Elementary Numerical Analysis

Bent E. Petersen

Summer 2004 – 20040806

Contents

This list of problems will be augmented during the quarter. Some 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, of other material on my web pages. Other problems may seem familiar because I used them as examples in my lectures.

Some of the problems here may be used with minor editing on future tests. Other problems, especially those which are pretty theoretical, are for instruction and amusement.

If you have difficulties with any of the problems then ask me for hints or clarification, but don't wait until the end of the term!

1 TAYLOR Polynomials

Problem 1.1. We know from calculus that EULER's number e is given by

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

Deduce $2.5 \leq \mathrm{e} \leq 2.75$. **Hint** Show $k! \geq 2(3^{k-2})$ if $k \geq 3$.

Problem 1.2. Suppose we use the TAYLOR sum

$$\sum_{k=0}^{n} \frac{1}{k!} \frac{1}{2^k}$$

to approximate $\sqrt{\mathrm{e}} = \mathrm{e}^{1/2}$. How large should we take n to guarantee the error is no larger than 10^{-12}? **Hint** You may find it useful to note $\mathrm{e}^{1/2} \leq 2$.

Problem 1.3. 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 1.4. 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 1.5. 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 1.6. 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 1.7. 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 1.8. 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 1.9. 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 1.10. 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 $\left| f^{(5)}(x) \right| \leq 0.005$ for $|\, x \,| \leq 0.5$ 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 1.11. Suppose $f(0) = 1$, $f'(0) = -2$, $f''(0) = -\frac{1}{5}$, $f'''(0) = \frac{2}{3}$ and $\left| f^{(4)}(x) \right| \leq 0.016$ for $0 \leq x \leq 2$. Estimate $f(1)$ by using a TAYLOR polynomial of degree 3. Compute a good bound for the absolute error.

Problem 1.12. 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 1.13. 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 1.14. We know

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

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 1.15. **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 1.16. 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 1.17. 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 1.18. 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 1.19. Show $n! \geq 2^{n-1}$ if $n \geq 2$. Then use

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

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

Problem 1.20. 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 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 1.21. 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

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

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 1.22. 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 1.23. 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 1.24. 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 1.25. 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 $-\frac{1}{2} \leq x \leq 0$.

---

2 Errors

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 called the *number of significant decimal digits in* x_a *as an approximation to* x_t or simply the number of *significant digits* when x_t and *"decimal"* are understood. It is clearly given by

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

Here $\lfloor \quad \rfloor$ indicates the *greatest integer function* otherwise known as *floor*. Note this definition of significant decimal digits may not agree with the count obtained by other methods, but it is more convenient, and equally useful.

Problem 2.1. 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 2.2. Imagine an (impractical) binary floating point representation with exponent ranging from -63 to 64 and with a mantissa of length 12 bits. Assume only normalized representations are used and there is no packing.

A.) What is the largest exactly representable positive number?

B.) What is the smallest exactly representable positive number?

C.) What is the unit round?

Answers:

A.) $2^{53}\left(2^{12} - 1\right) \approx 3.69 \times 10^{19}$

B.) $2^{-63} \approx 1.08 \times 10^{-19}$

C.) $2^{-11} \approx 4.88 \times 10^{-4}$

Problem 2.3. In the previous problem suppose we implement gradual underflow by allowing denormals when the exponent is -63. Suppose also we reserve the largest exponent, 64, to designate $\pm\infty$, projective infinity, NaN's and other special data. How does this affect the answers to the previous problem?

Problem 2.4. Imagine an (impractical) binary floating point representation with exponent ranging from -63 to 64 and with a mantissa of length 12 bits. Assume only normalized representations are used and the format is packed (that is, the most significant bit in the mantissa is not explicitly stored).

A.) What is the largest exactly representable positive number?

B.) What is the smallest exactly representable positive number?

C.) What is the unit round?

Problem 2.5. In the previous problem suppose we implement gradual underflow by allowing denormals when the exponent is -63. Suppose also we reserve the largest exponent, 64, to designate $\pm\infty$, projective infinity, NaN's and other special data. How does this affect the answers to the previous problem? Note that numbers with exponent -63 are not packed (to allow for denormals).

Problem 2.6. Imagine an (impractical) binary floating point representation with exponent ranging from -63 to 64 and with a mantissa of length 12 bits. Assume only normalized representations are used and there is no packing. Find the error that we make when storing 113.2 (decimal)? **Ans:** $1/80 = 0.0125$.

Problem 2.7. In the previous problem if the length of the mantissa is only 7 bits, what would be the error in storing 113.2?

