Introduction to Linear Algebra

Lecture Notes for the Course UMAT-302 at SSSIHL.

Reference: Linear Algebra by Larry Smith, Springer.

Compiled and Edited by: Hirak Doshi.

An Offering with Love and Gratitude at His Divine Lotus Feet.

Contents

1 Matrices **1**

1.1 Basic Operations 1

1.2 Special Types of Matrices 2

1.3 Determinants 3

2 Introduction to Vector Spaces **7**

2.1 Definitions and Examples 7

2.2 Linear Subspaces 8

2.3 Linear Span 9

2.4 Linear Dependence and Independence 10

3 Finite Dimensional Vector Spaces and Bases **13**

3.1 Definitions and Examples 13

3.2 Properties of Bases 13

3.3 Fundamental Results 15

4 Linear Transformations **18**

4.1 Definition and Examples 18

4.2 Properties of Linear Transformations 18

4.3 Kernels and Images 21

4.4 Isomorphisms 26

5 Representing Linear Transformations by Matrices **29**
5.1 Matrix Representation 29
5.2 Fundamental Results 30
5.3 Change of Basis 33

6 System of Linear Equations **35**
6.1 Existence Results 35
6.2 Affine Subspaces 36
6.3 Echelon Forms 39

7 Special Transformations **42**
7.1 Projection Maps 42
7.2 Nilpotent Transformations 43
7.3 Cyclic Transformations 44

8 The Theory of Eigen Values and Eigen Vectors **46**
8.1 Rank and Nullity 46
8.2 Eigen Values and Eigen Vectors 48
8.3 Characteristic Polynomial 50

1 Matrices

1.1 Basic Operations

1. **Matrix Addition:** Matrices of the same dimensions can be added by adding their corresponding entries.

   Example 1. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. The sum $A + B$ is calculated by adding corresponding entries:

   $$A + B = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}.$$

2. **Scalar Multiplication:** A matrix can be multiplied by a scalar by multiplying each entry of the matrix by the scalar.

   Example 2. Let $k = 2$ and $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. The scalar multiplication kA is calculated as:

   $$kA = 2 \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}.$$

3. **Matrix Transpose:** The transpose of a matrix A, denoted as A^T, is obtained by swapping its rows and columns.

   Example 3. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$. The transpose A^T is calculated by swapping rows and columns:

   $$A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

4. **Matrix Multiplication:** Matrix multiplication is defined as the product of two matrices $A = (a_{ij})$ and $B = (b_{ij})$, resulting in a matrix $C = (c_{ij})$, where

   $$c_{ij} = \sum_k a_{ik} b_{kj}.$$

   Example 4. Consider two matrices A and B where the number of columns in A

matches the number of rows in B:

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

The matrix product $C = AB$ is calculated as:

$$C = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

1.2 Special Types of Matrices

We look at some of the special types of matrices which will be useful in subsequent chapters.

1. **Identity Matrix (I):** The identity matrix is a square matrix with ones on the main diagonal and zeros elsewhere. It serves as a multiplicative identity in matrix multiplication.

   Example 5.
   $$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. **Scalar Matrix:** A scalar matrix is a diagonal matrix of the form $\mathbf{A} = e\mathbf{I}$ where $\mathbf{I}$ is the identity matrix.

   Example 6.
   $$A = 3 \cdot I = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}.$$

3. **Nilpotent Matrix:** A nilpotent matrix is a square matrix N such that $N^k = 0$ for some positive integer k. The integer k is the smallest such k called the index of nilpotence.

   Example 7.
   $$N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ is nilpotent as } N^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

4. **Idempotent Matrix:** An idempotent matrix is a square matrix P such that $P^2 = P$.

Example 8.

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

5. **Invertible Matrix (Non-Singular):** An invertible matrix, also known as a non-singular matrix, has an inverse A^{-1} such that $AA^{-1} = A^{-1}A = I$.

   Example 9.

   $$A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

6. **Triangular Matrix:** A triangular matrix is a matrix in which all entries above or below the main diagonal are zero. It can be upper triangular or lower triangular.

   Example 10 (Upper Triangular).

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

7. **Symmetric Matrix:** A symmetric matrix is a square matrix S such that $S^T = S$.

   Example 11.

   $$S = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}.$$

8. **Skew-Symmetric Matrix:** A skew-symmetric matrix is a square matrix K such that $K^T = -K$.

   Example 12.

   $$K = \begin{bmatrix} 0 & 2 & -3 \\ -2 & 0 & 4 \\ 3 & -4 & 0 \end{bmatrix}.$$

1.3 Determinants

Definition 1. Let $A = (a_{ij})$ be an $n \times n$ matrix. The (i, j) – **minor of** A denoted by M_{ij} is the matrix obtained by deleting the i^{th} row and j^{th} column.

Example 13. Let $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 3 \\ 2 & 0 & 7 \end{bmatrix}$. Then the minor matrices M_{13}, M_{22} of A are given as

$$M_{13} = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}, \quad M_{22} = \begin{bmatrix} 1 & 2 \\ 2 & 7 \end{bmatrix}.$$

Definition 2. The **determinant of** A denoted by $\det A$, is defined by the recursive formula:

$$\det A = a, \text{if } A = [a] \text{ is a } 1 \times 1 \text{ matrix}$$

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det M_{ij}, \text{if } A \text{ is a square matrix of size } n > 1.$$

Definition 3. The **cofactor of** a_{ij} is the number A_{ij} defined as

$$A_{ij} = (-1)^{i+j} \det M_{ij}.$$

Thus, we observe that

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} \text{ for each } i = 1, 2, \ldots$$

Example 14. Let $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 3 \\ 2 & 0 & 7 \end{bmatrix}$. Then, by the above definition:

$$\begin{aligned} \det A &= 1 \cdot \det \begin{bmatrix} 1 & 3 \\ 0 & 7 \end{bmatrix} - 0 \cdot \det \begin{bmatrix} -1 & 3 \\ 2 & 7 \end{bmatrix} + 2 \cdot \det \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix} \\ &= 1(1 \cdot 7 - 0 \cdot 3) - 0 + 2(-1 \cdot 0 - 1 \cdot 2) \\ &= 7 - 4 = 3. \end{aligned}$$

Theorem 1 (Properties of Determinants). Let A be a square matrix. Then the following properties hold:

1. **Linearity**:

$$\det(A+B) \neq \det(A) + \det(B).$$
$$\det(kA) = k^n \det(A).$$

2. **Product of Matrices**:

$$\det(AB) = \det(A)\det(B).$$

3. **Equality of Rows/Columns**:
   If any two rows/columns of A are equal, then $\det A = 0$.

4. **Transposition**:

$$\det(A^T) = \det(A), \text{ where } A^T \text{ is the transpose of the matrix } A.$$

5. **Row or Column Operations**:
   If B is obtained from A by swapping two rows, then $\det B = -\det A$.

6. **Triangular and Diagonal Matrices**:
   For a triangular matrix (upper or lower), the determinant is the product of the diagonal elements.

7. **Block Matrices**:

$$\det\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) = \det(A) \cdot \det(C), \text{ where } A, B, C \text{ are square matrices.}$$

Theorem 2.

$$\sum_j a_{ij} A_{kj} = 0 \text{ if } k \neq i.$$

Proof. If B is obtained from A by replacing the k^{th} row of A by the i^{th} row, then B has two equal rows and hence $\det B = 0$. □

Definition 4. The **cofactor matrix** of A, denoted by A^{cof}, is the matrix $A^{\text{cof}} = (a^*_{ij})$, where $a^*_{ij} = A_{ji}$.

Theorem 3.

$$A\,A^{\text{cof}} = (\det A)I.$$

Proof.
$$A\,A^{\mathrm{cof}} = \sum_j a_{ij} a^*_{jk} = \sum_j a_{ij} A_{kj} = \begin{cases} 0, & k \neq i \\ \det A, & k = i \end{cases} = \sum_j a_{ij} A_{ij} = (\det A) I. \qquad \square$$

Theorem 4. A matrix A is invertible if and only if $\det A \neq 0$.

Proof. Let A be invertible. Then $AA^{-1} = I$. Using the properties of determinants, we see that

$$\det AA^{-1} = \det A \det A^{-1} = \det I = 1.$$

Hence $\det A \neq 0$ and $\det A^{-1} = (\det A)^{-1}$.
Conversely, suppose that $\det A \neq 0$. Define the matrix B as

$$B = \frac{1}{\det A} A^{\mathrm{cof}}.$$

Then,

$$AB = \frac{1}{\det A} A\,A^{\mathrm{cof}} = \frac{1}{\det A} (\det A) I = I \text{ (by Theorem 3)}$$

and hence A is invertible. $\square$

2 Introduction to Vector Spaces

2.1 Definitions and Examples

A **vector space** $\mathcal{V}$ over a field $\mathbb{F}$ consists of the following:

1. A set of vectors $\mathcal{V}$.
2. Two binary operations:
   (a) **Vector addition**, denoted as $+$, which takes two vectors $\mathbf{A}$ and $\mathbf{B}$ in $\mathcal{V}$ and produces a vector $\mathbf{A}+\mathbf{B}$ in $\mathcal{V}$.
   (b) **Scalar multiplication**, denoted as $\cdot$, which takes a scalar α from the field $\mathbb{F}$ and a vector $\mathbf{A}$ in $\mathcal{V}$ and produces a vector $\alpha \cdot \mathbf{A}$ in $\mathcal{V}$.

A vector space $\mathcal{V}$ must satisfy the following axioms for all vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ in $\mathcal{V}$ and all scalars α and β in $\mathbb{F}$:

1. **Vector Addition Axioms**:
   (a) **Closure under Addition**: $\mathbf{A}+\mathbf{B}$ is in $\mathcal{V}$.
   (b) **Commutativity**: $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$.
   (c) **Associativity**: $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$.
   (d) **Identity Element**: There exists a vector $\mathbf{0}$ (called the zero vector) such that $\mathbf{A}+\mathbf{0} = \mathbf{A}$ for all $\mathbf{A}$ in $\mathcal{V}$.
   (e) **Inverse Elements**: For every vector $\mathbf{A}$ in $\mathcal{V}$, there exists a vector $-\mathbf{A}$ such that $\mathbf{A}+(-\mathbf{A}) = \mathbf{0}$.
2. **Scalar Multiplication Axioms**:
   (a) **Closure under Scalar Multiplication**: $\alpha \cdot \mathbf{A}$ is in $\mathcal{V}$.
   (b) **Compatibility with Field Multiplication**: $(\alpha\beta) \cdot \mathbf{A} = \alpha \cdot (\beta \cdot \mathbf{A})$.
   (c) **Identity Element**: There exists a scalar 1 such that $1 \cdot \mathbf{A} = \mathbf{A}$ for all $\mathbf{A}$ in $\mathcal{V}$.
   (d) **Distributivity of Scalars over Vector Addition**: $\alpha \cdot (\mathbf{A}+\mathbf{B}) = \alpha \cdot \mathbf{A} + \alpha \cdot \mathbf{B}$.
   (e) **Distributivity of Scalars over Field Addition**: $(\alpha+\beta) \cdot \mathbf{A} = \alpha \cdot \mathbf{A} + \beta \cdot \mathbf{A}$.

Here are some examples of vector spaces:

1. **Euclidean Space** $\mathbb{R}^n$: The set of all n-dimensional real vectors with vector addition and scalar multiplication defined as component-wise operations.

2. **Polynomial Space** $P_n(\mathbb{R})$: The set of all polynomials of degree at most n with real or complex coefficients, where vector addition is polynomial addition and scalar multiplication is scalar multiplication of the coefficients.

3. **Function Space** $C([a, b])$: The set of all continuous real-valued functions defined on the closed interval $[a, b]$, with vector addition defined as function addition and scalar multiplication defined as scalar multiplication of the function values.

4. **Matrix Space** $M_{m\times n}(\mathbb{F})$: The set of all $m \times n$ matrices with entries from the field $\mathbb{F}$, where vector addition is matrix addition and scalar multiplication is scalar multiplication of the matrix entries.

2.2 Linear Subspaces

Definition 5. A subset $\mathcal{U}$ of a vector space $\mathcal{V}$ is called a **linear subspace** if it is itself a vector space with respect to the vector addition and scalar multiplication operations defined in $\mathcal{V}$. In other words, $\mathcal{U}$ is a linear subspace of $\mathcal{V}$ if it satisfies the following properties:

1. **Closure under Vector Addition**: For all vectors $\mathbf{A}$ and $\mathbf{B}$ in $\mathcal{U}$, $\mathbf{A} + \mathbf{B}$ is also in $\mathcal{U}$.

2. **Closure under Scalar Multiplication**: For all vectors $\mathbf{A}$ in $\mathcal{U}$ and scalars α, $\alpha \cdot \mathbf{A}$ is also in $\mathcal{U}$.

3. **Containing the Zero Vector**: The zero vector $\mathbf{0}$ of $\mathcal{V}$ is in $\mathcal{U}$.

In other words, $\mathcal{U}$ is a linear subspace if it is a vector space in its own right. Here are some examples of linear subspaces:

Example 15. The Trivial Subspace: The set containing only the zero vector, $\{\mathbf{0}\}$, is a linear subspace of any vector space.

Example 16. The Whole Space $\mathcal{V}$: The entire vector space $\mathcal{V}$ itself is a linear subspace of itself.

Example 17. Let $\mathcal{V} = \mathbb{R}^3$. Then the set $\mathcal{U} = \Big\{(0, 0, z) : z \in \mathbb{R}\Big\}$ is a linear subspace of $\mathcal{V}$.

2.3 Linear Span

Definition 6. Let $\mathcal{V}$ be a vector space and $\mathbf{A}_1, \dots, \mathbf{A}_k$ be vectors in $\mathcal{V}$. Then, the **linear span** of $\mathbf{A}_1, \dots, \mathbf{A}_k$, denoted by $\mathcal{L}(\mathbf{A}_1, \dots, \mathbf{A}_k)$, is the set of all vectors in $\mathcal{V}$ which are linear combinations of $\mathbf{A}_1, \dots, \mathbf{A}_k$, i.e.

$$\mathcal{L}(\mathbf{A}_1, \dots, \mathbf{A}_k) = \left\{ \sum_{i=1}^{k} a_i \mathbf{A}_i : a_i \in \mathbb{R} \right\}.$$

Theorem 5. Let $E = \{\mathbf{A}_1, \dots, \mathbf{A}_k\}$ be a set of vectors in $\mathcal{V}$. Then $\mathcal{L}(E)$ is a linear subspace of $\mathcal{V}$.

Proof. Let $\mathbf{A}, \mathbf{B} \in \mathcal{L}(E)$. Then,

$$\begin{aligned} \alpha\mathbf{A} + \beta\mathbf{B} &= \alpha \sum_{i=1}^{k} a_i \mathbf{A}_i + \beta \sum_{i=1}^{k} b_i \mathbf{A}_i, \ \ a_i, b_i \in \mathbb{R} \\ &= \sum_{i=1}^{k} (\alpha a_i + \beta b_i) \mathbf{A}_i \\ &= \sum_{i=1}^{k} d_i \mathbf{A}_i \in \mathcal{L}(E), \end{aligned}$$

where $d_i = \alpha a_i + \beta b_i$. □

Theorem 6. Let $E \subseteq \mathcal{V}$. Then $E = \mathcal{L}(E)$ if and only if E is a linear subspace of $\mathcal{V}$.

