BITS Pilani
Pilani Campus

Mathematical Foundations for Data Science

MFDS Team

innovate achieve lead

BITS Pilani
Pilani Campus

DSECL ZC416, MFDS

Lecture No.2

Agenda

- Field
- Vector spaces and subspaces
- Linear independence and dependence
- Basis and dimensions
- Linear transformation

Field – Definition and Examples

Group : (G, '*') is a group if
i.* is closed
ii.* is associative
iii.* has an identity
iv.* has an inverse
Eg: < R , + > , < R, x >

< G, * > is **Abelian** if a * b = b * a ∀ a,b∈ G
Eg: < R, + > , < R, *> are Abelian

<F,+, .> is a **Field** if <F, +> and <F, .> are Abelian
Eg: < R, +, .> , < C, +, .>, < Q, +, .>

Vector Space

Real Vector Space

A nonempty set V of elements **a**, **b**, … is called a **real vector space** (or *real linear space*), and these elements are called **vectors** (regardless of their nature, which will come out from the context or will be left arbitrary) if, in V, there are defined two algebraic operations (called *vector addition* and *scalar multiplication*) as follows.

I. Vector addition associates with every pair of vectors **a** and **b** of V a unique vector of V, called the *sum* of **a** and **b** and denoted by $\mathbf{a} + \mathbf{b}$, such that the following axioms are satisfied.

Vector Space

Real Vector Space (continued 1)

I.1 *Commutativity.* For any two vectors **a** and **b** of V,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

I.2 *Associativity.* For any three vectors **a**, **b**, **c** of V,

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \qquad \text{(written } \mathbf{a} + \mathbf{b} + \mathbf{c}\text{)}.$$

I.3 There is a unique vector in V, called the *zero vector* and denoted by **0**, such that for every **a** in V,

$$\mathbf{a} + \mathbf{0} = \mathbf{a}.$$

I.4 For every **a** in V, there is a unique vector in V that is denoted by $-\mathbf{a}$ and is such that

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}.$$

Vector Space

Real Vector Space (continued 2)

II. Scalar multiplication. The real numbers are called **scalars**. Scalar multiplication associates with every $\mathbf{a}$ in V and every scalar c a unique vector of V, called the *product* of c and $\mathbf{a}$ and denoted by $c\mathbf{a}$ (or $\mathbf{a}c$) such that the following axioms are satisfied.

II.1 *Distributivity.* For every scalar c and vectors $\mathbf{a}$ and $\mathbf{b}$ in V,

$$c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}.$$

II.2 *Distributivity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$(c + k)\mathbf{a} = c\mathbf{a} + k\mathbf{a}.$$

II.3 *Associativity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$c(k\mathbf{a}) = (ck)\mathbf{a} \qquad \text{(written } ck\mathbf{a}\text{).}$$

II.4 For every $\mathbf{a}$ in V,

$$1\mathbf{a} = \mathbf{a}.$$

Subspace

By a **subspace** of a vector space *V* we mean

"a nonempty subset of *V* (including *V* itself) that forms a vector space with respect to the two algebraic operations (addition and scalar multiplication) defined for the vectors of *V*."

- Space $(W,+,\cdot)$:within a vector space
- $W \neq \Phi$ and $W \subseteq (V, +, .)$ over F is a subspace if
- $0 \in W$, $\alpha\, w_1 + w_2 \in W$
- Ex: $V = \{(x_1, x_2) \mid x_1, x_2 \in R\}$ over R, $W = \{(x_1, 0) \mid x_1 \in R\}$
- Set of singular matrices is not a subspace of M_{2x2}

Linear Dependence and Independence of Vectors

Given any set of m vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(m)}$ (with the same number of components), a **linear combination** of these vectors is an expression of the form

$$c_1\mathbf{a}_{(1)} + c_2\,\mathbf{a}_{(2)} + \dots + c_m\mathbf{a}_{(m)}$$

where $c_1, c_2, \dots, c_m$ are any scalars.