---

3 Numeric Differentiation

Problem 3.1. Here is an unusual second order method for approximating the first derivative

$$f'(a) = \frac{-5f(a) + 9f(a+2h) - 4f(a+3h)}{6h} + \mathcal{O}\left(h^2\right).$$

Show that the error term can be written as

$$cf^{(3)}(a)h^2 + \mathcal{O}\left(h^3\right)$$

and compute the constant c. **Hint:** Use Taylor expansions.

Problem 3.2. Here is an unusual first order method for approximating the second derivative

$$f''(a) = \frac{f(a) - 3f(a+2h) + 2f(a+3h)}{3h^2} + \mathcal{O}\left(h\right).$$

Show that the error term can be written as

$$cf^{(3)}(a)h + \mathcal{O}\left(h^2\right)$$

and compute the constatnt c. **Hint:** Use TAYLOR expansions.

Here are a few standard numeric differentiation formulæ:

FIRST DERIVATIVE

$$\begin{aligned}
f'(a) &= \frac{f(a+h)-f(a)}{h} + \mathcal{O}(h) \\
f'(a) &= \frac{f(a+h)-f(a-h)}{2h} + \mathcal{O}(h^2), \text{ central 3-point} \\
f'(a) &= \frac{-f(a+2h)+4f(a+h)-3f(a)}{2h} + \mathcal{O}(h^2), \text{ forward 3-point} \\
f'(a) &= \frac{-f(a+2h)+8f(a+h)-8f(a-h)+f(a-2h)}{12h} + \mathcal{O}(h^4), \text{ central 5-point}
\end{aligned}$$

SECOND DERIVATIVE

$$\begin{aligned}
f''(a) &= \frac{f(a+h)-2f(a)+f(a-h)}{h^2} + \mathcal{O}(h^2), \text{ central 3-point} \\
f''(a) &= \frac{-f(a+3h)+4f(a+2h)-5f(a+h)+2f(a)}{h^2} + \mathcal{O}(h^2), \text{ forward 4-point}
\end{aligned}$$

THIRD DERIVATIVE

$$\begin{aligned}
f'''(a) &= \frac{f(a+2h)-2f(a+h)+2f(a-h)-f(a-2h)}{2h^3} + \mathcal{O}(h^2), \text{ central 5-point} \\
f'''(a) &= \frac{-3f(a+4h)+14f(a+3h)-24f(a+2h)+18f(a+h)-5f(a)}{2h^3} \\
&\quad +\mathcal{O}(h^2), \text{ forward 5-point}
\end{aligned}$$

Problem 3.3. Given $f(1.8) = 2.34$, $f(1.9) = 2.21$, $f(2.0) = 2.12$ and $f(2.1) = 2.07$ estimate $f'(2.0)$ and $f''(2.0)$ by using central second order estimates.

Problem 3.4. 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 $f'(0.875)$ and $f''(0.875)$ by using second order central estimates.

Problem 3.5. 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 central 5-point method listed above with step size 0.125? Use the same step size and compute the estimate given by order 2 central 3-point method. Explain why there is no reason to expect one result to be better than the other for this problem.

---

4 Roots. Midpoint Method

Problem 4.1. 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 4.2. 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 4.3. 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 4.4. 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 4.5. 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 4.6. 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.

---

5 Roots. NEWTON's Method

This section deals with the NEWTON root estimate iteration

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

Problem 5.1. 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 5.2. 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 5.3. 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 5.4. 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 5.5. 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 5.6. 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 5.7. 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 $-1.627772774213681160\ldots$ is the only real root of $p(x)$ find the actual errors in the root estimates. Do they behave as you would expect?

Problem 5.8. 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 5.9. 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 5.10. 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.

In applying NEWTON's method we may not wish to compute the derivative at each step. Thus we could try "freezing it" - say

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}.$$

This idea is probably not a good one in the scalar case here, but it has advantages in dealing with solving nonlinear *systems*.

Problem 5.11. The polynomial $p(x) = x^3 - 3x^2 + 2x - 4$ has a simple root at approximately 2.7963219032594415351 . Use NEWTON's method with an initial guess $x_0 = 3.0$ and 5 iterations to estimate the root. Compute the error in each iterate and comment on the convergence.

Problem 5.12. Repeat the previous problem but using the modified NEWTON's method described above (with the derivative frozen at 3.0). Compare the apparent rates of convergence.

