Informal Notes about PROOFS

- **What is a PROOF?**

A proof is a finite logical sequence of TRUE statements culminating in the statement for which a proof was required.
(Thus the conclusion is also TRUE)

- **What types of Statements are acceptable (admissible) in a PROOF?**

In this case

- A statement is a sentence that could be either TRUE or FALSE. Sentences like 'command' are not acceptable.
- Statements which are acceptable as TRUE are:

(1) Assumptions or Hypotheses.

(2) Definitions

(3) Previously known results

(4) Statements which logically follow from earlier statements in the sequence.

Note: Most likely errors in proofs occur in (4).

Typical Propositional Forms

(1) Implication : This is the most common form.

Symbolic Notation : $p \Rightarrow q$

Language Expression : If p, then q

OR p implies q.

(2) The Biconditional :

Symbolic Notation : $p \Leftrightarrow q$

Language Expression : p if and only if q

OR : p is equivalent to q

Note: This is a way to combine two implications. It means $p \Rightarrow q$ and $q \Rightarrow p$ (The converse of $p \Rightarrow q$)

Types of Proofs (for $p \Rightarrow q$):

(1) Direct Proof :

$p \longrightarrow$ True (Assumption): Then a logical sequence of true statement follows.

$$\left.\begin{array}{l} p_1 \\ \vdots \\ p_n \end{array}\right\} \text{(All True)}$$

$q \longrightarrow$ The end : True (Conclusion)

(2) Contrapositive Proof:

- The contrapositive of $p \Rightarrow q$ is

$\sim q \Rightarrow \sim p$ (The symbol '$\sim$' stands for negation or not.)

Now $p \Rightarrow q$ and $\sim q \Rightarrow \sim p$ are equivalent. (By a Theorem of Logic)

This provides another type of proof

We start with: $\sim q \rightarrow$ True (Assumption)

$$\left.\begin{matrix} p_1 \\ \vdots \\ p_n \end{matrix}\right\} \text{(All True)}$$

$\sim p \rightarrow$ True (The desired conclusion)

Conclusion: Proving $\sim q \Rightarrow \sim p$ yields a proof of $p \Rightarrow q$

(3) Proof BY WAY OF CONTRADICTION (BWOC):

Here we prove p and $\sim q \Longrightarrow r$

(Notation: $p \wedge \sim q$ ($\wedge \equiv$ and))

where r is some proposition known to be false (NOT TRUE).

Thus we start with:

$p \to$ True (By the given assumption)

$\sim q \to$ New assumption ⋯⋯⋯ ⊛
(Introduced by the Prover)

⋮

$r \to$ follows from the proof

$\sim r \to$ True (Previously known)

So, r and $\sim r$ is true which is not possible.
Thus there is something wrong in the proof.
The only possible place for error is the assumption ⊛. Thus $\sim q$ is false.
So, q is true and hence $p \Rightarrow q$ is true which is what we wanted.

SOME TIPS FOR WRITING PROOFS

(1) Be clear about
- what is given (assumptions/hypotheses)

and
- What is the desired conclusion (to be proved)

You can write: GIVEN
RTP (Required to be proved)

at the start of the proof

- This is particularly necessary for 'if and only if' propositions.

(2) Pay attention to notation:

- Write down what each symbol stands for.
- Do not introduce a new symbol without stating what it stands for.
- Do not use the same symbol for two different objects.
- Clearly distinguish vectors and scalars.

(3) Check that each statement (step) in the proof is legitimate.

Note: A common error is to write a statement which is not a known result. It may be false or just as hard to prove as the desired conclusion.

(4) Use short simple statements as steps in the proof: with explanation if necessary (put in brackets if appropriate).

(5) Use precise mathematical language and symbols/equations as far as possible.

(6) Objectives of the proof:

Every proof is written for a certain class of readers and it has to be appropriate for them.

e.g. (1) Proofs in journal articles: By

experts for experts : Brief with many steps left out.

(2) Text books and Lecture notes: For learners:

— Usually steps are not left out and there are extra explanations

(3) By a student in a test: Your aim is to convince the examiner that you have understood the logic. Do not leave gaps and cite used results explicitly.

Linear Algebra Lecture 1

First consider an example:

$$
\begin{aligned}
x_1 - 7x_2 + 2x_3 - 5x_4 + 8x_5 &= 10 \\
x_2 - 3x_3 + 3x_4 + x_5 &= -5 \\
x_1 \qquad\qquad + x_4 - x_5 &= 4
\end{aligned}
$$

- The above is a system of 3 linear equations in 5 unknowns
- Such system is important in practical world.

Matrix Formulation:

$$
\underset{3\times 5}{\begin{bmatrix} 1 & -7 & 2 & -5 & 8 \\ 0 & 1 & -3 & 3 & 1 \\ 1 & 0 & 0 & 1 & -1 \end{bmatrix}}
\underset{5\times 1}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}}
=
\underset{3\times 1}{\begin{bmatrix} 10 \\ -5 \\ 4 \end{bmatrix}}
$$

Vector Formulation :

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -7 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -5 \\ 3 \\ 1 \end{bmatrix}$$

$$+ x_5 \begin{bmatrix} 8 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 10 \\ -5 \\ 4 \end{bmatrix}$$

System of Linear Equations :

- A system of equations of the form :

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

where a_{ij} and b_i are <u>scalars</u> and the x_j are unknown variables is called a system of <u>m linear</u> <u>equations in n unknowns.</u>

- Any ordered n tuple $(s_1, s_2, \ldots, s_n)$ of scalars which satisfies all the m equations is called a solution of the system.

- The set of all solutions is called the "solution set" of the system.

- A system of Linear Equation has either (1) No solution
  OR (2) Exactly one solution
  OR (3) infinitely many solutions

- A system of Linear Equation is said to be consistent if it has either one solution or infinitely many solutions

- A system is called inconsistent if it has no solution

Matrix Formulation :

A system of Linear Equation can be compactly expressed in matrix notation as :

$$\boxed{Ax = b}$$

where $A = [a_{ij}]_{m \times n}$ is called the coefficient matrix and

$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$ and $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}_{m \times 1}$

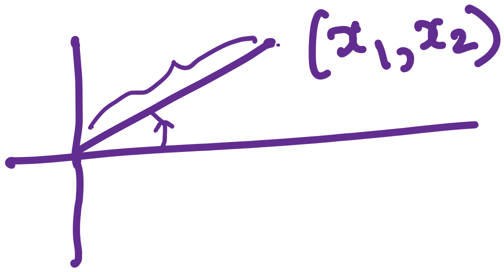
are vectors.

- A vector is an ordered k-tuple of scalars where k is any positive integer

Vectors are denoted as $\underbrace{(x_1, \ldots, x_k) \text{ or } [x_1, \ldots, x_k]}_{\text{(Row Vector)}}$

OR $\underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \text{ or } \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}}_{\text{(Column Vector)}}$

Ex: In two dimensional plane, vectors are ordered pairs.



—x———x———x———x———x

Vector Formulation:

A system of Linear Equations can also be written in a vector form:

$$x_1 V_1 + x_2 V_2 + \cdots + x_n V_n = b$$

where x_i's are the scalar unknowns and V_i's are the column vectors formed from the coefficients of the system of Linear equations.

Note (Explanation):

The system of equation $A_{m\times n} X_{n\times 1} = b_{m\times 1}$

Can be written as

$$x_1 V_1 + x_2 V_2 + \cdots + x_n V_n = b$$

Where $V_1, V_2, \ldots, V_n$ are the n columns of A ie. $A = [V_1 \; V_2 \cdots V_n]$

because

$$\boxed{AX} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}_{m\times 1}$$

$$= \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix}$$

$$= x_1 \underbrace{\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}}_{V_1} + x_2 \underbrace{\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}}_{V_2} + \cdots\cdots + x_n \underbrace{\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{V_n}$$

$$= \boxed{x_1 V_1 + x_2 V_2 + \cdots\cdots + x_n V_n}$$

(Please go to the next page)

Solving System of Linear Equations:

- Small systems of Linear Equations (with two or three variables) can be solved by a method of "elimination" or method of "substitution".

Our Goal: To obtain a more systematic strategy (i.e. an "algorithm") to solve a system.

Note: In this process, the variables play no role.
All the calculations are done with the coefficients and RHS scalars (R.H.S. := Right hand side)
Thus we should directly work with matrices and develope a matrix algorithm.
It has several applications.

Elementary Row Operations:

- Given any $m \times n$ matrix A, we define three elementary row operations:

(1) Multiplication of one row of A by a non-zero scalar c (Scaling)

(2) Replacement of one row of A by the sum of the row and a scalar multiple of a different row (Replacement).

[3) Interchange of two rows of A (Interchange)

Thus by applying an elementary row operation e to A, we get a new matrix which will be denoted by $e(A)$.

Note: To each elementary row operation e, there corresponds an elementary row operation e_1 of the same type such that $e_1(e(A)) = A$

Thus the process is reversible.

Example: Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{array}{l} \to R_1 \\ \to R_2 \\ \to R_3 \end{array}$

Scaling: $A \xrightarrow{R_1 \to 2R_1} \begin{bmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

The reverse operation:

$\xrightarrow{R_1 \to \frac{1}{2}R_1} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = A$

Interchange

$A \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}$

The reverse Operation:

$\xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = A$

Replacement:

$A \xrightarrow{R_3 \to R_3 + (-5)R_1} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -2 & -6 \end{bmatrix}$

The reverse operation :

$R_3 \to R_3 + 5R_1$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = A$$

MTH 100 : Lecture 2

Two Special type of Matrices :

Ex:

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 6} \longrightarrow \text{Echelon form}$$

Ex:

$$\begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 6} \longrightarrow \text{Reduced Row Echelon form (RREF)}$$

Note: The word Echelon came from the Latin word $\boxed{\text{Scala}}$ which means Ladder.

It came to French as "échelle" and then to English which now means Level or Step.

Definition:
An $m \times n$ matrix is said to be in <u>echelon form</u> if:
(1) All nonzero rows are above all zero rows.
(2) Each leading entry of a row (the first nonzero entry in a row) is to the right of the leading entry of the row above it.
(3) All entries in a column below a leading entry are zero.

<u>Note:</u> (3) is a consequence of (2). Still we have written it explicitly here in the interest of clarity.

Definition:
A $m \times n$ matrix is said to be in <u>Reduced Row Echelon form (RREF)</u> if
(1) All non zero rows are above all zero rows.
(2) Each leading (i.e. the first nonzero entry)

entry of a row is to the right of the leading entry of a row above it.

(3) The leading entry (i.e. the first nonzero entry) in each non zero row is 1.

(4) Each column which contains such a leading entry (necessarily 1) has all its other entries as zero

Note: RREF matrix is in Echelon form and has two further requirements (3) and (4).

Example: An Example of Row-Reduction (Gauss-Jordan Elimination)

$$A = \begin{bmatrix} 0 & 5 & 10 & 8 \\ 1 & 2 & 6 & 7 \\ 2 & 4 & 12 & 6 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 & 6 & 7 \\ 0 & 5 & 10 & 8 \\ 2 & 4 & 12 & 6 \end{bmatrix}$$

$R_3 \to R_3 + (-2)R_1$

(Echelon form) (Forward Phase)

$$\begin{bmatrix} 1 & 2 & 6 & 7 \\ 0 & 5 & 10 & 8 \\ 0 & 0 & 0 & -8 \end{bmatrix}$$

$$\xrightarrow[R_3 \to -\frac{1}{8} R_3]{R_2 \to \frac{1}{5} R_2} \begin{bmatrix} 1 & 2 & 6 & 7 \\ 0 & 1 & 2 & \frac{8}{5} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_1 \to R_1 + (-7) R_3]{R_2 \to R_2 + \left(-\frac{8}{5}\right) R_3} \begin{bmatrix} 1 & 2 & 6 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 + (-2) R_2} \begin{bmatrix} \boxed{1} & 0 & 2 & 0 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix} \to \text{RREF matrix}$$

(Backward Phase)

(Pivot columns)

Ex: Find the RREF form of

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 1 \end{bmatrix}$$

1

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 6R_1]{R_2 \to R_2 - 4R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -9 \\ 0 & -5 & -10 & -15 \end{bmatrix}$$

$$\xrightarrow{R_2 \to \left(-\frac{1}{3}\right) R_2} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & -5 & -10 & -15 \end{bmatrix} \xrightarrow{R_3 \to R_3 + 5R_2} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 - 2R_2} \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(RREF matrix)

(2)

$$\begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 1 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 5R_1]{R_2 \to R_2 - 3R_1} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -34 \end{bmatrix}$$

$$\xrightarrow{R_2 \to \left(-\frac{1}{4}\right) R_2} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & -8 & -16 & -34 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 + 8R_2} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & -10 \end{bmatrix}$$

$$\xrightarrow{R_3 \to \left(-\frac{1}{10}\right) R_3} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_2 \to R_2 - 3R_2]{R_1 \to R_1 - 7R_3} \begin{bmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 - 3R_2} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(RREF matrix)

Row Reduction Algorithm (Gauss-Jordan elimination):

- The input is an $m \times n$ matrix.
- We carry out elementary row operations on the input matrix.
- **Pivot Position:** A position corresponding to a leading position in an echelon form (a leading 1 in the RREF of the matrix). Its column is called a $\boxed{\text{Pivot column.}}$

- At the start move all zero rows to the bottom using interchange operations.

- **Step 1:**
  Start with Left most nonzero column. It will be the pivot column.

- **Step 2**
  Using interchange operation make the top element of the pivot column nonzero. (This will be the pivot position)

- **Step 3**
  Use replacement operations to make all entries in the pivot column below the pivot position as 0's.

- Step 4

Cover (ignore) the row containing pivot position and all rows above it (if any).
Repeat steps 1 to 4 for all rows below until all the non zero rows have been processed.

Note: Steps 1 – 4 is called $\boxed{\text{Forward Phase}}$ which produces a matrix in echelon form.
(This portion of the algorithm is referred to as Gaussian Reduction or Gaussian elimination)

- Step 5:

Use scaling operation to make all the pivot elements 1.

- Step 6:

Starting with the right-most pivot, creat zeros in the entire column above it, by using replacement operations.
Repeat this step moving leftward and upward.

Steps 5–6 is called the Backward Phase of the algorithm which produces an RREF matrix.

Note: The algorithm will stop after a finite number of steps and we will get a RREF matrix.

Conclusion:

Definition: If A and B are $m \times n$ matrices, we say that B is row equivalent to A if B can be obtained from A by a finite sequence of row operations.

Proposition 1: Given any $m \times n$ matrix A, there exists an RREF matrix which is row equivalent to A.

Proof: The proof is given by the above algorithm.

(i.e. we have given a constructive proof, rather than a pure existence proof.)

Note: In the example done in class, all the matrices in the intermediate steps were row equivalent to the original matrix A. However we obtained an RREF matrix only at the final step.

Proposition 2:

Row Equivalence is an equivalence relation on the set $R^{m \times n}$ of $m \times n$ matrices with entries from the field $\mathbb{R}$ (set) of real numbers. (m, n are fixed)

Proof: Exercise.

Note: If '$\sim$' denotes the relation of row equivalence, then
- $A \sim A$ (scaling first row by 1) for every $m \times n$ matrix A,
- $A \sim B \Rightarrow B \sim A$ (reversing the operation)
- $A \sim B$, $B \sim C \Rightarrow$ Combining the operations $A \sim C$

So, row equivalence is an equivalence relation.

Remark:
- Later on we will also work with matrices with complex entries. The proposition will continue to hold with $\mathbb{R}$ replaced by $\mathbb{C}$.
- Every equivalence relation induces a partition of the underlying set.

(The parts of the partition are called equivalence classes.)

The equivalence classes are pairwise disjoint subsets whose union is the whole set.

Conversely given any partition of a set, there exists a corresponding equivalence relation. (Justify yourself!)

<u>Note:</u> Define $x \sim y$ if and only if x, y belong to the same set of the partition.

This is an equivalence relation (Show!).

<u>Remark:</u> The RREF matrix of any given matrix is <u>unique</u>.

i.e. A matrix can't be row equivalent to two distinct RREF matrices.

Alternatively, two distinct RREF matrices can't be row-equivalent to each other.

<u>Thus</u>, inside each equivalence class for this equivalence relation, there is a distinctive member i.e. the one and only RREF matrix in it.

- This fact can be used to determine whether two matrices are row-equivalent to each other.

MTH 100 : Lecture 3

Application to Determinants: (Will be taken up later)

We note the following :

- If A is an $n \times n$ matrix and B is an echelon form $n \times n$ matrix obtained from A by Gaussian reduction without applying any scaling operation,

then $\det(A) = (-1)^k \det B = (-1)^k b_{11} b_{22} \cdots b_{nn}$

where k is the number of interchange operation.

- This is the preferred algorithm to calculate the determinant. That is why in software for matrix calculations, the two phases of the Row Reduction algorithm are carried out separately and we can obtain the determinant on the way.

Back to System of Linear Equations:

- Consider a system of Linear Equation in matrix form : $Ax = b$
- If $b = \bar{0}$, the system is called Homogeneous.
  A homogeneous system always has the trivial solution consisting of all zeros.
- If $b \neq \bar{0}$, the system is called non-homogeneous. A non-homogeneous system may or may not have any solution.
- A system which has atleast one solution is called consistent.
  Otherwise it is said to be inconsistent.

- Now for System of Linear Equations, we will directly work with matrices.
- For Homogeneous System, we will work with coefficient matrix A.
- For Non-homogeneous System we will work with the Augmented matrix $[A : b]$.
  It is obtained by putting a column

corresponding to b as an additional column (the $(n+1)$st column).

Observation :

If we obtain a row equivalent matrix to either the coefficient matrix (in the case of homogeneous system) or the augmented matrix (in the non-homogeneous case), then the solution sets of the two corresponding systems are same. In this case we say the systems are equivalent.

Homogeneous System :

Suppose that we have row-reduced the coefficient matrix A to an RREF matrix R.

- The leading variables in each nonzero row of R correspond to pivot columns. These are called Basic Variables. Remaining variables (if any) are called Free Variables.

- If we write the matrix equation $Rx = \bar{0}$ as a system of Linear equations, we can obtain the general solution of the system. (The system $Rx = \bar{0}$ is equivalent to original system $Ax = \bar{0}$)

<u>The general solution is best expressed in Vector terms.</u>

(To the next page)

Example: Homogeneous System:

(1)
$$\left.\begin{aligned} x_1 + 2x_2 - 3x_3 &= 0 \\ 2x_1 + 4x_2 - 2x_3 &= 0 \\ 3x_1 + 6x_2 - 4x_3 &= 0 \end{aligned}\right\}$$

The coefficient matrix $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -2 \\ 3 & 6 & -4 \end{bmatrix}$

Let us row reduce A

$$\begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -2 \\ 3 & 6 & -4 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 3R_1]{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 4 \\ 0 & 0 & 5 \end{bmatrix} \xrightarrow{R_2 \to \frac{1}{4}R_2} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 5 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 5R_2} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \to R_1 + 3R_2} \begin{bmatrix} \boxed{1} & 2 & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{bmatrix} = R \text{ (say)}$$

(RREF matrix)

Pivot columns

The basic variables are x_1 and x_3

The free variable is x_2

So, the system becomes $R\,x = \bar{0}$ i.e. $R\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$$\Rightarrow \left.\begin{array}{r} x_1 + 2x_2 = 0 \\ x_3 = 0 \end{array}\right\}$$

Let us express basic variables in terms of free variables and introduce a dummy equation:

$$x_1 = -2x_2$$
$$x_2 = x_2 \rightarrow \text{(dummy equation)}$$
$$x_3 = 0 = 0.x_2$$

We can write the solution in vector form:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2x_2 \\ x_2 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$ where x_2 acts as a parameter.

- There are <u>infinitely many solutions</u>.
- The solution set can be concisely described as:

$$S = \left\{ t\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} : t \in \mathbb{R} \right\} = \left\{ t\bar{u} : \bar{u} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, t \in \mathbb{R} \right\}$$

<u>Check</u>: $A\bar{u} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -2 \\ 3 & 6 & -4 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Ex ②: $AX = \bar{0}$ where $A = \begin{bmatrix} \boxed{1} & 2 & 0 & 3 \\ 0 & 0 & \boxed{1} & 1 \end{bmatrix}$

We note that A is a RREF matrix

$\left.\begin{array}{l}\text{The Basic Variables are } x_1 \text{ and } x_3 \\ \text{The free variables are } x_2 \text{ and } x_4\end{array}\right\}$

The system reduces to:

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_4 &= 0 \\ x_3 + x_4 &= 0 \end{aligned}\right\}$$

Expressing Basic Variables in terms of free variables:

$$\begin{aligned} x_1 &= -2x_2 - 3x_4 \\ x_2 &= x_2 + 0 \text{ (dummy equation)} \\ x_3 &= 0 - x_4 \\ x_4 &= 0 + x_4 \text{ (dummy equation)} \end{aligned}$$

So, in the vector form

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

The solution set is:

$$S = \left\{ t \underbrace{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{\bar{u}} + s \underbrace{\begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}}_{\bar{w}} : t, s \in \mathbb{R} \right\}$$

$= \{ t\bar{u} + s\bar{w} : t, s \in \mathbb{R} \}$

- There are infinitely many solutions.

- There are infinitely many solutions.

Remark :

It is possible to obtain the solution set in a different form :

e.g. : $S_1 = \{ r\bar{u}_1 + s\bar{u}_2 + t\bar{u}_3 : r, s, t \in \mathbb{R} \}$

where $\bar{u}_1 = \begin{bmatrix} -8 \\ 1 \\ -2 \\ 2 \end{bmatrix}$, $\bar{u}_2 = \begin{bmatrix} -11 \\ 1 \\ -3 \\ 3 \end{bmatrix}$, $\bar{u}_3 = \begin{bmatrix} 5 \\ 2 \\ 3 \\ -3 \end{bmatrix}$

We can verify (later) that $S = S_1$, regarded as a set of vectors.

So, to avoid such difficulties (which arises because the solution set is infinite), we will always obtain the solution via the RREF matrix.

The reasons are :

(1) The RREF matrix is unique.

(2) The number of vectors obtained

on RHS via the RREF is least possible. (i.e. any other method can not provide fewer vectors.)

— x ——— x ——— x ——— x

<u>Recall</u> Homogeneous System $Ax = \bar{0}$,
A is the coefficient matrix and assume that A is row reduced to an RREF matrix R by elementary row operations.

Then the system $Rx = \bar{0}$ is equivalent to the original system $Ax = \bar{0}$.

We can solve the system $Rx = \bar{0}$ by using Basic Variables and Free Variables. and <u>the solution is best expressed in vector form</u>:

<u>Observations</u>:

(1) If the number of nonzero rows r of R is less than the number of variables n, the system has a non-trivial solution

- We express the Basic Variables in terms of free variables

- Free variables behave like parameters; i.e. we can choose any values for them and each such choice gives a solution. This way we get infinitely many solutions.

(2) Special case of 1:

If A is an $m \times n$ matrix with $m < n$, then the homogeneous system $Ax = 0$ must have a nontrivial solution. (actually infinitely many solutions because in this case, there have to be free variables)

(3) If the number of non-zero rows of R is equal to the number of variables (i.e. the number of columns), then there are no free variables and the system has a unique solution (only the trivial solution of all zeros).

Proposition 3: If A is a square matrix, then A is row equivalent to the identity matrix if and only if the homogeneous system $Ax = 0$ has only the trivial solution.

Proof: Exercise.

Summary of Homogeneous Systems:

(1) System is always consistent.

(2) If the system has a unique solution then it is the trivial solution of all zeros: In this case the RREF is either the $n \times n$ identity matrix I_n or has I_n as its upper portion with only zero rows below.

(3) Else, the system contains free variables and has infinitely many solutions (one of which is the trivial solution): This happens when number of non-zero rows in the RREF is less than the number of variables.

(4) If number of equations is less than the number of variables,

then the system has infinitely many solutions. This is a special case of 3.

MTH100 : Lecture 4

Example on Proposition 3:

Consider the system of equations $AX = 0$

where $A = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 - R_1} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 5 & 6 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 5R_2} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_2 \to R_2 - R_3]{R_1 \to R_1 + 2R_3} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 + R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

(RREF matrix)

So, the corresponding equivalent system of equation is:

$$\left.\begin{aligned} x_1 &= 0 \\ x_2 &= 0 \\ x_3 &= 0 \end{aligned}\right\}$$

So, the system has a unique solution

$X = \bar{0}$ i.e. $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Non-homogeneous System:

For a Non-homogeneous system $Ax=b$, we work with the augmented matrix $[A:b]_{m\times(n+1)}$ and reduce it to an RREF matrix, say R.

