Lectures 1 and 2
System of Linear Equations

Dipankar Ghosh

Department of Mathematics
Indian Institute of Technology Hyderabad

January 03, 2020

Welcome!

- Welcome to my course Elementary Linear Algebra.
- We'll study Linear Algebra in the next six weeks.
- I have made the lecture notes. So you can follow that.
- 100% attendance is compulsory. Then only you are allowed to write the final exam.
- In case you need any further assistance, please get in touch with me.
- My office is Academic Block C, Room No. 313-D.
- Email ID is dghosh@iith.ac.in

What is Linear Algebra?

- Linear algebra is the branch of mathematics concerning linear equations such as

$$a_1x_1 + \cdots + a_nx_n = b$$

linear functions such as

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \ldots + a_nx_n$$

and their representations through matrices and vector spaces.

- It is central to almost all areas of mathematics.
- For instance, linear algebra is fundamental in modern presentations of geometry: for describing basic objects such as lines, planes and rotations.
- It is also used in most sciences and engineering areas, because it allows modeling many natural phenomena, and efficiently computing with such models. I will get back to this point later.

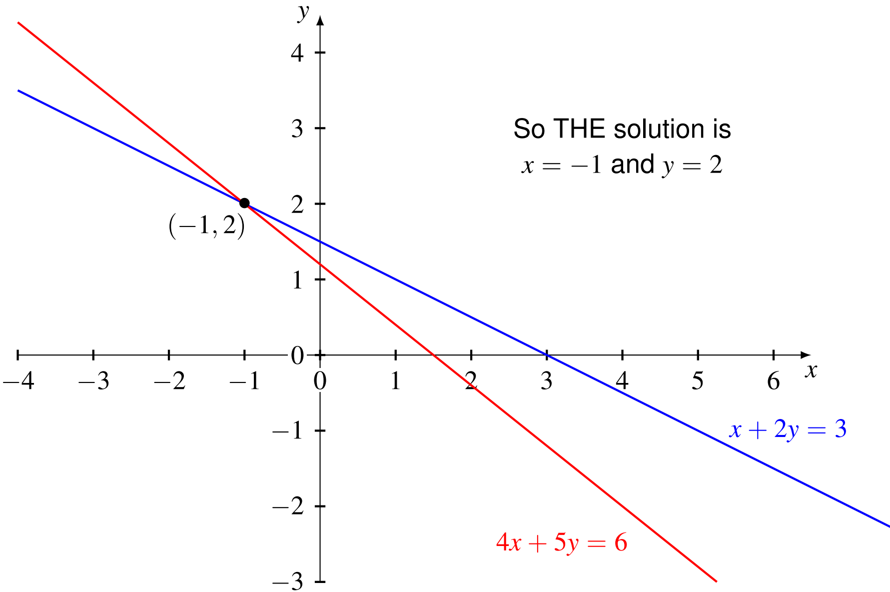
Solving linear equations

- One of the central problem of linear algebra is '**solving linear equations**'.
- Consider the following **system of linear equations:**

$$\begin{aligned} x + 2y &= 3 \qquad \text{(1st equation)} \\ 4x + 5y &= 6 \qquad \text{(2nd equation)}. \end{aligned}$$

- Here x and y are the unknowns. We want to solve this system, i.e., we want to find the values of x and y in $\mathbb{R}$ such that the equations are satisfied.

What does it mean geometrically?



So THE solution is
$x = -1$ and $y = 2$

How can we solve the system?

- We can solve the system by **Gaussian Elimination**. The original system is

$$\begin{aligned} x + 2y &= 3 \qquad \text{(1st equation)} \\ 4x + 5y &= 6 \qquad \text{(2nd equation)}. \end{aligned} \tag{1}$$

- We want to change it into an equivalent system, which is comparatively easy to solve.
- Eliminating x from the 2nd equation, we obtain a **triangulated** system:

$$\begin{aligned} x + 2y &= 3 \qquad \text{(equation1)} \\ -3y &= -6 \qquad \text{(equation 2) - 4(equation 1)}. \end{aligned} \tag{2}$$

- Both the systems have same solutions. We can solve the 2nd system by **Back-substitution**. What is it?
- In this case, the solution is $y = 2$, $x = -1$.