---

6 Roots. HALLEY's Method

In class we derived NEWTON's method from the tangent line approximation to a graph. We can also derive it directly from a TAYLOR expansion. If x_n is an approximation to a root and we write the root as $x_n + \delta$ we have

$$0 = f(x_n + \delta) = f(x_n) + f'(\xi_n)\delta$$

for some ξ_n between x_n and $x_n + \delta$. Thus

$$\delta = -\frac{f(x_n)}{f'(\xi_n)}$$

and we expect, when all goes well, $\delta \approx \delta_n$ where

$$\delta_n = -\frac{f(x_n)}{f'(x_n)}.$$

Accordingly, the next root estimate is

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

HALLEY's method results from taking a higher order TAYLOR expansion

$$0 = f(x_n + \delta) = f(x_n) + f'(x_n)\delta + \frac{1}{2} f''(\eta_n)\delta^2$$

form some η_n between x_n and $x_n + \delta$. This equation is quadratic in δ. We could use the quadratic formula, but we would probably get a headache. Instead note the δ which satisfies $f(x_n + \delta) = 0$ is given by

$$\delta = -\frac{f(x_n)}{f'(\xi_n)}$$

where of course we do not know ξ_n. If we substitute this expression for one of the factors of δ in the δ^2 term we obtain

$$0 = f(x_n) + f'(x_n)\delta - \frac{f(x_n)f''(\eta_n)}{2f'(\xi_n)}.$$

It follows that

$$\delta = -\frac{2f(x_n)f'(\xi_n)}{2f'(x_n)f'(\xi_n) - f(x_n)f''(\eta_n)}.$$

Thus, when all works well, we expect $\delta \approx \Delta_n$ where

$$\Delta_n = -\frac{2f(x_n)f'(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)}.$$

Setting $x_{n+1} = x_n + \Delta_n$ as the next root estimate we obtain HALLEY's iteration

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \left(1 - \frac{f(x_n)f''(x_n)}{2f'(x_n)^2}\right)^{-1}.$$

If we start sufficiently close to a simple root α then some heavy algebra and analysis shows HALLEY's method yields

$$K(\alpha - x_{n+1}) \approx (K(\alpha - x_n))^3$$

for a certain constant K, that is, HALLEY's method is of order 3. In more prosaic terms, if we start sufficiently close to a simple root, then after things settle down we will see roughly a tripling of the number of significant digits with each iteration!

Problem 6.1. The increasing function f defined by

$$f(x) = x - \cos(x)$$

has a simple root given approximately by

$$\alpha = 0.73908513321516064165531208767387340401341175890075746 5$$

Starting with an initial guess of $x_0 = 0.1$ compute a few NEWTON and HALLEY iterations and the error in each. Comment on your results. Are your results consistent with your expectations? **Note:** The error in x_4 for NEWTON's method is about -1.03×10^{-11} and for HALLEY's method it is about 2.37×10^{-49}. If you are using the limited precision of an ordinary calculator there is not much point in going past x_2 or x_3 in this example.

7 Roots. Secant Method

Problem 7.1. 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 7.2. 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 $-1.627772774213681160\ldots$ is the only real root of $p(x)$ find the actual errors in the root estimates. Do they behave as you would expect?

---

8 Roots. Miscellaneous

Problem 8.1. 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 8.2. 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 8.3. Let $f(x) = \exp(x) - x^2$. Prove that f has precisely one real root. Explain why the root must be in the interval $(-1, 0)$. **Hint:** Sign change.

Problem 8.4. 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 $\alpha'(0)$. Use your result to estimate a root of $p(x) = x^3 + 1.1x^2 - x - 1$.

---

9 Fixed Point Iteration

Problem 9.1. 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 9.2. 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 9.3. Plot $y = \cos(\sin(x))$ and $y = x$. Estimate the point where the graphs intersect as an estimate of a fixed point for $\cos(\sin(x))$ Deduce also that $\cos(\sin(x))$ has a unique fixed point by showing $x - \cos(\sin(x))$ is increasing. Do a few fixed point iterations to improve your fixed point estimate. **Remark** the fixed point is about $0.768169156\ldots$.

Problem 9.4. Plot $y = \cos(\cos(x))$ and $y = x$. Estimate the point where the graphs intersect as an estimate of a fixed point for $\cos(\cos(x))$. Deduce also that $\cos(\cos(x))$ has a unique fixed point by showing $x - \cos(\cos(x))$ is increasing. Do a few fixed point iterations to improve your fixed point estimate. **Remark** the fixed point is about $0.739085133\ldots$.

Problem 9.5. 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 a root finding iteration to approximate a root of $f(x) - x$. **Ans:** $1.93456321\ldots$.

---

10 NEWTON's Method for Systems

Warning: The problems in this section are very heavy computationally. Do not plan on zipping through them!

The "frozen coefficient matrix" method referred to below is just NEWTON's method with the partial derivatives evaluated just at the initial point and not updated in subsequent iterations.

Problem 10.1. 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 10.2. Repeat the previous problem, but using the "frozen coefficient matrix" algorithm. Do several more iterations. Comments?