Proposition 4 (Existence and Nature of solutions):

The system $Ax=\bar{b}$ is consistent if and only if the rightmost column of R is not a pivot column.

i.e. there is no row of the form $[0,0,\cdots,0,p]$ with p nonzero.

If the system is consistent, then it has either (1) a unique solution if there are no free variables

or (2) infinitely many solutions when there is atleast one free variable.

Proof: Exercise.

Examples of Non-homogeneous Systems

① Consider $Ax = b$ where

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix} \quad \text{and} \quad \bar{b} = \begin{bmatrix} 7 \\ 9 \\ 30 \end{bmatrix}$$

The Augmented matrix

$$[A : b] = \left[\begin{array}{ccc:c} 1 & 0 & 2 & 7 \\ 2 & -1 & 3 & 9 \\ 4 & 1 & 8 & 30 \end{array}\right] \xrightarrow[R_3 \to R_3 - 4R_1]{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc:c} 1 & 0 & 2 & 7 \\ 0 & -1 & -1 & -5 \\ 0 & 1 & 0 & 2 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 + R_2} \left[\begin{array}{ccc:c} 1 & 0 & 2 & 7 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & -1 & -3 \end{array}\right] \xrightarrow[R_3 \to (-1)R_3]{R_2 \to (-1)R_2} \left[\begin{array}{ccc:c} 1 & 0 & 2 & 7 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$$\xrightarrow[R_2 \to R_2 - R_3]{R_1 \to R_1 - 2R_3} \left[\begin{array}{ccc:c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$\Rightarrow$ The corresponding equivalent system of equation is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

So, the solution of the system is :

$$\left.\begin{aligned} x_1 &= 1 \\ x_2 &= 2 \\ x_3 &= 3 \end{aligned}\right\} \quad \text{i.e. } X = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

- So, we have a unique solution (There are no free variables)

<u>Check</u>: $AX = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 9 \\ 30 \end{bmatrix} = \bar{b}$ (as expected)

(To the next Page)

② $A x = \bar{b}$ where $A = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 8 & 16 \\ 8 & 20 & 40 \end{bmatrix}$

and $\bar{b} = \begin{bmatrix} 3 \\ 11 \\ 28 \end{bmatrix}$

$$[A:b] = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 3 & 8 & 16 & 11 \\ 8 & 20 & 40 & 28 \end{array}\right] \xrightarrow[R_3 \to R_3 - 8R_1]{R_2 \to R_2 - 3R_1} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 0 & 2 & 4 & 2 \\ 0 & 4 & 8 & 4 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 0 & 2 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_2 \to \frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_1 \to R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] = R \text{ (RREF matrix)}$$

(Pivot columns)

So, x_3 is a free variable.

The corresponding system is: $\left.\begin{aligned} x_1 &= 1 \\ x_2 + 2x_3 &= 1 \end{aligned}\right\}$

$$\Rightarrow \left.\begin{aligned} x_1 &= 1 + 0 \cdot x_3 \\ x_2 &= 1 - 2x_3 \\ x_3 &= 0 + x_3 \end{aligned}\right\} \longrightarrow \text{(dummy equation)}$$

So, $X = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\bar{u}} + x_3 \underbrace{\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}}_{\bar{w}}$ is a solution of $AX = \bar{b}$

<u>Check</u>: $A\bar{u} = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 8 & 16 \\ 8 & 20 & 40 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \\ 28 \end{bmatrix} = \bar{b}$

$A\bar{w} = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 8 & 16 \\ 8 & 20 & 40 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \bar{0}$

Thus the solution set is $S = \{ \bar{u} + t\bar{w} : t \in \mathbb{R} \}$ where $\bar{u}$ is a solution of the non-homogeneous system and $\bar{w}$ is a solution of the associated homogeneous system $AX = \bar{0}$

- Since x_3 (or t) acts as a parameter, we get <u>infinitely many solutions.</u>

③ Consider $A x = \bar{b}$ when

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 8 & 16 \\ 8 & 20 & 40 \end{bmatrix} \text{ and } \bar{b} = \begin{bmatrix} 4 \\ 11 \\ 28 \end{bmatrix}$$

The Augmented matrix

$$[A : b] = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 4 \\ 3 & 8 & 16 & 11 \\ 8 & 20 & 40 & 28 \end{array}\right] \xrightarrow[R_3 \to R_3 - 8R_1]{R_2 \to R_2 - 3R_1} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 4 \\ 0 & 2 & 4 & -1 \\ 0 & 4 & 8 & -4 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 4 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 \end{array}\right] \xrightarrow{R_2 \to \frac{1}{2} R_2} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 4 \\ 0 & 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 & -2 \end{array}\right]$$

$$\xrightarrow{R_1 \to R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 & -2 \end{array}\right] \xrightarrow{R_3 \to -\frac{1}{2} R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_1 \to R_1 - 5R_3]{R_2 \to R_2 + \frac{1}{2} R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

So, the last row is of the form
$[0, 0, \cdots, 0, p]$ when $p \neq 0$.

Thus the system is inconsistent and there is no solution. (By proposition ④).

- The corresponding system becomes

$$\left.\begin{aligned} x_1 &= 0 \\ x_2 + 2x_3 &= 0 \\ 0 &= 1 \longrightarrow \text{not true} \end{aligned}\right\}$$

So, the system is inconsistent.

Vector Interpretation of Solutions:

Let $Ax=b$ be a non-homogeneous system and let $Ax=0$ be its associated homogeneous system.
Assume that the non-homogeneous system is consistent so that it has atleast one solution u. By necessity $u \neq 0$.
Now the relationship between solutions

of the two systems is given in the following observation (R$):

- If a vector u is a given solution of $Ax=b$, then another vector is a solution of $Ax=b$ if and only if it is of the form $u+v$ where v is a solution of the associated homogeneous system.

- In case $Ax=0$ has only trivial solution (i.e. $v=0$), then there is a unique solution u. (Otherwise we have infinitely many solutions.)

Proof: (H.W.)

$\boxed{\text{Lecture : 5}}$

Observation (R S):

- If a vector u is a given solution of $Ax=\bar{b}$, then another vector is a solution of $Ax=\bar{b}$ if and only if it is of the form $u+v$ where v is a solution of the associated homogeneous system.

- In case $Ax=0$ has only trivial solution (i.e. $v=0$), then there is a unique solution u. (otherwise we have infinitely many solutions.)

Proof: $\Rightarrow$: given : Let u be a solution of $Ax=\bar{b}$ and let w be another solution of $Ax=\bar{b}$

want to show : w is of the form $u+v$ where v is a solution of $Ax=\bar{0}$

Let $v=w-u$, then $w=u+v$

and $Av = A(w-u) = Aw - Au = \bar{b}-\bar{b} = \bar{0}$

$\boxed{\text{So, } v \text{ is a solution of the associated homogeneous equation and } w=u+v}$

$\Leftarrow$:

Given: Let u be a solution of $AX=\bar{b}$ and let v be a solution of $AX=\bar{0}$

Want to show: $w=u+v$ is a solution of $AX=\bar{b}$

We have $Aw=A(u+v) = Au + Av = \bar{b}+\bar{0}=\bar{b}$

So, w is a solution of $AX=\bar{b}$

(Q.E.D.)

Example:

Consider the system of equation

$$x_1+x_2+x_3=1$$
$$2x_1-x_2+x_3=2$$

The Augmented Matrix $[A:b]$:

$$\left[\begin{array}{ccc:c} 1 & 1 & 1 & 1 \\ 2 & -1 & 1 & 2 \end{array}\right] \xrightarrow[R_2 \to R_2-2R_1]{} \left[\begin{array}{ccc:c} 1 & 1 & 1 & 1 \\ 0 & -3 & -1 & 0 \end{array}\right]$$

(columns 1–3: A; last column: $\bar{b}$)

$$\xrightarrow[R_2 \to \left(-\frac{1}{3}\right)R_2]{} \left[\begin{array}{ccc:c} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{1}{3} & 0 \end{array}\right] \xrightarrow[R_1 \to R_1-R_2]{} \underbrace{\left[\begin{array}{ccc:c} 1 & 0 & \frac{2}{3} & 1 \\ 0 & 1 & \frac{1}{3} & 0 \end{array}\right]}_{\text{(RREF matrix)}}$$

The corresponding system is:

$$\left.\begin{aligned} x_1 + \tfrac{2}{3}x_3 &= 1 \\ x_2 + \tfrac{1}{3}x_3 &= 0 \end{aligned}\right\} \Rightarrow \begin{aligned} x_1 &= 1 - \tfrac{2}{3}x_3 \\ x_2 &= -\tfrac{1}{3}x_3 \\ x_3 &= x_3 \end{aligned}$$

(dummy equation)

So, the solution is of the form:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{u} + t\underbrace{\begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix}}_{v} = u + tv$$

where $t \in \mathbb{R}$ (scalar)

where u is a solution of the given non-homogeneous system and v is a solution of the associated homogeneous system.

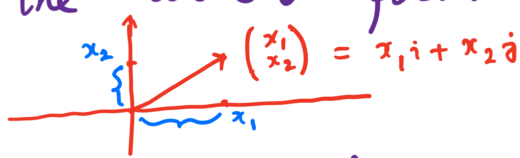
Check: $Au = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \bar{b}$

$$Av = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}\begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} - \frac{1}{3} + 1 \\ -\frac{4}{3} + \frac{1}{3} + 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$:

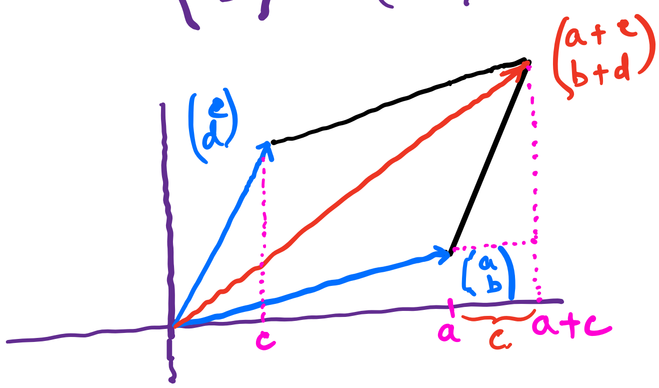
- A vector in $\mathbb{R}^2$ is an ordered pair of real numbers (written either as $\underbrace{\text{column}}$ or $\overbrace{\text{row}}$) (e.g. $\begin{pmatrix} a \\ b \end{pmatrix}$ or (a, b))
(In case of $\mathbb{R}^3$, it is a 3-tuple)

It gives us the geometric interpretation of the vector as the arrow pointing from $(0,0)$ to $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

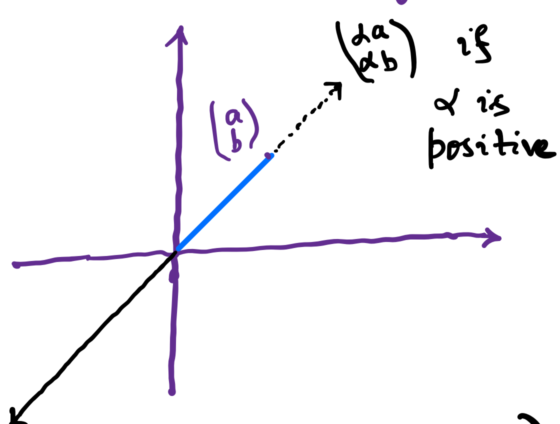


- We can add two vectors (coordinate wise) and multiply any vector by a real number and this is consistent with the geometric interpretation.

Thus $$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix} \quad \text{and} \quad \alpha \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \alpha a \\ \alpha b \end{pmatrix}$$
for any $\alpha \in \mathbb{R}$



(Addition of Vectors)



(Scalar Multiplication of a vector)

Note that addition and scalar multiplication satisfies the following properties.

- $\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix} = \begin{pmatrix} c+a \\ d+b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$

(commutative property)

Similarly

- $\left[\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} \right] + \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \left[\begin{pmatrix} c \\ d \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \right]$

(Associative property)

- $\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$
- $\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} -a \\ -b \end{pmatrix} = \begin{pmatrix} -a \\ -b \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Also $1 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1.a \\ 1.b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$

$\alpha \left(\beta \begin{pmatrix} a \\ b \end{pmatrix} \right) = (\alpha \beta) \begin{pmatrix} a \\ b \end{pmatrix}$ for any $\alpha, \beta \in \mathbb{R}$

$$(\alpha+\beta)\begin{pmatrix} a \\ b \end{pmatrix} = \alpha\begin{pmatrix} a \\ b \end{pmatrix} + \beta\begin{pmatrix} a \\ b \end{pmatrix}$$

for any $\alpha, \beta \in \mathbb{R}$

$$\alpha\left[\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}\right] = \alpha\begin{pmatrix} a \\ b \end{pmatrix} + \alpha\begin{pmatrix} c \\ d \end{pmatrix}$$

for any $\alpha \in \mathbb{R}$

- Note that all these properties follow from the corresponding properties of real numbers
  (Please verify them at home)

- Similarly we can define vector addition and scalar multiplication in $\mathbb{R}^3$ and also in $\mathbb{R}^n$ (for any positive integer n)

- <u>Geometrical interpretation of solutions:</u>

- In case we are looking with 2-tuples or 3 tuples, we can have a geometrical interpretation.
- Each vector correspond to a point either <u>in plane</u> (2-space) or <u>in space</u> (3-space)

- Then the solution of a homogeneous system is either the origin only or all the points on a line or, a plane through the origin.

- If a non-homogeneous system has even a single solution (i.e. a point in plane or space) then its entire solution set consists of either only that point or the line or plane through that point which is parallel to the solution of the associated homogeneous system.

From geometry:

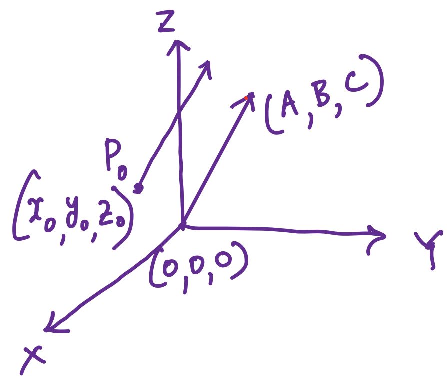
- Equation of the line through $P_0(x_0, y_0, z_0)$ parallel to a given vector

$$v = Ai + Bj + Ck$$

(i.e. the line segment from $(0,0,0)$ to (A,B,C))

is given by

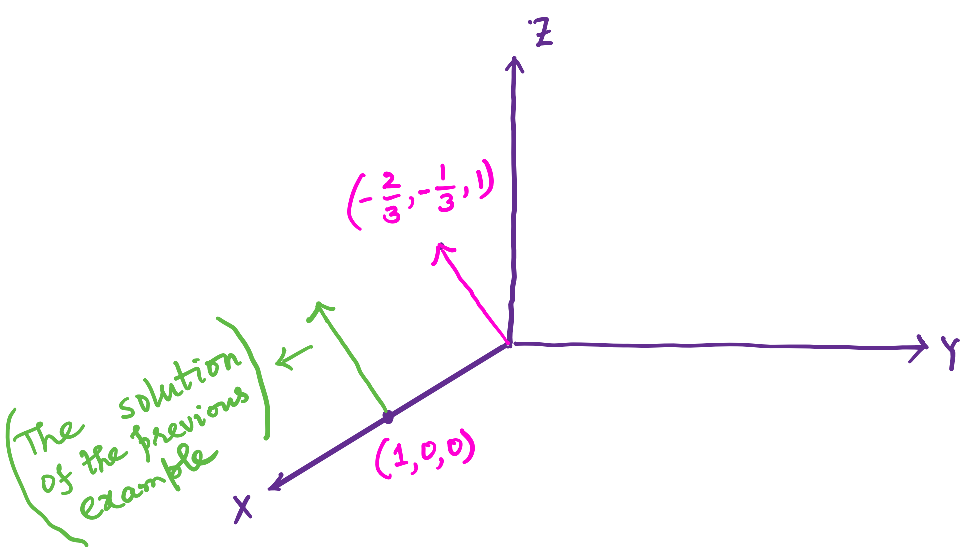
$$\left.\begin{aligned} x &= x_0 + tA \\ y &= y_0 + tB \\ z &= z_0 + tC \end{aligned}\right\} \text{ where } t \in \mathbb{R}$$



So, $(x,y,z) = (x_0,y_0,z_0) + t(A,B,C)$

$$= u + tv$$

- The solution we have obtained corresponds to the geometrical equation of the line through $(1,0,0)$ which is parallel to the vector determined by $\left(-\frac{2}{3}, -\frac{1}{3}, 1\right)$



Summary For Non-homogeneous System:

Associated Homogeneous System $Ax=\bar{0}$		Non-homogeneous System $Ax=\bar{b}$
Case 1: Unique Solution (trivial) $\updownarrow$ (No free Variable)	$\longrightarrow$	Inconsistent or Unique solution (non-zero)
Case 2: Infinitely many solutions $\updownarrow$ Atleast one free Variables	$\longrightarrow$	Inconsistent or Infinitely many solutions.

Note: $Ax=\bar{b}$ can be inconsistent in both cases. However, if it is consistent, nature of solutions corresponds to nature of solution of $Ax=\bar{0}$

MTH 100: Lecture 6

Invertible Matrices:

An $m\times m$ (square) matrix A is called invertible if there exists another square matrix B such that $\boxed{BA = AB = I_m}$ ($m\times m$ identity matrix)

B is called an inverse of A.

- **Another terminology:** Invertible matrices are also called Non singular.

Matrices which are not invertible are called Singular.

Observation ①: The inverse of A if it exists is unique. (Notation: A^{-1})

Let $AB = BA = I$, $AC = CA = I$ (i.e. B & C are two inverses of A). Then

$BAC = (BA)C = I.C = C$
$BAC = B(AC) = B.I = B$
$\Rightarrow B = C$. So, the inverse is unique.

Observation ②: If A is invertible, then so is A^{-1} and $(A^{-1})^{-1} = A$ (Since $A^{-1}A = AA^{-1} = I$)

Observation ③: If A and B are invertible, so is AB and $(AB)^{-1} = B^{-1}A^{-1}$

because $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A(I)A^{-1} = (AI)A^{-1} = AA^{-1} = I$

and $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}(I)B = (B^{-1}I)B = B^{-1}B = I$

Observation ④ (Generalization of ③):

The product of invertible matrices is invertible and the inverse of the product is the product of the inverses taken in reverse order.

In other words, if $A_1, A_2, \ldots A_n$ $(n \geq 2)$ are invertible matrices, then $C = A_1 A_2 \cdots A_n$ is an invertible matrix and $C^{-1} = A_n^{-1} \ldots A_2^{-1} A_1^{-1}$

Elementary Matrices:

- An $m \times m$ (square) matrix is said to be an <u>elementary matrix</u> if it is obtained <u>from the $m \times m$ identity matrix I_m</u> by an <u>elementary row operation.</u>

Proposition ⑤: If e is an elementary row operation and E is the $m \times m$ elementary matrix $e(I_m)$, then for every $m \times n$ matrix A

$$\boxed{e(A) = EA}$$

Thus applying an elementary row operation is the same as left multiplication by the corresponding elementary matrix.

Proof: Exercise

Note: The three types of elementary row operation have to be treated separately.

Ex:

(This is not a proof) Let $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

Let e be the replacement operation

$$e : R_3 \longrightarrow R_3 + 2R_1$$

Then $e(A) = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 9 & 12 & 15 \end{bmatrix}$

$$E = e(I_3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

$$EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 9 & 12 & 15 \end{bmatrix} = e(A)$$

- This is an illustration of proposition 5. (However, this is not a proof)
- You can also try with other types of operation.

Proposition ⑥: Every elementary matrix is invertible.

Proof: Let E be an elementary matrix. Let e be the corresponding row operation.

So, $e(I) = E$

- We know that there is another row operation of the same type (we call it f) that reverses the action of e.
- Let F be the elementary matrix corresponding to f i.e. $f(I) = F$

$$\text{Now, } FE = (FE)I = F(EI) = F(e(I)) \quad (\text{By proposition ⑤})$$
$$= f(e(I)) \quad (\text{By proposition ⑤})$$
$$= I \quad \left(\text{since } f \text{ is the reverse operation of } e\right)$$

$$EF = (EF)I = E(FI) = E(f(I)) \quad (\text{By proposition ⑤})$$
$$= e(f(I)) \quad (\text{By proposition ⑤})$$
$$= \left(I \quad \text{since } e \text{ is the reverse operation of } f\right)$$

- So, $EF = FE = I$

Thus E is invertible and $E^{-1} = F$

Note: The inverse of an elementary matrix is also an elementary matrix of the same type.

Ex: An example of finding the inverse of a matrix by row reduction:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$$

: We will take the enlarged matrix $[A : I]$

$$[A:I] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3 \to R_3 - 4R_1]{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right] \xrightarrow[R_3 \to (-1)R_3]{R_2 \to (-1)R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

$$\xrightarrow[R_1 \to R_1 - 2R_3]{R_2 \to R_2 - R_3} \left[\underbrace{\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}}_{I} \,\middle|\, \underbrace{\begin{array}{ccc} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{array}}_{\text{This will be } A^{-1}} \right]$$

Check:

$$\begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix} \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note: This method is preferable to the adjoint/ Determinant formula which requires approximately $(n!)$ calculations.
Gauss-Jordan elimination requires approximately $\left(\frac{3}{2}n^3\right)$ operations.

Theorem ①: The following are equivalent for an $m \times m$ square matrix A.

(a) A is invertible

(b) A is row equivalent to the identity matrix.

(c) The homogeneous system $Ax = 0$ has only the trivial solution.

(d) The system of equation $Ax = b$ has atleast one solution for every $b \in \mathbb{R}^m$

Proof:

MTH 100: Lecture 7

Theorem ①: The following are equivalent for a $n \times n$ square matrix A:

(a) A is invertible

(b) A is row equivalent to the identity matrix.

(c) The homogeneous system $AX = \bar{0}$ has only the trivial solution.

(d) The system of equation $AX = \bar{b}$ has atleast one solution for every $\bar{b} \in \mathbb{R}^n$

Proof: First we will prove (a) $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ (a).

$$\left(\begin{tikzcd} & (a) \arrow[dr, Rightarrow] & \\ (b) \arrow[ur, Rightarrow] & & (c) \arrow[ll, Rightarrow] \end{tikzcd}\right)$$

((a) $\Leftrightarrow$ (d) : Later)

- (a) $\Rightarrow$ (c):

Given: A is invertible

To show: $AX = \bar{0}$ has only the trivial solution.

- Let u be a solution of $AX = \bar{0}$

$$\Rightarrow Au = \bar{0}$$

$$\Rightarrow A^{-1}(Au) = A^{-1}(\bar{0}) \quad \left(\text{Since } A \text{ invertible } A^{-1} \text{ exists}\right)$$

$$\Rightarrow A^{-1}(Au) = \bar{0}$$

$$\Rightarrow (A^{-1}A)u = \bar{0}$$

$$\Rightarrow I.u = \bar{0} \Rightarrow \boxed{u = \bar{0}}$$

So, $AX = \bar{0}$ has only the trivial solution.

(c) $\Rightarrow$ (b): Given: $A x = \bar{0}$ has only the trivial solution.

To show: A is row equivalent to the identity matrix.

- Let R be the RREF matrix of A.
  Then $R x = \bar{0}$ also has only the trivial solution.
  $\Rightarrow$ R has no free variable
  $\Rightarrow$ R has only basic variables
  $\Rightarrow$ R has leading entry as 1 in each row (There are m rows)
  $\Rightarrow$ R has exactly one 1 in each column (There are m columns
  $\Rightarrow$ R is I $\Rightarrow$ $\boxed{\text{A is row equivalent to I}}$

(b) $\Rightarrow$ (a): Given: A is row equivalent to the identity matrix I

To show: A is invertible

- A is row equivalent to I
  $\Rightarrow$ There are elementary row operations $e_1, e_2, \ldots, e_{p-1}, e_p$ such that
  $$e_p(e_{p-1} \cdots\cdots (e_2(e_1(A))) \cdots\cdots) = I$$

Let E_i be the elementary matrix corresponding to e_i for $i = 1, 2, \ldots, p$
(i.e. $E_i = e_i(I)$)

Then $E_p(E_{p-1} \cdots\cdots (E_2(E_1 A)) \cdots\cdots) = I$ (By proposition 5)

$$\Rightarrow (E_p E_{p-1} \cdots E_2 E_1) A = I$$

Let $B = E_p E_{p-1} \cdots E_2 E_1$

Then B is invertible (By observation 4 and Proposition 6)

and we have $BA = I$

Multiplying both sides by B^{-1} from the left, we obtain

$$B^{-1}(BA) = B^{-1}.I$$

$$\Rightarrow (B^{-1}B) A = B^{-1}$$

$$\Rightarrow I.A = B^{-1} \Rightarrow \boxed{A = B^{-1}}$$

So, A is the inverse of an invertible matrix

$\Rightarrow$ $\boxed{A \text{ is invertible}}$ (By Observation 2)

Calculation of Inverse matrix:

Corollary (1.1): An invertible matrix A is a product of elementary matrices.

