

# Linear Algebra

Semester 1, 2021

*Dr Ravindra Pethiyagoda*

TARANG JANAWALKAR

This work is licensed under a Creative Commons  
“Attribution-NonCommercial-ShareAlike 4.0 International” license.



# Contents

<b>Contents</b>	<b>1</b>
<b>1 Euclidean Vector Spaces</b>	<b>3</b>
1.1 Vectors . . . . .	3
1.2 Position and Displacement Vectors . . . . .	3
1.3 Vector Addition . . . . .	4
1.4 Scalar Multiplication . . . . .	4
1.5 Norm of a Vector . . . . .	4
1.6 The Unit Vector . . . . .	4
1.7 The Dot Product . . . . .	4
1.8 The Cross Product . . . . .	4
<b>2 Vector Identities</b>	<b>5</b>
<b>3 Linear System of Equations</b>	<b>6</b>
3.1 Linear Equations . . . . .	6
3.2 Homogeneous Linear Equations . . . . .	6
3.3 Linear Systems . . . . .	6
3.4 Coefficient Matrices . . . . .	6
3.5 Augmented Matrices . . . . .	6
3.6 Elementary Row Operations . . . . .	7
3.7 Pivots . . . . .	7
3.8 Gaussian Elimination . . . . .	7
3.9 Row-Echelon Form . . . . .	7
3.10 Gauss-Jordan Elimination . . . . .	7
3.11 Solutions to Linear Systems . . . . .	8
<b>4 Matrices</b>	<b>8</b>
4.1 Matrix Addition . . . . .	8
4.2 Scalar Multiplication . . . . .	8
4.3 Matrix Multiplication . . . . .	8
4.4 The Identity Matrix . . . . .	9
4.5 The Inverse Matrix . . . . .	9
4.6 The Diagonal Matrix . . . . .	9
4.7 Matrix Transpose . . . . .	9
4.8 Matrix Trace . . . . .	9
<b>5 General Vector Spaces</b>	<b>10</b>
5.1 Real Vector Spaces . . . . .	10
5.2 Subspaces . . . . .	11
5.3 Spanning Sets . . . . .	12
5.4 Linear Independence . . . . .	12
5.5 Basis Vectors . . . . .	12
5.6 Dimension . . . . .	12

<b>6</b>	<b>Fundamental Subspaces</b>	<b>13</b>
6.1	The Four Fundamental Subspaces of a Matrix . . . . .	13
6.2	The General Solution of a System of Equations . . . . .	13
6.3	Row Equivalence . . . . .	14
6.4	Rank . . . . .	14
6.5	Nullity . . . . .	14
<b>7</b>	<b>Orthogonality</b>	<b>14</b>
7.1	Orthogonal Subspaces . . . . .	14
7.2	Orthogonal Complements . . . . .	15
7.3	Vector Projections . . . . .	15
7.4	Projection onto a Subspace . . . . .	16
7.5	Least Squares . . . . .	16
<b>8</b>	<b>Linear Maps</b>	<b>17</b>
8.1	Matrix Transformations . . . . .	17
8.2	General Linear Transformations . . . . .	17
8.3	Subspaces of Linear Transformations . . . . .	18
8.4	Constructing a Transformation Matrix . . . . .	19
<b>9</b>	<b>Determinants</b>	<b>19</b>
9.1	Properties of Determinants . . . . .	19
9.2	Matrix Minors . . . . .	19
9.3	Matrix Cofactors . . . . .	20
9.4	The Determinant of a Matrix . . . . .	20
9.5	The Cofactor Matrix . . . . .	20
9.6	The Adjugate of a Matrix . . . . .	20
9.7	The Inverse of a Matrix . . . . .	20
<b>10</b>	<b>Invariant Subspaces</b>	<b>20</b>
10.1	Trivial Invariant Subspaces . . . . .	21
10.2	Eigenspaces . . . . .	21
10.3	The Eigenvalue Problem . . . . .	21
10.4	Properties of Eigenvalues . . . . .	21
<b>11</b>	<b>Eigen Decomposition</b>	<b>22</b>
11.1	Similarity Transformations . . . . .	22
11.2	Matrix Diagonalisation . . . . .	22
11.3	Powers of a Matrix . . . . .	23
<b>12</b>	<b>System of Differential Equations</b>	<b>23</b>
12.1	First-Order Differential Equations . . . . .	23
12.2	First-Order System of Differential Equations . . . . .	23
12.3	Solution using Diagonalisation . . . . .	24
12.4	Principle of Superposition . . . . .	24
12.5	Higher-Order Differential Equations . . . . .	25

# 1 Euclidean Vector Spaces

## 1.1 Vectors

**Definition 1.1.** An  $n$ -dimensional **vector** is an ordered list of  $n$  numbers.

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$$

**Theorem 1.1.1.**  $\mathbb{R}^n$  is the set of all ordered  $n$ -tuples of real numbers.

$$\mathbb{R}^n = \{(v_1, v_2, \dots, v_n) : v_1, v_2, \dots, v_n \in \mathbb{R} : n \in \mathbb{N}\}$$

Notation:

1. Component form:  $\mathbf{v} = \langle v_1, v_2 \rangle = (v_1, v_2) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
2. Unit vector form:  $\mathbf{v} = v_1 \hat{\mathbf{i}} + v_2 \hat{\mathbf{j}}$ , where  $\hat{\mathbf{i}}$  and  $\hat{\mathbf{j}}$  are basis vectors along the  $x$  and  $y$  axes respectively.
3. Denotation:  $\mathbf{v} = v = \vec{v}$

## 1.2 Position and Displacement Vectors

**Definition 1.2.** The displacement vector  $\overrightarrow{AB}$  from  $\mathbf{a}$  to  $\mathbf{b}$  can be defined as  $\mathbf{b} - \mathbf{a}$ .