Proof. Let E be a linear subspace of $\mathcal{V}$. If $A_i \in E, i = 1, \dots, k$, then $\sum_{i=1}^{k} a_i A_i \in E$ since E is closed under vector addition and scalar multiplication so that $\mathcal{L}(E) \subseteq E$. By definition, $E \subseteq \mathcal{L}(E)$. Hence $E = \mathcal{L}(E)$.
Conversely, if $E = \mathcal{L}(E)$ then by Theorem 5, $\mathcal{L}(E)$ is a subspace of $\mathcal{V}$ and hence, E is a subspace of $\mathcal{V}$. □

Theorem 7. Let S and T be two subspaces of $\mathcal{V}$. Then

1. $S \cap T$ is a linear subspace of $\mathcal{V}$.

2. $S + T$ is a linear subspace of $\mathcal{V}$.

3. $S \cup T$ is a linear subspace of $\mathcal{V}$ if and only if $S \subseteq T$ or $T \subseteq S$.

Proof. Let S and T be two subspaces of $\mathcal{V}$.

1. Let $\mathbf{A}, \mathbf{B} \in S \cap T$. Then $\mathbf{A}, \mathbf{B} \in S$ and $\mathbf{A}, \mathbf{B} \in S \cap T$. Therefore $a\mathbf{A} + b\mathbf{B} \in S$ and $a\mathbf{A} + b\mathbf{B} \in T$. Hence $a\mathbf{A} + b\mathbf{B} \in S \cap T$.

2. Let $\mathbf{A}, \mathbf{B} \in S + T$. Then $\mathbf{A} = \mathbf{X}_1 + \mathbf{Y}_1, \mathbf{B} = \mathbf{X}_2 + \mathbf{Y}_2, X_i \in S, Y_i \in T$. Therefore $a\mathbf{A} + b\mathbf{B} = a\mathbf{X}_1 + b\mathbf{X}_2 + a\mathbf{Y}_1 + b\mathbf{Y}_2$. Note that $a\mathbf{X}_1 + b\mathbf{X}_2 \in S$ and $a\mathbf{Y}_1 + b\mathbf{Y}_2 \in T$. Hence $a\mathbf{A} + b\mathbf{B} \in S + T$.

3. Let $S \subseteq T$ or $T \subseteq S$. Then $S \cup T = T$ or $S \cup T = S$ respectively. □

2.4 Linear Dependence and Independence

Definition 7. A set of distinct vectors $\mathbf{A}_1, \ldots, \mathbf{A}_k$ is **linearly dependent** if there exists scalars $a_1, \ldots, a_k$ not all zero such that

$$a_1\mathbf{A}_1 + \cdots + a_k\mathbf{A}_k = 0.$$

Definition 8. A set of distinct vectors $\mathbf{A}_1, \ldots, \mathbf{A}_k$ is **linearly independent** if the relation

$$a_1\mathbf{A}_1 + \cdots + a_k\mathbf{A}_k = 0$$

implies $a_1 = \cdots = a_k = 0$.

Example 18. The set $\left\{(1,0,0), (0,1,0), (1,1,0)\right\}$ is linearly dependent since

$$(1,0,0) + (0,1,0) - (1,1,0) = 0.$$

Here $a_1 = a_2 = 1, a_3 = -1$.

Example 19. The set $\left\{1 + x, 1 - x\right\}$ is linearly independent in $P_1(\mathbb{R})$ since

$$a_1(1 + x) + a_2(1 - x) = 0$$

gives $a_1 + a_2 = 0$ and $a_1 - a_2 = 0$ whence we conclude $a_1 = a_2 = 0$.

Theorem 8. If a set contains $\mathbf{0}$, then it is linearly dependent.

Proof. The relation $1 \cdot \mathbf{0} = \mathbf{0}$ is a linear relation with a nonzero scalar 1. □

Theorem 9. Let $E = \{\mathbf{A}_1, \ldots, \mathbf{A}_k\}$ be a finite set of vectors. Then E is linearly dependent if and only if there is a vector $\mathbf{A}_i$ which is linearly dependent on the set $E \setminus \{\mathbf{A}_i\}$.

Proof. Suppose that there is a vector in E which is linearly dependent on the remaining vectors. Without any loss in generality, we may assume this vector to be $\mathbf{A}_1$, i.e. $\mathbf{A}_1 \in \mathcal{L}(E \setminus \{\mathbf{A}_1\})$. Then

$$\mathbf{A}_1 = a_2\mathbf{A}_2 + \ldots a_k\mathbf{A}_k.$$

But then

$$(-1)\mathbf{A}_1 + a_2\mathbf{A}_2 + \ldots a_k\mathbf{A}_k = 0$$

is a linear relation with a non-zero scalar -1. Hence E is linearly dependent.
On the other hand, let E be a linearly dependent set. Then, there are scalars, $a_1, \ldots, a_k$ not all zero such that, there is a linear relation

$$a_1\mathbf{A}_1 + \cdots + a_k\mathbf{A}_k = 0.$$

By rearranging, without any loss in generality we may choose $a_1 \neq 0$. Then,

$$\mathbf{A}_1 = -\frac{a_2}{a_1}\mathbf{A}_2 - \frac{a_3}{a_1}\mathbf{A}_3 - \cdots - \frac{a_k}{a_1}\mathbf{A}_k.$$

which shows that $\mathbf{A}_1 \in \mathcal{L}(E \setminus \{\mathbf{A}_1\})$. This proves the theorem. □

Theorem 10. If E is a finite set such that $\mathcal{L}(E) = \mathcal{U}$ where $\mathcal{U}$ is a linear subspace of $\mathcal{V}$. Then, there exists a linearly independent subset F of E such that $\mathcal{L}(F) = \mathcal{U} = \mathcal{L}(E)$.

Proof. If E is linearly dependent, then set $F = E$ and the proof is complete. Otherwise, suppose that it is linearly dependent. Then by Theorem 9 there is a vector $\mathbf{A}$ which is linearly dependent on the remaining vectors. Let $E_1 = E \setminus \mathbf{A}$. Then $\mathbf{A} \in E_1$. By construction, $\mathcal{L}(E_1) \subseteq \mathcal{L}(E)$. But

$$\mathcal{L}(E) = \mathcal{L}(E_1 \cup \{\mathbf{A}\}) \subseteq \mathcal{L}(\mathbf{A}) + \mathcal{L}(E_1) \subseteq \mathcal{L}(E_1) + \mathcal{L}(E_1) = \mathcal{L}(E_1).$$

Hence $\mathcal{L}(E_1) = \mathcal{L}(E)$. If $\mathcal{L}(E_1)$ is linearly independent then set $F = E_1$ and the proof is complete, otherwise continue the process. Since E is a finite set, the process terminates after finite number of times. □

Theorem 11. Let $\mathcal{V}$ be a vector space and $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3 \in \mathcal{V}$ such that $\mathbf{A}_3 \notin \mathcal{L}(\mathbf{A}_1, \mathbf{A}_2)$. Then $\mathbf{A}_1, \mathbf{A}_2$ are linearly independent if and only if $\mathbf{A}_1 + \mathbf{A}_3, \mathbf{A}_2 + \mathbf{A}_3$ are linearly independent.

Proof. Suppose that $\mathbf{A}_1, \mathbf{A}_2$ are linearly independent. Consider the relation,

$$a_1(\mathbf{A}_1 + \mathbf{A}_3) + a_2(\mathbf{A}_2 + \mathbf{A}_3) = 0. \tag{1}$$

Rearranging, we obtain

$$a_1\mathbf{A}_1 + a_2\mathbf{A}_2 + (a_1 + a_2)\mathbf{A}_3 = 0. \tag{2}$$

Since $\mathbf{A}_3 \notin \mathcal{L}(\mathbf{A}_1, \mathbf{A}_2)$, we must have $a_1 + a_2 = 0$, otherwise we would get

$$\mathbf{A}_3 = -\frac{a_1}{a_1 + a_2}\mathbf{A}_1 - \frac{a_2}{a_1 + a_2}\mathbf{A}_2.$$

Hence $a_2 = -a_1$. Therefore,

$$a_1(\mathbf{A}_1 - \mathbf{A}_2) = 0.$$

But since $\mathbf{A}_1, \mathbf{A}_2$ are linearly independent, we must have $a_1 = 0$ and hence $a_2 = 0$. Conversely, suppose that $\mathbf{A}_1 + \mathbf{A}_3, \mathbf{A}_2 + \mathbf{A}_3$ are linearly independent. Then from (2), we have

$$a_1\mathbf{A}_1 + a_2\mathbf{A}_2 = 0 \text{ and } a_1 + a_2 = 0.$$

The second relation holds since $\mathbf{A}_3 \notin \mathcal{L}(\mathbf{A}_1, \mathbf{A}_2)$. But then from (1), we have $a_1 = a_2 = 0$ since $\mathbf{A}_1 + \mathbf{A}_3, \mathbf{A}_2 + \mathbf{A}_3$ are linearly independent. Hence $\mathbf{A}_1, \mathbf{A}_2$ are linearly independent. □

3 Finite Dimensional Vector Spaces and Bases

3.1 Definitions and Examples

Definition 9. A vector space $\mathcal{V}$ is **finite-dimensional** if there exists a finite set of vectors E such that $\mathcal{L}(E) = \mathcal{V}$.

Example 20. The vector space $\mathbb{R}^n$ is finite-dimensional as $\mathbb{R}^n = \mathcal{L}(\mathbf{E}_1, \ldots, \mathbf{E}_n)$ where $\mathbf{E}_i = (0, \ldots, 1, \ldots, 0)$ has a 1 in the i^{th} place.

Example 21. The vector space $P_n(\mathbb{R})$ is finite-dimensional as $P_n(\mathbb{R}) = \mathcal{L}(1, x, \ldots, x^n)$, i.e. the polynomials $1, x, x^2, \ldots, x^n$ span the space $P_n(\mathbb{R})$.

Definition 10. A set of vectors E is a **basis for** $\mathcal{V}$ if E is linearly independent and $\mathcal{L}(E) = \mathcal{V}$.

Example 22. The set $\{\mathbf{E}_1, \ldots, \mathbf{E}_n\}$ is a basis for $\mathbb{R}^n$ since the vectors $\mathbf{E}_1, \ldots, \mathbf{E}_n$ are linearly independent in $\mathbb{R}^n$ and also span $\mathbb{R}^n$ as seen in Example 20.

Example 23. The set $\{1, x, \ldots, x^n\}$ is a basis for $P_n(\mathbb{R})$ since the polynomials $1, x, x^2, \ldots, x^n$ are linearly independent in $P_n(\mathbb{R})$ and also span $P_n(\mathbb{R})$ as seen in Example 21.

Example 24. The set $\Big\{1, x-1, (x-2)(x-1)\Big\}$ is a basis for $P_2(\mathbb{R})$. Suppose there are scalars a_1, a_2, a_3 such that

$$a_1 + a_2(x-1) + a_3(x-2)(x-1) = 0.$$

Then, we obtain $a_3 = 0, a_2 - 3a_3 = 0$ and $a_1 - a_2 + 2a_3 = 0$ which gives $a_1 = a_2 = a_3 = 0$. Hence, they are linearly independent. To check linear span, observe that if $p(x) = a_0 + a_1x + a_2x^2$ is an arbitrary polynomial in $P_2(\mathbb{R})$, then we want b_0, b_1, b_2 such that

$$a_0 + a_1x + a_2x^2 = b_0 + b_1(x-1) + b_2(x-2)(x-1).$$

Solving for b_i, we obtain $b_0 = a_0 + a_1 + a2, b_1 = a_1 + 3a_2, b_2 = a_2$. Hence, $\Big\{1, x-1, (x-2)(x-1)\Big\}$ is the spanning set.

3.2 Properties of Bases

Theorem 12. Let $E = \{\mathbf{A}_1, \ldots, \mathbf{A}_k\}$ be a finite set of vectors. Then E is linearly dependent if and only if

$$\mathbf{A}_m \in \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_{m-1})$$

for some $m \le k$.

Proof. Suppose that $\mathbf{A}_m \in \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_{m-1})$ for some $m \le k$. Then $\mathbf{A}_m$ is linearly dependent on the remaining vectors of E. Hence E is linearly dependent.
On the other hand, let E be a linearly dependent set. Then, there are scalars, $a_1, \ldots, a_k$ not all zero such that, there is a linear relation

$$a_1\mathbf{A}_1 + \cdots + a_k\mathbf{A}_k = 0.$$

Let $m = \max\{1, \ldots, k\}$ for which $a_m \neq 0$. By this construction, we ensure that $a_{m+1} = a_{m+2} = \cdots = 0$ so that the linear relation may be written as

$$a_1\mathbf{A}_1 + \cdots + a_k\mathbf{A}_m = 0.$$

But then,

$$\mathbf{A}_m = -\frac{a_1}{a_m}\mathbf{A}_1 - \frac{a_1}{a_m}\mathbf{A}_2 - \cdots - \frac{a_{m-1}}{a_m}\mathbf{A}_{m-1}.$$

which shows that $\mathbf{A}_m \in \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_{m-1})$. This proves the theorem. □

Theorem 13. Let $\mathcal{V}$ be a finite-dimensional vector space and $E = \{\mathbf{A}_1, \ldots, \mathbf{A}_k\}$ be a finite set of vectors. If F is linearly independent set of vectors in $\mathcal{L}(E)$, then F is finite and the number of elements in F is atmost k.

Proof. Let $H = \{\mathbf{B}_1, \ldots, \mathbf{B}_k\}$ be a finite subset of F. Then, H is linearly independent. Since V is finite-dimensional, $\mathcal{L}(E) = \mathcal{V}$ and hence F is finite. Consider the set

$$G_1 = \{\mathbf{B}_s, \mathbf{A}_1, \ldots, \mathbf{A}_k\}.$$

Since H is a linearly independent subset of $\mathcal{L}(E)$, we have $\mathbf{B}_s \in \mathcal{L}(E)$ and hence G_1 is linearly dependent. Therefore by Theorem 12, there is a vector in G_1 which is linearly dependent on the preceeding vectors. This vector cannot be $\mathbf{B}_s$, since $\mathbf{B}_s$ is a linearly independent vector and so $\mathbf{B}_s \neq \mathbf{0}$. So it has to be an $A_i, i = 1, \ldots, k$. Without any loss in generality, we may assume this vector to be $\mathbf{A}_k$, i.e.

$$\mathbf{A}_k \in \mathcal{L}(\mathbf{B}_s, \mathbf{A}_1, \ldots, \mathbf{A}_{k-1}). \tag{3}$$

Define $E_1 = \{\mathbf{B}_s, \mathbf{A}_1, \ldots, \mathbf{A}_{k-1}\}$. From (3) we see that,

$$\mathbf{A}_k = a_1\mathbf{A}_1 + \cdots + a_{k-1}\mathbf{A}_{k-1} + b_s\mathbf{B}_s. \tag{4}$$

But since $\mathbf{B}_s \in \mathcal{L}(E)$, we have,

$$\mathbf{B}_s = c_1\mathbf{A}_1 + \cdots + c_k\mathbf{A}_k. \tag{5}$$

Substituting (5) in (4), we get,

$$\mathbf{A}_k = a_1\mathbf{A}_1 + \cdots + a_{k-1}\mathbf{A}_{k-1} + b_s(c_1\mathbf{A}_1 + \cdots + c_k\mathbf{A}_k).$$

Rearranging the terms, we obtain,

$$\mathbf{A}_k = (a_1 + b_s c_1)\mathbf{A}_1 + \cdots + (a_{k-1} + b_s c_{k-1})\mathbf{A}_{k-1} + b_s c_k \mathbf{A}_k$$

which shows that $\mathbf{A}_k \in \mathcal{L}(E)$. Thus, we conclude $\mathcal{L}(E_1) = \mathcal{L}(E)$. Now consider,

$$G_2 = \{\mathbf{B}_{s-1}, \mathbf{B}_s, \mathbf{A}_1, \ldots, \mathbf{A}_{k-1}\}.$$

As before, there is a vector in G_2 which is linearly dependent on the preceeding vectors. This vector cannot be $\mathbf{B}_{s-1}$ since $\mathbf{B}_{s-1}$ is linearly independent, i.e. $\mathbf{B}_{s-1} \neq \mathbf{0}$. ALso it cannot be $\mathbf{B}_s$ since $\{\mathbf{B}_{s-1}, \mathbf{B}_s\}$ is linearly independent. Therefore, it must be an $\mathbf{A}_i$. As before we may choose this vector to be $\mathbf{A}_{k-1}$, i.e.

$$\mathbf{A}_{k-1} \in \mathcal{L}(\mathbf{B}_{s-1}, \mathbf{B}_s, \mathbf{A}_1, \ldots, \mathbf{A}_{k-2})$$

and as before, set $E_2 = \{\mathbf{B}_{s-1}, \mathbf{B}_s, \mathbf{A}_1, \ldots, \mathbf{A}_{k-2}\}$ to obtain $\mathcal{L}(E_2) = \mathcal{L}(E)$. We want to show $s \leq k$. Assume on the contrary that $s > k$. Then by repeating the above arguments k times, we can construct a set $E_k = \{\mathbf{B}_{s-(k-1)}, \mathbf{B}_{s-(k-2)}, \ldots, \mathbf{B}_s\}$ such that $\mathcal{L}(E_k) = \mathcal{L}(E)$. But then,

$$\mathbf{B}_{s-k} \in \mathcal{L}(E) = \mathcal{L}(E_k)$$

which shows that $H = \{\mathbf{B}_1, \ldots, \mathbf{B}_k\}$ is linearly dependent which is a contradiction. Hence $s \leq k$ as required. □

3.3 Fundamental Results

Theorem 14. A vector space $\mathcal{V}$ is finite-dimensional if and only if every linearly independent set of vectors in $\mathcal{V}$ is finite.