Another method to solve the system: Cramer's Rule

- The system can be written as

$$\begin{aligned} x + 2y &= 3 \\ 4x + 5y &= 6 \end{aligned} \quad \text{or} \quad \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

- The solution depends completely on those six numbers in the equations. There must be a formula for x and y in terms of those six numbers. Cramer's Rule provides the formula:

$$x = \frac{\begin{vmatrix} 3 & 2 \\ 6 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = \frac{3 \cdot 5 - 6 \cdot 2}{1 \cdot 5 - 4 \cdot 2} = \frac{3}{-3} = -1$$

$$y = \frac{\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = \frac{1 \cdot 6 - 4 \cdot 3}{1 \cdot 5 - 4 \cdot 2} = \frac{-6}{-3} = 2.$$

Which approach is better?

- The direct use of the **determinant formula** for large number of equations and variables would be very difficult.
- So the better method is **Gaussian Elimination**. Let's study it systematically.
- We understand the Gaussian Elimination method by examples.

How many solutions do exist for a given system?

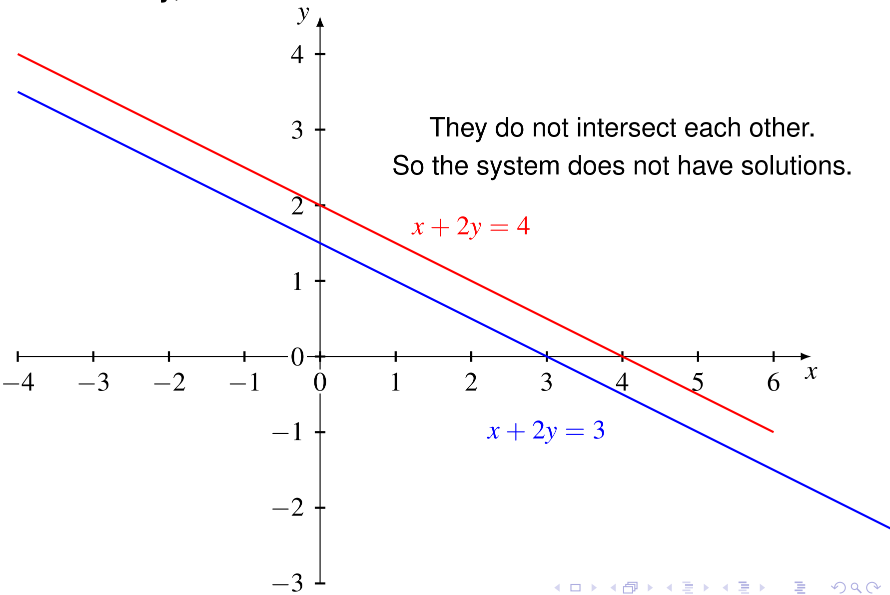
- A system may have only ONE solution. For example the system which we have already discussed.
- A system may NOT have a solution at all. For example

$$\left.\begin{aligned} x + 2y &= 3 \\ x + 2y &= 4 \end{aligned}\right\}. \quad \text{After Gaussian Elimination} \quad \begin{aligned} x + 2y \quad &= 3 \\ \color{red}{\mathbf{0}} \quad\quad &\color{red}{= \mathbf{1}} \end{aligned}$$

- This is absurd. So the system does not have solutions.

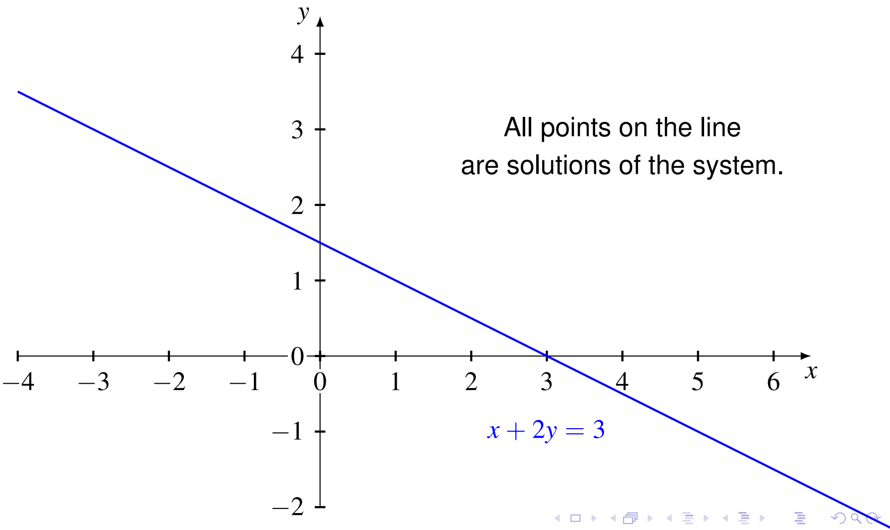
A system may NOT have a solution at all

Geometrically,



A system may have infinitely many solutions

$$\left.\begin{aligned} x + 2y &= 3 \\ 2x + 4y &= 6 \end{aligned}\right\}$$ equivalent to the system $x + 2y = 3$



Applications of system of linear equations

- Systems of linear equations appear in a wide variety of applications. Let us discuss a few here.
- We will discuss two such applications:
  1. Polynomial Curve Fitting.
  2. Networks and Kirchhoff's Laws for electricity.

