Linear Algebra

A Visual and Application-Driven Approach

for High School Students

Sanjay Dagam

To all students who dare to ask "why?"
and who find beauty in patterns and structure.

Preface

Welcome to the fascinating world of linear algebra! This textbook is designed specifically for high school students who have completed AP Calculus and are ready to explore one of the most powerful and beautiful branches of mathematics.

Why Linear Algebra?

Linear algebra is everywhere. When you search on Google, the PageRank algorithm uses eigenvalues. When you watch a Pixar movie, linear transformations animate every character. When Netflix recommends a show, matrix factorization analyzes your viewing patterns. From quantum mechanics to machine learning, from computer graphics to economics, linear algebra is the mathematical language of the modern world.

But beyond its applications, linear algebra offers something profound: it teaches us to think abstractly about structure and patterns. You'll learn to see connections between seemingly different problems—to recognize that solving systems of equations, understanding geometric transformations, and finding patterns in data are all manifestations of the same underlying mathematical principles.

How to Use This Book

This textbook takes a visual, application-driven approach. Each chapter includes:

- **Blue boxes** for definitions and key theorems
- **Green boxes** for insights, connections, and geometric interpretations
- **Red boxes** for common mistakes and important warnings
- **Worked examples** that show techniques step-by-step
- **Practice problems** at three levels: Basic, Intermediate, and Challenge
- **Applications** connecting abstract concepts to real-world uses

Prerequisites

This book assumes you have completed:

- Algebra I and II
- Geometry
- Precalculus
- AP Calculus AB or BC

We'll review key concepts as needed, but familiarity with functions, trigonometry, and basic calculus will help you focus on the new ideas.

A Note on Proofs

Mathematics is not just about computation—it's about understanding *why* things are true. Throughout this book, you'll encounter proofs that explain the reasoning behind results. Don't skip them! Reading and understanding proofs develops your mathematical maturity and deepens your understanding. Appendix B provides guidance on proof techniques.

Computational Tools

While we emphasize understanding concepts, computation is important too. Appendix D introduces tools like Python/NumPy, MATLAB, and graphing calculators that can handle large-scale computations and help you visualize abstract ideas.

Getting Help

- Work through examples carefully before attempting problems
- Form study groups—explaining concepts to others deepens your understanding
- Use the appendices for notation, proof techniques, and prerequisite review
- Consult the companion *Solutions Manual* to check your work
- Don't be afraid to struggle with difficult problems—persistence builds mathematical thinking

Acknowledgments

I would like to thank my math, physics, chemistry teachers and parents for their support and encouragement in writing this textbook.

I hope this book ignites your passion for mathematics and opens doors to new ways of thinking about the world. Linear algebra is a journey of discovery—enjoy every moment of it!

Sanjay Dagam

Contents

Preface **iv**

1 Introduction to Vectors **1**

1.1 What Are Vectors? 1
1.1.1 Introducing Vectors 1
1.1.2 Notation and Representation 2
1.1.3 Real-World Examples 2
1.2 Vector Notation and Representation 3
1.2.1 Component Form 3
1.2.2 Column Vector Notation 3
1.2.3 Vectors in Three Dimensions 3
1.2.4 Magnitude of a Vector 4
1.2.5 Direction of a Vector 4
1.2.6 Unit Vectors 5
1.3 Vector Operations 6
1.3.1 Vector Addition 6
1.3.2 Vector Subtraction 7
1.3.3 Scalar Multiplication 7
1.3.4 Properties of Vector Operations 8
1.4 The Dot Product 9
1.4.1 Definition and Computation 9
1.4.2 Geometric Interpretation 9
1.4.3 Finding Angles Between Vectors 10
1.4.4 Orthogonality 10
1.4.5 Properties of the Dot Product 11
1.4.6 Projections 11
1.4.7 Applications of the Dot Product 12
1.5 The Cross Product 13
1.5.1 Definition 13
1.5.2 Computing Cross Products 13
1.5.3 Geometric Interpretation 14
1.5.4 Area of a Parallelogram 14

1.5.5 Properties of the Cross Product . . . 15
1.5.6 Applications of the Cross Product . . . 15
1.6 Applications of Vectors . . . 15
1.6.1 Physics: Projectile Motion . . . 16
1.6.2 Navigation and GPS . . . 16
1.6.3 Computer Graphics: Transformations . . . 16
1.6.4 Engineering: Force Analysis . . . 17
1.7 Chapter Summary . . . 17
1.8 Practice Problems . . . 18
1.8.1 Basic Problems . . . 18
1.8.2 Intermediate Problems . . . 18
1.8.3 Challenge Problems . . . 19

2 Systems of Linear Equations **20**
2.1 Introduction to Linear Systems . . . 20
2.1.1 What is a Linear Equation? . . . 20
2.1.2 Systems of Linear Equations . . . 21
2.1.3 Geometric Interpretation . . . 22
2.1.4 Systems in Two and Three Variables . . . 23
2.2 Solving Systems by Elimination . . . 23
2.2.1 Elementary Row Operations . . . 23
2.2.2 The Gaussian Elimination Algorithm . . . 24
2.2.3 Consistent, Inconsistent, and Dependent Systems . . . 25
2.3 Matrix Representation of Systems . . . 26
2.3.1 Coefficient Matrices and Augmented Matrices . . . 27
2.3.2 Row Operations on Matrices . . . 27
2.4 Row Echelon and Reduced Row Echelon Form . . . 28
2.4.1 Row Echelon Form . . . 28
2.4.2 Reduced Row Echelon Form . . . 29
2.4.3 Gauss-Jordan Elimination . . . 29
2.4.4 Leading Variables and Free Variables . . . 31
2.5 Applications of Linear Systems . . . 32
2.5.1 Network Flow Problems . . . 32
2.5.2 Chemical Equation Balancing . . . 33
2.5.3 Economics: Supply and Demand . . . 34
2.5.4 Circuit Analysis . . . 34
2.6 Chapter Summary . . . 35
2.7 Practice Problems . . . 36
2.7.1 Basic Problems . . . 36
2.7.2 Intermediate Problems . . . 37

2.7.3 Challenge Problems . . . 38

3 Matrices **40**

3.1 Introduction to Matrices . . . 40
3.1.1 Matrix Notation and Terminology . . . 40
3.1.2 Types of Matrices . . . 41
3.1.3 Matrix Equality . . . 42
3.1.4 Real-World Contexts for Matrices . . . 42
3.2 Matrix Operations . . . 43
3.2.1 Matrix Addition and Subtraction . . . 43
3.2.2 Scalar Multiplication . . . 44
3.2.3 Properties of Matrix Addition and Scalar Multiplication . . . 44
3.2.4 Matrix Multiplication . . . 45
3.2.5 Why Matrix Multiplication Works This Way . . . 47
3.2.6 Properties of Matrix Multiplication . . . 47
3.2.7 Powers of Matrices . . . 48
3.3 Matrix Transpose . . . 49
3.3.1 Definition and Properties . . . 49
3.3.2 Symmetric Matrices . . . 50
3.4 Special Matrices . . . 51
3.4.1 Identity Matrix (Revisited) . . . 51
3.4.2 Diagonal Matrices . . . 51
3.4.3 Upper and Lower Triangular Matrices . . . 52
3.4.4 Introduction to Matrix Inverse . . . 52
3.5 Matrix Inverse . . . 53
3.5.1 Computing 2×2 Inverses . . . 53
3.5.2 Computing Inverses Using Row Operations . . . 54
3.5.3 Properties of Matrix Inverses . . . 56
3.5.4 When Does an Inverse Exist? . . . 56
3.5.5 Solving Systems Using Inverses . . . 56
3.6 Applications of Matrices . . . 57
3.6.1 Image Processing and Filters . . . 58
3.6.2 Graph Theory and Adjacency Matrices . . . 58
3.6.3 Markov Chains and Probability . . . 58
3.6.4 Leslie Matrices in Population Models . . . 59
3.6.5 Cryptography . . . 59
3.7 Chapter Summary . . . 60
3.8 Practice Problems . . . 61
3.8.1 Basic Problems . . . 61
3.8.2 Intermediate Problems . . . 62

3.8.3 Challenge Problems . . . 62

4 Determinants **64**

4.1 Introduction to Determinants . . . 64
4.1.1 Determinants of 2×2 Matrices . . . 64
4.1.2 Geometric Interpretation in 2D . . . 65
4.1.3 Connection to Invertibility . . . 66
4.2 Computing Determinants . . . 66
4.2.1 Determinant of 3×3 Matrices . . . 66
4.2.2 Cofactor Expansion (Laplace Expansion) . . . 67
4.2.3 Determinants of Triangular Matrices . . . 69
4.2.4 Row Operations and Determinants . . . 69
4.3 Properties and Applications of Determinants . . . 71
4.3.1 Fundamental Properties . . . 71
4.3.2 Determinant Product Rule: Proof Sketch . . . 72
4.3.3 Determinant and Volume in Higher Dimensions . . . 72
4.3.4 Cramer's Rule . . . 73
4.3.5 The Adjugate Matrix and Inverse Formula . . . 74
4.4 Applications of Determinants . . . 75
4.4.1 Area and Volume Calculations . . . 75
4.4.2 Linear Independence . . . 76
4.4.3 Eigenvalues (Preview) . . . 76
4.4.4 Differential Equations . . . 77
4.5 Chapter Summary . . . 77
4.6 Practice Problems . . . 78
4.6.1 Basic Problems . . . 78
4.6.2 Intermediate Problems . . . 79
4.6.3 Challenge Problems . . . 80

5 Vector Spaces **81**

5.1 Introduction to Vector Spaces . . . 81
5.1.1 Motivation: Beyond $\mathbb{R}^n$ . . . 81
5.1.2 Definition of a Vector Space . . . 82
5.1.3 Examples of Vector Spaces . . . 83
5.1.4 Verifying Vector Space Axioms . . . 83
5.1.5 Non-Examples . . . 84
5.2 Subspaces . . . 84
5.2.1 Definition and Examples . . . 85
5.2.2 Operations on Subspaces . . . 86
5.3 Linear Independence . . . 87

5.3.1 Definition and Intuition 87
5.3.2 Testing for Linear Independence 88
5.3.3 Properties of Linear Independence 89
5.4 Span and Basis 90
5.4.1 Span of a Set of Vectors 90
5.4.2 Spanning Sets 91
5.4.3 Basis 91
5.4.4 Coordinates Relative to a Basis 92
5.5 Dimension 93
5.5.1 The Dimension Theorem 93
5.5.2 Finding a Basis and Dimension 94
5.5.3 Dimension and Linear Independence/Spanning 94
5.6 Column Space, Row Space, and Null Space 95
5.6.1 Column Space 95
5.6.2 Null Space 96
5.6.3 Rank-Nullity Theorem 97
5.7 Chapter Summary 97
5.8 Practice Problems 97
5.8.1 Basic Problems 97
5.8.2 Intermediate Problems 98
5.8.3 Challenge Problems 99

6 Linear Transformations 100
6.1 Introduction to Linear Transformations 100
6.1.1 Functions Between Vector Spaces 100
6.1.2 Definition of Linear Transformation 101
6.1.3 Properties of Linear Transformations 103
6.2 Matrix Representation of Linear Transformations 103
6.2.1 The Standard Matrix 103
6.2.2 Geometric Linear Transformations in $\mathbb{R}^2$ 104
6.2.3 Composition of Transformations 106
6.3 Kernel and Range 106
6.3.1 Kernel (Null Space) 107
6.3.2 Range (Image) 108
6.3.3 Rank-Nullity Theorem (Revisited) 109
6.4 One-to-One and Onto Transformations 109
6.4.1 One-to-One (Injective) 109
6.4.2 Onto (Surjective) 110
6.4.3 Isomorphisms 110
6.5 Applications of Linear Transformations 111

6.5.1 Computer Graphics . . . 111
6.5.2 Differential Equations . . . 111
6.5.3 Data Transformations . . . 112
6.6 Chapter Summary . . . 112
6.7 Practice Problems . . . 113
6.7.1 Basic Problems . . . 113
6.7.2 Intermediate Problems . . . 113
6.7.3 Challenge Problems . . . 114

7 Eigenvalues and Eigenvectors **115**
7.1 Introduction to Eigenvalues and Eigenvectors . . . 115
7.1.1 Motivation: Special Directions . . . 115
7.1.2 Defining Eigenvalues and Eigenvectors . . . 116
7.1.3 The Eigenspace . . . 117
7.1.4 A Simple Example . . . 117
7.2 Finding Eigenvalues and Eigenvectors . . . 118
7.2.1 The Characteristic Equation . . . 118
7.2.2 The Characteristic Polynomial . . . 119
7.2.3 Finding Eigenvectors . . . 120
7.2.4 The Process Summarized . . . 121
7.2.5 A 3×3 Example . . . 121
7.2.6 Algebraic and Geometric Multiplicity . . . 122
7.3 Diagonalization . . . 123
7.3.1 What is Diagonalization? . . . 123
7.3.2 The Diagonalization Theorem . . . 123
7.3.3 How to Diagonalize a Matrix . . . 124
7.3.4 When is a Matrix Diagonalizable? . . . 125
7.3.5 Computing Powers of Matrices . . . 125
7.4 Applications of Eigenvalues and Eigenvectors . . . 126
7.4.1 Google PageRank Algorithm . . . 127
7.4.2 Population Dynamics . . . 128
7.4.3 Vibration Analysis and Natural Frequencies . . . 128
7.4.4 Principal Component Analysis (PCA) . . . 129
7.4.5 Differential Equations . . . 130
7.5 Symmetric Matrices and Orthogonal Diagonalization . . . 131
7.5.1 Special Properties of Symmetric Matrices . . . 131
7.5.2 Orthogonal Matrices . . . 132
7.5.3 Orthogonal Diagonalization . . . 132
7.5.4 Why Symmetric Matrices Matter . . . 133
7.6 Chapter Summary . . . 133

7.7 Practice Problems . . . 134
7.7.1 Basic Problems . . . 134
7.7.2 Intermediate Problems . . . 135
7.7.3 Challenge Problems . . . 135

8 Inner Product Spaces 137
8.1 Inner Products . . . 137
8.1.1 Generalizing the Dot Product . . . 137
8.1.2 Examples of Inner Products . . . 138
8.1.3 Norm and Distance . . . 139
8.1.4 Angle Between Vectors . . . 140
8.1.5 The Cauchy-Schwarz Inequality . . . 141
8.2 Orthogonality . . . 142
8.2.1 Orthogonal and Orthonormal Vectors . . . 142
8.2.2 Orthogonal Sets and Bases . . . 143
8.2.3 Orthogonal Complements . . . 144
8.3 Gram-Schmidt Process . . . 145
8.3.1 Motivation . . . 145
8.3.2 The Gram-Schmidt Algorithm . . . 145
8.3.3 Step-by-Step Example . . . 146
8.3.4 QR Factorization . . . 147
8.4 Orthogonal Projections . . . 148
8.4.1 Projection onto a Vector . . . 148
8.4.2 Projection onto a Subspace . . . 149
8.4.3 Best Approximation . . . 149
8.5 Least Squares Approximation . . . 150
8.5.1 The Least Squares Problem . . . 150
8.5.2 The Normal Equations . . . 150
8.5.3 Linear Regression: Fitting a Line . . . 151
8.5.4 Polynomial Fitting . . . 152
8.6 Chapter Summary . . . 152
8.7 Practice Problems . . . 153
8.7.1 Basic Problems . . . 153
8.7.2 Intermediate Problems . . . 154
8.7.3 Challenge Problems . . . 155

9 Applications and Advanced Topics 156
9.1 Markov Chains . . . 156
9.1.1 Introduction to Markov Chains . . . 156
9.1.2 Transition Matrices . . . 157

9.1.3 Steady-State Vectors . . . 157
9.1.4 Google PageRank . . . 158
9.2 Linear Programming (Introduction) . . . 159
9.2.1 What is Linear Programming? . . . 159
9.2.2 Geometric Interpretation in 2D . . . 160
9.2.3 The Simplex Method (Overview) . . . 161
9.3 Singular Value Decomposition (Introduction) . . . 161
9.3.1 What is SVD? . . . 161
9.3.2 Geometric Interpretation . . . 162
9.3.3 Computing SVD (Simple Example) . . . 162
9.3.4 Image Compression with SVD . . . 162
9.4 Linear Algebra in Data Science . . . 163
9.4.1 Data Matrices . . . 163
9.4.2 Dimensionality Reduction with PCA . . . 164
9.4.3 Machine Learning: Linear Models . . . 164
9.4.4 Covariance Matrices . . . 165
9.5 Linear Algebra in Computer Graphics . . . 165
9.5.1 3D Transformations . . . 165
9.5.2 Homogeneous Coordinates . . . 166
9.5.3 Viewing Transformations . . . 167
9.5.4 Lighting and Shading . . . 167
9.6 Differential Equations and Systems . . . 167
9.6.1 Systems of Linear Differential Equations . . . 167
9.6.2 Solution Using Eigenvalues . . . 168
9.6.3 Stability Analysis . . . 169
9.7 Chapter Summary . . . 169
9.8 Practice Problems . . . 170
9.8.1 Basic Problems . . . 170
9.8.2 Intermediate Problems . . . 171
9.8.3 Challenge Problems . . . 172

Mathematical Notation and Symbols **174**
.1 Sets and Numbers . . . 174
.2 Greek Letters . . . 175
.3 Vector and Matrix Notation . . . 176
.4 Summation and Product Notation . . . 176
.5 Logic and Proof Notation . . . 177
.6 Special Symbols . . . 177

Proof Techniques **178**

.7 Introduction to Proofs . . . 178
.7.1 Structure of a Proof . . . 178
.8 Direct Proof . . . 178
.9 Proof by Contrapositive . . . 179
.10 Proof by Contradiction . . . 179
.11 Proof by Induction . . . 180
.12 Existence Proofs . . . 181
.13 Counterexamples . . . 181

Review of Prerequisite Concepts **183**
.14 Algebra Review . . . 183
.14.1 Factoring . . . 183
.14.2 Exponent Rules . . . 183
.14.3 Logarithm Rules . . . 183
.15 Trigonometry Review . . . 184
.15.1 Trigonometric Functions . . . 184
.15.2 Special Angles . . . 184
.15.3 Pythagorean Identity . . . 184
.15.4 Angle Sum Formulas . . . 185
.15.5 Double Angle Formulas . . . 185
.16 Calculus Review . . . 185
.16.1 Derivatives . . . 185
.16.2 Integration . . . 186
.16.3 Taylor Series . . . 186
.17 Complex Numbers . . . 186
.17.1 Definition . . . 186
.17.2 Operations . . . 187
.17.3 Euler's Formula . . . 187

Computational Tools and Software **188**
.18 Calculator Techniques . . . 188
.18.1 Matrix Operations on Graphing Calculators . . . 188
.19 Python and NumPy . . . 188
.19.1 Getting Started . . . 189
.19.2 Creating Matrices and Vectors . . . 189
.19.3 Matrix Operations . . . 189
.19.4 Vector Operations . . . 190
.20 MATLAB . . . 190
.20.1 Basic Syntax . . . 190
.21 Online Tools . . . 191

.21.1 Wolfram Alpha 191
.21.2 Desmos Matrix Calculator 191
.21.3 Symbolab 192
.22 Tips for Computational Work 192

Solutions to Selected Exercises **193**
.23 About the Solutions 193
.24 How to Use These Solutions 193
.25 Notation in Solutions 194

Further Resources **195**
.26 Recommended Textbooks 195
.26.1 For Further Study 195
.26.2 For Different Perspectives 195
.27 Online Resources 196
.27.1 Video Lectures 196
.27.2 Interactive Tools 196
.27.3 Practice Problems 197
.28 Applications and Connections 197
.28.1 Computer Science 197
.28.2 Engineering 197
.28.3 Sciences 197
.29 Career Connections 198
.30 Advanced Topics to Explore 198

1

Introduction to Vectors

> *"'In mathematics, you don't understand things. You just get used to them."*
>
> — *John von Neumann*

Chapter Overview

Vectors are one of the most fundamental concepts in mathematics, physics, and engineering. Unlike the numbers you've worked with throughout your mathematical education, vectors carry both magnitude (size) and direction. In this chapter, we'll explore what vectors are, how to work with them, and why they're incredibly useful for modeling the world around us—from the motion of objects to the rendering of video game graphics.

1.1 What Are Vectors?

1.1.1 Introducing Vectors

Imagine you're giving directions to a friend. Simply saying "walk 5 blocks" isn't enough—your friend needs to know which direction to walk! This is the essence of a vector: information that includes both *how much* (magnitude) and *which way* (direction).

Definition: Vector

A **vector** is a mathematical object that has both magnitude (length or size) and direction. Vectors are often represented by arrows, where the length of the arrow represents the magnitude and the arrow points in the direction of the vector.

In contrast, a **scalar** is just a number—it has magnitude but no direction. Examples include temperature (75°F), mass (50 kg), or time (3 hours).

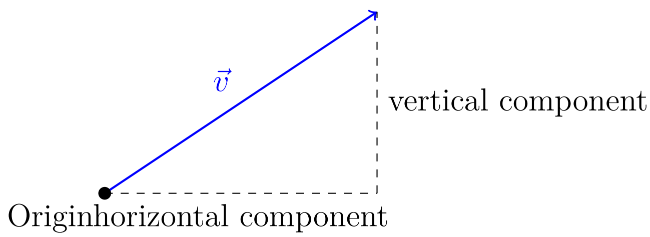


Figure 1.1: A vector $\vec{v}$ represented as an arrow

1.1.2 Notation and Representation

We denote vectors in several ways:

- **Boldface**: $\mathbf{v}$ or $\mathbf{u}$
- **Arrow notation**: $\vec{v}$ or $\vec{u}$
- **Component form**: $\langle v_1, v_2 \rangle$ or (v_1, v_2) in 2D

Throughout this textbook, we'll primarily use arrow notation ($\vec{v}$) for handwritten work and boldface ($\mathbf{v}$) in typed text.

Insight: Vectors vs. Points

It's important to distinguish between vectors and points. A point represents a location in space, while a vector represents a displacement or movement. However, we often represent vectors using the coordinates of their "tip" when their "tail" is at the origin.

1.1.3 Real-World Examples

Vectors appear everywhere in the real world:

- **Physics**: Velocity (speed + direction), force, acceleration, momentum
- **Navigation**: "Travel 50 miles northeast"
- **Computer Graphics**: Position of objects, movement in games
- **Engineering**: Stress and strain in materials, electric and magnetic fields
- **Economics**: Multi-dimensional data (price, quantity, time)

Example 1.1.1. *Hurricane Wind Velocity A hurricane has winds blowing at 120 mph toward the northwest. This is naturally described by a vector: the magnitude is 120 mph and the direction is northwest (typically 315° or $\frac{7\pi}{4}$ radians from east).*

1.2 Vector Notation and Representation

1.2.1 Component Form

In two dimensions, we can describe any vector by how far it moves in the horizontal (x) direction and vertical (y) direction. This is called **component form**.

Definition: Component Form (2D)

A vector $\vec{v}$ in two dimensions can be written as:

$$\vec{v} = \langle v_1, v_2 \rangle = v_1\hat{i} + v_2\hat{j}$$

where v_1 is the x-component, v_2 is the y-component, and $\hat{i}$ and $\hat{j}$ are unit vectors in the x and y directions.

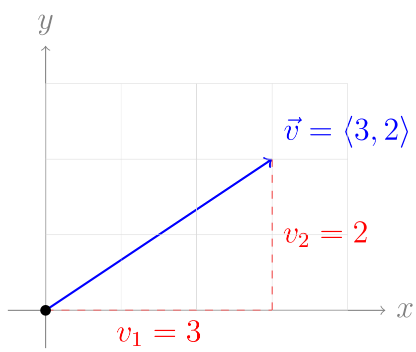


Figure 1.2: Vector $\vec{v} = \langle 3, 2 \rangle$ in component form

1.2.2 Column Vector Notation

Another common way to write vectors is as a column:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

This notation becomes especially useful when we study matrices and linear transformations later. For now, know that $\langle 3, 2 \rangle$ and $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ represent the same vector.

1.2.3 Vectors in Three Dimensions

Everything we've discussed extends naturally to three dimensions:

Definition: Component Form (3D)

A vector $\vec{v}$ in three dimensions can be written as:

$$\vec{v} = \langle v_1, v_2, v_3 \rangle = v_1\hat{i} + v_2\hat{j} + v_3\hat{k} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

where v_1, v_2, and v_3 are the x, y, and z components.

Example 1.2.1. *3D Vector The vector $\vec{w} = \langle 2, -1, 3 \rangle$ moves 2 units in the positive x direction, 1 unit in the negative y direction, and 3 units in the positive z direction.*

1.2.4 Magnitude of a Vector

The **magnitude** (or **length**) of a vector is the distance from its tail to its tip. We denote the magnitude of $\vec{v}$ as $\|\vec{v}\|$ or $|\vec{v}|$.

Formula: Magnitude

For a vector $\vec{v} = \langle v_1, v_2 \rangle$ in 2D:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2}$$

For a vector $\vec{v} = \langle v_1, v_2, v_3 \rangle$ in 3D:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

This formula comes directly from the Pythagorean theorem!

Example 1.2.2. *Finding Magnitude Find the magnitude of $\vec{v} = \langle 3, -4 \rangle$.*

Solution:

$$\begin{aligned} \|\vec{v}\| &= \sqrt{3^2 + (-4)^2} \\ &= \sqrt{9 + 16} \\ &= \sqrt{25} \\ &= 5 \end{aligned}$$

The vector $\vec{v}$ has magnitude 5.

1.2.5 Direction of a Vector

In 2D, the direction of a vector is often specified by the angle θ it makes with the positive x-axis.

Formula: Direction Angle

For a vector $\vec{v} = \langle v_1, v_2 \rangle$:

$$\theta = \arctan\left(\frac{v_2}{v_1}\right)$$

(Be careful with quadrants when using this formula!)

Common Mistake: Arctan and Quadrants

The arctan function only gives values between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$, which means it can't distinguish between all four quadrants. Always check the signs of v_1 and v_2 to determine the correct quadrant. Many calculators have an `atan2(y,x)` function that handles this automatically.

1.2.6 Unit Vectors

A **unit vector** is a vector with magnitude 1. Unit vectors are useful for indicating direction without specifying magnitude.

Definition: Unit Vector

A vector $\hat{u}$ is a unit vector if $\|\hat{u}\| = 1$. We often denote unit vectors with a "hat" symbol: $\hat{u}$.

The standard unit vectors in 2D and 3D are:

- $\hat{i} = \langle 1, 0 \rangle$ (points in the positive x direction)
- $\hat{j} = \langle 0, 1 \rangle$ (points in the positive y direction)
- $\hat{k} = \langle 0, 0, 1 \rangle$ (points in the positive z direction in 3D)

To convert any non-zero vector into a unit vector pointing in the same direction, we **normalize** it:

Formula: Normalization

The unit vector in the direction of $\vec{v}$ is:

$$\hat{v} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\|\vec{v}\|}\vec{v}$$

Example 1.2.3. *Finding a Unit Vector Find the unit vector in the direction of $\vec{v} = \langle 3, 4 \rangle$.*

Solution: *First, find the magnitude:*

$$\|\vec{v}\| = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = 5$$

Now normalize:

$$\hat{v} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{5}\langle 3, 4\rangle = \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle$$

We can verify: $\|\hat{v}\| = \sqrt{\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = \sqrt{\frac{9+16}{25}} = \sqrt{\frac{25}{25}} = 1$ ✓

1.3 Vector Operations

Now that we understand what vectors are, let's explore how to perform operations with them.

1.3.1 Vector Addition

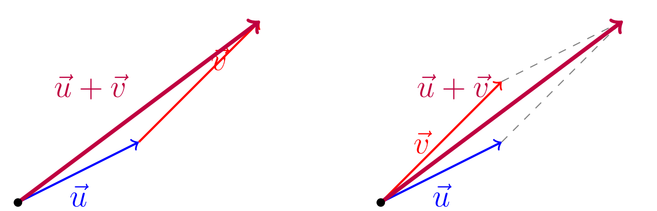
When we add two vectors, we combine their effects. Geometrically, we place the tail of the second vector at the tip of the first vector; the sum is the vector from the tail of the first to the tip of the second.

Definition: Vector Addition

For vectors $\vec{u} = \langle u_1, u_2\rangle$ and $\vec{v} = \langle v_1, v_2\rangle$:

$$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2\rangle$$

Simply add corresponding components.



Tip-to-tail method   Parallelogram method

Figure 1.3: Two geometric interpretations of vector addition

Example 1.3.1. *Adding Vectors Let* $\vec{u} = \langle 2, 3\rangle$ *and* $\vec{v} = \langle -1, 4\rangle$. *Find* $\vec{u} + \vec{v}$.

Solution:

$$\vec{u} + \vec{v} = \langle 2 + (-1), 3 + 4\rangle = \langle 1, 7\rangle$$

Physical Interpretation

If you walk 2 miles east and 3 miles north (vector $\vec{u}$), then walk 1 mile west and 4 miles north (vector $\vec{v}$), your net displacement is 1 mile east and 7 miles north ($\vec{u} + \vec{v}$).

1.3.2 Vector Subtraction

Vector subtraction can be thought of as adding the opposite vector.

Definition: Vector Subtraction

For vectors $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$:

$$\vec{u} - \vec{v} = \langle u_1 - v_1, u_2 - v_2 \rangle = \vec{u} + (-\vec{v})$$

Geometrically, $\vec{u} - \vec{v}$ is the vector from the tip of $\vec{v}$ to the tip of $\vec{u}$ (when both start at the origin).

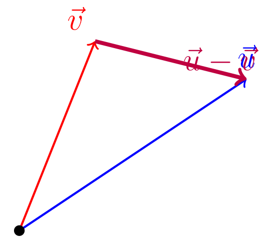


Figure 1.4: Geometric interpretation of $\vec{u} - \vec{v}$

1.3.3 Scalar Multiplication

Multiplying a vector by a scalar changes its magnitude and possibly its direction.

Definition: Scalar Multiplication

For a scalar c and vector $\vec{v} = \langle v_1, v_2 \rangle$:

$$c\vec{v} = \langle cv_1, cv_2 \rangle$$

Effects of scalar multiplication:

- If $c > 1$: stretches the vector (makes it longer)
- If $0 < c < 1$: shrinks the vector (makes it shorter)
- If $c < 0$: reverses the direction and scales by $|c|$
- If $c = 0$: produces the zero vector $\vec{0} = \langle 0, 0 \rangle$

Example 1.3.2. *Scalar Multiplication Let $\vec{v} = \langle 3, -2 \rangle$. Find $-2\vec{v}$ and $\frac{1}{2}\vec{v}$.*

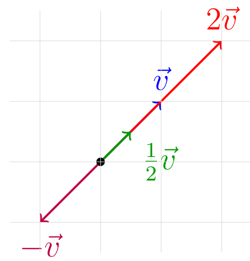


Figure 1.5: Scalar multiplication effects

Solution:

$$-2\vec{v} = -2\langle 3, -2\rangle = \langle -6, 4\rangle$$
$$\frac{1}{2}\vec{v} = \frac{1}{2}\langle 3, -2\rangle = \left\langle \frac{3}{2}, -1 \right\rangle$$

1.3.4 Properties of Vector Operations

Vector addition and scalar multiplication satisfy several important properties:

Properties of Vector Addition

For vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$:

1. **Commutativity**: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
2. **Associativity**: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
3. **Identity**: $\vec{v} + \vec{0} = \vec{v}$ (where $\vec{0}$ is the zero vector)
4. **Inverse**: $\vec{v} + (-\vec{v}) = \vec{0}$

Properties of Scalar Multiplication

For scalars c and d, and vectors $\vec{u}$ and $\vec{v}$:

1. **Associativity**: $c(d\vec{v}) = (cd)\vec{v}$
2. **Identity**: $1\vec{v} = \vec{v}$
3. **Distributivity over vector addition**: $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
4. **Distributivity over scalar addition**: $(c + d)\vec{v} = c\vec{v} + d\vec{v}$

These properties might seem obvious, but they form the foundation for vector spaces, which we'll study in Chapter 5.

Example 1.3.3. *Using Vector Properties Simplify:* $3(\vec{u} - 2\vec{v}) + 4\vec{v}$

Solution:

$$\begin{aligned} 3(\vec{u} - 2\vec{v}) + 4\vec{v} &= 3\vec{u} - 6\vec{v} + 4\vec{v} \\ &= 3\vec{u} + (-6\vec{v} + 4\vec{v}) \\ &= 3\vec{u} - 2\vec{v} \end{aligned}$$

1.4 The Dot Product

The dot product (also called the **scalar product** or **inner product**) is our first way of multiplying two vectors. Unlike addition, the dot product produces a scalar, not a vector.

1.4.1 Definition and Computation

Definition: Dot Product

For vectors $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$:

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2$$

In three dimensions, for $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$:

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Example 1.4.1. *Computing Dot Products Let $\vec{u} = \langle 2, -3 \rangle$ and $\vec{v} = \langle 4, 1 \rangle$. Find $\vec{u} \cdot \vec{v}$.*

Solution:

$$\vec{u} \cdot \vec{v} = (2)(4) + (-3)(1) = 8 - 3 = 5$$

1.4.2 Geometric Interpretation

The dot product has a beautiful geometric meaning:

Geometric Formula for Dot Product

$$\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos\theta$$

where θ is the angle between the vectors.

This formula connects algebra (component multiplication) with geometry (angles)!

