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

Contents **1**

1 Euclidean Vector Spaces **4**
1.1 Vectors . . . 4
1.2 Position and Displacement Vectors . . . 4
1.3 Vector Addition . . . 5
1.4 Scalar Multiplication . . . 5
1.5 Norm of a Vector . . . 5
1.6 The Unit Vector . . . 5
1.7 The Dot Product . . . 5
1.8 The Cross Product . . . 5

2 Vector Identities **7**

3 Linear System of Equations **8**
3.1 Linear Equations . . . 8
3.2 Homogeneous Linear Equations . . . 8
3.3 Linear Systems . . . 8
3.4 Coefficient Matrices . . . 8
3.5 Augmented Matrices . . . 8
3.6 Elementary Row Operations . . . 9
3.7 Pivots . . . 9
3.8 Gaussian Elimination . . . 9
3.9 Row-Echelon Form . . . 9
3.10 Gauss-Jordan Elimination . . . 9
3.11 Solutions to Linear Systems . . . 10

4 Matrices **11**
4.1 Matrix Addition . . . 11
4.2 Scalar Multiplication . . . 11
4.3 Matrix Multiplication . . . 11
4.4 The Identity Matrix . . . 11
4.5 The Inverse Matrix . . . 12
4.6 The Diagonal Matrix . . . 12
4.7 Matrix Transpose . . . 12
4.8 Matrix Trace . . . 12

5 General Vector Spaces **13**
5.1 Real Vector Spaces . . . 13
5.2 Subspaces . . . 14
5.3 Spanning Sets . . . 15
5.4 Linear Independence . . . 15
5.5 Basis Vectors . . . 15
5.6 Dimension . . . 15

6 Fundamental Subspaces **16**
6.1 The Four Fundamental Subspaces of a Matrix . . . 16
6.2 The General Solution of a System of Equations . . . 16
6.3 Row Equivalence . . . 17
6.4 Rank . . . 17
6.5 Nullity . . . 17

7 Orthogonality **18**
7.1 Orthogonal Subspaces . . . 18
7.2 Orthogonal Complements . . . 18
7.3 Vector Projections . . . 19
7.4 Projection onto a Subspace . . . 19
7.5 Least Squares . . . 20

8 Linear Maps **21**
8.1 Matrix Transformations . . . 21
8.2 General Linear Transformations . . . 21
8.3 Subspaces of Linear Transformations . . . 22
8.4 Constructing a Transformation Matrix . . . 23

9 Determinants **24**
9.1 Properties of Determinants . . . 24
9.2 Matrix Minors . . . 24
9.3 Matrix Cofactors . . . 24
9.4 The Determinant of a Matrix . . . 24
9.5 The Cofactor Matrix . . . 25
9.6 The Adjugate of a Matrix . . . 25
9.7 The Inverse of a Matrix . . . 25

10 Invariant Subspaces **26**
10.1 Trivial Invariant Subspaces . . . 26
10.2 Eigenspaces . . . 26
10.3 The Eigenvalue Problem . . . 26
10.4 Properties of Eigenvalues . . . 27

11 Eigen Decomposition **28**
11.1 Similarity Transformations . . . 28
11.2 Matrix Diagonalisation . . . 28
11.3 Powers of a Matrix . . . 29

12 System of Differential Equations **30**
12.1 First-Order Differential Equations . . . 30
12.2 First-Order System of Differential Equations . . . 30
12.3 Solution using Diagonalisation . . . 30
12.4 Principle of Superposition . . . 31
12.5 Higher-Order Differential Equations . . . 31

List of Figures **32**

References **33**

1 Euclidean Vector Spaces

1.1 Vectors

Definition 1.1.1. An n-dimensional **vector** is an ordered list of n numbers.

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

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

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

Notation:

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

1.2 Position and Displacement Vectors

Definition 1.2.1. The **displacement vector** $\overrightarrow{AB}$ from $\boldsymbol{a}$ to $\boldsymbol{b}$ can be defined as $\boldsymbol{b} - \boldsymbol{a}$.