Polynomial curve fitting

- Suppose n points are given in the xy-plane:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n).$$

  They represent a collection of data.

- For example, suppose x_i represents a time in a particular day, and y_i represents the number of cars at that time in a particular road. Suppose we have collected n such data for a particular day. Now we want to estimate the traffic of that road for the whole day with this data. One way we can do this by finding a polynomial function of degree $(n-1)$:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

  whose graph passes through these n points. To solve for the coefficients of $p(x)$, substitute each of the points into the polynomial function and obtain linear equations in variables $a_0, a_1, \ldots, a_{n-1}$.

Polynomial curve fitting

- In this way, we obtain n linear equations in n unknowns:

$$
\begin{aligned}
a_0 + a_1x_1 + a_2x_1^2 + \cdots + a_{n-1}x_1^{n-1} &= p(x_1) = y_1 \\
a_0 + a_1x_2 + a_2x_2^2 + \cdots + a_{n-1}x_2^{n-1} &= y_2 \\
&\vdots \\
a_0 + a_1x_n + a_2x_n^2 + \cdots + a_{n-1}x_n^{n-1} &= y_n
\end{aligned}
$$

- By solving this system of equations, we get $p(x)$.
- Then we can compute the (approximate) traffic $p(x')$ for the road at every time x'.

Kirchhoff's circuit laws: Physics concept

- Kirchhoff's laws are fundamental in circuit theory. They quantify how current flows through a circuit and how voltage varies around a loop in a circuit.
- Kirchhoff's current law (1st Law) states that current flowing into a node (or a junction) must be equal to current flowing out of it.

An example to understand Gaussian Elimination

- Consider the system:

$$\begin{aligned} v + w &= 5 \\ 4u - 6v &= -2 \\ -2u + 7v + 2w &= 9 \end{aligned}$$

- There is no harm to **interchange the positions of two equations**. This is called the row operation of type 1. So the original system is equivalent to the following system.

$$\begin{aligned} 4u - 6v &= -2 \\ v + w &= 5 \\ -2u + 7v + 2w &= 9 \end{aligned}$$

- What was the aim? To change the system so that the coefficient of u in the 1st equation becomes non-zero.

An example to understand Gaussian Elimination...

- So the system becomes

$$\begin{aligned} \mathbf{4}u - 6v &= -2 \\ v + w &= 5 \\ -2u + 7v + 2w &= 9 \end{aligned}$$

- We call the coefficient 4 as the **first pivot**.
- There is no harm if we **multiply an equation by a non-zero constant**. This is called the row operation of type 2. So we can always make the pivot element 1.
- We now eliminate u from the 3rd equation by adding $(1/2)$ times the 1st equation to the 3rd equation.

$$\begin{aligned} \mathbf{4}u - 6v &= -2 \\ \mathbf{1} \cdot v + w &= 5 \\ 4v + 2w &= 8 \end{aligned}$$

- This is called the **row operation of 3rd type**. We already got the 2nd pivot. In the last stage, we eliminate v from the 3rd equation. Apply (3rd eqn) $-$ 4 (2nd eqn).

Triangular system and back-substitution

- After the elimination process, we obtain a **triangular system**:

$$\begin{aligned} \color{red}{4}u - 6v + 0w &= -2 \\ \color{red}{1} \cdot v + w &= 5 \\ \color{red}{-2}w &= -12 \end{aligned}$$

- Now the system can be solved by **backward substitution**, bottom to top. The red colored coefficients are pivots.
- The last equation gives $w = 6$.
- Substituting $w = 6$ into the 2nd equation, we find $v = -1$.
- Substituting $w = 6$ and $v = -1$ into the 1st equation, we get $u = -2$.

