# Assignment 2

## General Instructions

- Solutions due by $17^{th}$ Feb 2026
- Hand calculations should be submitted together with the report.
- Allowed to work in teams of two for the programming part, yet write your own report
- Write your name and roll-number on your report

## Hand calculations

1. (15 points) In this problem, our objective is to figure out the rank of a given matrix $A$ using Gaussian Elimination and thereby comment on the solvability of $Ax = b$. The upper triangular matrix obtained upon Gaussian Elimination is denoted as $U$ below and the corresponding RHS vector is denoted as $y$. *There is no need to consider pivoting while answering this question.*

   Note the following:

   - Rank of a matrix is the number of linearly independent rows or columns in the matrix.
   - Upon performing Gaussian Elimination, (i) the number of rows which are **not entirely zeroes** in the resulting upper triangular matrix, is equal to the rank of the matrix A; (ii) the number of rows which are **not entirely zeroes** in the resulting augmented $(U|y)$ matrix, is equal to the rank of the augmented matrix $(A|b)$.
   - When the rank of A = rank of $(A|b)$ = # cols of A, then the linear system has a unique solution
   - When the rank of A = rank of $(A|b)$ < # cols of A, then the linear system has infinite solutions
   - When the rank of A $\neq$ rank of $(A|b)$, then the linear system is inconsistent and has no solution

   Solve the following systems using Gaussian Elimination, and classify if they have a unique, infinite, or no solution. Obtain the ranks of matrices $A$ and $(A|b)$ and justify your answer on solvability: unique/infinite/no-solution using the criteria based on ranks of the matrices $A$ and $(A|b)$ provided above.

   (a) $x + 2y + z = 1;\ \ 2x + 2y = 1;\ \ x + 3y + z = 1$

   (b) $x + 2y - z = 1;\ \ 2x + 2y = 1;\ \ x + 3y - 2z = 1$

   (c) $x + 2y - z = 3;\ \ 2x + 2y = 4;\ \ x + 3y - 2z = 4$

   (d) $x + 2y = 3;\ \ 2x + 3y = 5;\ \ 3x + 7y = 10$

   (e) $x + 2y = 3;\ \ 2x + 3y = 5;\ \ 3x + 7y = 11$

2. (15 points) We will examine the round-off errors in numerically solving linear systems using Gaussian elimination with/without pivoting techniques. The relative error between the exact and the numerical solution (using floating point arithmetic) is related to the *Growth factor* ($\rho$), defined as

   $$\rho = \frac{\max_{i,j} |u_{ij}|}{\max_{i,j} |a_{ij}|}.$$

   The smaller the growth factor, lesser will the number of digits you lose in the solution. Similarly, if the growth factor is larger, you will lose more digits in the solution, leading to higher error.

   To be precise, for a $m \times m$ matrix $A$, if $\rho \sim \mathcal{O}(2^m)$, then $\mathcal{O}(m)$ bits of precision are lost in the solution, which can be catastrophic for practical applications ($m$ is in hundreds or thousands). Use of pivoting techniques in practical cases, restricts the magnitude of the growth factor, and thus limits the loss of accuracy (and thereby lends stability to the solution procedure).
   Below is an example of this scenario:

(a) Consider the matrix:

$$A = \begin{pmatrix} \epsilon & 1 \\ 1 & 1 \end{pmatrix} \tag{1}$$

where $\epsilon$ is a very small number. Determine the growth factor using Gaussian Elimination (i) without pivoting and (ii) with partial pivoting; Compute the growth factor as $\epsilon \to 0$.

(b) As noted earlier, while pivoting restricts the magnitude of the growth factor, in a worst case scenario, the stability offered may not be enough and significant loss of accuracy in solution can still result. Below is an example: Consider the matrix:

$$B = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \tag{2}$$

Find the growth factor using Gaussian Elimination with partial pivoting. What is your estimate of growth factor for a similar matrix with size $m \times m$? How many bits of precision are lost if $m = 100$?

(c) *Partial vs Complete pivoting:*
In partial pivoting technique, while eliminating the $k^{th}$ column, row swaps (with row index $j \geq k$) are performed to get the largest element in magnitude within the $k^{th}$ column as the pivot. Whereas, in *complete pivoting* technique, while eliminating the $k^{th}$ column, rows with index $j \geq k$ and columns with index $j \geq k$ are swapped to get the largest element in magnitude as the pivot.

Thus, at every stage of Gaussian elimination, $(m-k)$ entries are compared to identify the pivot in the case of partial pivoting, as against $(m-k)^2$ entries in the case of complete pivoting. This explains the higher computational cost to the solution procedure when complete pivoting is employed. Nonetheless, the advantage to complete pivoting is that, it typically lends more stability to Gaussian elimination than partial pivoting.

To illustrate this, for the above matrix $B$, find the growth factor using Gaussian elimination with complete pivoting. What is your estimate of growth factor for a similar matrix with size $m \times m$? How many bits of precision are lost if $m = 100$?

3. (10 points) We learnt two ways to get $A = LU$ in class:

(a) From Gaussian Elimination and the multiplication factors

(b) Using Doolittle's algorithm.

Consider the matrix:

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{pmatrix} \tag{3}$$

Get $L$ and $U$ matrices both ways keeping $l_{ii} = 1$ for $i = \{1, 2, 3, 4\}$. Are they the same?

# Programming

1. (30 points) Write a program that performs Gaussian elimination for a square system of size $n \times n$. Consider writing it as a modular program with separate functions or subroutines that perform forward-elimination and backward-substitution. To test the program, solve the system $Ax = b$ of order $n$, with $A = [a_{ij}]$ defined by

$$a_{ij} = \max(i, j).$$

Also define $b = [1, 1, ..., 1]^T$. Solve the system to obtain the solution vector $x$, for $n = 32, 128, 512, 1024$. Plot a graph between $n$ and $1/\sum_{i=1}^{n} {x_i}^2$. (30 points)

2. (30 points) A large industrial furnace, as shown in the figure below, is supported on a long column of fireclay brick. During steady-state operation, installation is such that three surfaces of the column are maintained at 500 K while the remaining surface is exposed to an airstream at a temperature of 300 K. Owing to the symmetry of the problem the number of unknowns (temperature at the interior nodal locations) is reduced from 12 to 8. Using energy balance the following equations are obtained in terms of the nodal temperatures:

$$\begin{aligned}
&\text{Node 1:} && T_2 + T_3 + 1000 - 4T_1 = 0\\
&\text{Node 3:} && T_1 + T_4 + T_5 + 500 - 4T_3 = 0\\
&\text{Node 5:} && T_3 + T_6 + T_7 + 500 - 4T_5 = 0\\
&\text{Node 2:} && 2T_1 + T_4 + 500 - 4T_2 = 0\\
&\text{Node 4:} && T_2 + 2T_3 + T_6 - 4T_4 = 0\\
&\text{Node 6:} && T_4 + 2T_5 + T_8 - 4T_6 = 0\\
&\text{Node 7:} && 2T_5 + T_8 + 2000 - 9T_7 = 0\\
&\text{Node 8:} && 2T_6 + 2T_7 + 1500 - 9T_8 = 0.
\end{aligned}$$

Using the Gaussian elimination algorithm developed above, calculate and report the temperatures $T_1$ through $T_8$. Also, plot the variation of temperature as a function of $y$ at locations $x = 0.25$ and $x = 0.5$.