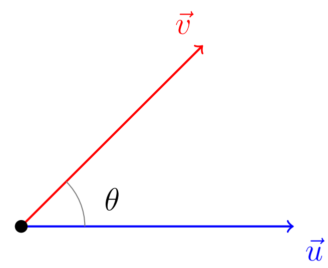


Figure 1.6: The angle θ between vectors $\vec{u}$ and $\vec{v}$

1.4.3 Finding Angles Between Vectors

We can rearrange the geometric formula to find angles:

Formula: Angle Between Vectors

$$\cos\theta = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}$$

Therefore:

$$\theta = \arccos\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}\right)$$

Example 1.4.2. *Finding the Angle Find the angle between* $\vec{u} = \langle 1, 2\rangle$ *and* $\vec{v} = \langle 3, 1\rangle$.

Solution: *First, compute the dot product:*

$$\vec{u}\cdot\vec{v} = (1)(3) + (2)(1) = 5$$

Next, find the magnitudes:

$$\|\vec{u}\| = \sqrt{1^2 + 2^2} = \sqrt{5}$$
$$\|\vec{v}\| = \sqrt{3^2 + 1^2} = \sqrt{10}$$

Now find the angle:

$$\cos\theta = \frac{5}{\sqrt{5}\cdot\sqrt{10}} = \frac{5}{\sqrt{50}} = \frac{5}{5\sqrt{2}} = \frac{1}{\sqrt{2}}$$
$$\theta = \arccos\left(\frac{1}{\sqrt{2}}\right) = 45\text{ř} = \frac{\pi}{4}\ \textit{radians}$$

1.4.4 Orthogonality

Two vectors are **orthogonal** (perpendicular) if the angle between them is 90°.

Test for Orthogonality

Vectors $\vec{u}$ and $\vec{v}$ are orthogonal if and only if:

$$\vec{u} \cdot \vec{v} = 0$$

This follows from the geometric formula: if $\theta = 90ř$, then $\cos 90ř = 0$, so $\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\|(0) = 0$.

Example 1.4.3. *Testing for Orthogonality Are $\vec{u} = \langle 2, 3\rangle$ and $\vec{v} = \langle -3, 2\rangle$ orthogonal?*

Solution:

$$\vec{u} \cdot \vec{v} = (2)(-3) + (3)(2) = -6 + 6 = 0$$

Yes! Since the dot product is zero, the vectors are orthogonal.

Insight: Why This Matters

Orthogonality is fundamental in linear algebra. Perpendicular directions are independent—moving in one direction doesn't affect your position in the perpendicular direction. This idea generalizes to higher dimensions and becomes crucial for understanding coordinate systems, projections, and decompositions.

1.4.5 Properties of the Dot Product

Properties of the Dot Product

For vectors $\vec{u}$, $\vec{v}$, $\vec{w}$ and scalar c:

1. **Commutativity**: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
2. **Distributivity**: $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$
3. **Scalar multiplication**: $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$
4. **Positive definiteness**: $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2 \geq 0$, with equality only when $\vec{v} = \vec{0}$

1.4.6 Projections

One of the most important applications of the dot product is computing **projections**. The projection of $\vec{v}$ onto $\vec{u}$ is the vector that represents the "shadow" of $\vec{v}$ in the direction of $\vec{u}$.

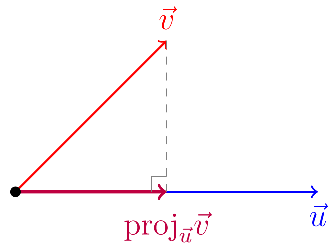


Figure 1.7: Projection of $\vec{v}$ onto $\vec{u}$

Formula: Vector Projection

The projection of $\vec{v}$ onto $\vec{u}$ is:

$$\text{proj}_{\vec{u}}\vec{v} = \frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|^2}\vec{u} = \frac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u}$$

The scalar projection (the length of the projection) is:

$$\text{comp}_{\vec{u}}\vec{v} = \frac{\vec{v} \cdot \vec{u}}{\|\vec{u}\|}$$

Example 1.4.4. *Computing a Projection Find the projection of $\vec{v} = \langle 3, 4 \rangle$ onto $\vec{u} = \langle 1, 0 \rangle$.*

Solution: *First, compute the necessary values:*

$$\vec{v} \cdot \vec{u} = (3)(1) + (4)(0) = 3$$
$$\vec{u} \cdot \vec{u} = (1)(1) + (0)(0) = 1$$

Now find the projection:

$$proj_{\vec{u}}\vec{v} = \frac{3}{1}\langle 1, 0 \rangle = \langle 3, 0 \rangle$$

This makes sense: we're projecting onto the x-axis, so we just keep the x-component!

1.4.7 Applications of the Dot Product

1. Work in Physics

In physics, if a constant force $\vec{F}$ moves an object through a displacement $\vec{d}$, the work done is:

$$W = \vec{F} \cdot \vec{d}$$

Example 1.4.5. *Calculating Work A force of $\vec{F} = \langle 10, 5 \rangle$ Newtons moves an object from the origin to point $(3, 2)$ meters. How much work is done?*

Solution: *The displacement is* $\vec{d} = \langle 3, 2 \rangle$. *The work is:*

$$W = \vec{F} \cdot \vec{d} = (10)(3) + (5)(2) = 30 + 10 = 40 \ Joules$$

2. Computer Graphics

In 3D graphics, the dot product is used for lighting calculations. The brightness of a surface depends on the angle between the surface normal (perpendicular vector) and the light direction.

3. Machine Learning

In data science, the dot product measures similarity between data vectors. Larger dot products indicate more similar data points.

1.5 The Cross Product

The cross product is defined only in three dimensions and produces a vector (unlike the dot product, which produces a scalar).

1.5.1 Definition

Definition: Cross Product

For vectors $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$:

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \langle u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1 \rangle$$

Important Note

The cross product is only defined in three dimensions! It doesn't exist for 2D vectors.

1.5.2 Computing Cross Products

The determinant formula might look intimidating, but it's systematic:

Example 1.5.1. *Computing a Cross Product Find* $\vec{u} \times \vec{v}$ *where* $\vec{u} = \langle 1, 2, 3 \rangle$ *and* $\vec{v} = \langle 4, 5, 6 \rangle$.

Solution: *Using the formula:*

$$\begin{aligned}\vec{u}\times\vec{v} &= \langle u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1\rangle \\ &= \langle (2)(6) - (3)(5), (3)(4) - (1)(6), (1)(5) - (2)(4)\rangle \\ &= \langle 12 - 15, 12 - 6, 5 - 8\rangle \\ &= \langle -3, 6, -3\rangle\end{aligned}$$

1.5.3 Geometric Interpretation

The cross product $\vec{u}\times\vec{v}$ has three important geometric properties:

1. **Direction**: $\vec{u}\times\vec{v}$ is perpendicular to both $\vec{u}$ and $\vec{v}$
2. **Magnitude**: $\|\vec{u}\times\vec{v}\| = \|\vec{u}\|\|\vec{v}\|\sin\theta$ where θ is the angle between the vectors
3. **Orientation**: The direction follows the right-hand rule

Right-Hand Rule

Point your right hand's fingers in the direction of $\vec{u}$, curl them toward $\vec{v}$, and your thumb points in the direction of $\vec{u}\times\vec{v}$.

1.5.4 Area of a Parallelogram

The magnitude of the cross product gives the area of the parallelogram formed by the two vectors!

Formula: Area via Cross Product

The area of the parallelogram with adjacent sides $\vec{u}$ and $\vec{v}$ is:

$$A = \|\vec{u}\times\vec{v}\|$$

Example 1.5.2. *Finding Area Find the area of the parallelogram with adjacent sides $\vec{u} = \langle 1, 0, 1\rangle$ and $\vec{v} = \langle 0, 1, 1\rangle$.*

Solution: *First, compute the cross product:*

$$\begin{aligned}\vec{u}\times\vec{v} &= \langle (0)(1) - (1)(1), (1)(0) - (1)(1), (1)(1) - (0)(0)\rangle \\ &= \langle -1, -1, 1\rangle\end{aligned}$$

Now find the magnitude:

$$\|\vec{u}\times\vec{v}\| = \sqrt{(-1)^2 + (-1)^2 + 1^2} = \sqrt{3}$$

The area is $\sqrt{3}$ square units.

1.5.5 Properties of the Cross Product

Properties of the Cross Product

1. **Anti-commutativity**: $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$
2. **Distributivity**: $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$
3. **Scalar multiplication**: $(c\vec{u}) \times \vec{v} = c(\vec{u} \times \vec{v}) = \vec{u} \times (c\vec{v})$
4. **Parallel vectors**: $\vec{u} \times \vec{v} = \vec{0}$ if and only if $\vec{u}$ and $\vec{v}$ are parallel
5. **Not associative**: $\vec{u} \times (\vec{v} \times \vec{w}) \neq (\vec{u} \times \vec{v}) \times \vec{w}$ in general

Common Mistake

Unlike the dot product, the cross product is NOT commutative. In fact, $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$. Order matters!

1.5.6 Applications of the Cross Product

1. Torque in Physics

Torque $\vec{\tau}$ (rotational force) is given by:

$$\vec{\tau} = \vec{r} \times \vec{F}$$

where $\vec{r}$ is the position vector and $\vec{F}$ is the applied force.

2. Normal Vectors

In computer graphics, we need to find vectors perpendicular to surfaces. If two vectors lie in a plane, their cross product gives a normal vector to that plane.

3. Angular Momentum

In physics, angular momentum is:

$$\vec{L} = \vec{r} \times \vec{p}$$

where $\vec{p}$ is linear momentum.

1.6 Applications of Vectors

Let's explore some real-world applications that showcase the power of vectors.

1.6.1 Physics: Projectile Motion

When an object is launched, its position, velocity, and acceleration can all be described using vectors.

Example 1.6.1. *Projectile Motion A ball is thrown with initial velocity* $\vec{v}_0 = \langle 20, 30 \rangle$ *m/s (20 m/s horizontally, 30 m/s vertically). Gravity causes an acceleration of* $\vec{a} = \langle 0, -9.8 \rangle$ *m/s². Find the velocity after 2 seconds.*

Solution: *Using the kinematic equation* $\vec{v}(t) = \vec{v}_0 + \vec{a}t$*:*

$$\begin{aligned} \vec{v}(2) &= \langle 20, 30 \rangle + 2\langle 0, -9.8 \rangle \\ &= \langle 20, 30 \rangle + \langle 0, -19.6 \rangle \\ &= \langle 20, 10.4 \rangle \ m/s \end{aligned}$$

The ball is still moving horizontally at 20 m/s but has slowed vertically to 10.4 m/s.

1.6.2 Navigation and GPS

GPS systems use vectors to calculate positions and directions. If you're at point A and want to reach point B, the displacement vector is $\vec{AB} = B - A$.

Example 1.6.2. *Navigation You're at position* $A = (3, 5)$ *km on a map and need to reach* $B = (8, 1)$ *km. Find the displacement vector and the distance.*

Solution: *The displacement vector is:*

$$\vec{AB} = \langle 8 - 3, 1 - 5 \rangle = \langle 5, -4 \rangle \ km$$

The distance is:

$$\|\vec{AB}\| = \sqrt{5^2 + (-4)^2} = \sqrt{25 + 16} = \sqrt{41} \approx 6.4 \ km$$

You need to travel 5 km east and 4 km south.

1.6.3 Computer Graphics: Transformations

In video games and animation, objects are represented by collections of vectors (vertices). Moving, rotating, or scaling objects involves vector operations.

Connection to Later Chapters

In Chapter 6, we'll see how matrices can represent complex geometric transformations like rotations and reflections, making it easy to manipulate 3D objects in games and films.

1.6.4 Engineering: Force Analysis

Engineers use vectors to analyze forces on structures. When multiple forces act on an object, the net force is the vector sum of all individual forces.

Example 1.6.3. *Force Equilibrium A sign is held by two cables. Cable 1 exerts force* $\vec{F}_1 = \langle 30, 40 \rangle$ *N and cable 2 exerts* $\vec{F}_2 = \langle -20, 50 \rangle$ *N. Find the total force on the sign.*

Solution:

$$\vec{F}_{total} = \vec{F}_1 + \vec{F}_2 = \langle 30, 40 \rangle + \langle -20, 50 \rangle = \langle 10, 90 \rangle \ N$$

The net force is 10 N horizontally and 90 N vertically upward.

1.7 Chapter Summary

In this chapter, we've explored the fundamental concepts of vectors:

- **Vectors** have both magnitude and direction, unlike scalars which have only magnitude
- Vectors can be represented in **component form**: $\vec{v} = \langle v_1, v_2, v_3 \rangle$
- **Vector addition** combines vectors tip-to-tail; **scalar multiplication** scales vectors
- The **dot product** $\vec{u} \cdot \vec{v}$ produces a scalar and relates to angles and projections
- The **cross product** $\vec{u} \times \vec{v}$ (3D only) produces a perpendicular vector
- Vectors are essential in physics, engineering, computer graphics, and many other fields

Looking Ahead

Vectors are the building blocks of linear algebra. In the next chapter, we'll see how systems of linear equations can be solved using vectors and matrices—powerful tools that extend the ideas we've learned here to solve complex problems with many variables.

1.8 Practice Problems

1.8.1 Basic Problems

1. Find the magnitude and direction (angle with positive x-axis) of $\vec{v} = \langle 3, 3 \rangle$.

2. Let $\vec{u} = \langle 2, -5 \rangle$ and $\vec{v} = \langle -3, 1 \rangle$. Compute:

   (a) $\vec{u} + \vec{v}$

   (b) $3\vec{u} - 2\vec{v}$

   (c) $\|\vec{u}\|$

3. Find a unit vector in the direction of $\vec{w} = \langle -4, 3 \rangle$.

4. Determine if $\vec{a} = \langle 1, 2 \rangle$ and $\vec{b} = \langle 4, -2 \rangle$ are orthogonal.

5. Compute $\vec{u} \cdot \vec{v}$ where $\vec{u} = \langle 5, -2, 3 \rangle$ and $\vec{v} = \langle 1, 4, -1 \rangle$.

6. Find the angle between $\vec{u} = \langle 1, 1 \rangle$ and $\vec{v} = \langle 0, 1 \rangle$.

7. Calculate $\vec{u} \times \vec{v}$ where $\vec{u} = \langle 2, 0, 1 \rangle$ and $\vec{v} = \langle 1, 3, 0 \rangle$.

8. Find the area of the parallelogram with adjacent sides $\vec{a} = \langle 1, 2, 3 \rangle$ and $\vec{b} = \langle 2, 1, 1 \rangle$.

1.8.2 Intermediate Problems

9. Find the projection of $\vec{v} = \langle 4, 2 \rangle$ onto $\vec{u} = \langle 3, 4 \rangle$.

10. A force $\vec{F} = \langle 10, 20, -5 \rangle$ N acts on an object that moves from point $A = (1, 2, 3)$ to point $B = (4, 1, 5)$ (distances in meters). How much work is done?

11. Show that $\vec{u} = \langle 1, 2, 2 \rangle$, $\vec{v} = \langle 2, 1, -2 \rangle$, and $\vec{w} = \langle -2, 2, -1 \rangle$ are mutually orthogonal (each pair is orthogonal).

12. Find a vector perpendicular to both $\vec{u} = \langle 1, 0, 1 \rangle$ and $\vec{v} = \langle 0, 1, 1 \rangle$.

13. An airplane flies with velocity $\vec{v}_p = \langle 200, 0 \rangle$ mph (east). A wind blows with velocity $\vec{v}_w = \langle -30, 40 \rangle$ mph. What is the airplane's actual velocity relative to the ground?

14. Prove that $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + 2(\vec{u} \cdot \vec{v}) + \|\vec{v}\|^2$.

1.8.3 Challenge Problems

15. Prove the triangle inequality: $\|\vec{u}+\vec{v}\| \leq \|\vec{u}\|+\|\vec{v}\|$ for any vectors $\vec{u}$ and $\vec{v}$.
16. Find all vectors $\vec{v}$ that are orthogonal to both $\vec{u}_1 = \langle 1,2,1\rangle$ and $\vec{u}_2 = \langle 2,1,0\rangle$.
17. Prove that $\vec{u}\cdot(\vec{v}\times\vec{w})$ represents the volume of the parallelepiped formed by $\vec{u}$, $\vec{v}$, and $\vec{w}$.
18. Show that $(\vec{u}\times\vec{v})\cdot\vec{w} = (\vec{v}\times\vec{w})\cdot\vec{u} = (\vec{w}\times\vec{u})\cdot\vec{v}$. This is called the scalar triple product.
19. The angle between $\vec{u}$ and $\vec{v}$ is $\frac{\pi}{3}$. If $\|\vec{u}\| = 4$ and $\|\vec{v}\| = 3$, find $\|\vec{u}-\vec{v}\|$.
20. Prove the Cauchy-Schwarz inequality: $|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\|\vec{v}\|$ for any vectors $\vec{u}$ and $\vec{v}$.

2

Systems of Linear Equations

"'Pure mathematics is, in its way, the poetry of logical ideas."

— Albert Einstein

Chapter Overview

In the real world, we rarely deal with just one equation at a time. Whether you're balancing chemical equations, analyzing electrical circuits, or optimizing a business strategy, you'll encounter multiple equations that must be satisfied simultaneously. In this chapter, we'll develop systematic methods for solving systems of linear equations—methods that are not only theoretically elegant but also computationally powerful. These techniques form the backbone of modern computational mathematics and have applications ranging from computer graphics to machine learning.

2.1 Introduction to Linear Systems

2.1.1 What is a Linear Equation?

Before we can solve systems of equations, we need to understand what makes an equation *linear*.

Definition: Linear Equation

A **linear equation** in variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form:

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

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

Key characteristics of linear equations:

- Each variable appears to the first power only (no x^2, $\sqrt{x}$, $\sin(x)$, etc.)

- Variables are not multiplied together (no xy, x_1x_2, etc.)
- Variables do not appear in denominators, exponents, or inside functions

Example 2.1.1. *Identifying Linear Equations Which of the following are linear equations?*
(a) $3x + 2y - z = 7$ *(b)* $x^2 + y = 5$ *(c)* $2x - 3y + 4z - w = 0$ *(d)* $xy + z = 1$

Solution:

(a) Linear ✓ *(all variables to first power, no products)*

(b) Not linear (contains x^2)

(c) Linear ✓ *(standard form with four variables)*

(d) Not linear (contains product xy)

2.1.2 Systems of Linear Equations

A **system of linear equations** is a collection of linear equations involving the same variables.

Definition: System of Linear Equations

A system of m linear equations in n variables has the form:

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

A **solution** to a system is an ordered set of values for the variables that satisfies all equations simultaneously.

Example 2.1.2. *Verifying Solutions Verify that $x = 2$, $y = -1$ is a solution to the system:*

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

Solution: *Substitute into the first equation:*

$$(2) + 2(-1) = 2 - 2 = 0✓$$

Substitute into the second equation:

$$3(2) - (-1) = 6 + 1 = 7\checkmark$$

Both equations are satisfied, so $(2, -1)$ *is indeed a solution.*

2.1.3 Geometric Interpretation

Linear systems have beautiful geometric interpretations:

Two variables (2D):

- Each linear equation represents a line in the plane
- A solution is a point where the lines intersect

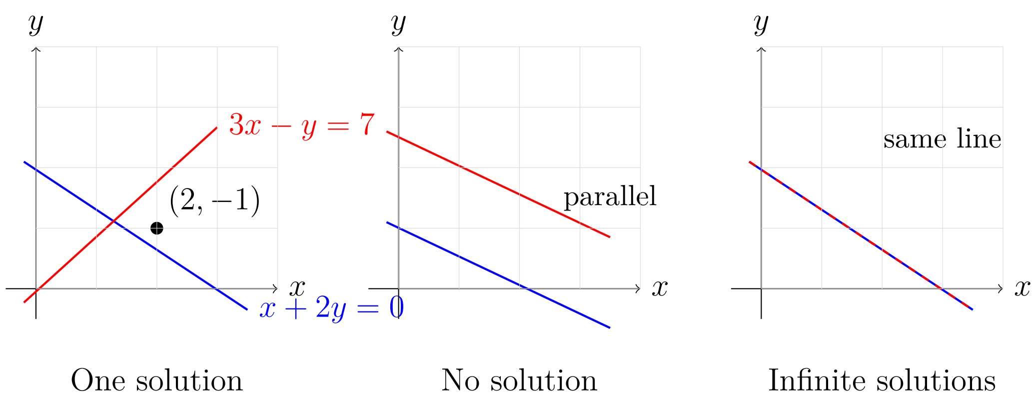


Figure 2.1: Geometric possibilities for 2×2 systems

Three variables (3D):

- Each linear equation represents a plane in 3D space
- A solution is a point where all planes intersect

Insight: Types of Solutions

A system of linear equations can have:

1. **Exactly one solution** (consistent and independent)
2. **No solution** (inconsistent) - equations contradict each other
3. **Infinitely many solutions** (consistent and dependent) - equations are redundant

2.1.4 Systems in Two and Three Variables

Let's start with familiar small systems before developing general methods.

Example 2.1.3. *Solving a 2×2 System by Substitution Solve the system:*

$$\begin{aligned} x + y &= 5 \\ x - y &= 1 \end{aligned}$$

Solution: *From the first equation:* $x = 5 - y$

Substitute into the second equation:

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

Now find x*:* $x = 5 - y = 5 - 2 = 3$

Solution: $(x, y) = (3, 2)$

For larger systems, substitution becomes cumbersome. We need a more systematic approach!

2.2 Solving Systems by Elimination

The **method of elimination** (also called **Gaussian elimination**) is a systematic procedure for solving any system of linear equations.

2.2.1 Elementary Row Operations

The key insight is that certain operations on equations preserve solutions:

Elementary Row Operations

The following operations transform a system into an equivalent system (same solution set):

1. **Swap two equations** (interchange)

2. **Multiply an equation by a nonzero constant** (scaling)

3. **Add a multiple of one equation to another** (replacement)

Why These Operations Work

These operations preserve solutions because:

- Swapping equations just reorders them
- Multiplying by a nonzero constant doesn't change the solution set
- If x satisfies both equations, it also satisfies their sum

2.2.2 The Gaussian Elimination Algorithm

The strategy is to use row operations to systematically eliminate variables, creating a triangular system that's easy to solve.

Example 2.2.1. *Gaussian Elimination for a 3×3 System Solve the system:*

$$\begin{aligned} x + 2y + z &= 9 \\ 2x + 4y + 3z &= 21 \\ 3x + 2y - z &= 8 \end{aligned}$$

Solution:

Step 1: *Eliminate x from equations 2 and 3.*

Replace equation 2 with (equation 2) - 2(equation 1):

$$\begin{aligned} (2x + 4y + 3z) - 2(x + 2y + z) &= 21 - 2(9) \\ 2x + 4y + 3z - 2x - 4y - 2z &= 21 - 18 \\ z &= 3 \end{aligned}$$

Replace equation 3 with (equation 3) - 3(equation 1):

$$\begin{aligned} (3x + 2y - z) - 3(x + 2y + z) &= 8 - 3(9) \\ 3x + 2y - z - 3x - 6y - 3z &= 8 - 27 \\ -4y - 4z &= -19 \end{aligned}$$

Our system is now:

$$
\begin{aligned}
x + 2y + z &= 9 \\
z &= 3 \\
-4y - 4z &= -19
\end{aligned}
$$

Step 2: *From equation 2, we have* $z = 3$.
Substitute into equation 3:

$$
\begin{aligned}
-4y - 4(3) &= -19 \\
-4y - 12 &= -19 \\
-4y &= -7 \\
y &= \frac{7}{4}
\end{aligned}
$$

Step 3: *Substitute* $y = \frac{7}{4}$ *and* $z = 3$ *into equation 1:*

$$
\begin{aligned}
x + 2\left(\frac{7}{4}\right) + 3 &= 9 \\
x + \frac{7}{2} + 3 &= 9 \\
x + \frac{13}{2} &= 9 \\
x &= \frac{5}{2}
\end{aligned}
$$

Solution: $(x, y, z) = \left(\frac{5}{2}, \frac{7}{4}, 3\right)$

2.2.3 Consistent, Inconsistent, and Dependent Systems

As we solve systems, we may encounter special cases:

Classification of Systems

- **Consistent**: Has at least one solution
- **Inconsistent**: Has no solution (leads to a contradiction like $0 = 5$)
- **Independent**: Has exactly one solution
- **Dependent**: Has infinitely many solutions (contains redundant equations)

Example 2.2.2. *Inconsistent System Solve:*

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

Solution: *Subtract the first equation from the second:*

$$(x + y) - (x + y) = 5 - 3 \implies 0 = 2$$

This is a contradiction! The system has no solution.

Geometric interpretation: *These are parallel lines that never intersect.*

Example 2.2.3. *Dependent System Solve:*

$$\begin{aligned} x + 2y &= 4 \\ 2x + 4y &= 8 \end{aligned}$$

Solution: *The second equation is just 2 times the first equation, so they represent the same line.*

From the first equation: $x = 4 - 2y$

The solutions are all points of the form $(4 - 2y, y)$ *where* y *can be any real number.*

We can write this as: $(x, y) = (4 - 2t, t)$ *for* $t \in \mathbb{R}$

Answer: *Infinitely many solutions:* $(4 - 2t, t)$ *for any* $t \in \mathbb{R}$

Common Mistake: Missing Infinite Solutions

When an equation reduces to $0 = 0$ during elimination, don't assume there's no solution! This means the equation was redundant, and the system has infinitely many solutions. You must express the solution in terms of free variables.

2.3 Matrix Representation of Systems

Writing out full equations with variables becomes tedious for large systems. Matrices provide a compact, efficient representation.

2.3.1 Coefficient Matrices and Augmented Matrices

Definition: Coefficient Matrix and Augmented Matrix

For the system:

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

The **coefficient matrix** is:

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

The **augmented matrix** includes the constants:

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

Example 2.3.1. *Writing an Augmented Matrix Write the augmented matrix for:*

$$\begin{aligned} 2x - y + 3z &= 8 \\ x + 2y - z &= 1 \\ 3x + y + 2z &= 5 \end{aligned}$$

Solution:

$$[A|b] = \left[\begin{array}{ccc|c} 2 & -1 & 3 & 8 \\ 1 & 2 & -1 & 1 \\ 3 & 1 & 2 & 5 \end{array}\right]$$

Each row represents one equation; each column (except the last) represents one variable.

2.3.2 Row Operations on Matrices

The elementary row operations from before translate directly to matrices:

Elementary Row Operations (Matrix Form)

1. **Row interchange**: Swap rows i and j (notation: $R_i \leftrightarrow R_j$)
2. **Row scaling**: Multiply row i by nonzero constant c (notation: cR_i)
3. **Row replacement**: Add c times row j to row i (notation: $R_i + cR_j$)

Example 2.3.2. *Performing Row Operations Given the matrix:*

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 11 \end{array}\right]$$

Perform the operation $R_2 - 3R_1$:

Solution:

$$\begin{aligned} R_2 - 3R_1 &= \begin{bmatrix} 3 & 4 & 11 \end{bmatrix} - 3\begin{bmatrix} 1 & 2 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 4 & 11 \end{bmatrix} - \begin{bmatrix} 3 & 6 & 15 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -2 & -4 \end{bmatrix} \end{aligned}$$

The resulting matrix is:

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 0 & -2 & -4 \end{array}\right]$$

2.4 Row Echelon and Reduced Row Echelon Form

To solve systems systematically, we transform matrices to special forms.

2.4.1 Row Echelon Form

Definition: Row Echelon Form (REF)

A matrix is in **row echelon form** if:

1. All nonzero rows are above rows of all zeros
2. The first nonzero entry in each nonzero row (called a **leading entry** or **pivot**) is 1
3. Each leading entry is to the right of the leading entry in the row above it

Example 2.4.1. *Examples of Row Echelon Form These matrices are in row echelon form:*

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 5 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

This matrix is NOT in row echelon form:

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix} \quad \textit{(first entry should be 1, not 0)}$$

2.4.2 Reduced Row Echelon Form

Definition: Reduced Row Echelon Form (RREF)

A matrix is in **reduced row echelon form** if:

1. It is in row echelon form
2. Each leading entry is 1 (automatically satisfied by REF definition)
3. Each leading 1 is the only nonzero entry in its column

Example 2.4.2. *Examples of RREF These matrices are in reduced row echelon form:*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Insight: Uniqueness of RREF

Every matrix has a unique reduced row echelon form! This means different sequences of row operations will always lead to the same RREF. However, there are many possible row echelon forms for the same matrix.

2.4.3 Gauss-Jordan Elimination

Gauss-Jordan elimination is the process of transforming a matrix to reduced row echelon form.

Example 2.4.3. *Complete Gauss-Jordan Elimination Solve the system by transforming the augmented matrix to RREF:*

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

Solution:

Start with the augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 2 & 3 & 3 & 7 \\ 1 & 1 & 2 & 3 \end{array}\right]$$

Step 1: *Create zeros below the first pivot.*

$R_2 - 2R_1$:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & -1 & 1 & -1 \\ 1 & 1 & 2 & 3 \end{array}\right]$$

$R_3 - R_1$:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & -1 & 1 & -1 \\ 0 & -1 & 1 & -1 \end{array}\right]$$

Step 2: *Make the second pivot equal to 1.*

$-R_2$:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \end{array}\right]$$

Step 3: *Create zeros below the second pivot.*

$R_3 + R_2$:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Step 4: *Create zeros above pivots (back substitution).*

$R_1 - 2R_2$:

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

This is in RREF!

Reading the solution: *The system is now:*

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

Variable z is free (no pivot in its column). Let $z = t$ where $t \in \mathbb{R}$.

Then:

$$x = 2 - 3t$$
$$y = 1 + t$$
$$z = t$$

Answer: $(x, y, z) = (2 - 3t, 1 + t, t)$ *for any $t \in \mathbb{R}$ (infinitely many solutions)*

2.4.4 Leading Variables and Free Variables

Definition: Leading and Free Variables

- A **leading variable** (or **basic variable**) corresponds to a column with a pivot
- A **free variable** corresponds to a column without a pivot
- Free variables can take any value; leading variables are determined by the free variables

Insight: Counting Solutions

For a system in RREF with n variables:

- If there are n pivots: exactly one solution
- If there are fewer than n pivots and no contradiction: infinitely many solutions
- If there's a pivot in the augmented column (row like $[0\ 0\ \cdots\ 0 \mid 1]$): no solution

Example 2.4.4. *Identifying Free Variables Given the RREF matrix:*

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Identify leading and free variables.

Solution: *Pivots are in columns 1, 3, and 4.*

- *Leading variables:* x_1, x_3, x_4
- *Free variables:* x_2, x_5

The solution is:

$$
\begin{aligned}
x_1 &= 4 - 2x_2 - 3x_5 \\
x_3 &= 2 - x_5 \\
x_4 &= 5 + 2x_5 \\
x_2 &= s \quad (\textit{free}) \\
x_5 &= t \quad (\textit{free})
\end{aligned}
$$

where $s, t \in \mathbb{R}$ *can be any real numbers.*

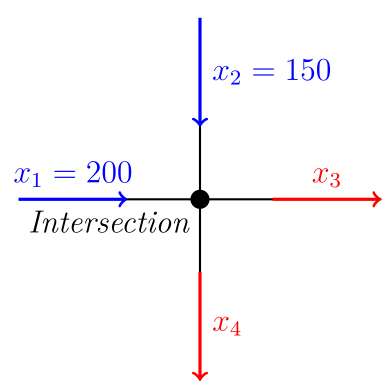
2.5 Applications of Linear Systems

Linear systems appear throughout science, engineering, and everyday life. Let's explore some applications.

2.5.1 Network Flow Problems

In network flow problems, we track how quantities (traffic, water, data) move through a network.

Example 2.5.1. *Traffic Flow Consider the intersection shown below, where the arrows indicate traffic flow (vehicles per hour):*



If traffic flow is conserved (cars in = cars out), find x_3 and x_4.

Solution: *By conservation of flow:*

$$x_1 + x_2 = x_3 + x_4$$

$$200 + 150 = x_3 + x_4$$
$$x_3 + x_4 = 350$$

Without additional information, we have infinitely many solutions. For example:

- *If $x_3 = 200$, then $x_4 = 150$*
- *If $x_3 = 180$, then $x_4 = 170$*

In general: $x_3 = t$ and $x_4 = 350 - t$ where $0 \leq t \leq 350$.

2.5.2 Chemical Equation Balancing

Balancing chemical equations is really solving a system of linear equations!

Example 2.5.2. *Balancing a Chemical Equation Balance the equation for the combustion of propane:*

$$C_3H_8 + O_2 \to CO_2 + H_2O$$

Solution: *Let x_1, x_2, x_3, x_4 be the coefficients:*

$$x_1C_3H_8 + x_2O_2 \to x_3CO_2 + x_4H_2O$$

Balance each element:

- *Carbon: $3x_1 = x_3$*
- *Hydrogen: $8x_1 = 2x_4$*
- *Oxygen: $2x_2 = 2x_3 + x_4$*

From the first equation: $x_3 = 3x_1$

From the second equation: $x_4 = 4x_1$

Substitute into the third:

$$2x_2 = 2(3x_1) + 4x_1$$
$$2x_2 = 10x_1$$
$$x_2 = 5x_1$$

Let $x_1 = 1$ (we want the smallest whole numbers):

$$x_1 = 1$$
$$x_2 = 5$$
$$x_3 = 3$$
$$x_4 = 4$$

Balanced equation: $C_3H_8 + 5O_2 \to 3CO_2 + 4H_2O$

2.5.3 Economics: Supply and Demand

Example 2.5.3. *Market Equilibrium The supply and demand for a product are given by:*

$$\begin{aligned} \textit{Supply:} \quad & S = 2p - 100 \\ \textit{Demand:} \quad & D = -3p + 400 \end{aligned}$$

where p is the price in dollars. Find the equilibrium price and quantity.

