to be converging? As a sanity check on your work, the 2-norm of the error in the third iterate appears to be about 0.06180215

Problem 17.2: Consider the linear system $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} -4 & -1 & 2 \\ 2 & 5 & 2 \\ 2 & 2 & -5 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$$

Starting with an initial "guess" $[0,0,0]^T$ compute the first 5 GAUSS–SEIDEL iterates.

Compute also the exact solution (any way you like) and use it to find the 2-norm of the error in each iterate. Do the GAUSS–SEIDEL iterates appear to be converging? As a sanity check on your work, the 2-norm of the error in the third iterate appears to be about 0.03885003

Problem 17.3: Consider the linear system $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

Starting with an initial "guess" $[0,0]^T$ compute the first 5 GAUSS–JACOBI iterates.

Compute also the exact solution (any way you like) and use it to find the 2-norm of the error in each iterate. Do the GAUSS–JACOBI iterates appear to be converging? As a sanity check on your work, the 2-norm of the error in the third iterate appears to be about 0.18034689

Problem 17.4: Consider the linear system $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

Starting with an initial "guess" $[0,0]^T$ compute the first 5 GAUSS–SEIDEL iterates.

Compute also the exact solution (any way you like) and use it to find the 2-norm of the error in each iterate. Do the GAUSS–SEIDEL iterates appear to be converging? As a sanity check on your work, the 2-norm of the error in the third iterate appears to be about .00116276

Problem 17.5: Consider the system of linear equations

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 2 \\ -3x_1 + 2x_2 + x_3 &= 3 \\ 2x_2 + 2x_3 &= 1 \end{aligned}$$

By separating out the diagonal convert this system to a (vector) fixed point problem. Perform one GAUSS–JACOBI iteration starting with an initial guess $x_0 = (0,0,0)^T$. What is your result? **Ans:** $(2, 3/2, 1/2)$

Problem 17.6: Consider the system of linear equations

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 2 \\ -3x_1 + 2x_2 + x_3 &= 3 \\ 2x_2 + 2x_3 &= 1 \end{aligned}$$

By separating out the diagonal convert this system to a (vector) fixed point problem. Perform one GAUSS-SEIDEL iteration starting with an initial guess $x_0 = (0,0,0)^T$. What is your result? **Ans:** $(2, 9/2, -1/2)$

18 NEWTON's iteration for nonlinear systems

Problem 18.1: Use NEWTON's method for systems to approximate a solution to the system of equations

$$\begin{aligned} x^2 - 2x - y &= -\frac{1}{2} \\ x^2 + 4y^2 &= 4. \end{aligned}$$

Use the initial point $[2.0, 0.3]$ and iterate twice. Given that the actual solution is $[1.900676726, 0.3112185654]$ compute the error in each of your iterates. Comment. What happens if you use the initial point $[-1.00, 0.25]$? Do three iterations.

Problem 18.2: Use NEWTON's method for systems to approximate a solution to the system of equations

$$\begin{aligned} x^3 - 3y &= -2 \\ y^3 - 2x^2 + 3x &= -4. \end{aligned}$$

Use the initial point $[-0.8, 0.4]$ and iterate three times. Given that the actual solution is $[-0.865034729, 0.450902472]$ compute the error in each of your iterates. Comment.

Problem 18.3: Use NEWTON's method for systems to approximate a solution to the system of equations

$$\begin{aligned} x^2 + y^2 + z^3 &= 3 \\ x + 2y - 2z &= 0 \\ x^3 + y^3 - 3z^2 &= 2. \end{aligned}$$

Use the initial point $[1.0, -1.0, 0]$ and iterate three times. Given that the actual solution is $[1.44825, -0.95663, -0.23250]$ compute the error in each of your iterates. Comment.

Problem 18.4: The solution of the system

$$\begin{aligned} x^2 + y^2 &= 1 \\ x^2 + y &= -1 \end{aligned}$$

is obviously $[0, -1]$. Use the initial point $[0.25, -2.0]$ and iterate 5 times. At each step compute the error in your iterate. What behavior of the errors do you observe?

Problem 18.5: Estimate the minimum point of

$$x^4 + 2y^4 + 3x^2y^2 + 2xy^3 + 3x - 2y^2 + 4$$

by setting the partial derivatives equal to 0 and using NEWTON's method with 3 iterations and an initial guess of $[-1.0, 1.0]$.

Problem 18.6: The system of equations

$$\begin{aligned} 2x^2 + y^2 &= 4 \\ x^2 - 4xy - y^2 &= 0 \end{aligned}$$

has four solutions. Use NEWTON's method with the initial point $[1.0, 1.0]$ and with four iterations to estimate one solution. What happens if you use the initial points $[-1.0, 1.0]$, $[1.0, -1.0]$ or $[-1.0, -1.0]$?

19 Interpolation polynomials

Problem 19.1: The function $f(x)$ satisfies $f(0) = 1$, $f(1) = 1$ and $f(2) = 2$. **(A)** Find the interpolation polynomial of degree at most 2 which interpolates $f(x)$ between these points. **(B)** Use the result of part (A) to estimate $f(\frac{3}{2})$. **(C)** What estimate do you get for $f(\frac{3}{2})$ if you use linear interpolation between successive points?

