Error Analysis: Propagated Error in Addition and Subtraction

Let

$$x_T = x_A + \epsilon \text{ and } \quad y_T = y_A + \eta$$

are positive real numbers.

The relative error $E_r(x_A \pm y_A)$ is given by

$$\begin{aligned}
E_r(x_A \pm y_A) &= \frac{(x_T \pm y_T) - (x_A \pm y_A)}{x_T \pm y_T} \\
&= \frac{(x_T \pm y_T) - (x_T - \epsilon \pm (y_T - \eta))}{x_T \pm y_T} \\
&= \frac{\epsilon \pm \eta}{x_T \pm y_T}.
\end{aligned}$$

This shows that relative error propagates slowly with addition, whereas amplifies drastically with subtraction when $x_T \approx y_T$.

Error Analysis: Propagated Error in Multiplication

The relative error $E_r(x_A \times y_A)$ is given by

$$
\begin{aligned}
E_r(x_A \times y_A) &= \frac{(x_T \times y_T) - (x_A \times y_A)}{x_T \times y_T} \\
&= \frac{(x_T \times y_T) - ((x_T - \epsilon) \times (y_T - \eta))}{x_T \times y_T} \\
&= \frac{\eta x_T + \epsilon y_T - \epsilon\eta}{x_T \times y_T} \\
&= \frac{\epsilon}{x_T} + \frac{\eta}{y_T} - \left(\frac{\epsilon}{x_T}\right)\left(\frac{\eta}{y_T}\right)
\end{aligned}
$$

$$\Rightarrow \quad E_r(x_A \times y_A) = E_r(x_A) + E_r(y_A) - E_r(x_A)E_r(y_A)$$

This shows that relative error propagates slowly with multiplication

Error Analysis: Propagated Error in Division

The relative error $E_r(x_A/y_A)$ is given by

$$
\begin{aligned}
E_r(x_A/y_A) &= \frac{(x_T/y_T) - (x_A/y_A)}{x_T/y_T} \\
&= \frac{(x_T/y_T) - ((x_T - \epsilon)/(y_T - \eta))}{x_T/y_T} \\
&= \frac{x_T(y_T - \eta) - y_T(x_T - \epsilon)}{x_T(y_T - \eta)} \\
&= \frac{\epsilon y_T - \eta x_T}{x_T(y_T - \eta)} \\
&= \frac{y_T}{y_T - \eta}(E_r(x_A) - E_r(y_A))
\end{aligned}
$$

Error Analysis: Propagated Error in Division

The relative error $E_r(x_A/y_A)$ is given by

$$
\begin{aligned}
E_r(x_A/y_A) &= \frac{(x_T/y_T) - (x_A/y_A)}{x_T/y_T} \\
&= \frac{(x_T/y_T) - ((x_T - \epsilon)/(y_T - \eta))}{x_T/y_T} \\
&= \frac{x_T(y_T - \eta) - y_T(x_T - \epsilon)}{x_T(y_T - \eta)} \\
&= \frac{\epsilon y_T - \eta x_T}{x_T(y_T - \eta)} \\
&= \frac{y_T}{y_T - \eta}\left(E_r(x_A) - E_r(y_A)\right)
\end{aligned}
$$

$$
\Rightarrow \quad E_r(x_A/y_A) = \frac{1}{1 - E_r(y_A)}\left(E_r(x_A) - E_r(y_A)\right).
$$

$\Longrightarrow$ relative error propagates slowly with division, unless $E_r(y_A) \approx 1$

General Form of Linear System (contd.)

These equations can be written in the matrix notation as

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

The last equation is usually written in the following short form

$$Ax = b,$$

where

- A stands for the $n \times n$ matrix with entries a_{ij}.
- $x = (x_1, x_2, \cdots, x_n)^T$

General Form of Linear System (contd.)

