Linear Algebra

An Introduction to Linear Algebra for Pre-Calculus Students

by

Tamara Anthony Carter
Adjunct Professor of Mathematics, University of North Texas

Richard A. Tapia
Noah Harding Professor, Computational and Applied Mathematics
Director, Center for Excellence and Equity in Education

Anne Papakonstantinou
Director, Rice University School Mathematics Project

Rice University

May 1995

Contents

Preface v

1 **Introduction to Matrices** **1**

2 **Addition of Matrices** **11**

3 **Multiplication of Matrices** **18**

4 **Equations** **33**

4.1 Coding 43

5 **Determinants** **54**

5.1 Expansion by Minors 56

5.2 Using Gaussian Elimination 58

5.3 Inverses and Solutions to Systems 60

5.4 Definition of the Determinant 62

6 **Consistent and Inconsistent Systems** **68**

First Review **75**

7 **Markov Chains** **80**

8 **Least Squares Approximation** **95**

8.1 Proof of Normal Equations 107

9 Eigenvalues and Eigenvectors **114**

10 Numerical Challenges **132**

10.1 Operation Counts 136

Second Review **140**

11 Tests **143**

11.1 Matrix Test 1A 143

11.2 Matrix Test 1B 145

11.3 Matrix Test 1C 147

11.4 Matrix Test 2A 150

11.5 Matrix Test 2B 152

12 Solutions **154**

12.1 Solutions to Introduction - Problems from page 8 154

12.2 Solutions to Addition - Problems from page 15 161

12.3 Solutions to Multiplication - Problems from page 30 170

12.4 Solutions to Systems of Equations - Problems from page 51 183

12.5 Solutions to Determinants - Problems from page 64 188

12.6 Solutions to Consistent and Inconsistent - Problems from page 74 192

12.7 Solutions to First Review from page 75 193

12.8 Solutions to Markov Chains - Problems from page 90 199

12.9 Solutions to Least Squares - Problems from page 110 202

12.10 Solutions to Eigenpairs - Problems from page 130 205

12.11 Solutions to Second Review from page 140 206

12.12 Solutions to Matrix Test 1A 208

12.13 Solutions to Matrix Test 1B 210
12.14 Solutions to Matrix Test 1C 212
12.15 Solutions to Matrix Test 2A 214
12.16 Solutions to Matrix Test 2B 215

Bibliography **217**

Index **218**

Preface

This book is an introduction to linear algebra for pre-calculus students. It is a stand-alone unit in the sense that no prior knowledge of matrices is assumed. Students with experience in general mathematics, up to and including Algebra I, should be able to comprehend the material. However, most students have not had experience with the topics in the latter chapters, so the pace of the course should allow for the students to spend extra time with these chapters. We begin with chapters that explain the matrix operations of addition, subtraction, scalar multiplication, and matrix multiplication. These topics are covered in most pre-calculus texts that are currently in use. This unit also allows the students to explore the notions of inverse, determinant, and consistent and inconsistent systems; these topics are covered in some pre-calculus text books. Our unit also provides the students with an introduction to Markov chains, curve fitting, eigenpairs, and some of the numerical challenges that are encountered when matrices are used to solve real-world problems. These latter topics are rarely addressed in pre-calculus texts. The unit was created from elementary principles with significant input from Rice University faculty and students. Various current texts, recommendations from the National Council of Teachers of Mathematics (NCTM), and the Texas essential elements were examined in order to determine which topics should and should not be included in this text.

The state of Texas significantly influences the content of pre-college text books, because a book that is approved for adoption in Texas has a large potential market. In order for a book to be adopted in Texas, all of the "essential elements" for that course must be covered in the text. Essential element 4B for Algebra II states that students should be able to "use augmented matrices by hand or by computer to solve two- or three-variable linear systems." Essential element 3I for Elementary Analysis states that students should be able to "solve matrix equations and real-world problems whose solutions involve matrix equations." According to essential element 3H for Elementary Analysis, students should be able to "solve a system of equations or inequalities using graphing techniques and apply [them] in real-world situations" (*Houston Independent School District Scope and Sequence Grades 9-12*, 1992). (This text does not address essential element 3H because only two-dimensional problems can be solved with graphing techniques and "real world problems" require many more dimensions.) Since these are the only requirements concerning matrices for Texas high-school mathematics books, many books meet these requirements but do not really give the students an adequate understanding of linear algebra. It is true that a college linear algebra text would contain ample detail, but few pre-calculus students have the mathematical maturity necessary to read these texts.

Since most pre-calculus texts only touch on the subject of matrices, one might question the need for a more in-depth study of linear algebra at the pre-calculus level. The NCTM has recognized this need and stated that "matrices and their applications" should receive "increased attention" in high school (*Curriculum and Evaluation Standards for School Mathematics*, 1989, p. 126). It also could be argued that linear algebra is as important as calculus to many engineers and other scientists. The introduction of linear algebra at the pre-calculus level would give the students a knowledge base on which to build when they study linear algebra in college. The arrays that are studied in linear algebra are of vital importance to computer programmers and computer users. Linear algebra is also central to the computational and mathematical sciences.

In *The Psychology of Learning Mathematics*, Richard Skemp states "...the learning of mathematics, especially in its early stages and for the average student, [is] very dependent on good teaching" (1987, p. 21). Unfortunately, many teachers have not had much experience, and do not feel entirely comfortable, with linear algebra, so it is difficult for them to teach more than just the procedures of matrix manipulation. Therefore, we have attempted to write this unit so that the students can directly access the material. Since discussions add to, and strengthen, one's understanding of a topic, thought-provoking questions and their answers are provided at the end of each chapter to spark class discussions. To help the teacher, complete solution steps have been provided in addition to the solutions where appropriate.

Skemp also states that "concepts of a higher order than those which people already have cannot be communicated to them by a definition [alone], but only by arranging for them to encounter a suitable collection of examples" (1987, p. 18). For this reason, most new concepts in this unit are presented with an example that builds on the intuition of the student. Then the formal definition is given, and other examples follow to clarify the concept. This helps motivate the students because they can immediately see a use for the concept. It also gives the concept a foundation in the mind of the student. Although the notion of building concepts in this manner seems logical, few text books utilize this approach.

Because many books teach procedures rather than concepts, the students do not receive enough information to expand beyond the examples in the book. For example, some books teach methods which apply only to the special case of 2 by 2 matrices when they address the notions of inverse and determinant. However, this text presents methods for finding inverses and determinants of square matrices of any size. Since the students learn the concepts and these general methods, their knowledge is not restricted by the examples in the book.

Because computers are essential to modern society, computer programming assignments are included at the end of the first three chapters as a means to help the students solidify their knowledge of matrices. In some of the later chapters, students are encouraged to use calculators to help them explore matrices so that they are not tied to problems that can be reasonably computed by hand. The students are not asked to use a particular computer language or a particular calculator, but the code for working programs are provided for the teachers in BASIC and PASCAL.

When new topics are introduced in this unit, they are tied, as much as possible, to previous topics. This is in an attempt to allow the students to appreciate linear algebra as a whole rather than view each chapter as a separate entity. For example, solutions to systems of equations are computed using Gaussian elimination, Gauss-Jordan elimination, and Cramer's rule. The text demonstrates to the students that a form of Gaussian elimination can be used, as an alternative to expansion by minors, to compute determinants. We use these two methods of computing the determinant to discuss efficiency of algorithms so that the students know that finding the correct answer is not the only concern. The students are also told that the determinant of a matrix is the same as the product of its eigenvalues. The relationship between the steady state of a transition matrix and eigenpairs is also demonstrated to the students. These ties and others provide the students with different perspectives from which to view problems.

Most texts do not mention Markov chains, even though they are well within the grasp of pre-calculus students. The NCTM believes that "in grades 9-12, the mathematics curriculum should include topics from discrete mathematics so that all students can represent graphs, matrices, sequences, and recurrence relations" (1989, p. 176). Hopefully, this chapter will catch the attention of the student because a Markov chain is a real application of matrices rather than a contrived book example. This chapter also offers the student a glance into the fascinating world of probability.

Curve fitting is another interesting application of matrix equations, and it can be used immediately in the life of a pre-calculus student. Most pre-calculus students take a laboratory science in which they could use curve fitting to analyze their data. This cross-discipline application also helps the students to view mathematics as a useful tool rather than just a subject to take in school. The NCTM believes that "in grades 9-12, the mathematics curriculum should include the continued study of data analysis and statistics so that all students can use curve fitting to predict from data" (1989, p. 167). Curve fitting is also a bridge between the fields of mathematics and statistics.

Eigenpairs are essentially never found in pre-calculus text books, but they have a wide range of physical applications that could interest students. The computational methods taught in this unit build naturally on previous topics in the text. Because the computation of eigenpairs quickly increases in difficultly as the size of the matrix increases, only simple examples are given in this unit. However, students are introduced to the concept of matrices and to many of their applications, so they will have a foundation on which to build when they study eigenpairs in college.

The chapter entitled "Numerical Challenges" is important to the students' overall knowledge even though the students are not asked to perform computations. This chapter reminds students that the world is not solely comprised of pretty book examples. It also helps dispel the notion that mathematics is only about learning what other people already know. It is good for students to know that many important challenges remain in mathematics and that bright young minds are needed to research these topics.

This entire unit was written so that pre-calculus teachers and students will have a text that clearly and accurately explains the introductory concepts of linear algebra. It explores the topics that are currently addressed in pre-calculus courses, but emphasizes concepts rather than than just procedures. This unit also provides students with many more real-world linear algebra topics to explore than are presented in current texts. It is hoped that this unit will not only help students understand linear algebra, but will also spark an interest in, and an appreciation for, the mathematical sciences.

> How can it be that mathematics, being after all a product of human thought which is independent of experience, is so admirably appropriate to the objects of reality? - Albert Einstein

It truly is beautiful that the abstract concepts of mathematics can be used to model the world around us. This is astounding because the laws of mathematics were not created with the universe, but have been defined by mankind over the centuries. These laws model the world so well, that people often fail to distinguish between the real situation and the mathematical model that is being used to study it. For example, because matrices can be used to represent a system of equations which model the real world, people often think that the solution to the system will also be the solution to the real-world problem. However, the solution is only as good as the model that was used to represent the problem. Since the world is so complex, mathematical models cannot accurately model every detail of the universe. However, they may come amazingly close and help illuminate many of the mysteries of the universe.

Chapter 1

Introduction to Matrices

If you were asked for your weight in pounds, you would use a real number such as 140 to answer the question. If you were asked for your height in inches, you would answer with another real number such as 66.5. If we asked these questions to everyone in the class, we would want some way to know which weight goes with which height. One way to organize this data is to use an ordered pair. We could represent your weight and height with the ordered pair (140, 66.5). This is called an **ordered pair** because we always list the information in the same order. In other words, we list weight first and then height in every pair of numbers, so (140, 66.5) would be different from (66.5, 140). The **elements** are the individual pieces of information. Elements are also referred to as entries or components. In this book, we will only use real numbers as elements. The elements of this ordered pair are 140 and 66.5. We could also ask you for your age in years and append that information so that we have the ordered triple (140, 66.5, 18). We could ask you for n pieces of information, where n is any counting number. If we arrange the n pieces of information in a specific order, we call it an ordered n-tuple. In general, lists of ordered information are called **vectors**. If we write them in rows, as we did above, we call them **row vectors**. If we write them in columns, such as $\begin{bmatrix} 140 \\ 66.5 \end{bmatrix}$ and $\begin{bmatrix} 140 \\ 66.5 \\ 18 \end{bmatrix}$, we call them **column vectors**.

Definition 1.1 A **real** $n-$**vector** is an ordered n-tuple of real numbers.

The real numbers are called the **elements** of the vector.

Since we are only working with real numbers in this book, we will drop the word real when referring to vectors. When it is not important to specify how many elements are in the vector, we drop the qualifier n.

> **Remark 1** Did you notice that we used parentheses on some vectors and brackets on others? Actually, both are accepted notations, but we will use brackets for consistency throughout the rest of the book.

> **Remark 2** Sometimes you will see the elements of a row vector separated by commas. Commas are not necessary unless confusion can arise without the use of commas.

If you were asked to add, subtract, or multiply real numbers, you would know what to do. If we are going to use vectors to help us organize our information, we also need rules for vectors so that when we add, subtract, or multiply vectors, we get the same solutions as if we had not organized our data this way. Remember that vectors are simply tools that we use to display information in an organized manner. Therefore, we do not want our solutions to change just because we organized our data into a vector. As we study this book, we will learn more about how to perform mathematical operations with vectors.

Consider the following information:

> The Cardinals win seven, lose six, and tie one. The Eagles win five, lose eight, and tie one. The Falcons win two, lose twelve, and have no ties. The Owls win nine, lose five, and have no ties.

We can represent this data using the four vectors $\begin{bmatrix} 7 & 6 & 1 \end{bmatrix}$, $\begin{bmatrix} 5 & 8 & 1 \end{bmatrix}$, $\begin{bmatrix} 2 & 12 & 0 \end{bmatrix}$, and $\begin{bmatrix} 9 & 5 & 0 \end{bmatrix}$. However, it would be nice if we could combine all these vectors together into one set of data. If we consider each vector as one row of an array, then we will have all our data in one arrangement.

Definition 1.2 A **real matrix** is an arrangement of real numbers into rows and columns.

The real numbers are called the **elements** of the matrix.

Since we are only working with real numbers in this book, we also will drop the word real when referring to matrices. Notice that a vector is a special matrix that has only one row or one column. When we organize our vectors into a matrix, it could look like this:

$$\begin{bmatrix} 7 & 6 & 1 \\ 5 & 8 & 1 \\ 2 & 12 & 0 \\ 9 & 5 & 0 \end{bmatrix}$$

Basically, we have all of our information organized into one arrangement called a matrix.

We have all the appropriate numbers in our matrix, but if we want to know which numbers correspond to which team, we have to look back at our paragraph. For this reason, we often label our **matrices** (plural of matrix). Labels are not a formal part of the matrix, but they are very useful. Our matrix could look like this after it has been labeled:

$$\begin{array}{c} \\ C \\ E \\ F \\ O \end{array} \begin{array}{c} \begin{array}{ccc} W & L & T \end{array} \\ \begin{bmatrix} 7 & 6 & 1 \\ 5 & 8 & 1 \\ 2 & 12 & 0 \\ 9 & 5 & 0 \end{bmatrix} \end{array}$$

Just by looking at this matrix, we can tell that the Owls won the most games and the Falcons lost the most. One of the advantages of matrices is that information is easier to see and compare than when it is not organized into a matrix.

This matrix is referred to as a **4 by 3 matrix** (often written 4×3) because there are 4 rows and 3 columns. Therefore, the **dimensions** of this matrix are 4 by 3. The dimensions of a matrix tell you the "size" of the matrix because they tell you the number of rows and columns in the matrix. By convention, we list the number of rows before the number of columns.

> **Definition 1.3** The **dimensions** of a matrix are the number of rows and columns (listed in that order) of the matrix.

Each element of the matrix is named according to its position. Typically, capital letters represent matrices and small letters with subscripts represent elements in the matrix. Since vectors can be considered to be matrices with only one row or one column, they could be labled with capital letters also. However, vectors are usually represented by small letters. If we name the above matrix A, the element 6 is in the position a_{12} (read a one two) because it is in row 1 and column 2. Also by convention, we list the row number of the element before the column number. An element in row i and column j would be denoted by a_{ij}. This gives us a compact way to refer to specific elements of a matrix.

> **Remark 3** Although some mathematicians make a distinction between a 1 by 1 matrix, a 1-vector, and a real number, we will not make any distinction between them and will treat them exactly the same.

Can you represent the same information as before in a 3 by 4 matrix? Yes, you can. It would look like the matrix B which follows.

$$A = \begin{array}{c} \\ C \\ E \\ F \\ O \end{array} \begin{array}{c} \begin{array}{ccc} W & L & T \end{array} \\ \begin{bmatrix} 7 & 6 & 1 \\ 5 & 8 & 1 \\ 2 & 12 & 0 \\ 9 & 5 & 0 \end{bmatrix} \end{array}$$

$$B = \begin{array}{c} \\ W \\ L \\ T \end{array} \begin{array}{c} \begin{array}{cccc} C & E & F & O \end{array} \\ \begin{bmatrix} 7 & 5 & 2 & 9 \\ 6 & 8 & 12 & 5 \\ 1 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Matrix B is the **transpose** of A, and A is the **transpose** of B. **Transposing** a matrix results in writing the columns as rows and the rows as columns, but what really happens is that element a_{ij} is placed in the position b_{ji} of the new matrix. Therefore, a_{12} moves to the position b_{21} when we form the transpose of A. The transpose of A is denoted by A^T (read A transpose). Therefore, matrix B is A^T.

Definition 1.4 By the **transpose** of the m by n matrix A, denoted by A^T, we mean the n by m matrix which has a_{ji} as its $(i, j)^{th}$ element.

Definition 1.5 We say that two m by n matrices, A and B are **equal** if their corresponding elements are equal.

In other words, $A = B$ if A and B have the same dimensions and $a_{11} = b_{11}$, $a_{12} = b_{12}$, etc. Is $A = A^T$? Usually not, but we have a special word for a matrix which satisfies $A = A^T$.

Definition 1.6 A matrix is said to be **symmetric** if $A = A^T$.

Observe that the following matrix is symmetric:

$$S = \begin{bmatrix} 9 & 2 & 5 & 1 \\ 2 & 7 & 0 & 8 \\ 5 & 0 & 4 & 6 \\ 1 & 8 & 6 & 3 \end{bmatrix}.$$

Notice that $a_{ij} = a_{ji}$ for all i and j; as is true for all symmetric matrices. Symmetric matrices are easy to spot because if you draw a line down the main diagonal (from 9 to 3 in this matrix), then the two halves are mirror images of each other. Symmetric matrices have many special qualities that will be used when you study matrices in more detail. The matrix S, given above, has another special property; it is a **square matrix** because S has the same number of rows as columns. Notice that S is a 4 by 4 square matrix. We said that the main diagonal for S runs from 9 to 3. For any square matrix, the **main diagonal** runs from the upper left corner to the lower right corner.

Definition 1.7 We say that an m by n matrix is **square** if $m = n$.

Questions

1. For a matrix A, what is the transpose of A^T?

2. Does a symmetric matrix have to be square?

3. Are all square matrices symmetric?

Answers

1. Let us choose a generic matrix. We need to be careful when choosing a generic matrix. Vectors and square matrices often have special properties, so we will

not use them unless they are specifically needed. Let us follow our rules for transposes on this generic matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

Now let's form the transpose of A^T using the same rules.

$$\left(A^T\right)^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Notice that $\left(A^T\right)^T = A$. This is true for all matrices, but we have only proven it for 2 by 3 matrices. For a general proof, let us follow a general element of the matrix A, a_{ij}. Initially, it is in position (i, j) of matrix A. It is in position (j, i) of A^T and in position (i, j) of matrix $\left(A^T\right)^T$. This is true for every element of every matrix, so $\left(A^T\right)^T = A$ is true in general.

2. Yes. When a matrix is transposed, the columns become rows and the rows become columns. If $A = A^T$, the matrix must have the same number of rows as columns.

> **Remark 4** A **counterexample** is an example that illustrates that the statement in question is false. When you want to prove that a statement is false, a counterexample is sufficient. However, an example is not sufficient to prove that a statement is true. For instance, you could use your father as an example of the statement that "all humans are male" because he is human and male. However, we know that there are humans that are not male; your mother is a good counterexample to the statement "all humans are male." Therefore, an

example cannot be used to prove that a statement is true. You would have to show that it is true for ALL cases. For our example, you would have to establish in some way that EVERY human is male.

3. No. A counterexample is $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Since $a_{12} \neq a_{21}$, A is not symmetric.

Problems

1. (a) Form a 4 by 5 matrix, B, such that $b_{ij} = i * j$, where $*$ represents multiplication.

   (b) What is B^T?

   (c) Is B symmetric? Why or why not?

2. Using matrix A below, spell words by replacing each element requested with the letter in that position of the matrix. For example, $a_{52}a_{21}a_{32}$ represents cat.

$$A = \begin{bmatrix} \text{z} & \text{e} & \text{l} & \text{i} & \text{g} \\ \text{a} & \text{h} & \text{p} & \text{r} & \text{w} \\ \text{k} & \text{t} & \text{y} & \text{f} & \text{n} \\ \text{o} & \text{x} & \text{s} & \text{u} & \text{j} \\ \text{b} & \text{c} & \text{m} & \text{v} & \text{d} \end{bmatrix}$$

   (a) $a_{53}a_{21}a_{32}a_{24}a_{14}a_{52}a_{12}a_{43}$ $\quad a_{21}a_{24}a_{12}$ $\quad a_{12}a_{21}a_{43}a_{33}$

   (b) $a_{34}a_{24}a_{14}a_{12}a_{35}a_{55}a_{43}$

   (c) $a_{52}a_{21}a_{13}a_{52}a_{44}a_{13}a_{21}a_{32}a_{41}a_{24}$

   (d) Make up a statement of your own using the information given in this matrix. Write the statement using matrix elements and translate it. You have every letter of the alphabet in the matrix except the letter q.

3. (a) Put the following information into a 3 by 2 matrix and attach labels:

   The Lions won 5 games and lost 8. The Tigers won 9 and lost 4. The Bears won 7 and lost 6.

   (b) Transpose the matrix from part (a) and attach labels.

4. (a) Each team played 15 games. They either won, lost, or tied each game. Put the following information into a 3 by 4 matrix and attach labels:

   The Snakes won 6 and lost 8. The Lizards won 8 and lost 7. The Frogs won 9 and tied 2. The Toads lost 9 and tied 1.

   (b) Transpose the matrix from part (a) and attach labels.

5. (a) Put the following information into a 5 by 3 matrix and attach labels:

   Amit has a 3.48 GPA and scored 160 on his PSAT. His SAT score is 1580. Perry scored 121 on his PSAT and 1320 on the SAT. He has a 3.65 GPA. Don's GPA is 2.76, his SAT score is 840, and his PSAT score is 102. Heather scored 1260 on her SAT and 99 on the PSAT. She maintains a 3.80 GPA. Shelly scored 980 on the SAT and 83 on the PSAT with a GPA of 3.01.

   (b) Transpose the matrix from part (a) and attach labels.

Computer Project

Write a computer program that will form a matrix from the numbers that the user enters. Make sure you specify how the user is to enter the information. You need to ask questions about the dimensions. Display the matrix and its transpose. Make this and all future programs user-friendly. Some high-level programming languages contain commands that will directly read, write, or manipulate a matrix for you.

Do not use any of these commands. However, you will need to use arrays. Include comments in your code to tell your teacher (and yourself later) what you did.

Chapter 2

Addition of Matrices

If the Cardinals won 7 games in the first half of the regular season and won 8 in the second half, how many games did they win during the regular season? You know that the answer is 15 because $7+8=15$. The Eagles lost 8 games in the first half and lost 6 in the second half of the season. How many games did the Eagles lose all season? They lost 14 games. We know how to answer these questions using real numbers because we have represented our data by real numbers, and addition, subtraction, and multiplication are all defined and well-known operations for real numbers. However, how would we add when our information is represented by matrices? Let the matrix A represent the statistics from the first half of the season, and let the matrix B represent the statistics from the second half of the season.

$$\begin{array}{c} \\ C \\ E \\ F \\ O \end{array}\begin{array}{c} \begin{array}{ccc} W & L & T \end{array} \\ \begin{bmatrix} 7 & 6 & 1 \\ 5 & 8 & 1 \\ 2 & 12 & 0 \\ 9 & 5 & 0 \end{bmatrix} \end{array} = A \qquad \begin{array}{c} \\ C \\ E \\ F \\ O \end{array}\begin{array}{c} \begin{array}{ccc} W & L & T \end{array} \\ \begin{bmatrix} 8 & 6 & 1 \\ 9 & 6 & 0 \\ 5 & 9 & 1 \\ 11 & 4 & 0 \end{bmatrix} \end{array} = B$$

Look carefully at how you answered the questions above. Then look at where those numbers appear in the matrices. How would you add $A+B$?

Take time to think before reading further!

Definition 2.1 Matrices of the same dimensions are **added** by adding corresponding elements.

For instance, a_{ij} corresponds to b_{ij} because they both lie in the i^{th} row and j^{th} column of their respective matrices. Therefore, we would add, $a_{ij} + b_{ij}$ to obtain the $(i, j)^{th}$ element of $A + B$.

$$A + B = \begin{bmatrix} 7 & 6 & 1 \\ 5 & 8 & 1 \\ 2 & 12 & 0 \\ 9 & 5 & 0 \end{bmatrix} + \begin{bmatrix} 8 & 6 & 1 \\ 9 & 6 & 0 \\ 5 & 9 & 1 \\ 11 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 7+8 & 6+6 & 1+1 \\ 5+9 & 8+6 & 1+0 \\ 2+5 & 12+9 & 0+1 \\ 9+11 & 5+4 & 0+0 \end{bmatrix} = \begin{bmatrix} 15 & 12 & 2 \\ 14 & 14 & 1 \\ 7 & 21 & 1 \\ 20 & 9 & 0 \end{bmatrix}$$

Think about the similarities between addition and subtraction. How do you think matrices are subtracted?

Definition 2.2 Matrices of the same dimensions are **subtracted** by subtracting corresponding elements.

Suppose Y represents the wins, losses, and ties for these teams for the entire season (regular season and the playoffs together). Consider the following data

$$\begin{array}{c} \\ C \\ E \\ F \\ O \end{array} \begin{array}{c} \begin{array}{ccc} W & L & T \end{array} \\ \begin{bmatrix} 17 & 13 & 2 \\ 15 & 15 & 1 \\ 7 & 21 & 1 \\ 23 & 9 & 0 \end{bmatrix} \end{array} = Y.$$

How would you find the number of wins, losses, and ties for the playoffs? We would subtract the number of wins, losses, and ties for the regular season from the number of wins, losses, and ties for the entire season.

$$Y-(A+B) = \begin{bmatrix} 17 & 13 & 2 \\ 15 & 15 & 1 \\ 7 & 21 & 1 \\ 23 & 9 & 0 \end{bmatrix} - \begin{bmatrix} 15 & 12 & 2 \\ 14 & 14 & 1 \\ 7 & 21 & 1 \\ 20 & 9 & 0 \end{bmatrix} = \begin{bmatrix} 17-15 & 13-12 & 2-2 \\ 15-14 & 15-14 & 1-1 \\ 7-7 & 21-21 & 1-1 \\ 23-20 & 9-9 & 0-0 \end{bmatrix}$$

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

Remark 5 Remember, to add (or subtract) matrices, add (or subtract) corresponding elements.

The addition property of zero for real numbers tells us that $r+0=0+r=r$. There is also an addition property of zero for matrices which states that $A+0=0+A=A$ where 0 represents the zero matrix of the same dimensions as A.

Definition 2.3 A **zero matrix** is a matrix which has the number 0 for each of its elements.

Remark 6 We say "a" zero matrix instead of "the" zero matrix because for different pairs of dimensions, we have different zero matrices. However, for a given pair of dimensions, the zero matrix is unique because zero is unique in the real number system. It is usual to merely say the zero matrix and not refer to its dimensions when no confusion can arise.