Proof. If $\mathcal{V}$ is finite-dimensional, then every linearly independent set of vectors is finite by Theorem 13. On the other hand, suppose that every linearly independent set of vectors is finite but $\mathcal{V}$ is not finite-dimensional. Then, $\mathcal{V}$ is not spanned by any finite set of vectors in $\mathcal{V}$. Let $\mathbf{0} \neq \mathbf{A}_1 \in \mathcal{V}$. Then $\{\mathbf{A}_1\}$ is a linearly independent set in $\mathcal{V}$. Since $\mathcal{V}$ is not finite-dimensional $\mathcal{L}(\mathbf{A}_1) \neq \mathcal{V}$. Choose a vector $\mathbf{A}_2 \in \mathcal{V}$ such that $\mathbf{A}_2 \notin \mathcal{L}(\mathbf{A}_1)$. Then the set $\{\mathbf{A}_1, \mathbf{A}_2\}$ is linearly independent. Again since $\mathcal{V}$ is not finite-dimensional, $\mathcal{L}(\mathbf{A}_1, \mathbf{A}_2) \neq \mathcal{V}$. But continuing this process, we can find $\mathbf{A}_{i+1} \notin \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_i)$ and an infinite set $\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_i, \ldots\}$ which is linearly independent which contradicts our hypothesis that every linearly independent set of vectors is finite. Hence $\mathcal{V}$ is finite-dimensional. $\square$

Definition 11. Let $\mathcal{V}$ be a finite-dimensional vector space. The **dimension of $\mathcal{V}$**, denoted by $\dim \mathcal{V}$ is the number of elements in the basis of $\mathcal{V}$.

Example 25. $\dim \mathbb{R}^n = n$ as there are n vectors $(1, \ldots, 0), \ldots, (0, \ldots, 1)$ in the basis.

Example 26. $\dim P_n(\mathbb{R}) = n + 1$ as there are $n + 1$ vectors $1, x, x^2, \ldots, x^n$ in the basis.

Theorem 15 (Basis Extension Theorem). Let $\mathcal{V}$ be a finite-dimensional vector space and $\mathbf{A}_1, \ldots, \mathbf{A}_m$ be linearly independent vectors in $\mathcal{V}$. Then there exist $n = \dim \mathcal{V} - m$ vectors $\mathbf{B}_1, \ldots, \mathbf{B}_n$ in $\mathcal{V}$ such that $\{\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1, \ldots, \mathbf{B}_n\}$ is a basis for $\mathcal{V}$.

Proof. If $\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m) = \mathcal{V}$, then $m = \dim \mathcal{V}$ and the proof is complete. So suppose, $\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m) \neq \mathcal{V}$. Choose a vector $\mathbf{B}_1 \in \mathcal{V}$ such that $\mathbf{B}_1 \notin \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m)$. Then the set

$$\{\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1\}$$

is a linearly independent set since $\mathbf{B}_1$ does not depend on the preceeding vectors by construction (see Theorem 12).

If $\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1) = \mathcal{V}$, then $m = \dim \mathcal{V} - 1$ and the proof is complete. So assume $\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1) \neq \mathcal{V}$. Then, we can find $\mathbf{B}_2 \in \mathcal{V}$ such that $\mathbf{B}_2 \notin \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1)$. Then the set

$$\{\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1, \mathbf{B}_2\}$$

is linearly independent. Continue this process k times such that $m + k = \dim \mathcal{V}$ to get a linearly independent set

$$\{\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1, \ldots, \mathbf{B}_k\}.$$

In this case $k = n$ and

$$\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_m, \mathbf{B}_1, \ldots, \mathbf{B}_n) = \mathcal{V}.$$

as required. This completes the proof. $\square$

The Basis Extension Theorem is a fundamental result whose implications are far-reaching. Some of the immediate consequences are listed below.

Theorem 16. Let $\mathcal{V}$ be a finite-dimensional vector space and $\mathcal{U}$ be a linear subspace of $\mathcal{V}$. Then $\mathcal{U}$ is finite-dimensional and $\dim \mathcal{U} \leq \dim \mathcal{V}$.

Theorem 17. Let $\mathcal{V}$ be a finite-dimensional vector space with $\dim \mathcal{V} = n$. If the vectors $\mathbf{A}_1, \ldots, \mathbf{A}_n$ are linearly indepedent in $\mathcal{V}$, then they are a basis.

Proof. If the vectors $\mathbf{A}_1, \ldots, \mathbf{A}_n$ are not a basis, then by Theorem 15, there exists $\mathbf{B}_1, \ldots, \mathbf{B}_m$ vectors such that $\{\mathbf{A}_1, \ldots, \mathbf{A}_n, \mathbf{B}_1, \ldots, \mathbf{B}_m\}$ would be a basis for $\mathcal{V}$. But then $\dim \mathcal{V} = m + n \neq n$ which is a contradiction. $\square$

Theorem 18. Let $\mathcal{V}$ be a finite-dimensional vector space with $\dim \mathcal{V} = n$. If the vectors $\mathbf{A}_1, \ldots, \mathbf{A}_n$ span $\mathcal{V}$, then they are a basis.

Theorem 19. Let $\mathcal{V}$ be a finite-dimensional vector space and $\mathcal{U}$ be a linear subspace of $\mathcal{V}$ with $\dim \mathcal{U} = n$. Then $\mathcal{U} = \mathcal{V}$.

Proof. Since $\dim \mathcal{U} = n$, there is a basis $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ for $\mathcal{U}$. By Theorem 15,

$$\mathcal{U} = \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_n) = \mathcal{V}. \qquad \square$$

Example 27. Let $\mathcal{V} = \mathbb{R}^3$ and $A_1 = (1, 2, 3)$. Using Theorem 15, we will extend the vector A_1 to form a basis for $\mathcal{V}$. Observe that $\mathcal{L}(A_1) \neq \mathcal{V}$. Choose $A_2 = (1, 0, 0) \in \mathcal{V}$. Since $A_2 \notin \mathcal{L}(A_1)$, we conclude that the set $\{A_1, A_2\}$ is linearly independent. Clearly $\mathcal{L}(A_1, A_2) \neq \mathcal{V}$. For example, $(0, 1, 0) \in \mathcal{V}$ but not in $\mathcal{L}(A_1, A_2)$. Consider $A_3 = (0, 1, 0) \in \mathcal{V}$. Then $\{A_1, A_2, A_3\}$ is linearly independent. Therefore by Theorem 17, $\{A_1, A_2, A_3\}$ is a basis for $\mathcal{V}$.

4 Linear Transformations

4.1 Definition and Examples

A **linear transformation** is a function that preserves vector addition and scalar multiplication. Formally, a function T from vector space $\mathcal{V}$ to vector space $\mathcal{W}$ is a linear transformation if it satisfies two properties:

1. **Additivity**: For all vectors $\mathbf{A}, \mathbf{B} \in \mathcal{V}$,

$$T(\mathbf{A} + \mathbf{B}) = T(\mathbf{A}) + T(\mathbf{B}).$$

2. **Scalar Multiplication**: For all vectors $\mathbf{A} \in \mathcal{V}$ and scalars c,

$$T(c\mathbf{A}) = cT(\mathbf{A}).$$

Example 28 (Dilation in 2D Space). Consider the 2D vector space $\mathbb{R}^2$. The transformation T that scales each vector by a constant factor k is a linear transformation.

$$T(\mathbf{A}) = k\mathbf{A}.$$

Example 29 (Projection onto a Line). In 2D space, consider the line L spanned by the vector $\mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The transformation T that projects any vector $\mathbf{A} \in \mathbb{R}^2$ onto L is a linear transformation.

$$T(\mathbf{A}) = \frac{\mathbf{A} \cdot \mathbf{B}}{||\mathbf{B}||^2}\mathbf{B}.$$

This transformation projects vectors onto the line L defined by $\mathbf{B}$.

Example 30 (Rotation in 2D Space). The transformation T that rotates vectors in 2D space counterclockwise by an angle θ is a linear transformation.

$$T(\mathbf{A}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\mathbf{A}.$$

4.2 Properties of Linear Transformations

Theorem 20. Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then $T(\mathbf{0}) = \mathbf{0}$.

Proof. $T(\mathbf{0}) = T(0\mathbf{0}) = 0T(\mathbf{0}) = \mathbf{0}$. □

Theorem 21. Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation. Suppose $\mathcal{U}$ is a linear subspace of $\mathcal{V}$. Then the set

$$T(\mathcal{U}) = \{T(\mathbf{A}) \in \mathcal{W} : \mathbf{A} \in \mathcal{U}\}$$

is a linear subspace of $\mathcal{W}$.

Proof. To prove this theorem, we need to show that the set $T(\mathcal{U})$ is a linear subspace of $\mathcal{W}$. For this, we need to verify the three properties of a linear subspace: closure under vector addition, closure under scalar multiplication, and containing the zero vector.

1. Closure under Vector Addition:
Let $\mathbf{A}', \mathbf{B}' \in T(\mathcal{U})$. Then, there exists $\mathbf{A}, \mathbf{B} \in \mathcal{U}$ such that

$$\mathbf{A}' = T(\mathbf{A}) \text{ and } \mathbf{B}' = T(\mathbf{B}).$$

Since $\mathcal{U}$ is a linear subspace, $\mathbf{A} + \mathbf{B} \in \mathcal{U}$. Since T is a linear transformation, we have $T(\mathbf{A} + \mathbf{B}) = \mathbf{A}' + \mathbf{B}' \in T(\mathcal{U})$.

2. Closure under Scalar Multiplication:
Let $\mathbf{A}' \in T(\mathcal{U})$ and c be a scalar. Again, since $\mathcal{U}$ is a linear subspace, $c\mathbf{A} \in \mathcal{U}$. Since T is a linear transformation, it preserves scalar multiplication, $T(c\mathbf{A}) = cT(\mathbf{A}) = c\mathbf{A}' \in T(\mathcal{U})$.

3. Containing the Zero Vector:
Since $\mathcal{U}$ is a linear subspace of $\mathcal{V}$, it contains the zero vector $\mathbf{0}$ of $\mathcal{V}$. Since T is a linear transformation, it maps the zero vector to the zero vector:

$$T(\mathbf{0}) = \mathbf{0} \in T(\mathcal{U})$$

Therefore, $T(\mathcal{U})$ contains the zero vector of $\mathcal{W}$. This completes the proof of the theorem. □

Theorem 22. Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation and E is a set of vectors in $\mathcal{V}$. Then,

$$T(\mathcal{L}(E)) = \mathcal{L}(T(E)).$$

Proof. We will prove this theorem by showing that $T(\mathcal{L}(E))$ is a subset of $\mathcal{L}(T(E))$ and vice versa.

Part 1: $T(\mathcal{L}(E)) \subseteq \mathcal{L}(T(E))$.
Let $\mathbf{A} \in \mathcal{L}(E)$. This means that $\mathbf{A}$ can be expressed as a linear combination of vectors in E:

$$\mathbf{A} = c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \ldots + c_n\mathbf{A}_n,$$

where $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ are vectors in E, and $c_1, c_2, \ldots, c_n$ are scalars.

Now, apply the linear transformation T to both sides of the equation:

$$T(\mathbf{A}) = T(c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \ldots + c_n\mathbf{A}_n).$$

Since T is a linear transformation, it preserves vector addition and scalar multiplication:

$$T(\mathbf{A}) = c_1T(\mathbf{A}_1) + c_2T(\mathbf{A}_2) + \ldots + c_nT(\mathbf{A}_n).$$

Each of the vectors $T(\mathbf{A}_1), T(\mathbf{A}_2), \ldots, T(\mathbf{A}_n)$ is in $T(E)$ because T maps vectors from E to $T(E)$. Therefore, the right-hand side of the equation is a linear combination of vectors in $T(E)$. Hence, $T(\mathbf{A})$ is an element of $\mathcal{L}(T(E))$. Since $\mathbf{A}$ was an arbitrary element of $\mathcal{L}(E)$, we have shown that $T(\mathcal{L}(E)) \subseteq \mathcal{L}(T(E))$.
Part 2: $\mathcal{L}(T(E)) \subseteq T(\mathcal{L}(E))$.
Let $\mathbf{B} \in \mathcal{L}(T(E))$. This means that $\mathbf{B}$ can be expressed as a linear combination of vectors in $T(E)$:

$$\mathbf{B} = d_1\mathbf{B}_1 + d_2\mathbf{B}_2 + \ldots + d_m\mathbf{B}_m,$$

where $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_m$ are vectors in $T(E)$, and $d_1, d_2, \ldots, d_m$ are scalars. Since each $\mathbf{B}_i$ is in $T(E)$, there exists a corresponding vector $\mathbf{A}_i$ in E such that $T(\mathbf{A}_i) = \mathbf{B}_i$. Now, we can express $\mathbf{B}$ as a linear combination of the original vectors in E:

$$\mathbf{B} = d_1T(\mathbf{A}_1) + d_2T(\mathbf{A}_2) + \ldots + d_mT(\mathbf{A}_m).$$

Again, since T is a linear transformation, it preserves vector addition and scalar multiplication:

$$\mathbf{B} = T(d_1\mathbf{A}_1) + T(d_2\mathbf{A}_2) + \ldots + T(d_m\mathbf{A}_m) = T\left(\sum_{i=1}^{m} d_i\mathbf{A}_i\right).$$

Define

$$\mathbf{A} = \sum_{i=1}^{m} d_i\mathbf{A}_i$$

So $\mathbf{A} \in \mathcal{L}(E)$ and for $\mathbf{B} \in \mathcal{L}(T(E))$, we have found an $\mathbf{A} \in \mathcal{L}(E)$ such that $T(\mathbf{A}) = \mathbf{B}$. This implies $\mathbf{B} \in T(\mathcal{L}(E))$ and hence $\mathcal{L}(T(E)) \subseteq T(\mathcal{L}(E))$. Combining both parts, we conclude that

$$T(\mathcal{L}(E)) = \mathcal{L}(T(E)),$$

which completes the proof. □

Theorem 23. Let $T : \mathcal{V} \to \mathcal{W}$ and $S : \mathcal{W} \to \mathcal{U}$ be linear transformations. Then, the composition,

$$S \cdot T : \mathcal{V} \to \mathcal{U}$$

is a linear transformation.