Now consider the equation

(1) $$c_1\mathbf{a}_{(1)} + c_2\,\mathbf{a}_{(2)} + \dots + c_m\mathbf{a}_{(m)} = \mathbf{0}$$

Clearly, *this vector equation (1) holds if we choose all c_j's zero, because then it becomes $\mathbf{0} = \mathbf{0}$.*

If this is the only m-tuple of scalars for which (1) holds, then our vectors $\mathbf{a}_{(1)} \dots, \mathbf{a}_{(m)}$ are said to form a *linearly independent set* or, more briefly, we call them **linearly independent**.

Linear Dependence and Independence of Vectors

Otherwise, if (1) also holds with scalars not all zero, we call these vectors **linearly dependent**.

This means that we can express at least one of the vectors as a linear combination of the other vectors. For instance, if (1) holds with, say, $c_1 \neq 0$, we can solve (1) for $\mathbf{a}_{(1)}$:

$$\mathbf{a}_{(1)} = k_2 \mathbf{a}_{(2)} + \ldots + k_m \mathbf{a}_{(m)} \text{ where } k_j = -c_j / c_1.$$

The **rank** of a matrix **A** is the maximum number of linearly independent row vectors of **A**.

It is denoted by rank **A**.

Linear Dependence and Independence of Vectors

Linear Independence and Dependence of Vectors

Consider p vectors that each have n components. Then these vectors are linearly independent if the matrix formed, with these vectors as row vectors, has rank p.

However, these vectors are linearly dependent if that matrix has rank less than p.

Linear Dependence of Vectors

Consider p vectors each having n components. If $n < p$, then these vectors *are linearly dependent.*

Linear Dependence and Independence

Let $S = \{v_1, v_2, \ldots v_n\} \in V$
Elements of S are LI if

$$\sum_{i=1}^{n} \alpha_i v_i = 0 \quad \Rightarrow \alpha_i = 0 \ \ \forall i \text{ is the only solution}$$

Elements of S are LD if

$$\sum_{i=1}^{n} \alpha_i v_i = 0 \quad \text{has atleast one non zero solution}$$