Questions

1. So far, we have worked with specific examples. In general, does $A+B=B+A$?

2. Is $(A+B)^T = A^T + B^T$?

Answers

1. Yes. Let's look at the two general matrices of dimensions 2 by 3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}.$$

$$\begin{aligned} A+B &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix} \\ & \text{and} \\ B+A &= \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \\ &= \begin{bmatrix} b_{11}+a_{11} & b_{12}+a_{12} & b_{13}+a_{13} \\ b_{21}+a_{21} & b_{22}+a_{22} & b_{23}+a_{23} \end{bmatrix}. \end{aligned}$$

Looking at these general matrices should indicate to us that $A+B = B+A$ because the commutative law of addition for real numbers tells us that $a_{ij}+b_{ij} = b_{ij}+a_{ij}$ for any i and any j. Therefore, $A+B = B+A$ is true when the operations are defined (ie., when the matrices have the same dimensions.) We proved this for 2 by 3 matrices and reasonsed that it would be true for matrices of other dimensions. We can prove that $A+B = B+A$ in general by looking at the general $(i,j)^{th}$ element of each side of the equation. The $(i,j)^{th}$ element of $A+B$ is $a_{ij}+b_{ij}$ and the $(i,j)^{th}$ element of $B+A$ is $b_{ij}+a_{ij}$. Therefore, using the commutative law of addition for real numbers, $A+B = B+A$.

2. Yes, $(A+B)^T = A^T + B^T$. Let us look at a generic element from each side of the equation. First, let's look at the left side of the equation. A generic element of $A+B$ would be $a_{ij} + b_{ij}$. When this matrix is transposed, the generic element is $a_{ji} + b_{ji}$. Therefore, a generic element of the left side of the equation is $a_{ji} + b_{ji}$, which is exactly the same as a generic element of the right side of the equation.

Problems

1. Using the following matrices, perform the operation indicated when it is defined and state that the operation is not defined for the particular matices when that is the case:

$$A = \begin{bmatrix} 7 & 4 & 8 & 6 \\ 9 & 3 & 0 & 2 \\ 1 & 5 & 6 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 8 & 7 & 4 & 0 \\ 9 & 6 & 2 & 5 \\ 1 & 4 & 7 & 2 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 9 & 0 & 2 \\ 7 & 4 & 6 & 5 \\ 3 & 8 & 7 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 7 & 6 & 3 \\ 9 & 5 & 1 \\ 0 & 2 & 4 \\ 7 & 6 & 1 \end{bmatrix} \quad E = \begin{bmatrix} 0 & 5 & 2 \\ 9 & 8 & 1 \\ 6 & 4 & 3 \\ 0 & 7 & 5 \end{bmatrix} \quad F = \begin{bmatrix} 9 & 2 & 1 \\ 0 & 4 & 3 \\ 7 & 6 & 5 \\ 4 & 1 & 8 \end{bmatrix}$$

(a) $A + C$ (b) $D + E$ (c) $F - D$ (d) $F + B$

(e) $B - (A + C)$ (f) $D - (E + F)$ (g) $B + C - B$ (h) $A - D$

(i) $A + D^T$ (j) $D + E - B^T$

2. What matrix would need to be added to A to produce the 3 by 5 zero matrix if

$$A = \begin{bmatrix} 2 & 0 & -8 & 7 & -9 \\ \frac{1}{2} & 5 & -6 & 4 & 1 \\ -2 & 10 & 3 & 13 & -7 \end{bmatrix}?$$

3. Matrix A represents the number of wins and losses for these teams in one year and B represents the number of wins and losses for the next year.

$$\begin{array}{l} \\ \text{Lions} \\ \text{Tigers} \\ \text{Bears} \end{array} \begin{array}{c} \text{W} \quad \text{L} \\ \begin{bmatrix} 5 & 8 \\ 9 & 4 \\ 7 & 6 \end{bmatrix} \end{array} = A \qquad \begin{array}{l} \\ \text{Lions} \\ \text{Tigers} \\ \text{Bears} \end{array} \begin{array}{c} \text{W} \quad \text{L} \\ \begin{bmatrix} 7 & 5 \\ 6 & 6 \\ 4 & 8 \end{bmatrix} \end{array} = B$$

(a) What are the teams' records for the two years combined?

(b) Write a sentence about what row 3 tells us.

(c) If the three season record for these teams is represented by C, how many games did each team win and lose in the third year?

$$\begin{array}{l} \\ \text{Lions} \\ \text{Tigers} \\ \text{Bears} \end{array} \begin{array}{c} \text{W} \quad \text{L} \\ \begin{bmatrix} 20 & 19 \\ 22 & 17 \\ 16 & 23 \end{bmatrix} \end{array} = C$$

4. Matrix A represents the points scored from three kinds of shots made by each team during the first period of a basketball game, B represents the same information from the second period, C represents the same information from the third period, and E represents the total number of points scored from each of the three kinds of shots in the game by each team. The column for free throws is labeled by FT, field goals by FG, and three-point shots by T.

(a) How many points of each kind were scored by each team in the fourth period? (There are 4 periods in a basketball game).

$$\begin{array}{l} \\ \text{Home} \\ \text{Visitor} \end{array} \begin{array}{c} \text{FT} \quad \text{FG} \quad \text{T} \\ \begin{bmatrix} 5 & 12 & 3 \\ 3 & 10 & 0 \end{bmatrix} \end{array} = A \qquad \begin{array}{l} \\ \text{Home} \\ \text{Visitor} \end{array} \begin{array}{c} \text{FT} \quad \text{FG} \quad \text{T} \\ \begin{bmatrix} 7 & 16 & 0 \\ 4 & 18 & 3 \end{bmatrix} \end{array} = B$$

$$\begin{array}{l} \\ \text{Home} \\ \text{Visitor} \end{array} \begin{array}{c} \begin{array}{ccc} \text{FT} & \text{FG} & \text{T} \end{array} \\ \begin{bmatrix} 3 & 12 & 3 \\ 3 & 16 & 3 \end{bmatrix} \end{array} = C \qquad \begin{array}{l} \\ \text{Home} \\ \text{Visitor} \end{array} \begin{array}{c} \begin{array}{ccc} \text{FT} & \text{FG} & \text{T} \end{array} \\ \begin{bmatrix} 21 & 58 & 6 \\ 15 & 60 & 9 \end{bmatrix} \end{array} = E$$

(b) Which team won the game? What was the final score for each team?

5. If A is a square matrix, is $A + A^T$ always symmetric? Explain.

6. If matrices A and B are symmetric and have the same dimensions, is $A - B$ symmetric? Explain.

Computer Program

Write a program that will add and subtract matrices. You may build this on the program that you wrote in Chapter 1. The user should enter the matrices and indicate whether the matrices are to be added or subtracted. You should be able to add or subtract as many matrices as you wish without having to re-enter the previous solution. For example, can your program handle $A+B-C+D$? Remember that your program should be user-friendly and should have comments in the code. Again, write this program without using any commands that directly read, write, or manipulate matrices.

Chapter 3

Multiplication of Matrices

We have three recipes for breakfast foods. Each recipe feeds three people. The ingredients are as follows:

Pancakes: 2 cups baking mix, 2 eggs, and 1 cup milk.

Biscuits: $2\frac{1}{4}$ cups baking mix and $\frac{3}{4}$ cups milk.

Waffles: 2 cups baking mix, 1 egg, $1\frac{1}{3}$ cups milk, and 2 tablespoons vegetable oil.

Let's write this in the form of a labeled matrix so that it is easier to read.

$$\begin{array}{c} \\ \text{P} \\ \text{B} \\ \text{W} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Bm} & \text{E} & \text{M} & \text{O} \end{array} \\ \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \end{array} = R$$

If we want to feed 6 people instead of 3, what do we need to do? We double each recipe. That means we need twice as much of each ingredient, so we multiply every element of the matrix by the number 2.

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

When we multiply a matrix by a real number, we call the real number a **scalar** and call the operation **scalar multiplication**. Scalar multiplication consists of multiplying each element of a matrix by a given scalar. We use the terms scalar and scalar multiplication because, in abstract algebra, we often have the need to consider more general scalars than real numbers. However, in this book, we restrict our attention to scalars that are real numbers.

Definition 3.1 If c is a real number and A is a matrix whose $(i,j)^{th}$ element is a_{ij}, then the **scalar product** cA is the matrix whose $(i,j)^{th}$ element is ca_{ij}.

For example, if $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$, then $cA = \begin{bmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{bmatrix}$.

The result is a matrix of the same dimensions as the original matrix. Notice that scalar multiplication is consistent with what you know about real numbers. For example, you learned that $x + x = 2x$. It is also true that

$$R + R = \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} + \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 & 2 & 0 \\ 4\frac{1}{2} & 0 & 1\frac{1}{2} & 0 \\ 4 & 2 & 2\frac{2}{3} & 4 \end{bmatrix} = 2R.$$

Now we know how much of each ingredient we need to serve pancakes, biscuits, and waffles to 6 people. (Remember that each recipe serves 3 people).

If we want to feed 3 people pancakes, 12 people biscuits, and 9 people waffles, how much baking mix will we need? We need to make one batch of pancakes, 4 batches of biscuits, and 3 batches of waffles. Let's represent this with the row vector

$$\begin{matrix} & \text{P} & \text{B} & \text{W} \\ s = & [\,1 & 4 & 3\,] \end{matrix}.$$

We could have written this as a column vector instead of a row vector, but a row vector will be useful in later problems. The vector describing how much baking mix we need is the first column of R. We will call it a. Therefore,

$$\begin{matrix} & \text{Bm} \\ \text{P} & 2 \\ \text{B} & 2\frac{1}{4} \\ \text{W} & 2 \end{matrix} = a.$$

We need $1 * 2 + 4 * 2\frac{1}{4} + 2 * 3 = 17$ cups of baking mix. The process by which we found the number of cups of baking mix needed is called finding the **inner product** of two vectors.

Definition 3.2 The **inner product** of n-vectors x and y, denoted by $\langle x, y \rangle$, is $x_1y_1 + x_2y_2 + \ldots + x_ny_n$ where n is the dimension of the vectors.

Notice that the definition of inner product requires the vectors to have the same dimension. The inner product of two vectors is a scalar. Therefore, $\langle s, a \rangle = 17$.

Remark 7 Some people refer to the inner product as the **dot product** and denote it $x \cdot y$.

What would we do if we wanted to know how much of each ingredient we need for 1 batch of pancakes, 4 batches of biscuits, and 3 batches of waffles? We would take the inner product of s and a particular column of R to find out how much of that particular ingredient we need. This procedure motivates our definition of matrix multiplication which will be described in detail later in this chapter.

Now let's multiply our row vector, s, by our recipe matrix, R.

$$s = \begin{array}{c} \begin{array}{ccc} \text{P} & \text{B} & \text{W} \end{array} \\ \begin{bmatrix} 1 & 4 & 3 \end{bmatrix} \end{array} \qquad \begin{array}{c} \\ \text{P} \\ \text{B} \\ \text{W} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Bm} & \text{E} & \text{M} & \text{O} \end{array} \\ \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \end{array} = R$$

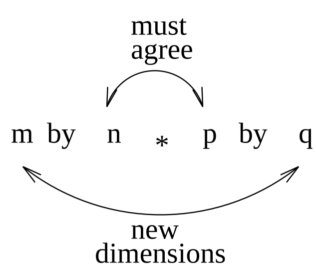
$$s * R = \begin{bmatrix} 1 & 4 & 3 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix}$$

$$
\begin{array}{lcccccccc}
\langle \text{column 1 of s, R} \rangle & = & 1*2 & + & 4*2\frac{1}{4} & + & 3*2 & = & 17 \\
\langle \text{column 2 of s, R} \rangle & = & 1*2 & + & 4*0 & + & 3*1 & = & 5 \\
\langle \text{column 3 of s, R} \rangle & = & 1*1 & + & 4*\frac{3}{4} & + & 3*1\frac{1}{3} & = & 8 \\
\langle \text{column 4 of s, R} \rangle & = & 1*0 & + & 4*0 & + & 3*2 & = & 6
\end{array}
$$

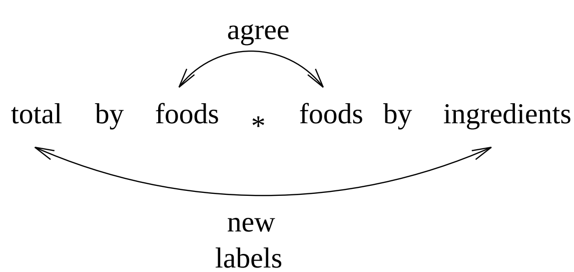
$$
s*R \;=\; \begin{array}{c} \begin{array}{cccc} \text{Bm} & \text{E} & \text{M} & \text{O} \end{array} \\ \begin{bmatrix} 17 & 5 & 8 & 6 \end{bmatrix} \end{array}
$$

We need 17 cups of baking mix, 5 eggs, 8 cups of milk, and 6 tablespoons of oil.

There are several interesting things to notice about matrix multiplication. We multiplied a 1 by 3 matrix by a 3 by 4 matrix and got a 1 by 4 matrix. This pattern will always hold when we multiply. The middle numbers must be the same (like the threes were in this case), when we multiply matrices. The resulting matrix will always have the dimensions of the outside numbers (1 by 4 in this case) when multiplication is defined. The following picture expresses the requirements on the dimensions:

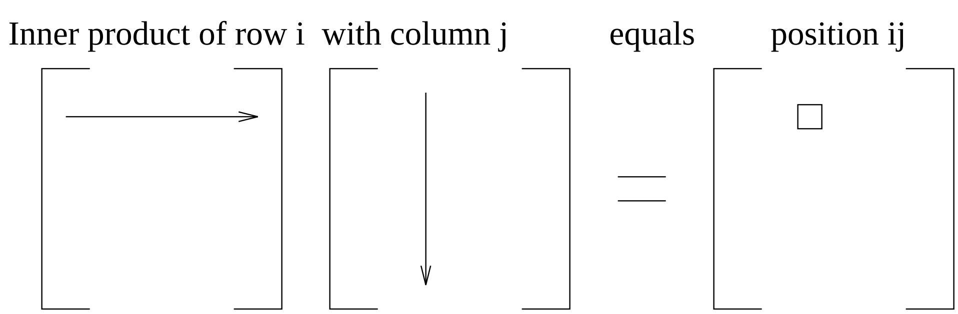


Even though the labels are not a formal part of the matrix, and are not always attached to a matrix, this also happens with the labels. The labels of the inside dimensions must agree if we want a meaningful product.



The label, total by ingredients, is meaningful because foods was the label for the inside dimensions of both matrices that we multiplied.

Let's also look closely at how we multiply the matrices because we will multiply matrices with larger dimensions later. This is a hands on activity. Take your left pointer finger and place it at the beginning of the first row of the first matrix (the only row we have in this case). Take your right pointer finger and place it on the first number of the first column of the second matrix. Multiply the two numbers to which you are pointing. Each time you move, your left hand will go across the row, and your right hand will go down the column. When you reach the end of the row and column, add the numbers you have obtained from the multiplications. This number goes in the first row and first column of your product matrix. This is the same as taking the inner product of the first row of the first matrix and the first column of the second matrix. Now you can move to the first row, second column doing the same thing. This number will go in the first row, second column of your product matrix. In short, position ij of your product matrix consists of the inner product of the i^{th} row of your first matrix and the j^{th} column of the second matrix. This is a lot easier to do than it is to describe! Your left hand will move across and your right hand will move down. Do this for every row and column combination to get your product matrix. No, you are not too old to do this. A lot of college students multiply matrices this way. After you do this enough times, your hands will not let you do it incorrectly ever again. This picture depicts the motions necessary to find a product:



Definition 3.3 Consider the m by p matrix A and the p by n matrix B. By the **matrix product** A times B, we mean the m by n matrix whose $(i,j)^{th}$ element is the inner product of the i^{th} row of A with the j^{th} column of B.

Since many of us watch our money closely, let's look at another example. How much does it cost to make each of these foods? First, we need to know how much each ingredient costs. We can find that information when we go to the grocery store. Baking mix costs 17 cents per cup; eggs are 8 cents each; milk costs 13 cents per cup; and oil is 4 cents per tablespoon.

Look at the dimensions of our matrices and the labels we have put on them. **THINK** about what dimensions and labels should be on our product matrix. This will tell you how we should organize the data about the cost of each food. **Write down your product matrix before you read further**. **Also write down each step of the multiplication and addition that you do to find the product matrix**. Compare that to the matrix products that follow. Look carefully at where each number that you used appears in the matrix.

Let's call our cost matrix C. Remember that since the dimensions of C are 4 by 1, C could also be considered to be a column vector of dimension 4.

$$\begin{array}{c} \\ \text{P} \\ \text{B} \\ \text{W} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Bm} & \text{E} & \text{M} & \text{O} \end{array} \\ \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \end{array} = R \qquad \begin{array}{c} \text{Bm} \\ \text{E} \\ \text{M} \\ \text{O} \end{array} \begin{array}{c} \text{Cents} \\ \begin{bmatrix} 17 \\ 8 \\ 13 \\ 4 \end{bmatrix} \end{array} = C$$

Now let's multiply to find out how much the ingredients of each recipe cost.

$$R * C = \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \begin{bmatrix} 17 \\ 8 \\ 13 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 2*17 & + & 2*8 & + & 1*13 & + & 0*4 \\ 2\frac{1}{4}*17 & + & 0*8 & + & \frac{3}{4}*13 & + & 0*4 \\ 2*17 & + & 1*8 & + & 1\frac{1}{3}*13 & + & 2*4 \end{bmatrix}$$

$$= \begin{matrix} \\ \text{P} \\ \text{B} \\ \text{W} \end{matrix} \begin{matrix} \text{Cents} \\ \begin{bmatrix} 63 \\ 48 \\ 67\frac{1}{3} \end{bmatrix} \end{matrix}$$

This gives us the cost of the ingredients needed for each food to feed three people. Biscuits are cheaper than pancakes which are slightly cheaper than waffles.

Remember to use your fingers the way we discussed earlier to remember which numbers to multiply. Let's look at the dimensions and labels for this example. The dimensions of matrix R are 3 by 4 and the dimensions of C are 4 by 1, so $R * C$ is 3 by 1. The labels also tell us that we set up the product correctly. We have food by ingredients multiplied by ingredients by cents to get food by cents. This is what we want!

What if we also want to know the calorie content of each recipe? If we know the calorie content of each ingredient, we can find the number of calories in each recipe. There are 510 calories per cup of baking mix, 70 calories in each egg, 90 calories in a cup of milk, and 120 calories in a tablespoon of oil. Look at the labels for matrix R and for the product matrix to decide how to organize this information. We will put this in a 4 by 1 matrix (also called a column vector of dimension 4) and name it K. To find out how many calories are in each recipe, we multiply $R * K$.

$$\begin{array}{c} \\ P \\ B \\ W \end{array} \begin{array}{c} \begin{array}{cccc} \text{Bm} & \text{E} & \text{M} & \text{O} \end{array} \\ \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \end{array} = R \qquad \begin{array}{c} \\ \text{Bm} \\ \text{E} \\ \text{M} \\ \text{O} \end{array} \begin{array}{c} \text{Calories} \\ \begin{bmatrix} 510 \\ 70 \\ 90 \\ 120 \end{bmatrix} \end{array} = K$$

$$\begin{aligned} R * K &= \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \begin{bmatrix} 510 \\ 70 \\ 90 \\ 120 \end{bmatrix} \\ &= \begin{bmatrix} 2 * 510 & + & 2 * 70 & + & 1 * 90 & + & 0 * 120 \\ 2\frac{1}{4} * 510 & + & 0 * 70 & + & \frac{3}{4} * 90 & + & 0 * 120 \\ 2 * 510 & + & 1 * 70 & + & 1\frac{1}{3} * 90 & + & 2 * 120 \end{bmatrix} \\ &= \begin{array}{c} \\ P \\ B \\ W \end{array} \begin{array}{c} \text{Calories} \\ \begin{bmatrix} 1250 \\ 1215 \\ 1450 \end{bmatrix} \end{array} \end{aligned}$$

Now we know that there are 1250 calories in a recipe of pancakes, 1215 calories in a recipe of biscuits, and 1450 calories in a recipe of waffles for three people. Therefore, biscuits have the fewest calories, pancakes are in the middle, and waffles have the most.

What would we do if we want to know both the cost of each food and the number of calories? We could use matrix multiplication twice like we did above, but we also have the ability to set up only one matrix multiplication to find both pieces of information. We can append K to C to form a single matrix. Let's call it F.

$$\begin{array}{c} \\ P \\ B \\ W \end{array} \begin{array}{c} \begin{array}{cccc} \text{Bm} & \text{E} & \text{M} & \text{O} \end{array} \\ \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \end{array} = R \qquad \begin{array}{c} \\ \text{Bm} \\ \text{E} \\ \text{M} \\ \text{O} \end{array} \begin{array}{c} \begin{array}{cc} \text{Cents} & \text{Calories} \end{array} \\ \begin{bmatrix} 17 & 510 \\ 8 & 70 \\ 13 & 90 \\ 4 & 120 \end{bmatrix} \end{array} = F$$

Now let's multiply $R * F$ to find the information we want.

$$R * F = \begin{bmatrix} 2 & 2 & 1 & 0 \\ 2\frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 2 & 1 & 1\frac{1}{3} & 2 \end{bmatrix} \begin{bmatrix} 17 & 510 \\ 8 & 70 \\ 13 & 90 \\ 4 & 120 \end{bmatrix}$$

$$= \begin{array}{c} \\ P \\ B \\ W \end{array} \begin{array}{c} \begin{array}{cc} \text{Cents} & \text{Calories} \end{array} \\ \begin{bmatrix} 63 & 1250 \\ 48 & 1215 \\ 67\frac{1}{3} & 1450 \end{bmatrix} \end{array}$$

Notice that with only one matrix multiplication, we are able to find the same products that we found in the previous two matrix multiplications. The power to combine information is one of the assets of matrix multiplication. Although the same number of operations are needed whether we use one matrix multiplication or two, it is easier to keep track of all of our information when we use one matrix multiplication.

Would we have been able to multiply $F * R$? No, the dimensions are wrong because matrix multiplication is defined only if the "inside dimensions" agree. What happens when you try to multiply these matrices? You run out of numbers in the row and column at different times. This should alert you to the fact that something is wrong.

Remark 8 A shortened way of writing $R * F$ is RF. When there is no sign between two matrices or two sets of parentheses, it is implied that you should multiply.

We learned in the last chapter that there is a matrix version of the addition property of zero. There is also a matrix version of the multiplication property of one. The real number version tells us that if a is a real number, then $a * 1 = 1 * a = a$. The matrix version of this property states that if A is a square matrix, then $A * I = I * A = A$, where I is the **identity matrix** of the same dimensions as A. If A is not square, then $A * I = A$ and $I * A = A$ where I in each case is the identity matrix with dimensions such that the multiplication would be a defined operation.

Definition 3.4 An **identity matrix** is a square matrix with ones along the main diagonal and zeros elsewhere.

The symbol I is used to represent an identity matrix when its dimensions are not necessary and when the dimensions can be determined from the context. The symbol I_n represents the identity matrix of dimension n *by* n. The matrix $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Notice that if A is m by p, then $A * I_p = A$ and $I_m * A = A$. The identity matrix gets its name because I is the multiplicative identity for matrices just as the number 1 is the multiplicative identity for real numbers.

Questions

This is a good place to use your calculator if it handles matrices. Do enough examples of each to convince yourself of your answer to each question. If your calculator does not handle matrices, or if you want a more mathematical argument, use generic matrices and carry out these operations like we did in the addition section.

Answer these questions on your own before you read beyond this paragraph. Remember to consider the dimensions of the matrices.

1. Consider $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$. Does $AB = BA$ for all B for which matrix multiplication is defined?

2. In general, does $AB = BA$?

3. Does $A(BC) = (AB)C$?

4. Does $A(B + C) = AB + AC$?

5. Does $(AB)^T = B^T A^T$?

6. Does $A - B = -(B - A)$?

7. For real numbers, if $ab = 0$, we know that either a or b must be zero. Is it true that $AB = 0$ implies that A or B is a zero matrix?

8. Are $A^T A$ and AA^T always symmetric?

Answers

1. Yes. Let's consider the generic 2 by 2 matrix $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$. Let's look at the left side of the equation; $AB = \begin{bmatrix} ab_{11} & ab_{12} \\ ab_{21} & ab_{22} \end{bmatrix}$. Now let's look at the right side of the equation; $BA = \begin{bmatrix} b_{11}a & b_{12}a \\ b_{21}a & b_{22}a \end{bmatrix}$. Since a and each element of B are scalars, the order of multiplication does not matter. Therefore, if $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, then $AB = BA$ for any B for which matrix multiplication is defined.

For the general case, $A = aI$, let's look at the left side of the equation. The product $AB = aIB = aB$ for any matrix B for which matrix multiplication is defined. Therefore, a general element of this matrix is ab_{ij}. Now, let's look at the right side of the equation. We need to remember that $a = a^T$ because a is a scalar. The product $BA = BaI = Ba = (a^T B^T)^T = (aB^T)^T$. Each element of the matrix aB^T is ab_{ji}, so the transpose of this matrix has elements ab_{ij}. Therefore, the two sides are equal.

2. No. Unlike multiplication with real numbers, $AB \neq BA$ in general. There are occasions when $AB = BA$, but these occasions are very rare. In fact, the only time that $AB = BA$ for every B is when A is a scalar multiple of the identity matrix. **It is very important to remember that AB is NOT, in general, equal to BA.**

3. If the dimensions are correct for multiplication, $A(BC) = (AB)C$. We call this the associative property of matrices. An example with correct dimensions is matrix A is 4 by 3, matrix B is 3 by 2, and matrix C is 2 by 5. This product results in a 4 by 5 matrix. The associative property of matrices becomes quite useful when you want to reduce the number of multiplications performed. Refer to problem 5 for an example. The number of multiplications performed becomes very important when you are dealing with large matrices.

4. Yes, $A(B + C) = AB + AC$. This means that the distributive property holds for matrices.

5. Yes, $(AB)^T = B^T A^T$. If matrix A is 4 by 3 and matrix B is 3 by 5, AB is 4 by 5, so $(AB)^T$ is 5 by 4. Just by looking at dimensions, we can tell that $(AB)^T \neq A^T B^T$, because the dimensions of $A^T B^T$ tell us that this multiplication cannot

be performed. The dimensions of B^TA^T are correct for matrix multiplication and give a resulting matrix that is 5 by 4. This is not a proof that $(AB)^T = B^TA^T$ is true, but it is a good indication. Work several examples to convince yourself that $(AB)^T = B^TA^T$.

6. Yes, $A - B = -(B - A)$ if the dimensions of A and B are the same so that subtraction is defined. This is true because $-(B-A) = -1(B-A) = -B+A = A - B$.

7. No. For example, if $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix}$, then $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, but neither A nor B is a zero matrix. **Remember that** $AB = 0$ **does NOT imply that either** $A = 0$ **or** $B = 0$.