(**Note:** Any sequence of row operations that reduces A to I also transforms I to A^{-1} (We are using Theorem 1 (b) here))

Proof: If A is invertible, then by Theorem 1(b), A is row equivalent to I.

So, there are some elementary row operations $e_1, e_2, \ldots, e_p$ such that

$$e_p(e_{p-1} \cdots (e_2(e_1(A))) \cdots) = I$$

Let $E_1, E_2, \ldots, E_p$ be the corresponding elementary matrices (i.e. $E_i = e_i(I)$),

then $E_p(E_{p-1} \cdots (E_2(E_1 A)) \cdots) = I$

$\Rightarrow (E_p E_{p-1} \cdots E_2 E_1) A = I$

$\Rightarrow A = (E_p E_{p-1} E_2 E_1)^{-1} I$

$\Rightarrow A = (E_1^{-1} E_2^{-1} \cdots E_{p-1}^{-1} E_p^{-1}) I$

$\Rightarrow \boxed{A = E_1^{-1} E_2^{-1} \cdots E_{p-1}^{-1} E_p^{-1}}$

So, A is a product of some elementary matrices

<u>Note:</u> We can say $A^{-1} = (E_1^{-1} E_2^{-1} \cdots E_p^{-1})^{-1}$

$\Rightarrow A^{-1} = (E_p^{-1})^{-1} \cdots (E_2^{-1})^{-1} (E_1^{-1})^{-1}$

$\Rightarrow A^{-1} = E_p \cdots E_2 E_1$

$\Rightarrow A^{-1} = (E_p \cdots E_2 E_1) I$

$\Rightarrow A^{-1} = E_p(\cdots E_2(E_1 I) \cdots)$

$\Rightarrow \boxed{A^{-1} = e_p(\cdots e_2(e_1(I)) \cdots)}$

Thus the same sequence of row operations that reduces A to I also reduces I to A^{-1}.

Method of obtaining A^{-1}:

Form the enlarged matrix $[A : I]$ and carry out elementary row operations till 'A' part becomes I. The final result has the form $[I : A^{-1}]$.

- Corollary (1.2): If A has a left inverse or a right inverse, then it has an inverse.

Note: (1) B is a left inverse of A if $BA = I$

(2) D is a right inverse of A if $AD = I$

MTH 100: Lecture 8

Corollary (1.2): If A has a left inverse or a right inverse then it has an inverse.

Note: ① B is a left inverse of A if $BA = I$
② D is a right inverse of A if $AD = I$

Proof: Case ①: Suppose A has a left inverse.

Then there exists a matrix B such that $BA = I$

Consider the homogeneous system $A\bar{x} = \bar{0}$

$$\Rightarrow B(A\bar{x}) = B.\bar{0}$$
$$\Rightarrow (BA)\bar{x} = \bar{0}$$
$$\Rightarrow I.\bar{x} = \bar{0} \Rightarrow \bar{x} = \bar{0}$$

So, $A\bar{x} = \bar{0}$ has only the trivial solution

By Theorem ①, A is invertible.

Furthermore, $BA = I$ (given)

(we have shown that A^{-1} exists)

$$\Rightarrow (BA)A^{-1} = I.A^{-1}$$
$$\Rightarrow B(AA^{-1}) = A^{-1}$$
$$\Rightarrow B.I = A^{-1}$$
$$\Rightarrow \boxed{B = A^{-1}}$$

Case 2: Suppose A has a right inverse.
Then there exists a matrix D such that

$$AD = I$$

So, A is a left inverse of D

Therefore by case 1, D is invertible

Now $(AD)D^{-1} = I.D^{-1}$

$$\Rightarrow A(DD^{-1}) = D^{-1}$$

$$\Rightarrow A.I = D^{-1}$$

$$\Rightarrow A = D^{-1}$$

Therefore A is the inverse of an invertible matrix D

So, A is invertible
and $A^{-1} = (D^{-1})^{-1}$

$$\Rightarrow \boxed{A^{-1} = D}$$

Corollary (1.3): Suppose a square matrix A can be factored as a product of square matrices i.e. $A = A_1 A_2 \cdots A_n$ with $n \geq 2$ (A_i's are all square matrices)

Then A is invertible if and only if each A_i is invertible.

Proof: '$\Leftarrow$':

If A_i's are all invertible then $A = A_1 A_2 \cdots A_n$ is also invertible and $A^{-1} = A_n^{-1} \cdots A_2^{-1} A_1^{-1}$ (By Observation 4)

'$\Rightarrow$': Given: A is invertible

To show: Each A_i is invertible.

First we will show that A_n is invertible

Consider the homogeneous system

$$A_n \bar{x} = \bar{0}$$

$$\Rightarrow (A_1 A_2 \cdots A_{n-1}) A_n \bar{x} \quad (A_1 A_2 \cdots A_{n-1}) \bar{0}$$

$$\Rightarrow (A_1 A_2 \cdots A_{n-1} A_n) \bar{x} = \bar{0}$$

$$\Rightarrow \quad A\bar{x} = \bar{0} \Rightarrow \bar{x} = \bar{0}$$ (By Theorem 1 since A is invertible)

So, the homogeneous system $A_n \overline{x} = \overline{0}$ has only the trivial solution.

Therefore by Theorem (1), A_n is invertible.

Now $A_1 A_2 \cdots A_{n-1} A_n = A$

$\Rightarrow (A_1 A_2 \cdots A_{n-1} A_n) A_n^{-1} = A A_n^{-1}$ (since A_n^{-1} exists)

$\Rightarrow (A_1 A_2 \cdots A_{n-1}) A_n A_n^{-1} = A A_n^{-1}$

$\Rightarrow (A_1 A_2 \cdots A_{n-1}) I = A A_n^{-1}$

$\Rightarrow A_1 A_2 \cdots A_{n-1} = A A_n^{-1}$

Let $B = A A_n^{-1}$

Then B is an invertible matrix and

$$B = A_1 A_2 \cdots A_{n-1}$$

Repeating the same argument, we conclude that A_{n-1} is invertible.

Continuing the same process we can show that each A_i is invertible.

Final Part of Theorem 1: (a) $\Leftrightarrow$ (d)

(a) $A_{m\times m}$ is invertible

(d) The non-homogeneous system $A\bar{x}=\bar{b}$ has atleast one solution for any choice of $\bar{b}\in\mathbb{R}^m$.

<u>(a) $\Rightarrow$ (d)</u> <u>Given</u>: A is invertible

<u>To show</u>: $A\bar{x}=\bar{b}$ has atleast one solution for any choice of $\bar{b}\in\mathbb{R}^m$.

Let $\bar{b}\in\mathbb{R}^m$ be any <u>arbitrary</u> but <u>fixed</u> vector.

Let $\bar{v}=A^{-1}\bar{b}$ (A^{-1} exists since A is invertible)

Since A^{-1} is a $m\times m$ matrix and $\bar{b}$ is a $m\times 1$ vector, $\bar{v}$ is a $m\times 1$ vector.

Now $A\bar{v} = A(A^{-1}\bar{b})$

$\Rightarrow A\bar{v} = (AA^{-1})\bar{b}$

$\Rightarrow A\bar{v} = I.\bar{b}$

$\Rightarrow A\bar{v} = \bar{b}$

So, $\bar{v}$ is a solution of the non-homogeneous system $A\bar{x}=\bar{b}$ as required.

(d) $\Rightarrow$ (a): Given: $A\bar{x} = \bar{b}$ has atleast one solution for any choice of $\bar{b} \in \mathbb{R}^m$

To show: A is invertible.

Let $\bar{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{m\times 1} \in \mathbb{R}^m$, $\bar{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}_{m\times 1} \in \mathbb{R}^m$, ... $\bar{e}_m = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^m$

By the given condition

$A\bar{x} = \bar{e}_i$ has atleast one solution for $i = 1, 2, \dots, m$

Let $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_m$ be the solutions.

i.e. $A\bar{v}_i = \bar{e}_i$ for $i = 1, 2, \dots, n$

Let $B = [\bar{v}_1, \bar{v}_2, \dots, \bar{v}_m]$

Then B is a $m\times m$ matrix

and
$$AB = A[\bar{v}_1, \bar{v}_2, \dots, \bar{v}_m]$$
$$= [A\bar{v}_1, A\bar{v}_2, \dots, A\bar{v}_m]$$
$$= [\bar{e}_1, \bar{e}_2, \dots, \bar{e}_m]$$

$$\Rightarrow AB = I_{m\times m}$$

So, B is a right inverse of A.
Now by corollary (1.2), A has an inverse and so, A is invertible. (QED)

<u>Note (Explanation):</u>

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}_{m\times m}$

and $B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{bmatrix}_{m\times m}$

Now $AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{bmatrix}$

$$= \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+\cdots+a_{1m}b_{m1} & a_{11}b_{12}+a_{12}b_{22}+\cdots+a_{1m}b_{m2} & \cdots & a_{11}b_{1m}+\cdots+a_{1m}b_{mm} \\ a_{21}b_{11}+a_{22}b_{21}+\cdots+a_{2m}b_{m1} & a_{21}b_{12}+a_{22}b_{22}+\cdots+a_{2m}b_{m2} & \cdots & a_{21}b_{1m}+\cdots+a_{2m}b_{mm} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}b_{11}+a_{m2}b_{21}+\cdots+a_{mm}b_{m1} & a_{m1}b_{12}+a_{m2}b_{22}+\cdots+a_{mm}b_{m2} & \cdots & a_{m1}b_{1m}+\cdots+a_{mm}b_{mm} \end{bmatrix}$$

$$= \left[\begin{bmatrix} a_{11} & a_{12} \cdots a_{1m} \\ a_{21} & a_{22} \cdots a_{2m} \\ \cdots & \cdots \\ a_{m1} & a_{m2} \cdots a_{mm} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \cdots a_{1m} \\ a_{21} & a_{22} \cdots a_{2m} \\ \cdots & \cdots \\ a_{m1} & a_{m2} \cdots a_{mm} \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \end{bmatrix} \cdots \begin{bmatrix} a_{11} & a_{12} \cdots a_{1m} \\ a_{21} & a_{22} \cdots a_{2m} \\ \cdots & \cdots \\ a_{m1} & a_{m2} \cdots a_{mm} \end{bmatrix} \begin{bmatrix} b_{1m} \\ b_{2m} \\ \vdots \\ b_{mm} \end{bmatrix} \right]$$

$$= \left[A \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{bmatrix} \quad A \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \end{bmatrix} \cdots\cdots A \begin{bmatrix} b_{1m} \\ b_{2m} \\ \vdots \\ b_{mm} \end{bmatrix} \right]$$

$$= \left[A\bar{v}_1 \quad A\bar{v}_2 \ \cdots \ A\bar{v}_m \right]$$

where $\bar{v}_i = \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{mi} \end{bmatrix} =$ The ith column vector of the matrix B

for $i = 1, 2, \ldots, m$

(i.e. $B = [\bar{v}_1 \ \bar{v}_2 \ \cdots \ \bar{v}_m]$)

So, $AB = [A\bar{v}_1 \quad A\bar{v}_2 \ \cdots \ A\bar{v}_m]$

Corollary (1.4) (Alternative version of Last equivalence of Theorem (1))

A matrix A is invertible if and only if the system $A\bar{x}=\bar{b}$ has a unique solution for any choice of vector $\bar{b} \in \mathbb{R}^m$

Note: The proof of the implication:

"The system $A\bar{x}=\bar{b}$ has a unique solution for any choice of vector $\bar{b} \in \mathbb{R}^m$
$\Rightarrow$ A is invertible"

is exactly same as the proof of (d) $\Rightarrow$ (a) in Theorem (1).

Now in the converse part, to prove that A is invertible $\Rightarrow$ $A\bar{x}=\bar{b}$ has a unique solution for any choice of vector $\bar{b} \in \mathbb{R}^m$, first we need to prove existence of a solution of $A\bar{x} = \bar{b}$ for any choice of $\bar{b} \in \mathbb{R}^m$

in exactly the same way as $(a) \Rightarrow (d)$ in Theorem 1.

<u>To prove</u> the uniqueness of the solution

assume $A\bar{v}_1 = \bar{b}$ and $A\bar{v}_2 = \bar{b}$ be two such solutions.

Then $A(\bar{v}_1 - \bar{v}_2) = A\bar{v}_1 - A\bar{v}_2 = \bar{b} - \bar{b} = \bar{0}$

$\Rightarrow A(\bar{v}_1 - \bar{v}_2) = \bar{0}$

i.e. $\bar{v}_1 - \bar{v}_2$ is a solution of the homogeneous system $A\bar{x} = \bar{0}$

Since A is invertible, by Theorem 1, the homogeneous system $A\bar{x} = \bar{0}$ has only the trivial solution.

Hence $\bar{v}_1 - \bar{v}_2 = \bar{0} \Rightarrow \bar{v}_1 = \bar{v}_2$

i.e. the solution is unique.

MTH 100 : Lecture 9

LU factorization of a matrix

Motivation: $Ax = b_1$, $Ax = b_2$, ..., $Ax = b_m$

(So, b_i's change but A remains fixed)

- One way of solving is to find A^{-1} and then find $A^{-1} b_i$ for $i = 1, 2, \ldots, m$

- However a more efficient way is to factor $A = LU$ which requires reducing to an echelon form only.
  Then the equation can be solved.

Definition: Suppose A is a $m \times n$ matrix which can be reduced to an echelon form (upper triangular) matrix without using row

interchange operations.
(So, only replacement operations are used in the forward phase of the row reduction algorithm)

Then A can be factorized as

$$A = LU$$

where L is an $m \times m$ Lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form matrix obtained from A by row reduction.

Any such factorization is called LU factorization of A.

- The matrix L is invertible and is called a Unit Lower triangular matrix.

- **Application to Solving Linear System**

Consider $A\bar{x} = \bar{b}$

Let $A = LU$

$A\bar{x} = \bar{b} \Rightarrow (LU)\bar{x} = \bar{b}$

$\Rightarrow L(U\bar{x}) = b \Rightarrow \boxed{U\bar{x} = L^{-1}\bar{b}}$

Note that

$$L(L^{-1}\bar{b}) = (LL^{-1})\bar{b} = \bar{b}$$

So, $\boxed{L^{-1}\bar{b} \text{ is a solution of } L\bar{y} = \bar{b}}$

- The solution of the $A\bar{x} = \bar{b}$ is replaced by the solution of

$$\left.\begin{aligned} L\bar{y} &= \bar{b} \\ \text{and} \quad U\bar{x} &= \bar{y} \end{aligned}\right\}$$

These systems are triangular and therefore easy to solve.

Ex: Let $A = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix}$

- We row reduce A to echelon form without interchanges or scaling and by adding multiples of a row to a lower row at every step:

$$\begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix} \xrightarrow[\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1}]{} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 5 & 6 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 - 5R_2} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$= U$ (say)

Note of caution: In general U doesn't always have 1's in the diagonal.

Let us write the row operations e_i and their inverses f_i:

$e_1 : R_2 \to R_2 - R_1$ $\qquad$ $f_1 : R_2 \to R_2 + R_1$

$e_2 : R_3 \to R_3 - 2R_1$ $\qquad$ $f_2 : R_3 \to R_3 + 2R_1$

$e_3 : R_3 \to R_3 - 5R_2$ $\qquad$ $f_3 : R_3 \to R_3 + 5R_2$

To get L from I, we find $(f_1 f_2 f_3)I$

Explanation:

Note: The same steps which take A to U will take L to I.

So, $I = e_3(e_2(e_1(L))) = E_3(E_2(E_1 L))$

$= (E_3 E_2 E_1)L$ where $E_i = e_i(I)$

$$\Rightarrow L = (E_3 E_2 E_1)^{-1} I = (E_1^{-1} E_2^{-1} E_3^{-1}) I$$

$$\Rightarrow L = f_1(f_2(f_3(I)))$$

So, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[f_3 : R_3 \to R_3 + 5R_2]{} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}$

$$\Big\downarrow f_2 : R_3 \to R_3 + 2R_1$$

$$\underbrace{\begin{bmatrix} \boxed{1} & 0 & 0 \\ 1 & \boxed{1} & 0 \\ 2 & 5 & \boxed{1} \end{bmatrix}_{3\times 3}}_{= \, L \text{ (say)}} \xleftarrow[f_1 : R_2 \to R_2 + R_1]{} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 5 & 1 \end{bmatrix}$$

Note that L is a lower triangular matrix with 1's on the diagonal i.e. L is a unit lower triangular matrix.

check: $A = LU$ (??)

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix}$$

$= A$ (as desired)

Ex: Solve $A\bar{x} = \bar{b}$ where $\bar{b} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

and $A = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix}$

Note that $A\bar{x} = \bar{b}$ is solved by solving $L\bar{y} = \bar{b}$ and then solving $U\bar{x} = \bar{y}$

We have $A = LU$ where $L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 5 & 1 \end{bmatrix}$

and $U = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (By the previous example)

Let $\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ and consider $L\bar{y} = \bar{b}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{aligned} y_1 &= 2 \Rightarrow & y_1 &= \boxed{2} \\ y_1 + y_2 &= -1 \Rightarrow & y_2 &= -1 - y_1 = -1-2 = \boxed{-3} \\ 2y_1 + 5y_2 + y_3 &= 1 \Rightarrow & y_3 &= 1 - 2y_1 - 5y_2 \\ & & &= 1 - 2(2) - 5(-3) = \boxed{12} \end{aligned}$$

(Forward substitution)

Now we solve $U\bar{x} = \bar{y}$

$$\Rightarrow \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 12 \end{bmatrix}$$

$$\Rightarrow \left.\begin{aligned} x_1 - x_2 - 2x_3 &= 2 \\ x_2 + x_3 &= -3 \\ x_3 &= 12 \end{aligned}\right\}$$

Then $x_3 = \boxed{12}$

$x_2 = -3 - x_3 = -3 - 12 = \boxed{-15}$

and $x_1 = 2 + x_2 + 2x_3 = 2 + (-15) + 2(12)$

$= \boxed{11}$

(Backward Substitution)

<u>Check:</u> <u>$A\bar{x} = \bar{b}$ (??)</u>

$$A\bar{x} = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 3 & 2 \end{bmatrix} \begin{bmatrix} 11 \\ -15 \\ 12 \end{bmatrix} = \begin{bmatrix} 11 + 15 - 24 = 2 \\ 11 - 12 = -1 \\ 22 - 45 + 24 = 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

$= \bar{b}$

(as desired)

Ex: Find the solution of $A\bar{x} = \bar{b}$ for general $\bar{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ with this given A.

Here after solving $L\bar{y} = \bar{b}$, we will get:

$$\begin{aligned} y_1 &= b_1 \\ y_2 &= b_2 - b_1 \\ y_3 &= b_3 - 2b_1 - 5(b_2 - b_1) \end{aligned}$$

Then $U\bar{x} = \bar{y} \Rightarrow x_3 = \boxed{y_3}$

and $x_2 = y_2 - x_3 = \boxed{y_2 - y_3}$

and $x_1 = y_1 + x_2 + 2x_3$

$\Rightarrow x_1 = y_1 + y_2 - y_3 + 2y_3$

$\Rightarrow x_1 = \boxed{y_1 + y_2 + y_3}$

So, from the entries of L and U, the system can be solved easily for any $\bar{b}$.

MTH : Lecture 10

LU algorithm :

- Input a $m \times n$ matrix A
- Step 1 : Row reduce A, if possible to an echelon form matrix U, using only row replacement operations that add a multiple of a row to a row below it.
- Step 2 : Place entries in L such that the same sequence of row operations reduces L to I.

Remark : (a) The computational efficiency of the LU approach depends on the fact that L is obtained without doing any significant extra work.

(b) Step 1 is not always possible but if it is, then the theoretical justification indicates why an LU factorization is then obtained.

Theoretical Justification :

- Suppose it is possible to row reduce A to an echelon form matrix U using only row replacement operations that add a multiple of a row to a row below it.
- Then there are unit lower triangular elementary matrices $E_1, E_2, \ldots, E_p$

Such that $E_p \cdots E_2 E_1 A = U$

$E_i = e_i(I)$ Note that E_i is unit lower triangular because e_i is a replacement operation.

- Hence $A = (E_p \cdots E_1)^{-1} U$

$$\Rightarrow A = LU$$

where $L = (E_p \cdots E_1)^{-1}$ is clearly invertible.

Also note that inverses and products of unit Lower triangular matrices are also unit Lower triangular matrices.

(Verify !!) (Exercise)

• Finally $L = (E_p \cdots E_1)^{-1}$

$$\Rightarrow L = (E_p \cdots E_1)^{-1} I$$

$$\Rightarrow L = (E_1^{-1} E_2^{-1} \cdots E_p^{-1}) I = (F_1 F_2 \cdots F_p) I$$
$$= f_1(f_2 \cdots f_p(I) \ldots)$$

Also $\boxed{I = (E_p \cdots E_1) L}$

i.e. the same sequence of row operations that reduces A to U reduces L to I.

LU factorization (General Case)

• In practical work, row interchanges are almost always used for computational stability.

So, this situation can be handled in nearly the same way except that the resultant L is not necessarily unit Lower triangular but is permuted unit Lower triangular.

(i.e. we can make L into a unit Lower triangular matrix by a permutation of the row.)

Ex: $A = \begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix}$

$$\begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix} \xrightarrow[R_3 \to R_3 + \frac{1}{2}R_1]{R_2 \to R_2 - 3R_1} \begin{bmatrix} 2 & -4 & 4 & -2 \\ 0 & 3 & -5 & 3 \\ 0 & -6 & 10 & -1 \end{bmatrix}$$

$$\xrightarrow{R_3 \to R_3 + 2R_2} \begin{bmatrix} 2 & -4 & 4 & -2 \\ 0 & 3 & -5 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix} = U \text{ (say)}$$

Row operations

$e_1 : \quad R_2 \to R_2 - 3R_1$

$e_2 : \quad R_3 \to R_3 + \frac{1}{2}R_1$

$e_3 : \quad R_3 \to R_3 + 2R_2$

Inverse operations

$f_1 : \quad R_2 \to R_2 + 3R_1$

$f_2 : \quad R_3 \to R_3 - \frac{1}{2}R_1$

$f_3 : \quad R_3 \to R_3 - 2R_2$

Now L will be a 3×3 unit lower triangular matrix.

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

$$\Big\downarrow \; R_3 \longrightarrow R_3 - \tfrac{1}{2} R_1$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -\frac{1}{2} & -2 & 1 \end{bmatrix} \xleftarrow{R_2 \longrightarrow R_2 + 3R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & -2 & 1 \end{bmatrix}$$

$$\| \\ L \text{ (say)}$$

So, $A = LU$ (check !!)

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -\frac{1}{2} & -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -4 & 4 & -2 \\ 0 & 3 & -5 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix} = A$$

Ex: Solve $A\bar{x} = \bar{b}$ where $\bar{b} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$

and A is given in the previous example.

