IIT MADRAS **BS DEGREE**

MATHEMATICS

Mathematics for Data Science-II

Sushmitha, Shubham, Subhasis and Subhajit

Linear Algebra

Copyright © IITM BS

About the author

Mr. Shubham Maurya

Mr. Shubham Maurya is an instructor in IITM BS degree program of IIT Madras. He have completed his graduation from Banaras Hindu University and received BHU Medal for securing first position, and Master of Science in mathematics from prestigious IIT Madras. He have a deep understanding of mathematical concept, theories, and applications, and have delivered his mathematical ideas in a clear and accessible manner to a wide range of audiences of IITM BS degree students.

In this book, he is try to explain his mathematical idea in a such a manner that every student can easily understand the concepts with more examples so that reader get benefited from his unique understanding and inside vision.

Contents

INDIAN INSTITUTE OF TECHNOLOGY MADRAS • सिद्धिर्भवति कर्मजा •

1. Vector and Matrices

"What has been affirmed without proof can also be denied without proof."

— Euclid

1.1 Introduction

Vectors are foundational elements of linear algebra. Often we encounter data in a table. For example in the following table we can get a complete view on India's GDP from 2000-01 to 2012-13 with sector wise break-ups.

Financial Year	Gross Domestic Product (in Rs. Cr) at 2004-05 Prices	Agriculture & Allied Services (in Rs. Cr.) at 2004-05 Prices	Agriculture (in Rs. Cr.) at 2004-05 Prices	Industry (in Rs. Cr.) at 2004-05 Prices	Mining and Quarrying (in Rs. Cr.) at 2004-05 Prices	Manufacturing (in Rs. Cr.) at 2004-05 Prices	Services (in Rs. Cr.) at 2004-05 Prices
2000-01	2342774	522755	439432	640043	69472	363163	1179976
2001-02	2472052	554157	467815	656737	70766	371408	1261158
2002-03	2570690	517559	429752	704095	76721	396912	1349035
2003-04	2777813	564391	476324	755625	78792	422062	1457797
2004-05	2971464	565426	476634	829783	85028	453225	1576255
2005-06	3253073	594487	502996	910413	86141	499020	1748173
2006-07	3564364	619190	523745	1021204	92578	570458	1923970
2007-08	3896636	655080	556956	1119995	95997	629073	2121561
2008-09	4158676	655689	555442	1169736	98055	656302	2333251
2009-10	4516071	660987	557715	1276919	103830	730435	2578165
2010-11	4937006	713477	606848	1393879	108938	801476	2829650
2011-12	5243582	739495	630540	1442498	108249	823023	3061589
2012-13	5503476	752746		1487533	108713	838541	3263196

Figure 1.1: India's GDP data from 2000-01 to 2012-13 with sector wise break-ups

In the following table we can view the average run scored by five players of Indian cricket team: V.Kohli, M.S.Dhoni, R. Sharma, K.L.Rahul and S.Dhawan against Australia, England, New Zealand, South Africa, Sri Lanka and Pakistan.

vs Teams	V.Kohli	M.S.Dhoni	R.Sharma	K.L.Rahul	S.Dhawan
Australia	54.57	44.86	61.33	45.75	45.80
England	45.30	46.84	50.44	6.60	32.45
New Zealand	59.91	49.47	33.47	68.33	32.72
South Africa	64.35	31.92	33.30	26.00	49.87
Sri Lanka	60.00	64.40	46.25	34.75	70.21
Pakistan	48.72	53.52	51.42	57.00	54.28

Table 1.1: Team-wise batting averages

A vector can be thought of as a list. In the context of the above examples, vectors could be columns or rows. If we choose the row corresponding to the year 2010-11 in the table given in Figure 1.1, then (4937006, 713477, 606848, 1393879, 108938, 801476, 2829650) is considered as a row vector. Similarly if we consider the averages of all the five players against South Africa given in Table 1.1, then (64.35, 31.92, 33.30, 26.00, 49.87) is also a row vector. Now if we choose the column corresponding to Gross domestic product (in Rs. Cr.) at 2004-05 prices in the table given in Figure 1.1, then

$$\begin{pmatrix} 2342774 \\ 2472052 \\ 2570690 \\ 2777813 \\ 2971464 \\ 3253073 \\ 3564364 \\ 3896636 \\ 4158676 \\ 4516071 \\ 4937006 \\ 5243582 \\ 5503476 \end{pmatrix}$$

is considered as a column vector. Similarly if we consider the averages of R.Sharma against Australia, England, New Zealand, South Africa, Sri Lanka and Pakistan given in Table 1.1, then

$$\begin{pmatrix} 61.33 \\ 50.44 \\ 33.47 \\ 33.30 \\ 46.25 \\ 51.42 \end{pmatrix}$$

is also a column vector.

1.2 Why vectors are important?

Vectors can be used to perform arithmetic operations on lists such as the table columns or rows e.g. suppose we want the average sectoral GDP across the years 2000-01 to 2009-10. One way to do this is by adding each elements of each column from 2000-01 to 2009-10 and divide the total by 10 to get the average. But the more efficient way to do it is by considering the row vectors corresponding to each year starting from 2000-01 till 2009-10 and adding those row vectors co-ordinate wise and then multiplying the resultant vector by $\frac{1}{10}$, i.e. multiplying each element of the resultant vector by $\frac{1}{10}$. Let us describe this process without using the particular numbers of the table as it will make it more cumbersome.

$$\begin{array}{c|ccccccc}
2000-01 & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\
2001-02 & a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
2008-09 & a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} \\
2009-10 & a_{10\ 1} & a_{10\ 2} & a_{10\ 3} & a_{10\ 4} & a_{10\ 5} & a_{10\ 6} & a_{10\ 7}
\end{array}$$

Adding row vectors corresponding to each year:

$$(a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{16}, a_{17}) + (a_{21}, a_{22}, a_{23}, a_{24}, a_{25}, a_{26}, a_{27}) +$$

$$\dots$$

$$+(a_{91}, a_{92}, a_{93}, a_{94}, a_{95}, a_{96}, a_{97}) + (a_{10\ 1}, a_{10\ 2}, a_{10\ 3}, a_{10\ 4}, a_{10\ 5}, a_{10\ 6}, a_{10\ 7})$$

we get,

$$\left(\Sigma_{i=1}^{10} a_{i1}, \Sigma_{i=1}^{10} a_{i2}, \dots, \Sigma_{i=1}^{10} a_{i6}, \Sigma_{i=1}^{10} a_{i7}\right)$$

Multiplying $\frac{1}{10}$ co-ordinatewise we get,

$$\frac{1}{10}\left(\Sigma_{i=1}^{10} a_{i1}, \Sigma_{i=1}^{10} a_{i2}, \dots, \Sigma_{i=1}^{10} a_{i6}, \Sigma_{i=1}^{10} a_{i7}\right) = \left(\frac{1}{10}\Sigma_{i=1}^{10} a_{i1}, \frac{1}{10}\Sigma_{i=1}^{10} a_{i2}, \dots, \frac{1}{10}\Sigma_{i=1}^{10} a_{i6}, \frac{1}{10}\Sigma_{i=1}^{10} a_{i7}\right)$$

Now observe that each element of the vector $\left(\frac{1}{10}\Sigma_{i=1}^{10} a_{i1}, \frac{1}{10}\Sigma_{i=1}^{10} a_{i2}, \ldots, \frac{1}{10}\Sigma_{i=1}^{10} a_{i6}, \frac{1}{10}\Sigma_{i=1}^{10} a_{i7}\right)$ denotes the average GDP in the corresponding sector.

Let us try to understand the usage of vectors in little more details through some examples.

Example 1.2.1. Suppose Arun has to buy 3 Kg Rice and 2 Kg dal and Neela has to buy 5 Kg Rice and 6 Kg dal. Then the vectors (3, 2) for Arun and (5, 6) for Neela represent their demands.

Items	Arun	Neela	Total
Rice in Kg	3	5	8
Dal in Kg	2	6	8

We can add this vectors to get $(3, 2) + (5, 6) = (8, 8)$ which is nothing but represents that together they have to buy 8 Kg Rice and 8 Kg dal.

Example 1.2.2. Stocks in Grocery Shops:

Items	In stock	Buyer A	Buyer B	Buyer C	New stock
Rice in Kg	150	8	12	3	100
Dal in Kg	50	8	5	2	75
Oil in Liter	35	4	7	5	30
Biscuits	70	10	10	5	80
Soap Bar	25	4	2	1	30

Taking stock of the items in the grocery shop can be done easily using vector representation : $(150, 50, 35, 70, 25) + (-8, -8, -4, -10 - 4) + (-12, -5, -7, -10, -2) + (-3, -2, -5, -5, -1) + (100, 75, 30, 80, 30) = (227, 110, 49, 125, 48)$. Note that we add corresponding entries of the vectors. This is an example of addition of vectors. If Buyer A comes the next day and buys the same items in the same quantities then we can add the vector two times or multiply each co-ordinate of the vector by 2.

$$(8, 8, 4, 10, 4) + (8, 8, 4, 10, 4) = (16, 16, 8, 20, 8) = 2.(8, 8, 4, 10, 4)$$

Multiplying a vector by a scalar (i.e. all its entries if it is a list) is called scalar multiplication.

1.2.1 Visualization

In this section we try to visualize the vectors in $\mathbb{R}^2$. Any point (a, b) on $\mathbb{R}^2$ is a vector (a, b) in $\mathbb{R}^2$, which can be visualized by the arrow joining the origin $(0, 0)$ to (a, b).

$$Point(a,b) \equiv Vector(a,b) \equiv a\hat{\imath} + b\hat{\jmath}$$

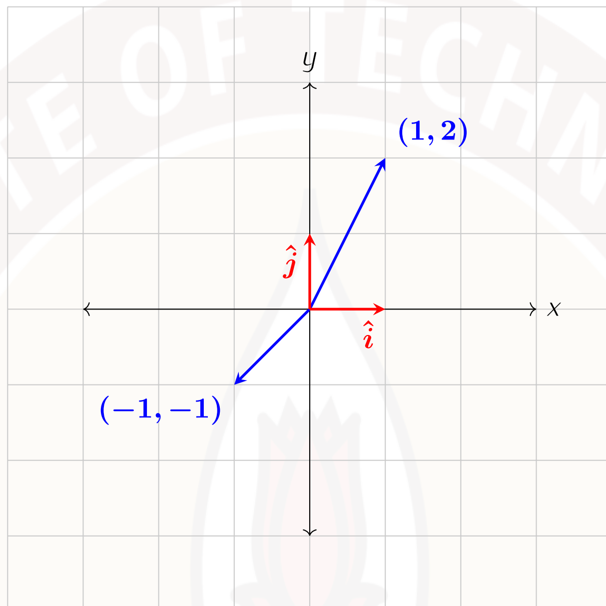


Figure 1.2: Vector visualization on $\mathbb{R}^2$

We can write $(1,2) = 1\hat{\imath} + 2\hat{\jmath}$ and similarly, $(-1,-1) = (-1)\hat{\imath} + (-1)\hat{\jmath}$.
In general, vectors in $\mathbb{R}^n$ are lists (or rows or columns) with n real entries.

$$\text{Vectors with } n \text{ entries} \equiv \text{Vectors in } \mathbb{R}^n \equiv \text{Points in } \mathbb{R}^n.$$

1.2.2 Visualization of vector addition

We can add two vectors by joining them head-to-tail or by parallelogram law.

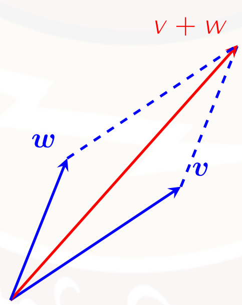


Figure 1.3: Visualization of vector addition in $\mathbb{R}^2$

If we want to add the two vectors $(1, 2)$ and $(2, 1)$, then we have to complete the diagram as shown in the Figure 1.4. The diagonal of the parallelogram denotes the vector resulting from the addition of these two vectors.

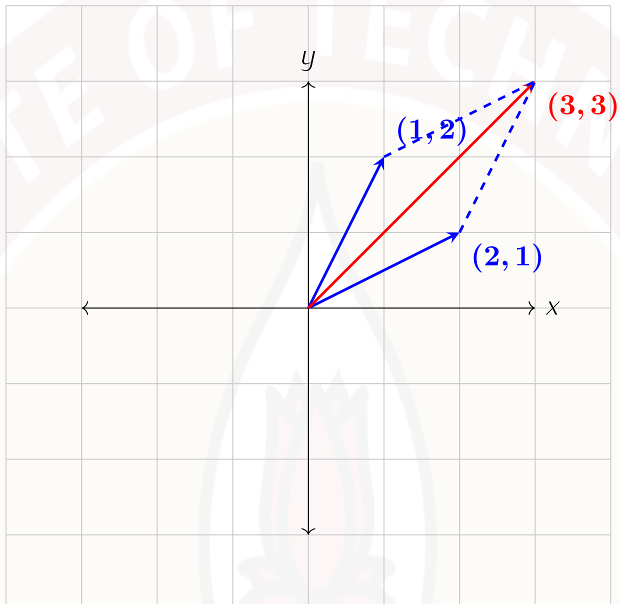


Figure 1.4: Visualization of vector addition in $\mathbb{R}^2$

So we have $(1, 2) + (2, 1) = (3, 3)$ as we can observe in the Figure 1.4.

1.2.3 Vectors in physical context

In high school physics we have studied vectors in quite details, but may be in a different approach. A vector has magnitude (size) and direction. The length of the line shows its magnitude and the arrowhead points in the direction. Although a vector has magnitude and direction, it does not have position. That is, as long as its length is not changed, a vector is not altered if it is displaced parallel to itself. Some examples of vectors which appear in physics :

- Velocity
- Acceleration
- Force

Example 1.2.3. A plane is flying towards the north and wind is blowing from the North-West.

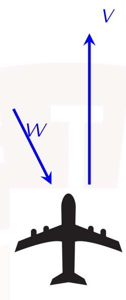


Figure 1.5: Caption

$v=$ velocity of the flight and $w=$ velocity of the wind

1.2.4 Exercise

Question 1. Choose the set of correct options using Figure 1.6.

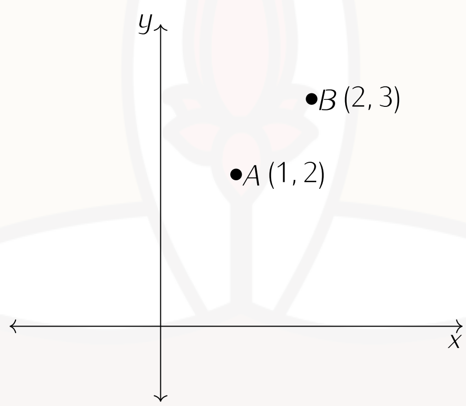


Figure 1.6:

[Hint: Recall that, vector addition and scalar multiplication are done coordinate-wise.]

- **Option 1:** $2A$ is the vector $(2, 4)$.
- **Option 2:** $3B$ is the vector $(6, 9)$.
- **Option 3:** $A + B$ is the vector $(3, 5)$.
- **Option 4:** $A - B$ is the vector $(-1, -1)$.

Question 2. The marks obtained by Karthika, Romy and Farzana in Quiz 1, Quiz 2 and End sem (with the maximum marks for each exam being 100) are shown in Table 1.2. Use the above information to answer the following questions:

	Quiz 1	Quiz 2	End sem
Karthika	51	50	61
Romy	33	41	45
Farzana	38	21	35

Table 1.2:

a) Choose the following set of correct options.

- **Option 1:** Marks obtained by Romy in Quiz 1, Quiz 2 and End sem represent a row vector.
- **Option 2:** Quiz 2 marks of Karthika, Romy and Farzana represent a column vector.
- Option 3: Number of components in column vector representing Quiz 2 marks are 9.
- **Option 4:** Number of components in row vector representing Romy's marks are 3.

b) In order to improve her marks, Farzana undertook project work and succeeded in increasing her marks. Her marks became doubled for each exam. Choose the correct options.

- **Option 1:** To obtain the marks obtained by Farzana after completion of the project, scalar multiplication has to be done by 2 to the row vector representing Farzana's marks.
- Option 2: To obtain the marks obtained by Farzana after completion of the project, scalar multiplication has to be done by 1 to the row vector representing Farzana's marks.
- **Option 3:** After completion of the project the row vector representing Farzana's marks is (76, 42, 70)
- Option 4: After completion of the project the row vector representing Farzana's marks is (76, 21, 35).
- Option 5: After completion of the project the row vector representing Farzana's marks is (66, 82, 90)

c) Following Farzana's improved marks due to her project (i.e her marks become doubled for each exam), all students were given bonus marks in Quiz 2, which is given by the column vector $\begin{pmatrix} 10 \\ 12 \\ 15 \end{pmatrix}$. What will be the column vector representing the final marks obtained in Quiz 2 by Karthika, Romy and Farzana?

- **Option 1:** $\begin{pmatrix} 60 \\ 53 \\ 57 \end{pmatrix}$
- Option 2: $\begin{pmatrix} 60 \\ 53 \\ 36 \end{pmatrix}$.
- Option 3: $\begin{pmatrix} 61 \\ 45 \\ 53 \end{pmatrix}$
- Option 4: $\begin{pmatrix} 71 \\ 57 \\ 85 \end{pmatrix}$

Question 3. Consider vectors $A(-1, 2)$ and $B(2, -2)$ in $\mathbb{R}^2$ as shown in Figure 1.7.

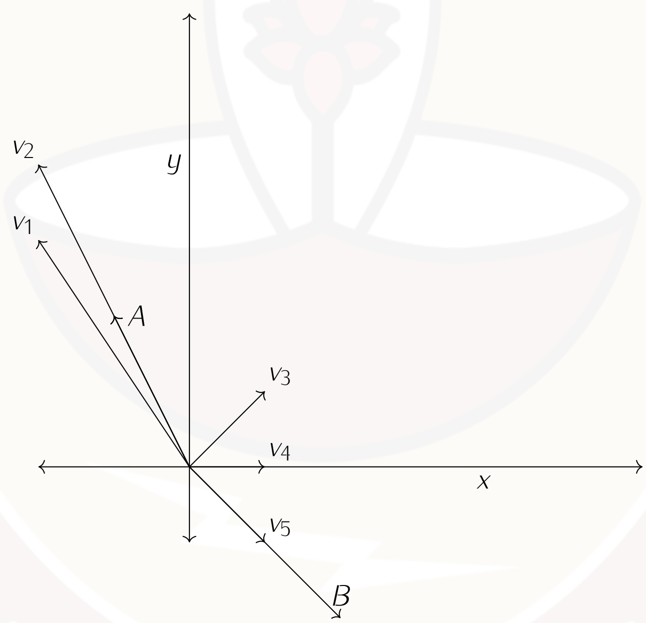


Figure 1.7:

Choose the set of correct options.
[Hint: Recall the geometric representation of vectors, scalar multiplication and vectors addition.]

- Option 1: v_1 represents a scalar multiple of A.
- **Option 2:** v_2 represents a scalar multiple of A.
- **Option 3:** v_5 represents a scalar multiple of B.
- Option 4: v_1 represents a scalar multiple of B.
- **Option 5:** v_4 represents a scalar multiple of $A+B$.
- Option 6: v_3 represents a scalar multiple $A+B$.

Question 4. Let $A = (1, 1, 1)$ and $B = (2, -1, 4)$ be two vectors. Suppose $cA + 3B = (4, j, k)$, where c, j, k are real numbers (scalars). Find the value of c. [Ans: -2]

Question 5. Let $A = (1, 1, 1)$ and $B = (2, -1, 4)$ be two vectors. Suppose $cA + 3B = (4, j, k)$, where c, j, k are real numbers (scalars). Find the value of $j + k$. [Ans: 5]

1.3 Matrices

Matrix, a set of numbers arranged in rows and columns so as to form a rectangular array. The numbers are called the elements, or entries, of the matrix. Matrices have wide applications in engineering, physics, economics, and statistics as well as in various branches of mathematics. The term matrix was introduced by the 19th-century English mathematician James Sylvester, but it was his friend the mathematician Arthur Cayley who developed the algebraic aspect of matrices in two papers in the 1850s. Cayley first applied them to the study of systems of linear equations, where they are still very useful.

1.3.1 What is a matrix?

Definition 1.3.1. A matrix is a rectangular array of numbers, arranged in rows and columns. (plural : matrices)

Two important notions:

- An $m \times n$ matrix has m rows and n columns.
- (i, j)-th entry of a matrix is the entry occuring in the i-th row and j-th column.

Let us try to understand these notions using the following example:

Example 1.3.1.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}_{2\times 3}$$

- This is a 2×3 matrix (It has 2 rows and 3 columns).
- $(1, 2)$-th entry is 2.
- $(2, 3)$-th entry is 4.

There are some special types of matrices.

Definition 1.3.2. A square matrix is a matrix in which the number of rows is the same as the number of columns.

Example 1.3.2.

$$\begin{bmatrix} 0.3 & 5 & -7 \\ 2.8 & 0 & 1 \\ 0 & -2.5 & -1 \end{bmatrix}_{3\times 3}$$

This is a 3×3 matrix (3 rows and 3 columns).

- The i-th diagonal entry of a square matrix is the (i, i)-th entry.
- The diagonal of a square matrix is the set of diagonal entries.
- $0.3, 0$ and -1 are the diagonal entries of the square matrix given above.

Definition 1.3.3. A square matrix in which all entries except the diagonal are 0 is called a diagonal matrix.

Example 1.3.3.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 4.2 \end{bmatrix}_{3\times 3}$$

This is a 3×3 diagonal matrix, where diagonal entries are $1, -3$ and 4.2

Definition 1.3.4. A diagonal matrix in which all the entries in the diagonal are equal is called a scalar matrix.

Example 1.3.4.

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}_{3\times 3}$$

This is a 3×3 diagonal matrix, where all the diagonal entries are -3.

Definition 1.3.5. The scalar matrix with all diagonal entries 1 is called the Identity matrix and is denoted by I.

Example 1.3.5.

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This is the 3×3 identity matrix.

1.3.2 Linear equations and matrices

One of the important applications of matrices is to study the system of linear equations. A set of linear equations can be represented in terms of matrices. We will study later how we can find solutions of the system using this representation. For now, let us take to example to understand how we can represent the systems using matrices.

Example 1.3.6. Set of linear equations can be represented in terms of matrices.

$$\begin{aligned} 3x + 4y &= 5 \\ 4x + 6y &= 10 \end{aligned}$$

can be represented by the matrix, $\left[\begin{array}{cc|c} 3 & 4 & 5 \\ 4 & 6 & 10 \end{array}\right]$

Example 1.3.7. Set of linear equations can be represented in terms of matrices.

$$\begin{aligned} x - 2y + 5z &= 10 \\ -x + 3y - 4z &= 0 \\ 2x + y + z &= 7 \end{aligned}$$

$$\text{can be represented by the matrix, } \left[\begin{array}{ccc|c} 1 & -2 & 5 & 10 \\ -1 & 3 & -4 & 0 \\ 2 & 1 & 1 & 7 \end{array}\right]$$

1.3.3 Addition of matrices

Matrix addition is the operation of adding two matrices by adding the corresponding entries together.

Example 1.3.8.

$$\begin{bmatrix} 1 & 9 \\ 0.6 & 7 \\ 4 & 1.5 \end{bmatrix}_{3\times 2} + \begin{bmatrix} 0 & 7 \\ 0.6 & -7 \\ 2.5 & 0.6 \end{bmatrix}_{3\times 2} = \begin{bmatrix} 1 & 16 \\ 1.2 & 0 \\ 6.5 & 2.1 \end{bmatrix}_{3\times 2}$$

Definition 1.3.6. The sum of two $m \times n$ matrix A and B is calculated entrywise : the (i, j)-th entry of the matrix $A + B$ is the sum of (i, j)-th entry of A and (i, j)-th entry of B

$$(AB)_{ij} = A_{ij} + B_{ij}$$

Remark 1.3.1. From the definition it is clear that if A is an $m \times n$ matrix and B is an $k \times l$ matrix, then addition of A and B is possible if and only if $m = k$ and $n = l$.

Example 1.3.9.

$$\begin{bmatrix} 1/2 & -3/4 & 3 \end{bmatrix}_{1\times 3} + \begin{bmatrix} 2 & -3 & -1 \end{bmatrix}_{1\times 3} = \begin{bmatrix} 5/2 & -15/4 & 2 \end{bmatrix}_{1\times 3}$$

1.3.4 Scalar multiplication (Multiplying a matrix by a number)

Multiplying a matrix by a number means, a number is multiplied with every other element of the matrix.

Example 1.3.10.

$$3\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix}_{2\times 3} = \begin{bmatrix}3 & 6 & 9\\ 12 & 15 & 18\end{bmatrix}_{2\times 3}$$

Definition 1.3.7. The product of a matrix A with a number c is denoted by cA and the (i, j)-th entry of cA is product of (i, j)-th entry of A with the number c.

$$(cA)_{ij} = c(A_{ij})$$

Let us consider another example.

Example 1.3.11.

$$4\begin{bmatrix}1 & 0\\ -5 & 3\end{bmatrix}_{2\times 2} = \begin{bmatrix}4 & 0\\ -20 & 12\end{bmatrix}_{2\times 2}$$

1.3.5 Multiplication of matrices

If $R = \begin{bmatrix}a & b & c\end{bmatrix}$ and $C = \begin{bmatrix}d\\ e\\ f\end{bmatrix}$, then the product RC is given by

$$\begin{bmatrix}ad + be + cf\end{bmatrix}$$

Suppose $A = \begin{bmatrix}R_1\\ R_2\end{bmatrix}$ and $B = \begin{bmatrix}C_1 & C_2 & C_3\end{bmatrix}$, where R_i's denote the rows of matrix A and C_i's denote the columns of matrix B. Moreover, assume that the number of columns of A and the number of rows of B are the same. The product AB is given by,

$$AB = \begin{bmatrix}R_1C_1 & R_1C_2 & R_1C_3\\ R_2C_1 & R_2C_2 & R_2C_3\end{bmatrix}$$

Let $A = \begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}$ and $B = \begin{bmatrix}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{bmatrix}$ be two 3×3 matrices. Then the product AB is given by

$$AB = \begin{bmatrix}a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33}\\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33}\\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33}\end{bmatrix}$$

Definition 1.3.8. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix, then their product AB will be an $m \times p$ matrix ($A_{m\times n}B_{n\times p} = (AB)_{m\times p}$) and the (i, j)-th entry of AB is defined as follows,

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}.B_{kj}$$

Remark 1.3.2. Multiplication of matrices A and B is defined only when the number of columns of A is the same as the number of rows of B.

Example 1.3.12. Let A be $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and B be $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix}$. Let is try to calculate AB.

For (1,1)-th entry:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 7 & & \\ & & \end{bmatrix}_{2\times 3}$$

For (1,2)-th entry:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 7 & 10 & \\ & & \end{bmatrix}_{2\times 3}$$

For (1,3)-th entry:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 7 & 10 & 13 \\ & & \end{bmatrix}_{2\times 3}$$

For (2,1)-th entry:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 7 & 10 & 13 \\ 15 & & \end{bmatrix}_{2\times 3}$$

For (2,2)-th entry:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 7 & 10 & 13 \\ 15 & 22 & \end{bmatrix}_{2\times 3}$$

For (2,3)-th entry:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 7 & 10 & 13 \\ 15 & 22 & 29 \end{bmatrix}_{2\times 3}$$

Hence we have,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix}_{2\times 3} = \begin{bmatrix} 7 & 10 & 13 \\ 15 & 22 & 29 \end{bmatrix}_{2\times 3}$$

Example 1.3.13.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \begin{bmatrix} 5 \\ 6 \end{bmatrix}_{2\times 1} = \begin{bmatrix} 17 \\ 39 \end{bmatrix}_{2\times 1}$$

Example 1.3.14.

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}_{1\times 3} \begin{bmatrix} 2 & 0.8 \\ 5 & 0.7 \\ 1/2 & -2 \end{bmatrix}_{3\times 2} = \begin{bmatrix} 13.5 & -3.8 \end{bmatrix}_{1\times 2}$$

Example 1.3.15.

$$\begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{bmatrix}_{3\times 3} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}_{3\times 2} = \begin{bmatrix} c & 2c \\ 3c & 4c \\ 5c & 6c \end{bmatrix}_{3\times 2} = c \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}_{3\times 2}$$

Remark 1.3.3. Scalar multiplication by c is multiplication by scalar matrix $cI_{n\times n}$ where $I_{n\times n}$ denotes the $n \times n$ identity matrix.

$$I_{n\times n}A_{n\times n} = A_{n\times n} = A_{n\times n}I_{n\times n}$$
$$I_{n\times n}A_{n\times k} = A_{n\times k}$$
$$A_{m\times n}I_{n\times n} = A_{m\times n}$$

where $I_{n\times n}$ denotes the $n \times n$ identity matrix.

1.3.6 Properties of matrix addition and multiplication

In this subsection we mention properties of matrix addition and multiplication without the proof. We want all of you to verify these on your own.

- $(A + B) + C = A + (B + C)$ (Associativity of addition)
- $(AB)C = A(BC)$ (Associativity of multiplication)
- $A + B = B + A$ (Commutativity of addition)
- In general $AB \neq BA$ (assuming both make sense)
- $\lambda(A + B) = \lambda A + \lambda B$ for some real number λ.
- $\lambda(AB) = (\lambda A)B = A(\lambda B)$ for some real number λ.
- $A(B + C) = AB + AC$
- $(A + B)C = AC + BC$

1.3.7 Exercise

Question 6. Suppose $A = \begin{bmatrix} 2 & 4 & 5 & 1 \\ 11 & -2 & 9 & -6 \\ -3 & 4 & 7 & 7 \end{bmatrix}$. Which of the following is true about the matrix A? [Hint: The (i, j)-th entry is the entry which is at the i-th row and j-th column.]

- Option 1: It is a 4×3 matrix.
- **Option 2:** It is a 3×4 matrix.
- Option 3: (2,3)-th entry of the matrix A is 4.
- **Option 4:** (2,3)-th entry of the matrix A is 9.

Question 7. Which of the following statements is(are) TRUE?

[Hint: Recall, the definitions of scalar matrix, diagonal matrix, and identity matrix.]

- Option 1: Any diagonal matrix is a scalar matrix.
- Option 2: Scalar matrices may not be square matrices.
- **Option 3:** Scalar matrices must be square matrices.
- Option 4: Any scalar matrix is an identity matrix.

Question 8. Suppose $P = \begin{bmatrix} 3 & -1 & 7 \\ 4 & 0 & 1 \\ 2 & -5 & 2 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 4 & -9 \end{bmatrix}$, $R = \begin{bmatrix} 0 & -3 & 10 \end{bmatrix}$, $D = \begin{bmatrix} -2 \\ 4 \\ 5 \end{bmatrix}$

[Hint: If A is a matrix of order $m \times n$ and B is a matrix of order $n \times p$, then the order of AB is $m \times p$.]

- **Option 1:** The matrix PD is of the order 3×1.
- Option 2: The matrix PD is of the order 1×3.
- Option 3: The matrix QD is of the order 3×3.
- **Option 4:** The matrix QD is of the order 1×1.
- **Option 5:** The matrix DQ is of the order 3×3.

- Option 6: The matrix DQ is of the order 1×1.
- Option 7: QD is not defined.
- **Option 8:** QR is not defined.
- **Option 9:** $P + Q$ is not defined.
- **Option 10:** $P + D$ is not defined.

Feedback:

- P is a 3×3 matrix, D is a 3×1 matrix. Think about the order of PD.
- Q is a 1×3 matrix, D is a 3×1 matrix. Think about the order of QD and DQ.
- Q is a 1×3 matrix, R is a 1×3 matrix. Think whether is it possible to define QR or not.
- Figure out which pair of matrices has the same number of rows as well as the same number of columns.

Question 9. Suppose $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ Which of the following options are true?

- **Option 1:** $A^2 = I$
- Option 2: $A^2 = A$
- Option 3: $B^2 = I$
- **Option 4:** $B^2 = 0$

Question 10. Suppose A is a 3×3 scalar matrix and (1,1)-th entry of the matrix A is 4. Suppose B is a 3×3 square matrix such that (i, j)-th entry is equal to $i^2 + j^2$. Find the (2,2)-th entry of the matrix $2A + B$. [Ans: 16]

Question 11. Suppose $A = \begin{bmatrix} -1 & 2 \\ -4 & 7 \end{bmatrix}$ and $A^2 - \alpha A + I = 0$ for some $\alpha \in \mathbb{R}$.
Find the value of α. [Ans: 6]

1.4 System of linear equations

We have mentioned earlier that one of the applications of matrices is to solve system of linear equations. In this section we will study in detail how we can use matrices to do so. Let us start with an example from our day to day life.

Example 1.4.1. Suppose the purchases of A, B and C are given in the following table.

Items	Buyer A	Buyer B	Buyer C
Rice (in Kg)	8	12	3
Dal (in Kg)	8	5	2
Oil (in Liter)	4	7	5

Suppose A paid $Rs.1960$, B paid $Rs.2215$ and C paid $Rs.1135$. We want to find the price of each items using this data. Suppose price of Rice is $Rs.x$ per kg., price of dal is $Rs.y$ per kg., price of oil is $Rs.z$ per liter. Hence we have the following system of linear equations:

$$\begin{aligned} 8x + 8y + 4z &= 1960 \\ 12x + 5y + 7z &= 2215 \\ 3x + 2y + 5z &= 1135 \end{aligned}$$

can be represented as $Ax = b$, where

$$A = \begin{bmatrix} 8 & 8 & 4 \\ 12 & 5 & 7 \\ 3 & 2 & 5 \end{bmatrix} x = \begin{bmatrix} x \\ y \\ z \end{bmatrix} b = \begin{bmatrix} 1960 \\ 2215 \\ 1135 \end{bmatrix}$$

Simple checking shows that, $x = 45, y = 125, z = 150$ satisfies the equations. We are not mentioning here how this solution can be obtained, that we will do in the next chapter. But for now, we want you to verify that these values satisfy all the three equations simultaneously.

So the main question which should be asked here is, what a linear equation is.

Definition 1.4.1. A linear equation is an equation that may be put in the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b = 0$$

where $x_1, x_2, \ldots, x_n$ are the variables (or unknown), and $b, a_1, a_2, \ldots, a_n$ are the coefficients, which are often real numbers.

Example 1.4.2. $2x + 3y + 5z + 9 = 0$, where x, y, z are variables and $2, 3, 5, 9$ are the coefficients.

Example 1.4.3. $x - 3z - 7 = 0$, where x, z are variables and $1, -3, -7$ are the coefficients.

A system of linear equations is a collection of one or more linear equations involving the same set of variable. Which is illustrated in the following example.

Example 1.4.4.

$$\begin{aligned} 3x + 2y + z &= 6 \\ x - \frac{1}{2}y + \frac{2}{3}z &= \frac{7}{6} \\ 4x + 6y - 10z &= 0 \end{aligned}$$

is a system of three equations in the three variables x, y, z. A solution to a linear system is an assignment of values to the variables such that all the equations are simultaneously satisfied. A solution to the system above is given by

$$x = 1, y = 1, z = 1$$

Definition 1.4.2. A general system of m linear equations with n unknowns can be written as

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

The system of linear equations is equivalent to a matrix equation of the form

$$Ax = b$$

where A is an $m \times n$ matrix, x is a column vector with n entries and b is a column vector with m entries.

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Example 1.4.5. We consider the same set of linear equations mentioned in the Example 1.4.4, which is as follows:

$$\begin{aligned} 3x + 2y + z &= 6 \\ x - \frac{1}{2}y + \frac{2}{3}z &= \frac{7}{6} \\ 4x + 6y - 10z &= 0 \end{aligned}$$

The matrix representation of the system of linear equations is $Ax = b$, where

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 1 & -\frac{1}{2} & \frac{2}{3} \\ 4 & 6 & -10 \end{bmatrix} x = \begin{bmatrix} x \\ y \\ z \end{bmatrix} b = \begin{bmatrix} 6 \\ \frac{7}{6} \\ 0 \end{bmatrix}$$

Example 1.4.6. We consider the same set of linear equations mentioned in the example 1.4.1, which is as follows:

$$8x + 8y + 4z = 1960$$
$$12x + 5y + 7z = 2215$$
$$3x + 2y + 5z = 1135$$

which can be represented as $Ax = b$, where

$$A = \begin{bmatrix} 8 & 8 & 4 \\ 12 & 5 & 7 \\ 3 & 2 & 5 \end{bmatrix} x = \begin{bmatrix} x \\ y \\ z \end{bmatrix} b = \begin{bmatrix} 1960 \\ 2215 \\ 1135 \end{bmatrix}$$

1.4.1 Solutions of linear system of equations

A linear system may behave in any one of three possible ways:

1) The system has infinitely many solutions.
2) The system has a single unique solution.
3) The system has no solution.

In this section we will illustrate these three cases by examples and their geometrical representations.

Example 1.4.7. Example of infinitely many solutions

Suppose the purchases of A and B are given in the following table:

Items	Buyer A	Buyer B
Rice in Kg	2	4
Dal in Kg	1	2

Suppose A paid $Rs.215$, B paid $Rs.430$. We want to find the price of each items using this data. Suppose price of Rice is $Rs.x$ per kg., price of dal is $Rs.y$ per kg. Hence we have the following system of linear equations:

$$2x + y = 215$$
$$4x + 2y = 430$$

There are infinitely many x and y satisfying both the equations.

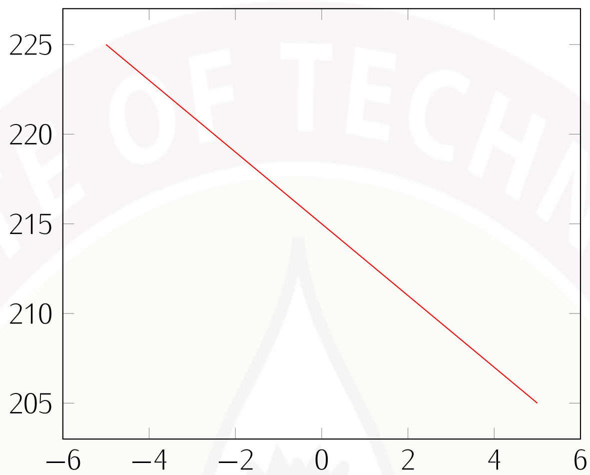


Figure 1.8:

Both the equations represent the same straight line in two dimensional plane as shown in Figure 1.8.

Example 1.4.8. Example of system of equations with no solution

Suppose as in the previous case both A and B bought the same amount of items as mentioned in Table 1.4.7. But for some reason the seller gave a discount to B. Let us assume, A paid $Rs.215$, B paid $Rs.400$. Now after returning home they decided to find out the price of each items by solving the linear system of equations as earlier cases. Suppose price of Rice is $Rs.x$ per kg., price of dal is $Rs.y$ per kg. Hence we have the following system of linear equations:

$$\begin{aligned} 2x + y &= 215 \\ 4x + 2y &= 400 \end{aligned}$$

But in this case there are no solution of this system of equations.

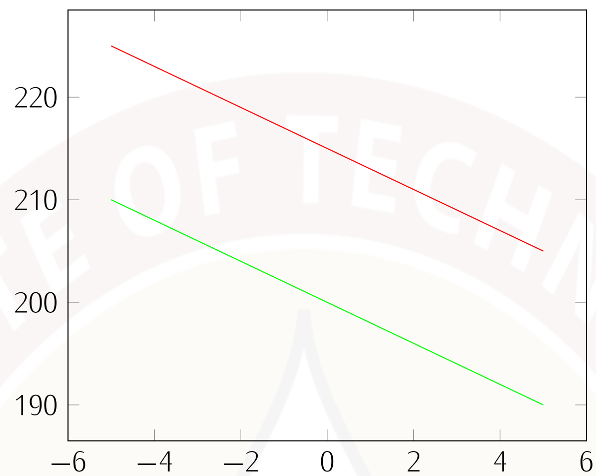


Figure 1.9:

The equations represent two parallel straight lines in two dimensional plane as shown in Figure 1.9.

Example 1.4.9. Example with unique solution

Suppose the purchases of A and B are given in the following table:

Items	Buyer A	Buyer B
Rice in Kg	2	3
Dal in Kg	1	1

Suppose A paid $Rs.215$, B paid $Rs.260$. We want to find the price of each items using this data. Suppose price of Rice is $Rs.x$ per kg., price of dal is $Rs.y$ per kg. Hence we have the following system of linear equations:

$$2x + y = 215$$
$$3x + y = 260$$

There are infinitely many x and y satisfying both the equations.

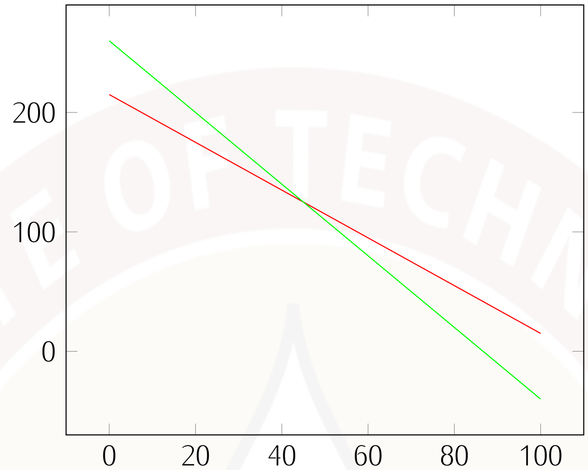


Figure 1.10:

The equations represent two straight lines intersecting with each other at one point as shown in Figure 1.10.

1.4.2 Exercise

Consider a system of linear equations (System 1):

$$\begin{aligned} -2x_1 + 3x_2 + x_3 &= 1 \\ -x_1 + x_3 &= 0 \\ 2x_2 &= 5 \end{aligned}$$

Answer questions 12 and 13 based on the above data.

Question 12. If the matrix representation of system (1) is $Ax = b$, where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then

- **Option 1:** $A = \begin{bmatrix} -2 & 3 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$
- Option 2: $b = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}$
- Option 3: $A = \begin{bmatrix} -2 & -1 & 0 \\ 3 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix}$, and $b = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$

- **Option 4:** $A = \begin{bmatrix} -2 & 3 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$, and $b = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$

Question 13. System (1) has
[Hint: Solve for x_1, x_2, and x_3.]

- **Option 1:** a unique solution.
- Option 2: no solution.
- Option 3: infinitely many solutions.
- Option 4: None of the above.

Question 14. The Plane 1 and Plane 2 in Figure M2W1AQ3, correspond to two different linear equations, which form a system of linear equations.

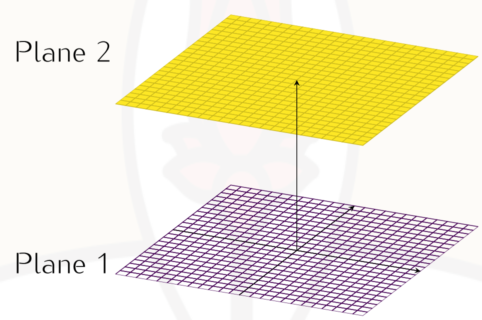


Figure 1.11: Figure M2W1AQ3

The above system of linear equations has
[Hint: If a system of linear equations has a solution, then the point corresponding to the solution must lie on each plane corresponding to each linear equation of the given system.]

- Option 1: a unique solution.
- **Option 2:** no solution.
- Option 3: infinitely many solutions.
- Option 4: None of the above.

Question 15. Consider the geometric representations (Figures $(a), (b),$ and (c)) of three systems of linear equations.

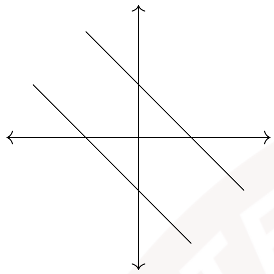


Figure 1.12: Figure (a)

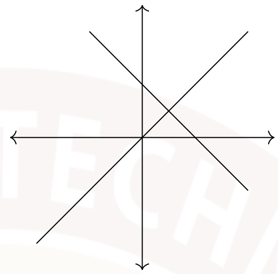


Figure 1.13: Figure (b)

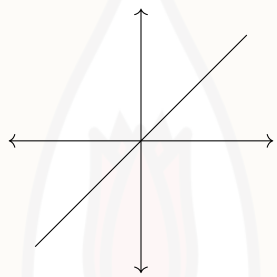


Figure 1.14: Figure (c)

Choose the set of correct options.

- **Option 1:** Figure (a) represents a system of linear equations which has no solution.
- Option 2: Figure (a) represents a system of linear equations which has infinitely many solutions.
- **Option 3:** Figure (b) represents a system of linear equations which has a unique solution.
- Option 4: Figure (b) represents a system of linear equations which has infinitely many solutions.
- **Option 5:** Figure (c) represents a system of linear equations which has infinitely many solutions.
- Option 6: Figure (c) represents a system of linear equations which has no solution.

Question 16. Consider a system of equations:

$$\begin{aligned} 2x_1 + 3x_2 &= 6 \\ -2x_1 + kx_2 &= d \\ 4x_1 + 6x_2 &= 12 \end{aligned}$$

Choose the set of correct options.

- **Option 1:** $Ax = b$ represents the above system, where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $A = \begin{bmatrix} 2 & 3 \\ -2 & k \\ 4 & 6 \end{bmatrix}$, and $b = \begin{bmatrix} 6 \\ d \\ 12 \end{bmatrix}$
- **Option 2:** The system has no solution if $k = -3$, $d = 0$.
- **Option 3:** The system has a unique solution if $k = 3$, $d = 0$.
- Option 4: The system has infinitely many solutions if $k = -3$, $d = 6$.
- **Option 5:** The system has infinitely many solutions if $k = -3$, $d = -6$.

Question 17. Let x_1 and x_2 be solutions of the system of linear equations $Ax = b$. Which of the following options are correct?

- Option 1: $x_1 + x_2$ is a solution of the system of linear equations $Ax = b$.
- **Option 2:** $x_1 + x_2$ is a solution of the system of linear equations $Ax = 2b$.
- Option 3: $x_1 - x_2$ is a solution of the system of linear equations $Ax = b$.
- **Option 4:** $x_1 - x_2$ is a solution of the system of linear equations $Ax = 0$.

Question 18. Let v be a solution of the systems of linear equations $A_1x = b$ and $A_2x = b$. Which of the following options are correct ?

- Option 1: v is a solution of the system of linear equations $(A_1 + A_2)x = b$.
- **Option 2:** v is a solution of the system of linear equations $(A_1 + A_2)x = 2b$.
- **Option 3:** v is a solution of the system of linear equations $(A_1 - A_2)x = 0$.
- Option 4: v is a solution of the system of linear equations $(A_1 - A_2)x = b$.

Question 19. Consider a system of equations:

$$\begin{aligned} x_1 - 3x_2 &= 4 \\ 3x_1 + kx_2 &= -12 \end{aligned}$$

Where $k \in \mathbb{R}$. If the given system has a unique solution, then k should not be equal to [Ans: -9]

1.5 Determinant

Every square matrix A has an associated number, called its determinant and denoted by $det(A)$ or $|A|$. It is used in :

- solving a system of linear equations,
- finding the inverse of a matrix,
- calculus and more.

1.5.1 First order determinant :

If $A = [a]$, a 1×1 matrix then $det(A) = a$

1.5.2 Second Order Determinant

Consider a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We calculate the determinant of the matrix as follows:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad det(A) = ad - bc$$

Example 1.5.1. $A = \begin{bmatrix} 2 & 3 \\ 6 & 10 \end{bmatrix} \qquad det(A) = 20 - 18 = 2$

Example 1.5.2. $A = \begin{bmatrix} 5 & 2/3 \\ 6 & 3/7 \end{bmatrix} \qquad det(A) = 15/7 - 4 = -13/7$

1.5.3 Third order determinant

Consider a 3×3 matrix A as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

We will calculate the determinant of matrix A. So first consider the element a_{11}. We have to consider the 2×2 submatrix ignoring the first row and first column, as shown below.

$$\begin{bmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

We calculate $(-1)^{1+1}a_{11} \times det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$

Now consider the element a_{12}. We have to consider the 2×2 submatrix ignoring the first row and the second column, as shown below.

$$\begin{bmatrix} a_{11} & \boxed{a_{12}} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

We calculate $(-1)^{1+2}a_{12} \times det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$

Now consider the element a_{13}. We have to consider the 2×2 submatrix ignoring the first row and the third column, as shown below.

$$\begin{bmatrix} a_{11} & a_{12} & \boxed{a_{13}} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

We calculate $(-1)^{1+3}a_{13} \times det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$

Calculating the above three terms, we are now in a position to calculate the determinant of matrix A.

$$\begin{aligned} det(A) &= (-1)^{1+1}a_{11} \times det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} + (-1)^{1+2}a_{12} \times det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + (-1)^{1+3}a_{13} \times det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

Example 1.5.3. Consider the following matrix:

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

$$\begin{aligned} det(A) &= 2 \times det \begin{bmatrix} 8 & 7 \\ 6 & 9 \end{bmatrix} - 4 \times det \begin{bmatrix} 3 & 7 \\ 5 & 9 \end{bmatrix} + 1 \times det \begin{bmatrix} 3 & 8 \\ 5 & 6 \end{bmatrix} \\ &= 2(72 - 42) - 4(27 - 35) + 1(18 - 40) \\ &= 2(30) - 4(-8) + 1(-22) \\ &= 60 + 32 - 22 \\ &= 70 \end{aligned}$$

Example 1.5.4. Consider the following matrix: (Observe that it is an upper triangular matrix)

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

$$\begin{aligned} det(A) &= 2 \times det \begin{bmatrix} 8 & 7 \\ 0 & 9 \end{bmatrix} - 4 \times det \begin{bmatrix} 0 & 7 \\ 0 & 9 \end{bmatrix} + 3 \times det \begin{bmatrix} 0 & 8 \\ 0 & 0 \end{bmatrix} \\ &= 2(72 - 0) - 4(0 - 0) + 3(0 - 0) \\ &= 2(72) - 4(0) + 3(0) \\ &= 144 \end{aligned}$$

Observer that, the determinant in this case is the product of the diagonal elements.

