Queens College, CUNY, Department of Computer Science

**Numerical Methods**

**CSCI 361 / 761**

**Spring 2018**

Instructor: Dr. Sateesh Mane



# Midterm 2 Spring 2018

## due Wednesday April 25, 2018, 11:59 pm

- **NOTE: It is the policy of the Computer Science Department to issue a failing grade to any student who either gives or receives help on any test.**
- This is an **open-book** test.
- **Any problem to which you give two or more (different) answers receives the grade of zero automatically.**
- This is a **take home exam.**
  Please submit your solution via email, as a file attachment, to `Sateesh.Mane@qc.cuny.edu`.
  The file name should have either of the formats:

  `StudentId_first_last_CS361_midterm2_Apr2018`

  `StudentId_first_last_CS761_midterm2_Apr2018`

  Acceptable file types are txt, doc/docx, pdf (also cpp, with text in comment blocks).
- **In all questions where you are asked to submit programming code, programs which display any of the following behaviors will receive an automatic F:**
  1. Programs which do not compile successfully (compiler warnings which are not fatal are excluded, e.g. use of deprecated features).
  2. Array out of bounds.
  3. Dereferencing of uninitialized variables (including null pointers).
  4. Operations which yield NAN or infinity, e.g. divide by zero, square root of negative number, etc. *Infinite loops.*
  5. Programs which do NOT implement the public interface stated in the question.
- **In addition, note the following:**
  1. Programs which compile and run successfully but have memory leaks will receive a poor grade (but not F).
  2. All debugging and/or output statements (e.g. `cout` or `printf`) will be commented out.
  3. Program performance will be tested solely on function return values and the values of output variable(s) in the function arguments.
  4. In other words, program performance will be tested solely via the public interface presented to the calling application. (I will write the calling application.)

## General information

- **You are permitted to copy and use the code in the online lecture notes.**
- **Value of $\pi$ to machine precision on any computer.**
  1. Some compilers support the constant `M_PI` for $\pi$, in which case you can write

     ```
     const double pi = M_PI;
     ```
  2. If your compiler does not support `M_PI`, the value of $\pi$ can be computed via

     ```
     const double pi = 4.0*atan2(1.0,1.0);
     ```
- **64–bit computers**
  1. The questions in this exam do not involve problems of overflow.
  2. Solutions involving the writing of algorithms will not be judged if they work on a 64–bit instead of a 32–bit computer.

- **If you submit code, put all your code in ONE cpp file.**
  **Else include all the code in your main docx or pdf or txt file.**
  **DO NOT SUBMIT MULTIPLE CPP FILES.**

# 1 Question 1 no code

- **Finite differences with unequal steps.**
- Suppose the forward and backward steps are not equal.
- Suppose the forward step is $h_1$ and the backward step is $h_2$ and $h_2 \neq h_1$.
- **Write the Taylor series for $f(x+h_1)$ and $f(x-h_2)$ up to $O(f''''(x))$.**
- **Derive a numerical expression for the first derivative as follows:**

$$\frac{f(x+h_1) - f(x-h_2)}{h_1 + h_2} = f'(x) + \text{first two error terms}\,. \tag{1.1}$$

- **Show that if $h_2 \neq h_1$ the leading error term in eq. (1.1) is $O((h_1 - h_2)f''(x))$.**
- **Using only $f(x)$, $f(x+h_1)$ and $f(x-h_2)$, derive a finite difference approximation for the first derivative $f'(x)$, where the leading error term is $O(f'''(x))$.**
  1. **Write your answer up to and including the term is $O(f''''(x))$.**
  2. **Simplify your expression in the special case $h_1 = h$ and $h_2 = 2h$.**
- **Using only $f(x)$, $f(x+h_1)$ and $f(x-h_2)$, derive a finite difference approximation for the second derivative $f''(x)$. The leading error term is $O(f'''(x))$ if $h1 \neq h_2$.**
  1. **Write your answer up to and including the term is $O(f''''(x))$.**
  2. **Simplify your expression in the special case $h_1 = h$ and $h_2 = 2h$.**
- *If you have done your work correctly, your answers should reduce to the expressions in the lectures if $h_1 = h_2 = h$.*

## 2 Question 2 show code, possible partial credit

- Instead of a unit circle described by the equation $x^2+y^2=1$, it is also possible to have curves described by the equation
$$|x|^\alpha + |y|^\beta = 1\,. \tag{2.1}$$
Here $\alpha, \beta > 0$ and $\alpha \neq \beta$ in general. The curve with $\alpha = 1.7$ and $\beta = 3.2$ is plotted in Fig. 1.