Let us now state a result concerning the solvability of the system

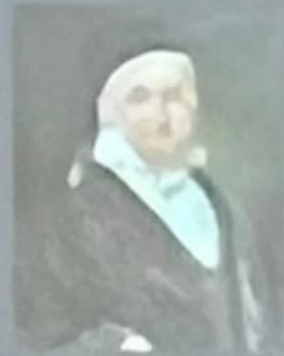
$$Ax = b.$$

Theorem

Let A be an $n \times n$ matrix and $\mathbf{b} \in \mathbb{R}^n$. Then the following statements concerning the system of linear equations $A\mathbf{x} = \mathbf{b}$ are equivalent.

1. $\det(A) \neq 0$
2. For each right hand side vector $\mathbf{b}$, the system $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}$.
3. For $\mathbf{b} = 0$, the only solution of the system $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = 0$.

Linear Systems: Naive Gaussian Elimination Method



Carl Friedrich Gauss (1777–1855) *German mathematician*

Consider the following system:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \quad \Longleftrightarrow \quad \text{Upper triangular System} \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3.
\end{aligned}
$$

Linear Systems: Naive Gaussian Elimination Method (contd.)

Step 1: If $a_{11} \neq 0$, then define

$$m_{21} = \frac{a_{21}}{a_{11}}, \quad m_{31} = \frac{a_{31}}{a_{11}}.$$

We will now obtain a new system that is equivalent to the given system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

- The first equation will be retained as it is.
- The second equation will be replaced by the equation $E_2 - m_{21}E_1$

Linear Systems: Naive Gaussian Elimination Method (contd.)

Step 1: If $a_{11} \neq 0$, then define

$$m_{21} = \frac{a_{21}}{a_{11}}, \quad m_{31} = \frac{a_{31}}{a_{11}}.$$

We will now obtain a new system that is equivalent to the given system

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3
\end{aligned}$$

- The first equation will be retained as it is.
- The second equation will be replaced by the equation $E_2 - m_{21}E_1$
- The third equation will be replaced by the equation $E_3 - m_{31}E_1$

Linear Systems: Naive Gaussian Elimination Method (contd.)

$$m_{21}E_1 \implies a_{21}x_1 + \frac{a_{12}a_{21}}{a_{11}}x_2 + \frac{a_{13}a_{21}}{a_{11}}x_3 = \frac{b_1 a_{21}}{a_{11}}$$

$$E_2 - m_{21}E_1 \implies 0 + (a_{22} - m_{21}a_{12})x_2 + (a_{23} - m_{21}a_{13})x_3 = b_2 - m_{21}b_1.$$

The new system is given by

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ 0 + a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 = b_2^{(2)} \\ 0 + a_{32}^{(2)}x_2 + a_{33}^{(2)}x_3 = b_3^{(2)}, \end{cases}$$

where the coefficients $a_{ij}^{(2)}$, and $b_k^{(2)}$ are given by

$$a_{ij}^{(2)} = a_{ij} - m_{i1}a_{1j}, \quad i, j = 2, 3$$

$$b_k^{(2)} = b_k - m_{k1}b_1, \quad k = 2, 3.$$

Naive Gaussian Elimination Method (contd.)

Step 2: If $a_{22}^{(2)} \neq 0$, then define $m_{32} = \dfrac{a_{32}^{(2)}}{a_{22}^{(2)}}$.

We will now obtain a new system that is equivalent to the system

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
m_{32} \quad 0 + a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 &= b_2^{(2)} \\
+ \quad 0 + a_{32}^{(2)}x_2 + a_{33}^{(2)}x_3 &= b_3^{(2)}.
\end{aligned}
$$

as follows:

- The first two equations are retained.
- The third equation will be replaced by the equation $E_3 - m_{32}E_2$.

Naive Gaussian Elimination Method (contd.)

Note the new system is given by

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
0 + a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 &= b_2^{(2)} \\
0 + 0 + a_{33}^{(3)}x_3 &= b_3^{(3)}.
\end{aligned}
$$

where the coefficient $a_{33}^{(3)}$, and $b_3^{(3)}$ are given by

$$
\begin{aligned}
a_{33}^{(3)} &= a_{33}^{(2)} - m_{32}a_{23}^{(2)}, \\
b_3^{(3)} &= b_3^{(2)} - m_{32}b_2^{(2)}.
\end{aligned}
$$

Naive Gaussian Elimination Method (contd.)