Proof. We need to verify the two properties of a linear transformation. Indeed,

$$\begin{aligned} S \cdot T(a\mathbf{A} + b\mathbf{B}) &= S(T(a\mathbf{A} + b\mathbf{B})) \\ &= S(aT(\mathbf{A}) + bT(\mathbf{B})) \\ &= aS \cdot T(\mathbf{A}) + bS \cdot T(\mathbf{B}). \end{aligned}$$ □

4.3 Kernels and Images

Definition 12. Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation. The **kernel** of T is defined as

$$\operatorname{Ker} T = \left\{\mathbf{A} \in \mathcal{V} : T(\mathbf{A}) = \mathbf{0}\right\}.$$

The **image** of T is defined as

$$\operatorname{Im} T = \left\{\mathbf{B} \in \mathcal{W} : \exists\, \mathbf{A} \in \mathcal{V} \text{ such that } T(\mathbf{A}) = \mathbf{B}\right\}.$$

Theorem 24. Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then,

1. $\operatorname{Ker} T$ is a linear subspace of $\mathcal{V}$.
2. $\operatorname{Im} T$ is a linear subspace of $\mathcal{W}$.

Proof. **Part 1:** $\operatorname{Ker} T$ is a linear subspace of $\mathcal{V}$.
To prove that $\operatorname{Ker} T$ is a linear subspace of $\mathcal{V}$, we need to show that it satisfies the three properties of a linear subspace:
1. **Closure under Vector Addition:** Let $\mathbf{A}_1, \mathbf{A}_2 \in \operatorname{Ker} T$. This means that $T(\mathbf{A}_1) = \mathbf{0}$

and $T(\mathbf{A}_2) = \mathbf{0}$. Therefore,

$$T(\mathbf{A}_1 + \mathbf{A}_2) = T(\mathbf{A}_1) + T(\mathbf{A}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Hence $\mathbf{A}_1 + \mathbf{A}_2 \in \operatorname{Ker} T$.

2. **Closure under Scalar Multiplication:** Let $\mathbf{A} \in \operatorname{Ker} T$ and c be a scalar. This means that $T(\mathbf{A}) = \mathbf{0}$. Therefore,

$$T(c\mathbf{A}) = cT(\mathbf{A}) = c\mathbf{0} = \mathbf{0}.$$

Thus, $T(c\mathbf{A}) = \mathbf{0}$, $c\mathbf{A} \in \operatorname{Ker} T$.

3. **Containing the Zero Vector:** Since $T(\mathbf{0}) = \mathbf{0}$, the zero vector $\mathbf{0}$ is in $\operatorname{Ker} T$.

Thus, we conclude that $\operatorname{Ker} T$ is a linear subspace of $\mathcal{V}$.

Part 2: $\operatorname{Im} T$ is a linear subspace of $\mathcal{W}$.

1. **Closure under Vector Addition:** Let $\mathbf{B}_1, \mathbf{B}_2 \in \operatorname{Im} T$. This means that there exist vectors $\mathbf{A}_1, \mathbf{A}_2 \in \mathcal{V}$ such that $T(\mathbf{A}_1) = \mathbf{B}_1$ and $T(\mathbf{A}_2) = \mathbf{B}_2$. Using the properties of a linear transformation:

$$T(\mathbf{A}_1 + \mathbf{A}_2) = T(\mathbf{A}_1) + T(\mathbf{A}_2) = \mathbf{B}_1 + \mathbf{B}_2.$$

Since $\mathbf{A}_1 + \mathbf{A}_2 \in \mathcal{V}$ and $T(\mathbf{A}_1 + \mathbf{A}_2) = \mathbf{B}_1 + \mathbf{B}_2$, we conclude that $\mathbf{B}_1 + \mathbf{B}_2 \in \operatorname{Im} T$.

2. **Closure under Scalar Multiplication:** Let $\mathbf{B} \in \operatorname{Im} T$ and c be a scalar. Thus, there exists a vector $\mathbf{A}$ in $\mathcal{V}$ such that $T(\mathbf{A}) = \mathbf{w}$. Using the properties of a linear transformation:

$$T(c\mathbf{A}) = cT(\mathbf{A}) = c\mathbf{B}.$$

Since $c\mathbf{A} \in \mathcal{V}$ and $T(c\mathbf{A}) = c\mathbf{B}$, we conclude that $c\mathbf{B} \in \operatorname{Im} T$.

3. **Containing the Zero Vector:** Since $T(\mathbf{0}) = \mathbf{0}$, the zero vector $\mathbf{0}$ is in $\operatorname{Im} T$.

Thus, we conclude that $\operatorname{Im} T$ is a linear subspace of $\mathcal{W}$. □

Example 31. Let $D : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ be the differentiation operator. For $f(x) \in P_n(\mathbb{R})$, $\ker D$ contains all $f(x)$ such that $D(f(x)) = f'(x) = 0$. The only polynomials that satisfy this condition are constant polynomials. Therefore, the kernel of D is the set of constant polynomials $P_0(\mathbb{R})$. Therefore,

$$\ker D = P_0(\mathbb{R}).$$

For $g(x) \in P_n(\mathbb{R})$, $\operatorname{Im} D$ contains all $g(x)$ such that there exists $f(x) \in P_n(\mathbb{R})$ with $D(f(x)) = g(x)$. Since the derivative of a polynomial $f(x)$ is $D(f(x)) = f'(x)$, any polynomial $g(x)$ in $P_n(\mathbb{R})$ can be in the image of D of degree of at most $n-1$. Therefore, $\operatorname{Im} D$ is the set of polynomials of degree at most $n-1$:

$$\operatorname{Im} D = P_{n-1}(\mathbb{R}).$$

Generalizing, we can show that $\operatorname{Ker} D^m = P_{m-1}(\mathbb{R}), \operatorname{Im} D^m = P_{n-m}(\mathbb{R})$.

Theorem 25 (Dimension Formula). Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation. If $\mathcal{V}$ is finite dimensional, so are $\operatorname{Ker} T$ and $\operatorname{Im} T$ and

$$\dim \mathcal{V} = \dim \operatorname{Ker} T + \dim \operatorname{Im} T.$$

Proof. Given $\mathcal{V}$ is finite dimensional and $\operatorname{Ker} T$ is a linear subspace of $\mathcal{V}$, so $\operatorname{Ker} T$ is finite dimensional. Let $\{\mathbf{A}_1, \dots, \mathbf{A}_s\}$ be a basis for $\operatorname{Ker} T$. Then by Basis Extension Theorem, we can find linearly independent vectors $\mathbf{B}_1, \dots \mathbf{B}_t$ such that $\{\mathbf{A}_1, \dots, \mathbf{A}_s, \mathbf{B}_1, \dots \mathbf{B}_t\}$ is a basis for $\mathcal{V}$. Now,

$$\begin{aligned}
\operatorname{Im} T = T(\mathcal{V}) &= T(\mathcal{L}(\mathbf{A}_1, \dots, \mathbf{A}_s, \mathbf{B}_1, \dots \mathbf{B}_t)) \\
&= \mathcal{L}(T(\mathbf{A}_1), \dots, T(\mathbf{A}_s), T(\mathbf{B}_1), \dots T(\mathbf{B}_t)) \quad \text{(by Theorem 22)} \\
&= \mathcal{L}(\mathbf{0}, \dots, \mathbf{0}, T(\mathbf{B}_1), \dots T(\mathbf{B}_t)) \\
&= \mathcal{L}(T(\mathbf{B}_1), \dots T(\mathbf{B}_t)).
\end{aligned}$$

This shows that $\operatorname{Im} T$ is finite dimensional. We claim that $\{T(\mathbf{B}_1), \dots T(\mathbf{B}_t)\}$ form a basis for $\operatorname{Im} T$. All that remains to be shown is, they are linearly independent. To that end, suppose there are scalars $b_1, \dots, b_t$ such that

$$b_1 T(\mathbf{B}_1) + \cdots + b_t T(\mathbf{B}_t) = 0.$$

By definition of a linear transformation, we have,

$$T\left(\sum_{i=1}^{t} b_i \mathbf{B}_i\right) = 0.$$

Let $\mathbf{B} = \sum_{i=1}^{t} b_i \mathbf{B}_i$. Then $T(\mathbf{B}) = 0$. This means $\mathbf{B} \in \operatorname{Ker} T$. Now, since $\{\mathbf{A}_1, \dots, \mathbf{A}_s\}$ is a basis for $\operatorname{Ker} T$, we have,

$$\mathbf{B} = a_1\mathbf{A}_1 + \cdots + a_s\mathbf{A}_s,$$

for scalars $a_1, \ldots, a_s$. But then,

$$a_1\mathbf{A}_1 + \cdots + a_s\mathbf{A}_s - b_1\mathbf{B}_1 - \cdots - b_t\mathbf{B}_t = 0.$$

Since $\{\mathbf{A}_1, \ldots, \mathbf{A}_s, \mathbf{B}_1, \ldots \mathbf{B}_t\}$ is a basis for $\mathcal{V}$, we conclude that $a_i = b_i = 0$ which shows that the $\{T(\mathbf{B}_1), \ldots T(\mathbf{B}_t)\}$ is linearly independent and hence a basis for $\operatorname{Im} T$. Thus,

$$\dim \operatorname{Ker} T = s, \dim \operatorname{Im} T = t \text{ and } \dim \mathcal{V} = n = s + t,$$

as required. This completes the proof. □

Theorem 26. Let T:$\mathcal{V} \to \mathcal{W}$ be a linear transformation. Let $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ be a basis for $\mathcal{V}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ be a set of vectors in $\mathcal{W}$. Then the linear extensio of

$$T(\mathbf{A}_i) = \mathbf{B}_i, i = 1, \ldots, n$$

is a linear transformation.

Proof. We first observe that T is well-defined since $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is a basis for $\mathcal{V}$. Let $\mathbf{A}, \mathbf{C} \in \mathcal{V}$. Then,

$$\mathbf{A} = \sum_{i=1}^{n} a_i\mathbf{A}_i,\ \mathbf{C} = \sum_{i=1}^{n} c_i\mathbf{A}_i,$$

Therefore,

$$\begin{aligned} T(a\mathbf{A} + c\mathbf{C}) &= a\sum_{i=1}^{n} a_iT(\mathbf{A}_i) + c\sum_{i=1}^{n} c_iT(\mathbf{C}_i) \\ &= a\sum_{i=1}^{n} a_i\mathbf{B}_i + c\sum_{i=1}^{n} c_i\mathbf{B}_i \\ &= aT(\mathbf{A}) + cT(\mathbf{C}). \end{aligned}$$

□

Theorem 27. If $S, T : \mathcal{V} \to \mathcal{W}$ be linear transformations, the so is $S + T : \mathcal{V} \to \mathcal{W}$.

Proof. Obvious. □

Theorem 28. Let $\mathbf{A}_1, \ldots, \mathbf{A}_n \in \mathcal{V}$ and $\mathbf{E}_1, \ldots, \mathbf{E}_n$ be the standard basis for $\mathbb{R}^n$. Let

$T : \mathbb{R}^n \to \mathcal{V}$ be the linear extension of

$$T(\mathbf{E}_i) = \mathbf{A}_i, \; i = 1, \dots, n.$$

Then,

1. $\mathbf{A}_1, \dots, \mathbf{A}_n$ are linearly independent if and only if $\operatorname{Ker} T = \{\mathbf{0}\}$.
2. $\mathbf{A}_1, \dots, \mathbf{A}_n$ span $\mathcal{V}$ if and only if $\operatorname{Im} T = \mathcal{V}$.

Proof. Let $T : \mathbb{R}^n \to \mathcal{V}$ be the linear extension of

$$T(\mathbf{E}_i) = \mathbf{A}_i, \; i = 1, \dots, n.$$

1. Suppose $\mathbf{A}_1, \dots, \mathbf{A}_n$ are linearly independent. Let $\mathbf{B} = (b_1, \dots, b_n) \in \mathbb{R}^n$. Then,

   $$\mathbf{B} = \sum_{i=1}^{n} b_i \mathbf{E}_i.$$

   Suppose $\mathbf{B} \in \operatorname{Ker} T$. Then $T(\mathbf{B}) = 0$, i.e.

   $$0 = T(\mathbf{B}) = T\left(\sum_{i=1}^{n} b_i \mathbf{E}_i \right) = \sum_{i=1}^{n} b_i T(\mathbf{E}_i) = \sum_{i=1}^{n} b_i \mathbf{A}_i.$$

   But since $\mathbf{A}_i$ are linearly independent, we have $b_i = 0, i = 1, \dots, n$. Hence $\mathbf{B} = \mathbf{0}$ and hence $\operatorname{Ker} T = \{\mathbf{0}\}$.
   Conversely, let $\operatorname{Ker} T = \{\mathbf{0}\}$. Then, for any $\mathbf{A} = (a_1, \dots, a_n) \in \operatorname{Ker} T$, we have $a_i = 0$, i.e. $\mathbf{A} = \mathbf{0}$. But then,

   $$\mathbf{0} = \mathbf{A} = \sum_{i=1}^{n} a_i \mathbf{E}_i.$$

   so that $\mathbf{0} = T(\mathbf{0}) = \sum_{i=1}^{n} a_i \mathbf{A}_i$. Hence the vectors $\mathbf{A}_1, \dots, \mathbf{A}_n$ are linearly independent trivially.

2. Suppose $\mathbf{A}_1, \dots, \mathbf{A}_n$ span $\mathcal{V}$. Then $\mathcal{L}(\mathbf{A}_1, \dots, \mathbf{A}_n) = \mathcal{V}$. By definition, $\operatorname{Im} T \subseteq \mathcal{V}$. Let $\mathbf{B} \in \mathcal{V}$. Then,

   $$\mathbf{B} = \sum_{i=1}^{n} a_i \mathbf{A}_i = \sum_{i=1}^{n} a_i T(\mathbf{E}_i) = T\left(\sum_{i=1}^{n} a_i \mathbf{E}_i \right).$$

Let $\mathbf{A} = (a_1, \ldots, a_n) = \sum_{i=1}^{n} a_i \mathbf{E}_i$. Then, $\mathbf{B} = T(\mathbf{A})$ and hence $\mathbf{B} \in \operatorname{Im} T$. Combining both the parts, we conclude that $\operatorname{Im} T = \mathcal{V}$.