Here we mention some important properties of determinant and leave those for the learners to check and verify.

Proposition 1.5.1. *The following are some important properties of determinant of matrices.*

- *Determinant of an identity matrix (of any order n) is 1. (Verify it for 2×2 and 3×3 matrices first).*
- $det(A^T) = det(A)$, *where A^T denotes the transpose of matrix A.*
- $det(AB) = det(A)det(B)$, *where both A and B are $n \times n$ matrices (Verify it for 2×2 and 3×3 matrices first).*

Proposition 1.5.2. *Determinant of inverse of a matrix*

Let is denote inverse of a matrix A by A^{-1}, then we have the $AA^{-1} = I$. Calculating determinant of both the sides we have, $det(AA^{-1}) = det(I)$, i.e., $det(A).det(A^{-1}) = 1$. Hence we get $det(A^{-1}) = \frac{1}{det(A)}$.

1.5.4 Invariance under elementary row and column operations

Important properties of the determinant include the following, which include invariance under elementary row and column operations.

1) Switching two rows or columns changes the sign.
2) Multiples of rows and columns can be added together without changing the determinant's value.

3) Scalar multiplication of a row by a constant t multiplies the determinant by t.

4) A determinant with a row or column of zeros has value 0. (we can calculate the determinant by expanding with respect to that row or column.)

Switching two rows or columns changes the sign:
For any $n \times n$ matrix, switching two rows or columns changes the sign. We will verify this when $n = 2$ and suggest the learners to verify it for 3×3 matrices.
For a 2×2 matrix, we will show this. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ Switching first and second row we get, $\tilde{A} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$ $det(A) = ad - bc$ and $det(\tilde{A}) = cb - da = -(ad - bc) = -det(A)$
Multiples of rows and columns can be added together without changing the determinant's value:
Again we will verify it for 2×2 matrix and suggest the learners to check for a 3×3 matrix.
Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix.
Let us construct $\tilde{A}$ by multiplying t with the second row of A and add it to the first row of A. Hence $\tilde{A} = \begin{bmatrix} a + tc & b + td \\ c & d \end{bmatrix}$

$$det(\tilde{A}) = (a + tc)d - (b + td)c = ad + tcd - bc - tdc = ad - bc = det(A)$$

Scalar multiplication of a row by a constant t multiplies the determinant by t:
A very simple checking shows this for a 2×2 matrix.

$$A = \begin{bmatrix} a & tb \\ c & td \end{bmatrix} = t \begin{bmatrix} a & b \\ c & d \end{bmatrix} = t\tilde{A}$$

where $\tilde{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$det(A) = atd - tbc = t(ad - bc) = t.det(\tilde{A})$$

1.5.5 Determinant in terms of minors

Definition 1.5.1. If A is a square matrix, then the minor of the entry in the i-th row and j-th column (also called the (i, j) minor, or a first minor) is the determinant of the submatrix formed by deleting the i-th row and j-th column. This number is often denoted M_{ij}.

Consider a 3×3 matrix A as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

- Minor of the entry at 1-st row and 1-st column, i.e., $(1, 1)$ minor is the determinant of the 2×2 submatrix obtained by ignoring the first row and first column, as shown below.

$$\begin{bmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

  So $(1, 1)$ minor of the given matrix is $(a_{22}a_{33} - a_{23}a_{32})$.

- Minor of the entry at 2-nd row and 3-rd column, i.e., $(2, 3)$ minor is the determinant of the 2×2 submatrix obtained by ignoring the second row and third column, as shown below.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & \boxed{a_{23}} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

  So $(2, 3)$ minor of the given matrix is $(a_{11}a_{32} - a_{12}a_{31})$.

Definition 1.5.2. The (i, j) cofactor is obtained by multiplying the minor by $(-1)^{i+j}$.

- $(1, 1)$ cofactor of the given matrix above is $(-1)^{1+1}(a_{22}a_{33} - a_{23}a_{32}) = (a_{22}a_{33} - a_{23}a_{32})$.

- $(2, 3)$ cofactor of the given matrix above is $(-1)^{2+3}(a_{11}a_{32} - a_{12}a_{31}) = -(a_{11}a_{32} - a_{12}a_{31})$.

Observe that for a 3×3 matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

- We can calculate the determinant expanding with respect to the first row as follows:

$$\begin{aligned} det(A) &= (-1)^{1+1}a_{11} \times det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} + (-1)^{1+2}a_{12} \times det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + (-1)^{1+3}a_{13} \times det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ &= (-1)^{1+1}a_{11}M_{11} + (-1)^{1+2}a_{12}M_{12} + (-1)^{1+3}a_{13}M_{13} \end{aligned}$$

- We can also calculate the determinant expanding with respect to the second column row as follows:

$$det(A) = (-1)^{1+2}a_{12} \times det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + (-1)^{2+2}a_{22} \times det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} + (-1)^{3+2}a_{32} \times det\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}$$
$$= (-1)^{1+2}a_{12}M_{12} + (-1)^{2+2}a_{22}M_{22} + (-1)^{3+2}a_{32}M_{32}$$

This can be generalised in for $n \times n$ matrices.
Let A be an $n \times n$ matrix, then

$$det(A) = \Sigma_{j=1}^{n}(-1)^{i+j}a_{ij}M_{ij}, \text{ for a fixed } i$$
$$= \Sigma_{i=1}^{n}(-1)^{i+j}a_{ij}M_{ij}, \text{ for a fixed } j$$

1.5.6 Exercise

Question 20. Let A be a 3×3 matrix with non-zero determinant. If $det(2A) = k\ det(A)$, then what will be the value of k? [Ans: 8]

[Hint: If a scalar (c) is multiplied with one row of a matrix A, then the determinant of the new matrix will be c times the determinant of A.]

Question 21. If $A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & -3 & 4 \\ -1 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 & -2 \\ 0 & 1 & -1 \\ 3 & 4 & 2 \end{bmatrix}$, then choose the set of correct options.

- Option 1: $det(A) = -9$ and $det(B) = -3$.
- Option 2: $det(A) = 9$ and $det(B) = 3$.
- **Option 3:** $det(A) = -9$ and $det(B) = 3$.
- Option 4: $det(AB) = -27$ and $det(BA) = 27$.
- **Option 5:** $det(AB) = det(BA) = -27$.

Question 22. Let A be a 2×2 matrix, which is given as $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Define the following matrices :

$$B = \begin{bmatrix} a_{11} - a_{21} & a_{12} - a_{22} \\ a_{21} & a_{22} \end{bmatrix}, C = \begin{bmatrix} a_{11} - a_{12} & a_{12} \\ a_{21} - a_{22} & a_{22} \end{bmatrix},$$

$$D = \begin{bmatrix} a_{11} + a_{21} & a_{12} - a_{22} \\ a_{21} & a_{22} \end{bmatrix}, E = \begin{bmatrix} a_{11} - a_{21} & a_{12} + a_{22} \\ a_{21} & a_{22} \end{bmatrix}$$

Which of the matrices among $B, C, D,$ and E have the same determinant as that of the matrix A, for any real numbers $a_{11}, a_{12}, a_{21}, a_{22}$?

- Option 1: B and D
- Option 2: B and E
- **Option 3:** B and C
- Option 4: D and E
- Option 5: C and E

Question 23. Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ ta_{11} - sa_{31} & ta_{12} - sa_{32} & ta_{13} - sa_{33} \\ ra_{31} & ra_{32} & ra_{33} \end{bmatrix}$ be a matrix and $r, s, t \neq 0$. Find $det(A)$ [Ans: 0]

Question 24. Suppose $A = \begin{bmatrix} -1 & 4 & -2 \\ 0 & 1 & 3 \\ 1 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & -1 \\ 4 & \alpha & 0 \\ 0 & -3 & -5 \end{bmatrix}$. If $det(AB) = 32$, then find the value of α. [Ans: 1]

Question 25. Let A be a square matrix of order 3 and B be a matrix that is obtained by adding 2 times the first row of A to the third row of A and adding 3 times the second row of A to the first row of A. What is the value of $det(6A^2B^{-1})$?

- Option 1: $6\ det(A)$
- Option 2: $6\ det(A)det(B)$
- **Option 3:** $6^3\ det(A)$
- Option 4: $6^3\ det(A)det(B)$

Question 26. Choose the set of correct options

- **Option 1:** If A is a real 3×3 matrix, then $det(A) = det(A^T)$.
- Option 2: If $A = [a_{ij}]$ is a real 4×4 matrix, then the order of the sub matrix obtained by deleting the i-th row and j-th column of A is 4×4.
- **Option 3:** If $A = [a_{ij}]$ is a real 4×4 matrix, then the order of the sub matrix obtained by deleting the i-th row and j-th column of A is 3×3.
- **Option 4:** If A is a real 3×3 matrix, then the orders of all possible square sub matrices of A are $1 \times 1, 2 \times 2,$ and 3×3.

Question 27. Let $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 1 & 1 \end{bmatrix}$ be a 3×3 matrix. Which of the following is(are) correct?

[Hint: Determinant can be calculated by expanding with respect to different rows or columns.]

- **Option A:**
  $det(A) = 3 \times det\left(\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\right) - 2 \times det\left(\begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}\right) + 2 \times det\left(\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}\right)$
- **Option B:**
  $det(A) = 3 \times det\left(\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\right) + 2 \times det\left(\begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}\right) + 2 \times det\left(\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}\right)$
- **Option C:**
  $det(A) = -2 \times det\left(\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\right) + 3 \times det\left(\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}\right) - 2 \times det\left(\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}\right)$
- Option D:
  $det(A) = 2 \times det\left(\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\right) - 3 \times det\left(\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}\right) + 2 \times det\left(\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}\right)$

Question 28. Suppose $A = \begin{bmatrix} 4 & -1 & 3 \\ 2 & 0 & 1 \\ 3 & -2 & 0 \end{bmatrix}$ and C_{ij} is the (i, j)-th cofactor of the matrix A.
Let B be 3×3 matrix where (i, j)-th entry is equal to C_{ij}. Find $det(B)$. [Ans: 49]

Question 29. Suppose A and B are two 3×3 matrices such that $det(A) = 4$ and $B = 3A$. Find the value of $\sqrt[3]{det(A^2B)}$. [Ans: 12]

2. Solving system of linear equations

"What we know is a drop, what we don't know is an ocean."

— Isaac Newton

We have already solved some equations for one variable , like, $x + 5 = 3$ etc. and also know that when a equation is said to be linear (an equation $f(x) = c$, $f(x)$ has degree one) or quadratic (an equation $f(x) = c$, $f(x)$ has degree two) or cubic (an equation $f(x) = c$, $f(x)$ has degree three) etc.

In previous section we have gone through vectors and matrices, in this section we are going solve equations which contains more than one variables with one or more than one equations and all equations are of one degree (called linear equations) (i.e. let a system of linear equations in two variables and two equations is $f(x, y) = c$ and $g(x, y) = c'$, where $c, c' \in \mathbb{R}$, and $f(x, y)$ and $g(x, y)$ are one degree polynomials). With the help of matrix, we will solve the system of linear equations and will know when a system of linear equations has a unique solution, no solution and infinitely many solutions and we will visualize it geometrically.

Lets understand it more precisely with an example.

Items	Buyer A	Buyer B	Buyer C
Rice (in Kg)	8	12	3
Dal (in Kg)	8	5	2
Oil (in Liter)	4	7	5

Example 2.0.1. Suppose there are three buyers A, B and C bought three items according to the given below table form a shop and buyer A paid ₹1960, buyer B paid ₹2215 and buyer C paid ₹1135 to the shopkeeper. Suppose buyer A wants to know the prices of rice and dal per kg. and oil per liter. He assumes the price of

rice is ₹x per kg., price of dal is ₹y per kg. and price of oil is ₹z per liter.

Question 30. How much amount paid by the buyer A to the shopkeeper if he bought 8 kg. rice, 8 kg. dal and 4 litre oil?
Total amount $=$ ₹$8x + 8y + 4z$, which is equal to ₹1960
i.e., $8x + 8y + 4z = 1960$.
This is an equation with three variables x, y and z and left side of the equation is a polynomial in three variables of degree one.

Question 31. How much amount paid by the buyer B to the shopkeeper if he bought 12 kg. rice, 5 kg. dal and 7 litre oil?
Total amount $=$ ₹$12x + 5y + 7z$, which is equal to ₹2215
i.e., $12x + 5y + 7z = 2215$.
This is also an equation with three variables x, y and z and left side of the equation is a polynomial in three variables of degree one.

Question 32. How much amount paid by the buyer C to the shopkeeper if he bought 3 kg. rice, 2 kg. dal and 5 litre oil?
Total amount $=$ ₹$3x + 2y + 5z$, which is equal to ₹1135
i.e., $12x + 5y + 7z = 1135$.
This is also an equation with three variables x, y and z and left side of the equation is a polynomial in three variable of degree one.

Hence we have a following system of linear equations

$$8x + 8y + 4z = 1960$$

$$12x + 5y + 7z = 2215$$

$$12x + 5y + 7z = 1135$$

Observe that all three equations have degree one and it involve three variables x, y and z.

2.1 Linear equation

A linear equation is an equation of the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

where $x_1, x_2, \ldots, x_n$ are the variables (or unknowns) and $a_1, a_2, \ldots, a_n$ are the coefficients which are real numbers, b is also a real number.

Example 2.1.1. $8x + 8y + 4z = 1960$ is a linear equation, where x, y and z are variables and 8, 8 and 4 are the coefficients.

Question 33. Which of the followings is/are a linear equation?

- Option 1: $2x^2 + y = 1$
- **Option 2:** $x + 2y = 1$
- **Option 3:** $x + y + z = 1$
- Option 4: $3x^2 + 3y^2 + z^2 = 1$

2.2 System of linear equations

A system of linear equations is a collections of one or more linear equations involving the set of variables, for example

$$8x + 8y + 4z = 1960$$

$$12x + 5y + 7z = 2215$$

$$12x + 5y + 7z = 1135$$

is a system of linear equations with three variable x, y and z. A solution of the above system of linear equations is an assignment of values to the variables such that all equations are simultaneously satisfied. For the above system of linear equations, $x = 45$, $y = 125$ and $z = 150$ satisfies all three equations simultaneously.

A general system of m linear equations with n variables can be written as

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$
$$\ldots$$
$$\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

2.2.1 Matrix representation of a system of linear equations

A general system of m linear equations with n variables

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

can be written as matrix equation $Ax = b$, where A is matrix of order $m \times n$ and called *coefficient matrix*, x is a column vector with n entries, and b is a column vector with m entries as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Example 2.2.1. Consider a system of linear equations:

$$
\begin{aligned}
-2x_1 + 3x_2 + x_3 &= 1 \\
-x_1 + x_3 &= 0 \\
2x_2 &= 5
\end{aligned}
$$

The matrix representation of system is $Ax = b$, where $A = \begin{bmatrix} -2 & 3 & 1 \\ -1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$ $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$

Question 34. Consider a system of equations:

$$
\begin{aligned}
2x_1 + 3x_2 &= 6 \\
-2x_1 + 2x_2 &= 3 \\
4x_1 + 6x_2 &= 12
\end{aligned}
$$

1) Find the matrix representation $Ax = b$, where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of system of linear equations.

2) What is the order of matrix A?

2.3 Solution of a system of linear equations

A system of linear equations may have

1. a unique solution.
2. infinitely many solutions.
3. no solution.

Let understand it with 3 examples

Example 2.3.1. Consider a system of linear equations

$$x + y = 1$$
$$x - y = 1$$

Observe that assigned values $x = 1$ and $y = 0$ is the only values of x and y satisfying the both equations simultaneously. In this case, we say that the above system of linear has a unique solution.
Also, there are only two variables so we can relate it with coordinate system ($X-$ axis and $Y-$ axis) and can visualize it geometrically as follow:

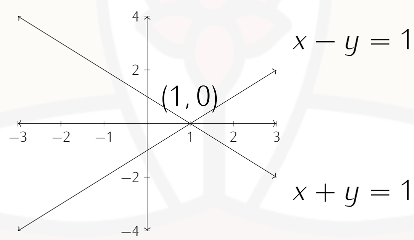


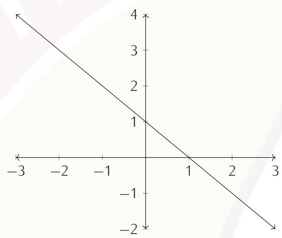
Figure 2.1:

In figure we see that these two lines are intersecting at only one point (1,0).
In general if there is a system of linear equations in two variables then we always related it with coordinate system (XY- plane) and equations represent lines in the plane. Suppose there is a system of linear equations in two variables with three equations or more than three equations and the system has a unique solution then all equations represents lines in the plane and all lines pass through only one point which will represent the solution point.

Example 2.3.2. Consider a system of linear equations

$$x + y = 1$$
$$2x + 2y = 2$$

Observe that assigned values $x = 1$ and $y = 0$ is a values of x and y satisfying the both equations simultaneously, also, $x = 2$ and $y = -1$ is another values of x and y satisfying the both equations simultaneously, similarly, there are infinitely values of x and y satisfying the both equations simultaneously. In this case, we say that the above system of linear has infinitely solutions.
We can visualize it geometrically. Observe that both equations represents the same line $x + y = 1$.



$x + y = 1$ or $2x + 2y = 2$

Figure 2.2:

In figure, we can see that both equations represent only one line $x + y = 1$ so any point on the line will be a solution of the system of linear equation and a line has infinitely many points so this system of linear has infinitely many solutions.
In general if there is a system of linear equations in two variables then we always related it with coordinate system (XY- plane) and equations represent lines in the plane. Suppose there is a system of linear equations in two variables with three equations or more than three equations and the system has infinitely many solutions then all equations represent only one line and this shows (as we argued above) that system of lines has infinitely many solutions.

Example 2.3.3. Consider a system of linear equations

$$x + y = 1$$
$$x + y = 0$$

Since left side of both equations are the same this implies $1 = 0$ which is an absurd. That means there is no possible values of x and y which satisfying both equations simultaneously. In this case, we say that the above system of linear has no solution. Lets see it geometrically.

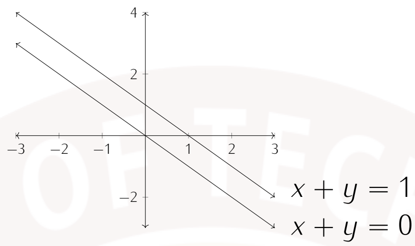


Figure 2.3:

In figure we see that these two lines are not intersecting i.e., there is no point which lines on both lines that means there is no possible value of x and y which satisfies both equations simultaneously and so there is no solution for the system of linear equations.

Example 2.3.4. Find solution of the system of linear equations

$$2x + 3y = 2$$
$$3x - 5y = 1$$

Multiply both sides of the equation $2x + 3y = 2$ by 3, we get equation (1) as

$$6x + 9y = 6 \ldots (1)$$

Multiply both sides of the equation $3x - 5y = 1$ by 2, we get equation (2) as

$$6x - 10y = 2 \ldots (2)$$

Subtract equ (1) from equ (2), we get

$$-19y = -4 \implies y = \frac{4}{19}$$

Substitute $y = \frac{4}{19}$ in equation (1) or (2), we get $x = \frac{78}{114}$

So, the above system of linear equations has a unique solution and solution is $x = \frac{78}{114}$, $y = \frac{4}{19}$

The matrix representation of the above system of linear equations is

$$\begin{bmatrix} 2 & 3 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

and solution is column vector $\begin{bmatrix} \frac{78}{119} \\ \\ \frac{4}{19} \end{bmatrix}$

Lets see some example when a system of linear equations has three variables with one or more than one equations

Example 2.3.5. Consider a system of linear equations

$$x + y + z = 1$$
$$x + y + z = 5$$

What can we say about the nature of solution of system of linear equations?

If there is a system of linear equation in three variables, then we can related it with coordinate system $X-$ axis, $Y-$ axis and $Z-$ axis, and each equation in coordinate system represents a plane as we can see in following figure.
equation $x + y + z = 1$ represent blue colored plane and equation $x + y + z = 5$ represents red colored plane in the coordinate system. As in figure, both planes are not intersecting that means the above system of linear equations has no solution.

Algebraically, we can see that left side of the both equations are equal in the system of linear equations, so we can equate right side of the both equations we get 1=5 which is absurd, so we can also conclude that the above system of linear equations no solution.

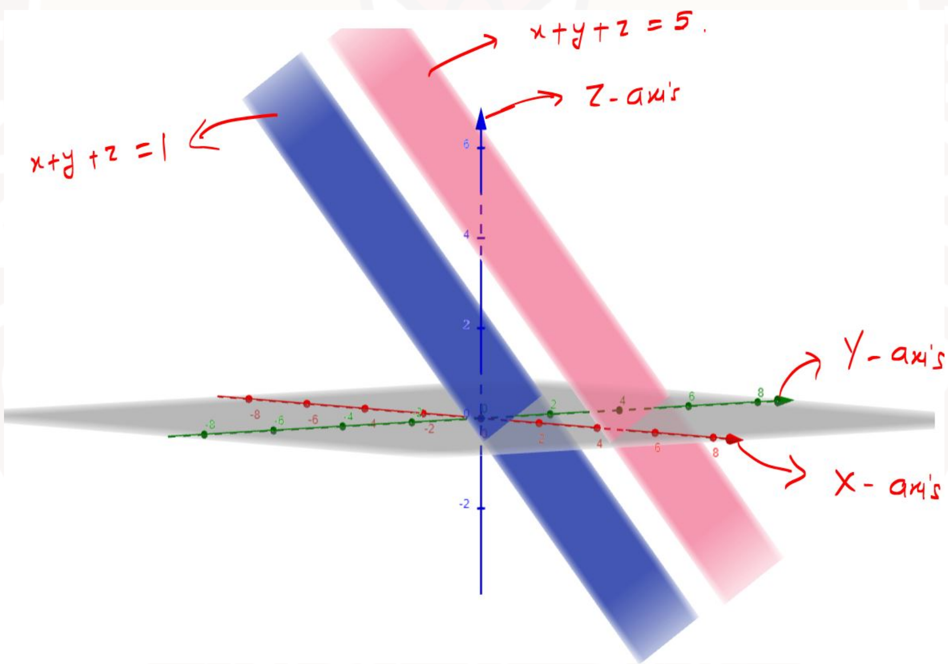


Figure 2.4:

Example 2.3.6. Consider a system of linear equations

$$x + y + z = 1$$
$$x - y + z = 5$$

What can we say about the nature solution of system of linear equations?

As we can see in figure, equation $x+y+z = 1$ represent blue colored plane and equation $x - y + z = 5$ represents red colored plane in the coordinate system. In figure, we can see that both planes are intersecting and intersection is represented by a line which is in yellow color i.e., there are infinitely many points on yellow colored line which also lie on both planes, algebraically we can say there are infinitely many values of x, y and z (points on the yellow colored line) which satisfy both equations simultaneously and so the above system of linear equation has infinitely many solutions.

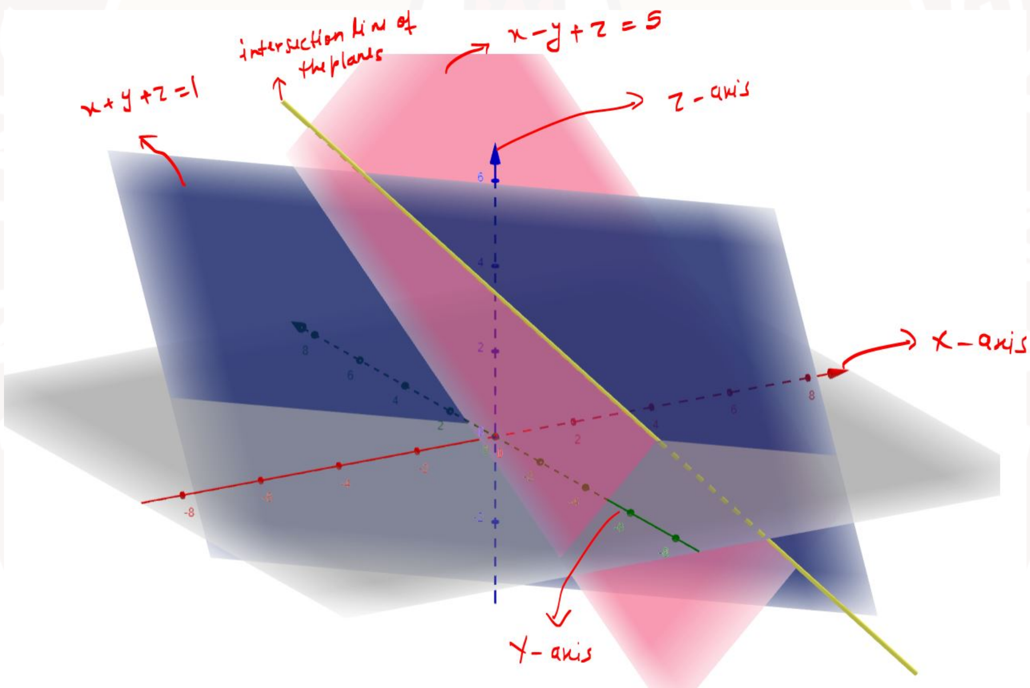


Figure 2.5:

We can see here also,

We can write the first equation as

$$x + y + z = 1 \implies y + z = 1 - x$$

substitute $y + z = 1 - x$ in second equation, we get,

$$x - (1 - x) = 5 \implies x = 3$$

So we get,

$$x = 3 \text{ and } y + z = 1$$

Hence if $y = 1$ we get $z = 0$ or $y = 0$ we get $z = 1$ and many more.

So $x = 3, y = 1$ and $z = 0$
and $x = 3, y = 0$ and $z = 1$ are the solution of the above system of linear equations, similarly we can get infinitely many solution of the above system of linear equations. (We will see it more precisely as we go ahead)

Example 2.3.7. Consider a system of linear equations

$$
\begin{aligned}
x + y + z &= 1 \\
y &= -2 \\
x &= 1
\end{aligned}
$$

What can we say about the nature solution of system of linear equations?

As we can see in figure, equation $x + y + z = 1$ represent blue colored plane, equation $x = 1$ represents red colored plane and $y = -2$ represents green colored plan in the coordinate system. In figure, we can see that all three planes are intersecting at point which represented by a black colored dot point i.e., there is only one point (1,-2,2) which lies on all three planes, algebraically we can say there is only one value of x, y and z (i.e., $x = 1, y = -2, z = 2$) which satisfy all three equations simultaneously and so the above system of linear equation has a unique solution.

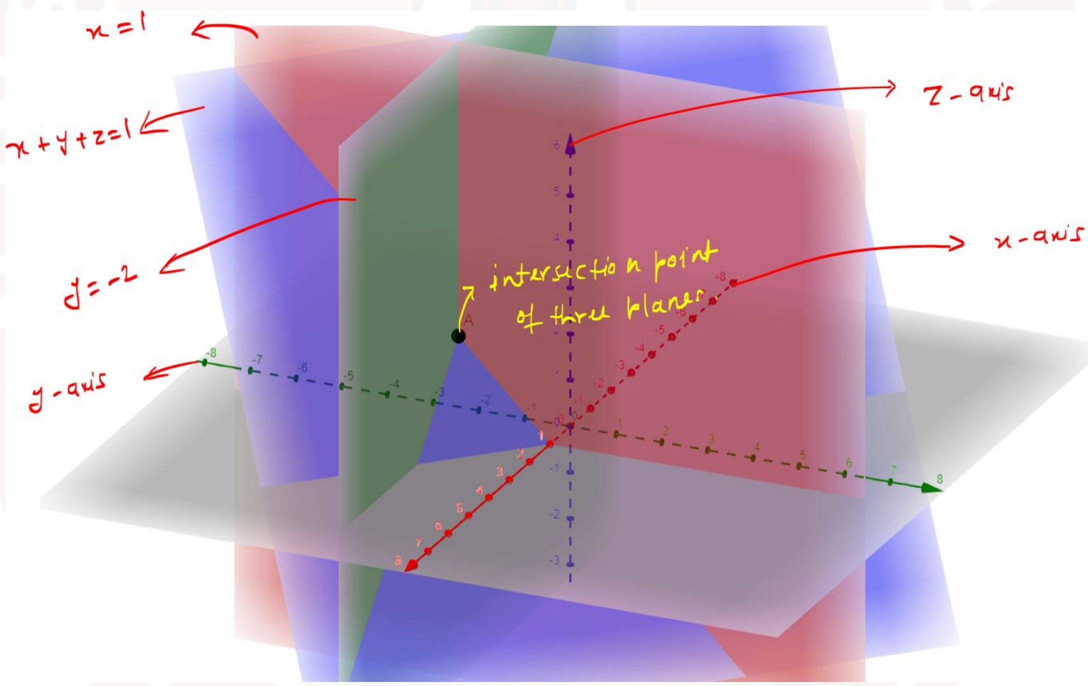


Figure 2.6:

we can see it here also
As system of linear equation, we have as a second equation $y = -2$ and third

equation as $x = 1$, substitute this in the first equation we get, $z = 2$.

So $x = 1, y = -2$ and $z = 2$ is a unique solution of the above system of linear equations.

Question 35. Consider a system of equations:

$$\begin{aligned} 2x_1 + 3x_2 &= 6 \\ -2x_1 + kx_2 &= d \\ 4x_1 + 6x_2 &= 12 \end{aligned}$$

Choose the set of correct options.

[Hint: Observe that third equation is a multiple of the first one (Dividing by 2 from both the sides of the third equation gives the first equation). So it is enough to check the solutions for the first and second equation.]

- **Option 1:** $Ax = b$ represents the above system of linear equations, where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $A = \begin{bmatrix} 2 & 3 \\ -2 & k \\ 4 & 6 \end{bmatrix}$, and $b = \begin{bmatrix} 6 \\ d \\ 12 \end{bmatrix}$
- **Option 2:** The system has no solution if $k = -3$, $d = 0$.
- **Option 3:** The system has a unique solution if $k = 3$, $d = 0$.
- Option 4: The system has infinitely many solutions if $k = -3$, $d = 6$.
- **Option 5:** The system has infinitely many solutions if $k = -3$, $d = -6$.

Feedback:

- **For Option 2:** If $k = -3$ and $d = 0$, then by comparing with the first equation we get $6 = 0$, which is absurd. So there can not be any x_1 and x_2, satisfying both the first and second equation together.
- **For Option 4:** If $k = -3$ and $d = 6$, then by comparing with the first equation we get $6 = -6$, which is absurd. So there can not be any x_1 and x_2, satisfying both the first and second equation together.
- **For Option 5:** If $k = -3$ and $d = -6$, then the second equation is a multiple of the first equation. The third equation is also a multiple of the first equation. So all of them have the same set of solutions and there are infinitely many values of x_1 and x_2 which satisfy the first equation.

2.3.1 Exercise

Question 36. Choose the set of correct options.
[Hint: Think of A as a 2×2 or 3×3 matrix, and b accordingly.]

- **Option 1:** Every system of linear equations has either a unique solution, no solution or infinitely many solutions.
- Option 2: If each equation of a system of linear equations is multiplied by a non-zero constant c, then the solution of the new system of equations is c times the solution of the old system of equations.
- **Option 3:** If $Ax = b$ is a system of linear equations which has a solution, then the system of linear equations $cAx = b$, where $c \neq 0$, will also have a solution.
- **Option 4:** If $Ax = b$ is a system of linear equations which has a solution, then $\frac{1}{c}Ax = b$, where $c \neq 0$, will also have a solution.

Question 37. The Plane 1 and Plane 2 in Figure below, correspond to two different linear equations, which form a system of linear equations.

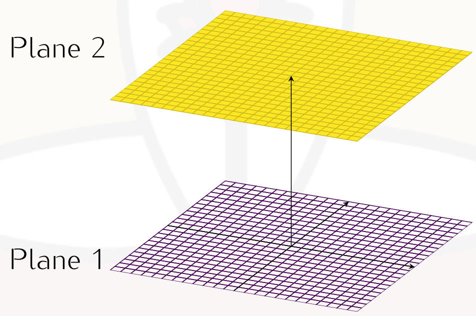


Figure 2.7:

The above system of linear equations has
[Hint: If a system of linear equations has a solution, then the point corresponding to the solution must lie on each plane corresponding to each linear equation of the given system.]

- Option 1: a unique solution.
- **Option 2:** no solution.
- Option 3: infinitely many solutions.
- Option 4: None of the above.

Question 38. Let x_1 and x_2 be solutions of the system of linear equations $Ax = b$. Which of the following options are correct?

- Option 1: $x_1 + x_2$ is a solution of the system of linear equations $Ax = b$.
- **Option 2:** $x_1 + x_2$ is a solution of the system of linear equations $Ax = 2b$.
- Option 3: $x_1 - x_2$ is a solution of the system of linear equations $Ax = b$.
- **Option 4:** $x_1 - x_2$ is a solution of the system of linear equations $Ax = 0$.

Question 39. Let v be a solution of the systems of linear equations $A_1x = b$ and $A_2x = b$. Which of the following options are correct ?

- Option 1: v is a solution of the system of linear equations $(A_1 + A_2)x = b$.
- **Option 2:** v is a solution of the system of linear equations $(A_1 + A_2)x = 2b$.
- **Option 3:** v is a solution of the system of linear equations $(A_1 - A_2)x = 0$.
- Option 4: v is a solution of the system of linear equations $(A_1 - A_2)x = b$.

Question 40. Consider a system of equations:

$$x_1 - 3x_2 = 4$$
$$3x_1 + kx_2 = -12$$

Where $k \in \mathbb{R}$. If the given system has a unique solution, then k should not be equal to [Ans: -9]

Question 41. Consider two system of linear equations:
System 1:

$$x_1 + 2x_2 = -5$$
$$0x_1 - x_2 = 5$$

System 2:

$$-2x_3 + x_4 = -5$$
$$3x_3 + x_4 = 5$$

Suppose there is another system of linear equations given by

$$(1 - 2)(x_1 + x_3) + (2 + 1)(x_2 + x_4) = m$$
$$(0 + 3)(x_1 + x_3) + (-1 + 1)(x_2 + x_4) = n$$

for some real values of m and n. Find the value of $n - m$. [Ans: 46]

2.4 Cramer's Rule

In previous section, I have gone through the matrices, types of matrices and determinant of a matrix. We have also gone through system of linear equations and know types of solution of system of linear equations i.e. no solution, unique solution and infinitely many solutions.

In this section we will try to solve a system of linear equations which has a unique solution Using Cramer's rule. Cramer's rule is an algorithm which will be used to find solve of system of linear equations. But every system of linear need not be solved using Cramer's rule. Cramer's rule can be applicable to some specific types of system of linear equations those are as follows:

1. Coefficient matrix of the system of linear equations should be a square matrix.
2. Determinant of the coefficient matrix of the system of linear equations should be non-zero i.e Coefficient matrix should be invertible.

2.4.1 Cramer's rule for invertible coefficient matrix of order 2

Consider a system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

matrix representation of the above system be $Ax = b$, where $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$.

Steps to find the solution of above system of linear equations:

Step-1 Find the determinant of the matrix A, $det(A) = a_{11}a_{22} - a_{12}a_{21}$

Step-2 Define A_{x_1}, where A_{x_1} be the matrix obtained from A by replacing the first column of A with the column matrix b, i.e.,

$$A_{x_1} = \begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}$$

and find the determinant of the matrix A_{x_1}, $det(A_{x_1}) = b_1a_{22} - b_2a_{12}$

Step-3 Define A_{x_2}, where A_{x_2} be the matrix obtained from A by replacing the second column of A with the column matrix b i.e.,

$$A_{x_2} = \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix}$$

and find the determinant of the matrix A_{x_2}, $det(A_{x_2}) = b_2a_{11} - b_1a_{21}$

Step-4 $x_1 = \dfrac{det(A_{x_1})}{det(A)} = \dfrac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}$ and $x_2 = \dfrac{det(A_{x_2})}{det(A)} = \dfrac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$

Example 2.4.1. Consider the system of linear equations

$$2x_1 + x_2 = 1$$
$$3x_1 + 4x_2 = -1$$

Find the solution of the system of linear equations.

Matrix representation of the above system of linear equations is $Ax = b$, where $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$det(A) = 2.4 - 1.3 = 8 - 3 = 5$$

Lets find A_{x_1},

$$A_{x_1} = \begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix}$$

so

$$det(A_{x_1}) = 4.1 - 1.(-1) = 4 + 1 = 5$$

Now, lets find A_{x_2},

$$A_{x_2} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}$$

so

$$det(A_{x_2}) = 2.(-1) - 3.1 = -2 - 3 = -5$$

So finally,

$$x_1 = \frac{det(A_{x_1})}{det(A)} = \frac{5}{5} = 1$$

and

$$x_2 = \frac{det(A_{x_2})}{det(A)} = \frac{-5}{5} = -1$$

2.4.2 Cramer's rule for invertible coefficient matrix of order 3

Consider a system of linear equations as follows:

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

Let the matrix representation of the above system be $Ax = b$, where $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

Steps to find the solution of above system of linear equations:

Step-1 Find the determinant of the matrix A.

Step-2 Define A_{x_1} , where A_{x_1} be the matrix obtained from A by replacing the first column of A with the column matrix b, i.e.,

$$A_{x_1} = \begin{bmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{bmatrix}$$

and find the determinant of the matrix A_{x_1}

Step-3 Define A_{x_2} , where A_{x_2} be the matrix obtained from A by replacing the second column of A with the column matrix b, i.e.,

$$A_{x_2} = \begin{bmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{bmatrix}$$

and find the determinant of the matrix A_{x_2}

Step-4 Define A_{x_3} , where A_{x_3} be the matrix obtained from A by replacing the third column of A with the column matrix b, i.e.,

$$A_{x_3} = \begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{bmatrix}$$

and find the determinant of the matrix A_{x_3}

step-5 $x_1 = \dfrac{det(A_{x_1})}{det(A)}$, $x_2 = \dfrac{det(A_{x_2})}{det(A)}$ and $x_3 = \dfrac{det(A_{x_3})}{det(A)}$

Example 2.4.2. Consider a system of linear equations

$$\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ 2x_1 + x_3 - 2x_3 &= -1 \\ -x_1 - 2x_2 + 4x_3 &= 1 \end{aligned}$$

Find the solution of the system of linear equations.

Matrix representation of the above system of linear equation is $Ax = b$, where

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -2 \\ -1 & -2 & 4 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } b = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Determinant of the matrix A, $det(A) = 3$

Lets find the A_{x_1},

$$A_{x_1} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -2 \\ 1 & -2 & 4 \end{bmatrix}$$

so

$$det(A_{x_1}) = -1$$

Now, lets find the A_{x_2},

$$A_{x_2} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -2 \\ -1 & 1 & 4 \end{bmatrix}$$

so

$$det(A_{x_2}) = -7$$

And

$$A_{x_3} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -1 \\ -1 & -2 & 1 \end{bmatrix}$$

and so

$$det(A_{x_3}) = -3$$

So finally

$$x_1 = \frac{det(A_{x_1})}{det(A)} == -\frac{1}{3}, \ x_2 = \frac{det(A_{x_2})}{det(A)} = -\frac{7}{3} \text{ and } x_3 = \frac{det(A_{x_3})}{det(A)} = \frac{-3}{3} = -1$$

2.4.3 Cramer's rule for invertible coefficient matrix of order n

Consider a system of linear equations

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

matrix representation of the above system of linear equation is $Ax = b$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Steps to find the solution of above system of linear equations:

Step-1 Find the determinant of the matrix A.

Step-2 Define A_{x_i} , where A_{x_i} be the matrix obtained from A by replacing the ith column of A with the column matrix b, and find the determinant of the matrix $A_{x_i}, i = 1, 2, \ldots, n$

step-3 $x_i = \dfrac{det(A_{x_i})}{det(A)}, i = 1, 2, \ldots, n$

2.4.4 Exercise

Question 42. Consider a system of linear equations

$$
\begin{aligned}
x_1 + x_3 &= 1\\
-x_1 + x_2 - x_3 &= 1\\
-x_2 + x_3 &= 1
\end{aligned}
$$

Let matrix representation of the above system be $Ax = b$, where $A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Let A_{x_i} be the matrix obtained by replacing the i-th column of A (i.e., $\begin{bmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \end{bmatrix}$) by b, for $i = 1, 2, 3$.

Use the above information to answer questions 1 and 2.

Question 1: Choose the set of correct options

- Option 1: $A_{x_1} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$
- **Option 2:** $A_{x_1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$
- Option 3: $A_{x_2} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$
- **Option 4:** $A_{x_3} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$

Question 2: Choose the set of correct options.

- **Option 1:** $x_1 = -2$.
- Option 2: $x_2 = -2$.
- **Option 3:** $x_3 = 3$.

- Option 4: None of the above.

Question 43. Consider the system of linear equations $Ax = b$, where $A = \begin{bmatrix} 1 & a & 0 \\ a & 1 & 2 \\ 1 & 0 & a \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ 0 \\ x_3 \end{bmatrix}$, and $b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, and the solution for x is partially known.

What is the value of a^2, if $a \neq 0, \sqrt{3}, -\sqrt{3}$ is given?

[Hint: Observe that the second row of the vector x is given as 0, which implies that x_2 is known, where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$] [Answer: 2]

Question 44. Consider the system of linear equations $Ax = b$, where $A = \begin{bmatrix} 1 & 2a+3 \\ 3a+2 & 1 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$, and $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and the solution for x is partially known.

What is the value of $3a$, if $det(A) \neq 0$? [Answer: -1]

2.5 Finding the solution of a system of linear equations with an invertible coefficient matrix

In previous section we just studied, finding a unique solution of a system of linear equations using Cramer's rule. In this section, we will learn another method to find a unique solution of a system of linear equations using inverse matrix of the coefficient matrix.

Inverse of the coefficient matrix method e can be applicable to some specific types of system of linear equations those are as follows:

1. Coefficient matrix of the system of linear equations should be a square matrix.
2. Determinant of the coefficient matrix of the system of linear equations should be non-zero i.e Coefficient matrix should be invertible.

Consider a system of linear of linear equations $Ax = b$ where A is an invertible matrix. Then solution of the system of linear equations can be obtain as follows:

$$Ax = b$$

Pre- multiplication of A^{-1} both sides, we get,

$$A^{-1}Ax = A^{-1}b \implies x = A^{-1}b$$

Steps to find the solution of above system of linear equations using inverse of the coefficient matrix :
Consider a system of linear equations

$$\begin{gathered}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\ldots \\
\ldots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n
\end{gathered}$$

matrix representation of the above system of linear equation is $Ax = b$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Step-1 Find the inverse of the matrix A.

Step-2 Find the matrix multiplication of A^{-1} with b, i.e. $A^{-1}b$ which is a column matrix.

step-3 Compare with the column matrix x and find the value of $x_1, x_2, \dots x_n$

Example 2.5.1. Consider a system of linear equations

$$x_1 + 2x_2 = 1$$
$$3x_1 + 5x_2 = 2$$

Find the solution of the above system of linear equations.

Matrix representation of the above system of linear equations is $Ax = b$, where $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Let's find the inverse of the coefficient matrix

$$adj(A) = \begin{bmatrix} 5 & -2 \\ -3 & 1 \end{bmatrix}$$

$$det(A) = -1$$

So

$$A^{-1} = \frac{adj(A)}{det(A)} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1}b = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Hence $x_1 = -1$ and $x_2 = 1$

Example 2.5.2. Consider a system of linear of linear equations

$$\begin{aligned} 8x_1 + 8x_2 + 4x_3 &= 1960 \\ 12x_1 + 5x_2 + 7x_3 &= 2215 \\ 3x_1 + 2x_2 + 5x_3 &= 1135 \end{aligned}$$

Matrix representation of the above system of linear equations is $Ax = b$, where

$$A = \begin{bmatrix} 8 & 8 & 4 \\ 12 & 5 & 7 \\ 3 & 2 & 5 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } b = \begin{bmatrix} 1960 \\ 2215 \\ 1135 \end{bmatrix}$$

As we see that $det(A) = -188 \neq 0$ so matrix A is invertible.
Inverse of the A is

$$A^{-1} = \frac{1}{-188} \begin{bmatrix} 11 & -32 & 36 \\ -39 & 28 & -8 \\ 9 & 8 & -56 \end{bmatrix}$$

Hence

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{-188} \begin{bmatrix} 11 & -32 & 36 \\ -39 & 28 & -8 \\ 9 & 8 & -56 \end{bmatrix} \begin{bmatrix} 1960 \\ 2215 \\ 1135 \end{bmatrix} = \begin{bmatrix} 45 \\ 125 \\ 150 \end{bmatrix}$$

So

$$x_1 = 45, x_2 = 125 \text{ and } x_3 = 150$$

2.5.1 Exercise

Question 45. Consider a system of linear equations

$$\begin{aligned} 2x_1 - x_2 &= 3 \\ x_1 - x_3 &= 3 \\ x_2 - x_3 &= 2 \end{aligned}$$

Let the matrix representation of the above system be $Ax = b$, where $A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $b = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}$. Use the above information to answer questions 1 and 2.

Question 1: Choose the set of correct options.

- Option 1: $A^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$
- **Option 2:** $A^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & 2 \\ 1 & -2 & 1 \end{bmatrix}$
- **Option 3:** Adjoint of the matrix A is $\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & 2 \\ 1 & -2 & 1 \end{bmatrix}$
- **Option 4:** $det(A) = 1$.

Question 2: Choose the set of correct options.

- Option 1: $x_1 = -2$.
- **Option 2:** $x_2 = 1$.
- **Option 3:** $x_3 = -1$
- Option 4: None of the above.

2.6 The Gauss elimination method

In previous sections, we have learned some method to solve a system of linear equations when a system of linear equations has a unique solution (i.e., when the coefficient matrix of the system of linear equations is an square matrix and has determinant non zero).

But what will be the case when a system of linear equations has following things?

1. when the coefficient matrix of the system of linear equations is an square matrix and has determinant zero.

2. A system of linear equations consists of m equations and n variables i.e. a system of linear equations whose coefficient matrix is an rectangular matrix.

Or how do we know an algebraic method to get that given system of linear equations has unique solution, no solutions or infinitely many solutions.

To solve such kind of system of linear equations there is a method called " Gauss elimination method"

To learn such method lets see some forms of matrices

1. The echelon form or row echelon form

2. The Reduced Row echelon form.

2.6.1 Homogeneous and Non-homogeneous system of linear equations

Homogeneous system of linear equations:

A system of linear equations of type

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0$$

$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0$$

$$\ldots$$

$$\ldots$$

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0$$

matrix representation of the above system of linear equation is $Ax = 0$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

is called homogeneous system of linear equations.

It is obvious that 0 is always a solution of the system of linear equations which is called **trivial** solution of the system.

For a homogeneous system of linear equations there are always two possibilities:

1. 0 is the unique solution.
   Reason: 1- If coefficient matrix of the homogeneous system of linear equations is invertible, then $Ax = 0 \implies A^{-1}A = A^{-1}0$
   (Pre-multiplication of A^{-1} both sides of the equation) $\implies x = 0$.
   2- In a homogeneous system of linear equations, if number of equations is greater than number of variables then the system of linear equations has a unique solution which is 0. This can be get using Gauss elimination method.

2. There are infinitely many solutions other than 0.
   Reason: In a homogeneous system of linear equations, if number of variables is greater than number of equations then the system of linear equations has infinitely many solutions other than 0. This can be get using Gauss elimination method.

Non-homogeneous system of linear equations:

A system of linear equations of type

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$
$$\ldots$$
$$\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

matrix representation of the above system of linear equation is $Ax = b$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mathrm{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

in which any $b_i \neq 0$, then the system of this type is called non-homogeneous system of linear equations.