8. Yes, A^TA and AA^T are always symmetric. Remember that $(AB)^T = B^TA^T$ and that a matrix is symmetric if it is equal to its transpose. Let's look at the transpose of A^TA; $(A^TA)^T = A^T(A^T)^T = A^TA$. Therefore, A^TA is symmetric. The same procedure proves that AA^T is symmetric; $(AA^T)^T = (A^T)^TA^T = AA^T$.

Problems

1. The matrix below expresses the approximate distance, in miles, between any of the following two cities: Houston, Los Angeles, New York, and Washington DC.

	H	LA	NY	DC
H	0	1540	1610	1370
LA	1540	0	2790	2650
NY	1610	2790	0	240
DC	1370	2650	240	0

(a) What special kind of matrix is this (other than square and 4 by 4)?

(b) If we want to know the same information in kilometers, what should we do? Remember, for our purposes here, one mile is equal to 1.6 kilometers.

(c) What is the resulting matrix when you perform the operation that you suggested in part (b)?

2. Perform the operations requested below if they are possible using these matrices.

$$A = \begin{bmatrix} 5 & 9 & 2 \\ 1 & 7 & 6 \\ 3 & 4 & 8 \end{bmatrix} \qquad B = \begin{bmatrix} 9 & 1 & 6 \\ 7 & 2 & 4 \\ 8 & 10 & 3 \end{bmatrix} \qquad C = \begin{bmatrix} 5 & 3 & 6 \end{bmatrix} \qquad D = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

(a) $4C$ (b) AD (c) DA (d) BC (e) $3CB$

(f) $C(A+B)$ (g) AB (h) BA (i) CAD (j) DBC

(k) $AD + (CB)^T$ (l) DC (m) CD

3. The matrix G represents the average score for each student on tests, quizzes, and homework. Tests are 50% of the grade, quizzes are 30% and homework is 20%.

$$\begin{array}{c} \\ \text{Amy} \\ \text{Bill} \\ \text{Chou} \\ \text{David} \\ \text{Erica} \end{array} \begin{array}{c} \begin{array}{ccc} \text{T} & \text{Q} & \text{H} \end{array} \\ \begin{bmatrix} 78 & 80 & 75 \\ 76 & 90 & 95 \\ 72 & 70 & 85 \\ 60 & 70 & 80 \\ 84 & 80 & 90 \end{bmatrix} \end{array} = G$$

(a) Write the vector P expressing the percentages that would be used to find the final grade for each student.

(b) Would GP or PG produce a matrix of the final grades?

(c) Using matrices, determine the final grade for each student. Please show your work.

4. Does $c\,(Ax) = A\,(cx)$ where c is a scalar, A is a 2 by 2 matrix, and x is a dimension 2 column vector? Explain your answer.

5. Place the parentheses where needed to minimize the number of multiplications performed to work this problem. How many simple multiplications did it take to find T?

$$T = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} d & e & f \end{bmatrix} \begin{bmatrix} g \\ h \\ k \end{bmatrix} \begin{bmatrix} l & p & q \end{bmatrix}$$

6. Does $A = \begin{bmatrix} 4 & 3 \\ 2 & 5 \end{bmatrix}$ satisfy the equation $3A^2 - 2A = \begin{bmatrix} 58 & 75 \\ 50 & 83 \end{bmatrix}$? Explain why A does or does not satisfy the equation. Note: For real numbers, a multiplied by itself n times can be written as a^n. Similarly, the matrix A multiplied by itself n times can be written as A^n. Therefore, A^2 means AA.

Computer Program

Make changes and additions to your program from Chapter 2 so that it can also multiply matrices. A warning message should be displayed if the matrices are not of correct dimensions for the operation requested. Your program should be able to handle $(AB - C + D)E$ and other similar problems. Remember that your program should be user-friendly and should have comments in the code. Again, write this program without using any commands that directly read, write, or manipulate matrices.

Chapter 4

Equations

Solving equations is an important part of mathematics. If we are working with more than one unknown at a time, we need to solve systems of equations. You may already know how to solve a system of linear equations, but matrices provide a more compact way to arrive at the solution. Matrices are also easier to manipulate on a computer or calculator. Both of these facts will become more important when you work with larger systems.

Let's look at a system of linear equations. The system

$$\begin{aligned} 5x_1 + 3x_2 &= 93 \\ -4x_1 - 2x_2 &= -66 \end{aligned}$$

can be written in matrix form as $AX = B$ where

$$A = \begin{bmatrix} 5 & 3 \\ -4 & -2 \end{bmatrix}, \; X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \text{ and B} = \begin{bmatrix} 93 \\ -66 \end{bmatrix}.$$

However, you will usually see $Ax = b$ rather than $AX = B$ because most authors use small letters to represent vectors. You can multiply this out to convince yourself that $AX = B$ does represent this system.

When you learned to solve systems of linear equations, you learned that

(a) you arrive at the same solution no matter which equation you write first;

(b) the solution doesn't change if you multiply an equation by a scalar other than zero; and

(c) you can replace an equation with the sum of that equation and another equation without changing the solution.

These may not be exactly the words you used when you were solving a system of linear equations, but you did all these things. **Experiment with the system above to convince yourself that these statements are true.**

We can also solve this system entirely in matrix form. We use the same rules, and we call them **elementary row operations (EROs)**. The EROs tell us that we can

(a) interchange any two rows;

(b) multiply any row by a non-zero scalar; and

(c) replace any row by the sum of that row and any other row.

Proper use of EROs will leave us with a system that has the same solution as our original system, but is much easier to solve. If you were presented the system

$$\begin{aligned} x_1 &= a \\ x_2 &= b \end{aligned}$$

you would be able to "solve" it instantly because you only have to read off the solution. If this system was written using matrix notation, it would look like this:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is the 2 by 2 identity matrix. Because you can just read off the solution when a system is in this form, our first goal is to transform our system into this form.

Let's solve the system above using matrices. We can represent this entire system with a 2 by 3 matrix which looks like this: $\left[\begin{array}{cc|c} 5 & 3 & 93 \\ -4 & -2 & -66 \end{array}\right]$. This is called an

augmented matrix because we combined 2 matrices (a matrix and a vector for this system). In this case, we combined the 2 by 2 coefficient matrix which is made of the coefficients for our unknowns and the 2 by 1 matrix from the right-hand side of the equations into one 2 by 3 matrix. In other words, we put A to the left of the bar and put b to the right of the bar. The application of an ERO to the augmented matrix does not change the solution set of the linear system that the augmented matrix represents because whatever you do to the left side of an equation, you also do to the right side. Therefore, we will arrive at the same solution whether we use augmented matrices or not, and augmented matrices are more compact to write. Using matrix notation, our goal is to transform our system into one that looks like the following:

$$\left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \middle| \begin{bmatrix} a \\ b \end{bmatrix}\right].$$

In other words, we want the identity matrix to the left of the bar and the solution to the right of the bar.

> **Remark 9** The bar is not a formal part of the matrix, so it is not necessary. It is placed there so that we can refer to the different parts of the augmented matrix and easily move back and forth between the augmented matrix and the linear system that it represents.

Let's use EROs to obtain a system of this form. It is a good idea to write notes to yourself about what you do in each step. This helps you locate and correct your mistake if you make one. It also helps you to explain your work. In this book, $r1$ represents row 1.

$$\left[\begin{array}{cc|c} 5 & 3 & 93 \\ -4 & -2 & -66 \end{array}\right] \quad \begin{array}{l} \text{Original} \\ \text{augmented matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 18.6 \\ -4 & -2 & -66 \end{array}\right] \quad r1 \div 5$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 18.6 \\ 0 & 0.4 & 8.4 \end{array}\right] \quad 4 * r1 + r2$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 18.6 \\ 0 & 1 & 21 \end{array}\right] \quad r2 \div 0.4$$

$$\left[\begin{array}{cc|c} 1 & 0 & 6 \\ 0 & 1 & 21 \end{array}\right] \quad -6 * r2 + r1$$

When we convert this from augmented matrix notation back to the algebraic notation for a system of equations, it looks like this:

$$\begin{aligned} 1x_1 + 0x_2 &= 6 \\ 0x_1 + 1x_2 &= 21 \end{aligned}$$

This tells us that $x_1 = 6$ and $x_2 = 21$. **Substitute this solution into the system to assure yourself that we are correct**. If we systematically use elementary row operations to obtain the identity matrix to the left of the bar, we call this the **Gauss-Jordan** elimination method.

Now, let's solve the system

$$\begin{aligned} 5x_1 + 3x_2 &= 70 \\ -4x_1 - 2x_2 &= -56 \end{aligned}$$

using Gauss-Jordan elimination.

$$\left[\begin{array}{cc|c} 5 & 3 & 70 \\ -4 & -2 & -56 \end{array}\right] \quad \begin{array}{l} \text{Original} \\ \text{augmented matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 14 \\ -4 & -2 & -56 \end{array}\right] \quad r1 \div 5$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 14 \\ 0 & 0.4 & 0 \end{array}\right] \quad 4 * r1 + r2$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 14 \\ 0 & 1 & 0 \end{array}\right] \quad r2 \div 0.4$$

$$\left[\begin{array}{cc|c} 1 & 0 & 14 \\ 0 & 1 & 0 \end{array}\right] \quad -0.6 * r2 + r1$$

Look back at the two systems of equations that we solved. How are they similar? We performed the same steps both times because the steps involved in solving a system of equations depend only on the matrix that is to the left of the bar. If we want to solve a system of equations with the same matrix A for different b vectors that we will be given at a later time, it would be nice if we did not have to do Gauss-Jordan elimination every time.

Let's look at the scalar version of this equation, $ax = b$, to help us find a general method for matrices. We know that $x = a^{-1}b$ if $a \neq 0$ because $a^{-1} = 1/a$ where a^{-1} is called the multiplicative inverse or the reciprocal. There is something analogous to this with matrices. It is also called the **inverse**. With scalars, $a^{-1}a = aa^{-1} = 1$.

Definition 4.1 The matrix A^{-1} (called A inverse) is the **inverse** of a square matrix$A$ if $A^{-1}A = AA^{-1} = I$ where I is the identity matrix.

Once we find A^{-1}, $Ax = b$ can be solved by matrix multiplication rather than Gauss-Jordan elimination. We follow the algebraic steps below to find an expression for x:

$$\begin{aligned} Ax &= b \\ A^{-1}Ax &= A^{-1}b \\ Ix &= A^{-1}b \end{aligned}$$

$$x = A^{-1}b$$

This means that if we find A^{-1}, we only need to multiply to solve systems with the same matrix A for different b vectors. Please remember that $A^{-1}b \neq bA^{-1}$, so you must multiply in the correct order.

> **Remark 10** If we have all the b vectors at the time when we wish to solve the system, we can simply augment all the b vectors together on the right side of the bar. Then the solution for each b vector will fall in the column that originally contained that b vector. For example, if we wished to solve $Ax = b$ and $Ax = c$ for the same A matrix, we could use the augmented matrix $\left[\begin{array}{cc|cc} a_1 & a_2 & b_1 & c_1 \\ a_3 & a_4 & b_2 & c_2 \end{array}\right]$. When the matrix to the left of the bar reaches the identity matrix by use of EROs, the solution to $Ax = b$ will be in the first column to the right of the bar, and the solution to $Ax = c$ will be in the second column to the right of the bar. Now you may wonder why we would ever need an inverse. If we do not have all the right-hand sides at the time when we solve the problem, we would find A^{-1} and multiply as indicated earlier. This situation often occurs when the solution to one system is the right-hand side of the next system.

Let's find A^{-1} for the same matrix that we have been using, $A = \begin{bmatrix} 5 & 3 \\ -4 & -2 \end{bmatrix}$. We can do this by solving the equation $AX = I$ for the n by n matrix X. Because we know that $AA^{-1} = I$, we know that our solution, X, is the same as A^{-1}.

$$\left[\begin{array}{cc|cc} 5 & 3 & 1 & 0 \\ -4 & -2 & 0 & 1 \end{array}\right] \quad \text{Original augmented matrix}$$

$$\left[\begin{array}{cc|cc} 1 & 0.6 & 0.2 & 0 \\ -4 & -2 & 0 & 1 \end{array}\right] \quad r1 \div 5$$

$$\left[\begin{array}{cc|cc} 1 & 0.6 & 0.2 & 0 \\ 0 & 0.4 & 0.8 & 1 \end{array}\right] \quad 4 * r1 + r2$$

$$\left[\begin{array}{cc|cc} 1 & 0.6 & 0.2 & 0 \\ 0 & 1 & 2 & 2.5 \end{array}\right] \quad r2 \div 0.4$$

$$\left[\begin{array}{cc|cc} 1 & 0 & -1 & -1.5 \\ 0 & 1 & 2 & 2.5 \end{array}\right] \quad -0.6 * r2 + r1$$

Notice that we used the exact same steps again. We now know that $A^{-1} = \begin{bmatrix} -1 & -1.5 \\ 2 & 2.5 \end{bmatrix}$.

Remark 11 In computational mathematics, the inverse is very seldom found because other methods exist that serve the same purpose and require fewer steps. However, the inverse will serve our needs at this level and is important in the theory of matrices.

Using the Gauss-Jordan elimination method, let's find A^{-1} where $A = \begin{bmatrix} 0 & 2 & 4 \\ 4 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$.

$$\left[\begin{array}{ccc|ccc} 0 & 2 & 4 & 1 & 0 & 0 \\ 4 & 2 & 3 & 0 & 1 & 0 \\ 1 & 3 & 6 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{l} \text{Original} \\ \text{augmented} \\ \text{matrix} \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 6 & 0 & 0 & 1 \\ 4 & 2 & 3 & 0 & 1 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \end{array}\right] \quad \begin{array}{l} \text{Switch } r1 \text{ and } r3 \text{ because we cannot} \\ \text{have a zero on the main diagonal, and} \\ \text{we would prefer 1 rather than 4.} \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 6 & 0 & 0 & 1 \\ 0 & -10 & -21 & 0 & 1 & -4 \\ 0 & 2 & 4 & 1 & 0 & 0 \end{array}\right] \quad -4 * r1 + r2$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 6 & 0 & 0 & 1 \\ 0 & 1 & 2.1 & 0 & -0.1 & 0.4 \\ 0 & 2 & 4 & 1 & 0 & 0 \end{array}\right] \quad r2 \div (-10)$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 6 & 0 & 0 & 1 \\ 0 & 1 & 2.1 & 0 & -0.1 & 0.4 \\ 0 & 0 & -0.2 & 1 & 0.2 & -0.8 \end{array}\right] \quad -2 * r2 + r3$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 6 & 0 & 0 & 1 \\ 0 & 1 & 2.1 & 0 & -0.1 & 0.4 \\ 0 & 0 & 1 & -5 & -1 & 4 \end{array}\right] \quad r3 \div (-0.2)$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 30 & 6 & -23 \\ 0 & 1 & 0 & 10.5 & 2 & -8 \\ 0 & 0 & 1 & -5 & -1 & 4 \end{array}\right] \quad \begin{array}{l} -6 * r3 + r1 \\ -2.1 * r3 + r2 \\ \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1.5 & 0 & 1 \\ 0 & 1 & 0 & 10.5 & 2 & -8 \\ 0 & 0 & 1 & -5 & -1 & 4 \end{array}\right] \quad -2 * r2 + r1$$

Therefore, $A^{-1} = \begin{bmatrix} -1.5 & 0 & 1 \\ 10.5 & 2 & -8 \\ -5 & -1 & 4 \end{bmatrix}$. **Multiply AA^{-1} and $A^{-1}A$ to convince yourself that they both multiply to I.**

Did you notice that there was a pattern to our elimination? **Look at the example with a 3 by 3 matrix to see if you can find the pattern.**

1. Begin with the first row. Let $i = 1$.

2. Check to see if the pivot for row i is zero. The **pivot** is the element of the main diagonal that is on the current row. For instance, if you are working with row

i, then the pivot element is a_{ii}. If the pivot is zero, exchange that row with a row below it that does not contain a zero in column i. If this is not possible, then an inverse to that matrix does not exist.

3. Divide every element of row i by the pivot.

4. For every row below row i, replace that row with the sum of that row and a multiple of row i so that each new element in column i below row i is zero.

5. Let $i = i + 1$. This means that you move to the next row and column. Repeat steps 2 through 5 until you have zeros for every element below the main diagonal. Now you have a matrix to the left of the bar that is called **upper triangular** because all the non-zero numbers fall in the triangle above and including the main diagonal.

6. Now we work to get zeros above the main diagonal. The index i should be equal to the number of rows.

7. For every row above row i, replace that row with the sum of that row and a multiple of row i so that each new element in column i above row i is zero. You will notice that the zeros below the main diagonal are still zeros.

8. Let $i = i - 1$. This means that you move to the left one column and up a row. Repeat steps 6-8 until you have zeros for every element above the main diagonal. Since the zeros below the main diagonal did not change, you now have a **diagonal matrix** to the left of the bar because all the non-zero elements lie on the main diagonal. Since all the elements along the diagonal of this diagonal matrix are the number one, this matrix is the identity matrix. Therefore, the matrix to the right of the bar is our solution.

Remark 12 Notice that we obtain all the zeros below the main diagonal before we work to get any zeros above the main diagonal. Other books tell you to obtain all the zeros needed for a column above and below the diagonal before you move to the next column. That method makes the problem easier to code on a computer, but the method that we used often requires fewer calculations.

WARNING: We know that a^{-1} is not defined when $a = 0$. It is also true that A^{-1} is not always defined. Is it possible to find a unique solution to the system if the matrix A does not have an inverse? No it is not. You will learn more about this in Chapter 6.

We know that we can use the Gauss-Jordan elimination method to solve a system of equations using matrices, but we don't really have to do all that work if we are only trying to solve a system of linear equations. It is true that it is easy to solve a system if the identity matrix is to the left of the bar because you can just read off the answer. However, it is also fairly easy if the matrix to the left of the bar is upper triangular because you can read the last element of the solution and substitute it into the previous equation to obtain another element. Repeated use of substitution will yield the entire solution. Therefore there is a method called **Gaussian** elimination that stops row operations after you have an upper triangular matrix to the left of the bar. At that point, you use **back-substitution** to find the remaining values of the solution. This is very similar to the way you learned to solve systems of equations algebraically. Once you find a solution, you substitute it in everywhere to decrease the size of your system. Let's go back to our original 2 by 2 matrix example in this section.

$$\left[\begin{array}{cc|c} 5 & 3 & 93 \\ -4 & -2 & -66 \end{array}\right] \qquad \begin{array}{l}\text{Original}\\ \text{augmented matrix}\end{array}$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 18.6 \\ -4 & -2 & -66 \end{array}\right] \quad r1 \div 5$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 18.6 \\ 0 & 0.4 & 8.4 \end{array}\right] \quad 4 * r1 + r2$$

$$\left[\begin{array}{cc|c} 1 & 0.6 & 18.6 \\ 0 & 1 & 21 \end{array}\right] \quad r2 \div 0.4$$

In Gaussian elimination, we can stop performing row operations now since we have an upper triangular matrix to the left of the bar. When we translate from the augmented matrix into a system of equations, we get

$$\begin{aligned} x_1 + 0.6x_2 &= 18.6 \\ x_2 &= 21 \end{aligned} .$$

We can read from the second equation that $x_2 = 21$. We substitute 21 for x_2 into the first equation to get $x_1 + 0.6(21) = 18.6$, so $x_1 = 6$. This is the same solution as before, and Gaussian elimination requires fewer operations than does Gauss-Jordan elimination. **Try this with the 3 by 3 matrix to see that you get the same solution.** You can see for a 3 by 3 or larger matrix that fewer steps are required. In fact, Gaussian elimination requires approximately $n^3/3$ steps and Gauss-Jordan elimination requires approximately $n^3/2$ steps. You can read more about this in the last chapter of this book.

4.1 Coding

Did you ever make up codes so that you could pass secret notes to your friends? See if you can figure out this coded phrase: 69 108 130 159 -50 -86 -96 -124 . Don't worry if you don't know it now; by the end of this chapter, you will be able to figure out the word. What sort of codes did you use? A very popular code is to give each letter of the alphabet a number.

A=1	J=10	S=19
B=2	K=11	T=20
C=3	L=12	U=21
D=4	M=13	V=22
E=5	N=14	W=23
F=6	O=15	X=24
G=7	P=16	Y=25
H=8	Q=17	Z=26
I=9	R=18	space=27

Unfortunately, this code is so well-known, that your message would not be very secretive. Some people choose to shift the code above so that A=10, B=11,. . ., R=1,. . ., Z=9 or something similar. However, since each letter is represented by a particular number and that number always stands for the same letter, this type of code is also easily broken. We need a code that is more difficult to break but is still easy to encode and decode. Let's look at one way to do this.

In order to send a secret message, you and your friend need to pick a matrix that has an inverse to be your secret coding matrix. For this example, let's use $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Therefore, $A^{-1} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}$.

Now we need to pick a message to send. Let's send the word "Smiles". We will derive our secret code by multiplying AB where B is our message. Since A has two columns, B must have two rows (in order for matrix multiplications to work). Therefore, B must be a 2 by 3 matrix

$$B = \begin{bmatrix} S & M & I \\ L & E & S \end{bmatrix}.$$

Notice that we chose to write our message across the rows. We need to let our friend know this when we choose the secret coding matrix because we could have just as easily written our letters down the columms. Since we want this message to be coded, we need to pick up numbers for each letter. We will stick with the standard A=1, B=2, etc. Therefore,

$$B = \begin{bmatrix} 19 & 13 & 9 \\ 12 & 5 & 19 \end{bmatrix}.$$

To code our message, we need to multiply AB. So

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 19 & 13 & 19 \\ 12 & 5 & 19 \end{bmatrix} = \begin{bmatrix} 43 & 23 & 47 \\ 105 & 59 & 103 \end{bmatrix}.$$

Since we want to add one more layer of secrecy to our code, we will write out the code in a line so that we don't give our "enemy" a clue that we used matrices to code and that our coding matrix had 2 rows. Now we can broadcast 43 23 47 105 59 103 in a public place if we want, and our message will be safe. When our friend receives our messages 43 23 47 105 59 103, she will want to decode it. She knows that our coding matrix was $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, and she knows our method of coding (writing across the rows, $AB =$ our code, and A=1, B=2, etc.) Since our coding matrix has two rows, our code, C, must also have twor rows. Therefore, she can convert our message back into the matrix

$$C = \begin{bmatrix} 43 & 23 & 47 \\ 105 & 59 & 103 \end{bmatrix}.$$

To solve $AB = C$, she multiplies on the left by A^{-1} to get $B = A^{-1}C$. Therefore,

$$B = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \begin{bmatrix} 43 & 23 & 47 \\ 105 & 59 & 103 \end{bmatrix} = \begin{bmatrix} 19 & 13 & 9 \\ 12 & 5 & 19 \end{bmatrix}.$$

She can convert this back into the matrix $\begin{bmatrix} S & M & I \\ L & E & S \end{bmatrix}$ to get SMILES. If we use the same coding matrix to code the work SMIRK, we get

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 19 & 13 & 9 \\ 18 & 11 & 27 \end{bmatrix} = \begin{bmatrix} 55 & 35 & 63 \\ 129 & 83 & 135 \end{bmatrix}.$$

If we look at all three words we have coded with this same coding matrix, we see

SMILES 43	23	47	105	59	103	
SMIRK	55	35	63	129	83	135
MILE	37	19	87	47		

Notice that although these words share many letters, their codes are not similar at all. Also notice that the number 47 represents the letter I in SMILES and the letter E in MILE. These are some of the features that make this sort of code so difficult to break. Even if our "enemy" knew that we we were coding using matrices, he would not know what size coding matrix we used or which numbers were in that coding matrix. Since the numbers in the coding matrix can be any real numbers (even negative numbers and fractions0, it would take a LONG time to guess the correct matrix even with the help of a very fast computer.

The matrix

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

was used to code the message 42 62 96 53 63 166 68 97 165. So to decode the message, we need to find A^{-1}. Since we computed it earlier in the chapter, we know that

$$A^{-1} = \begin{bmatrix} -1.5 & 0 & 1 \\ 10.5 & 2 & -8 \\ -5 & -1 & 4 \end{bmatrix}.$$

Because we coded using a 3 by 3 matrix, we need to break our message into 3 rows. Therefore, our message should be written as

$$B = \begin{bmatrix} 42 & 62 & 96 \\ 53 & 63 & 166 \\ 68 & 97 & 165 \end{bmatrix}.$$

When we multiply A^{-1}, we get

$$\begin{bmatrix} -1.5 & 0 & 1 \\ 10.5 & 2 & 8 \\ -5 & -1 & 4 \end{bmatrix} \begin{bmatrix} 42 & 62 & 96 \\ 53 & 63 & 166 \\ 68 & 97 & 165 \end{bmatrix} = \begin{bmatrix} 5 & 4 & 21 \\ 3 & 1 & 20 \\ 9 & 15 & 4 \end{bmatrix}$$

which translates into the word "education." Now can you decode the message from the beginning of this section? It was encoded using the same format as the other and we used the coding matrix

$$A = \begin{bmatrix} 5 & 3 \\ -4 & -2 \end{bmatrix}.$$

Coding is a fun way to use matrices and their inverses, but it also has important practical applications when governments andother organizations try to transmit private messages over public systems such as a radio or satellite.

Questions

1. There is a formula that can be used to find the inverse of a 2 by 2 matrix. Look at A and A^{-1} for several 2 by 2 matrices or perform Gauss-Jordan elimination on an augmented matrix formed from a general 2 by 2 matrix to find this formula.

2. Does a non-square matrix have an inverse?

3. An upper triangular matrix was described in the chapter. Describe and give an example of a **lower triangular** matrix.

4. Is $(A^{-1})^T = (A^T)^{-1}$?

Extension Questions

5. Is $x = A_L^{-1}b$ always a solution to $Ax = b$? You will probably need to read the answer to the second question before answering this one.

6. Is $x = A_L^{-1}b$ the only solution to $Ax = b$ when there is a solution?

Answers

1. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \quad \text{Original augmented matrix}$$

$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ c & d & 0 & 1 \end{array}\right] \quad r1 \div a$$

$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & \frac{ad-bc}{a} & \frac{-c}{a} & 1 \end{array}\right] \quad -c * r1 + r2$$

$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right] \quad r2 \div \frac{ad-bc}{a}$$

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right] \quad \frac{-b}{a} * r2 + r1$$

This can also be written as $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. This formula for the inverse of a 2 by 2 matrix is a good one to memorize.

2. No. Only square matrices can have an inverse such that $A^{-1}A = AA^{-1} = I$. However, there are one-sided inverses for some rectangular matrices, A, such

that $A_L^{-1}A = I$ or $AA_R^{-1} = I$. For example, if $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -2 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}$, then $AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $BA = \begin{bmatrix} -2 & -6 & 0 \\ 1 & 3 & 0 \\ 1 & 2 & 1 \end{bmatrix}$. This means that B is a right inverse for A. It is a "right" inverse because A multiplied by B on the right produces an identity matrix. It is "an" inverse rather than "the" inverse because it is not unique. The matrix

$$C = \begin{bmatrix} 2 & -5 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}$$

is also a right inverse of A. A left inverse can be represented by A_L^{-1} and a right inverse can be represented by A_R^{-1}.