Problem 19.2: The function $f(x)$ satisfies $f(-1) = 1$, $f(0) = 2$ and $f(2) = 1$. **(A)** Find the interpolation polynomial of degree at most 2 which passes through the given points on the graph of $f(x)$. **(B)** Use this interpolation polynomial to estimate $f(1)$. **(C)** If it is known that the third derivative of $f(x)$ on the interval $[-1, 2]$ is bounded in absolute value by 1.12 find a bound for the error in your estimate of $f(1)$.

Problem 19.3: Compute the interpolation polynomial of degree at most 2 for the function $f(x) = x^3$ with nodes at -1, 0 and 1. Compute also the interpolation polynomial in the case the nodes are 0, 1 and 2.

Problem 19.4: The function $f(x)$ satisfies $f(0) = 1$, $f(1) = 1$ and $f(2) = 2$. **(A)** Find the interpolation polynomial $P_2(x)$ of degree at most 2 which interpolates $f(x)$ between these points. **(B)** Suppose we know $|f^{(3)}(t)| \leq 2.0$ for $0 \leq t \leq 2$. Compute a good bound for the absolute error in $P_2(0.5)$ as an approximation to $f(0.5)$.

Problem 19.5: The function $f(x)$ satisfies $f(0) = 1$, $f(1) = 0$, $f(2) = 0$ and $f(3) = -1$. **(A)** Find the interpolation polynomial $P_3(x)$ of degree at most 3 which interpolates $f(x)$ between these points. **(B)** Suppose we know $|f^{(4)}(t)| \leq 1.502$ for $0 \leq t \leq 3$. Compute a good bound for the absolute error in $P_3(1.5)$ as an approximation to $f(1.5)$.

Problem 19.6: Compute the interpolation polynomial of degree at most 2 for the function $f(x) = x^3$ with nodes at -1, 0 and 1. Compute also the interpolation polynomial in the case the nodes are 0, 1 and 2.

Problem 19.7: The function $f(x)$ satisfies $f(0) = 1$, $f(1) = 1$ and $f(2) = 2$. **(A)** Find the interpolation polynomial $p(x)$ of degree at most 2 which interpolates $f(x)$ between these points. **(B)** Suppose we know $|f^{(3)}(t)| \le 2.0$ for $0 \le t \le 2$. Compute a good bound for the absolute error in $p(0.5)$ as an approximation to $f(0.5)$.

Problem 19.8: The function $f(x)$ satisfies $f(0) = 1$, $f(1) = 0$, $f(2) = 0$ and $f(3) = -1$. **(A)** Find the interpolation polynomial $p(x)$ of degree at most 3 which interpolates $f(x)$ between these points. **(B)** Suppose we know $|f^{(4)}(t)| \le 1.502$ for $0 \le t \le 3$. Compute a good bound for the absolute error in $p(1.5)$ as an approximation to $f(1.5)$.

Problem 19.9: The function $f(x)$ satisfies $f(0) = 3$, $f(1) = 1$ and $f(2) = 2$. **(A)** Find the interpolation polynomial of degree at most 2 which interpolates $f(x)$ between these points. **(B)** Use the result of part (A) to estimate $f(\frac{3}{2})$. **(C)** What estimate do you get for $f(\frac{3}{2})$ if you use linear interpolation between successive points?

Problem 19.10: The function $f(x)$ satisfies $f(-1) = 1$, $f(0) = 2$ and $f(2) = 1$. **(A)** Find the interpolation polynomial of degree at most 2 which passes through the given points on the graph of $f(x)$. **(B)** Use this interpolation polynomial to estimate $f(1)$. **(C)** If it is known that the third derivative of $f(x)$ on the interval $[-1, 2]$ is bounded in absolute value by 1.12 find a bound for the error in your estimate of $f(1)$.

Problem 19.11: Find the interpolation polynomial $p(x)$ of degree ≤ 3 through the points $(1, -1),\ (2, 3),\ (3, 2),\ (4, -2)$.

Problem 19.12: Find the interpolation polynomial of degree ≤ 3 through the points $(0, 2), (1, 0), (2, 3)$ and $(3, 1)$.
Ans: $-\frac{5}{3}x^3 + \frac{15}{2}x^2 - \frac{47}{6}x + 2$

20 Interpolation polynomials and NEWTON divided differences

DIVIDED DIFFERENCES

$$
\begin{aligned}
f[x_0, x_1] &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\
f[x_0, x_1, \cdots, x_k] &= \frac{f[x_1, \cdots, x_k] - f[x_0, \cdots, x_{k-1}]}{x_k - x_0} \\
P_n(x) &= f(x_0) + f[x_0, x_1](x - x_0) + \cdots + f[x_0, x_1, \cdots, x_n](x - x_0) \cdots (x - x_{n-1})
\end{aligned}
$$

Problem 20.1: Let $x_0 = 0$, $x_1 = 1.5$ and $x_2 = 2$ and suppose $f(x_0) = 1$, $f(x_1) = 0$ and $f(x_2) = -1$. Compute the NEWTON divided differences $f[x_0, x_1]$ and $f[x_0, x_1, x_2]$. Find the linear interpolation polynomial

$P_1(x)$ through the first two points and the quadratic interpolation polynomial $P_2(x)$ through all three points. What estimates do these polynomials give for $f(1)$?

Problem 20.2: Let $f(x) = x^3$. Let x_0, x_1 and x_2 be three distinct points. Compute the NEWTON divided difference $f[x_0, x_1, x_2]$.