For a non- homogeneous system of linear equations there are always three possibilities:

1. There is the unique solution.
2. There are infinitely many solutions.
3. There is no solution.

All the above solutions can be decided using Gauss elimination method.

2.6.2 The row echelon form and reduced row echelon form

A matrix is in row echelon form if :

- The first non-zero element (the leading entry) in a row is 1.
- The column containing the leading 1 of a row is to the right of the column containing the leading 1 of the row above it. In different words, all subsequent non-zero rows which also have their leading entries (i.e. first non-zero entries) as 1 and they should appear to the right of the leading entry in the previous row.
- Any non-zero rows are always above rows with all zeros.

A matrix is in reduced row echelon form if :

- The first non-zero element in the first row (the leading entry) is the number 1.
- The column containing the leading 1 of a row is to the right of the column containing the leading 1 of the row above it. In different words, all subsequent non-zero rows which also have their leading entries (i.e. first non-zero entries) as 1 and they should appear to the right of the leading entry in the previous row.

- The leading entry in each row must be the only non-zero number in its column.
- Any non-zero rows are always above rows with all zeros.

Note: Any matrix which is in reduced row echelon form is also in row echelon form.

Example 2.6.1. Choose the correct options for the matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

- Option 1: A is in row echelon form but not in reduced row echelon form.
- Option 2: A is in both row echelon form and reduced row echelon form.
- Option 3: A is neither in row echelon form nor reduced row echelon form.

Lets see some observations in the matrix:

1. First non-zero entry (element) in the first row is 1 which is the entry (1,1)-th (first row is the top most row and first column is left most column) and which is also called leading 1 of first row. $A = \begin{bmatrix} \boxed{1} & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
2. First non-zero entry in the second row is 1 which is entry (2, 3)-th and which is also called leading 1 of second row. $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{bmatrix}$
3. Third row is a zero row and so there is no non-zero entry or leading entry. And third row the below to all above non zero row.
4. Column containing leading 1 (entry (1, 1)-th) of the first row is the first column and this is the only non -zero entry in the first column. See the matrix and corresponding column $\begin{bmatrix} \boxed{1} & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
5. Column containing leading 1 (entry (2, 3)-th) of the second row is the third column and this is the only non -zero entry in the third column. See the matrix and corresponding column $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{bmatrix}$

6. Column containing leading 1 (entry (2, 3)-th) of the second row is the third column which is right of the column containing leading 1 (entry (1, 1)-th) of the first row. $\begin{bmatrix} \boxed{1} & 2 & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{bmatrix}$

These observations infer that matrix A satisfies all conditions of row and reduced row echelon form. so matrix is A is in row echelon form and reduced row echelon form.

Example 2.6.2. Choose the correct options for the matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

- Option 1: A is in row echelon form but not in reduced row echelon form.
- Option 2: A is in both row echelon form and reduced row echelon form.
- Option 3: A is neither in row echelon form nor reduced row echelon form.

Lets see some observations in the matrix:
Note: There is no non-zero row.

1. First non-zero entry (element) in the first row is 1 which is the entry (1,2)-th (first row is the top most row and first column is left most column) and which is also called leading 1 of first row . $\begin{bmatrix} 0 & \boxed{1} & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
2. First non-zero entry in the second row is 1 which is entry (2, 1)-th and which is also called leading 1 of second row. $\begin{bmatrix} 0 & 1 & 0 \\ \boxed{1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
3. First non-zero entry in the third row is 1 which is entry (3, 2)-th and which is also called leading 1 of third row. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & \boxed{1} & 0 \end{bmatrix}$
4. Column containing leading 1 (entry (2, 1)-th) of the second row is in the first column which is left of the column containing leading 1 (entry (1, 1)-th) of the first row. $\begin{bmatrix} 0 & \boxed{1} & 0 \\ \boxed{1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

Fourth observation infer that matrix A is not satisfying the second condition of row echelon form and reduced row echelon form. So matrix A is not in row echelon form and reduced row echelon form.

Example 2.6.3. Choose the correct options for the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- Option 1: A is in row echelon form but not in reduced row echelon form.
- Option 2: A is in both row echelon form and reduced row echelon form.
- Option 3: A is neither in row echelon form nor reduced row echelon form.

Lets see some observations in the matrix:
Note: There is no non-zero row.

1. First non-zero entry (element) in the first row is 1 which is the entry (1,2)-th (first row is the top most row and first column is left most column) and which is also called leading 1 of first row. $\begin{bmatrix} \boxed{1} & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
2. First non-zero entry in the second row is 1 which is entry (2, 1)-th and which is also called leading 1 of second row. $\begin{bmatrix} 1 & \boxed{1} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
3. First non-zero entry in the third row is 1 which is entry (3, 2)-th and which is also called leading 1 of third row. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}$
4. Conditions for row echelon form
   (a) The first non-zero element (the leading entry) in a row is 1. This is also followed by A.
   (b) The column containing the leading 1 of a row is to the right of the column containing the leading 1 of the row above it. In different words, all subsequent non-zero rows which also have their leading entries (i.e. first non-zero entries) as 1 and they should appear to the right of the leading entry in the previous row. This is also followed by A.
   (c) Any non-zero rows are always above rows with all zeros. There is no zero row in matrix A so no need to think about this.

5. **So matrix A is in row echelon form**

6. But observe in the second column $\begin{bmatrix} 1 & \boxed{1} & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & 1 \end{bmatrix}$, entry (2, 1)-th which is 1 and is the leading entry in the second row and so the column containing the leading entry is second column which is in red color but it is not the only non zero number in its column (entry (1,2)-th have 1 also) which does not follows the 3rd condition for reduced row echelon form (but column first and third columns follows the 3rd condition for reduced row echelon form).

Observation 5 and 6 infer that matrix A is in row echelon form and but not in reduced row echelon form.

2.6.3 Exercise

Question 46. Let $A = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ be a matrix (where the first row denotes the top most row, and the ordering of the rows is in the order top to bottom). Among the given set of options, identify the correct statements.

- Option 1: The first non-zero element in the first row is 3.
- **Option 2:** The first non-zero element in the second row is 1.
- Option 3: There is a non-zero element in the third row.
- **Option 4:** Since there is a row with all elements as zero, $det(A) = 0$.

Question 47. Let $I_{3\times3}$ denote the identity matrix of order 3. Answer questions 1 and 2 about the set S defined as :

$$S = \left\{ A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \right.$$

$$E = \begin{bmatrix} 0 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, F = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, G = \begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\left. I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, J = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \ K = I_{3\times3}, \ L = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \right\}.$$

Question 1: If S_1 is the subset of S consisting of all the matrices in S that are in row echelon form, then choose the correct option from the following.

- Option 1: $S_1 = S$
- Option 2: $S_1 = \{A, B, C, D, F, G, I\}$
- Option 3: $S_1 = \{B, C, D, E, F, G, H, I, J, K, L\}$
- Option 4: $S_1 = \{C, D, E, F, G, H, I, J, K\}$
- Option 5: $S_1 = \{C, D, E, F, G, H, J, K, L\}$
- **Option 6:** $S_1 = \{C, D, F, G, H, J, K, L\}$

- Option 7: Cardinality of S_1 is 7.
- **Option 8:** Cardinality of S_1 is 8.

Question 2: If S_2 is the subset of S consisting of all the matrices in S that are in reduced row echelon form, then choose the correct option from the following.

- Option 1: $S_2 = \{A, C, D, F, G, L\}$
- Option 2: $S_2 = \{B, C, D, E, F, G, H, J, K\}$
- Option 3: $S_2 = \{C, D, G, H, J, K, L\}$
- **Option 4:** $S_2 = \{C, D, H, J, K, L\}$
- Option 5: $S_2 = \{C, D, H, I, J, K, L\}$
- Option 6: Cardinality of S_2 is 7.
- **Option 7:** Cardinality of S_2 is 6.

2.6.4 Solution of $Ax = b$ when A is in reduced row echelon form

Let $Ax = b$ be a system of linear equations where A is in reduced row echelon form.

Note: Consider the case when i^{th} row of A (A is in reduced row echelon form) is zero row and b_i entry of matrix b is non zero, in that case system of linear equations $Ax = b$ has no solution.

Lets see the reason of above note, consider the system of linear equations where coefficient matrix is in reduced row echelon form and i^{th} row is zero row and i^{th} entry of b which is $b_i \neq 0$, see in the following example

$$\begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \vdots & \vdots & \vdots \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix}$$

Now the corresponding system of linear equations form, from the above matrix form has an equation (which we can obtain after multiplying the matrices)

$$0x_1 + 0x_2 + \ldots + 0x_n = b_i \neq 0 \implies 0 \neq 0$$

which is absurd.

Hence when i^{th} row of A (A is in reduced row echelon form) is zero row and b_i entry of matrix b is non zero, in that case system of linear equations $Ax = b$ has no solution (If a system of linear equation has no solution then that system of linear equations is called **inconsistent**).

Question 48. Consider the following system of linear equations:

$$\begin{aligned} 0x_1 + x_2 + 0x_3 + 0x_4 &= 1 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= 1 \\ x_1 + x_2 + 0x_3 + 0x_4 &= 1 \\ 0x_1 + 0x_2 + x_3 + x_4 &= 1. \end{aligned}$$

Choose the the set of correct options.

- Option 1: The system of linear of equations has a solution.

- **Option 2:** The system of linear equations has no solution.
- **Option 3:** $det(A) = 0$, where A is the coefficient matrix of the given system of linear equations.
- Option 4: None of the above.

Dependent variable and independent variable

Let $Ax = b$ be a system of linear equations, where A is in reduced row echelon form.
Assume that for each zero row of A, the corresponding entry of b also 0 (i.e., the system of linear equations has a solution) (If a system of linear equation has a solution (it does not matter it has a unique solution or infinitely many solutions) then that system of linear equations is called **consistent**).

- If the i-th column has the leading entry of some row, we call x_i a dependent variable.
- If the i-th column does not have the leading entry of any row, we call x_i an independent variable.

Example 2.6.4. Consider a system of linear equations:

$$0x_1 + x_2 + 0x_3 + 0x_4 = 1$$
$$0x_1 + 0x_2 + x_3 + 0x_4 = 1$$

Which of the variables are dependent and independent variables?

Matrix representation of the above system of linear equation is

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Observe that, in the coefficient matrix $\begin{bmatrix} 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is in reduced row echelon form and entry (1,2)-th which is 1, is the leading entry in its column which is second column and so corresponding variable is x_2 so x_2 is dependent variable.
Similarly, we can see in the third column $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \end{bmatrix}$, entry (2, 3)-th which

is one, is the leading entry in the its column which is the third column and so the corresponding variable X_3 is the dependent variable.
And we can see in coefficient matrix $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ which is in the reduced row echelon form, first and second column don't have leading entry so the variables x_1 and x_4 are independent variables.

Example 2.6.5. let $Ax = b$ be a system of linear equations, where $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, and $b = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$. Find the number of independent and dependent variables?

Observe that given coefficient matrix is in the reduced row echelon form and entries (1,1)-th and (2,3)-th in the matrix are the only leading entries in its column and so corresponding variables x_1 and x_3, resp., are dependent variables. Entry (2, 4)-th is not leading entry as entry (2,3)-th is leading the second row. Second and fourth columns don't have leading entries so corresponding variables x_2 and x_4 are independent variables.

Use of dependent variable and independent variable

Let $Ax = b$ be a system of linear equation such that the coefficient matrix A is in the reduced row echelon form and from there we can get which variables will be the dependent or independent variables. Let x_i and x_j (or many more as we get) be independent variables in the system of linear equations then we can write the rest variables as some linear function of x_i and x_j variables and as x_i and x_j are independent variables so we can give any value to x_i and x_j (or to many more variables as we get) and can get value of rest variables after substituting in the corresponding function. i.e., Once A is in reduced row echelon form then it makes easier to find to find all solution of the system of linear equations.

Lets see it with an example.

Example 2.6.6. let $Ax = b$ be a system of linear equations, where $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, and $b = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$.

Observe that coefficient matrix A is in the reduced row echelon form.
Now,

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_3 + 2x_4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

After comparing the entries we get, system of linear equations in equation form,

$$x_1 + 0x_2 + 0x_3 + 0x_4 = 3$$
$$0x_1 + 0x_2 + x_3 + 2x_4 = 2$$

As one of the above example we have seen that x_1 and x_3 are dependent variables and x_2 and x_4 are independent variables so we can write the dependent variable x_1 as some linear function of x_2 and x_4, similarly, the dependent variable x_3.

$$x_1 = 3 - 0x_2 - 0x_4 = 3$$
$$x_3 = 2 - 2x_4 - 0x_1 - 0x_2 = 2 - 2x_4$$

From above we can see $x_1 = 3$, let $x_2 = 1$, and let $x_4 = 0$ we get $x_3 = 2$ so $\begin{bmatrix} 3 \\ 2 \\ 0 \\ 2 \end{bmatrix}$ is a solution of the above system of linear equations.

Let $x_4 = 2$ then $x_3 = 0$ and let $x_2 = 0$ then $\begin{bmatrix} 3 \\ 0 \\ 0 \\ 2 \end{bmatrix}$ is the another solution.

similarly we have x_2 and x_4 are independent variables so as we change the values of x_2 and x_4, we can get infinitely many solutions of the above system of linear equations.

Note: If a system linear equations is consistent and it has at least one independent variable then the system of linear equations have infinitely many solutions

2.6.5 Exercise:

Question 49. Consider a system of linear equations:

$$0x_1 + x_2 + 0x_3 + 0x_4 = 1$$
$$0x_1 + 0x_2 + x_3 + 0x_4 = 1$$

Choose the set of correct options.

[Hint: Recall the definitions of independent and dependent variable with respect to reduced row echelon form.]

- Option 1: x_1 and x_2 are dependent variables.
- **Option 2:** x_2 is a dependent variable.
- Option 3: x_3 and x_4 are independent variables.
- **Option 4:** x_4 is an independent variable.

Question 50. Suppose a system of linear equations consists of only one equation and four variables as follows:

$$x_1 + x_2 + x_3 + x_4 = a$$

where a is a constant. Find out the number of independent variables. [Answer: 3]

[Hint:"Think about" how many variables can be expressed in terms of the other variables in the given system of linear equations?]

Question 51. Let $[A|b]$ denote the augmented matrix of the system of linear equations

$$2x_1 + x_2 = 3$$

$$x_1 + 3x_2 = 4$$

where, $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, and $b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Let the matrix

$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \end{array}\right]$$

denote the reduced row echelon form of the augmented matrix corresponding to the system linear equations above. Which of the following option(s) is (are) correct?

- Option 1: The values of a and b cannot be determined from the given information.

- Option 2: $a = b$ but their exact values cannot be determined from the given information.
- **Option 3:** $a = b = 1$
- Option 4: $a = 2$ and $b = 3$
- Option 5: The solutions for x_1 and x_2 are not unique.
- **Option 6:** $x_1 = x_2 = 1$, and the system of linear equations has a unique solution.
- Option 7: $x_1 = x_2 = 1$ is the solution. However, it is not possible to determine whether the system of equations has a unique solution or not from the given information.

Question 52. Let $Ax = b$ be a matrix representation of a system of linear equations, where A is a 4×4 matrix, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, and $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$.

Let the reduced row echelon form of A be $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Which of the following options are correct?
[Hint: Recall the definitions of independent and dependent variable with respect to reduced row echelon form.]

- Option 1: x_1 is dependent on x_2.
- **Option 2:** x_3 is dependent on x_4.
- **Option 3:** x_2 is an independent variable.
- Option 4: The solution of the system of linear equations (if it exists) is unique.
- **Option 5:** There exist infinitely many solutions for the given system of linear equations.

Question 53. Let $A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ be a matrix (where the first row denotes the top most row, and the ordering of the rows is from top to bottom). Consider the system of linear equations given by $Ax = b$ where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $b = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$. Answer the following questions 1 and 2.

Question 1: Find the number of independent variables. [Answer: 1]

Question 2: If $x_1 = 0$ is given, then find out the number of solutions of the given system of linear equations. [Answer: 1]

2.6.6 Elementary row operations

Let A be a matrix. If we do some operation on the rows of the matrix A, like interchange any two rows, multiply a row with a real number, adding a row to another row or multiply a row with a real number and then add to another row such kind of operations used to called elementary row operations.

In this section we will see mainly three different types of elementary row operations.

The three different types of elementary row operations that can be performed on a matrix are:

Note: R_i denotes i^{th} row of a matrix.

- **Type 1:** Interchanging two rows.
  For example, interchanging R_1 and R_2

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}$$

- **Type 2:** Multiplying a row with some constant.

  For example, Multiply R_2 with $c \in \mathbb{R}$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{c.R_2} \begin{bmatrix} 1 & 2 & 3 \\ 4c & 5c & 6c \\ 7 & 8 & 9 \end{bmatrix}$$

- **Type 3:** Adding a scalar multiple of a row to another row.

  For example, Multiply R_2 with $c \in \mathbb{R}$ and then add to R_3

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_3 + c.R_2} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7+4c & 8+5c & 9+6c \end{bmatrix}$$

Example 2.6.7. Let $A = \begin{bmatrix} 5 & 2 & 9 \\ 2 & 5 & 6 \\ 9 & 8 & -1 \end{bmatrix}$. Find the matrices after performing the elementary operations $R_2 - 5R_1$, $R_1 \leftrightarrow R_3$ and $2R_3$.

$$\begin{bmatrix} 5 & 2 & 9 \\ 2 & 5 & 6 \\ 9 & 8 & -1 \end{bmatrix} \xrightarrow{R_2 - 5R_1} \begin{bmatrix} 5 & 2 & 9 \\ -23 & -10 & -39 \\ 9 & 8 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 2 & 9 \\ 2 & 5 & 6 \\ 9 & 8 & -1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 9 & 8 & -1 \\ 2 & 5 & 6 \\ 5 & 2 & 9 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 2 & 9 \\ 2 & 5 & 6 \\ 9 & 8 & -1 \end{bmatrix} \xrightarrow{2R_3} \begin{bmatrix} 5 & 2 & 9 \\ 2 & 5 & 6 \\ 18 & 16 & -2 \end{bmatrix}$$

Question 54. The three different types of elementary row operations that can be performed on a matrix are:

- **Type 1:** Interchanging two rows.
- **Type 2:** Multiplying a row by some constant.
- **Type 3:** Adding a scalar multiple of a row to another row.

Consider the four matrices given below:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}, B = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 1 \\ -1 & -1 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 1 \\ 1 & -3 & 3 \end{bmatrix}.$$

Choose the set of correct options.

- **Option 1:** Matrix B is obtained from matrix A by an elementary row operation of Type 1.
- Option 2: Matrix C is obtained from matrix A by an elementary row operation of Type 1.
- **Option 3:** Matrix D is obtained from matrix C by an elementary row operation of Type 3.
- **Option 4:** Matrix A is obtained from matrix C by an elementary row operation of Type 2.

2.6.7 Row Reduction: (Reduced) Row echelon form

As we have seen above if system of linear equation $Ax = b$ has coefficient as in reduced row echelon form, then it makes easier to solve the system of linear equations.

Lets see some methods to change a matrix of any order $m \times n$ into row echelon form and the reduced row echelon form using elementary operations.

Steps to change a matrix in to row echelon form (till step-4, for row echelon form) and reduced row echelon form

Step-1: Find the left most non- zero column

For example: We have matrix $\begin{bmatrix} 0 & 3 & 2 \\ 2 & 0 & 2 \\ 2 & 3 & 0 \end{bmatrix}$, left most column is the first column

Step-2: Use the elementary row operation to get 1 in the top position of that column

For example: As we have above matrix, using Type-1 elementary operation we can interchange the first and second rows. And then using Type-2 elementary operation can make one in the top position of that column as follows:

$$\begin{bmatrix} 0 & 3 & 2 \\ 2 & 0 & 2 \\ 2 & 3 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 2 & 0 & 2 \\ 0 & 3 & 2 \\ 2 & 3 & 0 \end{bmatrix} \xrightarrow{R_1/2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 2 & 3 & 0 \end{bmatrix}$$

Step-3: Use Type-3 elementary operation to make the entries below the 1 into 0.

As we can see in the last matrix $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 2 & 3 & 0 \end{bmatrix}$, 1 is at left most and the top most position at the entry (1,1)-th (first row and firs column) and below of 1 in the second row is 0 itself i.e entry (2,1)-th, so we need to make 0 in the third row i.e entry (3,1)-th need to make it 0.

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 2 & 3 & 0 \end{bmatrix} \xrightarrow{R_3 - 2R_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 0 & 3 & 0 \end{bmatrix}$$

Step-4: If there are no non-zero rows below current row, the matrix is in row echelon form. Else, find the next non-zero row and and using the type-1 elementary operation, move all zero rows below to the all remaining non- zero rows. And repeat the same process in the sub matrix below to the current row to change the matrix in row echelon form.

As we have seen in the last matrix $\begin{bmatrix}1 & 0 & 1\\ 0 & 3 & 2\\ 0 & 3 & 0\end{bmatrix}$, 1 is at left most and the top most position at the entry (1,1)-th (first row and firs column) and below of 1 in the second row is 0 itself and made 0 in the third row, entry (3,1)-th. Now we will make the entry (2, 2)-th to 1 and below of it entries in that column to 0 and will repeat it for the next row. We can see the process below,

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

The obtained matrix is in row echelon form.

Step-5: (Continue for reduced row echelon form) Take the columns containing 1 in the leading position of some row, use type-3 operation to make all entries of those column to 0.

As we get last matrix $\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 2/3\\ 0 & 0 & 1\end{bmatrix}$, we can see entry (1,1)-th is 1 which is in leading position, and this is the only non-zero in its column. Similarly for the entry (2,2)-th is 1 and this is the only non zero entry in its column so no need to performs any operations but the entry (3,3) which is leading position (and 1 also) is not the only non-zero entry in its column so we have make all entries 0 in that above to 1. We can see it below,

$$\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 2/3\\ 0 & 0 & 1\end{bmatrix} \xrightarrow{R_2 - \frac{2}{3}R_3} \begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \xrightarrow{R_1 - R_3} \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

The obtain matrix is in reduced row echelon form.

Example 2.6.8. Find the row echalon form and reduced row echelon form of the the matrix $\begin{bmatrix}2 & 4 & 1\\ 3 & 8 & 7\\ 5 & 6 & 9\end{bmatrix}$.

$$\begin{bmatrix}2 & 4 & 1\\ 3 & 8 & 7\\ 5 & 6 & 9\end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix}1 & 2 & 1/2\\ 3 & 8 & 7\\ 5 & 6 & 9\end{bmatrix} \xrightarrow{R_2 - 3R_1,\, R_3 - 5R_1} \begin{bmatrix}1 & 2 & 1/2\\ 0 & 2 & 11/2\\ 0 & -4 & 13/2\end{bmatrix} \xrightarrow{R_2/2}$$

$$\begin{bmatrix} 1 & 2 & 1/2 \\ 0 & 1 & 11/4 \\ 0 & -4 & 13/2 \end{bmatrix} \xrightarrow{R_3 + 4R_2} \begin{bmatrix} 1 & 2 & 1/2 \\ 0 & 1 & 11/4 \\ 0 & 0 & 35/2 \end{bmatrix} \xrightarrow{\frac{2}{35}R_3} \begin{bmatrix} 1 & 2 & 1/2 \\ 0 & 1 & 11/4 \\ 0 & 0 & 1 \end{bmatrix} \text{(Row echelon form)}$$

$$\xrightarrow{R_1 - R_3/2,\, R_2 - \frac{11}{4}R_3} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{(Reduced row echelon form)}$$

Question 55. Let $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}$ be a square matrix of order 3. Which of the statements below are true for matrix A?

- **Option 1:** A can be transformed via elementary row operations into the matrix $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ which is in row echelon form.
- Option 2: The reduced row echelon form of matrix A is $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
- **Option 3:** The reduced row echelon form of matrix A is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
- **Option 4:** A can be transformed via elementary row operations into the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ which is in row echelon form.
- **Option 5:** A can be transformed via elementary row operations into the matrix $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ which is in row echelon form.

Question 56. Let the reduced row echelon form of a matrix A be

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Suppose the first and third columns of A are $\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ respectively. If the second column of the matrix A is given by $\begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$, then what is the value of $m_1 + m_2$.

[Answer: 1]

Question 57. Let $A = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 1 & -1 & -1 \end{bmatrix}$ be a matrix (where the first row denotes the top most row, and the ordering of the rows is from top to bottom). Consider the system of linear equations given by $Ax = b$ where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $b = \begin{bmatrix} 4 \\ 3 \\ -1 \end{bmatrix}$. Answer the following questions. **Q1:** If R be the reduced row echelon form of A, then find the number of non-zero rows of R. [Answer: 2]

Q2: If $x_1 = 0$ is given, then find out the value of x_2. [Answer: -0.5]

2.6.8 Effect of elementary row operations on the determinant of a matrix

- **Type 1:** Interchanging two rows.

  Let B be a matrix obtain from A by interchanging any two row then $det(B) = -det(A)$.

  For example, interchanging R_1 and R_2

  Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix} = B$

  Then $det(B) = -det(A)$

- **Type 2:** Multiplying a row with some constant.

  Let B be a matrix obtain from A by multiplying a row of A by $c \in \mathbb{R}$ then $det(B) = cdet(A)$.

  For example, Multiply R_2 with $c \in \mathbb{R}$

  Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{c.R_2} \begin{bmatrix} 1 & 2 & 3 \\ 4c & 5c & 6c \\ 7 & 8 & 9 \end{bmatrix} = B$

  Then $det(B) = cdet(A)$

- **Type 3:** Adding a scalar multiple of a row to another row.

  Let B be a matrix obtain from A by multiplying a row with an real number and then added to another row then $det(B) = det(A)$.

  For example, Multiply R_2 with $c \in \mathbb{R}$ and then add to R_3

  Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_3 + c.R_2} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7+4c & 8+5c & 9+6c \end{bmatrix} = B$

  Then $det(B) = det(A)$.

Question 58. Let A and B be square matrices of order 3. Consider the three equations below.

- **Equation 1:** $det(A) = -det(B)$
- **Equation 2:** $det(A) = -c\ det(B), c \in \mathbb{R}$
- **Equation 3:** $det(A) = det(B)$

Choose the set of correct options.

- **Option 1:** If matrix B is obtained from matrix A by an elementary row operation of type 1, then equation 1 is satisfied.
- **Option 2:** If matrix B is obtained from matrix A by an elementary row operation of type 1 followed by an elementary operation of type 2, then equation 2 is satisfied for some c.
- Option 3: If matrix B is obtained from A by an elementary row operation of type 2, then equation 3 is satisfied.
- **Option 4:** If matrix B is obtained from A by an elementary row operation of type 3, then equation 3 is satisfied.

Question 59. Let A be a 3×3 real matrix whose sum of entries of each column is 5 and sum of first two elements of each column is 3. Which of the following statements is (are) true? [Hint: Row operation: adding one row to other row.]

- **Option 1:** The determinant of matrix A is a multiple of 5.
- **Option 2:** The determinant of matrix A is a multiple of 3.

- **Option 3:** The determinant of matrix A is a multiple of 15.
- **Option 4:** The determinant of matrix A is a multiple of 2.
- **Option 5:** The determinant of matrix A is a multiple of 8.

2.6.9 Gauss elimination algorithm

As we have just learned to change a matrix in row echelon form and reduced row echelon form. Suppose the coefficient matrix of a given system of linear equations is in (reduced) row echelon form, then we can easily find the solution of the system of linear equations. Lets see it via an example.

Example 2.6.9. Consider a system of linear equation $Ax = b$, where $A = \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 6 & 7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 0 \\ 1 \\ -1 \end{bmatrix}$ Find the solution of the above system of linear equations.

Observe that the coefficient matrix is in row echelon form.

Lets write the equation form of the above matrix form of the system of linear equations, we get,

$$\begin{aligned} x_1 + 5x_4 &= 2 \\ x_2 + 6x_3 + 7x_4 &= 0 \\ x_3 &= 1 \\ x_4 &= -1 \end{aligned}$$

Observe that we have already got values of x_3 and x_4.
Substituting $x_3 = 1$ and $x_4 = -1$ in rest two equations we get, $x_2 = 1$ and $x_1 = 7$

2.6.10 Augmented matrix

Let $Ax = b$ be a system of linear equations where A is $m \times n$ matrix and b is $m \times 1$ matrix.

The augmented matrix of this system od linear equations is defined as the matrix of size $m \times (n+1)$ whose first n columns are the column of A and last column is b.

We denote the augmented matrix as $[A|b]$ put a vertical line between the first n columns and the last column b while writing it.

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

matrix representation of the above system of linear equation is $Ax = b$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mathrm{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

And so the augmented matrix of the above system of linear equations is

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right]$$

Example 2.6.10. Find the augmented matrix of the system of linear equations,

$$\begin{aligned} x_1 + 2x_2 + 5x_3 &= 2 \\ 3x_2 + x + 3 &= -1 \end{aligned}$$

The coefficient matrix of the above system linear equations $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

The augmented matrix of the above system of linear equations is

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & 3 & 1 & 1 \end{array}\right]$$

Steps to find the solution of a system of linear equations using Gauss elimination method

Consider a system of linear equations $Ax = b$.

Step-1: From the augmented matrix of the system $[A|b]$.

Step-2: Perform the same elementary row operations on the matrix $[A|b]$ that were used to bring the matrix A into into reduced row echelon form.

Step-4: After obtaining the reduced row echelon form the matrix $[A|b]$, let R be the submatrix of the obtained matrix of the first n columns and c be the submatrix of the obtained matrix consisting of the last column. We can write reduced row echelon form of $[A|b]$ to $[R|c]$.

Step-5: Form the corresponding the system of linear equations $Rx = c$.

Step-6: Find all the solutions of $Rx = c$, those solution will be the solution of $Ax = b$.

Example 2.6.11. Consider a system of linear equations:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 + x_4 &= 6 \\ x_1 + x_2 &= 2 \\ 7x_2 + x_3 + x_4 &= 8 \end{aligned}$$

The augmented matrix of the above system is

$$\left[\begin{array}{cccc|c} 3 & 2 & 1 & 1 & 6 \\ 1 & 1 & 0 & 0 & 2 \\ 0 & 7 & 1 & 1 & 8 \end{array}\right] \xrightarrow{R_3} \left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 1 & 1 & 0 & 0 & 2 \\ 0 & 7 & 1 & 1 & 8 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 0 & 1/3 & -1/3 & -1/3 & 0 \\ 0 & 7 & 1 & 1 & 8 \end{array}\right]$$

$$\xrightarrow{3R_3} \left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 7 & 1 & 1 & 8 \end{array}\right] \xrightarrow{R_3 - 7R_2} \left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 8 & 8 & 8 \end{array}\right] \xrightarrow{R_3/8} \left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 + R_3,\, R_1 - R_3/3} \left[\begin{array}{cccc|c} 1 & 2/3 & 0 & 0 & 5/3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] \xrightarrow{R_1 - 2R_2/3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

Matrix $\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$ is in reduced row echelon form of augmented matrix, from this matrix we can get form $Rx = c$, where $R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ (R is reduced row echelon form of the coefficient matrix) and $c = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Observe that, in R, entries (1,1)-th, (2,2)-th, (3,3)-th are the leading entries so the variables x_1, x_2 and x_3 are dependent variables and x_4 is the independent variable.

Equation form of $Rx = c$ is,

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 1 \\ x_3 + x_4 &= 1 \implies x_3 = 1 - x_4 \end{aligned}$$

So this system has infinitely many solutions.
Let $x_4 = c, c \in \mathbb{R}$, we get $x_3 = 1 - c$.
Hence $S = \{(1, 1, 1 - c, c) \mid c \in \mathbb{R}\}$ is the solution space of the system of linear equations.

Alternative method:

To find the solution of system of linear equations, we found the $[R|c]$ (reduced row echelon form) form of augmented matrix and then change $Rx = c$ in to equation form and decided the solution. But if we get only row echelon form of augmented matrix and then write the equation form, still we can decide the solution of the system of linear equations. lets see this:

Observe that $\left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$ is the row echelon form of augmented matrix. Lets find the equation form fro the row echelon form of the augmented matrix. We know that last column of $\left[\begin{array}{cccc|c} 1 & 2/3 & 1/3 & 1/3 & 2 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$ is right side of the all equations of the system. So, we can write

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

So,

$$\begin{aligned} x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3 + \frac{1}{3}x_4 &= 2 \dots (1) \\ x_2 - x_3 - x_4 &= 0 \dots (2) \\ x_3 + x_4 &= 1 \dots (3) \end{aligned}$$

After adding equations (2) and (3), we get,

$$x_2 = 1$$

.

Multiplying both sides of second equation by 1/3 and adding to (1), we get,

$$x_1 + x_2 = 1 \implies x_1 = 1 - x_2 = 1$$

(as we have $x_2 = 1$).

And form the equation (3), we get,

$$x_3 + x_4 = 1 \implies x_3 = 1 - x_4$$

Let $x_4 = c$, $c \in \mathbb{R}$, we get

$$x_3 = 1 - c$$

.

Hence $S = \{(1, 1, 1 - c, c) \mid c \in \mathbb{R}\}$ is the solution space of the system of linear equations.

Example 2.6.12. Consider a system of linear equations:

$$\begin{aligned} x_1 + x_2 + x_3 &= 2 \\ x_2 - 3x_3 &= 1 \\ 2x_1 + x_2 + 5x_3 &= 0 \end{aligned}$$

The augmented matrix of the above system is

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & -3 & 1 \\ 2 & 1 & 5 & 0 \end{array}\right] \xrightarrow{R_3 - 2R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & -1 & 3 & -4 \end{array}\right] \xrightarrow{R_3 + R_2}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & -3 \end{array}\right] \xrightarrow{R_1 - R_2,\, R_3/(-3)} \left[\begin{array}{ccc|c} 1 & 0 & 4 & 1 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Matrix $\left[\begin{array}{ccc|c} 1 & 0 & 4 & 1 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$ is in reduced row echelon form of augmented matrix, from this matrix we can get form $Rx = c$, where $R = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$ (R is reduced row echelon form of the coefficient matrix) and $c = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Equation form of $Rx = c$ is,

$$\begin{aligned} x_1 + 4x_3 &= 1 \\ x_2 - 3x_3 &= 1 \\ 0x_1 + 0x_2 + 0x_3 &= 1 \implies 0 = 1 \text{ (Which is absurd)} \end{aligned}$$

So this system has no solution.

Example 2.6.13. Consider a system of linear equations:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 6 \\ x_1 + x_2 &= 2 \\ 7x_2 + x_3 &= 8 \end{aligned}$$

The augmented matrix of the above system is

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 6 \\ 1 & 1 & 0 & 2 \\ 0 & 7 & 1 & 8 \end{array}\right] \xrightarrow{R_3} \left[\begin{array}{ccc|c} 1 & 2/3 & 1/3 & 2 \\ 1 & 1 & 0 & 2 \\ 0 & 7 & 1 & 8 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{ccc|c} 1 & 2/3 & 1/3 & 2 \\ 0 & 1/3 & -1/3 & 0 \\ 0 & 7 & 1 & 8 \end{array}\right]$$

$$\xrightarrow{3R_3} \left[\begin{array}{ccc|c} 1 & 2/3 & 1/3 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 7 & 1 & 8 \end{array}\right] \xrightarrow{R_3 - 7R_2} \left[\begin{array}{ccc|c} 1 & 2/3 & 1/3 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 8 & 8 \end{array}\right] \xrightarrow{R_3/8} \left[\begin{array}{ccc|c} 1 & 2/3 & 1/3 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 + R_3,\, R_1 - R_3/3} \left[\begin{array}{ccc|c} 1 & 2/3 & 0 & 5/3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_1 - 2R_2/3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Matrix $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$ is in reduced row echelon form of augmented matrix, from this matrix we can get form $Rx = c$, where $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (R is reduced row echelon form of the coefficient matrix) and $c = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Observe that, in R, entries (1,1)-th, (2,2)-th, (3,3)-th are the leading entries so the variables x_1, x_2 and x_3 all are dependent.

Equation form of $Rx = c$ is,

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 1 \\ x_3 &= 1 \end{aligned}$$

So this system has a unique solution.

Hence $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is the solution of the system.

2.6.11 Exercise

Question 60. Let $Ax = b$ be a matrix representation of a system of linear equations. Let $[R|c]$ be the reduced row echelon of the augmented matrix $[A|b]$ corresponding to the system. Choose the set of correct options.

[Hint: Recall reduced row echelon form.]

- **Option 1:** If the system $Rx = c$ has infinitely many solutions, then the system $Ax = b$ has infinite solutions.
- Option 2: If the system $Rx = c$ has no solutions, then the system $Ax = b$ has a unique solution.
- Option 3: If the system $Rx = c$ has a unique solution, then the system $Ax = b$ has no solution.
- **Option 4:** If the system $Rx = c$ has a unique solution, then the system $Ax = b$ has a unique solution.

Question 61. Consider a system of linear equations

$$\begin{aligned} 2x_1 + x_2 &= 1 \\ -x_1 + x_3 + x_4 &= -1 \\ x_1 + x_2 - x_3 + x_4 &= 2 \\ -x_1 + x_3 + x_4 &= 1. \end{aligned}$$

If the following matrix represents the augmented matrix of the system, then answer the Q1, Q2 and Q3:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & b_3 \\ a_{41} & a_{42} & a_{43} & a_{44} & b_4 \end{array}\right]$$

Q1: Find the value of a_{22}. [Answer: 0]

Q2: Find the sum of the elements of the 4-th row of the augmented matrix. [Answer: 2]

Q3: Find the sum of the elements of the 3-rd column of the augmented matrix. [Answer: 1]

Question 62. Let $Ax = b$ be a matrix representation of a system of linear equations and $b = 0$. Choose the set of correct options.

[Hint: Recall the definition of the trivial solution of a system of linear equations and applicability of Gauss Elimination Method.]

- Option 1: If A is an invertible matrix then the system has no solution.
- **Option 2:** If A is an invertible matrix then the system has a unique solution.
- **Option 3:** If A is an invertible matrix then the trivial solution is the only solution for the system.
- **Option 4 :** If $det(A) = 0$, then the system has infinitely many solutions.

Question 63. Let $Ax = b$ be a matrix representation of a system of linear equations and $b \neq 0$. Choose the set of correct options.

- Option 1: If A is an invertible matrix, then the system has no solution.
- **Option 2:** If A is an invertible matrix, then the system has a unique solution.

- Option 3: If A is an invertible matrix, then the trivial solution is a solution for the system.
- **Option 4 :** If $det(A) = 0$, then either the system has no solution or the system has infinitely many solutions.

Question 64. Consider the following systems of linear equations:

$$\text{System I:} \qquad \begin{aligned} x - y &= 3 \\ -y + 2z &= 1 \\ x + y + z &= 0 \end{aligned}$$

$$\text{System II:} \qquad \begin{aligned} x - y &= 3 \\ 2y + z &= 1 \\ 6y + 3z &= 0 \end{aligned}$$

$$\text{System III:} \qquad \begin{aligned} x - y &= 3 \\ 2y + z &= 1 \\ 6y + 3z &= 3 \end{aligned}$$

Choose the correct option.

- Option 1: All the three systems have a unique solution.
- **Option 2:** System I has a unique solution.
- **Option 3:** System II has no solution.
- **Option 4:** System III has infinitely many solutions.
- Option 5: Both the System II and III have infinitely many solutions.
- Option 6: Both the system II and III have no solution.

Question 65. Suppose P is a 3×3 real matrix as follows:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

The vector X_n is defined by the recurrence relation $PX_{n-1} = X_n$. If $X_2 = \begin{bmatrix}7\\8\\3\end{bmatrix}$, what is the sum of all the elements of X_0? [Answer: 6]

Question 66. Consider the system of linear equations given below:

$$\begin{aligned} x - y + z &= 2 \\ x + y - z &= 3 \\ -x + y + z &= 4. \end{aligned}$$

The system of linear equations has

[Hint: Use the relation between a system of linear equations and determinant of corresponding coefficient matrix.]

- Option 1: no solution.
- Option 2: infinitely many solutions.
- **Option 3:** a unique solution.
- Option 4: finitely many solutions.
- **Option 5:** 3 dependent variables.
- Option 6: 2 dependent and 1 independent variables.

Question 67. Consider the system of linear equations:

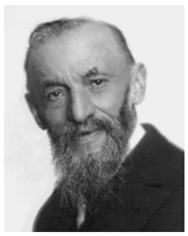
$$\begin{aligned} -x_1 + x_2 + 2x_3 &= 1 \\ 2x_1 + x_2 - 2x_3 &= -1 \\ 3x_2 + cx_3 &= d \end{aligned}$$

Choose the set of correct options.

[Hint: Recall the Gauss elimination method.]

- **Option 1:** If $c = 2$ and $d = 1$, then the system has infinitely many solutions.
- Option 2: If $c = 1$ and $d = 1$, then the system has infinitely many solutions.
- Option 3: If $c = 1$ and $d = 2$, then the system has no solution.
- **Option 4:** If $c = 3$ and $d = 2$, then the system has a unique solution.

3. Introduction to vector space



Certainly it is permitted to anyone to put forward whatever hypotheses he wishes, and to develop the logical consequences contained in those hypotheses. But in order that this work merit the name of Geometry, it is necessary that these hypotheses or postulates express the result of the more simple and elementary observations of physical figures.

— Giuseppe Peano

3.1 Introduction

In the last two chapters, we have studied a few methods to solve a system of linear equations. Vector spaces provide, among other things, a birds eye view of solutions to a system of linear equations. The idea of vector space was formalized by Peano in 1888. To motivate the abstract definition of a vector space, we will discuss its critical properties with an example. Consider a homogeneous system of linear equations

$$x + y - 2z = 0$$

$$x - 2y + z = 0$$

$$-2x + y + z = 0$$

The matrix representation of the above system of linear equation is

$$\begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.1}$$

Solving the above system is same as finding x, y, and z that solve the equation:

$$x \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + y \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + z \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Let A denote the matrix associated with the above system. Consider the set $S := \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^{3\times 1} \;\middle|\; A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$, which consists of all the solutions of the system (3.1). Observe that

$$v_1 := \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \text{ and } v_2 := \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix} \in S \implies v_1 + v_2 := \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \\ c_1 + c_2 \end{bmatrix} \in S,$$

and

$$v_1 := \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \in S \implies \alpha v_1 := \begin{bmatrix} \alpha a_1 \\ \alpha b_1 \\ \alpha c_1 \end{bmatrix} \in S \quad \text{for all } \alpha \in \mathbb{R}.$$

Equivalently, if v_1 and v_2 are in S then $\alpha_1 v_1 + \alpha_2 v_2$ is also in S, where α_1 and α_2 are real numbers. We call $\alpha_1 v_1 + \alpha_2 v_2$ a **linear combination** of v_1 and v_2. In more general, if $v_1, v_2, \ldots, v_n \in S$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$, then $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ (linear combination of $v_1, v_2, \ldots, v_n$) is also in S.

Roughly, we can think of a vector space V as a collection of objects that behave similar to the vectors in the set S. We can perform two operations on V;

- We can add the elements of V.
- We can multiply scalars ($\alpha \in \mathbb{R}$) with the elements of V.

These operations should satisfy a few conditions, the axioms for a vector space.

3.2 Vector space

Definition 3.2.1. A **vector space** $(V; +; \cdot; \mathbb{R})$ is a nonempty set V with two operations:

$$+(\textbf{addition}) : V \times V \to V \quad \text{and} \qquad \cdot\,(\textbf{scalar multiplication}) : \mathbb{R} \times V \to V.$$

which satisfy the following conditions for all u, v and w in V and for all scalars a and b in $\mathbb{R}$.

1. **Additive Closure:** $u + v \in V$.
2. **Multiplicative Closure:** $a \cdot u \in V$.
3. **Additive Commutativity:** $u + v = v + u$.
4. **Additive Associativity:** $(u + v) + w = u + (v + w)$.
5. **Existence of a zero vector:** There is a special vector $0_V \in V$, called the zero vector, such that $u + 0_V = u$.
6. **Additive Inverse:** For every $u \in V$, there is a vector u' , called the negative(additive inverse) of u, such that $u + u' = 0_V$.
7. **Unity:** $1 \cdot u = u$.
8. **Associativity of multiplication:** $(ab) \cdot u = a \cdot (b \cdot u)$.
9. **Distributivity:** $(a + b) \cdot u = a \cdot u + b \cdot u$ and $a \cdot (u + v) = a \cdot u + a \cdot v$.

Note:

- If just one of the vector space axiom fails to hold, then V not a vector space.
- Zero element 0_V of a vector space V is always unique.
- The real number $0 \in \mathbb{R}$ and the zero vector 0_V of a vector space V are commonly denoted by the symbol 0. One can always tell from the context whether 0 means the zero scalar ($0 \in \mathbb{R}$) or the zero vector ($0_V \in V$).
- It is standard to suppress $\cdot$ and only write au instead of $a \cdot u$.

$$\boxed{\text{A } \textbf{vector} \text{ is an element of a vector space.}}$$

Example 3.2.1. Let $V = \mathbb{R}^n$ be the set of all n-tuples $(x_1, x_2, \ldots, x_n)$ of real numbers. That is,

$$V = \left\{ (x_1, x_2, \ldots, x_n) \,\middle|\, x_1, x_2, \ldots, x_n \in \mathbb{R} \right\}.$$

Consider the **usual/standard vector addition and scalar multiplication** on V, defined as follows:

$$\left(x_1, x_2, \ldots, x_n\right) + \left(y_1, y_2, \ldots, y_n\right) = \left(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n\right)$$
$$c \cdot \left(x_1, x_2, \ldots, x_n\right) = \left(cx_1, cx_2, \ldots, cx_n\right)$$

Note: The addition and multiplication on the left hand side are defined in $\mathbb{R}^n$, whereas the addition and multiplication inside the bracket on the right-hand side occur in $\mathbb{R}$.

Let's verify all the conditions for being a vector space to know whether V is a vector space or not with respect to the given addition and scalar multiplication.

1. Additive Closure:

Clearly, If $u = \left(x_1, x_2, \ldots, x_n\right) \in V$ and $v = \left(y_1, y_2, \ldots, y_n\right) \in V$ then $u + v = \left(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n\right)$ is also in V. Hence V is closed under addition.

2. Multiplicative Closure:

Similarly, if $c \in \mathbb{R}$ and $u = \left(x_1, x_2, \ldots, x_n\right) \in V$ then $c \cdot u = \left(cx_1, cx_2, \ldots, cx_n\right)$ is also in V. Hence V is closed under scalar multiplication.

3. Additive Commutativity:

Consider two arbitrary vectors $u = \left(x_1, x_2, \ldots, x_n\right)$ and $v = \left(y_1, y_2, \ldots, y_n\right)$ of V, then

$$u + v = \left(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n\right) = \left(y_1 + x_1, y_2 + x_2, \ldots, y_n + x_n\right) = v + u.$$

Therefore, $u + v = v + u$ for all $u, v \in V$.

4. Additive Associativity:

Consider three arbitrary vectors $u = \left(x_1, x_2, \ldots, x_n\right)$, $v = \left(y_1, y_2, \ldots, y_n\right)$ and $w = \left(z_1, z_2, \ldots, z_n\right)$, then

$$(u + v) + w = \left((x_1 + y_1) + w_1, (x_2 + y_2) + w_2, \ldots, (x_n + y_n) + w_n\right)$$
$$u + (v + w) = \left(x_1 + (y_1 + w_1), x_2 + (y_2 + w_2), \ldots, x_n + (y_n + w_n)\right).$$

Since each x_i, y_i and w_i are real numbers, and $(x_i + y_i) + w_i = x_i + (y_i + w_i)$ for all $i = 1, 2, \ldots, n$, we have $(u + v) + w = u + (v + w)$ for all $u, v \in V$.

5. Existence of zero vector:

To get the zero vector, recall that the zero vector of a vector space V is a vector 0_V such that $u + 0_V = u$ for all $u \in V$. Let $u = \left(x_1, x_2, \ldots, x_n\right)$ and $0_V =$

$(m_1, m_2, \ldots, m_n)$, then we must have

$$\begin{aligned}
&(x_1, x_2, \ldots, x_n) + (m_1, m_2, \ldots, m_n) = (x_1, x_2, \ldots, x_n) \\
\implies &(x_1 + m_1, x_2 + m_2, \ldots, x_n + m_n) = (x_1, x_2, \ldots, x_n)
\end{aligned}$$

As a result, we obtain $x_i + m_i = x_i \implies m_i = 0$ for all $i = 1 \sim n$. Hence $0_V = (0, 0, \ldots, 0)$ is the zero vector of V with the property that $u + 0_V = u$ for all $u \in V$.