> **Remark 13** If someone asks if matrix A has an inverse, he or she is referring to a matrix A^{-1} such that $A^{-1}A = AA^{-1} = I$. Therefore, unless the matrix A is square, just $AB = I$ or $BA = I$ is not sufficient proof that B is the inverse of A; it is only proof that B is at least a one-sided inverse of A. However, if the matrix is square and $BA = I$ or $AB = I$, then $B = A^{-1}$and $A^{-1}A = AA^{-1} = I$. You can find a proof of this in a college linear algebra text.

3. A lower triangular matrix has all the non-zero numbers on and below the main diagonal. All the numbers above the main diagonal are zero. An example is

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

Notice that there can be zeros on and/or below the main diagonal, but all the numbers above the main diagonal MUST be zero.

4. Yes, if the inverse exists, $(A^{-1})^T = \left(A^T\right)^{-1}$. Let's look at a 2 by 2 matrix for an example.

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

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad (A^T)^{-1} = \frac{1}{ad-cb}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

Since $cb = bc$, these are equal. This only proves the 2 by 2 case. The following is a proof for the general case:

$$\begin{aligned} AA^{-1} &= I \\ \left(AA^{-1}\right)^T &= I^T = I \\ \left(A^{-1}\right)^T A^T &= I \end{aligned}$$

The last step follows because $(AB)^T = B^T A^T$ as we showed in Chapter 3. This proof proves our point because if $AB = I$, then A is the inverse of B and B is the inverse of A. Therefore, $(A^{-1})^T$ is the inverse of A^T.

5. No. Try to substitute this answer into the original equation $Ax = b$. You get $AA_L^{-1}b = b$. This is only true for all b if A is square because $AA_L^{-1} \neq I$ if A is not square. However, $x = A_R^{-1}b$ is always a solution to $Ax = b$ because $AA_R^{-1} = I$.

6. Yes, if there is a solution to $Ax = b$, $x = A_L^{-1}b$ will be the only solution. Suppose that there were two solutions, x_1 and x_2, to $Ax = b$. Then the following must be true:

$$Ax_1 = b \qquad \text{and} \qquad Ax_2 = b$$

$$\begin{aligned} A_L^{-1}Ax_1 &= A_L^{-1}b \quad \text{and} \quad A_L^{-1}Ax_2 = A_L^{-1}b \\ x_1 &= A_L^{-1}b \quad \text{and} \quad x_2 = A_L^{-1}b \end{aligned}$$

Therefore, $x_1 = x_2$. This follows directly from our original assumption, but contradicts it. This means that our original assumption must be wrong. This is an example of proof by contradiction.

Problems

1. If you know that $Ax = b$ where A is a matrix and x and b are vectors, can you say for sure that $x = A^{-1}b$? Why or why not?

2. If A^{-1} exists, what is $(AA^{-1})(A^{-1}A)^T$?

3. Use elementary row operations on augmented matrices to solve these systems of equations for x. Use Gauss-Jordan elimination for (a) and (c). Use Gaussian elimination and back-substitution for (b) and (d).

   (a) $3x_1 + 5x_2 = 2$ and $2x_1 + 4x_2 = 1$

   (b) $2x_1 + 9x_2 = -3$ and $x_1 + 3x_2 = 6$

   (c) $Ax = b$ where $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 3 \\ 3 & 8 & 5 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$

   (d) $Ax = b$ where $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 3 & 3 \\ 3 & 3 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 3 \\ 12 \\ 6 \end{bmatrix}$

4. Find the inverses of these matrices:

   (a) $A = \begin{bmatrix} 5 & 3 \\ 5 & 4 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix}$

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

5. Use your solutions to problem 4 and the letter to number translation that sets A=1, B=2, etc. to decode these messages.

   (a) 126 60 148 65 using $A = \begin{bmatrix} 5 & 3 \\ 5 & 4 \end{bmatrix}$

   (b) 90 45 260 145 using $A = \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix}$

   (c) 1 18 40 73 53 96 using $A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix}$

   (d) Choose a matrix that has an inverse and encode your own message. Use the letter to number translation that sets A=1, B=2, etc. and tell which matrix you used to encode your message.

6. Use your solutions to problem 4 to solve these systems for x if $Ax = b$.

   (a) $A = \begin{bmatrix} 5 & 3 \\ 5 & 4 \end{bmatrix}$ $\quad b = \begin{bmatrix} 10 \\ 15 \end{bmatrix}$

   (b) $A = \begin{bmatrix} 5 & 3 \\ 5 & 4 \end{bmatrix}$ $\quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

   (c) $A = \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix}$ $\quad b = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$

(d) $A = \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

(e) $A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$

(f) $A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$

(g) $A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 3 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix}$

7. What is $(A^{-1})^{-1}$? Hint: Think about the definition of the inverse of a matrix.

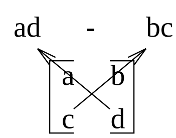
8. (a) What is the inverse of the matrix $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$ if $a, b, c \neq 0$?

   (b) Make a general statement about the inverse of a diagonal matrix.

Chapter 5

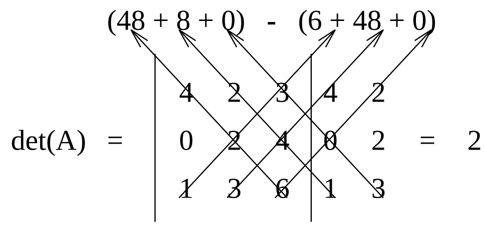
Determinants

The **determinant** of a square matrix is a real number that gives us valuable information about the matrix. Its definition is combersome, so it is in a special section at the end of this chapter. You will see some of the uses of the determinant in later chapters. For now, let's find out how to compute the determinant of a matrix so that we can use it later. The symbols $\det(A)$ and $|A|$ represent the determinant of A. In this case, the straight bars do NOT mean absolute value; they represent the determinant of the matrix. The determinant of a 1 by 1 matrix is simply the element of the matrix. If A is the 2 by 2 matrix, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) = ad - bc$ is found this way:



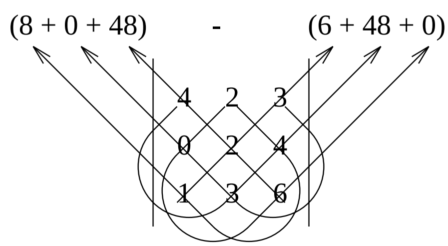
You may remember $ad - bc$ from the last chapter. If $\det(A) \neq 0$, then the inverse of the 2 by 2 matrix, A, is $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, which can also be written as $A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

We have already found the determinant for a 2 by 2 matrix, so let's look at the 3 by 3 matrix $A = \begin{bmatrix} 0 & 2 & 4 \\ 4 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$. In the 2 by 2 case, we subtracted products of the diagonals from each other beginning with the main diagonal. If we do that with the 3 by 3 case, we will be leaving out 4 of the 9 numbers. Let's rewrite the first two columns of A so that each number of the original matrix falls on a diagonal. It looks like this:



If we only use the diagonals that have 3 numbers on them, we will be using every number that was in our original matrix the same number of times. The diagonals towards the upper left give us $6 * 2 * 0 = 0$ and $1 * 3 * 2 = 6$ and $3 * 4 * 4 = 48$. If we add these together, we get 54. Now let's look at the diagonals towards the upper right. They give us $1 * 2 * 4 = 8$ and $3 * 4 * 0 = 0$ and $6 * 4 * 2 = 48$. If we add these together, we get 56. Let's subtract the upper right from the upper left like we did with the 2 by 2 case. We obtain $|A| = -2$.

The method of finding the determinant for the 3 by 3 matrix that we showed you is called "repeat the columns" for obvious reasons. Look at the numbers you used and where they came from. See if you can find a way to determine these same numbers without writing down the columns each time. It could look like this:



One of our favorite high school teachers likes to call this the "spaghetti method." It is also called the "basket weave method."

5.1 Expansion by Minors

Neither of the two previous methods will work for an n by n system if n is larger than three, so we will use another method called **expansion by minors**. Actually, this method works for any size square matrix, so let's use the same 3 by 3 example with the new method. First, we need to learn some new notation. The real number M_{ij} is the determinant of a submatrix of dimension $n-1$ by $n-1$ which contains everything except row i and column j of the original matrix. The number M_{ij} is called the **minor** for element ij of the matrix. For example, if $A = \begin{bmatrix} 4 & 2 & 3 \\ 0 & 2 & 4 \\ 1 & 3 & 6 \end{bmatrix}$ and we want to find M_{12}, we don't use row 1 or column 2 as shown below.

$$\begin{bmatrix} \not{4} & \not{2} & \not{3} \\ 0 & \not{2} & 4 \\ 1 & \not{3} & 6 \end{bmatrix}$$

Therefore, $M_{12} = \begin{vmatrix} 0 & 4 \\ 1 & 6 \end{vmatrix} = -4$. We will also need something to determine the sign. We set $s_{ij} = (-1)^{i+j}$ so that s_{ij} is always either positive one or negative one. For a 4 by 4 matrix, $S = \begin{bmatrix} +1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 \end{bmatrix}$, so s_{12} is negative one. The **cofactor** for a_{ij} is $C_{ij} = s_{ij}M_{ij}$. This means that $C_{12} = (-1)(-4) = 4$ for our example. Now we can write $\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$. This means that you take each element of a row or column and multiply it by its cofactor. When you add these terms together, you have the determinant of A. This is a lot easier to see than to explain, so let's find det(A) for $A = \begin{bmatrix} 4 & 2 & 3 \\ 0 & 2 & 4 \\ 1 & 3 & 6 \end{bmatrix}$. Choose any row or column to work with. We will use the first column for the example, but any of them will work. For each position in that column, we will have $a_{ij}s_{ij}M_{ij}$.

$$\begin{aligned} \det(A) &= a_{11}s_{11}M_{11} + a_{21}s_{21}M_{21} + a_{31}s_{31}M_{31} \\ &= 4(+1)\begin{vmatrix} 2 & 4 \\ 3 & 6 \end{vmatrix} + 0(-1)\begin{vmatrix} 2 & 3 \\ 3 & 6 \end{vmatrix} + 1(+1)\begin{vmatrix} 2 & 3 \\ 2 & 4 \end{vmatrix} \\ &= 4(12-12) - 0(12-9) + 1(8-6) \\ &= 4*0 - 0*3 + 1*2 \\ \det(A) &= 2 \end{aligned}$$

This gives us the same number for the determinant that we found before. Did you notice that by choosing the first column or the second row, we only had to find 2 minors because we know that 0 times anything is 0. Generally, choosing the row or column with the most zeros will save you a lot of work.

Now, let's find the determinant of the 4 by 4 matrix, $A = \begin{bmatrix} 5 & 4 & 6 & 3 \\ 0 & 2 & 1 & 0 \\ 9 & 7 & 4 & 6 \\ 2 & 8 & 1 & 3 \end{bmatrix}$. Let's expand along the second row since it contains two zeros.

$$\det(A) = 0(-1)\begin{vmatrix} 4 & 6 & 3 \\ 7 & 4 & 6 \\ 8 & 1 & 3 \end{vmatrix} + 2(+1)\begin{vmatrix} 5 & 6 & 3 \\ 9 & 4 & 6 \\ 2 & 1 & 3 \end{vmatrix} + 1(-1)\begin{vmatrix} 5 & 4 & 3 \\ 9 & 7 & 6 \\ 2 & 8 & 3 \end{vmatrix} + 0(+1)\begin{vmatrix} 5 & 4 & 6 \\ 9 & 7 & 4 \\ 2 & 8 & 1 \end{vmatrix}$$

$$= 0 + 2[(60+72+27) - (24+30+162)] - 1[(105+48+216) - (42+240+108)] + 0$$

$$= 2(-57) - (-21)$$

$$\det(A) = -93$$

Expansion by minors will allow you to find the determinant of a square matrix of any size. However, it requires a lot of operations when the matrix is large because each submatix used to determine a minor must be expanded. In fact, finding the determinant using this method requires $n!$ operations which is a very large number when n is large because $n! = n * (n-1) * (n-2) * \ldots * 2 * 1$.

5.2 Using Gaussian Elimination

The determinant can also be found using EROs in a manner similar to Gaussian elimination. In order to use EROs to find the determinant of a matrix, you must know a few facts about the determinant:

1. Interchanging two rows changes the sign of the determinant.
2. Multiplying a row by a scalar multiplies the determinant by that scalar.
3. Replacing any row by the sum of that row and any other row does NOT change the determinant.
4. The determinant of a triangular matrix (upper or lower triangular) is the product of the diagonal elements.

You can demonstrate these facts to yourself using a 2 by 2 matrix. Just as using EROs does not change the solution to a system, EROs combined with these rules will allow us to find the determinant of the original matrix. The use of EROs results in a system that is equivalent to the original system, so if we apply these rules to the determinant as we change the system, we will find the determinant to the original coefficient matrix. Let's use these rules to find the determinant of a 2 by 2 matrix. As we work this problem, we will let D_i represent the determinant of the current matrix yielded by EROs

$$A = \begin{bmatrix} 5 & 3 \\ -4 & -2 \end{bmatrix} \quad \text{Original matrix} \qquad D_1 = \det(A)$$

$$\begin{bmatrix} 1 & 0.6 \\ -4 & -2 \end{bmatrix} \quad r1 \div 5 \qquad D_2 = \det(A) \div 5$$

$$\begin{bmatrix} 1 & 0.6 \\ 0 & 0.4 \end{bmatrix} \quad 4 * r1 + r2 \qquad D_3 = \det(A) \div 5$$

$$D_3 = 0.4 \Rightarrow 0.4 = \det(A) \div 5 \Rightarrow \det(A) = 0.4 * 5 = 2$$

When we check this result using the formulas that we know, we get $\det(A) = 5(-2) - (-4)3 = 2$. Actually, we can solve for $\det(A)$ at any of the steps, but we work until we have a diagonal matrix because the determinant of a diagonal matrix is easy to find since it is simply the product of the diagonal elements.

Let's use this method to find the determinant of a 3 by 3 matrix.

$$A = \begin{bmatrix} 0 & 2 & 4 \\ 4 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix} \quad \text{Original matrix} \qquad D_1 = \det(A)$$

$$\begin{bmatrix} 1 & 3 & 6 \\ 4 & 2 & 3 \\ 0 & 2 & 4 \end{bmatrix} \quad \text{Switch } r1 \text{ and } r3 \qquad -D_2 = \det(A)$$

$$\begin{bmatrix} 1 & 3 & 6 \\ 0 & -10 & -21 \\ 0 & 2 & 4 \end{bmatrix} \quad -4 * r1 + r2 \qquad -D_3 = \det(A)$$

$$\begin{bmatrix} 1 & 3 & 6 \\ 0 & 1 & 2.1 \\ 0 & 2 & 4 \end{bmatrix} \quad r2 \div (-10) \qquad -D_4 = \det(A) \div (-10)$$

$$\begin{bmatrix} 1 & 3 & 6 \\ 0 & 1 & 2.1 \\ 0 & 0 & -0.2 \end{bmatrix} \quad -2*r2+r3$$

$$\begin{aligned} -D_5 &= \det(A) \div (-10) \\ D_5 &= -0.2 \\ \Rightarrow -(-0.2) &= \det(A) \div (-10) \\ \Rightarrow \det(A) &= (-10)*(0.2) = -2 \end{aligned}$$

We can verify this with either of the methods that we learned earlier to find the determinant of a matrix.

Using only these small examples might lead you to think that it is slower to find the determinant using elementary row operations. However, it actually requires fewer steps for larger problems than does expansion by minors. Therefore, this method is used more often to find the determinant in computational mathematics.

5.3 Inverses and Solutions to Systems

Determinants also provide another way to solve the system $Ax = b$. The method we are going to describe is called Cramer's rule. We need one more bit of notation. We will call the matrix A with the i^{th} column replaced by the vector b, B_i. Let's use the example that we worked with in Chapter 4. Matrix $A = \begin{bmatrix} 5 & 3 \\ -4 & -2 \end{bmatrix}$ and $b = \begin{bmatrix} 93 \\ -66 \end{bmatrix}$. Matrix $B_1 = \begin{bmatrix} 93 & 3 \\ -66 & -2 \end{bmatrix}$. Notice that the first column of A was replaced by the vector b. Replace the second column of A with b to get $B_2 = \begin{bmatrix} 5 & 93 \\ -4 & -66 \end{bmatrix}$. We need to find the determinants of each of these matrices. We find that $|A| = 2$, $|B_1| = 12$, and $|B_2| = 42$. The formula for Cramer's rule is $x_i = \frac{|B_i|}{|A|}$. Therefore,

$x_1 = \frac{12}{2} = 6$, and $x_2 = \frac{42}{2} = 21$. You should be happy to see that this is the same solution that Gauss-Jordan and Gaussian elimination gave us. Cramer's rule is essentially never used in computational mathematics because you are required to compute $n+1$ determinants, where n is the dimension of the square matrix, before you can find your solution for x. This requires a lot more work than Gaussian elimination, so Cramer's rule is usually used only to examine theoretical properties of matrices. You can read more about this in the last chapter of this book.

There is another way to find the inverse. We can use the cofactors and determinants that we used when we expanded by minors. If we place all the cofactors into a matrix and call it C, the formula for the inverse is $(A^{-1})_{ij} = \frac{C_{ji}}{\det(A)}$ or $A^{-1} = \frac{C^T}{\det(A)}$. Notice that in the first formula, i and j are reversed on the opposite sides of the equation. Transposing matrix C yields the same result in the second equation. Let's find the inverse of the matrix A that we used in Chapter 4, where $A = \begin{bmatrix} 0 & 2 & 4 \\ 4 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$. We have already found that $\det(A) = -2$, so let's find C^T. Since we know how to transpose a matrix, let's start with finding C. The element $c_{11} = (+1)\begin{vmatrix} 2 & 3 \\ 3 & 6 \end{vmatrix}$,

$c_{12} = (-1)\begin{vmatrix} 4 & 3 \\ 1 & 6 \end{vmatrix}$, $c_{13} = (+1)\begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix}$, $c_{21} = (-1)\begin{vmatrix} 2 & 4 \\ 3 & 6 \end{vmatrix}$, $c_{22} = (+1)\begin{vmatrix} 0 & 4 \\ 1 & 6 \end{vmatrix}$,

$c_{23} = (-1)\begin{vmatrix} 0 & 2 \\ 1 & 3 \end{vmatrix}$, $c_{31} = (+1)\begin{vmatrix} 2 & 4 \\ 2 & 3 \end{vmatrix}$, $c_{32} = (-1)\begin{vmatrix} 0 & 4 \\ 4 & 3 \end{vmatrix}$, $c_{33} = (+1)\begin{vmatrix} 0 & 2 \\ 4 & 2 \end{vmatrix}$.

Therefore, $C = \begin{bmatrix} 3 & -21 & 10 \\ 0 & -4 & 2 \\ -2 & 16 & -8 \end{bmatrix}$. That means that $C^T = \begin{bmatrix} 3 & 0 & -2 \\ -21 & -4 & 16 \\ 10 & 2 & -8 \end{bmatrix}$, so

$$A^{-1} = \begin{bmatrix} -1\frac{1}{2} & 0 & 1 \\ 10\frac{1}{2} & 2 & -8 \\ -5 & -1 & 4 \end{bmatrix}.$$ This method is not used often because it requires that you find n^2 determinants of dimension $n-1$ by $n-1$ and 1 determinant of dimension n by n. This means that this method would require approximately $\frac{n^5}{3}$ steps to compute the inverse if we computed the determinants with the best known algorithm. This is considerably more steps than are needed to compute the inverse using Gaussian elimination or Gauss-Jordan elimination. You can read more about this in Chapter 10.

5.4 Definition of the Determinant

We need some background knowledge before we can discuss the definition of the determinant. We want to form a product by choosing n elements where A is an n by n matrix. There will only be one element from each row and one element from each column in this product. For example, if one element of the product is a_{21}, then no other element in this product will be from row 2 or column 1. Let's look at a 3 by 3 matrix for an example.

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

We know that we will use an element from each column, so, for consistency, we will order the product this way: $a_{\_1}a_{\_2}a_{\_3}$. We can fill in the blanks with row numbers. If we choose to begin with a_{31}, then we can choose from rows 1 and 2 for the remaining positions. One possible product formed by these rules is $a_{31}a_{12}a_{23}$. Another possible product is $a_{11}a_{32}a_{23}$. There are 3!, or $3 * 2 * 1 = 6$, of these products. For our n by n

matrix, there are $n!$ possible products. All 6 possible products for this 3 by 3 matrix are: $a_{11}a_{22}a_{33}$, $a_{21}a_{32}a_{13}$, $a_{31}a_{12}a_{23}$, $a_{31}a_{22}a_{13}$, $a_{21}a_{12}a_{33}$, and $a_{11}a_{32}a_{23}$.

Now, we need to determine which sign ($+$ or $-$) should be attached to each product. To do this, you need to order the product with the column numbers increasing as we did above and look at the sequences of row numbers. For the product, $a_{11}a_{22}a_{33}$, we look at the row sequence (1,2,3). We are looking for **inversions**, or numbers that are out of order. Since 1 comes before 2, 1 comes before 3, and 2 comes before 3 in the sequence, there are no inversions in this sequence. In the sequence (2,3,1), which comes from the product $a_{21}a_{32}a_{13}$, there are two inversions because 2 is placed before 1 and 3 is placed before 1. There are also two inversions for the sequence from the product $a_{31}a_{12}a_{23}$. There are three inversions for (3,2,1) and one inversion each for (2,1,3) and (1,3,2). If the number of inversions is even, then the sign attached to the product is positive. If the number of inversions is odd, then the sign attached to the product is negative. Notice that we did not say that the product was positive or negative. We simply are determining whether the product will be added or subtracted.

Definition 5.1 The **determinant** of a square matrix is the sum of all the $n!$ possible signed products formed from the matrix using each row and each column only once for each product. The sign to be attached to the product is the same as the one determined by the formula $(-1)^N$ where N is the number of inversions as described above.

The determinant of the generic 3 by 3 matrix is: $a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$-$a_{31}a_{22}a_{13}$-$a_{21}a_{12}a_{33}$-$a_{11}a_{32}a_{23}$. For the matrix,

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

the determinant is $(0*2*6)+(4*3*4)+(1*2*3)-(1*2*4)-(4*2*6)-(0*3*3) = 0+48+6-8-48-0 = -2$ which is the same as we calculated at the beginning of the chapter.

Since this definition is cumbersome to follow, we generally do not compute the determinant by the definition, but it is good to know why the short cuts that we learned at the beginning of the chapter are valid. The determinant is also a good example of an abstract idea that has very important practical uses.

Questions

1. Can the determinant of a 2 by 2 matrix be found using expansion by minors?
2. Can you find the determinant of 5 by 5 or larger matrices? If so, how? If not, why not?

Answers

1. Yes, because the determinant of a 1 by 1 matrix is just the element of the matrix.
2. You can work larger examples, but you will have to expand the submatrices used to find the minors also because they will be larger than 3 by 3.

Problems

1. Find the determinants of these matrices. Show your work.

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

(b) $\begin{bmatrix} 5 & -2 \\ 3 & 6 \end{bmatrix}$

(c) $\begin{bmatrix} 3 & 0 & 4 \\ 6 & 2 & -1 \\ 5 & -7 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & -3 & 1 \\ 7 & 0 & 5 \\ -5 & 2 & 4 \end{bmatrix}$

(e) $\begin{bmatrix} 5 & 0 & 1 \\ 0 & 9 & 1 \\ 3 & -5 & 0 \end{bmatrix}$

(f) $\begin{bmatrix} 3 & -2 & 7 & 6 \\ -4 & 0 & 2 & 1 \\ 5 & 2 & 0 & -2 \\ 2 & 0 & -1 & 0 \end{bmatrix}$

(g) $\begin{bmatrix} 1 & 8 & -3 & 2 \\ 9 & 3 & 0 & 1 \\ -2 & 6 & 0 & -4 \\ 2 & -1 & 0 & 4 \end{bmatrix}$

(h) $\begin{bmatrix} 7 & 3 & 1 & 0 \\ -2 & -5 & 6 & 2 \\ 0 & 8 & -3 & 1 \\ 2 & 0 & 0 & -1 \end{bmatrix}$

(i) $\begin{bmatrix} 1 & 0 & 2 & 0 & 4 \\ 6 & 0 & 9 & 3 & 7 \\ 3 & 0 & 5 & 0 & 1 \\ 5 & 2 & 7 & 9 & 8 \\ 6 & 0 & 4 & 0 & 3 \end{bmatrix}$

2. Use Cramer's rule to solve these systems of equations:

(a)

$$\begin{aligned} 2x_1 + 3x_2 - 5x_3 &= -11 \\ -4x_1 - x_2 + 3x_3 &= 3 \\ 3x_1 - 2x_2 + x_3 &= 7 \end{aligned}$$

(b)

$$\begin{aligned} x_1 - 5x_2 + 7x_3 &= -10 \\ 9x_2 + 2x_3 &= 7 \\ x_1 + 3x_2 - x_3 &= 6 \end{aligned}$$

3. Find the inverse of matrix A using cofactors and determinants. Verify that you found the inverse by checking that I is the product matrix of AA^{-1} or $A^{-1}A$. Remember, since A is square, you do not have to check both because if AA^{-1}

or $A^{-1}A$ is the identity matrix, then so is the other.

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

4. Prove whether the following statements are true or false for 2 by 2 matrices. Remember that a counterexample establishes that a statement is false.

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

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

   (c) $\det(A + B) = \det(A) + \det(B)$

   (d) $\det(A^T) = \det(A)$

   Remark 14 In general, you may NOT assume that a statement is true for all matrices just because it is true for 2 by 2 matrices, but for the examples in this question, those that are true for 2 by 2 matrices are true for all matrices if the dimensions allow the operations to be performed.

5. Show that the determinant of an upper triangular matrix is the product of the diagonal entries.

Chapter 6

Consistent and Inconsistent Systems

When you solve a system of linear equations, what does your solution set (all of your solutions) describe geometrically? In each problem involving 2 by 2 matrices that we solved, our solution set was the point of intersection of the two lines represented by the equations in our system. In each problem with 3 by 3 matrices, the solution set was the point of intersection of 3 planes. However, a system of linear equations does not always have a point as the solution set.

If you solve the system

$$x_1 + 2x_2 = 4$$
$$2x_1 + 4x_2 = 8$$

algebraically, what do you get? You have an infinite number of solutions along the line $x_1 = 4 - 2x_2$ because any solution to the first equation, also solves the second equation. Therefore, you can choose any value for one of the variables, and you will be able to find a value for the other variable so that both equations are satisfied. This is called a **consistent** system because there is a solution. It is further categorized as an **underdetermined** system because there is not enough information to determine a unique solution.

Definition 6.1 A system is **consistent** if there is at least one solution.

Definition 6.2 A system is **underdetermined** when there are an infinite number of solutions.

