with Open Texts

**Linear Algebra with**

**Applications** W. Keith Nicholson

**Open Texts and Editorial**

Digital access to our high-quality texts is entirely FREE! All content is reviewed and updated regularly by subject matter experts, and the texts are adaptable; custom edi

tions can be produced by the Lyryx editorial team for those adopting Lyryx assessment. Access to the original source files is also open to anyone!

**Online Assessment**

**with Open Texts**

**Support**

Access to our in-house support team is available 7 days/week to provide prompt resolution to both student and instructor inquiries. In addition, we work one-on-one with instructors to provide a comprehensive system, cus

tomized for their course. This can include managing multiple sections, assistance with preparing online examinations, help with LMS integra

tion, and much more!

**Instructor Supplements**

Lyryx has been developing online for mative and summative assessment (homework and exams) for more than 20 years, and as with the textbook content, questions and problems are regularly reviewed and updated. Stu dents receive immediate personalized feedback to guide their learning. Stu dent grade reports and performance statistics are also provided, and full LMS integration, including gradebook synchronization, is available.

Additional instructor resources are also freely accessible. Product depen dent, these include: full sets of adapt able slides, solutions manuals, and multiple choice question banks with an exam building tool.

**Contact Lyryx Today!**

info@lyryx.com

Linear Algebra with Applications

Open Edition

Version 2021 — Revision A

**Be a champion of OER!**

Contribute suggestions for improvements, new content, or errata:

• A new topic

• A new example

• An interesting new question

• Any other suggestions to improve the material

**Contact Lyryx at info@lyryx.com with your ideas.**

****

**Creative Commons License (CC BY-NC-SA)**: This text, including the art and illustrations, are available under the Creative Commons license (CC BY-NC-SA), allowing anyone to reuse, revise, remix and redistribute the text.

To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/

Attribution

Linear Algebra with Applications Open Edition

Version 2021 — Revision A

To redistribute all of this book in its original form, please follow the guide below:

The front matter of the text should include a “License” page that includes the following statement.

This text is Linear Algebra with Applications by W. Keith Nicholson and Lyryx Learning Inc. View the text for free at https://lyryx.com/linear-algebra-applications/

To redistribute part of this book in its original form, please follow the guide below. Clearly indicate which content has been redistributed

The front matter of the text should include a “License” page that includes the following statement.

This text includes the following content from Linear Algebra with Applications by W. Keith Nicholson and Lyryx Learning Inc. View the entire text for free at https://lyryx.com/linear-algebra-applications/. <List of content from the original text>

The following must also be included at the beginning of the applicable content. Please clearly indicate which content has been redistributed from the Lyryx text.

This chapter is redistributed from the original Linear Algebra with Applications by W. Keith Nicholson and Lyryx Learning Inc. View the original text for free at https://lyryx.com/linear-algebra-applications/.

To adapt and redistribute all or part of this book in its original form, please follow the guide below. Clearly indicate which content has been adapted/redistributed and summarize the changes made.

The front matter of the text should include a “License” page that includes the following statement.

This text contains content adapted from the original Linear Algebra with Applications by W. Keith Nicholson and Lyryx Learning Inc. View the original text for free at https://lyryx.com/linear-algebra-applications/. <List of content and summary of changes>

The following must also be included at the beginning of the applicable content. Please clearly indicate which content has been adapted from the Lyryx text.

This chapter was adapted from the original Linear Algebra with Applications by W. Keith Nicholson and Lyryx Learning Inc. View the original text for free at https://lyryx.com/linear-algebra-applications/.

Citation

Use the information below to create a citation:

Author: W. Keith Nicholson

Publisher: Lyryx Learning Inc.

Book title: Linear Algebra with Applications

Book version: 2021A

Publication date: December 15, 2020

Location: Calgary, Alberta, Canada

Book URL: https://lyryx.com/linear-algebra-applications

For questions or comments please contact editorial@lyryx.com

Linear Algebra with Applications

Open Edition

**Base Text Revision History**

**Current Revision: Version 2021 — Revision A**

2021 A • Front matter has been updated including cover, Lyryx with Open Texts, copyright, and revision pages. Attribution page has been added.

• Typo and other minor fixes have been implemented throughout.

2019 A • New Section on Singular Value Decomposition (8.6) is included.

• New Example 2.3.2 and Theorem 2.2.4. Please note that this will impact the numbering of subsequent examples and theorems in the relevant sections.

• Section 2.2 is renamed as Matrix-Vector Multiplication.

• Minor revisions made throughout, including fixing typos, adding exercises, expanding explanations, and other small edits.

2018 B • Images have been converted to LaTeX throughout.

• Text has been converted to LaTeX with minor fixes throughout. Page numbers will differ from 2018A revision. Full index has been implemented.

2018 A • Text has been released with a Creative Commons license.

**Table of Contents**

Table of Contents iii Foreward vii Preface ix

1 Systems of Linear Equations 1 1.1 Solutions and Elementary Operations . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1 1.2 Gaussian Elimination . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9 1.3 Homogeneous Equations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20 1.4 An Application to Network Flow . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27 1.5 An Application to Electrical Networks . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29 1.6 An Application to Chemical Reactions . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31 Supplementary Exercises for Chapter 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

2 Matrix Algebra 35 2.1 Matrix Addition, Scalar Multiplication, and Transposition . . . . . . . . . . . . . . . . . . 35 2.2 Matrix-Vector Multiplication . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47 2.3 Matrix Multiplication . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64 2.4 Matrix Inverses . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79 2.5 Elementary Matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 94 2.6 Linear Transformations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103 2.7 LU-Factorization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 117 2.8 An Application to Input-Output Economic Models . . . . . . . . . . . . . . . . . . . . . 127 2.9 An Application to Markov Chains . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133 Supplementary Exercises for Chapter 2 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 141

3 Determinants and Diagonalization 143 3.1 The Cofactor Expansion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 143 3.2 Determinants and Matrix Inverses . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 156 3.3 Diagonalization and Eigenvalues . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 169 3.4 An Application to Linear Recurrences . . . . . . . . . . . . . . . . . . . . . . . . . . . . 190 3.5 An Application to Systems of Differential Equations . . . . . . . . . . . . . . . . . . . . 196 3.6 Proof of the Cofactor Expansion Theorem . . . . . . . . . . . . . . . . . . . . . . . . . . 202 Supplementary Exercises for Chapter 3 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 206

iii

iv Table of Contents

4 Vector Geometry 207 4.1 Vectors and Lines . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 207 4.2 Projections and Planes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 224 4.3 More on the Cross Product . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 242 4.4 Linear Operators on R3. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 248 4.5 An Application to Computer Graphics . . . . . . . . . . . . . . . . . . . . . . . . . . . . 255 Supplementary Exercises for Chapter 4 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 258

5 Vector Space Rn 261 5.1 Subspaces and Spanning . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 261 5.2 Independence and Dimension . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 269 5.3 Orthogonality . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 280 5.4 Rank of a Matrix . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 288 5.5 Similarity and Diagonalization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 296 5.6 Best Approximation and Least Squares . . . . . . . . . . . . . . . . . . . . . . . . . . . . 308 5.7 An Application to Correlation and Variance . . . . . . . . . . . . . . . . . . . . . . . . . 320 Supplementary Exercises for Chapter 5 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 325

6 Vector Spaces 327 6.1 Examples and Basic Properties . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 327 6.2 Subspaces and Spanning Sets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 336 6.3 Linear Independence and Dimension . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 343 6.4 Finite Dimensional Spaces . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 352 6.5 An Application to Polynomials . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 360 6.6 An Application to Differential Equations . . . . . . . . . . . . . . . . . . . . . . . . . . . 365 Supplementary Exercises for Chapter 6 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 370

7 Linear Transformations 371 7.1 Examples and Elementary Properties . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 371 7.2 Kernel and Image of a Linear Transformation . . . . . . . . . . . . . . . . . . . . . . . . 378 7.3 Isomorphisms and Composition . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 388 7.4 A Theorem about Differential Equations . . . . . . . . . . . . . . . . . . . . . . . . . . . 398 7.5 More on Linear Recurrences . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 401

8 Orthogonality 409 8.1 Orthogonal Complements and Projections . . . . . . . . . . . . . . . . . . . . . . . . . . 409 8.2 Orthogonal Diagonalization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 418 8.3 Positive Definite Matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 427 8.4 QR-Factorization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 431 8.5 Computing Eigenvalues . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 435

v

8.6 The Singular Value Decomposition . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 439 8.6.1 Singular Value Decompositions . . . . . . . . . . . . . . . . . . . . . . . . . . . 439 8.6.2 Fundamental Subspaces . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 445 8.6.3 The Polar Decomposition of a Real Square Matrix . . . . . . . . . . . . . . . . . 448 8.6.4 The Pseudoinverse of a Matrix . . . . . . . . . . . . . . . . . . . . . . . . . . . . 450

8.7 Complex Matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 454 8.8 An Application to Linear Codes over Finite Fields . . . . . . . . . . . . . . . . . . . . . . 465 8.9 An Application to Quadratic Forms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 479 8.10 An Application to Constrained Optimization . . . . . . . . . . . . . . . . . . . . . . . . . 489 8.11 An Application to Statistical Principal Component Analysis . . . . . . . . . . . . . . . . . 492

9 Change of Basis 495 9.1 The Matrix of a Linear Transformation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 495 9.2 Operators and Similarity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 504 9.3 Invariant Subspaces and Direct Sums . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 514

10 Inner Product Spaces 529 10.1 Inner Products and Norms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 529 10.2 Orthogonal Sets of Vectors . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 538 10.3 Orthogonal Diagonalization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 548 10.4 Isometries . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 555 10.5 An Application to Fourier Approximation . . . . . . . . . . . . . . . . . . . . . . . . . . 568

11 Canonical Forms 573 11.1 Block Triangular Form . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 573 11.2 The Jordan Canonical Form . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 581

A Complex Numbers 587 B Proofs 601 C Mathematical Induction 607 D Polynomials 613 Selected Exercise Answers 617 Index 653

**Foreward**

Mathematics education at the beginning university level is closely tied to the traditional publishers. In my opinion, it gives them too much control of both cost and content. The main goal of most publishers is profit, and the result has been a sales-driven business model as opposed to a pedagogical one. This results in frequent new “editions” of textbooks motivated largely to reduce the sale of used books rather than to update content quality. It also introduces copyright restrictions which stifle the creation and use of new pedagogical methods and materials. The overall result is high cost textbooks which may not meet the evolving educational needs of instructors and students.

To be fair, publishers do try to produce material that reflects new trends. But their goal is to sell books and not necessarily to create tools for student success in mathematics education. Sadly, this has led to a model where the primary choice for adapting to (or initiating) curriculum change is to find a different commercial textbook. My editor once said that the text that is adopted is often everyone’s third choice.

Of course instructors can produce their own lecture notes, and have done so for years, but this remains an onerous task. The publishing industry arose from the need to provide authors with copy-editing, edi torial, and marketing services, as well as extensive reviews of prospective customers to ascertain market trends and content updates. These are necessary skills and services that the industry continues to offer.

Authors of open educational resources (OER) including (but not limited to) textbooks and lecture notes, cannot afford this on their own. But they do have two great advantages: The cost to students is significantly lower, and open licenses return content control to instructors. Through editable file formats and open licenses, OER can be developed, maintained, reviewed, edited, and improved by a variety of contributors. Instructors can now respond to curriculum change by revising and reordering material to create content that meets the needs of their students. While editorial and quality control remain daunting tasks, great strides have been made in addressing the issues of accessibility, affordability and adaptability of the material.

For the above reasons I have decided to release my text under an open license, even though it was published for many years through a traditional publisher.

Supporting students and instructors in a typical classroom requires much more than a textbook. Thus, while anyone is welcome to use and adapt my text at no cost, I also decided to work closely with Lyryx Learning. With colleagues at the University of Calgary, I helped create Lyryx almost 20 years ago. The original idea was to develop quality online assessment (with feedback) well beyond the multiple-choice style then available. Now Lyryx also works to provide and sustain open textbooks; working with authors, contributors, and reviewers to ensure instructors need not sacrifice quality and rigour when switching to an open text.

I believe this is the right direction for mathematical publishing going forward, and look forward to being a part of how this new approach develops.

W. Keith Nicholson, Author

University of Calgary

vii

**Preface**

This textbook is an introduction to the ideas and techniques of linear algebra for first- or second-year students with a working knowledge of high school algebra. The contents have enough flexibility to present a traditional introduction to the subject, or to allow for a more applied course. Chapters 1–4 contain a one semester course for beginners whereas Chapters 5–9 contain a second semester course (see the Suggested Course Outlines below). The text is primarily about real linear algebra with complex numbers being mentioned when appropriate (reviewed in Appendix A). Overall, the aim of the text is to achieve a balance among computational skills, theory, and applications of linear algebra. Calculus is not a prerequisite; places where it is mentioned may be omitted.

As a rule, students of linear algebra learn by studying examples and solving problems. Accordingly, the book contains a variety of exercises (over 1200, many with multiple parts), ordered as to their difficulty. In addition, more than 375 solved examples are included in the text, many of which are computational in nature. The examples are also used to motivate (and illustrate) concepts and theorems, carrying the student from concrete to abstract. While the treatment is rigorous, proofs are presented at a level appropriate to the student and may be omitted with no loss of continuity. As a result, the book can be used to give a course that emphasizes computation and examples, or to give a more theoretical treatment (some longer proofs are deferred to the end of the Section).

Linear Algebra has application to the natural sciences, engineering, management, and the social sci ences as well as mathematics. Consequently, 18 optional “applications” sections are included in the text introducing topics as diverse as electrical networks, economic models, Markov chains, linear recurrences, systems of differential equations, and linear codes over finite fields. Additionally some applications (for example linear dynamical systems, and directed graphs) are introduced in context. The applications sec tions appear at the end of the relevant chapters to encourage students to browse.

**SUGGESTED COURSE OUTLINES**

This text includes the basis for a two-semester course in linear algebra.

• Chapters 1–4 provide a standard one-semester course of 35 lectures, including linear equations, ma trix algebra, determinants, diagonalization, and geometric vectors, with applications as time permits. At Calgary, we cover Sections 1.1–1.3, 2.1–2.6, 3.1–3.3, and 4.1–4.4 and the course is taken by all science and engineering students in their first semester. Prerequisites include a working knowledge of high school algebra (algebraic manipulations and some familiarity with polynomials); calculus is not required.

• Chapters 5–9 contain a second semester course including Rn, abstract vector spaces, linear trans formations (and their matrices), orthogonality, complex matrices (up to the spectral theorem) and applications. There is more material here than can be covered in one semester, and at Calgary we cover Sections 5.1–5.5, 6.1–6.4, 7.1–7.3, 8.1–8.7, and 9.1–9.3 with a couple of applications as time permits.

• Chapter 5 is a “bridging” chapter that introduces concepts like spanning, independence, and basis in the concrete setting of Rn, before venturing into the abstract in Chapter 6. The duplication is

ix

x Preface

balanced by the value of reviewing these notions, and it enables the student to focus in Chapter 6 on the new idea of an abstract system. Moreover, Chapter 5 completes the discussion of rank and diagonalization from earlier chapters, and includes a brief introduction to orthogonality in Rn, which creates the possibility of a one-semester, matrix-oriented course covering Chapter 1–5 for students not wanting to study the abstract theory.

**CHAPTER DEPENDENCIES**

The following chart suggests how the material introduced in each chapter draws on concepts covered in certain earlier chapters. A solid arrow means that ready assimilation of ideas and techniques presented in the later chapter depends on familiarity with the earlier chapter. A broken arrow indicates that some reference to the earlier chapter is made but the chapter need not be covered.

Chapter 1: Systems of Linear Equations

Chapter 2: Matrix Algebra

Chapter 3: Determinants and Diagonalization Chapter 4: Vector Geometry

Chapter 5: The Vector Space Rn

Chapter 6: Vector Spaces

Chapter 7: Linear Transformations Chapter 8: Orthogonality

Chapter 9: Change of Basis

Chapter 10: Inner Product Spaces Chapter 11: Canonical Forms

**HIGHLIGHTS OF THE TEXT**

• Two-stage definition of matrix multiplication. First, in Section 2.2 matrix-vector products are introduced naturally by viewing the left side of a system of linear equations as a product. Second, matrix-matrix products are defined in Section 2.3 by taking the columns of a product AB to be A times the corresponding columns of B. This is motivated by viewing the matrix product as compo sition of maps (see next item). This works well pedagogically and the usual dot-product definition follows easily. As a bonus, the proof of associativity of matrix multiplication now takes four lines.

• Matrices as transformations. Matrix-column multiplications are viewed (in Section 2.2) as trans formations Rn → Rm. These maps are then used to describe simple geometric reflections and rota tions in R2as well as systems of linear equations.

• Early linear transformations. It has been said that vector spaces exist so that linear transformations can act on them—consequently these maps are a recurring theme in the text. Motivated by the matrix transformations introduced earlier, linear transformations Rn → Rm are defined in Section 2.6, their standard matrices are derived, and they are then used to describe rotations, reflections, projections, and other operators on R2.

xi

• Early diagonalization. As requested by engineers and scientists, this important technique is pre sented in the first term using only determinants and matrix inverses (before defining independence and dimension). Applications to population growth and linear recurrences are given.