- You are required to numerically calculate the area enclosed by the curve $|x|^{1.7} + |y|^{3.2} = 1$.
  1. *Let's not panic and be stupid and think this is a two-dimensional integral.*
  2. Observe that the curve is symmetric around both the $x$ and $y$ axes.
  3. Hence we calculate the area in the first quadrant $x, y \geq 0$ and multiply the result by 4.
  4. **The Cartesian plane is divided into four quadrants.**
  5. See Fig. 1.
     (a) First quadrant: top right $x \geq 0$, $y \geq 0$.
     (b) Second quadrant: top left $x \leq 0$, $y \geq 0$.
     (c) Third quadrant: bottom left $x \leq 0$, $y \leq 0$.
     (d) Fourth quadrant: bottom right $x \geq 0$, $y \leq 0$.

- The area $A_1$ under the curve in the first quadrant is given by the following integral:
$$A_1 = \int_0^1 y(x)\,dx\,. \tag{2.2}$$

- **Write down the equation for $y(x)$ in the first quadrant.**
- **Prove that the integral in eq. (2.2) is a proper integral.**
- **Compute the value of $A_1$ using**
**(a) midpoint rule, (b) trapezoid rule, (c) Simpson's rule.**
  1. Use $n = 1000$ subintervals for all three calculations.
  2. **You are permitted to use the code displayed in the online lecture notes.**
- The total area $A$ is given by $A = 4A_1$.
- **State your computed values for the total area $A$ to four decimal places.**
- **Denote your values by $A_{\mathrm{M}}$, $A_{\mathrm{T}}$, $A_{\mathrm{S}}$.**
- *Do not worry if the three results you obtain are not equal to four decimal places.*

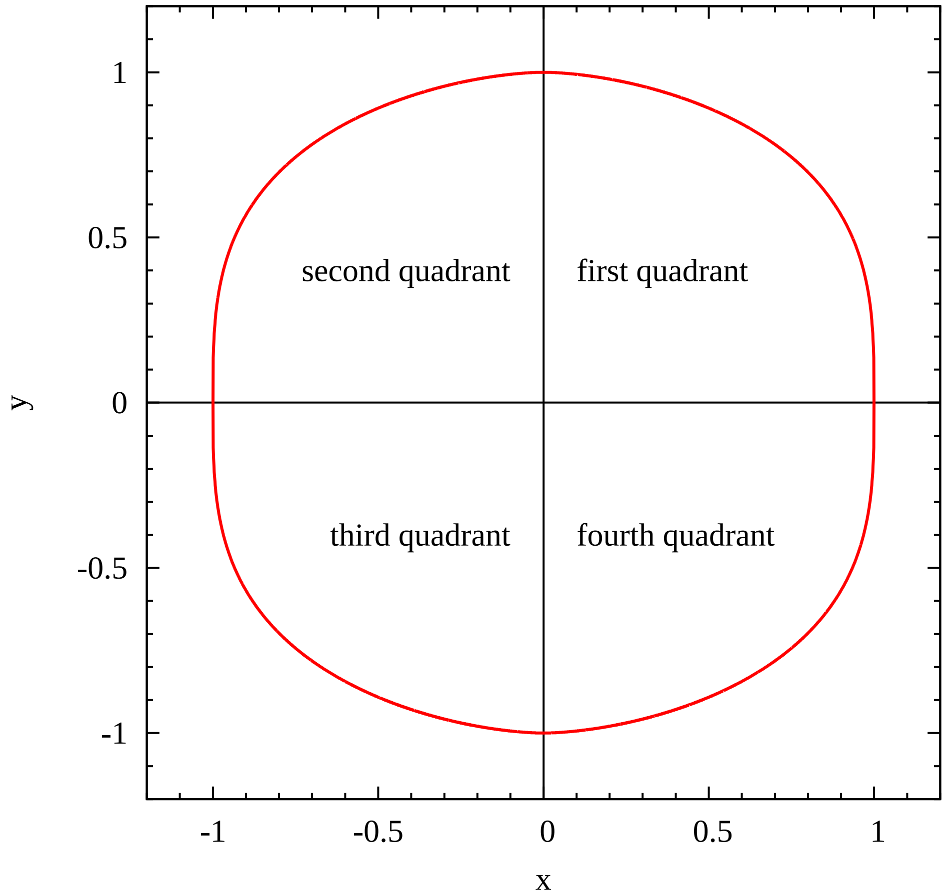


Figure 1: Graph of the curve $|x|^{1.7} + |y|^{3.2} = 1$ in Question 2.