Problem 20.3: For a certain function $f(x)$ we know the NEWTON divided differences: $f[-1] = 2,\ \ f[-1,1] = 1,\ \ f[-1,1,2] = -2,\ \ f[-1,1,2,4] = 2$. **(A)** Find the interpolation polynomial $P_3(x)$ for $f(x)$ of degree at most 3 with nodes at $-1,\ 1,\ 2,\ 4$. **(B)** Use the interpolation polynomial to estimate $f(0)$. **(C)** Find a good upper bound for the absolute error in the estimate of $f(0)$ given that $|f[-1,1,2,4,x]| \leq 1.4$ for all x in the interval $[-1,4]$.

Problem 20.4: Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial of degree n and let $x_0, x_1, \cdots, x_n$ be any $n+1$ distinct points. Use properties of the NEWTON divided differences to compute $f[x_0, x_1, \cdots, x_n]$.

Problem 20.5: Carefully fill in the missing entries in the following table of NEWTON divided differences.

k	0	1	2	3
x_k	-1	0	1	3
y_k	1	0	-1	2
$f[x_{k-1}, x_k]$	$\times$	-1	-1	
$f[x_{k-2}, x_{k-1}, x_k]$	$\times$	$\times$		
$f[x_{k-3}, x_{k-2}, x_{k-1}, x_k]$	$\times$	$\times$	$\times$	

Use the NEWTON divided differences to compute the interpolation polynomial of degree at most 3 for the 4 data points (x_k, y_k) $k = 0, \cdots, 3$.

Problem 20.6: Fill in the three missing entries in the following table of NEWTON divided differences.

k	0	1	2	3
x_k	-1	1	2	4
y_k	0	1	3	-1
$f[x_{k-1}, x_k]$	$\times$	$1/2$	2	
$f[x_{k-2}, x_{k-1}, x_k]$	$\times$	$\times$		$-4/3$
$f[x_{k-3}, x_{k-2}, x_{k-1}, x_k]$	$\times$	$\times$	$\times$	

Use the NEWTON divided differences to compute the interpolation polynomial of degree at most 3 for the 4 data points (x_k, y_k) $k = 0, \cdots, 3$.

Problem 20.7: Let $x_0 = 0$, $x_1 = 1.5$, $x_2 = 2$ and $x_3 = 3$ and suppose $f(x_0) = 1$, $f(x_1) = 1$, $f(x_2) = 1.5$ and $f(x_3) = 4$. Compute the NEWTON divided differences $f[x_0]$, $f[x_0, x_1]$, $f[x_0, x_1, x_2]$ and $f[x_0, x_1, x_2, x_3]$. Then find the cubic interpolation polynomial $p(x)$ for $f(x)$ with nodes x_0, x_1, x_2 and x_3.

Problem 20.8: Let $f(x) = x^3$. Let x_1, x_2, x_3 and x_4 be four distinct points. Compute the divided difference $f[x_1, x_2, x_3, x_4]$.

Problem 20.9: For a certain function f we have the NEWTON divided differences

$$f[1] = 2,\ f[1,2] = -3,\ f[1,2,4] = 2,\ f[1,2,4,5] = 5,\ f[1,2,4,5,6] = 4.$$

Find the interpolation polynomial $p(x)$ of degree ≤ 4 through the points $(k, f(k))$, $k = 1,\ 2,\ 4,\ 5,\ 6.$

21 ČEBYŠEV polynomials

Problem 21.1: Let $T_n(x)$ be the n^{th} ČEBYŠEV polynomial so $T_n(x) = \cos(n\theta)$ if $x = \cos(\theta)$. One can show

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1. \end{aligned}$$

Let $\widetilde{T}_n(x) = 2^{1-n} T_n(x)$ be the normalized *monic* ČEBYŠEV polynomial. Notice

$$\max_{|x|\leq 1} |\widetilde{T}_n(x)| = 2^{1-n}.$$

Let $P(x) = 2x^3 + 2x^2 - 3x + 1$. **[a]** Find the polynomial $Q(x)$ of degree ≤ 2 obtained by *economizing* $P(x)$. **[b]** Find the polynomial $R(x)$ of degree ≤ 1 obtained by *economizing* $Q(x)$. **[c]** Use facts about ČEBYŠEV polynomials to estimate

$$\max_{|x|\leq 1} |P(x) - R(x)|.$$

22 Piecewise interpolation polynomials. Splines

Problem 22.1: Suppose

$$s(x) = \begin{cases} x^3 + ax^2 - 4x + c, & 0 \leq x \leq 2 \\ -x^3 + 9x^2 + bx + 34, & 2 \leq x \leq 4 \end{cases}$$

Find constants a, b, c such the $s(x)$ is twice continuously differentiable on the interval $[0, 4]$.

Problem 22.2: Let $x_0 = 0$, $x_1 = 1.5$, $x_2 = 2$ and $x_3 = 2.5$ and suppose $f(x_0) = 1$, $f(x_1) = 1$, $f(x_2) = 1.5$ and $f(x_3) = 1$. Compute the NEWTON divided differences $f[x_0]$, $f[x_0, x_1]$, $f[x_0, x_1, x_2]$ and $f[x_0, x_1, x_2, x_3]$. Then find the cubic interpolation polynomial $P_3(x)$ for $f(x)$ with nodes x_0, x_1, x_2 and x_3.

Problem 22.3: Suppose

$$s(x) = \begin{cases} ax^3 + x, & -2 \le x \le 0 \\ x^3 + bx, & 0 \le x \le 2 \end{cases}.$$

Find *all* values of the constants a, b such the $s(x)$ is twice continuously differentiable on the interval $[-2, 2]$.