Eg: $V = R^n$ over R
LI and LD are related to rank

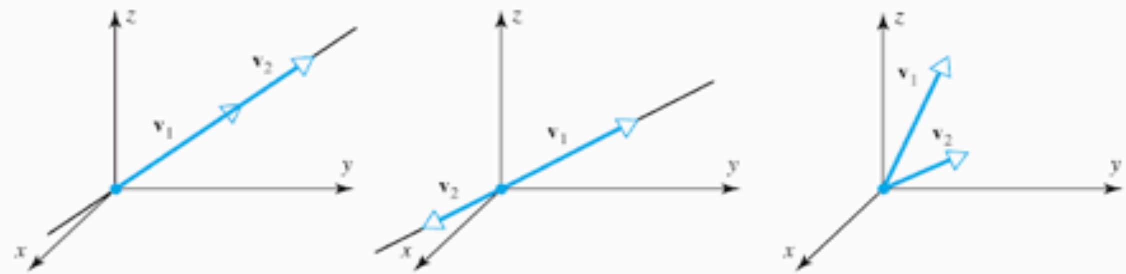
Linear Dependence and Independence

innovate achieve lead

A Geometric Interpretation of Linear Independence

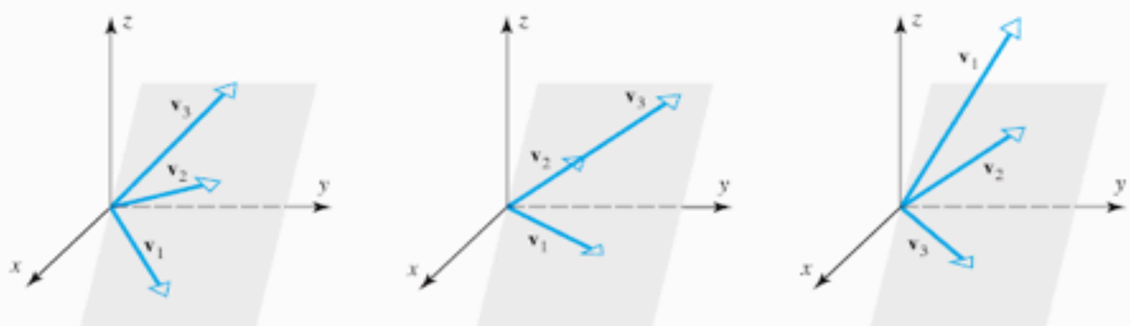
Linear independence has the following useful geometric interpretations in R^2 and R^3:

- Two vectors in R^2 or R^3 are linearly independent if and only if they do not lie on the same line when they have their initial points at the origin. Otherwise one would be a scalar multiple of the other (Figure 4.3.3).



▶ Figure 4.3.3 (*a*) Linearly dependent (*b*) Linearly dependent (*c*) Linearly independent

- Three vectors in R^3 are linearly independent if and only if they do not lie in the same plane when they have their initial points at the origin. Otherwise at least one would be a linear combination of the other two (Figure 4.3.4).



▶ Figure 4.3.4 (*a*) Linearly dependent (*b*) Linearly dependent (*c*) Linearly independent

Basis and Dimension

The maximum number of linearly independent vectors in V is called the **dimension** of V and is denoted by dim V.

A linearly independent set in V consisting of a maximum possible number of vectors in V is called a **basis** for V.
The number of vectors of a basis for V equals dim V.

The vector space $\boldsymbol{R}^n$ *consisting of all vectors with n components* (n *real numbers*) *has dimension n.*

- R over R → One dimensional vector space
- C over C → One dimensional vector space
- R over Q → Infinite dimensional

Basis and Dimension

The set of all linear combinations of given vectors $\mathbf{a}_{(1)}, \ldots,$ $\mathbf{a}_{(p)}$ with the same number of components is called the **span** of these vectors. Obviously, a span is a vector space. If in addition, the given vectors $\mathbf{a}_{(1)}, \ldots, \mathbf{a}_{(p)}$ are linearly independent, then they form a basis for that vector space.

This then leads to another equivalent definition of basis. A set of vectors is a **basis** for a vector space V if (1) the vectors in the set are linearly independent, and if (2) any vector in V can be expressed as a linear combination of the vectors in the set. If (2) holds, we also say that the set of vectors **spans** the vector space V.

Row Space and Column Space

- If A is an $m \times n$ matrix
  - the subspace of R^n spanned by the row vectors of A is called the row space of A
  - the subspace of R^m spanned by the column vectors is called the column space of A

The solution space of the homogeneous system of equation $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is called the nullspace of A.

$$A_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \cdots, \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Basis for Row Space and Column Space

If a matrix R is in row echelon form

- the row vectors with the leading 1's (i.e., the nonzero row vectors) form a basis for the row space of R
- the column vectors with the leading 1's of the row vectors form a basis for the column space of R

Basis for Row Space

Find a basis of row space of $A = \begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & 2 \end{bmatrix}$

Sol:

$$A = \begin{array}{c} \begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & 2 \end{bmatrix} \\ \begin{matrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{matrix} \end{array} \xrightarrow{\text{G.E.}} B = \begin{array}{c} \begin{array}{cc} \begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{matrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \\ \\ \\ \end{matrix} \end{array} \\ \begin{matrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 \end{matrix} \end{array}$$

a basis for $RS(A) =$ {the nonzero row vectors of B}

$= \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\} = \{(1, 3, 1, 3), (0, 1, 1, 0), (0, 0, 0, 1)\}$

Basis for Column Space

Find a basis for the column space of the matrix A.