# 3 Question 3 no code

## 3.1 Matrix 1

- Let $\alpha$, $\beta$ and $\theta$ be real numbers and $R$ be the following matrix:

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{3.1}$$

- **Calculate the trace and determinant of $R$.**
- **Calculate the inverse matrix $R^{-1}$.**
- **Solve the following equation for $x$ and $y$.**
  **Express your answer as a function of $\alpha$, $\beta$ and $\theta$.**

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{3.2}$$

- **Prove the following identity:**

$$x^2 + y^2 = \alpha^2 + \beta^2 \,. \tag{3.3}$$

- Relevant definitions and identities:

$$\begin{aligned} \cos^2\theta + \sin^2\theta = 1 \,. \\ \cos\theta &= \frac{e^{i\theta} + e^{-i\theta}}{2} \,. \\ \sin\theta &= \frac{e^{i\theta} - e^{-i\theta}}{2i} \,. \end{aligned}$$

### 3.2 Matrix 2

- Let $u$, $v$ and $\omega$ be real numbers and $B$ be the following matrix:

$$B = \begin{pmatrix} \cosh\omega & \sinh\omega \\ \sinh\omega & \cosh\omega \end{pmatrix} . \tag{3.4}$$

- **Calculate the trace and determinant of $B$.**

- **Calculate the inverse matrix $B^{-1}$.**

- **Solve the following equation for $x$ and $t$.**
  **Express your answer as a function of $u$, $v$ and $\omega$.**

$$\begin{pmatrix} \cosh\omega & \sinh\omega \\ \sinh\omega & \cosh\omega \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} . \tag{3.5}$$

- **Prove the following identity:**

$$t^2 - x^2 = u^2 - v^2 . \tag{3.6}$$

- Relevant definitions and identities:

$$\cosh^2\omega - \sinh^2\omega = 1 .$$
$$\cosh\omega = \frac{e^{\omega} + e^{-\omega}}{2} .$$
$$\sinh\omega = \frac{e^{\omega} - e^{-\omega}}{2} .$$

### 3.3 For your information (no calculations required):

- The matrix $R$ is a **rotation matrix**.
- The values of $x$ and $y$ (also $\alpha$ and $\beta$) are coordinates in a plane.
- The value of $x^2 + y^2$ is the squared length of the vector with coordinates $(x, y)$.
- The value of $\alpha^2 + \beta^2$ is the squared length of the vector with coordinates $(\alpha, \beta)$.
- A rotation does not change the length of a vector.

- The matrix $B$ is a **boost matrix**, from Einstein's Special Theory of Relativity.
- The values of $x$ and $t$ (also $u$ and $v$) are coordinates in space-time.
- The value of $t^2 - x^2$ is the squared **invariant separation** of the **space-time event** with space-time coordinates $(t, x)$.
- The value of $u^2 - v^2$ is the squared invariant separation of the space-time event with space-time coordinates $(u, v)$.
- A boost does not change the invariant separation.
- Note that the value of $t^2 - x^2$ can be **negative** even though it is called a "squared" invariant separation.
  1. If $t^2 - x^2 > 0$ it is called **timelike separation.**
  2. If $t^2 - x^2 < 0$ it is called **spacelike separation.**
  3. If $t^2 - x^2 = 0$ it is called **lightlike (or null) separation.**

## 4 Question 4 do by hand, code is optional for last step

- **Solve the following three sets of equations for $x_1$, $x_2$ and $x_3$:**

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= 3\,, \\ 2x_1 + 5x_2 + 7x_3 &= -6\,, \\ -3x_1 + 2x_2 + 4x_3 &= -1\,. \end{aligned} \tag{4.1}$$

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= 3\,, \\ 2x_1 + 5x_2 + 7x_3 &= 3\,, \\ -3x_1 + 2x_2 + 4x_3 &= -1\,. \end{aligned} \tag{4.2}$$