Solution: *At equilibrium, supply equals demand:*

$$\begin{aligned} 2p - 100 &= -3p + 400 \\ 5p &= 500 \\ p &= 100 \end{aligned}$$

The equilibrium quantity is:

$$S = 2(100) - 100 = 100 \text{ \textit{units}}$$

Answer: *Equilibrium price: \$100, Equilibrium quantity: 100 units*

2.5.4 Circuit Analysis

Kirchhoff's laws in electrical engineering lead to systems of linear equations.

Example 2.5.4. *Simple Circuit Consider a circuit with three resistors. Using Kirchhoff's voltage law (KVL) and current law (KCL), we might get:*

$$\begin{aligned} I_1 + I_2 &= I_3 \quad \textit{(current conservation)} \\ 5I_1 + 10I_2 &= 12 \quad \textit{(voltage loop 1)} \\ 10I_2 + 15I_3 &= 18 \quad \textit{(voltage loop 2)} \end{aligned}$$

where I_1, I_2, I_3 are currents in amperes.

Set up the augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 5 & 10 & 0 & 12 \\ 0 & 10 & 15 & 18 \end{array}\right]$$

(We would then use Gaussian elimination to solve for the currents.)

Real-World Connection

Modern circuit simulators (like SPICE) solve systems with thousands or millions of equations to analyze complex electronic circuits. The same Gaussian elimination algorithm you're learning scales up to handle these massive systems!

2.6 Chapter Summary

In this chapter, we developed powerful tools for solving systems of linear equations:

- **Linear equations** involve variables to the first power with no products
- **Systems** can have one solution, no solution, or infinitely many solutions
- **Gaussian elimination** uses row operations to solve systems systematically
- **Matrix representation** provides compact notation for large systems
- **Row echelon form (REF)** and **reduced row echelon form (RREF)** are standard forms that make solutions apparent
- **Free variables** parameterize infinite solution sets
- Applications include network flows, chemical reactions, economics, and circuit analysis

Looking Ahead

In Chapter 3, we'll dive deeper into matrices themselves—how to add, multiply, and manipulate them. Matrices are not just bookkeeping tools; they're powerful mathematical objects with rich algebraic structure. We'll see how matrix operations connect to linear transformations and geometric ideas.

2.7 Practice Problems

2.7.1 Basic Problems

1. Determine whether each equation is linear:

   (a) $2x - 3y + z = 7$

   (b) $x^2 + y = 4$

   (c) $\frac{x}{2} - \frac{y}{3} = 1$

   (d) $xy + z = 2$

2. Verify that $(x, y) = (3, -1)$ is a solution to:

$$\begin{aligned} 2x + y &= 5 \\ x - 3y &= 6 \end{aligned}$$

3. Solve by substitution:

$$\begin{aligned} x + y &= 7 \\ 2x - y &= 2 \end{aligned}$$

4. Write the augmented matrix for:

$$\begin{aligned} x - 2y + 3z &= 5 \\ 2x + y - z &= 1 \\ -x + 3y + 2z &= 4 \end{aligned}$$

5. Perform the row operation $R_2 - 3R_1$ on:

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 3 & 5 & 7 \end{array}\right]$$

6. Determine if each matrix is in row echelon form:

   (a) $\begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \end{bmatrix}$

7. Solve using Gaussian elimination:

$$\begin{aligned} x + y &= 3 \\ x - y &= 1 \end{aligned}$$

8. Solve:

$$\begin{aligned} x + 2y + z &= 6 \\ 2x + 3y + 2z &= 10 \\ x + y + z &= 4 \end{aligned}$$

2.7.2 Intermediate Problems

9. Transform to RREF and solve:

$$\left[\begin{array}{ccc|c} 2 & 4 & -2 & 6 \\ 1 & 3 & 1 & 5 \\ 3 & 5 & -1 & 8 \end{array}\right]$$

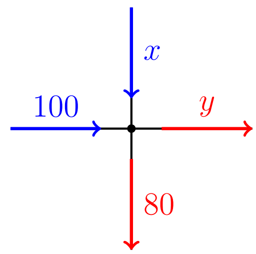
10. Determine whether the system is consistent and, if so, find all solutions:

$$\begin{aligned} x + 2y - z &= 3 \\ 2x + 4y - 2z &= 7 \\ 3x + 6y - 3z &= 9 \end{aligned}$$

11. Find all solutions (express free variables as parameters):

$$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 &= 2 \\ 2x_1 + 4x_2 - x_3 - 2x_4 &= -1 \\ -x_1 - 2x_2 + 2x_3 + 5x_4 &= 6 \end{aligned}$$

12. A traffic intersection has flows as shown. Find the relationship between x and y.



13. Balance the chemical equation:

$$Fe + O_2 \rightarrow Fe_2O_3$$

14. Find the equilibrium price and quantity if:

$$\text{Supply:} \quad S = 3p - 50$$
$$\text{Demand:} \quad D = -2p + 200$$

15. Solve the system and interpret geometrically:

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

2.7.3 Challenge Problems

16. Prove that if a system of linear equations has more than one solution, it must have infinitely many solutions.

17. Find all values of k for which the system has: (a) no solution, (b) exactly one solution, (c) infinitely many solutions.

$$x + 2y = 3$$
$$2x + 4y = k$$

18. For what value(s) of h does the system have infinitely many solutions?

$$x - 3y = 2$$
$$-2x + 6y = h$$

19. Solve the system with four variables:

$$\begin{aligned} x_1 + 2x_2 + x_3 - x_4 &= 3 \\ 2x_1 + 4x_2 + 3x_3 + x_4 &= 7 \\ x_1 + 2x_2 + 2x_3 + 2x_4 &= 4 \end{aligned}$$

20. A company produces three products A, B, and C. Each unit of A requires 2 hours of labor and 1 unit of raw material. Each unit of B requires 1 hour of labor and 2 units of raw material. Each unit of C requires 3 hours of labor and 2 units of raw material. If 100 hours of labor and 80 units of raw material are available, and the company wants to use all resources, set up and solve the system to find all possible production combinations.

21. Show that if the augmented matrix $[A|b]$ can be transformed to RREF $[I|c]$ where I is the identity matrix, then the system has the unique solution $x = c$.

3

Matrices

"'Mathematics is the art of giving the same name to different things."

— *Henri Poincaré*

Chapter Overview

In Chapter 2, we used matrices as convenient bookkeeping tools for solving systems of linear equations. But matrices are far more than notation—they're powerful mathematical objects in their own right. In this chapter, we'll explore the algebra of matrices: how to add them, multiply them, and work with special types of matrices. We'll discover that matrix multiplication, though initially strange, perfectly captures the composition of linear transformations. From computer graphics to cryptography, from quantum mechanics to Google's PageRank algorithm, matrices are the language in which much of modern science and technology is written.

3.1 Introduction to Matrices

3.1.1 Matrix Notation and Terminology

Definition: Matrix

A **matrix** is a rectangular array of numbers arranged in rows and columns. A matrix with m rows and n columns is called an $m \times n$ **matrix** (read "m by n").

General form:

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

The number a_{ij} is the **entry** (or **element**) in row i and column j.

Important terminology:

- The **size** (or **dimension**) of the matrix is $m \times n$
- A matrix with equal numbers of rows and columns ($m = n$) is called a **square matrix**
- The entries $a_{11}, a_{22}, a_{33}, \ldots$ form the **main diagonal**
- We often denote matrices by uppercase letters: $A, B, C, \ldots$
- The (i, j)-entry of matrix A can be written as a_{ij} or $(A)_{ij}$ or A_{ij}

Example 3.1.1. *Matrix Notation Consider the matrix:*

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

This is a 2×3 *matrix with:*

- $a_{11} = 2,\ a_{12} = -1,\ a_{13} = 3$
- $a_{21} = 0,\ a_{22} = 5,\ a_{23} = -2$

3.1.2 Types of Matrices

Special Types of Matrices

- **Row vector**: A $1 \times n$ matrix, e.g., $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$
- **Column vector**: An $m \times 1$ matrix, e.g., $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$
- **Zero matrix**: All entries are zero, denoted O or 0
- **Square matrix**: Number of rows equals number of columns ($n \times n$)
- **Diagonal matrix**: Square matrix with all off-diagonal entries equal to zero
- **Identity matrix**: Diagonal matrix with all diagonal entries equal to 1, denoted I_n or just I

Example 3.1.2. *Special Matrices Examples of special matrices:*

Zero matrix (2×3):

$$O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Diagonal matrix:

$$D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Identity matrix (3×3):

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Insight: Why Identity Matrix Matters

The identity matrix I plays the same role in matrix algebra that the number 1 plays in regular arithmetic: multiplying any matrix by I leaves it unchanged. This makes I the multiplicative identity for matrices.

3.1.3 Matrix Equality

Definition: Matrix Equality

Two matrices A and B are **equal** if:

1. They have the same dimensions
2. All corresponding entries are equal: $a_{ij} = b_{ij}$ for all i, j

Example 3.1.3. *Testing Matrix Equality Are these matrices equal?*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

Solution: $A = B$ *(same dimensions, all entries match)*

$A \neq C$ *(even though they contain the same numbers, the arrangement is different)*

3.1.4 Real-World Contexts for Matrices

Matrices naturally represent many types of data:

1. Data tables

$$\text{Test Scores} = \begin{array}{c} \begin{bmatrix} 85 & 92 & 78 \\ 90 & 88 & 95 \\ 76 & 84 & 88 \end{bmatrix} \end{array} \begin{array}{c} \text{Alice} \\ \text{Bob} \\ \text{Carol} \end{array}$$

Columns: Math, English, Science

2. Images

Digital images are matrices where each entry represents a pixel's brightness or color value.

3. Networks and graphs

An **adjacency matrix** represents connections in a network:

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

where $a_{ij} = 1$ if node i connects to node j, and $a_{ij} = 0$ otherwise.

4. Transformations

Matrices represent geometric transformations like rotations, reflections, and scaling (more on this in Chapter 6!).

3.2 Matrix Operations

Just as we can add and multiply numbers, we can add and multiply matrices—though with some important differences!

3.2.1 Matrix Addition and Subtraction

Definition: Matrix Addition

If A and B are both $m \times n$ matrices, their **sum** $A + B$ is the $m \times n$ matrix obtained by adding corresponding entries:

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

Important Restriction

You can only add or subtract matrices of the **same dimensions**. Adding a 2×3 matrix to a 3×2 matrix is undefined!

Example 3.2.1. *Matrix Addition Let* $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *and* $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. *Find* $A + B$.

Solution:

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

Matrix subtraction works the same way:

$$(A-B)_{ij}=a_{ij}-b_{ij}$$

3.2.2 Scalar Multiplication

Definition: Scalar Multiplication

If A is an $m\times n$ matrix and c is a scalar (number), then cA is the $m\times n$ matrix obtained by multiplying every entry of A by c:

$$(cA)_{ij}=c\cdot a_{ij}$$

Example 3.2.2. *Scalar Multiplication Let* $A=\begin{bmatrix}2&-1\\3&0\end{bmatrix}$. *Find* $3A$ *and* $-\frac{1}{2}A$.

Solution:

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

$$-\frac{1}{2}A=-\frac{1}{2}\begin{bmatrix}2&-1\\3&0\end{bmatrix}=\begin{bmatrix}-1&\frac{1}{2}\\-\frac{3}{2}&0\end{bmatrix}$$

3.2.3 Properties of Matrix Addition and Scalar Multiplication

These operations satisfy familiar algebraic properties:

Properties

For matrices A, B, C (of the same size) and scalars c, d:

1. **Commutativity**: $A+B=B+A$
2. **Associativity**: $(A+B)+C=A+(B+C)$
3. **Additive identity**: $A+O=A$ (where O is the zero matrix)
4. **Additive inverse**: $A+(-A)=O$
5. **Distributivity**: $c(A+B)=cA+cB$

6. **Distributivity**: $(c+d)A = cA + dA$
7. **Associativity**: $c(dA) = (cd)A$
8. **Identity**: $1A = A$

Connection to Vectors

Notice that these properties are identical to the properties of vector addition and scalar multiplication from Chapter 1! In fact, an $m \times n$ matrix can be viewed as a vector in $\mathbb{R}^{mn}$. This deep connection is why we study vector spaces in Chapter 5.

3.2.4 Matrix Multiplication

Matrix multiplication is more complex than addition. The definition might seem strange at first, but it's perfectly designed for composing linear transformations.

Definition: Matrix Multiplication

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. The **product** AB is the $m \times p$ matrix where:

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

In words: The (i, j)-entry of AB is the dot product of row i of A with column j of B.

Critical Dimension Requirement

To multiply $A \times B$:

- The number of **columns** of A must equal the number of **rows** of B
- If A is $m \times n$ and B is $n \times p$, then AB is $m \times p$

Mnemonic: $(m \times \mathbf{n})(\mathbf{n} \times p) = (m \times p)$ — the inner dimensions must match!

Example 3.2.3. *Matrix Multiplication (Small Example) Let* $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *and* $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. *Find* AB.

Solution:

Both are 2×2*, so the product is defined and will be* 2×2.

Entry (1,1): *Row 1 of A · Column 1 of B*

$$(1)(5) + (2)(7) = 5 + 14 = 19$$

Entry (1,2): *Row 1 of A · Column 2 of B*

$$(1)(6) + (2)(8) = 6 + 16 = 22$$

Entry (2,1): *Row 2 of A · Column 1 of B*

$$(3)(5) + (4)(7) = 15 + 28 = 43$$

Entry (2,2): *Row 2 of A · Column 2 of B*

$$(3)(6) + (4)(8) = 18 + 32 = 50$$

Therefore:

$$AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

Example 3.2.4. *Matrix Multiplication with Different Dimensions Let* $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ *(*2×3*) and* $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ *(*3×2*). Find* AB*.*

Solution:

A is 2×3 and B is 3×2, so AB will be 2×2.

Entry (1,1): $[1\ 2\ 3] \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1(1) + 2(0) + 3(1) = 4$

Entry (1,2): $[1\ 2\ 3] \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 1(0) + 2(1) + 3(1) = 5$

Entry (2,1): $[4\ 5\ 6] \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 4(1) + 5(0) + 6(1) = 10$

Entry (2,2): $[4\ 5\ 6] \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 4(0) + 5(1) + 6(1) = 11$

$$AB = \begin{bmatrix} 4 & 5 \\ 10 & 11 \end{bmatrix}$$

Visual Understanding

Think of matrix multiplication as a systematic way to combine rows of the first matrix with columns of the second. Each entry in the result summarizes how one row interacts with one column through the dot product.

3.2.5 Why Matrix Multiplication Works This Way

The definition of matrix multiplication isn't arbitrary—it perfectly captures composition of linear transformations!

Consider composing two transformations in $\mathbb{R}^2$:

- First, transform $\vec{x}$ by A: $\vec{y} = A\vec{x}$
- Then, transform $\vec{y}$ by B: $\vec{z} = B\vec{y}$

Combining these: $\vec{z} = B(A\vec{x}) = (BA)\vec{x}$

The matrix BA represents the composition of the two transformations. The way we defined multiplication ensures this works correctly! We'll explore this more deeply in Chapter 6.

3.2.6 Properties of Matrix Multiplication

Matrix multiplication has some familiar properties—and some surprises!

Properties of Matrix Multiplication

For appropriately sized matrices A, B, C and scalar c:

1. **Associativity**: $(AB)C = A(BC)$
2. **Left distributivity**: $A(B + C) = AB + AC$
3. **Right distributivity**: $(A + B)C = AC + BC$
4. **Scalar compatibility**: $c(AB) = (cA)B = A(cB)$
5. **Identity**: $IA = A$ and $AI = A$ (where I is the identity matrix)

Important: Matrix Multiplication is NOT Commutative!

In general, $AB \neq BA$. In fact:

- Sometimes AB is defined but BA is not (different dimensions)
- Even when both are defined, they usually give different results

- Sometimes $AB = BA$, but this is special, not typical

Example 3.2.5. *Non-Commutativity Let* $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$. *Compare* AB *and* BA.

Solution:

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$$

Clearly $AB \neq BA$*!*

Example 3.2.6. *When Order Matters in Real Life Consider transformations: "rotate by 90°" then "reflect across x-axis" gives a different result than doing them in the opposite order. The non-commutativity of matrix multiplication captures this reality!*

3.2.7 Powers of Matrices

For square matrices, we can define powers:

Definition: Matrix Powers

For a square matrix A and positive integer n:

$$A^n = \underbrace{A \cdot A \cdot A \cdots A}_{n \text{ times}}$$

By convention: $A^0 = I$ (the identity matrix)

Example 3.2.7. *Computing Matrix Powers Let* $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$. *Find* A^2 *and* A^3.

Solution:

$$A^2 = AA = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 0 & 1 \end{bmatrix}$$

$$A^3 = A^2 \cdot A = \begin{bmatrix} 4 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 7 \\ 0 & 1 \end{bmatrix}$$

3.3 Matrix Transpose

The transpose operation flips a matrix across its main diagonal.

3.3.1 Definition and Properties

Definition: Transpose

The **transpose** of an $m \times n$ matrix A, denoted A^T, is the $n \times m$ matrix obtained by interchanging rows and columns:

$$(A^T)_{ij} = A_{ji}$$

In other words, row i of A^T is column i of A.

Example 3.3.1. *Computing a Transpose Find the transpose of* $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$.

Solution:

A is 2×3, so A^T is 3×2:

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

Notice that:

- *Row 1 of A became column 1 of A^T*
- *Row 2 of A became column 2 of A^T*

Properties of Transpose

For matrices A and B (appropriately sized) and scalar c:

1. $(A^T)^T = A$ (transpose of transpose is original)
2. $(A + B)^T = A^T + B^T$
3. $(cA)^T = cA^T$
4. $(AB)^T = B^T A^T$ (order reverses!)

Common Mistake: Transpose of Product

The transpose of a product reverses the order: $(AB)^T = B^T A^T$, NOT $A^T B^T$. This is because of how row-column multiplication works.

Example 3.3.2. *Verifying* $(AB)^T = B^T A^T$ *Let* $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. *Verify that* $(AB)^T = B^T A^T$.

Solution:

First compute AB*:*

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$$

So:

$$(AB)^T = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$$

Now compute $B^T A^T$*:*

$$B^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$B^T A^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}$$

Indeed, $(AB)^T = B^T A^T$ ✓

3.3.2 Symmetric Matrices

Definition: Symmetric Matrix

A square matrix A is **symmetric** if $A^T = A$.
Equivalently: $a_{ij} = a_{ji}$ for all i, j (the matrix is symmetric across its main diagonal).

Example 3.3.3. *Symmetric Matrix The matrix* $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ *is symmetric.*

Notice: $a_{12} = a_{21} = 2$, $a_{13} = a_{31} = 3$, $a_{23} = a_{32} = 5$.

Symmetric matrices have special properties that make them particularly important in applications:

- All eigenvalues are real (we'll see this in Chapter 7)
- They can always be orthogonally diagonalized
- They naturally arise in physics (inertia tensors, covariance matrices)

3.4 Special Matrices

Certain types of matrices appear frequently and have special properties.

3.4.1 Identity Matrix (Revisited)

The identity matrix acts like the number 1:

Properties of Identity Matrix

For an $n \times n$ identity matrix I_n:

- $I_n A = A$ for any $n \times k$ matrix A
- $AI_n = A$ for any $k \times n$ matrix A
- $I_n^T = I_n$ (identity is symmetric)
- $I_n^k = I_n$ for any positive integer k

3.4.2 Diagonal Matrices

Definition: Diagonal Matrix

A square matrix D is **diagonal** if all off-diagonal entries are zero:

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

Notation: $D = \text{diag}(d_1, d_2, \dots, d_n)$

Diagonal matrices are easy to work with:

- Multiplying diagonal matrices: just multiply corresponding diagonal entries
- Powers are easy: $D^k = \text{diag}(d_1^k, d_2^k, \dots, d_n^k)$
- Transpose leaves them unchanged: $D^T = D$ (all diagonal matrices are symmetric)

Example 3.4.1. *Working with Diagonal Matrices Let* $D_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ *and* $D_2 =$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Find D_1D_2 and D_1^3.

Solution:

$$D_1D_2 = \begin{bmatrix} 2\cdot 1 & 0 & 0 \\ 0 & 3\cdot 2 & 0 \\ 0 & 0 & (-1)\cdot 4 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

$$D_1^3 = \begin{bmatrix} 2^3 & 0 & 0 \\ 0 & 3^3 & 0 \\ 0 & 0 & (-1)^3 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 27 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

3.4.3 Upper and Lower Triangular Matrices

Definition: Triangular Matrices

- An **upper triangular** matrix has all entries below the main diagonal equal to zero
- A **lower triangular** matrix has all entries above the main diagonal equal to zero

$$\text{Upper: } U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}, \quad \text{Lower: } L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$$

Triangular matrices are useful because:

- Systems with triangular coefficient matrices are easy to solve (back substitution)
- The product of upper (lower) triangular matrices is upper (lower) triangular
- The transpose of an upper triangular matrix is lower triangular and vice versa

3.4.4 Introduction to Matrix Inverse

Just as division is the inverse of multiplication for numbers, we have matrix inverses—but not every matrix has one!

Definition: Matrix Inverse

An $n \times n$ matrix A is **invertible** (or **nonsingular**) if there exists an $n \times n$ matrix B such that:

$$AB = BA = I_n$$

The matrix B is called the **inverse** of A, denoted A^{-1}.
If no such matrix exists, A is called **singular** or **noninvertible**.

Key Facts About Inverses

- Only square matrices can have inverses
- If A is invertible, its inverse is unique
- If A has an inverse, then $AA^{-1} = A^{-1}A = I$
- Not all square matrices are invertible

Example 3.4.2. *Verifying an Inverse Verify that* $B = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$ *is the inverse of* $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$.

Solution:

Compute AB:

$$AB = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 4-3 & -2+2 \\ 6-6 & -3+4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Compute BA:

$$BA = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 4-3 & 2-2 \\ -6+6 & -3+4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Since $AB = BA = I$, $B = A^{-1}$ ✓

3.5 Matrix Inverse

The matrix inverse is one of the most important concepts in linear algebra.

3.5.1 Computing 2×2 Inverses

For 2×2 matrices, there's a simple formula:

Formula: Inverse of a 2×2 Matrix

For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, if $\det(A) = ad - bc \neq 0$, then:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, the matrix is not invertible.

Example 3.5.1. *Finding a 2×2 Inverse Find the inverse of $A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$.*

Solution:

First, check if it's invertible:

$$\det(A) = (3)(2) - (1)(5) = 6 - 5 = 1 \neq 0$$

So A is invertible. Apply the formula:

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

Verify: $AA^{-1} = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ✓

3.5.2 Computing Inverses Using Row Operations

For larger matrices, we use row operations. The method: row-reduce $[A|I]$ to $[I|A^{-1}]$.

Algorithm: Finding the Inverse Using Row Reduction

To find A^{-1}:

1. Form the augmented matrix $[A|I]$
2. Use row operations to transform it to $[I|B]$
3. If successful, $B = A^{-1}$
4. If you get a row of zeros on the left side, A is not invertible

Example 3.5.2. *Finding Inverse by Row Reduction Find the inverse of $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.*

Solution:

Form $[A|I]$:

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

$R_3 - R_1$:

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

$R_3 + 2R_2$:

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

$\frac{1}{3}R_3$:

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

$R_2 - R_3$:

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

$R_1 - 2R_2$:

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

Therefore:

$$A^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

3.5.3 Properties of Matrix Inverses

Properties of Inverses

For invertible matrices A and B:

1. $(A^{-1})^{-1} = A$ (inverse of inverse is original)
2. $(AB)^{-1} = B^{-1}A^{-1}$ (order reverses, like transpose!)
3. $(A^T)^{-1} = (A^{-1})^T$
4. $(cA)^{-1} = \frac{1}{c}A^{-1}$ for $c \neq 0$
5. If A is invertible, so is A^T

3.5.4 When Does an Inverse Exist?

Not all square matrices have inverses. We'll explore this more in Chapter 4, but here are some key facts:

Conditions for Invertibility

An $n \times n$ matrix A is invertible if and only if:

- The determinant of A is nonzero (Chapter 4)
- The columns of A are linearly independent (Chapter 5)
- The rows of A are linearly independent
- A can be row-reduced to the identity matrix
- The equation $A\vec{x} = \vec{0}$ has only the trivial solution $\vec{x} = \vec{0}$
- The equation $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b}$

All these conditions are equivalent!

3.5.5 Solving Systems Using Inverses

If A is invertible, we can solve $A\vec{x} = \vec{b}$ easily:

Solving a Linear System Using the Inverse

If A is invertible:

$$A\vec{x} = \vec{b} \implies A^{-1}(A\vec{x}) = A^{-1}\vec{b} \implies \vec{x} = A^{-1}\vec{b}$$

Example 3.5.3. *Solving a System Using Inverse Solve the system using the inverse from our previous example:*

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

Solution:

The system is $A\vec{x} = \vec{b}$ *where* $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ *and* $\vec{b} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}$.

We found A^{-1} *earlier, so:*

$$\vec{x} = A^{-1}\vec{b} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{5}{3} - 2 + \frac{8}{3} \\ \frac{5}{3} + 1 - \frac{4}{3} \\ -\frac{5}{3} + 2 + \frac{4}{3} \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Answer: $(x, y, z) = (3, 2, 1)$

Important Computational Note

While solving systems using A^{-1} is theoretically elegant, it's actually less efficient than Gaussian elimination for large systems. Computing A^{-1} takes more operations than just row-reducing $[A|\vec{b}]$. Use inverses for theory and small systems; use elimination for large computations.

3.6 Applications of Matrices

Matrices are everywhere in modern science and technology. Let's explore some applications.

3.6.1 Image Processing and Filters

Digital images are matrices of pixel values. We can apply filters by matrix operations.

Example 3.6.1. *Image Brightness Adjustment To increase brightness by 20%, multiply the image matrix by 1.2:*

$$Image_{bright} = 1.2 \cdot Image_{original}$$

To invert an image (like a photo negative):

$$Image_{inverted} = 255 \cdot \mathbf{1} - Image_{original}$$

where **1** *is a matrix of all ones.*

3.6.2 Graph Theory and Adjacency Matrices

Networks (social networks, web pages, transportation systems) are represented by adjacency matrices.

Example 3.6.2. *Powers of Adjacency Matrices Consider a graph with adjacency matrix:*

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

The entry $(A^2)_{ij}$ counts the number of paths of length 2 from node i to node j!

$$A^2 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

For example, $(A^2)_{13} = 1$ means there's 1 path of length 2 from node 1 to node 3.

3.6.3 Markov Chains and Probability

Markov chains model systems that transition between states.

Example 3.6.3. *Weather Model Suppose tomorrow's weather depends only on today's weather:*

- *If sunny today: 70% chance sunny tomorrow, 30% chance rainy*

- *If rainy today: 40% chance sunny tomorrow, 60% chance rainy*

*The **transition matrix** is:*

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \begin{array}{l} \textit{from sunny} \\ \textit{from rainy} \end{array}$$

If today's state is $\vec{v}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *(sunny), then:*

- *Tomorrow:* $\vec{v}_1 = P\vec{v}_0 = \begin{bmatrix} 0.7 \\ 0.4 \end{bmatrix}$
- *Day after:* $\vec{v}_2 = P\vec{v}_1 = P^2\vec{v}_0$
- *After* n *days:* $\vec{v}_n = P^n\vec{v}_0$

3.6.4 Leslie Matrices in Population Models

Leslie matrices model age-structured populations.

Example 3.6.4. *Population Dynamics Consider a population divided into three age groups. The Leslie matrix:*

$$L = \begin{bmatrix} 0 & 1.5 & 0.8 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0 \end{bmatrix}$$

represents:

- *Row 1: birth rates from each age group*
- *Below diagonal: survival rates to next age group*

If the current population is $\vec{p}_0 = \begin{bmatrix} 100 \\ 80 \\ 40 \end{bmatrix}$*, next generation is:*

$$\vec{p}_1 = L\vec{p}_0 = \begin{bmatrix} 0 & 1.5 & 0.8 \\ 0.7 & 0 & 0 \\ 0 & 0.6 & 0 \end{bmatrix} \begin{bmatrix} 100 \\ 80 \\ 40 \end{bmatrix} = \begin{bmatrix} 152 \\ 70 \\ 48 \end{bmatrix}$$

3.6.5 Cryptography

Matrices can encode messages for secure communication.

Example 3.6.5. *Hill Cipher (Simplified) To encode a message, represent letters as numbers (A=0, B=1, ..., Z=25).*

Use an encoding matrix, say $E = \begin{bmatrix} 3 & 2 \\ 5 & 7 \end{bmatrix}$.

To encode "HI" (H=7, I=8):

$$\begin{bmatrix} 3 & 2 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 37 \\ 91 \end{bmatrix} \equiv \begin{bmatrix} 11 \\ 13 \end{bmatrix} \pmod{26}$$

Encrypted: "LN"

To decode, use E^{-1} *(computed modulo 26).*

3.7 Chapter Summary

In this chapter, we've explored matrices as mathematical objects:

- **Matrix addition** and **scalar multiplication** work entry-by-entry
- **Matrix multiplication** uses row-column dot products and is NOT commutative
- The **transpose** A^T flips a matrix across its main diagonal
- **Special matrices**: identity, diagonal, triangular, symmetric
- **Matrix inverse** A^{-1} satisfies $AA^{-1} = A^{-1}A = I$
- Not all square matrices have inverses; invertibility has many equivalent conditions
- Applications span image processing, networks, probability, population models, and cryptography

Looking Ahead

In Chapter 4, we'll introduce the determinant—a single number that encodes crucial information about a matrix. The determinant determines invertibility, gives geometric meaning (area/volume scaling), and is essential for eigenvalue calculations in Chapter 7.

3.8 Practice Problems

3.8.1 Basic Problems

1. Given $A = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$, compute:

   (a) $A + B$

   (b) $3A$

   (c) $2A - B$

2. Compute the products AB and BA where:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$$

   Are they equal?

3. Find the transpose of:

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

4. Determine if the following matrix is symmetric:

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

5. Find the inverse of $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ using the 2×2 formula.

6. Verify that $B = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$ is the inverse of $A = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$.

7. Compute A^2 and A^3 where $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

8. Which of the following matrices are invertible? (Use the 2×2 determinant test)

   (a) $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

   (b) $\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$

3.8.2 Intermediate Problems

9. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Verify that $(AB)^T = B^T A^T$.

10. Find the inverse of $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ using row reduction.

11. Solve the system $A\vec{x} = \vec{b}$ using A^{-1} where:

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 8 \\ 5 \end{bmatrix}$$

12. Let $D = \text{diag}(2, 3, -1)$. Compute D^4.

13. If $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, find a formula for A^n for any positive integer n.

14. Prove that if A and B are both invertible $n \times n$ matrices, then $(AB)^{-1} = B^{-1}A^{-1}$.

15. A Markov chain has transition matrix $P = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$. If the initial state is $\vec{v}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, find the state after 2 steps.

16. Show that if A is symmetric, then A^2 is also symmetric.

3.8.3 Challenge Problems

17. Prove that if A is invertible and $AB = AC$, then $B = C$. Does this hold if A is not invertible?

18. Find all 2×2 matrices A such that $A^2 = I$.

19. Prove that the transpose of an upper triangular matrix is lower triangular.

20. Let A be an $n \times n$ matrix. Prove that $A + A^T$ is always symmetric.

21. Find a 3×3 matrix A (other than I) such that $A^2 = A$ (called an idempotent matrix).

22. Prove that if A is an invertible matrix, then $(A^n)^{-1} = (A^{-1})^n$ for any positive integer n.

23. A matrix A is called **skew-symmetric** if $A^T = -A$.

(a) Show that the diagonal entries of a skew-symmetric matrix must be zero.

(b) Prove that any square matrix A can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

24. Consider the adjacency matrix of a graph A. Prove that $(A^k)_{ij}$ equals the number of paths of length k from vertex i to vertex j.

4

Determinants

"'In mathematics, the art of proposing a question must be held of higher value than solving it."

— *Georg Cantor*

Chapter Overview

Every square matrix has associated with it a single number called its **determinant**. This seemingly simple number encodes remarkable information: whether the matrix is invertible, how it scales areas and volumes, and what its eigenvalues are. The determinant is one of the most elegant concepts in linear algebra, connecting geometry, algebra, and computation. From solving systems of equations (Cramer's rule) to understanding how transformations distort space, determinants provide both theoretical insights and practical tools. In this chapter, we'll explore what determinants are, how to compute them efficiently, and why they matter.

4.1 Introduction to Determinants

4.1.1 Determinants of 2×2 Matrices

We begin with the simplest non-trivial case: 2×2 matrices.

Definition: Determinant of a 2×2 Matrix

For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the **determinant** of A is:

$$\det(A) = |A| = ad - bc$$

We use the notation $\det(A)$ or $|A|$ or $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Example 4.1.1. *Computing* 2×2 *Determinants Compute the determinant of* $A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$.

Solution:

$$\det(A) = (3)(4) - (1)(2) = 12 - 2 = 10$$

4.1.2 Geometric Interpretation in 2D

The determinant has a beautiful geometric meaning!