Problem 22.4: If $s(x) = 0$ for $x < 2$ and $s(x) = (x-2)^3$ for $x \ge 2$ is it true that s is a cubic spline? Justify your answer.

Problem 22.5: Let

$$s(x) = \begin{cases} 1 - x + ax^2 + x^3 & \text{if } 0 \le x \le 1 \\ 3 + bx + cx^2 - x^3 & \text{if } 1 < x \le 2 \end{cases}$$

Determine a, b and c so that $s(x)$ is a *natural* cubic spline on the interval $[0, 2]$.

Problem 22.6: Let

$$s(x) = \begin{cases} x^3 + ax^2 & \text{if } 0 \le x \le 1 \\ x^2 - bx + 1 & \text{if } 1 < x \le 2 \end{cases}$$

Determine a and b so that $s(x)$ is a cubic spline on the interval $[0, 2]$.

Problem 22.7: Let

$$s(x) = \begin{cases} 1 - x + ax^2 + x^3 & \text{if } 0 \le x \le 1 \\ 3 + bx + cx^2 - x^3 & \text{if } 1 < x \le 2 \end{cases}$$

Determine a, b and c so that $s(x)$ is a *natural* cubic spline on the interval $[0, 2]$.

Problem 22.8: (Deleted)

Problem 22.9: Let

$$s(x) = \begin{cases} 0 & \text{if } x \le 2 \\ (x-2)^3 & \text{if } 2 < x \end{cases}$$

Is $s(x)$ a cubic spline? Justify your answer.

Problem 22.10: (Deleted)

Problem 22.11: Let

$$s(x) = \begin{cases} x^3 + ax^2 & \text{if } 0 \le x \le 1 \\ x^2 - bx + 1 & \text{if } 1 < x \le 2 \end{cases}$$

Determine a and b so that $s(x)$ is a cubic spline on the interval $[0, 2]$.

Problem 22.12: **(A)** Write a careful definition of a cubic spline with nodes at $x_1 < x_2 < \cdots < x_n$. **(B)** Let

$$s(x) = \begin{cases} 6x - (x+1)^3 & \text{if } x \le -1 \\ 6x + (x+1)^3 & \text{if } -1 < x \end{cases}$$

Is $s(x)$ a cubic spline? Justify your answer.

Problem 22.13: Let $s(x)$ be the natural cubic interpolation spline with knots $(0,2)$, $(1,0)$, $(2,3)$ and $(3,1)$. Suppose

$$s(x) = \begin{cases} 2 - \frac{11}{3}x + \frac{5}{3}x^3 & \text{if } 0 \le x < 1 \\ \\ 7 - \frac{56}{3}x + 15x^2 - \frac{10}{3}x^3 & \text{if } 1 \le x < 2 \\ \\ -33 + \frac{124}{3}x + Ax^2 + Bx^3 & \text{if } 2 \le x \le 3 \end{cases}$$

Find A and B. **Ans:** $A = -15$, and $B = \frac{5}{3}$.

23 Linear least squares fitting

Problem 23.1: Consider a vehicle travelling at v mph, that is, $\frac{22}{15}v$ fps. Applying the brakes develops a coefficient of friction μ. If the vehicle stops in a distance of s feet, the work done by friction is μmgs where m is the mass of the vehicle and g is the acceleration of gravity, say 32 ft/sec^2. This work must be equal to the initial kinetic energy, $\frac{1}{2}m\left(v\frac{22}{15}\right)^2$. Thus if the speed is in mph then the stopping distance in feet is given by

$$s = \frac{121}{3600}\frac{v^2}{\mu}.$$

If the driver's recognition of the need to stop and reaction time to apply the brakes is t then a distance of $\frac{22}{15}vt$ is travelled before the brakes are even applied. It follows that the total stopping distance is

$$s = \frac{22}{15}tv + \frac{121}{3600}\frac{1}{\mu}v^2.$$

If the road surface is good, the brakes and tires are in good shape, and stopping is achieved without locking the brakes, a coefficient of friction μ of 0.7 to 0.8 may be achieved. The recognition–reaction time t is probably around 1 second, perhaps 2 or more seconds if the driver is tired.

Suppose we determine the following stopping distances experimentally (for a given driver and vehicle).

mph	20	30	40	50	60	70	80	90
ft	50	90	130	180	225	310	380	500

Use the method of least squares to fit this data to the relation

$$s = av + bv^2.$$

Once you have computed a and b use these values to determine the recognition–reaction time t and the coefficient of friction μ.

Problem 23.2: A thermistor is a device which can be used as a temperature sensor since its resistance to an electrical current changes with temperature. It is necessary to first calibrate it. For a thermistor the theoretical relation between the absolute temperature in Kelvins T (that is, the temperature in Celcius plus 273.15), and the resistance in Ohms R, is given by

$$R = A\mathrm{e}^{\frac{B}{T}}$$

or

$$\frac{1}{T} = a + b\log R$$

where $\log R$ is the natural logarithm, for certain constants A and B or a and b. In practice the relation is not so tidy. A popular relation used in practice is the *empirical* Steinhart-Hart equation

$$\frac{1}{T} = a + b\log R + c\left(\log R\right)^3$$

where c is usually quite small compared to a and b. To *calibrate* a thermistor means to determine the coefficients a, b and c experimentally by measuring the resistance at various temperatures and then fitting the Steinhart-Hart relation. While the method of least squares would be reasonable, often just 3 data points are used, and the corresponding linear equations are solved! The calibration results are usually provided by the manufacturer, who will probably also warn that the Steinhart-Hart relation is unreliable for temperatures outside the range used in calibration (a sure sign of *ad hoc* fitting).