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= 3\,, \\ 2x_1 + 5x_2 + 7x_3 &= 12\,, \\ -3x_1 + 2x_2 + 4x_3 &= -1\,. \end{aligned} \tag{4.3}$$

- *Let's not panic. It is only necessary to perform the LU decomposition once.*
- **Write down the matrix $A$ associated with eqs. (4.1), (4.2) and (4.3).**
- **Write out the steps in the LU decomposition of $A$.**
- **Display the final matrix in LU form.**
- **Also write down the final value of the array of the swap indices $S$.**
- **Also write down the total number of swaps performed.**
- **Calculate the determinant of the matrix $A$.**
- **Solve eqs. (4.1), (4.2) and (4.3) for $x_1$, $x_2$ and $x_3$.**
  1. **To answer this part of the question, you are permitted to use the functions displayed in the online lectures, for LU decomposition and backsubstitution.**
  2. **You do not need to display all the backsubstitution steps.**
  3. **Just state the answer.**

## 5 Question 5 no code

- You are given the following equations to solve for $x_1$, $x_2$ and $x_3$:

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= 3\,, \\ 2x_1 + 5x_2 + 7x_3 &= \boldsymbol{r_2}\,, \\ \boldsymbol{a_{31}} x_1 + 2x_2 + 4x_3 &= -1\,. \end{aligned} \tag{5.1}$$

- **Find the value of $a_{31}$ such that the LU decomposition encounters a zero pivot.**
- Denote that value of $a_{31}$ by $\alpha_{31}$.
- Set $a_{31} = \alpha_{31}$. **Then find the value of $r_2$ such that the equations are consistent.**
- **Note: Do NOT attempt to solve the resulting equations.**

# 6 Question 6 no code

- You are given the following set of equations for four unknowns $x_1, x_2, x_3, x_4$:

$$\begin{aligned} 4x_1 - x_2 &= 2\,, \\ 2x_1 + 4x_2 - x_3 &= 6\,, \\ 2x_2 + 4x_3 - x_4 &= 12\,, \\ 2x_3 + 4x_4 &= 40\,. \end{aligned} \tag{6.1}$$

- **Write the set of equations in eq. (6.1) in tridiagonal matrix form as follows:**

$$\begin{pmatrix} a_1 & c_1 & 0 & 0 \\ b_2 & a_2 & c_2 & 0 \\ 0 & b_3 & a_3 & c_3 \\ 0 & 0 & b_4 & a_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 12 \\ 40 \end{pmatrix} . \tag{6.2}$$

- If you have done your work correctly, the values of $a_i$, $b_i$ and $c_i$ will not depend on $i$.
- **Prove that the tridiagonal matrix in eq. (6.2) is strongly diagonally dominant.**
- **Solve eq. (6.2) for the unknowns $x_1, x_2, x_3, x_4$.**
  1. **Display the steps in your calculation.**
  2. If you do your work correctly, there should be about eight steps, four forward and four backsubstitution steps backward.

- **(Not a question.) For your information:**
- A matrix where all the elements are equal down the diagonals is called a **Toeplitz matrix.**
- Toeplitz matrices can be stored using less storage (only one number per diagonal).
- Technically, a Toeplitz matrix need not be square.
- However, the definitions are simpler for square matrices.
- There are efficient numerical algorithms to process Toeplitz matrices.

## 7 Question 7 no code

- Let $\mu$ be a real number and let $T$ be the following tridiagonal matrix:

$$T = \begin{pmatrix} 2+\mu^2 & \mu-1 & 0 & 0 & 0 \\ \mu-1 & 2+\mu^2 & \mu-1 & 0 & 0 \\ 0 & \mu-1 & 2+\mu^2 & \mu-1 & 0 \\ 0 & 0 & \mu-1 & 2+\mu^2 & \mu-1 \\ 0 & 0 & 0 & \mu-1 & 2+\mu^2 \end{pmatrix} \tag{7.1}$$