On the other hand, suppose that $\operatorname{Im} T = \mathcal{V}$. By definition, $\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_n) \subseteq \mathcal{V}$. Let $\mathbf{B} \in \operatorname{Im} T = \mathcal{V}$. Then there exists $\mathbf{A} = (a_1, \ldots, a_n) \in \mathbb{R}^n$ such that $T(\mathbf{A}) = \mathbf{B}$. Therefore,

$$\mathbf{B} = T(\mathbf{A}) = T\left(\sum_{i=1}^{n} a_i \mathbf{E}_i\right) = \sum_{i=1}^{n} a_i T(\mathbf{E}_i) = \sum_{i=1}^{n} a_i \mathbf{A}_i$$

showing that $\mathbf{B} \in \mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_n)$. Hence, $\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_n) = \mathcal{V}$. □

4.4 Isomorphisms

Definition 13. A linear transformation $T : \mathcal{V} \to \mathcal{W}$ is an **isomorphism** if there exists a linear transformation $S : \mathcal{W} \to \mathcal{V}$ such that

$$S \cdot T(\mathbf{A}) = \mathbf{A}, \quad \forall\, \mathbf{A} \in \mathcal{V}$$
$$T \cdot S(\mathbf{B}) = \mathbf{B}, \quad \forall\, \mathbf{B} \in \mathcal{W}.$$

S and T are called **inverse isomorphisms**. If there is a linear transformation $T : \mathcal{V} \to \mathcal{W}$, then we say $\mathcal{V}$ and $\mathcal{W}$ are **isomorphic**.

Theorem 29. Suppose that $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\operatorname{Ker} T = \{\mathbf{0}\}$. Then for each $\mathbf{B} \in \operatorname{Im} T$, there exists exactly one vector $\mathbf{A} \in \mathcal{V}$ such that $T(\mathbf{A}) = \mathbf{B}$.

Proof. Suppose that $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation and $\operatorname{Ker} T = \{\mathbf{0}\}$. Let $\mathbf{B} \in \operatorname{Im} T$ and suppose there exists two vectors $\mathbf{A}, \mathbf{C} \in \mathcal{V}$ such that $T(\mathbf{A}) = \mathbf{B}$ and $T(\mathbf{C}) = \mathbf{B}$. Then,

$$T(\mathbf{A} - \mathbf{C}) = T(\mathbf{A}) - T(\mathbf{C}) = \mathbf{B} - \mathbf{B} = \mathbf{0}$$

showing that $\mathbf{A} - \mathbf{C} \in \operatorname{Ker} T$. Since $\operatorname{Ker} T = \{\mathbf{0}\}$, we conclude $\mathbf{A} = \mathbf{C}$. □

Theorem 30. A linear transformation $T : \mathcal{V} \to \mathcal{W}$ is an isomorphism if and only if $\operatorname{Ker} T = \{\mathbf{0}\}$ and $\operatorname{Im} T = \mathcal{W}$.

Proof. Suppose that $\operatorname{Ker} T = \{\mathbf{0}\}$ and $\operatorname{Im} T = \mathcal{W}$. By Theorem 29, for each $\mathbf{B} \in \operatorname{Im} T$, there exists exactly one vector $\mathbf{A} \in \mathcal{V}$ such that $T(\mathbf{A}) = \mathbf{B}$. Define a map $S : \mathcal{W} \to \mathcal{V}$ by setting $S(\mathbf{B}) = \mathbf{A}$. By construction, such a map is well-defined since $\mathbf{A}$ is the

unique vector. We claim that S is a linear transformation.
Let $\mathbf{B}_1, \mathbf{B}_2 \in \operatorname{Im} T$. Then there exists unique vectors $\mathbf{A}_1, \mathbf{A}_2 \in \mathcal{V}$ such that

$$T(\mathbf{A}_1) = \mathbf{B}_1, \quad T(\mathbf{A}_2) = \mathbf{B}_2.$$

By definition, $S(\mathbf{B}_1) = \mathbf{A}_1$ and $S(\mathbf{B}_2) = \mathbf{A}_2$. Now $T(\mathbf{A}_1 + \mathbf{A}_2) = \mathbf{B}_1 + \mathbf{B}_2$. Here $\mathbf{A}_1 + \mathbf{A}_2$ is the unique vector which is mapped to $\mathbf{B}_1 + \mathbf{B}_2$ under T. Hence, by definition,

$$S(\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{A}_1 + \mathbf{A}_2 = S(\mathbf{B}_1) + S(\mathbf{B}_2).$$

Similarly, $S(a\mathbf{B}) = aS(\mathbf{B})$ whence we conclude that S is a linear transformation. Therefore,

$$S \cdot T(\mathbf{A}) = S(T(\mathbf{A})) = S(\mathbf{B}) = \mathbf{A}, \ \forall\, \mathbf{A} \in \mathcal{V}$$
$$T \cdot S(\mathbf{B}) = T(S(\mathbf{B})) = T(\mathbf{A}) = \mathbf{B}, \ \forall\, \mathbf{B} \in \mathcal{W}.$$

Hence T is an isomorphism.
On the other hand, suppose that T is an isomorphism and $\mathbf{A} \in \operatorname{Ker} T$. Therefore $T(\mathbf{A}) = 0$. But then,

$$\mathbf{A} = S \cdot T(\mathbf{A}) = S(\mathbf{0}) = \mathbf{0}.$$

Hence $\operatorname{Ker} T = \{\mathbf{0}\}$. Use Theorem 25 to conclude $\operatorname{Im} T = \mathcal{W}$. □

Theorem 31. Two finite-dimensional vector spaces $\mathcal{V}$ and $\mathcal{W}$ are isomorphic if and only if $\dim \mathcal{V} = \dim \mathcal{W}$.

Proof. Suppose $\dim \mathcal{V} = \dim \mathcal{W}$. Let $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ be a basis for $\mathcal{V}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ be a basis for $\mathcal{W}$. Let $T : \mathcal{V} \to \mathcal{W}$ be the linear transformation which is the linear extension of

$$T(\mathbf{A}_i) = \mathbf{B}_i, \ i = 1, \ldots, n.$$

Then $\operatorname{Im} T = \mathcal{W}$ since $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ is a basis for $\mathcal{W}$. Use Theorem 25 to show $\operatorname{Ker} T = \{\mathbf{0}\}$ and finally Theorem 30 to conclude that T is an isomorphism.
Conversely, suppose $\mathcal{V}$ and $\mathcal{W}$ are isomorphic. Then there is an isomorphism $T : \mathcal{V} \to \mathcal{W}$. Let $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ be a basis for $\mathcal{V}$. Then $\dim \mathcal{V} = n$. Since T is an isomorphism we will show that the set $\left\{T(\mathbf{A}_1), \ldots, T(\mathbf{A}_n)\right\}$ is a basis for $\mathcal{W}$ so that

$\dim \mathcal{W} = n = \dim \mathcal{V}$. Observe that

$$\begin{aligned}\mathcal{W} = \operatorname{Im} T = T(\mathcal{V}) &= T(\mathcal{L}(\mathbf{A}_1, \ldots, \mathbf{A}_n)) \\ &= \mathcal{L}(T(\mathbf{A}_1), \ldots, T(\mathbf{A}_n)) \quad \text{(by Theorem 22).}\end{aligned}$$

This shows that $\left\{T(\mathbf{A}_1), \ldots T(\mathbf{A}_n)\right\}$ span $\mathcal{W}$. All that remains to be shown is, they are linearly independent. To that end, suppose there are scalars $a_1, \ldots, a_n$ such that

$$a_1 T(\mathbf{A}_1) + \cdots + a_n T(\mathbf{A}_n) = 0.$$

By definition of a linear transformation, we have,

$$T\left(\sum_{i=1}^{n} a_i \mathbf{A}_i\right) = 0.$$

Let $\mathbf{A} = \sum_{i=1}^{n} a_i \mathbf{A}_i$. Then $T(\mathbf{A}) = 0$. This means $\mathbf{A} \in \operatorname{Ker} T$. But $\operatorname{Ker} T = \{\mathbf{0}\}$. Therefore,

$$a_1 \mathbf{A}_1 + \cdots + a_s \mathbf{A}_n = 0$$

But since $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is a basis for $\mathcal{V}$, we conclude that $a_i = 0$ which shows that the $\left\{T(\mathbf{A}_1), \ldots T(\mathbf{A}_n)\right\}$ is linearly independent and hence a basis for $\mathcal{W}$. □

One of the most important results for a finite-dimensional vector space is the following.

Theorem 32. A finite-dimensional vector space of dimension n is isomorphic to $\mathbb{R}^n$.

Proof. Hint: Note that $\dim \mathcal{V} = \dim \mathbb{R}^n = n$. Use Theorem 31 above and define a map $T : \mathcal{V} \to \mathbb{R}^n$ which takes basis vectors to basis vectors. □

5 Representing Linear Transformations by Matrices

A linear transformation T from a vector space V to a vector space W can be represented by a matrix with respect to some ordered basis.

5.1 Matrix Representation

Let $\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$ be an ordered basis for $\mathcal{V}$ and $\{\mathbf{B}_1, \dots, \mathbf{B}_m\}$ be an ordered basis for $\mathcal{W}$. Note that ordering is important otherwise the representation of the matrix will change.

Since $T(\mathbf{A}_1), \dots, T(\mathbf{A}_n)$ are elements in $\mathcal{W}$, they can be uniquely represented as a linear combination of the basis vectors $\mathbf{B}_1, \dots, \mathbf{B}_m$. Let $a_{ij}, i = 1, \dots, m, j = 1, \dots, n$ be scalars such that

$$\begin{aligned} T(A_1) &= a_{11}\mathbf{B}_1 + \cdots + a_{m1}\mathbf{B}_m, \\ &\vdots \\ T(A_n) &= a_{1n}\mathbf{B}_1 + \cdots + a_{mn}\mathbf{B}_m. \end{aligned}$$

Then the $m \times n$ matrix $A = (a_{ij})$ is called the **matrix of T relative to the ordered basis** $\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \dots, \mathbf{B}_m\}$.

Example 32. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as:

$$T(x, y) = (2x + 3y, 4x - y).$$

We want to find the matrix representation of T with respect to the following basis pair for $\mathbb{R}^2$:

$$\{(1, 0), (0, 1)\} \text{ and } \{(1, 0), (0, 1)\}.$$

Now,

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

Therefore, the matrix A of T relative to this basis pair is

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

5.2 Fundamental Results

Theorem 33. Let $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ be an ordered basis for $\mathcal{V}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$ be an ordered basis for $\mathcal{W}$. Then assigning to each linear transformation $T : \mathcal{V} \to \mathcal{W}$, its matrix relative to these ordered bases defines an isomorphism

$$M : \mathcal{L}(\mathcal{V}, \mathcal{W}) \to M_{m\times n}$$

where $\mathcal{L}(\mathcal{V}, \mathcal{W})$ is the vector space of all linear transformations from $\mathcal{V}$ to $\mathcal{W}$ and $M_{m\times n}$ is the vector space of all $m \times n$ matrices.

Proof. Refer to class notes and discussions. □

Theorem 34. Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional vector spaces. Then $\mathcal{L}(\mathcal{V}, \mathcal{W})$ finite-dimensional and $\dim \mathcal{L}(\mathcal{V}, \mathcal{W}) = \dim \mathcal{V} \cdot \dim \mathcal{W}$.

Proof. The result is an immediate consequence of Theorem 33. Since M is an isomorphism $\dim \mathcal{L}(\mathcal{V}, \mathcal{W}) = \dim M_{m\times n} = n \cdot m = \dim \mathcal{V} \cdot \dim \mathcal{W}$. □

Theorem 35. A linear transformation $T : \mathcal{V} \to \mathcal{V}$ is an isomorphism if and only if its matrix is invertible.

Proof. Let $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ be ordered bases for $\mathcal{V}$. Suppose that $T : \mathcal{V} \to \mathcal{V}$ is an isomorphism whose matrix relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ is A. Let $S : \mathcal{V} \to \mathcal{V}$ be the inverse of T whose matrix relative to the ordered bases $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is B. Then AB is the matrix of $T \cdot S : \mathcal{V} \to \mathcal{V}$ relative to the ordered bases $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$. Since T is an isomorphism, $T \cdot S(\mathbf{C}) = \mathbf{C}$ for all $\mathbf{C} \in \mathcal{V}$. In particular,

$$T \cdot S(\mathbf{B}_j) = \mathbf{B}_j = 0\mathbf{B}_1 + \cdots + 1\mathbf{B}_j + \cdots + 0\mathbf{B}_n$$

and the matrix of $T \cdot S$ relative to the ordered bases $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ is

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

Hence $AB = I$. Similarly, we can show that the matrix of $S \cdot T$ relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is I. Thus, $BA = I$. Thus, the matrix of T represented by A is invertible.

Conversely, let the matrix A of T be invertible. Then there is a matrix B such that $AB = BA = I$. Let $S : \mathcal{V} \to \mathcal{V}$ be the linear transformation whose matrix relative to the ordered bases $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is B. Then the matrix of $S \cdot T : \mathcal{V} \to \mathcal{V}$ relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is

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

Hence $S \cdot T$ and I have the same matrix relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$. Therefore, $S \cdot T = I$. Hence,

$$S \cdot T(\mathbf{A}) = \mathbf{A}, \quad \forall\, \mathbf{A} \in \mathcal{V}$$

and similarly,

$$T \cdot S(\mathbf{A}) = \mathbf{A}, \quad \forall\, \mathbf{A} \in \mathcal{V}.$$

so that S and T are inverse isomorphisms. □

Remark. It is important to understand that the definition of a linear transformation is independent of any basis representation and is more general in nature. By associating a matrix with a linear transformation we get a lot of additional information about the nature of the transformation. However, one should not blindly associate a matrix with a linear transformation without specifying the underlying basis pair. By doing so, one may commit a grave mistake as the following examples show.

Example 33. Let $I : \mathbb{R}^3 \to \mathbb{R}^3$ be the identity transformation. One would naturally expect that the matrix associated with I will be the identity matrix as discussed in Theorem 35. The reader may have observed that the matrix of $T \cdot S$ and $S \cdot T$

was an identity matrix relative to same ordered basis pair $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ and $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ respectively. If the basis is changed then the associated matrix also changes. To see this, consider the following basis pair of $\mathbb{R}^3$:

$$(1, 0, 0), (0, 1, 0), (0, 0, 1) \text{ and } (1, 1, 1), (1, 1, 0), (1, 0, 0).$$

Then the matrix of I relative to this basis pair is

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

which is not the identity matrix.

Example 34. Let $D : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ be the differentiation operator. It is well known that D is a nilpotent transformation and hence the matrix of D relative to the ordered bases $\{1, x, x^2\}$, $\{1, x, x^2\}$ is a nilpotent matrix. However, if we consider the following basis pair of $P_2(\mathbb{R})$,

$$1, x, x^2 \text{ and } x - 1, x + 1, (x - 1)^2$$

then the matrix of D relative to this basis pair is

$$M = \begin{bmatrix} 0 & -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Here $D^3 = 0$ but $M^3 \neq 0$.

Theorem 36. Let A and B be square matrices of size n such that $AB = I$. Then $BA = I$.