Problem 10.3. 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 10.4. Repeat the previous problem, but using the "frozen coefficient matrix" algorithm. Do several more iterations. Comments?

Problem 10.5. 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 10.6. Repeat the previous problem, but using the "frozen coefficient matrix" algorithm. Do several more iterations. Comments?

Problem 10.7. 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 10.8. Repeat the previous problem, but using the "frozen coefficient matrix" algorithm. Do several more iterations. Comments?

Problem 10.9. 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 10.10. Repeat the previous problem, but using the "frozen coefficient matrix" algorithm. Do several more iterations. Comments?

Problem 10.11. 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 $(-1.0, -1.0)$?

Problem 10.12. Repeat the previous problem, but using the "frozen coefficient matrix" algorithm. Do several more iterations. Comments?

11 Secant Method for Systems

The secant iteration for approximating roots can be extended to systems of equations. This method has the advantage of not requiring derivatives, but it does require us to supply 3 initial points (in the case of 2 variables).

To estimate a solution to the system

$$\begin{aligned} f(x,y) &= 0 \\ g(x,y) &= 0 \end{aligned}$$

we first choose 3 points (x_1, y_1), (x_2, y_2) and (x_3, y_3) which are close to a solution. By evaluating f at these 3 points we obtain 3 points in 3-space, $(x_n, y_n, f(x_n, y_n))$, $n = 1, 2, 3$. If these points are not collinear there is a unique plane P_1 through them. Likewise from g we obtain 3 points $(x_n, y_n, g(x_n, y_n))$, $n = 1, 2, 3$. If these points are not collinear there is a unique plane P_2 through them. If the planes P_1 and P_2 are not parallel they intersect in a line and if this line is not parallel to the (x, y)-plane it meets that plane in a point (x_4, y_4). Now use the points (x_2, y_2), (x_3, y_3) and (x_4, y_4) and repeat the construction to obtain the next iterate (x_5, y_5).

Problem 11.1. Use the secant method for systems to approximate a solution to the system of equations

$$\begin{aligned} x^3 - 3y + 2 &= 0 \\ y^3 - 2x^2 + 3x + 4 &= 0. \end{aligned}$$

Use the initial points $(-0.8, 0.4)$, $(-0.8, 0.5)$ and $(-0.9, 0.5)$ and iterate three times. Given that the actual solution is $(-0.865034729, 0.450902472)$ compute the error in each of your iterates. Comment. **Remark:** As a check on your work note the first iterate is $(-0.8646900900, 0.4492075016)$.

12 Interpolation polynomials

Problem 12.1. 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 12.2. 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 12.3. 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 12.4. 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 12.5. 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 12.6. 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 12.7. 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 $|f^{(3)}(t)| \leq 2.0$ for $0 \leq t \leq 2$. Compute a good bound for the absolute error in $p(0.5)$ as an approximation to $f(0.5)$.

Problem 12.8. 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 $|f^{(4)}(t)| \leq 1.502$ for $0 \leq t \leq 3$. Compute a good bound for the absolute error in $p(1.5)$ as an approximation to $f(1.5)$.

Problem 12.9. 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 12.10. 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 $f(1)$.

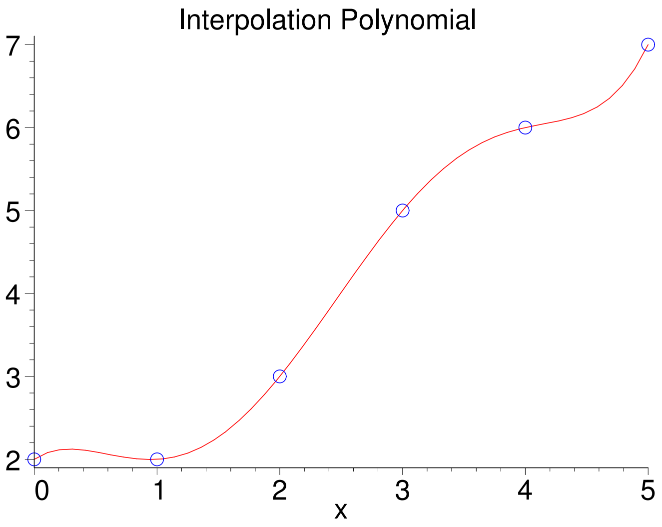
Problem 12.11. Find the interpolation polynomial $p(x)$ of degree ≤ 3 through the points $(1, -1)$, $(2, 3)$, $(3, 2)$, $(4, -2)$.