Geometric Meaning (2D)

For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

$|\det(A)|$ equals the **area** of the parallelogram formed by the column vectors $\begin{bmatrix} a \\ c \end{bmatrix}$ and $\begin{bmatrix} b \\ d \end{bmatrix}$.

The sign of $\det(A)$ indicates orientation:

- $\det(A) > 0$: vectors maintain orientation (counterclockwise)
- $\det(A) < 0$: vectors reverse orientation (clockwise)
- $\det(A) = 0$: vectors are collinear (parallelogram is flat)

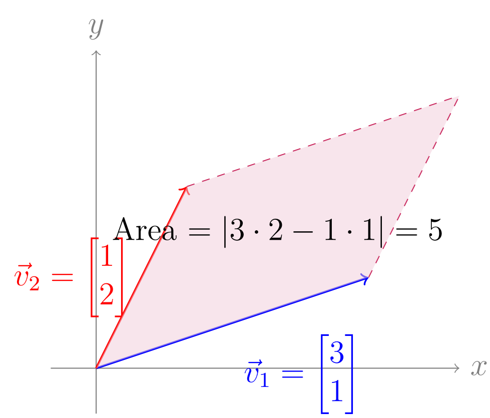


Figure 4.1: Determinant as area of parallelogram

Example 4.1.2. *Area Calculation Find the area of the parallelogram with adjacent sides* $\vec{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Solution:

Form the matrix with these vectors as columns:

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

$$\textit{Area} = |\det(A)| = |(3)(2) - (1)(1)| = |6 - 1| = 5$$

4.1.3 Connection to Invertibility

The determinant immediately tells us whether a matrix is invertible!

Invertibility Test

A square matrix A is invertible if and only if $\det(A) \neq 0$.

Why? If $\det(A) = 0$, the columns are linearly dependent (geometrically: they're collinear or coplanar), so the matrix maps all of space onto a lower-dimensional subspace. This transformation cannot be reversed.

Example 4.1.3. *Testing Invertibility Determine if* $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ *is invertible.*

Solution:

$$\det(A) = (2)(2) - (4)(1) = 4 - 4 = 0$$

Since $\det(A) = 0$*, the matrix is* ***not invertible*** *(singular).*

Notice: the second column is exactly twice the first column, confirming linear dependence.

4.2 Computing Determinants

For matrices larger than 2×2, we need systematic methods.

4.2.1 Determinant of 3×3 Matrices

Formula: 3×3 Determinant

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

$$\det(A) = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

This is called **cofactor expansion along the first row**.

The pattern: multiply each entry in the first row by the determinant of the 2×2 matrix that remains after deleting that entry's row and column, alternating signs.

Example 4.2.1. *Computing a* 3×3 *Determinant Compute* $\det(A)$ *where* $A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 1 \\ 5 & 2 & 1 \end{bmatrix}$.

Solution:

Expand along the first row:

$$\det(A) = 2\begin{vmatrix} 4 & 1 \\ 2 & 1 \end{vmatrix} - 1\begin{vmatrix} 0 & 1 \\ 5 & 1 \end{vmatrix} + 3\begin{vmatrix} 0 & 4 \\ 5 & 2 \end{vmatrix}$$

Compute each 2×2 *determinant:*

$$\begin{vmatrix} 4 & 1 \\ 2 & 1 \end{vmatrix} = (4)(1) - (1)(2) = 2$$

$$\begin{vmatrix} 0 & 1 \\ 5 & 1 \end{vmatrix} = (0)(1) - (1)(5) = -5$$

$$\begin{vmatrix} 0 & 4 \\ 5 & 2 \end{vmatrix} = (0)(2) - (4)(5) = -20$$

Therefore:

$$\det(A) = 2(2) - 1(-5) + 3(-20) = 4 + 5 - 60 = -51$$

4.2.2 Cofactor Expansion (Laplace Expansion)

We can expand along *any* row or column, not just the first row!

Definition: Minor and Cofactor

For an $n \times n$ matrix A:
Minor M_{ij}: The determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j.
Cofactor C_{ij}: The signed minor: $C_{ij} = (-1)^{i+j} M_{ij}$

The sign pattern for cofactors follows a checkerboard:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Cofactor Expansion Formula

The determinant of A can be computed by expanding along row i:

$$\det(A) = \sum_{j=1}^{n} a_{ij}C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

Or by expanding along column j:

$$\det(A) = \sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

Strategy: Choose Wisely!

Expand along the row or column with the most zeros to minimize computation. Each zero entry contributes nothing to the sum!

Example 4.2.2. *Strategic Cofactor Expansion Compute* $\det(A)$ *where* $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 4 \\ 5 & 6 & 7 \end{bmatrix}$.

Solution:

Notice column 2 has two zeros! Expand along column 2:

$$\det(A) = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32}$$

Since $a_{12} = 0$ *and* $a_{22} = 0$*:*

$$\det(A) = 0 \cdot C_{12} + 0 \cdot C_{22} + 6 \cdot C_{32} = 6C_{32}$$

Compute $C_{32} = (-1)^{3+2}M_{32}$*:*

$$M_{32} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = (1)(4) - (2)(3) = -2$$

$$C_{32} = (-1)^5(-2) = -(-2) = 2$$

Therefore:

$$\det(A) = 6(2) = 12$$

Much faster than expanding along the first row!

4.2.3 Determinants of Triangular Matrices

Triangular matrices have a wonderful property:

Determinant of Triangular Matrix

If A is upper triangular or lower triangular, then:

$$\det(A) = a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{nn}$$

The determinant is simply the **product of the diagonal entries**!

Example 4.2.3. *Triangular Matrix Determinant Compute* $\det(A)$ *where* $A = \begin{bmatrix} 2 & 3 & 1 & 5 \\ 0 & -1 & 4 & 2 \\ 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

Solution:

A is upper triangular, so:

$$\det(A) = (2)(-1)(3)(2) = -12$$

That's it! No cofactor expansion needed.

Special Case: Diagonal Matrices

For a diagonal matrix $D = \text{diag}(d_1, d_2, \ldots, d_n)$:

$$\det(D) = d_1 \cdot d_2 \cdots d_n$$

In particular, $\det(I) = 1$ (identity matrix).

4.2.4 Row Operations and Determinants

Row operations affect determinants in predictable ways:

Effect of Row Operations on Determinants

Let A' be the matrix obtained from A by a row operation:

1. **Row interchange**: $\det(A') = -\det(A)$
2. **Row scaling**: If row i is multiplied by c, then $\det(A') = c \cdot \det(A)$
3. **Row replacement**: If a multiple of one row is added to another, then $\det(A') = \det(A)$

This gives us an efficient algorithm for computing determinants!

Algorithm: Determinant by Row Reduction

To compute $\det(A)$:

1. Use row operations to reduce A to row echelon form (or triangular form)
2. Track how each operation affects the determinant
3. The determinant of the triangular form is the product of diagonal entries
4. Adjust for the row operations used

Example 4.2.4. *Determinant by Row Reduction Compute* $\det(A)$ *where* $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 5 & 3 \end{bmatrix}$.

Solution:

Perform $R_2 - 2R_1$*:*

$$A_1 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 3 & 5 & 3 \end{bmatrix}, \quad \det(A_1) = \det(A)$$

Perform $R_3 - 3R_1$*:*

$$A_2 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -1 & -6 \end{bmatrix}, \quad \det(A_2) = \det(A)$$

Perform $R_3 + R_2$*:*

$$A_3 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix}, \quad \det(A_3) = \det(A)$$

Now A_3 is upper triangular:

$$\det(A_3) = (1)(1)(-5) = -5$$

Therefore, $\det(A) = -5$.

4.3 Properties and Applications of Determinants

Determinants satisfy many elegant algebraic properties.

4.3.1 Fundamental Properties

Key Properties of Determinants

For $n \times n$ matrices A and B:

1. **Identity**: $\det(I) = 1$
2. **Transpose**: $\det(A^T) = \det(A)$
3. **Product**: $\det(AB) = \det(A) \cdot \det(B)$
4. **Inverse**: If A is invertible, $\det(A^{-1}) = \frac{1}{\det(A)}$
5. **Scalar multiple**: $\det(cA) = c^n \det(A)$ for $n \times n$ matrix A
6. **Zero row/column**: If A has a row or column of zeros, $\det(A) = 0$
7. **Identical rows/columns**: If two rows (or columns) are identical, $\det(A) = 0$

Example 4.3.1. *Using Properties Let A be a 3×3 matrix with* $\det(A) = 5$*. Find:*

(a) $\det(2A)$

(b) $\det(A^{-1})$

(c) $\det(A^T)$

(d) $\det(A^2)$

Solution:

(a) $\det(2A) = 2^3 \det(A) = 8 \cdot 5 = 40$

(b) $\det(A^{-1}) = \frac{1}{\det(A)} = \frac{1}{5}$

(c) $\det(A^T) = \det(A) = 5$

(d) $\det(A^2) = \det(A \cdot A) = \det(A) \cdot \det(A) = 5 \cdot 5 = 25$

Common Mistake: Sum of Determinants

In general, $\det(A + B) \neq \det(A) + \det(B)$!
The determinant is **not** linear in this sense. Only the product property holds.

4.3.2 Determinant Product Rule: Proof Sketch

Why does $\det(AB) = \det(A) \cdot \det(B)$?

Geometric intuition:

- A scales volumes by factor $|\det(A)|$
- B scales volumes by factor $|\det(B)|$
- Applying both transformations scales volumes by $|\det(A)| \cdot |\det(B)|$
- Therefore, $\det(AB) = \det(A) \cdot \det(B)$

This property is crucial: it means determinants respect composition of linear transformations!

4.3.3 Determinant and Volume in Higher Dimensions

The geometric interpretation extends to higher dimensions:

Geometric Meaning (General)

For an $n \times n$ matrix A:
$|\det(A)|$ equals the n-dimensional **volume** of the parallelepiped formed by the column vectors.

- 2D: area of parallelogram
- 3D: volume of parallelepiped
- 4D and beyond: hypervolume

Example 4.3.2. *Volume in 3D Find the volume of the parallelepiped with edges:*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Solution:

Form the matrix with these vectors as columns:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

This is upper triangular!

$$\det(A) = (1)(2)(3) = 6$$

$$\textit{Volume} = |\det(A)| = 6 \ \textit{cubic units}$$

4.3.4 Cramer's Rule

Determinants give an explicit formula for solving systems—though it's not computationally efficient for large systems.

Cramer's Rule

Consider the system $A\vec{x} = \vec{b}$ where A is $n \times n$ and invertible.
The solution is:

$$x_i = \frac{\det(A_i)}{\det(A)}$$

where A_i is the matrix obtained by replacing column i of A with $\vec{b}$.

Example 4.3.3. *Solving with Cramer's Rule Solve using Cramer's rule:*

$$\begin{aligned} 2x + y &= 5 \\ x + 3y &= 7 \end{aligned}$$

Solution:

The system is $A\vec{x} = \vec{b}$ where:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

Compute $\det(A)$*:*

$$\det(A) = (2)(3) - (1)(1) = 5$$

For x, replace column 1 with $\vec{b}$:

$$A_1 = \begin{bmatrix} 5 & 1 \\ 7 & 3 \end{bmatrix}, \quad \det(A_1) = (5)(3) - (1)(7) = 8$$

$$x = \frac{\det(A_1)}{\det(A)} = \frac{8}{5}$$

For y, replace column 2 with $\vec{b}$:

$$A_2 = \begin{bmatrix} 2 & 5 \\ 1 & 7 \end{bmatrix}, \quad \det(A_2) = (2)(7) - (5)(1) = 9$$

$$y = \frac{\det(A_2)}{\det(A)} = \frac{9}{5}$$

Solution: $(x, y) = \left(\frac{8}{5}, \frac{9}{5}\right)$

Computational Note

Cramer's rule is elegant theoretically but inefficient computationally. For an $n \times n$ system, it requires computing $n + 1$ determinants. Gaussian elimination is much faster for large systems!

4.3.5 The Adjugate Matrix and Inverse Formula

Determinants provide an explicit formula for matrix inverses.

Definition: Adjugate Matrix

The **adjugate** (or **classical adjoint**) of A, denoted $\text{adj}(A)$, is the transpose of the cofactor matrix:

$$\text{adj}(A) = [C_{ij}]^T$$

where C_{ij} are the cofactors of A.

Inverse Formula Using Determinants

If A is invertible:

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

Example 4.3.4. *Finding Inverse Using Adjugate Find the inverse of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ using the adjugate.*

Solution:

First, $\det(A) = (1)(4) - (2)(3) = -2 \neq 0$, *so A is invertible.*

Compute cofactors:

$$\begin{aligned} C_{11} &= (-1)^{1+1}(4) = 4 \\ C_{12} &= (-1)^{1+2}(3) = -3 \\ C_{21} &= (-1)^{2+1}(2) = -2 \\ C_{22} &= (-1)^{2+2}(1) = 1 \end{aligned}$$

Cofactor matrix:

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

Adjugate:

$$adj(A) = C^T = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$$

Inverse:

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

Verify: $AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ✓

4.4 Applications of Determinants

Determinants appear throughout mathematics and its applications.

4.4.1 Area and Volume Calculations

We've seen that determinants compute areas and volumes. This is useful in:

1. Calculus: Change of Variables

In multivariable calculus, when changing variables in integration, the Jacobian determinant appears:

$$\iint_R f(x,y)\,dx\,dy = \iint_S f(u(s,t), v(s,t))\,|\det(J)|\;ds\,dt$$

where J is the Jacobian matrix of partial derivatives.

2. Computer Graphics

Determinants help compute areas for:

- Culling back-facing polygons

- Point-in-polygon tests
- Computing triangle areas for texture mapping

4.4.2 Linear Independence

Determinant Test for Linear Independence

Vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in $\mathbb{R}^n$ are linearly independent if and only if the determinant of the matrix with these vectors as columns (or rows) is nonzero.

Example 4.4.1. *Testing Linear Independence Are the vectors* $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$ *linearly independent?*

Solution:

Form the matrix:

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

Compute $\det(A)$ *by row reduction:*

$R_2 - 2R_1$, $R_3 - 3R_1$:

$$\begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix}$$

$R_3 - 2R_2$:

$$\begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}$$

The last row is all zeros, so $\det(A) = 0$.

Therefore, the vectors are ***linearly dependent***.

(In fact, $\vec{v}_3 = 2\vec{v}_2 - \vec{v}_1$*)*

4.4.3 Eigenvalues (Preview)

In Chapter 7, we'll see that eigenvalues λ of a matrix A satisfy:

$$\det(A - \lambda I) = 0$$

This is called the **characteristic equation**. The determinant is essential for finding eigenvalues!

4.4.4 Differential Equations

In systems of differential equations, the Wronskian determinant tests whether solutions are linearly independent.

For functions $f_1(t), f_2(t), \ldots, f_n(t)$, the Wronskian is:

$$W(t) = \begin{vmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f_1'(t) & f_2'(t) & \cdots & f_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{vmatrix}$$

If $W(t) \neq 0$ at some point, the functions are linearly independent.

4.5 Chapter Summary

In this chapter, we explored the determinant—a powerful scalar invariant of square matrices:

- **Geometric meaning**: Determinants measure signed area/volume scaling
- **Invertibility**: A is invertible $\Leftrightarrow \det(A) \neq 0$
- **Computation**: Cofactor expansion, or more efficiently, row reduction
- **Properties**: $\det(AB) = \det(A)\det(B)$, $\det(A^T) = \det(A)$, $\det(A^{-1}) = 1/\det(A)$
- **Applications**: Cramer's rule, area/volume, linear independence, eigenvalues
- **Row operations**: Predictable effects on determinant values

Looking Ahead

In Chapter 5, we'll formalize the concept of vector spaces—abstract structures that generalize $\mathbb{R}^n$. We'll study bases, dimensions, and subspaces, building the theoretical framework that underlies all of linear algebra. The determinant will reappear when we discuss bases and dimension!

4.6 Practice Problems

4.6.1 Basic Problems

1. Compute the determinants:

   (a) $\begin{vmatrix} 3 & 2 \\ 1 & 4 \end{vmatrix}$

   (b) $\begin{vmatrix} -1 & 5 \\ 2 & 3 \end{vmatrix}$

   (c) $\begin{vmatrix} 6 & 3 \\ 4 & 2 \end{vmatrix}$

2. Find the area of the parallelogram with adjacent sides $\vec{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

3. Determine if the matrix $A = \begin{bmatrix} 5 & 10 \\ 2 & 4 \end{bmatrix}$ is invertible.

4. Compute $\det(A)$ using cofactor expansion along the first row:

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

5. Find $\det(A)$ where $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

6. Compute using cofactor expansion (choose strategically):

$$\begin{vmatrix} 1 & 0 & 2 \\ 3 & 4 & 0 \\ 5 & 6 & 7 \end{vmatrix}$$

7. Let $\det(A) = 3$ for a 3×3 matrix A. Find:

   (a) $\det(2A)$

   (b) $\det(A^{-1})$

   (c) $\det(A^T)$

8. Use row operations to compute:
$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 5 & 4 \\ 1 & 3 & 6 \end{vmatrix}$$

4.6.2 Intermediate Problems

9. Solve using Cramer's rule:
$$\begin{aligned} 3x + 2y &= 7 \\ x + 4y &= 5 \end{aligned}$$

10. Compute the determinant:
$$\begin{vmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$$

11. Find the volume of the parallelepiped with edges:
$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

12. If $\det(A) = 2$ and $\det(B) = -3$, find:
   (a) $\det(AB)$
   (b) $\det(A^2 B^{-1})$
   (c) $\det(3A)$ (where A is 3×3)

13. Determine if the vectors are linearly independent:
$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

14. Find the inverse of $A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ using the adjugate formula.

15. For what value(s) of k is the matrix singular?

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

16. Prove that if A is a 2×2 matrix with $\det(A) = 1$, then $\det(A^{100}) = 1$.

4.6.3 Challenge Problems

17. Prove that $\det(A^T) = \det(A)$ for any square matrix A.

18. Show that if A has a row of zeros, then $\det(A) = 0$.

19. Prove that $\det(AB) = \det(A)\det(B)$ for 2×2 matrices by direct calculation.

20. Compute the determinant:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$$

21. A matrix A is called **orthogonal** if $A^T A = I$. Prove that if A is orthogonal, then $\det(A) = \pm 1$.

22. Find all 2×2 matrices A such that $\det(A + I) = \det(A) + \det(I)$.

23. The determinant can be defined recursively. Prove that for an $n \times n$ matrix:

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

is independent of which row or column is chosen for expansion (this is non-trivial!).

24. Compute the determinant:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}$$

5

Vector Spaces

"'The essence of mathematics lies in its freedom."

— Georg Cantor

Chapter Overview

So far, we've worked primarily with vectors in $\mathbb{R}^n$—arrows with numerical components that we can add and scale. But the algebraic properties of vectors are shared by many other mathematical objects: polynomials, matrices, functions, and even solutions to differential equations. In this chapter, we'll abstract the essential properties of vectors to define **vector spaces**—algebraic structures that capture the essence of "vector-like" behavior. This abstraction is powerful: theorems we prove about abstract vector spaces apply simultaneously to geometry, algebra, calculus, and beyond. We'll explore subspaces, linear independence, bases, and dimension—foundational concepts that unify much of linear algebra.

5.1 Introduction to Vector Spaces

5.1.1 Motivation: Beyond $\mathbb{R}^n$

Consider these mathematical objects:

1. Polynomials: $p(x) = 2x^2 + 3x - 1$ and $q(x) = x^2 - x + 4$

We can add them: $(p + q)(x) = 3x^2 + 2x + 3$

We can scale them: $(5p)(x) = 10x^2 + 15x - 5$

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

We can add them: $A + B = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$

We can scale them: $3A = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$

3. Functions: $f(x) = \sin(x)$ and $g(x) = \cos(x)$

We can add them: $(f + g)(x) = \sin(x) + \cos(x)$

We can scale them: $(2f)(x) = 2\sin(x)$

All these objects behave like vectors! They satisfy the same algebraic rules. This observation leads us to the concept of a vector space.

5.1.2 Definition of a Vector Space

Definition: Vector Space

A **vector space** is a set V together with two operations:

- **Addition**: $\vec{u} + \vec{v} \in V$ for all $\vec{u}, \vec{v} \in V$
- **Scalar multiplication**: $c\vec{v} \in V$ for all $c \in \mathbb{R}$, $\vec{v} \in V$

that satisfy the following **axioms** for all $\vec{u}, \vec{v}, \vec{w} \in V$ and scalars $c, d \in \mathbb{R}$:

Addition axioms:

1. **Closure**: $\vec{u} + \vec{v} \in V$
2. **Commutativity**: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
3. **Associativity**: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
4. **Zero vector**: There exists $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$
5. **Additive inverse**: For each $\vec{v} \in V$, there exists $-\vec{v} \in V$ such that $\vec{v} + (-\vec{v}) = \vec{0}$

Scalar multiplication axioms:

6. **Closure**: $c\vec{v} \in V$
7. **Distributivity**: $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
8. **Distributivity**: $(c + d)\vec{v} = c\vec{v} + d\vec{v}$
9. **Associativity**: $c(d\vec{v}) = (cd)\vec{v}$
10. **Identity**: $1\vec{v} = \vec{v}$

Elements of V are called **vectors**, even if they're not geometric arrows!

Key Insight

The power of abstraction: Any theorem we prove using only these axioms applies to *all* vector spaces—whether we're working with arrows, polynomials, matrices, or functions. We prove it once, use it everywhere!

5.1.3 Examples of Vector Spaces

Example 5.1.1. $\mathbb{R}^n$ *The set of all n-tuples of real numbers with standard addition and scalar multiplication:*

$$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n \rangle$$

$$c\vec{v} = \langle cv_1, cv_2, \ldots, cv_n \rangle$$

This is the prototypical vector space we've been using all along!

Example 5.1.2. M_{mn}*: Matrices The set of all $m \times n$ matrices with matrix addition and scalar multiplication is a vector space.*

The zero vector is the $m \times n$ zero matrix.

Dimension: M_{mn} has dimension mn (we'll formalize this soon).

Example 5.1.3. P_n*: Polynomials The set of all polynomials of degree at most n:*

$$P_n = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n : a_i \in \mathbb{R}\}$$

Addition: $(p + q)(x) = p(x) + q(x)$

Scalar multiplication: $(cp)(x) = c \cdot p(x)$

Zero vector: The zero polynomial $p(x) = 0$

This is a vector space! (We'll verify the axioms below.)

Example 5.1.4. $C[a, b]$*: Continuous Functions The set of all continuous real-valued functions on interval $[a, b]$ with:*

Addition: $(f + g)(x) = f(x) + g(x)$

Scalar multiplication: $(cf)(x) = c \cdot f(x)$

Zero vector: The zero function $f(x) = 0$ for all x

This is an infinite-dimensional vector space!

5.1.4 Verifying Vector Space Axioms

Let's verify that P_2 (polynomials of degree ≤ 2) is a vector space.

Example 5.1.5. *Verifying P_2 is a Vector Space Let $p(x) = a_0 + a_1x + a_2x^2$ and $q(x) = b_0 + b_1x + b_2x^2$.*

Addition:

$$(p + q)(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

This is still a polynomial of degree ≤ 2, so closure holds. ✓

Commutativity:

$$(p+q)(x) = (a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2 = (b_0+a_0)+(b_1+a_1)x+(b_2+a_2)x^2 = (q+p)(x)$$

✓

Zero vector: *The zero polynomial $0(x) = 0$ satisfies $p(x) + 0(x) = p(x)$.* ✓

Scalar multiplication:

$$(cp)(x) = c(a_0 + a_1x + a_2x^2) = ca_0 + ca_1x + ca_2x^2$$

Still degree ≤ 2. ✓

The other axioms follow from properties of real numbers. Therefore, P_2 is a vector space.

5.1.5 Non-Examples

Not every set with addition and scaling is a vector space!

Example 5.1.6. *Non-Example: Positive Real Numbers Let $V = \{x \in \mathbb{R} : x > 0\}$ with standard addition and multiplication.*

*This is **not** a vector space because it fails closure under addition: If $\vec{u} = 2$ and $\vec{v} = 3$, then $\vec{u} + \vec{v} = 5 \in V$* ✓

But there's no zero vector! We'd need $\vec{0}$ with $\vec{v} + \vec{0} = \vec{v}$, meaning $\vec{0} = 0$. But $0 \notin V$ (not positive).

Fails axiom 4. $\times$

Example 5.1.7. *Non-Example: First Quadrant Let $V = \{(x, y) : x \geq 0, y \geq 0\}$ in $\mathbb{R}^2$.*

*This is **not** a vector space because scalar multiplication can take us outside V:*

If $\vec{v} = (1, 1) \in V$ and $c = -1$, then $c\vec{v} = (-1, -1) \notin V$.

Fails closure under scalar multiplication. $\times$

5.2 Subspaces

Often we want to study vector spaces that live inside larger vector spaces.

5.2.1 Definition and Examples

Definition: Subspace

A subset W of a vector space V is called a **subspace** of V if W is itself a vector space under the same operations (addition and scalar multiplication inherited from V).
Equivalently, W is a subspace if:

1. $\vec{0} \in W$ (contains the zero vector)
2. $\vec{u} + \vec{v} \in W$ for all $\vec{u}, \vec{v} \in W$ (closed under addition)
3. $c\vec{v} \in W$ for all $c \in \mathbb{R}$, $\vec{v} \in W$ (closed under scalar multiplication)

Subspace Test

To verify W is a subspace, you only need to check these three conditions! The other axioms are automatically inherited from V.

Example 5.2.1. *Subspaces of* $\mathbb{R}^3$ ***1. Lines through the origin:***
$W = \{t\vec{v} : t \in \mathbb{R}\}$ *for some fixed* $\vec{v} \neq \vec{0}$
This is a subspace:

- $\vec{0} = 0\vec{v} \in W$ ✓
- *If* $\vec{u}_1 = t_1\vec{v}$ *and* $\vec{u}_2 = t_2\vec{v}$, *then* $\vec{u}_1 + \vec{u}_2 = (t_1 + t_2)\vec{v} \in W$ ✓
- *If* $\vec{u} = t\vec{v}$, *then* $c\vec{u} = (ct)\vec{v} \in W$ ✓

2. Planes through the origin:
$W = \{s\vec{v}_1 + t\vec{v}_2 : s, t \in \mathbb{R}\}$ *for fixed* $\vec{v}_1, \vec{v}_2$
This is a subspace (verification similar).
3. All of $\mathbb{R}^3$ ***itself***
Trivially a subspace.
4. The zero subspace $\{\vec{0}\}$
Contains only the zero vector. This is the smallest subspace.

Geometric Insight

Subspaces of $\mathbb{R}^3$ are:

- $\{\vec{0}\}$ (the origin)
- Lines through the origin

- Planes through the origin
- All of $\mathbb{R}^3$

Key: Subspaces must pass through the origin!

Example 5.2.2. *Subspace of Polynomials Let* $W = \{p(x) \in P_3 : p(0) = 0\}$ *(polynomials with no constant term).*

Is W *a subspace of* P_3*?*

Check:

1. *Zero vector: The zero polynomial satisfies* $0(0) = 0$*, so* $\vec{0} \in W$ ✓
2. *Closure under addition: If* $p(0) = 0$ *and* $q(0) = 0$*, then* $(p+q)(0) = p(0) + q(0) = 0 + 0 = 0$ ✓
3. *Closure under scaling: If* $p(0) = 0$*, then* $(cp)(0) = c \cdot p(0) = c \cdot 0 = 0$ ✓

Yes, W *is a subspace!*

Explicitly: $W = \{a_1x + a_2x^2 + a_3x^3 : a_i \in \mathbb{R}\}$

Example 5.2.3. *Non-Subspace Let* $W = \{(x, y) \in \mathbb{R}^2 : x + y = 1\}$ *(a line not through the origin).*

Is W *a subspace of* $\mathbb{R}^2$*?*

Check: *Does* $\vec{0} = (0, 0) \in W$*?*

We'd need $0 + 0 = 1$*, which is false.*

Therefore W *is* ***not a subspace****.* ×

Geometric note: *Lines not through the origin are never subspaces!*

5.2.2 Operations on Subspaces

Definition: Intersection and Sum

If U and W are subspaces of V:
Intersection: $U \cap W = \{\vec{v} \in V : \vec{v} \in U \text{ and } \vec{v} \in W\}$
Sum: $U + W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$

Theorem: Subspace Properties

- The intersection $U \cap W$ is always a subspace
- The sum $U + W$ is always a subspace

- The union $U \cup W$ is generally **not** a subspace

Example 5.2.4. *Intersection of Subspaces Let U be the xy-plane in $\mathbb{R}^3$: $U = \{(x, y, 0) : x, y \in \mathbb{R}\}$*

Let W be the xz-plane: $W = \{(x, 0, z) : x, z \in \mathbb{R}\}$

Then $U \cap W$ is the x-axis: $U \cap W = \{(x, 0, 0) : x \in \mathbb{R}\}$

This is a subspace (a line through the origin). ✓

5.3 Linear Independence

Which vectors are "essential" for describing a vector space?

5.3.1 Definition and Intuition

Definition: Linear Independence

Vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ in a vector space V are **linearly independent** if the only solution to

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$$

is the trivial solution $c_1 = c_2 = \cdots = c_k = 0$.
If there exists a nontrivial solution (some $c_i \neq 0$), the vectors are **linearly dependent**.

Intuition: Vectors are linearly independent if none can be expressed as a combination of the others. They're all "essential"—removing any one reduces what we can represent.

Example 5.3.1. *Independent Vectors in $\mathbb{R}^3$ Are $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ linearly independent?*

Solution:

Consider $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$:

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This gives:

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $c_1 = c_2 = c_3 = 0$ is the only solution.

*The vectors are **linearly independent**.* ✓

These are the standard basis vectors!

Example 5.3.2. *Dependent Vectors Are $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ linearly independent?*

Solution:

Consider $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$:

$$c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Notice $\vec{v}_2 = 2\vec{v}_1$, so we can choose $c_1 = 2, c_2 = -1$:

$$2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

*Since there's a nontrivial solution, the vectors are **linearly dependent**.* $\times$

5.3.2 Testing for Linear Independence

For vectors in $\mathbb{R}^n$, we can use matrices:

Matrix Method for Testing Independence

To test if vectors $\vec{v}_1, \ldots, \vec{v}_k$ in $\mathbb{R}^n$ are linearly independent:

1. Form the matrix $A = [\vec{v}_1 \ \vec{v}_2 \ \cdots \ \vec{v}_k]$ with vectors as columns
2. Row-reduce to find solutions to $A\vec{x} = \vec{0}$
3. Linearly independent $\Leftrightarrow$ only trivial solution $\Leftrightarrow$ all variables are leading variables

If $k = n$ (square matrix), can also check: independent $\Leftrightarrow$ $\det(A) \neq 0$

Example 5.3.3. *Testing Independence with Matrices Determine if* $\vec{v}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}4\\5\\6\end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix}7\\8\\9\end{bmatrix}$ *are linearly independent.*

Solution:

Form the matrix:

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

From Chapter 4, we know $\det(A) = 0$ *(we computed this!).*

Since the determinant is zero, the vectors are ***linearly dependent***. $\times$

In fact, $\vec{v}_3 = 2\vec{v}_2 - \vec{v}_1$.

5.3.3 Properties of Linear Independence

Important Facts

1. A set containing the zero vector is always linearly dependent
2. A set with just one nonzero vector is linearly independent
3. Two vectors are linearly dependent if and only if one is a scalar multiple of the other
4. Any subset of a linearly independent set is linearly independent
5. Any superset of a linearly dependent set is linearly dependent
6. In $\mathbb{R}^n$, any set of more than n vectors is linearly dependent

Common Mistake

Linear independence is about the *set* of vectors, not individual vectors. It's a relationship among all the vectors together. One vector alone can't be "independent"—it's the set $\{\vec{v}\}$ that's independent (if $\vec{v} \neq \vec{0}$).

5.4 Span and Basis

5.4.1 Span of a Set of Vectors

Definition: Span

The **span** of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ is the set of all linear combinations:

$$\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\} = \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k : c_i \in \mathbb{R}\}$$

The span is the "space" generated by these vectors—everything you can reach by combining them.

Key Theorem

$\text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ is always a subspace of V.

Example 5.4.1. *Span in* $\mathbb{R}^3$ *What is Span* $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$*?*

Solution:

Any linear combination is:

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ 0 \end{bmatrix}$$

This is the xy-plane! Span$\{\vec{v}_1, \vec{v}_2\}$ $=$ *xy-plane.*

Example 5.4.2. *Span of Polynomials What is Span*$\{1, x, x^2\}$ *in the space of polynomials?*

Solution:

Any linear combination is:

$$c_1 \cdot 1 + c_2 \cdot x + c_3 \cdot x^2 = c_1 + c_2 x + c_3 x^2$$

This is exactly P_2*, all polynomials of degree at most 2.*

So Span$\{1, x, x^2\} = P_2$.

5.4.2 Spanning Sets

Definition: Spanning Set

A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ **spans** a vector space V if every vector in V can be written as a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$.
That is: $V = \text{Span}\{\vec{v}_1, \ldots, \vec{v}_k\}$

Example 5.4.3. *Spanning $\mathbb{R}^2$ Does $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ span $\mathbb{R}^2$?*

Solution:
Any vector $\begin{bmatrix} a \\ b \end{bmatrix}$ can be written as:

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Yes, these two vectors span $\mathbb{R}^2$. ✓

5.4.3 Basis

The most important spanning sets are those that are also linearly independent.