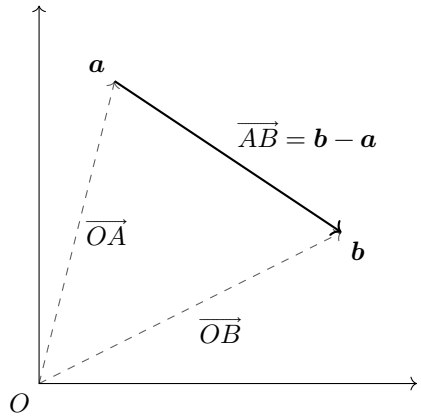

Figure 1: Displacement vector between two points.

1.3 Vector Addition

Definition 1.3.1. Vector addition is performed by adding the corresponding components of two vectors of the same dimension.

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

1.4 Scalar Multiplication

Definition 1.4.1. Scalar multiplication is performed by multiplying each element of the vector by the scalar.

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

1.5 Norm of a Vector

Definition 1.5.1. The **norm** of a vector $\boldsymbol{v}$, denoted by $\|\boldsymbol{v}\|$, is the *length* or *magnitude* of $\boldsymbol{v}$.

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

1.6 The Unit Vector

Definition 1.6.1. A **unit vector** is a vector, denoted $\hat{\boldsymbol{v}}$, that has a length of 1 in the direction of $\boldsymbol{v}$.

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

1.7 The Dot Product

Definition 1.7.1. The **dot product** is a function that associates each pair of vectors $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{R}^n$ a real number $\boldsymbol{v} \cdot \boldsymbol{w}$.

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

where θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$.

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

1.8 The Cross Product

Definition 1.8.1. The **cross product** is a function that associates each ordered pair of vectors $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{R}^3$ a vector $\boldsymbol{v} \times \boldsymbol{w} \in \mathbb{R}^3$.

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

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

2 Vector Identities

Theorem 2.0.1. *Commutativity of vector addition.*

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

Theorem 2.0.2.

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

Theorem 2.0.3. *Commutativity of dot products.*

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

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

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

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

$$(r\boldsymbol{a}) \cdot \boldsymbol{b} = r\,(\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot (r\boldsymbol{b})$$

Theorem 2.0.6. *Bilinearity of dot products.*

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

Theorem 2.0.7.

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

Theorem 2.0.8. *Anticommutativity of cross products.*

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

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

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

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

$$(r\boldsymbol{a}) \times \boldsymbol{b} = \boldsymbol{a} \times (r\boldsymbol{b}) = r\,(\boldsymbol{a} \times \boldsymbol{b})$$

Theorem 2.0.11.

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

Theorem 2.0.12.

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

3 Linear System of Equations

3.1 Linear Equations

Definition 3.1.1. A **linear equation** in n variables $x_1, x_2, \ldots, x_n$ can be expressed in the form

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

where the *coefficients* $a_1, a_2, \ldots, a_n$ and the *constant term* b are constants.

3.2 Homogeneous Linear Equations

Definition 3.2.1. In the special case where $b = 0$, the linear equation is called a **homogeneous linear equation**.

3.3 Linear Systems

Definition 3.3.1. A **linear system of equations** is a set of linear equations, where the variables x_i are called *unknowns*. The general linear system of m equations with n unknowns can be written as

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

A **solution** to the system is an n-tuple $\langle x_1, x_2, \ldots, x_n \rangle$ that satisfies each equation.

3.4 Coefficient Matrices

Definition 3.4.1. The coefficients of the variables in each equation can be placed inside the systems **coefficient matrix**.

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

3.5 Augmented Matrices

Definition 3.5.1. The information of a system can be contained in its **augmented matrix**.

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

Definition 3.5.2. An array having m rows and n columns, is an $m \times n$ **matrix**. This matrix may be denoted as a_{ij}, where a_{ij} is the entry in ith row and jth column of the matrix $\boldsymbol{A}$.

$$m \text{ rows} \left\{ \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{n \text{ columns}} \right.$$

3.6 Elementary Row Operations

Definition 3.6.1. 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 3.7.1. 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 3.8.1. 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 3.9.1. 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 3.10.1. 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**. matrix is in **reduced row-echelon form**.

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

3.11 Solutions to Linear Systems

Definition 3.11.1. A **consistent system** of equations has at least one solution, and an **inconsistent system** has no solution.