Problem 12.12. Find the interpolation polynomial of degree ≤ 3 through the points $(0,2)$, $(1,0)$, $(2,3)$ and $(3,1)$. **Ans:** $-\frac{5}{3}x^3 + \frac{15}{2}x^2 - \frac{47}{6}x + 2$

Problem 12.13. Given the following data

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

compute the interpolation polynomial of degree ≤ 5 with the given nodes. As a check on your work, compare your result with my graph below.



---

13 Interpolation polynomials and NEWTON divided differences

DIVIDED DIFFERENCES

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

where $P_k(x)$ is the interpolation polynomial (of degree $\leq k$) for the function f with nodes at $x_0, x_1, \cdots, x_k$.

A remarkable identity which holds for any function f is:

$$f(x) = P_n(x) + f[x_0, x_1, \cdots, x_n, x](x - x_0)(x - x_1)\cdots(x - x_n).$$

A consequence of this identity is if we view $P_n(x)$ as an approximation to $f(x)$ then the error is given by

$$\frac{1}{(n+1)!}\, f^{(n+1)}(\xi)\,(x - x_0)(x - x_1)\cdots(x - x_n)$$

for some ξ between $\min(x_0, x_1, \cdots, x_n, x)$ and $\max(x_0, x_1, \cdots, x_n, x)$.

Problem 13.1. 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 13.2. 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 13.3. 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 13.4. 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 13.5. 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
$f[x_{k-1}, x_k]$	×	-1	-1	
$f[x_{k-2}, x_{k-1}, x_k]$	×	×		
$f[x_{k-3}, x_{k-2}, x_{k-1}, x_k]$	×	×	×	

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 13.6. 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
$f[x_{k-1}, x_k]$	×	$1/2$	2	
$f[x_{k-2}, x_{k-1}, x_k]$	×	×		$-4/3$
$f[x_{k-3}, x_{k-2}, x_{k-1}, x_k]$	×	×	×	

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 13.7. 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 13.8. 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 13.9. 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 $(k, f(k))$, $k = 1,\ 2,\ 4,\ 5,\ 6$.

---

14 Interpolation and ČEBYŠEV polynomials

the n^{th} ČEBYŠEV polynomial $T_n(x)$ is given by

$$T_n(x) = \cos(n\theta) \text{ if } x = \cos(\theta).$$

One can show

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

One can show

$$T_{n+1}(x) = 2x\, T_n(x) - T_{n-1}(x).$$

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

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

The roots of $T_n(x)$ all lie in the interval $[-1, 1]$.

The nth ČEBYŠEV *nodes* in $[-1, 1]$, $x_0, x_1, \cdots, x_n$ are the roots of $T_{n+1}(x)$. They are given by

$$x_k = -\cos\left(\frac{2k+1}{2n+2}\pi\right), \qquad k = 0, 1, \cdots, n.$$

For other intervals we dilate and translate appropriately. Thus for the interval $[a,b]$ the ČEBYŠEV nodes are given by

$$x_k = \frac{a+b}{2} + \frac{a-b}{2}\cos\left(\frac{2k+1}{2n+2}\pi\right), \qquad k = 0, 1, \cdots, n.$$

Problem 14.1. Let

$$f(x) = \frac{1}{1+60x^2}.$$

Construct the interpolation polynomial for $f(x)$ with 13 equispaced nodes in the interval $[-1,1]$ and with 13 ČEBYŠEV nodes (so $n = 12$) relative to the interval $[-1,1]$. View each interpolation polynomial as an approximation of $f(x)$ and estimate the maximum absolute error in $[-1,1]$ graphically.
Ans. As a check on your work here is the interpolation polynomial with equispaced nodes,

$$\begin{aligned} &\frac{922640625}{465796}x^{12} - \frac{4695215625}{931592}x^{10} + \frac{34832075625}{7452736}x^{8} - \frac{7326034875}{3726368}x^{6} \\ + &\frac{2848365225}{7452736}x^{4} - \frac{59056665}{1863184}x^{2} + 1. \end{aligned}$$

If you use a calculator your result may contain some odd degree terms with small coefficients due to roundoff. The error in the equispaced case is about 8.0. In the ČEBYŠEV case the error is much smaller.