6. Additive Inverse:

To find the additive inverse of an element u in V, we need to find an element $u' \in V$ such that $u + u' = 0_V$.
Let $u = (x_1, x_2, \ldots, x_n)$ and $u' = (a_1, a_2, \ldots, a_n)$ in V, hence we must have

$$\begin{aligned}
&(x_1, x_2, \ldots, x_n) + (a_1, a_2, \ldots, a_n) = 0_V = (0, 0, \ldots, 0) \\
\implies &(x_1 + a_1, x_2 + a_2, \ldots, x_n + a_n) = (0, 0, \ldots, 0).
\end{aligned}$$

This gives $x_i + a_i = 0 \implies a_i = -x_i$. Hence $(-x_1, -x_2, \ldots, -x_n)$ is the additive inverse u.

7. Unity:

Take $u = (x_1, x_2, \ldots, x_n) \in V$ and $1 \in \mathbb{R}$, then

$$1 \cdot u = (x_1, x_2, \ldots, x_n) = u \quad \text{for all } u \in V.$$

8. Associativity of multiplication:

Let $a, b \in \mathbb{R}$ and $u = (x_1, x_2, \ldots, x_n) \in V$. Then

$$\begin{aligned}
(ab) \cdot u &= ((ab)x_1, (ab)x_2, \ldots, (ab)x_n) \\
&= (a(bx_1), a(bx_2), \ldots, a(bx_n)) \\
&= a \cdot ((bx_1), (bx_2), \ldots, (bx_n)) \\
&= a \cdot (b \cdot (x_1, x_2, \ldots, x_n)) \\
&= a \cdot (b \cdot u)
\end{aligned}$$

Hence $(ab) \cdot u = a \cdot (b \cdot u)$ for all $a, b \in \mathbb{R}$ and $u \in V$.

9. Distributivity:

$u = (x_1, x_2, \ldots, x_n)$, $v = (y_1, y_2, \ldots, y_n) \in V$ and $a, b \in \mathbb{R}$. Then

$$\begin{aligned}
\bullet\,(a+b)\cdot u &= \big((a+b)x_1, (a+b)x_2, \ldots, (a+b)x_n\big)\\
&= (ax_1+bx_1, ax_2+bx_2, \ldots, ax_n+bx_n)\\
&= (ax_1, ax_2, \ldots, ax_n) + (bx_1, bx_2, \ldots, bx_n)\\
&= a\cdot\big(x_1, x_2, \ldots, x_n\big) + b\cdot\big(x_1, x_2, \ldots, x_n\big)\\
&= a\cdot u + b\cdot v.
\end{aligned}$$

$$\begin{aligned}
\bullet\,a\cdot(u+v) &= a\cdot\big(x_1+y_1, x_2+y_2, \ldots, x_n+y_n\big)\\
&= \big(ax_1+ay_1, ax_2+ay_2, \ldots, ax_n+ay_n\big)\\
&= (ax_1, ax_2, \ldots, ax_n) + (ay_1, ay_2, \ldots, ay_n)\\
&= a\cdot\big(x_1, x_2, \ldots, x_n\big) + a\cdot\big(y_1, y_2, \ldots, y_n\big)\\
&= a\cdot u + a\cdot v.
\end{aligned}$$

All the axioms of vector space are satisfied. Hence V is a vector space with respect to the given operations.

Example 3.2.2. Let $V = M_{2\times 3}(\mathbb{R})$ be the set of all 2×3 matrices of real numbers.

$$V = \left\{ \begin{pmatrix} a_{11}, a_{12}, a_{13} \\ a_{21}, a_{22}, a_{23} \end{pmatrix} \,\middle|\, a_{ij} \in \mathbb{R} \right\}.$$

The **usual/standard vector addition and scalar multiplication** on V are defined as follows:

$$\begin{pmatrix} a_{11}, a_{12}, a_{13} \\ a_{21}, a_{22}, a_{23} \end{pmatrix} + \begin{pmatrix} b_{11}, b_{12}, b_{13} \\ b_{21}, b_{22}, b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11}, a_{12}+b_{12}, a_{13}+b_{13} \\ a_{21}+b_{21}, a_{22}+b_{22}, a_{23}+b_{23} \end{pmatrix}$$

$$c\cdot \begin{pmatrix} a_{11}, a_{12}, a_{13} \\ a_{21}, a_{22}, a_{23} \end{pmatrix} = \begin{pmatrix} ca_{11}, ca_{12}, ca_{13} \\ ca_{21}, ca_{22}, ca_{23} \end{pmatrix}$$

Existence of zero vector:

One can verify that the vector $0_V = \begin{pmatrix} 0,0,0 \\ 0,0,0 \end{pmatrix}$ is the zero vector of the vector space V with respect to the given operations.

Additive Inverse:

For any vector $u = \begin{pmatrix} a_{11}, a_{12}, a_{13} \\ a_{21}, a_{22}, a_{23} \end{pmatrix} \in V$, the element $-u = \begin{pmatrix} -a_{11}, -a_{12}, -a_{13} \\ -a_{21}, -a_{22}, -a_{23} \end{pmatrix} \in V$ is the additive inverse of u.

Just like in the previous example, one can check the other axioms to show that V is a vector space with respect to the given operations.

Example 3.2.3. Consider a real matrix A of order $m \times n$. Let V be the set of solutions of the homogeneous system **Ax=0**. That is,

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\}.$$

Addition and scalar multiplication on V are defined as follows:

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix} \qquad c \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{pmatrix}$$

Additive Closure:

Here, we want to show that if $u, v \in V$, then $u + v$ is also in V. To show that $u + v \in V$, we need to prove that $A(u + v) = 0$.

Let $u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$, $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in V$. By the definition of V, we have $Au = 0$, $Av = 0$.

As a result

$$A(u + v) = Au + Av = 0 + 0 = 0.$$

Therefore, $u + v \in V$, and V is closed under addition.

Multiplicative Closure:

Here, we want to show that if $u \in V$ and $c \in \mathbb{R}$, then $c \cdot u$ is also in V. If $u \in V$ then $Au = 0$.

$$A(c \cdot u) = c \cdot (Au) = c \cdot 0 = 0$$

Therefore, $c \cdot u \in V$, and V is closed under scalar multiplication.

Checking other axioms of a vector space is straightforward. Hence V is a vector space with respect to the given operations.

Example 3.2.4. Let $V = \mathbb{R}^2 = \{(x_1, x_2) \,|\, x_1, x_2 \in \mathbb{R}\}$, and the addition and scalar multiplication on V are defined as follows:

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 - y_2)$$
$$c \cdot (x_1, x_2) = (cx_1, cx_2).$$

For example:

$$\begin{aligned}(1,1)+(2,3)=&(1+2,1-3)=(3,-2)\\(2,3)+(1,1)=&(2+1,3-1)=(3,2)\\(4,1)+(1,4)=&(4+1,1-4)=(5,-3)\end{aligned}$$

We want to check whether V is a vector space or not with respect to the given operations.

Additive Closure and Multiplicative Closure:

If (x_1, x_2), $(y_1, y_2) \in V$ then

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 - y_2) \in V$$

Hence V is closed under addition. Similarly one can check that V is closed under scalar multiplication.

Additive Commutativity:

Let (x_1, x_2), $(y_1, y_2) \in V$.

$$\begin{aligned}(x_1,x_2)+(y_1,y_2)=&(x_1+y_1,x_2-y_2)\\(y_1,y_2)+(x_1,x_2)=&(y_1+x_1,y_2-x_2)\end{aligned}$$

It is clear that $(x_1, x_2) + (y_1, y_2) = (y_1, y_2) + (x_1, x_2)$ does not hold for all (x_1, x_2), $(y_1, y_2) \in V$. Hence "Additive Commutativity" fails to hold. So V is **not a vector space** with respect to the given operations.

Note: In fact, one can check that many of the axioms that a vector space must satisfy do not hold in this set with respect to the given operations.

Example 3.2.5. Consider the set $V = \{(x, 1) \,|\, x \in \mathbb{R}\}$. The addition and scalar multiplication on V is defined as follows:

$$\begin{aligned}(x,1)+(y,1):=&(x+y,1)\\c\cdot(x,1):=&(cx,1)\end{aligned}$$

For example:

$$\begin{aligned}(1,1)+(2,1)=&(1+2,1)=(3,1)\\(2,1)+(-2,1)=&(2-2,1)=(0,1)\\4\cdot(1,1)=&(4,1)\end{aligned}$$

We want to check whether V is a vector space or not with respect to the given operations. For that, we need to verify all the conditions one by one.

1. Additive Closure:
If $(x, 1), (y, 1) \in V$, then $(x, 1) + (y, 1) := (x + y, 1)$ is also in V. Hence V is closed under addition.

2. Multiplicative Closure:
Similarly, if $c \in \mathbb{R}$ and $(x, 1) \in V$, then $c \cdot (x, 1) = (cx, 1)$ is also in V. Hence V is closed under scalar multiplication.

3. Additive Commutativity:
Let $u = \big(x, 1\big)$ and $v = \big(y, 1\big)$ be in V, then

$$u + v = \big(x + y, 1\big) = \big(y + x, 1\big) = v + u,$$

Therefore, $u + v = v + u$ for all $u, v \in V$.

4. Additive Associativity:
Let $u = \big(x, 1\big)$, $v = \big(y, 1\big)$ and $w = \big(z, 1\big)$ be in V, then

$$(u + v) + w = ((x + y) + z, 1) = (x + (y + z), 1) = u + (v + w)$$

3. Existence of a zero vector:
We are looking for an element $0_V \in V$ such that for all $u \in V$, we have $u + 0_v = u$.

Let $u = (x, 1)$ and $0_V = (m, 1)$ in V, hence we must have

$$\begin{aligned} &(x, 1) + (m, 1) = (x, 1) \\ \implies\ &(x + m, 1) = (x, 1) \quad \text{(using the addition defined on } V). \end{aligned}$$

This gives $x + m = x \implies m = 0$. Hence the zero vector is $0_V = (0, 1)$.

6. Additive Inverse:
To find the additive inverse of an element u in V, we need to find an element $u' \in V$ such that $u + u' = 0_V$.

Let $u = (x, 1)$ and $u' = (a, 1)$ in V, hence we must have

$$\begin{aligned} &(x, 1) + (a, 1) = 0_V = (0, 1) \\ \implies\ &(x + a, 1) = (0, 1) \quad \text{(using the addition defined on } V). \end{aligned}$$

This gives $x + a = 0 \implies a = -x$. Hence $(-x, 1)$ is the additive inverse $u = (x, 1)$.

7. Unity:

Take $u = \big(x, 1\big) \in V$ and $1 \in \mathbb{R}$, then

$$1 \cdot u = \big(x, 1\big) = u \quad \text{for all } \ u \in V.$$

8. Associativity of multiplication:

Let $a, b \in \mathbb{R}$ and $u = \big(x, 1\big) \in V$. Then

$$\begin{aligned}(ab) \cdot u &= \big((ab)x, 1\big) \\ &= \big(a(bx), 1\big) \\ &= a \cdot \big(bx, 1\big) \\ &= a \cdot (b \cdot (x, 1)) \\ &= a \cdot (b \cdot u)\end{aligned}$$

Hence $(ab) \cdot u = a \cdot (b \cdot u)$ for all $a, b \in \mathbb{R}$ and $u \in V$.

9. Distributivity:

$u = \big(x, 1\big)$, $v = \big(y, 1\big) \in V$ and $a, b \in \mathbb{R}$. Then

$$\begin{aligned}\bullet\, (a+b) \cdot u &= \big((a+b)x, 1\big) \\ &= (ax + bx, 1) \\ &= (ax, 1) + (bx, 1) \\ &= a \cdot \big(x, 1\big) + b \cdot \big(x, 1\big) \\ &= a \cdot u + b \cdot v.\end{aligned}$$

$$\begin{aligned}\bullet\, a \cdot (u+v) &= a \cdot \big(x+y, 1\big) \\ &= \big(ax + ay, 1\big) \\ &= (ax, 1) + (ay, 1) \\ &= a \cdot \big(x, 1\big) + a \cdot \big(y, 1\big) \\ &= a \cdot u + a \cdot v.\end{aligned}$$

All the axioms of vector space are true for the given set V. Hence V is a vector space with respect to the given operations.

Example 3.2.6. Consider the set $V = \mathbb{R}^2 = \{(x, y) \,|\, x, y \in \mathbb{R}\}$. The addition and scalar multiplication on V are defined as follows:

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2 + 1, y_1 + y_2) \\ c \cdot (x, y) &:= (cx + c - 1, cy)\end{aligned}$$

For example:

$$\begin{aligned}(1,1)+(2,1) =&(1+2+1,1+1)=(4,2)\\(2,1)+(-2,3) =&(2-2+1,3+1)=(1,4)\\4\cdot(1,1) =&(4+4-1,4)=(7,4)\end{aligned}$$

1. Existence of a zero vector:

Let $u=(x,y)$ and $0_V=(m,n)$ in V, hence we must have

$$\begin{aligned}(x,y)+(m,n)&=(x,y)\\\implies (x+m+1,y+n)&=(x,y).\end{aligned}$$

This gives $x+m+1=x \implies m=-1$ and $y+n=y \implies n=0$. Hence the zero vector is $0_V=(-1,0)$.

2. Additive Inverse:

To find the additive inverse of an element u in V, we need to find an element $u'\in V$ such that $u+u'=0_V$.
Let $u=(x,y)$ and $u'=(a,b)$ in V, hence we must have

$$\begin{aligned}(x,y)+(a,b)&=0_V=(-1,0)\\\implies (x+a+1,y+b)&=(-1,0).\end{aligned}$$

This gives $x+a+1=0 \implies a=-x-1$ and $b=-y$. Hence $(-x-1,-y)$ is the additive inverse $u=(x,y)$.

3. Unity:

Take $u=\big(x,y\big)\in V$ and $1\in\mathbb{R}$, then

$$1\cdot u=\big(x+1-1,y\big)=(x,y)=u \quad \text{for all } u\in V.$$

4. Associativity of multiplication:

Let $a,b\in\mathbb{R}$ and $u=\big(x,y\big)\in V$. Then

$$\begin{aligned}(ab)\cdot u&=\big(abx+ab-1,aby\big)\\&=a\cdot\big(bx+b-1,by\big)\\&=a\cdot(b\cdot(x,y))\\&=a\cdot(b\cdot u)\end{aligned}$$

Hence $(ab)\cdot u=a\cdot(b\cdot u)$ for all $a,b\in\mathbb{R}$ and $u\in V$.

5. Distributivity:

$u = \big(x_1, y_1\big)$, $v = \big(x_2, y_2\big) \in V$ and $a, b \in \mathbb{R}$. Then

$$\bullet\,(a+b)\cdot(x_1,y_1) = \big(ax_1 + bx_1 + a + b - 1, ay_1 + by_1\big)$$
$$= (ax_1 + a - 1, ay_1) + (bx_1 + b - 1, by_1)$$
$$= a\cdot\big(x_1,y_1\big) + b\cdot\big(x_1,y_1\big)$$
$$= a\cdot u + b\cdot v.$$

$$\bullet\, a\cdot(u+v) = a\cdot\big(x_1 + x_2 + 1, y_1 + y_2\big)$$
$$= \big(ax_1 + ax_2 + a + a - 1, ay_1 + ay_2\big)$$
$$= (ax_1 + a - 1, ay_1) + (ax_2 + a - 1, ay_2)$$
$$= a\cdot\big(x_1,y_1\big) + a\cdot\big(x_2,y_2\big)$$
$$= a\cdot u + a\cdot v.$$

Checking other axioms are easy. Hence V is a vector space with respect to the given operations.

Note: You can see that $(0,0)$ is not always a zero vector of $\mathbb{R}^2$. The zero vector is an element of the set you are working with which satisfies the corresponding axioms.

3.2.1 Exercise

Question 68. Is the set $V = \{(x,y)\,|\,x,y \in \mathbb{R}\}$ a vector space under these operations?

- $(x_1,y_1) + (x_2,y_2) = (x_1 + x_2, 0)$ and $r(x,y) = (rx, y)$.
- $(x_1,y_1) + (x_2,y_2) = (x_1 + y_1, x_2 + y_2)$ and $r(x,y) = (rx, 0)$.

Question 69. Consider the set

$$V = \left\{ \begin{pmatrix} 1, a \\ b, 1 \end{pmatrix} \middle|\, a, b \in \mathbb{R} \right\}.$$

The vector addition and scalar multiplication on V are defined as follows:

$$\begin{pmatrix} 1, a \\ b, 1 \end{pmatrix} + \begin{pmatrix} 1, c \\ d, 1 \end{pmatrix} = \begin{pmatrix} 1, a+c \\ b+d, 1 \end{pmatrix}$$
$$c\cdot\begin{pmatrix} 1, a \\ b, 1 \end{pmatrix} = \begin{pmatrix} 1, ca \\ cb, 1 \end{pmatrix}$$

Is the set V a vector space?

3.3 Properties of vector spaces

Theorem 3.3.1 (Cancellation law of vector addition). *Let V be a vector space and v_1, v_2 and v_3 are in V.*

$$\textit{If } v_1 + v_3 = v_2 + v_3, \textit{ then } v_1 = v_2.$$

Proof. Since V is a vector space, the element $v_3 \in V$ has an additive inverse v_3' such that $v_3 + v_3' = 0_V$. Therefore,

$$\begin{aligned} & v_1 + v_3 = v_2 + v_3 \quad \text{(By assumption)} \\ \implies & (v_1 + v_3) + v_3' = (v_2 + v_3) + v_3' \quad \text{(By adding } v_3' \text{ to both sides)} \\ \implies & v_1 + (v_3 + v_3') = v_2 + (v_3 + v_3') \quad \text{(By "additive associativity")} \\ \implies & v_1 + 0_V = v_2 + 0_V \quad (v_3' \text{ is the additive inverse of } v_3) \\ \implies & v_1 = v_2 \quad (0_V \text{ is the zero vector } V). \end{aligned}$$

□

Note: Similarly, one can show that in a vector space V if $v_3 + v_1 = v_3 + v_2$, then $v_1 = v_2$.

Corollary 3.3.2. *(i) Zero vector of a vector space V is unique, that is, if there are vectors 0_V and $0_V'$ satisfy the property of a zero vector then $0_V = 0_V'$.*

(ii) For each vector $u \in V$ the additive inverse u' of u is unique.

Proof. (i) Suppose 0_V and $0_V'$ are two additive identities of V. Since 0_V is an additive identity,

$$0_V' + 0_V = 0_V' \tag{3.2}$$

Since $0_V'$ is also an additive identity,

$$0_V' + 0_V' = 0_V' \tag{3.3}$$

Therefore, from equations (3.2) and (3.3);

$$\begin{aligned} & 0_V' + 0_V = 0_V' + 0_V' \\ \implies & 0_V' = 0_V \quad \text{(By cancellation law of vector addition)} \end{aligned}$$

(ii) Suppose u' and w are two additive inverse of u. Then

$$u + u' = 0_V = u + w.$$

By Cancellation law of vector addition, we obtain $u' = w$. Hence additive inverse of each vector $u \in V$ is unique.

Note: We denote the additive inverse of u as $-u$. □

Corollary 3.3.3. *(i) For all $u \in V$, we have $0 \cdot u = 0_V$, where 0 is the zero scalar and 0_V is the zero vector.*

(ii) For all $c \in \mathbb{R}$, we have $c \cdot 0_V = 0_V$.

(iii) For each $c \in \mathbb{R}$ and for each $u \in V$, we have $(-c) \cdot u = -(c \cdot u) = c \cdot (-u)$.

Proof. (i) Since $0 = 0 + 0$, we obtain

$$0 \cdot u = (0 + 0) \cdot u.$$

Since scalar multiplication is distributive over addition (Distributive property),

$$(0 + 0) \cdot u = 0 \cdot u + 0 \cdot u.$$

From the last two equalities, we get

$$\begin{aligned} & 0 \cdot u + 0 \cdot u = 0 \cdot u \\ \implies & 0 \cdot u + 0 \cdot u = 0 \cdot u + 0_V \quad (0_V \text{ is the zero vector } V) \\ \implies & 0 \cdot u = 0_V \quad (\text{Cancellation law of vector addition}) \end{aligned}$$

(ii) Proof of this is similar to the proof (i).

(iii) Since $c + (-c) = 0$, we obtain

$$0_V = 0 \cdot u = (c + (-c)) \cdot u = c \cdot u + (-c) \cdot u.$$

Hence, $c \cdot u$ is the additive inverse of $(-c) \cdot u$, that is, $(-c) \cdot u = -(c \cdot u)$. Similarly, one can show that $c \cdot (-u) = -(c \cdot u)$.

□

3.4 Subspaces of Vector spaces

Definition 3.4.1. A non-empty subset W of a vector space V is called a **subspace** of V if W is a vector space under the operations addition and scalar multiplication defined in V.

To show that a non empty set W is a vector subspace, one doesn't need to check all the vector space axioms. It is enough to check two axioms of a vector space.

Theorem 3.4.1. *If W is a non empty subset of a vector space V, then W is a subspace of V if and only if the following conditions hold*

(1) *If w_1 and w_2 are in W, then $w_1 + w_2 \in W$.*

(2) *For all $c \in \mathbb{R}$ and for all $w_1 \in W$, $c \cdot w_1 \in W$.*

A subspace W of a vector space V is called a **proper subspace** if $W \subsetneq V$.

Note: Every vector space V over $\mathbb{R}$ has two trivial subspaces:

- V itself is a subspace of V.
- The subset consisting of the zero vector $\{0_V\}$ of V is also a subspace of V.

Question 70. Show that V and $\{0_V\}$ are subspaces of a vector space V.

Example 3.4.1. Consider the vector space $V = \mathbb{R}^2 = \{(x, y) \,|\, x, y \in \mathbb{R}\}$, with respect to the usual addition and scalar multiplication:

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2)\\ c \cdot (x_1, y_1) &= (cx_1, cy_1).\end{aligned}$$

We want to check whether $W = \{(x, y) \,|\, x + y = 0\} \subset \mathbb{R}^2$ is a vector subspace of $V = \mathbb{R}^2$ or not.

Note that W is non empty and a proper subset of V because $(1, 1) \in V$, but $(1, 1) \notin W$. By Theorem 3.4.1, it is enough to show that W is closed under addition and scalar multiplication. Let $w_1 = (x_1, y_1)$, $w_2 = (x_2, y_2) \in W$. Then by the definition of W, we have $x_1+y_1 = 0$ and $x_2+y_2 = 0$. To show that $(x_1, y_1)+(x_2, y_2) = (x_1 + x_2, y_1 + y_2) \in W$, it is enough to show that $x_1 + x_2 + y_1 + y_2 = 0$, which is obvious.

Next we want to show that $c \cdot (x_1, y_1) = (cx_1, cy_1) \in W$, that is, $cx_1 + cy_1 = 0$. Since $x_1 + y_1 = 0$, we have $cx_1 + cy_1 = 0$ for all $c \in \mathbb{R}$. Hence W is a **subspace/proper subspace** of V.

Geometrically, the set W represents a straight line in $\mathbb{R}^2$. The geometrical representation of W is given below.

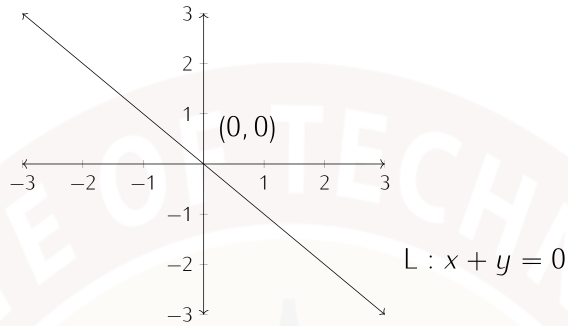


Figure 3.1:

From the above graph, it is clear that the line L passes through the origin. In general, a line in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$ if and only if it passes through the origin.

Question 71. A line in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$ if and only if it passes through the origin, where $\mathbb{R}^3$ is a vector space with respect to the usual addition and scalar multiplication.

Example 3.4.2. Consider the parabola $W = \{(x, y) \mid y = x^2\} \subset V = \mathbb{R}^2$. We want to check whether W is a vector subspace of V or not, where V is a vector space with respect to the usual addition and scalar multiplication. It is easy to see that W is not closed under addition and scalar multiplication. To prove this, It is enough to find one counterexample.

- $(1, 1), (-1, 1) \in W$, but $(1, 1) + (-1, 1) = (0, 2) \notin W$ (W is not closed under addition).
- $(-1, 1) \in W$, but $2 \cdot (-1, 1) = (-2, 2) \notin W$ (W is not closed under addition).

Hence W is not a subspace of V.

Example 3.4.3. Check whether the set $W = \{A \in M_{2\times2}(\mathbb{R}) \mid A = A^T\}$ of all 2×2 real symmetric matrices is a subspace of $M_{2\times2}(\mathbb{R})$ or not with standard addition and scalar multiplication.

$$A, B \in W \implies A^T = A, B^T = B \implies (A + B)^T = A^T + B^T = A + B.$$

Therefore $A+B$ is symmetric and W is closed under addition. Similarly, If $c \in \mathbb{R}$ and $A \in W$, then $(cA)^T = cA^T = cA$. Hence W is closed under scalar multiplication.

Example 3.4.4. Check whether the set W of 2×2 invertible matrices with real entries with standard addition and scalar multiplication is a subspace of $M_{2\times2}(\mathbb{R})$

or not.

Observe that W is not closed under addition:

$$\begin{pmatrix}1,0\\0,1\end{pmatrix}, \begin{pmatrix}-1,0\\0,-1\end{pmatrix} \in W \quad \text{but} \begin{pmatrix}1,0\\0,1\end{pmatrix} + \begin{pmatrix}-1,0\\0,-1\end{pmatrix} = \begin{pmatrix}0,0\\0,0\end{pmatrix} \notin W.$$

Example 3.4.5. Show that $W = \{(0, y, z) : y, z \in \mathbb{R}\}$ is a subspace of real vector space $\mathbb{R}^3$, where $\mathbb{R}^3$ is a vector space with respect to the usual addition and scalar multiplication.

Clearly, W is non empty and a proper subset of $\mathbb{R}^3$. If $(0, y_1, z_1)$ and $(0, y_2, z_2)$ in W then $(0, y_1 + y_2, z_1 + z_2)$ and $(0, cy_1, cz_1)$ are in W for all $c \in \mathbb{R}$. Hence W is a vector subspace.

3.4.1 Exercise

Question 72. Consider a system of m linear equations with n variables

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

A solution $(x_1, x_2, \ldots, x_n)$ of the above system is an element of $\mathbb{R}^n$. The set of all solutions W of the above system is a subspace of $\mathbb{R}^n$ iff $b = (b_1, b_2, \ldots, b_m) = (0, 0, \ldots, 0)$. Here we are considering $\mathbb{R}^n$ as a vector space with respect to the usual addition and scalar multiplication.

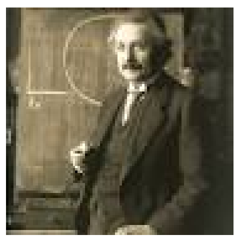
Hint: Check example 3.2.3.

Question 73. Intersection of two vector subspaces W_1 and W_2 of a vector subspace V is also a vector subspace.

Question 74. Show that the plane $W = \{(x, y, z) \,|\, x + y + z = 1\} \subset \mathbb{R}^3$ is not a vector subspace of $\mathbb{R}^3$, where $\mathbb{R}^3$ is a vector space with respect to the usual addition and scalar multiplication.

Question 75. Let $W = \{(x, y, z) \,|\, x \geq z\}$ be a subset of the vector space $\mathbb{R}^3$ (with respect to the usual addition and scalar multiplication). Then show that W is not a vector subspace of $\mathbb{R}^3$.

4. Basis and dimension



"Pure Mathematics is, in its way,
the poetry of logical ideas."

— Albert Einstein

4.1 Introduction

In the previous chapter, the concept of a vector space was formally introduced. As noticed, a non-zero vector space has infinitely many elements or vectors as we call. In this chapter, we will try to look at a smaller set with finitely many elements, called a basis, that describes the vector space completely. Before defining a basis, we study the concepts of linear dependence and span of a set.

4.2 Linear dependence and independence

We begin this section by defining what a linear combination is.

Definition 4.2.1. Let V be a vector space and $v_1, v_2, \ldots, v_n \in V$. Then $\sum_{i=1}^{n} \alpha_i v_i$ is said to be a linear combination of the vectors $v_1, v_2, \ldots, v_n$ with coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$.

Note that the linear combination of a set of vectors is another vector in V, since a vector space is closed under addition and scalar multiplication.

In $\mathbb{R}^2$, we can geometrically find the new vector which is a linear combination of other vectors. This is obtained by using the parallelogram law (the sum of two vectors is the diagonal of the parallelogram got by taking the vectors as two adjacent sides). An example is given in the following figure.

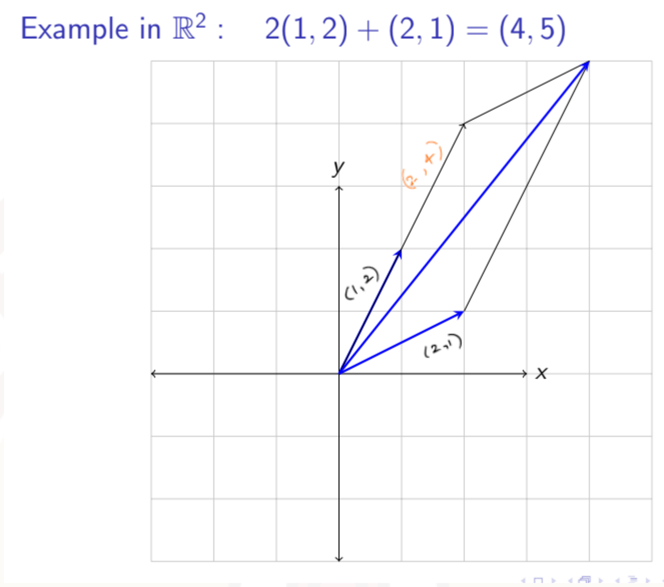


Figure 4.1:

Observe that $2(1,2)$ is obtained by scaling the vector $(1,2)$. Two adjacent sides of the parallelogram are $2(1,2)=(2,4)$ and $(2,1)$. The resulting vector is the diagonal of the parallelogram, $(4,5)$. The same may be obtained algebraically by adding the two vectors $(2,4)$ and $(2,1)$. We get the same result either way. In this case, $(4,5)$ is expressed as a linear combination of two vectors $v_1=(1,2)$ and $v_2=(2,1)$. The scalars used are $\alpha_1=2$ and $\alpha_2=1$, i.e., $2(1,2)+(2,1)=(4,5)$.

Also, we can see that any one of the above vectors can be written as a linear combination of the other two.

$$(4,5)=2(1,2)+(2,1)$$
$$(2,1)=2(1,2)-(4,5)$$
$$(1,2)=\frac{1}{2}(4,5)-\frac{1}{2}(2,1)$$

By taking all the three vectors to one side of the equation, we get $2(1,2)+(2,1)-(4,5)=(0,0)$. That is, the zero vector is written as a linear combination of the three vectors with non-zero coefficients. This is when we say that a set of vectors is linearly dependent. We shall formally define linear dependence of vectors after the following example which depicts linear combination of vectors in $\mathbb{R}^3$.

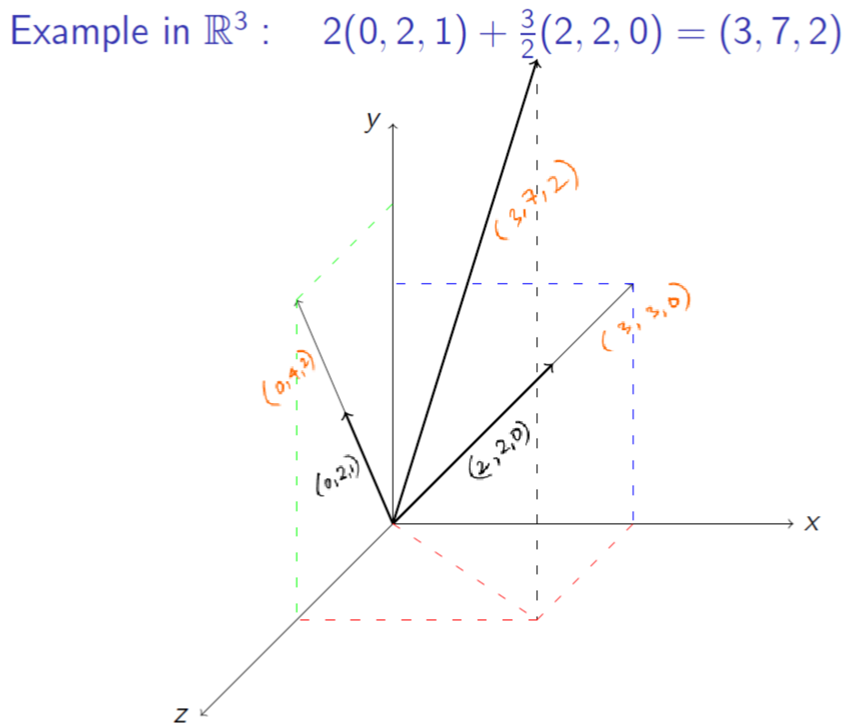


Figure 4.2:

In this example, we can see that $(3,7,2)$ is written as a linear combination of two vectors $(0,2,1)$ and $(2,2,0)$. Again, note that any one of the vectors can be written as a linear combination of the other two vectors.

$$(3,7,2)=2(0,2,1)+\frac{3}{2}(2,2,0)$$
$$(0,2,1)=\frac{1}{2}(3,7,2)-\frac{3}{4}(2,2,0)$$
$$(2,2,0)=\frac{2}{3}(3,7,2)-\frac{4}{3}(0,2,1)$$

In this case too, the zero vector can be written as a linear combination of the three vectors $(0,2,1)$, $(2,2,0)$ and $(3,7,2)$ with non-zero coefficients. We can easily verify that all the three vectors lie on the same plane, namely the plane $2x-2y+4z=0$.

Suppose we choose a vector that does not lie in this plane, for example $(1,1,1)$. In this case, we can show that this vector $(1,1,1)$ cannot be written as a linear combination of the above vectors. If suppose this is possible, that is, say $(1,1,1)$ is a linear combination of $(0,2,1)$ and $(2,2,0)$ then

$$(1,1,1)=a(0,2,1)+b(2,2,0)$$

for some $a, b \in \mathbb{R}$. Then we have

$$2b = 1, 2a + 2b = 1, a = 1.$$

But this is not possible. Thus $(1, 1, 1)$ cannot be written as a linear combination of these vectors.

From this, we can conclude that $\alpha(1, 1, 1) + \beta(0, 2, 1) + \gamma(2, 2, 0) = (0, 0, 0)$ is possible if and only if $\alpha = \beta = \gamma = 0$. Here, zero is a linear combination of these vectors if and only if all the coefficients are zero.

Now we are set to define linear dependence and independence of vectors.

4.2.1 Linear dependence

Definition 4.2.2. A set of vectors $\{v_1, v_2, ..., v_n\}$ is said to be linearly dependent, if there exist scalars $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}$ not all zero, such that $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$.

In other words, a set of vectors is linearly dependent if the zero vector can be written as a linear combination of those vectors with at least one of the coefficients being non-zero. Geometrically, any two vectors that lie on the same line, any three vectors that lie on the same plane, etc., are linearly dependent.

From our earlier discussion, $\{(4, 5), (2, 1), (1, 2)\}$, $\{(3, 7, 2), (0, 2, 1), (2, 2, 0)\}$ are linearly dependent sets. Notice that both the sets contain 3 vectors that lie on the same plane.

Example 4.2.1. Consider the set $\{(1, 2, 4), (2, -1, 2), (5, 0, 8)\}$. We shall now show that this set is linearly dependent. Suppose there exist $a, b, c \in \mathbb{R}$ such that

$$a(1, 2, 4) + b(2, -1, 2) + c(5, 0, 8) = (0, 0, 0).$$

Then we have the homogeneous system of linear equations

$$\begin{aligned} a + 2b + 5c &= 0 \\ 2a - b &= 0 \\ 4a + 2b + 8c &= 0 \end{aligned}$$

Solving, we get non-zero solutions to the system. For eg., $a = -1, b = -2, c = 1$ satisfies the system. Thus we have $-1(1, 2, 4) - 2(2, -1, 2) + 1(5, 0, 8) = (0, 0, 0)$, that is the zero vector can be written as a linear combination of the elements in the set with non-zero coefficients. Hence the set is linearly dependent.

Now, if we append another vector to this set say, $(1, 1, 1)$, the new set is $\{(1, 2, 4), (2, -1, 2), (5, 0, 8), (1, 1, 1)\}$. Clearly, this is linearly dependent, because $-1(1, 2, 4) - 2(2, -1, 2) + 1(5, 0, 8) + 0(1, 1, 1) = (0, 0, 0)$.

The above example brings clarity to the fact that every superset of a linearly dependent set is linearly dependent. It is easy to prove this statement and hence is omitted.

Remark 4.2.1. Suppose a set $\{v_1, v_2, ..., v_n\}$ contains the zero vector. Say $v_i = 0$. Then taking $\alpha_i = 1$ and all other coefficients as zero, we can write zero as a linear combination of these vectors with at least one coefficient being non-zero. Thus any set containing the zero vector is linearly dependent.

4.2.2 Linear independence

We shall begin by defining a linearly independent set.

Definition 4.2.3. A set of vectors $v_1, v_2, ..., v_n$ is said to be linearly independent if they are not linearly dependent. In other words, if there exist $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}$ such that $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$, then $\alpha_i = 0$ for all i.

If a set is linearly independent, then the only linear combination of these vectors that can yield the zero vector is when all the coefficients are zero.

Example 4.2.2. Consider the set of vectors $\{(1, 1), (1, -1)\}$. Suppose there exist $a, b \in \mathbb{R}$ such that $a(1, 1) + b(1, -1) = (0, 0)$. Then we have the following homogeneous system of linear equations.

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

Solving this, we get $a = b = 0$. That is, the only way zero can be written as a linear combination of $(1, 1)$ and $(1, -1)$ is when the coefficients are zero. Thus $\{(1, 1), (1, -1)\}$ is a linearly independent set.

Notice that a set of two vectors in linearly dependent if and only if one is a scalar multiple of the other. In the above example, $(1, 1)$ is not a scalar multiple of $(1, -1)$ and hence the set is not linearly dependent.

We had stated earlier that if two vectors lie on the same line, they are linearly dependent. Let us elaborate it here.

Remark 4.2.2. Suppose we have a set of two non-zero vectors v_1 and v_2. For this set to be linearly dependent, $\alpha_1 v_1 + \alpha_2 v_2 = 0$ with at least one of α_1 or α_2 being non-zero. But if exactly one of them is zero, say $\alpha_1 = 0$ and $\alpha_2 \neq 0$, then we have $\alpha_2 v_2 = 0$. Since $\alpha_2 \neq 0$, $v_2 = 0$, a contradiction. So, in this case, both α_1 and α_2 should be non-zero. Thus we have $v_1 = -\frac{\alpha_2}{\alpha_1} v_2$. v_1 and v_2 are scalar multiples of each other, that is they lie on the same line.

From the above remark, it is clear that two non-zero vectors are linearly independent if and only if they are not scalar multiples of each other.

Question 76. What can be said about the linear independence of 3 non-zero vectors?

We can argue along similar lines and show that three non-zero vectors are linearly independent if and only if none of them is a linear combination of the others.

We will end this section with an example of a set of linearly independent vectors.

Example 4.2.3. Consider the set $\{(1, 1, 2), (-1, 0, 1), (2, 1, 2)\}$. Suppose there exist $a, b, c \in \mathbb{R}$ such that $a(1, 1, 2) + b(-1, 0, 1) + c(2, 1, 1) = (0, 0, 0)$. We have the following homogeneous system of linear equations.

$$\begin{aligned} a - b + 2c &= 0 \\ a + c &= 0 \\ 2a + b + 2c &= 0 \end{aligned}$$

Solving this, we get a unique solution $a = b = c = 0$. Thus the given set is linearly independent.

4.2.3 Other ways to check linear independence

Let us recall what we do to check whether a set $\{v_1, v_2, ..., v_n\}$ is linearly independent or not.

We solve the equation $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$ for arbitrary real numbers $\alpha_1, \alpha_2, ..., \alpha_n$. Expanding this system, we get a homogeneous system of equations. Thus checking linear independence reduces to solving a system of linear equations, where the coefficients of the vectors are the unknowns. We know that a homogeneous system is always consistent, that is, a homogeneous system always has a solution. The zero vector is always a solution, but the question is, are there other non-zero solutions. If $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$ reduces to a system which has only the trivial solution, then the set is linearly independent. If we get a non-zero solution, then the set is linearly dependent. Note that we are not interested in the solution, we are only interested in finding whether the system has a unique solution (the trivial solution) or infinitely many solutions.

To conclude, to check whether the set $\{v_1, v_2, ..., v_n\}$ is linearly independent or not, we need to verify whether the homogeneous system $Vx = 0$ has only the trivial solution or not, where V is the matrix whose j^{th} column is the vector v_j. Note that if the matrix V is square, then the only thing we need to verify is whether the determinant of V is non-zero or not. (Why?)

Example 4.2.4. Consider the set $\{(1, 1), (2, 0)\}$. Clearly this set of vectors is linearly independent because they are not scalar multiples of each other. Now, we get a

2×2 matrix V since we have 2 vectors with 2 components each. We need to solve the equation $a(1,1) + b(2,0) = (0,0)$. This reduces to

$$\begin{aligned} a + 2b &= 0 \\ a &= 0 \end{aligned}$$

The matrix V is nothing but, $\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$. V is an invertible matrix and hence the vectors are linearly independent.

Example 4.2.5. Consider the set $\{(1,2,3),(2,-1,4)\}$. To form the coefficient matrix, we put the vectors as columns of V. We get a 3×2 matrix since we have 2 vectors and each vector has 3 components. The homogeneous system we get is

$$\begin{aligned} a + 2b &= 0 \\ 2a - b &= 0 \\ 3a + 4b &= 0 \end{aligned}$$

Using Gaussian elimination, we can verify whether that the system has a unique solution and hence the set is linearly independent.

Example 4.2.6. Consider the set $\{(1,1),(1,2),(3,4)\}$. Writing this as a homogeneous system of linear equations, we have

$$\begin{aligned} a + b + 3c &= 0 \\ a + 2b + 4c &= 0 \end{aligned}$$

Using Gaussian elimination, we can verify that the system has infinitely many solutions. This is obvious because we have more unknowns than the number of equations. Hence the set is linearly dependent.

Example 4.2.7. Consider the set $\{(1,1,1),(1,2,-1),(0,1,0)\}$. The homogeneous system that has to be solved is

$$\begin{aligned} a + b &= 0 \\ a + 2b + c &= 0 \\ a - b &= 0 \end{aligned}$$

The system has a unique solution (the trivial solution) if the determinant of the coefficient matrix is non-zero. It may be verified that the set is linearly independent.

From the above examples, let us make an observation. The case when we had 3 vectors in $\mathbb{R}^2$, we had infinitely many solutions.

Question 77. What can be said about the linear independence a set with 4 vectors in $\mathbb{R}^2$?

In general, a set with n vectors in $\mathbb{R}^2$, $n \geq 3$, is linearly dependent. The reasoning is exactly the same as what we had seen in Example 4.2.6.

More generally, a set with k vectors in $\mathbb{R}^n$, $k \geq n+1$, is always linearly dependent. In this case, once we do Gaussian elimination, we end up with some independent variables (recall independent and dependent variables).

4.2.4 Exercise

Question 78. Suppose the set $\{(2,3,0),(0,1,0),v\}$ is a linearly dependent set in the vector space $\mathbb{R}^3$ with usual addition and scalar multiplication. Which of the following vectors are possible as a candidate for v?

- Option 1: $(2,3,1)$
- **Option 2:** $(1,2,0)$
- **Option 3:** $(1,3,0)$
- Option 4: $(0,1,1)$
- **Option 5:** $(\pi,e,0)$

Question 79. If $(9,3,1)$ is a linear combination of the vectors $(1,1,1)$, $(1,-1,1)$ and (x,y,z) as $2(1,1,1)+3(1,-1,1)+4(x,y,z)=(9,3,1)$, then find the value of $x+y+2z$. [Ans : 0]

Question 80. Let V be a vector space and v_1, v_2, v_3 and $v_4 \in V$. If v_1 is a linear combination of $v_i, i=2,3,4$ *i.e.* $v_1 = av_2 + bv_3 + cv_4$ then

- Option 1: v_2 is a linear combination of $v_i, i=1,3,4$.
- Option 2: v_3 is a linear combination of $v_i, i=1,4$.
- **Option 3:** v_2 is a linear combination of v_2.
- **Option 4:** v_4 is a linear combination of v_4.

Question 81. Let S be a subset of $\mathbb{R}^3$ which is linearly dependent. Which of the following options are true?

- **Option 1:** $S \cup \{(1,0,0)\}$ must be linearly dependent.
- **Option 2:** $S \cup \{v\}$ must be linearly dependent for any $v \in \mathbb{R}^3$.

- Option 3: $S \setminus \{v\}$ must be linearly dependent for any $v \in \mathbb{R}^3$.
- **Option 4:** There may exist some $v \in S$, such that $S \setminus \{v\}$ is still linearly dependent.

Question 82. Choose the correct statement(s).

- Option 1: $0(0,0)+0(1,1) = (0,0)$ implies that the set $\{(0,0),(1,1)\}$ is a linearly independent subset of $\mathbb{R}^2$.
- **Option 2:** $2(1,0)+2(0,1) = (2,2)$ implies that the set $\{(1,0),(0,1),(2,2)\}$ is a linearly dependent subset of $\mathbb{R}^2$.
- Option 3: $2(1,0)+2(0,1) = (2,2)$ implies that the set $\{(1,0),(2,2)\}$ is a linearly dependent subset of $\mathbb{R}^2$.

Question 83. Choose the correct set of options.

- **Option 1:** Union of two distinct linearly independent sets may be linearly dependent.
- Option 2: Union of two distinct linearly independent sets must be linearly dependent.
- **Option 3:** Non-empty intersection of two linearly independent sets must be linearly independent.
- Option 4: Non-empty intersection of two linearly independent sets must be linearly dependent.

Question 84. If vectors v_1, v_2, v_3 are linearly independent and $\alpha v_1 + \beta v_2 + \gamma v_3 = 0$, then find the value of $\alpha + \beta + \gamma$. [Ans: 0]

Question 85. Let S be the solution set of a system of homogeneous linear equations with 3 variables and 3 equations, whose matrix representation is as follows:

$$Ax = 0$$

where A is the 3×3 coefficient matrix and x denotes the column vector $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Choose the set of correct options.

- **Option 1:** If v_1 and v_2 are in S, then any linear combination of v_1 and v_2 will also be in S.

- **Option 2:** The set S will be a subspace of $\mathbb{R}^3$, with respect to usual addition and scalar multiplication as in $\mathbb{R}^3$.
- **Option 3:** The set $\{v_1, v_2, v_1 - v_2\}$ is a linearly dependent subset in S.
- **Option 4:** The set $\{v_1, v_2\}$ is a linearly independent subset in S if v_1 is not a scalar multiple of v_2.

Question 86. For how many values of a is the set $\{(1, 2, a), (a, 0, a,), (-1, a^2, 8), (0, 1, -3a)\}$ linearly independent? [Answer: 0]

4.3 Spanning sets

Let us begin this section by defining the span of a set. The span of a set is roughly all the vectors that can be got by using the given set of vectors.

Definition 4.3.1. The span of a set S, denoted by span(S) is the set of all finite linear combinations of the elements of the set S. That is,

$$span(S) = \{\sum_{i=1}^{n} \alpha_i v_i : v_i \in S, \alpha_i \in \mathbb{R}\}$$

Note that span(S) is a vector subspace. (Verify!)

Example 4.3.1. Let $S = \{(1, 0)\} \subset \mathbb{R}^2$. The span of S will contain all possible finite linear combinations of the elements of S.

$$\begin{aligned} span(S) &= \{\alpha_i(1, 0) : \alpha_i \in \mathbb{R}\} \\ &= \{(\alpha, 0) : \alpha \in \mathbb{R}\} \end{aligned}$$

Thus the span of $\{(1, 0)\}$ is the x-axis. We already know that the x-axis is a subspace of $\mathbb{R}^2$.

Question 87. What is the span of the set $\{(1, 1)\}$?

Question 88. What is the span of $\{(x, y) : (x, y) \neq (0, 0)\}$?

Question 89. What is the span of $\{(a, b, c) : (a, b, c) \neq (0, 0, 0)\}$?

Example 4.3.2. Let $S = \{(1, 1, 0), (1, 0, 0)\}$. Then span of S will be the set

$$\begin{aligned} span(S) &= \{a(1, 1, 0) + b(1, 0, 0) : a, b \in \mathbb{R}\} \\ &= \{(a + b, a, 0) : a, b \in \mathbb{R}\} \\ &= \{(c, a, 0) : c, a \in \mathbb{R}\} \end{aligned}$$

This is nothing but the xy-plane in $\mathbb{R}^3$.