$$\begin{array}{c} A = \begin{bmatrix} 1 & 2 & -1 & -2 & 0 \\ 2 & 4 & -1 & 1 & 0 \\ 3 & 6 & -1 & 4 & 1 \\ 0 & 0 & 1 & 5 & 0 \end{bmatrix} \\ \quad\; a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \end{array}$$

Reduce A to the reduced row- echelon form

$$E = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [e_1 \quad e_2 \quad e_3 \quad e_4 \quad e_5]$$

$e_2 = 2e_1 \rightarrow a_2 = 2a_1$

$e_4 = 3e_1 + 5e_3 \rightarrow a_4 = 3a_1 + 5a_3$

$\{a_1, a_3, a_5\}$ is a basis for column space of A

Solution Space/ Null Space

- Determine a basis and the dimension of the solution space of the homogeneous system

$$\begin{aligned} 2x_1 + 2x_2 - x_3 \qquad\quad + x_5 &= 0 \\ -x_1 + x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 \qquad\quad - x_5 &= 0 \\ x_3 + \ x_4 + x_5 &= 0 \end{aligned}$$

  - The general solution of the given system is

$$x_1 = -s-t,\ \ x_2 = s,$$

$$x_3 = -t,\ \ x_4 = 0,\ \ x_5 = t$$

  - Therefore, the solution vectors can be written as

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Solution Space/ Null Space

Find the solution space of a homogeneous system $A\mathbf{x} = \mathbf{0}$.

$$A = \begin{bmatrix} 1 & 2 & -2 & 1 \\ 3 & 6 & -5 & 4 \\ 1 & 2 & 0 & 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 & -2 & 1 \\ 3 & 6 & -5 & 4 \\ 1 & 2 & 0 & 3 \end{bmatrix} \xrightarrow{G.J.E} \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$x_1 = -2s - 3t,\ x_2 = s,\ x_3 = -t,\ x_4 = t$$

$$\Rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2s-3t \\ s \\ -t \\ t \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix} = s\mathbf{v}_1 + t\mathbf{v}_2$$

$$\Rightarrow NS(A) = \{s\mathbf{v}_1 + t\mathbf{v}_2 \mid s, t \in R\}$$

Rank of a matrix

Theorem : (Row and column space have equal dimensions)

- If A is an $m \times n$ matrix, then the row space and the column space of A have the same dimension.

$$\dim(RS(A)) = \dim(CS(A))$$

- **Rank:**

The dimension of the row (or column) space of a matrix A is called the rank of A.

$$\operatorname{rank}(A) = \dim(RS(A)) = \dim(CS(A))$$

Nullity of Matrix

- **Nullity:** The dimension of the nullspace of A is called the nullity of A **nullity(A) = dim(NS(A))**

- Notes: $\text{rank}(A^T) = \dim(RS(A^T)) = \dim(CS(A)) = \text{rank}(A)$

  Therefore $\text{rank}(A^T) = \text{rank}(A)$

- **Theorem :** (Dimension of the solution space)

  If A is an $m \times n$ matrix of rank r, then the dimension of the solution space of $A\mathbf{x} = \mathbf{0}$ is $n - r$. That is

$$\textbf{nullity}(A) = n - \textbf{rank}(A) = n - r$$

$$\boxed{n = \textbf{rank}(A) + \textbf{nullity}(A)}$$

Rank Nullity Theorem for Matrix

Rank and Nullity of Matrix

Notes: (n = #variables= #leading variables + #nonleading variables)

rank(A): The number of leading variables in the solution of Ax=0.

(The number of nonzero rows in the row-echelon form of A)

nullity (A): The number of free variables (non leading variables) in the solution of Ax = 0.

If A is an $m \times n$ matrix and rank(A) = r, then

Fundamental Space	Dimension
$\boldsymbol{RS(A)=CS(A^T)}$	$\boldsymbol{r}$
$CS(A)=RS(A^T)$	r
$\boldsymbol{NS(A)}$	$\boldsymbol{n-r}$
$NS(A^T)$	$m-r$

Rank and Nullity of Matrix

- Find the rank and nullity of the matrix

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

- Solution:
  - The reduced row-echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  - Since there are two nonzero rows, the row space and column space are both two-dimensional, so rank(A) = 2.