15 Interpolation Splines

Problem 15.1. Suppose

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

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

Problem 15.2. 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 15.3. Suppose

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

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

Problem 15.4. 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 15.5. 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 15.6. 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 15.7. 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 15.8. (Deleted)

Problem 15.9. 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 15.10. (Deleted)

Problem 15.11. 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 15.12. **(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 15.13. Let $s(x)$ be the natural cubic interpolation spline with knots $(0, 2)$, $(1, 0)$, $(2, 3)$ and $(3, 1)$. Suppose

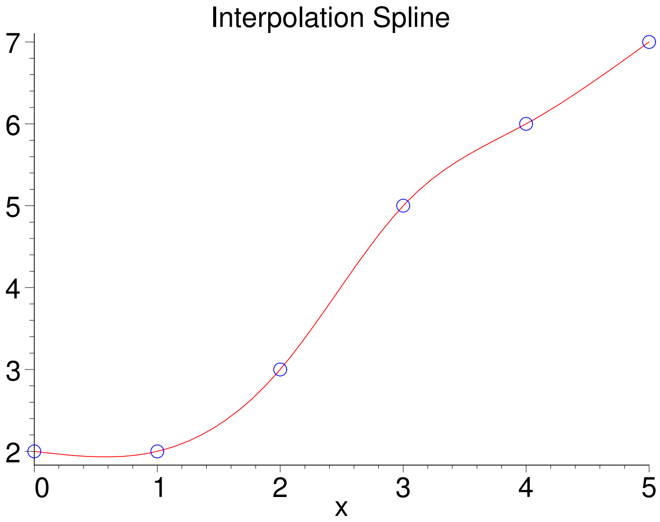
$$s(x) = \begin{cases} 2 - \frac{11}{3}x + \frac{5}{3}x^3 & \text{if } 0 \le x < 1 \\ \\ 7 - \frac{56}{3}x + 15x^2 - \frac{10}{3}x^3 & \text{if } 1 \le x < 2 \\ \\ -33 + \frac{124}{3}x + Ax^2 + Bx^3 & \text{if } 2 \le x \le 3 \end{cases}$$

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

Problem 15.14. Given the following data

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

do a natural cubic spline fit to the data. (Note the calculations can be done by hand, but are rather lengthy. You can do the problem easily in Maple, but of course you have to find out how first. Have fun.) As a check on your work, compare your result with my graph below.



---

16 Least Squares Fitting

Problem 16.1. 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 16.2. 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

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

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 16.3. 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 16.4. 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 16.5. 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 16.6. 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 16.7. 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 16.8. 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 16.9. 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.

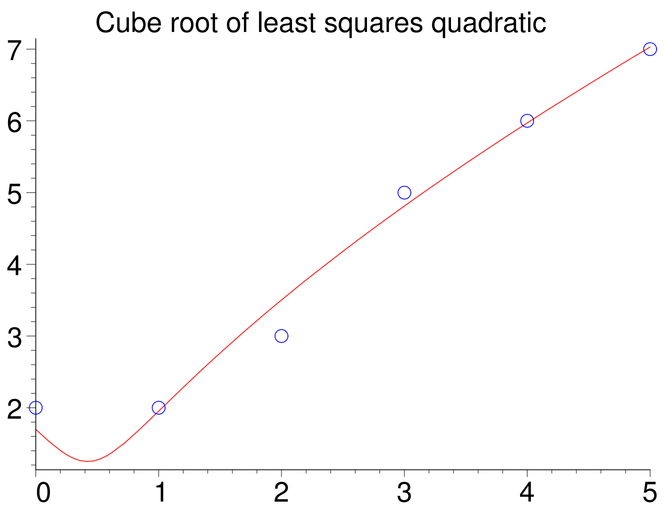
Problem 16.10. Given the following data

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 16.11. 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:** $a = 1.95201$, $b = 0.78422$

---

17 Compound Newton-Cotes Quadrature Methods

NEWTON-COTES numeric quadrature to approximate

$$\int_a^b f(x)\ \mathrm{d}x$$

of degree m with n subintervals, where n must be divisible by m may be described as follows. Let the *step-size* h be given by

$$h = \frac{b-a}{n}$$

and let

$$x_k = a + kh, \quad k = 0, 1, \ldots, n.$$

Now let $P_j(x)$ be the interpolation polynomial of degree $\le m$ with nodes

$$\left(x_{(j-1)m+i},\ f(x_{(j-1)m+i}\right), \quad i = 0, 1, \cdots, m$$

for $j = 1, \cdots, \frac{n}{m}$. Then the compoundNEWTON-COTES approximation is given by

$$NC(n, m, f) = \sum_{j=1}^{n/m} \int_{x_{(j-1)m}}^{x_{jm}} P_j(x)\ \mathrm{d}x.$$

Note when $m = n$ the word *compound* is dropped. If the error is $\mathcal{O}(h^p)$ then we say that the method has *order* p. If the method is exact for polynomials of degree q then we say the method is *degree q exact*. From the construction it is clear that $NC(n, m, f)$ is degree m exact, but in fact, if m is even then the method is degree $m+1$ exact! It follows then that $NC(n, m, f)$ is of order $m+1$ if m is odd and of order $m+2$ if m is even.