• Early dynamical systems. These are introduced in Chapter 3, and lead (via diagonalization) to applications like the possible extinction of species. Beginning students in science and engineering can relate to this because they can see (often for the first time) the relevance of the subject to the real world.

• Bridging chapter. Chapter 5 lets students deal with tough concepts (like independence, spanning, and basis) in the concrete setting of Rn before having to cope with abstract vector spaces in Chap ter 6.

• Examples. The text contains over 375 worked examples, which present the main techniques of the subject, illustrate the central ideas, and are keyed to the exercises in each section.

• Exercises. The text contains a variety of exercises (nearly 1175, many with multiple parts), starting with computational problems and gradually progressing to more theoretical exercises. Select solu tions are available at the end of the book or in the Student Solution Manual. There is a complete Solution Manual is available for instructors.

• Applications. There are optional applications at the end of most chapters (see the list below). While some are presented in the course of the text, most appear at the end of the relevant chapter to encourage students to browse.

• Appendices. Because complex numbers are needed in the text, they are described in Appendix A, which includes the polar form and roots of unity. Methods of proofs are discussed in Appendix B, followed by mathematical induction in Appendix C. A brief discussion of polynomials is included in Appendix D. All these topics are presented at the high-school level.

• Self-Study. This text is self-contained and therefore is suitable for self-study.

• Rigour. Proofs are presented as clearly as possible (some at the end of the section), but they are optional and the instructor can choose how much he or she wants to prove. However the proofs are there, so this text is more rigorous than most. Linear algebra provides one of the better venues where students begin to think logically and argue concisely. To this end, there are exercises that ask the student to “show” some simple implication, and others that ask her or him to either prove a given statement or give a counterexample. I personally present a few proofs in the first semester course and more in the second (see the Suggested Course Outlines).

• Major Theorems. Several major results are presented in the book. Examples: Uniqueness of the reduced row-echelon form; the cofactor expansion for determinants; the Cayley-Hamilton theorem; the Jordan canonical form; Schur’s theorem on block triangular form; the principal axes and spectral theorems; and others. Proofs are included because the stronger students should at least be aware of what is involved.

xii Preface

**CHAPTER SUMMARIES**

**Chapter 1: Systems of Linear Equations.**

A standard treatment of gaussian elimination is given. The rank of a matrix is introduced via the row echelon form, and solutions to a homogeneous system are presented as linear combinations of basic solu tions. Applications to network flows, electrical networks, and chemical reactions are provided.

**Chapter 2: Matrix Algebra.**

After a traditional look at matrix addition, scalar multiplication, and transposition in Section 2.1, matrix vector multiplication is introduced in Section 2.2 by viewing the left side of a system of linear equations as the product Ax of the coefficient matrix A with the column x of variables. The usual dot-product definition of a matrix-vector multiplication follows. Section 2.2 ends by viewing an m×n matrix A as a transformation Rn → Rm. This is illustrated for R2 → R2 by describing reflection in the x axis, rotation of R2through π2, shears, and so on.

In Section 2.3, the product of matrices A and B is defined by AB =Ab1 Ab2 ··· Abn, where the bi are the columns of B. A routine computation shows that this is the matrix of the transformation B followed by A. This observation is used frequently throughout the book, and leads to simple, conceptual proofs of the basic axioms of matrix algebra. Note that linearity is not required—all that is needed is some basic properties of matrix-vector multiplication developed in Section 2.2. Thus the usual arcane definition of matrix multiplication is split into two well motivated parts, each an important aspect of matrix algebra. Of course, this has the pedagogical advantage that the conceptual power of geometry can be invoked to illuminate and clarify algebraic techniques and definitions.

In Section 2.4 and 2.5 matrix inverses are characterized, their geometrical meaning is explored, and block multiplication is introduced, emphasizing those cases needed later in the book. Elementary ma trices are discussed, and the Smith normal form is derived. Then in Section 2.6, linear transformations Rn → Rm are defined and shown to be matrix transformations. The matrices of reflections, rotations, and projections in the plane are determined. Finally, matrix multiplication is related to directed graphs, matrix LU-factorization is introduced, and applications to economic models and Markov chains are presented.

**Chapter 3: Determinants and Diagonalization.**

The cofactor expansion is stated (proved by induction later) and used to define determinants inductively and to deduce the basic rules. The product and adjugate theorems are proved. Then the diagonalization algorithm is presented (motivated by an example about the possible extinction of a species of birds). As requested by our Engineering Faculty, this is done earlier than in most texts because it requires only deter minants and matrix inverses, avoiding any need for subspaces, independence and dimension. Eigenvectors of a 2 × 2 matrix A are described geometrically (using the A-invariance of lines through the origin). Di agonalization is then used to study discrete linear dynamical systems and to discuss applications to linear recurrences and systems of differential equations. A brief discussion of Google PageRank is included.

xiii

**Chapter 4: Vector Geometry.**

Vectors are presented intrinsically in terms of length and direction, and are related to matrices via coordi nates. Then vector operations are defined using matrices and shown to be the same as the corresponding intrinsic definitions. Next, dot products and projections are introduced to solve problems about lines and planes. This leads to the cross product. Then matrix transformations are introduced in R3, matrices of pro jections and reflections are derived, and areas and volumes are computed using determinants. The chapter closes with an application to computer graphics.

**Chapter 5: The Vector Space** Rn**.**

Subspaces, spanning, independence, and dimensions are introduced in the context of Rnin the first two sections. Orthogonal bases are introduced and used to derive the expansion theorem. The basic properties of rank are presented and used to justify the definition given in Section 1.2. Then, after a rigorous study of diagonalization, best approximation and least squares are discussed. The chapter closes with an application to correlation and variance.

This is a “bridging” chapter, easing the transition to abstract spaces. Concern about duplication with Chapter 6 is mitigated by the fact that this is the most difficult part of the course and many students welcome a repeat discussion of concepts like independence and spanning, albeit in the abstract setting. In a different direction, Chapter 1–5 could serve as a solid introduction to linear algebra for students not requiring abstract theory.

**Chapter 6: Vector Spaces.**

Building on the work on Rnin Chapter 5, the basic theory of abstract finite dimensional vector spaces is developed emphasizing new examples like matrices, polynomials and functions. This is the first acquain tance most students have had with an abstract system, so not having to deal with spanning, independence and dimension in the general context eases the transition to abstract thinking. Applications to polynomials and to differential equations are included.

**Chapter 7: Linear Transformations.**

General linear transformations are introduced, motivated by many examples from geometry, matrix theory, and calculus. Then kernels and images are defined, the dimension theorem is proved, and isomorphisms are discussed. The chapter ends with an application to linear recurrences. A proof is included that the order of a differential equation (with constant coefficients) equals the dimension of the space of solutions.

**Chapter 8: Orthogonality.**

The study of orthogonality in Rn, begun in Chapter 5, is continued. Orthogonal complements and pro jections are defined and used to study orthogonal diagonalization. This leads to the principal axes theo rem, the Cholesky factorization of a positive definite matrix, QR-factorization, and to a discussion of the singular value decomposition, the polar form, and the pseudoinverse. The theory is extended to Cnin Section 8.7 where hermitian and unitary matrices are discussed, culminating in Schur’s theorem and the spectral theorem. A short proof of the Cayley-Hamilton theorem is also presented. In Section 8.8 the field Zp of integers modulo p is constructed informally for any prime p, and codes are discussed over any finite field. The chapter concludes with applications to quadratic forms, constrained optimization, and statistical principal component analysis.

xiv Preface

**Chapter 9: Change of Basis.**

The matrix of general linear transformation is defined and studied. In the case of an operator, the rela tionship between basis changes and similarity is revealed. This is illustrated by computing the matrix of a rotation about a line through the origin in R3. Finally, invariant subspaces and direct sums are introduced, related to similarity, and (as an example) used to show that every involution is similar to a diagonal matrix with diagonal entries ±1.

**Chapter 10: Inner Product Spaces.**

General inner products are introduced and distance, norms, and the Cauchy-Schwarz inequality are dis cussed. The Gram-Schmidt algorithm is presented, projections are defined and the approximation theorem is proved (with an application to Fourier approximation). Finally, isometries are characterized, and dis tance preserving operators are shown to be composites of a translations and isometries.

**Chapter 11: Canonical Forms.**

The work in Chapter 9 is continued. Invariant subspaces and direct sums are used to derive the block triangular form. That, in turn, is used to give a compact proof of the Jordan canonical form. Of course the level is higher.

**Appendices**

In Appendix A, complex arithmetic is developed far enough to find nth roots. In Appendix B, methods of proof are discussed, while Appendix C presents mathematical induction. Finally, Appendix D describes the properties of polynomials in elementary terms.

**LIST OF APPLICATIONS**

• Network Flow (Section 1.4)

• Electrical Networks (Section 1.5)

• Chemical Reactions (Section 1.6)

• Directed Graphs (in Section 2.3)

• Input-Output Economic Models (Section 2.8)

• Markov Chains (Section 2.9)

• Polynomial Interpolation (in Section 3.2)

• Population Growth (Examples 3.3.1 and 3.3.12, Section 3.3)

• Google PageRank (in Section 3.3)

• Linear Recurrences (Section 3.4; see also Section 7.5)

• Systems of Differential Equations (Section 3.5)

• Computer Graphics (Section 4.5)

xv

• Least Squares Approximation (in Section 5.6)

• Correlation and Variance (Section 5.7)

• Polynomials (Section 6.5)

• Differential Equations (Section 6.6)

• Linear Recurrences (Section 7.5)

• Error Correcting Codes (Section 8.8)

• Quadratic Forms (Section 8.9)

• Constrained Optimization (Section 8.10)

• Statistical Principal Component Analysis (Section 8.11)

• Fourier Approximation (Section 10.5)

**ACKNOWLEDGMENTS**

Many colleagues have contributed to the development of this text over many years of publication, and I specially thank the following instructors for their reviews of the 7th edition:

Robert Andre

University of Waterloo

Dietrich Burbulla

University of Toronto

Dzung M. Ha

Ryerson University

Mark Solomonovich

Grant MacEwan

Fred Szabo

Concordia University

Edward Wang

Wilfred Laurier

Petr Zizler

Mount Royal University

It is also a pleasure to recognize the contributions of several people. Discussions with Thi Dinh and Jean Springer have been invaluable and many of their suggestions have been incorporated. Thanks are also due to Kristine Bauer and Clifton Cunningham for several conversations about the new way to look at matrix multiplication. I also wish to extend my thanks to Joanne Canape for being there when I had technical questions. Thanks also go to Jason Nicholson for his help in various aspects of the book, partic ularly the Solutions Manual. Finally, I want to thank my wife Kathleen, without whose understanding and cooperation, this book would not exist.

As we undertake this new publishing model with the text as an open educational resource, I would also like to thank my previous publisher. The team who supported my text greatly contributed to its success.

xvi Preface

Now that the text has an open license, we have a much more fluid and powerful mechanism to incorpo rate comments and suggestions. The editorial group at Lyryx invites instructors and students to contribute to the text, and also offers to provide adaptations of the material for specific courses. Moreover the LaTeX source files are available to anyone wishing to do the adaptation and editorial work themselves!

W. Keith Nicholson

University of Calgary

**Chapter 1**

**Systems of Linear Equations**

**1.1 Solutions and Elementary Operations**

Practical problems in many fields of study—such as biology, business, chemistry, computer science, eco nomics, electronics, engineering, physics and the social sciences—can often be reduced to solving a sys tem of linear equations. Linear algebra arose from attempts to find systematic methods for solving these systems, so it is natural to begin this book by studying linear equations.

If a, b, and c are real numbers, the graph of an equation of the form

ax+by = c

is a straight line (if a and b are not both zero), so such an equation is called a linear equation in the variables x and y. However, it is often convenient to write the variables as x1, x2, ..., xn, particularly when more than two variables are involved. An equation of the form

a1x1 +a2x2 +···+anxn = b

is called a linear equation in the n variables x1, x2, ..., xn. Here a1, a2, ..., an denote real numbers (called the coefficients of x1, x2, ..., xn, respectively) and b is also a number (called the constant term of the equation). A finite collection of linear equations in the variables x1, x2, ..., xn is called a system of linear equations in these variables. Hence,

2x1 −3x2 +5x3 = 7

is a linear equation; the coefficients of x1, x2, and x3 are 2, −3, and 5, and the constant term is 7. Note that each variable in a linear equation occurs to the first power only.

Given a linear equation a1x1 +a2x2 +···+anxn = b, a sequence s1, s2, ..., sn of n numbers is called a solution to the equation if

a1s1 +a2s2 +···+ansn = b

that is, if the equation is satisfied when the substitutions x1 = s1, x2 = s2, ..., xn = sn are made. A sequence of numbers is called a solution to a system of equations if it is a solution to every equation in the system.

For example, x = −2, y = 5, z = 0 and x = 0, y = 4, z = −1 are both solutions to the system

x + y + z = 3

2x + y + 3z = 1

A system may have no solution at all, or it may have a unique solution, or it may have an infinite family of solutions. For instance, the system x+y = 2, x+y = 3 has no solution because the sum of two numbers cannot be 2 and 3 simultaneously. A system that has no solution is called inconsistent; a system with at least one solution is called consistent. The system in the following example has infinitely many solutions.

1

2 Systems of Linear Equations

Example 1.1.1

Show that, for arbitrary values of s and t,

x1 = t −s+1

x2 = t +s+2

x3 = s

x4 = t

is a solution to the system

x1 − 2x2 +3x3 +x4 = −3

2x1 − x2 +3x3 −x4 = 0

Solution. Simply substitute these values of x1, x2, x3, and x4 in each equation.

x1 −2x2 +3x3 +x4 = (t −s+1)−2(t +s+2) +3s+t = −3

2x1 −x2 +3x3 −x4 = 2(t −s+1)−(t +s+2) +3s−t = 0

Because both equations are satisfied, it is a solution for all choices of s and t.

The quantities s and t in Example 1.1.1 are called parameters, and the set of solutions, described in this way, is said to be given in parametric form and is called the general solution to the system. It turns out that the solutions to every system of equations (if there are solutions) can be given in parametric form (that is, the variables x1, x2, ... are given in terms of new independent variables s, t, etc.). The following example shows how this happens in the simplest systems where only one equation is present.

Example 1.1.2

Describe all solutions to 3x−y+2z = 6 in parametric form.

Solution. Solving the equation for y in terms of x and z, we get y = 3x+2z−6. If s and t are arbitrary then, setting x = s, z = t, we get solutions

x = s

y = 3s+2t −6 s and t arbitrary

z = t

Of course we could have solved for x: x =13(y−2z+6). Then, if we take y = p, z = q, the solutions are represented as follows:

x =13(p−2q+6)

y = p p and q arbitrary

z = q

The same family of solutions can “look” quite different!

y

y y

x−y = 1

x+y = 3

P(2, 1)

x

(a) Unique Solution (x = 2, y = 1)

x+y = 4

x+y = 2

x

(b) No Solution

−6x+2y = −8

3x−y = 4

x

1.1. Solutions and Elementary Operations 3

When only two variables are involved, the solutions to systems of lin ear equations can be described geometrically because the graph of a lin ear equation ax + by = c is a straight line if a and b are not both zero. Moreover, a point P(s, t) with coordinates s and t lies on the line if and only if as + bt = c—that is when x = s, y = t is a solution to the equa tion. Hence the solutions to a system of linear equations correspond to the points P(s, t) that lie on all the lines in question.

In particular, if the system consists of just one equation, there must be infinitely many solutions because there are infinitely many points on a line. If the system has two equations, there are three possibilities for the corresponding straight lines:

1. The lines intersect at a single point. Then the system has a unique solution corresponding to that point.

2. The lines are parallel (and distinct) and so do not intersect. Then the system has no solution.

3. The lines are identical. Then the system has infinitely many solutions—one for each point on the (common) line.

These three situations are illustrated in Figure 1.1.1. In each case the graphs of two specific lines are plotted and the corresponding equations are indicated. In the last case, the equations are 3x−y = 4 and −6x+2y = −8, which have identical graphs.

With three variables, the graph of an equation ax+by+cz = d can be shown to be a plane (see Section 4.2) and so again provides a “picture” of the set of solutions. However, this graphical method has its limitations: When more than three variables are involved, no physical image of the graphs (called hyperplanes) is possible. It is necessary to turn to a more

(c) Infinitely many solutions (x = t, y = 3t −4)

Figure 1.1.1

“algebraic” method of solution.

Before describing the method, we introduce a concept that simplifies the computations involved. Consider the following system

3x1 + 2x2 − x3 + x4 = −1

2x1 − x3 + 2x4 = 0

3x1 + x2 + 2x3 + 5x4 = 2

of three equations in four variables. The array of numbers1



3 2 −1 1 −1 

2 0 −1 2 0 3 1 2 5 2

 

occurring in the system is called the augmented matrix of the system. Each row of the matrix consists of the coefficients of the variables (in order) from the corresponding equation, together with the constant

1A rectangular array of numbers is called a matrix. Matrices will be discussed in more detail in Chapter 2.