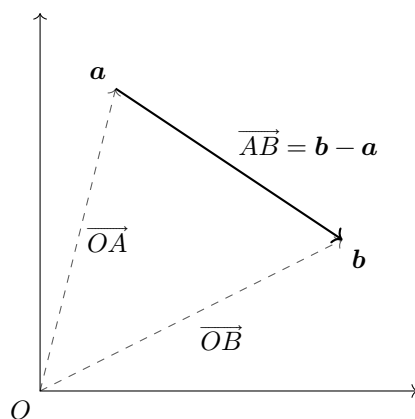


Figure 1: Displacement vector between two points.

### 1.3 Vector Addition

**Definition 1.3.** **Vector addition** is performed by adding the corresponding components of two vectors of the same dimension.

$$\mathbf{a} + \mathbf{b} = \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

### 1.4 Scalar Multiplication

**Definition 1.4.** **Scalar multiplication** is performed by multiplying each element of the vector by the scalar.

$$a\mathbf{v} = \begin{bmatrix} av_1 \\ \vdots \\ av_n \end{bmatrix}$$

### 1.5 Norm of a Vector

**Definition 1.5.** The **norm** of a vector  $\mathbf{v}$ , denoted by  $\|\mathbf{v}\|$ , is the *length* or *magnitude* of  $\mathbf{v}$ .

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

### 1.6 The Unit Vector

**Definition 1.6.** A **unit vector** is a vector, denoted  $\hat{\mathbf{v}}$ , that has a length of 1 in the direction of  $\mathbf{v}$ .

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

### 1.7 The Dot Product

**Definition 1.7.** The **dot product** is a function that associates each pair of vectors  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  a real number  $\mathbf{v} \cdot \mathbf{w}$ .

$$\begin{aligned} \mathbf{v} \cdot \mathbf{w} &= v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \\ &= \|\mathbf{v}\| \|\mathbf{w}\| \cos(\theta) \end{aligned}$$

where  $\theta$  is the angle between  $\mathbf{v}$  and  $\mathbf{w}$ .

**Theorem 1.7.1.** *If  $\mathbf{v} \cdot \mathbf{w} = 0$  then  $\mathbf{v}$  and  $\mathbf{w}$  are orthogonal.*

### 1.8 The Cross Product

**Definition 1.8.** The **cross product** is a function that associates each ordered pair of vectors  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$  a vector  $\mathbf{v} \times \mathbf{w} \in \mathbb{R}^3$ .

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\ &= \|\mathbf{v}\| \|\mathbf{w}\| \sin(\theta) \hat{\mathbf{n}} \end{aligned}$$

where  $\hat{\mathbf{n}}$  is the normal vector given by the right-hand rule.

## 2 Vector Identities

Let  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$  with  $r \in \mathbb{R}$ .

**Theorem 2.0.1.** *Commutativity of vector addition.*

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

**Theorem 2.0.2.**

$$\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$$

**Theorem 2.0.3.** *Commutativity of dot products.*

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

**Theorem 2.0.4.** *Distributivity of dot products over vector addition.*

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

**Theorem 2.0.5.** *Associativity of dot products over scalar multiplication.*

$$(r\mathbf{a}) \cdot \mathbf{b} = r(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (r\mathbf{b})$$

**Theorem 2.0.6.** *Bilinearity of dot products.*

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

**Theorem 2.0.7.**

$$\mathbf{a} \times \mathbf{a} = \mathbf{0}$$

**Theorem 2.0.8.** *Anticommutativity of cross products.*

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$$

**Theorem 2.0.9.** *Distributivity of cross products over vector addition.*

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

**Theorem 2.0.10.** *Associativity of cross products over scalar multiplication.*

$$(r\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (r\mathbf{b}) = r(\mathbf{a} \times \mathbf{b})$$

**Theorem 2.0.11.**

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$$

**Theorem 2.0.12.**

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$$



### 3.6 Elementary Row Operations

**Definition 3.7.** A linear system can be solved using the following **elementary row operations**:

1. **scalar multiplication**: multiplying any row by a constant
2. **row addition**: adding a multiple of one row to another
3. **row exchange**: exchanging any two rows

### 3.7 Pivots

**Definition 3.8.** The first non-zero entry of the row in a matrix is called the **pivot** of the row.

**Theorem 3.7.1.** *If a row apart from the first has a pivot, then this pivot must be to the right of the pivot in the preceding row.*

### 3.8 Gaussian Elimination

**Definition 3.9. Gaussian elimination** is a method for solving linear systems. These systems can be solved by composing the augmented matrix of a system, and performing elementary row operations, to put the matrix in the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{mn} \end{bmatrix}$$

### 3.9 Row-Echelon Form

**Definition 3.10.** A matrix that has undergone Gaussian elimination is in **row-echelon form** if the pivots of the augmented matrix are all 1.

$$\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ & 1 & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

### 3.10 Gauss-Jordan Elimination

**Definition 3.11. Gauss-Jordan elimination** extends Gaussian elimination so that the entries in a column containing a pivot are zeros, and the pivots are all 1. This new augmented matrix is then in **reduced row-echelon form**.

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



### 3.11 Solutions to Linear Systems

**Definition 3.12.** A **consistent system** of equations has at least one solution, and an **inconsistent system** has no solution.

## 4 Matrices

**Definition 4.1.** A **matrix** is an array of numbers arranged into *rows* and *columns*, and can be used to represent a linear transformation.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

### 4.1 Matrix Addition

**Definition 4.2.** **Matrix addition** is performed by adding the corresponding components of two matrices of the same dimension.

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

### 4.2 Scalar Multiplication

**Definition 4.3.** **Scalar multiplication** is performed by multiplying each element of a matrix by a scalar.

$$c\mathbf{A} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

### 4.3 Matrix Multiplication

**Definition 4.4.** **Matrix multiplication** is performed by multiplying each row in the first matrix by the columns of the second matrix.

$$\mathbf{AB} = \mathbf{C}$$

$$\begin{bmatrix} \text{---} & \mathbf{a}_1 & \text{---} \\ \text{---} & \mathbf{a}_2 & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a}_m & \text{---} \end{bmatrix} \begin{bmatrix} \left| \begin{array}{c} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{array} \right| & \left| \begin{array}{c} \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{array} \right| & \cdots & \left| \begin{array}{c} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{array} \right| \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_n \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_n \end{bmatrix}$$

**Theorem 4.3.1.** A matrix product is defined if and only if the number of columns in the first matrix is equal to the number of rows in the second matrix.