4 Matrices

Definition 4.0.1. A **matrix** is an array of numbers arranged into *rows* and *columns*, and can be used to represent a linear transformation.

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

4.1 Matrix Addition

Definition 4.1.1. Matrix addition is performed by adding the corresponding components of two matrices of the same dimension.

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

4.2 Scalar Multiplication

Definition 4.2.1. Scalar multiplication is performed by multiplying each element of a matrix by a scalar.

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

4.3 Matrix Multiplication

Definition 4.3.1. Matrix multiplication is performed by multiplying each row in the first matrix by the columns of the second matrix.

$$\boldsymbol{AB} = \boldsymbol{C}$$

$$\begin{bmatrix} — & \boldsymbol{a}_1 & — \\ — & \boldsymbol{a}_2 & — \\ & \vdots & \\ — & \boldsymbol{a}_m & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1\boldsymbol{b}_1 & \boldsymbol{a}_1\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1\boldsymbol{b}_n \\ \boldsymbol{a}_2\boldsymbol{b}_1 & \boldsymbol{a}_2\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2\boldsymbol{b}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_m\boldsymbol{b}_1 & \boldsymbol{a}_m\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_m\boldsymbol{b}_n \end{bmatrix}$$

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 4.4.1. The **identity matrix** is the simplest nontrivial **diagonal matrix**, denoted $\mathbb{1}$, such that

$$\mathbb{1}\boldsymbol{A} = \boldsymbol{A}$$

written explicitly as

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

4.5 The Inverse Matrix

Definition 4.5.1. The **inverse** of a **square matrix** is a matrix $\boldsymbol{A}^{-1}$, such that

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \mathbb{1}$$

Theorem 4.5.1. *The inverse of a* 2×2 *matrix is given by*

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

Theorem 4.5.2. *The inverse of an* $n \times n$ *matrix can be determined by solving* $\left[\begin{array}{c|c} \boldsymbol{A} & \mathbb{1} \end{array}\right]$.

4.6 The Diagonal Matrix

Definition 4.6.1. A **diagonal** matrix, denoted $\operatorname{diag}(d_{11},\ d_{22},\ \ldots,\ d_{nn})$, is an $n \times n$ matrix $\boldsymbol{D}$ in which entries outside the main diagonal are all zero.

$$\boldsymbol{D} = \operatorname{diag}(d_{11}, d_{22}, \ldots, d_{nn}) = \begin{bmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{nn} \end{bmatrix}$$

4.7 Matrix Transpose

Definition 4.7.1. The **transpose** of a matrix, denoted by $\boldsymbol{A}^{\top}$, is obtained by replacing all a_{ij} elements with a_{ji}, so that the matrix $\boldsymbol{A}$ is flipped over its main diagonal.

4.8 Matrix Trace

Definition 4.8.1. The **trace** of an $n \times n$ matrix $\boldsymbol{A}$, denoted $\operatorname{Tr}(\boldsymbol{A})$, is defined as

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

5 General Vector Spaces

5.1 Real Vector Spaces

Definition 5.1.1. 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 $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w} \in V$, and all scalars k and m, then V is a **vector space**, and the objects in V are vectors.*

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

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

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

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

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

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

Axiom 4 (Additive identity)**.**

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

Axiom 5 (Additive inverse)**.**

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

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

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

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

$$k\left(\boldsymbol{u} + \boldsymbol{v}\right) = k\boldsymbol{u} + k\boldsymbol{v}$$

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

$$\left(k + m\right)\boldsymbol{u} = k\boldsymbol{u} + m\boldsymbol{u}$$

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

$$k\left(m\boldsymbol{u}\right) = \left(km\right)\boldsymbol{u}$$

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

$$1\boldsymbol{u} = \boldsymbol{u}$$

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 $\boldsymbol{v} \in V$, and k is a scalar.*

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

5.2 Subspaces

Definition 5.2.1. 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 $\boldsymbol{M}_{nn}$.*