Gaussian Elimination process in short

$$
\begin{aligned}
&\text{Original System} \\
&\quad\Downarrow \text{ Forward Elimination} \\
&\text{Triangular System} \\
&\quad\Downarrow \text{ Backward Substitution} \\
&\text{Solution}
\end{aligned}
$$

In the Gaussian Elimination process, the 3 types of row operations are called **elementary row operations.**

Augmented matrix of the system

- Consider the system:

$$\begin{aligned} v + w &= 5 \\ 4u - 6v &= -2 \\ -2u + 7v + 2w &= 9 \end{aligned}$$

- The **coefficient matrix** of the system is given by

$$\begin{bmatrix} 0 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$$

- The **augmented matrix** of the system is given by

$$\begin{bmatrix} 0 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix}$$

The forward elimination steps

- The forward elimination steps can be described as follows.

$$\begin{bmatrix} 0 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \overset{R1 \leftrightarrow R2}{\Longrightarrow} \begin{bmatrix} 4 & -6 & 0 & -2 \\ 0 & 1 & 1 & 5 \\ -2 & 7 & 2 & 9 \end{bmatrix} \overset{R3 \to R3 + (1/2)R1}{\Longrightarrow}$$

$$\begin{bmatrix} 4 & -6 & 0 & -2 \\ 0 & 1 & 1 & 5 \\ 0 & 4 & 2 & 8 \end{bmatrix} \overset{R3 \to R3 - 4 \cdot R2}{\Longrightarrow} \begin{bmatrix} 4 & -6 & 0 & -2 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & -2 & -12 \end{bmatrix}$$

- Now one can solve the corresponding system by back substitution. This is the reason we call the operations in the Gaussian Elimination process as elementary row operations.
- In this case, where we have a full set of 3 pivots, there is only one solution.

When we have less pivots than 3

- When we have less pivots than 3, i.e., if a zero appears in a pivot position, then the system may not have solution at all, or it can have infinitely many solutions.
- For example, if the augmented matrix corresponding to a system has the form

$$\begin{bmatrix} 1 & 1 & 1 & * \\ 2 & 2 & 5 & * \\ 4 & 4 & 8 & * \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 1 & 1 & * \\ 0 & \mathbf{0} & 3 & * \\ 0 & 0 & 4 & * \end{bmatrix}.$$

- Now consider some particular values of $*$.

When we have less pivots than 3 contd...

- $\begin{bmatrix} 1 & 1 & 1 & * \\ 0 & 0 & 3 & * \\ 0 & 0 & 4 & * \end{bmatrix}$. Considering some particular values of $*$,
- Example 1: $\begin{bmatrix} 1 & 1 & 1 & * \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 4 & 7 \end{bmatrix} \xrightarrow{R3-(4/3)R2} \begin{bmatrix} 1 & 1 & 1 & * \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.
- The corresponding system is

$$\begin{aligned} u + v + w &= * \\ 3w &= 6 \\ 0 &= -1 \end{aligned}$$

- This system does not have solution.

When we have less pivots than 3 contd...

- $\begin{bmatrix} 1 & 1 & 1 & * \\ 0 & 0 & 3 & * \\ 0 & 0 & 4 & * \end{bmatrix}$. Considering some particular values of $*$,
- Example 2: $\begin{bmatrix} 1 & 1 & 1 & * \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 4 & 8 \end{bmatrix} \xrightarrow{R3-(4/3)R2} \begin{bmatrix} 1 & 1 & 1 & * \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.
- The corresponding system is

$$\begin{aligned} u + \mathbf{v} + w &= * \\ 3w &= 6 \end{aligned}$$

- This system has infinitely many solutions.
- From the last equation, we get $w = 2$.
- Substituting $w = 2$ to the 1st equation, we have $u + v = *$, which has infinitely many solutions. We call v a free variable.

System of linear equations (in general)

- Consider a system of m linear equations in n variables $x_1, \ldots, x_n$.

$$
\begin{aligned}
A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\
A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\
&\vdots \\
A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m
\end{aligned}
$$

- Here $A_{ij}, b_i \in \mathbb{R}$, and $x_1, \ldots, x_n$ are unknown. We try to find the values of $x_1, \ldots, x_n$ in $\mathbb{R}$ satisfied by the system.
- Any n tuple $(x_1, \ldots, x_n)$ of elements of $\mathbb{R}$ which satisfies the system (i.e., which satisfies every equation of the system) is called a **solution** of the system.
- If $b_1 = \cdots = b_m = 0$, then it is called a **homogeneous system**.
- Every homogeneous system has a trivial solution $x_1 = \cdots = x_n = 0$. What about non-homogeneous system?