Definition: Basis

A **basis** for a vector space V is a set of vectors that:

1. Spans V (every vector in V is a linear combination)
2. Is linearly independent (no redundancy)

A basis is a "minimal spanning set" or a "maximal independent set."

Why Bases Matter

A basis gives us a **coordinate system** for the vector space. Every vector can be uniquely expressed in terms of the basis vectors. Bases are the vector space analog of axes in geometry!

Example 5.4.4. *Standard Basis for $\mathbb{R}^3$ The **standard basis** for $\mathbb{R}^3$ is:*

$$\left\{ \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Spans: *Any* $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\vec{e}_1 + b\vec{e}_2 + c\vec{e}_3$ ✓

Independent: *Only* $0\vec{e}_1 + 0\vec{e}_2 + 0\vec{e}_3 = \vec{0}$ ✓

This is a basis!

Example 5.4.5. *Standard Basis for P_2 The standard basis for P_2 is $\{1, x, x^2\}$.*

Any $p(x) = a_0 + a_1x + a_2x^2 = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2$

These are independent: if $c_0 \cdot 1 + c_1 \cdot x + c_2 \cdot x^2 = 0$ (the zero polynomial), then $c_0 = c_1 = c_2 = 0$.

Example 5.4.6. *Alternative Basis for $\mathbb{R}^2$ Is $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ a basis for $\mathbb{R}^2$?*

Check independence: $c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

This gives: $c_1 + c_2 = 0$ and $c_1 - c_2 = 0$, so $c_1 = c_2 = 0$. Independent! ✓

Check spanning: *Can we write* $\begin{bmatrix} a \\ b \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$?

This gives: $c_1 + c_2 = a$ and $c_1 - c_2 = b$.

Solving: $c_1 = \frac{a+b}{2}$, $c_2 = \frac{a-b}{2}$.

For any a, b, we can find c_1, c_2. Spans! ✓

Yes, this is a basis for $\mathbb{R}^2$. It's just a different coordinate system!

5.4.4 Coordinates Relative to a Basis

Definition: Coordinate Vector

If $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for V, and

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

then the **coordinate vector** of $\vec{v}$ relative to $\mathcal{B}$ is:

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

The coefficients $c_1, \dots, c_n$ are the **coordinates** of $\vec{v}$ in basis $\mathcal{B}$.

Example 5.4.7. *Finding Coordinates Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ be a basis for $\mathbb{R}^2$.*

Find $[\vec{v}]_{\mathcal{B}}$ *where* $\vec{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

Solution:

We need: $\begin{bmatrix} 3 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

This gives: $c_1 + c_2 = 3$ *and* $c_1 - c_2 = 1$.

Solving: $c_1 = 2$, $c_2 = 1$.

Therefore:

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

In the standard basis, $\vec{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. *In basis* $\mathcal{B}$*, it's* $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

5.5 Dimension

5.5.1 The Dimension Theorem

A remarkable fact: all bases for a given vector space have the same number of vectors!

Theorem: Unique Dimension

If V is a vector space with a finite basis, then all bases for V have the same number of elements.
This number is called the **dimension** of V, denoted $\dim(V)$.

Why This Matters

Dimension is an intrinsic property of the space itself, not dependent on which basis we choose. It measures the "degrees of freedom" in the space—how many independent directions exist.

Example 5.5.1. *Dimensions of Common Spaces*

- $\dim(\mathbb{R}^n) = n$ *(basis: n standard vectors)*
- $\dim(P_n) = n + 1$ *(basis: $\{1, x, x^2, \ldots, x^n\}$)*
- $\dim(M_{mn}) = mn$ *(basis: matrices with single 1, rest 0s)*
- $\dim(\{\vec{0}\}) = 0$ *(empty set is basis)*
- $\dim(C[a, b]) = \infty$ *(infinite-dimensional!)*

5.5.2 Finding a Basis and Dimension

Example 5.5.2. *Basis for a Subspace Find a basis for* $W = \{(x, y, z) \in \mathbb{R}^3 : x+2y-z = 0\}$.

Solution:

This is a plane through the origin (a subspace).

Solve for one variable: $z = x + 2y$

So $(x, y, z) = (x, y, x + 2y) = x(1, 0, 1) + y(0, 1, 2)$

Every vector in W *is a combination of* $\vec{v}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix}0\\1\\2\end{bmatrix}$.

These are linearly independent (not scalar multiples).

Basis: $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\2\end{bmatrix}\right\}$

Dimension: $\dim(W) = 2$

Geometric interpretation: W *is a 2-dimensional plane in 3-dimensional space.*

5.5.3 Dimension and Linear Independence/Spanning

Important Theorems

Let V be a vector space with $\dim(V) = n$.

1. Any set of n linearly independent vectors in V is a basis
2. Any set of n vectors that spans V is a basis
3. Any linearly independent set can be extended to a basis
4. Any spanning set contains a basis
5. Any set of more than n vectors is linearly dependent
6. Any set of fewer than n vectors does not span V

Example 5.5.3. *Using Dimension In* $\mathbb{R}^3$, *is* $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ *a basis?*

Solution:

Since $\dim(\mathbb{R}^3) = 3$ *and we have 3 vectors, we only need to check either independence or* *spanning (not both!).*

Check independence using determinant:

$$\det \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = 1 \neq 0$$

Independent! Since we have 3 independent vectors in a 3-dimensional space, this is automatically a basis. ✓

5.6 Column Space, Row Space, and Null Space

For matrices, several important subspaces arise naturally.

5.6.1 Column Space

Definition: Column Space

For an $m \times n$ matrix A, the **column space** $\text{Col}(A)$ is the span of the column vectors of A.
$\text{Col}(A)$ is a subspace of $\mathbb{R}^m$.
The **rank** of A is $\text{rank}(A) = \dim(\text{Col}(A)) =$ number of pivot columns.

Example 5.6.1. *Finding Column Space Find a basis for the column space of* $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 6 & 7 \end{bmatrix}$.

Solution:

Row reduce to identify pivot columns:

$$A \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Pivots in columns 1 and 3.

A basis for Col(A) *consists of the pivot columns from the original matrix:*

$$\text{Basis: } \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}$$

rank$(A) = 2$

5.6.2 Null Space

Definition: Null Space

For an $m \times n$ matrix A, the **null space** $\text{Nul}(A)$ is the set of all solutions to $A\vec{x} = \vec{0}$:

$$\text{Nul}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$$

$\text{Nul}(A)$ is a subspace of $\mathbb{R}^n$.
The **nullity** of A is $\text{nullity}(A) = \dim(\text{Nul}(A))$ = number of free variables.

Example 5.6.2. *Finding Null Space Find a basis for the null space of* $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$.

Solution:

Solve $A\vec{x} = \vec{0}$*:*

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Row reduce:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

This gives: $x_1 + 2x_2 + 3x_3 = 0$*, so* $x_1 = -2x_2 - 3x_3$

Free variables: x_2, x_3

General solution:

$$\vec{x} = \begin{bmatrix} -2x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Basis: $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$

nullity$(A) = 2$

5.6.3 Rank-Nullity Theorem

Rank-Nullity Theorem

For an $m \times n$ matrix A:

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

That is: (number of pivot columns) + (number of free variables) = (total columns)

This is one of the fundamental theorems of linear algebra!

5.7 Chapter Summary

In this chapter, we developed the abstract theory of vector spaces:

- **Vector spaces** are sets with addition and scalar multiplication satisfying 10 axioms
- **Subspaces** are vector spaces within vector spaces (must contain $\vec{0}$)
- **Linear independence**: No vector is a combination of the others
- **Span**: All possible linear combinations of a set of vectors
- **Basis**: A linearly independent spanning set—a coordinate system
- **Dimension**: The number of vectors in any basis (intrinsic to the space)
- **Column space, null space**: Important subspaces associated with matrices
- **Rank-Nullity Theorem**: Relates dimensions of column space and null space

Looking Ahead

In Chapter 6, we'll study linear transformations—functions between vector spaces that preserve the vector space structure. We'll see how matrices represent linear transformations and how bases affect this representation. The abstract theory we've developed will become concrete and computational!

5.8 Practice Problems

5.8.1 Basic Problems

1. Determine if $W = \{(x, y, 0) : x, y \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$.

2. Is $W = \{(x, y) \in \mathbb{R}^2 : xy = 0\}$ a subspace of $\mathbb{R}^2$?

3. Determine if the vectors are linearly independent:

$$\begin{bmatrix}1\\2\end{bmatrix}, \quad \begin{bmatrix}3\\6\end{bmatrix}$$

4. Find Span $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$ and describe it geometrically.

5. Determine if $\{1, x, x^2\}$ is a basis for P_2.

6. Find the dimension of $M_{2\times 3}$ (the space of 2×3 matrices).

7. Is $\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$ a basis for $\mathbb{R}^3$?

8. Find a basis for the null space of $A = \begin{bmatrix}1 & 2 & 3\\0 & 0 & 0\end{bmatrix}$.

5.8.2 Intermediate Problems

9. Find a basis for $W = \{(x, y, z, w) \in \mathbb{R}^4 : x + y = 0, z - w = 0\}$.

10. Determine if $\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}, \begin{bmatrix}7\\8\\9\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\}$ is linearly independent.

11. Let $\mathcal{B} = \left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\5\end{bmatrix}\right\}$ be a basis for $\mathbb{R}^2$. Find $[\vec{v}]_{\mathcal{B}}$ where $\vec{v} = \begin{bmatrix}7\\16\end{bmatrix}$.

12. Find a basis for the column space of:

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

13. Verify the Rank-Nullity Theorem for $A = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}$.

14. Find a basis for P_2 consisting of polynomials that all satisfy $p(1) = 0$.

15. Extend $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ to a basis for $\mathbb{R}^3$.

5.8.3 Challenge Problems

16. Prove that if $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is linearly independent, then so is $\{\vec{v}_1, \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_2 + \vec{v}_3\}$.

17. Let U and W be subspaces of a vector space V. Prove that $U \cap W$ is also a subspace.

18. Prove that $U \cup W$ is a subspace if and only if $U \subseteq W$ or $W \subseteq U$.

19. If $\dim(V) = n$ and $\{\vec{v}_1, \dots, \vec{v}_n\}$ spans V, prove that it's a basis.

20. Let A be an $m \times n$ matrix. Prove that $\text{rank}(A) = \text{rank}(A^T)$.

21. Find the dimension of the subspace of $M_{3\times3}$ consisting of all symmetric matrices.

22. Prove the Rank-Nullity Theorem: If A is $m \times n$, then $\text{rank}(A) + \text{nullity}(A) = n$.

23. Let V be the vector space of all polynomials. Show that V is infinite-dimensional.

6

Linear Transformations

"'Mathematics is the art of reducing any problem to linear algebra."

— *William Stein*

Chapter Overview

Functions are everywhere in mathematics—they describe how one quantity depends on another. In linear algebra, we study a special class of functions called **linear transformations**: functions between vector spaces that preserve the vector space structure. These transformations include rotations, reflections, projections, and many other geometric operations. The remarkable connection: every linear transformation can be represented by a matrix, and conversely, every matrix defines a linear transformation. This chapter bridges abstract vector spaces with concrete matrix computations, showing how geometry, algebra, and analysis unite in the theory of linear transformations.

6.1 Introduction to Linear Transformations

6.1.1 Functions Between Vector Spaces

Recall that a function $T : V \to W$ assigns to each element $\vec{v} \in V$ a unique element $T(\vec{v}) \in W$.

- V is the **domain**
- W is the **codomain**
- $T(\vec{v})$ is the **image** of $\vec{v}$
- The set of all images is the **range** of T

Example 6.1.1. *Functions Between Vector Spaces* ***1.*** $T : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by* $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$ *(scaling)*

2. $T : P_2 \to P_1$ *defined by* $T(p(x)) = p'(x)$ *(differentiation)*

3. $T : \mathbb{R}^3 \to \mathbb{R}$ *defined by* $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = x + y + z$ *(projection onto a line)*

6.1.2 Definition of Linear Transformation

Not all functions between vector spaces are linear. What makes a transformation "linear"?

Definition: Linear Transformation

A function $T : V \to W$ between vector spaces is a **linear transformation** if for all $\vec{u}, \vec{v} \in V$ and all scalars $c \in \mathbb{R}$:
1. Additivity: $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
2. Homogeneity: $T(c\vec{v}) = cT(\vec{v})$
These can be combined into one condition: $T(c_1\vec{u} + c_2\vec{v}) = c_1T(\vec{u}) + c_2T(\vec{v})$

Intuition: Linearity Preserves Structure

A linear transformation preserves:

- Lines remain lines (don't curve)
- The origin stays fixed: $T(\vec{0}) = \vec{0}$
- Parallel lines remain parallel
- Ratios of distances along lines are preserved

Linear transformations are the "nicest" kind of function for vector spaces.

Example 6.1.2. *Verifying Linearity Is* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by* $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + y \\ 2x \end{bmatrix}$ *linear?*

Solution:

Check additivity:

$$
\begin{aligned}
T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) &= T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) \\
&= \begin{bmatrix} (x_1 + x_2) + (y_1 + y_2) \\ 2(x_1 + x_2) \end{bmatrix} \\
&= \begin{bmatrix} x_1 + y_1 + x_2 + y_2 \\ 2x_1 + 2x_2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 + y_1 \\ 2x_1 \end{bmatrix} + \begin{bmatrix} x_2 + y_2 \\ 2x_2 \end{bmatrix} \\
&= T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \quad \checkmark
\end{aligned}
$$

Check homogeneity:

$$
\begin{aligned}
T\left(c\begin{bmatrix} x \\ y \end{bmatrix}\right) &= T\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right) \\
&= \begin{bmatrix} cx + cy \\ 2cx \end{bmatrix} \\
&= c\begin{bmatrix} x + y \\ 2x \end{bmatrix} \\
&= cT\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) \quad \checkmark
\end{aligned}
$$

Both conditions hold, so T is linear.

Example 6.1.3. *Non-Linear Transformation Is $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + 1 \\ y \end{bmatrix}$ linear?*

Solution:

Check if $T(\vec{0}) = \vec{0}$:

$$T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 + 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

*Since $T(\vec{0}) \neq \vec{0}$, this transformation is **not linear**.* $\times$

Geometric interpretation: *This is a translation, which shifts everything by 1 unit right. Translations (except by zero) are not linear.*

Common Mistake: Adding Constants

Any transformation with a constant term added (like $T(x) = 2x + 3$) is not linear. Linear transformations must map zero to zero!

6.1.3 Properties of Linear Transformations

Important Properties

If $T : V \to W$ is linear, then:

1. $T(\vec{0}) = \vec{0}$ (zero vector maps to zero vector)
2. $T(-\vec{v}) = -T(\vec{v})$ (negatives are preserved)
3. $T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v})$ (differences are preserved)
4. $T(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1T(\vec{v}_1) + \cdots + c_nT(\vec{v}_n)$ (all linear combinations are preserved)

Property 4 is crucial: a linear transformation is completely determined by what it does to a basis!

6.2 Matrix Representation of Linear Transformations

6.2.1 The Standard Matrix

Every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be represented by a matrix.

Theorem: Standard Matrix Representation

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
There exists a unique $m \times n$ matrix A such that:

$$T(\vec{x}) = A\vec{x} \quad \text{for all } \vec{x} \in \mathbb{R}^n$$

The matrix A is called the **standard matrix** for T, and its columns are:

$$A = [T(\vec{e}_1)\ T(\vec{e}_2)\ \cdots\ T(\vec{e}_n)]$$

where $\vec{e}_1, \ldots, \vec{e}_n$ are the standard basis vectors of $\mathbb{R}^n$.

Key Insight

To find the matrix for a linear transformation, just see what it does to the standard basis vectors!

Example 6.2.1. *Finding the Standard Matrix Find the standard matrix for $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by:*

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - y \\ 2x + y \\ 3y \end{bmatrix}$$

Solution:

Compute $T(\vec{e}_1)$ and $T(\vec{e}_2)$:

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 - 0 \\ 2(1) + 0 \\ 3(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 - 1 \\ 2(0) + 1 \\ 3(1) \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}$$

The standard matrix is:

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

Verify: $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ 2x + y \\ 3y \end{bmatrix}$ ✓

6.2.2 Geometric Linear Transformations in $\mathbb{R}^2$

Let's explore important geometric transformations.

Example 6.2.2. *Rotation Rotation counterclockwise by angle θ about the origin has matrix:*

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Derivation: *Where do the standard basis vectors go?*

$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *rotates to* $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$

$\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *rotates to* $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$

Example: *Rotation by* 90ř*:*

$$R_{90ř} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Check: $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ✓

Example 6.2.3. *Reflection Reflection across the x-axis:*

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

Reflection across the y-axis:

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

Reflection across the line $y = x$:

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

Example 6.2.4. *Scaling Scaling by factor k:*

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

Different scaling in x and y directions:

$$\begin{bmatrix} k_x & 0 \\ 0 & k_y \end{bmatrix}$$

Example 6.2.5. *Projection Projection onto the x-axis:*

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

This "flattens" the plane onto the x-axis by setting $y = 0$.

Example 6.2.6. *Shear Horizontal shear (shifts points horizontally by amount proportional to y):*

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

Transforms $\begin{bmatrix} x \\ y \end{bmatrix}$ *to* $\begin{bmatrix} x + ky \\ y \end{bmatrix}$

6.2.3 Composition of Transformations

Theorem: Composition and Matrix Multiplication

If $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ has matrix A and $T_2 : \mathbb{R}^m \to \mathbb{R}^p$ has matrix B, then:
The composition $T_2 \circ T_1 : \mathbb{R}^n \to \mathbb{R}^p$ (first apply T_1, then T_2) has matrix BA.
That is: $(T_2 \circ T_1)(\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$

Why Matrix Multiplication Works This Way

Matrix multiplication is defined precisely to make composition work! The formula AB was designed so that composing transformations corresponds to multiplying matrices.

Example 6.2.7. *Composition Example Let T_1 be rotation by 45ř and T_2 be reflection across the x-axis.*

$$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The composition $T_2 \circ T_1$ (rotate then reflect) has matrix:

$$BA = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

Note: Order matters! $T_1 \circ T_2$ (reflect then rotate) would give $AB \neq BA$.

6.3 Kernel and Range

Just as functions have domains and ranges, linear transformations have important associated subspaces.

6.3.1 Kernel (Null Space)

Definition: Kernel

The **kernel** (or **null space**) of a linear transformation $T : V \to W$ is:

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

The kernel is the set of all vectors that map to zero.

Theorem: Kernel is a Subspace

$\ker(T)$ is a subspace of the domain V.

Proof: 1. $T(\vec{0}) = \vec{0}$, so $\vec{0} \in \ker(T)$ ✓

2. If $\vec{u}, \vec{v} \in \ker(T)$, then $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) = \vec{0} + \vec{0} = \vec{0}$, so $\vec{u} + \vec{v} \in \ker(T)$ ✓

3. If $\vec{v} \in \ker(T)$ and $c \in \mathbb{R}$, then $T(c\vec{v}) = cT(\vec{v}) = c\vec{0} = \vec{0}$, so $c\vec{v} \in \ker(T)$ ✓

Example 6.3.1. *Finding a Kernel Find* $\ker(T)$ *where* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *is given by:*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2y - z \\ 2x + 4y - 2z \end{bmatrix}$$

Solution:

The standard matrix is:

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

Find $\ker(T)$ *by solving* $A\vec{x} = \vec{0}$*:*

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Row reduce:

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

This gives: $x + 2y - z = 0$*, so* $x = -2y + z$

Free variables: y, z

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

$$\ker(T) = Span \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

This is a 2-dimensional plane through the origin in $\mathbb{R}^3$*.*

6.3.2 Range (Image)

Definition: Range

The **range** (or **image**) of a linear transformation $T : V \to W$ is:

$$\text{range}(T) = \{T(\vec{v}) : \vec{v} \in V\} = \{w \in W : w = T(\vec{v}) \text{ for some } \vec{v} \in V\}$$

The range is the set of all possible outputs.

Theorem: Range is a Subspace

range(T) is a subspace of the codomain W.

For $T : \mathbb{R}^n \to \mathbb{R}^m$ with matrix A:

$$\text{range}(T) = \text{Col}(A)$$

Example 6.3.2. *Finding Range Find the range of* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *with matrix:*

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

Solution:

The range is the column space of A*.*

Row reduce to find pivot columns:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Pivot in column 1 only.

$$\textit{range}(T) = \textit{Span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$$

This is a line through the origin in $\mathbb{R}^2$*.*

6.3.3 Rank-Nullity Theorem (Revisited)

Rank-Nullity Theorem for Linear Transformations

If $T : V \to W$ is a linear transformation and V is finite-dimensional, then:

$$\dim(\ker(T)) + \dim(\text{range}(T)) = \dim(V)$$

Also written as: $\text{nullity}(T) + \text{rank}(T) = \dim(V)$

This is the same theorem we saw for matrices, now in the language of transformations!

6.4 One-to-One and Onto Transformations

6.4.1 One-to-One (Injective)

Definition: One-to-One

A linear transformation $T : V \to W$ is **one-to-one** (or **injective**) if different inputs produce different outputs:
If $T(\vec{u}) = T(\vec{v})$, then $\vec{u} = \vec{v}$.
Equivalently: T is one-to-one if $\ker(T) = \{\vec{0}\}$.

Test for One-to-One

T is one-to-one $\Leftrightarrow \ker(T) = \{\vec{0}\} \Leftrightarrow T(\vec{x}) = \vec{0}$ has only the trivial solution
For matrix A: one-to-one $\Leftrightarrow$ columns are linearly independent

Example 6.4.1. *Testing One-to-One Is* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *with matrix* $A = \begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\\0 & 0 & 0\end{bmatrix}$ *one-to-one?*

Solution:

Check if $\ker(T) = \{\vec{0}\}$ *by solving* $A\vec{x} = \vec{0}$*:*

The matrix is already in row echelon form with pivot in columns 1 and 2.

Column 3 is free, so there are nontrivial solutions to $A\vec{x} = \vec{0}$.

Therefore, $\ker(T) \neq \{\vec{0}\}$, *and* T *is* ***not one-to-one***. $\times$

6.4.2 Onto (Surjective)

Definition: Onto

A linear transformation $T : V \to W$ is **onto** (or **surjective**) if every element of W is the image of some element in V:
For every $\vec{w} \in W$, there exists $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$.
Equivalently: T is onto if $\text{range}(T) = W$.

Test for Onto

$T : \mathbb{R}^n \to \mathbb{R}^m$ with matrix A is onto $\Leftrightarrow$ $\text{range}(T) = \mathbb{R}^m \Leftrightarrow \text{rank}(A) = m \Leftrightarrow$ every row has a pivot

Example 6.4.2. *Testing Onto Is* $T : \mathbb{R}^2 \to \mathbb{R}^3$ *with matrix* $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ *onto?*

Solution:

For T *to be onto, we need range*$(T) = \mathbb{R}^3$.

But range$(T) =$ *Col*$(A) =$ *Span* $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

This is the xy-plane in $\mathbb{R}^3$, *which is only 2-dimensional.*

Since range$(T) \neq \mathbb{R}^3$, T *is* ***not onto***. $\times$

Observation: *A transformation from a lower-dimensional space to a higher-dimensional space can never be onto!*

6.4.3 Isomorphisms

Definition: Isomorphism

A linear transformation $T : V \to W$ is an **isomorphism** if it is both one-to-one and onto.
Two vector spaces are **isomorphic** if there exists an isomorphism between them.

Isomorphic spaces are "essentially the same"—they have the same dimension and structure, just different labels.

Theorem: Isomorphism and Dimension

If V and W are finite-dimensional, then V and W are isomorphic if and only if $\dim(V) = \dim(W)$.

Example 6.4.3. *Isomorphic Spaces* $\mathbb{R}^3$ *and* P_2 *are isomorphic!*

Define $T : P_2 \to \mathbb{R}^3$ *by:*

$$T(a_0 + a_1 x + a_2 x^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

This is an isomorphism (verify: one-to-one and onto).

Both spaces are 3-dimensional, so they have the same structure.

6.5 Applications of Linear Transformations

6.5.1 Computer Graphics

Linear transformations are fundamental in computer graphics for rendering 3D scenes.

Typical pipeline: 1. Model transformation (place objects in scene) 2. View transformation (position camera) 3. Projection transformation (create 2D image) 4. Viewport transformation (fit to screen)

Homogeneous coordinates: To handle translations (which aren't linear), we use 4D vectors:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Then translation by (t_x, t_y, t_z) becomes multiplication by:

$$\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

6.5.2 Differential Equations

The derivative operator $D : P_n \to P_{n-1}$ defined by $D(p) = p'$ is a linear transformation!

Linearity: $(af + bg)' = af' + bg'$

This is why superposition works in differential equations: if y_1 and y_2 are solutions, so is $c_1y_1 + c_2y_2$.

6.5.3 Data Transformations

In data science, linear transformations include:

- Standardization (centering and scaling data)
- Principal Component Analysis (PCA) for dimension reduction
- Feature transformations

6.6 Chapter Summary

In this chapter, we connected abstract vector spaces with concrete matrix operations:

- **Linear transformations** preserve vector space structure: $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ and $T(c\vec{v}) = cT(\vec{v})$
- Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ has a **standard matrix** A
- **Composition** of transformations corresponds to matrix multiplication
- **Kernel**: vectors mapping to zero (generalizes null space)
- **Range**: all possible outputs (generalizes column space)
- **One-to-one**: $\ker(T) = \{\vec{0}\}$ (no information lost)
- **Onto**: $\text{range}(T) = W$ (all outputs achievable)
- **Isomorphisms**: one-to-one and onto (spaces are "the same")
- Applications in graphics, differential equations, and data science

Looking Ahead

In Chapter 7, we'll study eigenvalues and eigenvectors—special vectors that linear transformations simply scale rather than rotate. These provide insight into the fundamental behavior of transformations and have applications from Google's PageRank to quantum mechanics!

6.7 Practice Problems

6.7.1 Basic Problems

1. Determine if $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ x - y \end{bmatrix}$ is linear.

2. Find the standard matrix for $T : \mathbb{R}^2 \to \mathbb{R}^3$ where:

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + 2y \\ 3x \\ y - x \end{bmatrix}$$

3. Find the matrix for rotation by 180ř in $\mathbb{R}^2$.

4. What transformation does the matrix $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ represent?

5. Find $\ker(T)$ where $T : \mathbb{R}^3 \to \mathbb{R}$ is defined by:

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = x + 2y - z$$

6. Find the range of $T : \mathbb{R}^2 \to \mathbb{R}^2$ with matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

7. Is $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$ one-to-one? Onto?

8. If $T : \mathbb{R}^3 \to \mathbb{R}^2$ with $\dim(\ker(T)) = 1$, find $\dim(\text{range}(T))$.

6.7.2 Intermediate Problems

9. Let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation by 45ř and $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection across the x-axis. Find the matrix for $T_1 \circ T_2$ and $T_2 \circ T_1$.

10. Find a basis for $\ker(T)$ and $\text{range}(T)$ where:

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + y \\ 2x + 2y \\ x + y - z \end{bmatrix}$$

11. Prove that $T : P_2 \to P_2$ defined by $T(p(x)) = p(x) + p'(x)$ is linear.

12. Show that the derivative operator $D : P_3 \to P_2$ defined by $D(p) = p'$ is onto but not one-to-one.

13. Find all linear transformations $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2\\3\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-1\\4\end{bmatrix}$.

14. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ have matrix $A = \begin{bmatrix}1 & 2 & 3\\0 & 1 & 4\\0 & 0 & 1\end{bmatrix}$. Is T an isomorphism?

15. Find the matrix for projection onto the line $y = 2x$ in $\mathbb{R}^2$.

6.7.3 Challenge Problems

16. Prove that if $T : V \to W$ is linear and one-to-one, then T maps linearly independent sets to linearly independent sets.

17. Let $T : V \to W$ be linear. Prove that T is one-to-one if and only if the only solution to $T(\vec{v}) = \vec{0}$ is $\vec{v} = \vec{0}$.

18. Prove the Rank-Nullity Theorem: If $T : V \to W$ is linear with $\dim(V) = n$, then:

$$\dim(\ker(T)) + \dim(\text{range}(T)) = n$$

19. Show that composition of linear transformations is associative: $(T_3 \circ T_2) \circ T_1 = T_3 \circ (T_2 \circ T_1)$.

20. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear. Prove that T is onto if and only if T is one-to-one.

21. Find the matrix for reflection across the line $y = mx$ in $\mathbb{R}^2$.

22. Prove that if V and W are finite-dimensional vector spaces with $\dim(V) = \dim(W)$, then V and W are isomorphic.

23. Let $T : V \to W$ be an isomorphism. Prove that $T^{-1} : W \to V$ is also a linear transformation.

7

Eigenvalues and Eigenvectors

"'Eigenvalues are the DNA of a matrix."

— Gilbert Strang

Chapter Overview

Imagine you have a transformation that stretches, rotates, and skews space in complicated ways. Are there any special directions that remain unchanged—directions that only get stretched or compressed, but don't rotate? These special directions are called eigenvectors, and the amount they get stretched by is called eigenvalues. Despite their seemingly abstract nature, eigenvalues and eigenvectors are arguably the most important concepts in applied linear algebra. They appear in Google's PageRank algorithm, facial recognition software, vibrating bridges, population dynamics, quantum mechanics, and countless other applications. In this chapter, we'll discover what makes these quantities so special and why they deserve to be called the "DNA of a matrix."

7.1 Introduction to Eigenvalues and Eigenvectors

7.1.1 Motivation: Special Directions

When we apply a linear transformation to a vector, the result is usually a vector pointing in a completely different direction. However, for certain special vectors, the transformation only stretches or compresses them—their direction remains unchanged.

Notice how $\vec{v}_1$ and $\vec{v}_2$ only change in length (they stretch along their original direction), while $\vec{w}$ changes direction entirely. The vectors $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of the transformation.

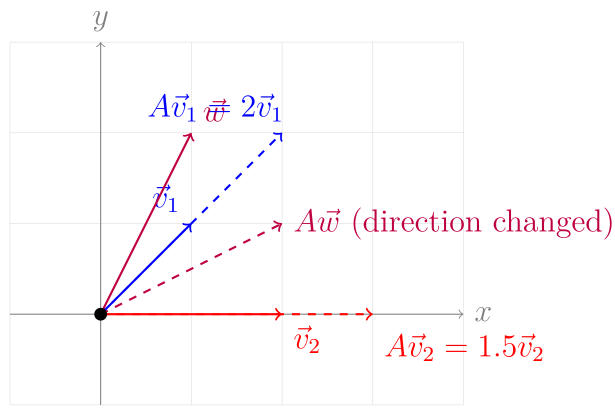


Figure 7.1: Some vectors maintain their direction under transformation A, while others don't

7.1.2 Defining Eigenvalues and Eigenvectors

Definition: Eigenvalue and Eigenvector

Let A be an $n \times n$ matrix. A nonzero vector $\vec{v}$ is called an **eigenvector** of A if there exists a scalar λ such that:

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

The scalar λ is called an **eigenvalue** of A corresponding to the eigenvector $\vec{v}$.

Let's unpack this definition:

- When we multiply matrix A by eigenvector $\vec{v}$, we get back a multiple of $\vec{v}$
- The eigenvector $\vec{v}$ must be nonzero (otherwise every λ would work!)
- The eigenvalue λ tells us the *scale factor*—how much $\vec{v}$ is stretched or compressed
- If $\lambda > 1$: the vector stretches
- If $0 < \lambda < 1$: the vector compresses
- If $\lambda < 0$: the vector reverses direction
- If $\lambda = 0$: the vector collapses to the zero vector

Example 7.1.1. *Verifying an Eigenvector Let* $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ *and* $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. *Verify that* $\vec{v}$ *is an eigenvector and find the corresponding eigenvalue.*

Solution: *We compute* $A\vec{v}$*:*

$$A\vec{v} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 3\vec{v}$$

Since $A\vec{v} = 3\vec{v}$, the vector $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 3$.

Insight: Geometric Interpretation

Eigenvectors represent directions that are preserved by the linear transformation. The eigenvalue tells you how much stretching or compression occurs along that direction. This makes eigenvectors the "natural axes" for understanding what a transformation really does.

7.1.3 The Eigenspace

If $\vec{v}$ is an eigenvector with eigenvalue λ, then so is any nonzero scalar multiple $c\vec{v}$ (as long as $c \neq 0$). This means eigenvectors come in families.

Definition: Eigenspace

The **eigenspace** corresponding to eigenvalue λ is the set of all eigenvectors with eigenvalue λ, together with the zero vector. It is denoted E_λ and consists of all vectors $\vec{v}$ satisfying:

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

Equivalently, $E_\lambda = \text{Null}(A - \lambda I)$.

The eigenspace is actually a subspace (it's closed under addition and scalar multiplication), which is why we can call it a "space."

7.1.4 A Simple Example

Example 7.1.2. *Finding Eigenvalues and Eigenvectors by Inspection Consider the diagonal matrix $D = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix}$. Find its eigenvalues and eigenvectors.*

Solution: *For diagonal matrices, this is straightforward! Let's try the standard basis vectors:*

For $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$:

$$D\vec{v}_1 = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} = 5\vec{v}_1$$

For $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$*:*

$$D\vec{v}_2 = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix} = -2\vec{v}_2$$

Therefore:

- $\lambda_1 = 5$ *with eigenvector* $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
- $\lambda_2 = -2$ *with eigenvector* $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Insight: Diagonal Matrices

For diagonal matrices, the eigenvalues are simply the diagonal entries, and the eigenvectors are the standard basis vectors. This is one reason why diagonalization (Section 7.3) is so powerful—if we can represent a transformation using a diagonal matrix, everything becomes simple!