We will use the result of the previous example : $A = LU$

First we solve $L\bar{y} = \bar{b}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -\frac{1}{2} & -2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$$

$\Rightarrow \quad y_1 = \boxed{2}$

and $3y_1 + y_2 = 1 \Rightarrow y_2 = 1 - 3y_1 = 1 - 3(2)$
$= \boxed{-5}$

and $-\frac{1}{2}y_1 - 2y_2 + y_3 = 4 \Rightarrow y_3 = 4 + \frac{1}{2}y_1 + 2y_2$
$= 4 + \frac{1}{2}(2) + 2(-5)$
$= 5 - 10 = \boxed{-5}$

So, $\bar{y} = \begin{bmatrix} 2 \\ -5 \\ -5 \end{bmatrix}$

Now we solve $U\bar{x} = \bar{y}$

$$\begin{bmatrix} 2 & -4 & 4 & -2 \\ 0 & 3 & -5 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \\ -5 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 2x_1 - 4x_2 + 4x_3 - 2x_4 = 2 \\ 3x_2 - 5x_3 + 3x_4 = -5 \\ x_3 = x_3 \rightarrow \text{(dummy equation)} \\ 5x_4 = -5 \end{cases}$$

Note that x_3 is a free variable here.

$$\Rightarrow \quad x_4 = -1 = \boxed{-1 + 0.x_3}$$

and $x_3 = \boxed{0 + 1.x_3}$

and $3x_2 = -5 + 5x_3 - 3x_4$

$$\Rightarrow 3x_2 = -5 + 5x_3 - 3(-1)$$

$$\Rightarrow 3x_2 = -2 + 5x_3$$

$$\Rightarrow \boxed{x_2 = -\frac{2}{3} + \frac{5}{3}x_3}$$

and $2x_1 = 2 + 4x_2 - 4x_3 + 2x_4$

$$\Rightarrow 2x_1 = 2 + 4\left(-\frac{2}{3} + \frac{5}{3}x_3\right) - 4x_3 + 2(-1)$$

$$\Rightarrow 2x_1 = 2 - \frac{8}{3} + \frac{20}{3}x_3 - 4x_3 - 2$$

$$\Rightarrow 2x_1 = -\frac{8}{3} + \frac{8}{3}x_3$$

$$\Rightarrow \boxed{x_1 = -\frac{4}{3} + \frac{4}{3}x_3}$$

So, $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ 0 \\ -1 \end{bmatrix}}_{\bar{u}} + x_3 \underbrace{\begin{bmatrix} \frac{4}{3} \\ \frac{5}{3} \\ 1 \\ 0 \end{bmatrix}}_{\bar{v}}$

So, the set of solutions can be written as:

$$S = \{ \bar{u} + t\bar{v} \; : \; t \in \mathbb{R} \}$$

<u>check:</u>

$$A\bar{u} = \begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix} \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ 0 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{8}{3} + \frac{8}{3} + 0 + 2 \\ -8 + 6 + 3 \\ \frac{4}{3} + \frac{8}{3} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} = \bar{b}$$

and

$$A\bar{v} = \begin{bmatrix} 2 & -4 & 4 & -2 \\ 6 & -9 & 7 & -3 \\ -1 & -4 & 8 & 0 \end{bmatrix} \begin{bmatrix} \frac{4}{3} \\ \frac{5}{3} \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{8}{3} - \frac{20}{3} + 4 + 0 \\ 8 - 15 + 7 + 0 \\ -\frac{4}{3} - \frac{20}{3} + 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \bar{0}$$

Ex: Let $A = \begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix}$

$$\begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix} \xrightarrow[\substack{R_2 \to R_2 + 2R_1 \\ R_3 \to R_3 - \frac{3}{2}R_1 \\ R_4 \to R_4 + 3R_1 \\ R_5 \to R_5 - 4R_1}]{} \begin{bmatrix} 2 & -6 & 6 \\ 0 & -7 & 5 \\ 0 & 14 & -10 \\ 0 & -14 & 10 \\ 0 & 21 & -15 \end{bmatrix}$$

$R_3 \to R_3 + 2R_2$
$R_4 \to R_4 - 2R_2$
$R_5 \to R_5 + 3R_2$

$$\downarrow$$

$$\begin{bmatrix} 2 & -6 & 6 \\ 0 & -7 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = U \text{(say)}$$

Now L will be a 5×5 Lower triangular matrix.

Row operations	Inverse Operations
$e_1 : R_2 \longrightarrow R_2 + 2R_1$	$f_1 : R_2 \longrightarrow R_2 - 2R_1$
$e_2 : R_3 \longrightarrow R_3 - \frac{3}{2}R_1$	$f_2 : R_3 \longrightarrow R_3 + \frac{3}{2}R_1$
$e_3 : R_4 \longrightarrow R_4 + 3R_1$	$f_3 : R_4 \longrightarrow R_4 - 3R_1$
$e_4 : R_5 \longrightarrow R_5 - 4R_1$	$f_4 : R_5 \longrightarrow R_5 + 4R_1$
$e_5 : R_3 \longrightarrow R_3 + 2R_2$	$f_5 : R_3 \longrightarrow R_3 - 2R_2$
$e_6 : R_4 \longrightarrow R_4 - 2R_2$	$f_6 : R_4 \longrightarrow R_4 + 2R_2$
$e_7 : R_5 \longrightarrow R_5 + 3R_2$	$f_7 : R_5 \longrightarrow R_5 - 3R_2$

$$I_{5\times 5} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[\substack{R_5 \to R_5 - 3R_2 \\ R_4 \to R_4 + 2R_2 \\ R_3 \to R_3 - 2R_2}]{} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & -3 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[\substack{R_5 \to R_5 + 4R_1 \\ R_4 \to R_4 - 3R_1 \\ R_3 \to R_3 + \frac{3}{2}R_1 \\ R_2 \to R_2 - 2R_1}]{} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ \frac{3}{2} & -2 & 1 & 0 & 0 \\ -3 & 2 & 0 & 1 & 0 \\ 4 & -3 & 0 & 0 & 1 \end{bmatrix} = L$$

So, $A = LU$ where L and U are given above.

Let $$\bar{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

and we solve for $\bar{y}$ in $L\bar{y} = \bar{b}$

$$L\bar{y} = \bar{b} \Rightarrow \boxed{y_1 = b_1}$$

and $-2y_1 + y_2 = b_2 \Rightarrow y_2 = b_2 + 2y_1$

$$\Rightarrow \boxed{y_2 = b_2 + 2b_1}$$

and $\frac{3}{2}y_1 - 2y_2 + y_3 = b_3 \Rightarrow y_3 = b_3 - \frac{3}{2}y_1 + 2y_2$

$$\Rightarrow y_3 = b_3 - \frac{3}{2}b_1 + 2(b_2 + 2b_1)$$

$$\Rightarrow \boxed{y_3 = b_3 + 2b_2 + \frac{5}{2}b_1}$$

and $-3y_1 + 2y_2 + y_4 = b_4 \Rightarrow y_4 = b_4 + 3y_1 - 2y_2$

$$\Rightarrow y_4 = b_4 + 3b_1 - 2(b_2 + 2b_1)$$

$$\Rightarrow \boxed{y_4 = b_4 - 2b_2 - b_1}$$

and $4y_1 - 3y_2 + y_5 = b_5$

$$\Rightarrow y_5 = b_5 - 4y_1 + 3y_2$$

$$\Rightarrow y_5 = b_5 - 4b_1 + 3(b_2 + 2b_1)$$

$$\Rightarrow \boxed{y_5 = b_5 + 3b_2 + 2b_1}$$

So, $$\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 + 2b_1 \\ b_3 + 2b_2 + \frac{5}{2}b_1 \\ b_4 - 2b_2 - b_1 \\ b_5 + 3b_2 + 2b_1 \end{bmatrix}$$

- First we choose $b_1 = 2$, $b_2 = 3$, $b_3 = 5$, $b_4 = 8$, $b_5 = -13$

Then $y_1 = \boxed{2}$, $y_2 = 3 + 2(2) = \boxed{7}$, $y_3 = 5 + 2(3) + \frac{5}{2}(2) = \boxed{16}$

$y_4 = 8 - 2(3) - 2 = \boxed{0}$, $y_5 = -13 + 3(3) + 2(2) = \boxed{0}$

Now $U\bar{x} = \bar{y} \Rightarrow \begin{bmatrix} 2 & -6 & 6 \\ 0 & -7 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 16 \\ 0 \\ 0 \end{bmatrix}$

$$\Rightarrow \left.\begin{aligned} 2x_1 - 6x_2 + 6x_3 &= 2 \\ -7x_2 + 5x_3 &= 7 \\ 0 &= 16 \leftarrow \text{ a Contradiction} \\ 0 &= 0 \\ 0 &= 0 \end{aligned}\right\}$$

So, for this choice of $\bar{b}$, the system doesn't have any solution.

- Now we choose $b_1 = 2$, $b_2 = 3$, $b_3 = -11$, $b_4 = 8$, $b_5 = -13$

Then $y_1 = \boxed{2}$, $y_2 = 3 + 2(2) = \boxed{7}$

$y_3 = -11 + 2(3) + \frac{5}{2}(2) = -11 + 11 = \boxed{0}$

$y_4 = 8 - 2(3) - 2 = 8 - 8 = \boxed{0}$

$y_5 = -13 + 3(3) + 2(2) = -13 + 13 = \boxed{0}$

Now $U\bar{x} = \bar{y} \Rightarrow \begin{bmatrix} 2 & -6 & 6 \\ 0 & -7 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$$\Rightarrow \left.\begin{aligned} 2x_1 - 6x_2 + 6x_3 &= 2 \\ -7x_2 + 5x_3 &= 7 \\ 0 &= 0 \\ 0 &= 0 \\ 0 &= 0 \end{aligned}\right\}$$

$$\Rightarrow \quad -7x_2 = 7 - 5x_3$$

(Here x_3 will be a free variable)

$$\Rightarrow \quad x_2 = -1 + \frac{5}{7}x_3$$

and $\quad 2x_1 = 2 + 6x_2 - 6x_3$

$$\Rightarrow 2x_1 = 2 + 6\left(-1 + \frac{5}{7}x_3\right) - 6x_3$$

$$\Rightarrow 2x_1 = -4 - \frac{12}{7}x_3$$

$$\Rightarrow x_1 = -2 - \frac{6}{7}x_3$$

$$\Rightarrow \quad \begin{aligned} x_1 &= -2 - \frac{6}{7}x_3 \\ x_2 &= -1 + \frac{5}{7}x_3 \\ x_3 &= \phantom{-1 + \frac{5}{7}} x_3 \end{aligned}$$

$\longrightarrow$ (dummy equation)

$$\Rightarrow \quad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix}}_{\bar{u}} + x_3 \underbrace{\begin{bmatrix} -\frac{6}{7} \\ \frac{5}{7} \\ 1 \end{bmatrix}}_{\bar{v}}$$

Then the set of solutions can be written as:

$$S = \{ \bar{u} + t\bar{v} \text{ where } t \in \mathbb{R} \}$$

<u>Check:</u>

$$A\bar{u} = \begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix} \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4+6=2 \\ 8-5=3 \\ -6-5=-11 \\ 12-4=8 \\ -16+3=-13 \end{bmatrix}$$

$$\Rightarrow A\bar{u} = \begin{bmatrix} 2 \\ 3 \\ -11 \\ 8 \\ -13 \end{bmatrix} = \boxed{\bar{b}}$$

and $$A\bar{v} = \begin{bmatrix} 2 & -6 & 6 \\ -4 & 5 & -7 \\ 3 & 5 & -1 \\ -6 & 4 & -8 \\ 8 & -3 & 9 \end{bmatrix} \begin{bmatrix} -\frac{6}{7} \\ \frac{5}{7} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{12}{7} - \frac{30}{7} + 6 = 0 \\ \frac{24}{7} + \frac{25}{7} - 7 = 0 \\ -\frac{18}{7} + \frac{25}{7} - 1 = 0 \\ \frac{36}{7} + \frac{20}{7} - 8 = 0 \\ -\frac{48}{7} - \frac{15}{7} + 9 = 0 \end{bmatrix}$$

$$\Rightarrow A\bar{v} = \boxed{\bar{0}}$$

MTH 100 : Lecture 11

- We would like to study matrices further.

Instead of looking matrices as a rectangular array of numbers, we will look at a matrix as

<u>a collection of rows</u> (row vectors)

<u>a collection of columns</u> (column vectors)

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \equiv \left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \cdots \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right)$$

or

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- Also, if we multiply a matrix with a $(n \times 1)$ column we get a $(m \times 1)$ column

i.e. $$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \underbrace{\begin{pmatrix} - \\ - \\ \vdots \\ - \end{pmatrix}}_{n \times 1} = \underbrace{\begin{pmatrix} - \\ - \\ \vdots \\ - \end{pmatrix}}_{m \times 1}$$

Thus a <u>transformation</u> is taking place here.

- <u>Earlier</u> (<u>Ref:</u> Lecture 5) we have looked at rows or columns with 2 or 3 entries as vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ and looked at their properties like <u>addition and scalar multiplication</u>.

- The important aspect is <u>these properties and not the set it self</u>.

- We can take <u>some other set</u> and observe the <u>same properties</u>.

The set can be : a collection of numbers
a collection of row vectors
a collection of column vectors
a collection of some matrices
a collection of some sequence of numbers
a collection of some functions.

- This leads to Axiomatic approach to Mathematics.

Axioms : Some Rules to be followed by the elements of a set.

Thus we now have a Set and Axioms/Rules.

(Exact nature of the elements of the set is not important)

Any proof about the set uses only the axioms and logical reasoning.

This leads to Abstract system or structure.

- Different sets following the same set of axioms give the same system.
- Different axioms leads to different systems.

- Advantages of Axiomatic approach:
- Proofs are more rigorous.
- It unifies different Mathematical objects.
- Ease of proof in abstract setting.
  (compared to concrete or specific Mathematical objects.)

  (compared to concrete or specific mathematical objects)
- Novelty
  This approach leads to different branches of Mathematics.

Examples:

- Discrete Mathematics: Graphs, Lattices, Posets.
- Algebra: Groups, Rings, Fields.

- Measure Theory and Probability : Measure space, Measurable set.
- Topology : Metric Spaces and Topological Spaces.
- Functional Analysis : Banach Spaces or Hilbert Spaces.
- Linear Algebra : Vector Space over a field.
- Elements of the vector space are called vectors and elements of the field are called scalars.

- Let us look at the set of real numbers and observe its properties such as addition, subtraction, multiplication, division etc.
- We can add two real numbers and get another real number.

We can multiply two real numbers and get another real number. i.e. $a+b \in \mathbb{R}\ \forall\, a,b \in \mathbb{R}$ and $a.b \in \mathbb{R}\ \forall\, a,b \in \mathbb{R}$

(Note: $\forall \equiv$ for all, $\in \equiv$ belongs to)

- Addition and multiplication satisfies associative and commutative properties.

$$\left.\begin{array}{l} a+(b+c)=(a+b)+c \\ a(bc)=(ab)c \end{array}\right\} \forall\, a,b,c \in \mathbb{R}$$

$$\left.\begin{array}{l} a+b=b+a \\ ab=ba \end{array}\right\} \forall\, a,b \in \mathbb{R}$$

- We have real number 0 (additive identity) such that $a+0=0+a=a\ \forall a \in \mathbb{R}$

and real number 1 (multiplicative identity) such that $1.a=a.1=a\ \forall\, a \in \mathbb{R}$

- For every real number, we have its negative (additive inverse). For every $a \in \mathbb{R}$, there exists $-a \in \mathbb{R}$ such that $a+(-a)=0$

For every nonzero real number, we have its reciprocal (multiplicative inverse). For every $a \in \mathbb{R}$ $(a \neq 0)$, there exists $a^{-1} \in \mathbb{R}$ such that $a^{-1}a = aa^{-1} = 1$

- Real numbers satisfies Distributive properties: $a(b+c)=ab+ac\ \forall\, a,b,c \in \mathbb{R}$

This leads to the concept of a Field.

Informal Definition: An algebraic system with addition and multiplication of elements in which universal addition, subtraction, multiplication and division except that division by zero element (denoted by 0) is not possible.

Formal Definition:

A field is a nonempty set F with two binary operations called addition and multiplication which satisfies the following properties.

(Usual symbols: '$+$' for addition and $\cdot$ or $*$ for multiplication sometimes omitted altogether for multiplication.
Thus: $a+b$; $a \cdot b$ or $a*b$ or ab will be written)

(A) Closure Under the operation:

For every element $a, b \in F$,
$a+b \in F$ and $a \cdot b \in F$

(B) The following properties hold:

(a) Associative Property:

$$\left.\begin{aligned} (a+b)+c &= a+(b+c) \\ (a\cdot b)\cdot c &= a\cdot(b\cdot c) \end{aligned}\right\} \text{ for every } a, b, c \in F$$

(b) Commutative Property:

$$\left.\begin{aligned} a+b &= b+a \\ a\cdot b &= b\cdot a \end{aligned}\right\} \text{ for every } a, b \in F$$

(c) Zero property and identity (unit) property:

There exists a zero element '0' and an identity or unit element '1' (not the same as zero element which satisfies

$$\left.\begin{aligned} a+0 &= 0+a = a \\ \text{and} \quad 1.a &= a.1 = a \end{aligned}\right\} \text{ for every } a \in F$$

(d) For any $a \in F$, it will have an additive inverse b and multiplicative inverse c (provided $a \neq 0$) such that

$$a+b = 0$$

and $\quad a.c = 1 \quad$ (provided $a \neq 0$)

(e) Distributive Property:

$$a.(b+c) = a.b + a.c \quad \text{for every } a, b, c \in F$$

Examples:

(1) We have shown that set of real numbers $\mathbb{R}$ is a field with respect to usual addition and multiplication

(2) Set of rational numbers $\mathbb{Q}$ is a field with respect to usual addition and multiplication.

(3) Set of Complex numbers $\mathbb{C}$ is also a field with respect to usual addition and multiplication.

Examples:

(1) Let $\mathbb{Z}$ be the set of integers.

There is no multiplicative inverse for any non-zero integer.

So, $\mathbb{Z}$ is not a field with respect to usual addition and multiplication.

(2) Let $\mathbb{R}^{2\times 2}$ be the set of all 2×2 matrices with real entries.

Now commutative property of multiplication doesnot hold.

Therefore $\mathbb{R}^{2\times 2}$ is not a field.

(3) Let us consider the set of all 2×2 invertible matrices with real entries.

Now sum of two invertible matrices may not be an invertible matrix.

e.g. both $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ are invertible but

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$ is not invertible.

So, the above set is not closed under matrix addition and therefore it is not a field.

<u>Properties:</u> The following properties hold for every field F.

- Since $1 \neq 0$, F must have atleast two elements. (This is an axiom)
- The zero element is unique and the unit element is unique.
- The additive inverse of every $c \in F$ is unique and is usually denoted by $-c$.
- The multiplicative inverse of every $c \in F$ $(c \neq 0)$ is unique and is usually denoted by c^{-1}.
- $0.c = 0$ for every $c \in F$

<u>Proof:</u> $0 + 0.c = 0.c = (0+0).c = 0.c + 0.c$ (By Distributive Property)

$\Rightarrow$ adding additive inverse of $0.c$ to both sides (from the right)

we obtain $\boxed{0 = 0.c}$

- F has no <u>zero divisors.</u>

(Definition: An element $c \in F$ $(c \neq 0)$ is said to be a zero divisor if there exists an element $d \in F$ $(d \neq 0)$ such that $c.d = 0$)

<u>Proof:</u> If F has a zero divisor $c \in F$ $(c \neq 0)$, then there exists $d \in F$ $(d \neq 0)$ such that $c.d = 0$

$\Rightarrow$ Multiplying both sides by d^{-1} from the right we get

$(c.d).d^{-1} = 0.d^{-1} \Rightarrow c.(d.d^{-1}) = 0$ (By the previous property)

$\Rightarrow c.1 = 0 \Rightarrow c = 0$, a contradiction.

Vector Space:

A vector space is a non-empty set V of objects called vectors together with an associated field F of scalars with two operations called addition and scalar multiplication which satisfies the following properties:

(A) Closure under operations:

$$u + v \in V \text{ for every } u, v \in V$$

and $cu \in V$ for every $u \in V$ and every $c \in F$

(B) The following properties hold for addition:

(a) associative property:

$(u+v)+w = u+(v+w)$ for every $u, v, w \in V$

(b) Identity property:

There exists a 'zero vector' 0 such that

$0 + u = u + 0 = u$ for every vector $u \in V$

(c) Every vector $u \in V$ has an additive inverse $v \in V$ such that

$$u + v = v + u = 0$$

(d) Commutative property:

$u + v = v + u$ for every $u, v \in V$

Moreover:

(C) (a) $c(u+v) = cu + cv$ for every $u, v \in V$ and every $c \in F$

(b) $(c+d)u = cu + du$ for every $u \in V$ and every $c, d \in F$

(c) $c(du) = (cd)u$ for every $u \in V$ and every $c, d \in F$

(d) $1.u = u$ for every $u \in V$ where 1 is the unit element of F.

Note: (1) In lecture 5, we have shown that $\mathbb{R}^2$ is a vector space over the field $\mathbb{R}$.

(Although the definition of vector space was not mentioned there, all the properties were shown)

(2) In the same way we can show that

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_1, x_2, \ldots, x_n \in \mathbb{R} \right\}$$

is a vector space over the base field $\mathbb{R}$.

Note that addition and scalar multiplication in $\mathbb{R}^n$ is defined by:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix} \quad \text{and} \quad c\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix} \text{ for every } c \in \mathbb{R}$$

MTH 100 : Lecture 12

Last time we have defined field and Vector space over a field

Examples of Vector spaces

Ex ①: The space

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : x_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots n \right\}$$

(For any $n \geq 1$

The base field is $\mathbb{R}$

Addition:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

Scalar Multiplication:

For any $c \in \mathbb{R}$

$$c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx \\ \vdots \\ cx_n \end{bmatrix}$$

Want to show that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$.

Closure Property:

Let $u = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $v = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$

and Let $c \in \mathbb{R}$

$$u + v = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \in \mathbb{R}^n$$

$$c\,u = \begin{bmatrix} c\,x_1 \\ \vdots \\ c\,x_n \end{bmatrix} \in \mathbb{R}^n$$

- Let $w = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{R}^n$

$$(u+v)+w = \left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) + \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} (x_1 + y_1) + z_1 \\ \vdots \\ (x_n + y_n) + z_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + (y_1 + z_1) \\ \vdots \\ x_n + (y_n + z_n) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 + z_1 \\ \vdots \\ y_n + z_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \left(\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \right)$$

$$= u + (v + w)$$

- $\bar{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$ and

$$\bar{0} + u = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} 0 + x_1 \\ \vdots \\ 0 + x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = u$$

for every $u \in \mathbb{R}^n$

Similarly $u + \bar{0} = u$ for every $u \in \mathbb{R}^n$

- For every $u = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$

$$\begin{bmatrix} -x_1 \\ \vdots \\ -x_n \end{bmatrix} \in \mathbb{R}^n$$

$$= (-u \text{ (say)})$$

such that $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} -x_1 \\ \vdots \\ -x_n \end{bmatrix} = \begin{bmatrix} x_1 + (-x_1) \\ \vdots \\ x_n + (-x_n) \end{bmatrix}$

$$= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

- $u + v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

$$= \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} y_1 + x_1 \\ \vdots \\ y_n + x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= v + u$$

To show:

$$c(u+v) = cu + cv$$

$$c\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}\right) = c\begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

$$= \begin{bmatrix} c(x_1+y_1) \\ \vdots \\ c(x_n+y_n) \end{bmatrix} = \begin{bmatrix} cx_1+cy_1 \\ \vdots \\ cx_n+cy_n \end{bmatrix}$$

$$= \begin{bmatrix} cx_1 \\ \vdots \\ cx_n \end{bmatrix} + \begin{bmatrix} cy_1 \\ \vdots \\ cy_n \end{bmatrix}$$

$$= c\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + c\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$= cu + cv$$

- To show that $(c+d)u = cu + du$:

$$(c+d)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} (c+d)x_1 \\ \vdots \\ (c+d)x_n \end{bmatrix}$$

$$= \begin{bmatrix} cx_1+dx_1 \\ \vdots \\ cx_n+dx_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ \vdots \\ cx_n \end{bmatrix} + \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix}$$

$$= c\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + d\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = cu + du$$

- To show that $c(du) = (cd)u$

$$c(du) = c\left(d\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = c\begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix}$$

$$= \begin{bmatrix} c(dx_1) \\ \vdots \\ c(dx_n) \end{bmatrix} = \begin{bmatrix} (cd)x_1 \\ \vdots \\ (cd)x_n \end{bmatrix} = (cd)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (cd)u$$

To show:

- $1.u = u$ where $1 \in \mathbb{R}$

$$1.\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1.x_1 \\ \vdots \\ 1.x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = u$$

We can now conclude that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$

Ex: Is $\mathbb{R}^n$ a vector space over the base field $\mathbb{Q}$?

Ex: Is $\mathbb{R}^n$ a vector space over the base field $\mathbb{C}$?

- $\mathbb{R}^n$ is frequently referred to as Euclidean space.