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

5.3 Spanning Sets

Definition 5.3.1. If the vector $\boldsymbol{w}$ is in a vector space V, then $\boldsymbol{w}$ is a **linear combination** of the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$, if $\boldsymbol{w}$ can be expressed in the form

$$\boldsymbol{w} = k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n$$

Theorem 5.3.1. *If $S = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ 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 5.4.1. 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*

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n = \boldsymbol{0}$$

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

5.5 Basis Vectors

Definition 5.5.1. 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 5.6.1. 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

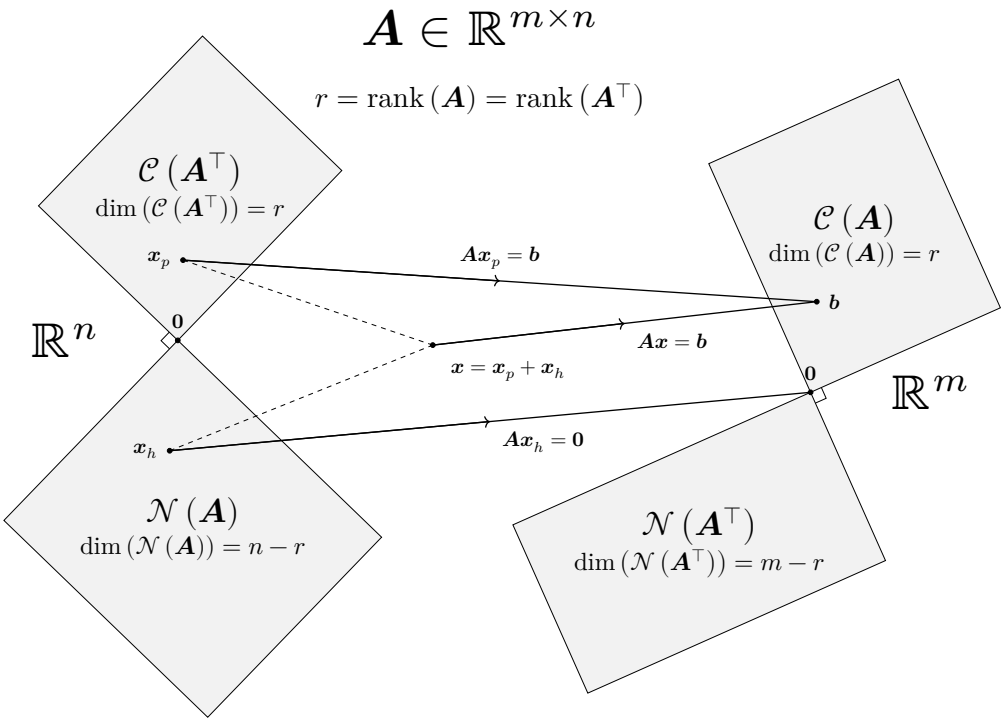

Figure 2: The Four Fundamental Subspaces of a Matrix.

6.1 The Four Fundamental Subspaces of a Matrix

Definition 6.1.1. If $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is an $m \times n$ matrix, then:

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

6.2 The General Solution of a System of Equations

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

$\mathcal{C}\left(\boldsymbol{A}^{\top}\right)$, and the homogeneous solution is the solution to $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$, or $\mathcal{N}\left(\boldsymbol{A}\right)$.

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

6.3 Row Equivalence

Definition 6.3.1. 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 6.4.1. The **rank** of a matrix, denoted by $\operatorname{rank}\left(\boldsymbol{A}\right)$, is given by $\dim\left(\mathcal{C}\left(\boldsymbol{A}\right)\right)$.

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

$$\operatorname{rank}\left(\boldsymbol{A}\right)=\dim\left(\mathcal{C}\left(\boldsymbol{A}\right)\right)=\dim\left(\mathcal{C}\left(\boldsymbol{A}^{\top}\right)\right)$$

6.5 Nullity

Definition 6.5.1. The **nullity** of a matrix, denoted by $\operatorname{null}\left(\boldsymbol{A}\right)$, is given by $\dim\left(\mathcal{N}\left(\boldsymbol{A}\right)\right)$.

7 Orthogonality

Definition 7.0.1. Two vectors are **orthogonal** if the following holds.

