1. (10 points) Suppose a system function for \(e^x\) is *not* available, but you know the Taylor expansion
\[
e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots = \sum_{k=0}^{\infty} \frac{x^k}{k!}.
\]
A possibility is to approximate \(e^x\) with a partial sum \(\sum_{k=0}^{n} \frac{x^k}{k!}\) for an appropriately large value of \(n\). Discuss approximating \(e^{-x}\) with such a partial sum in floating-point arithmetic when \(x\) is large and positive. (You may assume that \(n\) is taken so large that the partial sum would be sufficiently accurate in exact arithmetic.) What are the concerns with evaluating \(\sum_{k=0}^{n} \frac{(-x)^k}{k!}\) in floating point arithmetic? Is there a better way to approximate \(e^{-x}\) using a partial sum that can be more accurately evaluated?

Remark: Approximating \(e^x\) with \(\sum_{k=0}^{n} \frac{x^k}{k!}\) is fairly straightforward when \(x \geq 0\), even in floating-point arithmetic. The only consideration is that, for greatest accuracy, one should accumulate the sum from the smallest values of the summand to the largest, i.e., from \(k = n\) down to \(k = 0\).

2. (10 points) A matrix \(A \in \mathbb{R}^{n \times n}\) with entries \(a_{ij}\) is said to be *banded* if there are non-negative integers \(m_\ell\) and \(m_u\) (the *lower* and *upper bandwidths*, resp.) such that \(a_{ij} \neq 0\) only if \(i - m_\ell \leq j \leq i + m_u\). For example, \(A\) is *tridiagonal* if it is banded and \(m_\ell = m_u = 1\). For banded matrices, Gaussian elimination can be implemented with economies in both arithmetic and storage. As usual, pivoting is necessary in general but isn’t needed in some special cases, e.g., if \(A\) is symmetric positive-definite. To keep this problem simple, we consider only the algorithm without pivoting:

**Banded Gaussian Elimination (without pivoting):**

For \(k = 1, \ldots, n - 1\)

For \(i = k + 1, \ldots, \min\{k + m_\ell, n\}\)

\(a_{ik} \leftarrow -a_{ik}/a_{kk}\)

For \(j = k + 1, \ldots, \min\{k + m_u, n\}\)

\(a_{ij} \leftarrow a_{ij} + a_{ik}a_{kj}\)

About how many multiplications does this algorithm require when \(1 \leq m_\ell \ll n\) and \(1 \leq m_u \ll n\)? You may simplify the multiplication count by replacing \(\min\{k + m_\ell, n\}\) by \(k + m_\ell\) and \(\min\{k + m_u, n\}\) by \(k + m_u\) wherever they appear.
3. (30 points) In class and in the text (see Th. 8.5, p. 536), we had bounds on the relative error in the solution of $Ax = b$, $A \in \mathbb{R}^{n\times n}$, computed using Gaussian elimination with pivoting. These bounds depend on a “growth factor” $\rho$, which is a bound on the largest matrix entry (in absolute value) encountered during the elimination. When complete pivoting is used, $\rho$ can be shown to grow slowly with $n$; thus Gaussian elimination with complete pivoting is always stable. In contrast, when partial pivoting is used, the best provable bound is $\rho \leq 2^{n-1}$, and so Gaussian elimination with partial pivoting is potentially unstable, despite being virtually always stable in practice. One cannot obtain a better theoretical bound on $\rho$ than $\rho \leq 2^{n-1}$; indeed, there are examples for which $\rho = 2^{n-1}$, as seen in the following.

Suppose $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$ are given by

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ -1 & 1 & \ddots & \vdots \\ \vdots & -1 & \ddots & 0 & 1 \\ \vdots & \vdots & \ddots & 1 & 1 \\ -1 & -1 & \cdots & -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 2 \\ 1 \end{pmatrix}.$$

Applying Gaussian elimination with partial pivoting to $Ax = b$ results in $Ux = c$, where

$$U = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ 0 & 1 & \ddots & \vdots & 2 \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & \ddots & 1 & 2^{n-2} \\ 0 & 0 & \cdots & 0 & 2^{n-1} \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2^{n-3} \\ 2^{n-2} + 1 \\ 2^{n-1} + 1 \end{pmatrix}.$$

(It turns out there is no pivoting in this example, i.e., pivoting isn’t invoked during the elimination, but this isn’t important for the problem.)

a. What solution will be computed in finite precision arithmetic if $2^{n-2} > 1/\epsilon$, where $\epsilon$ is machine epsilon (unit roundoff error)? (Reason this out; don’t use the computer.)

b. With $n = 56$, first compute the solution in MATLAB using the “backslash” command, i.e., “\texttt{soln1 = A\backslash b}”. Then compute it again using the “\texttt{qr}” command as follows: “\texttt{[Q,R] = qr(A); soln2 = R\backslash (Q'*b);}”. (Don’t hand in printouts of these solutions.) Compute the errors in these solutions by forming their differences with the exact solution $(-1/2^{n-1}, -1/2^{n-2}, \ldots, -1/4, 1/2, 1 + 1/2^{n-1})^T$. Use the “\texttt{norm}” command to compute the 2-norms of these errors; print out and hand in the 2-norms you obtain.
4. (25 points) (Adapted from J. B. Keller, *Probability of a shutout in racquetball*, SIAM Review, 26 (1984), pp. 267–268.) The probability $P$ that Player A will shut out (i.e., win by a score of 21-0) Player B in a game of racquetball is given by

$$P = \frac{1 + p}{2} \left( \frac{p}{1 - p + p^2} \right)^{21},$$

where $p$ is the probability that Player A will win any one rally (independent of the server). Use the MATLAB “*fzero*” command to determine the value of $p$ that will result in Player A expecting to shut out Player B in half the games they play.

5. (25 points) (From problem 7, p. 118 of Atkinson.) Suppose that an amount of $P_1$ dollars is put into an account at the beginning of years 1, 2, $\ldots$, $N_1$ and that the account accumulates interest at a fractional rate $r$. (The percentage rate is $r \times 100$; for example, $r = .05$ corresponds to 5% interest.) Suppose also that, at the beginning of years $N_1 + 1$, $N_1 + 2$, $\ldots$, $N_1 + N_2$, an amount of $P_2$ dollars is withdrawn from the account and that the account balance is exactly zero after the withdrawal at year $N_1 + N_2$. Then the variables satisfy the following equation:

$$P_1 \left[ (1 + r)^{N_1} - 1 \right] = P_2 \left[ 1 - (1 + r)^{-N_2} \right].$$

If $N_1 = 30$, $N_2 = 20$, $P_1 = 2000$, and $P_2 = 8000$, then what is $r$? You may use any method or software you like to find the answer.