<u>Ex ②:</u> The space $\mathbb{R}^{m \times n}$ of $m \times n$ matrices with real entries.

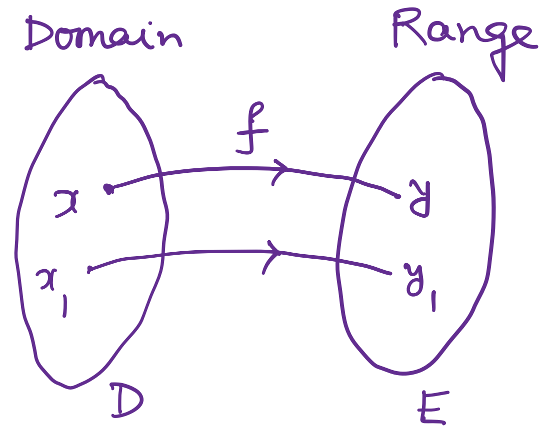
This is a vector space over $\mathbb{R}$. (Show that!)

<u>Note:</u> This vector space is useful in <u>image processing</u>:

$\downarrow$

Function :

A function is a correspondence between two sets called Domain and Range such that for every element in the domain, there is a corresponding element in the range.



$y = f(x)$

(y is called the image of x under f)

- More than one element can correspond to the same element in the range but one element can not correspond to more than one element in the range.

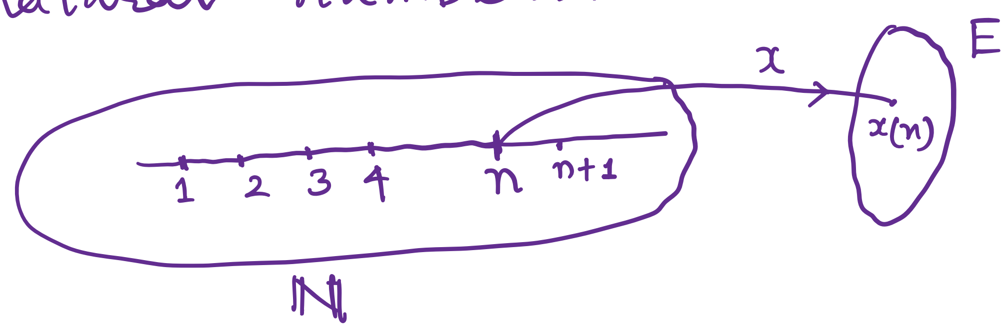
- If the range is a subset of $\mathbb{R}$ (or $\mathbb{C}$) we call it real valued (complex valued) function.

  In that case if the domain D is also a subset of $\mathbb{R}$ (or $\mathbb{C}$), we can talk about continuity and differentiability of the functions.

• Note: Two functions f and g are equal if the values (images) of f and g are equal in every point of the domain.

Definition: If the domain of a function is $\mathbb{N}$ (the set of natural numbers) we call it a sequence

Thus a sequence is a function of natural numbers.



If x is a sequence, the image of n under x is often denoted by

$$x(n) = x_n$$

Then we can denote a sequence by $\{x_1, x_2, x_3, \dots\}$ or by $\{x_n\}$

Thus we can count the terms of a sequence one by one (it is countably infinite)

- If we have a sequence $\{x_n\}$ we can see if $\lim_{n \to \infty} x_n$ exists.

If $\lim_{n \to \infty} x_n$ exists, it is called a convergent sequence.

Ex: ① Let $\{x_n\} = \{\frac{1}{n}\} = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$

$$\lim_{n \to \infty} x_n = \lim_{n \to \infty} \frac{1}{n} = 0$$

So, the sequence $\{x_n\}$ converges to 0

② Let $\{x_n\} = \{n+1\} = \{2, 3, 4 \dots\}$

$\{x_n\}$ doesn't converge

(3) Let $\{x_n\} = \{(-1)^n\} = \{-1, 1, -1, 1, -1, 1, \dots\}$

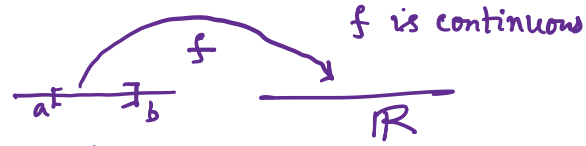
$\{x_n\}$ doesn't converge.

Ex ③: Let $[a, b]$ be a fixed closed interval on $\mathbb{R}$.

Let $C[a, b]$ be the set of all continuous functions from $[a, b]$ to $\mathbb{R}$.

This is a vector space over $\mathbb{R}$.

Vector addition :



If f and $g \in C[a,b]$

We define

$$f+g \text{ by } (f+g)(x) = f(x)+g(x)$$

for every $x \in [a,b]$

Scalar multiplication

If $c \in \mathbb{R}$, $f \in C[a,b]$

We define

$$cf \text{ by } (cf)(x) = c.f(x)$$

for every $x \in [a,b]$

Show that $C[a,b]$ is a vector space over $\mathbb{R}$.

Note: Often we take $[a,b]$ as $[0,1]$ or $[0,2\pi]$

This vector space is important in Signal and System.

Ex ④: The space $\mathbb{R}^\infty$ of real sequences is a vector space over $\mathbb{R}$.

$$\mathbb{R}^\infty = \left\{ \{a_n\} : \{a_n\} \text{ is a sequence of real numbers} \right\}$$

Addition: $\{a_n\} + \{b_n\} = \{a_n + b_n\}$

scalar multiplication: $c\{a_n\} = \{c\,a_n\}$

- Note: This is useful in discrete or digital signals.

- Of more interest is

  $C \subset \mathbb{R}^\infty$

  Set of convergent sequences (It is a subset of $\mathbb{R}^\infty$)

- C is also a vector space over $\mathbb{R}$

MTH:100 : Lecture 13

<u>Question:</u> Is $\mathbb{R}^n$ a vector space over $\mathbb{Q}$?

We know $\mathbb{R}^n$ is a vector space over the field $\mathbb{R}$

- The additive properties are satisfied.
- The properties of scalar multiplication is satisfied for all scalar in $\mathbb{R}$ and so is satisfied for all rational numbers (since $\mathbb{Q} \subset \mathbb{R}$). So, $\mathbb{R}^n$ is a vector space over $\mathbb{Q}$.

<u>Question:</u> Is $\mathbb{R}^n$ a vector space over $\mathbb{C}$?

Let $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$, $\underset{\text{(scalar)}}{i \in \mathbb{C}}$ $\quad i \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} i \\ i \\ \vdots \\ i \end{pmatrix} \notin \mathbb{R}^n$

So, it is not closed under multiplication by complex numbers. Hence $\mathbb{R}^n$ is not a vector space over $\mathbb{C}$.

<u>Ex:</u>

Let $C[a,b] = \{ f : [a,b] \to \mathbb{R} : f \text{ is Continuous} \}$

Then $C[a,b]$ is a vector space over the base field $\mathbb{R}$.

<u>Note:</u> Two functions are equal if they have the same values at all points in their common domain. ($\forall \equiv$ for all)

- The addition and scalar multiplication in $C[a,b]$ is defined as:
- For $f, g \in C[a,b]$, $f+g$ is defined as $\boxed{(f+g)(x) = f(x) + g(x) \quad \forall x \in [a,b]}$

• and for $c \in \mathbb{R}$ and $f \in C[a,b]$, cf is defined as $\boxed{(cf)(x) = c(f(x)) \ \forall\, x \in [a,b]}$

First we check closure property:

We know that sum of two continuous function is continuous.

So, if $f, g \in C[a,b]$, $f+g \in C[a,b]$

Also constant multiple of a continuous function is continuous.

So, if $f \in C[a,b]$ and $c \in \mathbb{R}$, then $cf \in C[a,b]$.

• Now let $f, g, h \in C[a,b]$.

$$\begin{aligned}
[(f+g)+h](x) &= (f+g)(x) + h(x) \\
&= [f(x)+g(x)] + h(x) \\
&= f(x) + [g(x)+h(x)] \quad \text{(By associative property of real numbers)} \\
&= [f+(g+h)](x) \quad \forall\, x \in [a,b]
\end{aligned}$$

So, $(f+g)+h = f+(g+h)$

Now the function $\bar{0}$ defined by

$$\bar{0}(x) = 0 \quad \forall \ x \in [a, b]$$

is a continuous function and so

$\bar{0} \in C[a, b]$.

Now $(f + \bar{0})(x) = f(x) + \bar{0}(x) = f(x) + 0$

$= f(x) \ \forall x \in [a, b]$

$(\bar{0} + f)(x) = \bar{0}(x) + f(x)$

$= 0 + f(x) = f(x) \ \forall x \in [a, b]$

So, $\quad f + \bar{0} = \bar{0} + f = f \quad \forall f \in [a, b]$

Now for any $f \in C[a, b]$,

define $-f$ as $(-f)(x) = -f(x)$

$\forall x \in [a, b]$

Then $-f \in C[a, b]$

and $[f + (-f)](x) = f(x) + (-f)(x)$

$= f(x) - f(x) = 0$

$= \bar{0}(x)$

$[(-f) + f](x) = (-f)(x) + f(x)$

$= -f(x) + f(x) = 0 = \bar{0}(x)$

$\forall x \in [a, b]$

So, $f + (-f) = (-f) + f = \bar{0}$
$\forall f \in C[a,b]$

Now $(f+g)(x) = f(x) + g(x)$ (By commutative property of addition of real numbers)
$= g(x) + f(x)$
$= (g+f)(x) \;\forall x \in [a,b]$

So, $f+g = g+f$

So, Commutative property is satisfied

Now for $c, d \in \mathbb{R}$ and $f, g \in C[a,b]$

$[c(f+g)](x) = c\,(f+g)(x)$
$= c\,[f(x) + g(x)]$
$= c f(x) + c g(x) = (cf)(x) + (cg)(x)$
$= (cf + cg)(x) \;\forall x \in [a,b]$

$\Rightarrow c(f+g) = cf + cg$

Now $[(c+d)f](x) = (c+d).f(x)$
$= c f(x) + d f(x) = (cf)(x) + (df)(x)$

$$= (cf + df)(x) \quad \forall\, x \in [a, b]$$

$$\Rightarrow \quad (c + d)f = cf + df$$

Now, $[c(df)](x) = c\,(df)(x)$

$$= c\,[d\,f(x)] = (cd)\,f(x)$$

$$= [(cd)f](x) \quad \forall\, x \in [a, b]$$

So, $\quad c(df) = (cd)f$

Also $(1.f)(x) = 1\,f(x) = f(x)$

$$\forall\, x \in [a, b]$$

$$\Rightarrow \quad \boxed{1.f = f}$$

(1 is the unit element of the field $\mathbb{R}$)

Therefore $C[a, b]$ is a vector space over $\mathbb{R}$.

Ex: $\quad \mathbb{R}^{\infty} = \{ (a_n) : (a_n) \text{ is a sequence of real numbers} \}$

This is a vector space over $\mathbb{R}$

$$(a_n) + (b_n) = (a_n + b_n)$$

$$c\,(a_n) = (c\,a_n) \quad \text{where } c \in \mathbb{R}.$$

$\mathbb{N} = \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6$

$(a_n): \quad a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad a_6 \quad \cdots$

$(b_n): \quad b_1 \quad b_2 \quad b_3 \quad b_4 \quad b_5 \quad b_6 \quad \cdots$

$\boxed{(a_n)+(b_n)} \rightarrow \{ a_1+b_1 \quad a_2+b_2 \quad a_3+b_3, \cdots \}$

$\boxed{c\,(a_n)} \rightarrow \{ ca_1 \quad ca_2 \quad ca_3, \cdots \}$

Ex: Let $R_n(t)$ be the set of all polynomials (in variable t) of degree $\leq n$ with with real coefficients.

e.g. For $n=3$, $R_3(t)$ is the set of polynomials (in variable t) of degree ≤ 3.

Note: $a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$ is a polynomial (in variable of t) of degree n

Show that $R_n(t)$ is a vector space over the field $\mathbb{R}$.

(The zero polynomial is regarded as an element of $R_n(t)$ for $n = 0, 1, 2, \dots$)

Note: The set of all polynomials of degree n with real coefficients is not a vector space over $\mathbb{R}$.
(closure property of addition is not satisfied)

Ex: The set $R(t)$ of all polynomials with real coefficients is a vector space over $\mathbb{R}$.

- Note that $R_n(t) \subset R(t)$ for any positive integer n.

Note: Vector spaces can also be defined over $\mathbb{C}$ or any other field.

Back to field:

Consider $\{0, 1\} = \mathbb{Z}_2$ (notation)

Define: addition and multiplication by:

$\oplus$	0	1
0	0	1
1	1	0

$*$	0	1
0	0	0
1	0	1

(Arithmatic modulo 2)

- $\mathbb{Z}_2$ is a field. (check!)

So, we can consider $\mathbb{Z}_2^n$: The set of n-tuples whose entries are from $\mathbb{Z}_2$.

$$= \{ (x_1, x_2, \dots x_n) : x_i = 0 \text{ or } 1 \}$$

A typical element of $\mathbb{Z}_2^n$ is the n-tuple $(0, 1, 1, 0, \dots \dots 0)$

$\mathbb{Z}_2^n$ is extremely important in coding.

Ex: $1.((1+0)+1)+1$

$$= 1.(1+1)+1 = 1.0+1 = 0+1$$
$$= \boxed{1}$$

Ex: Using modular arithmatic (mod 2) find

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Modular Arithmatic :

Let n be a positive integer.

For any integer a,

define $a \pmod n$ = The remainder when a is divided by n

Note that $0 \le \text{remainder} < n$

Ex: $10 \pmod 3 = 1$

$7 \pmod 4 = 3$

Now for any positive integer n

define $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$

Define : $\left.\begin{aligned} a \oplus b &= (a+b) \bmod n \\ a * b &= (a \cdot b) \bmod n \end{aligned}\right\}$ for all $a, b \in \mathbb{Z}_n$

So, $\mathbb{Z}_2 = \{0, 1\}$ In $\mathbb{Z}_2$, $1+1 = 2 = 0$ since $2 \pmod 2 = 0$

$\mathbb{Z}_3 = \{0, 1, 2\}$

In $\mathbb{Z}_3$, $2+1 = 3 = 0$ (since $3 \pmod 3 = 0$)

$2+2 = 4 = 1$ (since $4 \pmod 3 = 1$)

Ex: Show that $\mathbb{Z}_2$ and $\mathbb{Z}_3$ are fields.

Ex: $\mathbb{Z}_4 = \{0, 1, 2, 3\}$

Show that $\mathbb{Z}_4$ is not a field.

Proposition: $\mathbb{Z}_p$ is a field if and only if p is a prime.

Note: One direction '$\Rightarrow$' will be proved.
The other direction '$\Leftarrow$' will not be proved since it will need more modular arithmatic.

Ex: We have seen examples of fields such as: $\underbrace{\mathbb{Q} \subset \mathbb{R}} \subset \mathbb{C}$

Question: Are there any field between $\mathbb{Q}$ and $\mathbb{R}$?

Define:

$$\mathbb{Q}(\sqrt{2}) = \{ a + b\sqrt{2} \; : \; a, b \in \mathbb{Q} \}$$

clearly $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{R}$

- Show that $\mathbb{Q}(\sqrt{2})$ is a field with respect to usual addition and multiplication of real numbers.

MTH 100 : Lecture 14

Ex: $\mathbb{Z}_2 = \{0, 1\}$ is a field.

$\left(\begin{array}{c} 1 * 1 = 1 \\ 1^{-1} = 1 \text{ in } \mathbb{Z}_2 \end{array} \right)$

Ex: $\mathbb{Z}_3 = \{0, 1, 2\}$ is a field.

$1 \times 1 = 1 \pmod 3 = 1$
So, $1^{-1} = 1$ in $\mathbb{Z}_3$
$2 * 2 = 4 \pmod 3 = 1$
So, $2^{-1} = 2$ in $\mathbb{Z}_3$

Ex: $\mathbb{Z}_4 = \{0, 1, 2, 3\}$

Now $2 \in \mathbb{Z}_4$, $2 \neq 0$

and $2 * 2 = 4 \pmod 4 = 0$

So, $2 * 2 = 0$ in $\mathbb{Z}_4$.

So, 2 is a zero divisor in $\mathbb{Z}_4$.

So, $\mathbb{Z}_4$ can not be a field.

Ex: $\mathbb{Z}_6$ is not a field.

Ex: Let n be a composite integer. Then $\mathbb{Z}_n$ is not a field.

$\left(\begin{array}{c} n = r k \\ 1 < r < n \\ 1 < k < n \end{array} \right)$

Zero divisor:

A zero divisor is a non zero element $a (\neq 0)$ such that there exists $b (\neq 0)$ satisfying

$$a * b = 0$$

- A field cannot have a zero divisor

Proof: Let F be a field and $a \in F$, $a \neq 0$ is a zero divisor in F.

So, there exists $b \in F$ $(b \neq 0)$ such that

$$a * b = 0$$

Since $b \neq 0$ and $b \in F$, $b^{-1} \in F$

Now from the above, $(a * b) * b^{-1} = 0 * b^{-1}$

$\Rightarrow a * (b * b^{-1}) = 0$ (By property of field)

$\Rightarrow a * (e) = 0$ (where e = multiplicative identity in F)

$\Rightarrow a = 0$, a contradiction

So, $\boxed{F \text{ doesnot have a zero divisor}}$

<u>Theorem:</u> $\mathbb{Z}_p$ is a field iff (if and only if) p is a prime.

(Note: will prove one part of the theorem)

<u>Proof:</u> '$\Longrightarrow$': <u>Given:</u> $\mathbb{Z}_p$ is a field.

<u>Want to show:</u> p is a prime.

We will prove it BWOC (BWOC $\equiv$ By way of contradiction)

Let us assume that p is not a prime.

Then $p = rk$ where $1 < r < p$
$1 < k < p$

$$r * k = rk \pmod p = p \pmod p = 0$$

So, r and k are both divisors of zero in $\mathbb{Z}_p$

Thus $\mathbb{Z}_p$ can not be a field, a contradiction.

Hence p has to be a prime.

<u>'$\Longleftarrow$':</u> (This requires more of modular arithmetic).

Consequences of the Vector Space definition:

Proposition: Let V be a vector space over a field F.

Then (a) The zero vector is unique.

(b) The additive inverse vector of any vector u is unique

(We use the notation $-u$ for the inverse vector of u.)

(c) $0.u = \bar{0}$ $\forall$ $u \in V$

(d) $c.\bar{0} = \bar{0}$ $\forall$ $c \in F$

(e) $-u = (-1)u$ $\forall$ $u \in V$

(f) Cancellation Law:

If $u + v = u + w$

then $v = w$ $\forall$ $u, v, w \in V$

Proof: Exercise

① let $\bar{0}$ and $\bar{0}_1$ be zero vectors in V

$\bar{0} + u = u + \bar{0} = u$ $\forall$ $u \in V$ ---- ①

$\bar{0}_1 + u = u + \bar{0}_1 = u$ $\forall$ $u \in V$ --- ②

Let $u = \bar{0}_1$ in ①. Then $\bar{0} + \bar{0}_1 = \bar{0}_1$

Now let $u = \bar{0}$ in ②. Then $\bar{0} + \bar{0}_1 = \bar{0}$, Combining we get $\boxed{\bar{0} = \bar{0}_1}$

② Let $u \in V$, Let v and v_1 be two additive inverses of u

So, $u + v = v + u = \bar{0}$ ....①

and $u + v_1 = v_1 + u = \bar{0}$ .....②

From ② $(u + v_1) + v = \bar{0} + v = v$

By ② $(v_1 + u) + v = v$

$\Rightarrow v_1 + (u + v) = v$

$\Rightarrow v_1 + \bar{0} = v$ (By ①)

$\Rightarrow \boxed{v_1 = v}$

So, additive inverse is unique.

(c), (d), (e) : Exercise

(f) Given $u + v = u + w$

Now $-u \in V$

and so, $(-u) + (u + v) = (-u) + (u + w)$

$\Rightarrow ((-u) + u) + v = ((-u) + u) + w$

$\Rightarrow 0 + v = 0 + w$

$\Rightarrow \boxed{v = w}$

- When we gave examples of vector spaces, we noticed some subsets:

e.g. $C \subset R^\infty$, $\quad C^1[a,b] \subset C[a,b]$

$R_0(t) \subset R_1(t) \subset R_2(t) \subset \dots \subset R_n(t) \subset R(t)$

Subspace:

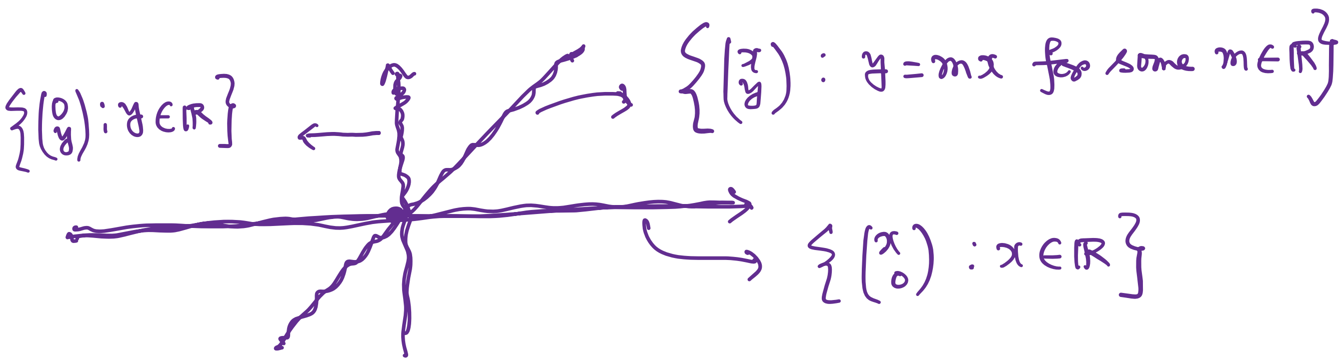
Let V be a vector space over the field F.

A (vector) subspace of V is a nonempty subset of V which is also a vector space over F with the operations of vector addition and scalar multiplication taken from V.

Ex: If V is any vector space over F, then $\{0\}$ (zero subspace) and V are always subspaces of V.

Subspaces other than V and $\{0\}$ are known as proper subspaces.

<u>Ex:</u> $\mathbb{R}^2$ is a vector space over $\mathbb{R}$.



- The set $\left\{\begin{pmatrix}x\\0\end{pmatrix} : x \in \mathbb{R}\right\}$ is a proper subspace of $\mathbb{R}^2$.

- The set $\left\{\begin{pmatrix}0\\y\end{pmatrix} : y \in \mathbb{R}\right\}$ is a proper subspace of $\mathbb{R}^2$.

- The set $\left\{\begin{pmatrix}x\\y\end{pmatrix} : y = mx \text{ for some } m \in \mathbb{R}\right\}$ is a proper subspace of $\mathbb{R}^2$.

Question: Is $\mathbb{R}$ a subspace of $\mathbb{R}^2$?

$\mathbb{R}$ is not even a subset of $\mathbb{R}^2$

$$\mathbb{R}^2 = \left\{ \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\text{(two tuple)}} : x, y \in \mathbb{R} \right\}$$

$$\mathbb{R} = \left\{ \underbrace{x}_{\text{(one tuple)}} : x \in \mathbb{R} \right\}$$

Now $\left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$ is a subspace of $\mathbb{R}^2$

It behaves very much like $\mathbb{R}$, but is logically distinct from $\mathbb{R}$

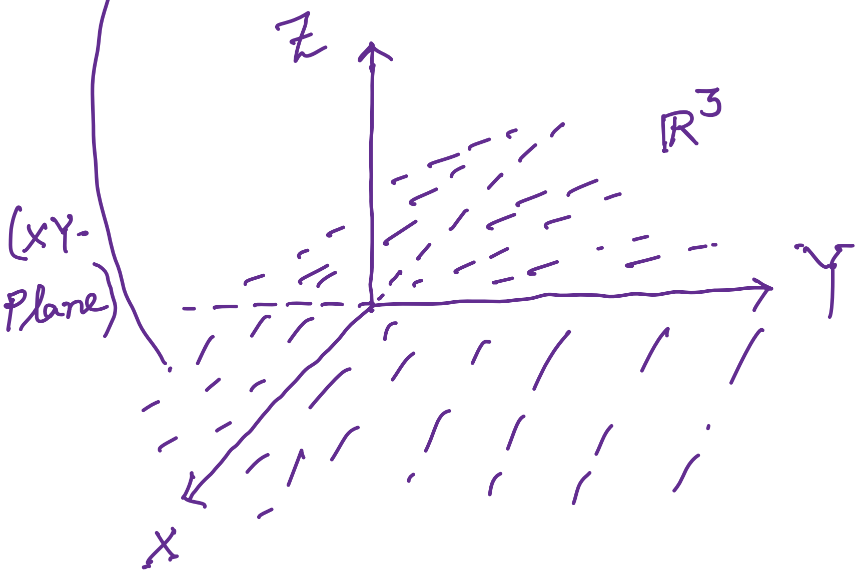
Example: $\mathbb{R}^3$ is a vector space over $\mathbb{R}$.