Proof. We will prove the result for a finite-dimensional vector space $\mathcal{V}$ where $\dim \mathcal{V} = n$. One may as well consider $\mathbb{R}^n$ and the proof remains valid due to Theorem 32. Let $T, S : \mathcal{V} \to \mathcal{V}$ whose matrices relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ are A and B respectively. Now AB is the matrix of $T \cdot S$ relative to the ordered bases $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}, \{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$. Since $AB = I$, we see that $T \cdot S$ is the identity transformation I. We will first show that the vectors $S(\mathbf{B}_1), \ldots, S(\mathbf{B}_n)$ are linearly independent. Suppose on the contrary they are linearly dependent. Then there exists scalars $a_1, \dots, a_n$ not all zero such that

$$a_1 S(\mathbf{B}_1) + \cdots + a_n S(\mathbf{B}_n) = \mathbf{0}.$$

Applying T both sides we get

$$a_1 T \cdot S(\mathbf{B}_1) + \cdots + a_n T \cdot S(\mathbf{B}_n) = \mathbf{0}.$$

Since $T \cdot S = I$, we get the relation,

$$a_1 \mathbf{B}_1 + \cdots + a_n \mathbf{B}_n = \mathbf{0}$$

which is a contradiction since $\mathbf{B}_1, \dots, \mathbf{B}_n$ is a basis for $\mathcal{V}$. Hence $S(\mathbf{B}_1), \dots, S(\mathbf{B}_n)$ are linearly independent and therefore by Theorem 17, they form a basis for $\mathcal{V}$. So any vector $\mathbf{C} \in \mathcal{V}$ can be written as:

$$\mathbf{C} = c_1 S(\mathbf{B}_1) + \cdots + c_n S(\mathbf{B}_n) = S\left(\sum_{i=1}^{n} c_i \mathbf{B}_i\right) = S(\mathbf{D}) \tag{6}$$

where $\mathbf{D} = \sum_{i=1}^{n} c_i \mathbf{B}_i$. To show $BA = I$ we will show that $S \cdot T$ is the identity transformation, i.e. $S \cdot T(\mathbf{C}) = \mathbf{C}$ for all $\mathbf{C} \in \mathcal{V}$. But then from (6), we see that

$$S \cdot T(\mathbf{C}) = (S \cdot T) \cdot S(\mathbf{D}) = S \cdot (T \cdot S)(\mathbf{D}) = S \cdot I(\mathbf{D}) = S(\mathbf{D}) = \mathbf{C}.$$

and hence $BA = I$ as required. □

5.3 Change of Basis

Theorem 37. Let A, B be $m \times n$ matrices, $\dim \mathcal{V} = n$ and $\dim \mathcal{W} = m$. Then A, B represent the same linear transformation T relative to different pair of ordered bases if and only if there exists nonsingular matrices P and Q such that

$$A = PBQ^{-1}$$

where P is a $m \times m$ matrix and Q is a $n \times n$ matrix.

Proof. Let the matrix of T relative to the ordered bases $\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \dots, \mathbf{B}_m\}$ be A and the matrix of T relative to the ordered bases $\{\mathbf{C}_1, \dots, \mathbf{C}_n\}$ and $\{\mathbf{D}_1, \dots, \mathbf{D}_m\}$ be B. Let P be the matrix of the identity transformation $I : \mathcal{W} \to \mathcal{W}$ relative to

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
\underset{\{\mathbf{A}_1,\ldots,\mathbf{A}_n\}}{\mathcal{V}} \arrow[r, "A", "T"'] \arrow[d, "Q^{-1}"', "I_{\mathcal{V}}"] & \underset{\{\mathbf{B}_1,\ldots,\mathbf{B}_m\}}{\mathcal{W}} \\
\underset{\{\mathbf{C}_1,\ldots,\mathbf{C}_n\}}{\mathcal{V}} \arrow[r, "B", "T"'] & \underset{\{\mathbf{D}_1,\ldots,\mathbf{D}_m\}}{\mathcal{W}} \arrow[u, "P", "I_{\mathcal{W}}"']
\end{tikzcd}
$$

Figure 1: Visual representation of change of basis.

the ordered bases $\{\mathbf{D}_1, \ldots, \mathbf{D}_m\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$. Similarly, let Q be the matrix of the identity transformation $I : \mathcal{V} \to \mathcal{V}$ relative to the ordered bases $\{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$ and $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$. Then by Theorem 35 both P and Q are invertible. Then PB is the matrix of $T : \mathcal{V} \to \mathcal{W}$ relative to the ordered bases $\{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$. Subsequently, PBQ^{-1} is the matrix of $T : \mathcal{V} \to \mathcal{W}$ relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$. But A is the matrix of $T : \mathcal{V} \to \mathcal{W}$ relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$, see Figure 1 for a better visualization. Hence,

$$A = PBQ^{-1}.$$

To prove the converse, suppose P and Q be invertible matrices such that $A = PBQ^{-1}$. Let the matrix of T relative to the ordered bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$ be A. Define

$$
\begin{aligned}
\mathbf{C}_1 &= Q^{-1}(\mathbf{A}_1), \ldots, \mathbf{C}_n = Q^{-1}(\mathbf{A}_n),\\
\mathbf{D}_1 &= P^{-1}(\mathbf{B}_1), \ldots, \mathbf{D}_m = P^{-1}(\mathbf{B}_m).
\end{aligned}
$$

This definition is well-defined as P and Q are invertible. We need to show that the matrix of T relative to the ordered bases $\{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$ and $\{\mathbf{D}_1, \ldots, \mathbf{D}_m\}$ is B. Since $A = PBQ^{-1}$, we can compute and show that $B = P^{-1}AQ$. From the above construction AQ is the matrix of $T : \mathcal{V} \to \mathcal{W}$ relative to the ordered bases $\{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_m\}$. Therefore $P^{-1}AQ$ is the matrix of $T : \mathcal{V} \to \mathcal{W}$ relative to the ordered bases $\{\mathbf{C}_1, \ldots, \mathbf{C}_n\}$ and $\{\mathbf{D}_1, \ldots, \mathbf{D}_m\}$ which is B as required. $\square$

6 System of Linear Equations

6.1 Existence Results

We are interested in the solution of the problem

$$AX = B$$

where

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation whose matrix relative the standard bases is A. Then,

$$T(X) = T(x_1, \dots, x_n) = AX = \left(\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \dots, \sum_{j=1}^{n} a_{mj}x_j \right). \tag{7}$$

Let $\mathbf{B} = (b_1, \dots, b_m) \in \mathbb{R}^m$.

Definition 14. A **solution** of the system (7) is a vector $\mathbf{S} = (s_1, \dots, s_n) \in \mathbb{R}^n$ such that $T(\mathbf{S}) = \mathbf{B}$, i.e.

$$T(s_1, \dots, s_n) = (b_1, \dots, b_m).$$

Theorem 38. The system (7) has a solution if and only $\mathbf{B} \in \operatorname{Im} T$.

Proof.

$$\begin{aligned} & \text{The system (7) has a solution} \\ \iff & \text{There exists a vector } \mathbf{S} \text{ such that } T(\mathbf{S}) = \mathbf{B} \\ \iff & \mathbf{B} \in \operatorname{Im} T. \end{aligned}$$

□

Definition 15. The **column space** of A, denoted as $C(A)$, is the linear span of the column vectors of A, i.e. $C(A) = \mathcal{L}(\mathbf{A}_{(1)}, \dots, \mathbf{A}_{(n)})$ where $\mathbf{A}_{(i)} = (a_{1i}, a_{2i}, \dots, a_{mi}), i = 1, \dots, n$.

Theorem 39. The system (7) has a solution if and only $\mathbf{B} \in C(A)$.

Proof. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation whose matrix relative the standard bases is A. Observe that, by definition of the representation of T by it's matrix A, we have $\mathbf{A}_{(i)} = T(\mathbf{E}_i), i = 1, \dots, n$. Therefore,

$$\operatorname{Im} T = T(\mathcal{L}(\mathbf{E}_1, \dots, \mathbf{E}_i)) = \mathcal{L}(T(\mathbf{E}_1), \cdots, T(\mathbf{E}_n)) = \mathcal{L}(\mathbf{A}_{(1)}, \dots, \mathbf{A}_{(n)}) = C(A)$$

and by Theorem 38, $\mathbf{B} \in \operatorname{Im} T = C(A)$. $\square$

Example 35. The following system

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

has no solution since $C(A) = \mathcal{L}((1,1,2),(1,0,1))$ and the vector $\mathbf{B} = (1,1,0) \notin C(A)$.

Definition 16. A system $AX = B$ is called **homogeneous** if $B = \mathbf{0}$.

Theorem 40. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation whose matrix relative the standard bases is A. Then $\mathcal{S} = \operatorname{Ker} T$ where $\mathcal{S}$ is the set of all solutions for $AX = \mathbf{0}$.

Proof. $\mathbf{S} \in \mathcal{S} \iff T(\mathbf{S}) = \mathbf{0} \iff A\mathbf{S} = \mathbf{0} \iff \mathbf{S}$ is a solution of $AX = 0$. Hence $\mathcal{S} = \operatorname{Ker} T$. $\square$

6.2 Affine Subspaces

Definition 17. Let $\mathcal{V}$ be a vector space and $\mathcal{U}$ be a linear subspace of $\mathcal{V}$. Then the set defined as

$$\mathbf{A} + \mathcal{U} = \left\{ \mathbf{A} + \mathbf{X} : \mathbf{X} \in \mathcal{U} \right\}$$

is called a **parallel translate** or a **parallel** of $\mathcal{U}$ in $\mathcal{V}$. A parallel of some linear subspace of $\mathcal{V}$ is called an **affine subspace** of $\mathcal{V}$.

Example 36. Let $\mathcal{V} = \mathbb{R}^2$ and $\mathcal{U} = \left\{ (x,y) : x + y = 0 \right\}$ be the linear subspace. Let $\mathbf{A} = (3,1)$. A visualization of a parallel of $\mathcal{U}$ in $\mathcal{V}$ is shown in Figure 2.

Theorem 41. The following results hold:

1. $\mathbf{A} \in \mathbf{A} + \mathcal{U}$.

2. If $\mathbf{B} \in \mathbf{A} + \mathcal{U}$, then $\mathbf{B} + \mathcal{U} = \mathbf{A} + \mathcal{U}$.

3. Two parallels have no common vectors.

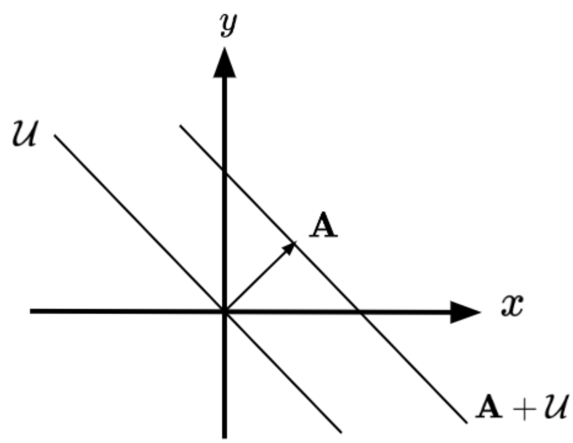

Figure 2: Parallel translate of $\mathcal{U}$ in $\mathcal{V}$

4. If $\mathbf{B}, \mathbf{C} \in \mathbf{A} + \mathcal{U}$ then $\mathbf{B} - \mathbf{C} \in \mathcal{U}$.

Proof. Let us prove each of the claims.

1. $\mathbf{0} \in \mathcal{U}$ and $\mathbf{A} = \mathbf{A} + \mathbf{0}$.

2. Let $\mathbf{B} \in \mathbf{A} + \mathcal{U}$. Then,

$$\begin{aligned}
\mathbf{B} + \mathcal{U} &= \left\{\mathbf{B} + \mathbf{Y} : \mathbf{Y} \in \mathcal{U}\right\}, \\
&= \left\{\mathbf{A} + \mathbf{X} + \mathbf{Y} : \mathbf{X}, \mathbf{Y} \in \mathcal{U}\right\}, \ (\because \mathbf{B} \in \mathbf{A} + \mathcal{U}) \\
&= \left\{\mathbf{A} + \mathbf{C} : \mathbf{C} \in \mathcal{U}\right\}, \ (\mathbf{C} = \mathbf{X} + \mathbf{Y}) \\
&= \mathbf{A} + \mathcal{U}.
\end{aligned}$$

3. Let $\mathbf{C} \in \mathbf{A} + \mathcal{U}$ and $\mathbf{C} \in \mathbf{B} + \mathcal{U}$. Then $\mathbf{C} + \mathcal{U} = \mathbf{A} + \mathcal{U} = \mathbf{B} + \mathcal{U}$ by above relation.

4. Let $\mathbf{B}, \mathbf{C} \in \mathbf{A} + \mathcal{U}$. Then $\mathbf{B} = \mathbf{A} + \mathbf{X}$ and $\mathbf{C} = \mathbf{A} + \mathbf{Y}$ so that $\mathbf{B} - \mathbf{C} = \mathbf{X} - \mathbf{Y} \in \mathcal{U}$. □

Theorem 42. Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation and $\mathbf{C} \in \operatorname{Im} T$. Define

$$\mathcal{P} = \left\{\mathbf{D} \in \mathcal{V} : T(\mathbf{D}) = \mathbf{C}\right\}.$$

If $\mathbf{A}$ is any vector in $\mathcal{V}$ with $T(\mathbf{A}) = \mathbf{C}$ then $\mathcal{P} = \mathbf{A} + \operatorname{Ker} T$.

Proof. Let $\mathbf{A}$ be any vector in $\mathcal{V}$ with $T(\mathbf{A}) = \mathbf{C}$ and let $\mathbf{B} \in \mathbf{A} + \operatorname{Ker} T$. Then $\mathbf{B} = \mathbf{A} + \mathbf{X}, \mathbf{X} \in \operatorname{Ker} T$. Then,

$$T(\mathbf{B}) = T(\mathbf{A}) = \mathbf{C}$$

showing $\mathbf{B} \in \mathcal{P}$ so that $\mathbf{A} + \operatorname{Ker} T \subseteq \mathcal{P}$.
Conversely, suppose $\mathbf{B} \in \mathcal{P}$. Then $T(\mathbf{B}) = \mathbf{C}$. Then,

$$T(\mathbf{B} - \mathbf{A}) = \mathbf{C} - \mathbf{C} = \mathbf{0}$$

showing $\mathbf{B} - \mathbf{A} \in \operatorname{Ker} T$ i.e. $\mathbf{B} \in \mathbf{A} + \operatorname{Ker} T$ so that $\mathcal{P} \subseteq \mathbf{A} + \operatorname{Ker} T$. $\square$

Theorem 43. Let $\mathcal{S}$ be the solution set of the system $AX = B$. Then $\mathcal{S}$ is either empty or an affine subspace of $\mathbb{R}^n$. Precisely, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation whose matrix relative to the standard bases is A, then either $\mathcal{S} = \emptyset$ or $\mathcal{S}$ is a parallel translate of $\operatorname{Ker} T$ by some vector $\mathbf{S}$.

Remark. To solve a system of equations $AX = B$ completely, an affine subspace of $\mathbb{R}^n$ needs to be specified. As this affine subspace is a parallel translate of $\operatorname{Ker} T$, we may consider the system

$$AX = 0$$

and obtain a basis for the solution of this system (called the homogeneous system). Denote this by $\mathcal{S}_c$, the complementary solution space. Next find a vector $\mathbf{S}_p \in \mathbb{R}^n$ for which $A\mathbf{S}_p = B$. This is a particular solution of the system $AX = B$. Therefore the solution set $\mathcal{S}$ of the system $AX = B$ is completely determined by the relation

$$\mathcal{S} = \mathbf{S}_p + \mathcal{S}_c$$

where $\mathcal{S}_c$ is the basis for $\operatorname{Ker} T$.