Equivalent systems of linear equations

- Consider a system of linear equations:

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\ &\vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m \end{aligned}$$

- Suppose the following system is obtained by applying elementary row operations on the 1st system.

$$\begin{aligned} B_{11}x_1 + B_{12}x_2 + \cdots + B_{1n}x_n &= b'_1 \\ B_{21}x_1 + B_{22}x_2 + \cdots + B_{2n}x_n &= b'_2 \\ &\vdots \\ B_{m1}x_1 + B_{m2}x_2 + \cdots + B_{mn}x_n &= b'_m \end{aligned}$$

- Since the elementary row operations are invertible with inverses of same types, the 1st system can also be obtained from the 2nd system by applying elementary row operations.
- In this case, we call that the **two systems are equivalent**.

Equivalent systems have same solutions set

- Consider two **equivalent systems:**

$$\begin{aligned}
A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\
A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\
&\vdots \\
A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m
\end{aligned}$$

and

$$\begin{aligned}
B_{11}x_1 + B_{12}x_2 + \cdots + B_{1n}x_n &= b'_1 \\
B_{21}x_1 + B_{22}x_2 + \cdots + B_{2n}x_n &= b'_2 \\
&\vdots \\
B_{m1}x_1 + B_{m2}x_2 + \cdots + B_{mn}x_n &= b'_m
\end{aligned}$$

- Then they have the same set of solutions.

Writing a system of linear equations by matrices

- Consider a system:

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\ &\vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m \end{aligned}$$

- We write the system by matrices as follows: $Ax = b$, where

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

- For a non-homogeneous system, we apply **elementary row operations** on the augmented matrix $(A \mid b)$.

A homogeneous system of linear equations

- For a homogeneous system:

$$
\begin{aligned}
A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= 0 \\
A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= 0 \\
&\vdots \\
A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= 0
\end{aligned}
$$

- i.e., when the system is $Ax = 0$, then it is enough to consider the coefficient matrix A.
- So, in this case, we apply elementary row operations on A.

The elementary row operations (total three)

1. Interchange of two rows of A, say rth and sth rows.
2. Multiplication of one row of A by a non-zero scalar $c \in \mathbb{R}$.
3. Replacement of the rth row of A by
$$(r\text{th row} + c \cdot s\text{th row}),$$
where $c \in \mathbb{R}$ and $r \neq s$.

All the above three **operations are invertible**, and each has inverse operation of the **same type**.

4. The 1st one is it's own inverse.
5. For the 2nd one, inverse operation is 'multiplication of that row of A by $1/c \in \mathbb{R}$'.
6. For the 3rd one, inverse operation is 'replacement of the rth row of A by (rth row $- c \cdot s$th row)'.

Row equivalence of matrices

- Let A and B be two $m \times n$ matrices over $\mathbb{R}$.
- We say that B **is row equivalent to** A if B can be obtained from A by a finite sequence of elementary row operations, i.e.,

$$B = e_r \cdots e_2 e_1(A),$$

  where e_i are some elementary row operations.
- 'Row equivalence' is an '**equivalence relation**':
- 'Row equivalence' is reflexive, i.e., A is row equivalent to A.
- 'Row equivalence' is symmetric, i.e.,

$$B = e_r \cdots e_2 e_1(A) \implies A = (e_1)^{-1}(e_2)^{-1} \cdots (e_r)^{-1}(B).$$

  In this case, we say that A and B are row equivalent.
- 'Row equivalence' is transitive, i.e., if B is row equivalent to A and C is row equivalent to B, then C is row equivalent to A.

Row equivalence of two homogeneous systems

- Among three elementary row operations, considering row operation of each type, we observe that we observe that:
- Two matrices A and B are row equivalent

  if and only if

  the corresponding homogeneous systems $Ax = 0$ and $Bx = 0$ are equivalent.
- In this case, both the systems have exactly the same solutions.
- For non-homogeneous systems, the augmented matrices $(A|b)$ and $(B|c)$ are row equivalent

  if and only if

  the corresponding systems $Ax = b$ and $Bx = c$ are equivalent.
- In this case, both $Ax = b$ and $Bx = c$ have the same solutions.

Row reduced matrix

Definition

An $m \times n$ matrix A over $\mathbb{R}$ is called **row reduced** if
(1) the 1st non-zero entry in each non-zero row of A is equal to 1;
(2) each column of A which contains the leading non-zero entry of some row has all its other entries 0.

Example

(i) $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ $\surd$ (ii) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ $\times$

(iii) $\begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$ $\times$ (iv) $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$ $\surd$

Here (ii) and (iii) are not row reduced matrices.