- **You will need to consider the cases $\mu \geq 1$ and $\mu < 1$ separately.**
  1. If $\mu \geq 1$ then $|\mu - 1| = \mu - 1$.
  2. If $\mu < 1$ then $|\mu - 1| = -(\mu - 1) = 1 - \mu$.
  3. **Remember to pay attention to the special cases in the first and last rows.**
- **Prove that the matrix $T$ in eq. (7.1) is strongly diagonally dominant for all $\mu \geq 1$.**
- **Find all values of $\mu$ for which the matrix $T$ in eq. (7.1) is:**
  1. Strongly diagonally dominant.
  2. Weakly **but not strongly** diagonally dominant.
  3. **Not diagonally dominant.**
- For your information, the matrix $T$ in eq. (7.1) is a symmetric tridiagonal Toeplitz matrix.
- **Solve the following matrix equation for the unknowns $x_1, x_2, x_3, x_4, x_5$, for $\mu = 1$:**

$$T_{(\mu=1)} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} . \tag{7.2}$$

- *Don't laugh.*

## 8 Question 8 submit code, possible partial credit

- See eq. (4.1). Let $(\gamma_1, \gamma_2, \gamma_3)$ denote the solutions for $(x_1, x_2, x_3)$ of the equations below:

$$\begin{aligned} -x_1 + 2x_2 + x_3 &= 3\,, \\ 2x_1 + 5x_2 + 7x_3 &= -6\,, \\ -3x_1 + 2x_2 + 4x_3 &= -1\,. \end{aligned} \tag{8.1}$$

- **For each case $i = 1, 2, 3$, determine if the integral in eq. (8.2) is proper.**

$$I(\gamma_i) = \int_{-1}^{1} x^{2\gamma_i} \cos(x^2)\, dx\,. \tag{8.2}$$

- If the integral in eq. (8.2) is proper, **compute its value numerically using the extended trapezoid rule with $n = 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024$ subintervals.**
- Denote the results by $R(j,0)$, where $n = 2^j$, so $j = 0, 1, \ldots, 10$.
- Then use first order Romberg integration. Define $R(j,1)$ as follows:

$$R(j,1) = \frac{4R(j,0) - R(j-1,0)}{3}\,, \qquad j = 1, 2, \ldots, 10\,. \tag{8.3}$$

- Then use second order Romberg integration. Define $R(j,2)$ as follows:

$$R(j,2) = \frac{16R(j,1) - R(j-1,1)}{15}\,, \qquad j = 2, 3, \ldots, 10\,. \tag{8.4}$$

- **Fill the table below with values to six decimal places.**

| $n$ | $j$ | $R(j,0)$ | $R(j,1)$ | $R(j,2)$ |
|---|---|---|---|---|
| 1 | 0 | 6 d.p. | | |
| 2 | 1 | 6 d.p. | 6 d.p. | |
| 4 | 2 | 6 d.p. | 6 d.p. | 6 d.p. |
| 8 | 3 | 6 d.p. | 6 d.p. | 6 d.p. |
| 16 | 4 | 6 d.p. | 6 d.p. | 6 d.p. |
| 32 | 5 | 6 d.p. | 6 d.p. | 6 d.p. |
| 64 | 6 | 6 d.p. | 6 d.p. | 6 d.p. |
| 128 | 7 | 6 d.p. | 6 d.p. | 6 d.p. |
| 256 | 8 | 6 d.p. | 6 d.p. | 6 d.p. |
| 512 | 9 | 6 d.p. | 6 d.p. | 6 d.p. |
| 1024 | 10 | 6 d.p. | 6 d.p. | 6 d.p. |

- **State the smallest value of $j$ for which $R(j,2)$ converges to six decimal places.**
- **State the value of the integral $I(\gamma_i)$ to six decimal places.**

# 9 Question 9 submit code

## 9.1 Complete elliptic integral of the second kind

- There are three kinds of complete elliptic integrals (and three incomplete elliptic integrals).
- The **complete elliptic integral of the second kind** is given by the following integral:

$$E(x) = \int_0^1 \frac{\sqrt{1-x^2t^2}}{\sqrt{1-t^2}}\,dt \qquad (0 \le x \le 1)\,. \tag{9.1}$$

- For your information, for an ellipse with semi–major axis $a$ and semi–minor axis $b$ and eccentricity $e_{\rm ell} = \sqrt{1-b^2/a^2}$, the circumference of the ellipse, say $c$, is given by $c = 4aE(e_{\rm ell})$.
- The function $E(x)$ can be expressed as power series as follows:

$$E(x) = \frac{\pi}{2}\left[1 - \left(\frac{1}{2}\right)^2\frac{x^2}{1} - \left(\frac{1\cdot 3}{2\cdot 4}\right)^2\frac{x^4}{3} - \left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^2\frac{x^6}{5} - \cdots\right]. \tag{9.2}$$