Finally, we obtained the system

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
0 + a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 &= b_2^{(2)} \\
0 + 0 + a_{33}^{(3)}x_3 &= b_3^{(3)}
\end{aligned}
\quad \Longleftrightarrow \quad \text{Given system}
$$

- $a_{33}^{(3)} \neq 0 \Rightarrow x_3$
- Put x_3 in (E_2) gives x_2
- Put x_2 and x_3 in (E_1) gives x_1.

This solution phase is called
backward substitution phase.

Naive Gaussian Elimination Method (contd.)

Gaussian elimination $\Rightarrow$ LU-factorization.

$$U = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} \\ 0 & 0 & a_{33}^{(3)} \end{pmatrix}.$$

Define a lower triangular matrix L by

$$L = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix}$$

It is easy to verify that $LU = A$.

Naive Gaussian Elimination Method (contd.)

Remark:

Why Naive?

Naive Gaussian Elimination Method (contd.)

Example:

Consider the linear system

$$6x_1 + 2x_2 + 2x_3 = -2$$
$$2x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3 = 1$$
$$x_1 + 2x_2 - x_3 = 0.$$

Let us compute using 4-digit rounding.

Naive Gaussian Elimination Method (contd.)

The system to be solved is

$$\begin{aligned}
6.000x_1 + 2.000x_2 + 2.000x_3 &= -2.000\\
2.000x_1 + 0.6667x_2 + 0.3333x_3 &= 1.000\\
1.000x_1 + 2.000x_2 - 1.000x_3 &= 0.000
\end{aligned}$$

After eliminating x_1 from the second and third equations, we get (with $m_{21} = 0.3333$, $m_{31} = 0.1667$).

$$\begin{aligned}
6.000x_1 + 2.000x_2 + 2.000x_3 &= -2.000\\
0.000x_1 + 0.0001x_2 - 0.3333x_3 &= 1.667\\
0.000x_1 + 1.667x_2 - 1.333x_3 &= 0.3334
\end{aligned}$$

Naive Gaussian Elimination Method (contd.)

After eliminating x_2 from the third equation, we get (with $m_{32} = 16670$)

$$6.000x_1 + 2.000x_2 + 2.000x_3 = -2.000$$
$$0.000x_1 + 0.0001x_2 - 0.3333x_3 = 1.667$$
$$0.000x_1 + 0.0000x_2 + 5555x_3 = -27790$$

Using back substitution, we get

$$x_1 = 1.335,\ x_2 = 0 \text{ and } x_3 = -5.003,$$

whereas the actual solution is

$$x_1 = 2.6,\ x_2 = -3.8 \text{ and } x_3 = -5.$$

Reason?

The coefficient of x_2 in (E_2) should have been zero, but rounding approximation prevented it.

Naive Gaussian Elimination Method (contd.)

The above examples highlight the inadequacy of the Naive Gaussian elimination method. These inadequcies can be overcome by modifying the procedure of Naive Gaussian elimination method. There are many kinds of modification. We will discuss one of the most popular modified methods which is called

Modified Gaussian elimination method with partial pivoting

Modified Gaussian elimination method with partial pivoting

Consider the following system of three linear equations in three variables x_1, x_2, x_3:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{cases}$$

For convenience, we call the first, second, and third equations by names E_1, E_2, and E_3 respectively.

Modified Gaussian elimination method with partial pivoting

Step 1: Define $s_1 = \max\{\,|a_{11}|, |a_{21}|, |a_{31}|\,\}$.

Note that $s_1 \not= 0$: **why?**

Let k be the least number such that $s_1 = |a_{k1}|$.