$NC(n, 1, f)$ is called the *compound trapezoidal rule* or the *compound 2-point rule*. We have

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

$NC(n, 2, f)$ (n even) is called the *compound* SIMPSON*'s rule* or the *comound 3-point rule*. We have

$$NC(n, 2, f) = \frac{h}{3}\ \left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 4f(x_{n-1}) + f(x_n)\right].$$

$NC(n,3,f)$ (n divisible by 3) is called the *compound* SIMPSON*'s 3/8's rule* or the *compound 4-point rule*. We have

$$\begin{aligned} NC(n,3,f) &= \frac{3h}{8}\,[f(x_0)+3f(x_1)+3f(x_2)+2f(x_3)+3f(x_4)+3f(x_5) \\ &+ \cdots + 3f(x_{n-1})+f(x_n)]. \end{aligned}$$

$NC(n,4,f)$ (n divisible by 4) is called the *compound* BOOLE*'s rule* or the *compound 5-point rule*. We have

$$\begin{aligned} NC(n,4,f) &= \frac{2h}{45}\,[7f(x_0)+32f(x_1)+12f(x_2)+32f(x_3)+14f(x_4) \\ &+ \cdots + 32f(x_{n-1})+7f(x_n)]. \end{aligned}$$

$NC(n,5,f)$ (n divisible by 5) is called the *compound 6-point rule*. We have

$$\begin{aligned} NC(n,5,f) &= \frac{5h}{288}\,[19f(x_0)+75f(x_1)+50f(x_2)+50f(x_3)+75f(x_4)+38f(x_5)+75f(x_6) \\ &+ \cdots + 75f(x_{n-1})+19f(x_n)]. \end{aligned}$$

Since the formulae above after the first few become somewhat unwieldy it is preferable to simply give the sequence of coefficients for the non-compound method. Thus

trapezoidal rule	$\frac{h}{2}[1,1]$
SIMPSON's rule	$\frac{h}{3}[1,4,1]$
SIMPSON's 3/8's rule	$\frac{3h}{8}[1,3,3,1]$
BOOLE's rule	$\frac{2h}{45}[7,32,12,32,7]$
6-point	$\frac{5h}{288}[19,75,50,50,75,19]$
7-point	$\frac{h}{140}[41,216,27,272,27,216,41]$
8-point	$\frac{7h}{17280}[751,3577,1323,2989,2989,1323,3577,751]$
9-point	$\frac{4h}{14175}[989,5888,-928,10496,-4540,10496,-928,5888,989]$

Note the minus signs in the 9-point formula.

Problem 17.1. We know

$$\int_0^3 \mathrm{e}^x \,\mathrm{d}x \approx 19.0855369231876677409\ldots.$$

Estimate the integral using the trapezoidal, SIMPSON's, SIMPSON's 3/8's, Boole and 7-point rules, each with 12 subintervals. That is, compute each of the estimates $NC(12,1,f)$, $NC(12,2,f)$, $NC(12,3,f)$, $NC(12,4,f)$ and $NC(12,6,f)$ where $f(x) = \mathrm{e}^x$ on the interval $[0,3]$. Then compute the error in each estimate. Comments? **Remark** As a check on your work, note the error in $NC(12,6,f)$ is -2.888341×10^{-7} and the other errors are larger in absolute value.

Problem 17.2. Use SIMPSON's rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 17.3. Use the trapezoidal rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 17.4. **(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 17.5. 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 17.6. Use SIMPSON's rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 17.7. Use the trapezoidal rule with 5 function evaluations (4 intervals) to estimate $\int_0^\pi \sin(x)\,dx$. Find the actual error.

Problem 17.8. 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 17.9. 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 17.10. Use T_2 to estimate $\int_1^3 x^4\,\mathrm{d}x$. **Ans:** 57

Problem 17.11. Use S_2 to estimate $\int_2^4 x^5\,\mathrm{d}x$. **Ans:** 676

18 RICHARDSON Extrapolation and ROMBERG Qudrature

Problem 18.1. 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 BOOLE'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 18.2. 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 18.3. 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 18.4. **(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 18.5. 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 BOOLE'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 18.6. 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 18.7. 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 18.8. **(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 18.9. 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 BOOLE'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 18.10. 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 18.11. 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 BOOLE's rule in the form

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

Suppose $\displaystyle\int_0^1 f(x)\ \mathrm{d}x = 0.49979926915939992782 58\cdots$ 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 18.12. If we approximate the integral $\displaystyle\int_a^b f(x)\ \mathrm{d}x$ 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

19 GAUSS-LEGENDRE Quadrature

Problem 19.1. 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 19.2. 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 19.3. 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 19.4. 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:** $\frac{5}{12}$

---

20 Miscellaneous Quadrature Problems

Problem 20.1. We know

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

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 20.2. 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 20.3. 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 20.4. 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?