Row reduced echelon matrix

Definition

An $m \times n$ matrix A over $\mathbb{R}$ is called **row reduced echelon** matrix if
(1) A is row reduced;
(2) every zero row (?) of A occurs below every non-zero row (?);
(3) if rows $1, \ldots, r$ are the non-zero rows, and if the leading non-zero entry of row i occurs in column k_i for $1 \leqslant i \leqslant r$,
then $k_1 < k_2 < \cdots < k_r$.

In this case, (i, k_i) are called the **pivot positions**, and x_{k_i} are called the **pivot variables**.

Example

(i) $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ $\surd$ (ii) $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix}$ $\times$

The matrix in (ii) is row reduced, but NOT row reduced echelon.

Every matrix is row equivalent to a row reduced echelon matrix

Theorem

Every $m \times n$ matrix over $\mathbb{R}$ is **row equivalent** *to a row reduced echelon matrix.*

Example

$$\begin{bmatrix} 0 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xRightarrow{Row\ \ operations} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 6 \end{bmatrix}.$$

Exercise. Let A be an $n \times n$ row reduced echelon matrix over $\mathbb{R}$. Show that A is invertible if and only if A is the identity matrix.

Example: A matrix $\to$ Row reduced echelon matrix

$$\begin{bmatrix} 0 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xRightarrow{R1 \leftrightarrow R2} \begin{bmatrix} 4 & -6 & 0 & -2 \\ 0 & 1 & 1 & 5 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xRightarrow{R1 \to (1/4)R1}$$

$$\begin{bmatrix} 1 & -3/2 & 0 & -1/2 \\ 0 & 1 & 1 & 5 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xRightarrow{R3 \to R3 + 2 \cdot R1} \begin{bmatrix} 1 & -3/2 & 0 & -1/2 \\ 0 & 1 & 1 & 5 \\ 0 & 4 & 2 & 8 \end{bmatrix} \xRightarrow{R3 \to R3 - 4 \cdot R2}$$

$$\begin{bmatrix} 1 & -3/2 & 0 & -1/2 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & -2 & -12 \end{bmatrix} \xRightarrow{R3 \to (-1/2)R3} \begin{bmatrix} 1 & -3/2 & 0 & -1/2 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & 6 \end{bmatrix}$$

(Triangular system with pivot entries 1)

$$\xRightarrow{R2 \to R2 - R3} \begin{bmatrix} 1 & -3/2 & 0 & -1/2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 6 \end{bmatrix} \xRightarrow{R1 \to R1 + (3/2)R2} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 6 \end{bmatrix}$$

So it is just combination of forward and backward eliminations.

Solution of a system corr. to a row reduced echelon

Example

$$\begin{bmatrix} 0 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \xRightarrow{Row\ operations} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 6 \end{bmatrix}.$$

Considering the corresponding system, we have the solution $u = -2$, $v = -1$ and $w = 6$.

Solution of a system corr. to a row reduced echelon

Consider the homogeneous system corr. to the coefficient matrix

$$\begin{bmatrix} 0 & 1 & -3 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ (which is a row reduced echelon matrix).

Here $(1,2)$ and $(2,4)$ are the pivot positions. So x_2 and x_4 are the pivot variables. The remaining variables are called **free variables**.

$$\begin{array}{rrrrcl} x_2 & -3x_3 & & +(1/2)x_5 & = & 0 \\ & & x_4 & +2x_5 & = & 0 \end{array}$$

which yields that

$$\begin{aligned} x_2 &= 3x_3 - (1/2)x_5 \\ x_4 &= -2x_5 \end{aligned}$$

The values of x_1, x_3 and x_5 can be chosen freely.

Solution of a system corr. to a row reduced echelon

- Consider $Ax = 0$, where A is a row reduced echelon matrix.
- Let rows $1, \dots, r$ be non-zero, and the leading non-zero entry of row i occurs in column k_i.
- The system $Ax = 0$ then consists of r non-trivial equations.
- The variables $x_{k_1}, x_{k_2}, \dots, x_{k_r}$ are the pivot variables.
- Let $u_1, \dots, u_{n-r}$ denote the remaining $n - r$ (free) variables.
- Then the r non-trivial equations of $Ax = 0$ can be written as

$$x_{k_1} + \sum_{j=1}^{n-r} C_{1j} u_j = 0$$

$$\dots$$

$$x_{k_r} + \sum_{j=1}^{n-r} C_{rj} u_j = 0$$

- We may assign any values to $u_1, \dots, u_{n-r}$. Then $x_{k_1}, \dots, x_{k_r}$ are determined uniquely by those assigned values.