4 Systems of Linear Equations

term. For clarity, the constants are separated by a vertical line. The augmented matrix is just a different way of describing the system of equations. The array of coefficients of the variables

 

3 2 −1 1 2 0 −1 2 3 1 2 5

 





is called the coefficient matrix of the system and **Elementary Operations**

−1 0

2



 is called the constant matrix of the system.

The algebraic method for solving systems of linear equations is described as follows. Two such systems are said to be equivalent if they have the same set of solutions. A system is solved by writing a series of systems, one after the other, each equivalent to the previous system. Each of these systems has the same set of solutions as the original one; the aim is to end up with a system that is easy to solve. Each system in the series is obtained from the preceding system by a simple manipulation chosen so that it does not change the set of solutions.

As an illustration, we solve the system x + 2y = −2, 2x + y = 7 in this manner. At each stage, the corresponding augmented matrix is displayed. The original system is

x + 2y = −2 2x + y = 7

1 2 −2 2 1 7

First, subtract twice the first equation from the second. The resulting system is

x + 2y = −2 − 3y = 11

1 2 −2 0 −3 11

which is equivalent to the original (see Theorem 1.1.1). At this stage we obtain y = −113by multiplying the second equation by −13. The result is the equivalent system

x+2y = −2 y = −113

1 2 −2 0 1 −113

Finally, we subtract twice the second equation from the first to get another equivalent system.

x =163 y = −113



1 0 163 0 1 −113

 

Now this system is easy to solve! And because it is equivalent to the original system, it provides the solution to that system.

Observe that, at each stage, a certain operation is performed on the system (and thus on the augmented matrix) to produce an equivalent system.

1.1. Solutions and Elementary Operations 5

Definition 1.1 Elementary Operations

The following operations, called elementary operations, can routinely be performed on systems of linear equations to produce equivalent systems.

I. Interchange two equations.

II. Multiply one equation by a nonzero number.

III. Add a multiple of one equation to a different equation.

Theorem 1.1.1

Suppose that a sequence of elementary operations is performed on a system of linear equations. Then the resulting system has the same set of solutions as the original, so the two systems are equivalent.

The proof is given at the end of this section.

Elementary operations performed on a system of equations produce corresponding manipulations of the rows of the augmented matrix. Thus, multiplying a row of a matrix by a number k means multiplying every entry of the row by k. Adding one row to another row means adding each entry of that row to the corresponding entry of the other row. Subtracting two rows is done similarly. Note that we regard two rows as equal when corresponding entries are the same.

In hand calculations (and in computer programs) we manipulate the rows of the augmented matrix rather than the equations. For this reason we restate these elementary operations for matrices.

Definition 1.2 Elementary Row Operations

The following are called elementary row operations on a matrix.

I. Interchange two rows.

II. Multiply one row by a nonzero number.

III. Add a multiple of one row to a different row.

In the illustration above, a series of such operations led to a matrix of the form

1 0 ∗ 0 1 ∗

where the asterisks represent arbitrary numbers. In the case of three equations in three variables, the goal

is to produce a matrix of the form 

1 0 0 ∗ 0 1 0 ∗ 0 0 1 ∗

 

This does not always happen, as we will see in the next section. Here is an example in which it does happen.

6 Systems of Linear Equations

Example 1.1.3

Find all solutions to the following system of equations.

3x + 4y + z = 1

2x + 3y = 0

4x + 3y − z = −2

Solution. The augmented matrix of the original system is

 

3 4 1 1 2 3 0 0 4 3 −1 −2

 

To create a 1 in the upper left corner we could multiply row 1 through by 13. However, the 1 can be obtained without introducing fractions by subtracting row 2 from row 1. The result is

 

1 1 1 1 2 3 0 0 4 3 −1 −2

 

The upper left 1 is now used to “clean up” the first column, that is create zeros in the other positions in that column. First subtract 2 times row 1 from row 2 to obtain

 

1 1 1 1 0 1 −2 −2 4 3 −1 −2

 

Next subtract 4 times row 1 from row 3. The result is



1 1 1 1 

0 1 −2 −2 0 −1 −5 −6

 

This completes the work on column 1. We now use the 1 in the second position of the second row to clean up the second column by subtracting row 2 from row 1 and then adding row 2 to row 3. For convenience, both row operations are done in one step. The result is

 

1 0 3 3 0 1 −2 −2 0 0 −7 −8

 

Note that the last two manipulations did not affect the first column (the second row has a zero there), so our previous effort there has not been undermined. Finally we clean up the third column. Begin by multiplying row 3 by −17to obtain

 

1 0 3 3 0 1 −2 −2 0 0 1 87

 

1.1. Solutions and Elementary Operations 7

Now subtract 3 times row 3 from row 1, and then add 2 times row 3 to row 2 to get



1 0 0 −37

0 1 0 27 0 0 1 8~~7~~

 

The corresponding equations are x = −37, y =27, and z =87, which give the (unique) solution.

Every elementary row operation can be reversed by another elementary row operation of the same type (called its inverse). To see how, we look at types I, II, and III separately:

Type I Interchanging two rows is reversed by interchanging them again.

Type II Multiplying a row by a nonzero number k is reversed by multiplying by 1/k.

Type III Adding k times row p to a different row q is reversed by adding −k times row p to row q (in the new matrix). Note that p 6= q is essential here.

To illustrate the Type III situation, suppose there are four rows in the original matrix, denoted R1, R2, R3, and R4, and that k times R2 is added to R3. Then the reverse operation adds −k times R2, to R3. The following diagram illustrates the effect of doing the operation first and then the reverse:



R1

R2

R3

R4



→



R1

R2

R3 +kR2 R4



→



R1

R2

(R3 +kR2)−kR2 R4



=



R1

R2

R3

R4

 

The existence of inverses for elementary row operations and hence for elementary operations on a system of equations, gives:

Proof of Theorem 1.1.1. Suppose that a system of linear equations is transformed into a new system by a sequence of elementary operations. Then every solution of the original system is automatically a solution of the new system because adding equations, or multiplying an equation by a nonzero number, always results in a valid equation. In the same way, each solution of the new system must be a solution to the original system because the original system can be obtained from the new one by another series of elementary operations (the inverses of the originals). It follows that the original and new systems have the

same solutions. This proves Theorem 1.1.1.

|  |
| --- |

8 Systems of Linear Equations

**Exercises for 1.1**

Exercise 1.1.1 In each case verify that the following are solutions for all values of s and t.

Exercise 1.1.7 Write the augmented matrix for each of the following systems of linear equations.

a. x + 2y = 0

a. x = 19t −35

x − 3y = 5 2x + y = 1

b.

y = 1

c. x + y = 1

y = 25−13t z = t

is a solution of

x − y + z = 2 x − z = 1

y + 2x = 0

d.

y + z = 0

z − x = 2

2x + 3y + z = 5 5x + 7y − 4z = 0

Exercise 1.1.8 Write a system of linear equations that has each of the following augmented matrices.



1 −1 6 0





2 −1 0 −1



a.  

b. 

b. x1 = 2s+12t +13 x2 = s

0 1 0 3 2 −1 0 1



−3 2 1 0 0 1 1 3

x3 = −s−3t −3 x4 = t

is a solution of

Exercise 1.1.9 Find the solution of each of the following systems of linear equations using augmented matrices.

a. x + 2y = 1

2x1 + 5x2 + 9x3 + 3x4 = −1

x − 3y = 1 2x − 7y = 3

b.

3x + 4y = −1

x1 + 2x2 + 4x3 = 1

c. 3x + 4y = 1

2x + 3y = −1 3x + 4y = 2

d.

4x + 5y = −3

Exercise 1.1.2 Find all solutions to the following in parametric form in two ways.

Exercise 1.1.10 Find the solution of each of the follow ing systems of linear equations using augmented matri ces.

a. 2x + y + z = −1

a. 2 3x+y = 2 b. x+3y = 1 c. 3x−y+2z = 5 d. x−2y+5z = 1

x + y + 2z = −1 2x + y + 3z = 0 − 2y + z = 2

b.

x + 2y + z = 0 3x − 2z = 5

Exercise 1.1.3 Regarding 2x = 5 as the equation 2x + 0y = 5 in two variables, find all solutions in para

Exercise 1.1.11 Find all solutions (if any) of the follow ing systems of linear equations.

a. 3x−2y = 5

metric form.

Exercise 1.1.4 Regarding 4x − 2y = 3 as the equation

3x−2y = 5 −12x+8y = −20

b.

−12x+8y = 16

4x − 2y + 0z = 3 in three variables, find all solutions in

Exercise 1.1.12 Show that the system



parametric form.

Exercise 1.1.5 Find all solutions to the general system



ax = b of one equation in one variable (a) when a = 0

x + 2y − z = a 2x + y + 3z = b x − 4y + 9z = c

and (b) when a 6= 0.

Exercise 1.1.6 Show that a system consisting of exactly one linear equation can have no solution, one solution, or infinitely many solutions. Give examples.

is inconsistent unless c = 2b−3a.

Exercise 1.1.13 By examining the possible positions of lines in the plane, show that two equations in two vari ables can have zero, one, or infinitely many solutions.

Exercise 1.1.14 In each case either show that the state ment is true, or give an example2showing it is false.

a. If a linear system has n variables and m equations, then the augmented matrix has n rows.

b. A consistent linear system must have infinitely many solutions.

c. If a row operation is done to a consistent linear system, the resulting system must be consistent.

d. If a series of row operations on a linear system re sults in an inconsistent system, the original system is inconsistent.

Exercise 1.1.15 Find a quadratic a+bx+cx2such that the graph of y = a+bx+cx2contains each of the points (−1, 6), (2, 0), and (3, 2).

3x + 2y = 5

1.2. Gaussian Elimination 9

Exercise 1.1.17 Find a, b, and c such that x2−x+3

(x2+2)(2x−1)=ax+b

x2+2+c

2x−1

[Hint: Multiply through by (x2 + 2)(2x− 1) and equate coefficients of powers of x.]

Exercise 1.1.18 A zookeeper wants to give an animal 42 mg of vitamin A and 65 mg of vitamin D per day. He has two supplements: the first contains 10% vitamin A and 25% vitamin D; the second contains 20% vitamin A and 25% vitamin D. How much of each supplement should he give the animal each day?

Exercise 1.1.19 Workmen John and Joe earn a total of $24.60 when John works 2 hours and Joe works 3 hours. If John works 3 hours and Joe works 2 hours, they get $23.90. Find their hourly rates.

Exercise 1.1.20 A biologist wants to create a diet from

Exercise 1.1.16 Solve the system x = 5x′ − 2y′

7x + 5y = 1by

fish and meal containing 183 grams of protein and 93 grams of carbohydrate per day. If fish contains 70% pro

changing variables

y = −7x′ + 3y′and solving the re

tein and 10% carbohydrate, and meal contains 30% pro

sulting equations for x′and y′.

**1.2 Gaussian Elimination**

tein and 60% carbohydrate, how much of each food is required each day?

The algebraic method introduced in the preceding section can be summarized as follows: Given a system of linear equations, use a sequence of elementary row operations to carry the augmented matrix to a “nice” matrix (meaning that the corresponding equations are easy to solve). In Example 1.1.3, this nice matrix

took the form 

1 0 0 ∗ 0 1 0 ∗ 0 0 1 ∗

 

The following definitions identify the nice matrices that arise in this process.

2Such an example is called a counterexample. For example, if the statement is that “all philosophers have beards”, the existence of a non-bearded philosopher would be a counterexample proving that the statement is false. This is discussed again in Appendix B.

10 Systems of Linear Equations

Definition 1.3 Row-Echelon Form (Reduced)

A matrix is said to be in row-echelon form (and will be called a row-echelon matrix) if it satisfies the following three conditions:

1. All zero rows (consisting entirely of zeros) are at the bottom.

2. The first nonzero entry from the left in each nonzero row is a 1, called the leading 1 for that row.

3. Each leading 1 is to the right of all leading 1s in the rows above it.

A row-echelon matrix is said to be in reduced row-echelon form (and will be called a reduced row-echelon matrix) if, in addition, it satisfies the following condition:

4. Each leading 1 is the only nonzero entry in its column.

The row-echelon matrices have a “staircase” form, as indicated by the following example (the asterisks

indicate arbitrary numbers).



0 1 ∗ ∗ ∗ ∗ ∗



0 0 0 1 ∗ ∗ ∗

0 0 0 0 1 ∗ ∗

0 0 0 0 0 0 1

0 0 0 0 0 0 0

The leading 1s proceed “down and to the right” through the matrix. Entries above and to the right of the leading 1s are arbitrary, but all entries below and to the left of them are zero. Hence, a matrix in row echelon form is in reduced form if, in addition, the entries directly above each leading 1 are all zero. Note that a matrix in row-echelon form can, with a few more row operations, be carried to reduced form (use row operations to create zeros above each leading one in succession, beginning from the right).

Example 1.2.1

The following matrices are in row-echelon form (for any choice of numbers in ∗-positions).

1 ∗ ∗ 0 0 1

 

0 1 ∗ ∗ 0 0 1 ∗ 0 0 0 0

 

 

1 ∗ ∗ ∗ 0 1 ∗ ∗ 0 0 0 1

 

 

1 ∗ ∗ 0 1 ∗ 0 0 1

 

The following, on the other hand, are in reduced row-echelon form.

1 ∗ 0 0 0 1

 

0 1 0 ∗ 0 0 1 ∗ 0 0 0 0

 

 

1 0 ∗ 0 0 1 ∗ 0 0 0 0 1

 

 

1 0 0 0 1 0 0 0 1

 

The choice of the positions for the leading 1s determines the (reduced) row-echelon form (apart from the numbers in ∗-positions).

The importance of row-echelon matrices comes from the following theorem.

1.2. Gaussian Elimination 11

Theorem 1.2.1

Every matrix can be brought to (reduced) row-echelon form by a sequence of elementary row operations.

In fact we can give a step-by-step procedure for actually finding a row-echelon matrix. Observe that while there are many sequences of row operations that will bring a matrix to row-echelon form, the one we use is systematic and is easy to program on a computer. Note that the algorithm deals with matrices in general, possibly with columns of zeros.

Gaussian3Algorithm4

Step 1. If the matrix consists entirely of zeros, stop—it is already in row-echelon form.

Step 2. Otherwise, find the first column from the left containing a nonzero entry (call it a), and move the row containing that entry to the top position.

Step 3. Now multiply the new top row by 1/a to create a leading 1.

Step 4. By subtracting multiples of that row from rows below it, make each entry below the leading 1 zero.

This completes the first row, and all further row operations are carried out on the remaining rows. Step 5. Repeat steps 1–4 on the matrix consisting of the remaining rows.

The process stops when either no rows remain at step 5 or the remaining rows consist entirely of zeros.

Observe that the gaussian algorithm is recursive: When the first leading 1 has been obtained, the procedure is repeated on the remaining rows of the matrix. This makes the algorithm easy to use on a computer. Note that the solution to Example 1.1.3 did not use the gaussian algorithm as written because the first leading 1 was not created by dividing row 1 by 3. The reason for this is that it avoids fractions. However, the general pattern is clear: Create the leading 1s from left to right, using each of them in turn to create zeros below it. Here are two more examples.

Example 1.2.2

Solve the following system of equations.

3x + y − 4z = −1

x + 10z = 5

4x + y + 6z = 1

3Carl Friedrich Gauss (1777–1855) ranks with Archimedes and Newton as one of the three greatest mathematicians of all time. He was a child prodigy and, at the age of 21, he gave the first proof that every polynomial has a complex root. In 1801 he published a timeless masterpiece, Disquisitiones Arithmeticae, in which he founded modern number theory. He went on to make ground-breaking contributions to nearly every branch of mathematics, often well before others rediscovered and published the results.

4The algorithm was known to the ancient Chinese.

12 Systems of Linear Equations

Solution. The corresponding augmented matrix is

 

3 1 −4 −1 1 0 10 5 4 1 6 1

 

Create the first leading one by interchanging rows 1 and 2

 

1 0 10 5 3 1 −4 −1 4 1 6 1

 

Now subtract 3 times row 1 from row 2, and subtract 4 times row 1 from row 3. The result is



1 0 10 5 

0 1 −34 −16 0 1 −34 −19

 

Now subtract row 2 from row 3 to obtain



1 0 10 5 

0 1 −34 −16 0 0 0 −3

 

This means that the following reduced system of equations

x + 10z = 5

y − 34z = −16

0 = −3

is equivalent to the original system. In other words, the two have the same solutions. But this last system clearly has no solution (the last equation requires that x, y and z satisfy 0x+0y+0z = −3, and no such numbers exist). Hence the original system has no solution.