For examples of real Steinhart-Hart coefficients see

`http://www.atpsensor.com/ntc/steinhart/steinhart.html`

Now both R and T have a distinguished 0 point so we do not need to worry about translation invariance, but it is hard to believe that Ohms and degrees Kelvin have some special physical significance. In other words, the form of the equation ought to be invariant under a change of units. A change of units, for example Ohms (Ω) to kilo-Ohms (kΩ), in the Steinhart-Hart relation will introduce a $(\log R)^2$ term. On the principle that the form of the relation should be independent of the choice of units we could introduce this extra term from the beginning. Thus we have the *modified* Steinhart-Hart relation

$$\frac{1}{T} = c_1 + c_2\log R + c_3(\log R)^2 + c_4(\log R)^3.$$

In general, nothing is really gained by this refinement since the Steinhart-Hart relation is simply a conventional (deeply ingrained) way to interpolate (approximately) between measured data points (and gives good results in practice).

Given the following experimental data for a given thermistor calibrate it, that is, use least squares fitting to determine the Steinhart-Hart coefficients a, b and c.

Temperature in Kelvins	329.19	314.82	307.41	302.10	298.15	295.02	288.37
Resistance in Ohms	8,000	15,000	20,000	25,000	30,000	35,000	50,000

Note the resistance at $25°$ C ($298.15°$ K), in the above case 30 kΩ, is called the *reference* resistance of the thermistor.

In the next five *drill* problems you are asked to compare the values at the nodes of a least squares approximation with the given data. It would be best to do the comparison graphically if possible.

Problem 23.3: Given the following data

x	0	1	2	3	4	5
y	2	3	4	5	1	2

do a least squares fit to the equation

$$y = a + bx.$$

Compare the values of this least squares polynomial at the nodes with the given values.

Problem 23.4: Given the following data

x	0	1	2	3	4	5
y	2	3	4	5	1	2

do a least squares fit to the equation

$$y = a + bx + cx^2.$$

Compare the values of this least squares polynomial at the nodes with the given values.

Problem 23.5: Given the following data

x	0	1	2	3	4	5
y	2	3	4	5	1	2

do a least squares fit to the equation

$$y = a + bx + cx^2 + dx^3.$$

Compare the values of this least squares polynomial at the nodes with the given values.

Problem 23.6: Given the following data

x	0	1	2	3	4	5
y	2	3	4	5	1	2

do a least squares fit to the equation

$$y = ax + bx^3 + cx^5.$$

Compare the values of this least squares polynomial at the nodes with the given values.

Problem 23.7: Given the following data

x	0	1	2	3	4	5
y	2	3	4	5	1	2

do a least squares fit to the equation

$$y = a + bx + cx^2 + dx^3 + ex^4 + fx^5.$$

Compare the values of this least squares polynomial at the nodes with the given values. What happened? Why?

Problem 23.8: Suppose we wish to fit the data

x	1.0	2.0	3.0	4.0	5.0	6.0	7.0
y	5.2	6.4	7.9	9.8	12.1	14.9	18.4

to the equation

$$y = M\mathrm{e}^{Ax}$$

by using the method of least squares. We find that we need to solve an unpleasant system of nonlinear equations. To obtain a more tractable problem we transform the equation to

$$\log y = \log M + Ax$$

and solve the corresponding *linear* least squares problem - don't forget to transform the data too!

x	1.0	2.0	3.0	4.0	5.0	6.0	7.0
$\log y$	1.65	1.86	2.07	2.28	2.49	2.70	2.91

This method is known as *transformation*. It is important to realize that it is an act of desperation! The solution of the transformed problem in general will not agree with the solution of the original problem (assuming we can even find it).

Solve the transformed problem, exponentiate to obtain the desired relation and compare the values at the nodes with the given (original) data.

Problem 23.9: Given the following data

x	0	1	2	3	4	5
y	4	3	2	1	1	1

do a least squares fit to the equation

$$y = \frac{1}{a + bx}$$

by first transforming the problem to a linear least squares problem. Compare the values of the least squares relation at the nodes with the given values, graphically if possible.

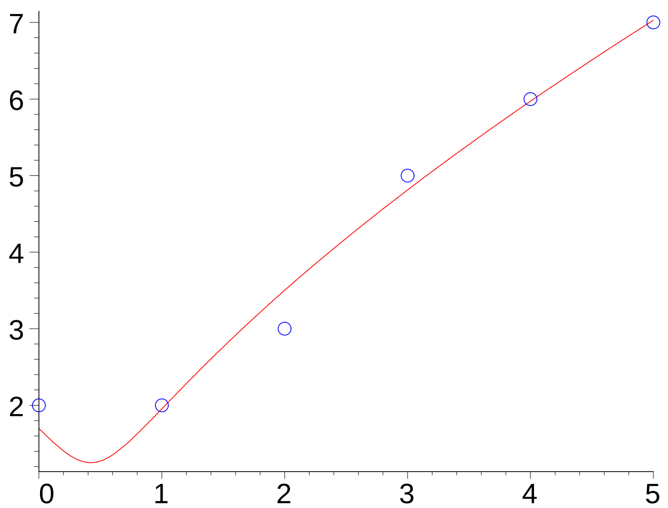
Problem 23.10: Given the following data

x	0	1	2	3	4	5
y	2	2	3	5	6	7

do a least squares fit to the equation

$$y^3 = a + bx + cx^2$$

by first transforming the problem to a linear least squares problem. Compare the values provided by the least squares equation at the nodes with the given values, graphically if possible. As a check on your work, compare your result with my graph below.