For a linear system, if there are two or more solutions, then there are an infinite number of solutions. These solutions all lie on the same line. Geometrically, this system represents a line because both equations are representations of the same line. Try to solve this system with Gaussian elimination. What do you get? We get $\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & 0 & 0 \end{array}\right]$, because the second equation is a multiple of the first. The second equation requires that $0x_1 + 0x_2 = 0$ which is always true. Therefore, our second equation made no additional requirements beyond what the first equation requires. When Gaussian elimination on a system with a square coefficient matrix leaves you with zeros across an entire row of the augmented matrix and there are no rows with zeros to the left of the bar and a non-zero number to the right of the bar, you know that you have a consistent system that is underdetermined. We move the zero row or rows to the bottom of the matrix. When we try to get zeros above the main diagonal in Gauss-Jordan elimination, we do not try to get zeros in a column if the diagonal element of that column is zero. **Try this next system before reading further!**

$$\begin{aligned} 2x_1 + 4x_2 + 5x_3 &= 47 \\ 3x_1 + 10x_2 + 11x_3 &= 104 \\ 3x_1 + 2x_2 + 4x_3 &= 37 \end{aligned}$$

What is the solution to the system? Our work and solution are below.

$$\left[\begin{array}{ccc|c} 2 & 4 & 5 & 47 \\ 3 & 10 & 11 & 104 \\ 3 & 2 & 4 & 37 \end{array}\right] \begin{array}{l} \text{Original} \\ \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 2.5 & 23.5 \\ 3 & 10 & 11 & 104 \\ 3 & 2 & 4 & 37 \end{array}\right] \begin{array}{l} r1 \div 2 \\ \\ \\ \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 2.5 & 23.5 \\ 0 & 4 & 3.5 & 33.5 \\ 0 & -4 & -3.5 & -33.5 \end{array}\right] \begin{array}{l} \\ -3 * r1 + r2 \\ -3 * r1 + r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 2.5 & 23.5 \\ 0 & 1 & .875 & 8.375 \\ 0 & -4 & -3.5 & -33.5 \end{array}\right] \begin{array}{l} \\ r2 \div 4 \\ \\ \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 2.5 & 23.5 \\ 0 & 1 & .875 & 8.375 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ 4 * r2 + r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0.75 & 6.75 \\ 0 & 1 & 0.875 & 8.375 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} -2 * r2 + r1 \\ \\ \\ \end{array}$$

This tells us that $x_1 = 6.75 - .75x_3$ and $x_2 = 8.375 - .875x_3$. That means that we can choose whatever number we want for one element of x and get corresponding valid solutions for the other two. For instance, if we choose x_3 to be 1, then $x_1 = 6$ and $x_2 = 7.5$. Instead, we may choose $x_1 = 6.75$ then $x_2 = 8.375$ and $x_3 = 0$. There are an infinite number of solutions that we can find in this manner. Notice that this system is also consistent and underdetermined.

What do you get if you try to solve

$$x_1 + 2x_2 = 4$$
$$2x_1 + 4x_2 = 9$$

algebraically? The result is a requirement that you know cannot be satisfied. We get $0 = 1$ (you may arrive at a different contradictory requirement). This system is called **inconsistent**.

Definition 6.3 A system is **inconsistent** if it has no solutions.

If you graph these lines, you will see that they are parallel. Try to solve this system using Gaussian elimination. Our work follows:

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 2 & 4 & 9 \end{array}\right] \quad \begin{array}{l} \text{Original} \\ \text{Augmented Matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & 0 & 1 \end{array}\right] \quad \begin{array}{l} \\ -2 * r1 + r2 \end{array}$$

This requires that $0x_1 + 0x_2 = 1$. We know that this cannot be true. When using Gaussian elimination on a system, if you have zeros in an entire row to the left of the bar and do not have a zero to the right of the bar on that row, you know that you have an inconsistent system. Therefore, there is no point where the lines (or planes if you are in higher dimensions) intersect, so the system does not have a solution.

With underdetermined and inconsistent systems, you will never be able to get an identity matrix to the left of the bar of the augmented matrix. Therefore, we will not be able to find an inverse for the coefficient matrix. The only coefficient matrices that have inverses are those that have a unique point as the solution to the system, and the only coefficient matrices that have a unique point as the solution to the system are those that have inverses. These systems are **consistent** because they have a solution and **uniquely determined** because there is exactly one solution. This type of system was explored in Chapter 4.

Definition 6.4 A system is **uniquely determined** if there is exactly one solution to the system.

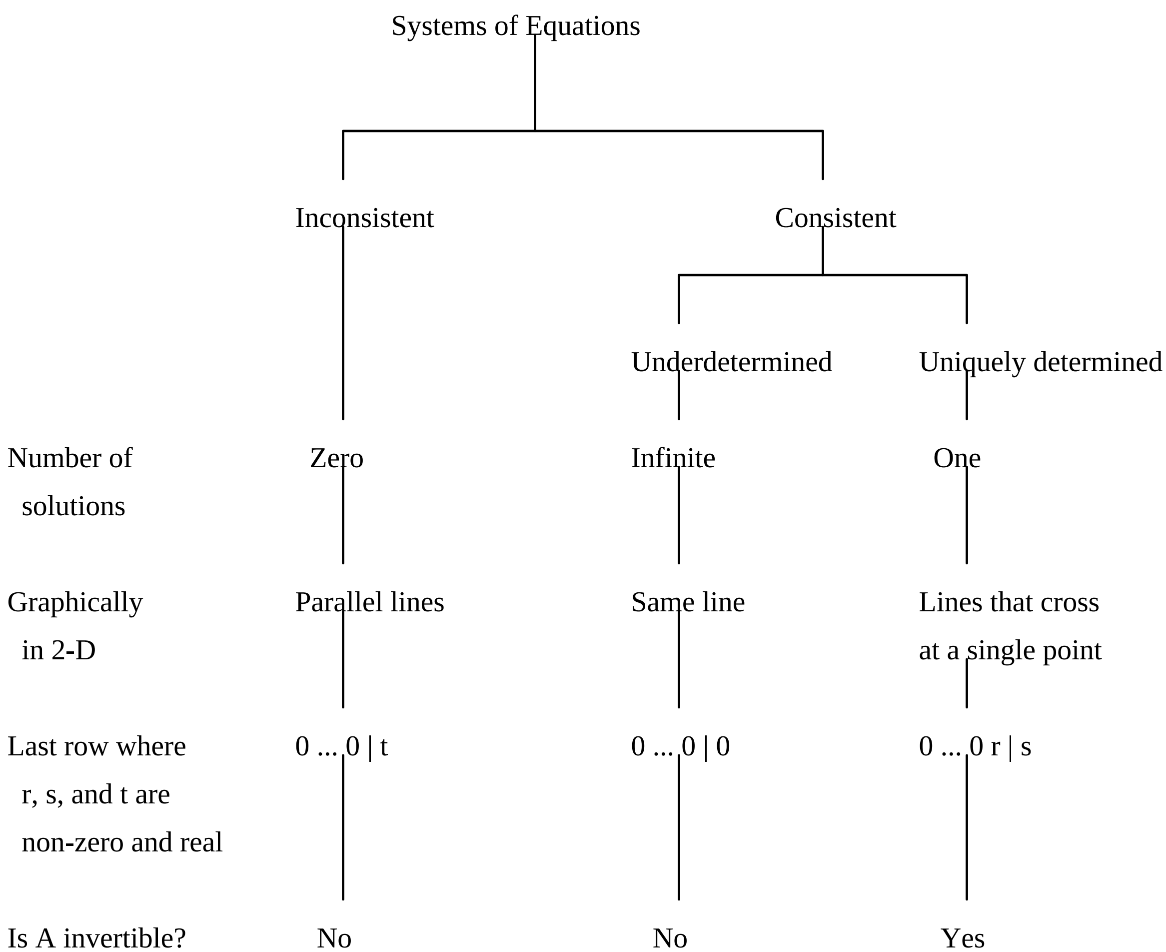
Remark 15 Many pre-calculus texts refer to underdetermined systems as **dependent** systems and to uniquely determined systems as **independent** systems.

For the examples above, find the determinant of A in each case. Can you draw a conclusion from this data that relates the determinant of A to whether or not the system has a unique solution? If not, solve some additional systems and try again to draw a conclusion.

Using determinants, we can check to see if a system is uniquely determined. If the determinant of the coefficient matrix is not equal to zero, the system is uniquely determined. Thinking of the systems that involve 2 by 2 matrices, can you tell why systems with determinants equal to zero are not uniquely determined? Remember that the formula for the inverse of a 2 by 2 matrix is $\frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Does $\frac{1}{\det(A)}$ have a meaning if $\det(A) = 0$? No, it does not. This is another illustration of the fact that the square matrix represented by A in $Ax = b$ for an inconsistent or underdetermined systems does not have an inverse. You will see that this holds true for all the problems that you do. **If** $\det(A) \neq 0$, **the system is uniquely determined and A is invertible. If** $\det(A) = 0$**, the system is either underdetermined or inconsistent; therefore, A is not invertible.** In practice, the determinant is not used to test systems in this manner because other methods require fewer steps. However, these relationships are very important to the theory of matrices.

> **Remark 16** In this section, occasionally, we refer to systems that have square coefficient matrices. This is only for simplicity. All of these classifications also apply to matrices with non-square coefficient matrices. However, we cannot take the determinant of a non-square matrix. Also, we cannot rely entirely upon the appearance of the last row of the system after Gaussian elimination eventhough that information is still valuable.

Here is a visual breakdown of the information that you have been given. This breakdown assumes a square coefficient matrix.



Note: A system is inconsistent if ANY row of the matrix has zeros left of the bar and a non-zero number right of the bar. Therefore, the matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 6 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

would be inconsistent eventhough the entire last row contains only zeros.

Problems

1. Classify these systems as either consistent or inconsistent. If the system is consistent, further categorize it as underdetermined or uniquely determined. Explain why the system fits into that category. Also, explain what this means graphically for each system.

   (a) $2x_1 + 3x_2 = 9$ and $3x_1 + 4\frac{1}{2}x_2 = 13$

   (b) $3x_1 + 4x_2 = 7$ and $9x_1 + 12x_2 = 21$

   (c) $2x_1 + 3x_2 = 8$ and $3x_1 + 4x_2 = 11$

   (d) $Ax = b$ where $A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 1 & 6 \\ 1 & -7 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 5 \\ 6 \\ -3 \end{bmatrix}$

   (e) $Ax = b$ where $A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 2 & -1 \\ 0 & -11 & 18 \end{bmatrix}$ and $b = \begin{bmatrix} 5 \\ 7 \\ 11 \end{bmatrix}$

   (f) $Ax = b$ where $A = \begin{bmatrix} 4 & 2 & 6 \\ 1 & \frac{1}{2} & 1\frac{1}{2} \\ 6 & 3 & 9 \end{bmatrix}$ and $b = \begin{bmatrix} -2 \\ -\frac{1}{2} \\ -3 \end{bmatrix}$

   (g) $Ax = b$ where $A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ 3 & 4\frac{1}{2} & 6 \end{bmatrix}$ and $b = \begin{bmatrix} 9 \\ 6 \\ 13 \end{bmatrix}$

2. Which of the matrices in problem 1 are invertible?

First Review

1. What are the dimensions of matrix A?

$$A = \begin{bmatrix} 3 & 2 & 5 \\ 7 & 1 & 0 \end{bmatrix}$$

2. Construct a generic 3 by 4 matrix, B, using the correct subscripted notation. (Hint: The element in the upper left corner is b_{11}.)

3. (a) Put the following information into a 2 by 3 matrix and label it.

   Keith scored 94 on his test and had a 99 homework average. Juan received 75 on his test and averaged 80 on homework. Yolanda's homework average was 90, but she scored 70 on the test.

   (b) Transpose the matrix in problem 3a and attach labels.

4. Consider the matrices

$$A = \begin{bmatrix} 4 & 3 & 9 \\ 0 & -1 & 5 \\ 10 & 2 & -8 \end{bmatrix} \text{ and } B = \begin{bmatrix} 9 & -2 & 6 \\ 8 & 0 & 1 \\ -3 & 12 & 7 \end{bmatrix}.$$

   (a) What is $A + B$?

   (b) What is $A^T + B$?

5. If matrix B is symmetric does $A + B = A + B^T$? Why or why not?

6. To raise money, our high school band decided to sell candy. They sold candy with nuts (N) and plain chocolate (P). Hoping to inspire people to sell candy, the band director held a contest among the grade levels to see which grade

would sell the most candy. The contest lasted for 3 weeks. Here are the results from the first two weeks. The numbers represent packages sold:

$$\begin{array}{c} \\ \text{Fresh.} \\ \text{Soph.} \\ \text{Jr.} \\ \text{Sr.} \end{array} \begin{array}{c} \begin{array}{cc} \text{N} & \text{P} \end{array} \\ \begin{bmatrix} 400 & 350 \\ 300 & 350 \\ 350 & 300 \\ 200 & 250 \end{bmatrix} \end{array} = W1 \qquad \begin{array}{c} \\ \text{Fresh.} \\ \text{Soph.} \\ \text{Jr.} \\ \text{Sr.} \end{array} \begin{array}{c} \begin{array}{cc} \text{N} & \text{P} \end{array} \\ \begin{bmatrix} 300 & 300 \\ 200 & 250 \\ 300 & 200 \\ 250 & 200 \end{bmatrix} \end{array} = W2$$

(a) How much of each kind of candy had each grade level sold by the end of the second week?

(b) Which grade level was leading the contest?

(c) By the end of the third week, the totals were:

$$\begin{array}{c} \\ \text{Fresh.} \\ \text{Soph.} \\ \text{Jr.} \\ \text{Sr.} \end{array} \begin{array}{c} \begin{array}{cc} \text{N} & \text{P} \end{array} \\ \begin{bmatrix} 1000 & 850 \\ 700 & 750 \\ 900 & 700 \\ 600 & 600 \end{bmatrix} \end{array} = T$$

How much of each kind of candy did each grade level sell during the third week?

(d) How many packages of plain chocolate were sold during the three-week period by the band?

(e) If the band makes 30 cents profit from each package of candy with nuts and makes 20 cents profit from each package of plain chocolate, how much profit did each grade level make? Answer this question using matrices.

(f) How much total profit did the band make from this venture? Please write your answer in dollars.

7. Consider the following matrices:

$$A = \begin{bmatrix} 4 & 7 & 8 \\ -3 & 1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 5 & -2 & 6 \\ 0 & 9 & 3 \end{bmatrix}$$

   (a) Is AB defined? If yes, what is it? If no, why not?

   (b) Using two of A, A^T, B, and B^T form a product that is a 2 by 2 matrix. For instance, AA^T is a 2 by 2 matrix. Find two more examples of this.

8. For

$$A = \begin{bmatrix} 2 & 5 & 7 & 8 \\ 0 & -8 & 3 & -1 \\ 9 & 1 & -2 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 5 & 7 & 10 \\ 6 & 3 & -1 \\ 0 & 4 & 2 \\ 1 & 8 & 9 \end{bmatrix}$$

   find AB.

9. Solve this system of equations using Gauss-Jordan elimination:

$$\begin{aligned} x_1 + 3x_2 - x_3 &= -2 \\ 2x_1 + 3x_2 + x_3 &= 2 \\ 3x_1 + 6x_2 + x_3 &= 3 \end{aligned}$$

10. Solve this system of equations using Gaussian elimination:

$$\begin{aligned} 4x_1 + 2x_2 - x_3 &= -8 \\ 3x_1 - x_2 + 2x_3 &= -3 \\ x_1 + 5x_3 &= 8 \end{aligned}$$

11. Find the inverse of this matrix: $\begin{bmatrix} 1 & 8 & 4 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \end{bmatrix}$

12. Does the inverse of a matrix always exist? Explain.

13. Find the determinant of these matrices. Show your work when possible.

(a) $\begin{bmatrix} 4 & 6 \\ -1 & -2 \end{bmatrix}$

(b) $\begin{bmatrix} 4 & 8 & 2 \\ -1 & 0 & 3 \\ 9 & 6 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 7 & 9 & 2 & -6 \\ -5 & 4 & 0 & 4 \\ 2 & 1 & 1 & 6 \\ 3 & 7 & 0 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 6 & 6 & 0 & 0 & 1 \\ -9 & -8 & 0 & 4 & 7 \\ 5 & 6 & 2 & 5 & 9 \\ 0 & 3 & 0 & 0 & 0 \\ 1 & 10 & 0 & 0 & 1 \end{bmatrix}$

14. Identify the following as consistent or inconsistent. If the system is consistent, further categorize it as underdetermined or uniquely determined. Explain.

(a) $Ax = b$ where $A = \begin{bmatrix} 2 & 1 & 3 \\ 7 & 6 & 11 \\ 3 & 4 & 5 \end{bmatrix}$ and $b = \begin{bmatrix} 4 \\ 10 \\ 2 \end{bmatrix}$

(b)

$$\begin{aligned} 2x_1 + 3x_2 + \ x_3 &= 10 \\ 4x_1 + 2x_2 \qquad &= 10 \end{aligned}$$

$$3x_1 + 2x_2 + 4x_3 = 20$$

(c)

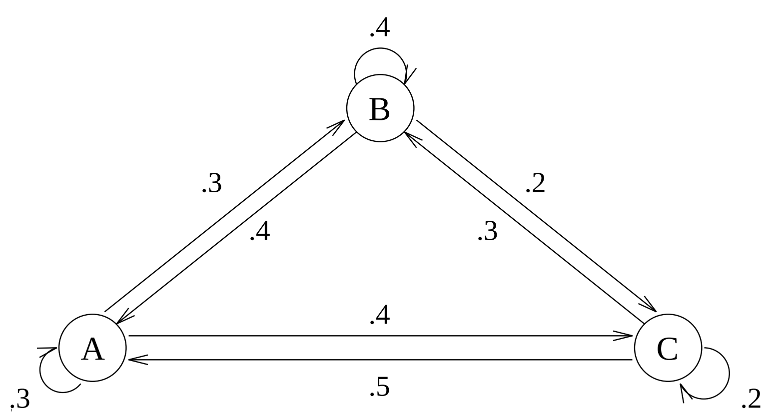
$$\begin{aligned} 7x_1 + 7x_2 + 11x_3 &= 10 \\ x_1 + 2x_2 + 3x_3 &= 4 \\ 4x_1 + x_2 + 2x_3 &= 3 \end{aligned}$$

(d) $Ax = b$ where $A = \begin{bmatrix} 5 & -1 & 3 \\ 2 & 1 & 4 \\ 3 & 5 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 13 \\ 8 \\ 7 \end{bmatrix}$

Chapter 7

Markov Chains

In College Town, Pizza Company gets a lot of business. They get so many calls each night that they have to operate three kitchens to fill all the orders. They decided to spread the kitchens out so that each one is near one of the housing sections of the university. Since the same people own all three branches of Pizza Company, they only hired one set of delivery drivers to serve all three kitchens. After a driver makes a delivery, he or she goes to the nearest kitchen to pick up the next order. Therefore, the location of a delivery person's next order depends only on his or her present location. The kitchens are logically named according to their area of campus. Of the calls to kitchen A, 30% are delivered in area A, 30% go to area B, and 40% go to area C. Of the orders placed at kitchen B, 40% go to area A, 40% go to area B, and 20% go to area C. Of the calls to kitchen C, 50% go to area A, 30% go to area B, and 20% go to area C. The picture below depicts the situation.



As you might guess, this information will be easier to read if we write it in matrix form. We will call this matrix S because it expresses the probability of movement (transition) from one **state** to the next. A **state** is the condition or location of an object in the system at a particular time. Therefore, our diagram is called a state diagram. Matrices of this type are called **transition** matrices. Our labeled transition matrix looks like this:

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} .3 & .3 & .4 \\ .4 & .4 & .2 \\ .5 & .3 & .2 \end{bmatrix} \end{array} = S$$

For each element, S_{ij}, i represents the starting location, and j represents the ending location for that move. This means that the row is the beginning location, and the column is the ending location after one move. We will want to learn things about what will happen in the future to Pizza Company, and this situation has the attributes necessary for what is called a Markov chain. Therefore, we will model the problem as a Markov chain in order to obtain information about the future.

A problem can be considered a (homogeneous) **Markov chain** if it has the following properties:

(a) For each time period, every object (person) in the system is in exactly one of the defined states. At the end of each time period, each object either moves to a new state or stays in that same state for another period.

(b) The objects move from one state to the next according to the transition probabilities which depend only on the current state (they do not take any previous history into account). The total probability of movement from a state (movement from a state to the same state does count as movement) must equal one.

(c) The transition probabilities do not change over time (the probability of going from state A to state B today is the same as it will be at any time in the future).

Remark 17 Requirement (c) is not a requirement of Markov chains in general. It is a requirement for a special kind of Markov chain called a homogeneous Markov chain. We will be studying only homogeneous Markov chains in this book, so we will use the term Markov chain to refer to a homogeneous Markov chain.

The transition matrix used to model the Markov chain will have the following properties:

(a) Each element of the transition matrix is a probability; therefore, each is a number between 0 and 1, inclusive.

(b) The elements of each row of the transition matrix sum to 1. This is due to property (b) of a Markov Chain.

(c) The transition matrix must be square because it has a row and a column for each state.

We will assume that it takes each delivery person the same amount of time to make one delivery. Therefore, after one delivery, of the cars that began in A, 30% will again be in A, 30% will be in B, and 40% will be in C. Since we only have three locations, and each delivery person must be somewhere after each delivery, the probability that a car is in one of those three locations must be one. This is why each row sums to 1. Because we are dealing with probabilities, each entry must be between 0 and 1, inclusive. The most important fact that lets us model this situation as a Markov chain is that the next location for delivery depends only on the current location, not previous history. It is also true that our matrix of probabilities does not change during the time we are observing.

Remark 18 Some assumptions are not completely accurate, but if we did not make assumptions to generalize the problem, we would not have the ability to approximate a solution to the problem. We just need to make sure that our assumptions are reasonable. The assumption that each delivery takes the same amount of time is reasonable if you consider that the average delivery times should be nearly equal.

Now that you know the background, we can begin the fun part. Do you know what the probability matrix would look like that describes where a car would be after exactly 2 deliveries? What about 3, 4, or 5 deliveries? Can you predict the probability matrix for the cars after a night of deliveries?

Well, lets start with a simpler question. If you begin at kitchen C, what is the probability that you will be in area B after 2 deliveries? Think about how you can get to B. We can go from C to C, then from C to B. We can go from C to B, then from B to B. We can go from C to A, then from A to B. We will let P(CB) represent the probability of going from C to B in one delivery. Let's write this probability in our shorthand notation. Do you remember how probabilities work? If two (or more) independent events must both (all) happen, we multiply their probabilities together. If there are two (or more) distinct events that would both (all) work, we add the probabilities of those events together.

Remark 19 Notice that we said we can multiply probabilities together if the events are independent. We know that our events are independent of one another because someone in area A is equally likely to order a pizza whether or not someone in area B or C ordered a pizza. If our events were not independent, we would not be able to simply multiply the probabilities.

This gives us P(CA)P(AB) + P(CB)P(BB) + P(CC)P(CB) for the probability that a delivery person goes from C to B in 2 deliveries. Let us substitute numbers into our formula. We get $(.5)(.3) + (.3)(.4) + (.2)(.3) = .33$. This tells us that if we begin at kitchen C, we have a 33% chance of being in kitchen B after 2 deliveries. Let us try this for another pair. If we begin at kitchen B, what is the probability of being at kitchen B after 2 deliveries? **Try this before you read further!** The probability of going from kitchen B to kitchen B in two deliveries is P(BA)P(AB) + P(BB)P(BB) + P(BC)P(CB) $= (.4)(.3) + (.4)(.4) + (.2)(.3) = .34$. Now it wasn't so bad calculating where you would be after 2 deliveries, but what if you need to know where you will be after 5 deliveries? That would take a LONG time. There must be an easier way, right? Look carefully at where these numbers come from. Do they fall into rows or columns? As you might suspect, they are the result of matrix multiplication. Going from C to B in 2 deliveries is the same as taking the inner product of row 3 and column 2. Going from B to B in 2 deliveries is the same as taking the inner product of row 2 and column 2. If you multiply S by S, you will get the same answers that you got for these two questions and the rest of the probability matrix after 2 deliveries.

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} .41 & .33 & .26 \\ .38 & .34 & .28 \\ .37 & .33 & .30 \end{bmatrix} \end{array} = S^2$$

You will notice that the elements on each row still add to 1 and each element is between 0 and 1, inclusive. Since we are modeling our problem with a Markov chain, this is essential. This matrix indicates the probabilities of going from kitchen i to kitchen j in exactly 2 deliveries.

Now that we have this matrix, it should be easier to find where we will be after 3 deliveries. We will let p(AB) represent the probability of going from A to B in 2 deliveries. Let's find the probability of going from C to B in 3 deliveries. The probability is p(CA)P(AB) + p(CB)P(BB) + p(CC)P(CB) $= (.37)(.3) + (.33)(.4) + (.3)(.3) = .333$. Where do these numbers come from? This probability is the inner product of row three of S^2 and column two of S. Therefore, if we multiply S^2 by S, we will get the probability matrix for 3 deliveries.

$$S^2S = SSS = S^3 = \begin{bmatrix} .385 & .333 & .282 \\ .390 & .334 & .276 \\ .393 & .333 & .274 \end{bmatrix}$$

By now, you probably know how we find the matrix of probabilities for 4, 5 or more deliveries. Notice that the elements on each row still add to 1. Therefore, it is important that you do not round your answers. Keep as many decimal places as possible to retain accuracy.

$$S^4 = \begin{bmatrix} .3897 & .3333 & .2770 \\ .3886 & .3334 & .2780 \\ .3881 & .3333 & .2786 \end{bmatrix}$$

$$S^5 = \begin{bmatrix} .38873 & .33333 & .27794 \\ .38894 & .33334 & .27772 \\ .38905 & .33333 & .27762 \end{bmatrix}$$

$$S^6 = \begin{bmatrix} .388921 & .333333 & .277746 \\ .388878 & .333334 & .277788 \\ .388857 & .333333 & .277810 \end{bmatrix}$$

$$S^7 = \begin{bmatrix} .3888825 & .3333333 & .2777842 \\ .3888910 & .3333334 & .2777756 \\ .3888953 & .3333333 & .2777714 \end{bmatrix}$$

What do you notice about these matrices as we take into account more and more deliveries? The numbers in each column seems to be converging to a particular number. **Think about what this tells us about our long term probabilities**. This tells us that after a large number of deliveries, it no longer matters which kitchen we were in when we started. At the end of the evening, we have a $38.\overline{8}\%$ chance of being at kitchen A, $33.\overline{3}\%$ chance of being at kitchen B, and $27.\overline{7}\%$ chance of being in kitchen C. This convergence will happen with most of the transition matrices that we consider.

Remark 20 If all the entries of the transition matrix are between 0 and 1 EXCLUSIVELY, then convergence is guaranteed to take place. Convergence may take place when 0 and 1 are in the transition matrix, but convergence is no longer guaranteed. For an example, look at the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Think about the situation that this matrix represents in order to understand why A^k oscillates as k grows.