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

Theorem 7.0.1. **0** *is orthogonal to every vector in* V*.*

Theorem 7.0.2. **0** *is the only vector in* V*, that is orthogonal to itself.*

Theorem 7.0.3.

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

Theorem 7.0.4.

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

7.1 Orthogonal Subspaces

Definition 7.1.1. 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.

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

7.2 Orthogonal Complements

Definition 7.2.1. 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.

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

Theorem 7.2.1.

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

Theorem 7.2.2.

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

7.3 Vector Projections

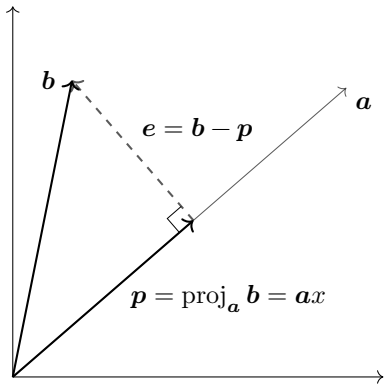

Figure 3: Vector Projection of $\boldsymbol{b}$ onto $\boldsymbol{a}$.

Definition 7.3.1. Let the **vector projection** of $\boldsymbol{b}$ onto $\boldsymbol{a}$, denoted as $\text{proj}_{\boldsymbol{a}}\,\boldsymbol{b}$, be the orthogonal projection of $\boldsymbol{b}$ in the direction of $\boldsymbol{a}$, that minimises the error vector: $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$.

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

$$\text{proj}_{\boldsymbol{a}}\,\boldsymbol{b} = \boldsymbol{a}x = \boldsymbol{a}\left(\boldsymbol{a}^\top\boldsymbol{a}\right)^{-1}\boldsymbol{a}^\top\boldsymbol{b}$$

alternatively

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

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

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

□

7.4 Projection onto a Subspace

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

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

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

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

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

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

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

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

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

which gives the following equation

$$\boldsymbol{A}^\top \left(\boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}}\right) = \boldsymbol{0}$$

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

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

□

7.5 Least Squares

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

8 Linear Maps

8.1 Matrix Transformations

Definition 8.1.1. A **matrix transformation** $T_{\boldsymbol{A}} : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping of the form

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

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

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

8.2 General Linear Transformations

Theorem 8.2.1. *If $T : V \to W$ 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. $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$
2. $T(k\boldsymbol{u})\, kT(\boldsymbol{u})$