## 4.4 The Identity Matrix

**Definition 4.5.** The **identity matrix** is the simplest nontrivial **diagonal matrix**, denoted  $\mathbf{I}$ , such that

$$\mathbf{I}\mathbf{A} = \mathbf{A}$$

written explicitly as

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

## 4.5 The Inverse Matrix

**Definition 4.6.** The **inverse** of a **square matrix** is a matrix  $\mathbf{A}^{-1}$ , such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

**Theorem 4.5.1.** *The inverse of a  $2 \times 2$  matrix is given by*

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

**Theorem 4.5.2.** *The inverse of an  $n \times n$  matrix can be determined by solving  $[\mathbf{A} \mid \mathbf{I}]$ .*

## 4.6 The Diagonal Matrix

**Definition 4.7.** A **diagonal matrix**, denoted  $\text{diag}(d_{11}, d_{22}, \dots, d_{nn})$ , is an  $n \times n$  matrix  $\mathbf{D}$  in which entries outside the main diagonal are all zero.

$$\mathbf{D} = \text{diag}(d_{11}, d_{22}, \dots, d_{nn}) = \begin{bmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{nn} \end{bmatrix}$$

## 4.7 Matrix Transpose

**Definition 4.8.** The **transpose** of a matrix, denoted by  $\mathbf{A}^\top$ , is obtained by replacing all  $a_{ij}$  elements with  $a_{ji}$ , so that the matrix  $\mathbf{A}$  is flipped over its main diagonal.

## 4.8 Matrix Trace

**Definition 4.9.** The **trace** of an  $n \times n$  matrix  $\mathbf{A}$ , denoted  $\text{Tr}(\mathbf{A})$ , is defined as

$$\text{Tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$$

## 5 General Vector Spaces

### 5.1 Real Vector Spaces

**Definition 5.1.** A **vector space** is a set that is closed under vector addition and scalar multiplication.

**Theorem 5.1.1.** *If the following axioms are satisfied by all objects  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ , and all scalars  $k$  and  $m$ , then  $V$  is a **vector space**, and the objects in  $V$  are vectors.*

**Axiom 1** (Closure under addition).

$$\mathbf{u} + \mathbf{v} \in V$$

**Axiom 2** (Commutativity of vector addition).

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

**Axiom 3** (Associativity of vector addition).

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$$

**Axiom 4** (Additive identity).

$$\mathbf{u} + \mathbf{0} = \mathbf{u}$$

**Axiom 5** (Additive inverse).

$$\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$$

**Axiom 6** (Closure under scalar multiplication).

$$k\mathbf{u} \in V$$

**Axiom 7** (Distributivity of vector addition).

$$k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$$

**Axiom 8** (Distributivity of scalar addition).

$$(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$$

**Axiom 9** (Associativity of scalar multiplication).

$$k(m\mathbf{u}) = (km)\mathbf{u}$$

**Axiom 10** (Scalar multiplication identity).

$$1\mathbf{u} = \mathbf{u}$$

To identify that a set with two operations is a vector space:

1. Identify the set  $V$  of objects that will become vectors.
2. Identify the addition and scalar multiplication operations on  $V$ .

3. Verify Axioms 1 and 6.
4. Confirm that Axioms 2, 3, 4, 5, 7, 8, 9, and 10 hold.

**Theorem 5.1.2.** *Let  $V$  be a vector space. If  $\mathbf{v} \in V$ , and  $k$  is a scalar.*

1.  $0\mathbf{v} = \mathbf{0}$
2.  $k\mathbf{0} = \mathbf{0}$
3.  $(-1)\mathbf{v} = -\mathbf{v}$
4. *If  $k\mathbf{v} = \mathbf{0}$ , then  $k = 0$  or  $\mathbf{v} = \mathbf{0}$*

## 5.2 Subspaces

**Definition 5.2.** A subset  $W$  of a vector space  $V$  is called a **subspace** of  $V$  if  $W$  is itself a vector space under the addition and scalar multiplication operations defined on  $V$ .

**Theorem 5.2.1.** *Let  $W$  be a subspace of the vector space  $V$ , then the following axioms must be satisfied.*

1. **Axiom 1:** *Closure under addition*
2. **Axiom 6:** *Closure under scalar multiplication*

**Theorem 5.2.2.** *Every vector space has at least two subspaces, itself and its zero subspace.*

**Theorem 5.2.3.** *Subspaces of  $\mathbb{R}^2$ .*

1.  $\{\mathbf{0}\}$
2. *Lines through the origin*
3.  $\mathbb{R}^2$

**Theorem 5.2.4.** *Subspaces of  $\mathbb{R}^3$ .*

1.  $\{\mathbf{0}\}$
2. *Lines through the origin*
3. *Planes through the origin*
4.  $\mathbb{R}^3$

**Theorem 5.2.5.** *Subspaces of  $\mathbf{M}_{nn}$ .*

1. *Upper triangular matrices*
2. *Lower triangular matrices*
3. *Diagonal matrices*
4.  $\mathbf{M}_{nn}$

### 5.3 Spanning Sets

**Definition 5.3.** If the vector  $\mathbf{w}$  is in a vector space  $V$ , then  $\mathbf{w}$  is a **linear combination** of the vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ , if  $\mathbf{w}$  can be expressed in the form

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \dots + k_n\mathbf{v}_n$$

**Theorem 5.3.1.** If  $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$  is a nonempty set of vectors in a vector space  $V$ , then the set  $W$  of all possible linear combinations of the vectors in  $S$  is a subspace of  $V$ . The subspace  $W$  is called the subspace of  $V$  **spanned** by  $S$  and the vectors in  $S$  **span**  $W$ . If a vector in  $S$  can be expressed as the linear combination of any vectors in  $S$  then the set is **linearly dependent**.