Example 1.2.3

Solve the following system of equations.

x1 − 2x2 − x3 + 3x4 = 1

2x1 − 4x2 + x3 = 5

x1 − 2x2 + 2x3 − 3x4 = 4

Solution. The augmented matrix is 

1 −2 −1 3 1 2 −4 1 0 5 1 −2 2 −3 4

 

1.2. Gaussian Elimination 13

Subtracting twice row 1 from row 2 and subtracting row 1 from row 3 gives



1 −2 −1 3 1 

0 0 3 −6 3 0 0 3 −6 3

 

Now subtract row 2 from row 3 and multiply row 2 by 13to get



1 −2 −1 3 1 

0 0 1 −2 1 0 0 0 0 0

 

This is in row-echelon form, and we take it to reduced form by adding row 2 to row 1:



1 −2 0 1 2 

0 0 1 −2 1 0 0 0 0 0

 

The corresponding reduced system of equations is

x1 − 2x2 + x4 = 2

x3 − 2x4 = 1

0 = 0

The leading ones are in columns 1 and 3 here, so the corresponding variables x1 and x3 are called leading variables. Because the matrix is in reduced row-echelon form, these equations can be used to solve for the leading variables in terms of the nonleading variables x2 and x4. More precisely, in the present example we set x2 = s and x4 = t where s and t are arbitrary, so these equations become

x1 −2s+t = 2 and x3 −2t = 1

Finally the solutions are given by

x1 = 2+2s−t

x2 = s

x3 = 1+2t

x4 = t

where s and t are arbitrary.

The solution of Example 1.2.3 is typical of the general case. To solve a linear system, the augmented matrix is carried to reduced row-echelon form, and the variables corresponding to the leading ones are called leading variables. Because the matrix is in reduced form, each leading variable occurs in exactly one equation, so that equation can be solved to give a formula for the leading variable in terms of the nonleading variables. It is customary to call the nonleading variables “free” variables, and to label them by new variables s, t, ..., called parameters. Hence, as in Example 1.2.3, every variable xiis given by a formula in terms of the parameters s and t. Moreover, every choice of these parameters leads to a solution

14 Systems of Linear Equations

to the system, and every solution arises in this way. This procedure works in general, and has come to be called

Gaussian Elimination

To solve a system of linear equations proceed as follows:

1. Carry the augmented matrix to a reduced row-echelon matrix using elementary row operations.

2. If a row0 0 0 ··· 0 1

occurs, the system is inconsistent.

3. Otherwise, assign the nonleading variables (if any) as parameters, and use the equations corresponding to the reduced row-echelon matrix to solve for the leading variables in terms of the parameters.

There is a variant of this procedure, wherein the augmented matrix is carried only to row-echelon form. The nonleading variables are assigned as parameters as before. Then the last equation (corresponding to the row-echelon form) is used to solve for the last leading variable in terms of the parameters. This last leading variable is then substituted into all the preceding equations. Then, the second last equation yields the second last leading variable, which is also substituted back. The process continues to give the general solution. This procedure is called back-substitution. This procedure can be shown to be numerically more efficient and so is important when solving very large systems.5

Example 1.2.4

Find a condition on the numbers a, b, and c such that the following system of equations is consistent. When that condition is satisfied, find all solutions (in terms of a, b, and c).

x1 + 3x2 + x3 = a

−x1 − 2x2 + x3 = b

3x1 + 7x2 − x3 = c

Solution. We use gaussian elimination except that now the augmented matrix



1 3 1 a 

−1 −2 1 b 3 7 −1 c

 

has entries a, b, and c as well as known numbers. The first leading one is in place, so we create

zeros below it in column 1: 

1 3 1 a 0 1 2 a+b 0 −2 −4 c−3a

 

5With n equations where n is large, gaussian elimination requires roughly n3/2 multiplications and divisions, whereas this number is roughly n3/3 if back substitution is used.

1.2. Gaussian Elimination 15

The second leading 1 has appeared, so use it to create zeros in the rest of column 2:



1 0 −5 −2a−3b 

0 1 2 a+b 0 0 0 c−a+2b

 

Now the whole solution depends on the number c−a+2b = c−(a−2b). The last row corresponds to an equation 0 = c−(a−2b). If c 6= a−2b, there is no solution (just as in Example 1.2.2). Hence:

The system is consistent if and only if c = a−2b.

In this case the last matrix becomes



1 0 −5 −2a−3b 

0 1 2 a+b 0 0 0 0

 

Thus, if c = a−2b, taking x3 = t where t is a parameter gives the solutions

x1 = 5t −(2a+3b) x2 = (a+b)−2t x3 = t.

**Rank**

It can be proven that the reduced row-echelon form of a matrix A is uniquely determined by A. That is, no matter which series of row operations is used to carry A to a reduced row-echelon matrix, the result will always be the same matrix. (A proof is given at the end of Section 2.5.) By contrast, this is not true for row-echelon matrices: Different series of row operations can carry the same matrix A to different

row-echelon matrices. Indeed, the matrix A =

1 −1 4 2 −1 2

can be carried (by one row operation) to

the row-echelon matrix

1 −1 4 0 1 −6

, and then by another row operation to the (reduced) row-echelon

matrix

1 0 −2 0 1 −6

. However, it is true that the number r of leading 1s must be the same in each of

these row-echelon matrices (this will be proved in Chapter 5). Hence, the number r depends only on A and not on the way in which A is carried to row-echelon form.

Definition 1.4 Rank of a Matrix

The rank of matrix A is the number of leading 1s in any row-echelon matrix to which A can be carried by row operations.

16 Systems of Linear Equations Example 1.2.5





Compute the rank of A =

1 1 −1 4 2 1 3 0 0 1 −5 8



.

Solution. The reduction of A to row-echelon form is

A =

 

1 1 −1 4 2 1 3 0 0 1 −5 8



 →



1 1 −1 4 

0 −1 5 −8 0 1 −5 8



 →

 

1 1 −1 4 0 1 −5 8 0 0 0 0

 

Because this row-echelon matrix has two leading 1s, rank A = 2.

Suppose that rank A = r, where A is a matrix with m rows and n columns. Then r ≤ m because the leading 1s lie in different rows, and r ≤ n because the leading 1s lie in different columns. Moreover, the rank has a useful application to equations. Recall that a system of linear equations is called consistent if it has at least one solution.

Theorem 1.2.2

Suppose a system of m equations in n variables is consistent, and that the rank of the augmented matrix is r.

1. The set of solutions involves exactly n−r parameters.

2. If r < n, the system has infinitely many solutions.

3. If r = n, the system has a unique solution.

Proof. The fact that the rank of the augmented matrix is r means there are exactly r leading variables, and hence exactly n−r nonleading variables. These nonleading variables are all assigned as parameters in the gaussian algorithm, so the set of solutions involves exactly n − r parameters. Hence if r < n, there is at least one parameter, and so infinitely many solutions. If r = n, there are no parameters and so a unique

solution.

|  |
| --- |

Theorem 1.2.2 shows that, for any system of linear equations, exactly three possibilities exist: 1. No solution. This occurs when a row 0 0 ··· 0 1 occurs in the row-echelon form. This is the case where the system is inconsistent.

2. Unique solution. This occurs when every variable is a leading variable.

3. Infinitely many solutions. This occurs when the system is consistent and there is at least one nonleading variable, so at least one parameter is involved.

1.2. Gaussian Elimination 17

Example 1.2.6

Suppose the matrix A in Example 1.2.5 is the augmented matrix of a system of m = 3 linear equations in n = 3 variables. As rank A = r = 2, the set of solutions will have n−r = 1 parameter. The reader can verify this fact directly.

Many important problems involve linear inequalities rather than linear equations. For example, a condition on the variables x and y might take the form of an inequality 2x−5y ≤ 4 rather than an equality 2x−5y = 4. There is a technique (called the simplex algorithm) for finding solutions to a system of such inequalities that maximizes a function of the form p = ax+by where a and b are fixed constants.

**Exercises for 1.2**



Exercise 1.2.1 Which of the following matrices are in 

reduced row-echelon form? Which are in row-echelon

b.

form?

1 −2 0 2 0 1 1 0 0 1 5 0 −3 −1 0 0 0 0 1 6 1

 



1 −1 2



2 1 −1 3

0 0 0 0 0 0 0 



a. 

1 2 1 3 1 1



0 0 0 0 0 1

b.

0 0 0 0



c.

0 1 −1 0 1 1



1 −2 3 5



1 0 0 3 1



0 0 0 1 −1 0

d. 

0 0 0 0 0 0

c.

0 0 0 1

 

0 0 0 1 1 0 0 0 0 1



1 −1 2 4 6 2



1 1

0 0 1





0 1 2 1 −1 −1



f. 

d.

0 0 0 1 0 1

e.

0 1



0 0 1 0 0 1

0 0 0 0 0 0

Exercise 1.2.2 Carry each of the following matrices to reduced row-echelon form.

Exercise 1.2.4 Find all solutions (if any) to each of the following systems of linear equations.



0 −1 2 1 2 1 −1



a. 3x − y = 0

a.



0 1 −2 2 7 2 4 0 −2 4 3 7 1 0



x − 2y = 1 4y − x = −2

b.

2x − 3y = 1

c. 3x − y = 2

0 3 −6 1 6 4 1 



2x + y = 5 3x + 2y = 6

d.

2y − 6x = −4

0 −1 3 1 3 2 1

e. 2x − 3y = 5

b.





0 −2 6 1 −5 0 −1 0 3 −9 2 4 1 −1

3x − y = 4 2y − 6x = 1

f.

3y − 2x = 2

0 1 −3 −1 3 0 1

Exercise 1.2.3 The augmented matrix of a system of linear equations has been carried to the following by row

Exercise 1.2.5 Find all solutions (if any) to each of the following systems of linear equations.

a. −2x + 3y + 3z = −9

operations. In each case solve the system.

x + y + 2z = 8

b.



1 2 0 3 1 0 −1



3x − y + z = 0 −x + 3y + 4z = −4

3x − 4y + z = 5 −5x + 7y + 2z = −14



0 0 1 −1 1 0 2



c. x + 2y − z = 2

a.

0 0 0 0 0 1 3 0 0 0 0 0 0 0

x + y − z = 10 −x + 4y + 5z = −5 x + 6y + 3z = 15

d.

2x + 5y − 3z = 1 x + 4y − 3z = 3

18 Systems of Linear Equations

a. 2x + y − z = a

e. 3x − 2y + z = −2

3x + y − z = a

b.

5x + y = 2 3x − y + 2z = 1 x + y − z = 5

f.

x − y + 3z = 5 −x + y + z = −1

x − y + 2z = b 5x + 3y − 4z = c

2y + 3z = b

x − z = c

g. x + 2y − 4z = 10

x + y + z = 2

h.

c. x+ay = 0

x + z = 1 2x + 5y + 2z = 7

2x − y + 2z = 5 x + y − 2z = 7

−x + 3y + 2z = −8 x + z = 2 3x + 3y + az = b

d.

y+bz = 0

z+cx = 0

e.

Exercise 1.2.6 Express the last equation of each system as a sum of multiples of the first two equations. [Hint: Label the equations, use the gaussian algorithm.] f.

a. x1 + 2x2 − 3x3 = −3

3x − y + 2z = 3

x + y − z = 2

2x − 2y + 3z = b

x + ay − z = 1

x1 + x2 + x3 = 1 2x1 − x2 + 3x3 = 3 x1 − 2x2 + 2x3 = 2

b.

x1 + 3x2 − 5x3 = 5 x1 − 2x2 + 5x3 = −35

−x + (a−2)y + z = −1 2x + 2y + (a−2)z = 1

Exercise 1.2.7 Find all solutions to the following sys tems.

a. 3x1 + 8x2 − 3x3 − 14x4 = 2

Exercise 1.2.10 Find the rank of each of the matrices in Exercise 1.2.1.

Exercise 1.2.11 Find the rank of each of the following matrices.

2x1 + 3x2 − x3 − 2x4 = 1



1 1 2





−2 3 3



x1 − 2x2 + x3 + 10x4 = 0

a. 

b. 



x1 + 5x2 − 2x3 − 12x4 = 1



b. x1 − x2 + x3 − x4 = 0

3 −1 1

−1 3 4

1 1 −1 3





3 −4 1

−5 7 2



3 −2 1 −2



−x1 + x2 + x3 + x4 = 0

c.  

d. 

x1 + x2 − x3 + x4 = 0 x1 + x2 + x3 + x4 = 0



−1 4 5 −2 1 6 3 4

1 2 −1 0





1 −1 3 5 −1 1 1 −1

e. 



c. x1 − x2 + x3 − 2x4 = 1

−x1 + x2 + x3 + x4 = −1

0 a 1−a a2 +1 1 2−a −1 −2a2



−x1 + 2x2 + 3x3 − x4 = 2

1 1 2 a2



x1 − x2 + 2x3 + x4 = 1

f. 



d. x1 + x2 + 2x3 − x4 = 4

3x2 − x3 + 4x4 = 2

1 1−a 2 0 2 2−a 6−a 4

x1 + 2x2 − 3x3 + 5x4 = 0

x1 + x2 − 5x3 + 6x4 = −3

Exercise 1.2.8 In each of the following, find (if possi ble) conditions on a and b such that the system has no solution, one solution, and infinitely many solutions.

a. x + by = −1

Exercise 1.2.12 Consider a system of linear equations with augmented matrix A and coefficient matrix C. In each case either prove the statement or give an example showing that it is false.

a. If there is more than one solution, A has a row of zeros.

x − 2y = 1 ax + by = 5

b.

ax + 2y = 5

b. If A has a row of zeros, there is more than one

c. ax + y = 1

solution.

x − by = −1 x + ay = 3

d.

2x + y = b

c. If there is no solution, the reduced row-echelon

Exercise 1.2.9 In each of the following, find (if possi ble) conditions on a, b, and c such that the system has no solution, one solution, or infinitely many solutions.

form of C has a row of zeros.

d. If the row-echelon form of C has a row of zeros, there is no solution.

e. There is no system that is inconsistent for every choice of constants.

f. If the system is consistent for some choice of con stants, it is consistent for every choice of con stants.

Now assume that the augmented matrix A has 3 rows and 5 columns.

g. If the system is consistent, there is more than one solution.

h. The rank of A is at most 3.

i. If rank A = 3, the system is consistent.

j. If rank C = 3, the system is consistent.

Exercise 1.2.13 Find a sequence of row operations car rying

1.2. Gaussian Elimination 19

a. (−2, 1), (5, 0), and (4, 1)

b. (1, 1), (5, −3), and (−3, −3)

Exercise 1.2.17 Three Nissans, two Fords, and four Chevrolets can be rented for $106 per day. At the same rates two Nissans, four Fords, and three Chevrolets cost $107 per day, whereas four Nissans, three Fords, and two Chevrolets cost $102 per day. Find the rental rates for all three kinds of cars.

Exercise 1.2.18 A school has three clubs and each stu dent is required to belong to exactly one club. One year the students switched club membership as follows: Club A. 410 remain in A, 110 switch to B, 510 switch to C. Club B. 710 remain in B, 210 switch to A, 110 switch to C. Club C. 610 remain in C, 210 switch to A, 210 switch to B.

If the fraction of the student population in each club is unchanged, find each of these fractions.



b1 +c1 b2 +c2 b3 +c3 

c1 +a1 c2 +a2 c3 +a3 a1 +b1 a2 +b2 a3 +b3



to

 

a1 a2 a3 b1 b2 b3 c1 c2 c3



Exercise 1.2.19 Given points (p1, q1), (p2, q2), and (p3, q3) in the plane with p1, p2, and p3 distinct, show



that they lie on some curve with equation y = a + bx + cx2. [Hint: Solve for a, b, and c.]

Exercise 1.2.14 In each case, show that the reduced row-echelon form is as given.

Exercise 1.2.20 The scores of three players in a tour nament have been lost. The only information available

a.

 

p 0 a b 0 0 q c r



 with abc 6= 0;

 

1 0 0 0 1 0 0 0 1

is the total of the scores for players 1 and 2, the total for



players 2 and 3, and the total for players 3 and 1.



a. Show that the individual scores can be rediscov

b.

 

1 a b+c 1 b c+a 1 c a+b



 where c 6= a or b 6= a;

ered.

b. Is this possible with four players (knowing the to

 

1 0 ∗ 0 1 ∗ 0 0 0





az + by + cz = 0

tals for players 1 and 2, 2 and 3, 3 and 4, and 4 and 1)?

Exercise 1.2.21 A boy finds $1.05 in dimes, nickels,

Exercise 1.2.15 Show that

a1x + b1y + c1z = 0al ways has a solution other than x = 0, y = 0, z = 0.

Exercise 1.2.16 Find the circle x2 +y2 +ax+by+c = 0 passing through the following points.

and pennies. If there are 17 coins in all, how many coins of each type can he have?

Exercise 1.2.22 If a consistent system has more vari ables than equations, show that it has infinitely many so lutions. [Hint: Use Theorem 1.2.2.]

20 Systems of Linear Equations

**1.3 Homogeneous Equations**

A system of equations in the variables x1, x2, ..., xn is called homogeneous if all the constant terms are zero—that is, if each equation of the system has the form

a1x1 +a2x2 +···+anxn = 0

Clearly x1 = 0, x2 = 0, ..., xn = 0 is a solution to such a system; it is called the trivial solution. Any solution in which at least one variable has a nonzero value is called a nontrivial solution. Our chief goal in this section is to give a useful condition for a homogeneous system to have nontrivial solutions. The following example is instructive.

Example 1.3.1

Show that the following homogeneous system has nontrivial solutions.

x1 − x2 + 2x3 − x4 = 0

2x1 + 2x2 + x4 = 0

3x1 + x2 + 2x3 − x4 = 0

Solution. The reduction of the augmented matrix to reduced row-echelon form is outlined below.