Rank and Nullity of Matrix

- The corresponding system of equations will be

$$x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 = 0$$

$$x_2 - 2x_3 - 12x_4 - 16\,x_5 + 5\,x_6 = 0$$

- It follows that the general solution of the system is

$$x_1 = 4r + 28s + 37t - 13u,\ x_2 = 2r + 12s + 16t - 5u,$$

$$x_3 = r,\ x_4 = s,\ x_5 = t,\ x_6 = u$$

or

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus, nullity$(A) = 4$

Linear Transformation

Let X and Y be any vector spaces. To each vector $\mathbf{x}$ in X we assign a unique vector $\mathbf{y}$ in Y. Then we say that a **mapping** (or **transformation** or **operator**) of X into Y is given.

Such a mapping is denoted by a capital letter, say F. The vector $\mathbf{y}$ in Y assigned to a vector $\mathbf{x}$ in X is called the **image** of $\mathbf{x}$ under F and is denoted by $F(x)$ [or $F\mathbf{x}$, without parentheses].

F is called a **linear mapping** or **linear transformation** if, for all vectors $\mathbf{v}$ and $\mathbf{x}$ in X and scalars c,

$$F(\mathbf{v}+\mathbf{x}) = F(\mathbf{v}) + F(\mathbf{x}) \tag{10}$$

$$F(c\mathbf{x}) = cF(\mathbf{x}).$$

Linear Transformation of Space R^n into Space R^m

From now on we let $X = R^n$ and $Y = R^m$. Then any real $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ gives a transformation of R^n into R^m,

(11) $$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

Since $\mathbf{A}(\mathbf{u} + \mathbf{x}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{x}$ and $\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x}$, this transformation is linear.

If $\mathbf{A}$ in (11) is square, $n \times n$, then (11) maps R^n into R^n. If this $\mathbf{A}$ is nonsingular, so that $\mathbf{A}^{-1}$ exists (see Sec. 7.8), then multiplication of (11) by $\mathbf{A}^{-1}$ from the left and use of $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ gives the **inverse transformation**

(14) $$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}.$$

It maps every $\mathbf{y} = \mathbf{y}_0$ onto that $\mathbf{x}$, which by (11) is mapped onto $\mathbf{y}_0$. *The inverse of a linear transformation is itself linear,* because it is given by a matrix, as (14) shows.

Range and Kernel

- Let T : V →W be linear transformation
- Range(T) is subspace of W
- Kernel(T) is subspace of V
- Nullity (T) = dim(Kernel(T))
- Rank(T) = dim(Range(T))
- Rank Nullity Theorem for Linear Transformation

dim(Kernel(T) + dim(Range(T)) = dim V

- Theorem with examples from matrices and linear transformations

Rank Nullity Theorem Example

$T : R^2 \rightarrow R^2$, $T(x) = Ax$ where $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ Find the rank and nullity of linear transformation and verify the Rank Nullity Theorem

Let $v = [x , y]$ be a vector of R^2, $v \in \ker(T)$

$Av = 0$

$$\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Augmented Matrix $\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & -1 & 0 \end{array}\right]$

$x = 0$, $y = 0 \rightarrow \ker T = [0, 0]$

Range Space of T is $R(T) = \text{col}(A)$

Rank Nullity Theorem Example

Range Space of T is R(T) = col (A)

$$= \left\{ x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} \,\middle|\, x_1, x_2 \in R \right\}$$

Nullity T = 0

Basis for R(T) = $\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$

Rank T = dim [R(T)] = 2

Rank T + nullity T = 2 + 0 = 2 = dim V

dim V = dim R^2 = 2

Hence Rank Nullity Theorem is verified