- We shall compute the value of $E(x)$ in two ways, by using the integral in eq. (9.1) and by approximating the infinite series in eq. (9.2) by a finite sum.
- The integral in eq. (9.1) is improper because the integrand diverges at $t = 1$ (unless $x = 1$, in which case we obtain $0/0$, which a computer also cannot evaluate).
  1. However, we can compute the integral in eq. (9.1) using the midpoint rule.
  2. **Let $I_n(x)$ denote the value of the integral in eq. (9.1) using the midpoint rule with $n$ subintervals.**
- Let us approximate the infinite series in eq. (9.2) by a finite sum.
  1. **Let $S_m(x)$ denote the value of the series in eq. (9.2), when the series is terminates at the term in $x^{2m}$.**

$$S_m(x) = \frac{\pi}{2}\left[1-\left(\frac{1}{2}\right)^2\frac{x^2}{1}-\left(\frac{1\cdot 3}{2\cdot 4}\right)^2\frac{x^4}{3}-\cdots-\left(\frac{1\cdot 3\cdot 5\ldots(2m-1)}{2\cdot 4\cdot 6\ldots(2m)}\right)^2\frac{x^{2m}}{2m-1}\right]. \tag{9.3}$$

  2. **Formulate an efficient algorithm to compute the sum for $S_m(x)$ in eq. (9.3).**
- **(Optional/bonus) Prove using the sum in eq. (9.2) that $E(x)$ is a decreasing function of $x$, i.e. if $x_2 > x_1$ then $E(x_2) < E(x_1)$.**
- **(Not a question.) For your information:**
  1. Observe from the series in eq. (9.2) that $E(0) = \frac{1}{2}\pi \simeq 1.57$ for $x = 0$.
  2. Observe from the integral in eq. (9.1) that $E(1) = 1$ for $x = 1$.
  3. Hence $E(x)$ decreases monotonically from $\frac{1}{2}\pi$ to 1 as $x$ increases from 0 to 1.

## 9.2 Computation

- **Set $x = 0$. We know $E(0) = \frac{1}{2}\pi$ is the exact value.**
- Obviously the sum in eq. (9.2) will yield $S_m(0) = \frac{1}{2}\pi$ for all values of $m$.
- **Compute the value of $I_n(0)$ and fill in the table below.**

| $n$ | $I_n(1)$ | $\lvert I_n(1) - \frac{1}{2}\pi\rvert$ |
|---|---|---|
| $10^1$ | 6 d.p. | |
| $10^2$ | 6 d.p. | |
| $10^3$ | 6 d.p. | |
| $10^4$ | 6 d.p. | |
| $10^5$ | 6 d.p. | |
| $10^6$ | 6 d.p. | |

- **Set $x = 1$. We know $E(1) = 1$ is the exact value.**
- **Compute the values of $I_n(1)$ and $S_m(1)$ and fill in the table below.**

| $n$ | $I_n(1)$ | $\lvert I_n(1) - 1\rvert$ | $m$ | $S_m(1)$ | $\lvert S_m(1) - 1\rvert$ |
|---|---|---|---|---|---|
| $10^1$ | 6 d.p. | | $10^1$ | 6 d.p. | |
| $10^2$ | 6 d.p. | | $10^2$ | 6 d.p. | |
| $10^3$ | 6 d.p. | | $10^3$ | 6 d.p. | |
| $10^4$ | 6 d.p. | | $10^4$ | 6 d.p. | |
| $10^5$ | 6 d.p. | | $10^5$ | 6 d.p. | |
| $10^6$ | 6 d.p. | | $10^6$ | 6 d.p. | |