1 −1 2 −1 0 

2 2 0 1 0 3 1 2 −1 0



 →



1 −1 2 −1 0 

0 4 −4 3 0 0 4 −4 2 0



 →



1 0 1 0 0 

0 1 −1 0 0 0 0 0 1 0

 

The leading variables are x1, x2, and x4, so x3 is assigned as a parameter—say x3 = t. Then the general solution is x1 = −t, x2 = t, x3 = t, x4 = 0. Hence, taking t = 1 (say), we get a nontrivial solution: x1 = −1, x2 = 1, x3 = 1, x4 = 0.

The existence of a nontrivial solution in Example 1.3.1 is ensured by the presence of a parameter in the solution. This is due to the fact that there is a nonleading variable (x3 in this case). But there must be a nonleading variable here because there are four variables and only three equations (and hence at most three leading variables). This discussion generalizes to a proof of the following fundamental theorem.

Theorem 1.3.1

If a homogeneous system of linear equations has more variables than equations, then it has a nontrivial solution (in fact, infinitely many).

Proof. Suppose there are m equations in n variables where n > m, and let R denote the reduced row-echelon form of the augmented matrix. If there are r leading variables, there are n−r nonleading variables, and so n−r parameters. Hence, it suffices to show that r < n. But r ≤ m because R has r leading 1s and m rows,

and m < n by hypothesis. So r ≤ m < n, which gives r < n.

|  |
| --- |

Note that the converse of Theorem 1.3.1 is not true: if a homogeneous system has nontrivial solutions, it need not have more variables than equations (the system x1 + x2 = 0, 2x1 + 2x2 = 0 has nontrivial solutions but m = 2 = n.)

1.3. Homogeneous Equations 21

Theorem 1.3.1 is very useful in applications. The next example provides an illustration from geometry.

Example 1.3.2

We call the graph of an equation ax2 +bxy+cy2 +dx+ey+ f = 0 a conic if the numbers a, b, and c are not all zero. Show that there is at least one conic through any five points in the plane that are not all on a line.

Solution. Let the coordinates of the five points be (p1, q1), (p2, q2), (p3, q3), (p4, q4), and (p5, q5). The graph of ax2 +bxy+cy2 +dx+ey+ f = 0 passes through (pi, qi) if

ap2i +bpiqi +cq2i +d pi +eqi + f = 0

This gives five equations, one for each i, linear in the six variables a, b, c, d, e, and f . Hence, there is a nontrivial solution by Theorem 1.3.1. If a = b = c = 0, the five points all lie on the line with equation dx+ey+ f = 0, contrary to assumption. Hence, one of a, b, c is nonzero.

**Linear Combinations and Basic Solutions**

As for rows, two columns are regarded as equal if they have the same number of entries and corresponding entries are the same. Let x and y be columns with the same number of entries. As for elementary row operations, their sum x+y is obtained by adding corresponding entries and, if k is a number, the scalar product kx is defined by multiplying each entry of x by k. More precisely:

If x =



x1

x2...

xn



and y =



y1

y2...

yn



then x+y =



x1 +y1

x2 +y2 ...

xn +yn



and kx =



kx1

kx2...

kxn



.

A sum of scalar multiples of several columns is called a linear combination of these columns. For example, sx+ty is a linear combination of x and y for any choice of numbers s and t.

Example 1.3.3

If x =

3 −2

and y =

−1 1

then 2x+5y =

6 −4

+

−5 5

=

1 1

.

Example 1.3.4

Let x =

 

1 0 1



, y =

 

2 1 0



 and z =

 

3 1 1



. If v =

 

0

−1 2



 and w =

 

1 1 1



, determine whether v

and w are linear combinations of x, y and z.

22 Systems of Linear Equations

Solution. For v, we must determine whether numbers r, s, and t exist such that v = rx+sy+tz, that is, whether

 

0

−1 2



 = r

 

1 0 1



+s

 

2 1 0



+t

 

3 1 1



 =

 

r +2s+3t s+t

r +t

 

Equating corresponding entries gives a system of linear equations r +2s+3t = 0, s+t = −1, and r +t = 2 for r, s, and t. By gaussian elimination, the solution is r = 2−k, s = −1−k, and t = k where k is a parameter. Taking k = 0, we see that v = 2x−y is a linear combination of x, y, and z. Turning to w, we again look for r, s, and t such that w = rx+sy+tz; that is,

 

1 1 1



 = r

 

1 0 1



+s

 

2 1 0



+t

 

3 1 1



 =

 

r +2s+3t s+t

r +t

 

leading to equations r +2s+3t = 1, s+t = 1, and r +t = 1 for real numbers r, s, and t. But this time there is no solution as the reader can verify, so w is not a linear combination of x, y, and z.

Our interest in linear combinations comes from the fact that they provide one of the best ways to describe the general solution of a homogeneous system of linear equations. When solving such a system



x1

x2...

with n variables x1, x2, ..., xn, write the variables as a column6 matrix: x =



. The trivial solution





xn

is denoted 0 =

00...

. As an illustration, the general solution in Example 1.3.1 is x1 = −t, x2 = t, x3 = t, 0

and x4 = 0, where t is a parameter, and we would now express this by saying that the general solution is

x =



−ttt 0



, where t is arbitrary.

Now let x and y be two solutions to a homogeneous system with n variables. Then any linear combi nation sx+ty of these solutions turns out to be again a solution to the system. More generally:

Any linear combination of solutions to a homogeneous system is again a solution. (1.1) In fact, suppose that a typical equation in the system is a1x1 +a2x2 +···+anxn = 0, and suppose that

x =



x1

x2...

xn



, y =





y1

are solutions. Then a1x1 +a2x2+···+anxn = 0 and a1y1+a2y2 +···+anyn = 0.

y2...

yn

6The reason for using columns will be apparent later.

1.3. Homogeneous Equations 23

Hence sx+ty =



sx1 +ty1 sx2 +ty2



is also a solution because

...

sxn +tyn

a1(sx1 +ty1) +a2(sx2 +ty2) +···+an(sxn +tyn)

= [a1(sx1) +a2(sx2) +···+an(sxn)] +[a1(ty1) +a2(ty2) +···+an(tyn)]

= s(a1x1 +a2x2 +···+anxn) +t(a1y1 +a2y2 +···+anyn)

= s(0) +t(0)

= 0

A similar argument shows that Statement 1.1 is true for linear combinations of more than two solutions. The remarkable thing is that every solution to a homogeneous system is a linear combination of certain particular solutions and, in fact, these solutions are easily computed using the gaussian algorithm. Here is an example.

Example 1.3.5

Solve the homogeneous system with coefficient matrix

A =



1 −2 3 −2 

−3 6 1 0 −2 4 4 −2

 

Solution. The reduction of the augmented matrix to reduced form is



1 −2 3 −2 0 

−3 6 1 0 0 −2 4 4 −2 0



 →



1 −2 0 −150

0 0 1 −350 0 0 0 0 0

 

so the solutions are x1 = 2s+15t, x2 = s, x3 =35t, and x4 = t by gaussian elimination. Hence we can write the general solution x in the matrix form

x =



x1

x2

x3

x4



=



2s+15t

s

35tt



= s



210 0



+t

 

150 351



= sx1 +tx2.

Here x1 =



210 0



and x2 =

 



150

are particular solutions determined by the gaussian algorithm. 351

The solutions x1 and x2 in Example 1.3.5 are denoted as follows:

24 Systems of Linear Equations

Definition 1.5 Basic Solutions

The gaussian algorithm systematically produces solutions to any homogeneous linear system, called basic solutions, one for every parameter.

Moreover, the algorithm gives a routine way to express every solution as a linear combination of basic solutions as in Example 1.3.5, where the general solution x becomes

x = s



210 0



+t

 

150 351



= s



210 0



+15t



103 5

 

Hence by introducing a new parameter r = t/5 we can multiply the original basic solution x2 by 5 and so eliminate fractions. For this reason:

Convention:

Any nonzero scalar multiple of a basic solution will still be called a basic solution.

In the same way, the gaussian algorithm produces basic solutions to every homogeneous system, one for each parameter (there are no basic solutions if the system has only the trivial solution). Moreover every solution is given by the algorithm as a linear combination of these basic solutions (as in Example 1.3.5). If A has rank r, Theorem 1.2.2 shows that there are exactly n−r parameters, and so n−r basic solutions. This proves:

Theorem 1.3.2

Let A be an m×n matrix of rank r, and consider the homogeneous system in n variables with A as coefficient matrix. Then:

1. The system has exactly n−r basic solutions, one for each parameter.

2. Every solution is a linear combination of these basic solutions.

Example 1.3.6

Find basic solutions of the homogeneous system with coefficient matrix A, and express every solution as a linear combination of the basic solutions, where

A =



1 −3 0 2 2

−2 6 1 2 −5 3 −9 −1 0 7 −3 9 2 6 −8

 

1.3. Homogeneous Equations 25

Solution. The reduction of the augmented matrix to reduced row-echelon form is



1 −3 0 2 2 0

−2 6 1 2 −5 0 3 −9 −1 0 7 0 −3 9 2 6 −8 0



→



1 −3 0 2 2 0

0 0 1 6 −1 0 0 0 0 0 0 0 0 0 0 0 0 0

 

so the general solution is x1 = 3r −2s−2t, x2 = r, x3 = −6s+t, x4 = s, and x5 = t where r, s, and t are parameters. In matrix form this is

x =



x1

x2

x3

x4

x5



=



3r −2s−2t

r

−6s+t

s

t



= r



3100 0



+s



−20

−6

1

0



+t



−2010 1

 

Hence basic solutions are

x1 =

**Exercises for 1.3**



3100 0



, x2 =



−20

−6

1

0



, x3 =



−2010 1

 

Exercise 1.3.1 Consider the following statements about a system of linear equations with augmented matrix A. In each case either prove the statement or give an example for which it is false.

a. If the system is homogeneous, every solution is trivial.

b. If the system has a nontrivial solution, it cannot be homogeneous.

c. If there exists a trivial solution, the system is ho mogeneous.

d. If the system is consistent, it must be homoge neous.

f. If there exists a solution, there are infinitely many solutions.

g. If there exist nontrivial solutions, the row-echelon form of A has a row of zeros.

h. If the row-echelon form of A has a row of zeros, there exist nontrivial solutions.

i. If a row operation is applied to the system, the new system is also homogeneous.

Exercise 1.3.2 In each of the following, find all values of a for which the system has nontrivial solutions, and determine all solutions in each case.

a. x + 2y + z = 0

Now assume that the system is homogeneous.

x − 2y + z = 0 x + ay − 3z = 0 −x + 6y − 5z = 0

b.

x + 3y + 6z = 0 2x + 3y + az = 0

c. ax + y + z = 0

e. If there exists a nontrivial solution, there is no triv ial solution.

x + y − z = 0 ay − z = 0

x + y + az = 0

d.

x + y − z = 0 x + y + az = 0

26 Systems of Linear Equations





Exercise 1.3.3 Let x =

2

1

−1



, y =

 

1 0 1



, and

Exercise 1.3.6

a. Does Theorem

z =

 

1

1

−2



. In each case, either write v as a linear com

1.3.1 imply that the system −z+3y = 0

2x−6y = 0has nontrivial solutions? Explain. b. Show that the converse to Theorem 1.3.1 is not

bination of x, y, and z, or show that it is not such a linear combination.

true. That is, show that the existence of nontrivial solutions does not imply that there are more vari



0





4



ables than equations.

a.  v =



b. 

v =



1

−3



3

−4

Exercise 1.3.7 In each case determine how many solu

3





3



tions (and how many parameters) are possible for a ho

c.  v =

mogeneous system of four linear equations in six vari



d. 

v =

1 0



0

ables with augmented matrix A. Assume that A has

3

nonzero entries. Give all possibilities.

Exercise 1.3.4 In each case, either express y as a linear combination of a1, a2, and a3, or show that it is not such a linear combination. Here:

a. Rank Rank A = 2. b. A = 1. c. A has a row of zeros.

a1 =

 

−1 3

0

1



, a2 =

 

3 1 2 0



, and a3 =

 

1 1 1 1

 

d. The row-echelon form of A has a row of zeros. Exercise 1.3.8 The graph of an equation ax+by+cz = 0

 

1 2

 

 

−1 9

is a plane through the origin (provided that not all of a, 

b, and c are zero). Use Theorem 1.3.1 to show that two 

planes through the origin have a point in common other

a. y =

y =

4 0

b.

2 6

than the origin (0, 0, 0). Exercise 1.3.9

Exercise 1.3.5 For each of the following homogeneous systems, find a set of basic solutions and express the gen eral solution as a linear combination of these basic solu tions.

a. x1 + 2x2 − x3 + 2x4 + x5 = 0

x1 + 2x2 + 2x3 + x5 = 0

2x1 + 4x2 − 2x3 + 3x4 + x5 = 0

b. x1 + 2x2 − x3 + x4 + x5 = 0

−x1 − 2x2 + 2x3 + x5 = 0

−x1 − 2x2 + 3x3 + x4 + 3x5 = 0

c. x1 + x2 − x3 + 2x4 + x5 = 0

x1 + 2x2 − x3 + x4 + x5 = 0

2x1 + 3x2 − x3 + 2x4 + x5 = 0

4x1 + 5x2 − 2x3 + 5x4 + 2x5 = 0

d. x1 + x2 − 2x3 − 2x4 + 2x5 = 0

2x1 + 2x2 − 4x3 − 4x4 + x5 = 0

x1 − x2 + 2x3 + 4x4 + x5 = 0

−2x1 − 4x2 + 8x3 + 10x4 + x5 = 0

a. Show that there is a line through any pair of points in the plane. [Hint: Every line has equation ax+by+c = 0, where a, b, and c are not all zero.]

b. Generalize and show that there is a plane ax+by+ cz+d = 0 through any three points in space.

Exercise 1.3.10 The graph of

a(x2 +y2) +bx+cy+d = 0

is a circle if a 6= 0. Show that there is a circle through any three points in the plane that are not all on a line.

Exercise 1.3.11 Consider a homogeneous system of lin ear equations in n variables, and suppose that the aug mented matrix has rank r. Show that the system has non trivial solutions if and only if n > r.

Exercise 1.3.12 If a consistent (possibly nonhomoge neous) system of linear equations has more variables than equations, prove that it has more than one solution.

1.4. An Application to Network Flow 27

**1.4 An Application to Network Flow**

There are many types of problems that concern a network of conductors along which some sort of flow is observed. Examples of these include an irrigation network and a network of streets or freeways. There are often points in the system at which a net flow either enters or leaves the system. The basic principle behind the analysis of such systems is that the total flow into the system must equal the total flow out. In fact, we apply this principle at every junction in the system.

Junction Rule

At each of the junctions in the network, the total flow into that junction must equal the total flow out.

This requirement gives a linear equation relating the flows in conductors emanating from the junction.

Example 1.4.1

A network of one-way streets is shown in the accompanying diagram. The rate of flow of cars into intersection A is 500 cars per hour, and 400 and 100 cars per hour emerge from B and C, respectively. Find the possible flows along each street.

Solution. Suppose the flows along the streets are f1, f2, f3, f4,

f1

500 400 A B

f5, and f6 cars per hour in the directions shown.

Then, equating the flow in with the flow out at each intersection,

f2

f3

D

C

f4

f5

100

we get

Intersection A 500 = f1 + f2 + f3

f6

Intersection B f1 + f4 + f6 = 400

Intersection C f3 + f5 = f6 +100

Intersection D f2 = f4 + f5

These give four equations in the six variables f1, f2, ..., f6.

f1 + f2 + f3 = 500

f1 + f4 + f6 = 400

f3 + f5 − f6 = 100

f2 − f4 − f5 = 0

The reduction of the augmented matrix is



1 1 1 0 0 0 500

1 0 0 1 0 1 400 0 0 1 0 1 −1 100 0 1 0 −1 −1 0 0



→





1 0 0 1 0 1 400 

0 1 0 −1 −1 0 0 0 0 1 0 1 −1 100 0 0 0 0 0 0 0

Hence, when we use f4, f5, and f6 as parameters, the general solution is f1 = 400− f4 − f6 f2 = f4 + f5 f3 = 100− f5 + f6

This gives all solutions to the system of equations and hence all the possible flows.

28 Systems of Linear Equations

Of course, not all these solutions may be acceptable in the real situation. For example, the flows f1, f2, ..., f6 are all positive in the present context (if one came out negative, it would mean traffic flowed in the opposite direction). This imposes constraints on the flows: f1 ≥ 0 and f3 ≥ 0 become

f4 + f6 ≤ 400 f5 − f6 ≤ 100

Further constraints might be imposed by insisting on maximum values on the flow in each street. **Exercises for 1.4**

20

Exercise 1.4.1 Find the possible flows in each of the fol

B

lowing networks of pipes.

f1

f3

55

15

A

a.

C

f2

40

b.

50

f1 f2 f3

f4 f5 50

25

f1 f2

f4 f5

D

20

60

a. Find the possible flows.

b. If canal BC is closed, what range of flow on AD

must be maintained so that no canal carries a flow

of more than 30?