Example 37. Specify the solution for the following system of equations

$$\begin{aligned} x_1 + x_2 + x_3 &= 3, \\ x_1 - 2x_2 &= -1. \end{aligned} \tag{8}$$

We first consider the associated homogeneous system $AX = 0$ which is

$$\begin{aligned} x_1 + x_2 + x_3 &= 0, \\ x_1 - 2x_2 &= 0. \end{aligned}$$

From these we get

$$x_2 = \frac{1}{2}x_1,$$
$$x_3 = -\frac{3}{2}x_1.$$

Set $x_1 = 2$. Then $x_2 = 1, x_3 = -3$. Thus the solution set for the homogeneous system of (8) is spanned by the vector $(2, 1, -3)$. Therefore,

$$\mathcal{S}_c = \mathcal{L}((2, 1, -3)).$$

A particular solution of (8) is $\mathbf{S}_p = (1, 1, 1)$. Thus the complete solution $\mathcal{S}$ of (8) is given as

$$\mathcal{S} = (1, 1, 1) + \mathcal{L}((2, 1, -3)).$$

6.3 Echelon Forms

Definition 18. An **echelon matrix** is a matrix $A = (a_{ij})$ where the leading coefficient (the first nonzero entry) in each row is a 1 and it appears to the right of the leading coefficient in the previous row.

Example 38. The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ is an echelon matrix.

Definition 19. $A = (a_{ij})$ is called a **reduced echelon matrix** if all the entries below and above the leading coefficient are zero, i.e. the first non-zero entry in each row is the only non-zero entry in that column.

Example 39. The matrix $\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ is a reduced echelon matrix.

Theorem 44. Consider the system

$$AX = B.$$

By a finite sequence of the operations:

1. interchanging two equations,

2. multiplying an equation by a non-zero scalar,

3. adding two equations

we may obtain a system of equations

$$\bar{A}\bar{X} = \bar{B}$$

where $\bar{A}$ is in reduced-echelon form.

Definition 20. The **augmented matrix** of the system $AX = B$ is the matrix $[A \mid B]$, i.e. the matrix

$$\begin{bmatrix} a_{11} & \dots & a_{1n} & \mid & b_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & \mid & b_m \end{bmatrix}.$$

Example 40. Solve the following system of equations using reduced echelon form:

$$2x_1 + 3x_2 - x_3 = 7,$$
$$3x_1 + 2x_2 + 2x_3 = 5,$$
$$4x_1 - x_2 + 3x_3 = 1.$$

To solve this system, we can represent it as an augmented matrix and perform row operations to reduce it to its reduced echelon form:

$$\begin{bmatrix} 2 & 3 & -1 & \mid & 7 \\ 3 & 2 & 2 & \mid & 5 \\ 4 & -1 & 3 & \mid & 1 \end{bmatrix}.$$

We'll start by applying row operations to create zeros below the leading coefficients:

1. $R_1 \leftarrow \frac{1}{2}R_1$
$$\begin{bmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \mid & \frac{7}{2} \\ 3 & 2 & 2 & \mid & 5 \\ 4 & -1 & 3 & \mid & 1 \end{bmatrix}.$$

2. $R_2 \leftarrow R_2 - 3R_1$
$$\begin{bmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \mid & \frac{7}{2} \\ 0 & -\frac{5}{2} & \frac{7}{2} & \mid & -\frac{11}{2} \\ 4 & -1 & 3 & \mid & 1 \end{bmatrix}.$$

3. $R_3 \leftarrow R_3 - 4R_1$

$$\begin{bmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & | & \frac{7}{2} \\ 0 & -\frac{5}{2} & \frac{7}{2} & | & -\frac{11}{2} \\ 0 & -7 & 5 & | & -13 \end{bmatrix}.$$

4. $R_2 \leftarrow -\frac{2}{5}R_2,\ R_1 \leftarrow R_1 - \frac{3}{2}R_1,\ R_3 \leftarrow R_3 + 7R_2$

$$\begin{bmatrix} 1 & 0 & -\frac{8}{5} & | & \frac{1}{5} \\ 0 & 1 & -\frac{7}{5} & | & \frac{11}{5} \\ 0 & 0 & -\frac{24}{5} & | & \frac{12}{5} \end{bmatrix}.$$

5. $R_3 \leftarrow -\frac{5}{24}R_3,\ R_1 \leftarrow R_1 - \frac{8}{5}R_3$

$$\begin{bmatrix} 1 & 0 & 0 & | & \frac{1}{5} \\ 0 & 1 & -\frac{7}{5} & | & \frac{11}{5} \\ 0 & 0 & 1 & | & -\frac{1}{2} \end{bmatrix}.$$

6. $R_2 \leftarrow R_2 + \frac{7}{5}R_3$

$$\begin{bmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & \frac{3}{2} \\ 0 & 0 & 1 & | & -\frac{1}{2} \end{bmatrix}.$$

Now, the system is in reduced echelon form and we can read the solution to the system of equations as $x_1 = 1$, $x_2 = \frac{3}{2}$, and $x_3 = -\frac{1}{2}$.

7 Special Transformations

7.1 Projection Maps

Definition 21. A linear transformation $P : \mathcal{V} \to \mathcal{V}$ is called a **projection** if $P^2 = P$, i.e. $P(P(\mathbf{A})) = \mathbf{A}$ for all $\mathbf{A} \in \mathcal{V}$.

Theorem 45. Let $\mathcal{V}$ be a finite-dimensional vector space of dimension n and $S : \mathcal{V} \to \mathcal{V}$ be a projection. Then there is a basis $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ for $\mathcal{V}$ such that,

$$S(\mathbf{A}_i) = \begin{cases} \mathbf{A}_i, \ 1 \le i \le r, \\ 0, \quad r+1 \le i \le n, \end{cases}$$

where $r = \dim \operatorname{Im} S$ and hence the matrix of S relative to the basis $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{bmatrix}.$$

Proof. Let $\{\mathbf{B}_1, \ldots, \mathbf{B}_r\}$ be a basis for $\operatorname{Im} S$ and $\{\mathbf{C}_1, \ldots, \mathbf{C}_s\}$ be a basis for $\operatorname{Ker} S$. Therefore $\dim \mathcal{V} = n = r + s$. We will show that the set $\{\mathbf{B}_1, \ldots, \mathbf{B}_r, \mathbf{C}_1, \ldots, \mathbf{C}_s\}$ is a basis for $\mathcal{V}$. Observe that

$$S(C_i) = 0, i = 1, \ldots, s. \tag{9}$$

Since $\mathbf{B}_i \in \operatorname{Im} S$, there exists $\mathbf{D}_i \in \mathcal{V}$ such that $S(\mathbf{D}_i) = \mathbf{B}_i$. Therefore,

$$S(\mathbf{B}_i) = S(S(\mathbf{D}_i)) = S^2(\mathbf{D}_i) = S(\mathbf{D}_i) = \mathbf{B}_i. \tag{10}$$

for $i = 1, \ldots, r$ since S is a projection. By Theorem 17, we only need to show that the set $\{\mathbf{B}_1, \ldots, \mathbf{B}_r, \mathbf{C}_1, \ldots, \mathbf{A}_s\}$ is linearly independent. To that end, consider the linear relation,

$$b_1\mathbf{B}_1 + \cdots + b_r\mathbf{B}_r + c_1\mathbf{C}_1 + \cdots + c_s\mathbf{C}_s = 0.$$

Applying S on both sides and using (9), (10) we obtain,

$$b_1\mathbf{B}_1 + \cdots + b_r\mathbf{B}_r = 0.$$

Since $\{\mathbf{B}_1, \ldots, \mathbf{B}_r\}$ is a basis for $\operatorname{Im} S$, we conclude that $b_i = 0, i = 1, \ldots, r$. We are now left with

$$c_1\mathbf{C}_1 + \cdots + c_s\mathbf{C}_s = 0.$$

Again since $\{\mathbf{C}_1, \ldots, \mathbf{C}_s\}$ is a basis for $\operatorname{Ker} S$, so $c_i = 0, i = 1, \ldots, s$. Thus, we have shown that the set $\{\mathbf{B}_1, \ldots, \mathbf{B}_r, \mathbf{C}_1, \ldots, \mathbf{C}_s\}$ is a basis for $\mathcal{V}$. Now, set $\mathbf{A}_1 = \mathbf{B}_1, \ldots, \mathbf{A}_r = \mathbf{B}_r, \mathbf{A}_{r+1} = \mathbf{C}_1, \ldots, \mathbf{A}_n = \mathbf{C}_s$. The, we will have,

$$S(\mathbf{A}_i) = \begin{cases} \mathbf{A}_i, & 1 \leq i \leq r, \\ 0, & r+1 \leq i \leq n, \end{cases}$$

as required. □

7.2 Nilpotent Transformations

Definition 22. A linear transformation $T : \mathcal{V} \to \mathcal{V}$ is **nilpotent of index** k if $T^k = \mathbf{0}$ and $T^{k-1} \neq \mathbf{0}$, i.e. $T^k(\mathbf{A}) = \mathbf{0}$ for all $\mathbf{A} \in \mathcal{V}$ but there exists at least one $\mathbf{B} \in \mathcal{V}$ such that $T^{k-1}(\mathbf{B}) \neq \mathbf{0}$.

Example 41. Let $D : P_n(\mathbb{R}) \to P_n(\mathbb{R})$ be the differentiation operator. Then the matrix of D relative to the standard bases is

$$\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

which is nilpotent of index $n + 1$.

Theorem 46. Let $T : \mathcal{V} \to \mathcal{V}$ be a nilpotent transformation of index k. Let $\mathbf{B} \in \mathcal{V}$ be a vector such that $T^{k-1}(\mathbf{B}) \neq \mathbf{0}$. Then the set

$$\left\{\mathbf{B}, T(\mathbf{B}), T^2(\mathbf{B}), \ldots, T^{k-1}(\mathbf{B})\right\}$$

is linearly independent. Hence $k \leq n = \dim \mathcal{V}$.

Proof. Consider the linear relation

$$b_0\mathbf{B} + b_1T(\mathbf{B}) + b_2T^2(\mathbf{B}) + \cdots + b_{k-1}T^{k-1}(\mathbf{B}) = \mathbf{0}.$$

Since $T^k(\mathbf{B}) = 0$, we have $T^{k+1}(\mathbf{B}) = T^{k+2}(\mathbf{B}) = \cdots = 0$. Applying T^{k-1} both sides in the above relation, we obtain:

$$b_0T^{k-1}(\mathbf{B}) + b_1T^k(\mathbf{B}) + b_2T^{k+1}(\mathbf{B}) + \cdots + b_{k-1}T^{2k-2}(\mathbf{B}) = \mathbf{0}$$

which gives $b_0T^{k-1}(\mathbf{B}) = \mathbf{0}$ so that $b_0 = 0$ since $T^{k-1}(\mathbf{B}) \neq \mathbf{0}$.
Similarly, applying $T^{k-i}, i = 2, \ldots, k-1$ we can show $b_{i-1} = 0$ which shows linear independence. □

Theorem 47. If $k = \dim \mathcal{V}$, then

$$\left\{\mathbf{B}, T(\mathbf{B}), T^2(\mathbf{B}), \ldots, T^{k-1}(\mathbf{B})\right\}$$

is a basis for $\mathcal{V}$.

Proof. Apply Theorem 46 and Theorem 17. □

7.3 Cyclic Transformations

Definition 23. A linear transformation $T : \mathcal{V} \to \mathcal{V}$ is **cyclic** if there exists a vector $\mathbf{A} \in \mathcal{V}$ such that the collection

$$\left\{\mathbf{A}, T(\mathbf{A}), T^2(\mathbf{A}), \ldots\right\}$$

spans $\mathcal{V}$. The vector $\mathbf{A}$ is called the **cyclic vector** for T.

An immediate consequence of the definition is the following result.

Theorem 48. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation. A vector $\mathbf{A}$ is cyclic vector of T if and only if $T(\mathbf{A}) \notin \mathcal{L}(\mathbf{A})$.

Theorem 49. Let $T : \mathcal{V} \to \mathcal{V}$ be a cyclic transformation and $\dim \mathcal{V} = n$. Let $\mathbf{A} \in \mathcal{V}$ be a cyclic vector of T. Then,

$$\left\{\mathbf{A}, T(\mathbf{A}), T^2(\mathbf{A}), \ldots, T^{n-1}(\mathbf{A})\right\}$$

is a basis for $\mathcal{V}$.

Proof. By Theorem 17, it suffices to show that the vectors $\mathbf{A}, T(\mathbf{A}), T^2(\mathbf{A}), \ldots, T^{n-1}(\mathbf{A})$ are linearly independent. Suppose on the contrary that they are linearly dependent.

By Theorem 9, there exists $m \leq n-1$ such that

$$T^m(\mathbf{A}) \in \mathcal{L}(\mathbf{A}, T(\mathbf{A}), \dots, T^{m-1}(\mathbf{A})).$$

Applying T both sides, we get

$$T^{m+1}(\mathbf{A}) \in \mathcal{L}(T(\mathbf{A}), T^2(\mathbf{A}), \dots, T^m(\mathbf{A})) \subseteq \mathcal{L}(\mathbf{A}, T(\mathbf{A}), \dots, T^{m-1}(\mathbf{A})).$$

Applying T again, we see that

$$T^{m+2}(\mathbf{A}) \in \mathcal{L}(T^2(\mathbf{A}), T^3(\mathbf{A}), \dots, T^{m+1}(\mathbf{A})) \subseteq \mathcal{L}(\mathbf{A}, T(\mathbf{A}), \dots, T^{m-1}(\mathbf{A})).$$

Continuing this process, we observe that

$$T^m(\mathbf{A}), T^{m+1}(\mathbf{A}), T^{m+2}(\mathbf{A}), \dots \in \mathcal{L}(\mathbf{A}, T(\mathbf{A}), \dots, T^{m-1}(\mathbf{A}))$$

so that

$$\mathcal{V} = \mathcal{L}(\mathbf{A}, T(\mathbf{A}), \dots, T^m(\mathbf{A}), T^{m+1}(\mathbf{A}), \dots) \subseteq \mathcal{L}(\mathbf{A}, T(\mathbf{A}), \dots, T^{m-1}(\mathbf{A}))$$

showing that $\dim \mathcal{V} = m \leq n$ which is a contradiction. □

Theorem 50. A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is cyclic if and only if $T \neq eI$ where $e \in \mathbb{R}$ and $I : \mathbb{R}^2 \to \mathbb{R}^2$ is the identity transformation.

Proof.

$$\begin{aligned}
&\text{T is not cyclic} \\
\iff & T(\mathbf{A}) \in \mathcal{L}(\mathbf{A}),\ \forall\, \mathbf{A} \in \mathcal{V}, \text{ by Theorem 48} \\
\iff & T(1,0) = e_1(1,0),\ \ T(0,1) = e_2(0,1),\ \ T(1,1) = e(1,1) \\
\iff & (e,e) = e_1(1,0) + e_2(0,1) = (e_1, e_2) \implies e = e_1 = e_2 \\
\iff & T(1,0) = (e,0),\ \ T(0,1) = (0,e) \\
\iff & \text{Matrix of } T \text{ relative to the standard basis is} \begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix} = eI \\
\iff & T = eI.
\end{aligned}$$

□

8 The Theory of Eigen Values and Eigen Vectors

Throughout this chapter, $\mathcal{V}$ will denote a finite-dimensional vector space.

8.1 Rank and Nullity

Definition 24. An **endomorphism** is a linear transformation $T : \mathcal{V} \to \mathcal{V}$ from a vector space $\mathcal{V}$ to itself.

Theorem 51. Let $T : \mathcal{V} \to \mathcal{V}$ be an endomorphism. Then there exists bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ for $\mathcal{V}$ such that the matrix of T is

$$\begin{array}{c} \left[\begin{array}{cccccc} 1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 1\end{array}}\right\} k \\ \left.\vphantom{\begin{array}{c}0\\ \ddots\\ 0\end{array}}\right\} n-k \end{array}$$

The integer k is called the **rank** of the linear transformation and the integer $n-k$ is called the **nullity** of the linear transformation.