Now we are set to define a spanning set for a vector space. It is nothing but a set of vectors which generates all the elements of the vector space.

Definition 4.3.2. Let V be a vector space and $S \subset V$. S is said to be a spanning set if $\text{span}(S) = V$.

Example 4.3.3. Let $S = \{(1,0),(0,1)\}$. Then $\text{span}(S) = \mathbb{R}^2$. Thus S is a spanning set for $\mathbb{R}^2$. That is any element of $\mathbb{R}^2$ can be generated using elements of the set S. Any (x,y) can be written as $x(1,0) + y(0,1)$.

Example 4.3.4. Let $S = \{(1,0),(0,1),(1,1)\}$. Then S spans $\mathbb{R}^2$.

Clearly, every superset of a spanning set is also a spanning set.

Example 4.3.5. $S = \{(1,1),(1,0)\}$ spans $\mathbb{R}^2$ because $(x,y) = y(1,1) + (x-y)(1,0)$.

Question 90. If $T \subset S$, what can be said about the relationship between $\text{span}(T)$ and $\text{span}(S)$?

Question 91. What is $\text{span}(\text{span}(S))$?

4.3.1 Building spanning sets

We may append vectors to a set to build spanning sets for a vector space. Consider $\mathbb{R}^3$.

- Step 1: Start with $S_0 = \emptyset$. $\text{Span}(S_0) = \{(0,0,0)\}$. (Why!?)
- Step 2: Since $\text{span}(S_0) \neq \mathbb{R}^3$, append a vector, say $(1,1,0)$ to S_0. Now $S_1 = S_0 \cup \{(1,1,0)\} = \{(1,1,0)\}$. $\text{span}(S_1)$ is a line in $\mathbb{R}^3$, which still doesn't cover the entire space $\mathbb{R}^3$.
- Step 3: Now choose a vector outside $\text{span}(S_1)$. Let $S_2 = S_1 \cup \{(1,1,1)\} = \{(1,1,0),(1,1,1)\}$. $\text{Span}(S_2)$ is the plane $x = y$, which still doesn't cover $\mathbb{R}^3$.
- Step 4: We now choose a vector outside $\text{span}(S_2)$. Let $S_3 = S_2 \cup \{(1,0,0)\} = \{(1,1,0),(1,1,1),(1,0,0)\}$. Now, it is easy to verify that S_3 spans $\mathbb{R}^3$.

Notice that at each stage we added a vector which was not in the span of the previous vectors.

Question 92. Can you construct a spanning set for $\mathbb{R}^3$ starting with the set $\{(1,2,3)\}$?

4.3.2 Exercise

Question 93. Choose the set of correct options.

- **Option 1:** If S is a spanning set of the vector space V, then $S \cup \{v\}$ must be a spanning set of V, for all $v \in V$.
- **Option 2:** Span of an empty set is the zero vector space.
- Option 3: If S is a spanning set of the vector space V, then $S \setminus \{v\}$ must be a spanning set of V, for all $v \in S$.
- Option 4: If S is a spanning set of the vector space V, then $S \cup \{v\}$ may not be a spanning set of V, for all $v \in V$.

Question 94. Let S be the set of matrices where $S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$. Choose the correct statement(s):

- Option 1: $Span(S)$ is the vector space consisting of only lower triangular square matrices of order 2.
- **Option 2:** $Span(S)$ is the vector space consisting of only upper triangular square matrices of order 2.
- Option 3: $Span(S)$ is the vector space consisting of all the square matrices of order 2.
- Option 4: $Span(S)$ is the vector space consisting of only scalar matrices of order 2.

Question 95. Let S be the set of matrices where $S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$. Choose the correct statement(s):

- Option 1: $Span(S)$ is the vector space consisting of only the Identity matrix of order 2.
- **Option 2:** $Span(S)$ is the vector space consisting of only diagonal matrices of order 2.
- Option 3: $Span(S)$ is the vector space consisting of all the square matrices of order 2.
- Option 4: $Span(S)$ is the vector space consisting of only scalar matrices of order 2.

4.4 Basis of a vector space

With the idea of a spanning set and a linearly independent set, we can now define what a basis is. In layman's terms, a basis is the smallest set that gives complete information about the vector space.

Definition 4.4.1. A basis B of a vector space V is a linearly independent set that also spans V.

In the algorithm that we used to build a spanning set, note that since we wanted the newly added element to not be in the span of the previous elements, we were also building a linearly independent set, which will ultimately be a basis. Thus, we can construct a basis for every vector space.

So, the goal here is to keep only the necessary vectors so that we get all the information about the vector space. For example, in $\mathbb{R}^n$, any set with more than n vectors are linearly dependent. So a basis for $\mathbb{R}^n$ can have at most n elements.

Example 4.4.1. Let e_i denote the vector in $\mathbb{R}^n$ whose i^{th} coordinate is 1 and all other coordinates are 0. This forms a basis for $\mathbb{R}^n$. This is called the standard basis. (Verify that this set is actually a basis: Recall that you need to check whether the set is linearly independent and if it spans $\mathbb{R}^n$.)

Question 96. Show that $\{(1,1),(1,0)\}$ is a basis for $\mathbb{R}^2$.

It is clear from the previous example and question that there may be multiple bases for the same vector space. We will see a lot more examples a little later in this section.

Before proving the next theorem, let us define two terms.

Definition 4.4.2. A set S is said to be a maximal linearly independent set, if S is linearly independent and any superset of S is linearly dependent. In other words, appending any vector to S makes it linearly dependent.

Definition 4.4.3. A set S is said to be a minimal spanning subset of a vector space V, if S spans V and any subset of S does not span V. In other words, by removing any vector from S, the new set no longer spans V.

Theorem 4.4.1. *The following conditions are equivalent for a set B to be a basis of a vector space V:*

1. *B is linearly independent and $\text{span}(B) = V$.*

2. *B is a maximal linearly independent set.*

3. *B is a minimal spanning set.*

Why should these conditions be equivalent? The first condition is the definition of a basis. Let us first see why the first two statements are equivalent.

Suppose B is a basis, then B is linearly independent. Suppose $B' = B \cup \{v\}$. Since B is a basis, it spans V and hence v can be written as a linear combination of elements of B, hence making B' linearly dependent. Thus B is a maximal linearly independent set.
To prove the converse, if B is a maximal linearly independent set, then clearly it is linearly independent. It remains to show that B spans V. Suppose $v \in V$ is such that $v \notin B$. Then the set $B \cup \{v\}$ is a linearly independent set, a contradiction to B being maximal linearly independent set.

In a very similar fashion, the equivalence of the first and third statements may be established.

Now, it is clearly established that a basis is a minimal spanning set as well as a maximal linearly independent set. So we now have two ways to find a basis for a vector space.

1. Start with $\emptyset$ and keep appending vectors till the set spans the vector space V.

2. Start with a spanning set and keep deleting the vectors that are dependent on other vectors in the set till the set becomes linearly independent.

Example 4.4.2. Let $V = \mathbb{R}^2$. To get a basis for V, let us start with $\emptyset$ and append a non-zero vector to it, say $(1, 1)$. Now, $\{(1, 1)\}$ does not span $\mathbb{R}^2$ and hence we shall append a vector that does not belong to the span of $\{(1, 1)\}$, say $\{(1, 2\}$. Now, the set $\{(1, 1), (1, 2)\}$ is a linearly independent set (by the choice of the vectors) and also it is easy to verify that any vector in $\mathbb{R}^2$ can be written as a linear combination of these vectors. Thus $\{(1, 1), (1, 2)\}$ spans $\mathbb{R}^2$ and hence is a basis for $\mathbb{R}^2$.

Example 4.4.3. Let $V = \mathbb{R}^2$. It is easy to check that $\{(1, 2), (2, 3), (2, 4), (4, 5)\}$ spans V. Now, we shall delete those vectors which can be written as a linear combination of the other vectors. Clearly, $(2, 4) = 2(1, 2)$. So, we can remove one of those. Let us remove $(2, 4)$. Now the set that remains is $\{(1, 2), (2, 3), (4, 5)\}$. Note that any of the vectors in this set can be written as a linear combination of the other two vectors. For instance, $(1, 2) = \frac{3}{2}(2, 3) - \frac{1}{2}(4, 5)$. Removing any one of the vectors, we get a linearly independent set that also spans V and hence a basis for V.

Observe that irrespective of what set that we start with or which method we follow, we always end up with a basis for $\mathbb{R}^2$ which has exactly two elements.

Question 97. Can you obtain a basis for $\mathbb{R}^3$ starting from the set $\{(1, 1, 1), (2, 4, 0), (0, 1, 1), (2, 2, 4), (3, 4, 5)\}$?

In general, any basis for $\mathbb{R}^n$ has **exactly** n elements.

4.4.1 Exercise

Question 98. Let V be a vector space which is defined as follows:

$$V = \{(x, y, z, w) \mid x + z = y + w\} \subseteq \mathbb{R}^4$$

with usual addition and scalar multiplication. Which of the following set forms a basis of V?

- Option 1: $\{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 0, 1)\}$.
- Option 2: $\{(1, 1, 0, 0), (0, 1, -1, 0), (0, -1, 0, 1)\}$.
- **Option 3:** $\{(1, 0, -1, 0), (1, 1, 0, 0), (1, 0, 0, 1)\}$.
- **Option 4:** $\{(1, 1, 0, 0), (0, 1, 1, 0), (0, -1, 0, 1)\}$.

Question 99. If $\{v_1, v_2, v_3\}$ forms a basis of $\mathbb{R}^3$, then which of the following are true?

- **Option 1:** $\{v_1, v_2, v_1 + v_3\}$ forms a basis of $\mathbb{R}^3$.
- **Option 2:** $\{v_1, v_1 + v_2, v_1 + v_3\}$ forms a basis of $\mathbb{R}^3$.
- **Option 3:** $\{v_1, v_1 + v_2, v_1 - v_3\}$ forms a basis of $\mathbb{R}^3$.
- **Option 4:** $\{v_1, v_1 - v_2, v_1 - v_3\}$ forms a basis of $\mathbb{R}^3$.

4.5 Dimension of a vector space

It is indeed an exciting fact that any two bases of a vector space has the same number of elements! This unique number is called the dimension of the vector space. Let us define it formally.

Definition 4.5.1. The dimension of a vector space V, is the size/cardinality of a basis B of V. It is denoted by $\dim(V)$.

(Prof. Sarang has used the terminology rank of a vector space in the lecture videos. We shall not use that terminology here, instead we shall stick to dimension of a vector space throughout the book.)

As already mentioned, any basis for $\mathbb{R}^n$ has exactly n elements and hence $\dim(\mathbb{R}^n) = n$.

Now let us calculate the dimensions of certain subspaces.

Example 4.5.1. Let W be the subspace of $\mathbb{R}^3$ spanned by $\{(1, 0, 0), (0, 1, 0), (3, 5, 0)\}$. To calculate the dimension of W, we first construct a basis for W. Clearly $\{(1, 0, 0), (0, 1, 0), (3, 5, 0)\}$ spans W and hence we need to check whether it is linearly independent or not. But $\{(1, 0, 0), (0, 1, 0)\}$ is linearly independent and also spans W (why?). Thus $\{(1, 0, 0), (0, 1, 0)\}$ is a basis for W and hence dimension of W is 2.

Note that the three vectors in the previous example lie on the xy-plane. We know to get a basis for the xy-plane and that is what we have obtained in the previous example too.

Next, we shall do the same example in terms of matrices, which might be an easier approach to follow. The steps to be followed are given below:

- Step 1: Write down the vectors as rows of a matrix A.
- Step 2: Apply row reduction on A to obtain a row echelon form.
- Step 3: The non-zero rows of the reduced matrix form a basis for W.

The number of non-zero rows in the reduced matrix is the dimension of the subspace W.

Note that in the above mentioned method, the basis vectors need not be in the set that we started with. Now, we propose another method by which we get a basis from the given set of vectors.

The vectors are arranged as columns of a matrix and the matrix, say A, is reduced to row echelon/reduced row echelon form, say R. The columns of A corresponding to the columns of R containing pivots form a basis for the column space of A. We illustrate this through examples.

Example 4.5.2. Let W be the subspace of $\mathbb{R}^3$ spanned by $\{(1, 1, 2), (2, -1, 4), (3, 0, 6)\}$. Let us now find a basis for W.

We first form a matrix A whose columns are the given vectors. Then $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 0 \\ 2 & 4 & 6 \end{pmatrix}$.

After row reduction, we get the row echelon form $R = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Note that the pivots are present in the first and second columns and hence the first and second columns of the matrix A form a basis for the subspace W. Thus a basis for W is given by $\{(1, 1, 2), (2, -1, 4)\}$.

Example 4.5.3. Let W be the subspace of $\mathbb{R}^4$ spanned by the set of vectors $\{(1, 1, 0, 0), (2, 1, -1, 4), (3, 4, 1, -4), (3, 2, -1, 4)\}$.

Here the matrix A is $\begin{pmatrix} 1 & 2 & 3 & 3 \\ 1 & 1 & 4 & 2 \\ 0 & -1 & 1 & -1 \\ 0 & 4 & -4 & 4 \end{pmatrix}$. By converting A into row echelon form, we get the matrix $R = \begin{pmatrix} 1 & 2 & 3 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. The pivots are present in the first and second columns and hence the first two vectors form a basis for W. Thus $\{(1,1,0,0),(2,1,-1,4)\}$ forms a basis for W.

4.5.1 Exercise

Question 100. Choose the set of correct options.

- **Option 1:** The dimension of the vector space $M_{1\times 2}(\mathbb{R})$ is 2.
- Option 2: The dimension of the vector space $M_{2\times 1}(\mathbb{R})$ is 1.
- Option 3: The dimension of the vector space $M_{3\times 3}(\mathbb{R})$ is 3.
- **Option 4:** A basis of $M_{2\times 2}(\mathbb{R})$ is the set $\left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$.

Question 101. Consider the following sets:

- $V_1 = \{A \mid A \in M_{2\times 2}(\mathbb{R})$ and A is a symmetric matrix, i.e., $A = A^T\}$
- $V_2 = \{A \mid A \in M_{2\times 2}(\mathbb{R})$ and A is a scalar matrix$\}$
- $V_3 = \{A \mid A \in M_{2\times 2}(\mathbb{R})$ and A is a diagonal matrix$\}$
- $V_4 = \{A \mid A \in M_{2\times 2}(\mathbb{R})$ and A is an upper triangular matrix$\}$
- $V_5 = \{A \mid A \in M_{2\times 2}(\mathbb{R})$ and A is a lower triangular matrix$\}$

All $V_i, i = 1,2,3,4,5$ are subspaces of the vector space $M_{2\times 2}(\mathbb{R})$.
Choose the set of correct options.

- **Option 1:** The dimension of V_1 is 3.
- Option 2: The dimension of V_2 is 3.
- Option 3: The dimension of V_3 is 1.
- **Option 4:** The dimension of V_4 is 3.
- **Option 5:** The dimension of V_5 is 3.

Question 102. Find the dimension of the vector space

$$V = \{A \mid \text{ sum of entries in each row is 0, and } A \in M_{3\times 2}(\mathbb{R})\}.$$

[Answer: 3]

Question 103. Find the dimension of the vector space

$$V = \{(x,y,z,w) \mid x+y = z+w, x+w = y+z, \text{ and } x,y,z,w \in \mathbb{R}\}.$$

[Answer: 2]

5. Rank and Nullity of a matrix

"If only I had the Theorems! Then I should find the proofs easily enough"

— Bernhard Riemann

5.1 Introduction

In this chapter, we will define what the rank of a matrix is and outline the procedure to calculate the rank of a matrix. Then, we move on to define the concept of null space of a matrix and outline a procedure to calculate its basis and dimension, namely, the nullity of the matrix. We end this chapter by stating the very important rank-nullity theorem.

5.2 Rank of a matrix

Suppose a person X starts moving from position A. After 10 minutes, he arrives at a position B that is 3km to the west of A and 4km to the south of A. This information helps us to identify the position of X after 10 minutes. In addition to this, if we also say that X is 5km to the south west of the starting point A, we still arrive at the same location B. Hence this information is redundant. This is because in a two dimensional plane, with two directions we can identify the position of an object completely (recall that the dimension of $\mathbb{R}^2$ is 2). The role of linearly independent vectors is highlighted here. In this sense, if we look to find the number of "directions" needed to completely understand a matrix, we arrive at the concept of the rank of a matrix.

Definition 5.2.1. Let A be an $m \times n$ matrix. The rank of the matrix A is the number of linearly independent columns of A.

The subspace spanned by the columns of A is called the **column space** of A and the dimension of the column space is called the **column rank** of A. Similarly, the subspace spanned by the rows of A is called the **row space** of A and the dimension of the row space is called the **row rank** of A.

It is a very important fact that the row rank and the column rank of a matrix are always equal and is called the rank of the matrix.

Example 5.2.1. Let $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -1 \\ 4 & -2 & 2 \end{pmatrix}$. The column space of A is the subspace spanned by $\{(1, -1, 4), (2, 0, -2), (3, -1, 2)\}$. Note that third vector is the sum of the first two vectors and hence the column rank (=dimension of the column space) of A is 2.

The row space is the subspace spanned by $\{(1, 2, 3), (-1, 0, -1), (4, -2, 2)\}$. Clearly, $(4, -2, 2) = -(1, 2, 3) - 5(-1, 0, -1)$ and hence the row rank (=dimension of the row space) of A is also 2.

A simple procedure to calculate the rank of a matrix is by converting the matrix into row echelon form or the reduced row echelon form and counting the number of non-zero rows of the matrix. Performing elementary row operations does not affect the number of linearly independent rows and hence the rank of the matrix.

The row echelon form of the matrix in the above example is $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Thus the rank, which is equal to the number of non-zero rows of the row echelon form of A is 2.

Example 5.2.2. Let $A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 5 & 3 & 0 & 0 \\ -1 & 0 & 1 & 2 \end{pmatrix}$. The row echelon form of A is $\begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & -\frac{5}{7} & 0 \\ 0 & 0 & 1 & \frac{7}{5} \end{pmatrix}$. Hence the rank (the number of non-zero rows in the row echelon form) of A is 3. Note that the column space is a subset of $\mathbb{R}^3$ and the row space is a subset of $\mathbb{R}^4$.

Example 5.2.3. If rank of the matrix $\begin{bmatrix} 2 & -3 & 4 \\ 0 & 1 & -2 \\ 1 & -3 & a \end{bmatrix}$ is 2 then what is the value of a? Note that if the rank of the matrix is 3, then the dimension of the column space is 3 and hence the column vectors are linearly independent. This happens when the determinant of the matrix is non-zero. But, now since it is given that the rank is 2, we have only two linearly independent vectors and one the vectors is dependent on the other two. This forces the determinant of the matrix to be 0. By equating the determinant of the matrix to 0, we are making the rank to be atmost 2. $det(A) = 2a - 10$ and for rank to be less than 3, $det(A) = 0$ and thus we get $a = 5$. By putting $a = 5$ and reducing A to REF, verify that the rank$(A) = 2$.

From the above explanation of the rank of a matrix, it is clear that the rank cannot exceed the number of columns or the number of rows of a matrix. Thus

- $\text{rank}(A) \leq \min\{m, n\}$, where A is an $m \times n$ matrix.
- $\text{rank}(AB) \leq \min\{\text{ rank}(A), \text{rank}(B)\}$.

5.2.1 Exercise

Question 104. Choose the set of correct options from the following.

- **Option 1:** Row rank and column rank of a matrix is always the same.
- **Option 2:** The rank of a zero matrix is always 0.
- **Option 3:** The rank of a matrix, all of whose entries are the same non-zero real number, must be 1.
- Option 4: Rank of the $n \times n$ identity matrix is 1 for any $n \in \mathbb{N}$.

Question 105. Find the rank of $A = \begin{bmatrix} 0 & 1 & 3 & 1 \\ 1 & 0 & 0 & 1 \\ 3 & 2 & 0 & 3 \end{bmatrix}$. [Answer: 3]

Question 106. Consider the following upper triangular matrix to choose the correct options.

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$$

where $a, b, c, d, e, f \in \mathbb{R}$.

- **Option 1:** If $f = 0$, then the rank of the matrix must be less than or equal to 2.
- Option 2: If $f = 0$, then the rank of the matrix must be exactly 2.
- **Option 3:** If a, b, c, d, e, f are all non-zero then the rank of the matrix must be 3.
- **Option 4:** If a, d, f are non-zero then the rank of the matrix must be 3.

Question 107. What is the rank of the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{bmatrix}$, where x, y, z are distinct non-zero integers? [Answer: 3]

5.3 Nullity of a matrix

Let A be an $m \times n$ matrix. The subspace $\{x \in \mathbb{R}^n : Ax = 0\}$ is called the null space of the matrix A. This is nothing but the solution space of the homogeneous system of linear equations $Ax = 0$. (One can verify quickly that the null space is actually a subspace.) In other words, the null space contains those vectors of $\mathbb{R}^n$ that are "killed" by the matrix A, i.e., taken to the zero vector by the matrix A. The dimension of the null space of A is the nullity of the matrix A.

Example 5.3.1. Let $A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 4 & -1 \\ 3 & 3 & -1 \end{pmatrix}$. The null space of A is nothing but the solutions of the system $Ax = 0$. Recall that we can solve a system of linear equations using Gaussian elimination. Since we have a homogeneous system, the last column of the augmented matrix is not going to be affected by any row operations and will remain zero throughout and hence can be ignored. Since this is our first example, we shall anyway include the last column too. The augmented matrix is $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 2 & 4 & -1 & 0 \\ 3 & 3 & -1 & 0 \end{pmatrix}$. The reduced row echelon form is $\begin{pmatrix} 1 & 0 & -\frac{1}{6} & 0 \\ 0 & 1 & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. x_3 is an independent variable as there is no pivot in the third column of the reduced row echelon form. Thus the solution set (null space of A) is $\{(-\frac{k}{6}, -\frac{k}{6}, k) : k \in \mathbb{R}\}$. The dimension of the null space is 1, that is the nullity of A is 1.

Note that we had one independent variable in the reduced row echelon form in the previous example and the nullity was equal to 1. It is in fact true in general that the nullity of the matrix is equal to the number of independent variables in the reduced row echelon form of the matrix.

We got the complete solution set of the null space by doing the following:

- Assign an arbitrary value k_i to the independent variables.
- Compute the value of the dependent variables in terms of these k_i's from the unique row they occur in.
- The set of all solutions is obtained by letting k_i's vary over $\mathbb{R}$.

Once we have the subspace, it is easy to get a basis for the null space. By substituting $k_i = 1$ and $k_j = 0$ for all $j \neq i$, as i varies, we get a basis for the null space of A.

Example 5.3.2. Consider the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$. To find the null space of A, we find the solutions of $Ax = 0$. We do row operations on the augmented matrix (but recall the last column of zeros are not going to be affected by any of the row operations and hence can be ignored) and obtain the reduced form as $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. There is only one column with a pivot. x_1 is dependent and x_2 and x_3 are independent. So the nullity is 2. Recall the procedure given above. We just assign arbitrary values to the independent variables x_2 and x_3, say $x_2 = t$ and $x_3 = s$. Thus we have, from the reduced system, $x_1 = -2t - 3s$. Hence the null space of A is $\{(-2t-3s, t, s) : t, s \in \mathbb{R}\}$. We can easily get a basis, by putting $t = 1$ and $s = 0$ and a second vector by putting $t = 0$ and $s = 1$, i.e., $\{(-2, 1, 0), (-3, 0, 1)\}$ forms a basis for the null space of A.

Example 5.3.3. Let $A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 2 & 1 & 1 & 0 \\ 3 & 3 & 1 & 4 \end{pmatrix}$. First we find the reduced form of the matrix. $\begin{pmatrix} 1 & 0 & -\frac{2}{3} & -\frac{4}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{8}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$. x_1 and x_2 are dependent, whereas x_3 and x_4 are independent. We assign $x_3 = t$ and $x_4 = s$. Thus the null space is $\{(-\frac{2}{3}t + \frac{4}{3}s, \frac{1}{3}t - \frac{8}{3}s, t, s : t, s \in \mathbb{R}\}$. Clearly nullity is 2 since there are two independent variables. A basis can be got by putting $t = 1$, $s = 0$ and another vector by putting $t = 0$, $s = 1$. $\{(-\frac{2}{3}, \frac{1}{3}, 1, 0), (\frac{4}{3}, -\frac{8}{3}, 0, 1)\}$ is a basis for the null space obtained by the above procedure.

5.3.1 Exercise

Question 108. The null space of a matrix $A_{4\times 3}$ is

- Option 1: the subspace $W = \{x \in \mathbb{R}^4 \mid Ax = 0\}$ of $\mathbb{R}^4$.
- **Option 2**: the subspace $W = \{x \in \mathbb{R}^3 \mid Ax = 0\}$ of $\mathbb{R}^3$.
- Option 3: the subspace $W = \{x \in \mathbb{R}^2 \mid Ax = 0\}$ of $\mathbb{R}^2$.
- Option 4: the first column of the matrix $A_{4\times 3}$.

Question 109. Nullity of a matrix $A_{3\times 4}$ is

- Option 1: 3
- Option 2: 4

- **Option 3:** the number of independent variables in the system of linear equations $Ax = 0$, where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.
- Option 4: the number of dependent variables in the system of linear equations $Ax = 0$, where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

Question 110. Choose the correct set of options from the following.

- Option 1: The nullity of a non-zero scalar matrix of order 3 must be 3.
- **Option 2:** The nullity of a non-zero scalar matrix of order 3 must be 0.
- Option 3: The nullity of a non-zero diagonal matrix of order 3 must be 3.
- **Option 4:** The nullity of a non-zero diagonal matrix of order 3 can be at most 2.

Question 111. Consider the coefficient matrix A of the following system of linear equations:

$$\begin{aligned} x_1 + x_2 + x_4 &= 0 \\ x_2 + x_3 &= 0 \\ x_1 - x_3 + x_4 &= 0 \end{aligned}$$

Which one of the following vector spaces represents the null space of A?

- Option 1: $\{(t_1 + t_2, t_1, t_1, t_2) \mid t_1, t_2 \in \mathbb{R}\}$
- **Option 2:** $\{(t_1 - t_2, -t_1, t_1, t_2) \mid t_1, t_2 \in \mathbb{R}\}$
- Option 1: $\{(t_1 + t_2, -t_1, t_1, t_2) \mid t_1, t_2 \in \mathbb{R}\}$
- Option 1: $\{(t_1 - t_2, -t_1, t_1, -t_2) \mid t_1, t_2 \in \mathbb{R}\}$

Also, find a basis for the null space.

5.4 The rank-nullity theorem

In this section, we state one of the very important theorems of linear algebra, namely, the rank-nullity theorem. Here, we state the theorem for matrices. A similar version for linear transformations will be stated in Chapter 6.

Theorem 5.4.1. *Let A be an $m \times n$ matrix. Then rank(A)+nullity(A) = n.*

We shall not prove the theorem, but it can be easily see from the following observation. The rank of the matrix is nothing but the number of non-zero rows in the (reduced) row echelon form of the matrix. This also denotes the number of dependent variables in the reduced form of the homogeneous system $Ax = 0$, because the number of non-zero rows will be equal to the number of rows with pivot elements. We also noticed earlier that the nullity of the matrix is equal to the number of independent variables of the reduced form of the homogeneous system $Ax = 0$. Thus the sum of dependent variables (rank) and independent variables (nullity) gives the total number of variables (which equals the number of columns of the matrix).

This theorem simplifies our job of finding the rank and nullity of a matrix. Given a matrix, once we know the rank, we get the nullity using this theorem and vice versa.

Example 5.4.1. Let nullity of the matrix $A_{3\times5}$ be 2. Then the rank of A can be got using the rank-nullity theorem. rank(A) + nullity(A) = 5. Thus rank(A) = 2.

Checking linear dependence of n vectors in the vector space $\mathbb{R}^n$

Recall that the dimension of $\mathbb{R}^n$ is n and we have calculated different bases for $\mathbb{R}^n$. The most commonly used one is the standard basis consisting of n vectors with the i^{th} vector having 1 at the i^{th} position and zero at all other positions.

In general, if a set of n vectors in $\mathbb{R}^n$ is linearly independent, then it must be a basis for $\mathbb{R}^n$. Recall the procedure to check whether or not a set of vectors is linearly independent. We try to equate the linear combination of those vectors to 0 and solve for the coefficients. Basically, we try to solve a homogeneous system whose coefficient matrix A is the matrix whose columns are the given vectors. If the system has a unique solution (that is the zero solution), this means that the coefficients are all zero and hence set of vectors in independent. In the case of infinitely many solutions, the set of vectors become linearly dependent. Note that since we are dealing with a set of n vectors in $\mathbb{R}^n$, the coefficient matrix of the system will be a square matrix. Thus the homogeneous system $Ax = 0$ has a unique solution if $det(A) \neq 0$. In this case, the set of vectors are linearly independent and hence form a basis for $\mathbb{R}^n$. In the case when $det(A) = 0$, the set of vectors are linearly dependent and do not form a basis for $\mathbb{R}^n$.

Example 5.4.2. Consider the set of vectors $\{(1,2,3),(1,0,1),(4,5,7)\}$. Since we have a set of 3 vectors, it forms a basis for $\mathbb{R}^3$ only if it is linearly independent (in this case, it will be maximal linearly independent!). We check if this forms a basis for $\mathbb{R}^3$ by finding the determinant of the matrix A whose columns are the given vectors. That is $A = \begin{pmatrix} 1 & 1 & 4 \\ 2 & 0 & 5 \\ 3 & 1 & 7 \end{pmatrix}$. $det(A) = 4 \neq 0$ and hence the set is linearly independent. But this is enough to check that it is a basis for $\mathbb{R}^3$ (Why?).

5.4.1 Exercise

Question 112. Which of the following options are correct for a square matrix A of order $n \times n$, where n is any natural number?

- **Option 1:** If the determinant is non-zero, then the nullity of A must be 0.
- Option 2: If the determinant is non-zero, then the nullity of A may be non-zero.
- **Option 3:** If the nullity of A is non-zero, then the determinant of A must be 0.
- Option 4: If the nullity of A is non-zero, then the determinant of A may be non-zero.

Question 113. Choose the set of correct statements.

- **Option 1:** If nullity of a 3×3 matrix is c for some natural number c, $0 \leq c \leq 3$, then the nullity of $-A$ will also be c.
- Option 2: $nullity(A+B) = nullity(A) + nullity(B)$.
- **Option 3:** Nullity of the zero matrix of order $n \times n$, is n.
- Option 4: Nullity of the zero matrix of order $n \times n$, is 0.
- **Option 5:** There exist square matrices A and B of order $n \times n$, such that nullity of both A and B is 0, but the nullity of $A+B$ is n.

Question 114. If A is a 3×4 matrix, then which of the following options are true?

- **Option 1:** $rank(A)$ must be less than or equal to 3.
- **Option 2:** $nullity(A)$ must be greater than or equal to 1.
- **Option 3:** If A has 2 columns which are non-zero and not multiples of each other, while the remaining columns are linear combinations of these 2 columns, then $nullity(A) = 2$.

- Option 4: If A has 2 columns which are non-zero and not multiples of each other, while the remaining columns are linear combinations of these 2 columns, then $nullity(A) = 1$.

Question 115. Let $Ax = 0$ be a homogeneous system of linear equations which has infinitely many solutions, where A is an $m \times n$ matrix ($m > 1, n > 1$). Which of the following statements are possible?

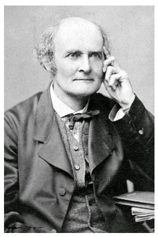
- **Option 1:** $rank(A) = m$ and $m < n$.
- Option 2: $rank(A) = m$ and $m > n$.
- Option 3: $rank(A) = m$ and $m = n$.
- **Option 4:** $nullity(A) = n$.
- **Option 5:** $nullity(A) \neq 0$.

Question 116. Let $Ax = 0$ be a homogeneous system of linear equations which has a unique solution, where $A \in \mathbb{R}^{n \times n}$. What is the nullity of A? [Answer : 0]

Question 117. Suppose $x_1 \neq 0$ solves $Ax = 0$, where $A \in \mathbb{R}^{2 \times 4}$. What is the minimum number of elements in a linearly independent subset of null space of A that also spans the set of solutions of $Ax = 0$? [Answer : 2]

INDIAN INSTITUTE OF TECHNOLOGY MADRAS • सिद्धिर्भवति कर्मजा •

6. Linear Transformation



"As for everything else, so for a mathematical theory: beauty can be perceived but not explained"

— Arthur Cayley

6.1 Linear Mapping

In previous course of mathematics we have studied about the function, like $f(x) = x^2 + 3x$ e.t.c., where the function's domain and codomain are either $\mathbb{R}$ or subset of $\mathbb{R}$.
In this section we are going to study some special type of function whose domain and codomain are vector spaces or a subspaces of a vector space with some specific properties.

Let see it with an example

Suppose there are 3 shops in a locality, the original shop A and two other shops B and C. The price of rice, dal and oil in each of these shops are as given in the table below:

	Rice (per kg)	Dal (per kg)	Oil (per kg)
Shop A	45	125	150
Shop B	40	120	170
Shop C	50	130	160

Based on these prices, how will we decide from which shop to buy your groceries?

To get this idea, lets write the expression for the total cost of buying x_1 kg of rice, x_2 kg of dal and x_3 kg of oil for each of the three shops and try compare them. So total cost of buying x_1 kg of rice, x_2 kg of dal and x_3 kg of oil from shop A is $45x_1 + 125x_2 + 150x_3$.

This can be thought as a function C_A and viewed as a matrix multiplication.

$$C_A(x_1, x_2, x_3) = 45x_1 + 125x_2 + 150x_3 = \begin{bmatrix} 45 & 125 & 150 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Now, observe that the function C_A holds special properties:

$$\begin{aligned} C_A(\alpha(x_1, x_2, x_3)) &= (\alpha x_1, \alpha x_2, \alpha x_3) \\ &= 45\alpha x_1 + 125\alpha x_2 + 150\alpha x_3 \\ &= \alpha(45x_1 + 125x_2 + 150x_3) \\ &= \alpha C_A(x_1, x_2, x_3) \end{aligned}$$

$$\begin{aligned} C_A((x_1, x_2, x_3) + (y_1, y_2, y_3)) &= (x_1 + y_1, x_2 + y_2, x_3 + y_3) \\ &= 45(x_1 + y_1) + 125(x_2 + y_2) + 150(x_3 + y_3) \\ &= (45x_1 + 125x_2 + 150x_3) + (45y_1 + 125y_2 + 150y_3) \\ &= C_A(x_1, x_2, x_3) + C_A(y_1, y_2, y_3) \end{aligned}$$

These properties called linearity property of the function C_A. This property can be verified using above matrix multiplication expression of C_A also.

Similarly, for shops B and C , we get function C_B and C_C (holds the same properties as function C_A) whose expression and matrix form are:

$$C_B(x_1, x_2, x_3) = 40x_1 + 120x_2 + 170x_3 = \begin{bmatrix} 40 & 120 & 170 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$C_C(x_1, x_2, x_3) = 50x_1 + 130x_2 + 160x_3 = \begin{bmatrix} 50 & 130 & 160 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Comparing these expression, it is clear that for any quantities x_1, x_2, x_3 that one would buy (i.e., when x_1, x_2, x_3 are positive), the third expression yield larger values than fist one.

However, the comparison between the second expression and the others depends on the quantities of the item bought, i.e., on x_1, x_2, x_3.

A natural way to make this comparison would be the to create a vector of costs i.e., $(C_A(x_1, x_2, x_3), C_B(x_1, x_2, x_3), C_C(x_1, x_2, x_3))$.

We can think of the cost vector a function c from $\mathbb{R}^3$ to $\mathbb{R}^3$ by setting these expression as the coordinates in $\mathbb{R}^3$ i.e.,

$$c(x_1, x_2, x_3) = (C_A(x_1, x_2, x_3), C_B(x_1, x_2, x_3), C_C(x_1, x_2, x_3))$$

$$c(x_1, x_2, x_3) = (45x_1 + 125x_2 + 150x_3, 40x_1 + 120x_2 + 170x_3, 50x_1 + 130x_2 + 160x_3)$$

We can use matrix multiplication to express the cost function c in a compact form and extract its properties:

$$c(x_1, x_2, x_3) = \begin{bmatrix} 45 & 125 & 150 \\ 40 & 120 & 170 \\ 50 & 130 & 160 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

As we have seen the linearity of the cost function C_A, C_B, C_C, the cost function $c : \mathbb{R}^3 \to \mathbb{R}^3$ also follows linearity property:

$$\begin{aligned} c(\alpha(x_1, x_2, x_3) + (y_1, y_2, y_3)) &= c((\alpha x_1 + y_1, \alpha x_2 + y_2, \alpha x_3 + y_3)) \\ &= c(C_A(\alpha x_1 + y_1, \alpha x_2 + y_2, \alpha x_3 + y_3), C_B(\alpha x_1 + y_1, \alpha x_2 + y_2, \alpha x_3 + y_3), \\ &\quad C_C(\alpha x_1 + y_1, \alpha x_2 + y_2, \alpha x_3 + y_3)) \\ &= c(\alpha C_A(x_1, x_2, x_3) + C_A(y_1, y_2, y_3), \alpha C_B(x_1, x_2, x_3) + C_B(y_1, y_2, y_3), \\ &\quad \alpha C_C(x_1, x_2, x_3) + C_C(y_1, y_2, y_3)) \\ &= \alpha c(C_A(x_1, x_2, x_3), C_B(x_1, x_2, x_3), C_C(x_1, x_2, x_3)) \\ &\quad + c(C_A(y_1, y_2, y_3), C_B(y_1, y_2, y_3), C_C(y_1, y_2, y_3)) \\ &= \alpha c(x_1, x_2, x_3) + c(y_1, y_2, y_3) \end{aligned}$$

We can verify it using above matrix multiplication expression of c.

So we can say that the cast function c is linear mapping.

6.1.1 Linear Mapping: The formal definition

Definition 6.1.1. A linear mapping f from $\mathbb{R}^3 \to \mathbb{R}^3$ can be defined as follows:

$$f(x_1, x_2, \ldots x_n) = \left(\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \ldots, \sum_{j=1}^{n} a_{nj}x_j \right)$$

Where the coefficient a_{ij} are real numbers (scalars). A linear mapping can be thought of as a collection of linear combinations.

We can write the expression on the RHS in matrix form as Ax, where

$$A = \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{m1} & a_{m2} \dots a_{mn} \end{bmatrix} \text{ and } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

So using the matrix multiplication expression or directly from the expression $f(x_1, x_2, \dots x_n) = \left(\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \dots, \sum_{j=1}^{n} a_{nj}x_j\right)$, we can easily check that the function f holds linearity:

$$f(\alpha(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)) = A(\alpha x + y)$$

$$= \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{m1} & a_{m2} \dots a_{mn} \end{bmatrix} \left(\begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right)$$

$$= \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{m1} & a_{m2} \dots a_{mn} \end{bmatrix} \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{m1} & a_{m2} \dots a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \alpha f((x_1, x_2, \dots, x_n) + f(y_1, y_2, \dots, y_n))$$

6.1.2 Exercises

A book shop is organizing an year end sale. Price of any Bengali, Hindi, Tamil, and Urdu book is fixed as ₹200, ₹180, ₹230, and ₹250, respectively. Let $T(x, y, z, w)$ denote the total price of x number of Bengali books, y number of Hindi books, z number of Tamil books, and w number of Urdu books. Table 6.1 shows the numbers of books of different languages purchased by some customers.

	Bengali	Hindi	Tamil	Urdu
Samprita	3	0	0	2
Srinivas	0	1	2	1
Anna	0	1	0	3
Tiyasha	2	2	0	1
Hasan	2	2	1	1

Table 6.1:

Answer the below questions from the given data.

(1) What will be the correct expression for $T(x, y, z, w)$?

- Option 1: $T(x, y, z, w) = (200 + 180 + 230 + 250)(x + y + z + w)$
- Option 2: $T(x, y, z, w) = x + y + z + w$
- Option 3: $T(x, y, z, w) = 200x + 230y + 180z + 250w$
- Option 4: $T(x, y, z, w) = 200x + 180y + 230z + 250w$

(2) Which of the following expressions represents the total price of the books purchased by Samprita?

- Option 1: $T(3, 0, 0, 2)$
- Option 2: $T(3, 0, 0, 2, 2)$
- Option 3: $T(3, 2)$
- Option 4: $T(5)$

(3) What will be total price (in ₹) of the books purchased by Tiyasha?

(4) What will be total price (in ₹) of the books purchased by Hasan?

(5) Which of the following expressions represent the total price of the books purchased by Srinivas?

- Option 1: $2T(0, 1, 1, 0) + T(0, 0, 0, 1)$
- Option 2: $T(0, 1, 0, 0) + T(0, 0, 1, 1)$
- Option 3: $T(0, 1, 0, 0) + 2T(0, 0, 1, 0) + T(0, 0, 0, 1)$
- Option 4: $T(0, 1, 1, 0) + T(0, 0, 1, 1)$

(6) Which of the following expressions represents the total price of the books purchased by Anna?

- Option 1: $T(0, 1, 0, 0) + T(0, 0, 0, 1)$

- Option 2: $T(0,1,0,0) + T(0,0,1,0)$
- Option 3: $T(0,1,0,0) + 3T(0,0,0,1)$
- Option 4: $T(0,1,1,0) + 2T(0,0,1,0)$

(7) Which of the following expressions represent the total price of the books purchased by Srinivas and Anna together?

- Option 1: $T(0,2,2,4)$
- Option 2: $T(2,2,4)$
- Option 3: $T(0,2,2,1)$
- Option 4: $T(0,1,2,1) + T(0,1,0,3)$

(8) Which of the following expressions represent the difference between the total price of the books purchased by Samprita and Tiyasha?

- Option 1: $|T(5,2,0,3)|$
- Option 2: $|T(1,-2,0,1)|$
- Option 3: $|T(3,0,0,2) - T(2,2,0,1)|$
- Option 4: $|T(3,0,0,2) + T(-2,-2,0,-1)|$

6.2 Linear Transformation

Definition 6.2.1. A function $T : V \to W$ between two vector spaces V and W is said to be a linear transformation if for any two vectors v_1 and v_2 in the vector space V and for any $c \in \mathbb{R}$ (scalar) the following conditions hold:

- $T(v_1 + v_2) = T(v_1) + T(v_2)$
- $T(cv_1) = cT(v_1)$

Form the definition of linear transformation it is clear that linear transformation is linear mapping.

Example 6.2.1. Consider a mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined as

$$T(x, y) = (2x, y)$$

Let $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$, then

$$T(v_1 + v_2) = T((x_1, y_1) + (x_2, y_2)) = T(x_1 + x_2, y_1 + y_2)$$

$$\implies T(v_1 + v_2) = (2(x_1 + x_2), (y_1 + y_2)) = ((2x_1 + 2x_2), (y_1 + y_1))$$

$$\implies T(v_1 + v_2) = (2x_1, y_1) + (2x_2, y_2) = T(v_1) + T(v_2)$$

Let $c \in \mathbb{R}$, then

$$T(cv_1) = T(c(x_1, y_1)) = T(cx_1, cy_1) = (2cx_1, cy_1) = c(2x_1, y_1) = cT(v_1)$$

Hence, T is a linear transformation.

We now state the following proposition which is quite evident from the definition of linear transformations.

Proposition 6.2.1. *A linear transformation $T : V \to W$, where V, W are vector spaces, $T(0) = 0$ i.e., Image of the zero vector of V is the zero vector of W.*

Proof. Since T is a linear transformation so we can write

$$T(0 + 0) = T(0) + T(0)$$

$$\implies T(0) = T(0) + T(0)$$

$$\implies T(0) = 0$$

□

6.2.1 Exercises:

Check which of the following mapping is linear transformation

1. Consider a mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that
$$T(x, y) = (2x, 0)$$

2. Consider a mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that
$$T(x, y, z) = (\frac{x}{2}, 3y, 5z)$$

3. Consider a mapping $T : \mathbb{R}^3 \to \mathbb{R}^4$ such that
$$T(x, y, z) = (4y - z, 3y + \frac{11}{19}z, 5x - 2z, 23y)$$

4. Consider a mapping $T : \mathbb{R} \to \mathbb{R}^3$ such that
$$T(t) = (t, 3t, \frac{23}{89}t)$$

5. Consider a mapping $T : \mathbb{R}^2 \to \mathbb{R}$ such that
$$T(x, y) = x$$

6.2.2 Images of the vectors in the basis of a vector space

Example 6.2.2. Let us choose the standard basis $\{(1,0),(0,1)\}$ of $\mathbb{R}^2$. Define the linear transformation as follows:

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T((1,0)) = (2,0)$$
$$T((0,1)) = (0,1)$$

Using the information we have the following:

$$\begin{aligned} T(x,y) &= T(x(1,0)+y(0,1)) \\ &= T(x(1,0)) + f(y(0,1)) \\ &= xT(1,0) + yT(0,1) \\ &= x(2,0) + y(0,1) \\ &= (2x,y) \end{aligned}$$

Hence the explicit definition of T is given by

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$

$$T(x,y) = (2x,y)$$

Example 6.2.3. If we choose a different basis for $\mathbb{R}^2$, then we will get different linear transformation. Let us choose $\{(1,0),(1,1)\}$ to be a basis for $\mathbb{R}^2$. Define the linear transformation as we have defined earlier:

$$f : \mathbb{R}^2 \to \mathbb{R}^2$$
$$f((1,0)) = (2,0)$$
$$f((1,1)) = (1,1)$$

Now we have,

$$\begin{aligned} f(x,y) &= f((x-y)(1,0)+y(1,1)) \\ &= f((x-y)(1,0)) + f(y(1,1)) \\ &= (x-y)f(1,0) + yf(1,1) \\ &= (x-y)(2,0) + y(1,1) \\ &= (2x-2y,0) + (y,y) \\ &= (2x-y,y) \end{aligned}$$

Hence the explicit definition of T is given by

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$

$$T(x,y) = (2x-y,y)$$

We end this section with the following remark which is visualised by the above examples.

Remark 6.2.1. Let $T : V \to W$ be a linear transformation. It is enough to know the image of the basis elements of V to get the explicit definition of T.

6.2.3 Exercises

1 Choose the set of correct options.

- Option 1: Let $u = (3, 1, 0)$, $v = (0, 1, 7)$, and $w = (3, 0, -7)$. There is a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $T(u) = T(v) = (0, 0, 0)$ and that $T(w) = (5, 1, 0)$.
- **Option 2:** Let $u = (2, 1, 0)$, $v = (1, 0, 1)$, and $w = (3, 5, 6)$. There is a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ such that $T(u) = (1, 0)$, $T(v) = (0, 1)$ and that $T(w) = (5, 6)$.
- Option 3: Let $u = (1, 0, 0)$, $v = (1, 0, 1)$, and $w = (0, 1, 0)$. There is a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $T(u) = (0, 0, 0)$, $T(v) = (0, 0, 0)$, $T(w) = (0, 0, 0)$, and $T(1, 4, 2) = (2, 4, 1)$.
- **Option 4:** Let $u = (1, 0)$, and $v = (0, 1)$,. There is a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(u) = (\pi, 1)$, and $T(v) = (1, e)$.

6.3 Injective and surjective linear transformations

In our earlier courses we have seen that a function is called injective if there are no two elements from the domain which map to a same image and a function is called surjective if every element of the codomain of the function has a pre-image. Similarly,

1. a linear transformation $T : V_1 \to W$, where V, W are vector spaces, is called a **monomorphism**, if T is an injective map from V to W, i.e., $T(v_1) = T(v_2) \implies v_1 = v_2$.
2. a linear transformation $T : V_1 \to W$, where V, W are vector spaces, is called an **epimorphism**, if T is a surjective map from V to W, i.e., for every $w \in W$ there exists $v \in V$ such that $T(v) = w$.
3. a linear transformation $T : V \to W$ is **isomorphism** if it is both injective (monomorphism) and surjective (epimorphism).

Example 6.3.1. Consider the following linear transformation T defined as follows:

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(x, y) = (2x, y)$$

Checking of injectivity: If $T(x_1, y_1) = T(x_2, y_2)$, then $(2x_1, y_1) = (2x_2, y_2)$. Hence $x_1 = x_2$ and $y_1 = y_2$, i.e., $(x_1, y_1) = (x_2, y_2)$. So, T is an injective linear transformation.

Checking of surjectivity: If $(u, v) \in \mathbb{R}^2$, then $(\frac{u}{2}, v) \in \mathbb{R}^2$ such that $T(\frac{u}{2}, v) = (u, v)$. So T is a surjective linear transformation.

Hence T is a bijective linear transformation, hence an isomorphism.