Sometimes, you will be given a vector of initial distributions to describe how many or what part of the objects are in each state in the beginning. Using this vector, you can find out how many or what part of the objects are in each state at any later time. If the initial distribution vector is a vector of decimals, it tells what part of the total number of objects are in each state in the beginning. It contains only numbers between 0 and 1, inclusive, and the elements in the row sum to one. Alternatively, the vector of initial distributions could contain the actual number of objects or people in each state in the beginning. In this case, all the elements will be nonnegative and the elements in each row will add to the total number of objects or people in the entire system. For our example, the vector of initial distributions can tell you what part of the drivers originally begin in each area. If we start out with a uniform distribution, we will have $\frac{1}{3}$ of our delivery cars in each area. After one delivery, the distribution will be 40% of our deliveries in area A, $33.\overline{3}\%$ in area B, and $26.\overline{6}\%$ in area C. We find this by multiplying our initial distribution matrix by our transition matrix.

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

After the entire evening, we said that the fractions would converge to particular numbers so that the area from which we start doesn't matter. After many deliveries, we will obtain the same right-hand side no matter with which initial distribution we start. For example,

$$\begin{bmatrix} .\overline{3} & .\overline{3} & .\overline{3} \end{bmatrix} \begin{bmatrix} .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \end{bmatrix} = \begin{bmatrix} .3\overline{8} & .3\overline{3} & .2\overline{7} \end{bmatrix}.$$

Notice that the right-hand side is the same as one of the rows of our transition matrix after many deliveries. This is exactly what we expected because we said that $38.\overline{8}\%$ of the people will be in area A after many deliveries regardless of what percentage of the people were in area A in the initial distribution. Check this with several initial distributions to convince yourself of the truth of this statement.

If the initial distribution indicates the actual number of people in the system, our system can be represented by the following after one delivery:

$$\begin{bmatrix} 18 & 18 & 18 \end{bmatrix} \begin{bmatrix} .3 & .3 & .4 \\ .4 & .4 & .2 \\ .5 & .3 & .2 \end{bmatrix} = \begin{bmatrix} 21.6 & 18 & 14.4 \end{bmatrix}$$

Did you notice that we now have a fractional number of people in areas A and C after one delivery? We know that this cannot happen, but this gives us a good idea of approximately how many delivery people are in each area. After many deliveries, the right-hand side of this equality will also be very close to a particular vector. For example,

$$\begin{bmatrix} 18 & 18 & 18 \end{bmatrix} \begin{bmatrix} .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \end{bmatrix} = \begin{bmatrix} 21 & 18 & 15 \end{bmatrix}$$

The particular vector that the product converges to is the total number of people in the system (54 in this case) times any row of the matrix that A^k converges to as k grows,

$$54 \begin{bmatrix} .3\overline{8} & .3\overline{3} & .2\overline{7} \end{bmatrix} = \begin{bmatrix} 21 & 18 & 15 \end{bmatrix}.$$

Try some examples to convince yourself that the vector indicating the number of people in each area after many deliveries will not change if people are moved from one state to another in the initial distribution. Also notice that the number of people in the entire system never changes. People move from place to place, but the system never loses or gains people.

Remark 21 It is usual to associate the word vector with a column vector, so a row vector is a transposed vector. Therefore, we will write

$$\begin{bmatrix} 18 & 18 & 18 \end{bmatrix} \begin{bmatrix} .3 & .3 & .4 \\ .4 & .4 & .2 \\ .5 & .3 & .2 \end{bmatrix} = \begin{bmatrix} 21.6 & 18 & 14.4 \end{bmatrix} \text{ as } x^T A = b^T \text{ and}$$

$$\begin{bmatrix} 18 & 18 & 18 \end{bmatrix} \begin{bmatrix} .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \end{bmatrix} = \begin{bmatrix} 21 & 18 & 15 \end{bmatrix} \text{ as } x^T A^k = b^T \text{ where } k$$

is a large whole number.

Remark 22 Some authors set up transition matrices so that j represents the starting location and i represents the ending location. In these cases, the columns add to one. For this case, the entire equation is transposed, so instead of $x^T A^k = b^T$ where x is the column vector of initial distribution and b is the column vector of distributions after k steps, the equation is $(A^k)^T x = b$.

Questions

1. Using the data from our example, if Pizza Company has the money to enlarge one of their kitchens, which kitchen should they enlarge?

2. If they have to close a kitchen, which one should they close?

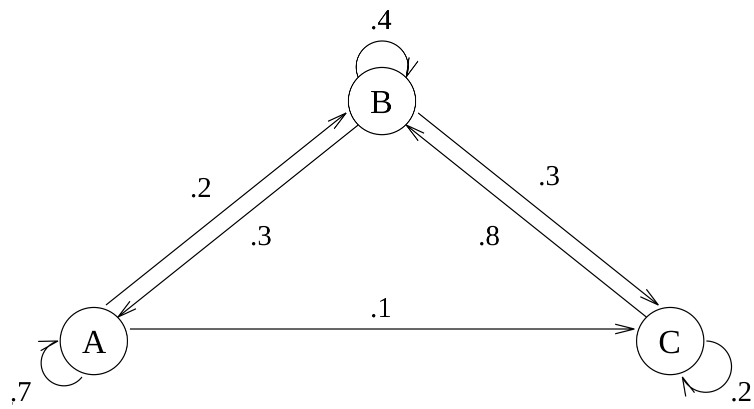
Answers

1. Since almost 39% of their business comes from area A, Pizza Company should enlarge kitchen A.

2. Since the least amount of their business comes from area C, that should be the first to close.

Problems

1. Set up a matrix, similar to the matrices we used in this chapter, that corresponds to this state diagram:



2. (a) Draw a picture corresponding to this transition matrix:

	A	B	C	D
A	.25	.15	.2	.4
B	0	.5	.5	0
C	0	0	1	0
D	.3	.4	.1	.2

   (b) Look closely at C in your picture. What do you notice that is strange about the way information flows near C? What effect do you think this might have on the long-range distribution of matter in this system?

3. Which of the following are transition matrices? Explain.

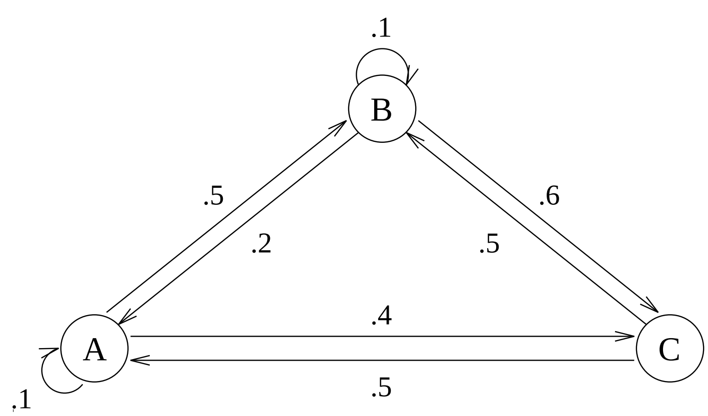
   (a) $\begin{bmatrix} .4 & .3 & .3 \\ .2 & .4 & .4 \\ .6 & .1 & .3 \end{bmatrix}$

(b) $\begin{bmatrix} .2 & .3 & .5 \\ .6 & .1 & .2 \\ .7 & .1 & .3 \end{bmatrix}$

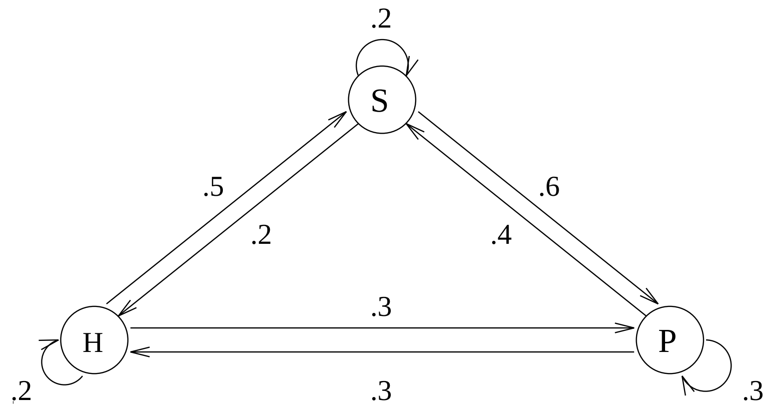
(c) $\begin{bmatrix} .25 & .15 & .3 & .4 \\ .5 & 0 & .15 & .3 \\ .15 & .35 & .4 & .2 \\ .1 & .5 & .2 & .2 \end{bmatrix}$

4. Which of these situations can be modeled by a homogeneous Markov chain? If they cannot be modeled by a Markov chain, explain why not.

   (a) The picture represents the probability that a delivery truck that is currently in region i (A, B, or C) will be in region j (A, B, or C) for the next time period.



   (b) The picture represents the probability that a person eating meal i (H, S, or P) for lunch today will eat meal j (H, S, or P) for lunch tomorrow. The letter H stands for hamburger, S stands for salad, and P stands for pizza.



5. Look at the questions in problem 4. Write a sample transition matrix for the problem or problems that can be modeled using a Markov chain.
6.

Assume S is the transition matrix

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \left[\begin{array}{ccc} .2 & .3 & .5 \\ .4 & .4 & .2 \\ .4 & .6 & 0 \end{array}\right] \end{array} = S$$

(a) What is the probability of going from state A to state B in one step?

(b) What is the probability of going from state B to state C in exactly two steps?

(c) What is the probability of going from state C to state A in exactly three steps?

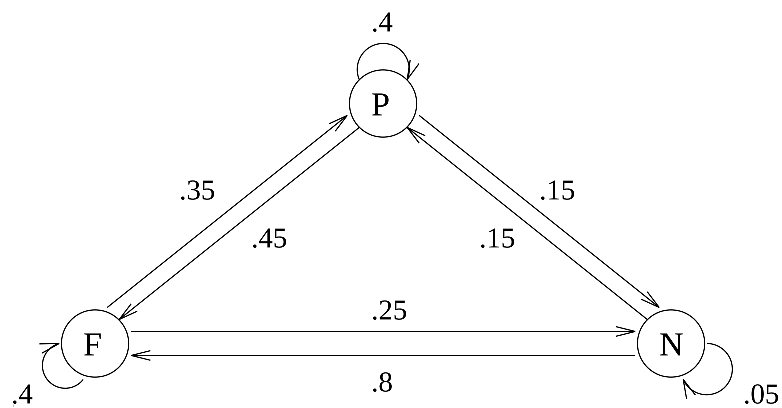
(d) Give the transition matrix, S_2, for two steps (S_2 would give the probabilities of going from state i to state j in exactly 2 steps).

(e) Give the transition matrix for three steps.

(f) Give the transition matrix for four steps.

(g) To what matrix do these transition matrices appear to converge after a large number of steps? Your solution should be accurate to two decimal places.

7. A math teacher, not wanting to be predictable, decided to assign homework based on probabilities. On the first day of class, she drew this picture on the board to tell the students whether to expect a full assignment, a partial assignment, or no assignment the next day.



(a) Construct and label the transition matrix that corresponds to this drawing. Label it A.

(b) If students have a full assignment today, what is the probability that they will have a full assignment again tomorrow.

(c) If students have no assignment today, what is the probability that they will have no assignment again tomorrow.

(d) Today is Wednesday and students have a partial assignment. What is the probability that they will have no homework on Friday?

(e) Matrix A is the transition matrix for one day. Find the transition matrix for two days (for example, if today is Monday, what are the chances of getting each kind of assignment on Wednesday?).

(f) Find the transition matrix for three days.

(g) If you have no homework this Friday, what is the is the probability that you will have no homework next Friday (since we are only considering school days, there are only 5 days in a week)? Give your answer accurate to two decimal places.

(h) Find, to two decimal places, the matrix to which matrix A would appear to converge after many days.

(i) Explain the meaning of your solution to problem 7h.

Chapter 8

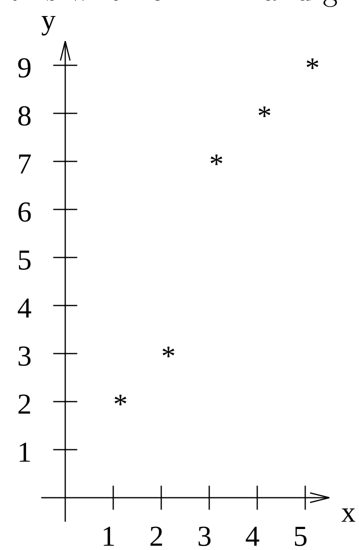
Least Squares Approximation

Suppose you are in a science class and you receive these instructions:

> Find the temperature of the water (in degrees Celsius) at the times 1, 2, 3, 4, and 5 seconds after you have applied heat to the container. Conduct your experiment carefully. Graph each data point with time on the x-axis and temperature on the y-axis. Your data should follow a straight line. Find the equation of this line.

The data from the experiment looks like this when charted and graphed:

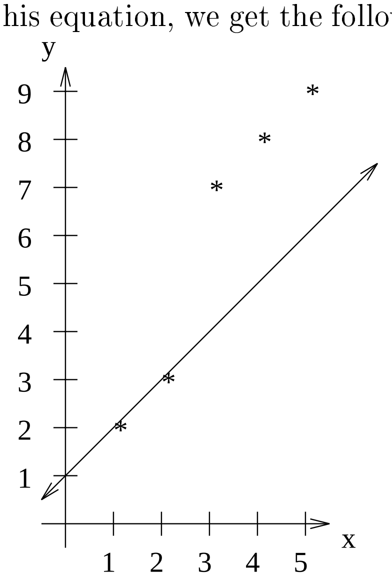
x	y
1	2
2	3
3	7
4	8
5	9



Notice that our data points don't fall exactly on a straight line as they were supposed to, so how are we going to find the slope and intercept of the line?

This is a common problem with experimental sciences because the data points that we measure seldom fall on a straight line. Therefore, scientists try to find an approximation. In this case, they would try to find the line that best fits the data in some sense. The first problem is to define "best fit." It is convenient to define an error as a distance from the actual value of y for x (the value that was measured in the experiment) to the predicted value of y for x. Therefore, it seems reasonable that the "best fit" line would somehow minimize the errors, but how? You could minimize the sum of the absolute values of the errors; this is called the ℓ_1 fit. It would also be reasonable to find the biggest error for each line and choose the line that minimizes this quantity; this is called the ℓ_∞ fit. However, the fit that is used most often is the ℓ_2 fit which is called the **least squares fit**. This method is called the least squares fit because it finds the line that minimizes the sum of the squares of the errors. Gauss developed this method to solve a problem when he was a young man (about the age of a high-school senior) to help his friend solve a chemistry problem. This is the fit that is most often used because it is the only one that can be found by solving a system of linear equations.

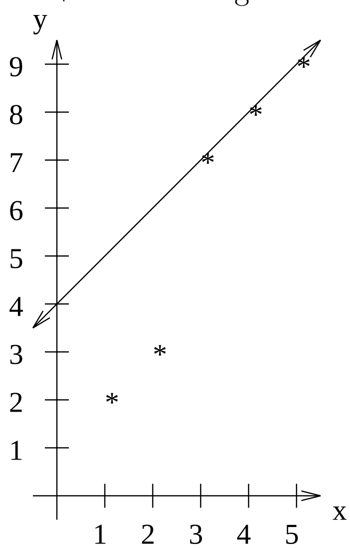
You have just read a lot of new information, so let's illustrate the concepts with our example. We have the graph of the data above. Now we need to guess which line best fits our data. If we assume that the first two points are correct and choose the line that goes through them, we get the line $y = 1 + x$. If we substitute our points into this equation, we get the following chart. The points and line are graphed below.



x	y	predicted y	error	$(\text{error})^2$
1	2	2	0	0
2	3	3	0	0
3	7	4	3	9
4	8	5	3	9
5	9	6	3	9

Therefore, the sum of the squares of the errors is 27. Do you think that we can do better than this?

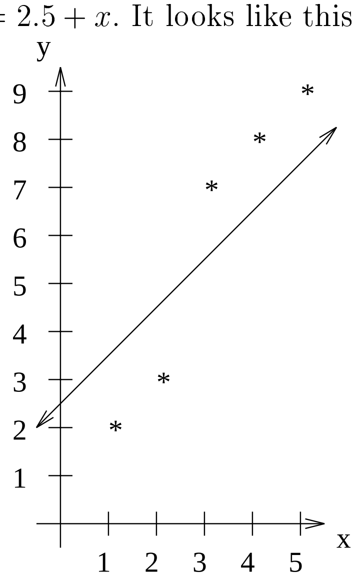
If we choose the line that goes through the points when $x = 3$ and 4, we get the line $y = 4 + x$. Will we get a better fit? Let's look at it.



x	y	predicted y	error	$(\text{error})^2$
1	2	5	-3	9
2	3	6	-3	9
3	7	7	0	0
4	8	8	0	0
5	9	9	0	0

The sum of the squares of the error is 18. That is a better fit, but can we do even better?

Let's try the line that is half way between these two lines. The equation would be $y = 2.5 + x$. It looks like this:



x	y	predicted y	error	$(\text{error})^2$
1	2	3.5	-1.5	2.25
2	3	4.5	-1.5	2.25
3	7	5.5	1.5	2.25
4	8	6.5	1.5	2.25
5	9	7.5	1.5	2.25

The sum of the squares of the error is 11.25 with this line, so this is the best line yet. Can we do better? It doesn't seem very scientific or efficient to keep guessing at which line would give the best fit. Surely there is a methodical way to determine the best fit line. Let's think about what we want.

A line in slope-intercept form looks like $c_0 + c_1 x = y$ where c_0 is the y-intercept and c_1 is the slope. We want to find c_0 and c_1 such that $c_0 + c_1 x_i = y_i$ is true for all our data points:

$$\begin{aligned} c_0 + 1c_1 &= 2 \\ c_0 + 2c_1 &= 3 \\ c_0 + 3c_1 &= 7 \\ c_0 + 4c_1 &= 8 \\ c_0 + 5c_1 &= 9 \end{aligned}$$

We know that there may not exist c_0 and c_1 that fit all these equations, so we try to find the best fit. We can write these equations in the form $Xc = y$ (these are just new letters for our familiar equation $Ax = b$) where

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}, \ c = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}, \ \text{and } y = \begin{bmatrix} 2 \\ 3 \\ 7 \\ 8 \\ 9 \end{bmatrix}.$$

In general, we cannot solve this system because the system is usually inconsistent because it is overdetermined. In other words, we have more equations than unknowns (the unknowns are the two variables, c_0 and c_1, for which we are trying to solve). There is a system of equations called the **normal equations** that can be used to find least squares solution to systems with more equations than unknowns.

Theorem 8.1 For the system $Xc = y$, c^* is a least-squares solution (ie., it minimizes the sum of the squares of the errors), if and only if c^* is a solution of the normal equations $X^TXc = X^Ty$.

Remark 23 It is important to remember that the solution to the normal equations is only an approximate solution for $Xc = y$. In general, it is not an exact solution because $Xc = y$ may be inconsistent, so it may not have a solution. In other words, there may not exist a vector, c, that makes $Xc = y$ a true statement. Therefore, we use the normal equations to find an approximate solution.

The normal equations will give us the "best fit" line (or curve) every time according to the way we defined "best fit." The proof of this is at the end of this chapter. Let's try applying the normal equations to our system. First, we multiply so that we have a system that we can solve.

$$
\begin{aligned}
X^TXc &= X^Ty \\
\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}
\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}
&=
\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}
\begin{bmatrix} 2 \\ 3 \\ 7 \\ 8 \\ 9 \end{bmatrix} \\
\begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}
\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}
&=
\begin{bmatrix} 29 \\ 106 \end{bmatrix}
\end{aligned}
$$

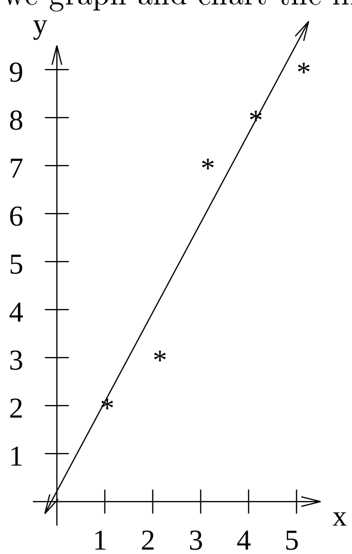
Now we can work with the augmented matrix and use Gauss-Jordan elimination to find the solution of the normal equations. This solution will be the coefficients of the line which give the best fit in the least squares sense.

$$\left[\begin{array}{cc|c} 5 & 15 & 29 \\ 15 & 55 & 106 \end{array}\right] \begin{array}{l}\text{Augmented} \\ \text{System}\end{array} \Longrightarrow \left[\begin{array}{cc|c} 1 & 3 & 5.8 \\ 15 & 55 & 106 \end{array}\right] \; r1 \div 5$$

$$\left[\begin{array}{cc|c} 1 & 3 & 5.8 \\ 0 & 10 & 19 \end{array}\right] \; -15 * r1 + r2 \Longrightarrow \left[\begin{array}{cc|c} 1 & 3 & 5.8 \\ 0 & 1 & 1.9 \end{array}\right] \; r2 \div 10$$

$$\left[\begin{array}{cc|c} 1 & 0 & 0.1 \\ 0 & 1 & 1.9 \end{array}\right] \; -3 * r2 + r1 \Longrightarrow \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0.1 \\ 1.9 \end{bmatrix}$$

When we graph and chart the line $y = 0.1 + 1.9x$, we get:



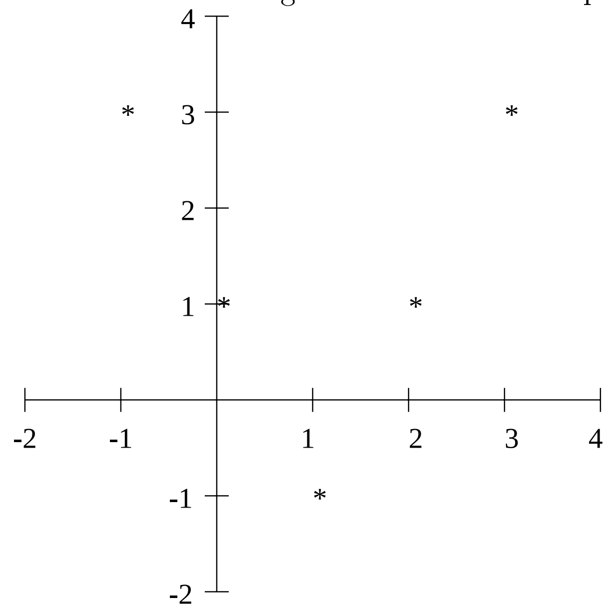
x	y	predicted y	error	$(\text{error})^2$
1	2	2.0	0	0
2	3	3.9	$-.9$	.81
3	7	5.8	1.2	1.44
4	8	7.7	.3	.09
5	9	9.6	$-.6$	.36

The sum of the squares of the error is 2.7. This is a great improvement over our guesses and we know that we cannot do any better. In general, if we have n data points, we solve $X^T X c = X^T y$ with $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{n-1} \\ 1 & x_n \end{bmatrix}$, $c = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}$.

The ellipse marks (written as $\vdots, \cdots,$ or $\ddots$) tell you to continue in the same pattern.

What if we are told that our data is not supposed to fit a straight line, but instead falls in the shape of a parabola? Consider the following data from another experiment:

x	y
-1	3
0	1
1	-1
2	1
3	3



We can find the curve that best fits our data in a similar manner. The general equation for a parabola is $c_0 + c_1 x + c_2 x^2 = y$. Therefore, we want to find the values of the coefficients, c_1, c_2, and c_3, so that the curve we find best fits these equations:

$$
\begin{aligned}
c_0 - 1c_1 + 1c_2 &= 3 \\
c_0 + 0c_1 + 0c_2 &= 1 \\
c_0 + 1c_1 + 1c_2 &= -1 \\
c_0 + 2c_1 + 4c_2 &= 1 \\
c_0 + 3c_1 + 6c_2 &= 3
\end{aligned}
$$

Let us use the normal equations with $X = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}$, $c = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$, and $y = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \\ 3 \end{bmatrix}$.

$$X^T X c = X^T y$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 5 & 15 \\ 5 & 15 & 35 \\ 15 & 35 & 99 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 33 \end{bmatrix}$$

Now we can augment the matrix and solve using Gaussian elimination.

$$\left[\begin{array}{ccc|c} 5 & 5 & 15 & 7 \\ 5 & 15 & 35 & 7 \\ 15 & 35 & 99 & 33 \end{array}\right] \begin{array}{l} \text{Augmented} \\ \text{Matrix} \end{array} \implies \left[\begin{array}{ccc|c} 1 & 1 & 3 & 1\frac{2}{5} \\ 5 & 15 & 35 & 7 \\ 15 & 35 & 99 & 33 \end{array}\right] \; r1 \div 5$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 1\frac{2}{5} \\ 0 & 10 & 20 & 0 \\ 0 & 20 & 54 & 12 \end{array}\right] \begin{array}{c} \\ -5*r1+r2 \\ -15*r1+r3 \end{array} \Longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 3 & 1\frac{2}{5} \\ 0 & 1 & 2 & 0 \\ 0 & 20 & 54 & 12 \end{array}\right] r2 \div 10$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 1\frac{2}{5} \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 14 & 12 \end{array}\right] \begin{array}{c} \\ \\ -20*r2+r3 \end{array} \Longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 3 & 1\frac{2}{5} \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & \frac{6}{7} \end{array}\right] r3 \div 14$$

Back-substitution yields the coefficients

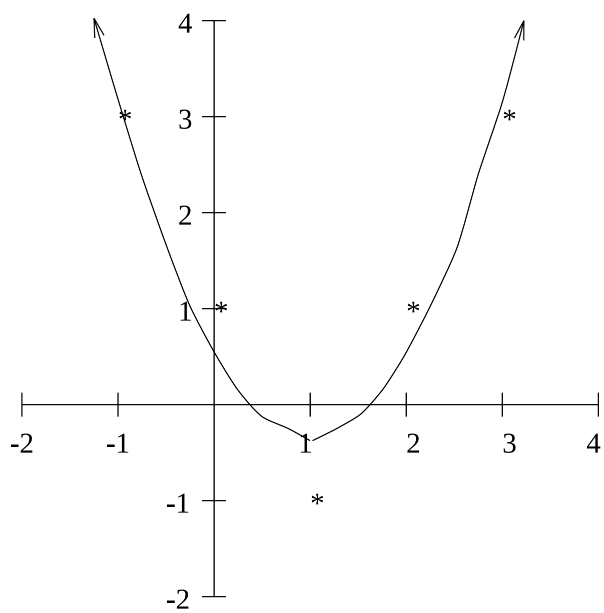
$$c_2 = \frac{6}{7}$$

$$c_1 + 2\left(\frac{6}{7}\right) = 0 \Rightarrow c_1 = -1\frac{5}{7}$$

$$c_0 + \left(-1\frac{5}{7}\right) + 3\left(\frac{6}{7}\right) = 1\frac{2}{5} \Rightarrow c_0 = \frac{19}{35}$$

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \frac{19}{35} \\ -1\frac{5}{7} \\ \frac{6}{7} \end{bmatrix}$$

These coefficients indicate that the curve we want is $y = \frac{19}{35} - 1\frac{5}{7}x + \frac{6}{7}x^2$. Let's graph this curve and fill in our chart:



x	y	expected y	error	$(\text{error})^2$
-1	3	$3\frac{4}{35}$	$\frac{-4}{35}$	$\frac{16}{1225}$
0	1	$\frac{19}{35}$	$\frac{16}{35}$	$\frac{256}{1225}$
1	-1	$\frac{-11}{35}$	$\frac{-24}{35}$	$\frac{576}{1225}$
2	1	$\frac{19}{35}$	$\frac{16}{35}$	$\frac{256}{1225}$
3	3	$3\frac{4}{35}$	$\frac{-4}{35}$	$\frac{16}{1225}$

We find that the sum of the squared errors is $\frac{32}{35}$. Using our definition of least squares "best fit," you will not be able to find a parabola that fits the data better than this one. In general, to find the parabola that best fits the data, you use the normal equations $X^TXc = X^Ty$ with

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 \\ 1 & x_n & x_n^2 \end{bmatrix}, \text{c} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}, \text{ and y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}.$$

Notice that the normal equations used to find the best fit line and the best fit parabola have the same form. Do you think that we could expand this to higher degree polynomials? Yes, we can. In general, we use the normal equations $X^TXc = X^Ty$ with

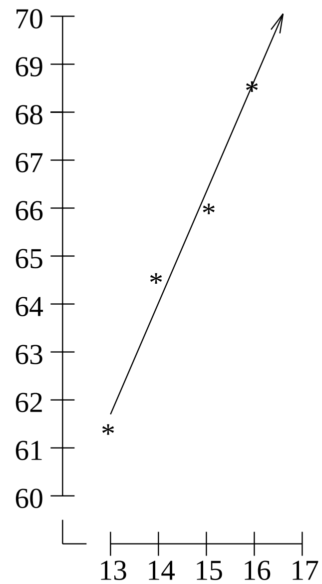
$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^m \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}, \text{ c} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{m-1} \\ c_m \end{bmatrix}, \text{ and y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix},$$

where m represents the degree of the polynomial curve that you wish to fit and n represents the number of data points. The least squares "best fit" curve for these equations is $c_0 + c_1x + c_2x^2 + \ldots + c_{m-1}x^{m-1} + c_mx^m$. Remember that the degree is the highest power of the variable in your equation. A line is a first degree polynomial and a parabola is a second degree polynomial.