Solution to a homogeneous system (when $m < n$)

Theorem

Let A be an $m \times n$ matrix over $\mathbb{R}$ with $m < n$. Then the homogeneous system $Ax = 0$ has a non-trivial solution. In fact (over $\mathbb{R}$) it has infinitely many solutions.

Proof. The matrix A is row equivalent to a row reduced echelon matrix B. Then $Ax = 0$ and $Bx = 0$ have the same solutions. If r is number of non-zero rows, then $r \leqslant m < n$. The system $Bx = 0$ can be expressed as

$$x_{k_1} + \sum_{j=1}^{n-r} C_{1j} u_j = 0, \qquad \ldots, \qquad x_{k_r} + \sum_{j=1}^{n-r} C_{rj} u_j = 0$$

where $u_1, \ldots, u_{n-r}$ are the free variables. Now assign any values to $u_1, \ldots, u_{n-r}$ to get infinitely many solutions.

Solution to a homogeneous system (when $m = n$)

Theorem

Let A be an $n \times n$ matrix over $\mathbb{R}$. Then A is row equivalent to the $n \times n$ identity matrix if and only if the system $Ax = 0$ has only the trivial solution.

Proof.

The matrix A is row equivalent to a row reduced echelon matrix B. Then $Ax = 0$ and $Bx = 0$ have the same solutions. If r is number of non-zero rows, then $r \leqslant n$. The system $Bx = 0$ can be expressed as

$$x_{k_1} + \sum_{j=1}^{n-r} C_{1j} u_j = 0, \ldots, x_{k_r} + \sum_{j=1}^{n-r} C_{rj} u_j = 0,$$

where $u_1, \ldots, u_{n-r}$ are the free variables. Hence it can be observed that $B = I_n$ is the identity matrix if and only if $r = n$ if and only if the system has the trivial solution. $\square$

Solution to a non-homogeneous system $Ax = b$

- Consider the augmented matrix $(A \,|\, b)$ corr. to $Ax = b$.
- Apply elementary row operations on $(A \,|\, b)$ to get row reduced echelon form $(B \,|\, c)$.
- The systems $Ax = b$ and $Bx = c$ are equivalent, and hence they have the same solutions.
- Let $1, \ldots, r$ be the non-zero rows of B, and the leading non-zero entry of row i occurs in column k_i.
- The variables $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ are the pivot variables.
- Let $u_1, \ldots, u_{n-r}$ denote the remaining $n - r$ (free) variables.
- Then the system $Bx = c$ can be written as

$$x_{k_1} + \sum_{j=1}^{n-r} C_{1j} u_j = c_1, \ \ldots, \ x_{k_r} + \sum_{j=1}^{n-r} C_{rj} u_j = c_r$$

$$0 = c_{r+1}$$

$$\ldots$$

$$0 = c_m$$

- We may assign any values to $u_1, \ldots, u_{n-r}$. Then $x_{k_1}, \ldots, x_{k_r}$ are determined uniquely by those assigned values.

Solution to a non-homogeneous system $Ax = b$ contd...

- The systems $Ax = b$ and $Bx = c$ are equivalent, and hence they have the same solutions.
- The system $Bx = c$ can be written as

$$x_{k_1} + \sum_{j=1}^{n-r} C_{1j}u_j = c_1, \ \ldots, \ x_{k_r} + \sum_{j=1}^{n-r} C_{rj}u_j = c_r$$
$$0 = c_{r+1}$$
$$\ldots$$
$$0 = c_m$$

- Thus the system $Ax = b$ (equivalenly, $Bx = c$) has a solution if and only if $c_{r+1} = \cdots = c_m = 0$. IN THIS CASE:
- $r = n$ if and only if the system has a unique solution.
- $r < n$ if and only if the system has infinitely many solutions.