---

21 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.

The ideas described above are actually applied to linear systems even though there are deterministic ways to solve linear systems. The reason is that for very large sparse systems the iterative methods remain practical whereas the deterministic methods become unmanageable. For the sake of description let's use a small system to illustrate:

The system of linear equations

$$\begin{aligned} a_1x + b_1y + c_1z &= e_1 \\ a_2x + b_2y + c_2z &= e_2 \\ a_3x + b_3y + c_3z &= e_3 \end{aligned}$$

may be converted to a fixed point problem by pulling out the diagonal terms:

$$\begin{aligned} x &= \frac{1}{a_1}\,(e_1 - b_1y - c_1x) \\ y &= \frac{1}{b_2}\,(e_2 - a_2x - c_2z) \\ z &= \frac{1}{c_3}\,(e_3 - a_3x - b_3y)\,. \end{aligned}$$

If we apply the method of simultaneous displacements to this fixed-point problem then we obtain the GAUSS-JACOBI iteration for the original linear system. If we apply the method of successive displacements then we obtain the GAUSS-SEIDEL iteration for the linear system. Note there is another way to describe the GAUSS-SEIDEL iteration in terms of triangualr matrices which is more convenient for theoretical work, but is equivalent to our description here.

Problem 21.1. The fixed point problem

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

where

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

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. 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 21.2. Repeat the previous problem but this time use the method of successive displacements. Compare the errors to the ones observed in the previous problem.

Problem 21.3. Consider the fixed-point problem

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

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

Problem 21.5. The system of equations

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

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 21.6. 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 simultaneous displacements with 5 iterations to approximate a solution. Compute the 2-norm of the actual error in each iteration. Comment.

Problem 21.7. 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.

---

22 Vectors. Inner Products and Norms

Problem 22.1. Assume the canonical inner product in $\mathbb{R}^5$ and compute the angle (in radians) between the vectors

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

Problem 22.2. 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 (carefully in English).

Problem 22.3. 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 22.4. 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 22.5. 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 22.6. 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 $\vec{v}_1, \vec{v}_2, \vec{v}_3$?

Problem 22.7. There are several ways of establishing that any 5 vectors in $\mathbb{R}^4$ are linearly dependent. Provide one or more arguments.

---

23 Matrices

Problem 23.1. 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 23.2. 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 23.3. If $A = (a_{ij})$ is an $n \times n$ matrix the *trace* or *spur* tr(A) of A is defined by

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

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, tr(AB) and tr(BA). Do you care to make a conjecture?

Problem 23.4. 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 23.5. 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 23.6. 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 23.7. 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 23.8. 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 23.9. 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

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

Problem 23.10. 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 23.11. 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.

---

24 GAUSS-JORDAN Row Reduction

Problem 24.1. Find the complete row reduced row echelon GAUSS–JORDAN canonical form of the matrix A (henceforth abbreviated $\operatorname{rref} A$) where

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

Problem 24.2. 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 $\operatorname{rref} B$.

Problem 24.3. 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,

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

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 $\operatorname{row}(A)$, though they may change $\operatorname{col}(A)$. However, since $\dim \operatorname{row}(A) = \dim \operatorname{col}(A)$ we see that elementary row operations do not change $\dim \operatorname{col}(A)$.

Problem 24.4. 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 24.5. **Part A.** Find the row reduced echelon canonical form $\operatorname{rref} A$ 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?

---

25 LU Facrorization

Problem 25.1. 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 25.2. 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).

---

26 Positive Definite Matrices

The $n \times n$ matrix A is said to be *positive definite* if it is symmetric and if

$$x^T Ax > 0$$

for each n-vector $x \neq 0$. One can show that an $n \times n$ symmetric matrix A is positive definite if and only its principal cofactors are positive.

Problem 26.1. Show, in two different ways, that the matrix

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

is positive definite.

Problem 26.2. Find a positive definite matrix B such that $B^2 = A$ where

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

---

27 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 diagonally dominant and 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 just the *method of successive displacements* applied to the fixed point problem obtained by pulling out the diagonal. The GAUSS–JACOBI method is just the *method of simultaneous displacements* applied to the fixed point problem obtained by pulling out the diagonal.

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 27.1. 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 27.2. 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 27.3. 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 27.4. 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 27.5. 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 27.6. 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:** $(2, 9/2, -1/2)$

Bent E. Petersen
Department of Mathematics
Oregon State University
Corvallis, OR 97331-4605

24 hour phone numbers
office (541) 737-5163
fax (541) 737-0517

email: petersen@math.oregonstate.edu
web: http://oregonstate.edu/~peterseb