If we can find the best fit curve for any degree polynomial, why don't we always use a higher degree polynomial and fit the data better? After all, if we have n data points and fit them to a polynomial of degree $n-1$, we will have a perfect fit every time because our systems would not be inconsistent. However, our goal is not just to find a curve that fits the data closely. Usually, we want the curve to predict what would happen between our data points. If we choose a curve that exactly fits all our data points, we are incorporating the error in our measurements into our model unless the model fits the data exactly (which occurs only rarely.) Unfortunately, there is no set rule for deciding what degree polynomial should be used to fit the data. However, first and second degree polynomials provide the simplest models and should fit most of your data until you start modeling more complicated systems.

If you notice, we said that we usually fit a curve so that we can predict what would happen between our data points. Predicting an outcome between data points is called **interpolation**. Why didn't we say anything about predicting the behavior beyond our data points? Predicting an outcome beyond the data is called **extrapolation**. It is usually dangerous to extrapolate much beyond the data because we have no indication that the data will continue to follow the same curve since our curve was only fit to the data. For example, we measured the height of a teenage boy every year for a few years and charted his growth. The growth appeared linear, so we fit a line to the data and got $y = 32 + 2.25x$. We have graphed the data with age on the x-axis and height on the y-axis.

Age	Height in inches
13	61.5
14	64.5
15	66.0
16	68.5



If we extrapolate back several years, this young man was over two and a half feet tall when he was born. According to this model, he will never stop growing, so he will be 8 feet 4 inches tall by the time he is 30 and almost 14 feet tall by the time he is 60. Do you think that this is an accurate prediction?

If the temperature at the airport on the 4th of July was in the 90's for two years in a row, would it be reasonable to predict that the temperature in January between those years was also in the 90's? No, it would not. We have two problems with this model. One problem is that we only have 2 data points. You can always find a line that fits the two points, but there is no reason to believe that the relationship between the day of the year and the temperature is a linear relationship. Also, we didn't take into account other factors that could affect our model such as the pattern of the seasons. These are problems that can arise when you model a situation. When we start modeling situations and using least squares to make predictions, we are entering the world of statistics. That means that we must think about what the data represents rather than just apply the normal equations. There are many interesting applications of statistics that you can explore in another course. However, using matrices, you already know one way to find a "best fit" curve for your data.

8.1 Proof of Normal Equations

We are interested in the fact that a solution to the normal equations is a least squares solution. We will prove this fact, but will not prove the theorem in the other direction. In this proof, y represents the solution of the normal equations (solution of $A^TAy = A^Tb$) and x is any n-vector.

The length of the vector x is defined to be $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, so the squared length of the vector is $\|x\|^2 = \langle x, x\rangle = x^Tx = x_1^2 + x_2^2 + \cdots + x_n^2$. The square of the length of the vector $Ax - b$, $\|Ax - b\|^2$, is the sum of the squares of the errors since each element of $Ax - b$ for data fitting represents the error at a point. We want to prove that the sum of the squares of the error for y is less than or equal to the sum of the squares of the error for x. In other words, we want to prove that $\|Ay - b\|^2 \leq \|Ax - b\|^2$ Notice that the steps of the proof are numbered. Explanations for each step follow the proof.

$$\text{To begin with, we have}$$

$$0 \leq \|Ax - Ay\|^2 \tag{1}$$

$$= (Ax - Ay)^T(Ax - Ay) \tag{2}$$

$$= (x^TA^T - y^TA^T)(Ax - Ay) \tag{3}$$

$$= x^TA^TAx - x^TA^TAy - y^TA^TAx + y^TA^TAy. \tag{4}$$

$$\text{so}$$

$$x^TA^TAy + y^TA^TAx \leq x^TA^TAx + y^TA^TAy \tag{5}$$

$$\text{equivalently,}$$

$$2x^TA^TAy \leq x^TA^TAx + y^TA^TAy \tag{6}$$

$$\text{Now, by adding and subtracting like terms, we obtain}$$

$$2x^TA^TAy - 2y^TA^TAy - 2x^TA^Tb + 2y^TA^Tb$$

$$\leq x^TA^TAx + y^TA^TAy - 2y^TA^TAy - 2x^TA^Tb + 2y^TA^Tb + b^Tb - b^Tb \tag{7}$$

which we can rewrite as

$$2((x^T - y^T)(A^TAy) - (x^T - y^T)(A^Tb))$$

$$\leq (x^TA^TAx - 2x^TA^Tb + b^Tb) - (y^TA^TA - 2y^TA^Tb + b^Tb) \tag{8}$$

or

$$2(x^T - y^T)(A^TAy - A^Tb) \leq (Ax - b)^T(Ax - b) - (Ay - b)^T(Ay - b). \tag{9}$$

It follows that

$$0 \leq (Ax - b)^T(Ax - b) - (Ay - b)^T(Ay - b). \tag{10}$$

Hence

$$(Ay - b)^T(Ay - b) \leq (Ax - b)^T(Ax - b) \tag{11}$$

and

$$\|Ay - b\|^2 \leq \|Ax - b\|^2 . \tag{12}$$

1. This is the sum of numbers that have been squared. Squared numbers can not be negative, so the sum of them can not be negative either.

2. $\|x\|^2 = x^Tx$

3. $(M - N)^T = M^T - N^T$ and $(RS)^T = S^TR^T$

4. Multiply binomials but remember that the order of multiplication matters. FOIL is one method to multiply binomials.

5. Rearrange the terms

6. All the terms are 1 by 1 matrices, which we consider to be the same as real numbers. Transposing a real number does not change it. $(x^TA^TAy)^T = y^TA^TAx$

7. We subtracted $2y^TA^TAy$ and $2x^TA^Tb$ from both sides and added $2y^TA^Tb$ to both sides of the inequality. Since we added and subtracted the same terms from both sides of the inequality, we did not change the inequality. We also added zero to the right-hand side of the inequality in the form of $b^Tb - b^Tb$. We did all this so that we can factor to get what we want. This is similar to completing the square.

8. Factor out the two from every term of the left-hand side. Then factor out $(x^T - y^T)$ from the first two and last two terms of the left-hand side. On the right-hand side, rearrange the terms.

9. On the left-hand side, factor out $(x^T - y^T)$. On the right-hand side, factor both trinomials.

10. The left-hand side reduces to zero because $A^TAy = A^Tb$.

11. Rearrange the inequality

12. $x^Tx = \|x\|^2$

Question

So far, we have used the normal equations only to fit data to polynomials. Are the normal equations restricted to polynomials?

Answer

No, the normal equations are not restricted to polynomials. They can be used for any function of one variable. For example, we could find the curve that best fits the function $y = c_0 + \frac{c_1}{x}$ for certain data points. We would solve $X^TXc = X^Ty$ with

$$X = \begin{bmatrix} 1 & \frac{1}{x_1} \\ 1 & \frac{1}{x_2} \\ \vdots & \vdots \\ 1 & \frac{1}{x_{n-1}} \\ 1 & \frac{1}{x_n} \end{bmatrix}, \ c = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \text{ and } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}.$$ The procedures would be the same as fitting data to a straight line except that we use $\frac{1}{x}$ instead of x.

Problems

1. (a) Graph the following points. Find and graph the line (accurate to one decimal place) that best fits the data according to the least squares definition presented in this chapter. Find the sum of the squared errors.

x	y
-2	6
0	3
2	0

(b) Follow the directions for part (a) with the following data:

x	y
-2	6
0	4
2	0

(c) Follow the directions for part (a) with the following data:

x	y
-2	6
-1	5
0	4
1	1
2	0

2. Find the line that best fits this data:

x	y
-2	8
0	6
1	5
2	3

3. Find the line that best fits this data:

x	y
1	3
3	10
4	11
6	19
8	24

4. Find the parabola that best fits this data and give the sum of the squared errors (use two decimal places):

x	y
-2	16
0	0
2	11
3	26

5. Find the parabola that best fits this data:

x	y
-2	10
-1	6
0	4
1	1
2	3

6. Using the following data, find the best fit straight line and the best fit parabola. Your solutions should be accurate to four decimal places.

x	y
-3	0
-1	3
0	6
1	8
2	9

7. Find the matrix X and the vector y that would be used in the normal equations to find the best fit cubic (third degree) polynomial to the following data:

x	y
-2	44
-1	11
0	3
1	1
3	-91

8. On a sunny day, measure the height of 5 objects and record those as x values. Measure the shadows of your objects and record those as y values. It is best if you take the measurements all at the same time and not around noon. Label your data and graph it. Find the best fit line through the data.

Chapter 9

Eigenvalues and Eigenvectors

Have you ever heard the words eigenvalue and eigenvector? They are derived from the German word "eigen" which means "proper" or "characteristic." An eigenvalue of a square matrix is a scalar that is usually represented by the Greek letter λ (pronounced lambda). As you might suspect, an eigenvector is a vector. Moreover, we require that an eigenvector be a non-zero vector, in other words, an eigenvector can not be the zero vector. We will denote an eigenvector by the small letter x. All eigenvalues and eigenvectors satisfy the equation $Ax = \lambda x$ for a given square matrix, A.

> **Remark 24** Remember that, in general, the word scalar is not restricted to real numbers. We are only using real numbers as scalars in this book, but eigenvalues are often complex numbers.

> **Definition 9.1** Consider the square matrix A. We say that λ is an **eigenvalue** of A if there exists a non-zero vector x such that $Ax = \lambda x$. In this case, x is called an **eigenvector** (corresponding to λ), and the pair (λ, x) is called an **eigenpair** for A.

Let's look at an example of an eigenvalue and eigenvector. If you were asked if $x = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue $\lambda = 0$ for $A = \begin{bmatrix} 6 & 3 \\ -2 & -1 \end{bmatrix}$, you could find out by substituting x, λ, and A into the equation $Ax = \lambda x$.

$$\begin{aligned} Ax &= \lambda x \\ \begin{bmatrix} 6 & 3 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} &= 0 \begin{bmatrix} 1 \\ -2 \end{bmatrix} \end{aligned}$$

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

Therefore, λ and x are an eigenvalue and an eigenvector, respectively, for A.

Now that you have seen an eigenvalue and an eigenvector, let's talk a little more about them. Why did we require that an eigenvector not be zero? If the eigenvector was zero, the equation $Ax = \lambda x$ would yield $0 = 0$. Since this equation is always true, it is not an interesting case. Therefore, we define an eigenvector to be a non-zero vector that satisfies $Ax = \lambda x$. However, as we showed in the previous example, an eigenvalue can be zero without causing a problem. We usually say that x is an eigenvector corresponding to the eigenvalue λ if they satisfy $Ax = \lambda x$. Since each eigenvector is associated with an eigenvalue, we often refer to an x and λ that correspond to one another as an **eigenpair**. Did you notice that we called x "an" eigenvector rather than "the" eigenvector corresponding to λ? This is because any non-zero, scalar multiple of an eigenvector is also an eigenvector. If you let c represent a scalar, then we can prove this fact through the following steps.

$$A(cx) = cAx = c\lambda x = \lambda(cx)$$

Since any non-zero, scalar multiple of an eigenvector is also an eigenvector, $\begin{bmatrix} 2 \\ -4 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ are also eigenvectors corresponding to $\lambda = 0$ when $A = \begin{bmatrix} 6 & 3 \\ -2 & -1 \end{bmatrix}$.

You have already computed eigenvectors in this course. When we studied Markov chains, you computed an eigenvector corresponding to A^T when you found the matrix to which the probabilities seemed to converge after many steps. Any row of that matrix is an eigenvector for A^T because all the rows of that matrix are the same. We write that row as a column vector when we use it as an eigenvector. The eigenvector that you found is called the **dominant eigenvector**

Definition 9.2 The **dominant eigenvector** of a matrix is an eigenvector corresponding to the eigenvalue of largest magnitude (for real numbers, largest absolute value) of that matrix.

Although we only found one eigenvector, we found a very important eigenvector. Many of the "real world" applications are primarily interested in the dominant eigenpair. The method that you used to find this eigenvector is called the **power method**. The power method will be explained later in this chapter. An eigenvector corresponding to the transpose of a transition matrix is the transpose of any row of the matrix that A^k converges to as k grows, these rows are all the same. The dominant eigenvalue is always 1 for a transition matrix. Let's look at the example that we used in the Markov chain chapter. Consider the matrix $A = \begin{bmatrix} .3 & .3 & .4 \\ .4 & .4 & .2 \\ .5 & .3 & .2 \end{bmatrix}$.

If k is large, $A^k \approx \begin{bmatrix} .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \\ .3\overline{8} & .3\overline{3} & .2\overline{7} \end{bmatrix}$. Therefore, an eigenvector corresponding to the dominant eigenvalue, λ =1, is $\begin{bmatrix} .3\overline{8} \\ .3\overline{3} \\ .2\overline{7} \end{bmatrix}$. Let's see if $A^T x = \lambda x$ holds true.

$$\begin{bmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{bmatrix} \begin{bmatrix} .3\overline{8} \\ .3\overline{3} \\ .2\overline{7} \end{bmatrix} = 1 \begin{bmatrix} .3\overline{8} \\ .3\overline{3} \\ .2\overline{7} \end{bmatrix}.$$

Yes, the equation holds, so we have found an eigenpair corresponding to the transpose of the transition matrix.

Remark 25 For a transition matrix, the dominant eigenvalue is always 1. An eigenvector corresponding to $\lambda = 1$ for A^T is the transpose of any row of A^* where A^* is the matrix to which A^k converges as k grows. The matrix for which we are finding an eigenpair must have been set up so that the columns (not the rows) add to 1 for the eigenvector to be read from A^*; this is why we are dealing with A^T instead of A. These rules will require some modification if we are not dealing with a transition matrix.

If we do not have a transition matrix, can we still use the power method? Yes we can, but we need to modify the steps a bit because the dominant eigenvalue will not necessarily be the number one. Let us explain how to use the power method. An example follows the remarks to help clarify these steps.

1. Choose a vector and call it x_0. Set $i = 0$.

2. Multiply to get the next approximation for x using the formula $x_{i+1} = Ax_i$.

3. Divide every term in x_{i+1} by the last element of the vector and call the new vector $x_{i+1}^{'}$.

4. Repeat steps 2 and 3 until $x_i^{'}$ and $x_{i+1}^{'}$ agree to the desired number of digits.

5. The vector obtained in step 4 is an approximate eigenvector corresponding to the dominant eigenvalue. We will call it x.

6. An approximation to the dominant eigenvalue is $\frac{x^T Ax}{x^T x}$. This is called the **Rayleigh quotient** of x.

Remark 26 Because any constant multiple of an eigenvector is an eigenvector, we did not have to divide by the last element in the vector in step

3. We could have divided by any element or not divided at all. We divided so that our vector would not grow too large and we could tell when we had converged. We divided by the last element of the vector so that we would have a well-defined algorithm for using the power method. The choice of the last element over any of the others was arbitrary. Therefore, if the last element is zero, divide by another element of the vector for that entire problem. When people program the power method on a computer, they usually divide by $||x||_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, which is defined as the length of the vector, so that they don't have to worry about whether or not an element is zero.

Remark 27 Some calculators will not let you divide a vector by a constant. On those calculators, you can multiply by the multiplicative inverse (reciprocal) of the constant.

Remark 28 You will probably not be able to directly input the Rayleigh quotient into your calculator. It will consider the numerator and denominator as 1 by 1 matrices. We consider 1 by 1 matrices to be the same as real numbers, but your calculator may not consider them the same. Since you cannot divide matrices, your calculator will probably give you an error message.

You have seen the steps to the power method. Let's demonstrate those steps on the matrix $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$. For step 1, we arbitrarily chose $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Let's make a chart for steps 2 and 3.

$$x_1 = Ax_0 = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 5 \end{bmatrix} \Rightarrow x_1^{'} = \begin{bmatrix} 1.8 \\ 1 \end{bmatrix}$$

$$x_2 = Ax_1 = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1.8 \\ 1 \end{bmatrix} = \begin{bmatrix} 11.4 \\ 5.8 \end{bmatrix} \Rightarrow x_2^{'} = \begin{bmatrix} 1.965517241 \\ 1 \end{bmatrix}$$

$$x_3 = Ax_2 = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1.965517241 \\ 1 \end{bmatrix} = \begin{bmatrix} 11.89655172 \\ 5.96551724 \end{bmatrix} \Rightarrow x_3^{'} = \begin{bmatrix} 1.994219653 \\ 1 \end{bmatrix}$$

$$x_4 = Ax_3 = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1.994219653 \\ 1 \end{bmatrix} = \begin{bmatrix} 11.98265896 \\ 5.99421965 \end{bmatrix} \Rightarrow x_4^{'} = \begin{bmatrix} 1.99903568 \\ 1 \end{bmatrix}$$

Therefore, it looks like an eigenvector is $x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The corresponding eigenvalue is:

$$\frac{x^T Ax}{x^T x} = \frac{\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}} = \frac{\begin{bmatrix} 30 \end{bmatrix}}{\begin{bmatrix} 5 \end{bmatrix}} \Rightarrow \lambda = 6$$

Let's try to find the dominant eigenpair of another matrix. Consider the matrix $A = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix}$. Again, we will choose $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, but we could have chosen any vector of dimension 2. Let's look at the chart for steps 2 and 3.

$$x_1 = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \end{bmatrix} \Rightarrow x_1^{'} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We divided by 8 to get $x_1^{'}$. We are allowed to do this because, as Remark 26 states, we can divide by any element in the vector as long as we are consistent with our choice throughout the problem.

$$x_2 = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \Rightarrow x_2^{'} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Notice that $x_2^{'} = x_0$. This means that our vectors will just continue in a cycle and never converge. The power method only works if there is one eigenvalue whose absolute value is strictly larger than the absolute value of the other eigenvalues. Even when the power method works, convergence might be slow.

The power method found one very important eigenpair, but what should we do if we want to find all the eigenpairs? We know that $Ax = \lambda x$ for all eigenpairs. We can transform this equation into a form that will help us. When we learned about the identity matrix, we learned that $Ix = x$ for any x. Therefore, $Ax = \lambda Ix$. We can use algebra steps from here.

$$\begin{aligned} Ax &= \lambda I x \\ Ax - \lambda I x &= 0 \\ (A - \lambda I)x &= 0 \end{aligned}$$

We know that $x = 0$ would solve this equation, but we defined an eigenvector to be non-zero, so if there is an eigenvector solution to the equation $(A - \lambda I)x = 0$, then there must be more than one solution to the equation. We learned in Chapter 6 that the system has a unique solution if $\det(A - \lambda I) \neq 0$. Therefore, we know that if there is a non-zero solution to $(A - \lambda I)x = 0$, then $\det(A - \lambda I) = 0$. The equation $\det(A - \lambda I) = 0$ even has a name. It is called the **characteristic equation**. We can solve the characteristic equation to find all the eigenvalues of certain matrices. There will be as many eigenvalues as there are rows in the matrix (or columns since the matrix must be square), but some of the eigenvalues might be identical to each other.

Let's find both of the eigenvalues of the matrix $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$.

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} 3 - \lambda & 6 \\ 1 & 4 - \lambda \end{bmatrix} \\ \det(A - \lambda I) &= (3 - \lambda)(4 - \lambda) - 6 \end{aligned}$$

$$= \lambda^2 - 7\lambda + 6$$
$$= (\lambda - 6)(\lambda - 1)$$

Therefore, $\lambda = 6$ or $\lambda = 1$. We now know our eigenvalues. Remember that all eigenvalues are paired with an eigenvector. Therefore, we can substitute our eigenvalues, one at a time, into the formula $(A - \lambda I)x = 0$ and solve to find a corresponding eigenvector.

Let's find an eigenvector corresponding to $\lambda = 6$.

$$\begin{aligned} (A - \lambda I)x &= 0 \\ \left(\begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \right) x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} -3 & 6 \\ 1 & -2 \end{bmatrix} x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

$$\left[\begin{array}{cc|c} -3 & 6 & 0 \\ 1 & -2 & 0 \end{array}\right] \begin{array}{l} \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & -2 & 0 \\ 1 & -2 & 0 \end{array}\right] \begin{array}{l} r1 \div (-3) \\ \end{array}$$

$$\left[\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ -1 * r1 + r2 \end{array}$$

Notice that this system is underdetermined. Therefore, there are an infinite number of solutions. So, any vector that solves the equation $x_1 - 2x_2 = 0$ is an eigenvector corresponding to $\lambda = 6$ when $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$. To have a consistent method for finding an eigenvector, let's choose the solution in which $x_2 = 1$. We can use back-substitution to find that $x_1 - 2(1) = 0$ which implies that $x_1 = 2$. This tells us that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = 6$ when $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$. This is the same solution that we found when we used the power method to find the dominant eigenpair.

Let's find an eigenvector corresponding to $\lambda = 1$.

$$\begin{aligned} (A - \lambda I)x &= 0 \\ \left(\begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix} x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

$$\left[\begin{array}{cc|c} 2 & 6 & 0 \\ 1 & 3 & 0 \end{array}\right] \quad \begin{array}{l} \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ 1 & 3 & 0 \end{array}\right] \quad r1 \div 2$$

$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad -1 * r1 + r2$$

Notice that this system is underdetermined. This will always be true when we are finding an eigenvector using this method. So, any vector that solves the equation $x_1 + 3x_2 = 0$ is an eigenvector corresponding to the eigenvalue $\lambda = 1$ when $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$. Again, let's choose the eigenvector in which the last element of x is 1. Therefore, $x_2 = 1$ and $x_1 + 3(1) = 0$, so $x_1 = -3$. This tells us that $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = 1$ when $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$. Using the characteristic equation and Gaussian elimination, we are able to find all the eigenvalues to the matrix and corresponding eigenvectors.

Let's find the eigenpairs for the matrix $A = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix}$ for which the power method fails.

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} 2-\lambda & 6 \\ 2 & -2-\lambda \end{bmatrix} \\ \det(A - \lambda I) &= (2-\lambda)(-2-\lambda) - 12 \\ &= \lambda^2 - 16 \\ &= (\lambda - 4)(\lambda + 4) \end{aligned}$$

Therefore, $\lambda = 4$ or $\lambda = -4$. The power method does not work because $|4| = |-4|$. In other words, there is not a unique dominant eigenvalue.

Let's find an eigenvector corresponding to $\lambda = 4$.

$$\begin{aligned} (A - \lambda I)x &= 0 \\ \left(\begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \right) x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} -2 & 6 \\ 2 & -6 \end{bmatrix} x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

$$\left[\begin{array}{cc|c} -2 & 6 & 0 \\ 2 & -6 & 0 \end{array}\right] \begin{array}{l} \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & -3 & 0 \\ 2 & -6 & 0 \end{array}\right] \; r1 \div (-2)$$

$$\left[\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ -2 * r1 + r2 \end{array}$$

Since the system is underdetermined, we have an infinite number of solutions. Let's choose the solution in which $x_2 = 1$. We can use back-substitution to find that $x_1 - 3(1) = 0$ which implies that $x_1 = 3$. This tells us that $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = 4$ when $A = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix}$.

Let's find an eigenvector corresponding to $\lambda = -4$.

$$\begin{aligned}
(A - \lambda I)x &= 0 \\
\left(\begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix} - \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix} \right) x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\begin{bmatrix} 6 & 6 \\ 2 & 2 \end{bmatrix} x &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\end{aligned}$$

$$\left[\begin{array}{cc|c} 6 & 6 & 0 \\ 2 & 2 & 0 \end{array}\right] \quad \text{Augmented Matrix}$$

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 2 & 2 & 0 \end{array}\right] \quad r1 \div 6$$

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad -2 * r1 + r2$$

Again, let's choose the eigenvector in which the last element of x is 1. Therefore, $x_2 = 1$ and $x_1 + 1(1) = 0$, so $x_1 = -1$. This tells us that $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = -4$ when $A = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix}$. Using the characteristic equation and Gaussian elimination, we are able to find both of the eigenvalues to the matrix and corresponding eigenvectors.

We can find eigenpairs for larger systems using this method, but the characteristic equation gets impossible to solve directly when the system gets too large. We could use approximations that get close to solving the characteristic equation, but there are better ways to find eigenpairs that you will study in the future. However, these two methods give you an idea of how to find eigenpairs.

Another matrix for which the power method will not work is the matrix $A = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$, because the eigenvalues are both the real number 5. The method that we showed you earlier will yield the eigenvector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to correspond to the eigenvalue $\lambda = 5$. Other methods will reveal, and you can check, that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is also an eigenvector of A corresponding to $\lambda = 5$. Notice that these two eigenvectors are not multiples of one another. If the same eigenvalue is repeated p times for a particular matrix, then there can be as many as p different eigenvectors that are not multiples of each other that correspond to that eigenvalue.