### 5.4 Linear Independence

**Definition 5.4.** If  $S$  is a set of two or more vectors in a vector space  $V$ , then  $S$  is **linearly independent** if no vector in  $S$  can be expressed as a linear combination of the others.

**Theorem 5.4.1.** A set  $S$  is linearly independent if and only if there is one solution to the equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \dots + k_n\mathbf{v}_n = \mathbf{0}$$

where the coefficients satisfying this equation are  $k_1 = 0, k_2 = 0, \dots, k_n = 0$ .

### 5.5 Basis Vectors

**Definition 5.5.** If  $S$  is a set of vectors in a vector space  $V$ , then  $S$  is called a **basis** for  $V$  if

1.  $S$  spans  $V$ .
2.  $S$  is linearly independent.

### 5.6 Dimension

**Definition 5.6.** The **dimension** of a finite-dimensional vector space  $V$ , denoted  $\dim(V)$ , is the number of vectors in a basis for  $V$ .

**Theorem 5.6.1.** The zero vector space is defined to have dimension zero.

## 6 Fundamental Subspaces

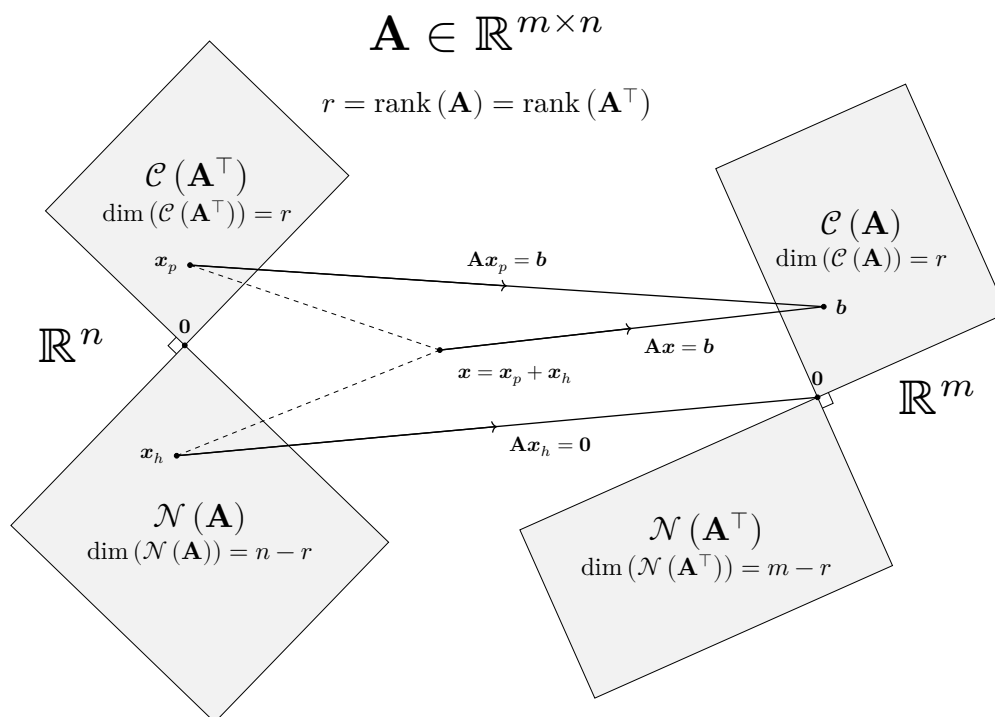


Figure 2: The Four Fundamental Subspaces of a Matrix.

### 6.1 The Four Fundamental Subspaces of a Matrix

**Definition 6.1.** If  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is an  $m \times n$  matrix, then:

1. The subspace spanned by the *column vectors* of  $\mathbf{A}$ , is the **column space** of  $\mathbf{A}$ , denoted  $\mathcal{C}(\mathbf{A})$ .
2. The subspace spanned by the *row vectors* of  $\mathbf{A}$ , is the **row space** of  $\mathbf{A}$ , denoted  $\mathcal{C}(\mathbf{A}^\top)$ .
3. The subspace spanned by the *solution space* of the equation  $\mathbf{A}\mathbf{x} = \mathbf{0}$ , is the **null space** of  $\mathbf{A}$ , denoted  $\mathcal{N}(\mathbf{A})$ .
4. The subspace spanned by the *solution space* of the equation  $\mathbf{A}^\top\mathbf{y} = \mathbf{0}$  (or  $\mathbf{y}^\top\mathbf{A} = \mathbf{0}$ ), is the **left null space** of  $\mathbf{A}$ , denoted  $\mathcal{N}(\mathbf{A}^\top)$ .

### 6.2 The General Solution of a System of Equations

**Theorem 6.2.1.** The *general solution* to a matrix equation  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , can be given by adding the *particular* and *homogeneous* solutions, where the particular solution is the solution to  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , or

$\mathcal{C}(\mathbf{A}^\top)$ , and the homogeneous solution is the solution to  $\mathbf{Ax} = \mathbf{0}$ , or  $\mathcal{N}(\mathbf{A})$ .

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$$

### 6.3 Row Equivalence

**Definition 6.2.** Two matrices are **row equivalent** if each can be obtained from the other by elementary row operations. These matrices have the same row space and null space.

### 6.4 Rank

**Definition 6.3.** The **rank** of a matrix, denoted by  $\text{rank}(\mathbf{A})$ , is given by  $\dim(\mathcal{C}(\mathbf{A}))$ .