7.2 Finding Eigenvalues and Eigenvectors

7.2.1 The Characteristic Equation

For most matrices, we can't just guess the eigenvalues. We need a systematic method. Let's start with the defining equation:

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

Rearranging:

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

We can't factor out $\vec{v}$ directly because A is a matrix and λ is a scalar. However, we can write:

$$A\vec{v} - \lambda I\vec{v} = \vec{0}$$

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

This is a homogeneous system! For a nonzero solution $\vec{v}$ to exist, the matrix $(A - \lambda I)$ must be singular (non-invertible), which means:

The Characteristic Equation

The eigenvalues of matrix A are the solutions to:

$$\det(A - \lambda I) = 0$$

This equation is called the **characteristic equation** of A.

7.2.2 The Characteristic Polynomial

When we expand $\det(A - \lambda I)$, we get a polynomial in λ.

Definition: Characteristic Polynomial

The **characteristic polynomial** of an $n \times n$ matrix A is:

$$p(\lambda) = \det(A - \lambda I)$$

This is a polynomial of degree n in the variable λ.

Example 7.2.1. *Finding Eigenvalues of a 2×2 Matrix Find the eigenvalues of* $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$.

Solution: *First, form* $A - \lambda I$*:*

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

Now compute the determinant:

$$\begin{aligned} \det(A - \lambda I) &= (4-\lambda)(3-\lambda) - (1)(2) \\ &= 12 - 4\lambda - 3\lambda + \lambda^2 - 2 \\ &= \lambda^2 - 7\lambda + 10 \end{aligned}$$

Set the characteristic polynomial equal to zero:

$$\lambda^2 - 7\lambda + 10 = 0$$

Factor:

$$(\lambda - 5)(\lambda - 2) = 0$$

Therefore, the eigenvalues are $\lambda_1 = 5$ *and* $\lambda_2 = 2$.

7.2.3 Finding Eigenvectors

Once we have the eigenvalues, we find the corresponding eigenvectors by solving $(A-\lambda I)\vec{v} = \vec{0}$.

Example 7.2.2. *Finding Eigenvectors Find the eigenvectors corresponding to the eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 2$ from the previous example.*

Solution:

For $\lambda_1 = 5$***:***

Solve $(A - 5I)\vec{v} = \vec{0}$*:*

$$\begin{bmatrix} 4-5 & 1 \\ 2 & 3-5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives us the system:

$$-v_1 + v_2 = 0$$
$$2v_1 - 2v_2 = 0$$

Both equations simplify to $v_2 = v_1$*. Choosing* $v_1 = 1$*, we get:*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 2$***:***

Solve $(A - 2I)\vec{v} = \vec{0}$*:*

$$\begin{bmatrix} 4-2 & 1 \\ 2 & 3-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives us:

$$2v_1 + v_2 = 0$$
$$2v_1 + v_2 = 0$$

So $v_2 = -2v_1$*. Choosing* $v_1 = 1$*, we get:*

$$\vec{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Therefore:

- $\lambda_1 = 5$ *has eigenvector* $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *(or any scalar multiple)*
- $\lambda_2 = 2$ *has eigenvector* $\vec{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ *(or any scalar multiple)*

Common Mistake: Forgetting Scalar Multiples

Remember that eigenvectors are not unique! If $\vec{v}$ is an eigenvector, so is $c\vec{v}$ for any nonzero scalar c. Different textbooks or calculators might give different eigenvectors for the same eigenvalue—they're all correct as long as they're scalar multiples of each other.

7.2.4 The Process Summarized

Algorithm: Finding Eigenvalues and Eigenvectors

To find eigenvalues and eigenvectors of an $n \times n$ matrix A:
Step 1: Compute the characteristic polynomial $\det(A - \lambda I)$.
Step 2: Solve $\det(A - \lambda I) = 0$ to find eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.
Step 3: For each eigenvalue λ_i, solve the homogeneous system $(A - \lambda_i I)\vec{v} = \vec{0}$ to find corresponding eigenvectors.

7.2.5 A 3×3 Example

Example 7.2.3. *Eigenvalues of a 3×3 Matrix Find the eigenvalues of* $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Solution: *Form* $A - \lambda I$*:*

$$A - \lambda I = \begin{bmatrix} 2-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{bmatrix}$$

Since this is upper triangular, the determinant is the product of diagonal entries:

$$\det(A - \lambda I) = (2-\lambda)(2-\lambda)(3-\lambda) = (2-\lambda)^2(3-\lambda)$$

Setting this equal to zero:

$$(2-\lambda)^2(3-\lambda) = 0$$

The eigenvalues are $\lambda_1 = 2$ *(with multiplicity 2) and* $\lambda_2 = 3$ *(with multiplicity 1).*

Insight: Triangular Matrices

For triangular matrices (upper or lower), the eigenvalues are simply the diagonal entries! This makes finding eigenvalues much easier and is another reason why certain matrix forms are so useful.

7.2.6 Algebraic and Geometric Multiplicity

When an eigenvalue appears as a repeated root of the characteristic polynomial, we need to distinguish between two types of multiplicity.

Definition: Multiplicity

Let λ be an eigenvalue of matrix A.

- The **algebraic multiplicity** of λ is the number of times λ appears as a root of the characteristic polynomial.
- The **geometric multiplicity** of λ is the dimension of the eigenspace E_λ, which equals the number of linearly independent eigenvectors corresponding to λ.

Important Relationship

For any eigenvalue λ:

$$1 \leq \text{geometric multiplicity} \leq \text{algebraic multiplicity}$$

The geometric multiplicity is always at least 1 (there's at least one eigenvector), and it never exceeds the algebraic multiplicity.

Example 7.2.4. *Different Multiplicities Consider* $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$. *Find the eigenvalues and their multiplicities.*

Solution: *The characteristic polynomial is:*

$$\det(A - \lambda I) = (2 - \lambda)^2 = 0$$

So $\lambda = 2$ *with algebraic multiplicity 2.*

To find the geometric multiplicity, solve $(A - 2I)\vec{v} = \vec{0}$*:*

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives $v_2 = 0$ and v_1 is free. The eigenspace is one-dimensional:

$$E_2 = span\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$$

Therefore, the geometric multiplicity is 1.

For this matrix: algebraic multiplicity = 2, geometric multiplicity = 1.

7.3 Diagonalization

7.3.1 What is Diagonalization?

Diagonal matrices are wonderful to work with—their eigenvalues are obvious, and computing powers is trivial. The question is: can we represent other matrices in diagonal form?

Definition: Diagonalizable Matrix

An $n \times n$ matrix A is **diagonalizable** if there exists an invertible matrix P and a diagonal matrix D such that:

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

Equivalently:

$$P^{-1}AP = D$$

We say that P **diagonalizes** A.

The matrix P performs a change of basis, D represents the transformation in the new basis, and P^{-1} converts back.

7.3.2 The Diagonalization Theorem

Theorem: Diagonalization

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

When A is diagonalizable:

- The columns of P are eigenvectors of A
- The diagonal entries of D are the corresponding eigenvalues of A

Specifically, if $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are linearly independent eigenvectors with eigenvalues

$\lambda_1, \lambda_2, \ldots, \lambda_n$, then:

$$P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

7.3.3 How to Diagonalize a Matrix

Algorithm: Diagonalizing a Matrix

To diagonalize an $n \times n$ matrix A:
Step 1: Find all eigenvalues of A.
Step 2: Find n linearly independent eigenvectors of A.

- If you can't find n linearly independent eigenvectors, A is not diagonalizable.

Step 3: Form P using the eigenvectors as columns.
Step 4: Form D using the corresponding eigenvalues on the diagonal.
Step 5: Verify that $AP = PD$ (optional but recommended).

Example 7.3.1. *Diagonalizing a 2×2 Matrix Diagonalize* $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$.

Solution: *From our earlier work, we know:*

- $\lambda_1 = 5$ *with eigenvector* $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- $\lambda_2 = 2$ *with eigenvector* $\vec{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

Since we have 2 linearly independent eigenvectors for a 2×2 *matrix,* A *is diagonalizable.*

Form P *and* D*:*

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

Let's verify: First compute P^{-1}*:*

$$P^{-1} = \frac{1}{-3} \begin{bmatrix} -2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}$$

Now check PDP^{-1}:

$$PDP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} = A \quad \checkmark$$

7.3.4 When is a Matrix Diagonalizable?

Not all matrices are diagonalizable. Here are some useful criteria:

Sufficient Conditions for Diagonalizability

1. If A is an $n \times n$ matrix with n **distinct** eigenvalues, then A is diagonalizable.
2. If A is a **symmetric** matrix ($A^T = A$), then A is diagonalizable (we'll explore this more in Section 7.5).

Example 7.3.2. *Non-Diagonalizable Matrix Show that $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ is not diagonalizable.*

Solution: *From an earlier example, we know A has eigenvalue $\lambda = 2$ with algebraic multiplicity 2 but geometric multiplicity 1. This means there is only one linearly independent eigenvector.*

Since we need 2 linearly independent eigenvectors to diagonalize a 2×2 matrix, A is not diagonalizable.

Common Mistake

Having repeated eigenvalues does NOT automatically mean a matrix is non-diagonalizable. What matters is whether there are enough linearly independent eigenvectors. For example, the identity matrix I has only one eigenvalue ($\lambda = 1$), but it's already diagonal!

7.3.5 Computing Powers of Matrices

One of the most powerful applications of diagonalization is computing matrix powers efficiently.

Why Diagonalization Helps

If $A = PDP^{-1}$, then:

$$A^k = (PDP^{-1})^k = PD^kP^{-1}$$

And computing D^k for a diagonal matrix is trivial:

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}$$

Example 7.3.3. *Computing A^{10} Let $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$. Compute A^{10}.*

Solution: *From our diagonalization:*

$$A = PDP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}$$

Therefore:

$$A^{10} = PD^{10}P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 5^{10} & 0 \\ 0 & 2^{10} \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}$$

Computing:

$$5^{10} = 9{,}765{,}625, \quad 2^{10} = 1{,}024$$

$$D^{10} = \begin{bmatrix} 9{,}765{,}625 & 0 \\ 0 & 1{,}024 \end{bmatrix}$$

$$\begin{aligned} A^{10} &= \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 9{,}765{,}625 & 0 \\ 0 & 1{,}024 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix} \\ &= \begin{bmatrix} 6{,}510{,}759 & 3{,}255{,}034 \\ 6{,}510{,}068 & 3{,}255{,}375 \end{bmatrix} \end{aligned}$$

Try computing this by multiplying A ten times—diagonalization is much more efficient!

7.4 Applications of Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors appear in an astonishing variety of applications. Let's explore some of the most important ones.

7.4.1 Google PageRank Algorithm

One of the most famous applications of eigenvalues is Google's PageRank algorithm, which ranks web pages in search results.

The Idea: A web page is important if many important pages link to it. This circular definition is resolved using eigenvectors!

We create a matrix where entry a_{ij} represents the probability of moving from page j to page i. The importance scores form the dominant eigenvector (the eigenvector corresponding to the largest eigenvalue, which is $\lambda = 1$ for these matrices).

Example 7.4.1. *Simplified PageRank Consider 3 web pages with the following link structure:*

- *Page 1 links to Pages 2 and 3*
- *Page 2 links to Page 3*
- *Page 3 links to Page 1*

The transition matrix (assuming equal probability of following each link) is:

$$M = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{bmatrix}$$

The PageRank vector $\vec{r}$ satisfies $M\vec{r} = \vec{r}$, meaning $\vec{r}$ is an eigenvector with eigenvalue $\lambda = 1$.

Solving $(M - I)\vec{r} = \vec{0}$ gives (after normalization):

$$\vec{r} = \begin{bmatrix} 0.4 \\ 0.2 \\ 0.4 \end{bmatrix}$$

This tells us Pages 1 and 3 are equally important (40

Connection: Markov Chains

PageRank is an example of a Markov chain—a system that transitions between states with given probabilities. The long-term behavior is determined by the eigenvector corresponding to eigenvalue 1, called the steady-state vector.

7.4.2 Population Dynamics

Eigenvalues help us understand long-term population behavior in ecology.

Example 7.4.2. *Leslie Matrix Model Consider a population divided into age groups. A Leslie matrix L describes how the population changes each year:*

$$L = \begin{bmatrix} 0 & 1.5 & 0.8 \\ 0.5 & 0 & 0 \\ 0 & 0.6 & 0 \end{bmatrix}$$

where entry ℓ_{ij} represents the contribution of age group j to age group i in the next time period.

The population vector at time t is $\vec{p}_t = L^t \vec{p}_0$. The long-term growth rate is determined by the largest eigenvalue!

For this matrix, the dominant eigenvalue is approximately $\lambda \approx 0.93$, which is less than 1. This means the population is declining at about 7% per year.

7.4.3 Vibration Analysis and Natural Frequencies

In physics and engineering, eigenvalues determine the natural frequencies of vibrating systems.

The Scenario: Consider a system of masses connected by springs. The motion is governed by:

$$M\ddot{\vec{x}} = -K\vec{x}$$

where M is the mass matrix and K is the stiffness matrix.

Looking for solutions of the form $\vec{x}(t) = \vec{v}\cos(\omega t)$, we get:

$$K\vec{v} = \omega^2 M\vec{v}$$

This is a generalized eigenvalue problem! The eigenvalues ω^2 give the squared natural frequencies, and the eigenvectors describe the modes of vibration.

Example 7.4.3. *Two-Mass System Two equal masses connected by three identical springs have the system:*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ddot{\vec{x}} = -\begin{bmatrix} 2k & -k \\ -k & 2k \end{bmatrix} \vec{x}$$

The eigenvalues of $K = \begin{bmatrix} 2k & -k \\ -k & 2k \end{bmatrix}$ are $\lambda_1 = k$ and $\lambda_2 = 3k$.

The natural frequencies are:

$$\omega_1 = \sqrt{k}, \quad \omega_2 = \sqrt{3k}$$

The corresponding eigenvectors show:

- *Mode 1: Both masses move together (in phase)*
- *Mode 2: Masses move in opposite directions (out of phase)*

Real-World Impact

Understanding natural frequencies is crucial in engineering. The Tacoma Narrows Bridge collapsed in 1940 because wind excited one of its natural frequencies. Engineers now use eigenvalue analysis to ensure bridges, buildings, and aircraft avoid dangerous resonances.

7.4.4 Principal Component Analysis (PCA)

In data science and machine learning, PCA uses eigenvalues to reduce the dimensionality of data while preserving the most important information.

The Idea: Given a dataset with many variables, PCA finds new axes (principal components) that capture the most variance in the data. These axes are the eigenvectors of the covariance matrix!

Example 7.4.4. *PCA in 2D Suppose we have data about students' math and physics scores. The scores are highly correlated (good math students tend to be good at physics). The covariance matrix might be:*

$$C = \begin{bmatrix} 100 & 80 \\ 80 & 100 \end{bmatrix}$$

The eigenvalues are $\lambda_1 = 180$ *and* $\lambda_2 = 20$.

The eigenvector for λ_1 *points in the direction of maximum variance—roughly along the line where math and physics scores are equal. This single principal component captures* $180/(180 + 20) = 90\%$ *of the variance, allowing us to represent each student with essentially one number instead of two!*

Modern Applications

PCA is used everywhere in modern data science:

- Image compression (representing images with fewer variables)
- Facial recognition (finding key features in faces)
- Genomics (identifying patterns in gene expression)

- Finance (finding common factors driving stock prices)

7.4.5 Differential Equations

Eigenvalues are crucial for solving systems of differential equations.

Consider the system:

$$\frac{d\vec{x}}{dt} = A\vec{x}$$

If A is diagonalizable with $A = PDP^{-1}$, then the solution is:

$$\vec{x}(t) = Pe^{Dt}P^{-1}\vec{x}_0$$

where e^{Dt} is the diagonal matrix with entries $e^{\lambda_i t}$.

Example 7.4.5. *Coupled Differential Equations Solve the system:*

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Solution: *The eigenvalues of* $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ *are* $\lambda_1 = -1$ *and* $\lambda_2 = -3$.

The eigenvectors are $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

The general solution is:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = c_1 e^{-t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3t}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Using the initial condition:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

This gives $c_1 = 1/2$ *and* $c_2 = 1/2$, *so:*

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \frac{1}{2}e^{-t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}e^{-3t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(e^{-t} + e^{-3t}) \\ \frac{1}{2}(e^{-t} - e^{-3t}) \end{bmatrix}$$

Stability Analysis

The eigenvalues tell us about stability:

- If all eigenvalues have negative real parts: the system is stable (solutions decay to zero)

- If any eigenvalue has positive real part: the system is unstable (solutions grow without bound)
- If eigenvalues are purely imaginary: the system oscillates

7.5 Symmetric Matrices and Orthogonal Diagonalization

7.5.1 Special Properties of Symmetric Matrices

Symmetric matrices (where $A^T = A$) have remarkable properties that make them especially important in applications.

Theorem: Eigenvalues of Symmetric Matrices

If A is a real symmetric matrix, then:

1. All eigenvalues of A are real numbers
2. Eigenvectors corresponding to distinct eigenvalues are orthogonal
3. A is always diagonalizable

Example 7.5.1. *Symmetric Matrix Let* $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. *Find its eigenvalues and eigenvectors.*

Solution: *Characteristic equation:*

$$\det(A - \lambda I) = (3 - \lambda)^2 - 1 = \lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2) = 0$$

Eigenvalues: $\lambda_1 = 4$ *and* $\lambda_2 = 2$ *(both real* ✓*)*

For $\lambda_1 = 4$*:*

$$(A - 4I)\vec{v} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \vec{0} \implies \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 2$*:*

$$(A - 2I)\vec{v} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \vec{0} \implies \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Check orthogonality:

$$\vec{v}_1 \cdot \vec{v}_2 = (1)(1) + (1)(-1) = 0 \quad ✓$$

7.5.2 Orthogonal Matrices

Definition: Orthogonal Matrix

A square matrix Q is **orthogonal** if its columns form an orthonormal set, which is equivalent to:

$$Q^T Q = I \quad \text{or equivalently} \quad Q^{-1} = Q^T$$

Orthogonal matrices represent rotations and reflections—they preserve lengths and angles.

7.5.3 Orthogonal Diagonalization

Theorem: Spectral Theorem (Informal)

Every real symmetric matrix A can be orthogonally diagonalized. That is, there exists an orthogonal matrix Q and a diagonal matrix D such that:

$$A = QDQ^T$$

The columns of Q are orthonormal eigenvectors of A, and the diagonal entries of D are the corresponding eigenvalues.

This is called the **spectral decomposition** of A.

Example 7.5.2. *Orthogonal Diagonalization Orthogonally diagonalize* $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$.

Solution: *From the previous example, we have eigenvectors* $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

These are orthogonal but not unit vectors. Normalize them:

$$\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Form Q and D:

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

Verify:

$$Q^T = Q^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$QDQ^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = A \quad \checkmark$$

7.5.4 Why Symmetric Matrices Matter

Symmetric matrices appear constantly in applications:

- **Covariance matrices** in statistics (used in PCA)
- **Moment of inertia tensors** in physics
- **Hessian matrices** in optimization (second derivatives)
- **Graph Laplacian matrices** in network analysis
- **Stress and strain tensors** in engineering

The fact that they can be orthogonally diagonalized means we can always find natural coordinate systems in which these problems become simple.

Geometric Interpretation

Orthogonal diagonalization reveals the principal axes of a quadratic form. For example, the equation $3x^2 + 2xy + 3y^2 = 1$ describes an ellipse in a tilted coordinate system. The eigenvectors point along the major and minor axes of the ellipse!

7.6 Chapter Summary

In this chapter, we've explored eigenvalues and eigenvectors—arguably the most important concepts in linear algebra:

- **Eigenvectors** are special directions preserved by a linear transformation
- **Eigenvalues** measure the scaling factor along those directions
- The **characteristic equation** $\det(A - \lambda I) = 0$ gives us the eigenvalues
- A matrix is **diagonalizable** if it has n linearly independent eigenvectors
- **Diagonalization** makes computing matrix powers and understanding transformations much easier
- **Applications** are everywhere: PageRank, population dynamics, vibration analysis, PCA, differential equations, and more
- **Symmetric matrices** have special properties: real eigenvalues, orthogonal eigenvectors, and guaranteed diagonalizability

Looking Ahead

In the next chapter, we'll generalize the dot product to create inner product spaces. This will allow us to extend concepts like length, angle, and orthogonality to more abstract settings, and we'll develop powerful tools like the Gram-Schmidt process and least squares approximation. The eigenvectors we found in this chapter for symmetric matrices are actually orthogonal with respect to an inner product!

7.7 Practice Problems

7.7.1 Basic Problems

1. Determine if $\vec{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$. If so, find the eigenvalue.

2. Find the eigenvalues of each matrix:

   (a) $A = \begin{bmatrix} 5 & 0 \\ 0 & -3 \end{bmatrix}$

   (b) $B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$

   (c) $C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

3. For $A = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$:

   (a) Find the characteristic polynomial

   (b) Find all eigenvalues

   (c) Find an eigenvector for each eigenvalue

4. Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$.

5. Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. Find the eigenvalues and their algebraic multiplicities.

6. Determine if $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is diagonalizable. If so, find matrices P and D such that $A = PDP^{-1}$.

7. If $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, compute A^{10} using diagonalization.

8. Show that if λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2.

7.7.2 Intermediate Problems

9. For $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, find all eigenvalues and a basis for each eigenspace.

10. Diagonalize $A = \begin{bmatrix} 7 & 2 \\ -2 & 3 \end{bmatrix}$ and use it to compute A^5.

11. Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Find the eigenvalues and show that A is diagonalizable.

12. Prove that if A is invertible and λ is an eigenvalue of A, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

13. Show that the trace of a matrix (sum of diagonal entries) equals the sum of its eigenvalues, and the determinant equals the product of its eigenvalues.

14. Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ and orthogonally diagonalize it.

15. Suppose A has eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -1$ with corresponding eigenvectors $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Find $A^{100} \begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

16. Solve the system of differential equations:

$$\frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

7.7.3 Challenge Problems

17. Prove that eigenvectors corresponding to distinct eigenvalues are linearly independent.

18. Let A be a 3×3 matrix with eigenvalues $2, 3, 5$. What are the possible values of $\det(A)$ and $\text{tr}(A)$?

19. Prove that if A is a real symmetric matrix, then eigenvectors corresponding to distinct eigenvalues are orthogonal.

20. A matrix A satisfies $A^2 = A$ (called idempotent). Show that the only possible eigenvalues are 0 and 1.

21. Consider the Fibonacci matrix $F = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Diagonalize F and use it to derive a formula for the nth Fibonacci number.

22. Find the eigenvalues of the circulant matrix:

$$C = \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}$$

23. Suppose A is a 5×5 matrix with characteristic polynomial $p(\lambda) = (\lambda - 2)^3(\lambda + 1)^2$. What can you conclude about the diagonalizability of A?

24. A population is modeled by $\vec{p}_{n+1} = A\vec{p}_n$ where $A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$ and $\vec{p}_n$ represents the population distribution between two locations. Find the long-term steady-state distribution.

25. Prove the Cayley-Hamilton theorem for 2×2 matrices: every matrix satisfies its own characteristic equation. That is, if $p(\lambda) = \det(A - \lambda I)$, then $p(A) = 0$.

26. For the quadratic form $Q(x, y) = 5x^2 + 4xy + 5y^2$, find the matrix A such that $Q(x, y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix}$. Orthogonally diagonalize A and identify the type of conic section described by $Q(x, y) = 1$.

8

Inner Product Spaces

“‘The introduction of the cipher 0 or the group concept was general nonsense too, and mathematics was more or less stagnating for thousands of years because nobody was around to take such childish steps...”

— Alexander Grothendieck

Chapter Overview

You’ve been using the dot product since Chapter 1—multiplying corresponding components and adding them up to get a single number. But what makes the dot product so special? Why does it tell us about angles and lengths? In this chapter, we’ll discover that the dot product is just one example of a more general concept called an *inner product*. By abstracting the essential properties of the dot product, we can define notions of length, distance, and angle in spaces far beyond $\mathbb{R}^n$—including spaces of polynomials, functions, and matrices. This generalization leads to powerful tools like the Gram-Schmidt process for creating orthogonal bases and the method of least squares for fitting data. These techniques are fundamental to everything from signal processing to machine learning.

8.1 Inner Products

8.1.1 Generalizing the Dot Product

Recall that the dot product in $\mathbb{R}^n$ is defined as:

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

This operation has several important properties:

- It’s always a real number (a scalar)

- $\vec{v} \cdot \vec{v} \geq 0$, and equals zero only when $\vec{v} = \vec{0}$
- It's symmetric: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- It's linear: $\vec{u} \cdot (c\vec{v} + \vec{w}) = c(\vec{u} \cdot \vec{v}) + \vec{u} \cdot \vec{w}$

These properties are so useful that we want to extend them to other vector spaces.

Definition: Inner Product

Let V be a vector space. An **inner product** on V is a function that assigns to each pair of vectors $\vec{u}, \vec{v} \in V$ a real number $\langle \vec{u}, \vec{v} \rangle$ satisfying:

1. **Positivity:** $\langle \vec{v}, \vec{v} \rangle \geq 0$ for all $\vec{v} \in V$
2. **Definiteness:** $\langle \vec{v}, \vec{v} \rangle = 0$ if and only if $\vec{v} = \vec{0}$
3. **Symmetry:** $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$ for all $\vec{u}, \vec{v} \in V$
4. **Linearity:** $\langle \vec{u}, c\vec{v} + \vec{w} \rangle = c\langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$ for all $\vec{u}, \vec{v}, \vec{w} \in V$ and scalar c

A vector space with an inner product is called an **inner product space**.

Notation Note

We use angle brackets $\langle \vec{u}, \vec{v} \rangle$ to denote a general inner product. When working specifically with the dot product in $\mathbb{R}^n$, we may still write $\vec{u} \cdot \vec{v}$. The dot product is the standard inner product on $\mathbb{R}^n$.

8.1.2 Examples of Inner Products

Example 8.1.1. *The Dot Product in $\mathbb{R}^n$ The standard dot product is an inner product:*

$$\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Let's verify one property (positivity):

$$\langle \vec{v}, \vec{v} \rangle = v_1^2 + v_2^2 + \cdots + v_n^2 \geq 0$$

since it's a sum of squares. The other properties can be verified similarly.

Example 8.1.2. *Weighted Inner Product in $\mathbb{R}^n$ For positive constants $w_1, w_2, \ldots, w_n$, we can define:*

$$\langle \vec{u}, \vec{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n$$

This gives different "weights" to different components. For instance, if $w_1 = 2$, $w_2 = 1$, $w_3 = 3$*:*

$$\langle \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \rangle = 2(1)(2) + 1(2)(1) + 3(3)(0) = 4 + 2 + 0 = 6$$

Example 8.1.3. *Inner Product on Polynomial Space Let* P_2 *be the space of polynomials of degree at most 2. Define:*

$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$$

For example, let $p(x) = x$ *and* $q(x) = x^2$*:*

$$\langle p, q \rangle = \int_0^1 x \cdot x^2\,dx = \int_0^1 x^3\,dx = \left[\frac{x^4}{4}\right]_0^1 = \frac{1}{4}$$

This is indeed an inner product (we can verify all four axioms).

Example 8.1.4. *Inner Product on Function Space Let* $C[a, b]$ *be the space of continuous functions on the interval* $[a, b]$*. Define:*

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx$$

For $f(x) = \sin(x)$ *and* $g(x) = \cos(x)$ *on* $[0, \pi]$*:*

$$\langle f, g \rangle = \int_0^\pi \sin(x)\cos(x)\,dx = \int_0^\pi \frac{1}{2}\sin(2x)\,dx = \left[-\frac{1}{4}\cos(2x)\right]_0^\pi = 0$$

So $\sin(x)$ *and* $\cos(x)$ *are orthogonal on* $[0, \pi]$*!*

Common Mistake: Not All Products Are Inner Products

The function $\langle \vec{u}, \vec{v} \rangle = u_1 v_1$ in $\mathbb{R}^2$ is NOT an inner product because it fails definiteness: $\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rangle = 0$ even though the vector is nonzero.

8.1.3 Norm and Distance

Once we have an inner product, we can define length and distance.

Definition: Norm

The **norm** (or **length**) of a vector $\vec{v}$ in an inner product space is:

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

Definition: Distance

The **distance** between vectors $\vec{u}$ and $\vec{v}$ is:

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{\langle \vec{u} - \vec{v}, \vec{u} - \vec{v} \rangle}$$

Example 8.1.5. *Norm in $\mathbb{R}^3$ With the standard dot product, the norm of $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ is:*

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{1^2 + 2^2 + 2^2} = \sqrt{9} = 3$$

Example 8.1.6. *Norm of a Polynomial Using the inner product $\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$, find the norm of $p(x) = x$.*

Solution:

$$\|p\| = \sqrt{\langle p, p \rangle} = \sqrt{\int_0^1 x^2\,dx} = \sqrt{\left[\frac{x^3}{3}\right]_0^1} = \sqrt{\frac{1}{3}} = \frac{1}{\sqrt{3}}$$

Properties of the Norm

For any vectors $\vec{u}, \vec{v}$ and scalar c:

1. $\|\vec{v}\| \geq 0$, with equality if and only if $\vec{v} = \vec{0}$
2. $\|c\vec{v}\| = |c|\|\vec{v}\|$
3. **Triangle Inequality:** $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

8.1.4 Angle Between Vectors

In $\mathbb{R}^n$, we know that:

$$\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos\theta$$

where θ is the angle between the vectors. This motivates the general definition:

Definition: Angle

The **angle** θ between nonzero vectors $\vec{u}$ and $\vec{v}$ in an inner product space is given by:

$$\cos\theta = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|\|\vec{v}\|}$$

This is well-defined because the Cauchy-Schwarz inequality (which we'll see shortly) guarantees that:

$$-1 \leq \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|\|\vec{v}\|} \leq 1$$

Example 8.1.7. *Angle Between Polynomials Find the angle between $p(x) = 1$ and $q(x) = x$ using $\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx$.*

Solution: *First compute the inner product:*

$$\langle p, q\rangle = \int_0^1 1 \cdot x\,dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}$$

Compute the norms:

$$\|p\| = \sqrt{\int_0^1 1^2\,dx} = \sqrt{1} = 1$$

$$\|q\| = \sqrt{\int_0^1 x^2\,dx} = \sqrt{\frac{1}{3}} = \frac{1}{\sqrt{3}}$$

Therefore:

$$\cos\theta = \frac{1/2}{1 \cdot 1/\sqrt{3}} = \frac{\sqrt{3}}{2}$$

$$\theta = \arccos\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{6} = 30ř$$

8.1.5 The Cauchy-Schwarz Inequality

One of the most important inequalities in mathematics:

Theorem: Cauchy-Schwarz Inequality

For any vectors $\vec{u}, \vec{v}$ in an inner product space:

$$|\langle\vec{u}, \vec{v}\rangle| \leq \|\vec{u}\|\|\vec{v}\|$$

Equality holds if and only if $\vec{u}$ and $\vec{v}$ are linearly dependent (one is a scalar multiple of the other).

Proof Sketch: If $\vec{v} = \vec{0}$, both sides are zero. Otherwise, consider the vector:

$$\vec{w} = \vec{u} - \frac{\langle\vec{u}, \vec{v}\rangle}{\langle\vec{v}, \vec{v}\rangle}\vec{v}$$

By construction, $\vec{w}$ is orthogonal to $\vec{v}$ (we can verify $\langle\vec{w}, \vec{v}\rangle = 0$). Since $\|\vec{w}\|^2 \geq 0$, expanding and simplifying leads to the Cauchy-Schwarz inequality.

Why This Matters

The Cauchy-Schwarz inequality ensures that $\frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|\|\vec{v}\|}$ always lies between -1 and 1, making the angle definition valid. It's also the foundation for proving the triangle inequality and appears throughout analysis, probability, and optimization.

8.2 Orthogonality

8.2.1 Orthogonal and Orthonormal Vectors

Definition: Orthogonal Vectors

Vectors $\vec{u}$ and $\vec{v}$ are **orthogonal** (or **perpendicular**) if:

$$\langle \vec{u}, \vec{v} \rangle = 0$$

We write $\vec{u} \perp \vec{v}$.

Definition: Orthonormal Vectors

A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is **orthonormal** if:

1. Each vector has norm 1: $\|\vec{v}_i\| = 1$ for all i
2. The vectors are mutually orthogonal: $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ for all $i \neq j$

Equivalently: $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta (1 if $i = j$, 0 otherwise).

Example 8.2.1. *Standard Basis in $\mathbb{R}^3$ The standard basis vectors are orthonormal:*

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Each has norm 1, and $\vec{e}_i \cdot \vec{e}_j = 0$ for $i \neq j$.