Exercise 1.4.3 A traffic circle has five one-way streets,

and vehicles enter and leave as shown in the accompany

ing diagram.

f4

75 60

f5

35 25

50

f3 f4

f5f6f7

50

E D

A

f3

C

40

Exercise 1.4.2 A proposed network of irrigation canals is described in the accompanying diagram. At peak de

f1f2

B

30

a. Compute the possible flows.

40

mand, the flows at interchanges A, B, C, and D are as shown.

b. Which road has the heaviest flow?

1.5. An Application to Electrical Networks 29

**1.5 An Application to Electrical Networks**7

In an electrical network it is often necessary to find the current in amperes (A) flowing in various parts of the network. These networks usually contain resistors that retard the current. The resistors are indicated by a symbol ( ), and the resistance is measured in ohms (Ω). Also, the current is increased at various points by voltage sources (for example, a battery). The voltage of these sources is measured in volts (V), and they are represented by the symbol ( ). We assume these voltage sources have no resistance. The flow of current is governed by the following principles.

Ohm’s Law

The current I and the voltage drop V across a resistance R are related by the equation V = RI.

Kirchhoff’s Laws

1. (Junction Rule) The current flow into a junction equals the current flow out of that junction.

2. (Circuit Rule) The algebraic sum of the voltage drops (due to resistances) around any closed circuit of the network must equal the sum of the voltage increases around the circuit.

When applying rule 2, select a direction (clockwise or counterclockwise) around the closed circuit and then consider all voltages and currents positive when in this direction and negative when in the opposite direction. This is why the term algebraic sum is used in rule 2. Here is an example.

Example 1.5.1

Find the various currents in the circuit shown.

Solution.

20Ω

A

10 V

5 V

10Ω

I3

20 V

First apply the junction rule at junctions A, B, C, and D to obtain

Junction A I1 = I2 +I3

Junction B I6 = I1 +I5

B

I1 I6 5Ω

I2 5Ω C

10 V

I5

Junction C I2 +I4 = I6

Junction D I3 +I5 = I4

D

I4

Note that these equations are not independent

(in fact, the third is an easy consequence of the other three). Next, the circuit rule insists that the sum of the voltage increases (due to the sources) around a closed circuit must equal the sum of the voltage drops (due to resistances). By Ohm’s law, the voltage

loss across a resistance R (in the direction of the current I) is RI. Going counterclockwise around three closed circuits yields

7This section is independent of Section 1.4

30 Systems of Linear Equations

Upper left 10 + 5 = 20I1

Upper right −5 + 20 = 10I3 +5I4

Lower −10 = −5I5 −5I4

Hence, disregarding the redundant equation obtained at junction C, we have six equations in the six unknowns I1, ..., I6. The solution is

I1 =1520 I4 =2820

I2 =−1

20 I5 =1220

I3 =1620 I6 =2720

The fact that I2 is negative means, of course, that this current is in the opposite direction, with a magnitude of 120 amperes.

**Exercises for 1.5**

In Exercises 1 to 4, find the currents in the circuits.

Exercise 1.5.1

20 V

6Ω I1

I2

4Ω

10 V

Exercise 1.5.3 5 VI1

10Ω

I3

20Ω

5 V

20 V

5 V

I6

2ΩI3

I2

10 V

10Ω I4

I5

20Ω

Exercise 1.5.2

5 V

I2

I1

5Ω

10Ω

Exercise 1.5.4 All resistances are 10Ω.

I6 I4 I2

10 V

I3

I5

Exercise 1.5.5

5Ω I3 10 V

I1

20 V 2Ω

Find the voltage x such that the current I1 = 0.

I1

1Ω

x V

I2

1Ω

2 V

I3

5 V

1.6. An Application to Chemical Reactions 31

**1.6 An Application to Chemical Reactions**

When a chemical reaction takes place a number of molecules combine to produce new molecules. Hence, when hydrogen H2 and oxygen O2 molecules combine, the result is water H2O. We express this as

H2 +O2 → H2O

Individual atoms are neither created nor destroyed, so the number of hydrogen and oxygen atoms going into the reaction must equal the number coming out (in the form of water). In this case the reaction is said to be balanced. Note that each hydrogen molecule H2 consists of two atoms as does each oxygen molecule O2, while a water molecule H2O consists of two hydrogen atoms and one oxygen atom. In the above reaction, this requires that twice as many hydrogen molecules enter the reaction; we express this as follows:

2H2 +O2 → 2H2O

This is now balanced because there are 4 hydrogen atoms and 2 oxygen atoms on each side of the reaction.

Example 1.6.1

Balance the following reaction for burning octane C8H18 in oxygen O2:

C8H18 +O2 → CO2 +H2O

where CO2 represents carbon dioxide. We must find positive integers x, y, z, and w such that xC8H18 +yO2 → zCO2 +wH2O

Equating the number of carbon, hydrogen, and oxygen atoms on each side gives 8x = z, 18x = 2w and 2y = 2z+w, respectively. These can be written as a homogeneous linear system

8x − z = 0

18x − 2w = 0

2y − 2z − w = 0

which can be solved by gaussian elimination. In larger systems this is necessary but, in such a simple situation, it is easier to solve directly. Set w = t, so that x =19t, z =89t, 2y =169t +t =259t. But x, y, z, and w must be positive integers, so the smallest value of t that eliminates fractions is 18. Hence, x = 2, y = 25, z = 16, and w = 18, and the balanced reaction is

2C8H18 +25O2 → 16CO2 +18H2O

The reader can verify that this is indeed balanced.

It is worth noting that this problem introduces a new element into the theory of linear equations: the insistence that the solution must consist of positive integers.

32 Systems of Linear Equations

**Exercises for 1.6**

In each case balance the chemical reaction.

Exercise 1.6.1 CH4 + O2 → CO2 + H2O. This is the burning of methane CH4.

Exercise 1.6.2 NH3 + CuO → N2 + Cu + H2O. Here NH3 is ammonia, CuO is copper oxide, Cu is copper, and N2 is nitrogen.

Exercise 1.6.3 CO2 + H2O → C6H12O6 + O2. This is called the photosynthesis reaction—C6H12O6 is glu cose.

Exercise 1.6.4 Pb(N3)2 + Cr(MnO4)2 → Cr2O3 + MnO2 +Pb3O4 +NO.

**Supplementary Exercises for Chapter 1**

a b

Exercise 1.1 We show in Chapter 4 that the graph of an

Exercise 1.5 If ad 6= bc, show that

has re

equation ax+by+cz = d is a plane in space when not all of a, b, and c are zero.

duced row-echelon form

1 0 0 1

.

c d

a. By examining the possible positions of planes in space, show that three equations in three variables can have zero, one, or infinitely many solutions.

b. Can two equations in three variables have a unique solution? Give reasons for your answer.

Exercise 1.2 Find all solutions to the following systems of linear equations.

a. x1 + x2 + x3 − x4 = 3

3x1 + 5x2 − 2x3 + x4 = 1

−3x1 − 7x2 + 7x3 − 5x4 = 7

Exercise 1.6 Find a, b, and c so that the system

x + ay + cz = 0

bx + cy − 3z = 1

ax + 2y + bz = 5

has the solution x = 3, y = −1, z = 2. Exercise 1.7 Solve the system

x + 2y + 2z = −3

2x + y + z = −4

x − y + iz = i

where i2 = −1. [See Appendix A.]

Exercise 1.8 Show that the real system

x1 + 3x2 − 4x3 + 3x4 = −5

b. x1 + 4x2 − x3 + x4 = 2 3x1 + 2x2 + x3 + 2x4 = 5

 

x + y + z = 5 2x − y − z = 1 −3x + 2y + 2z = 0

x1 − 6x2 + 3x3 = 1

x1 + 14x2 − 5x3 + 2x4 = 3

Exercise 1.3 In each case find (if possible) conditions on a, b, and c such that the system has zero, one, or in finitely many solutions.

a. x + y + 3z = a

has a complex solution: x = 2, y = i, z = 3 − i where i2 = −1. Explain. What happens when such a real sys tem has a unique solution?

Exercise 1.9 A man is ordered by his doctor to take 5 units of vitamin A, 13 units of vitamin B, and 23 units of vitamin C each day. Three brands of vitamin pills are available, and the number of units of each vitamin per

x + 2y − 4z = 4 3x − y + 13z = 2 4x + y + a2z = a+3

b.

ax + y + 5z = 4 x + ay + 4z = a

pill are shown in the accompanying table.

| Brand | Vitamin | | |
| --- | --- | --- | --- |
| A | B | C |
| 1  2  3 | 1  1  0 | 2  1  1 | 4  3  1 |

Exercise 1.4 Show that any two rows of a matrix can be interchanged by elementary row transformations of the other two types.

a. Find all combinations of pills that provide exactly the required amount of vitamins (no partial pills allowed).

b. If brands 1, 2, and 3 cost 3¢, 2¢, and 5¢ per pill, respectively, find the least expensive treatment.

Exercise 1.10 A restaurant owner plans to use x tables seating 4, y tables seating 6, and z tables seating 8, for a total of 20 tables. When fully occupied, the tables seat 108 customers. If only half of the x tables, half of the y tables, and one-fourth of the z tables are used, each fully occupied, then 46 customers will be seated. Find x, y, and z.

Exercise 1.11

a. Show that a matrix with two rows and two columns that is in reduced row-echelon form must have one of the following forms:

1.6. An Application to Chemical Reactions 33

[Hint: The leading 1 in the first row must be in column 1 or 2 or not exist.]

b. List the seven reduced row-echelon forms for ma trices with two rows and three columns.

c. List the four reduced row-echelon forms for ma trices with three rows and two columns.

Exercise 1.12 An amusement park charges $7 for adults, $2 for youths, and $0.50 for children. If 150 peo ple enter and pay a total of $100, find the numbers of adults, youths, and children. [Hint: These numbers are nonnegative integers.]

Exercise 1.13 Solve the following system of equations for x and y.

x2 + xy − y2 = 1

2x2 − xy + 3y2 = 13

x2 + 3xy + 2y2 = 0

1 0 0 1

0 1 0 0

0 0 0 0

1 ∗ 0 0

[Hint: These equations are linear in the new variables x1 = x2, x2 = xy, and x3 = y2.]

**Chapter 2**

**Matrix Algebra**

In the study of systems of linear equations in Chapter 1, we found it convenient to manipulate the aug mented matrix of the system. Our aim was to reduce it to row-echelon form (using elementary row oper ations) and hence to write down all solutions to the system. In the present chapter we consider matrices for their own sake. While some of the motivation comes from linear equations, it turns out that matrices can be multiplied and added and so form an algebraic system somewhat analogous to the real numbers. This “matrix algebra” is useful in ways that are quite different from the study of linear equations. For example, the geometrical transformations obtained by rotating the euclidean plane about the origin can be viewed as multiplications by certain 2×2 matrices. These “matrix transformations” are an important tool in geometry and, in turn, the geometry provides a “picture” of the matrices. Furthermore, matrix algebra has many other applications, some of which will be explored in this chapter. This subject is quite old and was first studied systematically in 1858 by Arthur Cayley.1

**2.1 Matrix Addition, Scalar Multiplication, and**

**Transposition**

A rectangular array of numbers is called a matrix (the plural is matrices), and the numbers are called the entries of the matrix. Matrices are usually denoted by uppercase letters: A, B, C, and so on. Hence,

A =

1 2 −1 0 5 6

B =

1 −1 0 2

C =

 

1 3 2

 

are matrices. Clearly matrices come in various shapes depending on the number of rows and columns. For example, the matrix A shown has 2 rows and 3 columns. In general, a matrix with m rows and n columns is referred to as an m ×n matrix or as having size m ×n. Thus matrices A, B, and C above have sizes 2×3, 2×2, and 3×1, respectively. A matrix of size 1×n is called a row matrix, whereas one of size m×1 is called a column matrix. Matrices of size n×n for some n are called square matrices.

Each entry of a matrix is identified by the row and column in which it lies. The rows are numbered from the top down, and the columns are numbered from left to right. Then the (i, j)-entry of a matrix is the number lying simultaneously in row i and column j. For example,

The (1, 2)-entry of

1 −1 0 1

is −1.

The (2, 3)-entry of

1 2 −1 0 5 6

is 6.

1Arthur Cayley (1821-1895) showed his mathematical talent early and graduated from Cambridge in 1842 as senior wran gler. With no employment in mathematics in view, he took legal training and worked as a lawyer while continuing to do mathematics, publishing nearly 300 papers in fourteen years. Finally, in 1863, he accepted the Sadlerian professorship in Cam bridge and remained there for the rest of his life, valued for his administrative and teaching skills as well as for his scholarship. His mathematical achievements were of the first rank. In addition to originating matrix theory and the theory of determinants, he did fundamental work in group theory, in higher-dimensional geometry, and in the theory of invariants. He was one of the most prolific mathematicians of all time and produced 966 papers.

35

36 Matrix Algebra

A special notation is commonly used for the entries of a matrix. If A is an m × n matrix, and if the (i, j)-entry of A is denoted as ai j, then A is displayed as follows:

A =



a11 a12 a13 ··· a1n

a21 a22 a23 ··· a2n ............ am1 am2 am3 ··· amn

 

This is usually denoted simply as A =ai j. Thus ai j is the entry in row i and column j of A. For example, a 3×4 matrix in this notation is written

A =



a11 a12 a13 a14 

a21 a22 a23 a24 a31 a32 a33 a34

 

It is worth pointing out a convention regarding rows and columns: Rows are mentioned before columns. For example:

• If a matrix has size m×n, it has m rows and n columns.

• If we speak of the (i, j)-entry of a matrix, it lies in row i and column j.

• If an entry is denoted ai j, the first subscript i refers to the row and the second subscript j to the column in which ai j lies.

Two points (x1, y1) and (x2, y2) in the plane are equal if and only if2they have the same coordinates, that is x1 = x2 and y1 = y2. Similarly, two matrices A and B are called equal (written A = B) if and only if:

1. They have the same size.

2. Corresponding entries are equal.

If the entries of A and B are written in the form A =ai j, B =bi j, described earlier, then the second condition takes the following form:

A =ai j=bi jmeans ai j = bi j for all i and j

Example 2.1.1

Given A =

a b c d

, B =

1 2 −1 3 0 1

and C =

1 0 −1 2

discuss the possibility that A = B,

B = C, A = C.

Solution. A = B is impossible because A and B are of different sizes: A is 2×2 whereas B is 2×3. Similarly, B = C is impossible. But A = C is possible provided that corresponding entries are

2If p and q are statements, we say that p implies q if q is true whenever p is true. Then “p if and only if q” means that both p implies q and q implies p. See Appendix B for more on this.

equal:

a b c d

=

1 0 −1 2

2.1. Matrix Addition, Scalar Multiplication, and Transposition 37

means a = 1, b = 0, c = −1, and d = 2.

**Matrix Addition**

Definition 2.1 Matrix Addition

If A and B are matrices of the same size, their sum A+B is the matrix formed by adding corresponding entries.

If A =ai jand B =bi j, this takes the form

A+B =ai j +bi j

Note that addition is not defined for matrices of different sizes.

Example 2.1.2

If A =

2 1 3 −1 2 0

and B =

1 1 −1 2 0 6

, compute A+B.

Solution.

A+B =

Example 2.1.3

2+1 1+1 3−1 −1+2 2+0 0+6

=

3 2 2 1 2 6

Find a, b, and c if a b c +c a b =3 2 −1.

Solution. Add the matrices on the left side to obtain

a+c b+a c+b=3 2 −1

Because corresponding entries must be equal, this gives three equations: a+c = 3, b+a = 2, and c+b = −1. Solving these yields a = 3, b = −1, c = 0.

If A, B, and C are any matrices of the same size, then

A+B = B+A (commutative law)

A+ (B+C) = (A+B) +C (associative law)

In fact, if A =ai jand B =bi j, then the (i, j)-entries of A+B and B+A are, respectively, ai j +bi j and bi j +ai j. Since these are equal for all i and j, we get

A+B =ai j +bi j =bi j +ai j = B+A

38 Matrix Algebra

The associative law is verified similarly.

The m×n matrix in which every entry is zero is called the m×n zero matrix and is denoted as 0 (or 0mn if it is important to emphasize the size). Hence,

0+X = X

holds for all m×n matrices X. The negative of an m×n matrix A (written −A) is defined to be the m×n matrix obtained by multiplying each entry of A by −1. If A =ai j, this becomes −A =−ai j. Hence, A+ (−A) = 0

holds for all matrices A where, of course, 0 is the zero matrix of the same size as A. A closely related notion is that of subtracting matrices. If A and B are two m × n matrices, their difference A−B is defined by

Note that if A =ai jand B =bi j, then

A−B = A+ (−B)

A−B =ai j+−bi j=ai j −bi j

is the m×n matrix formed by subtracting corresponding entries. Example 2.1.4

Let A =

3 −1 0 1 2 −4

, B =

1 −1 1 −2 0 6

, C =

1 0 −2 3 1 1

. Compute −A, A−B, and

A+B−C.

Solution.

−A =

−3 1 0 −1 −2 4

A−B =

3−1 −1−(−1) 0−1 1−(−2) 2−0 −4−6

=

2 0 −1 3 2 −10

A+B−C =

Example 2.1.5

3+1−1 −1−1−0 0+1−(−2) 1−2−3 2+0−1 −4+6−1

=

3 −2 3 −4 1 1

Solve

3 2 −1 1

+X =

1 0 −1 2

where X is a matrix.