**Theorem 6.4.1.** The column space and row space have the same dimension so that

$$\text{rank}(\mathbf{A}) = \dim(\mathcal{C}(\mathbf{A})) = \dim(\mathcal{C}(\mathbf{A}^\top))$$

### 6.5 Nullity

**Definition 6.4.** The **nullity** of a matrix, denoted by  $\text{null}(\mathbf{A})$ , is given by  $\dim(\mathcal{N}(\mathbf{A}))$ .

## 7 Orthogonality

**Definition 7.1.** Two vectors are **orthogonal** if the following holds.

$$\mathbf{v} \cdot \mathbf{w} = 0 \iff \mathbf{v}^\top \mathbf{w} = 0$$

**Theorem 7.0.1.**  $\mathbf{0}$  is orthogonal to every vector in  $V$ .

**Theorem 7.0.2.**  $\mathbf{0}$  is the only vector in  $V$ , that is orthogonal to itself.

**Theorem 7.0.3.**

$$\|\mathbf{v}\|^2 = \mathbf{v}^\top \mathbf{v}$$

**Theorem 7.0.4.**

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}^\top \mathbf{v}}$$

### 7.1 Orthogonal Subspaces

**Definition 7.2.** Two subspaces  $U$  and  $W$  of a vector space  $V$ , are **orthogonal subspaces** iff every vector in  $U$  is orthogonal to every vector in  $W$ .

$$\forall \mathbf{u} \in U : \forall \mathbf{w} \in W : \mathbf{u}^\top \mathbf{w} = 0$$

## 7.2 Orthogonal Complements

**Definition 7.3.** If  $U$  is a subspace of  $V$ , then the **orthogonal complement** of  $U$ , denoted  $U^\perp$ , is the set of all vectors in  $V$  that are orthogonal to every vector in  $U$ .

$$U^\perp = \{\forall \mathbf{u} \in U : \mathbf{v} \in V : \mathbf{v}^\top \mathbf{u} = 0\}$$

**Theorem 7.2.1.**

$$(U^\perp)^\perp = U$$

**Theorem 7.2.2.**

$$\dim U + \dim U^\perp = \dim V$$

## 7.3 Vector Projections

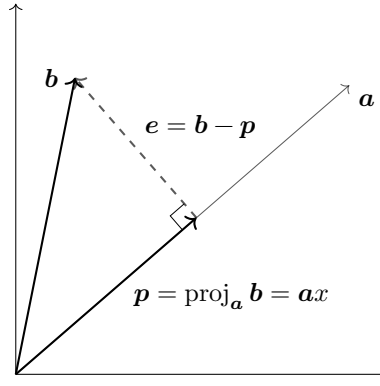


Figure 3: Vector Projection of  $\mathbf{b}$  onto  $\mathbf{a}$ .

**Definition 7.4.** Let the **vector projection** of  $\mathbf{b}$  onto  $\mathbf{a}$ , denoted as  $\text{proj}_{\mathbf{a}} \mathbf{b}$ , be the *orthogonal projection* of  $\mathbf{b}$  in the direction of  $\mathbf{a}$ , that minimises the error vector:  $\mathbf{e} = \mathbf{b} - \mathbf{p}$ .

**Theorem 7.3.1.** The projection of  $\mathbf{b}$  onto  $\mathbf{a}$  is given by

$$\text{proj}_{\mathbf{a}} \mathbf{b} = \mathbf{a}x = \mathbf{a}(\mathbf{a}^\top \mathbf{a})^{-1} \mathbf{a}^\top \mathbf{b}$$

alternatively

$$\text{proj}_{\mathbf{a}} \mathbf{b} = \mathbf{a}x = \mathbf{a} \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} = \mathbf{a} \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}}$$

*Proof.* As  $\mathbf{p}$  lies on line through  $\mathbf{a}$ ,  $\mathbf{p} = \mathbf{a}x$ , so that  $\mathbf{e} = \mathbf{b} - \mathbf{a}x$ . As  $\mathbf{e}$  is orthogonal to  $\mathbf{a}$ , we can construct the following relationship.

$$\begin{aligned} \mathbf{a}^\top \mathbf{e} &= 0 \\ \mathbf{a}^\top (\mathbf{b} - \mathbf{a}x) &= 0 \\ \mathbf{a}^\top \mathbf{b} - \mathbf{a}^\top \mathbf{a}x &= 0 \\ \mathbf{a}^\top \mathbf{a}x &= \mathbf{a}^\top \mathbf{b} \\ x &= (\mathbf{a}^\top \mathbf{a})^{-1} \mathbf{a}^\top \mathbf{b} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \end{aligned}$$



□

## 7.4 Projection onto a Subspace

**Theorem 7.4.1.** *Let  $W$  be a subspace of the vector space  $V$  such that if  $\mathbf{b} \in V$ , then  $\mathbf{p} = \text{proj}_W \mathbf{b}$  is the **best approximation** of  $\mathbf{b}$  on  $W$ , so that*

$$\|\mathbf{b} - \mathbf{p}\| < \|\mathbf{b} - \mathbf{w}\|$$

for all  $\mathbf{w} \in W$ , where  $\mathbf{w} \neq \mathbf{p}$ .

**Theorem 7.4.2.** *The projection of  $\mathbf{b}$  onto the vector space  $W$  is given by*

$$\text{proj}_W \mathbf{b} = \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$$

*Proof.* As  $\mathbf{p} \in W$ ,  $\mathbf{p}$  can be represented as the linear combination of the basis vectors  $\mathbf{a}_i$  that span  $W$ .

$$\begin{aligned} \mathbf{p} &= \hat{x}_1 \mathbf{a}_1 + \hat{x}_2 \mathbf{a}_2 + \cdots + \hat{x}_n \mathbf{a}_n \\ &= \begin{bmatrix} | & | & \cdots & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{bmatrix} \\ &= \mathbf{A}\hat{\mathbf{x}} \end{aligned}$$