Example: Solution to a non-homogeneous system

- Consider a system $Ax = b$, where $A = \begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & 1 \\ 0 & 5 & -1 \end{pmatrix}$.
- The corr. augmented matrix is $(A \mid b) = \begin{pmatrix} 1 & -2 & 1 & b_1 \\ 2 & 1 & 1 & b_2 \\ 0 & 5 & -1 & b_3 \end{pmatrix}$.
- Applying elementary row operations on $(A \mid b)$, we get $\begin{pmatrix} 1 & 0 & \frac{3}{5} & \frac{1}{5}(b_1 + 2b_2) \\ 0 & 1 & -\frac{1}{5} & \frac{1}{5}(b_2 - 2b_1) \\ 0 & 0 & 0 & (b_3 - b_2 + 2b_1) \end{pmatrix}$.
- The system $Ax = b$ has a solution if and only if $b_3 - b_2 + 2b_1 = 0$. In this CASE,
- $x_1 = -\frac{3}{5}x_3 + \frac{1}{5}(b_1 + 2b_2)$ and $x_2 = \frac{1}{5}x_3 + \frac{1}{5}(b_2 - 2b_1)$.
- Assign any value to x_3, and compute x_1, x_2.

Elementary matrices

Definition

An $m \times m$ matrix is called an **elementary matrix** if it can be obtained from the $m \times m$ identity matrix by applying a SINGLE elementary row operation.

Example

(i) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. (obtained by applying 1st type elementary row oper.).

(ii) $\begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}$, $c \neq 0$. (iii) $\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}$, $c \neq 0$. (2nd type).

(iv) $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$, $c \in \mathbb{R}$. (v) $\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$, $c \in \mathbb{R}$. (3rd type).

These are all the 2×2 elementary matrices.

Elementary matrices vs elementary row operation

- Consider a matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$.
- $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$. $(R1 \leftrightarrow R2.)$
- $\begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} c & 2c & 3c \\ 4 & 5 & 6 \end{pmatrix}$. $(R1 \to c \cdot R1.)$
- $\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4+c & 5+2c & 6+3c \end{pmatrix}$.
  $(R2 \to R2 + c \cdot R1.)$
- So applying an *elementary row operation* on a matrix is same as left multiplying by the corresponding *elementary matrix*.

Theorem on elementary matrices and elementary row operation

Theorem

Let e be an elementary row operation. Let E be the corresponding $m \times m$ elementary matrix, i.e., $E = e(I_m)$, where I_m is the $m \times m$ identity matrix. Then, for every $m \times n$ matrix A,

$$EA = e(A).$$

Corollary

Let A and B be two $m \times n$ matrices. Then
A and B are equivalent

if and only if

$B = PA$, where P is a product of some $m \times m$ elementary matrices.

Elementary matrices are invertible

Theorem

Every elementary matrix is invertible.

Proof.

Let E be an elementary matrix corresponding to the elementary row operation e, i.e., $E = e(I)$. Note that e has an inverse operation, say e'. Set $E' := e'(I)$. Then

$$EE' = e(E') = e(e'(I)) = I \text{ and}$$
$$E'E = e'(E) = e'(e(I)) = I.$$

$\square$

Invertible matrices

Theorem

Let A be an $n \times n$ matrix. Then the following are equivalent:
(1) A is invertible.
(2) A is row equivalent to the $n \times n$ identity matrix.
(3) A is a product of some elementary matrices.

Proof. Let A be row-equivalent to a row-reduced echelon matrix B. Then

$$B = E_k \cdots E_2 E_1 A \tag{3}$$

Since elementary matrices are invertible, we have

$$E_1^{-1} E_2^{-1} \cdots E_k^{-1} B = A. \tag{4}$$

Hence A is invertible if and only if B is invertible if and only if $B = I$ (since B is row-reduced echelon) if and only if $E_1^{-1} E_2^{-1} \cdots E_k^{-1} = A$.

Invertible matrices

Theorem

Let A be an $n \times n$ invertible matrix. If a sequence of elementary row operations reduces A to the identity I, then that same sequence of operations when applied to I yields A^{-1}.

Proof.

Note that if $E_k \cdots E_2 E_1 A = I$, then

$$A^{-1} = (E_k \cdots E_2 E_1) = E_k \cdots E_2 E_1 (I).$$

$\square$

Example: How to compute inverse of a matrix

Let $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$. Want to compute A^{-1}. Consider $(A|I_2) =$

$$\begin{pmatrix} 2 & -1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix} \xrightarrow{\frac{1}{2}R1} \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix} \xrightarrow{R2 \to R2 - R1}$$

$$\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{7}{2} & -\frac{1}{2} & 1 \end{pmatrix} \xrightarrow{\frac{2}{7}R2}$$

$$\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} \end{pmatrix} \xrightarrow{R1 \to R1 + \frac{1}{2}R2} \begin{pmatrix} 1 & 0 & \frac{3}{7} & \frac{1}{7} \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} \end{pmatrix}$$

So $A^{-1} = \begin{pmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{pmatrix}$.

Thank You!