Problem 23.11: Find a and b so that the relation

$$y = a\,x + b\,\log(x)$$

best-fits the data points

$$[1, 2],\ [2, 4],\ [3, 7]$$

in the sense of least squares. **Ans:** $a = 1.95201, b = 0.78422$

24 Numeric quadrature. Compound NEWTON–COTES methods

Problem 24.1: Use SIMPSON's rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 24.2: Use the trapezoidal rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 24.3: **(A)** Use SIMPSON's rule with 2 subintervals (3 function evaluations) to estimate $\int_0^2 x^4\,dx$. **(B)** Use SIMPSON's rule with 4 subintervals to estimate the same integral. **(C)** RICHARDSON's extrapolation suggests that $R_4 = \frac{1}{15}(16S_4 - S_2)$ may be a better estimate. Compute R_4. **(D)** Compute the actual error in each of the above parts.

Problem 24.4: We know

$$\pi = 4\int_0^1 \frac{dx}{1+x^2}.$$

Use SIMPSON's rule with 4 subintervals to obtain a rational approximation to π, i.e., an approximation of the form $\pi \approx ????/2550$. Estimate the error.

Problem 24.5: Use SIMPSON's rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 24.6: Use the trapezoidal rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 24.7: The function f has the values shown in the table below.

x	0	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000
$f(x)$	2.802	2.873	2.911	3.109	3.002	2.999	2.981	2.786	2.544

Estimate the integral $\int_0^1 f(x)\,dx$ as follows: (A) Using the trapezoidal rule with 4 subintervals, (B) Using Simpson's rule with 2 subintervals, (C) Using Simpson's rule with 4 subintervals. In all four cases round your answer to 3 decimal places.

Problem 24.8: Let T_n be the (compound) trapezoidal rule with n subintervals (so $n+1$ nodes). Let S_n be the compound SIMPSON's rule with n subintervals (n even). Suppose for a certain function f on an interval $[a,b]$ we have $T_3(f) = 8.812138$, $T_6(f) = 5.893109$ and $T_{12}(f) = 5.109739$. Compute $S_6(f)$ and $S_{12}(f)$. **Ans:** 4.920099, 4.848616

Problem 24.9: Use T_2 to estimate $\int_1^3 x^4 \, \mathrm{d}x$. **Ans:** 57

Problem 24.10: Use S_2 to estimate $\int_2^4 x^5 \, \mathrm{d}x$. **Ans:** 676

25 Numeric quadrature. RICHARDSON extrapolation and ROMBERG quadrature

Problem 25.1: Consider the elementary integral

$$\int_0^1 x^4 \, dx = 0.2$$

Let T_1, T_2 and T_4 be the trapezoidal estimates of this integral for 1,2 and 4 subintervals, respectively. If we apply RICHARDSON's extrapolation we obtain the SIMPSON's rule estimates S_2 and S_4. If we extrapolate once again we obtain BODE's rule B_4. **[a]** Compute T_1, T_2, T_4, S_2, S_4 and B_4 and the corresponding actual errors. **[b]** Comment on the observed errors.

Problem 25.2: Consider the integral

$$\int_{-\pi/2}^{\pi/2} \cos(x) \, dx = 2.0000$$

Let T_2 and T_4 be the trapezoidal estimates of this integral for 2 and 4 subintervals, respectively. If we apply RICHARDSON's extrapolation we obtain the SIMPSON's rule estimate S_4. Compute T_2, T_4 and S_4.

Problem 25.3: Use SIMPSON's rule with 3 function evaluations to obtain an estimate S_2 of the integral $\int_0^\pi \sin(x) \, dx$. Use SIMPSON's rule with 5 function evaluations to obtain an estimate S_4 of the integral $\int_0^\pi \sin(x) \, dx$. RICHARDSON's extrapolation suggests that $R_4 = \frac{1}{15}(16S_4 - S_2)$ may be a better estimate than S_4. Compute R_4 and the error in each of S_4 and R_4.

Problem 25.4: **(A)** Use SIMPSON's rule with 2 subintervals (3 function evaluations) to estimate $\int_0^2 x^5 \, dx$. **(B)** Use SIMPSON's rule with 4 subintervals to estimate the same integral. **(C)** RICHARDSON's extrapolation suggests that $R_4 = \frac{1}{15}(16S_4 - S_2)$ may be a better estimate. Compute R_4. **(D)** Compute the actual error in each of the above parts. **(E)** Do your results agree with your expectation?

Problem 25.5: Consider the elementary integral

$$\int_0^1 x^4 \, dx = 0.2$$

Let T_1, T_2 and T_4 be the (ccompound) trapezoidal estimates of this integral for 1,2 and 4 subintervals, respectively. If we apply Richardson's extrapolation we obtain the Simpson's rule estimates S_2 and S_4. If we extrapolate once again we obtain Bode's rule B_4. **[A]** Compute T_1, T_2, T_4, S_2, S_4 and B_4 and the corresponding actual errors. **[B]** Comment on the observed errors.