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

8.3 Subspaces of Linear Transformations

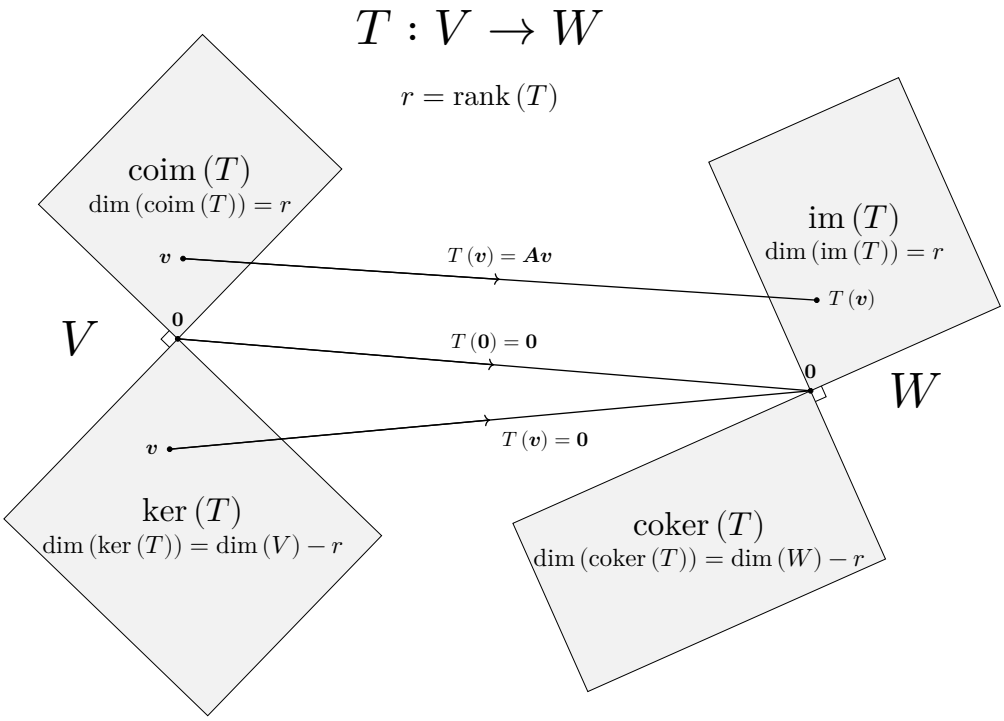

Figure 4: Subspaces of a Linear Transformation.

Definition 8.3.1. If $T : V \to W$ 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.

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

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

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

8.4 Constructing a Transformation Matrix

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

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

where

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

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

9 Determinants

9.1 Properties of Determinants

1. $\det(\mathbb{1}) = 1$.
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 $\boldsymbol{A}$ has two equal rows, then $\det(\boldsymbol{A}) = 0$.
5. Adding a scalar multiple of one row to another does not change the determinant of a matrix.
6. If $\boldsymbol{A}$ has a row of zeros, then $\det(\boldsymbol{A}) = 0$.
7. If $\boldsymbol{A}$ is triangular, then $\det(\boldsymbol{A}) = \prod_{i=1}^{n} a_{ii}$.
8. If $\boldsymbol{A}$ is singular, then $\det(\boldsymbol{A}) = 0$.
9. $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.
10. $\det(\boldsymbol{A}^\top) = \det(\boldsymbol{A})$.

9.2 Matrix Minors

Definition 9.2.1. The **minor** of a_{ij} in $\boldsymbol{A}$, denoted M_{ij}, is the determinant of the submatrix formed by deleting the ith row and jth column of $\boldsymbol{A}$.

9.3 Matrix Cofactors

Definition 9.3.1. The **cofactor** of a_{ij} in $\boldsymbol{A}$ is defined as

$$C_{ij} = (-1)^{i+j} M_{ij}$$

9.4 The Determinant of a Matrix

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

$$\det(\boldsymbol{A}) = \sum_{j=1}^{n} a_{ij} C_{ij} = \sum_{i=1}^{n} a_{ij} C_{ij}$$

where a_{ij} is the entry in the ith row and jth column of $\boldsymbol{A}$.

9.5 The Cofactor Matrix

Definition 9.5.1. The **cofactor matrix** of an $n \times n$ matrix $\boldsymbol{A}$, denoted $\boldsymbol{C}$, is defined as the matrix of the cofactors of $\boldsymbol{A}$.

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

9.6 The Adjugate of a Matrix

Definition 9.6.1. The **adjugate** (or *classical adjoint*) of a square matrix $\boldsymbol{A}$, denoted $\operatorname{adj}(\boldsymbol{A})$, is the transpose its cofactor matrix.

$$\operatorname{adj}(\boldsymbol{A}) = \boldsymbol{C}^{\top}$$

9.7 The Inverse of a Matrix

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

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \operatorname{adj}(\boldsymbol{A})$$

10 Invariant Subspaces

Definition 10.0.1. Consider the subspace $\mathcal{V}$ of the linear mapping $T : V \to V$ 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 $T : V \to V$ from a vector space V to itself, then*

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

10.1 Trivial Invariant Subspaces

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

10.2 Eigenspaces

Definition 10.2.1. 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 $T : V \to V$, then*

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

*where λ is the **eigenvalue** associated with the **eigenvector** $\boldsymbol{q}$.*

10.3 The Eigenvalue Problem

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

$$\det(\boldsymbol{A} - \lambda \mathbb{1}) = \mathbf{0}$$

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

$$(\boldsymbol{A} - \lambda \mathbb{1})\,\boldsymbol{q} = \mathbf{0}$$

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

$$\begin{aligned} \boldsymbol{A}\boldsymbol{q} &= \lambda \boldsymbol{q} \\ \boldsymbol{A}\boldsymbol{q} - \lambda \boldsymbol{q} &= \mathbf{0} \\ (\boldsymbol{A} - \lambda \mathbb{1})\,\boldsymbol{q} &= \mathbf{0} \end{aligned}$$

The linear system $(\boldsymbol{A} - \lambda \mathbb{1})\,\boldsymbol{q} = \mathbf{0}$ has a nontrivial solution iff $\boldsymbol{A} - \lambda \mathbb{1}$ is singular. □

10.4 Properties of Eigenvalues

Theorem 10.4.1.