Question: Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?

No: $\mathbb{R}^2$ is not even a subset of $\mathbb{R}^3$

The set $\left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$

is a subspace of $\mathbb{R}^3$ which behaves very much like $\mathbb{R}^2$, but is logically distinct from $\mathbb{R}^2$.



Test for Subspaces

Proposition: Let V be a vector space over a field F. A subset W of V is a subspace if and only if it satisfies the following three properties:

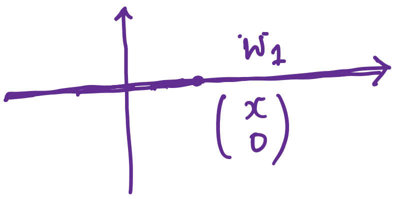
(1) The zero vector 0 is in W

(2) W is closed under addition
i.e. $u+v \in W \quad \forall\, u, v \in W$

(3) W is closed under scalar multiplication
i.e. $cu \in W \quad \forall\, c \in F$ and $\forall\, u \in W$

Note: (1) can be replaced by (1')
(1'): W is nonempty.

Ex(1) $\mathbb{R}^2$ is a vector space over $\mathbb{R}$



$$W_1 = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$$

is a subspace.

Check:

(1) $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in W_1$

(2) If $\begin{pmatrix} x_1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} x_2 \\ 0 \end{pmatrix} \in W_1$, then

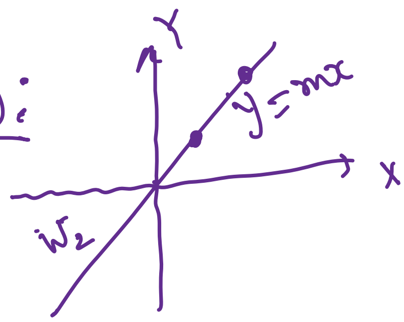
$$\begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} \in W_1$$

(3) If $\begin{pmatrix} x \\ 0 \end{pmatrix} \in W_1$ and $c \in \mathbb{R}$,

then $c \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} cx \\ 0 \end{pmatrix} \in W_1$

Hence, W_1 is a subspace of $\mathbb{R}^2$

Ex(2):



$$W_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : \begin{array}{l} y = mx \\ x, y \in \mathbb{R} \end{array} \right\}$$

Check

(1) $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in W_2$

(2) If $\begin{pmatrix} x_1 \\ mx_1 \end{pmatrix}$ and $\begin{pmatrix} x_2 \\ mx_2 \end{pmatrix} \in W_2$

then $\begin{pmatrix} x_1 \\ mx_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ mx_2 \end{pmatrix} = \begin{pmatrix} x_1+x_2 \\ mx_1+mx_2 \end{pmatrix}$

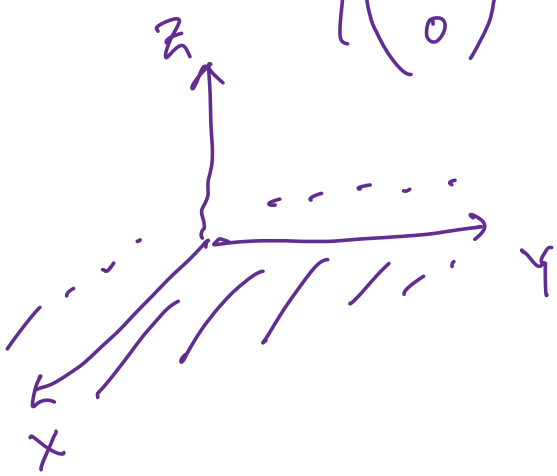
$$= \begin{pmatrix} x_1+x_2 \\ m(x_1+x_2) \end{pmatrix} \in W_2$$

(3) If $c \in \mathbb{R}$ and $\begin{pmatrix} x \\ mx \end{pmatrix} \in W_2$

then $c\begin{pmatrix} x \\ mx \end{pmatrix} = \begin{pmatrix} cx \\ c(mx) \end{pmatrix} = \begin{pmatrix} cx \\ m(cx) \end{pmatrix} \in W_2$

Hence W_2 is a subspace of $\mathbb{R}^2$.

Ex: Let $W = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\} \subset \mathbb{R}^3$



Show that W is a subspace of $\mathbb{R}^3$

MTH100: Lecture 15

Subspace: Let V be a vector space over a field F.
A (vector) subspace of V is a nonempty subset of V which is also a vector space over F with respect to the same operations of vector addition and scalar multiplication taken from V.

Test for Subspaces

Proposition: Let V be a vector space over a field F.
A subset W of V is a subspace if and only if it satisfies the following three properties:

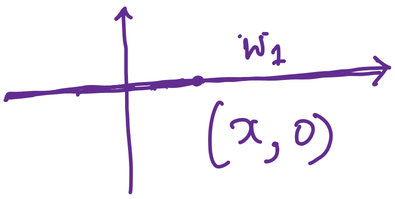
(1) The zero vector $\bar{0}$ is in W

(2) W is closed under addition
i.e. $u+v \in W \quad \forall\, u, v \in W$

(3) W is closed under scalar multiplication
i.e. $cu \in W \quad \forall\, c \in F$ and $\forall\, u \in W$

Note: (1) can be replaced by (1')
(1') : W is nonempty.

Ex(1) $\mathbb{R}^2$ is a vector space over $\mathbb{R}$



$$W_1 = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$$

is a subspace.

Check:

(1) $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in W_1$

(2) If $\begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_2 \\ 0 \end{pmatrix} \in W_1$ then

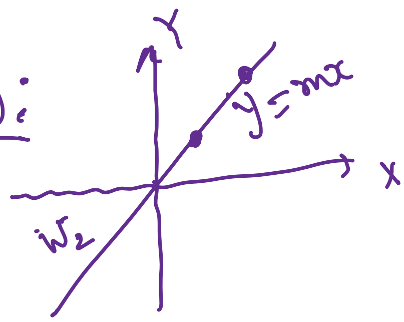
$$\begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} \in W_1$$

(3) If $\begin{pmatrix} x \\ 0 \end{pmatrix} \in W_1$ and $c \in \mathbb{R}$,

then $c \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} cx \\ 0 \end{pmatrix} \in W_1$

Hence, W_1 is a subspace of $\mathbb{R}^2$

Ex(2):



$$W_2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : \begin{matrix} y = mx \\ x, y \in \mathbb{R} \end{matrix} \right\}$$

Check

(1) $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in W_2$ (since $0 = m.0$)

(2) If $\begin{pmatrix} x_1 \\ mx_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ mx_2 \end{pmatrix} \in W_2$

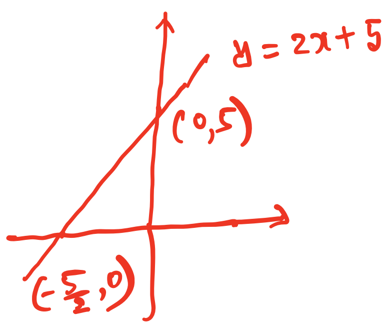
then $\begin{pmatrix} x_1 \\ mx_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ mx_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ mx_1 + mx_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ m(x_1 + x_2) \end{pmatrix} \in W_2$

(3) If $\begin{pmatrix} x \\ mx \end{pmatrix} \in W_2$ and $c \in \mathbb{R}$

then $c\begin{pmatrix} x \\ mx \end{pmatrix} = \begin{pmatrix} cx \\ c(mx) \end{pmatrix} = \begin{pmatrix} cx \\ m(cx) \end{pmatrix} \in W_2$

Hence W_2 is a subspace of $\mathbb{R}^2$

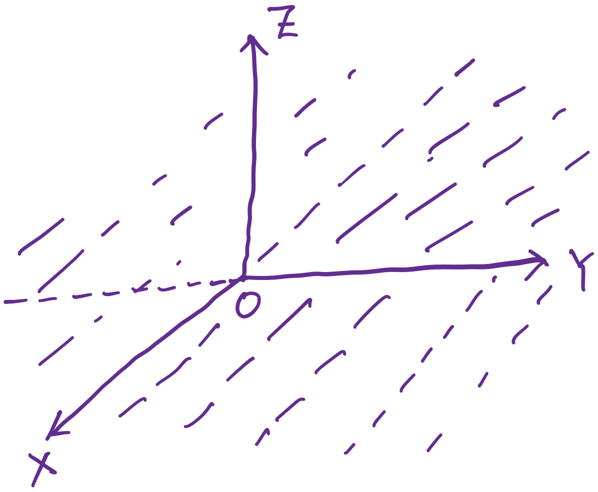
Ex ③: Is $W_3 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : y = 2x + 5, \ x, y \in \mathbb{R} \right\}$ a subspace of $\mathbb{R}^2$



$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \notin W_3$ So, W_3 is not a subspace of $\mathbb{R}^2$.

Ex 4: Let $W = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\} \subset \mathbb{R}^3$

Show that W is a subspace of $\mathbb{R}^3$



(1) $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in W$

(2) If $\begin{pmatrix} x_1 \\ y_1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix} \in W$

then $\begin{pmatrix} x_1 \\ y_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ 0 \end{pmatrix} \in W$

(3) If $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in W$ and $c \in \mathbb{R}$

then $c \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} cx \\ cy \\ 0 \end{pmatrix} \in W$

So, W is a subspace of $\mathbb{R}^3$

Ex 5: Consider the set W of all solutions of the system of equations $A\bar{x} = \bar{0}$ where A is a $m \times n$ matrix.

- W is a subset of $\mathbb{R}^n$
- W is also a subspace of $\mathbb{R}^n$

(1) $\bar{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in W$ because $A\bar{0} = \bar{0}$

(2) If $\bar{x}$ and $\bar{y} \in W$ then $A\bar{x} = \bar{0}$, $A\bar{y} = \bar{0}$

So, $A(\bar{x} + \bar{y}) = A\bar{x} + A\bar{y} = \bar{0} + \bar{0} = \bar{0}$

So, $\bar{x} + \bar{y} \in W$

(3) If $\bar{x} \in W$ and $c \in \mathbb{R}$

then $A\bar{x} = \bar{0}$

Now $A(c\bar{x}) = c(A\bar{x}) = c.\bar{0} = \bar{0}$

So, $c\bar{x} \in W$

So, W is a subspace of $\mathbb{R}^n$.

Ex ⑥: Let W be the set of all solutions of the system $A\bar{x} = \bar{b}$ where A is a $m \times n$ matrix and $\bar{b}_{m\times 1} \neq \bar{0}$

Is this a subspace of $\mathbb{R}^n$?

$$A\bar{0} = \bar{0} \neq \bar{b} \text{ (given)}$$

So, $\bar{0} \notin W$

Hence W is not a subspace of $\mathbb{R}^n$.

Ex ⑦: Let $\mathbb{R}^{n\times n}$ be the set of all $n\times n$ matrices over $\mathbb{R}$.

Then $\mathbb{R}^{n\times n}$ is a vector space over $\mathbb{R}$ with respect to matrix addition and scalar multiplication.

Let W be the set of all $n\times n$ symmetric matrices over $\mathbb{R}$

i.e. $W = \{ A \in \mathbb{R}^{n\times n} : A^t = A \}$ (A^t = Transpose of A)

- W is a subspace of $\mathbb{R}^{n\times n}$

Check:

① $\mathbf{0}_{n\times n} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \end{pmatrix}_{n\times n} \in W$ (since $\mathbf{0}^t = \mathbf{0}$)

② Let $A, B \in W$; Then $A^t = A$ and $B^t = B$

$$\Rightarrow (A+B)^t = A^t + B^t = A + B$$

$$\Rightarrow A + B \in W$$

③ Let $A \in W$ and $c \in \mathbb{R}$; Then $A^t = A$

Now $(cA)^t = cA^t = cA$

So, $cA \in W$

Thus W is a subspace of $\mathbb{R}^{n\times n}$.

Ex: $\mathbb{R}^{\infty}$ (The set of all sequences with real entries) is a vector space over $\mathbb{R}$.

Let C be the set of all convergent sequences with real entries.

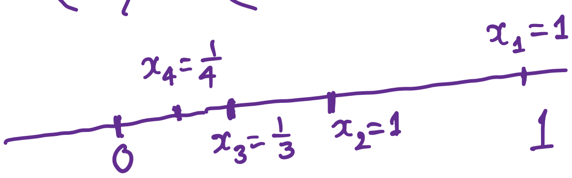
Note: Let (x_n) be a sequence of real numbers $(x_1, x_2, \ldots)$.

If there exists $l \in \mathbb{R}$ such that

$\lim\limits_{n \to \infty} x_n = l$, then (x_n) is called convergent and we say x_n converges to l.

Ex:

Let $(x_n) = \left(\frac{1}{n}\right)$



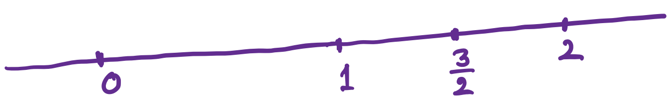
$\lim\limits_{n \to \infty} x_n = \lim\limits_{n \to \infty} \frac{1}{n}$

$= 0$

x_n converges to 0

Ex: Let $(x_n) = \left(1 + \frac{1}{n}\right)$

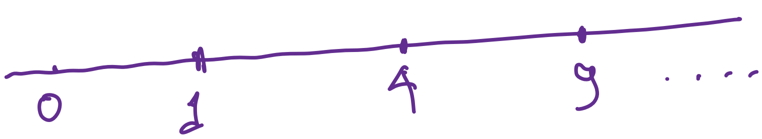
$\lim\limits_{n \to \infty} x_n = 1.$



x_n converges to 1.

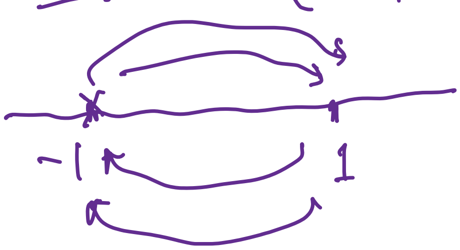
Ex: $(x_n) = (n^2)$

Not Convergent



Ex: $(x_n) = (-1)^n$

Not Convergent



$x_1 = -1$

$x_2 = 1$

$x_3 = -1$

$x_4 = 1$

(In fact it is oscillating)

Ex: We know $C \subset \mathbb{R}^\infty$. C is also a subspace of $\mathbb{R}^\infty$.

check:

(1) $(0, 0, 0, \ldots) \in C$

(2) $(x_n), (y_n) \in C$

Thus (x_n) is convergent, (y_n) is convergent.

Then $(x_n + y_n)$ is also convergent.

(If $x_n \to l$ and $y_n \to m$, then $x_n + y_n \to l + m$)

(3) Let $d \in \mathbb{R}$, $(x_n) \in C$

So, (x_n) is convergent. Then (dx_n) is also convergent.

(If $x_n \to l$, then $dx_n \to dl$)

Hence C is a subspace of $\mathbb{R}^\infty$.

Ex: $R_0(t) < R_1(t) < R_2(t) < \quad < \underset{\sim\sim}{R_n(t)} < \quad < R(t)$

Show that $R_n(t)$ is a subspace of $R(t)$.

Ex: Find some subspaces of $C[a,b]$

$C^1[a,b]$?? $C^1[a,b] = \{f : [a,b] \to \mathbb{R} : f$ is differentiable and f' is continuous$\}$

$C^\infty[a,b]$?? Similarly $C^2[a,b], \ldots, C^\infty[a,b]$ can be defined.

Another Test for Subspaces

Proposition 9: A non-empty subset W of V (V is a vector space over a field F) is a subspace if and only if for each u and v in W and each scalar c in F,

$$cu + v \in W$$

(i.e. $cu + v \in W \ \forall\, u, v \in W$ and $\forall c \in F$)

Exercise: Show that the two tests are equivalent.

Addendum to Lecture 15

Proposition 3: If A is a square matrix, then A is row equivalent to the identity matrix iff the homogeneous system $A\bar{x}=\bar{0}$ has only the trivial solution.

Proof: $\Rightarrow$: If $A_{n\times n}$ is a square matrix and A is row equivalent to the identity matrix I_n, then after row operations, A will be row reduced to the RREF matrix I_n.

Hence the system of equation will be $I_n \bar{x}_{n\times 1} = \bar{0}$

$$\Rightarrow \quad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

i.e. the system $A\bar{x}=\bar{0}$ has only the trivial solution.

$\Leftarrow$: Conversely assume that $A\bar{x}=\bar{0}$ has only the trivial solution.

Then if we row reduce A to its RREF matrix, there will be only basic variables and no free variable (because free variable will give non-trivial solution).

Thus RREF matrix will have only pivot columns. Since it is a square matrix there will be no zero rows and it will be the identity matrix I_n.

i.e. A is row equivalent to I_n

Proposition 4: The system $A\bar{x} = \bar{b}$ is consistent iff the rightmost column of R (where R is the RREF matrix corresponding to A) is not a pivot column.
i.e. there is no row of the form
$[0, 0, \ldots, 0, p]$ with p nonzero.

Proof: $\Rightarrow$:

If there is a row of the form
$[0, 0, \ldots, 0, p]$ with $p \neq 0$,
then if we write the reduced set of equations explicitly, one of the equation will be

$$0.x_1 + 0.x_2 + \cdots + 0.x_n = p$$

$$\Rightarrow \quad 0 = p \text{ where } p \neq 0$$

a contradiction
and $A\bar{x} = \bar{b}$ is inconsistent.

Thus $A\bar{x} = \bar{b}$ consistent $\Rightarrow$ There is no row of the form $[0, 0, \ldots, p]$ with $p \neq 0$

$\Leftarrow$ If there is no row of the form
$[0, 0, \ldots, p]$ with $p \neq 0$,
then if we write the reduced set of equations explicitly, there will be basic variables and possibly some free variables. The basic variables can be solved in terms of free variables and so the system $A\bar{x} = \bar{b}$ is consistent.

Observation 4:

If $A_1, A_2, \cdots A_n \,(n \geq 2)$ are invertible matrices then $C = A_1 A_2 \cdots\cdots A_n$ is invertible and

$$C^{-1} = A_n^{-1} \cdots A_2^{-1} A_1^{-1}$$

Proof: For $n=2$,

$$(A_2^{-1} A_1^{-1})(A_1 A_2) = A_2^{-1}(A_1^{-1} A_1)A_2 = A_2^{-1} I A_2$$
$$= A_2^{-1} A_2 = I$$

and $(A_1 A_2)(A_2^{-1} A_1^{-1}) = A_1 (A_2 A_2^{-1}) A_1^{-1} = A_1 I A_1^{-1}$

$$= A_1 A_1^{-1} = I$$

Thus for $n=2$, $(A_1 A_2)^{-1} = A_2^{-1} A_1^{-1}$

Now assume that

$$(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots\cdots A_2^{-1} A_1^{-1}$$

Now $(A_{k+1}^{-1} A_k^{-1} \cdots A_1^{-1})(A_1 \cdots A_k A_{k+1})$

$$= A_{k+1}^{-1} (A_k^{-1} \cdots A_1^{-1})(A_1 \cdots A_k) A_{k+1}$$
$$= A_{k+1}^{-1} I A_{k+1} = A_{k+1}^{-1} A_{k+1} = I$$

and $(A_1 \cdots A_k A_{k+1})(A_{k+1}^{-1} A_k^{-1} \cdots A_1^{-1})$

$$= (A_1 \cdots\cdots A_k)(A_{k+1} A_{k+1}^{-1})(A_k^{-1} \cdots\cdots A_1^{-1})$$
$$= (A_1 \cdots A_k)\, I\, (A_k^{-1} \cdots A_1^{-1}) = (A_1 \cdots A_k)(A_k^{-1} \cdots\cdots A_1^{-1})$$
$$= I$$

Thus the formula is true for $n = k+1$ if it is true for $n=k$.

Since the formula is proved for $n=2$, by the principle of mathematical induction it is true for all positive integer n.

Proposition 5: If e is an elementary row operation and E is the $m\times m$ elementary matrix $e(I_m)$, then for every $m\times n$ matrix A,

$$e(A) = EA$$

Proof: Please see solution of Worksheet 3. (problem 13)

Properties: Let V be a vector space over a field F.

Then (c) $0.u = \bar{0} \quad \forall\ u\in V$

(d) $c.\bar{0} = \bar{0} \quad \forall\ c\in F$

(e) $-u = (-1)u \quad \forall\ u\in V$

Proof: (c) $\bar{0} + 0.u = 0.u = (0+0).u = 0.u + 0.u$

$$\Rightarrow \bar{0} + 0.u + (-0.u) = 0.u + 0.u + (-0.u)$$

$$\Rightarrow \bar{0} + (0.u + (-0.u)) = 0.u + (0.u + (-0.u))$$

$$\Rightarrow \bar{0}+\bar{0} = 0.u + \bar{0} \Rightarrow \bar{0} = 0.u \Rightarrow \boxed{0.u = \bar{0}}$$

(d) $\bar{0} + c.\bar{0} = c\bar{0} = c(\bar{0}+\bar{0}) = c\bar{0} + c\bar{0}$

$$\Rightarrow \bar{0} + c.\bar{0} + (-c\bar{0}) = c\bar{0} + c\bar{0} + (-c\bar{0})$$

$$\Rightarrow \bar{0} + (c.\bar{0} + (-c\bar{0})) = c\bar{0} + (c\bar{0} + (-c\bar{0}))$$

$$\Rightarrow \quad \bar{0} + \bar{0} = c\bar{0} + \bar{0} \Rightarrow \bar{0} = c\bar{0} \Rightarrow \boxed{c\bar{0} = \bar{0}}$$

(e) $\bar{0} = 0.u = ((-1) + 1)u = (-1)u + 1u = (-1)u + u$

$$\Rightarrow \bar{0} + (-u) = (-1)u + u + (-u)$$

$$\Rightarrow \bar{0} + (-u) = (-1)u + (u + (-u))$$

$$\Rightarrow -u = (-1)u + \bar{0} \Rightarrow \boxed{-u = (-1)u}$$

Proposition: Test 1: Let V be a vector space over a field F. Then a nonempty subset W is a subspace of V

$\Leftrightarrow$ (1) $\bar{0} \in W$

(2) $u + v \in W \ \forall \ u, v \in W$

(3) $cu \in W \ \forall \ c \in F$ and $\forall \ u \in W$

Proof: $\Rightarrow$ If W is a subspace of V, then W is a vector space over the field F.

Then closure properties of addition & scalar multiplication are satisfied in W

$\Rightarrow \quad u + v \in W \ \forall \ u, v \in W$ and $cu \in W \ \forall c \in F$ and $\forall u \in W$

Also W must have the zero element and so $\bar{0} \in W$

$\Leftarrow$: Assume that W satisfies (1), (2) and (3)

(2) and (3) $\Rightarrow$ closure properties of addition and scalar multiplication are satisfied in W.

(1) $\Rightarrow \bar{0} \in W$

Now $c=-1$ in (3) $\Rightarrow$ $(-1)u = -u \in W$ $\forall\, u \in W$

i.e. additive inverse exists in W $\forall\, u \in W$.

Now associative property of addition, commutative property of addition and all other properties of scalar multiplication are hereditary and hence are satisfied in W (since they are satisfied in V).

Therefore W is a subspace of V.

<u>Note:</u> (1) can be replaced by $(1')$: $W \neq \phi$

i.e. (1), (2), (3) $\Leftrightarrow$ $(1')$, (2), (3).

<u>$\Rightarrow$:</u> Since $\bar{0} \in W$, $W \neq \phi$ and $(1')$ holds.

<u>$\Leftarrow$:</u> Since $W \neq \phi$, there exists $u \in W$

By (3) $(-1)u = -u \in W$ (By taking $c=-1$)

and By (2) $u + (-u) \in W$ $\Rightarrow$ $\bar{0} \in W$ and so (1) holds.

<u>Proposition: Test (2):</u> Let V be a vector space over a field F.