Example 6.3.2. Consider the following linear transformation T defined as follows:

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(x, y) = (2x, 0)$$

Checking of injectivity: If $T(x_1, y_1) = T(x_2, y_2)$ implies $(2x_1, 0) = (2x_2, 0)$, hence $x_1 = x_2$. But it does not guarantee that $y_1 = y_2$. As for example, $(1, 2)$ and $(1, 3)$ have the same image $(2, 0)$. Hence T is not injective.

Checking of surjectivity: There is no pre-image for the vector (u, v), where v is non-zero. So T is not surjective.

6.3.1 Null space and Range space of a linear transformation

Definition 6.3.1. Kernel or nullspace of a linear transformation: Let $T : V \to W$ be a linear transformation. We define kernel of T (denoted by $ker(T)$) to be the set of all vectors v in V such that $T(v) = 0$.

$$Ker(T) = \{v \in V \mid T(v) = 0\}$$

Definition 6.3.2. Image or range space of a linear transformation: Let $T : V \to W$ be a linear transformation. We define image of T (denoted by $Im(T)$) as follows:

$$Im(T) = \{w \in W \mid \exists v \in V \text{ for which } T(v) = w\}$$

Example 6.3.3. Let us consider the same example we considered before.

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(x, y) = (2x, y)$$

Let $v = (x, y)$ such that $T(v) = 0$

$$\implies T(x, y) = (0, 0)$$

$$\implies (2x, y) = (0, 0)$$

$$\implies x = 0 \text{ and } y = 0$$

$$\implies v = (x, y) = (0, 0)$$

$Ker(T) = \{(0, 0)\}$. Hence $ker(T)$ is a singleton set which contains only the zero vector the space $\mathbb{R}^2$.
$Im(T) = \mathbb{R}^2$. Hence $Im(T)$ is the whole space $\mathbb{R}^2$, as we have observed earlier it is surjective.

Theorem 6.3.1. *Kernel or nullspace of a linear transformation $T : V \to W$ is a vector subspace of V.*

Proof. Let $v, v' \in Ker(T)$. We have $T(v + v') = T(v) + T(v') = 0 + 0 = 0$. So, $v + v' \in Ker(T)$.
Moreover, $T(cv) = cT(v) = c0 = 0$, so $cv \in Ker(T)$ for all $c \in \mathbb{R}$.
Hence, $Ker(T)$ is a vector subspace of V.

□

Theorem 6.3.2. *Image or range space of a linear transformation $T : V \to W$ is a vector subspace of W.*

Proof. Let $w, w' \in Im(T)$. Then there exists $v, v' \in V$ such that, $T(v) = w$ and $T(v') = w'$. So we have $T(v+v') = T(v)+T(v') = w+w'$. Hence $w+w' \in Im(T)$.
Similarly, $T(cv) = cT(v) = cw$, for all $c \in \mathbb{R}$. Hence $cw \in Im(T)$ for all $c \in \mathbb{R}$.
So, $Im(T)$ is a vector subspace of W. □

Definition 6.3.3. Dimension of kernel or nullspace of linear transformation T is defined as $Nullity(T)$ and dimension of image or range space of linear transformation T is defined as $Rank(T)$.

Theorem 6.3.3. *T is an injective linear transformation if and only if $Ket(T) = \{0\}$.*

Proof. First assume that T is injective. Let $v \in Ker(T)$. Hence $T(v) = 0 = T(0)$, as T is injective then $v = 0$. Hence we can conclude $Ker(T) = \{0\}$.
Conversely, Let $Ker(T) = \{0\}$. Let for the vectors v_1 and v_2 in V, we have $T(v_1) = T(v_2)$. Then $T(v_1 - v_2) = T(v_1) - T(v_2) = 0$. Hence $v_1 - v_2 \in ker(T)$. Which implies $v_1 - v_2 = 0$ i.e., $v_1 = v_2$. Hence T is injective.

□

Remark 6.3.1. A linear transformation $T : V \to W$ is surjective if and only if $Im(T) = W$

Corollary 6.3.4. *$T : V \to W$ is an injective (resp. surjective) linear transformation if and only if $Nullity(T) = 0$ (resp. $Rank(T) = dim(W)$).*

Definition 6.3.4. We define two vector spaces V and W are isomorphic to each other iff there exists an isomorphism $T : V \to W$.

At this point we mention an important theorem with the sketch of the proof of it. We want the readers to complete the proof.

Theorem 6.3.5. *Any n dimensional vector space is isomorphic to $\mathbb{R}^n$.*

Proof. Let V be a vector space of dimension n. Consider a basis $\{v_1, v_2, \ldots, v_n\}$ of V. We can define a linear transformation $T : V \to \mathbb{R}^n$ as follows:

$$T : V \to \mathbb{R}^n$$
$$T(v_i) = e_i$$

Where $e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$, i.e. 1 at the i-th coordinate and 0 elsewhere. Let $v \in V$ such that $T(v) = 0$. As $v \in V$, then v can be written as $v = a_1v_1 + a_2v_2 + \ldots + a_nv_n$, for some a_i's in $\mathbb{R}$. Hence $T(v) = 0$ implies

$$T(a_1v_1 + a_2v_2 + \ldots + a_nv_n) = 0$$

$$a_1T(v_1) + a_2T(v_2) + \ldots + a_nT(v_n) = 0$$
$$a_1e_1 + a_2e_2 + \ldots + a_ne_n = 0$$
$$(a_1, a_2, \ldots, a_n) = 0$$

i.e, $a_i = 0$ for all $i = 1, 2, \ldots, n$. Hence, $v = 0$.
So $Ker(T) = \{0\}$, which implies that T is injective.
We leave the checking of surjectivity to the readers. □

6.4 Matrix representation of linear transformation

We begin this section with some examples.

Example 6.4.1. Let us consider the linear transformation

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(x, y) = (2x, y)$$

Now if we consider a basis $\{(1, 0), (0, 1)\}$ for $\mathbb{R}^2$ both for the domain and the codomain. Then we can represent it as a matrix as follows:

$$T(1, 0) = (2, 0) = 2(1, 0) + 0(0, 1)$$
$$T(0, 1) = (0, 1) = 0(1, 0) + 1(0, 1)$$

The matrix representation of the linear transformation is given by,

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Example 6.4.2. Now for the same transformation T, suppose we consider a different basis $\{(1, 0), (1, 1)\}$ of $\mathbb{R}^2$, both for the domain and the codomain. Then we can represent it as a matrix as follows:

$$T(1,0) = (2,0) = 2(1,0) + 0(1,1)$$
$$T(1,1) = (2,1) = 1(1,0) + 1(1,1)$$

Hence the matrix representation of the linear transformation is given by,

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

Note: It is important to note that changing basis gives us different matrices corresponding to same linear transformation.

Definition 6.4.1. Let $T : V \to W$ be a linear transformation. Let $\beta = \{v_1, v_2, \dots, v_n\}$ be a basis of V and $\gamma = \{w_1, w_2, \dots, w_m\}$ be a basis of W. So each $T(v_i)$ can be uniquely written as linear combination of w_j's, where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$.

$$\begin{aligned} T(v_1) &= a_{11}w_1 + a_{21}w_2 + \dots + a_{m1}w_m \\ T(v_2) &= a_{12}w_1 + a_{22}w_2 + \dots + a_{m2}w_m \\ &\vdots \\ T(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \dots + a_{mn}w_m \end{aligned}$$

The matrix corresponding to the linear transformation f with respect to the bases β and γ is given by,

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

Example 6.4.3. Consider the following linear transformation:

$$T : \mathbb{R}^3 \to \mathbb{R}^3$$
$$T(x, y, z) = (x - z, 2x + 3y + z, 3y + 3z)$$

We consider the standard ordered basis of $\mathbb{R}^3$ for both the domain and codomain.

$$T((1,0,0)) = (1,2,0) = 1(1,0,0) + 2(0,1,0) + 0(0,0,1)$$
$$T((0,1,0)) = (0,3,3) = 0(1,0,0) + 3(0,1,0) + 3(0,0,1)$$
$$T((0,0,1)) = (-1,1,3) = -1(1,0,0) + 1(0,1,0) + 3(0,0,1)$$

The matrix representation of T with respect to the standard ordered basis of $\mathbb{R}^3$ for both the domain and codomain.

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

Example 6.4.4. Consider the following linear transformation:

$$T : \mathbb{R}^3 \to \mathbb{R}^3$$
$$T(x, y, z) = (x - z, 2x + 3y + z, 3y + 3z)$$

We consider the ordered basis $\{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$ of $\mathbb{R}^3$ for both the domain and codomain.

$$\begin{aligned} T((1, 1, 0)) &= (1, 5, 3) = \left(\frac{3}{2}\right)(1, 1, 0) + \left(\frac{7}{2}\right)(0, 1, 1) + \left(-\frac{1}{2}\right)(1, 0, 1) \\ T((0, 1, 1)) &= (-1, 4, 6) = \left(-\frac{3}{2}\right)(1, 1, 0) + \left(\frac{11}{2}\right)(0, 1, 1) + \left(\frac{1}{2}\right)(1, 0, 1) \\ T((1, 0, 1)) &= (0, 3, 3) = 0(1, 1, 0) + 3(0, 1, 1) + 0(1, 0, 1) \end{aligned}$$

The matrix representation of T with respect to the standard ordered basis of $\mathbb{R}^3$ for both the domain and codomain.

$$\begin{bmatrix} \frac{3}{2} & -\frac{3}{2} & 0 \\ \frac{7}{2} & \frac{11}{2} & 3 \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Example 6.4.5. Consider the following linear transformation:

$$T : \mathbb{R}^3 \to \mathbb{R}^3$$
$$T(x, y, z) = (x - z, 2x + 3y + z, 3y + 3z)$$

We consider the ordered basis $\{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$ of $\mathbb{R}^3$ for the domain and the standard ordered basis for the codomain.

$$\begin{aligned} T((1, 1, 0)) &= (1, 5, 3) = 1(1, 0, 0) + 5(0, 1, 0) + 3(0, 0, 1) \\ T((0, 1, 1)) &= (-1, 4, 6) = -1(1, 0, 0) + 4(0, 1, 0) + 6(0, 0, 1) \\ T((1, 0, 1)) &= (0, 3, 3) = 0(1, 0, 0) + 3(0, 1, 0) + 3(0, 0, 1) \end{aligned}$$

The matrix representation of T with respect to the standard ordered basis of $\mathbb{R}^3$ for both the domain and codomain.

$$\begin{bmatrix} 1 & -1 & 0 \\ 5 & 4 & 3 \\ 3 & 6 & 3 \end{bmatrix}$$

6.4.1 Exercises

1) Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, such that $T(x, y) = (x, 0)$. Which of the following options are correct?

- Option 1: The matrices corresponds to T with respect to the standard ordered basis of $\mathbb{R}^2$, i.e., $\{(1,0),(0,1)\}$, for both the domain and co-domain is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
- **Option 2:** The matrices corresponds to T with respect to the standard ordered basis of $\mathbb{R}^2$, i.e., $\{(1,0),(0,1)\}$, for both the domain and co-domain is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.
- **Option 3:** T is neither one-one nor onto.
- Option 4: T is one-one but not onto.

Let $W = \{(x,y,z) \mid x = 2y+z\}$ be a subspace of $\mathbb{R}^3$. Let $\beta = \{(2,1,0),(1,0,1)\}$ be a basis of W. Let $T : W \to \mathbb{R}^2$ be a linear transformation such that $T(2,1,0) = (1,0)$ and $T(1,0,1) = (0,1)$. Answer the questions 2, 3 and 4 using the given information.

2) Which of the following is the appropriate definition of T?

- Option 1: $T(x,y,z) = (x,y)$
- Option 2: $T(x,y,z) = (x,z)$
- **Option 3:** $T(x,y,z) = (y,z)$
- **Option 4:** $T(x,y,z) = (x-y-z, x-2y)$

3) Choose the correct options.

- Option 1: T is one to one but not onto.
- Option 2: T is onto but not one to one.
- Option 3: T is neither one to one nor onto.
- **Option 4:** T is an isomorphism.

4) What will be the matrix representation of T with respect to the basis β for W and $\gamma = \{(1,1),(1,-1)\}$ for $\mathbb{R}^2$?

- **Option 1:** $\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
- Option 2: $\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- Option 3: $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
- Option 4: $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

5) Which option represents the kernel and image of the following linear transformation?

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$
$$T(x, y) = (x, 0)$$

- Option 1: $ker(T) = Span\{(1, 0)\}, Im(T) = Span\{(1, 0)\}$.
- Option 2: $ker(T) = Span\{(1, 0)\}, Im(T) = Span\{(0, 1)\}$.
- **Option 3:** $ker(T) = Span\{(0, 1)\}, Im(T) = Span\{(1, 0)\}$.
- Option 4: $ker(T) = Span\{(0, 1)\}, Im(T) = Span\{(0, 1)\}$.

6.5 Finding basis for null space and range space by Row reduced echelon form

Let $T : V \to W$ be a linear. Follow the steps below to find bases for the null space and the range space of T.

Step 1: Find the matrix A corresponding to T with respect to some standard ordered bases $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$ for V and W respectively.

Step 2: Use row reduction on A to obtain the matrix R which is in reduced row echelon form.

Step 3: The basis of the solution space of $Rx = 0$ is the basis of null space of matrix A and can be obtained by finding the pivot and non-pivot columns (dependent and independent variables) as seen earlier.

Step 4: The vectors $\begin{bmatrix} c_{11} \\ c_{12} \\ . \\ . \\ . \\ c_{1n} \end{bmatrix}, \begin{bmatrix} c_{21} \\ c_{22} \\ . \\ . \\ . \\ c_{2n} \end{bmatrix}, \ldots, \begin{bmatrix} c_{k1} \\ c_{k2} \\ . \\ . \\ . \\ c_{kn} \end{bmatrix}$ form a basis of the null space of A precisely when the vectors $v'_1, v'_2, \ldots, v'_k \in Ker(T)$ where $v'_i = \sum_{j=1}^{n} c_{ij} v_j$, form a basis for ker$(T)$. Use the basis obtained in step 3 to thus get a basis for $Ker(T)$.

Step 5: Recall that if $i_1, i_2, \ldots, i_r$ are the columns of R containing the pivot elements, then the same columns of A form a basis for the column space of A.

Step 6: The vectors $\begin{bmatrix} d_{11} \\ d_{12} \\ . \\ . \\ . \\ d_{1m} \end{bmatrix}, \begin{bmatrix} d_{21} \\ d_{22} \\ . \\ . \\ . \\ d_{2m} \end{bmatrix}, \ldots, \begin{bmatrix} d_{r1} \\ d_{r2} \\ . \\ . \\ . \\ d_{rm} \end{bmatrix}$ form a basis of the column space of A precisely when the vectors $w'_1, w'_2, \ldots, w'_r \in im(T)$ where $w'_i = \sum_{j=1}^{m} d_{ij} w_j$, form a basis for $Im(T)$. Use the basis obtained in step 5 to thus get a basis for $Im(T)$.

Using these steps we will find out the basis of nullspace and range space of a linear transformation in the following example:

Example 6.5.1. Consider the following linear transformation:

$$T : \mathbb{R}^4 \to \mathbb{R}^3$$

$$T(x_1, x_2, x_3, x_4) = (2x_1 + 4x_2 + 6x_3 + 8x_4, x_1 + 3x_2 + 5x_4, x_1 + x_2 + 6x_3 + 3x_4)$$

The matrix corresponding to the standard basis is as follows:

$$\begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 3 & 0 & 5 \\ 1 & 1 & 6 & 3 \end{bmatrix}$$

The row reduced echelon form of the above matrix is,

$$\begin{bmatrix} 1 & 0 & 9 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The required null space will the set of all vectors such that the following holds:

$$\begin{bmatrix} 1 & 0 & 9 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$$

Which gives us $x_1 = -9x_3 - 2x_4$ and $x_2 = 3x_3 - x_4$

Hence the null space is spanned by the vectors $\{(-9, 3, 1, 0), (-2, -1, 0, 1)\}$. Moreover, the first and second coloumn of the row reduced echelon form contains the pivot elements. Hence the range space is spanned by the vectors $\{(2, 1, 1), (4, 3, 1)\}$.

6.6 Rank-nullity theorem

We state the theorem without any proof.

Theorem 6.6.1. *Let $T : V \to W$ be a linear transformation. Then*

$$Rank(T) + Nullity(T) = dim(V).$$

We end this chapter with stating some immediate corollaries of Rank nullity theorem.

Corollary 6.6.2. *Let $T : V \to W$ be a linear transformation.*

- *If T is injective, then $Rank(T) = dim(V)$.*
- *If T is an isomorphism, then $dim(W) = dim(V)$.*

Proof. If T is injective, then $Nullity(T) = 0$. Hence from rank nullity theorem, $Rank(T) + 0 = dim(V)$, i.e., $Rank(T) = dim(V)$.
Moreover if T is surjective then, $dim(W) = Rank(T)$. Hence if T is an isomorphism, i.e., T is both injective and surjective, then $dim(W) = Rank(T) = dim(V)$. □

Example 6.6.1. Find the rank and nullity of the the linear transformation

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x, y) & \mapsto & (0, x). \end{array}$$

and verify the rank nullity theorem.

Solution: To find the rank and nullity, we need to find the range and kernel, respectively. By definition, range space of T is

$$\begin{aligned} \text{Range}(T) :&= \{T(x, y) \mid x, y \in \mathbb{R}\} \\ &= \{(0, x) \mid x \in \mathbb{R}\} \\ &= \{x(0, 1) \mid x \in \mathbb{R}\} \\ &= \text{Span}\{(0, 1)\}. \end{aligned}$$

Since $\{(0, 1)\}$ is a basis for Range(T), the Rank(T) = dim(Range(T)) is 1.
Similarly, by definition, kernel of T is

$$\begin{aligned} \text{Ker}(T) :&= \{(x, y) \in \mathbb{R}^2 \mid T(x, y) = (0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid (0, x) = (0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid x = 0\} \\ &= \{(0, y) \mid y \in \mathbb{R}\} \\ &= \text{Span}\{(0, 1)\}. \end{aligned}$$

Since $\{(0,1)\}$ is also a basis for Ker(T), the Nullity(T) = dim(Ker(T)) is 1. Hence we have the following equality:

$$\text{Rank}(T) + \text{Nullity}(T) = 1 + 1 = \dim(V) = \dim(\mathbb{R}^2) = 2.$$

Example 6.6.2. Consider a linear transformation $T : V \to W$ such that $\dim(V) = 5$, $\dim(W) = 7$ and T is one-to-one. Find Rank(T) and Nullity(T).

Solution: Since T is one-to-one, one concludes that $\text{Ker}(T) = \{0\}$. Therefore the Nullity(T) = dim(Ker(T)) is 0. From the rank-nullity theorem, we obtain;

$$\begin{aligned} \dim(V) &= \text{Nullity}(T) + \text{Rank}(T) = 0 + \text{Rank}(T) \\ \Rightarrow \text{Rank}(T) &= \dim(V) = 5. \end{aligned}$$

Example 6.6.3. Consider a linear transformation $T : V \to W$ such that $\dim(V) = 5$, $\dim(W) = 3$ and dim(Ker(T)) is 2. Then show that T is onto.

Solution: By rank nullity theorem, we have;

$$\begin{aligned} &\dim(V) = \text{Nullity}(T) + \text{Rank}(T) = \dim(\text{Ker}(T)) + \text{Rank}(T) = 2 + \text{Rank}(T) \\ &\Rightarrow \text{Rank}(T) = \dim(V) - 2 = 3. \end{aligned}$$

Since Rank(T) is same as the dimension of W, the range of T is same as W. Therefore T is on-to.

Example 6.6.4. Validate the statement: There is a linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ with kernel $\{0\}$.

Solution: Since image of T is a subspace of $\mathbb{R}^3$, Rank(T)= dim(Range (T)) ≤ 3. Now, by rank nullity:

$$\begin{aligned} \text{Nullity}(T) + \text{Rank}(T) &= \dim(\mathbb{R}^4) = 4 \\ \Rightarrow \text{Nullity}(T) &= 4 - \text{Rank}(T) \geq 1 \quad (\because \text{Rank}(T) \leq 3). \end{aligned}$$

Therefore T can not be one-one, and hence the given statement is false.

6.6.1 Exercises

Consider the following linear transformation:

$$\begin{gathered} T : \mathbb{R}^3 \to \mathbb{R}^2 \\ T(x, y, z) = (2x + 3z, 4y + z) \end{gathered}$$

Answer questions 1,2,3 and 4, using the information given above.

1) Which of the following matrices corresponds to the given linear transformation T with respect to the standard ordered basis for $\mathbb{R}^3$ and the standard ordered basis for $\mathbb{R}^2$?

- Option 1: $\begin{bmatrix} 2 & 3 & 0 \\ 4 & 1 & 0 \end{bmatrix}$
- Option 2: $\begin{bmatrix} 2 & 0 \\ 0 & 4 \\ 3 & 1 \end{bmatrix}$
- **Option 3:** $\begin{bmatrix} 2 & 0 & 3 \\ 0 & 4 & 1 \end{bmatrix}$
- Option 4: $\begin{bmatrix} 2 & 4 \\ 3 & 1 \\ 0 & 0 \end{bmatrix}$

2) Which of the following represents a basis of the kernel of T?

- Option 1: $\{(-\frac{3}{2}, 0, 1), (0, -\frac{1}{4}, 1)\}$.
- **Option 2:** $\{(-\frac{3}{2}, -\frac{1}{4}, 1)\}$.
- Option 3: $\{(-\frac{3}{2}, -\frac{1}{4}, 2)\}$.
- Option 4: $\{(2, 0, 3), (0, 4, 1)\}$.

3) What will be the dimension of the subspace $Im(T)$? [Answer: 2]

4) Choose the correct option.

- Option 1: T is an isomorphism.
- Option 2: T is one to one but not onto.
- **Option 3:** T is onto but not one to one.
- Option 4: T is neither one to one, nor onto.

5) Choose the set of correct options.

- **Option 1:** Nullity and rank of the identity transformation on a vector space of dimension n are 0 and n respectively.
- Option 2: Nullity and rank of the identity transformation on a vector space of dimension n are 1 and $n-1$ respectively.
- Option 3: Nullity and rank of the identity transformation on a vector space of dimension n are n and 0 respectively.
- Option 4: Nullity and rank of an isomorphism between two vector spaces V and W (both of dimension n) are n and 0 respectively.

– **Option 5:** Nullity and rank of an isomorphism between two vector spaces V and W (both of dimension n) are 0 and n respectively.
– **Option 6:** There cannot exist an isomorphism between two vector spaces whose dimensions are not the same.

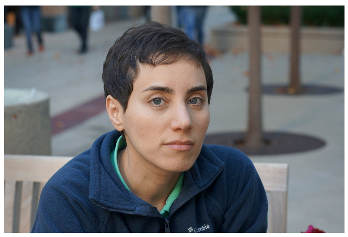
6) Choose the set of correct options.

– **Option 1:** Any injective linear transformation between any two vector spaces which have the same dimensions, must be an isomorphism.
– **Option 2:** Any surjective linear transformation between any two vector spaces which have the same dimensions, must be an isomorphism.
– **Option 3:** There does not exist any surjective linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$.
– Option 4: There does not exist any injective linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$.

7) Choose the set of correct options.

– **Option 1:** There exists a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that Image(T)= Kernel(T).
– Option 2: There exists a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that Image(T)= Kernel(T).
– Option 3: If $T : V \to V$ is a linear transformation and v_1,$v_2 \in V$ are linearly independent then $T(v_1)$, $T(v_2)$ are also linearly independent.
– **Option 4:** If $T : V \to V$ is a linear transformation and $T(v_1)$, $T(v_2)$ are linearly independent then v_1, $v_2 \in V$ are also linearly independent.

7. Equivalence and similarity of matrices



"The beauty of mathematics only shows itself to more patient followers."

— Maryam Mirzakhani

7.1 Equivalence of Matrices

Definition 7.1.1. Let A and B be two matrices of order $m \times n$. We say that A is equivalent to B if there is an invertible matrix P of order $n \times n$ and an invertible matrix Q of order $m \times m$ such that:

$$B = QAP.$$

Let $M_{m\times n}(\mathbb{R})$ denote the set of all $m\times n$ matrices with entries from $\mathbb{R}$. Equivalence of matrices on $M_{m\times n}(\mathbb{R})$ is an equivalence relation, that is, if A,B and $C \in M_{m\times n}(\mathbb{R})$, then

- A is equivalent to itself.

  Take $Q = I_{m\times m}$ to be the identity matrix of order m and $P = I_{n\times n}$ to be the identity matrix of order n. Then we can write

  $$A = I_{m\times m}\, A\, I_{n\times n}.$$

  That is, A is equivalent to itself.

- If A is equivalent to B then B is equivalent to A.

  If A is equivalent to B, then we know that there are two invertible matrices P and Q of order n and m, respectively, such that

  $$B = QAP.$$

We can rewrite the above equality as:

$$A = Q^{-1}BP^{-1}.$$

Since P and Q are invertible, P^{-1} and Q^{-1} are also invertible. Therefore B is equivalent to A.

- If A is equivalent to B and B is equivalent to C then A is equivalent to C.

  If A is equivalent to B and B is equivalent to C, then we can write

  $$B = QAP \quad \text{and} \quad C = Q'BP',$$

  where Q,Q' are invertible matrices of order m and P,P' are invertible matrices of order n. Using the above relation, we can write C as

  $$C = (Q'Q)A(PP').$$

  Note that both $Q'Q$ and PP' are invertible matrices. Therefore A is equivalent to C.

Example 7.1.1. Show that $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 7 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 2 & 7 \end{bmatrix}$ are equivalent.

Solution: Take $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, both are invertible matrices. Then we can write

$$B = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 7 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Therefore A and B are similar.

Note: If it is given that A and B are equivalent then finding P and Q can be very challenging in most of the cases. Later we will see a method to find P and Q in a particular case.

Example 7.1.2. Suppose you know that $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ are equivalent matrices. If it is given that $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then find the number possible choices for P.

Solution: Let $P = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Then using the relation $B = QAP$, we obtain

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$\implies \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$\implies \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} d+g & e+h & f+i \\ a-d & b-e & c-f \end{bmatrix}.$$

By equating each component of both the matrices, we have

$$\begin{aligned} d+g=1, \quad & a-d=1 \quad \text{(by comparing the 1st column)} \quad (1) \\ e+h=-1, \quad & b-e=2 \quad \text{(by comparing the 2nd column)} \quad (2) \\ f+i=0, \quad & c-f=1 \quad \text{(by comparing the 3rd column)} \quad (3). \end{aligned}$$

Now we can see that the above three systems are independent of each other (each system has different variables), so we can solve them independently.

From system (1), we get $d = 0, a = 1, g = 1$ as a solution.
From system (2), we get $e = 0, b = 2, h = -1$ as a solution.
From system (3), we get $f = 0, c = 1, i = 0$ as a solution.

Now you can see that $P = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}$ is invertible and satisfies the equality $B = QAP$. **Note that P is not unique (because there are infinitely many solutions of the above three systems).**

Finding the matrices P and Q in a particular case:

Consider a linear transformation $T : V \to W$. Let $\beta_1 := \{v_1, v_2, \cdots, v_n\}$ and $\beta_2 := \{u_1, u_2, \cdots, u_n\}$ be two ordered bases of V, and $\gamma_1 = \{w_1, w_2, \cdots, w_m\}$ and $\gamma_2 = \{x_1, x_2, \cdots, x_m\}$ be the two ordered bases of W.

- Let A be the matrix representation of T with respect the bases β_1 of V and γ_1 of W.

- Let B be the matrix representation of T with respect the bases β_2 of V and γ_2 of W.

Then A and B are equivalent, and satisfy the equality

$$B = QAP,$$

where P and Q are defined as follows:

- For P, express the elements of the ordered basis β_2 in terms of the ordered basis β_1, that is,

$$\begin{aligned} u_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ u_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ &\dots\dots\dots\dots\dots\dots \\ u_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n. \end{aligned}$$

The matrix P is given by $P = \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix}$.

- For Q, express the elements of the ordered basis γ_1 in terms of the ordered basis γ_2, that is,

$$\begin{aligned} w_1 &= b_{11}x_1 + b_{21}x_2 + \cdots + b_{m1}x_m \\ w_2 &= b_{12}x_1 + b_{22}x_2 + \cdots + a_{m2}x_m \\ &\dots\dots\dots\dots\dots\dots \\ w_m &= b_{1m}x_1 + b_{2m}x_2 + \cdots + b_{mm}x_m. \end{aligned}$$

The matrix Q is given by $Q = \begin{bmatrix} b_{11} & b_{12} \dots b_{1m} \\ b_{21} & b_{22} \dots b_{2m} \\ \vdots & \vdots \dots \vdots \\ b_{m1} & b_{m2} \dots b_{mm} \end{bmatrix}$.

They satisfy the relation

$$B = \begin{bmatrix} b_{11} & b_{12} \dots b_{1m} \\ b_{21} & b_{22} \dots b_{2m} \\ \vdots & \vdots \dots \vdots \\ b_{m1} & b_{m2} \dots b_{mm} \end{bmatrix} A \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix}.$$

Example 7.1.3. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1).$$

- Let A be the matrix representation of the linear transformation T with respect to the ordered bases $\beta_1 = \{(1,0,0),(0,1,0),(0,0,1)\}$ for the domain and $\gamma_1 = \{(1,0),(0,1)\}$ for the co-domain.
- Let B be the matrix representation of the linear transformation T with respect to the ordered bases $\beta_2 = \{(1,0,-1),(1,1,1),(1,0,0)\}$ for the domain and $\gamma_2 = \{(0,1),(1,0)\}$ for the co-domain.

Let Q and P be matrices such that $B = QAP$. Then find all matrices A,B,P and Q.

Solution: First we will find A and B. Since

$$\begin{aligned}
T(1,0,0) &= (1,-1) = 1(1,0) + (-1)(0,1)\\
T(0,1,0) &= (1,0) = 1(1,0) + 0(0,1)\\
T(0,0,1) &= (0,2) = 0(1,0) + 2(0,1),
\end{aligned}$$

we see that the matrix of T relative to β_1, γ_1 is

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

Similarly for the matrix B:

$$\begin{aligned}
T(1,0,-1) &= (1,-3) = (-3)(0,1) + 1(1,0)\\
T(1,1,1) &= (2,1) = 1(0,1) + 2(1,0)\\
T(1,0,0) &= (1,-1) = -1(0,1) + 1(1,0),
\end{aligned}$$

the matrix of T relative to β_2, γ_2 is

$$B = \begin{bmatrix} -3 & 1 & -1\\ 1 & 2 & 1 \end{bmatrix}.$$

Next, we compute Q and P. For P, we express the elements of the ordered basis β_2 in terms of the ordered basis β_1:

$$\begin{aligned}
(1,0,-1) &= 1(1,0,0) + 0(0,1,0) + (-1)(0,0,1)\\
(1,1,1) &= 1(1,0,0) + 1(0,1,0) + 1(0,0,1)\\
(1,0,0) &= 1(1,0,0) + 0(0,1,0) + 0(0,0,1).
\end{aligned}$$

Therefore the matrix P is

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{bmatrix}.$$

For Q, we express the elements of the ordered basis γ_1 in terms of the ordered basis γ_2:

$$(1,0) = 0(0,1) + 1(1,0)$$
$$(0,1) = 1(0,1) + 0(1,0)$$

Therefore the matrix Q is

$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Hence we have the relation:

$$B = QAP$$

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

There are a few relatively easier methods to check whether two matrices are equivalent or not.

Other Characterization of Equivalent matrices:

(1) Two matrices A and B are equivalent if A can be transformed in to B by a combination of elementary row and column operations.

(2) Two matrices A and B are equivalent if Rank(A)=Rank(B).

Example 7.1.4. Show that A= $\begin{bmatrix} 4 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$ and B= $\begin{bmatrix} 5 & 1 & -1 \\ 6 & 5 & 1 \\ 4 & 2 & 0 \end{bmatrix}$ are equivalent.

Solution:
We will use the above two methods to show that A and B are equivalent.

Method-1:

$$A = \begin{bmatrix} 4 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \xrightarrow{R_1:=R_1+R_2} \begin{bmatrix} 5 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \xrightarrow{R_2:=R_2+2R_3} \begin{bmatrix} 5 & 1 & -1 \\ -1 & 3 & 2 \\ -1 & 1 & 1 \end{bmatrix} \xrightarrow{R_3:=R_3+R_1}$$
$$\begin{bmatrix} 5 & 1 & -1 \\ -1 & 3 & 2 \\ 4 & 2 & 0 \end{bmatrix} \xrightarrow{R_2:=R_2+R_1} \begin{bmatrix} 5 & 1 & -1 \\ 4 & 4 & 1 \\ 4 & 2 & 0 \end{bmatrix} \xrightarrow{R_2:=R_2+\frac{1}{2}R_3} B = \begin{bmatrix} 5 & 1 & -1 \\ 6 & 5 & 1 \\ 4 & 2 & 0 \end{bmatrix}$$

We obtain B by performing a few elementary operations on A. Hence A and B are equivalent.

Method-2:
Here we will compute rank of A and B. The row reduce echelon form of both A and B are $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Therefore, rank$(A)$ = rank(B)=3, and hence A and B are equivalent matrices.

Example 7.1.5. Show that A= $\begin{bmatrix} 2 & 6 & -3 \\ 1 & 3 & 2 \end{bmatrix}$ and B= $\begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & -7 \end{bmatrix}$ are equivalent.

Solution:
Method-1:

$$A = \begin{bmatrix} 2 & 6 & -3 \\ 1 & 3 & 2 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & -3 \end{bmatrix} \xrightarrow{R_2 := R_2 + (-2)R_1} B = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & -7 \end{bmatrix}$$

We obtain B by performing elementary operations on A. Hence A and B are equivalent.

Method-2:
The row reduce echelon form of A and B are $\begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Therefore, rank$(A)$ = rank(B)=2, and hence A and B are equivalent matrices.

Example 7.1.6. Check whether $A = \begin{bmatrix} 2 & 0 & 0 \\ 7 & -1 & 0 \\ 6 & 8 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 2 \\ -2 & 0 & -1 \\ 1 & 3 & 5 \end{bmatrix}$ are equivalent or not.

Solution: The row reduced echelon form of the matrix A is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and row reduced echelon form of B is $\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}$. Therefore rank$(A) \neq$ rank(B), and **hence A and B are not equivalent.**

Example 7.1.7. Check whether $A = \begin{bmatrix} 2 & 0 & 0 \\ 7 & -1 & 0 \\ 6 & 8 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ -2 & 0 \end{bmatrix}$ are equivalent or not.

Solution: Two matrices of different order can not be equivalent. **Therefore A and B are not equivalent.**

Example 7.1.8. Check whether $A = \begin{bmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 0 \end{bmatrix}$ are equivalent or not.

Solution: One can check that the rank$(A) = 1$ and rank$(B) = 2$. **Therefore A and B are not equivalent.**

Question 118. If A and B are equivalent matrices of order $m \times n$, investigate whether the following are true?

(1) A^T and B^T are equivalent.

(2) A^2 and B^2 are equivalent.

(3) AB and BA are equivalent.

Solution. (1) Since A, B are equivalent, it holds that there is a $n \times n$ invertible matrix P and a $m \times m$ invertible matrix Q such that:

$$B = QAP.$$

From the above equality, we have:

$$B^T = (QAP)^T = P^T A^T Q^T \quad \left(\because (AB)^T = B^T A^T\right).$$

Since P and Q are invertible, P^T and Q^T are also invertible. **Therefore A^T and B^T are equivalent.**

(2) In general, **It is not true that A^2, B^2 are always equivalent.**

Consider two matrices $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Note that rank$(A)$ and rank(B) are same. Therefore A and B are equivalent.

But $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $B^2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. We notice that rank$(A^2) = 0$, but rank$(B^2) = 1$. **Therefore A^2 and B^2 are not equivalent.**

(3) In general, AB, BA **are not always equivalent.** Take $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Notice that

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } BA = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Clearly, rank$(AB) \neq$ rank(BA). **As a result, AB and BA are not equivalent.**

7.2 Similar Matrices

Definition 7.2.1. Two matrices A and B of order $n \times n$ are said to be similar if there exists an invertible matrix P such that

$$B = P^{-1}AP.$$

Note: We check similarity only for **square matrices** of same order. If it is given that A and B are similar then finding P can be very difficult for higher order matrices. Later we will see a method to find P in a particular case.

Question 119. Show that similarity defines an equivalence relation between square matrices of same order.

Example 7.2.1. Check whether $A = \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} -10 & 7 \\ -13 & 9 \end{bmatrix}$ are similar or not.

Solution: Suppose A and B are similar. By definition this mean we have an invertible matrix P such that

$$\begin{aligned} B &= P^{-1}AP \\ \implies AP &= PB. \end{aligned}$$

Let $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then the equality $AP = PB$ becomes

$$\begin{aligned} \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} -10 & 7 \\ -13 & 9 \end{bmatrix} \\ \implies \begin{bmatrix} a+3c & b+3d \\ -a-2c & -b-2d \end{bmatrix} &= \begin{bmatrix} -10a-13b & 7a+9b \\ -10c-13d & 7c+9d \end{bmatrix}. \end{aligned}$$

By comparing the entries of the two matrices, we have a homogeneous system:

$$\begin{aligned} 11a + 13b + 3c &= 0 \\ 7a + 8b - 3d &= 0 \\ a - 8c - 13d &= 0 \\ b + 7c + 11d &= 0. \end{aligned}$$

The augmented matrix of the above system is

$$\left[\begin{array}{cccc|c} 11 & 13 & 3 & 0 & 0 \\ 7 & 8 & 0 & -3 & 0 \\ 1 & 0 & -8 & -13 & 0 \\ 0 & 1 & 7 & 11 & 0 \end{array}\right].$$

Row reduced echelon form of the above matrix is

$$\left[\begin{array}{cccc|c} 1 & 0 & -8 & -13 & 0 \\ 0 & 1 & 7 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The new system becomes;

$$a - 8c - 13d = 0$$
$$b + 7c + 11d = 0.$$

Since c and d are free variables there are infinitely many solution of the above system. Therefore there can be infinitely possible choices for P. In particular, if $c = -1$ and $d = 1$, we get $a = 5$ and $b = -4$. As a result, we get

$$P = \begin{bmatrix} 5 & -4 \\ -1 & 1 \end{bmatrix}.$$

Note that $\det(P) \neq 0$ and P satisfies the condition

$$\begin{bmatrix} -10 & 7 \\ -13 & 9 \end{bmatrix} = \begin{bmatrix} 5 & -4 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5 & -4 \\ -1 & 1 \end{bmatrix}.$$

Therefore A and B are similar.
Note: We need to choose the values of a,b,c, and d such that P becomes an invertible matrix.

Lemma 7.2.1. *If two matrices A and B are similar then they are also equivalent.*

Proof. Since A, B are similar, there is an invertible P such that:

$$B = P^{-1}AP$$

We can rewrite the above equality as:

$$B = QAP \quad \text{where } Q = P^{-1}.$$

Therefore A and B are equivalent. □

Note that the converse of the above theorem is not true, that is, if two square matrices A and B are equivalent that does not imply that A and B are similar. We will see this with an example.

Consider the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Note that Rank(A) =Rank(B) =2. Therefore A and B are equivalent. But A **and** B **are not similar.** We will show this using method of contradiction.

Suppose A and B are similar. Then by definition, we should get a invertible matrix P such that

$$B = P\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}P^{-1}.$$

But $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ is an identity matrix, it commutes with all 2×2 real matrices, that is,

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}M = M\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \quad \text{for all } M \in M_{2\times 2}(\mathbb{R}).$$

Now,

$$\begin{aligned}B &= P\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}P^{-1}\\ &= \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}PP^{-1}\\ &= \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.\end{aligned}$$

This is a contradiction because $B \neq \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$. Therefore our assumption is wrong and hence A and B are not similar.

Example 7.2.2. Check whether $A = \begin{bmatrix}5 & 3\\3 & 1\end{bmatrix}$ and $B = \begin{bmatrix}-1 & 9\\1 & -9\end{bmatrix}$ are similar or not.

Solution: Note that rank(A) $\neq$ rank(B). Therefore A and B are not equivalent. **Since** A **and** B **are not equivalent,** A **and** B **can not be similar**.

Lemma 7.2.2. *(1) If A or B is invertible, then AB is similar to BA.*

(2) If A and B are similar, then A^n is similar to B^n.

(3) If A and B are similar, then A^T is similar to B^T.

Proof. (1) First, assume that A is invertible, hence the inverse matrix A^{-1} exists. Now we can write

$$BA = A^{-1}(AB)A = P^{-1}(AB)P, \quad \text{where } P = A.$$

Therefore AB and BA are similar.

Similarly, if B is invertible, then B^{-1} exists. Then

$$BA = B(AB)B^{-1} = P^{-1}(AB)P, \quad \text{where } P = B^{-1}.$$

Hence AB and BA are similar.

(2) If A and B are similar, then we have invertible matrix P such that

$$B = P^{-1}AP.$$

Then for a positive integer n, we have

$$\begin{aligned} B^n &= (P^{-1}AP)^n \\ &= \underbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)}_{n \text{ times}} \\ &= P^{-1}A^nP. \end{aligned}$$

Therefore A^n and B^n are similar.

(3) If A and B are similar, then we have an invertible matrix P such that

$$B = P^{-1}AP.$$

From the above equality, we have:

$$\begin{aligned} B^T &= \left(P^{-1}AP\right)^T \\ &= P^TA^T(P^{-1})^T \qquad \left(\because (AB)^T = B^TA^T\right) \\ &= P^TA^T(P^T)^{-1} \qquad \left(\because (P^{-1})^T = (P^T)^{-1}\right) \end{aligned}$$

Hence A^T and B^T are similar. □

Finding the matrix P for a particular case:

Consider a linear transformation $T : V \to V$. Let $\beta := \{v_1, v_2, \cdots, v_n\}$ and $\gamma = \{u_1, u_2, \cdots, u_n\}$ be two ordered bases of V.

- Let A be the matrix representation of T with respect the basis β for both domain and co-domain.
- Let B be the matrix representation of T with respect the basis γ for both domain and co-domain.

Then A and B are similar, and satisfy the equality

$$B = P^{-1}AP,$$

where P and P^{-1} are defined as follows:

- For P, express the elements of the ordered basis γ in terms of the ordered basis β, that is,

$$\begin{aligned} u_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ u_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ u_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n. \end{aligned}$$

  The matrix P is given by $P = \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix}$.

- For P^{-1}, express the elements of the ordered basis β in terms of the ordered basis γ (or one can compute P^{-1} directly after computing P), that is,

$$\begin{aligned} v_1 &= b_{11}u_1 + b_{21}u_2 + \cdots + b_{n1}u_n \\ v_2 &= b_{12}u_1 + b_{22}u_2 + \cdots + a_{n2}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ v_n &= b_{1n}u_1 + b_{2n}u_2 + \cdots + b_{nn}u_n. \end{aligned}$$

  The matrix P^{-1} is given by $P^{-1} = \begin{bmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2n} \\ \vdots & \vdots \dots \vdots \\ b_{n1} & b_{n2} \dots b_{nn} \end{bmatrix}$.

  They satisfy the relation

$$B = P^{-1}AP = \begin{bmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & b_{22} \dots b_{2n} \\ \vdots & \vdots \dots \vdots \\ b_{n1} & b_{n2} \dots b_{nn} \end{bmatrix} A \begin{bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \vdots \dots \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{bmatrix}.$$

Example 7.2.3. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation defined by

$$T(x_1, x_2) = (-x_2, x_1).$$

- Let A be the matrix representation of the linear transformation T with respect to the ordered basis $\beta = \{(1,0),(0,1)\}$ for both domain co-domain.
- Let B be the matrix representation of the linear transformation T with respect to the ordered basis $\gamma = \{(1,2),(1,-1)\}$ for both domain co-domain.

Let P a matrix such that $B = P^{-1}AP$. Then find all matrices A,B,P and P^{-1}.

Solution: First we will find A and B. Since

$$T(1,0) = (0,1) = 0(1,0) + 1(0,1)$$
$$T(0,1) = (-1,0) = (-1)(1,0) + 0(0,1),$$

therefore matrix of T relative to β is

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Similarly for the matrix B:

$$T(1,2) = (-2,1) = \frac{-1}{3}(1,2) + \frac{-5}{3}(1,-1)$$
$$T(1,-1) = (1,1) = \frac{2}{3}(1,2) + \frac{1}{3}(1,-1),$$

the matrix of T relative to γ is

$$B = \begin{bmatrix} \frac{-1}{3} & \frac{2}{3} \\ \frac{-5}{3} & \frac{1}{3} \end{bmatrix}.$$

Next we compute P and P^{-1}. For P, we express the elements of the ordered basis γ in terms of the ordered basis β:

$$(1,2) = 1(1,0) + 2(0,1)$$
$$(1,-1) = 1(1,0) + (-1)(0,1).$$

Therefore the matrix P is

$$P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}.$$

For P^{-1}, we express the elements of the ordered basis β in terms of the ordered basis γ:

$$(1,0) = \frac{1}{3}(1,2) + \frac{2}{3}(1,-1)$$
$$(0,1) = \frac{1}{3}(1,2) + \frac{-1}{3}(1,-1).$$

Therefore the matrix P^{-1} is

$$P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-1}{3} \end{bmatrix}.$$

Hence we have the relation:

$$B = P^{-1}AP$$
$$\begin{bmatrix} \frac{-1}{3} & \frac{2}{3} \\ \frac{-5}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-1}{3} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}.$$

Example 7.2.4. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

- Let A be the matrix representation of the linear transformation T with respect to the ordered basis $\beta = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for both domain co-domain.
- Let B be the matrix representation of the linear transformation T with respect to the ordered basis $\gamma = \{(1, 0, 1), (-1, 2, 1), (2, 1, 1)\}$ for both domain co-domain.

Let P a matrix such that $B = P^{-1}AP$. Then find all matrices A,B,P and P^{-1}.

Solution: First we will find A and B. Since

$$\begin{aligned} T(1, 0, 0) &= (3, -2, -1) = 3(1, 0, 0) + (-2)(0, 1, 0) + (-1)(0, 0, 1) \\ T(0, 1, 0) &= (0, 1, 2) = 0(1, 0, 0) + 1(0, 1, 0) + 2(0, 0, 1) \\ T(0, 0, 1) &= (1, 0, 4) = 1(1, 0, 0) + 0(0, 1, 0) + 4(0, 0, 4), \end{aligned}$$

therefore matrix of T relative to β is

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

Similarly for the matrix B:

$$\begin{aligned} T(1, 0, 1) &= (4, -2, 3) = \frac{17}{4}(1, 0, 1) + \frac{-3}{4}(-1, 2, 1) + \frac{-1}{2}(2, 1, 1) \\ T(-1, 2, 1) &= (-2, 4, 9) = \frac{35}{4}(1, 0, 1) + \frac{15}{4}(-1, 2, 1) + \frac{-7}{2}(2, 1, 1) \\ T(2, 1, 1) &= (7, -3, 4) = \frac{11}{2}(1, 0, 1) + \frac{-3}{2}(-1, 2, 1) + 0(2, 1, 1), \end{aligned}$$

the matrix of T relative to γ is

$$B = \begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ \frac{-3}{4} & \frac{15}{4} & \frac{-3}{2} \\ \frac{-1}{2} & \frac{-7}{2} & 0 \end{bmatrix}.$$

Next we compute P and P^{-1}. For P, we express the elements of the ordered basis γ in terms of the ordered basis β:

$$\begin{aligned} (1,0,1) &= 1(1,0,0) + 0(0,1,0) + 1(0,0,1) \\ (-1,2,1) &= -1(1,0,0) + 2(0,1,0) + 1(0,0,1) \\ (2,1,1) &= 2(1,0,0) + 1(0,1,0) + 1(0,0,1) \end{aligned}$$

Therefore the matrix P is

$$P = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

For P^{-1}, we express the elements of the ordered basis β in terms of the ordered basis γ:

$$\begin{aligned} (1,0,0) &= \frac{-1}{4}(1,0,1) + \frac{-1}{4}(-1,2,1) + \frac{1}{2}(2,1,1) \\ (0,1,0) &= \frac{-3}{4}(1,0,1) + \frac{1}{4}(-1,2,1) + \frac{1}{2}(2,1,1) \\ (0,0,1) &= \frac{5}{4}(1,0,1) + \frac{1}{4}(-1,2,1) + \frac{-1}{2}(2,1,1), \end{aligned}$$

Therefore the matrix P^{-1} is

$$P^{-1} = \begin{bmatrix} \frac{-1}{4} & \frac{-3}{4} & \frac{5}{4} \\ \frac{-1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{bmatrix}.$$

Hence we have the relation:

$$B = P^{-1}AP$$

$$\begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ \frac{-3}{4} & \frac{15}{4} & \frac{-3}{2} \\ \frac{-1}{2} & \frac{-7}{2} & 0 \end{bmatrix} = \begin{bmatrix} \frac{-1}{4} & \frac{-3}{4} & \frac{5}{4} \\ \frac{-1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

7.3 Properties of similar matrices

Before going into the properties of similar matrices, we will mention a few facts without their proofs.