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

Theorem 10.4.2.

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

11 Eigen Decomposition

11.1 Similarity Transformations

Definition 11.1.1. A **similarity transformation** is a linear mapping of the form

$$\boldsymbol{A} \to \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$$

in which the matrices $\boldsymbol{A}$ and $\boldsymbol{Q}$ are $n \times n$ invertible matrices. Here we say, "$\boldsymbol{A}$ is similar to $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$".

11.2 Matrix Diagonalisation

Definition 11.2.1. The matrix $\boldsymbol{A}$ is a **diagonalisable** matrix if it is similar to a diagonal matrix. That is, there exists an invertible matrix $\boldsymbol{Q}$, and diagonal matrix $\boldsymbol{\Lambda}$, such that

$$\boldsymbol{\Lambda} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$$

Theorem 11.2.1. *Let $\boldsymbol{A}$ be an $n \times n$ matrix with n linearly independent eigenvectors, then $\boldsymbol{A}$ is diagonalisable if $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $\boldsymbol{Q}$ is a matrix composed of the eigenvectors of $\boldsymbol{A}$. Explicitly,*

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad \text{and} \quad \boldsymbol{Q} = \begin{bmatrix} | & | & & | \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \\ | & | & & | \end{bmatrix}$$

where $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$ are the eigenvectors of $\boldsymbol{A}$.

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

$$\begin{cases} \boldsymbol{A}\boldsymbol{q}_1 = \lambda_1 \boldsymbol{q}_1 \\ \boldsymbol{A}\boldsymbol{q}_2 = \lambda_2 \boldsymbol{q}_2 \\ \quad \vdots \quad \vdots \ \vdots \\ \boldsymbol{A}\boldsymbol{q}_n = \lambda_n \boldsymbol{q}_n \end{cases}$$

which we can rewrite as

$$\begin{aligned} \boldsymbol{A}\boldsymbol{Q} &= \boldsymbol{Q}\boldsymbol{\Lambda} \\ \boldsymbol{\Lambda} &= \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} \end{aligned} \tag{1}$$

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

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}$$

□

11.3 Powers of a Matrix

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

$$\boldsymbol{A}^k = \boldsymbol{Q}\boldsymbol{\Lambda}^k\boldsymbol{Q}^{-1}$$

Proof.

$$\begin{aligned}
\boldsymbol{A}^k &= \left(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\right)^k \\
&= \underbrace{\left(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\right)\left(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\right)\cdots\left(\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\right)}_{k \text{ times}} \\
&= \underbrace{\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\cdots\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}}_{k \text{ times}} \\
&= \underbrace{\boldsymbol{Q}\boldsymbol{\Lambda}\cancel{\boldsymbol{Q}^{-1}\boldsymbol{Q}}\boldsymbol{\Lambda}\cancel{\boldsymbol{Q}^{-1}\cdots\boldsymbol{Q}}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}}_{k \text{ times}} \\
&= \boldsymbol{Q}\underbrace{\boldsymbol{\Lambda}\boldsymbol{\Lambda}\cdots\boldsymbol{\Lambda}}_{k \text{ times}}\boldsymbol{Q}^{-1} \\
&= \boldsymbol{Q}\boldsymbol{\Lambda}^k\boldsymbol{Q}^{-1}
\end{aligned}$$

□

Theorem 11.3.2. *The eigenvalues of $\boldsymbol{A}^k$, $\forall k \in \mathbb{N}$ are λ_1^k, λ_2^k, ..., λ_n^k.*

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

12 System of Differential Equations

12.1 First-Order Differential Equations

Definition 12.1.1. 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 12.2.1. A **first-order system of differential equations** is of the form

$$\begin{cases} x_1' = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ x_2' = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \quad\quad \vdots \quad\quad \vdots \quad\quad\quad \vdots \\ x_n' = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases}$$