Then a nonempty subset W is a subspace of V

$\Leftrightarrow$ $cu + v \in W$ $\forall\, u, v \in W$ and $\forall\, c \in F$

The two tests are equivalent:

Test (1) $\Rightarrow$ Test (2)

If $u, v \in W$ and $c \in F$, then

$cu \in W$ (By ③ of Test(1))

$\Rightarrow \quad cu + v \in W$ (By (2) of Test(1))

Test (2) $\Rightarrow$ Test(1)

Since $W \neq \phi$, there exists an element $u \in W$

Taking $c = -1$ and $v = u$

we have $(-1)u + u = -u + u = 0 \in W$

$\boxed{\text{i.e. (1) of Test (1) holds.}}$

Now taking $c = 1$, we get $(1)u + v = u + v \in W$ $\forall\, u, v \in W$

$\boxed{\text{i.e. (2) of Test (1) holds.}}$

Taking $v = 0$ we get $cu + 0 = cu \in W$ $\forall\, u \in V$ and $\forall\, c \in F$

$\boxed{\text{i.e. (3) of Test (1) holds}}$

MTH 100 : Lecture 16

Span of a set of Vectors:

Let V be a vector space over a field F.

- Then a linear combination of finitely many given vectors is any sum of scalar multiples of the vectors.
- Thus if $\{v_1, v_2, \ldots, v_p\}$ is a finite set of vectors in V, then

  $c_1 v_1 + c_2 v_2 + \cdots + c_p v_p$ where $c_1, c_2, \ldots, c_p$

  are any set of scalars

  is a linear combination of $v_1, v_2, \ldots, v_p$.

Definition:

Let $S = \{v_1, v_2, \ldots, v_p\}$ be a finite set of vectors in a vector space V over a field F.

The span of S is defined as :

$$\text{Span } S = \{ c_1 v_1 + c_2 v_2 + \cdots + c_p v_p : c_1, c_2, \ldots, c_p \in F \}$$

- Clearly $v_i = 1.v_i = 0.v_1 + 0.v_2 + \cdots + 0.v_{i-1} + 1.v_i + 0.v_{i+1} + \cdots + 0.v_p$

  so, $v_i \in \text{Span } S$ for $i = 1, 2, \ldots, n$

Thus Span S is a subset of V

- Span S is a subspace of V.

Proof: (1) $\bar{0} = 0v_1 + 0.v_2 + \cdots + 0.v_p \in \text{Span}\, S$

(2) Let $w_1, w_2 \in \text{Span}\, S$

So, there exist scalars $c_1, c_2, \ldots, c_p \in F$

such that $w_1 = c_1 v_1 + c_2 v_2 + \cdots + c_p v_p$

and there exist scalars $d_1, d_2, \ldots, d_p \in F$

such that $w_2 = d_1 v_1 + d_2 v_2 + \cdots + d_p v_p$

Now $w_1 + w_2 = (c_1 v_1 + c_2 v_2 + \cdots + c_p v_p) + (d_1 v_1 + d_2 v_2 + \cdots + d_p v_p)$

$= c_1 v_1 + d_1 v_1 + c_2 v_2 + d_2 v_2 + \cdots + c_p v_p + d_p v_p$

$= (c_1 + d_1) v_1 + (c_2 + d_2) v_2 + \cdots + (c_p + d_p) v_p$

Since $c_1 + d_1, c_2 + d_2, \ldots, c_p + d_p \in F$

So, $w_1 + w_2 \in \text{Span}\, S$

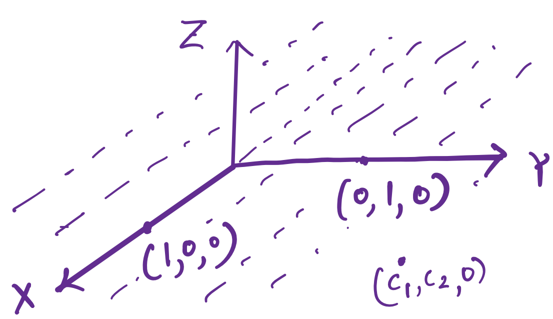
Let $u \in \text{Span}\, S$ and $c \in F$, There exists scalars $c_1, c_2, \ldots, c_p \in F$ such that $u = c_1 v_1 + c_2 v_2 + \cdots + c_p v_p \Rightarrow cu = c(c_1 v_1 + c_2 v_2 + \cdots + c_p v_p)$

$= (c c_1) v_1 + (c c_2) v_2 + \cdots + (c c_p) v_p \in \text{span}\, S$

Hence Span S is a subspace of V.

Ex: Let $V = \mathbb{R}^3$

Let $S = \{(1,0,0), (0,1,0)\}$

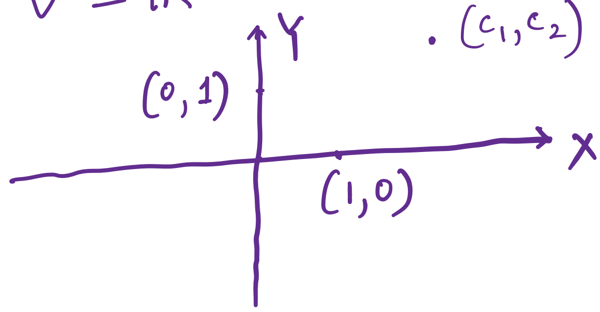


Since for any $c_1, c_2 \in \mathbb{R}$,

$$c_1(1,0,0) + c_2(0,1,0) = (c_1, c_2, 0)$$

$\text{Span } S = \{(x,y,0) : x, y \in \mathbb{R}\}$ is the XY plane.

Ex: Let $V = \mathbb{R}^2$



Let $S = \{(1,0), (0,1)\}$

Now for $c_1, c_2 \in \mathbb{R}$

$c_1(1,0) + c_2(0,1) = (c_1, c_2)$

Hence $\text{Span } S = \mathbb{R}^2$

Ex: Suppose W_1 and W_2 are two subspaces of a vector space V over a field F.

Prove that $W_1 \cap W_2$ is a subspace of V

Proof: (1) $\bar{0} \in W_1$, $\bar{0} \in W_2$ (Since W_1 & W_2 are subspaces of V)

$\Rightarrow \bar{0} \in W_1 \cap W_2$

(2) $u, v \in W_1 \cap W_2 \Rightarrow u, v \in W_1$ and $u, v \in W_2$

$\Rightarrow u+v \in W_1$ (W_1 is a subspace of V)

$\Rightarrow u+v \in W_2$ (W_2 is a subspace of V)

$\Rightarrow u+v \in W_1 \cap W_2$

(3) Let $c \in F$, $u \in W_1 \cap W_2$

Then $u \in W_1$ and $u \in W_2 \Rightarrow cu \in W_1$ and $cu \in W_2 \Rightarrow cu \in W_1 \cap W_2$

Hence $W_1 \cap W_2$ is a subspace of V.

Note: In the same way we can show that intersection of any family of subspaces is a subspace of V.

Note: $W_1 \cup W_2$ may not be a subspace of V.

Example: Let $V = \mathbb{R}^2$

Let $W_1 = \{(x, 0) : x \in \mathbb{R}\}$

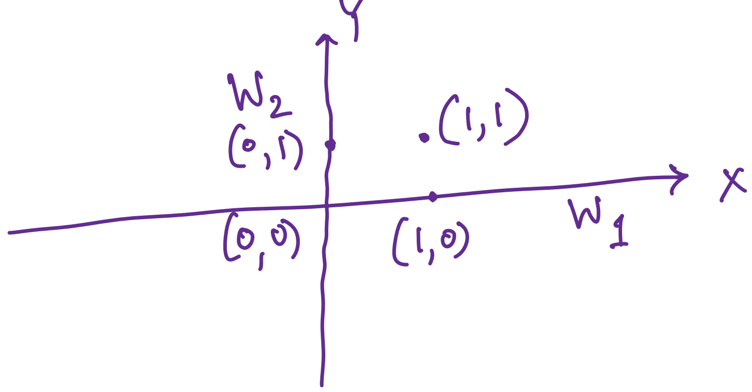
$W_2 = \{(0, y) : y \in \mathbb{R}\}$

Now, $(1,0) \in W_1$, $(0,1) \in W_2$

So, $(1,0) \in W_1 \cup W_2$, $(0,1) \in W_1 \cup W_2$

but $(1,0) + (0,1) = (1,1) \notin W_1 \cup W_2$

So, $W_1 \cup W_2$ is not a subspace of $V = \mathbb{R}^2$



Remarks:

(1) Span S is the smallest subspace of V containing S.

$$S \subset \underset{\text{(subspace)}}{\mathrm{Span}(S)} \subset \dots \subset V$$

S: (subset not necessarily a subspace)

Clearly $S \subset \mathrm{Span}(S)$

If $S = \{v_1, v_2, \dots, v_p\}$ then $v_i = 1 . v_i$ for $i = 1, 2, \dots, p$ and so $v_i \in \mathrm{Span}(S)$

Also if W is a subspace of V such that $S \subset W$ then $\mathrm{span}\, S \subset W$

Proof: Let $u \in \mathrm{Span}\, S$.

Since $S = \{v_1, v_2, \dots, v_p\}$, there exist scalars $c_1, c_2, \dots, c_p \in F$ such that $u = c_1 v_1 + c_2 v_2 + \dots + c_p v_p$

Now $v_1, v_2, \dots, v_p \in S$ and $S \subset W \Rightarrow v_1, v_2, \dots, v_p \in W$

$\Rightarrow c_1 v_1, c_2 v_2, \dots, c_p v_p \in W$ (Since W is a subspace of V)

$\Rightarrow c_1 v_1 + c_2 v_2 + \dots + c_p v_p \in W$ (Since W is a subspace of V)

$\Rightarrow u \in W$

Therefore $\mathrm{span}\, S \subset W$

(2) Span S is the intersection of all subspaces of V containing S. $(S \subset \mathrm{Span}\, S \subset \dots \subset V)$

Proof: Let B be the intersection of all Subspaces of V containing S

Since we have shown that Span S is a subspace of V containing S, $\boxed{B \subset \text{Span} S}$ ...... (α)

On the other hand, B is also a subspace of V and $S \subset B$,

by Remark(1), $\boxed{\text{Span } S \subset B}$ ...... (β)

Combining (α) and (β), we obtain

Span S = B = intersection of all subspaces of V containing S.

<u>Ex:</u> Let $S = \{v_1, v_2, v_3\} \subset \mathbb{R}^3$

where $v_1 = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 6 \\ 12 \\ 3 \end{bmatrix}$ and $v_3 = \begin{bmatrix} 3 \\ 25 \\ 9 \end{bmatrix}$

Let $d = \begin{bmatrix} 4 \\ 46 \\ 17 \end{bmatrix}$

<u>Question:</u> Is d in the span $\{v_1, v_2, v_3\}$?

Let us solve: $c_1 v_1 + c_2 v_2 + c_3 v_3 = d$ (c_1, c_2, c_2 are the unknown scalars)

$$\Rightarrow c_1 \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 6 \\ 12 \\ 3 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ 25 \\ 9 \end{bmatrix} = \begin{bmatrix} 4 \\ 46 \\ 17 \end{bmatrix}$$

<u>The Augmented matrix:</u>

$$[A : d] = \left[\begin{array}{ccc|c} 2 & 6 & 3 & 4 \\ 4 & 12 & 25 & 46 \\ 1 & 3 & 9 & 17 \end{array}\right] \xrightarrow[R_1 \leftrightarrow R_3]{} \left[\begin{array}{ccc|c} 1 & 3 & 9 & 17 \\ 4 & 12 & 25 & 46 \\ 2 & 6 & 3 & 4 \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 - 4R_1} \left[\begin{array}{ccc|c} 1 & 3 & 9 & 17 \\ 0 & 0 & -11 & -22 \\ 0 & 0 & -15 & -30 \end{array}\right]$$

$$\xrightarrow[R_3 \to (-\frac{1}{15} R_3)]{R_2 \to (-\frac{1}{11} R_2)} \left[\begin{array}{ccc|c} 1 & 3 & 9 & 17 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

$\Big\downarrow R_3 \to R_3 - R_2$

$$\left[\begin{array}{ccc|c} 1 & 3 & 9 & 17 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow[R_1 \to R_1 - 9R_2]{} \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] = \text{RREF matrix}$$

Since the last column is not a pivot column, the system of equations is consistent.

<u>Solving the system:</u>

$$\left.\begin{array}{l} c_1 + 3c_2 = -1 \\ c_3 = 2 \end{array}\right\} \Rightarrow \left.\begin{array}{l} c_1 = -1 - 3c_2 \\ c_3 = 2 \end{array}\right\}$$

There are infinitely many solutions.

One solution is $\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$
(By taking $c_2 = 0$)

So, $(-1)\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 6 \\ 12 \\ 3 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 25 \\ 9 \end{bmatrix} = \begin{bmatrix} 4 \\ 46 \\ 17 \end{bmatrix} = d$

Hence $d \in \text{span}\{v_1, v_2, v_3\}$

<u>Ex:</u> Let $d_1 = \begin{bmatrix} 4 \\ 46 \\ 18 \end{bmatrix} \in \mathbb{R}^3$

<u>Question:</u> Is $d_1 \in \text{span}\{v_1, v_2, v_3\}$ where v_1, v_2, v_3 are given in the previous example?

We perform the same sequence of row operations on d_1.

$$d_1 = \begin{bmatrix} 4 \\ 46 \\ 18 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_3]{} \begin{bmatrix} 18 \\ 46 \\ 4 \end{bmatrix} \xrightarrow[\substack{R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 2R_1}]{} \begin{bmatrix} 18 \\ -26 \\ -32 \end{bmatrix} \xrightarrow[\substack{R_2 \to (-\frac{1}{11}R_2) \\ R_3 \to (-\frac{1}{15}R_3)}]{} \begin{bmatrix} 18 \\ \frac{26}{11} \\ \frac{32}{15} \end{bmatrix}$$

$\Big|$

$$\xrightarrow{R_3 \to R_3 - R_2} \begin{bmatrix} 18 \\ \frac{26}{11} \\ -\frac{38}{165} \end{bmatrix} \xrightarrow{R_1 - 9R_2} \begin{bmatrix} -\frac{36}{11} \\ \frac{26}{11} \\ -\frac{38}{165} \end{bmatrix}$$

So, the RREF matrix corresponding to the Augmented matrix $[A : d_1]$ is

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & -\frac{36}{11} \\ 0 & 0 & 1 & \frac{26}{11} \\ 0 & 0 & 0 & -\frac{38}{165} \end{array}\right]$$

Since the last column is a pivot column (there is a row $[0, 0, 0, -\frac{38}{165}]$), the system of equations is inconsistent.

Hence $d_1 \notin \text{Span}\{v_1, v_2, v_3\}$

MTH 100 : Lecture 17

__Ex:__ Let $u_1 = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}$, $u_2 = \begin{bmatrix} 6 \\ 12 \\ 4 \end{bmatrix}$, $u_3 = \begin{bmatrix} 3 \\ 24 \\ 9 \end{bmatrix}$

Show that $\text{Span}\{u_1, u_2, u_3\} = \mathbb{R}^3$.

Let $A = [u_1 \ u_2 \ u_3] = \begin{bmatrix} 2 & 6 & 3 \\ 4 & 12 & 24 \\ 1 & 4 & 9 \end{bmatrix}$

Let us row reduce A

$$\begin{bmatrix} 2 & 6 & 3 \\ 4 & 12 & 24 \\ 1 & 4 & 9 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 4 & 9 \\ 4 & 12 & 24 \\ 2 & 6 & 3 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 2R_1]{R_2 \to R_2 - 4R_1} \begin{bmatrix} 1 & 4 & 9 \\ 0 & -4 & -12 \\ 0 & -2 & -15 \end{bmatrix}$$

$$\xrightarrow{R_2 \to \left(-\frac{1}{4}R_2\right)} \begin{bmatrix} 1 & 4 & 9 \\ 0 & 1 & 3 \\ 0 & -2 & -15 \end{bmatrix} \xrightarrow{R_3 \to R_3 + 2R_2} \begin{bmatrix} 1 & 4 & 9 \\ 0 & 1 & 3 \\ 0 & 0 & -9 \end{bmatrix} \xrightarrow{R_3 \to \left(-\frac{1}{9}R_3\right)} \begin{bmatrix} 1 & 4 & 9 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_2 \to R_2 - 3R_3]{R_1 \to R_1 - 9R_3} \begin{bmatrix} 1 & 4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 - 4R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

So, A is row equivalent to the identity matrix I_3.

Hence $A\bar{x} = \bar{b}$ has a solution for every $b \in \mathbb{R}^3$.

Therefore any vector $\bar{b} \in \mathbb{R}^3$ can be written as a linear combination of the columns of A.
(Viz. u_1, u_2, u_3)

$$\Rightarrow \quad b \in \text{span}\{u_1, u_2, u_3\}$$

Hence $\boxed{\text{Span}\{u_1, u_2, u_3\} = \mathbb{R}^3}$

<u>Note:</u>

$$A\bar{x} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_1 u_1 + x_2 u_2 + x_3 u_3$$

is a linear Combination of u_1, u_2 and u_3 where the scalars $x_1, x_2, x_3 \in \mathbb{R}$

- Linear independence and dependence:

Definition: Let $v_1, v_2, \ldots, v_p$ be a finite list of vectors in a vector space V over a field F. Then the vectors are said to be linearly dependent if there exist scalars $c_1, c_2, \ldots, c_p \in F$ not all zero such that $c_1 v_1 + c_2 v_2 + \cdots + c_p v_p = \bar{0}$

Definition: If a list of vectors is not linearly dependent, they are called linearly independent.

Thus if $\{v_1, v_2, \ldots, v_p\}$ is linearly independent and $c_1 v_1 + c_2 v_2 + \cdots + c_p v_p = \bar{0}$ then $c_1 = c_2 = \cdots = c_p = 0$

Ex: Consider the following elements of $\mathbb{R}^{2\times 2}$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Question: Are A, B, C linearly independent ?

Let $c_1 A + c_2 B + c_3 C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ where $c_1, c_2, c_3 \in \mathbb{R}$

$$c_1 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} c_1 + c_2 + c_3 & c_1 + c_3 \\ c_1 & c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \left.\begin{array}{lll} c_1 + c_2 + c_3 = 0 & & \\ c_1 = 0 & \Rightarrow & c_1 = 0 \\ c_1 + c_3 = 0 & \Rightarrow & 0 + c_3 = 0 \Rightarrow c_3 = 0 \\ c_1 + c_2 = 0 & \Rightarrow & 0 + c_2 = 0 \Rightarrow c_2 = 0 \end{array}\right\}$$

So, $c_1 A + c_2 B + c_3 C = \bar{0}$ (zero matrix)

$$\Rightarrow c_1 = c_2 = c_3 = 0$$

Hence A, B are C are linearly independent.

Ex: Let $V = C[0, 2\pi]$ (This is a vector space over $\mathbb{R}$)

Let $f_1(x) = 1$, $f_2(x) = \sin x$, $f_3(x) = \sin(2x)$

Question: Are f_1, f_2 and f_3 linearly independent?

Let $c_1 f_1 + c_2 f_2 + c_3 f_3 = \bar{0}(x)$ (The zero function of $C[0, 2\pi]$)

$(\bar{0}(x) = 0 \ \forall \ x \in [0, 2\pi])$

$$\Rightarrow \quad c_1(1) + c_2 \sin x + c_3 \sin 2x = \bar{0}(x) = 0 \ \forall \ x \in [0, 2\pi]$$

Let $x = 0$: Then $c_1 + c_2 \sin(0) + c_3 \sin(2 \times 0) = 0$

$$\Rightarrow c_1 + c_2 \times 0 + c_3 \times 0 = 0 \Rightarrow \boxed{c_1 = 0}$$

Let $x = \frac{\pi}{2}$: Then $c_2 \sin\left(\frac{\pi}{2}\right) + c_3 \sin\left(2 \times \frac{\pi}{2}\right) = 0$

$$\Rightarrow c_2 \times 1 + c_3 \times 0 = 0$$

$$\Rightarrow \boxed{c_2 = 0}$$

Let $x = \frac{\pi}{4}$: $c_3 \sin\left(2 \times \frac{\pi}{4}\right) = 0 \Rightarrow c_3 \times 1 = 0$

$$\Rightarrow \boxed{c_3 = 0}$$

Thus $c_1 f_1 + c_2 f_2 + c_3 f_3 = \bar{0}(x) \ \forall \ x \in [0, 2\pi] \Rightarrow c_1 = c_2 = c_3 = 0$

Hence f_1, f_2 and f_3 are linearly independent.

Note: We can use other points in $[0, 2\pi]$ to solve for the scalars c_1, c_2 and c_3.

Remark 1: Any list which contains the zero vector has to be linearly dependent.

Suppose $v_1, v_2, \ldots, v_p$ is a list of vectors such that $v_i = \bar{0}$ for some $1 \leq i \leq p$

Now $0.v_1 + 0.v_2 + \cdots + 0.v_{i-1} + 1.v_i + 0.v_{i+1} + \cdots + 0.v_p = v_i = \bar{0}$

So, we have the above linear combination of $v_1, v_2, \ldots, v_p$ to be a zero vector where one of the scalar is $1 \neq 0$

Hence $v_1, v_2, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_p$ is linearly dependent.

Remark 2: A single non zero vector is linearly independent.

Suppose $v \neq \bar{0}$ and $c.v = \bar{0}$ where the scalar $c \in F$

Then $c = 0$

(Suppose $c \neq 0$. Then $c^{-1} \in F$ and $c^{-1}(cv) = c^{-1}(\bar{0})$

$\Rightarrow (c^{-1}c)\, v = \bar{0}$

$\Rightarrow 1.v = \bar{0} \Rightarrow v = \bar{0}$, a contradiction)

Remark 3: A list of two non zero vector is linearly dependent only if one of the vectors is a scalar multiple of the other.

Suppose two non zero vectors v_1 and v_2 are linearly dependent. Then there exists scalars $c_1, c_2 \in F$ (atleast one of them is non zero) such that

$$c_1 v_1 + c_2 v_2 = \bar{0}$$

Without any loss of generality (WLOG), we assume that $c_1 \neq 0$

Then $c_1 v_1 + c_2 v_2 = \bar{0} \Rightarrow c_1 v_1 = -c_2 v_2 \Rightarrow c_1^{-1}(c_1 v_1) = c_1^{-1}(-c_2 v_2)$ (Since $c_1 \neq 0$, $c_1^{-1} \in F$)

$\Rightarrow (c_1^{-1}c_1)\, v_1 = -(c_1^{-1}c_2)\, v_2 \Rightarrow 1.v_1 = (-c_1^{-1}c_2)\, v_2$

$\Rightarrow v_1 = (-c_1^{-1}c_2)\, v_2$. Thus v_1 is a scalar multiple of v_2.

MTH 100 : Lecture 18

Remark 4: A list of non-zero vectors is linearly dependent if and only if atleast one of the vectors is a linear combination of the others.

Proof of Remark 4:

$\Rightarrow$: **given:** A list of vectors $v_1, v_2, \ldots, v_p$ is linearly dependent.

To show that: At least one of the vector is a linear combination of the others.

Since $v_1, v_2, \ldots, v_p$ are linearly dependent, there exist scalars $c_1, c_2, \ldots, c_p \in F$ not all zero such that $c_1 v_1 + c_2 v_2 + \cdots + c_p v_p = \bar{0}$

Let us assume $c_k \neq 0$ where $1 \leq k \leq p$

$$c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1} + c_k v_k + c_{k+1} v_{k+1} + \cdots + c_p v_p = \bar{0}$$

$$\Rightarrow c_k v_k = -c_1 v_1 - c_2 v_2 - \cdots - c_{k-1} v_{k-1} - c_{k+1} v_{k+1} - \cdots - c_p v_p$$

Now $c_k \neq 0$ and $c_k \in F$; so, $c_k^{-1} \in F$

Then $c_k^{-1}(c_k v_k) = -c_k^{-1}(c_1 v_1) - c_k^{-1}(c_2 v_2) - \cdots - c_k^{-1}(c_{k-1} v_{k-1})$
$- c_k^{-1}(c_{k+1} v_{k+1}) - \cdots - c_k^{-1}(c_p v_p)$

$\Rightarrow (c_k^{-1} c_k) v_k = (-c_k^{-1} c_1) v_1 - (c_k^{-1} c_2) v_2 - \cdots - (c_k^{-1} c_{k-1}) v_{k-1}$
$- (c_k^{-1} c_{k+1}) v_{k+1} - \cdots - (c_k^{-1} c_p) v_p$

$\Rightarrow v_k = (-c_k^{-1} c_1) v_1 + (-c_k^{-1} c_2) v_2 + \cdots + (-c_k^{-1} c_{k-1}) v_{k-1}$
$+ (-c_k^{-1} c_{k+1}) v_{k+1} + \cdots + (-c_k^{-1} c_p) v_p$

where $(-c_k^{-1} c_1), (-c_k^{-1} c_2), \ldots, (-c_k^{-1} c_{k-1}), (-c_k^{-1} c_{k+1}) \cdots (-c_k^{-1} c_p) \in F$

$\Leftarrow$: Given: Atleast one of the vector is a linear combination of the rest of the vectors

To show: The list is linearly dependent.