Consider the error vector  $\mathbf{e} = \mathbf{b} - \mathbf{p}$ . As  $\mathbf{e}$  is orthogonal to  $W$ , it will also be orthogonal to the vectors that span  $W$ . Therefore

$$\begin{cases} \mathbf{a}_1^\top (\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) = 0 \\ \mathbf{a}_2^\top (\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) = 0 \\ \vdots \\ \mathbf{a}_n^\top (\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) = 0 \end{cases}$$

which gives the following equation

$$\mathbf{A}^\top (\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) = \mathbf{0}$$

where we solve for  $\hat{\mathbf{x}}$

$$\begin{aligned} \mathbf{A}^\top \mathbf{b} - \mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} &= \mathbf{0} \\ \mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} &= \mathbf{A}^\top \mathbf{b} \\ \hat{\mathbf{x}} &= (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b} \end{aligned}$$

□

## 7.5 Least Squares

**Theorem 7.5.1.** *Suppose  $\mathbf{Ax} = \mathbf{b}$  is an inconsistent linear system. The **least squares** solution of  $\mathbf{Ax} = \mathbf{b}$  is given by the orthogonal projection  $\text{proj}_{C(\mathbf{A})} \mathbf{b}$ .*

## 8 Linear Maps

### 8.1 Matrix Transformations

**Definition 8.1.** A **matrix transformation**  $T_{\mathbf{A}} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a mapping of the form

$$T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

where  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . As this transformation is linear, the following linearity properties hold.

1.  $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
2.  $T(k\mathbf{u}) = kT(\mathbf{u})$

### 8.2 General Linear Transformations

**Theorem 8.2.1.** *If  $T : V \rightarrow W$  is a mapping between two vector spaces  $V$  and  $W$ , then  $T$  is the **linear transformation** from  $V$  to  $W$ , and the following properties hold.*

1.  $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
2.  $T(k\mathbf{u}) = kT(\mathbf{u})$

**Theorem 8.2.2.** *When  $V = W$ , the linear map is called a **linear operator**.*

### 8.3 Subspaces of Linear Transformations

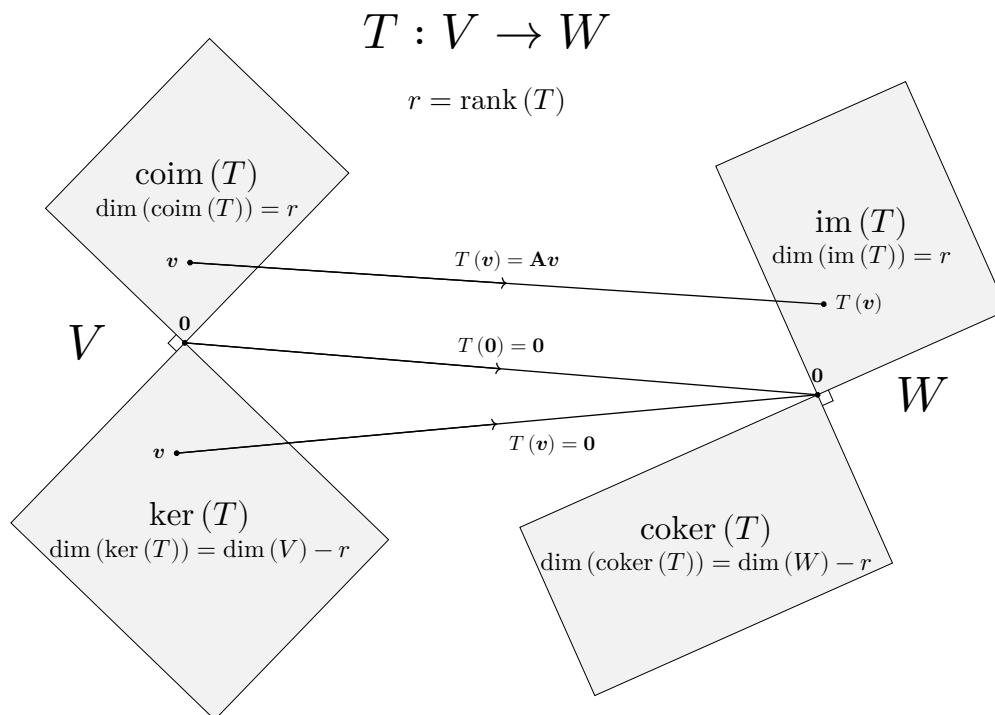


Figure 4: Subspaces of a Linear Transformation.

**Definition 8.2.** If  $T : V \rightarrow W$  is a linear transformation between two vector spaces  $V$  and  $W$ , then:

1. The vector space  $V$  is the **domain** of  $T$ .
2. The vector space  $W$  is the **codomain** of  $T$ .
3. The **image** (or **range**) of  $T$  is the set of vectors the linear transformation maps to.

$$\text{im}(T) = T(V) = \{T(\mathbf{v}) : \mathbf{v} \in V\} \subset W$$

4. The **kernel** of  $T$  is the set of vectors that map to the zero vector.

$$\text{ker}(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$$

## 8.4 Constructing a Transformation Matrix

**Theorem 8.4.1.** *The standard matrix for a linear transformation is given by the formula:*

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}$$

where

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

are the *standard basis vectors* for  $\mathbb{R}^n$ .

## 9 Determinants

### 9.1 Properties of Determinants

1.  $\det(\mathbf{I}) = 1$ .
2. Exchanging two rows of a matrix reverses the sign of its determinant.
3. Determinants are multilinear, so that
  - (a)  $\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$
  - (b)  $\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$
4. If  $\mathbf{A}$  has two equal rows, then  $\det(\mathbf{A}) = 0$ .
5. Adding a scalar multiple of one row to another does not change the determinant of a matrix.
6. If  $\mathbf{A}$  has a row of zeros, then  $\det(\mathbf{A}) = 0$ .
7. If  $\mathbf{A}$  is triangular, then  $\det(\mathbf{A}) = \prod_{i=1}^n a_{ii}$ .
8. If  $\mathbf{A}$  is singular, then  $\det(\mathbf{A}) = 0$ .
9.  $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ .
10.  $\det(\mathbf{A}^\top) = \det(\mathbf{A})$ .