Problem 25.6: Consider the integral

$$\int_{-\pi/2}^{\pi/2} \cos(x) \, dx = 2.0000$$

Let T_2 and T_4 be the (compound) trapezoidal estimates of this integral for 2 and 4 subintervals, respectively. If we apply Richardson's extrapolation we obtain the SIMPSON's rule estimate S_4. Compute T_2, T_4 and S_4.

Problem 25.7: Use SIMPSON's rule with 3 function evaluations to obtain an estimate S_2 of the integral $\int_0^\pi \sin(x)\, dx$. Use SIMPSON's rule with 5 function evaluations to obtain an estimate S_4 of the integral $\int_0^\pi \sin(x)\, dx$. RICHARDSON's extrapolation suggests that $B_4 = \frac{1}{15}(16S_4 - S_2)$ may be a better estimate than S_4. Compute B_4 and the error in each of S_4 and B_4.

Problem 25.8: **(A)** Use SIMPSON's rule with 2 subintervals (3 function evaluations) to estimate $\int_0^2 x^4\, dx$. **(B)** Use SIMPSON's rule with 4 subintervals to estimate the same integral. **(C)** RICHARDSON's extrapolation suggests that $B_4 = \frac{1}{15}(16S_4 - S_2)$ may be a better estimate. Compute B_4. **(D)** Compute the actual error in each of the above parts.

Problem 25.9: Denote by T_n the compound trapezoidal rule with n subintervals (for some function f on some interval $[a, b]$). If n is even we can apply RICHARDSON extrapolation to obtain SIMPSON's rule $S_n = \frac{4}{3}T_n - \frac{1}{3}T_{n/2}$. If n is divisible by 4 we can extrapolate once again to obtain BODE's rule $B_n = \frac{16}{15}S_n - \frac{1}{15}S_{n/2}$. Given

$$T_1 = 1.1063582750, \quad T_2 = 0.8005457781, \quad T_3 = 0.7517873921, \quad T_4 = 0.7354090855, \quad T_5 = 0.7279356090, \cdots$$

compute B_4.

Problem 25.10: For a certain function f we have trapzoidal estimates of $\displaystyle\int_1^2 f(x)\, \mathrm{d}x$ as follows: $T_2 = 0.33189$, $T_4 = 0.35947$ and $T_8 = 0.36648$. Use extrapolation to compute the SIMPSON's estimate S_4.

Problem 25.11: Let $T_n(f)$ denote the trapezoidal quadrature formula for the function f, with n equal length subintervals. In class we saw if n is even and we perform a certain Richardson extrapolation we obtain SIMPSON's rule in the form

$$S_n(f) = \frac{4}{3}T_n(f) - \frac{1}{3}T_{n/2}(f).$$

If n is divisible by 4 we can repeat the process and we obtain BODE's rule in the form

$$B_n(f) = \frac{16}{15}S_n(f) - \frac{1}{15}S_{n/2}(f).$$

Suppose $\int_0^1 f(x)\,\mathrm{d}x = 0.49979926915939992782 58\cdots$ and suppose $T_1(f) = 0.85826285$, $T_2(f) = 0.57944883$, and $T_4(f) = 0.51924378$. **Part (A):** Compute $S_4(f)$ and $B_4(f)$. **Part (b):** Compute and compare the errors in T_4, S_4 and B_4.

Problem 25.12: If we approximate the integral $\int_a^b f(x)\,\mathrm{d}x$ of a certain function f on the interval $[a, b]$ by the compound trapezoidal rule with n subintervals, T_n, for $n = 12$ and $n = 24$ we find

$$T_{12} = 3.56782, \quad T_{24} = 3.82993.$$

Use this information to compute the SIMPSON's rule estimate S_n with $n = 24$ subintervals.

Ans: 3.48045

26 Numeric quadrature. Miscellaneous questions

Problem 26.1: We know

$$x - \frac{x^3}{6} \le \sin(x) \le x - \frac{x^3}{6} + \frac{x^5}{120}, \quad \text{for} \quad x \ge 0.$$

Use this inequality (and only this inequality) to obtain an estimate of

$$\int_0^1 \frac{\sin(x)}{x}\,dx$$

with an error no bigger than 10^{-3}. Be sure to explain what you are doing and to back up your claim that the error is no bigger than 10^{-3}.

Problem 26.2: Determine a, b and c such that the quadrature formula

$$Q(f) = a\,f(0) + b\,f\left(\frac{1}{2}\right) + c\,f(1)$$

is exact for the integral

$$\int_0^1 f(x)\,x^3\,\mathrm{d}x$$

if $f(x)$ is a polynomial of degree ≤ 2, that is, is exact for $f(x) = 1$, $f(x) = x$ and $f(x) = x^2$. Find the error in $Q(x^4)$, as an approximation to $\int_0^1 x^7 \, \mathrm{d}x$.

Problem 26.3: Determine c_1, c_2, c_3 and c_4 such that the qudrature formula

$$Q(f) = c_1 \, f(0) + c_2 \, f(1) + c_3 \, f(3) + c_4 \, f(4)$$

is exact for the integral

$$\int_0^\infty f(x) \, \mathrm{e}^{-x} \, \mathrm{d}x$$

if $f(x)$ is a polynomial of degree ≤ 3, that is, is exact for $f(x) = 1$, $f(x) = x$, $f(x) = x^2$ and $f(x) = x^3$.