Interchange the first row and the k^{th} row. Rewritten system is

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 &= b_1^{(1)}\\
a_{21}^{(1)}x_1 + a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 &= b_2^{(1)}\\
a_{31}^{(1)}x_1 + a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 &= b_3^{(1)}
\end{aligned}
$$

Modified Gaussian elimination method with partial pivoting

Now eliminate the x_1 variable from the second and third equations of the system

$$\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 &= b_1^{(1)}\\
a_{21}^{(1)}x_1 + a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 &= b_2^{(1)}\\
a_{31}^{(1)}x_1 + a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 &= b_3^{(1)}.
\end{aligned}$$

Proceed as in naive Gaussian elimination method.

Modified Gaussian elimination method with partial pivoting

Step 2: Define $s_2 = \max\left\{ |a_{22}^{(2)}|, |a_{32}^{(2)}| \right\}$.

Let l be the least number such that $s_l = |a_{l2}^{(2)}|$.

Interchange the second row and the l^{th} rows. The rewritten system is

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 &= b_1^{(1)} \\
0 + a_{22}^{(3)}x_2 + a_{23}^{(3)}x_3 &= b_2^{(3)} \\
0 + a_{32}^{(3)}x_2 + a_{33}^{(3)}x_3 &= b_3^{(3)}.
\end{aligned}
$$

Modified Gaussian elimination method with partial pivoting

Note the new system is given by

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 &= b_1^{(1)} \\
0 + a_{22}^{(3)}x_2 + a_{23}^{(3)}x_3 &= b_2^{(3)} \\
0 + 0 + a_{33}^{(4)}x_3 &= b_3^{(4)},
\end{aligned}
$$

where the coefficient $a_{33}^{(4)}$, and $b_3^{(4)}$ are given by

$$
\begin{aligned}
a_{33}^{(4)} &= a_{33}^{(3)} - m_{32}a_{23}^{(3)}, \\
b_3^{(4)} &= b_3^{(3)} - m_{32}b_2^{(3)}.
\end{aligned}
$$

Note that the variable x_2 has been eliminated from the last equation. This phase is called Forward elimination phase with partial

Modified Gaussian elimination method with partial pivoting

The reduced system is

$$
\begin{aligned}
a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 &= b_1^{(1)} \\
0 + a_{22}^{(3)}x_2 + a_{23}^{(3)}x_3 &= b_2^{(3)} \\
0 + 0 + a_{33}^{(4)}x_3 &= b_3^{(4)}.
\end{aligned}
$$

Now do the Backward substitution phase.

Modified Gaussian elimination method with partial pivoting

Recall the system

$$
\begin{aligned}
6x_1 + 2x_2 + 2x_3 &= -2 \\
2x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3 &= 1 \\
x_1 + 2x_2 - x_3 &= 0.
\end{aligned}
$$

was solved using Gaussian Elimination method with four digit rounding arithmetic. Recall the reduced system after Step 1 was

$$
\begin{aligned}
6.000x_1 + 2.000x_2 + 2.000x_3 &= -2.000 \\
0.000x_1 + 0.0001x_2 - 0.3333x_3 &= 1.667 \\
0.000x_1 + 1.667x_2 - 1.333x_3 &= 0.3334
\end{aligned}
$$

Modified Gaussian elimination method with partial pivoting

The final system is (with $m_{32} = 0.00005999$)

$$
\begin{aligned}
6.000x_1 + 2.000x_2 + 2.000x_3 &= -2.000 \\
0.000x_1 + 1.667x_2 - 1.333x_3 &= 0.3334 \\
0.000x_1 + 0.0000x_2 - 0.3332x_3 &= 1.667
\end{aligned}
$$

with back substitution, we obtain the approximate solution as

$$x_1 = 2.602,\ x_2 = -3.801 \text{ and } x_3 = -5.003.$$

Recall, the solution obtained without pivoting was

$$x_1 = 1.335,\ x_2 = 0 \text{ and } x_3 = -5.003,$$

whereas the actual solution is

$$x_1 = 2.6,\ x_2 = -3.8 \text{ and } x_3 = -5.$$