Proof. Let $\{\mathbf{C}_1, \ldots, \mathbf{C}_m\}$ be a basis for $\operatorname{Ker} T$. By Theorem 15, we may find vectors $\mathbf{C}_{m+1}, \ldots, \mathbf{C}_n$ such that $\{\mathbf{C}_1, \ldots, \mathbf{C}_m, \mathbf{C}_{m+1}, \ldots, \mathbf{C}_n\}$ is a basis for $\mathcal{V}$ as $\dim \mathcal{V} = n$.
We claim that the vectors $T(\mathbf{C}_{m+1}), \ldots, T(\mathbf{C}_n)$ are linearly independent. To that end, consider the linear relation

$$c_{m+1}T(\mathbf{C}_{m+1}) + \cdots + c_n T(\mathbf{C}_n) = \mathbf{0}.$$

Using the properties of linear transformations and setting $\mathbf{C} = c_{m+1}\mathbf{C}_{m+1} + \cdots + c_n\mathbf{C}_n$, we see $T(\mathbf{C}) = \mathbf{0}$ implying $\mathbf{C} \in \operatorname{Ker} T$. Since $\{\mathbf{C}_1, \ldots, \mathbf{C}_m\}$ is a basis for $\operatorname{Ker} T$ we may write

$$\mathbf{C} = c_1\mathbf{C}_1 + \cdots + c_m\mathbf{C}_m.$$

Therefore,

$$c_1\mathbf{C}_1 + \cdots + c_m\mathbf{C}_m - c_{m+1}\mathbf{C}_{m+1} - \cdots - c_n\mathbf{C}_n = \mathbf{0}.$$

Since $\{\mathbf{C}_1, \ldots, \mathbf{C}_m, \mathbf{C}_{m+1}, \ldots, \mathbf{C}_n\}$ is a basis for $\mathcal{V}$, we must have $c_i = 0, i = 1, \ldots, m, m+$

$1, \ldots, n$. Now,

$$\begin{aligned}\operatorname{Im} T = T(\mathcal{V}) &= T(\mathcal{L}(\mathbf{C}_1, \ldots, \mathbf{C}_m, \mathbf{C}_{m+1}, \ldots \mathbf{C}_n)) \\ &= \mathcal{L}(T(\mathbf{C}_1), \ldots, T(\mathbf{C}_m), T(\mathbf{C}_{m+1}), \ldots T(\mathbf{C}_n)) \quad \text{(by Theorem 22)} \\ &= \mathcal{L}(\mathbf{0}, \ldots, \mathbf{0}, T(\mathbf{C}_{m+1}), \ldots T(\mathbf{C}_n)) \\ &= \mathcal{L}(T(\mathbf{C}_{m+1}), \ldots T(\mathbf{C}_n)).\end{aligned}$$

Hence the vectors $T(\mathbf{C}_{m+1}), \ldots, T(\mathbf{C}_n)$ form a basis for $\operatorname{Im} T$. Again by Theorem 15 there exists m vectors $\mathbf{D}_1, \ldots, \mathbf{D}_m$ such that $\{\mathbf{D}_1, \ldots, \mathbf{D}_m, T(\mathbf{C}_{m+1}), \ldots, T(\mathbf{C}_n)\}$ form a basis for $\mathcal{V}$. Set

$$\begin{array}{ccccccc} \mathbf{A}_1 & = & \mathbf{C}_{m+1}, & & \mathbf{B}_1 & = & T(\mathbf{C}_{m+1}) \;, \\ \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{A}_{n-m} & = & \mathbf{C}_n & & \mathbf{B}_{n-m} & = & T(\mathbf{C}_n), \\ \mathbf{A}_{n-m+1} & = & \mathbf{C}_1, & & \mathbf{B}_{n-m+1} & = & \mathbf{D}_1, \\ \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{A}_n & = & \mathbf{C}_m, & & \mathbf{B}_n & = & \mathbf{D}_m. \end{array}$$

Then by this construction,

$$T(\mathbf{A}_i) = \begin{cases} \mathbf{B}_i, & i = 1, \ldots, n-m, \\ 0, & i = n-m+1, \ldots, n. \end{cases}$$

Set $k = n - m$. Therefore, the matrix of T relative to the bases $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ and $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ is

$$\left[\begin{array}{cccccc} 1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 1\end{array}}\right\} k \\ \left.\vphantom{\begin{array}{c}0\\ \ddots\\ 0\end{array}}\right\} n-k \end{array} .$$

Since $\dim \operatorname{Im} T = n - m$ and $\dim \operatorname{Ker} T = m$, we see that $k = n - m = \dim \operatorname{Im} T$ is the rank of the linear transformation. □

Remark. From the above conclusion, it is clear that

$$\dim \mathcal{V} = \text{rank} + \text{nullity}.$$

Definition 25. For a matrix A, it's **row (column) rank** is defined as the maximum number of linearly independent rows (columns) of A.

Theorem 52. For a matrix A, row rank = column rank.

Proof. Obvious from Theorem 51. ◻

Example 42. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 0 \\ 2 & 4 & 6 \end{bmatrix}$. Then row rank = column rank = rank = 2.

8.2 Eigen Values and Eigen Vectors

Matrix representations of a linear transformation $T : \mathcal{V} \to \mathcal{V}$ do provide some useful information about the nature of the transformation but do not reveal the complete structure. The complete information can be obtained if we were to find a basis with respect to which the matrix of T is diagonal. Needless to say, the basis for both domain and range must be the same.
The behaviour is in some sense a "non-cyclic" one. That is, given a basis $\{\mathbf{E}_1, \ldots, \mathbf{E}_n\}$ for $\mathcal{V}$, i.e. $\dim \mathcal{V} = n$, the matrix of T relative to this basis will be a diagonal matrix only if $T(\mathbf{E}_i) \in \mathcal{L}(\mathbf{E}_i)$, (opposite to what we saw in Theorem 48). This means, there are numbers e_i such that

$$T(\mathbf{E}_i) = e_i \mathbf{E}_i.$$

This motivates the following definition.

Definition 26. Let $T : \mathcal{V} \to \mathcal{V}$ be an endomorphism. A number e is an **eigen value** of T if there exists a non-zero vector $\mathbf{E}$ such that

$$T(\mathbf{E}) = e\mathbf{E}.$$

Such a vector $\mathbf{E}$ is called an **eigen vector of T associated with the eigen value** e.

Theorem 53. Let $T : \mathcal{V} \to \mathcal{V}$ be an ednomorphism. Then T is represented by a diagonal matrix if and only if $\mathcal{V}$ has a basis consisting of eigen vectors of T.

Proof. Let $\{\mathbf{E}_1, \ldots, \mathbf{E}_n\}$ be a basis for $\mathcal{V}$. Then

T is represented by a diagonal matrix

$\iff$ The matrix of T relative to the basis $\{\mathbf{E}_1, \ldots, \mathbf{E}_n\}$ is of the form $\begin{bmatrix} e_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & e_n \end{bmatrix}$

$\iff T(\mathbf{E}_i) = e_i \mathbf{E}_i$

$\iff \{\mathbf{E}_1, \ldots, \mathbf{E}_n\}$ are eigen vectors of T. □

Definition 27. An endomorphism $T : \mathcal{V} \to \mathcal{V}$ is **diagonalizable** if there exists a basis with respect to which the matrix of T is diagonal.

Theorem 54. The following results hold:

1. The only eigen values of a projection map P are 0 and 1.

2. The only eigen value of a nilpotent transformation T is 0.

Proof. Let e denote an eigen value and $\mathbf{E}$ it's corresponding eigen vector.

1. Then $P(\mathbf{E}) = e\mathbf{E}$. Also since P is a projection, we see that

$$P(E) = P(P(\mathbf{E})) = P(e\mathbf{E}) = eP(\mathbf{E}) = e^2\mathbf{E}.$$

Subtracting, we obtain $(e^2 - e)\mathbf{E} = 0$. Since $\mathbf{E} \neq \mathbf{0}$, we conclude $e = 0, 1$.

2. Then $T(\mathbf{E}) = e\mathbf{E}$. Therefore

$$T^2(\mathbf{E}) = T(T(\mathbf{E})) = T(e\mathbf{E}) = eT(\mathbf{E}) = e^2\mathbf{E}.$$

Continuing this way k times, $0 = T^k(\mathbf{E}) = e^k\mathbf{E}$ whence $e = 0$. □

Definition 28. An endomorphism $T : \mathcal{V} \to \mathcal{V}$ is **singular** if it is not an isomorphism.

Theorem 55. Let $T : \mathcal{V} \to \mathcal{V}$ be an endomorphism. Then e is an eigen value of T if and only if $T - eI$ is singular.

Proof.

$$\begin{aligned}
& e \text{ is an eigen value of } T \\
\iff & T(\mathbf{E}) = e\mathbf{E}, \mathbf{E} \neq \mathbf{0} \\
\iff & (T - eI)(\mathbf{E}) = \mathbf{0} \\
\iff & \mathbf{E} \in \operatorname{Ker}(T - eI) \\
\iff & \operatorname{Ker}(T - eI) \neq \{\mathbf{0}\} \\
\iff & T - eI \text{ is not an isomorphism} \\
\iff & T - eI \text{ is singular.}
\end{aligned}$$

□

Definition 29. Let $T : \mathcal{V} \to \mathcal{V}$ be an endomorphism. Let e be an eigen value of T. For each e define

$$\mathcal{V}_e = \left\{\mathbf{E} : T(\mathbf{E}) = e\mathbf{E}\right\}.$$

The set $\mathcal{V}_e$ is called the **eigen space of** T associated with e.

Theorem 56. If e is an eigen value of $T : \mathcal{V} \to \mathcal{V}$ then $\mathcal{V}_e = \operatorname{Ker}(T - eI)$.

Proof.

$$\begin{aligned}
\mathcal{V}_e &= \left\{\mathbf{E} : T(\mathbf{E}) = e\mathbf{E}\right\} \\
&= \left\{\mathbf{E} : (T - eI)(\mathbf{E}) = \mathbf{0}\right\} \\
&= \operatorname{Ker}(T - eI).
\end{aligned}$$

□

8.3 Characteristic Polynomial

Definition 30. Let $T : \mathcal{V} \to \mathcal{V}$ be a linear transformation whose matrix relative to the basis $\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ is A. The **characteristic polynomial** of T is the polynomial $\Delta(t)$ of degree n defined as $\Delta(t) = \det(A - tI)$.

Remark. The definition of the characteristic polynomial is independent of the choice of basis. Indeed, if $\{\mathbf{B}_1, \ldots, \mathbf{B}_n\}$ is another basis of $\mathcal{V}$ for which the matrix of T is B, then by Theorem 37, there exists an invertible matrix P such that $B = PAP^{-1}$. Then,

$$B - tI = PAP^{-1} - tI = PAP^{-1} - tPIP^{-1} = P(A - tI)P^{-1}.$$

Therefore,

$$\det(B - tI) = \det P \det(A - tI) \det P^{-1} = \det(A - tI).$$

Definition 31. A value e is said to be a **root** of $\Delta(t)$ if $\Delta(e) = 0$.

Theorem 57 (Necessary and Sufficient Condition for Eigen Values)**.** Let $\Delta(t)$ be the characteristic polynomial of $T : \mathcal{V} \to \mathcal{V}$. Then e is an eigen value of T if and only if e is a root of the characteristic polynomial $\Delta(t)$.

Proof.

$$
\begin{aligned}
& e \text{ is an eigen value of } T \\
\iff & T - eI \text{ is singular, (Theorem 35)} \\
\iff & \det(A - eI) = 0 \text{ (Theorem 4)} \\
\iff & \Delta(e) = 0 \text{ (by Definition)} \\
\iff & e \text{ is a root of the characteristic polynomial } \Delta(t).
\end{aligned}
$$

□

Theorem 58. Let $T : \mathcal{V} \to \mathcal{V}$ be an endomorphism. Let $e_1, \ldots, e_m$ be distinct eigen values of T and $\mathbf{F}_1, \ldots, \mathbf{F}_m$ be the corresponding eigen vectors. Then $\mathbf{F}_1, \ldots, \mathbf{F}_m$ are linearly independent.

Proof. Suppose on the contrary that $\mathbf{F}_1, \ldots, \mathbf{F}_m$ are linearly dependent. Then by Theorem 12, there exists $\mathbf{F}_k$ such that $\mathbf{F}_k \in \mathcal{L}(\mathbf{F}_1, \ldots, \mathbf{F}_{k-1})$. Therefore,

$$\mathbf{F}_k = a_1\mathbf{F}_1 + \cdots + a_{k-1}\mathbf{F}_{k-1}. \tag{11}$$

Applying T both sides,

$$T(\mathbf{F}_k) = a_1 T(\mathbf{F}_1) + \cdots + a_{k-1} T(\mathbf{F}_{k-1}).$$

Since F_i are eigen vectors associated with the eigen values e_i, we obtain

$$e_k\mathbf{F}_k = a_1 e_1 \mathbf{F}_1 + \cdots + a_{k-1} e_{k-1} \mathbf{F}_{k-1}. \tag{12}$$

Multiplying Equation (11) by e_k and subtracting from Equation (12) we obtain

$$\mathbf{0} = a_1(e_1 - e_k)\mathbf{F}_1 + \cdots + a_{k-1}(e_{k-1} - e_k)\mathbf{F}_{k-1}.$$

But since $\mathbf{F}_1, \ldots, \mathbf{F}_{k-1}$ are linearly independent, we have

$$a_i(e_i - e_k) = 0, \ i = 1, \ldots, k-1.$$

But since the eigen values are distinct, $e_i \neq e_j, i \neq j$. Therefore $a_i = 0, i = 1, \ldots, k-1$. This shows that $\mathbf{F}_k = \mathbf{0}$ which is a contradiction. Hence the result. □

Theorem 59. Let $T : \mathcal{V} \to \mathcal{V}$ be an endomorphism and $\dim \mathcal{V} = n$. Let the characteristic polynomial has n distinct real roots. Then there is a basis for $\mathcal{V}$ consisting of the eigen vectors $\mathbf{F}_1, \ldots, \mathbf{F}_n$ of T.

Proof. Apply Theorem 57, Theorem 58 and Theorem 17. □

Example 43. Diagonalize the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined as:

$$T(x, y) = (x + 2y, 4x + 3y).$$

We observe that

$$T(1, 0) = (1, 4),$$
$$T(0, 1) = (2, 3).$$

Therefore, the matrix of T relative to the standard bases is

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}.$$

The characteristic polynomial $\Delta(t)$ is given as

$$\Delta(t) = \det(A - tI) = (t - 5)(t + 1).$$

Equating this to 0 gives $t = 5, -1$. Therefore the eigen values are real and distinct. Let us compute the eigen vectors.

1. $t = 5$. In this case, we have the equation

   $$2x - y = 0.$$

   Therefore the eigen space is computed as

   $$\mathcal{V}_5 = \mathcal{L}((1, 2)).$$

   Set $\mathbf{F}_1 = (1, 2)$.

2. $t = -1$. In this case, we have the equation

   $$x + y = 0.$$

Therefore the eigen space is computed as

$$\mathcal{V}_{-1} = \mathcal{L}((1, -1)).$$

Set $\mathbf{F}_2 = (1, -1)$.

The vectors $\mathbf{F}_1, \mathbf{F}_2$ are linearly independent and hence a basis for $\mathbb{R}^2$ relative to which the matrix of T is

$$\begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}.$$