Solution. We solve a numerical equation a+x = b by subtracting the number a from both sides to

obtain x = b−a. This also works for matrices. To solve

3 2 −1 1

+X =

1 0 −1 2

simply

subtract the matrix

3 2 −1 1

2.1. Matrix Addition, Scalar Multiplication, and Transposition 39

from both sides to get

X =

1 0 −1 2

−

3 2 −1 1

=

1−3 0−2 −1−(−1) 2−1

=

−2 −2 0 1

The reader should verify that this matrix X does indeed satisfy the original equation.

The solution in Example 2.1.5 solves the single matrix equation A+X = B directly via matrix subtrac tion: X = B−A. This ability to work with matrices as entities lies at the heart of matrix algebra. It is important to note that the sizes of matrices involved in some calculations are often determined by

the context. For example, if

A+C =

1 3 −1 2 0 1

then A and C must be the same size (so that A+C makes sense), and that size must be 2×3 (so that the sum is 2 × 3). For simplicity we shall often omit reference to such facts when they are clear from the context.

**Scalar Multiplication**

In gaussian elimination, multiplying a row of a matrix by a number k means multiplying every entry of that row by k.

Definition 2.2 Matrix Scalar Multiplication

More generally, if A is any matrix and k is any number, the scalar multiple kA is the matrix obtained from A by multiplying each entry of A by k.

If A =ai j, this is

kA =kai j

Thus 1A = A and (−1)A = −A for any matrix A.

The term scalar arises here because the set of numbers from which the entries are drawn is usually referred to as the set of scalars. We have been using real numbers as scalars, but we could equally well have been using complex numbers.

Example 2.1.6

If A =

3 −1 4 2 0 6

and B =

1 2 −1 0 3 2

compute 5A,12B, and 3A−2B.

Solution.

5A =

15 −5 20 10 0 30

,12B =

121 −12 0321

3A−2B =

9 −3 12 6 0 18

−

2 4 −2 0 6 4

=

7 −7 14

6 −6 14

40 Matrix Algebra

If A is any matrix, note that kA is the same size as A for all scalars k. We also have 0A = 0 and k0 = 0

because the zero matrix has every entry zero. In other words, kA = 0 if either k = 0 or A = 0. The converse of this statement is also true, as Example 2.1.7 shows.

Example 2.1.7

If kA = 0, show that either k = 0 or A = 0.

Solution. Write A =ai jso that kA = 0 means kai j = 0 for all i and j. If k = 0, there is nothing to do. If k 6= 0, then kai j = 0 implies that ai j = 0 for all i and j; that is, A = 0.

For future reference, the basic properties of matrix addition and scalar multiplication are listed in Theorem 2.1.1.

Theorem 2.1.1

Let A, B, and C denote arbitrary m×n matrices where m and n are fixed. Let k and p denote arbitrary real numbers. Then

1. A+B = B+A.

2. A+ (B+C) = (A+B) +C.

3. There is an m×n matrix 0, such that 0+A = A for each A.

4. For each A there is an m×n matrix, −A, such that A+ (−A) = 0.

5. k(A+B) = kA+kB.

6. (k + p)A = kA+ pA.

7. (kp)A = k(pA).

8. 1A = A.

Proof. Properties 1–4 were given previously. To check Property 5, let A =ai jand B =bi jdenote matrices of the same size. Then A+B =ai j +bi j, as before, so the (i, j)-entry of k(A+B) is k(ai j +bi j) = kai j +kbi j

But this is just the (i, j)-entry of kA + kB, and it follows that k(A + B) = kA + kB. The other Properties

can be similarly verified; the details are left to the reader.

|  |
| --- |

The Properties in Theorem 2.1.1 enable us to do calculations with matrices in much the same way that numerical calculations are carried out. To begin, Property 2 implies that the sum

(A+B) +C = A+ (B+C)

2.1. Matrix Addition, Scalar Multiplication, and Transposition 41

is the same no matter how it is formed and so is written as A+B+C. Similarly, the sum A+B+C +D

is independent of how it is formed; for example, it equals both (A+B) + (C +D) and A+ [B+ (C +D)]. Furthermore, property 1 ensures that, for example,

B+D+A+C = A+B+C +D

In other words, the order in which the matrices are added does not matter. A similar remark applies to sums of five (or more) matrices.

Properties 5 and 6 in Theorem 2.1.1 are called distributive laws for scalar multiplication, and they extend to sums of more than two terms. For example,

k(A+B−C) = kA+kB−kC

(k + p−m)A = kA+ pA−mA

Similar observations hold for more than three summands. These facts, together with properties 7 and 8, enable us to simplify expressions by collecting like terms, expanding, and taking common factors in exactly the same way that algebraic expressions involving variables and real numbers are manipulated. The following example illustrates these techniques.

Example 2.1.8

Simplify 2(A+3C)−3(2C−B)−3[2(2A+B−4C)−4(A−2C)] where A, B, and C are all matrices of the same size.

Solution. The reduction proceeds as though A, B, and C were variables.

2(A+3C)−3(2C−B)−3[2(2A+B−4C)−4(A−2C)]

= 2A+6C −6C +3B−3[4A+2B−8C−4A+8C]

= 2A+3B−3[2B]

= 2A−3B

**Transpose of a Matrix**

Many results about a matrix A involve the rows of A, and the corresponding result for columns is derived in an analogous way, essentially by replacing the word row by the word column throughout. The following definition is made with such applications in mind.

Definition 2.3 Transpose of a Matrix

If A is an m×n matrix, the transpose of A, written AT columns of A in the same order.

, is the n×m matrix whose rows are just the

In other words, the first row of ATis the first column of A (that is it consists of the entries of column 1 in order). Similarly the second row of ATis the second column of A, and so on.

42 Matrix Algebra

Example 2.1.9

Write down the transpose of each of the following matrices.

A =

 





1

 B =5 2 6 C =



3

2

1 2 3 4 5 6



 D =

 

3 1 −1 1 3 2 −1 2 1

 

Solution.



5

AT =1 3 2 , BT =



2

6



, CT =

1 3 5 2 4 6

, and DT = D.

If A =ai jis a matrix, write AT =bi j. Then bi j is the jth element of the ith row of ATand so is the jth element of the ith column of A. This means bi j = aji, so the definition of ATcan be stated as follows: If A =ai j, then AT =aji. (2.1)

This is useful in verifying the following properties of transposition.

Theorem 2.1.2

Let A and B denote matrices of the same size, and let k denote a scalar.

1. If A is an m×n matrix, then AT 2. (AT)T = A.

is an n×m matrix.

3. (kA)T = kAT .

4. (A+B)T = AT +BT .

Proof. Property 1 is part of the definition of AT, and Property 2 follows from (2.1). As to Property 3: If A =ai j, then kA =kai j, so (2.1) gives

(kA)T =kaji= kaji= kAT

Finally, if B =bi j, then A+B =ci jwhere ci j = ai j +bi j Then (2.1) gives Property 4: (A+B)T =ci j T =cji=aji +bji=aji+bji= AT +BT

|  |
| --- |

2.1. Matrix Addition, Scalar Multiplication, and Transposition 43

There is another useful way to think of transposition. If A =ai jis an m × n matrix, the elements a11, a22, a33, ... are called the main diagonal of A. Hence the main diagonal extends down and to the right from the upper left corner of the matrix A; it is outlined in the following examples:

 

a11 a12 a21 a22 a31 a32

 

a11 a12 a13 a21 a22 a23

 

a11 a12 a13 a21 a22 a23 a31 a32 a33

 

a11 a21

Thus forming the transpose of a matrix A can be viewed as “flipping” A about its main diagonal, or as “rotating” A through 180◦about the line containing the main diagonal. This makes Property 2 in Theorem 2.1.2 transparent.

Example 2.1.10

Solve for A if

2AT −3

1 2 −1 1

T

=

2 3 −1 2

.

Solution. Using Theorem 2.1.2, the left side of the equation is

2AT −3

1 2 −1 1

T

= 2AT T−31 2 −1 1

T

= 2A−3

1 −1 2 1

Hence the equation becomes

2A−3

1 −1 2 1

=

2 3 −1 2

Thus 2A =

2 3 −1 2

+3

1 −1 2 1

=

5 0 5 5

, so finally A =12

5 0 5 5

=52

1 0 1 1

.

Note that Example 2.1.10 can also be solved by first transposing both sides, then solving for AT, and so obtaining A = (AT)T. The reader should do this.

The matrix D =

1 2 2 5

in Example 2.1.9 has the property that D = DT. Such matrices are important;

a matrix A is called symmetric if A = AT. A symmetric matrix A is necessarily square (if A is m×n, then ATis n×m, so A = ATforces n = m). The name comes from the fact that these matrices exhibit a symmetry about the main diagonal. That is, entries that are directly across the main diagonal from each other are

equal.





For example,

a b c b′ d e c′e′f



 is symmetric when b = b′, c = c′, and e = e′.

44 Matrix Algebra

Example 2.1.11

If A and B are symmetric n×n matrices, show that A+B is symmetric.

Solution. We have AT = A and BT = B, so, by Theorem 2.1.2, we have (A+B)T = AT +BT = A+B. Hence A+B is symmetric.

Example 2.1.12

Suppose a square matrix A satisfies A = 2AT. Show that necessarily A = 0.

Solution. If we iterate the given equation, Theorem 2.1.2 gives A = 2AT = 22AT T= 22(AT)T = 4A

Subtracting A from both sides gives 3A = 0, so A =13(0) = 0. **Exercises for 2.1**

Exercise 2.1.1 Find a, b, c, and d if

Exercise 2.1.2 Compute the following:

a.

3 2 1 5 1 0

−5

3 0 −2 1 −1 2

a b

c−3d −d

b.

3

3

−5

6

+7

1

a.

c d

=

2a+d a+b

c.

−1

−2 1 3 2

−4

2

1 −2 0 −1

−1

+3

2 −3 −1 −2

a−b b−c

1 1

3 −1 2 −29 3 4 +3 11 −6 d.

b.

c−d d −a

= 2

−3 1

T

e.

1 −5 4 0



T



f.

0 −1 2 1 0 −4

 

2 1 0 6

−2 4 0

a

b

1

T

g.

3 −1 2 1

−2

1 −2 1 1

c. 3

b

+2

a

=

2

h.

3

2 1 −1 0

T

−2

1 −1 2 3

d.

a b

=

b c

Exercise 2.1.3 Let A =

2 1 0 −1

,

c d

d a

B =

3 −1 2 0 1 4

, C =

3 −1 2 0

,

D =

 

1 3

−1 0 1 4



, and E =

1 0 1 0 1 0

.

Compute the following (where possible).

2.1. Matrix Addition, Scalar Multiplication, and Transposition 45 Exercise 2.1.9 If A is any 2×2 matrix, show that:

a. 5 3A−2B b. C

a. A = a

1 0 0 0

+ b

0 1 0 0

+ c

0 0 1 0

+

3ET

c. d. B+D

d

4A e. T −3C (A+C)T

f.

0 0 0 1

for some numbers a, b, c, and d.

g. 2B−3E h. A−D

b. A = p

1 0 0 1

+ q

1 1 0 0

+ r

1 0 1 0

+

i.

(B−2E)T

s

0 1 1 0

for some numbers p, q, r, and s.

Exercise 2.1.4 Find A if:

a. 5A−

1 0 2 3

= 3A−

5 2 6 1

Exercise 2.1.10 Let A =1 1 −1,

B =0 1 2 , and C =3 0 1 . If rA + sB +tC = 0 for some scalars r, s, and t, show that

b. 3A−

2 1

= 5A−2

3 0

necessarily r = s = t = 0.

Exercise 2.1.11

a. If Q+A = A holds for every m×n matrix A, show

Exercise 2.1.5 Find A in terms of B if:

a. 2 A+B = 3A+2B b. A−B = 5(A+2B)

Exercise 2.1.6 If X, Y, A, and B are matrices of the same size, solve the following systems of equations to obtain X and Y in terms of A and B.

a. 4X +3Y = A

that Q = 0mn.

b. If A is an m×n matrix and A+A′ = 0mn, show that A′ = −A.

Exercise 2.1.12 If A denotes an m×n matrix, show that A = −A if and only if A = 0.

Exercise 2.1.13 A square matrix is called a diagonal matrix if all the entries off the main diagonal are zero. If A and B are diagonal matrices, show that the following matrices are also diagonal.

5X +3Y = A 2X +Y = B

b.

5X +4Y = B

a. A+B b. A−B

Exercise 2.1.7 Find all matrices X and Y such that: a. 2X −5Y =1 2

c. kA for any number k

Exercise 2.1.14 In each case determine all s and t such

3X −2Y =3 −1

b.

that the given matrix is symmetric:

Exercise 2.1.8 Simplify the following expressions

a.

1 s −2 t

b.

s t st 1

where A, B, and C are matrices.



s 2s st





2 s t



c.  

d. 

a. 2[9(A−B) +7(2B−A)]

−2[3(2B+A)−2(A+3B)−5(A+B)]

t −1 s t s2s



2s 0 s+t 3 3 t

Exercise 2.1.15 In each case find the matrix A.

b. 5[3(A−B+2C)−2(3C−B)−A] +2[3(3A−B+C) +2(B−2A)−2C]

a.

A+3

1 −1 0 1 2 4

T

=

 

2 1 0 5 3 8

 

46 Matrix Algebra

b.

3AT +2

1 0 0 2

T

=

8 0 3 1

Exercise 2.1.20 A square matrix W is called skew symmetric if WT = −W. Let A be any square matrix.

c. 2A−31 2 0  T= 3AT +2 1 −1 T

a. Show that A−ATis skew-symmetric.

d.

2AT −5

1 0 −1 2

T

= 4A−9

1 1 −1 0

b. Find a symmetric matrix S and a skew-symmetric matrix W such that A = S+W.

Exercise 2.1.16 Let A and B be symmetric (of the same size). Show that each of the following is symmetric.

a. (A−B) b. kA for any scalar k

Exercise 2.1.17 Show that A+ATis symmetric for any square matrix A.

Exercise 2.1.18 If A is a square matrix and A = kAT where k 6= ±1, show that A = 0.

Exercise 2.1.19 In each case either show that the state ment is true or give an example showing it is false.

a. If A+B = A+C, then B and C have the same size. b. If A+B = 0, then B = 0.

c. If the (3, 1)-entry of A is 5, then the (1, 3)-entry of ATis −5.

d. A and AT have the same main diagonal for every matrix A.

e. If B is symmetric and AT = 3B, then A = 3B.

f. If A and B are symmetric, then kA + mB is sym metric for any scalars k and m.

c. Show that S and W in part (b) are uniquely deter mined by A.

Exercise 2.1.21 If W is skew-symmetric (Exer cise 2.1.20), show that the entries on the main diagonal are zero.

Exercise 2.1.22 Prove the following parts of Theo rem 2.1.1.

a. (k+ p)A = kA+ pA b. (kp)A = k(pA)

Exercise 2.1.23 Let A, A1, A2, ..., An denote matrices of the same size. Use induction on n to verify the follow ing extensions of properties 5 and 6 of Theorem 2.1.1.

a. k(A1 +A2 +···+An) = kA1 +kA2 +···+kAn for any number k

b. (k1 + k2 + ··· + kn)A = k1A + k2A + ··· + knA for any numbers k1, k2, ..., kn

Exercise 2.1.24 Let A be a square matrix. If A = pBT and B = qATfor some matrix B and numbers p and q, show that either A = 0 = B or pq = 1.

[Hint: Example 2.1.7.]

2.2. Matrix-Vector Multiplication 47

**2.2 Matrix-Vector Multiplication**

Up to now we have used matrices to solve systems of linear equations by manipulating the rows of the augmented matrix. In this section we introduce a different way of describing linear systems that makes more use of the coefficient matrix of the system and leads to a useful way of “multiplying” matrices.

**Vectors**

It is a well-known fact in analytic geometry that two points in the plane with coordinates (a1, a2) and (b1, b2) are equal if and only if a1 = b1 and a2 = b2. Moreover, a similar condition applies to points (a1, a2, a3) in space. We extend this idea as follows.

An ordered sequence (a1, a2, ..., an) of real numbers is called an ordered n-tuple. The word “or dered” here reflects our insistence that two ordered n-tuples are equal if and only if corresponding entries are the same. In other words,

(a1, a2, ..., an) = (b1, b2, ..., bn) if and only if a1 = b1, a2 = b2, ..., and an = bn. Thus the ordered 2-tuples and 3-tuples are just the ordered pairs and triples familiar from geometry.

Definition 2.4 The set Rn of ordered n-tuples of real numbers

Let R denote the set of all real numbers. The set of all ordered n-tuples from R has a special notation:

Rn

denotes the set of all ordered n-tuples of real numbers.

There are two commonly used ways to denote the n-tuples in Rn: As rows (r1, r2, ..., rn) or columns 



r1r2...

; the notation we use depends on the context. In any event they are called vectors or n-vectors and rn

will be denoted using bold type such as x or v. For example, an m×n matrix A will be written as a row of

columns:

A =a1 a2 ··· anwhere aj denotes column j of A for each j.