Facts:

- If M is an invertible matrix of order n then
$$\text{Rank}(AM) = \text{Rank}(MA) = \text{Rank}(A),$$
for any arbitrary matrix of A order n.
- If A and B are two matrices of order n then
$$\text{Trace}(AB) = \text{Trace}(BA) \quad \text{and} \quad \text{Det}(AB) = \text{Det}(A)\text{Det}(B).$$

Lemma 7.3.1. *If two matrices A and B are similar, then they have the same rank.*

Proof. Let A and B be similar such that $B = P^{-1}AP$. Then
$$\begin{aligned}\text{Rank}(B) &= \text{Rank}(P^{-1}AP)\\ &= \text{Rank}(AP)\\ &= \text{Rank}(A).\end{aligned}$$
□

Lemma 7.3.2. *If two matrices A and B are similar, then they have the same trace.*

Proof. Let A and B be similar such that $B = P^{-1}AP$. Then
$$\begin{aligned}\text{Trace}(B) &= \text{Trace}(P^{-1}AP)\\ &= \text{Trace}(AP^{-1}P)\\ &= \text{Trace}(A).\end{aligned}$$
□

Lemma 7.3.3. *If two matrices A and B are similar, then they have the same determinant.*

Proof. Let A and B be similar such that $B = P^{-1}AP$. Then
$$\begin{aligned}\text{Det}(B) &= \text{Det}(P^{-1}AP)\\ &= \text{Det}(P^{-1})\text{Det}(A)\text{Det}(P)\\ &= \text{Det}(A).\end{aligned}$$
□

Note: If two matrices are similar then they have same rank, trace and determinant but the converse is not true. In particular, if two matrices A and B have same rank, trace and determinant that does not imply that they are similar.

Take $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Note that A and B have same rank, determinant and trace but they are not similar (for the explanation, check the paragraph after Lemma 7.2.1).

Example 7.3.1. Check whether $A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ are similar or not.

Solution: Trace$(A) = 5 \neq$ Trace$(B) = 4$. Since trace(A) and trace(B) are not equal, A and B are not similar.

Example 7.3.2. Check whether $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 11 \\ 0 & 4 \end{bmatrix}$ are similar or not.

Solution: Det$(A) = 6 \neq$ Det$(B) = 4$. Since det(A) and det(B) are not equal, A and B are not similar.

Example 7.3.3. Check whether $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ are similar or not.

Solution: Suppose A and B are similar. By definition this mean we have an invertible matrix P such that

$$\begin{aligned} B &= P^{-1}AP \\ \implies AP &= PB. \end{aligned}$$

Let $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then the equality $AP = PB$ becomes

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\ \implies \begin{bmatrix} a & b \\ a+c & b+d \end{bmatrix} &= \begin{bmatrix} a & a+b \\ c & c+d \end{bmatrix}. \end{aligned}$$

By comparing the entries of the two matrices, we have a homogeneous system:

$$\begin{aligned} a &= 0 \\ b - c &= 0. \end{aligned}$$

Since c and d are free variables there are infinitely many solution of the above system. Therefore there can be infinitely possible choices for P. In particular, if $c = 1$ and $d = 0$, we get $a = 0$ and $b = 1$. As a result, we get

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Note that $\det(P) \neq 0$ and P satisfies the condition

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Therefore A and B are similar.

7.4 Exercises

1) Choose the set of correct options.

   - Option 1: If a 3×3 matrix A is similar to the identity matrix of order 3, then A must be the identity matrix of order 3.
   - Option 2: If a 3×3 matrix A is similar to a diagonal matrix of order 3, then A must be a diagonal matrix of order 3.
   - Option 3: If A and B are similar matrices, then they are also equivalent matrices.
   - Option 4: If A and B are equivalent matrices, then they are also similar matrices.

   Consider two ordered bases

   $$\beta_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}, \text{ and } \beta_2 = \{(1, 0, 1), (0, 1, 1), (1, 1, 0)\}$$

   of $\mathbb{R}^3$ and consider the standard ordered basis $\gamma = \{(1, 0), (0, 1)\}$ of $\mathbb{R}^2$. Let A and B be the matrices corresponding to the linear transformation T with respect to the bases β_1 and β_2 for the domain, respectively and γ for the co-domain. Let Q and P be matrices such that $B = QAP$.

   Answer the questions 2,3, and 4 using the above information.

2) Which of the following matrices represents A?

   - Option 1: $\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

– Option 2: $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

– Option 3: $\begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}$

– Option 4: $\begin{bmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix}$

3) Which of the following matrices can be P?

– Option 1: There does not exist any matrix P which satisfies the property $B = QAP$.

– Option 2: $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}$

– Option 3: $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$

– Option 4: $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

4) Which of the following statements is true for Q?

– Option 1: Q can be the identity matrix of order 3.

– Option 2: Q can be the identity matrix of order 2.

– Option 3: Whatever the matrix P we choose, Q is always unique.

– Option 4: Q can be the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

5) Choose the set of correct options.

– Option 1: If two square matrices of the same order have the same determinants, then they must be similar to each other.

– Option 2: Similar matrices have the same rank.

– Option 3:If A and B are similar matrices, and $Ax = b$ has a unique solution, then $Bx = b$ also has a unique solution.

– Option 4: If A and B are similar matrices, and $Ax = b$ has a unique solution, then $Bx = b$ also has a unique solution.

6) Which of the following options are correct?

- Option 1: Two invertible matrices of the same order are equivalent.
- Option 2: $A = \begin{bmatrix} 1 & 7 & 6 \\ 2 & 1 & -1 \\ 5 & 2 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 & -2 \\ 1 & 1 & 4 \\ 1 & 2 & -6 \end{bmatrix}$ are equivalent.
- Option 3: $A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ are equivalent.
- Option 4: Consider two linear transformations $T_1(x, y, z) = (x + y - z, x + 4y - 2z, -2x + y + z)$ and $T_2(x, y, z) = (x + y, x - z, y + z)$. Let A and B denote the matrix representation of T_1 and T_2 with respect to the standard basis of $\mathbb{R}^3$. Then A and B are equivalent.

8. Affine subspaces and Affine Mapping

"In these days the angel of topology and the devil of abstract algebra fight for the soul of each individual mathematical domain."

— Hermann Weyl

8.1 Introduction

Definition 8.1.1. Let V be a vector space. An affine subspace of V is a subset L such that there exists $v \in V$ and a vector subspace $U \subseteq V$ such that

$$L = v + U := \{v + u \mid u \in U\}$$

Example 8.1.1. Consider the vector subspace $U = \{(x, 0) \mid x \in \mathbb{R}\}$ of the vector space $V = \mathbb{R}^2$ with usual addition and scalar multiplication. Let $v = (0, 1) \in V$. We define

$$L = v + U = \{(0, 1) + (x, 0) \mid x \in \mathbb{R}\} = \{(x, 1) \mid x \in \mathbb{R}\}$$

L is an affine space and the corresponding vector subspace is U.

The point to be noted here is the vector v is not unique in the representation of $L = v + U$. Consider a vector $v' = (2, 1)$. Then

$$v' + U = \{(2, 1) + (x, 0) \mid x \in \mathbb{R}\} = \{(x + 2, 1) \mid x \in \mathbb{R}\}$$

Observe that $\{(x + 2, 1) \mid x \in \mathbb{R}\} = \{(x, 1) \mid x \in \mathbb{R}\} = L$.

In fact in the representation $L = v + U$, v can be any vector in L. For example, consider an arbitrary vector $v'' = (a, 1) \in L$ for some $a \in \mathbb{R}$.

$$v'' + U = \{(a, 1) + (x, 0) \mid x \in \mathbb{R}\} = \{(x + a, 1) \mid x \in \mathbb{R}\}$$

Now observe that $\{(x + a, 1) \mid x \in \mathbb{R}\} = \{(x, 1) \mid x \in \mathbb{R}\} = L$.

Proposition 8.1.1. *The subspace U corresponding to an affine subspace $L = v + U$ is unique.*

Proof. Suppose the affine subspace L can be defined as $L = v+U$ and $L = v'+U'$.
Claim: $U = U'$.
Let $u \in U$, then $v + u \in L$. Hence $v + u \in v' + U'$. So there exists $u' \in U$, such that, $v + u = v' + u'$.
Again $v' = v' + 0 \in L$, as $0 \in U$. So $v' = v + u_1$ for some $u_1 \in U'$, i.e., $v' - v = u_1$ for some $u_1 \in U'$, i.e., $v' - v \in U'$. Hence we have,

$$\begin{aligned} v + u &= v' + u' \\ u &= (v' - v) + u' \end{aligned}$$

As $v' - v \in U'$ and $u' \in U'$, we have $u \in U'$. So $U \subseteq U'$.
By similar argument we have $U' \subseteq U$ and it concludes that $U = U'$. $\square$

Remark 8.1.1. Putting the discussion in the example and the theorem above, we can summarize that the subspace U corresponding to an affine subspace is unique. However, the vector v is not unique and, in fact, can be any vector in L. Affine subspaces are thus translations of a vector subspace of V. If $v = 0 \in V$, then $L = 0 + U = U$. This implies that a vector subspace is also an affine subspace.

Definition 8.1.2. Dimension of Affine subspace: We say an affine subspace is n-dimensional if the corresponding subspace U is n-dimensional.

8.1.1 Two dimensional affine subspaces:

First, let us recall all the possible subspaces of $\mathbb{R}^2$ with the usual addition and scalar multiplication.

1) Zero subspace, i.e., $\{(0, 0)\}$.

2) Any straight line passing through origin with slope m, i.e., $W_m = \{(x, mx) \mid x \in \mathbb{R}\}$.

3) The whole vector space $\mathbb{R}^2$ with usual addition and scalar multiplication.

The translations of these three types of vector subspaces are the all possible affine subspaces of $\mathbb{R}^2$.

1) Point affine subspace, i.e., $\{(a, b) + (0, 0)\} = \{(a, b)\}$, for some real numbers a and b .

2) Any straight line , i.e., for some real numbers a and b ,

$$\begin{aligned}
(a,b)+W_m &= \{(a,b)+(x,mx) \mid x\in\mathbb{R}\} && \text{for some } m\in\mathbb{R}\\
&= \{(x+a, mx+b) \mid x\in\mathbb{R}\} && \text{for some } m\in\mathbb{R}\\
&= \{(x', m(x'-a)+b) \mid x\in\mathbb{R}\} && \text{for some } m\in\mathbb{R}\\
&= \{(x', mx'+(b-ma)) \mid x\in\mathbb{R}\} && \text{for some } m\in\mathbb{R}\\
&= \{(x', mx'+c) \mid x\in\mathbb{R},\ \text{where } c=b-ma\} && \text{for some } m\in\mathbb{R}
\end{aligned}$$

3) The whole vector space $\mathbb{R}^2$ with usual addition and scalar multiplication.

8.1.2 Three dimensional affine subspaces:

First, let us recall all the possible subspaces of $\mathbb{R}^3$ with the usual addition and scalar multiplication.

1) Zero subspace, i.e., $\{(0,0,0)\}$.

2) Any straight line passing through origin, i.e., $Span\{v_1\}$ for some nonzero vector $v_1\in\mathbb{R}^3$.

3) Any plane passing through origin, i.e., $Span\{v_1, v_2\}$ for some nonzero vectors $v_1, v_2\in\mathbb{R}^3$, where the set $\{v_1, v_2\}$ is linearly independent.

4) The whole vector space $\mathbb{R}^3$ with usual addition and scalar multiplication.

The translates of these four types of vector subspaces are the possible affine subspaces of $\mathbb{R}^3$.

1) Point affine subspace, i.e., $\{(a,b,c)+(0,0,0)\} = \{(a,b,c)\}$, for some real numbers a, b and c .

2) Any straight line, i.e., $v+Span\{v_1\}$ for some nonzero vector $v_1\in\mathbb{R}^3$ and some vector (may be zero) $v\in\mathbb{R}^3$.

3) Any plane, i.e., $v+Span\{v_1, v_2\}$ for some nonzero vectors $v_1, v_2\in\mathbb{R}^3$, where the set $\{v_1, v_2\}$ is linearly independent and some vector (may be zero) $v\in\mathbb{R}^3$.

4) The whole vector space $\mathbb{R}^3$ with usual addition and scalar multiplication.

8.1.3 Visual representation

- Click on the following link to get a visual representation of Affine subspaces of $\mathbb{R}^2$. In the diagram, we get different straight lines for different values of a, b and m. Once you fix the value of a and b, then corresponding to the different values of m, we will get different affine subspaces, and the corresponding vector subspaces for all those affine subspaces are $\{(x, y) \mid ax + by = 0, \text{ where } x, y \in \mathbb{R}\}$.

- Click on the following link to get a visual representation of Affine subspaces of $\mathbb{R}^3$. In the diagram, we get different planes for different values of a, b, c and m. Once you fix the value of a, b and c, then corresponding to the different values of m, we will get different affine subspaces, and the corresponding vector subspaces for all those affine subspaces are $\{(x, y, z) \mid ax + by + cz = 0, \text{ where } x, y, z \in \mathbb{R}\}$.

8.1.4 Addition and scalar multiplication on affine subspaces

Let $L = v + U$ be an affine subspace. Let l and l' be two arbitrary elements of L. Hence $l = v + u$ and $l' = v + u'$ for some vectors u and u' in the vector subspace U. We define

$$l + l' = (v + u) + (v + u') = v + (u + u')$$

and

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

Example 8.1.2. Consider the affine subspace L of $\mathbb{R}^2$ defined as

$$(0, 1) + U = \{(x, 1) \mid x \in \mathbb{R}\}$$

where $U = \{(x, 0) \mid x \in \mathbb{R}\}$.

- **Addition:** Let $l_1 = (0, 1) + (x_1, 0) = (x_1, 1)$ and $l_2 = (0, 1) + (x_2, 0) = (x_2, 1)$.

$$l_1 + l_2 = ((0, 1) + (x_1, 0)) + ((0, 1) + (x_2, 0)) = (0, 1) + (x_1 + x_2, 0)$$

  i.e.,

$$l_1 + l_2 = (x_1, 1) + (x_2, 1) = (x_1 + x_2, 1)$$

- **Scalar multiplication:** Let $l = (0, 1) + (x, 0) = (x, 1)$.

$$cl = c((0, 1) + (x, 0)) = (0, 1) + c(x, 0) = (0, 1) + (cx, 0)$$

  i.e.,

$$cl = c(x, 1) = (cx, 1)$$

Example 8.1.3. Consider the affine subspace L of $\mathbb{R}^3$ defined as

$$\begin{aligned}
L &= \{(x, y, z) \mid x + 2y - z = 5, \text{ where } x, y, z \in \mathbb{R}\} \\
&= \{(x, y, z) \mid x + 2y - 5 = z, \text{ where } x, y, z \in \mathbb{R}\} \\
&= \{(x, y, x + 2y - 5) \mid \text{ where } x, y \in \mathbb{R}\} \\
&= \{(0, 0, -5) + (x, y, x + 2y) \mid \text{ where } x, y \in \mathbb{R}\} \\
&= (0, 0, -5) + \{(x, y, x + 2y) \mid \text{ where } x, y \in \mathbb{R}\} \\
&= (0, 0, -5) + \{(x, y, z) \mid x + 2y = z, \text{ where } x, y, z \in \mathbb{R}\} \\
&= (0, 0, -5) + \{(x, y, z) \mid x + 2y - z = 0, \text{ where } x, y, z \in \mathbb{R}\} \\
&= (0, 0, -5) + U
\end{aligned}$$

Where $U = \{(x, y, z) \mid x + 2y - z = 0, \text{ where } x, y, z \in \mathbb{R}\}$

- **Addition:** Let $l_1, l_2 \in L$.

$$l_1 = (x_1, y_1, z_1) = (x_1, y_1, x_1 + 2y_1 - 5) = (0, 0, -5) + (x_1, y_1, x_1 + 2y_1)$$

and

$$l_2 = (x_2, y_2, z_2) = (x_2, y_2, x_2 + 2y_2 - 5) = (0, 0, -5) + (x_2, y_2, x_2 + 2y_2)$$

$$\begin{aligned}
l_1 + l_2 &= (x_1, y_1, z_1) + (x_2, y_2, z_2) \\
&= (x_1, y_1, x_1 + 2y_1 - 5) + (x_2, y_2, x_2 + 2y_2 - 5) \\
&= ((0, 0, -5) + (x_1, y_1, x_1 + 2y_1)) + ((0, 0, -5) + (x_2, y_2, x_2 + 2y_2)) \\
&= (0, 0, -5) + (x_1 + x_2, y_1 + y_2, x_1 + 2y_1 + x_2 + 2y_2) \\
&= (0, 0, -5) + (x_1 + x_2, y_1 + y_2, x_1 + x_2 + 2(y_1 + y_2)) \\
&= (x_1 + x_2, y_1 + y_2, x_1 + x_2 + 2(y_1 + y_2) - 5)
\end{aligned}$$

- **Scalar multiplication:** Let $l \in L$.

$$l = (x, y, z) = (x, y, x + 2y - 5) = (0, 0, -5) + (x, y, x + 2y)$$

$$\begin{aligned}
cl &= c(x, y, z) \\
&= c(x, y, x + 2y - 5) \\
&= c((0, 0, -5) + (x, y, x + 2y)) \\
&= (0, 0, -5) + c(x, y, x + 2y) \\
&= (0, 0, -5) + (cx, cy, c(x + 2y)) \\
&= (cx, cy, c(x + 2y) - 5)
\end{aligned}$$

8.2 The solution set to a system of linear equations

Let $Ax = b$ be a system of linear equations.

Case 1: $b = 0$: In this case, it is a homogeneous system, and as seen before, the solution set is a subspace of $\mathbb{R}^n$, namely the null space of A.

Case 2: $b \notin$ column space A: In this case, $Ax = b$ does not have a solution, so the solution set is the empty set.

Case 3: $b \in$ column space A: Let v be a solution of the equation $Ax = b$ and $u \in$ nullspace(A). Hence we have,

$$A(v + u) = Av + Au = b + 0 = b.$$

So $v + u$ is also a solution of $Ax = b$. Hence, in this case, the solution set L is an affine subspace of $\mathbb{R}^n$. Specifically, it can be described as

$$L = v + \text{nullspace}(A),$$

where v is any solution of the equation $Ax = b$.

8.3 Affine mappings of affine subspaces

Let L and L' be affine subspaces of V and W respectively. Let $f : L \to L'$ be a function. Consider any vector $v \in L$ and the unique subspace $U \subseteq V$ such that $L = v + U$. Note that $f(v) \in L'$ and hence $L' = f(v) + U'$ where U' is the unique subspace of W corresponding to L'. Then f is an affine mapping from L to L' if the function $g : U \to U'$ defined by $g(u) = f(v + u) - f(v)$ is a linear transformation.

Let $l_1 = v + u_1$ and $l_2 = v + u_2$ be in L.

$$\begin{aligned}
f(l_1 + l_2) &= f((v + u_1) + (v + u_2)) \\
&= f(v + (u_1 + u_2)) \\
&= f(v + u_1 + u_2) - f(v) + f(v) \\
&= g(u_1 + u_2) + f(v) && \text{using the definition of } g \\
&= g(u_1) + g(u_2) + f(v) && \text{since } g \text{ is a linear transformation} \\
&= f(v + u_1) - f(v) + f(v + u_2) - f(v) + f(v) && \text{using the definition of } g \\
&= f(l_1) + f(l_2) - f(v)
\end{aligned}$$

Hence we have $f(l_1) + f(l_2) = f(v) + f(l_1 + l_2)$.

Let $l = v + u$ be in L.

$$\begin{aligned}
f(cl) &= f(c(v+u)) \\
&= f(v+cu) \\
&= f(v+cu) - f(v) + f(v) \\
&= g(cu) + f(v) && \text{using the definition of } g \\
&= cg(u) + f(v) && \text{since } g \text{ is a linear transformation} \\
&= c(f(v+u) - f(v)) + f(v) && \text{using the definition of } g \\
&= c(f(l) - f(v)) + f(v) \\
&= cf(l) + (1-c)f(v)
\end{aligned}$$

Hence we have $cf(l) = (c-1)f(v) + f(cl)$.

8.3.1 Affine mapping Corresponding to a linear transformation

For a linear transformation $T : U \to U'$ and fixed vectors $v \in L$ and $v' \in L'$, an affine mapping f can be obtained by defining $f(v+u) = v' + T(u)$, and in fact every affine mapping is obtained in this way.

Remark 8.3.1. Let f' be an arbitrary affine mapping from $L = v+U$ to $L' = v'+U'$. As f is an affine mapping then there exist a linear transformation $g : U \to U'$ defined by $g(u) = f(v+u) - f(v)$. We define $f(v) = v'$. Hence the affine map $f : L \to L'$ can be obtained by,

$$f(v+u) = v' + g(u)$$

where $g : U \to U'$ is a linear transformation. This concludes the fact that every affine mapping is obtained in the way mentioned above.

Example 8.3.1. Let $L = \{(x,1) \mid x \in \mathbb{R}\}$ and $L' = \{(x,y,5) \mid x,y \in \mathbb{R}\}$. The checking that, L' is an affine subspace of $\mathbb{R}^3$ is left to reader.

$$L = (0,1) + \{(x,0) \mid x \in \mathbb{R}\} = (0,1) + U$$

$$L' = (0,0,5) + \{(x,y,0) \mid x,y \in \mathbb{R}\} = (0,0,5) + U'$$

where $U = \{(x,0) \mid x \in \mathbb{R}\}$ and $U' = \{(x,y,0) \mid x,y \in \mathbb{R}\} = (0,0,5) + U'$
Hence according to the notation used in the definition of affine map $v = (0,1)$ and We define a map $f : L \to L'$ as $f(x,1) = (x,2x,5)$. Now we can define $g : U \to U'$ as $g(u) = f(v+u) - f(v)$, i.e., $g(x,0) = f(x,1) - f(0,1) = (x,2x,5) - (0,0,5) = (x,2x,0)$. Clearly, g is a linear transformation from U to U', and hence f is an affine mapping from L to L'.

Example 8.3.2. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be a map defined by, $f(x,y,z) = (2x+3y+2, 4x-5y+3)$. We define $T : \mathbb{R}^3 \to \mathbb{R}^2$ as $T(x,y,z) = (2x+3y, 4x-5y)$. Clearly T is a linear transformation and $f(x,y,z) = (2,3) + T(x,y,z)$. Hence f is an affine mapping from $\mathbb{R}^3$ to $\mathbb{R}^2$.

8.3.2 Exercises

1) Let L be an affine subspace of $\mathbb{R}^3$ defined as $L = \{(x, y, z) \mid x + y + 2z = 6\}$, then which of the following subspaces of $\mathbb{R}^3$ corresponds to the affine subspace L?

   – Option 1: $\{(x, y, z) \mid x + y + z = 5\}$
   – **Option 2:** $\{(x, y, z) \mid x + y + 2z = 0\}$
   – Option 3: $\{(x, y, z) \mid x + 2y + z = 0\}$
   – Option 4: $\{(x, y, z) \mid 2x + y + z = 0\}$

2) Which of the following subsets of $\mathbb{R}^2$ represent an affine subspace of $\mathbb{R}^2$?

   – **Option 1:** $\{(x, y) \mid x^2 + y^2 = 0\}$
   – Option 2: $\{(x, y) \mid x^2 + y^2 = 4\}$
   – Option 3: $\{(x, y) \mid x^5 + 1 = y\}$
   – **Option 4:** $\{(x, y) \mid 3x + y = 1\}$
   – **Option 5:** $\{(x, y) \mid x = 1\}$
   – Option 6: $\{(x, y) \mid y = 5x^2\}$
   – Option 7: $\{(x, y) \mid y = x^3 + 2\}$
   – **Option 8:** $\{(1,2)\}$

3) Let U be a subspace of the vector space $\mathbb{R}^3$ and a basis of U is given by $\{(0, 1, -3), (1, 0, -1)\}$. Then which of the following subsets of $\mathbb{R}^3$ is an affine subspace of $\mathbb{R}^3$ such that the corresponding vector subspace is U?

   – **Option 1:** $L = \{(x, y, z) \mid x + 3y + z = 3\}$
   – Option 2: $L = \{(x, y, z) \mid x + 3y + 2z = 0\}$
   – **Option 3:** $L = \{(x, y, z) \mid x + 3y + z = 0\}$
   – **Option 4:** $L = \{(x, y, z) \mid x + 3y + z = 8\}$

4) Consider a system of linear equations

$$\begin{aligned} -x + y - z &= 1 \\ x - y + z &= -1 \\ x + z &= 0 \end{aligned}$$

   Which of the following is an affine subspace L of $\mathbb{R}^3$ such that if $v \in L$, then v is a solution of the system of linear equations?

- Option 1: $\{(t, -1, t) \mid t \in \mathbb{R}\}$
- **Option 2:** $\{(-t, 1, t) \mid t \in \mathbb{R}\}$
- Option 3: $\{(-t, 1, -t) \mid t \in \mathbb{R}\}$
- Option 4: $\{(-t, -1, t) \mid t \in \mathbb{R}\}$

5) Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is an affine mapping such that $f(2, 3) = (3, -1)$,$f(2, 1) = (1, 2)$ and $f(1, 0) = (0, 1)$. If $f(-2, 5) = (a, b)$, then value of $a + b$ is [Ans: -9].

INDIAN INSTITUTE OF TECHNOLOGY MADRAS
सिद्धिर्भवति कर्मजा

9. Inner product space

"Mathematics knows no races or geographical boundaries; for mathematics, the cultural world is one country."

— David Hilbert

In previous chapter we have studied the notion of the vector space. We have also know that $\mathbb{R}^n$, for any finite $n \in \mathbb{N}$ forms a vector space usual addition and scalar multiplication for a fixed n. In this section we are going to impose some more condition (an inner product) on a vector space which forms another space, we will call it as inner product space.

We have already know that $\mathbb{R}^2$ or $\mathbb{R}^3$ forms a vector space with usual addition and scalar multiplication. In mathematics for data science -1, we know how to find distance between two points or distance of a point from the origin. We are going to study similar kind to things in a inner product space.

Before to study inner product, lets see some notion:

9.1 The dot product of two vectors in Euclidean space of dimension 2

Let $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$ be two vectors form the vector space $\mathbb{R}^2$, the dot product of these two vectors is the scalar computed as follows:

$$(x_1, y_1) \cdot (x_2, y_2) = x_1x_2 + y_1y_2$$

And dot product sign between two vectors is denoted by " $\cdot$ ".

Example 9.1.1. Find the dot product of two vectors $(2, 4)$ and $(3, 5)$ from the vector space $\mathbb{R}^2$.

$$(2,4)\dot{(}3,5) = 2 \times 3 + 4 \times 5 = 6 + 20 = 26$$

9.1.1 The length of a vector in Euclidean space of dimension 2

Let $(3,4)$ be a vector in $\mathbb{R}^2$

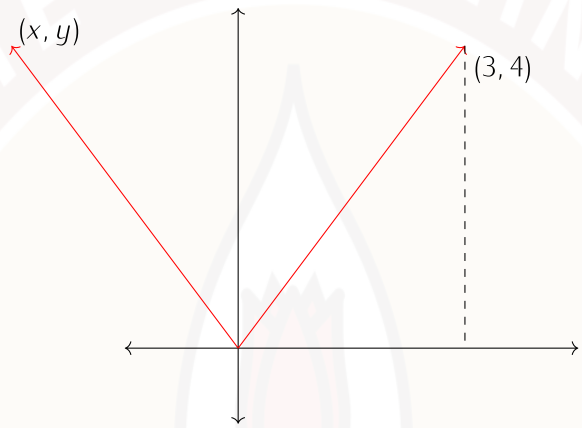


Figure 9.1:

Using Pythagoras' theorem the length of the vector (3,4) is $\sqrt{3^2+4^2} = 5$ unit

9.1.2 The relation between length and dot product in Euclidean space of dimension 2

Let $(3,4)$ be a vector in $\mathbb{R}^2$ and as we found the length of the vector is $\sqrt{3^2+4^2} = 5$ unit, this can be get using the dot product as follows:

Length of the vector $(3,4) = \sqrt{(3,4)\cdot(3,4)} = \sqrt{3^2+4^2} = 5$

More generally, the length of the vector $(x,y) \in \mathbb{R}^2$ is $\sqrt{x^2+y^2} = \sqrt{(x,y)\cdot(x,y)}$

9.1.3 The dot product and the angle between two vectors in Euclidean space of dimension 2

The angle between the vectors u and v measures how far the direction is of v from u (or vice versa). e.g. θ is the angle between $u = (3, 4)$ and $v = (1, 5)$ and it is measured in degree (between 0 and 360) or radians (between 0 and 2π).

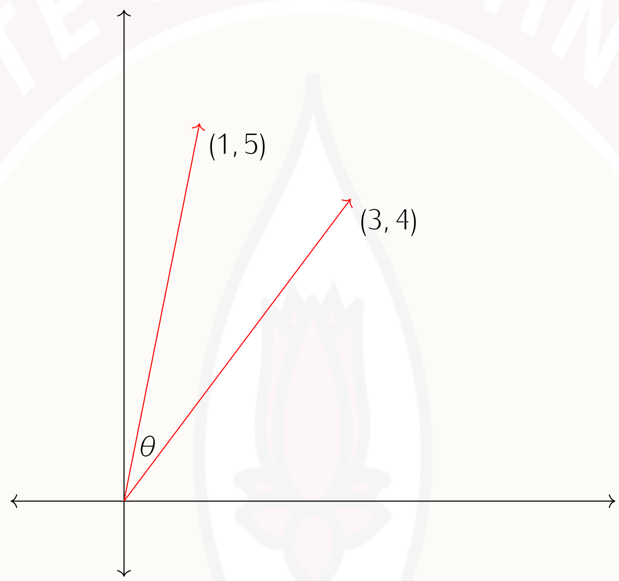


Figure 9.2:

Let u and v be two vectors in $\mathbb{R}^2$. Then we can compute the angle θ between the vectors $u = (u_1, u_2)$ and $v = (v_1, v_2)$ using the dot products as:

$$\cos\theta = \frac{u \cdot v}{\sqrt{(v \cdot v) \times (u \cdot u)}} = \frac{u_1 v_1 + u_2 v_2}{\sqrt{(v_1^2 + v_2^2) \times (u_1^2 + u_2^2)}}$$

i.e.

$$\theta = \cos^{-1} \frac{u \cdot v}{\sqrt{(v \cdot v) \times (u \cdot u)}} = \cos^{-1} \frac{u_1 v_1 + u_2 v_2}{\sqrt{(v_1^2 + v_2^2) \times (u_1^2 + u_2^2)}}$$

9.2 The dot product of two vectors in Euclidean space of dimension 3

Consider the vectors (1, 2, 3) and (2, 0, 1) in $\mathbb{R}^3$. The dot product of these two vectors gives us a scalar as follows:

$$(1,2,3)\cdot(2,0,1) = 1\times 2 + 2\times 0 + 3\times 1 = 2+0+3 = 5$$

For two general vectors (x_1, y_1, z_1) and (x_2, y_2, z_2) in $\mathbb{R}^3$, the dot product of these two vectors is the scalar computed as follows:

$$(x_1, y_1, z_1)\cdot(x_2, y_2, z_2) = x_1x_2 + y_1y_2 + z_1z_2$$

9.2.1 Length of a vector in Euclidean space of dimension 3

Let us find the length of the vector (4, 3, 3) in $\mathbb{R}^3$.

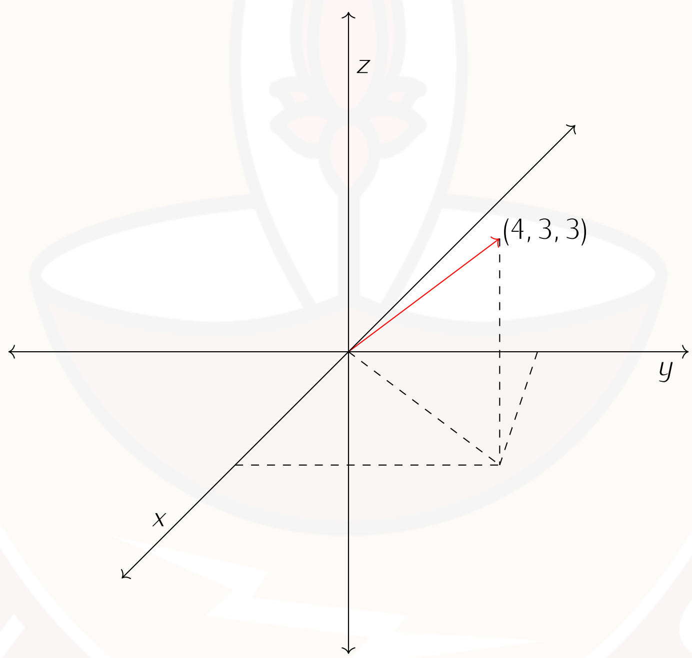


Figure 9.3:

By using the Pythagoras' theorem, the length of (4, 3, 3) is $\sqrt{4^2+3^2+3^2} = \sqrt{34}$ unit.

9.2.2 The length and dot product in Euclidean space of dimension 3

As we have done in $\mathbb{R}^2$, similarly, we can you do in $\mathbb{R}^3$ so dot product $(4, 3, 3) \cdot (4, 3, 3) = 4^2 + 3^3 + 3^2$ and hence the length of (4, 3, 3) can be expressed as the square root of dot product of the vector with itself.

$$\text{Length of the vector } (4, 3, 3) = \sqrt{(4, 3, 3) \cdot (4, 3, 3)} = \sqrt{4^2 + 3^3 + 3^2} = \sqrt{34}$$

Observe that $(4, 3, 3) \cdot (4, 3, 3) = 4^2 + 3^2 + 3^2$ and hence the lenght of (4,3,3) can be expressed as the square root of dot product of the vector with self.

$$\text{length of the vector } (4, 3, 3) = \sqrt{(4, 3, 3) \cdot (4, 3, 3)} = \sqrt{4^2 + 3^2 + 3^2} = \sqrt{34}$$

More generally, the length of the vector $(x, y, z) \in \mathbb{R}^3$ is $\sqrt{x^2 + y^2 + z^2} = \sqrt{(x, y, z) \cdot (x, y, z)}$

9.2.3 The angle between two vectors in the Euclidean space of dimension 3 and the dot product

The angle between the vectors u and v in $\mathbb{R}^3$ is the angle between them computed by passing a plane through them.

It measures how far the direction is of v from u(or vice versa) on that plane.

Let u and v be twp vectors in $\mathbb{R}^3$. Then we can compute the angle θ between the vectors u and v using the dot product as:

$$\cos\theta = \frac{u \cdot v}{\sqrt{(u \cdot c) \times (v \cdot v)}}$$

$$\theta = \cos^{-1}\left(\frac{u \cdot v}{\sqrt{(u \cdot c) \times (v \cdot v)}}\right)$$

Example 9.2.1. Find the angle between the vectors (1, 0,0) and (1, 0 ,1).

First find the dot product

$$(1, 0, 0) \cdot (1, 0, 1) = 1$$

$$(1, 0, 0) \cdot (1, 00) = 1$$

$$(1, 0, 1) \cdot (1, 0, 1) = 2$$

$$\theta = \cos^{-1} \frac{1}{\sqrt{1}\sqrt{2}} = \cos^{-1} \frac{1}{\sqrt{2}} = \frac{\pi}{4} \text{ radian or } 45^\circ$$

9.3 Dot product in Euclidean space of dimension n: length and angle

- Some definitions on the vector space $\mathbb{R}^n$, where $n \in \mathbb{N}$

  i) **Dot product:** Let $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ then the dot product of two vectors of $\mathbb{R}^n$ defined as $u \cdot v = u_1v_1 + u_2v_2 + \ldots + u_nv_n$ which is a scalar.

  ii) **Length of a vector:** Let $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$, then the length of the vector u is $||u||= \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$.

  iii) **Relation between the dot product and length of a vector:** $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$, then the length of the vector $||u||= \sqrt{u \cdot u} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$.

  iv) **The angle between two vectors in $\mathbb{R}^n$:** Let $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ then the angle θ between two vectors of $\mathbb{R}^n$ is given by $\theta = \cos^{-1}\left(\frac{u \cdot v}{\sqrt{(u \cdot u)(v \cdot v)}}\right) = \cos^{-1}\left(\frac{u_1v_1+u_2v_2+\ldots+u_nv_n}{\sqrt{(u_1^2+u_2^2+\ldots+u_n^2)(v_1^2+v_2^2+\ldots+v_n^2)}}\right)$

<u>Exercise:</u>

Question 120. $u = (1, 0, 1)$ and $v = (0, 1, 1)$ be two vectors from the vector space $\mathbb{R}^3$. Then the angle between u and v is

- $\frac{\pi}{2}$
- π
- $\frac{1}{2}$
- $\frac{\pi}{3}$

Question 121. Consider vector u, v, w and s from $\mathbb{R}^3$ such that the dot products $u \cdot v = 1, u \cdot w = -1$ and $u \cdot s = 0$, where $v = (1, 2, 3), w = (1, 1, -1)$ and $s = (1, 0, 1)$. If $u = (a, b, c)$ them find the value of $a + b + c$

Question 122. If $||a||= 3$, $||b||= 2\sqrt{2}$ and angle between a and b is $45°$, then the value of $a.b$ is [Ans: 6]

Question 123. Given that $||a|| = 3$ and $\left\|b\right\| = 5$ and $a \cdot b = 7.5$, for two vectors a and b in $\mathbb{R}^n$. Find the angle (in degrees) between the two vectors a, b. Note: $\cos 60° = 1/2, \cos 30° = \sqrt{3}/2$.

- Option 1: 30°
- **Option 2:** 60°
- Option 3: 120°
- Option 4: 150°

Question 124. Consider two vectors $a = (3, 4, -1), b = (2, -1, 2)$ in $\mathbb{R}^3$. Choose the correct options.

- **Option 1:** Length/Norm of a is $\sqrt{26}$ and that of b is 3.
- Option 2: Length/Norm of a is 5 and that of b is 3.
- Option 3: Length/Norm of a is $\sqrt{24}$ and that of b is $\sqrt{7}$.
- **Option 4:** Angle between the two vectors is 90 degrees.
- Option 5: Angle between the two vectors is approximately 0 degrees.

Question 125. If the dot product of two vectors is zero, then choose the correct option.

-
- **Option 1:** The two vectors are perpendicular to each other.
- Option 2: The two vectors are parallel to each other.
- Option 3: The two vectors are exactly the same.
- Option 4: The length of the two vectors must be the same.

Question 126. If $a = (6, -1, 3), b = (4, c, 2)$ where $a, b \in \mathbb{R}^3$ and a and b are perpendicular to each other, then find the value of c? (Answer: 30)

Question 127. Consider two vectors $a = (1, 2)$, $b = (2, 2)$ and θ is the angle between them. The sum of the two vectors is given by $c = a + b$. Choose the correct options.

- **Option 1:** Length/Norm of c is 5.
- Option 2: Length/Norm of c is 25.
- Option 3: Length/Norm of c is 4.
- **Option 4:** $\cos\theta = \frac{3}{\sqrt{10}}$.
- Option 5: $\cos\theta = \frac{3}{\sqrt{5}}$.

Question 128. Consider 3 vectors a, b, c in $\mathbb{R}^3$ and a scalar λ in $\mathbb{R}$. Choose the set of correct options.

Note: (.) represents the dot product.

-
- **Option 1:** $\lambda(a \cdot b) = (\lambda a) \cdot b$
- **Option 2:** $\lambda(a \cdot b) = (\lambda b) \cdot a$
- Option 3: $\lambda(a \cdot b) = (\lambda a) \cdot (\lambda b)$
- **Option 4:** $(a + c) \cdot b = a \cdot b + c \cdot b$
- **Option 5:** $(a + c) \cdot b = a \cdot c + c \cdot b$
- **Option 6:** $a \cdot a = 0, b \cdot b = 0$ if and only if a, b are null vectors.

Question 129. Consider two vectors $a = (1, 3, 5, 7, 9)$ and $b = (2, 4, 6, 8, 10)$. Choose the set of correct options.

- **Option 1:** The length of b is more than a.
- Option 2: The length of a is more than b.
- **Option 3:** The length of $(a - b)$ is $\sqrt{5}$.
- Option 4: The length of $(a - b)$ is $\sqrt{50}$.

9.4 Inner product on a vector space

As we have seen in the above sections some geometrical concepts like angle between two vectors, length of a vector e.t.c. in vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ and also with the help of dot product we connect this geometrical concepts algebraically i.e. numerically we are trying to find angle between two vectors, length of a vector e.t.c. . We can extend these concepts in more general form i.e., we can extend these concepts in different vector spaces, in fact, in all vector spaces using inner product.

We are already aware of with the concept functions on the vector spaces in previous sections like linear transformation. Inner product is also a function but is defined on the Cartesian product of the same vector space which having some more properties. Lets see the definition itself:

Definition 9.4.1. Inner product: An inner product on a vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying the following conditions:

Let $u, v, w \in V$,

i) $\langle v, v \rangle > 0$ for all $v \in V \setminus \{0\}$; $\langle v, v \rangle = 0$ if and only if $v = 0$.

ii) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.

iii) $\langle u, v \rangle = \langle v, u \rangle$.

iv) $\langle cu, v \rangle = c\langle u, v \rangle$

Note: A vector space together with an inner product $\langle \cdot, \cdot \rangle$ is called an inner product space.

Example 9.4.1. The dot product is an example of an inner product

Recall the dot product in the vector space $\mathbb{R}^n$, let $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ then the dot product of two vectors of $\mathbb{R}^n$ defined as

$$u \cdot v = u_1v_1 + u_2v_2 + \ldots + u_nv_n$$

which is a scalar.

We can think this dot product as a function on $\langle ., . \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined as

$$\langle u, v \rangle = u \cdot v$$

Lets verify the conditions of the inner product.

i) $\langle v, v\rangle > 0$ for all $v \in V \setminus \{0\}$; $\langle v, v\rangle = 0$ if and only if $v = 0$.

First assume $v \neq 0$,
$\langle v, v\rangle = v \cdot v = v_1v_1 + v_2v_2 + \ldots + v_nv_n = v_1^2 + v_2^2 + \ldots + v_n^2$
Since each term in $v_1^2 + v_2^2 + \ldots + v_n^2$ is square of some non zero real number so it will always grater than 0. Hence $\langle v, v\rangle > 0$

Now assume v is any vector in $\mathbb{R}^n$ such that

$$\begin{aligned} \langle v, v\rangle &= 0 \\ \implies v_1^2 + v_2^2 + \ldots + v_n^2 &= 0 \\ \implies v_1 = v_2 = \ldots = v_n &= 0 \\ \implies v &= (0, \ldots, 0) \end{aligned}$$

i.e. v is zero vector in $\mathbb{R}^n$.

Conversely, if v is zero vector in $\mathbb{R}^n$ i.e., $v = (0, \ldots, 0)$, then $\langle (0, \ldots, 0), (0, \ldots, 0)\rangle = 0 + 0 + \ldots + 0 = 0$

First condition is satisfied.

ii) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.
let $u = (u_1, u_2, \ldots, u_n), v = (v_1, v_2, \ldots, v_n), w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$

$\langle u + v, w\rangle$
$= (u + v) \cdot w$
$= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \cdot (w_1, w_2, \ldots, w_n)$
$= (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \ldots + (u_n + v_n)w_n$
$=u_1w_1 + v_1w_1 + u_2w_2 + v_2w_2 + \ldots + u_nw_n + v_nw_n$
$= (u_1w_1 + u_2w_2 + \ldots + u_nw_n) + (v_1w_1 + v_2w_2 + \ldots + v_nw_n)$
$= \langle u, w\rangle + \langle v, w\rangle$

Second condition is satisfied.

iii) $\langle u, v\rangle = \langle v, u\rangle$.

$\langle u, v\rangle = u \cdot v = v \cdot u = \langle v, u\rangle$

Third condition is satisfied.

iv) $\langle cu, v\rangle = c\langle u, v\rangle$

$$\langle cu, v\rangle = cu \cdot v = c\langle u, v\rangle$$

Fourth condition is satisfied.

So dot product is an inner product on the vector space $\mathbb{R}^n$.

Note: This inner product is also called Standard Inner Product on the on $\mathbb{R}^n$.

Example 9.4.2. Consider the following function

$$\langle ., .\rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

such that

$$\langle u, v\rangle = u_1v_1 - (u_1v_2 + u_2v_1) + 2u_2v_2,$$

where $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are the vectors of $\mathbb{R}^2$.

Observe that

$$\langle u, v\rangle = u_1v_1 - (u_1v_2 + u_2v_1) + 2u_2v_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Now we can easily check (as we checked for the dot product) that conditions of the inner product follow by the given function. So the given function is an inner product on the vector $\mathbb{R}^2$. Hence $\mathbb{R}^2$ is an inner product space with respect to the given inner product.

9.5 Norm on a vector space

In coordinate system we are very much familiar with the term distance between two points. Norm is also a function on a vector space which is usually used to find the length of vector in the vector space i.e., distance of the vector form the origin.

Definition 9.5.1. A norm on vector space V is a function

$$||\cdot||: V \to \mathbb{R}$$

$$x \to x$$

satisfying the following conditions:

i) $||x + y|| \leq ||x|| + ||y||$ for all $x, y \in V$.

ii) $||cx|| = |c|||x||$ for all $c \in \mathbb{R}$ and for all $x \in V$.

iii) $||x|| \geq 0$ for all $x \in V$; $||x||=0$ if and only if $x = 0$.

Example 9.5.1. Length is an example of a norm on the vector space $\mathbb{R}^n$

Let $u = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$, then the length of the vector u is

$$||u|| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}.$$

Now, we will check that length function $\mathbb{R}^n \to \mathbb{R}$ is a actually a norm on $\mathbb{R}^n$.

i) $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in \mathbb{R}^n$.

We have $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$
So $u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$

Therefore,
$||u + v||$
$= \sqrt{(u_1 + v_1)^2 + (u_2 + v_2)^2 + \ldots + (u_n + v_n)^2}$

$= \sqrt{(u_1^2 + v_1^2 + 2u_1v_1) + (u_2^2 + v_2^2 + 2u_2v_2) + \ldots + (u_n^2 + v_n^2 + 2u_nv_n)}$

$= \sqrt{(u_1^2 + u_2^2 + \ldots + u_n^2) + (v_1^2 + v_2^2 + \ldots + v_n^2) + 2(u_1v_1 + u_2v_2 + \ldots + u_nv_n)}$

$\leq \sqrt{(u_1^2 + u_2^2 + \ldots + u_n^2)} + \sqrt{(v_1^2 + v_2^2 + \ldots + v_n^2)}$ (Using the Cauchy-Schwartz inequality)
$= ||u|| + ||v||$

ii) $||cu|| = |c|||u||$ for all $c \in \mathbb{R}$ and for all $u \in \mathbb{R}^n$.

We have,

$$||cu|| = ||(cu_1, cu_2, \ldots, cu_n)||$$

$$= \sqrt{((cu_1)^2 + (cu_2)^2 + \ldots + (cu_n)^2)}$$

$$= |c|\sqrt{(u_1^2 + u_2^2 + \ldots + u_n^2)}$$

$$= |c|||u||$$

iii) $||u|| > 0$ for all $u \in \mathbb{R}^n$; $||u|| = 0$ if and only if $u = 0$.

This is obvious as length function is positive square root of $(u_1^2 + u_2^2 + \ldots + u_n^2)$ and length of vector of zero vector is and vice versa.

Example 9.5.2. Consider the following function $||.||: \mathbb{R}^n \to \mathbb{R}$ defined as

$$||u|| = |u_1| + |u_2| + \ldots + |u_n|$$

is a norm.

9.6 Norm induced by inner product

Theorem 9.6.1. *Let V be an inner product space with inner product $\langle ., . \rangle$. Then the function $||.||: V \to \mathbb{R}$ defined by $||v|| = \sqrt{\langle v, v \rangle}$ is a norm on V.*

Proof. Let $u, vw \in V$, lets check the first condition norm,

$$\begin{aligned} ||u+v||^2 &= \langle u+v, u+v \rangle \\ &= \langle u+v, u \rangle + \langle u+v, v \rangle \\ &= \langle u, u \rangle + \langle v, u \rangle + \langle u, v \rangle + \langle v, v \rangle \\ &= \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle \end{aligned}$$

Using the Cauchy-Schwartz inequality,

$$\begin{aligned} &\leq ||u||^2 + 2||u||||v|| + ||v||^2 \\ &= (||u|| + ||v||)^2 \end{aligned}$$

Hence,

$$||u+v|| \leq ||u|| + ||v||$$

lets check second condition of norm,

$$\begin{aligned} ||cu|| &= \sqrt{\langle cu, cu \rangle} \\ &= \sqrt{c^2 \langle u, u \rangle} \\ &= |c|\sqrt{\langle u, u \rangle} \end{aligned}$$

Hence,

$$||cu|| = |c|||u||$$

Now lets check third condition,

Let $u \neq 0$ and $||u||= \sqrt{\langle u, u, \rangle}$ which is positive square root of non zero real number, so it will be always positive.