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

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

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

12.3 Solution using Diagonalisation

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

$$\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{u}$$

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

$$\boldsymbol{u}' = \boldsymbol{\Lambda}\boldsymbol{u}$$

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

Proof.

$$\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$$
$$(\boldsymbol{Q}\boldsymbol{u})' = (\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1})(\boldsymbol{Q}\boldsymbol{u})$$

$$\begin{aligned}\boldsymbol{Q}\boldsymbol{u}' &= \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{u}\\ \cancel{\boldsymbol{Q}}\boldsymbol{u}' &= \cancel{\boldsymbol{Q}}\boldsymbol{\Lambda}\cancel{\boldsymbol{Q}^{-1}}\cancel{\boldsymbol{Q}}\boldsymbol{u}\\ \boldsymbol{u}' &= \boldsymbol{\Lambda}\boldsymbol{u}\end{aligned}$$

□

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

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

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 nth-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 $\boldsymbol{x} = \begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}^\top$. Then the first-order linear system of differential equations can be written as

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

List of Figures

1 Displacement vector between two points. 4
2 The Four Fundamental Subspaces of a Matrix. 16
3 Vector Projection of $\boldsymbol{b}$ onto $\boldsymbol{a}$. 19
4 Subspaces of a Linear Transformation. 22

References

[1] Howard Anton, Chris Rorres, and Anton Kaul. *Elementary Linear Algebra: Applications Version, 12th Australia and New Zealand Edition*. John Wiley & Sons Australia, 12th edition, 2019.

[2] Paul Dawkins. Solutions To Systems. URL: https://tutorial.math.lamar.edu/Classes/DE/SolutionsToSystems.aspx.

[3] Arash Komeili. The Superposition Principle and Undetermined Coefficients Revisited, 2020. URL: https://math.berkeley.edu/~arash/54/notes/n4_5.pdf.

[4] Glenn Lahodny Jr. Converting Higher-Order Equations to First-Order Systems, 2018. URL: https://www.math.tamu.edu/~glahodny/Math469/Section%204.6.pdf.

[5] Isaiah Lankham, Bruno Nachtergaele, and Anne Schilling. Eigenvalues and Eigenvectors, 2007. URL: https://www.math.ucdavis.edu/~anne/WQ2007/mat67-Li-Eigenvectors.pdf.

[6] John McCuan. Invariant Subspaces and Eigenspaces, 2008. URL: https://people.math.gatech.edu/~mccuan/courses/1502/linear3.pdf.

[7] Gilbert Strang. 18.06SC Linear Algebra, 2011. MIT OpenCourseWare. URL: https://ocw.mit.edu/courses/mathematics/18-06sc-linear-algebra-fall-2011.

[8] Gilbert Strang. Properties of determinants, 2011. MIT OpenCourseWare. URL: https://ocw.mit.edu/courses/mathematics/18-06sc-linear-algebra-fall-2011/least-squares-determinants-and-eigenvalues/properties-of-determinants/MIT18_06SCF11_Ses2.5sum.pdf.

[9] Gilbert Strang. The Four Fundamental Subspaces: 4 Lines, 2011. MIT OpenCourseWare. URL: http://web.mit.edu/18.06/www/Essays/newpaper_ver3.pdf.

[10] Terrence Tao. Axioms of real vector spaces, 2000. URL: https://www.math.ucla.edu/~tao/resource/general/121.1.00s/vector_axioms.html.

[11] Eric W. Weisstein. Eigen Decomposition. From MathWorld–A Wolfram Web Resource. URL: https://mathworld.wolfram.com/EigenDecomposition.html.

[12] Eric W. Weisstein. Matrix Diagonalization. From MathWorld–A Wolfram Web Resource. URL: https://mathworld.wolfram.com/MatrixDiagonalization.html.

[13] Eric W. Weisstein. Matrix Power. From MathWorld–A Wolfram Web Resource. URL: https://mathworld.wolfram.com/MatrixPower.html.

[14] Eric W. Weisstein. Similarity Transformation. From MathWorld–A Wolfram Web Resource. URL: https://mathworld.wolfram.com/SimilarityTransformation.html.