If x and y are two n-vectors in Rn, it is clear that their matrix sum x + y is also in Rnas is the scalar multiple kx for any real number k. We express this observation by saying that Rnis closed under addition and scalar multiplication. In particular, all the basic properties in Theorem 2.1.1 are true of these n-vectors. These properties are fundamental and will be used frequently below without comment. As for matrices in general, the n×1 zero matrix is called the zero n-vector in Rnand, if x is an n-vector, the n-vector −x is called the negative x.

Of course, we have already encountered these n-vectors in Section 1.3 as the solutions to systems of linear equations with n variables. In particular we defined the notion of a linear combination of vectors and showed that a linear combination of solutions to a homogeneous system is again a solution. Clearly, a linear combination of n-vectors in Rnis again in Rn, a fact that we will be using.

48 Matrix Algebra

**Matrix-Vector Multiplication**

Given a system of linear equations, the left sides of the equations depend only on the coefficient matrix A and the column x of variables, and not on the constants. This observation leads to a fundamental idea in linear algebra: We view the left sides of the equations as the “product” Ax of the matrix A and the vector x. This simple change of perspective leads to a completely new way of viewing linear systems—one that is very useful and will occupy our attention throughout this book.

To motivate the definition of the “product” Ax, consider first the following system of two equations in three variables:ax1 + bx2 + cx3 = b1



a′x1 + b′x2 + c′x3 = b1(2.2) 

and let A =

a b c a′ b′c′

, x =



x1 x2 x3

, b =

b1 b2

denote the coefficient matrix, the variable matrix, and

the constant matrix, respectively. The system (2.2) can be expressed as a single vector equation

ax1 + bx2 + cx3

a′x1 + b′x2 + c′x3

which in turn can be written as follows:

=

b1 b2

x1

a a′

+x2

b b′

+x3

c c′

=

b1 b2

Now observe that the vectors appearing on the left side are just the columns

a1 =

a a′

, a2 =

b b′

, and a3 =

c c′

of the coefficient matrix A. Hence the system (2.2) takes the form

x1a1 +x2a2 +x3a3 = b (2.3)

This shows that the system (2.2) has a solution if and only if the constant matrix b is a linear combination3 of the columns of A, and that in this case the entries of the solution are the coefficients x1, x2, and x3 in this linear combination.

Moreover, this holds in general. If A is any m×n matrix, it is often convenient to view A as a row of columns. That is, if a1, a2, ..., an are the columns of A, we write

A =a1 a2 ··· an

and say that A =a1 a2 ··· anis given in terms of its columns.

Now consider any system of linear equations with m × n coefficient matrix A. If b is the constant

matrix of the system, and if x =



x1

x2...

xn



is the matrix of variables then, exactly as above, the system can

be written as a single vector equation

x1a1 +x2a2 +···+xnan = b (2.4)

3Linear combinations were introduced in Section 1.3 to describe the solutions of homogeneous systems of linear equations. They will be used extensively in what follows.

2.2. Matrix-Vector Multiplication 49

Example 2.2.1

Write the system

 

3x1 + 2x2 − 4x3 = 0 x1 − 3x2 + x3 = 3 x2 − 5x3 = −1

in the form given in (2.4).

Solution.

x1

 

3 1 0



+x2

 

2

−3 1



+x3

 

−4 1

−5



 =

 

0

3

−1

 

As mentioned above, we view the left side of (2.4) as the product of the matrix A and the vector x. This basic idea is formalized in the following definition:

Definition 2.5 Matrix-Vector Multiplication

Let A =



x1

a1 a2 ··· an 



be an m×n matrix, written in terms of its columns a1, a2, ..., an. If

x =

x2

is any n-vector, the product Ax is defined to be the m-vector given by: .

.

.

xn

Ax = x1a1 +x2a2 +···+xnan

In other words, if A is m×n and x is an n-vector, the product Ax is the linear combination of the columns of A where the coefficients are the entries of x (in order).

Note that if A is an m×n matrix, the product Ax is only defined if x is an n-vector and then the vector Ax is an m-vector because this is true of each column aj of A. But in this case the system of linear equations with coefficient matrix A and constant vector b takes the form of a single matrix equation

Ax = b

The following theorem combines Definition 2.5 and equation (2.4) and summarizes the above discussion. Recall that a system of linear equations is said to be consistent if it has at least one solution.

Theorem 2.2.1

1. Every system of linear equations has the form Ax = b where A is the coefficient matrix, b is the constant matrix, and x is the matrix of variables.

2. The system Ax = b is consistent if and only if b is a linear combination of the columns of A.

3. If a1, a2, ..., an are the columns of A and if x =



x1

x2

.

.

.

xn





, then x is a solution to the linear

50 Matrix Algebra

system Ax = b if and only if x1, x2, ..., xn are a solution of the vector equation x1a1 +x2a2 +···+xnan = b

A system of linear equations in the form Ax = b as in (1) of Theorem 2.2.1 is said to be written in matrix form. This is a useful way to view linear systems as we shall see.

Theorem 2.2.1 transforms the problem of solving the linear system Ax = b into the problem of ex pressing the constant matrix B as a linear combination of the columns of the coefficient matrix A. Such a change in perspective is very useful because one approach or the other may be better in a particular situation; the importance of the theorem is that there is a choice.

Example 2.2.2

If A =



2 −1 3 5 

0 2 −3 1 −3 4 1 2



 and x =



210 −2



, compute Ax.

Solution. By Definition 2.5: Ax = 2 Example 2.2.3

 

2

0

−3



+1

 

−1 2

4



+0

 

3

−3 1



−2

 

5 1 2



 =

 

−7 0

−6



.

Given columns a1, a2, a3, and a4 in R3, write 2a1 −3a2 +5a3 +a4 in the form Ax where A is a matrix and x is a vector.

Solution. Here the column of coefficients is x =



2

−3

5

1



. Hence Definition 2.5 gives

Ax = 2a1 −3a2 +5a3 +a4

where A =a1 a2 a3 a4is the matrix with a1, a2, a3, and a4 as its columns.

Example 2.2.4





2

Let A =a1 a2 a3 a4be the 3×4 matrix given in terms of its columns a1 =



,

0



1





3





3



−1

a2 =



1 1

, a3 =



−1 −3

, and a4 =



. In each case below, either express b as a linear

1

0

2.2. Matrix-Vector Multiplication 51

combination of a1, a2, a3, and a4, or show that it is not such a linear combination. Explain what your answer means for the corresponding system Ax = b of linear equations.

a. b =

 







1

4

 b. b = 



2

2

3

1

Solution. By Theorem 2.2.1, b is a linear combination of a1, a2, a3, and a4 if and only if the system Ax = b is consistent (that is, it has a solution). So in each case we carry the augmented matrix [A|b] of the system Ax = b to reduced form.

a. Here



2 1 3 3 1 

0 1 −1 1 2 −1 1 −3 0 3



 →



1 0 2 1 0 

0 1 −1 1 0 0 0 0 0 1



, so the system Ax = b has no

solution in this case. Hence b is not a linear combination of a1, a2, a3, and a4.

b. Now



2 1 3 3 4 

0 1 −1 1 2 −1 1 −3 0 1



 →



1 0 2 1 1 

0 1 −1 1 2 0 0 0 0 0



, so the system Ax = b is consistent.

Thus b is a linear combination of a1, a2, a3, and a4 in this case. In fact the general solution is x1 = 1−2s−t, x2 = 2+s−t, x3 = s, and x4 = t where s and t are arbitrary parameters. Hence



4



2

x1a1 +x2a2 +x3a3 +x4a4 = b = 1



 for any choice of s and t. If we take s = 0 and t = 0, this

becomes a1 +2a2 = b, whereas taking s = 1 = t gives −2a1 +2a2 +a3 +a4 = b.

Example 2.2.5

Taking A to be the zero matrix, we have 0x = 0 for all vectors x by Definition 2.5 because every column of the zero matrix is zero. Similarly, A0 = 0 for all matrices A because every entry of the zero vector is zero.

Example 2.2.6

If I =

 

1 0 0 0 1 0 0 0 1



, show that Ix = x for any vector x in R3.

Solution. If x =

 

x1 x2 x3



 then Definition 2.5 gives











Ix = x1

 

1 0 0



+x2

 

0 1 0



+x3

 

0 0 1



 =

 

x1 0

0



+

 

0

x2 0

+



0

0

x3

 =



x1 x2 x3

 = x

52 Matrix Algebra

The matrix I in Example 2.2.6 is called the 3×3 identity matrix, and we will encounter such matrices again in Example 2.2.11 below. Before proceeding, we develop some algebraic properties of matrix-vector multiplication that are used extensively throughout linear algebra.

Theorem 2.2.2

Let A and B be m×n matrices, and let x and y be n-vectors in Rn 1. A(x+y) = Ax+Ay.

2. A(ax) = a(Ax) = (aA)x for all scalars a.

3. (A+B)x = Ax+Bx.

. Then:

Proof. We prove (3); the other verifications are similar and are left as exercises. Let A =a1 a2 ··· an

and B =b1 b2 ··· bnbe given in terms of their columns. Since adding two matrices is the same as adding their columns, we have

A+B =a1 +b1 a2 +b2 ··· an +bn

If we write x =



x1

x2...

xn



Definition 2.5 gives

(A+B)x = x1(a1 +b1) +x2(a2 +b2) +···+xn(an +bn)

= (x1a1 +x2a2 +···+xnan) + (x1b1 +x2b2 +···+xnbn)

= Ax+Bx

|  |
| --- |

Theorem 2.2.2 allows matrix-vector computations to be carried out much as in ordinary arithmetic. For example, for any m×n matrices A and B and any n-vectors x and y, we have:

A(2x−5y) = 2Ax−5Ay and (3A−7B)x = 3Ax−7Bx

We will use such manipulations throughout the book, often without mention.

**Linear Equations**

Theorem 2.2.2 also gives a useful way to describe the solutions to a system

Ax = b

of linear equations. There is a related system

Ax = 0

2.2. Matrix-Vector Multiplication 53

called the associated homogeneous system, obtained from the original system Ax = b by replacing all the constants by zeros. Suppose x1 is a solution to Ax = b and x0 is a solution to Ax = 0 (that is Ax1 = b and Ax0 = 0). Then x1 +x0 is another solution to Ax = b. Indeed, Theorem 2.2.2 gives

A(x1 +x0) = Ax1 +Ax0 = b+0 = b

This observation has a useful converse.

Theorem 2.2.3

Suppose x1 is any particular solution to the system Ax = b of linear equations. Then every solution x2 to Ax = b has the form

x2 = x0 +x1

for some solution x0 of the associated homogeneous system Ax = 0.

Proof. Suppose x2 is also a solution to Ax = b, so that Ax2 = b. Write x0 = x2 −x1. Then x2 = x0 +x1 and, using Theorem 2.2.2, we compute

Ax0 = A(x2 −x1) = Ax2 −Ax1 = b−b = 0

Hence x0 is a solution to the associated homogeneous system Ax = 0.

|  |
| --- |

Note that gaussian elimination provides one such representation.

Example 2.2.7

Express every solution to the following system as the sum of a specific solution plus a solution to the associated homogeneous system.

x1 − x2 − x3 + 3x4 = 2

2x1 − x2 − 3x3 + 4x4 = 6

x1 − 2x3 + x4 = 4

Solution. Gaussian elimination gives x1 = 4+2s−t, x2 = 2+s+2t, x3 = s, and x4 = t where s and t are arbitrary parameters. Hence the general solution can be written



x = 



x1

x2

x3

x4



=



4+2s−t

2+s+2t s

t



=



420 0



+



s



211 0



+t 



−120 1



 

 





Thus x1 =

420 0

is a particular solution (where s = 0 = t), and x0 = s

211 0

+t

−120 1

gives all

solutions to the associated homogeneous system. (To see why this is so, carry out the gaussian elimination again but with all the constants set equal to zero.)

The following useful result is included with no proof.

54 Matrix Algebra

Theorem 2.2.4

Let Ax = b be a system of equations with augmented matrixA b

. Write rank A = r.

1. rank A b

is either r or r +1.

2. The system is consistent if and only if rank A b= r.

3. The system is inconsistent if and only if rank A b= r +1.

**The Dot Product**

Definition 2.5 is not always the easiest way to compute a matrix-vector product Ax because it requires that the columns of A be explicitly identified. There is another way to find such a product which uses the matrix A as a whole with no reference to its columns, and hence is useful in practice. The method depends on the following notion.

Definition 2.6 Dot Product in Rn

If (a1, a2, ..., an) and (b1, b2, ..., bn) are two ordered n-tuples, their dot product is defined to be the number

a1b1 +a2b2 +···+anbn

obtained by multiplying corresponding entries and adding the results.

To see how this relates to matrix products, let A denote a 3×4 matrix and let x be a 4-vector. Writing

x =



x1

x2

x3

x4



and A =



a11 a12 a13 a14 

a21 a22 a23 a24 a31 a32 a33 a34

 

in the notation of Section 2.1, we compute

Ax =

 

a11 a12 a13 a14 a21 a22 a23 a24 a31 a32 a33 a34

 



x1

x2

x3

x4



= x1 

 

a11 a21 a31



+x2

 

a12 a22 a32



+x3

 

a13 a23 a33 



+x4

 

a14 a24 a34

 

=

a11x1 +a12x2 +a13x3 +a14x4 

a21x1 +a22x2 +a23x3 +a24x4 a31x1 +a32x2 +a33x3 +a34x4



From this we see that each entry of Ax is the dot product of the corresponding row of A with x. This computation goes through in general, and we record the result in Theorem 2.2.5.

2.2. Matrix-Vector Multiplication 55

Theorem 2.2.5: Dot Product Rule

Let A be an m×n matrix and let x be an n-vector. Then each entry of the vector Ax is the dot product of the corresponding row of A with x.

This result is used extensively throughout linear algebra.

If A is m × n and x is an n-vector, the computation of Ax by the dot product rule is simpler than using Definition 2.5 because the computation can be carried out directly with no explicit reference to the columns of A (as in Definition 2.5). The first entry of Ax is the dot product of row 1 of A with x. In hand calculations this is computed by going across row one of A, going down the column x, multiplying corresponding entries, and adding the results. The other entries of Ax are computed in the same way using the other rows of A with the column x.

In general, compute entry i of Ax as follows (see the diagram):

A x Ax 



 

 



 =

 



Go across row i of A and down column x, multiply corre 

sponding entries, and add the results.

row i entry i Example 2.2.8

As an illustration, we rework Example 2.2.2 using the dot product rule instead of Definition 2.5.

If A =



2 −1 3 5 

0 2 −3 1 −3 4 1 2



 and x =



210 −2



, compute Ax.

Solution. The entries of Ax are the dot products of the rows of A with x:

Ax =



2 −1 3 5 

0 2 −3 1 −3 4 1 2

 



210 −2



=



2 · 2 + (−1)1 + 3 · 0 + 5(−2) 

0 · 2 + 2 · 1 + (−3)0 + 1(−2) (−3)2 + 4 · 1 + 1 · 0 + 2(−2)



 =

 

−7 0

−6

 

Of course, this agrees with the outcome in Example 2.2.2.

Example 2.2.9

Write the following system of linear equations in the form Ax = b.

5x1 − x2 + 2x3 + x4 − 3x5 = 8

x1 + x2 + 3x3 − 5x4 + 2x5 = −2

−x1 + x2 − 2x3 + − 3x5 = 0

56 Matrix Algebra





Solution. Write A =

5 −1 2 1 −3 1 1 3 −5 2 −1 1 −2 0 −3





, b = 

 

8

−2 0



, and x =



x1

x2

x3

x4

x5



. Then the dot



product rule gives Ax =

5x1 − x2 + 2x3 + x4 − 3x5 x1 + x2 + 3x3 − 5x4 + 2x5 −x1 + x2 − 2x3 − 3x5

, so the entries of Ax are the left sides of

the equations in the linear system. Hence the system becomes Ax = b because matrices are equal if and only corresponding entries are equal.

Example 2.2.10

If A is the zero m×n matrix, then Ax = 0 for each n-vector x.

Solution. For each k, entry k of Ax is the dot product of row k of A with x, and this is zero because row k of A consists of zeros.

Definition 2.7 The Identity Matrix

For each n > 2, the identity matrix In is the n×n matrix with 1s on the main diagonal (upper left to lower right), and zeros elsewhere.

The first few identity matrices are

I2 =

1 0 0 1

, I3 =

 

1 0 0 0 1 0 0 0 1



, I4 =



1 0 0 0

0 1 0 0 0 0 1 0 0 0 0 1



, ...

In Example 2.2.6 we showed that I3x = x for each 3-vector x using Definition 2.5. The following result shows that this holds in general, and is the reason for the name.

Example 2.2.11

For each n ≥ 2 we have Inx = x for each n-vector x in Rn. Solution. We verify the case n = 4. Given the 4-vector x =



x1

x2

x3



the dot product rule gives











x4







I4x =

1 0 0 0

0 1 0 0 0 0 1 0 0 0 0 1



x1

x2

x3

x4

=

x1 +0+0+0

0+x2 +0+0 0+0+x3 +0 0+0+0+x4

=

x1

x2

x3

x4

= x