Again let,

$$||u||= 0$$
$$\iff \sqrt{\langle u, u\rangle} = 0$$
$$\iff \langle u, u\rangle = 0$$
$$\iff v = 0$$

Hence proved.

□

Exercise:

Question 130. Consider a function $f : V \times V \to \mathbb{R}$ where $V \subseteq \mathbb{R}^2$ defined by $f(v, w) = 2v_1w_1 + 5v_2w_2$, where $v = (v_1, v_2), w = (w_1, w_2)$. Choose the set of correct options.

- **Option 1:** f satisfies the symmetry condition of the inner product.
- **Option 2:** f satisfies the bilinearity condition of the inner product.
- **Option 3:** f satisfies the positivity condition of the inner product.
- **Option 4:** f is an inner product.
- Option 5: f is not an inner product.

Question 131. Consider a function $f : V \times V \to \mathbb{R}$ where $V \subseteq \mathbb{R}^2$ defined by $f(v, w) = v_1w_1 - v_1w_2 - v_2w_1 + 4v_2w_2$, where $v = (v_1, v_2), w = (w_1, w_2)$. Choose the set of correct options.

- **Option 1:** f satisfies the symmetry condition of the inner product.
- **Option 2:** f satisfies the bilinearity condition of the inner product.
- **Option 3:** f satisfies the positivity condition of the inner product.
- **Option 4:** f is an inner product.
- Option 5: f is not an inner product.

Question 132. Consider two vectors $a = (0.4, 1.3, -2.2), b = (2, 3, -5)$ in $\mathbb{R}^3$. Choose the set of correct options.

-

- **Option 1:** The two vectors satisfy the triangle inequality given by $\left\|a+b\right\| \leq \|a\| + \left\|b\right\|$.
- Option 2: The two vectors do not satisfy the triangle inequality.
- **Option 3:** The two vectors satisfy the Cauchy-Schwarz inequality given by $|< a, b >| \leq \|a\| \left\|b\right\|$.
- Option 4: The two vectors do not satisfy the Cauchy-Schwarz inequality.

Question 133. Consider a function $f : V \times V \to \mathbb{R}$ where $V \subseteq \mathbb{R}^2$ defined by $f(v, w) = v_1^2 w_1^2 + v_1 w_2^2 + v_2^2 w_1$, where $v = (v_1, v_2)$, $w = (w_1, w_2)$. Choose the set of correct options.

-
- **Option 1:** f satisfies the symmetry condition of the inner product.
- Option 2: f satisfies the positivity condition of the inner product.
- Option 3: f is an inner product.
- **Option 4:** f is not an inner product.

Question 134. Consider a vector (a_1, b_1, c_1) in $\mathbb{R}^3$. Which of these is (are) possible candidates for a norm?

- **Option 1:** $\sqrt{a_1^2 + b_1^2 + c_1^2}$
- Option 2: $a_1 - b_1$
- Option 3: $a_1 + b_1 + c_1$
- Option 4: $\max\{a_1, b_1, c_1\}$
- Option 5: $\min\{a_1, b_1, c_1\}$
- **Option 5:** $\max\{|a_1|, |b_1|, |c_1|\}$

Question 135. Choose the set of correct statements.

- **Option 1:** $V = \mathbb{R}^2$ and the function $< .,. >: V \times V \longrightarrow \mathbb{R}$ defined as $< (x_1, x_2), (y_1, y_2) >= 5x_1y_1 + 8x_2y_2 - 6x_1y_2 - 6x_2y_1$ is an inner product on V.
- Option 2: $V = \mathbb{R}^2$ and the function $< .,. >: V \times V \longrightarrow \mathbb{R}$ defined as $< (x_1, x_2), (y_1, y_2) >= |x_1| + 2|y_2|$ is an inner product on V.

- **Option 3:** $V = \mathbb{R}^2$ and the function $< .,. >: V \times V \longrightarrow \mathbb{R}$ defined as $< (x_1, x_2), (y_1, y_2) >= 3x_1y_1 + 5x_2y_2$ is an inner product on V.

- Option 4: $V = \mathbb{R}^2$ and the function $< .,. >: V \times V \longrightarrow \mathbb{R}$ defined as $< (x_1, x_2), (y_1, y_2) >= 3x_1y_1 - 5x_2y_2$ is an inner product on V.

9.7 Orthogonality and linear Independence

We have already studied how to find the angle between two vectors using the dot product in the vector space $\mathbb{R}^n$ and also know that the dot product and the length of vector are the special cases of inner product and a norm on $\mathbb{R}^n$.

Definition 9.7.1. Orthogonal vector: Two vectors u and v of an inner product space V are said to be orthogonal if $\langle u, v\rangle = 0$.

Example 9.7.1. Consider inner product space $\mathbb{R}^2$ with inner product

$$\langle u, v\rangle = u_1v_1 - (u_1v_2 + u_2v_1) + 2u_2v_2,$$

where $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are the vectors of $\mathbb{R}^2$.

Lets check vectors (1, 1) and (1,0) are orthogonal to each.

$\langle (1, 1), (1, 0)\rangle = 1 - (0 + 1) + 0 = 0$

So vectors (1, 1) and (1, 0) are orthogonal to each other.

Definition 9.7.2. Orthogonal set of vectors: An orthogonal set of vectors of an inner product space V is a set of vectors whose elements are mutually orthogonal.

Explicitly if $S = \{v_1, v_1, \dots, v_k\} \subseteq V$, then S is an orthogonal set of vectors if

$$\langle v_i, v_j\rangle = 0 \text{ for } i, j \in \{1, 2, \dots, k\} \text{ and } i \neq j.$$

Example 9.7.2. Consider the set of vectors $S = \{(4, 3, -2), (-3, 2, -3), (-5, 18, 17)\}$. Set S is an orthogonal set of vectors in the inner product space $\mathbb{R}^3$ with respect to dot product.

$\langle (4, 3, -2), (-3, 2, -3)\rangle = 4 \times (-3) + 3 \times 2 + (-2) \times (-3) = 0$

Similarly, $\langle (-3, 2, -3), (-5, 18, 17)\rangle = 0$ and $\langle (4, 3, -2), (-5, 18, 17)\rangle = 0$

Theorem 9.7.1. *Let $\{v_1, v_2, \dots, v_k\}$ be an orthogonal set of vectors in the inner product space V. Then $\{v_1, v_2, \dots, v_k\}$ is a linearly independent set of vectors*

Proof. Lets assume $\sum_{i=1}^{k} c_iv_i = 0$ where $c_i \in \mathbb{R}$ for all $i \in \{1, 2, \dots, k\}$

Then,
$\langle \sum_{i=1}^{k} c_iv_i, v_1\rangle = \langle 0, v_1\rangle = 0$

$\implies \sum_{i=1}^{k} \langle c_i v_i, v_1 \rangle = 0$
$\implies \sum_{i=1}^{k} c_i \langle v_1, v_1 \rangle = 0$

$\implies c_1 \langle v_i, v_1 \rangle = 0$ (since each set S is an orthogonal set)
$\implies c_1 = 0$

Similarly we can find all $c_i = 0$ for all $i \in \{2, \dots, k\}$

Hence, orthogonal set is an linearly independent set.

□

Definition 9.7.3. Orthogonal basis: Let V be an inner product space. A basis consisting of mutually orthogonal vectors is called an orthogonal basis.

We know that basis is a maximal linearly independent set and orthogonal set of vectors is already linearly independent so orthogonal set is a basis precisely when it a **maximal orthogonal set** (i.e., there is no orthogonal set which is strictly contains this set).

Note: If $dim(V) = n$, then

Orthogonal basis is nothing but the orthogonal set of n vectors

Lets see some example of orthogonal basis,

i) The standard basis of of inner product space $\mathbb{R}^n$ with respect to dot product.

ii) Just now we have seen the set $S = \{(4, 3, -2), (-3, 2, -3), (-5, 18, 17)\}$ is an orthogonal subset of $\mathbb{R}^3$ with respect to dot product and $dim(\mathbb{R}^3) = 3$ so the set S is an orthogonal basis of inner product $\mathbb{R}^3$.

iii) We have seen that vectors (1, 1) and (1, 0) are orthogonal vectors in the inner product space $\mathbb{R}^2$ with inner product

$$\langle u, v \rangle = u_1 v_1 - (u_1 v_2 + u_2 v_1) + 2 u_2 v_2,$$

where $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are the vectors of $\mathbb{R}^2$.

So the set $\{(1, 1), (1, 0)\}$ is an orthogonal basis of the inner product space $\mathbb{R}^3$.

Definition 9.7.4. Orthonormal set: An orthonormal set of vectors of an inner product space V is an orthogonal set of vectors such that the norm of each vector of the set is 1.

Explicitly if $S = \{v_1, v_1, \ldots, v_k\} \subseteq V$, then S is an orthonormal set of vectors if

$\langle v_i, v_j\rangle = 0$ for $i, j \in \{1, 2, \ldots, k\}$ and $i \neq j$ and $||v_i||= 1$ for all $i \in \{1, 2, \ldots, k\}$

Definition 9.7.5. Orthonormal basis: An orthonormal basis is an orthonormal set of vectors which form a basis.

Note: An orthonormal basis is an orthogonal basis where the norm if each vector is 1.

Note: An orthonormal basis is a maximal orthonormal set in the inner product space.

For example, the standard basis of $\mathbb{R}^n$ with respect to dot product forms an orthonormal basis.

Example 9.7.3. Consider $\mathbb{R}^3$ with the usual inner product (i.e. dot product) and the set $S = \{\frac{1}{3}(1, 2, 2), \frac{1}{3}(-2, -1, 2), \frac{1}{3}(2, -2, 1)\}$. Then S forms an orthonormal basis of $\mathbb{R}^3$.

$||\frac{1}{3}(1, 2, 2)||= \langle \frac{1}{3}(1, 2, 2), \frac{1}{3}(1, 2, 2)\rangle = 1$

Similarly, others vectors has norm 1.

Now, lets check orthogonality,

$\langle \frac{1}{3}(1, 2, 2), \frac{1}{3}(-2, -1, 2)\rangle = 0$

So vectors $\frac{1}{3}(1, 2, 2)$ and $\frac{1}{3}(-2, -1, 2)$ are orthogonal vectors.

Similarly, we can check others.

Hence set S is an orthonormal set and as $dim(\mathbb{R}^3) = 3$ so S is an orthonormal basis of $\mathbb{R}^3$.

9.7.1 Obtaining orthonormal set from orthogonal set

If $\gamma = \{v_1, v_2, \ldots, v_n\}$ is an orthogonal set of vectors then we can obtain an orthonormal set of vectors β from γ by dividing each vector by its norm i.e.,

$$\beta = \left\{\frac{v_1}{||v_1||}, \frac{v_2}{||v_2||}, \ldots, \frac{v_n}{||v_n||}\right\}.$$

Example 9.7.4. Consider $\mathbb{R}^2$ with usual inner product and the orthogonal basis $S = \{(1,3), (-3,1)\}$.

Then $S_1 = \{\frac{1}{\sqrt{10}}(1,3), \frac{1}{\sqrt{10}}(-3,1)\}$ is an orthonormal basis of $\mathbb{R}^2$.

9.7.2 Importance of orthonormal basis

Suppose $S = \{v_1, v_2, \dots, v_n\}$ is an orthonormal basis of an inner product space V and let $v \in V$.

Then v can be written as

$$v = c_1v_1 + c_2v_2 + \dots + c_nv_n$$

How do we find $c_1, c_2, \dots, c_n$? For any basis, this means writing a system of linear equation and solving it.

But since S is an orthonormal, we can use inner product and can compute $c_i = \langle v, v_i \rangle$, $i \in \{1, 2, 3, \dots, n\}$,

as

$$\begin{aligned}\langle v, v_i \rangle &= \langle c_1v_1 + c_2v_2 + \dots + c_nv_n, v_i \rangle \\ &= \langle c_1v_1, v_i \rangle + \langle c_2v_2, v_i \rangle + \dots + \langle c_nv_n, v_i \rangle \\ &= c_i \langle v_i, v_i \rangle \\ &= c_i ||v_i||^2 = c_i\end{aligned}$$

<u>Exercise:</u>

Question 136. Consider two vectors $a = (2,0,3,0,8), b = (3,2,-2,4,0)$ in $\mathbb{R}^5$. Choose the set of correct options.

- **Option 1:** a and b are orthogonal.
- Option 2: a and b are not orthogonal.
- Option 3: $(a - b) \cdot a = 0$.
- **Option 4:** $(a - b) \cdot a = 77$.

Question 137. Choose the set of correct statements.

- Option 1: In an orthogonal set, the norms of all the vectors are equal.
- **Option 2:** In an orthogonal set, the vectors are linearly independent.

- Option 3: In an orthogonal set, the vectors are linearly dependent.
- **Option 4:** If the columns of an $n \times n$ coefficient matrix A comprises the individual vectors of an orthogonal set in $\mathbb{R}^n$, then there must be a unique solution to the system $AX = b$, where X, b are $n \times 1$ vectors.
- Option 5: If the columns of an $n \times n$ coefficient matrix A comprises the individual vectors of an orthogonal set in $\mathbb{R}^n$, then there are no solutions to the system $AX = b$, where X, b are $n \times 1$ vectors.
- Option 6: The determinant of a square matrix formed by a set of orthogonal vectors in $\mathbb{R}^n$ is zero.
- **Option 7:** A set of n vectors can never form an orthogonal basis in $\mathbb{R}^{n-1}$.

Question 138. Which of the following is an orthogonal basis of the given vector spaces with respect to the standard inner product (dot product)?

- **Option 1:** $\{(1, 0), (0, 1)\}$ is an orthogonal basis of $\mathbb{R}^2$.
- **Option 2:** $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is an orthogonal basis of $\mathbb{R}^3$.
- Option 3: $\{(3, 4), (4, -3), (2, -3)\}$ is an orthogonal basis of $\mathbb{R}^2$.
- Option 4: $\{(2, 1, -1), (-1, 1, -1), (3, -3, 3)\}$ is an orthogonal basis of $\mathbb{R}^3$.

Question 139. Find a vector in $\mathbb{R}^4$ that is orthogonal to the subspace spanned by $(1, 1, 0, 0)$ and $(0, 1, 1, 0)$ with respect to the dot product as the inner product.

- **Option 1:** $(1, -1, 1, 0)$
- Option 2: $(2, 3, 4, 5)$
- **Option 3:** $(1, -1, 1, 1)$
- Option 4: $(1, 1, 1, 0)$

Question 140. Which of the following are orthogonal to the vector $(1, 2)$ in $\mathbb{R}^2$ with respect to the inner product $< v, w >= v_1w_1 - v_1w_2 - v_2w_1 + 4v_2w_2$?

- Option 1: $(-7, 1)$
- Option 2: $(6, -1)$
- **Option 3:** $(7, 1)$
- Option 4: $(-9, 1)$

- **Option 5:** $(14, 2)$

Question 141. Consider the system of linear equations:

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 1 \\ 2x_1 + x_2 &= 5 \\ 3x_1 - 6x_2 - 5x_3 &= 9. \end{aligned}$$

Number of solution of the above system is [Ans: 1]

Question 142. Let A be the coefficient matrix of the system of linear equations:

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 1 \\ 2x_1 + x_2 &= 5 \\ 3x_1 - 6x_2 - 5x_3 &= 9. \end{aligned}$$

Which one of the following is true about the matrix AA^T?

- Option 1: A scalar matrix.
- Option 2: The identity matrix.
- **Option 3:** A diagonal matrix.
- **Option 4:** A lower triangular matrix.
- **Option 5:** An upper triangular matrix.
- Option 6: None of the above.

Question 143. Which of these sets form an orthonormal basis of $\mathbb{R}^3$ with respect to the dot product as the inner product in $\mathbb{R}^3$?

- Option 1: $\{(1, 0, 0), (0, 1, 0)\}$
- Option 2: $\{(2, 0, 0), (0, 2, 0), (0, 0, 2)\}$
- **Option 3:** $\{(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}), (\frac{-3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}), (\frac{1}{\sqrt{66}}, \frac{7}{\sqrt{66}}, \frac{-4}{\sqrt{66}})\}$
- Option 4: $\{(2, 1, -1), (-1, -1, 1), (1, 2, -2)\}$
- **Option 5:** $\{(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}), (0, 1, 0), (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})\}$

Question 144. Consider two orthogonal vectors a, b in $\mathbb{R}^2$. If $a + b$ and $a - b$ are orthogonal, then choose the correct option.

- Option 1: $||a|| = ||b|| = 1$
- **Option 2:** $||a|| = ||b||$
- Option 3: $||a|| = 2\,||b||$
- Option 4: $2\,||a|| = ||b||$

Question 145. Consider $a = (1, 1), b = (1, -1)$. Let $V = \text{Span}\{a, b\}$. Choose the correct options by considering the standard inner product (dot product).

-
- **Option 1:** Vectors a, b form an orthogonal basis for V.
- Option 2: Vectors a, b form an orthonormal basis for V.
- **Option 3:** There exist scalar multiples of a, b which form an orthonormal basis
- Option 4: There do not exist scalar multiples of a, b which form an orthonormal basis

Question 146. Choose the set of correct statements.

- **Option 1:** The determinant of a matrix formed by 3 orthonormal vectors in $\mathbb{R}^3$ is ± 1.
- Option 2: The determinant of a matrix formed by 3 orthonormal vectors in $\mathbb{R}^3$ is 0.
- **Option 3:** The determinant of a matrix formed by 2 orthonormal vectors in $\mathbb{R}^2$ is ± 1.
- Option 4: The determinant of a matrix formed by 2 orthonormal vectors in $\mathbb{R}^2$ is 0.

Question 147. Consider a system of linear equations:

$$\begin{aligned} 2x_1 + 2x_2 + 7x_3 &= b_1 \\ 2x_1 + x_2 - 10x_3 &= b_2 \\ 3x_1 - 2x_2 + 2x_3 &= b_3. \end{aligned}$$

Let A be the coefficient matrix of the given system of linear equations. Let a matrix B contain the column vectors of A, which are normalized by their respective norms, as its columns (i.e. first column vector of A normalized by its norm is the first column of B). Which of the following statements are true ?

- **Option 1:** The determinant of BB^T is 1.
- **Option 2:** BB^T is an identity matrix.
- **Option 3:** BB^T is a scalar matrix.
- **Option 4:** BB^T is a diagonal matrix.

Question 148. Choose the correct option(s).

- **Option 1:** The vectors in an orthonormal set are linearly independent.
- Option 2: A set of linearly dependent vectors can be orthonormal.
- Option 3: A set of linearly independent vectors is always orthonormal.
- Option 4: A set of linearly independent vectors is always orthogonal but not orthonormal.
- **Option 5:** The vectors in an orthogonal set are linearly independent.

Consider the inner product $\langle u, v\rangle = 3u_1v_1 + 2u_2v_2$ where $u = (u_1, u_2)$, $v = (v_1, v_2)$ are vectors in $\mathbb{R}^2$. Answer following questions based on this information.

Question 149. Which of the following is an orthonormal basis with respect to the inner product defined above?

- Option 1: $\{(2, 3), (-4, 4)\}$
- **Option 2:** $\{\frac{1}{\sqrt{30}}(2, 3), \frac{1}{\sqrt{80}}(-4, 4)\}$
- Option 3: $\{\frac{1}{\sqrt{13}}(2, 3), \frac{1}{\sqrt{32}}(-4, 4)\}$
- Option 4: $\{(2, 3), (-3, 2)\}$
- Option 5: $\{\frac{1}{\sqrt{13}}(2, 3), \frac{1}{\sqrt{13}}(-3, 2)\}$

Question 150. Use the orthonormal basis $\{u, v\}$ obtained in question 7 with respect to the defined inner product. Express the vector $(4, 0)$ as a linear combination of the basis vectors u and v, as $(4, 0) = c_1u + c_2v$. Which of the following gives the coefficients of the linear combination?

- **Option 1:** $c_1 = \frac{24}{\sqrt{30}}, c_2 = \frac{-48}{\sqrt{80}}$
- Option 2: $c_1 = \frac{24}{\sqrt{13}}, c_2 = \frac{-48}{\sqrt{32}}$
- Option 3: $c_1 = \frac{8}{\sqrt{13}}, c_2 = \frac{-16}{\sqrt{32}}$
- Option 4: $c_1 = \frac{24}{\sqrt{30}}, c_2 = \frac{48}{\sqrt{80}}$

9.8 Projections using inner products

In this section, we shall discuss projections.

9.8.1 Projection of a vector along another vector

Let A and B be two points in $\mathbb{R}^2$. Suppose we want to find the point nearest to B on the line l passing through A and the origin.

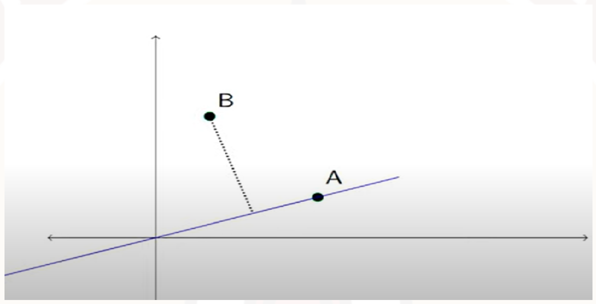


Figure 9.4:

As we already know, the nearest point will be the foot of the perpendicular drawn from the point B to the line l. In this section, we shall try to solve this problem in the perspective of vectors using the concept of inner products.

For this, we shall draw the vectors corresponding to the points A and B.

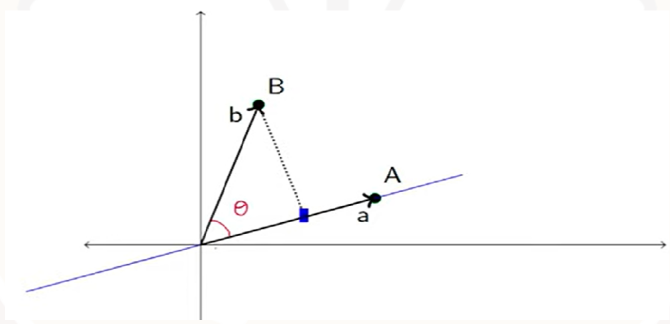


Figure 9.5:

We want to find the length of the vector v which is the foot of the perpendicular from B to the line l. Once we know the length, $||v||$, we can determine the vector v. This is because, $v = \alpha a$ (since v lies along the vector a). Hence $||v|| = \alpha ||a||$. From this we get the value of α to be $\dfrac{||v||}{||a||}$. Thus $v = \dfrac{||v||}{||a||} a$. Note this can also be written as $v = ||v||\dfrac{a}{||a||}$, that is, find the unit vector in the direction of a and multiply it with the length of the vector v. Since the vector a is know, it only remains to find $||a||$. We know $||v||$ from the right angled triangle. $||v|| = ||b||cos\ \theta$ (θ is the angle between the two vectors a and b) and we know $cos\ \theta$ from the earlier section as $\dfrac{\langle a, b\rangle}{||a||\ ||b||}$. Thus $v = ||b||\dfrac{\langle a, b\rangle}{||a||\ ||b||}\dfrac{a}{||a||} = \dfrac{\langle a, b\rangle}{||a||^2}a = \dfrac{\langle a, b\rangle}{\langle a, a\rangle}a$. So, if we know the vectors a and b, we can find the point along the vector a that is nearest to the vector b. This vector v, is called the projection of the vector b along the vector a.

Similarly, we can talk about shortest distances in $\mathbb{R}^3$. Suppose b is a point whose shortest distance from the plane generated by $(1, 0, 0)$ and $(0, 1, 0)$ needs to be calculated.

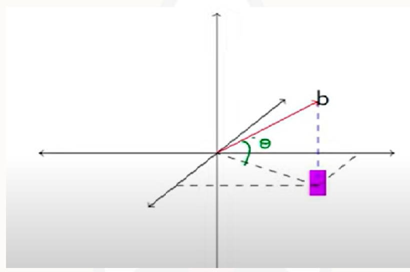


Figure 9.6:

We use the exact same idea as in the case of $\mathbb{R}^2$. Instead of a single vector a, we have a plane, namely the XY-plane in $\mathbb{R}^3$. We shall see projections in a general vector space and the case of $\mathbb{R}^3$ will be a particular case of the general vector space.

9.8.2 Projection of a vector onto a subspace

Let V be an inner product space, $v \in V$ and $W \subseteq V$ be a subspace. Then the projection of v onto W is the vector in W, denoted by $proj_W(v)$, computed as follows:

- Find an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$ for W.
- Define $proj_W(v) = \sum_{i=1}^{n}\langle v, v_i\rangle v_i$. (Note that $||v_i|| = 1\ \forall i$ and hence we are not dividing by its norm.)

Observe that $\{v_1, v_2, ..., v_n\}$ is a randomly chosen orthonormal basis for W and hence $proj_W(v)$ is independent of the chosen orthonormal basis. The projection of a vector onto a subspace does not change with the choice of a basis (as expected!). The projection of a vector v onto a subspace W is the nearest vector $v' \in W$, that is v' satisfies $||v - v'|| \leq ||v - w||$ for all $w \in W$.

Example 9.8.1. Let $V = \mathbb{R}^2$ and W be the line $y = x$. The projection of $(1, 2)$ onto W is $\frac{\langle (1,2),(1,1)\rangle}{\langle (1,1),(1,1)\rangle(1,1)} = (\frac{3}{2}, \frac{3}{2})$. This can be done by taking any vector along the line $y = x$ as a instead of $(1, 1)$. We can also solve the same problem by finding an orthonormal basis for W and using the formula for the projection. An orthonormal basis for W can be $\{\frac{1}{\sqrt{2}}(1, 1)\}$.

Note that when W is a one-dimensional subspace like the one given in the above example, then $proj_w(v) = proj_W(v)$, where W is the subspace generated by the vector w. Projection of a vector v along another vector w can also be visualised as the projection of v onto the subspace generated by w.

Example 9.8.2. Let $V = \mathbb{R}^3$ and $W = \{(x, y, z) : x+y = 0\}$. Clearly $\{(1, -1, 0), (0, 0, 1)\}$ is a basis for W. Also this is an orthogonal basis. We can convert it into an orthonormal basis by dividing each vector by its norm. Thus $\{\frac{1}{\sqrt{2}}(1, -1, 0), (0, 0, 1)\}$ is an orthonormal basis for W.

What is the projection of $(2, -2, 0)$? (Think!)

What is the projection of $(2, 1, 3)$? For this, we use the orthonormal basis of W obtained earlier.

$$\begin{aligned} proj_W(2,1,3) &= \langle (2,1,3), \frac{1}{\sqrt{2}}(1,-1,0)\rangle \frac{1}{\sqrt{2}}(1,-1,0) + \langle (2,1,3),(0,0,1)\rangle(0,0,1) \\ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(1,-1,0) + 3(0,0,1) \\ &= (\frac{1}{2}, -\frac{1}{2}, 3). \end{aligned}$$

Observe that $proj_W(v)$ lies in W (for obvious reasons!).

Now suppose we should project the vector v onto a subspace W for which an orthogonal basis is given instead of an orthonormal basis, we convert the orthogonal basis into an orthonormal one as illustrated in the previous example.

9.8.3 Projection as a linear transformation

Let V be an inner product space and W be a subspace. Define $T : V \to V$ defined by $T(v) = proj_W(v)$ for all $v \in V$.

Question 151.

- Verify that T is a linear transformation.
- What is the range of T?
- What is the rank of T?

- What is the kernel of T?
- What is $T(v)$ for any $v \in W$?

This linear transformation is called the projection map from V to W and is denoted by P_W.

Note that P_W sends vectors from V onto W and hence range of $P_W = W$. Also, the rank of P_W is the dimension of W (since $R(P_W) = W$).

Now, coming to the kernel of P_W. Suppose $\{v_1, v_2, ..., v_n\}$ is an orthonormal basis for W. The kernel of P_W is all those vectors that satisfy $P_W(v) = proj_W(v) = \sum_{i=1}^{n} \langle v, v_i \rangle v_i = 0$. Realise that we are looking for all those vectors v that are orthogonal to each of the v_i's. The set $W^\perp = \{v \in V : \langle v, w \rangle = 0, \ \forall \ w \in W\}$ is called the orthogonal complement of W and the kernel of P_W is precisely $W^\perp$.

Suppose $v \in W$, then $P_W(v) = v$ (why?). From this, it is easy to conclude that $P_W^2 = P_W$.

Also, $||P_W(v)|| \leq ||v||$ (why?). The projection of a vector cannot be longer than the vector itself.

Example 9.8.3. Consider the inner product $\langle a, b \rangle = a_1b_1 - a_1b_2 - a_2b_1 + 4a_2b_2$ where $a = (a_1, a_2)$, $b = (b_1, b_2)$ are vectors in $\mathbb{R}^2$.

1. Let $x = (1, 2)$. Find the projection of x in the direction of $(3, 4)$ using the inner product defined above.

$$\begin{aligned} proj_{(3,4)}(1, 2) &= \frac{\langle (1, 2), (3, 4) \rangle}{\langle (3, 4), (3, 4) \rangle}(3, 4) \\ &= \frac{(3)(1) - (3)(2) - (4)(1) + 4(4)(2)}{(3)(3) - (3)(4) - (4)(3) + 4(4)(4)(3, 4)}(3, 4) \\ &= \frac{25}{49}(3, 4) \end{aligned}$$

2. Let $x = (1, 2)$, find the projection of x in a direction perpendicular to $(3, 4)$.
   Direction (v_1, v_2), perpendicular to $(3, 4)$ can be got by solving $\langle (v_1, v_2), (3, 4) \rangle = 0$. Thus $3v_1 - 4v_1 - 3v_2 + 16v_2 = 0$, that is $13v_2 = v_1$. The direction perpendicular to $(3, 4)$ is $(13, 1)$.
   $proj_{(13,1)}(1, 2)$ can be got in a similar fashion.

With an orthogonal basis, we can find an orthonormal basis. But if the given basis is not orthogonal basis, then the procedure by which we convert the given basis to an orthonormal basis is called the Gram-Schmidt orthonormalization process that is discussed in the next section.

9.8.4 Exercises

Question 152. Consider an orthonormal basis $\{(1,0,0),(0,1,0)\}$ for a subspace W in $\mathbb{R}^3$. If $x = (1,2,3)$ is a vector in $\mathbb{R}^3$, then which of the following represents a vector in W whose distance from x is the least? Consider dot product as the standard inner product.

- Option 1: $(2,4,0)$
- Option 2: $(3,4,0)$
- Option 3: $(4,5,0)$
- **Option 4:** $(1,2,0)$

Question 153. Suppose W_1 and W_2 are subspaces of a vector space V. Let P_{W_1} and P_{W_2} denote the projection from V to W_1 and V to W_2 respectively and consider the following statements:

- **Statement P:** If $P_{W_1} + P_{W_2}$ is a projection from V to $W_1 + W_2$, then $P_{W_1} \circ P_{W_2} + P_{W_2} \circ P_{W_1} = 0$.
- **Statement Q:** $P_{W_1}^2 := P_{W_1} \circ P_{W_1}$ is not a projection from V to W_1.
- **Statement R:** If A is the matrix representation of P_{W_2}, then A can not be a symmetric matrix.
- **Statement S:** $P_{W_1} - P_{W_2}$ is a projection from V to $W_1 + W_2$.

Find the number of correct statements [Ans: 1]

Question 154. Consider an orthogonal basis $\{(1,2,1),(-2,0,2)\}$ of a subspace W, of the inner product space $\mathbb{R}^3$ with respect to the dot product. If $y = (1,2,3) \in \mathbb{R}^3$, then find $Proj_W(y)$.

- **Option 1:** $(\frac{1}{3}, \frac{8}{3}, \frac{7}{3})$
- Option 2: $(\frac{18}{\sqrt{6}} - \frac{32}{\sqrt{8}}, \frac{36}{\sqrt{8}}, \frac{18}{\sqrt{6}} + \frac{32}{\sqrt{8}})$
- Option 3: $(\frac{1}{3}, -\frac{8}{3}, \frac{7}{3})$
- Option 4: $(\frac{18}{\sqrt{6}} - \frac{32}{\sqrt{8}}, \frac{36}{\sqrt{8}}, -\frac{18}{\sqrt{6}} + \frac{32}{\sqrt{8}})$

9.9 Gram-Schmidt orthonormalization

As already mentioned, in any inner product space, the Gram-Schmidt process converts any basis into an orthonormal one. Let us begin by an example to understand the idea behind the process.

Example 9.9.1. Consider the basis $\{u_1, u_2, u_3\} = \{(1,2,2), (-1,0,2), (0,0,1)\}$ for $\mathbb{R}^3$ with the usual inner product. Note that this basis is not an orthonormal basis ($\langle (1,2,2), (-1,0,2)\rangle = 3 \neq 0$, thus the set is not even orthogonal). First we convert this basis into an orthogonal basis and then divide each vector by its norm to get an orthonormal basis.

First, let $v_1 = u_1(1,2,2)$. Now, we want to consider vectors orthogonal to v_1, that is vectors in $\langle v_1\rangle^{\perp}$. For this, we use the projection P_{v_1} on v_2.

Define v_2 as $u_2 - P_{v_1}(u_2)$. Thus we get

$$\begin{aligned}
v_2 &= (-1,0,2) - P_{v_1}((-1,0,2)) \\
&= (-1,0,2) - \frac{\langle(-1,0,2),(1,2,2)\rangle}{\langle(1,2,2),(1,2,2)}(1,2,2) \\
&= (-1,0,2) - \frac{3}{9}(1,2,2) \\
&= (-\frac{4}{3}, -\frac{2}{3}, \frac{4}{3})
\end{aligned}$$

Note that, at this stage, $\langle v_1, v_2\rangle = 0$. Note that the way we have defined v_2 made sure that $\langle v_1, v_2\rangle = 0$. This is because

$$\begin{aligned}
\langle v_1, v_2\rangle &= \langle v_1, u_2 - P_{v_1}(u_2)\rangle \\
&= \langle v_1, u_2 - \frac{\langle v_1, u_2\rangle}{\langle v_1, v_1} v_1\rangle \\
&= \langle v_1, u_2\rangle - \frac{\langle v_1, u_2\rangle}{\langle v_1, v_1\rangle}\langle v_1, v_1\rangle \\
&= \langle v_1, u_2\rangle - \langle v_1, u_2\rangle = 0
\end{aligned}$$

As already stated, $W^{\perp}$ is the null space of P_W. Here, we are using a vector from $W^{\perp}$, namely $(I - P_W)(v)$. Clearly $P_W((I - P_W)(v)) = 0$ for all v.

In the next step, we want v_3 to be in the orthogonal complement of span $\{v_1, v_2\}$, that is in the set span $(\{v_1, v_2\})^{\perp}$. That is, we want v_3 to be orthogonal to both v_1

and v_2. Define $v_3 = u_3 - P_{v_1}(u_3) - P_{v_2}(u_3)$.

$$\begin{aligned}
v_3 &= (0,0,1) - P_{v_1}(0,0,1) - P_{v_2}(0,0,1)\\
&= (0,0,1) - \frac{\langle (0,0,1),(1,2,2)\rangle}{\langle (1,2,2),(1,2,2)\rangle}(1,2,2) - \frac{\langle (0,0,1),(-\frac{4}{3},-\frac{2}{3},\frac{4}{3})\rangle}{\langle (-\frac{4}{3},-\frac{2}{3},\frac{4}{3}),(-\frac{4}{3},-\frac{2}{3},\frac{4}{3})\rangle}(-\frac{4}{3},-\frac{2}{3},\frac{4}{3})\\
&= (0,0,1) - \frac{2}{9}(1,2,2) - \frac{\frac{4}{3}}{\frac{36}{9}}(-\frac{4}{3},-\frac{2}{3},\frac{4}{3})\\
&= (0,0,1) - (\frac{2}{9},\frac{4}{9},\frac{4}{9}) - (-\frac{4}{9},-\frac{2}{9},\frac{4}{9})\\
&= (\frac{2}{9},-\frac{2}{9},\frac{1}{9})
\end{aligned}$$

Verify that $\{v_1, v_2, v_3\}$ forms an orthogonal set and hence is linearly independent and hence a basis for $\mathbb{R}^3$. From this, we can divide each of the vectors by its norm and construct an orthonormal basis.

9.9.1 The Gram-Schmidt process

Let $\gamma = \{v_1, v_2, \ldots, v_n\}$ be a given ordered basis of a vector space. The steps in the process described below yields an orthonormal basis $\beta = \{u_1, u_2, \ldots, u_n\}$.

Step-1 $w_1 = v_1$, and $u_1 = \frac{w_1}{||w_1||}$.

Step-2 $w_2 = v_2 - \langle v_2, u_1\rangle u_1$ and $u_2 = \frac{w_2}{||w_2||}$.

Step-3 $w_3 = v_3 - \langle v_3, u_1\rangle u_1 - \langle v_3, u_2\rangle u_2$ and $u_3 = \frac{w_3}{||w_3||}$.

$\vdots$

Step-i $w_i = v_i - \sum_{j=1}^{i-1}\langle v_i, u_j\rangle u_j$ and $u_i = \frac{w_i}{||w_i||}$.

$\vdots$

Step-n $w_n = v_n - \sum_{j=1}^{n-1}\langle v_n, u_j\rangle u_j$ and $u_n = \frac{w_n}{||w_n||}$.

Note that $\{w_1, w_2, \ldots, w_n\}$ is an orthogonal basis. At each stage, span $(\{v_1, v_2, \ldots, v_i\})=$ span $(\{u_1, u_2, \ldots, u_i\})=$ span $(\{w_1, w_2, \ldots, w_i\})$.

From this process, we conclude this section with an important theorem for finite dimensional inner product spaces.

Theorem 9.9.1. *Any finite dimensional inner product space has an orthonormal basis.*

9.9.2 Exercises

Question 155. Let W be a subspace of the inner product space $\mathbb{R}^4$ with respect to dot product and $\{v_1, v_2, v_3\}$ be an ordered basis of W, where $v_1 = (1,1,1,1), v_2 = (1,1,1,0), v_3 = (1,1,0,0)$. Which of these is the orthonormal basis of W obtained using the Gram-Schmidt process?

- **Option 1:** $\{\frac{1}{2}(1,1,1,1), \frac{1}{2\sqrt{3}}(1,1,1,-3), \frac{1}{\sqrt{6}}(1,1,-2,0)\}$
- Option 2: $\{\frac{1}{2}(1,1,1,1), \frac{1}{2\sqrt{3}}(1,1,1,-3), \frac{1}{\sqrt{6}}(1/3,1/3,2/3,0)\}$
- Option 3: $\{\frac{1}{2}(1,1,1,1), \frac{1}{2\sqrt{3}}(1,1,1,-3), \frac{1}{\sqrt{6}}(3,1,-2,0)\}$
- Option 4: $\{\frac{1}{2}(1,1,1,1), \frac{1}{2\sqrt{3}}(1,1,1,-3), \frac{1}{\sqrt{6}}(3,1,2,-2)\}$

Question 156. Let $a = (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$ be a vector from the inner product space $\mathbb{R}^3$ with respect to dot product and $W = \{(x,y,z) \in \mathbb{R}^3 \mid \langle (x,y,z), (\frac{2}{3}, \frac{2}{3}, \frac{1}{3}) \rangle = 0\}$ be a subspace of $\mathbb{R}^3$. Then which of the following is (are) a basis of W?

- Option 1: $\{(1,0,2), (0,1,2)\}$
- **Option 2:** $\{(1,0,-2), (0,1,-2)\}$
- **Option 3:** $\{(-1,0,2), (0,-1,2)\}$
- Option 4: $\{(1,0,2), (0,1,-2)\}$

Question 157. Consider a set $W = \{(1,1,1)\}$ from the inner product space $\mathbb{R}^3$. Find the dimension of the subspace $W^\perp$, where $W^\perp$ is the collection of the vectors which are orthogonal to the vector (1,1,1). [Ans: 2]

9.10 Orthogonal Transformations and Rotations

9.10.1 Orthogonal Transformations

Let V be an inner product space and T be a linear transformation from V to V. T is said to be an orthogonal transformation if

$$\langle Tv, Tw \rangle = \langle v, w \rangle, \ \forall v, w \in V.$$

That is, an orthogonal transformation preserves the inner product.

If the inner product space under consideration is $\mathbb{R}^n$ with the dot product as the inner product, then an orthogonal transformation preserves lengths and angles. (Why?)

Example 9.10.1. Consider the inner product space $\mathbb{R}^2$ with the dot product. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as $T(x, y) = (\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}})$. We shall now show that T is an orthogonal linear transformation. It is clear that T is a linear transformation (is it?!). To show that T is orthogonal, we need to show that

$$\langle T(x_1, y_1), T(x_2, y_2)\rangle = \langle (x_1, y_1), (x_2, y_2)\rangle \ \forall\ (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2.$$

$$\begin{aligned}
\langle T(x_1, y_1), T(x_2, y_2)\rangle &= \langle \left(\frac{x_1 + y_1}{\sqrt{2}}, \frac{x_1 - y_1}{\sqrt{2}}\right), \left(\frac{x_2 + y_2}{\sqrt{2}}, \frac{x_2 - y_2}{\sqrt{2}}\right)\rangle \\
&= \left(\frac{x_1 + y_1}{\sqrt{2}}\right)\left(\frac{x_2 + y_2}{\sqrt{2}}\right) + \left(\frac{x_1 - y_1}{\sqrt{2}}\right)\left(\frac{x_2 - y_2}{\sqrt{2}}\right) \\
&= \frac{1}{2}(x_1x_2 + x_1y_2 + y_1x_2 + y_1y_2 + x_1x_2 - x_1y_2 - y_1x_2 + y_1y_2) \\
&= x_1x_2 + y_1y_2 \\
&= \langle (x_1, y_1), (x_2, y_2)\rangle
\end{aligned}$$

Thus T is an orthogonal transformation.

Example 9.10.2. Consider the inner product spaces $\mathbb{R}^3$ and $\mathbb{R}^2$ with the dot product. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined as $T(x, y, z) = (x + y + z, x + y)$. It can be shown that T is not an orthogonal transformation.

$$\begin{aligned}
\langle T(1, 0, 0), T(0, 1, 0)\rangle &= \langle (1, 1), (1, 1)\rangle = 2 \\
\langle (1, 0, 0), (0, 1, 0)\rangle &= 0
\end{aligned}$$

Since $\langle T(1, 0, 0), T(0, 1, 0)\rangle \neq \langle (1, 0, 0), (0, 1, 0)\rangle$, T is not an orthogonal transformation.

9.10.2 Rotation Matrices

The rotation matrices rotate a vector by an angle θ. Let us calculate the rotation matrix in $\mathbb{R}^2$. Consider the standard basis of $\mathbb{R}^2$, $\{(1, 0), (0, 1)\}$. Rotate the plane by an angle θ. The matrix of the linear transformation is got from the vectors obtained after rotation.

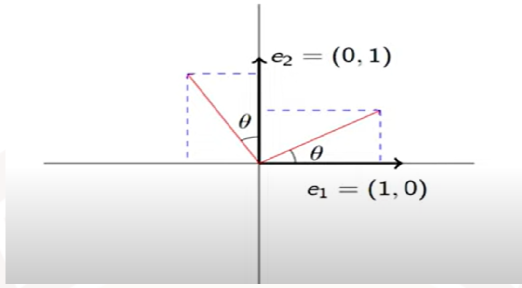


Figure 9.7:

We can see that the vector $(1, 0)$ gets rotated to the vector that makes an angle θ with the positive x-axis, that is the vector $(cos\ \theta, sin\ \theta)$. Similarly, the vector $(0, 1)$ gets rotated to the vector that makes an angle $\frac{\pi}{2} + \theta$ with the positive x-axis, that is the vector $(cos(\frac{\pi}{2} + \theta), sin(\frac{\pi}{2} + \theta))$ which is nothing but $(-sin\ \theta, cos\ \theta)$. Now it is easy to get the matrix of the linear transformation with the information in hand.

$$T(1, 0) = (cos\ \theta, sin\ \theta) = (cos\ \theta)(1, 0) + (sin\ \theta)(0, 1)$$
$$T(0, 1) = (-sin\ \theta, cos\ \theta) = (-sin\ \theta)(1, 0) + (cos\ \theta)(0, 1)$$

Thus the matrix of this linear transformation is $\begin{pmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{pmatrix}$. This transformation is denoted by R_θ and the matrix of R_θ is the rotation matrix (rotation by an angle θ) in $\mathbb{R}^2$.

Observe that $R_\theta^T = R_{-\theta}$ and $R_\theta^T R_\theta = R_\theta R_\theta^T = I$.

Also, the rotated vectors are also of length one and are orthogonal to each other and hence form an orthonormal basis for $\mathbb{R}^2$.

In $\mathbb{R}^3$, we get different rotation matrices based on the axis of rotation. Consider the rotations about the axes in $\mathbb{R}^3$. Since they preserve lengths and angles, they are orthogonal transformations.

Again, with respect to the standard ordered basis, the rotation matrix corresponding to the rotation about z-axis by angle θ can be got by keeping $(0, 0, 1)$ exactly same and rotate $\{(1, 0, 0), (0, 1, 0)\}$ in a fashion similar to the one in $\mathbb{R}^2$. Thus the matrix obtained denoted by $T_3(\theta)$ is $\begin{pmatrix} cos\ \theta & -sin\ \theta & 0 \\ sin\ \theta & cos\ \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Similarly we get the matrix corresponding to the rotation about x-axis, where the rotation happens on the yz-plane. $T_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\ \theta & -sin\ \theta \\ 0 & sin\ \theta & cos\ \theta \end{pmatrix}$.

Again, we get the matrix corresponding to the rotation about y-axis, where the rotation happens on the xz-plane. $T_2(\theta) = \begin{pmatrix} cos\ \theta & 0 & -sin\ \theta \\ 0 & 1 & 0 \\ sin\ \theta & 0 & cos\ \theta \end{pmatrix}$.

9.10.3 Orthogonal Matrices

Suppose A is the matrix of an orthogonal linear transformation $T : V \to V$ with respect to an orthonormal basis for V. Then A satisfies $AA^T = A^T A = I$. A matrix A satisfying $AA^T = A^T A = I$ is called an orthogonal matrix.

Note that the rows of a matrix A are orthonormal if and only if $AA^T = I$ and the columns of a matrix A are orthonormal if and only if $A^T A = I$. Thus for an orthogonal matrix, the rows form an orthonormal set and the columns form an orthonormal set.

9.10.4 Exercises

Question 158. Choose the correct option(s).

- Option 1: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation, where $\mathbb{R}^2$ is the inner product space with respect to the dot product. Then $Tu \cdot Tv = u \cdot v$.

- **Option 2:** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be an orthogonal linear transformation, where $\mathbb{R}^2$ is the inner product space with respect to the dot product. Then $Tu \cdot Tv = u \cdot v$.

- Option 3: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation, where $\mathbb{R}^2$ is the inner product space with respect to the inner product given by $< a, b >= 2a_1b_1 + 5a_2b_2$. Then $\langle Tu, Tv \rangle = \langle u, v \rangle$.

- **Option 4:** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be an orthogonal linear transformation, where $\mathbb{R}^2$ is the inner product space with respect to the inner product given by $< a, b >= 2a_1b_1 + 5a_2b_2$. Then $\langle Tu, Tv \rangle = \langle u, v \rangle$.

Question 159. Which of the following option(s) is/are true?

- **Option 1:** Let A and B be orthogonal matrices of order 3. Then A and B are equivalent matrices.

- Option 2: Let A and B be orthogonal matrices. Then A and B are similar matrices.

- **Option 3:** Let A be a rotation matrix corresponding to the anti-clock wise rotation of the XY-plane about the Z-axis with angle θ. Then nullity of the matrix A is 0.

- **Option 4:** Let A be a rotation matrix corresponding to the anti-clock wise rotation of the XY-plane about the Z-axis with angle θ and B be a rotation matrix corresponding to the anti- clock wise rotation of the YZ-plane about the X -axis with angle β. Then A and B are equivalent matrices.

Question 160. If A is the matrix representation of an orthogonal transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, then choose the set of correct statements.

- Option 1: The column vectors of A are orthogonal but not orthonormal.
- **Option 2:** The column vectors of A are orthonormal.
- Option 3: The row vectors of A are orthogonal but not orthonormal.
- **Option 4:** The row vectors of A are orthonormal.

Question 161. Which of the following matrices is (are) orthogonal?

- **Option 1:** $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.
- **Option 2:** $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.
- Option 3: $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ -1 & -2 & 1 \end{bmatrix}$.
- **Option 4:** $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{bmatrix}$.