### 9.2 Matrix Minors

**Definition 9.1.** The **minor** of  $a_{ij}$  in  $\mathbf{A}$ , denoted  $\mathbf{M}_{ij}$ , is the determinant of the submatrix formed by deleting the  $i$ th row and  $j$ th column of  $\mathbf{A}$ .

### 9.3 Matrix Cofactors

**Definition 9.2.** The **cofactor** of  $a_{ij}$  in  $\mathbf{A}$  is defined as

$$\mathbf{C}_{ij} = (-1)^{i+j} \mathbf{M}_{ij}$$

### 9.4 The Determinant of a Matrix

**Theorem 9.4.1.** The determinant of an  $n \times n$  matrix  $\mathbf{A}$  is given by

$$\det(\mathbf{A}) = \sum_{j=1}^n a_{ij} \mathbf{C}_{ij} = \sum_{i=1}^n a_{ij} \mathbf{C}_{ij}$$

where  $a_{ij}$  is the entry in the  $i$ th row and  $j$ th column of  $\mathbf{A}$ .

### 9.5 The Cofactor Matrix

**Definition 9.3.** The **cofactor matrix** of an  $n \times n$  matrix  $\mathbf{A}$ , denoted  $\mathbf{C}$ , is defined as the matrix of the cofactors of  $\mathbf{A}$ .

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

### 9.6 The Adjugate of a Matrix

**Definition 9.4.** The **adjugate** (or *classical adjoint*) of a square matrix  $\mathbf{A}$ , denoted  $\text{adj}(\mathbf{A})$ , is the transpose of its cofactor matrix.

$$\text{adj}(\mathbf{A}) = \mathbf{C}^\top$$

### 9.7 The Inverse of a Matrix

**Theorem 9.7.1.** The *inverse* of a nonsingular matrix  $\mathbf{A}$  is given by

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A})$$

## 10 Invariant Subspaces

**Definition 10.1.** Consider the subspace  $\mathcal{V}$  of the linear mapping  $T : V \rightarrow V$  from a vector space  $V$  to itself, then  $\mathcal{V}$  is an **invariant subspace** of  $T$  if

$$T(\mathcal{V}) \subseteq \mathcal{V}$$

**Theorem 10.0.1.** If  $\mathcal{V}$  is an invariant subspace of a linear mapping  $T : V \rightarrow V$  from a vector space  $V$  to itself, then

$$\forall \mathbf{v} \in \mathcal{V} \implies T(\mathbf{v}) \in \mathcal{V}$$

### 10.1 Trivial Invariant Subspaces

1.  $V$ .
2.  $\{0\}$ .
3.  $\ker(T)$ .
4.  $\operatorname{im}(T)$ .
5. Any linear combination of invariant subspaces.

### 10.2 Eigenspaces

**Definition 10.2.** If an invariant subspace is one-dimensional, then the subspace is called an **eigenspace** of the linear transformation.

**Theorem 10.2.1.** If  $\mathcal{V}$  is an eigenspace of the linear mapping  $T : V \rightarrow V$ , then

$$\mathcal{V} = \{\forall \mathbf{v} \in \mathcal{V} : \exists \lambda \in \mathbb{C} : T(\mathbf{v}) = \lambda \mathbf{v}\}$$

where  $\lambda$  is the **eigenvalue** associated with the **eigenvector**  $\mathbf{v}$ .

### 10.3 The Eigenvalue Problem

**Theorem 10.3.1.** The eigenvalues  $\lambda$  of an invertible square matrix  $\mathbf{A}$ , are the solutions to

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

**Theorem 10.3.2.** The eigenvectors associated with each eigenvalue, of an invertible square matrix  $\mathbf{A}$ , are the solutions to

$$(\mathbf{A} - \lambda \mathbf{I}) \mathbf{v} = 0$$

*Proof.* The eigenvalues and associated eigenvectors of a square matrix  $\mathbf{A}$ , are the solutions to  $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$ .

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$$

$$\mathbf{A}\mathbf{v} - \lambda \mathbf{v} = 0$$

$$(\mathbf{A} - \lambda \mathbf{I}) \mathbf{v} = 0$$

The linear system  $(\mathbf{A} - \lambda \mathbf{I}) \mathbf{v} = 0$  has a nontrivial solution iff  $\mathbf{A} - \lambda \mathbf{I}$  is singular. □

### 10.4 Properties of Eigenvalues

**Theorem 10.4.1.**

$$\operatorname{Tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$$

**Theorem 10.4.2.**

$$\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i$$

## 11 Eigen Decomposition

### 11.1 Similarity Transformations

**Definition 11.1.** A **similarity transformation** is a linear mapping of the form

$$\mathbf{A} \rightarrow \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$$

in which the matrices  $\mathbf{A}$  and  $\mathbf{V}$  are  $n \times n$  invertible matrices. Here we say, “ $\mathbf{A}$  is similar to  $\mathbf{V}^{-1}\mathbf{A}\mathbf{V}$ ”.

### 11.2 Matrix Diagonalisation

**Definition 11.2.** The matrix  $\mathbf{A}$  is a **diagonalisable** matrix if it is similar to a diagonal matrix. That is, there exists an invertible matrix  $\mathbf{V}$ , and diagonal matrix  $\mathbf{D}$ , such that

$$\mathbf{D} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$$

**Theorem 11.2.1.** Let  $\mathbf{A}$  be an  $n \times n$  matrix with  $n$  linearly independent eigenvectors, then  $\mathbf{A}$  is diagonalisable if  $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  and  $\mathbf{V}$  is a matrix composed of the eigenvectors of  $\mathbf{A}$ . Explicitly,

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}$$

where  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  are the eigenvectors of  $\mathbf{A}$ .