Example 8.2.2. *Orthonormal Functions The functions $f(x) = \frac{1}{\sqrt{\pi}} \sin(x)$ and $g(x) = \frac{1}{\sqrt{\pi}} \cos(x)$ are orthonormal on $[0, \pi]$ with inner product $\langle f, g \rangle = \int_0^\pi f(x)g(x)\,dx$.*
Check orthogonality:

$$\langle f, g \rangle = \frac{1}{\pi} \int_0^\pi \sin(x)\cos(x)\,dx = \frac{1}{\pi} \cdot 0 = 0 \quad \checkmark$$

Check normalization:

$$\|f\|^2 = \frac{1}{\pi}\int_0^\pi \sin^2(x)\,dx = \frac{1}{\pi}\cdot\frac{\pi}{2} = \frac{1}{2}$$

Wait, this should be 1 for orthonormality. Let me recalculate with the correct normalization constant:

Actually, for $\sin(x)$ *and* $\cos(x)$ *to be orthonormal on* $[0,\pi]$*, we need:*

$$f(x) = \sqrt{\frac{2}{\pi}}\sin(x), \quad g(x) = \sqrt{\frac{2}{\pi}}\cos(x)$$

8.2.2 Orthogonal Sets and Bases

Theorem: Orthogonal Sets Are Linearly Independent

If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is a set of nonzero orthogonal vectors (i.e., $\langle \vec{v}_i, \vec{v}_j\rangle = 0$ for $i \neq j$), then the set is linearly independent.

Proof: Suppose $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$.

Take the inner product of both sides with $\vec{v}_i$:

$$\langle c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k, \vec{v}_i\rangle = \langle \vec{0}, \vec{v}_i\rangle = 0$$

By linearity:

$$c_1\langle \vec{v}_1, \vec{v}_i\rangle + c_2\langle \vec{v}_2, \vec{v}_i\rangle + \cdots + c_k\langle \vec{v}_k, \vec{v}_i\rangle = 0$$

Since the vectors are orthogonal, all inner products are zero except $\langle \vec{v}_i, \vec{v}_i\rangle$:

$$c_i\langle \vec{v}_i, \vec{v}_i\rangle = 0$$

Since $\vec{v}_i \neq \vec{0}$, we have $\langle \vec{v}_i, \vec{v}_i\rangle > 0$, so $c_i = 0$.

This holds for all i, proving linear independence. □

Why Orthogonal Bases Are Wonderful

In an orthogonal basis, finding coordinates is trivial! If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is an orthogonal basis and we want to express $\vec{u} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$, we simply compute:

$$c_i = \frac{\langle \vec{u}, \vec{v}_i\rangle}{\langle \vec{v}_i, \vec{v}_i\rangle}$$

No need to solve systems of equations!

Example 8.2.3. *Coordinates in an Orthogonal Basis Let* $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *(these are orthogonal). Express* $\vec{u} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ *in this basis.*

Solution: *Compute:*

$$c_1 = \frac{\vec{u} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} = \frac{5(1) + 3(1)}{1^2 + 1^2} = \frac{8}{2} = 4$$

$$c_2 = \frac{\vec{u} \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} = \frac{5(1) + 3(-1)}{1^2 + (-1)^2} = \frac{2}{2} = 1$$

Therefore: $\vec{u} = 4\vec{v}_1 + 1\vec{v}_2$.

Verify: $4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ ✓

8.2.3 Orthogonal Complements

Definition: Orthogonal Complement

Let W be a subspace of an inner product space V. The **orthogonal complement** of W, denoted $W^\perp$, is the set of all vectors in V that are orthogonal to every vector in W:

$$W^\perp = \{\vec{v} \in V : \langle \vec{v}, \vec{w} \rangle = 0 \text{ for all } \vec{w} \in W\}$$

Example 8.2.4. *Orthogonal Complement in* $\mathbb{R}^3$ *Let* W *be the xy-plane in* $\mathbb{R}^3$*:* $W = \textit{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$.

The orthogonal complement $W^\perp$ *is the z-axis:* $W^\perp = \textit{span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Every vector in the z-axis is perpendicular to every vector in the xy-plane.

Theorem: Properties of Orthogonal Complements

Let W be a subspace of a finite-dimensional inner product space V. Then:

1. $W^\perp$ is a subspace of V
2. $(W^\perp)^\perp = W$
3. $W \cap W^\perp = \{\vec{0}\}$
4. $\dim(W) + \dim(W^\perp) = \dim(V)$

5. Every vector $\vec{v} \in V$ can be uniquely written as $\vec{v} = \vec{w} + \vec{w}^{\perp}$ where $\vec{w} \in W$ and $\vec{w}^{\perp} \in W^{\perp}$

The last property is called the **orthogonal decomposition theorem**.

8.3 Gram-Schmidt Process

8.3.1 Motivation

We've seen that orthogonal bases are extremely convenient. But how do we create one? Given an arbitrary basis, can we transform it into an orthogonal (or orthonormal) basis for the same subspace?

The answer is yes, using the **Gram-Schmidt process**—one of the most important algorithms in linear algebra.

8.3.2 The Gram-Schmidt Algorithm

Theorem: Gram-Schmidt Process

Let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ be a linearly independent set in an inner product space. The Gram-Schmidt process produces an orthogonal set $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ that spans the same subspace:

$$\begin{aligned}
\vec{u}_1 &= \vec{v}_1 \\
\vec{u}_2 &= \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle}\vec{u}_1 \\
\vec{u}_3 &= \vec{v}_3 - \frac{\langle \vec{v}_3, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle}\vec{u}_1 - \frac{\langle \vec{v}_3, \vec{u}_2 \rangle}{\langle \vec{u}_2, \vec{u}_2 \rangle}\vec{u}_2 \\
&\vdots \\
\vec{u}_k &= \vec{v}_k - \sum_{i=1}^{k-1} \frac{\langle \vec{v}_k, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle}\vec{u}_i
\end{aligned}$$

To obtain an orthonormal set, normalize each vector:

$$\hat{u}_i = \frac{\vec{u}_i}{\|\vec{u}_i\|}$$

Geometric Intuition

At each step, we're subtracting off the projections of $\vec{v}_k$ onto all previous orthogonal vectors $\vec{u}_1, \ldots, \vec{u}_{k-1}$. What remains is the component of $\vec{v}_k$ that's orthogonal to the subspace spanned by $\{\vec{u}_1, \ldots, \vec{u}_{k-1}\}$.

8.3.3 Step-by-Step Example

Example 8.3.1. *Gram-Schmidt in $\mathbb{R}^3$ Apply the Gram-Schmidt process to the basis:*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Solution:

Step 1: *Set* $\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

Step 2: *Compute* $\vec{u}_2$*:*

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1$$

Calculate:

$$\vec{v}_2 \cdot \vec{u}_1 = 1(1) + 0(1) + 1(0) = 1$$

$$\vec{u}_1 \cdot \vec{u}_1 = 1^2 + 1^2 + 0^2 = 2$$

Therefore:

$$\vec{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 - 1/2 \\ 0 - 1/2 \\ 1 - 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$$

(We can multiply by 2 to clear fractions: $\vec{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$*)*

Step 3: *Compute* $\vec{u}_3$*:*

$$\vec{u}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2$$

Calculate:

$$\vec{v}_3 \cdot \vec{u}_1 = 0(1) + 1(1) + 1(0) = 1$$

$$\vec{v}_3 \cdot \vec{u}_2 = 0(1) + 1(-1) + 1(2) = 1$$

$$\vec{u}_2 \cdot \vec{u}_2 = 1^2 + (-1)^2 + 2^2 = 6$$

Therefore:

$$\vec{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/6 \\ -1/6 \\ 2/6 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

(Multiply by 3: $\vec{u}_3 = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}$ *or* $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$*)*

The orthogonal basis is:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

To make it orthonormal, normalize each vector (divide by its norm).

Common Mistake: Using Original Vectors

When computing $\vec{u}_k$, make sure to subtract projections onto the *orthogonal* vectors $\vec{u}_1, \dots, \vec{u}_{k-1}$, NOT the original vectors $\vec{v}_1, \dots, \vec{v}_{k-1}$. The process builds on itself—each new vector must be orthogonal to all previously computed orthogonal vectors.

8.3.4 QR Factorization

An important consequence of Gram-Schmidt is the QR factorization.

Theorem: QR Factorization

Any $m \times n$ matrix A with linearly independent columns can be factored as:

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal columns, and R is an $n \times n$ upper triangular matrix with positive diagonal entries.

The columns of Q are the orthonormalized versions of the columns of A (via Gram-Schmidt), and R records the coefficients from the Gram-Schmidt process.

Example 8.3.2. *QR Factorization Factor* $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ *into QR.*

Solution: *The columns are* $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

Apply Gram-Schmidt:

$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, *so* $\hat{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$

$\|\vec{u}_2\| = \sqrt{1/4 + 1/4 + 1} = \sqrt{3/2}$, *so* $\hat{u}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

Thus:

$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{6} \\ 0 & 2/\sqrt{6} \end{bmatrix}$$

To find R*, use* $R = Q^T A$*:*

$$R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} \end{bmatrix}$$

(The detailed computation is left as an exercise.)

8.4 Orthogonal Projections

8.4.1 Projection onto a Vector

Recall from Chapter 1 that the projection of $\vec{v}$ onto $\vec{u}$ is:

$$\text{proj}_{\vec{u}}(\vec{v}) = \frac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u}$$

With general inner products, this becomes:

$$\text{proj}_{\vec{u}}(\vec{v}) = \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u}$$

8.4.2 Projection onto a Subspace

Definition: Orthogonal Projection onto a Subspace

Let W be a subspace of an inner product space V with orthogonal basis $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$. The **orthogonal projection** of $\vec{v}$ onto W is:

$$\text{proj}_W(\vec{v}) = \sum_{i=1}^{k} \frac{\langle \vec{v}, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle} \vec{u}_i$$

If the basis is orthonormal ($\|\vec{u}_i\| = 1$ for all i), this simplifies to:

$$\text{proj}_W(\vec{v}) = \sum_{i=1}^{k} \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i$$

The projection $\text{proj}_W(\vec{v})$ is the vector in W closest to $\vec{v}$. The vector $\vec{v} - \text{proj}_W(\vec{v})$ is orthogonal to W.

Example 8.4.1. *Projection onto a Plane Let W be the plane spanned by $\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ (the xy-plane). Find the projection of $\vec{v} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$ onto W.*

Solution: *Since $\{\vec{u}_1, \vec{u}_2\}$ is orthonormal:*

$$proj_W(\vec{v}) = (\vec{v} \cdot \vec{u}_1)\vec{u}_1 + (\vec{v} \cdot \vec{u}_2)\vec{u}_2 = 2\vec{u}_1 + 3\vec{u}_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$$

This is exactly the vector $\vec{v}$ with its z-component removed—as expected for projection onto the xy-plane!

8.4.3 Best Approximation

Theorem: Best Approximation Theorem

Let W be a subspace of an inner product space V, and let $\vec{v} \in V$. Then $\text{proj}_W(\vec{v})$ is the unique vector in W that is closest to $\vec{v}$. That is, for any $\vec{w} \in W$ with $\vec{w} \neq \text{proj}_W(\vec{v})$:

$$\|\vec{v} - \text{proj}_W(\vec{v})\| < \|\vec{v} - \vec{w}\|$$

This theorem is the foundation of least squares approximation.

8.5 Least Squares Approximation

8.5.1 The Least Squares Problem

Many real-world problems lead to inconsistent systems $A\vec{x} = \vec{b}$—systems with no exact solution. For example, when fitting a line to data points, there's usually no line passing exactly through all points.

Instead of solving $A\vec{x} = \vec{b}$ exactly, we find the "best" approximate solution $\hat{x}$ that minimizes the error $\|\vec{b} - A\vec{x}\|$.

Definition: Least Squares Solution

A **least squares solution** of $A\vec{x} = \vec{b}$ is a vector $\hat{x}$ that minimizes:

$$\|\vec{b} - A\vec{x}\|$$

over all vectors $\vec{x}$.

Geometric Interpretation

The vector $A\hat{x}$ is the orthogonal projection of $\vec{b}$ onto the column space of A. We're finding the vector in $\mathrm{Col}(A)$ that's closest to $\vec{b}$.

8.5.2 The Normal Equations

Theorem: Normal Equations

The least squares solution $\hat{x}$ satisfies:

$$A^T A\hat{x} = A^T\vec{b}$$

These are called the **normal equations**. If $A^T A$ is invertible (which happens when A has linearly independent columns), then:

$$\hat{x} = (A^T A)^{-1} A^T\vec{b}$$

Why? The error vector $\vec{b} - A\hat{x}$ must be orthogonal to the column space of A. This means:

$$A^T(\vec{b} - A\hat{x}) = \vec{0}$$

Expanding: $A^T\vec{b} - A^T A\hat{x} = \vec{0}$, which gives the normal equations.

8.5.3 Linear Regression: Fitting a Line

Example 8.5.1. *Least Squares Line Fitting Find the line $y = mx + c$ that best fits the data points $(1, 1)$, $(2, 2)$, $(3, 4)$.*

Solution: *We want to solve:*

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

This system is inconsistent (three equations, two unknowns), so we use least squares.

Compute $A^T A$:

$$A^T A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$$

Compute $A^T\vec{b}$:

$$A^T\vec{b} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17 \\ 7 \end{bmatrix}$$

Solve $A^T A\hat{x} = A^T\vec{b}$:

$$\begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 17 \\ 7 \end{bmatrix}$$

Using row reduction or the inverse:

$$(A^T A)^{-1} = \frac{1}{14(3) - 6(6)} \begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix} = \begin{bmatrix} 1/2 & -1 \\ -1 & 7/3 \end{bmatrix}$$

Therefore:

$$\begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 1/2 & -1 \\ -1 & 7/3 \end{bmatrix} \begin{bmatrix} 17 \\ 7 \end{bmatrix} = \begin{bmatrix} 17/2 - 7 \\ -17 + 49/3 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -2/3 \end{bmatrix}$$

The best-fit line is:

$$y = \frac{3}{2}x - \frac{2}{3}$$

Applications of Least Squares

Least squares is everywhere in modern science and engineering:

- **Statistics**: Linear regression, curve fitting
- **Machine Learning**: Training linear models

- **Signal Processing**: Filtering and noise reduction
- **Computer Vision**: Image reconstruction
- **Control Theory**: System identification

8.5.4 Polynomial Fitting

We can fit higher-degree polynomials using the same technique.

Example 8.5.2. *Fitting a Quadratic Find the parabola* $y = ax^2 + bx + c$ *that best fits* $(0, 1)$, $(1, 2)$, $(2, 3)$, $(3, 5)$.

Solution: *Set up the system:*

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}$$

Form and solve the normal equations $A^T A\hat{x} = A^T\vec{b}$ *(computation omitted for brevity).*

The result is approximately $a = 0.2$, $b = 1.1$, $c = 1.0$, *giving:*

$$y \approx 0.2x^2 + 1.1x + 1.0$$

8.6 Chapter Summary

In this chapter, we've extended the concept of the dot product to create inner product spaces:

- An **inner product** generalizes the dot product while preserving key properties (positivity, definiteness, symmetry, linearity)
- Inner products allow us to define **norm** (length), **distance**, and **angle** in abstract vector spaces
- The **Cauchy-Schwarz inequality** is fundamental: $|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\|\|\vec{v}\|$
- **Orthogonal** vectors satisfy $\langle \vec{u}, \vec{v} \rangle = 0$; orthogonal sets are linearly independent
- The **Gram-Schmidt process** transforms any basis into an orthogonal (or orthonormal) basis

- **QR factorization** decomposes a matrix as $A = QR$ where Q has orthonormal columns
- **Orthogonal projection** finds the closest vector in a subspace
- **Least squares** solves inconsistent systems by minimizing error, with applications in data fitting and regression

Looking Ahead

In the next chapter, we'll explore a wide range of applications that bring together everything we've learned. We'll see how eigenvalues, inner products, and projections combine to solve real-world problems in data science, differential equations, computer graphics, and more. Linear algebra is the language of modern applied mathematics!

8.7 Practice Problems

8.7.1 Basic Problems

1. Verify that $\langle \vec{u}, \vec{v} \rangle = 2u_1v_1 + 3u_2v_2$ defines an inner product on $\mathbb{R}^2$.

2. Using the inner product $\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$ on P_2:

   (a) Compute $\langle x, x^2 \rangle$

   (b) Find $\|1 + x\|$

   (c) Are $p(x) = x$ and $q(x) = 1 - 2x$ orthogonal?

3. Show that $\vec{u} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ are orthogonal.

4. Find the angle between $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.

5. Determine if $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is an orthogonal set. Is it orthonormal?

6. Find the projection of $\vec{v} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ onto $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

7. Apply the Gram-Schmidt process to $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}\right\}$.

8. Find an orthonormal basis for the subspace spanned by $\begin{bmatrix}1\\0\\1\end{bmatrix}$ and $\begin{bmatrix}0\\1\\1\end{bmatrix}$.

8.7.2 Intermediate Problems

9. Prove that if $\vec{u}$ and $\vec{v}$ are orthogonal, then $\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$ (Pythagorean theorem).

10. Apply the Gram-Schmidt process to find an orthogonal basis for the column space of:
$$A = \begin{bmatrix}1 & 0 & 1\\1 & 1 & 2\\0 & 1 & 1\end{bmatrix}$$

11. Find the QR factorization of $A = \begin{bmatrix}1 & 1\\1 & 2\\1 & 3\end{bmatrix}$.

12. Let $W = \text{span}\left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}\right\}$. Find the projection of $\vec{v} = \begin{bmatrix}1\\2\\3\\4\end{bmatrix}$ onto W.

13. Show that if $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis for $\mathbb{R}^3$, then for any $\vec{v} \in \mathbb{R}^3$:
$$\|\vec{v}\|^2 = |\langle \vec{v}, \vec{u}_1\rangle|^2 + |\langle \vec{v}, \vec{u}_2\rangle|^2 + |\langle \vec{v}, \vec{u}_3\rangle|^2$$
(This is Parseval's identity.)

14. Find the least squares solution to:
$$\begin{bmatrix}1 & 1\\2 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}2\\3\\4\end{bmatrix}$$

15. Find the best-fit line $y = mx + c$ through the points $(1, 2)$, $(2, 3)$, $(3, 5)$, $(4, 4)$.

16. Using $\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx$, show that $\sin(x)$ and $\cos(x)$ are orthogonal on $[-\pi, \pi]$.

8.7.3 Challenge Problems

17. Prove the Cauchy-Schwarz inequality: for any vectors $\vec{u}, \vec{v}$ in an inner product space, $|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\|\|\vec{v}\|$.

18. Prove the triangle inequality: $\|\vec{u}+\vec{v}\| \leq \|\vec{u}\|+\|\vec{v}\|$ using the Cauchy-Schwarz inequality.

19. Show that the distance function $d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$ satisfies:

    (a) $d(\vec{u}, \vec{v}) \geq 0$ with equality iff $\vec{u} = \vec{v}$

    (b) $d(\vec{u}, \vec{v}) = d(\vec{v}, \vec{u})$

    (c) $d(\vec{u}, \vec{w}) \leq d(\vec{u}, \vec{v}) + d(\vec{v}, \vec{w})$ (triangle inequality)

20. Let W be a subspace with orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_k\}$. Show that the projection matrix $P = \vec{u}_1\vec{u}_1^T + \cdots + \vec{u}_k\vec{u}_k^T$ satisfies $P^2 = P$ and $P^T = P$.

21. Find the polynomial $p(x) = a + bx$ that best approximates $f(x) = e^x$ on $[0, 1]$ using the inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$.

22. Suppose $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a linearly independent set. Show that after applying Gram-Schmidt, the resulting orthogonal set $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ satisfies:

    $$\text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \text{span}\{\vec{u}_1, \ldots, \vec{u}_k\}$$

    for $k = 1, 2, 3$.

23. Let A be an $m \times n$ matrix with linearly independent columns. Show that A^TA is invertible.

24. Find the distance from the point $(1, 1, 1, 1)$ to the subspace $W = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

25. Using Fourier series ideas, find the best approximation to $f(x) = x$ on $[-\pi, \pi]$ using $\text{span}\{1, \sin(x), \cos(x)\}$ with inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx$.

26. Prove that if Q is an $n \times n$ orthogonal matrix (i.e., $Q^TQ = I$), then multiplication by Q preserves inner products: $\langle Q\vec{u}, Q\vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle$ for all $\vec{u}, \vec{v} \in \mathbb{R}^n$.

9

Applications and Advanced Topics

"'The purpose of computation is insight, not numbers."

— Richard Hamming

Chapter Overview

Throughout this book, we've built a powerful toolkit of linear algebra concepts: vectors, matrices, eigenvalues, inner products, and projections. Now it's time to see how these abstract ideas come together to solve real-world problems. In this final chapter, we'll explore some of the most important applications of linear algebra in modern science, engineering, and technology. From predicting the future of populations to ranking web pages, from analyzing data to solving differential equations, linear algebra is the mathematical foundation of the modern world. These applications aren't just academic exercises—they're the tools that power search engines, enable self-driving cars, compress your photos, and help scientists understand complex systems.

9.1 Markov Chains

9.1.1 Introduction to Markov Chains

A Markov chain is a mathematical system that transitions from one state to another with given probabilities. The key property: the future depends only on the present state, not on how we got there.

Definition: Markov Chain

A **Markov chain** is a sequence of random variables $X_0, X_1, X_2, \ldots$ where the probability of moving to the next state depends only on the current state, not on previous states. This is called the **Markov property**:

$$P(X_{n+1} = j \mid X_n = i, X_{n-1}, \ldots, X_0) = P(X_{n+1} = j \mid X_n = i)$$

9.1.2 Transition Matrices

We represent Markov chains using transition matrices.

Definition: Transition Matrix

A **transition matrix** (or **stochastic matrix**) P is a square matrix where:

- Entry p_{ij} represents the probability of moving from state j to state i
- Each column sums to 1 (all probabilities from a given state sum to 1)
- All entries are non-negative: $p_{ij} \geq 0$

Example 9.1.1. *Weather Model Suppose tomorrow's weather depends only on today's weather:*

- *If sunny today: 80% chance sunny tomorrow, 20% chance rainy*
- *If rainy today: 40% chance sunny tomorrow, 60% chance rainy*

The transition matrix is:

$$P = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$$

If we represent the state as $\vec{x} = \begin{bmatrix} P(\textit{sunny}) \\ P(\textit{rainy}) \end{bmatrix}$*, then tomorrow's state is:*

$$\vec{x}_{n+1} = P\vec{x}_n$$

Starting with a sunny day ($\vec{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$*):*

Day 1: $\vec{x}_1 = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix}$

Day 2: $\vec{x}_2 = P\vec{x}_1 = \begin{bmatrix} 0.72 \\ 0.28 \end{bmatrix}$

And so on.

9.1.3 Steady-State Vectors

A key question: what happens in the long run?

Definition: Steady-State Vector

A **steady-state vector** (or **equilibrium vector**) $\vec{v}$ satisfies:

$$P\vec{v} = \vec{v}$$

In other words, $\vec{v}$ is an eigenvector of P with eigenvalue $\lambda = 1$.

Why Eigenvalue 1?

Transition matrices always have $\lambda = 1$ as an eigenvalue! This is because the columns sum to 1. The steady-state vector tells us the long-term probabilities of being in each state.

Example 9.1.2. *Finding Steady-State Distribution Find the steady-state for the weather model:* $P = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$.

Solution: *Solve* $(P - I)\vec{v} = \vec{0}$*:*

$$\begin{bmatrix} -0.2 & 0.4 \\ 0.2 & -0.4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \vec{0}$$

From the first equation: $-0.2v_1 + 0.4v_2 = 0 \Rightarrow v_1 = 2v_2$.

Since probabilities sum to 1: $v_1 + v_2 = 1 \Rightarrow 2v_2 + v_2 = 1 \Rightarrow v_2 = \frac{1}{3}$.

Therefore $v_1 = \frac{2}{3}$ *and the steady-state is:*

$$\vec{v} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

In the long run, it's sunny 2/3 of the time and rainy 1/3 of the time, regardless of the starting state!

9.1.4 Google PageRank

One of the most famous applications of Markov chains is Google's PageRank algorithm.

The Idea: A web page is important if many important pages link to it. We model web surfing as a Markov chain:

- States = web pages
- Transitions = following links (with equal probability among all links on a page)
- The steady-state vector gives the "importance" of each page

Example 9.1.3. *Mini PageRank Consider 4 web pages with links:*

- *Page 1 links to pages 2 and 3*
- *Page 2 links to page 4*
- *Page 3 links to pages 1 and 4*
- *Page 4 links to pages 1, 2, and 3*

The transition matrix (with equal probability for each link):

$$P = \begin{bmatrix} 0 & 0 & 1/2 & 1/3 \\ 1/2 & 0 & 0 & 1/3 \\ 1/2 & 0 & 0 & 1/3 \\ 0 & 1 & 1/2 & 0 \end{bmatrix}$$

Finding the steady-state vector (eigenvalue 1) gives the PageRank scores. After computation:

$$\vec{v} \approx \begin{bmatrix} 0.31 \\ 0.21 \\ 0.21 \\ 0.27 \end{bmatrix}$$

Page 1 has the highest rank (31%), followed by Page 4 (27%).

The Real PageRank

The actual PageRank algorithm includes modifications:

- A "damping factor" (usually 0.85) to handle pages with no outgoing links
- Random jumps to any page with small probability
- Weighted links based on anchor text and other factors

But the core idea—finding the steady-state of a massive Markov chain with billions of states—remains the same!

9.2 Linear Programming (Introduction)

9.2.1 What is Linear Programming?

Linear programming solves optimization problems with linear constraints.

Definition: Linear Programming Problem

A **linear programming problem** seeks to:
Maximize (or minimize): $z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ (objective function)
Subject to:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\leq b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leq b_m \\ x_1, x_2, \ldots, x_n &\geq 0 \end{aligned}$$

9.2.2 Geometric Interpretation in 2D

In two variables, constraints define half-planes, and their intersection is the **feasible region**. The optimal solution occurs at a vertex (corner point) of this region.

Example 9.2.1. *Production Problem A factory produces chairs and tables:*

- *Each chair requires 2 hours of labor and yields \$30 profit*
- *Each table requires 3 hours of labor and yields \$50 profit*
- *18 hours of labor available per day*
- *Materials limit: at most 5 chairs or 4 tables*

Let x = number of chairs, y = number of tables.

Objective: *Maximize profit* $P = 30x + 50y$

Constraints:

$$\begin{aligned} 2x + 3y &\leq 18 \quad \textit{(labor)} \\ x &\leq 5 \quad \textit{(chair limit)} \\ y &\leq 4 \quad \textit{(table limit)} \\ x, y &\geq 0 \end{aligned}$$

Solution: *Graph the feasible region. The vertices are:*

- $(0, 0)$*:* $P = 0$
- $(5, 0)$*:* $P = 150$

- $(0, 4)$*:* $P = 200$
- $(5, 2.67)$*: Not feasible (exceeds labor)*
- $(3, 4)$*:* $P = 30(3) + 50(4) = 290$

The maximum profit is $290, achieved by producing 3 chairs and 4 tables.

9.2.3 The Simplex Method (Overview)

For larger problems, we use the **simplex algorithm**, which efficiently moves from vertex to vertex, always improving the objective function, until reaching the optimum.

The simplex method uses matrix operations and is one of the most important algorithms in optimization. While the details are beyond our scope, the key insight is that it systematically searches the vertices of the feasible region using linear algebra.

Applications of Linear Programming

- **Business**: Production planning, resource allocation
- **Transportation**: Route optimization, logistics
- **Finance**: Portfolio optimization
- **Engineering**: Circuit design, network flow
- **Agriculture**: Crop planning, diet optimization

9.3 Singular Value Decomposition (Introduction)

9.3.1 What is SVD?

The Singular Value Decomposition is one of the most important matrix factorizations in applied linear algebra.

Theorem: Singular Value Decomposition

Every $m \times n$ matrix A can be factored as:

$$A = U\Sigma V^T$$

where:

- U is an $m \times m$ orthogonal matrix (left singular vectors)

- Σ is an $m \times n$ diagonal matrix with non-negative entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$ (singular values)
- V is an $n \times n$ orthogonal matrix (right singular vectors)

9.3.2 Geometric Interpretation

The SVD describes how A transforms space:

1. V^T rotates/reflects the input space
2. Σ stretches along coordinate axes (by singular values)
3. U rotates/reflects to the output space

Connection to Eigenvalues

The singular values of A are the square roots of the eigenvalues of A^TA (or AA^T). The columns of V are eigenvectors of A^TA, and the columns of U are eigenvectors of AA^T.

9.3.3 Computing SVD (Simple Example)

Example 9.3.1. *SVD of a 2×2 Matrix Find the SVD of* $A = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$.

Solution: *This matrix is already diagonal with entries 3 and -2. However, SVD requires non-negative diagonal entries.*

We can write:

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

So: $U = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$, $V = I$.

The singular values are $\sigma_1 = 3$ *and* $\sigma_2 = 2$.

9.3.4 Image Compression with SVD

One powerful application is image compression.

An image can be represented as an $m \times n$ matrix (pixel intensities). The SVD gives:

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

To compress, keep only the largest k singular values:

$$A_k = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T + \cdots + \sigma_k \vec{u}_k \vec{v}_k^T$$

This gives the best rank-k approximation to A (in terms of minimizing $\|A - A_k\|$).

Example 9.3.2. *Compression Ratio For a* 1000×1000 *image:*

- *Original: 1,000,000 values to store*
- *Rank-50 approximation: Need to store 50 singular values + 50 left vectors + 50 right vectors* $= 50 + 50(1000) + 50(1000) = 100,050$ *values*
- *Compression ratio:* $\frac{1,000,000}{100,050} \approx 10 : 1$

With $k = 50$*, we capture most of the important information while reducing storage by 90%!*

Other SVD Applications

- **Principal Component Analysis**: Finding directions of maximum variance in data
- **Recommender Systems**: Netflix prize, collaborative filtering
- **Natural Language Processing**: Latent semantic analysis
- **Signal Processing**: Noise reduction
- **Control Theory**: System identification

9.4 Linear Algebra in Data Science

9.4.1 Data Matrices

In data science, we typically represent datasets as matrices where:

- Each row is an observation (data point)
- Each column is a feature (variable)

For example, a dataset of houses might have columns for: square footage, number of bedrooms, age, price.

9.4.2 Dimensionality Reduction with PCA

When we have many features, visualization and analysis become difficult. Principal Component Analysis (PCA) reduces dimensionality while preserving variance.

The PCA Algorithm:

1. Center the data: subtract the mean of each feature
2. Compute the covariance matrix $C = \frac{1}{n-1}X^T X$
3. Find eigenvalues and eigenvectors of C
4. The eigenvectors (principal components) are the new axes
5. The eigenvalues measure variance along each component
6. Project data onto the top k principal components

Example 9.4.1. *PCA on 2D Data Consider student scores:* $X = \begin{bmatrix} 85 & 88 \\ 90 & 92 \\ 78 & 80 \\ 92 & 95 \end{bmatrix}$ *(columns: Math, Physics)*

Step 1: *Center the data: Mean:* $\bar{x} = [86.25, 88.75]$

$X_{centered} = \begin{bmatrix} -1.25 & -0.75 \\ 3.75 & 3.25 \\ -8.25 & -8.75 \\ 5.75 & 6.25 \end{bmatrix}$

Step 2: *Covariance matrix:*

$$C = \frac{1}{3}X_{centered}^T X_{centered} = \begin{bmatrix} 31.19 & 32.19 \\ 32.19 & 33.58 \end{bmatrix}$$

Step 3-4: *The largest eigenvalue's eigenvector points in the direction of maximum variance (roughly the line $y = x$).*

Most of the variance is explained by one principal component, showing that Math and Physics scores are highly correlated.

9.4.3 Machine Learning: Linear Models

Many machine learning algorithms are fundamentally linear algebra:

Linear Regression: We saw this in Chapter 8 as least squares: $\hat{\beta} = (X^T X)^{-1} X^T y$

Logistic Regression: Classification using $P(y = 1) = \frac{1}{1+e^{-\vec{w}^T\vec{x}}}$

Neural Networks: Each layer computes $\vec{y} = \sigma(W\vec{x} + \vec{b})$ where W is a weight matrix and σ is an activation function.

Why Linear Algebra in ML?

- Efficiently handles high-dimensional data
- Matrix operations are highly optimized (GPUs)
- Provides geometric intuition (hyperplanes, distances)
- Enables techniques like regularization (ridge, lasso)

9.4.4 Covariance Matrices

The covariance matrix is fundamental in statistics and machine learning.

Definition: Covariance Matrix

For centered data matrix X (with n observations and p features):

$$C = \frac{1}{n-1}X^T X$$

Entry c_{ij} measures how features i and j vary together:

- c_{ii} = variance of feature i
- c_{ij} for $i \neq j$ = covariance between features i and j
- C is symmetric and positive semi-definite

The eigenvalues of C measure variance along principal directions, and eigenvectors point in those directions.

9.5 Linear Algebra in Computer Graphics

9.5.1 3D Transformations

Computer graphics relies heavily on matrix transformations.

Basic 3D Transformations

Rotation about z-axis by angle θ:

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Scaling by factors s_x, s_y, s_z:

$$S = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix}$$

Translation by $\vec{t} = \langle t_x, t_y, t_z \rangle$: Translation is not a linear transformation! We use homogeneous coordinates.

9.5.2 Homogeneous Coordinates

To handle translation with matrices, we add a fourth coordinate.

A 3D point (x, y, z) becomes $(x, y, z, 1)$ in homogeneous coordinates.