Problem 26.4: Consider the quadrature method

$$I(f) \; = \; \frac{5}{18} f\left(\frac{1}{2}\,\frac{\sqrt{5}-\sqrt{3}}{\sqrt{5}}\right) \; + \; \frac{4}{9} f\left(\frac{1}{2}\right) \; + \; \frac{5}{18} f\left(\frac{1}{2}\,\frac{\sqrt{5}+\sqrt{3}}{\sqrt{5}}\right)$$

for approximating the integral $\int_0^1 f(x) \; \mathrm{d}x$. **Part (A):** Given the table

$\frac{1}{2}\,\frac{\sqrt{5}\pm\sqrt{3}}{\sqrt{5}} = \frac{1}{2} \pm \frac{1}{10}\sqrt{15}$	$\left(\frac{1}{2}\,\frac{\sqrt{5}\pm\sqrt{3}}{\sqrt{5}}\right)^4 = \frac{31}{100} \pm \frac{2}{25}\sqrt{15}$
$\left(\frac{1}{2}\,\frac{\sqrt{5}\pm\sqrt{3}}{\sqrt{5}}\right)^2 = \frac{2}{5} \pm \frac{1}{10}\sqrt{15}$	$\left(\frac{1}{2}\,\frac{\sqrt{5}\pm\sqrt{3}}{\sqrt{5}}\right)^5 = \frac{11}{40} \pm \frac{71}{1000}\sqrt{15}$
$\left(\frac{1}{2}\,\frac{\sqrt{5}\pm\sqrt{3}}{\sqrt{5}}\right)^3 = \frac{7}{20} \pm \frac{9}{10}\sqrt{15}$	$\left(\frac{1}{2}\,\frac{\sqrt{5}\pm\sqrt{3}}{\sqrt{5}}\right)^6 = \frac{61}{250} \pm \frac{63}{1000}\sqrt{15}$

compute the exact error in $I\left(x^p\right)$ for $p = 0, 1, 2, 3, 4, 5, 6$. (Simplify your calculations by noting some cancellations.) **Part (B):** If $P(x)$ is a polynomial of degree $\leq m$ what is the largest value of m for which you can guarantee that $I(P) = \int_0^1 P(x) \; \mathrm{d}x$, that is, for which $I(P)$ is exact?

27 GAUSS–LEGENDRE quadrature

Problem 27.1: GAUSS quadrature on $[-1, 1]$ with three nodes is given by

$$G_3(f) = \frac{5}{9}\, f\left(\frac{-\sqrt{15}}{5}\right) + \frac{8}{9}\, f(0) + \frac{5}{9}\, f\left(\frac{\sqrt{15}}{5}\right).$$

Use G_3 to estimate $\int_{-1}^{1} x^6 + 9,793,238,462,643,383,279,502,884,197\,x^5 \;\mathrm{d}x$ **Ans:** $\frac{6}{25}$

Problem 27.2: The 3−point GAUSS quadrature method normalized for the interval $[0,4]$ is given by

$$G_3(f) = \frac{10}{9}\,f\left(2-\frac{2}{5}\sqrt{15}\right)+\frac{16}{9}\,f(2)+\frac{10}{9}\,f\left(2+\frac{2}{5}\sqrt{15}\right).$$

Without doing too much calculation compute $G_3\left(x^5\right)$.

Problem 27.3: Let f be continuous on $[a,b]$, $n \geq 1$ and $h = \dfrac{b-a}{n}$. Let $x_k = a + k\,h$ for $k = 0, 1, \cdots, n$. The (compound) trapezoidal rule is given by

$$T_n(f) = \frac{h}{2}\left[\,f(x_0)+2f(x_1)+2f(x_2)+\,\cdots\,+2f(x_{n-1})+f(x_n)\right].$$

Simpson's rule (n even) is given by

$$\begin{aligned} S_n(f) &= \frac{1}{3}\left[\,4\,T_n(f) - T_{\frac{n}{2}}(f)\right] \\ &= \frac{h}{3}\left[\,f(x_0)+4\,f(x_1)+2\,f(x_2)+4\,f(x_3)+\,\cdots\,+4\,f(x_{n-1})+f(x_n)\right] \end{aligned}$$

Part (A): Use SIMPSON's rule with 2 subintervals to estimate $\int_0^2 x^5\,dx$. **Part (B):** Use SIMPSON's rule with 4 subintervals to estimate the same integral. **Part (C):** RICHARDSON's extrapolation suggests that BODE's rule $B_4 = \frac{1}{15}(16S_4 - S_2)$ may be a better estimate than S_4. Compute B_4. **Part (D):** Compute the actual error in each of the above parts and comment on it. **Part (E):** Using only mental calculation deduce the approximation provided by GAUSS quadrature with 3 nodes. Explain.

Problem 27.4: GAUSS quadrature with 3 nodes re-scaled to the interval $[0,1]$ is given by

$$G(f) = \frac{5}{18}\,f\left(\frac{5-\sqrt{15}}{10}\right)\,+\,\frac{4}{9}\,f\left(\frac{1}{2}\right)\,+\,\frac{5}{18}\,f\left(\frac{5+\sqrt{15}}{10}\right).$$

Find $G(x^5 + x^3)$ without doing much computing (think first). Simplify. **Ans:** $\frac{5}{12}$

28 Miscellaneous Problems

Problem 28.1: Write a brief technical essay, in decent English, describing the topic in Mth 351 which you most enjoyed. Your essay should contain some relevant technical facts or equations.