*Proof.* Let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  be linearly independent eigenvectors of  $\mathbf{A}$ , and  $\lambda_1, \lambda_2, \dots, \lambda_n$ , the associated eigenvalues. By definition of an eigenspace, we have

$$\begin{cases} \mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1 \\ \mathbf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2 \\ \vdots \\ \mathbf{A}\mathbf{v}_n = \lambda_n\mathbf{v}_n \end{cases}$$

which we can rewrite as

$$\begin{aligned} \mathbf{A}\mathbf{V} &= \mathbf{V}\mathbf{D} \\ \mathbf{D} &= \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \end{aligned} \tag{1}$$

by rearranging Equation 1, we have  $\mathbf{A}$  in terms of its eigenvalues and eigenvectors.

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$$

□

### 11.3 Powers of a Matrix

**Theorem 11.3.1.** *Let  $\mathbf{A}$  be a diagonalisable matrix, then for all  $k \in \mathbb{N}_0$*

$$\mathbf{A}^k = \mathbf{V}\mathbf{D}^k\mathbf{V}^{-1}$$

*Proof.*

$$\begin{aligned}\mathbf{A}^k &= (\mathbf{V}\mathbf{D}\mathbf{V}^{-1})^k \\ &= \underbrace{(\mathbf{V}\mathbf{D}\mathbf{V}^{-1})(\mathbf{V}\mathbf{D}\mathbf{V}^{-1}) \dots (\mathbf{V}\mathbf{D}\mathbf{V}^{-1})}_{k \text{ times}} \\ &= \underbrace{\mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{V}\mathbf{D}\mathbf{V}^{-1} \dots \mathbf{V}\mathbf{D}\mathbf{V}^{-1}}_{k \text{ times}} \\ &= \underbrace{\mathbf{V}\cancel{\mathbf{D}\mathbf{V}^{-1}}\mathbf{V}\cancel{\mathbf{D}\mathbf{V}^{-1}} \dots \mathbf{V}\cancel{\mathbf{D}\mathbf{V}^{-1}}}_{k \text{ times}} \\ &= \mathbf{V}\underbrace{\mathbf{D}\mathbf{D} \dots \mathbf{D}}_{k \text{ times}}\mathbf{V}^{-1} \\ &= \mathbf{V}\mathbf{D}^k\mathbf{V}^{-1}\end{aligned}$$

□

**Theorem 11.3.2.** *The eigenvalues of  $\mathbf{A}^k$ ,  $\forall k \in \mathbb{N}$  are  $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ .*

**Theorem 11.3.3.** *The eigenvectors of  $\mathbf{A}$  are equal to the eigenvectors of  $\mathbf{A}^k$ .*

## 12 System of Differential Equations

### 12.1 First-Order Differential Equations

**Definition 12.1.** A **first-order differential equation** is a differential equation where the highest derivative is of order one.

$$x' = ax$$

**Theorem 12.1.1.** *The general solution to a first-order linear differential equation is of the form*

$$x(t) = c_1 e^{at}$$

where  $c_1$  is an arbitrary constant.

### 12.2 First-Order System of Differential Equations

**Definition 12.2.** A **first-order system of differential equations** is of the form

$$\begin{cases} x'_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots & \vdots \\ x'_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{cases}$$



where  $x_1 = x_1(t)$ ,  $x_2 = x_2(t)$ , ...,  $x_n = x_n(t)$  are the functions to be determined. In matrix form, the system can be written as

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\mathbf{x}' = \mathbf{A}\mathbf{x}$$

### 12.3 Solution using Diagonalisation

**Theorem 12.3.1.** *The first-order system of differential equations  $\mathbf{x}' = \mathbf{A}\mathbf{x}$  can be solved using the following substitution*

$$\mathbf{x} = \mathbf{V}\mathbf{u}$$

where  $\mathbf{u}$  is a vector to be determined, and  $\mathbf{V}$  is the matrix that diagonalises  $\mathbf{A}$ .  $\mathbf{u}$  is determined by solving

$$\mathbf{u}' = \mathbf{D}\mathbf{u}$$

where  $\mathbf{D}$  is the diagonal similarity transformation of  $\mathbf{A}$ . This substitution uncouples the system of differential equations so that each equation can be solved as a first-order differential equation.

*Proof.*

$$\begin{aligned} \mathbf{x}' &= \mathbf{A}\mathbf{x} \\ (\mathbf{V}\mathbf{u})' &= (\mathbf{V}\mathbf{D}\mathbf{V}^{-1})(\mathbf{V}\mathbf{u}) \\ \mathbf{V}\mathbf{u}' &= \mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{V}\mathbf{u} \\ \cancel{\mathbf{V}}\mathbf{u}' &= \cancel{\mathbf{V}}\mathbf{D}\cancel{\mathbf{V}^{-1}\mathbf{V}}\mathbf{u} \\ \mathbf{u}' &= \mathbf{D}\mathbf{u} \end{aligned}$$

□

**Theorem 12.3.2.** *If  $\mathbf{A}$  is a diagonalisable matrix, then the general solution of  $\mathbf{x}' = \mathbf{A}\mathbf{x}$  can be expressed as*

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t} \mathbf{v}_1 + c_2 e^{\lambda_2 t} \mathbf{v}_2 + \cdots + c_n e^{\lambda_n t} \mathbf{v}_n$$

### 12.4 Principle of Superposition

**Theorem 12.4.1.** *If  $x_1$  and  $x_2$  are two solutions to a linear differential equation, then*

$$x = c_1 x_1 + c_2 x_2$$

*is also a solution to the differential equation.*

## 12.5 Higher-Order Differential Equations

**Theorem 12.5.1.** A *higher-order linear differential equation* can be solved by first converting it to a first-order linear system. Consider the  $n$ th-order differential equation

$$x^{(n)} + a_1 x^{(n-1)} + \cdots + a_{n-1} x' + a_n x = 0$$

We then define

$$\begin{aligned} x_1 &= x \\ x_2 &= x' \\ &\vdots \\ x_n &= x^{(n-1)} \end{aligned}$$

Let  $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^\top$ . Then the first-order linear system of differential equations can be written as

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$