Translation Matrix

$$T = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then:

$$\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ z + t_z \\ 1 \end{bmatrix}$$

Example 9.5.1. *Composite Transformation To rotate a point about $(1, 0, 0)$ instead of the origin:*

1. *Translate by $(-1, 0, 0)$ to move rotation center to origin*
2. *Rotate*
3. *Translate back by $(1, 0, 0)$*

The combined transformation is: $M = T(1, 0, 0) \cdot R \cdot T(-1, 0, 0)$

This is why matrix multiplication order matters!

9.5.3 Viewing Transformations

Rendering a 3D scene involves several coordinate transformations:

1. **Model coordinates** → object's local space
2. **World coordinates** → global scene space (via model matrix)
3. **Camera coordinates** → view from camera (via view matrix)
4. **Clip coordinates** → perspective projection (via projection matrix)
5. **Screen coordinates** → final 2D image (viewport transform)

Each transformation is a matrix multiplication:

$$\vec{v}_{\text{screen}} = M_{\text{viewport}} \cdot M_{\text{projection}} \cdot M_{\text{view}} \cdot M_{\text{model}} \cdot \vec{v}_{\text{object}}$$

Why This Matters

Video games and CGI films render millions of vertices per frame. GPU hardware is optimized for matrix-vector multiplication, making these linear algebra operations incredibly fast. Understanding the math helps create better graphics engines, shaders, and special effects.

9.5.4 Lighting and Shading

The dot product determines lighting:

Lambert's Law: Light intensity is proportional to $\vec{n} \cdot \vec{l}$ where:

- $\vec{n}$ is the surface normal (perpendicular to surface)
- $\vec{l}$ is the direction to the light source

When $\vec{n} \perp \vec{l}$ (dot product $= 0$), the surface is perpendicular to light and appears dark.

9.6 Differential Equations and Systems

9.6.1 Systems of Linear Differential Equations

Many physical systems are described by systems of differential equations.

Linear System

A system of first-order linear differential equations:

$$\frac{d\vec{x}}{dt} = A\vec{x}$$

where A is an $n \times n$ matrix and $\vec{x}(t)$ is a vector-valued function.

9.6.2 Solution Using Eigenvalues

If A has eigenvalues $\lambda_1, \ldots, \lambda_n$ with eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$, the general solution is:

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

Example 9.6.1. *Coupled System Solve:*

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \quad \vec{x}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Solution: *Eigenvalues:* $\lambda_1 = 1$, $\lambda_2 = 3$ *(diagonal entries for triangular matrix)*

For $\lambda_1 = 1$*:*

$$(A - I)\vec{v} = \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}\vec{v} = \vec{0} \Rightarrow \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For $\lambda_2 = 3$*:*

$$(A - 3I)\vec{v} = \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix}\vec{v} = \vec{0} \Rightarrow \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

General solution:

$$\vec{x}(t) = c_1 e^{t}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^{3t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Using $\vec{x}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$*:*

$$c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow c_1 = 0, c_2 = 1$$

Therefore:

$$\vec{x}(t) = e^{3t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} e^{3t} \\ e^{3t} \end{bmatrix}$$

9.6.3 Stability Analysis

The eigenvalues determine the behavior of solutions:

Stability Criteria

For $\frac{d\vec{x}}{dt} = A\vec{x}$:

- If all eigenvalues have $\text{Re}(\lambda) < 0$: system is **stable** (solutions decay to zero)
- If any eigenvalue has $\text{Re}(\lambda) > 0$: system is **unstable** (solutions grow)
- If eigenvalues are purely imaginary: system **oscillates**
- If eigenvalues are complex: solutions spiral

Example 9.6.2. *Predator-Prey Model The Lotka-Volterra equations (simplified linear version):*

$$\frac{d}{dt}\begin{bmatrix} R \\ F \end{bmatrix} = \begin{bmatrix} a & -b \\ c & -d \end{bmatrix}\begin{bmatrix} R \\ F \end{bmatrix}$$

where R = rabbits (prey), F = foxes (predators), and $a, b, c, d > 0$.

The eigenvalues determine if populations oscillate, grow, or die out. For certain parameter values, we get purely imaginary eigenvalues, leading to oscillating populations—matching real ecological observations!

Applications in Engineering

- **Electrical circuits**: RLC circuits, filter design
- **Mechanical systems**: Mass-spring-damper systems, vibrations
- **Control theory**: Feedback systems, stability analysis
- **Chemical kinetics**: Reaction networks
- **Economics**: Dynamic models of markets

9.7 Chapter Summary

In this final chapter, we've seen how linear algebra powers modern applications:

- **Markov chains** model systems that transition between states probabilistically, with applications from weather prediction to Google PageRank

- **Linear programming** optimizes linear objectives subject to linear constraints, used in business, logistics, and resource allocation
- **Singular Value Decomposition** is a powerful matrix factorization enabling image compression, dimensionality reduction, and data analysis
- **Data science** uses linear algebra for PCA, machine learning, and statistical analysis via covariance matrices
- **Computer graphics** transforms and renders 3D scenes using matrices, with applications in games, films, and virtual reality
- **Differential equations** describing physical systems can be solved using eigenvalues and eigenvectors

Looking Back and Forward

You've now completed a journey through linear algebra—from simple vectors to powerful applications. The concepts you've learned are the foundation of modern mathematics, science, and engineering. Whether you pursue pure mathematics, applied science, computer science, data science, or engineering, linear algebra will be an essential tool in your mathematical toolkit. The abstract ideas we studied—bases, eigenvalues, orthogonality, projections—are not just theoretical constructs; they're the language in which we describe and solve real-world problems.

9.8 Practice Problems

9.8.1 Basic Problems

1. For the transition matrix $P = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$:
   (a) Verify that it's a valid transition matrix
   (b) Find the steady-state vector
   (c) If you start in state 1, what's the probability of being in state 2 after 2 transitions?
2. A factory produces products A and B. Product A requires 2 hours and yields \$20 profit. Product B requires 3 hours and yields \$25 profit. With 30 hours available, set up and solve a linear programming problem to maximize profit.

3. Find the singular values of $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$.

4. Given data points $(1, 2)$, $(2, 4)$, $(3, 7)$, $(4, 8)$, compute the covariance between the x and y coordinates.

5. Write the 4×4 homogeneous coordinate matrix for translating by $(2, -1, 3)$ in 3D space.

6. Solve the system $\frac{d\vec{x}}{dt} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \vec{x}$ with $\vec{x}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

7. For lighting calculation, if the surface normal is $\vec{n} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and light direction is $\vec{l} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$, what is the light intensity (using $\vec{n} \cdot \vec{l}$)?

8. Determine the stability of $\frac{d\vec{x}}{dt} = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix} \vec{x}$.

9.8.2 Intermediate Problems

9. A Markov chain has transition matrix $P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.6 & 0.4 \\ 0.3 & 0.1 & 0.4 \end{bmatrix}$. Find the steady-state distribution.

10. A company produces three products with constraints on labor and materials. Set up the linear programming problem:

    - Product 1: 2 labor hours, 1 material unit, \$15 profit
    - Product 2: 3 labor hours, 2 material units, \$20 profit
    - Product 3: 1 labor hour, 1 material unit, \$10 profit
    - Available: 100 labor hours, 60 material units

11. Find the SVD of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ by computing eigenvalues of $A^T A$.

12. For the dataset:

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

Find the principal component (eigenvector of covariance matrix with largest eigenvalue).

13. Create the composite transformation matrix that rotates by 90° about the z-axis, then translates by $(1, 2, 0)$.

14. Solve $\frac{d\vec{x}}{dt} = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix} \vec{x}$ with $\vec{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

15. In a simple PageRank example with 3 pages where page 1 links to pages 2 and 3, page 2 links to page 1, and page 3 links to pages 1 and 2, find the PageRank scores.

9.8.3 Challenge Problems

16. Prove that every transition matrix has $\lambda = 1$ as an eigenvalue. (Hint: Consider what $P^T\vec{1}$ equals, where $\vec{1}$ is the vector of all 1's.)

17. For the feasible region defined by $x + y \leq 5$, $2x + y \leq 8$, $x, y \geq 0$, find all vertices and determine which maximizes $3x + 2y$.

18. Show that the singular values of A are the square roots of the eigenvalues of $A^T A$.

19. Prove that the covariance matrix is always symmetric and positive semi-definite.

20. Explain why the product of rotation matrices is another rotation matrix. What property ensures this?

21. For the system $\frac{d\vec{x}}{dt} = A\vec{x}$ where A has complex eigenvalues $\lambda = a \pm bi$, show that solutions spiral inward when $a < 0$ and spiral outward when $a > 0$.

22. A more realistic PageRank includes a damping factor $d = 0.85$:

$$\vec{v} = d(P\vec{v}) + \frac{1-d}{n}\vec{1}$$

Explain why this modification is necessary (consider pages with no outlinks).

23. In PCA, prove that projecting data onto the first k principal components minimizes the reconstruction error $\|X - X_k\|^2$.

24. Show that homogeneous coordinates can represent perspective projection, which makes distant objects appear smaller.

25. For the predator-prey model $\frac{d\vec{x}}{dt} = \begin{bmatrix} a & -b \\ c & -d \end{bmatrix} \vec{x}$ with $a, b, c, d > 0$, find conditions on the parameters for oscillatory behavior (purely imaginary eigenvalues).

Mathematical Notation and Symbols

.1 Sets and Numbers

Symbol	Meaning
$\mathbb{N}$	Natural numbers: $\{1, 2, 3, \ldots\}$
$\mathbb{Z}$	Integers: $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{Q}$	Rational numbers (fractions)
$\mathbb{R}$	Real numbers
$\mathbb{C}$	Complex numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\in$	"is an element of" or "belongs to"
$\notin$	"is not an element of"
$\subset$	"is a subset of"
$\subseteq$	"is a subset of or equal to"
$\emptyset$ or $\{\}$	Empty set
$\cup$	Union (elements in either set)
$\cap$	Intersection (elements in both sets)

Example .1.1. *Set Notation Let $A = \{1, 2, 3\}$ and $B = \{2, 3, 4, 5\}$. Then:*

- $2 \in A$ *(2 is in set A)*
- $6 \notin B$ *(6 is not in set B)*
- $A \cup B = \{1, 2, 3, 4, 5\}$
- $A \cap B = \{2, 3\}$

.2 Greek Letters

Letter	**Name**	**Letter**	**Name**
α	alpha	ν	nu
β	beta	ξ	xi
γ	gamma	π	pi
δ	delta	ρ	rho
ϵ or ε	epsilon	σ	sigma
ζ	zeta	τ	tau
η	eta	ϕ or φ	phi
θ or ϑ	theta	χ	chi
ι	iota	ψ	psi
κ	kappa	ω	omega
λ	lambda		
μ	mu		

Capital Greek Letters: Γ (Gamma), Δ (Delta), Θ (Theta), Λ (Lambda), Ξ (Xi), Π (Pi), Σ (Sigma), Φ (Phi), Ψ (Psi), Ω (Omega)

.3 Vector and Matrix Notation

Symbol	Meaning
$\vec{v}$ or $\mathbf{v}$	Vector (boldface or arrow notation)
$\langle v_1, v_2, \ldots, v_n \rangle$	Vector in component form
$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$	Column vector
$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$	Row vector
$\lVert \vec{v} \rVert$	Norm (magnitude) of vector $\vec{v}$
$\vec{u} \cdot \vec{v}$	Dot product of $\vec{u}$ and $\vec{v}$
$\vec{u} \times \vec{v}$	Cross product of $\vec{u}$ and $\vec{v}$ (3D only)
$\langle \vec{u}, \vec{v} \rangle$	Inner product of $\vec{u}$ and $\vec{v}$
A, B, C	Matrices (capital letters)
a_{ij}	Entry in row i, column j of matrix A
A^T	Transpose of matrix A
A^{-1}	Inverse of matrix A
$\det(A)$ or $\lvert A \rvert$	Determinant of matrix A
$\mathrm{tr}(A)$	Trace of matrix A (sum of diagonal entries)
I or I_n	Identity matrix (size $n \times n$)
O or $\vec{0}$	Zero matrix or zero vector

.4 Summation and Product Notation

Summation:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + a_3 + \cdots + a_n$$

Example .4.1. *Summation* $\sum_{i=1}^{4} i^2 = 1^2 + 2^2 + 3^2 + 4^2 = 1 + 4 + 9 + 16 = 30$

Product:

$$\prod_{i=1}^{n} a_i = a_1 \cdot a_2 \cdot a_3 \cdots a_n$$

Example .4.2. *Product* $\prod_{i=1}^{4} i = 1 \cdot 2 \cdot 3 \cdot 4 = 24 = 4!$

.5 Logic and Proof Notation

Symbol	Meaning
$\Rightarrow$	"implies" or "if...then"
$\Leftrightarrow$	"if and only if" (iff)
$\forall$	"for all" (universal quantifier)
$\exists$	"there exists" (existential quantifier)
$\therefore$	"therefore"
$\because$	"because"
$\neg$	"not" (negation)
$\wedge$	"and" (conjunction)
$\vee$	"or" (disjunction)
$\equiv$	"is equivalent to"
$\square$ or QED	End of proof

.6 Special Symbols

Symbol	Meaning
$\approx$	"approximately equal to"
$\neq$	"not equal to"
$\leq$	"less than or equal to"
$\geq$	"greater than or equal to"
$\ll$	"much less than"
$\gg$	"much greater than"
$\propto$	"proportional to"
$\perp$	"perpendicular to" or "orthogonal to"
$\parallel$	"parallel to"
∞	infinity
$\pm$	"plus or minus"
$\cdots$	horizontal ellipsis
$\vdots$	vertical ellipsis
$\ddots$	diagonal ellipsis

Proof Techniques

.7 Introduction to Proofs

A mathematical proof is a logical argument that establishes the truth of a statement beyond any doubt. Proofs are the foundation of mathematics—they transform observations into certainties.

.7.1 Structure of a Proof

A typical proof has the following structure:

1. **State what you want to prove** (the theorem or claim)
2. **List assumptions** (given information, definitions)
3. **Present logical steps** (using definitions, previously proven results)
4. **Reach the conclusion** (what was to be proved)
5. **Mark the end** (with □, QED, or "This completes the proof")

.8 Direct Proof

A direct proof starts with assumptions and uses logical steps to reach the conclusion.

Direct Proof Template

Goal: Prove "If P, then Q" (written $P \Rightarrow Q$)
Method:

1. Assume P is true
2. Use definitions, axioms, and previously proven theorems
3. Show that Q must be true

Example .8.1. *Direct Proof* ***Claim:*** *If n is an even integer, then n^2 is even.*

Proof: *Assume n is even. By definition, there exists an integer k such that $n = 2k$.*

Then:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

Since $2k^2$ is an integer, we can write $n^2 = 2m$ where $m = 2k^2$. Therefore, n^2 is even by definition. □

.9 Proof by Contrapositive

To prove "$P \Rightarrow Q$", we instead prove "$\neg Q \Rightarrow \neg P$" (the contrapositive).

Why This Works

A statement and its contrapositive are logically equivalent:

$$(P \Rightarrow Q) \equiv (\neg Q \Rightarrow \neg P)$$

Example .9.1. *Proof by Contrapositive* ***Claim:*** *If n^2 is even, then n is even.*

Proof: *(by contrapositive) We'll prove: If n is odd, then n^2 is odd.*

Assume n is odd. Then $n = 2k + 1$ for some integer k.

Therefore:

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

Since $2k^2 + 2k$ is an integer, n^2 is odd. This proves the contrapositive, so the original claim is true. □

.10 Proof by Contradiction

Assume the opposite of what you want to prove, then show this leads to a logical contradiction.

Proof by Contradiction Template

Goal: Prove statement P

Method:

1. Assume $\neg P$ (the opposite of P)
2. Use logical reasoning
3. Arrive at a contradiction (something impossible)
4. Conclude that $\neg P$ must be false, so P is true

Example .10.1. *Proof by Contradiction* ***Claim:*** *$\sqrt{2}$ is irrational.*

Proof: *Assume, for contradiction, that $\sqrt{2}$ is rational. Then we can write:*

$$\sqrt{2} = \frac{a}{b}$$

where a and b are integers with no common factors (in lowest terms) and $b \neq 0$.

Squaring both sides: $2 = \frac{a^2}{b^2}$, so $a^2 = 2b^2$.

This means a^2 is even, which implies a is even (proven earlier). So $a = 2k$ for some integer k.

Substituting: $(2k)^2 = 2b^2 \Rightarrow 4k^2 = 2b^2 \Rightarrow 2k^2 = b^2$.

This means b^2 is even, so b is even.

But if both a and b are even, they have a common factor of 2, contradicting our assumption that the fraction is in lowest terms.

This contradiction shows our assumption was wrong. Therefore, $\sqrt{2}$ is irrational. $\square$

.11 Proof by Induction

Mathematical induction proves statements about all natural numbers.

Principle of Mathematical Induction

To prove a statement $P(n)$ is true for all $n \geq n_0$:
Step 1 (Base Case): Prove $P(n_0)$ is true
Step 2 (Inductive Step): Assume $P(k)$ is true for some $k \geq n_0$ (inductive hypothesis), and prove $P(k+1)$ is true
Conclusion: By induction, $P(n)$ is true for all $n \geq n_0$

Example .11.1. *Proof by Induction* ***Claim:*** *For all $n \geq 1$, $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.*

Proof:

Base case ($n = 1$):

$$\sum_{i=1}^{1} i = 1 \quad \textit{and} \quad \frac{1(1+1)}{2} = \frac{2}{2} = 1$$

The formula holds for $n = 1$. ✓

Inductive step: *Assume the formula holds for $n = k$:*

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}$$

We must show it holds for $n = k + 1$:

$$\begin{aligned}
\sum_{i=1}^{k+1} i &= \left(\sum_{i=1}^{k} i\right) + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1) \quad \textit{(by inductive hypothesis)} \\
&= \frac{k(k+1) + 2(k+1)}{2} \\
&= \frac{(k+1)(k+2)}{2} \\
&= \frac{(k+1)((k+1)+1)}{2}
\end{aligned}$$

This is exactly the formula for $n = k + 1$.

By mathematical induction, the formula holds for all $n \geq 1$. □

.12 Existence Proofs

To prove "there exists an x such that $P(x)$", you can either:

1. **Constructive:** Explicitly find such an x
2. **Non-constructive:** Prove one must exist without finding it

Example .12.1. *Existence Proof* ***Claim:*** *There exists a matrix A such that $A^2 = -I$.*

Proof: *(constructive) Consider* $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Then:

$$A^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I$$

Therefore, such a matrix exists. □

.13 Counterexamples

To disprove a universal statement ("for all x, $P(x)$"), find one counterexample.

Example .13.1. *Disproof by Counterexample* ***Claim:*** *For all matrices A and B, $(A + B)^2 = A^2 + 2AB + B^2$.*

Disproof: *Let* $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Then:

$$(A+B)^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

But:

$$\begin{aligned} A^2 + 2AB + B^2 &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Since $(A+B)^2 \neq A^2 + 2AB + B^2$*, the claim is false.* $\square$

Note: *The correct formula requires commutativity:* $(A+B)^2 = A^2 + AB + BA + B^2$.

Review of Prerequisite Concepts

.14 Algebra Review

.14.1 Factoring

Common patterns:

- Difference of squares: $a^2 - b^2 = (a-b)(a+b)$
- Perfect square trinomial: $a^2 + 2ab + b^2 = (a+b)^2$
- Sum/difference of cubes: $a^3 \pm b^3 = (a \pm b)(a^2 \mp ab + b^2)$
- Quadratic formula: $ax^2 + bx + c = 0 \Rightarrow x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$

.14.2 Exponent Rules

For $a, b > 0$ and real numbers m, n:

- $a^m \cdot a^n = a^{m+n}$
- $\frac{a^m}{a^n} = a^{m-n}$
- $(a^m)^n = a^{mn}$
- $(ab)^n = a^n b^n$
- $a^0 = 1$ (if $a \neq 0$)
- $a^{-n} = \frac{1}{a^n}$
- $a^{1/n} = \sqrt[n]{a}$

.14.3 Logarithm Rules

For $a, b > 0$ and $a, b \neq 1$:

- $\log_a(xy) = \log_a(x) + \log_a(y)$

- $\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$
- $\log_a(x^r) = r\log_a(x)$
- $\log_a(a) = 1$
- $\log_a(1) = 0$
- $a^{\log_a(x)} = x$
- Change of base: $\log_b(x) = \frac{\log_a(x)}{\log_a(b)}$

.15 Trigonometry Review

.15.1 Trigonometric Functions

For angle θ in a right triangle:

$$\sin(\theta) = \frac{\text{opposite}}{\text{hypotenuse}}, \quad \cos(\theta) = \frac{\text{adjacent}}{\text{hypotenuse}}, \quad \tan(\theta) = \frac{\text{opposite}}{\text{adjacent}}$$

.15.2 Special Angles

θ	$\sin(\theta)$	$\cos(\theta)$	$\tan(\theta)$
0ř or 0	0	1	0
30ř or $\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
45ř or $\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
60ř or $\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
90ř or $\frac{\pi}{2}$	1	0	undefined

.15.3 Pythagorean Identity

$$\sin^2(\theta) + \cos^2(\theta) = 1$$

Variants:

- $1 + \tan^2(\theta) = \sec^2(\theta)$
- $1 + \cot^2(\theta) = \csc^2(\theta)$

.15.4 Angle Sum Formulas

- $\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$
- $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$
- $\tan(\alpha + \beta) = \frac{\tan(\alpha)+\tan(\beta)}{1-\tan(\alpha)\tan(\beta)}$

.15.5 Double Angle Formulas

- $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$
- $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 2\cos^2(\theta) - 1 = 1 - 2\sin^2(\theta)$
- $\tan(2\theta) = \frac{2\tan(\theta)}{1-\tan^2(\theta)}$

.16 Calculus Review

.16.1 Derivatives

Basic derivatives:

- $\frac{d}{dx}(c) = 0$ (constant)
- $\frac{d}{dx}(x^n) = nx^{n-1}$ (power rule)
- $\frac{d}{dx}(e^x) = e^x$
- $\frac{d}{dx}(\ln(x)) = \frac{1}{x}$
- $\frac{d}{dx}(\sin(x)) = \cos(x)$
- $\frac{d}{dx}(\cos(x)) = -\sin(x)$

Rules:

- Sum rule: $(f + g)' = f' + g'$
- Product rule: $(fg)' = f'g + fg'$
- Quotient rule: $\left(\frac{f}{g}\right)' = \frac{f'g-fg'}{g^2}$
- Chain rule: $(f(g(x)))' = f'(g(x)) \cdot g'(x)$

.16.2 Integration

Basic integrals:

- $\int x^n \, dx = \frac{x^{n+1}}{n+1} + C$ (for $n \neq -1$)
- $\int \frac{1}{x} \, dx = \ln |x| + C$
- $\int e^x \, dx = e^x + C$
- $\int \sin(x) \, dx = -\cos(x) + C$
- $\int \cos(x) \, dx = \sin(x) + C$

Fundamental Theorem of Calculus:

$$\int_a^b f(x) \, dx = F(b) - F(a)$$

where $F'(x) = f(x)$ (i.e., F is an antiderivative of f).

.16.3 Taylor Series

For certain functions, we can write:

- $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
- $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$
- $\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$

.17 Complex Numbers

.17.1 Definition

A complex number has the form $z = a + bi$ where:

- a is the real part: $\text{Re}(z) = a$
- b is the imaginary part: $\text{Im}(z) = b$
- i is the imaginary unit: $i^2 = -1$

.17.2 Operations

Addition: $(a+bi)+(c+di)=(a+c)+(b+d)i$

Multiplication: $(a+bi)(c+di)=(ac-bd)+(ad+bc)i$

Complex conjugate: $\overline{z}=a-bi$

Modulus: $|z|=\sqrt{a^2+b^2}$

.17.3 Euler's Formula

$$e^{i\theta}=\cos(\theta)+i\sin(\theta)$$

Special case ($\theta=\pi$):

$$e^{i\pi}+1=0$$

(Euler's identity—connecting five fundamental constants!)

Polar form:

$$z=r(\cos(\theta)+i\sin(\theta))=re^{i\theta}$$

where $r=|z|$ and $\theta=\arg(z)$ (the argument or angle).

Computational Tools and Software

.18 Calculator Techniques

.18.1 Matrix Operations on Graphing Calculators

Most graphing calculators (TI-83/84, TI-89, etc.) can perform matrix operations:

To enter a matrix:

1. Press `2nd` → `MATRIX` (or `MATRX`)
2. Select `EDIT`
3. Choose a matrix name (e.g., [A])
4. Enter dimensions (rows × columns)
5. Enter each element

Basic operations:

- Addition: `[A] + [B]`
- Multiplication: `[A] * [B]`
- Scalar multiplication: `5 * [A]`
- Transpose: `[A]^T` (in MATRIX MATH menu)
- Inverse: `[A]^(-1)`
- Determinant: `det([A])` (in MATRIX MATH menu)
- RREF: `rref([A])` (in MATRIX MATH menu)

.19 Python and NumPy

Python with the NumPy library is excellent for linear algebra computations.

.19.1 Getting Started

Installing NumPy:

```
pip install numpy
```

Importing:

```
import numpy as np
```

.19.2 Creating Matrices and Vectors

```
# Create a matrix
A = np.array([[1, 2], [3, 4]])

# Create a vector
v = np.array([1, 2, 3])

# Zero matrix
Z = np.zeros((3, 3))

# Identity matrix
I = np.eye(4)

# Random matrix
R = np.random.rand(2, 3)
```

.19.3 Matrix Operations

```
# Matrix addition
C = A + B

# Matrix multiplication
C = np.dot(A, B)  # or A @ B (Python 3.5+)

# Element-wise multiplication
C = A * B

# Transpose
A_T = A.T
```

```
# Inverse
A_inv = np.linalg.inv(A)

# Determinant
det_A = np.linalg.det(A)

# Eigenvalues and eigenvectors
eigenvalues, eigenvectors = np.linalg.eig(A)

# Solve linear system Ax = b
x = np.linalg.solve(A, b)

# Rank of a matrix
rank = np.linalg.matrix_rank(A)
```

.19.4 Vector Operations

```
# Dot product
dot_product = np.dot(u, v)

# Cross product (3D only)
cross_product = np.cross(u, v)

# Norm (magnitude)
magnitude = np.linalg.norm(v)

# Normalize a vector
v_normalized = v / np.linalg.norm(v)
```

.20 MATLAB

MATLAB (Matrix Laboratory) is designed specifically for matrix computations.

.20.1 Basic Syntax

```
% Create a matrix
A = [1 2; 3 4];
```

```
% Create a vector
v = [1; 2; 3];  % column vector
w = [1 2 3];    % row vector

% Matrix operations
C = A + B;      % addition
C = A * B;      % multiplication
C = A .* B;     % element-wise multiplication
A_T = A';       % transpose
A_inv = inv(A); % inverse
det_A = det(A); % determinant

% Eigenvalues and eigenvectors
[V, D] = eig(A);  % V = eigenvectors, D = eigenvalues

% Solve Ax = b
x = A \ b;

% Reduced row echelon form
R = rref(A);
```

.21 Online Tools

.21.1 Wolfram Alpha

Visit `www.wolframalpha.com` and type natural language queries:

```
"eigenvalues of {{1,2},{3,4}}"
"solve [1,2;3,4][x;y]=[5;6]"
"determinant [[1,2,3],[4,5,6],[7,8,9]]"
"inverse of {{1,2},{3,4}}"
```

.21.2 Desmos Matrix Calculator

Desmos (`www.desmos.com/matrix`) provides a visual matrix calculator for:

- Matrix arithmetic
- Determinants

- Inverses
- RREF
- Eigenvalues

.21.3 Symbolab

`www.symbolab.com` offers step-by-step solutions for:

- Systems of equations
- Matrix operations
- Eigenvalues and eigenvectors
- Diagonalization

.22 Tips for Computational Work

1. **Check dimensions:** Before multiplying matrices, verify that dimensions are compatible
2. **Numerical precision:** Computers use floating-point arithmetic, which can introduce small errors
3. **Singular matrices:** If a matrix is singular (non-invertible), computational tools may give errors or very large numbers
4. **Verify results:** Always check computational results with hand calculations for small examples
5. **Visualization:** Use graphing tools to visualize vectors, transformations, and eigenspaces
6. **Documentation:** Read the documentation for your chosen tool to understand its capabilities and limitations

Solutions to Selected Exercises

.23 About the Solutions

This appendix contains solutions to selected odd-numbered problems from each chapter. Solutions are provided for:

- All odd-numbered Basic Problems
- Selected odd-numbered Intermediate Problems
- Selected odd-numbered Challenge Problems

Full solutions are available in the separate Solutions Manual for instructors.

.24 How to Use These Solutions

Important Note

Try solving problems on your own first! The learning happens in the struggle. Only consult solutions after:

1. You've made a genuine attempt
2. You're stuck and need a hint
3. You want to verify your answer
4. You want to see an alternative approach

Study Tips

- If you need to look at a solution, read just enough to get unstuck, then try to finish on your own
- After solving a problem, compare your method to the solution—there are often multiple valid approaches
- If you don't understand a step in the solution, review the relevant section of the chapter

- Practice problems with similar structure to ones you found difficult

.25 Notation in Solutions

Solutions use the following conventions:

- **Given:** Lists the information provided in the problem
- **Goal:** Clearly states what we're trying to find or prove
- **Solution:** Presents the step-by-step work
- **Answer:** Highlights the final result
- $\square$ or **QED**: Marks the end of a proof
- ✓: Indicates verification of a result

Further Resources

.26 Recommended Textbooks

.26.1 For Further Study

- **Introduction to Linear Algebra** by Gilbert Strang
  - Clear explanations with emphasis on understanding
  - Excellent for self-study
  - Free lecture videos available from MIT OpenCourseWare
- **Linear Algebra and Its Applications** by David C. Lay
  - Application-focused approach
  - Great examples from engineering and sciences
  - Strong computational emphasis
- **Linear Algebra Done Right** by Sheldon Axler
  - More abstract, proof-based approach
  - Minimal use of determinants (postponed to later)
  - Excellent for students planning to study pure mathematics

.26.2 For Different Perspectives

- **3Blue1Brown's "Essence of Linear Algebra" series** (YouTube)
  - Outstanding visual explanations
  - Intuitive geometric understanding
  - Highly recommended supplement
- **No Bullshit Guide to Linear Algebra** by Ivan Savov
  - Concise, no-nonsense approach
  - Focus on computation and understanding
  - Good for quick reference

.27 Online Resources

.27.1 Video Lectures

- **MIT OpenCourseWare: 18.06 Linear Algebra**
  - Professor Gilbert Strang's legendary course
  - Complete video lectures, notes, and problem sets
  - `ocw.mit.edu`
- **Khan Academy: Linear Algebra**
  - Short videos on specific topics
  - Interactive practice problems
  - `khanacademy.org`
- **3Blue1Brown**
  - Visual, intuitive explanations
  - "Essence of Linear Algebra" series
  - `youtube.com/3blue1brown`

.27.2 Interactive Tools

- **GeoGebra**
  - Visualize vectors, transformations, eigenspaces
  - Free online and desktop versions
  - `geogebra.org`
- **Desmos Matrix Calculator**
  - Simple, visual matrix computations
  - `desmos.com/matrix`
- **Matrix Calculator**
  - Comprehensive matrix operations
  - Shows step-by-step solutions
  - `matrixcalc.org`

.27.3 Practice Problems

- **Paul's Online Math Notes**
  - Excellent notes with many examples
  - `tutorial.math.lamar.edu`
- **MIT OCW Problem Sets**
  - Challenging problems with solutions
  - Real MIT exam questions

.28 Applications and Connections

.28.1 Computer Science

- **Computer Graphics**: Transformations, lighting, rendering
- **Machine Learning**: Neural networks, dimensionality reduction (PCA)
- **Computer Vision**: Image processing, feature detection
- **Algorithms**: PageRank, recommendation systems

.28.2 Engineering

- **Electrical Engineering**: Circuit analysis, signal processing
- **Mechanical Engineering**: Stress analysis, vibrations
- **Control Theory**: Feedback systems, stability analysis
- **Robotics**: Kinematics, path planning

.28.3 Sciences

- **Physics**: Quantum mechanics, relativity, electromagnetism
- **Chemistry**: Molecular modeling, quantum chemistry
- **Biology**: Population dynamics, genetics
- **Economics**: Input-output analysis, optimization

.29 Career Connections

Linear algebra is essential for careers in:

- Data Science and Analytics
- Artificial Intelligence / Machine Learning
- Software Engineering (graphics, games, simulations)
- Quantitative Finance
- Aerospace Engineering
- Robotics
- Research Mathematics
- Physics and Engineering
- Operations Research
- Cryptography and Cybersecurity

.30 Advanced Topics to Explore

After mastering this textbook, consider exploring:

- **Abstract Algebra**: Groups, rings, fields
- **Numerical Linear Algebra**: Computational methods, iterative solvers
- **Functional Analysis**: Infinite-dimensional vector spaces
- **Differential Equations**: Applications of eigenvalues
- **Optimization**: Convex optimization, linear programming
- **Tensor Analysis**: Generalizations of matrices
- **Lie Groups and Lie Algebras**: Continuous symmetries

Final Thoughts

Linear algebra is one of the most useful branches of mathematics. The concepts you've learned—vectors, matrices, transformations, eigenvalues—appear throughout mathematics, science, and engineering. As you continue your mathematical journey, you'll find these ideas popping up again and again, each time revealing new connections and applications.

Keep practicing, stay curious, and remember: mathematics is not about memorizing formulas—it's about understanding patterns, making connections, and developing your ability to think clearly about abstract structures.

Good luck in your continued studies!