- Hence observe that the sum $S_m(x)$ converges more rapidly than $I_n(x)$ for small values $x \simeq 0$ and the integral $I_n(x)$ converges more rapidly than $S_m(x)$ for large values $x \simeq 1$.
- **Fill the table below for $I_n(x)$ and $S_m(x)$. Use $m = n = 1000$.
Write your answers to 4 decimal places.**

| $x$ | $I_n(x)$ | $S_m(x)$ |
|---|---|---|
| 0 | 4 d.p. | 4 d.p. |
| 0.1 | 4 d.p. | 4 d.p. |
| 0.2 | 4 d.p. | 4 d.p. |
| 0.3 | 4 d.p. | 4 d.p. |
| 0.4 | 4 d.p. | 4 d.p. |
| 0.5 | 4 d.p. | 4 d.p. |
| 0.6 | 4 d.p. | 4 d.p. |
| 0.7 | 4 d.p. | 4 d.p. |
| 0.8 | 4 d.p. | 4 d.p. |
| 0.9 | 4 d.p. | 4 d.p. |
| 1.0 | 4 d.p. | 4 d.p. |

- **(Optional) Plot a graph of $S_m(x)$ for $x = 0, 0.1, \ldots, 1.0$.**

## 9.3 Taylor series: remainder term

- **Let us estimate the remainder term of the Taylor series.**
- If we sum eq. (9.2) to $m$ terms, i.e. use $S_m(x)$, the remainder term $R_m(x)$ is

$$R_m(x) = -\left(\frac{1\cdot 3\cdots(2m+1)}{2\cdot 4\cdots(2m+2)}\right)^2 \frac{x^{2m+2}}{2m+1} - \left(\frac{1\cdot 3\cdots(2m+1)(2m+3)}{2\cdot 4\cdots(2m+2)(2m+4)}\right)^2 \frac{x^{2m+4}}{2m+3} - \cdots \tag{9.4}$$

- All the terms are negative (i.e. same sign) and the coefficients decrease in magnitude.
- Hence an upper bound on the magnitude of the remainder term is $|R_m(x)| \le U_m(x)$, where

$$\begin{aligned} U_m(x) &= \left(\frac{1\cdot 3\cdots(2m+1)}{2\cdot 4\cdots(2m+2)}\right)^2 \frac{x^{2m+2}}{2m+1}\left(1+x^2+x^4+\cdots\right) \\ &= \frac{1}{2m+1}\left(\frac{1\cdot 3\cdots(2m+1)}{2\cdot 4\cdots(2m+2)}\right)^2 \frac{x^{2m+2}}{1-x^2}\,. \end{aligned} \tag{9.5}$$

- Observe that $U_m(x) = 0$ for $x = 0$ and $U_m(x) \to \infty$ for $x \to 1$.
- We saw that the sum $S_m(x)$ is reliable for small $x \simeq 0$ and not so accurate for large $x \simeq 1$.
- **Set $x = 0.5$ and calculate the value of $U_m(0.5)$ and fill the table below.**
  **State your results in the 'scientific notation' form $a.bc \times 10^{-d}$.**

| $m$ | $U_m(0.5)$ |
|---|---|
| 2 | |
| 3 | |
| 4 | |
| 5 | |
| 6 | |

- *If you have done your work correctly, then to obtain accuracy for $E(0.5)$ to 4 decimal places it should be sufficient to use $m = 5$ or 6.*

### 9.4 Root finding

- **Find $x$ such that $E(x) = 1.48$.**
- **Use bisection and use $I_n(x)$ with $n = 1000$.**

  1. **Use numbers from the previous table.**
  2. **State the shortest initial bracket which encloses the root.**
  3. Denote the initial iterates by $x_0$ for $x_{\text{low}}$ and $x_1$ for $x_{\text{high}}$
  4. Iterate until the value of the root converges to 4 decimal places.

| $i$ | $x_i$ | $I_n(x) - 1.48$ |
|---|---|---|
| 0 | $x_0$ | 5 d.p. |
| 1 | $x_1$ | 5 d.p. |
| $\vdots$ | $\vdots$ | |
| | converged 4 d.p. | |

- **Use the secant method and use $S_m(x)$ with $m = 5$.**

  1. **For the two initial iterates, use the same two values for $x_0$ and $x_1$ which were employed above for bisection.**
  2. Iterate until the value of the root converges to 4 decimal places.

| $i$ | $x_i$ | $I_n(x) - 1.48$ |
|---|---|---|
| 0 | $x_0$ | 5 d.p. |
| 1 | $x_1$ | 5 d.p. |
| $\vdots$ | $\vdots$ | |
| | converged 4 d.p. | |

- *If you have done your work correctly, the values for the root, using $I_n(x)$ and $S_m(x)$, will* **not** *be equal.*
- This illustrates the difficulty when trying to compute roots (or function values in general) using numerical algorithms.
- The answers we get depend on external parameters such as $m$ and $n$.
- **Explain why the answer computed using $S_5(x)$ is reliable to 4 decimal places.**
- For $I_n(x)$, we need to go up to about $n \simeq 5 \times 10^6$ (five million) for the value of the root to agree with the calculation using $S_5(x)$.