Let us assume that v_k is a linear combination of the rest of the vectors

Then there exist scalars $c_1, c_2, \ldots, c_{k-1}, c_{k+1}, \ldots c_p \in F$

such that

$$v_k = c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1} + c_{k+1} v_{k+1} + \cdots + c_p v_p$$

$$\Rightarrow c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1} - v_k + c_{k+1} v_{k+1} + \cdots + c_p v_p = \bar{0}$$

$$\Rightarrow c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1} + (-1) v_k + c_{k+1} v_{k+1} + \cdots + c_p v_p = \bar{0}$$

where atleast one of the scalar is $(-1) \neq 0$

So, $v_1, v_2, \ldots, v_{k-1}, v_k, v_{k+1}, \ldots, v_p$ are linearly dependent.

Remark 5: Consequently any list of vectors which contains a repeated vector must be linearly dependent. A list which is linearly independent corresponds to a set.

Remark 6: Any list which contains a linearly dependent list is linearly dependent.

Remark 7: Any subset of a linearly independent set is linearly independent.

Proof of Remark 6:

Suppose $v_1, v_2, \ldots, v_p$ is a list of vectors that contains a list

$v_1, v_2, \ldots, v_k$ (where $1 < k \leq p$), which is linearly dependent.

(WLOG, can assume that the linearly dependent vectors are at the beginning of the list.)

Then there exist scalars $c_1, c_2, \ldots, c_k \in F$ (not all of them zeros)

such that $c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = \overline{0}$

Then $c_1 v_1 + c_2 v_2 + \cdots + c_k v_k + 0.v_{k+1} + \cdots + 0.v_p = \overline{0}$

Hence $v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_p$ are linearly dependent.

Note: Proofs of Remark 5 and Remark 7 are left as exercises.

Ex: Let $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $v_2 = \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}$, $v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Question: Are v_1, v_2, v_3 linearly independent?

Consider $c_1 v_1 + c_2 v_2 + c_3 v_3 = \bar{0}$ where $c_1, c_2, c_3 \in \mathbb{R}$

$$\Rightarrow c_1 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \left.\begin{aligned} c_1 + 5c_2 + c_3 &= 0 \\ 2c_1 + 6c_2 + c_3 &= 0 \\ 3c_1 + 7c_2 + c_3 &= 0 \\ 4c_1 + 8c_2 + c_3 &= 0 \end{aligned}\right\}$$

The coefficient matrix:

$$\begin{bmatrix} 1 & 5 & 1 \\ 2 & 6 & 1 \\ 3 & 7 & 1 \\ 4 & 8 & 1 \end{bmatrix} \xrightarrow[\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 3R_1 \\ R_4 \to R_4 - 4R_1}]{} \begin{bmatrix} 1 & 5 & 1 \\ 0 & -4 & -1 \\ 0 & -8 & -2 \\ 0 & -12 & -3 \end{bmatrix}$$

$$\xrightarrow[\substack{R_3 \to R_3 - 2R_2 \\ R_4 \to R_4 - 3R_2}]{} \begin{bmatrix} 1 & 5 & 1 \\ 0 & -4 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow[R_2 \to \left(-\frac{1}{4}R_2\right)]{} \begin{bmatrix} 1 & 5 & 1 \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\downarrow R_1 \to R_1 - 5R_2$$

$$\begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$ (Here c_3 is a free variable).

So, the solution is $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = c_3 \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ 1 \end{bmatrix}$ where c_3 is any real number.

If $c_3 = 4$, one possible solution is $\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$

Therefore $\boxed{v_1 - v_2 + 4v_3 = 0}$

Hence v_1, v_2 and v_3 are linearly dependent.

Ex: Let $V = C[0, 2\pi]$

Let $f(x) = 1$, $g(x) = \sin x$, $h(x) = \cos x$

Question: Are f, g, h linearly independent?

$$c_1 f + c_2 g + c_3 h = \bar{0}(x)$$

$$\Rightarrow \quad c_1 \times 1 + c_2 \sin x + c_3 \cos x = 0 \quad \forall\, x \in [0, 2\pi]$$

If $x = 0 \Rightarrow c_1 + 0 + c_3 \times 1 = 0 \Rightarrow c_1 + c_3 = 0$

If $x = \frac{\pi}{2} \Rightarrow c_1 + c_2 \times 1 + 0 = 0 \Rightarrow c_1 + c_2 = 0$

If $x = \pi \Rightarrow c_1 + 0 - c_3 = 0 \Rightarrow c_1 - c_3 = 0$

The coefficient matrix is

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \xrightarrow[R_3 \to R_3 - R_1]{R_2 \to R_2 - R_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \end{bmatrix} \xrightarrow{R_3 \to (-\frac{1}{2} R_3)} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_2 \to R_2 + R_3]{R_1 \to R_1 - R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \text{RREF matrix}$$

So, the system has only the trivial solution $c_1 = 0$, $c_2 = 0$ and $c_3 = 0$

Hence f, g and h are linearly independent.

<u>Ex:</u> Let $u_1 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$, $u_2 = \begin{bmatrix} 6 \\ 12 \\ 4 \end{bmatrix}$, $u_3 = \begin{bmatrix} 3 \\ 24 \\ 9 \end{bmatrix}$

Let $A = [u_1 \; u_2 \; u_3]_{3\times 3}$

$$A = \begin{bmatrix} 2 & 6 & 3 \\ 4 & 12 & 24 \\ 6 & 4 & 9 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 3R_1]{R_2 \to R_2 - 2R_1} \begin{bmatrix} 2 & 6 & 3 \\ 0 & 0 & 18 \\ 0 & -14 & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 2 & 6 & 3 \\ 0 & -14 & 0 \\ 0 & 0 & 18 \end{bmatrix} \xrightarrow[\substack{R_2 \to \left(-\frac{1}{14} R_2\right) \\ R_3 \to \left(\frac{1}{18} R_3\right)}]{R_1 \to \left(\frac{1}{2} R_1\right)} \begin{bmatrix} 1 & 3 & \frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 - \frac{3}{2} R_3} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 - 3R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \text{RREF matrix}$$

Therefore $c_1 u_1 + c_2 u_2 + c_3 u_3 = 0 \Rightarrow c_1 = c_2 = c_3 = 0$

So, u_1, u_2 and u_3 are linearly independent.

Furthermore for any vector $\bar{b} \in \mathbb{R}^3$, the equation $A\bar{x} = \bar{b}$ has a unique solution and so b can be written as a linear combination of u_1, u_2 and u_3.

Thus $\bar{b} \in \text{Span}\{u_1, u_2, u_3\} \ \forall \bar{b} \in \mathbb{R}^3$

Hence $\boxed{\text{Span}\{u_1, u_2, u_3\} = \mathbb{R}^3}$

Basis and Dimension:

Definition: Let V be a vector space over a field F.

A Basis for V is a linearly independent set S of vectors such that $\boxed{V = \text{Span } S}$

Ex: In the previous example, $\{u_1, u_2, u_3\}$ is a basis of $\mathbb{R}^3$.

Ex: Consider the vector space $\mathbb{R}^n$ over the field $\mathbb{R}$.

Consider the vectors $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, ..., $e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$

- $e_1, e_2, \ldots, e_n$ are linearly independent.

$$c_1 e_1 + c_2 e_2 + \cdots + c_n e_n = \overline{0}$$

$$\Rightarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \Rightarrow \quad c_1 = c_2 = \cdots = c_n = 0$$

- Any vector $\overline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ can be written as

$$\overline{x} = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

Conclusion: $\{e_1, e_2, \ldots, e_n\}$ forms a basis of $\mathbb{R}^n$

Note: In one of the examples considered before we have shown that $\{u_1, u_2, u_3\}$ is a basis of $\mathbb{R}^3$.

Applying the last example for $n=3$, we can also say that $\{e_1, e_2, e_3\}$ is another basis of $\mathbb{R}^3$.

(Note: Plural of Basis = Bases)

Definition: • A vector space which has a finite basis is called Finite dimensional.

• A vector space which doesn't have a finite basis is called infinite dimensional.

Note: The above definition would also apply to subspaces of V.

Ex: $\mathbb{R}^n$ is a finite dimensional vector space.

Example of an infinite dimensional vector space:

Let $R[t]$ be the vector space of all polynomials (in t) with real coefficients.

$R[t]$ is a vector space over $\mathbb{R}$.

We will show that $R[t]$ is infinite dimensional.

Suppose BWOC that $R[t]$ is finite dimensional.

Then it must have a finite basis,

say $B = \{ p_1(t), p_2(t), \dots, p_n(t) \}$

Let $N = \max \{ \deg p_1(t), \deg p_2(t), \dots, \deg p_n(t) \}$

where $\deg p_k(t)$ = degree of $p_k(t)$ (the k-th polynomial in B)

Let $p(t) = t^{N+1}$

Then $p(t) \notin \operatorname{Span} B$, a contradiction.

So, $R[t]$ is infinite dimensional.

MTH 100 : Lecture 19

Last time:

We defined Basis for a vector space and finite & infinite dimensional vector space.

Recall that: A basis for a vector space V is a linearly independent set S of vectors such that $\text{span}(S) = V$

Alternative Definition for Basis:

Proposition: $B = \{v_1, v_2, \ldots, v_n\}$ is a basis of the vector space V if and only if every vector $v \in V$ is uniquely expressible as a linear combination of the elements of B.

Note: In some books, the above is used as the definition of a Basis, and then it is shown that a Basis is a linearly independent spanning set.

Proof: exercise

Proposition (Steinitz Exchange Lemma):

Suppose $v_1, v_2, \dots, v_n$ are linearly independent vectors in a vector space V and suppose $V = \text{Span}\{w_1, w_2, \dots, w_m\}$

Then (a) $n \leq m$

(b) $\{v_1, v_2, \dots, v_n, w_{n+1}, w_{n+2}, \dots, w_m\}$ span V after reordering the w's if necessary.

Proposition: If V is a finite dimensional vector space, then any two bases of V have the same number of elements.

Proof: Let B_1 and B_2 be two bases of V with k_1 and k_2 vectors respectively.

Want to show: $k_1 = k_2$

Now B_1 is a linearly independent set of vectors in V and B_2 is a spanning set of V.

So, by Steinitz exchange lemma, $k_1 \leq k_2$

Now B_2 is a linearly independent set of vectors in V and B_1 is a spanning set of V.

So, by Steinitz exchange lemma, $k_1 \leq k_2$. Hence $\boxed{k_1 = k_2}$

Definition: The dimension of a finite dimensional vector space is the number of elements in a basis for V. This is written as $\boxed{\dim(V)}$

Note: The above proposition ensures that this is a proper definition.

Example: $\boxed{\dim(\mathbb{R}^n) = n}$ (Recall: $e_1, e_2, \dots, e_n$ is a basis of $\mathbb{R}^n$)

Special case: The dimension of the zero subspace of any vector space is taken as zero.
(It doesn't have a basis)

Proof of Steinitz Exchange Lemma

Given: V is a vector space
$v_1, v_2, \dots, v_n$ is a linearly independent set of vectors
and $V = \text{Span}\{w_1, w_2, \dots, w_m\}$

Step I: Since $v_1 \in V$, we can write

$$v_1 = c_1 w_1 + c_2 w_2 + \cdots + c_m w_m \quad \dots\dots (1)$$

where $c_1, c_2, \dots, c_m \in F$

If $c_i = 0 \ \forall i$, then $v_1 = 0$ which is not possible since $v_1, v_2, \ldots, v_n$ are linearly independent.

So, $c_i \neq 0$ for atleast one i

Renumbering if necessary, we can assume $c_1 \neq 0$.

Then (1) $\Rightarrow c_1 w_1 = v_1 - c_2 w_2 - \cdots - c_m w_m$

$\Rightarrow \quad c_1^{-1} c_1 w_1 = c_1^{-1} v_1 - c_1^{-1} c_2 w_2 - \cdots - c_1^{-1} c_m w_m$

$\Rightarrow \quad w_1 = d_1 v_1 + d_2 w_2 + \cdots + d_m w_m \quad \cdots\cdots$ (2)

where $d_1, d_2, \ldots, d_m$ are scalars.

From here we can conclude

$$\text{Span}\{v_1, w_2, w_3, \ldots, w_m\} = \text{span}\{w_1, w_2, \ldots, w_m\} = V$$

Let $x \in V$.

Then $x = f_1 w_1 + f_2 w_2 + \cdots + f_m w_m$ for scalars $f_1, f_2, \ldots, f_m \in F$.

$$\begin{aligned} &= f_1 (d_1 v_1 + d_2 w_2 + \cdots + d_m w_m) + f_2 w_2 + \cdots + f_m w_m \\ &= f_1 d_1 v_1 + (f_1 d_2 + f_2) w_2 + \cdots + (f_1 d_m + f_m) w_m \\ &= h_1 v_1 + h_2 w_2 + \cdots + h_m w_m \\ &\in \text{span}\{v_1, w_2, \ldots, w_m\} \end{aligned}$$

$$\text{Span}\{v_1, w_2, \ldots, w_m\} = V$$

Step II: Since $v_2 \in V = \text{Span}\{v_1, w_2, \ldots, w_m\}$, we can write $v_2 = l_1 v_1 + l_2 w_2 + \cdots + l_m w_m$

where $l_1, l_2, \ldots, l_m \in F$

Now atleast one of $l_2, l_3, \dots, l_m$ is nonzero.
Otherwise $v_2 = l_1 v_1$ that contradicts the fact that $v_1, v_2, \dots, v_n$ are linearly independent.
Renumbering if necessary, we can assume that $l_2 \neq 0$

Then $l_2 w_2 = -l_1 v_1 + v_2 - l_3 w_3 - \dots - l_m w_m$

$$\Rightarrow w_2 = -l_2^{-1} l_1 v_1 + l_2^{-1} v_2 - l_2^{-1} l_3 w_3 - \dots l_2^{-1} l_m w_m$$

Proceeding as before we can conclude

$$\text{Span}\{v_1, v_2, w_3, \dots, w_m\} = \text{Span}\{v_1, w_2, w_3, \dots, w_m\}$$
$$= \text{Span}\{w_1, w_2, w_3, \dots, w_m\} = V$$

This process will stop after the nth step atmost (Since there are only n vectors $v_1, v_2, \dots, v_n$)

Now we can think of two situations.

<u>Case1:</u> $n \leq m$

Then we are in the following situation:

$$\{\overset{v_1}{\overset{\downarrow}{w_1}}, \overset{v_2}{\overset{\downarrow}{w_2}}, \dots, \overset{v_n}{\overset{\downarrow}{w_n}}, w_{n+1}, \dots w_m\}$$

We have replaced n of the w-vectors and renumbering if necessary we get

$$\text{span}\{v_1, v_2, \dots, v_n, w_{n+1}, \dots, w_m\} = V$$

In this case we proved the lemma.

<u>Note:</u> If $n = m$, then the vectors $w_{n+1}, \dots$ etc. are not there in the original spanning set.

Case 2: $n > m$

Then we are in the following situation:

$$\begin{array}{ccccc} v_1 & v_2 & \cdots & v_m, & v_{m+1}, \cdots, v_n \\ \downarrow & \downarrow & & \downarrow & \\ \{w_1, & w_2, & \cdots, & w_m\} & \end{array}$$

Now, $\{v_1, v_2, \cdots, v_m\}$ is a spanning set for V.

Then $v_{m+1} \in \operatorname{Span}\{v_1, v_2, \cdots, v_m\}$

i.e. $v_{m+1} = p_1 v_1 + p_2 v_2 + \cdots + p_m v_m$

where $p_1, p_2, \cdots, p_m$ are scalars.

But this contradicts the linear independence of $\{v_1, v_2, \cdots, v_m, v_{m+1}, \cdots, v_n\}$.

Thus Case 2 cannot happen and in this case $n \leq m$ and the lemma is proved.

$\boxed{\text{MTH 100 : Lecture 20}}$

How to create Bases:

Proposition: Suppose $S = \{v_1, v_2, \dots, v_n\}$ is a linearly independent set in a vector space V. Suppose v is a vector which is not in the span S. Then the set obtained by adjoining v to S is linearly independent.

Proof: Need to show that

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n + c v = \bar{0} \quad \cdots\cdots (1)$$

$$\Rightarrow c_1 = c_2 = \cdots\cdots = c_n = c = 0$$

Suppose $c \neq 0$. Then $c^{-1} \in F$ (where F is the field of scalars.)

Then (1) $\Rightarrow cv = -c_1 v_1 - c_2 v_2 - \cdots\cdots - c_n v_n$

$$\Rightarrow v = -c^{-1} c_1 v_1 - c^{-1} c_2 v_2 - \cdots\cdots - c^{-1} c_n v_n$$

$$\Rightarrow v = (-c^{-1} c_1) v_1 + (-c^{-1} c_2) v_2 + \cdots\cdots + (-c^{-1} c_n) v_n$$

$\Rightarrow v \in \text{Span}\{v_1, v_2, \dots, v_n\}$ which contradicts the assumption that $v \notin \text{Span}\{v_1, v_2, \dots, v_n\}$

So, $c = 0$. Then (1) $\Rightarrow c_1 v_1 + c_2 v_2 + \cdots\cdots + c_n v_n = \bar{0}$

$\Rightarrow c_1 = c_2 = \cdots = c_n = 0$ (Since $v_1, v_2, \dots, v_n$ are linearly independent)

Hence $v_1, v_2, \dots, v_n, v$ are linearly independent.

Proposition: Any linearly independent set S in a finite dimensional vector space can be expanded to a basis.

Proof: Exercise

Hint: Use the previous proposition repeatedly. By Steinitz Exchange Lemma, the process has to stop and at that stage, a basis is obtained.

Proposition: Any finite spanning set S in a nonzero vector space can be contracted to a basis.

Proof: Exercise

Note: If a non-zero vector space V has a finite spanning set S, then it must be finite dimensional.

Summary:

Proposition: Let V be a non-zero finite dimensional vector space with dimension n.

Then,

- Any linear independent set of vectors must have $\leq n$ vectors.
  If a linearly independent set has n-vectors, then it must be a basis.
  i.e. it must also be a spanning set for V.
- Any spanning set for V must have $\geq n$ vectors. If a spanning set has n vectors, it must be a basis.
  i.e. it must also be linearly independent.

Note: Thus we can regard a basis as either a $\boxed{\text{maximal linearly independent set}}$ or as a $\boxed{\text{minimal spanning set.}}$

Ex: Let $V = \mathbb{R}^4$, $W = \text{span}(S)$, where $S = \{w_1, w_2, w_3\}$

Insert v_1 and v_2 into S replacing suitable w's to get a new spanning set for W applying the method of Steinitz Exchange Lemma.

$$v_1 = \begin{bmatrix} 2 \\ 3 \\ 7 \\ 9 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 3 \\ 4 \\ 8 \\ 12 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 9 \\ 11 \\ 19 \\ 33 \end{bmatrix},$$

$$w_3 = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -3 \end{bmatrix}$$

Need to Solve:

$$x_1 w_1 + x_2 w_2 + x_3 w_3 = v_1$$

$$\Rightarrow \quad x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 9 \\ 11 \\ 19 \\ 33 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 7 \\ 9 \end{bmatrix}$$

The Augmented matrix $[w_1, w_2, w_3 \mid v_1]$

$$= \left[\begin{array}{ccc|c} 1 & 9 & -1 & 2 \\ 1 & 11 & -1 & 3 \\ 1 & 19 & -1 & 7 \\ 1 & 33 & -3 & 9 \end{array}\right] \xrightarrow[\substack{R_3 \to R_3 - R_1 \\ R_4 \to R_4 - R_1}]{} \left[\begin{array}{ccc|c} 1 & 9 & -1 & 2 \\ 0 & 2 & 0 & 1 \\ 0 & 10 & 0 & 5 \\ 0 & 24 & -2 & 7 \end{array}\right]$$

$R_2 \to R_2 - R_1$

$$\xrightarrow{R_2 \to \frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & 9 & -1 & 2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 10 & 0 & 5 \\ 0 & 24 & -2 & 7 \end{array}\right] \xrightarrow[R_4 \to R_4 - 24R_2]{R_3 \to R_3 - 10R_2} \left[\begin{array}{ccc|c} 1 & 9 & -1 & 2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & -5 \end{array}\right]$$

$$\xrightarrow{R_3 \leftrightarrow R_4} \left[\begin{array}{ccc|c} 1 & 9 & -1 & 2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & -2 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_3 \to \left(-\frac{1}{2}R_3\right)} \left[\begin{array}{ccc|c} 1 & 9 & -1 & 2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{5}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_1 \to R_1 + R_3} \left[\begin{array}{ccc|c} 1 & 9 & 0 & \frac{9}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{5}{2} \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \to R_1 - 9R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{5}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus the solution is $x_1 = 0$, $x_2 = \frac{1}{2}$, $x_3 = \frac{5}{2}$

Hence $0 \cdot w_1 + \frac{1}{2} w_2 + \frac{5}{2} w_3 = v_1$

We can replace either w_2 or w_3 to get a new spanning set (w_1 can't be replaced).

Let $S_1 = \{w_1, v_1, w_3\}$ be our new spanning set.

Let us solve : $x_1 w_1 + x_2 v_1 + x_3 w_3 = v_2$

$$\Rightarrow x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 3 \\ 7 \\ 9 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 8 \\ 12 \end{bmatrix}$$

The augmented matrix $[w_1, v_1, w_3 \,\vdots\, v_2]$

$$= \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 1 & 3 & -1 & 4 \\ 1 & 7 & -1 & 8 \\ 1 & 9 & -3 & 12 \end{array}\right] \xrightarrow[\substack{R_3 \to R_3 - R_1 \\ R_4 \to R_4 - R_1}]{R_2 \to R_2 - R_1} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 5 & 0 & 5 \\ 0 & 7 & -2 & 9 \end{array}\right]$$

$$\xrightarrow[R_4 \to R_4 - 7R_2]{R_3 \to R_3 - 5R_2} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 \end{array}\right] \xrightarrow{R_3 \leftrightarrow R_4} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 \to (-\frac{1}{2} R_3)} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \to R_1 + R_3} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\xrightarrow{R_1 \to R_1 - 2R_2}$

$$\left[\begin{array}{ccc:c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So, $x_1 = 0$, $x_2 = 1$, $x_3 = -1$

$\Rightarrow \quad 0.w_1 + 1.v_1 + (-1) w_3 = v_2$

Thus w_3 can be replaced by v_2

(w_1 can't be replaced)

Hence the new spanning set will be

$$\boxed{\{w_1, v_1, v_2\}}$$

Dimension of Subspaces:

Definition: A proper subspace of a vector space is a subspace different from the zero subspace and the entire space.

Proposition: If W is a proper subspace of a finite-dimensional space V, then W is also finite dimensional and $0 < \dim W < \dim V$.

Proof: Since W is a proper subspace of V,

$$W \neq \{0\}$$

So, there exists $w_1 \in W$ $(w_1 \neq 0)$

If $\operatorname{span}\{w_1\} = W$, then W is finite dimensional. $(\dim W = 1)$

If $\operatorname{span}\{w_1\} \neq W$, there exists $w_2 \in W$ $(w_2 \neq 0)$ such that $w_2 \notin \operatorname{span}\{w_1\}$

By adjoining w_2 to w_1, we get a linearly independent set $\{w_1, w_2\}$.

Continuing in this way, we get a basis of W with atmost $\dim V$ elements (By Steinitz Exchange lemma)

Hence W is finite dimensional and $0 < \dim W \leq \dim V$

Since W is a proper subspace of V, there exists $v \in V$ $(v \neq 0)$ such that $v \notin W$. Adjoining v to any basis of W, we will have a linearly independent set in V. Hence $\dim W$ is strictly less than $\dim V$.

i.e. $\boxed{\dim W < \dim V}$