**See next page.**

- From eq. (9.1), for $x = 0.5$, the value of the integral is

$$E(0.5) = \int_0^1 \frac{\sqrt{1-\frac{1}{4}t^2}}{\sqrt{1-t^2}}\,dt\,. \tag{9.6}$$

- The integrand diverges to $\infty$ for $t \to 1$.

- **Calculate the value of $\sqrt{(1-(t^2/4))/(1-t^2)}$ and fill in the table below.**

| $n$ | $h = 1/n$ | $t = 1-h$ | $\sqrt{(1-(t^2/4))/(1-t^2)}$ | $h \times \sqrt{(1-(t^2/4))/(1-t^2)}$ |
|---|---|---|---|---|
| $10^3$ | $10^{-3}$ | 0.999 | 4 d.p. | 4 d.p. |
| $10^4$ | $10^{-4}$ | 0.9999 | 4 d.p. | 4 d.p. |
| $10^5$ | $10^{-5}$ | 0.99999 | 4 d.p. | 4 d.p. |
| $10^6$ | $10^{-6}$ | 0.999999 | 4 d.p. | 4 d.p. |

- *If you have done your work correctly, the final number in the last column should be about* 0.0006.

- Because of the divergence of the integrand as $t \to 1$, it requires $n > 10^6$ subintervals to compute the integral in eq. (9.1) to 4 decimal places, for $x = 0.5$.

- There are of course other ways of computing the integral.

- We can transform the integral to obtain

$$E(x) = \int_0^{\pi/2} \sqrt{1 - x^2 \sin^2\theta}\,d\theta\,. \tag{9.7}$$

- In this form the integrand is bounded for all values of $\theta$. It is a proper integral for $|x| \le 1$.

## 10 Question 10 no code

- **This is a question about a question in Midterm 1.**
- **You <u>do not</u> have to compute numbers or submit code for this question.**
- The Bessel function $J_0(x)$ can be computed by evaluating the following integral:

$$J_0(x) = \frac{1}{\pi}\int_0^{\pi} \cos(x\sin\theta)\,d\theta\,. \tag{10.1}$$

- **Write an expression (sum of terms) to compute the integral in eq. (10.1) using the trapezoid rule with $n$ subintervals.**
- In midterm 1, I approximated the integral in eq. (10.1) using the following sum (here I use $n$ instead of $N$ as I did in Midterm 1):

$$S(x) = \frac{1}{n}\sum_{j=0}^{n-1} \cos\Bigl(x\,\sin\frac{j\pi}{n}\Bigr)\,. \tag{10.2}$$

- **State the difference between the trapezoid rule formula and my sum in eq. (10.2).**
- **Explain why the difference does not matter and the sum in eq. (10.2) yields the same result as the trapezoid rule.**

## 11 Question 11 (bonus question) no code

- Let $I_n$ be the unit matrix of size $n \times n$.
- **Calculate the trace of $I_n$.**
- Let $M$ be an arbitrary square matrix of size $n \times n$.
- Define a matrix $D$ as follows:

$$D = \frac{\mathrm{trace}(M)}{n} I_n \,. \tag{11.1}$$

- **Prove that $D$ is diagonal. Also prove the following:**

$$\mathrm{trace}(D) = \mathrm{trace}(M) \,. \tag{11.2}$$

- Define a matrix $A$ as follows:

$$A = \tfrac{1}{2}(M - M^T) \,. \tag{11.3}$$

- **Prove that $A$ is antisymmetric.**
- Define a matrix $T$ as follows:

$$T = \tfrac{1}{2}(M + M^T) - \frac{\mathrm{trace}(M)}{n} I_n \,. \tag{11.4}$$

- **Prove that $T$ is symmetric and traceless.**
- **Prove the following:**

$$M = D + A + T \,. \tag{11.5}$$

- For your information, the matrices $D$, $A$ and $T$ are the **irreducible components** of the matrix $M$.
- Every square matrix can be decomposed this way.