We said that eigenvalues are often complex numbers. However, if the matrix A is symmetric, then the eigenvalues will always be real numbers. As you can see from the problems that we worked, eigenvalues can also be real when the matrix is not symmetric, but keep in mind that they are not guaranteed to be real.

Did you know that the determinant of a matrix is related to the eigenvalues of the matrix? The product of the eigenvalues of a square matrix is equal to the determinant of that matrix. Let's look at the two matrices that we have been working with. For $A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$,

$$\text{Product of eigenvalues} = \det(A)$$

$$\begin{aligned} 6 * 1 &= 12 - 6 \\ 6 &= 6 \end{aligned}$$

For $A = \begin{bmatrix} 2 & 6 \\ 2 & -2 \end{bmatrix}$,

$$\begin{aligned} \text{Product of eigenvalues} &= \det(A) \\ 4 * (-4) &= -4 - 12 \\ -16 &= -16 \end{aligned}$$

You can use this as a check to see that you have the correct eigenvalues and determinant for the matrix A.

Now that we know how to find eigenpairs, we might want to know what uses they have. The interesting uses come from larger systems, so we will just discuss them rather than solve them. Have you ever seen the video of the collapse of the Tacoma Narrows Bridge? The Tacoma Bridge was built in 1940. From the beginning, the bridge would form small waves like the surface of a body of water. This accidental behavior of the bridge brought many people who wanted to drive over this moving bridge. Most people thought that the bridge was safe despite the movement. However, about four months later, the oscillations (waves) became bigger. At one point, one edge of the road was 28 feet higher than the other edge. Finally, this bridge crashed into the water below. One explanation for the crash is that the oscillations of the bridge were caused by the frequency of the wind being too close to the natural frequency of the bridge. The natural frequency of the bridge is the eigenvalue of smallest magnitude of a system that models the bridge. This is why eigenvalues are very important to engineers when they analyze structures. (*Differential Equations and Their Applications*, 1983, pp. 171-173).

Remark 29 The eigenvalue of smallest magnitude of a matrix is the same as the inverse (reciprocal) of the dominant eigenvalue of the inverse of the matrix. Since most applications of eigenvalues need the eigenvalue of smallest magnitude, the inverse matrix is often solved for its dominant eigenvalue. This is why the dominant eigenvalue is so important.

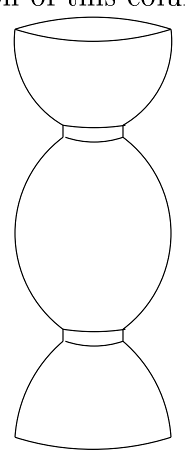
Also, a bridge in Manchester, England collapsed in 1831 because of conflicts between frequencies. However, this time, the natural frequency of the bridge was matched by the frequency caused by soldiers marching in step. Large oscillations occurred and the bridge collapsed. This is why soldiers break cadence when crossing a bridge.

Frequencies are also used in electrical systems. When you tune your radio, you are changing the resonant frequency until it matches the frequency at which your station is broadcasting. Engineers used eigenvalues when they designed your radio.

Frequencies are also vital in music performance. When instruments are tuned, their frequencies are matched. It is the frequency that determines what we hear as music. Although musicians do not study eigenvalues in order to play their instruments better, the study of eigenvalues can explain why certain sounds are pleasant to the ear while others sound “flat” or “sharp.” When two people sing in harmony, the frequency of one voice is a constant multiple of the other. That is what makes the sounds pleasant. Eigenvalues can be used to explain many aspects of music from the initial design of the instrument to tuning and harmony during a performance. Even the concert halls are analyzed so that every seat in the theater receives a high quality sound.

Car designers analyze eigenvalues in order to damp out the noise so that the occupants have a quiet ride. Eigenvalue analysis is also used in the design of car stereo systems so that the sounds are directed correctly for the listening pleasure of the passengers and driver. When you see a car that vibrates because of the loud booming music, think of eigenvalues. Eigenvalue analysis can indicate what needs to be changed to reduce the vibration of the car due to the music.

Eigenvalues are not only used to explain natural occurrences, but also to discover new and better designs for the future. Some of the results are quite surprising. If you were asked to build the strongest column that you could to support the weight of a roof using only a specified amount of material, what shape would that column take? Most of us would build a cylinder like most other columns that we have seen. However, Steve Cox of Rice University and Michael Overton of New York University proved, based on the work of J. Keller and I. Tadjbakhsh, that the column would be stronger if it was largest at the top, middle, and bottom. At the points $\frac{1}{4}$ of the way from either end, the column could be smaller because the column would not naturally buckle there anyway. A cross-section of this column would look like this:



Does that surprise you? This new design was discovered through the study of the eigenvalues of the system involving the column and the weight from above. Note that this column would not be the strongest design if any significant pressure came from the side, but when a column supports a roof, the vast majority of the pressure comes directly from above.

Eigenvalues can also be used to test for cracks or deformities in a solid. Can you imagine if every inch of every beam used in construction had to be tested? The problem is not as time consuming when eigenvalues are used. When a beam is struck, its natural frequencies (eigenvalues) can be heard. If the beam “rings,” then it is not flawed. A dull sound will result from a flawed beam because the flaw causes the eigenvalues to change. Sensitive machines can be used to “see” and “hear” eigenvalues more precisely.

Oil companies frequently use eigenvalue analysis to explore land for oil. Oil, dirt, and other substances all give rise to linear systems which have different eigenvalues, so eigenvalue analysis can give a good indication of where oil reserves are located. Oil companies place probes around a site to pick up the waves that result from a huge truck used to vibrate the ground. The waves are changed as they pass through the different substances in the ground. The analysis of these waves directs the oil companies to possible drilling sites.

There are many more uses for eigenvalues, but we only wanted to give you a sampling of their uses. When you study science or engineering in college, you will become quite familiar with eigenvalues and their uses. There are also numerical difficulties that can arise when data from real-world problems are used. Some of these difficulties are discussed in Chapter 10.

Problems

1. Use the power method to find the dominant eigenpair for the transpose of the matrix $A = \begin{bmatrix} .2 & .3 & .5 \\ .4 & .4 & .2 \\ .4 & .6 & 0 \end{bmatrix}$ which was used in problem 6 of Chapter 7. Give your solution accurate to two decimal places.

2. Use the power method to find the dominant eigenpair for the transpose of the matrix $A = \begin{bmatrix} .4 & .35 & .25 \\ .45 & .4 & .15 \\ .8 & .15 & .05 \end{bmatrix}$ which was used in problem 7 of Chapter 7. Give your solution accurate to two decimal places.

3. Use the power method to find the dominant eigenpair for $A = \begin{bmatrix} 1 & 0 \\ 4 & 8 \end{bmatrix}$.

4. Use the power method to find the dominant eigenpair for $A = \begin{bmatrix} 3 & 2.5 \\ 1 & 1.5 \end{bmatrix}$.

5. Use the characteristic equation to find all the eigenvalues of $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$ and a corresponding eigenvector for each.

6. Use the characteristic equation to find all the eigenvalues of $A = \begin{bmatrix} 5 & 0 \\ 2 & -2 \end{bmatrix}$ and a corresponding eigenvector for each.

7. Use the characteristic equation to find all the eigenvalues of $A = \begin{bmatrix} 3 & 7 \\ 2 & -2 \end{bmatrix}$ and a corresponding eigenvector for each.

8. If $\begin{bmatrix} -3 \\ 2 \\ -1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = 5$, find 3 more eigenvectors that correspond to $\lambda = 5$.

Chapter 10

Numerical Challenges

Since you have successfully worked so many problems with matrices, you may be wondering why people still need to study them. The methods for dealing with matrices that you have learned work nicely for most small matrices, but difficulties are encountered when the matrices are larger. Can you imagine solving by hand a system of equations in which A was of dimension 10 by 10? That would take a dreadfully long time. However, a computer can solve this system quickly. Does that mean that a computer can solve all systems? What if the system involved a matrix with dimensions of 100,000 by 100,000? The computer could have difficulty with this system. You might be thinking that this is a ridiculous size for a system of equations, but systems this size and larger are needed in most industrial applications. For example, the major airlines need systems of this size to schedule flights and crews. Also, large systems are needed when an oil or water reservoir is simulated to learn about how to best pump the liquid or perform an environmental cleanup.

One problem is that the system might be too large for the computer's memory to hold all the information at one time. This means that new algorithms must be found that require less of the problem to be stored in memory at one time. These new algorithms still need to arrive at the correct solution in a reasonable amount of time.

The amount of work (the number of individual additions, subtractions, multiplications, or divisions) needed to solve a system also depends on the size of the system (among other properties such as the percentage of zeros in the matrix and where the non-zero numbers are in the matrix). Even if the computer can solve this system, will it be able to do it in a reasonable amount of time? In many ways, Gaussian elimination is our method of choice for solving linear systems. Yet, even Gaussian elimination may take days, months, or even years to solve very large linear systems on the world's fastest computers. Researchers are working to find algorithms that will solve systems with less work. They are also working in the field of parallel computing in which the problem is broken down into parts. Parts that do not depend on one another can be sent to separate processing units so that those parts of the problem can be processed at the same time. Computers are already extremely fast, so the room that is left for improvement lies mostly in improving algorithms. There are also many problems that still cannot be solved. These are reasons why researchers are needed in the mathematical sciences.

Roundoff error is probably the biggest detriment to effective computation because it can affect small problems as well as large ones. If you try to turn the fraction $\frac{1}{3}$ into a decimal, you can carry out the approximation as far as you want. However, a computer has only a set number of digits that it can store for each number. Therefore, the fraction $\frac{1}{3}$ has to be rounded at some point. For a single number, this rounding is not too significant, but it becomes important when operations are performed with rounded numbers. Most computers store numbers in a manner similar to scientific notation called normal notation. In normal notation, a number is written as $m \times 10^a$ where $0.1 \leq |m| < 1$. (In scientific notation, $1 \leq |m| < 10$). In this equation, m stands for mantissa and a stands for abscissa. The number 1234 is represented as 0.1234×10^4. (Since most computers work in the binary system rather than the decimal system, this is not entirely accurate, but it will suffice to demonstrate roundoff error). The mantissa of this example is 0.1234. The abscissa is 4. For demonstration purposes, let us pretend that our computer stores 4-digit mantissas and 2-digit abscissas (the sign of the number is not included in the 4 digits of the mantissa on our fictional computer).

Let's add $10 + .00001$. The exact arithmetic solution is 10.00001. Our computer stores this as $0.1000 \times 10^2 + 0.1000 \times 10^{-4}$. When added, this yields 0.1000001×10^2. However, our computer will only store a 4 digit mantissa, so this rounds to 0.1000×10^2 which is the first addend of our problem. This means that our solution does not reflect that we added 0.1000×10^{-4} at all. Our solution is only off by 0.1000×10^{-4} which may not be much, but suppose that we want to add the number 0.1000×10^{-4} ten thousand times to the number 10 rather than just once. The exact solution is 11, but our computer still gives the solution as 10 because each time that we add 0.1000×10^{-4} to our solution, it doesn't make a difference. The source of this problem is that the two numbers that we are adding are of such different magnitudes. We can correct this problem by summing 0.1000×10^{-4} ten thousand times before we add the result to 10. This will yield a much closer solution. Roundoff problems can be minimized by paying attention to details like the order of magnitude of the numbers, but they cannot be eliminated.

Subtractive cancellation is another big problem when dealing with computers. When you subtract two numbers that are very close to each other, an error is introduced. Suppose we want to perform this operation on our fictional computer: $0.5555 - 0.5554$. Of course, the accurate solution is 0.0001, but that is not necessarily what the computer obtains. The computer represents this as $0.5555 \times 10^0 - 0.5554 \times 10^0$. The subtraction yields 0.0001×10^0, but the computer stores this number in normal form, so it should be stored as 0.1×10^{-3}. However, the computer fills all 4 digits of the mantissa. Since there is no accurate information for the other 3 digits of the mantissa, some computers fill the slots with random digits rather than zeros. Therefore, the solution is stored as $0.1ddd \times 10^{-3}$, where the d's are random digits.

Because of this error, we try to avoid subtracting numbers that are too close to each other.

We can also run into problems when we solve systems of equations. Let's look at the system

$$\begin{aligned} 0.0001x_1 + 10x_2 &= 10 \\ 10x_1 + 10x_2 &= 20 \end{aligned}$$

When we try to solve this system on our fictional computer, we run into problems.

$$\left[\begin{array}{cc|c} 0.0001 & 10 & 10 \\ 10 & 10 & 20 \end{array}\right] \begin{array}{l} \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & 100,000 & 100,000 \\ 10 & 10 & 20 \end{array}\right] \; r1 \div 0.0001$$

$$\left[\begin{array}{cc|c} 1 & 100,000 & 100,000 \\ 0 & -999,990 & -999,980 \end{array}\right] \; -10 * r1 + r2$$

$$\left[\begin{array}{cc|c} 1 & 100,000 & 100,000 \\ 0 & 0.99999 & 0.99998 \end{array}\right] \; r2 \div (-1,000,000)$$

$$\left[\begin{array}{cc|c} 1 & 100,000 & 100,000 \\ 0 & 1.000 & 1.000 \end{array}\right] \begin{array}{l} \text{Rounded to} \\ \text{4 digits of accuracy} \end{array}$$

$$x_2 = 1$$

$$x_1 + 100,000(1) = 100,000 \Rightarrow x_1 = 0$$

If you substitute this into the original equation, you can see that we definitely have a wrong solution. However, if we switch the order of the rows, we will get the right solution.

$$\left[\begin{array}{cc|c} 0.0001 & 10 & 10 \\ 10 & 10 & 20 \end{array}\right] \begin{array}{l} \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{cc|c} 10 & 10 & 20 \\ 0.0001 & 10 & 10 \end{array}\right] \begin{array}{l} \text{Switch the order} \\ \text{of the rows} \end{array}$$

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0.0001 & 10 & 10 \end{array}\right] \; r1 \div 10$$

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 9.9999 & 9.9998 \end{array}\right] \; -0.0001 * r1 + r2$$

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & .9999899999 \end{array}\right] \; r2 \div 9.9999$$

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 1.000 \end{array}\right] \begin{array}{l} \text{Rounded to} \\ \text{4 digits of accuracy} \end{array}$$

$$x_2 = 1$$

$$x_1 + 1(1) = 2 \Rightarrow x_1 = 1$$

You can check, by substituting into the original equation, that this yields the correct solution. The first time that we worked this problem, we magnified the rounding errors in our data because we divided by .0001, which is the same as multiplying by 10,000. This is what caused our error.

You have seen many ways that errors can be introduced into a solution. Although very little error is introduced in each step, the errors accumulate when calculations are performed with rounded data. At each step, the error could increase. Therefore, an error of only 0.1×10^{-6} per step could easily make our solution inaccurate if we perform enough steps before arriving at the solution.

10.1 Operation Counts

We mentioned earlier that we would be interested in the number of steps that it takes to solve a system with a particular algorithm. Since each step is an arithmetic operation, when we count the steps, we are counting the operations. Addition/subtraction steps are a lot quicker to compute than multiplication/division steps. Also, there are usually approximately the same number of addition/subtraction steps as multiplication/division steps in an algorithm. Therefore, since operation counts are only an approximation, we will only consider the multiplication/division steps.

We told you that Gaussian elimination requires fewer steps than Gauss-Jordan elimination, so we would like to show you how big of a difference that can make. Gaussian elimination requires approximately $\frac{n^3}{3}+n^2-\frac{n}{3}$ multiplication/division steps. Gauss-Jordan requires approximately $\frac{n^3}{2}+\frac{n^2}{2}$ multiplication/division steps (*Elementary Linear Algebra*, 1974, p. 38-39). Let's compute the approximate number of multiplcation steps needed for several values of n.

n	Gauss	Gauss − Jordan
2	6	6
3	17	18
4	36	40
5	65	75
6	106	126
10	430	550
100	343, 300	505, 000
1000	334, 333, 000	500, 500, 000
large n	$\frac{n^3}{3}$	$\frac{n^3}{2}$

The last line indicates an approximation for large n because the terms that have powers less than 3 are small compared to the third order term. For small systems, the two methods do not differ much, but the difference is drastic for large systems.

We could also look at the efficiency of Cramer's rule, but it depends on how we compute the determinant. We must compute $n+1$ determinants, but the method of computation makes a difference. The computation of the determinant for an n by n matrix using expansion by minors requires $n!$ multiplications/divisions because there are n minors required for the first expansion and each submatrix needed to determine the minor must also be expanded in the same manner until the submatrices are 3 by 3 or smaller. If we use expansion by minors to compute the determinant, then Cramer's rule requires approximately $(n+1)!$ multiplications/divisions. If we compute the determinant using Gaussian elimination, then Cramer's rule requires approximately $\frac{n^4}{3}$ multiplications/divisions because Gaussian elimination for each determinant requires approximately $\frac{n^3}{3}$ multiplications/divisions and $n+1$ determinants are needed. This is why we said that Cramer's rule is a theoretical tool, not a computational tool. Let's look at the number of multiplications/divisions required for several values of n for each method.

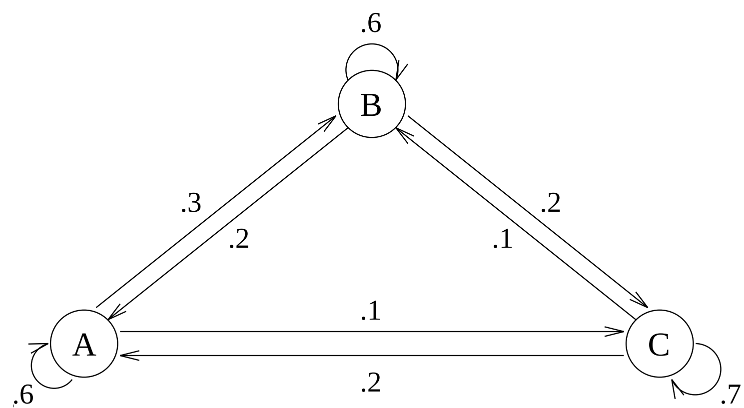
n	Expansion by Minors	Using Gaussian Elimination
2	2	6
3	6	27
4	24	86
5	120	208
6	720	432
10	$518,400$	3334
50	$3 * 10^{64}$	$2,083,333$

Factorials grow so quickly, that our calculators will not even store enough digits to compute the number of operations needed to use expansion by minors in Cramer's rule for $n = 100$. We certainly don't want to use an algorithm that requires that many operations when we have better options.

Although this chapter presents many problems that can occur when solving systems or using computers, it was not intended to discourage you. The methods that you learned in the previous chapters will work for most matrices that are not too large. However, you do need to be aware of some of the pitfalls of calculations so that you can avoid them. These problems are part of the reason for the national need for researchers in the field of computational mathematics. Hopefully, your eyes have been opened to an exciting possible career.

Second Review

1. A school district has 3 high schools. At the end of each year, teachers can be moved from one school to another. The picture below depicts the probability that a teacher will move from one school to another this year.



   (a) Construct and label the transition matrix that corresponds to this picture. Name the matrix A.

   (b) If a teacher works in school A this year, what is the probability that he or she will work in school C next year?

   (c) If a teacher works in school C this year, what is the probability that he or she will work in school B in the year after next (i.e., 2 years from now)?

   (d) Matrix A is the transition matrix for one year. Find the transition matrix for two years.

   (e) Find the transition matrix for three years.

   (f) Find, to two decimal places, the matrix to which A appears to converge after many years.

   (g) Explain the meaning of your solution to problem 1f.

2. (a) Using the following data and the normal equation, find the "best fit" straight line, accurate to one decimal place, to the points.

x	y
1	3
2	5
4	10
5	13
7	14

(b) Using the following data and the normal equation, find the "best fit" straight line, accurate to one decimal place, to the points.

x	y
-1	-4
1	1
2	4
3	7
4	11

3. (a) Using the following data and the normal equation, find the "best fit" parabola, accurate to one decimal place, to the points.

x	y
-3	8
-1	-3
0	-1
2	13
4	40

(b) Using the following data and the normal equation, find the "best fit" parabola, accurate to one decimal place, to the points.

x	y
-2	18
-1	6
0	0
2	10
4	40

4. (a) Use the power method to find the dominant eigenpair of the matrix A from problem 1.

(b) Use the power method to find the dominant eigenpair of the matrix
$$\begin{bmatrix} 5 & -6 \\ -2 & 1 \end{bmatrix}.$$

(c) Use the characteristic equation to find both of the eigenpairs of the matrix
$$\begin{bmatrix} 2 & -4 \\ 1 & -3 \end{bmatrix}.$$

Matrix Test 1A

1. On average, Heather spends 3 hours on homework, 4 hours watching TV, and 8 hours sleeping each day. Reena watches 1 hour of TV, sleeps 6 hours, and spends 5 hours doing homework. Edwin works on homework for an hour and sleeps for 9 hours. Mark sleeps 7 hours a day, watches 3 hours of TV, and spends 2 hours on homework.

   (a) Put the information into a 4 by 3 matrix and label it.

   (b) Transpose the matrix from problem 1a and attach labels.

   (c) If Edwin spends 5 hours a day with his computer and no one else works on a computer daily, convert the matrix from problem 1a into a 4 by 4 matrix and attach labels.

2. What is the sum $\begin{bmatrix} 4 & -1 \\ 7 & 2 \end{bmatrix} + \begin{bmatrix} -6 & 5 \\ -3 & -1 \end{bmatrix}$?

3. What is the product $\begin{bmatrix} 3 & -7 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 5 & -6 \\ 4 & 2 \end{bmatrix}$?

4. Find the inverse of the matrix $\begin{bmatrix} 4 & 6 \\ 1 & 2 \end{bmatrix}$

5. Give the two major steps needed to find the inverse of this matrix $\begin{bmatrix} 3 & 4 & 1 \\ 1 & 2 & 1 \\ -2 & 2 & 3 \end{bmatrix}$ OR actually find the inverse. (Only answer one of the questions. Both questions are worth the same number of points, so it doesn't matter which you answer).

6. Construct a symmetric matrix and explain why it is symmetric.

7. Solve this system using Gaussian elimination or Gauss-Jordan elimination and tell which method you used.

$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 11 \\ 4x_1 - x_2 + 2x_3 &= \ \ 5 \\ 3x_2 + 2x_3 &= 13 \end{aligned}$$

8. Find the determinant of the matrix $\begin{bmatrix} 7 & 2 & 5 & 3 \\ 1 & 0 & 2 & 0 \\ 8 & 0 & -2 & 4 \\ -3 & 0 & 4 & 1 \end{bmatrix}$.

9. Label each of these systems as consistent or inconsistent. If the system is consistent, further categorize it as underdetermined or uniquely determined. Explain why each system is categorized as it is.

   (a) $Ax = b$ where $A = \begin{bmatrix} 2 & 1 \\ 3 & -2 \end{bmatrix}$ and $b = \begin{bmatrix} 10 \\ 1 \end{bmatrix}$

   (b) $3x_1 + 4.5x_2 = 6$

   $2x_1 + \quad 3x_2 = 4$

Matrix Test 1B

1. On average, Gail spends 3 hours a day at her job, 2 hours studying, an hour watching TV, and 8 hours sleeping. Kerry spends 2 hours watching TV, 3 hours studying, 2 hours working, and 7 hours sleeping. Brad spends 5 hours a day working, 4 hours studying, and 6 hours sleeping.

   (a) Put the information into a 3 by 4 matrix and label it.

   (b) Transpose the matrix from problem 1a and attach labels.

   (c) Adam works 12 hours a day and sleeps 7 hours. Convert the matrix from problem 1a into a 4 by 4 matrix and attach labels.

2. What is the sum $\begin{bmatrix} 2 & 5 \\ 6 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix}$?

3. What is the product $\begin{bmatrix} 2 & -5 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} 7 & -4 \\ 0 & 5 \end{bmatrix}$?

4. Find the inverse of the matrix $\begin{bmatrix} 5 & 4 \\ -2 & -1 \end{bmatrix}$.

5. Give the two major steps needed to find the inverse of this matrix $\begin{bmatrix} 1 & 2 & 1 \\ 3 & -1 & 2 \\ 1 & 3 & 1 \end{bmatrix}$ OR actually find the inverse. (Only answer one of the questions. Both questions are worth the same number of points, so it doesn't matter which you answer).

6. Construct a symmetric matrix and explain why it is symmetric.

7. Solve this system using Gaussian elimination or Gauss-Jordan elimination and tell which method you used.

$$\begin{aligned} 3x_1 - x_2 + 4x_3 &= 11 \\ 2x_2 - x_3 &= 3 \\ x_1 - 3x_2 + 2x_3 &= -1 \end{aligned}$$

8. Find the determinant of the matrix $\begin{bmatrix} 3 & -4 & 0 & 3 \\ 2 & 5 & 0 & 5 \\ 5 & 1 & 2 & 6 \\ 1 & 0 & 0 & 2 \end{bmatrix}$.

9. Label each of these systems as consistent or inconsistent. If the system is consistent, further categorize it as underdetermined or uniquely determined. Explain why each system is categorized as it is.

   (a) $Ax = b$ where $A = \begin{bmatrix} 3 & 7.5 \\ 2 & 5 \end{bmatrix}$ and $b = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$

   (b) $5x_1 - x_2 = 9$

   $3x_1 + x_2 = 7$

Matrix Test 1C

1. Complete the following using matrices. When you create a matrix, label it. If a matrix does not clearly answer the question, please write a sentence to explain your answer.

   Shelly and Clint make dog houses. They call their company Canine Cabins. Shelly builds the walls. Clint builds the roof and floor. Then Clint attaches the roof and floor to the walls and Shelly paints it. Canine Cabins come in two sizes. The walls of a small house require 25 square feet of wood and 1 hour of labor. The roof and floor require 15 square feet of wood and an hour of labor. The walls of a large house require 70 square feet of wood and an hour of labor. The roof and floor of a large house use 30 square feet of wood and an hour of labor.

   (a) Compile this information into two 2 by 3 matrices (remember that we have not used paint yet, but we will). You will have one matrix for walls and one for roof and floor.

   (b) Using the matrices, determine how much wood, paint, and labor are used for each house before they are assembled if we build one large and one small house.

   (c) A completed small houses requires 40 square feet of wood, 1 pint of paint, and 3.5 hours of labor. A large house requires 100 square feet of wood, 1.5 pints of paint, and 4 hours labor. Form the 2 by 3 matrix that represents the total units of each material required to complete each size of house.

(d) Determine how much wood, paint, and time are required to attach the roof and floor to the walls and paint the house for each size.

(e) If wood costs \$0.50 per square foot, paint costs \$6 per pint, and Shelly and Clint earn \$8 per hour, how much does each size house cost to produce?

(f) This month, Canine Cabins had orders for 30 small houses and 20 large houses. How much of each material do they need to fill the orders?

(g) How much money will they need to buy the material to fill the orders?

2. If A is a 3 by 4 matrix, B is 2 by 4, and C is 2 by 3, list all the ways using (A or A^T), (B or B^T), and (C or C^T) that you can multiply these three matrices together. Each matrix or its transpose must be used exactly once in each multiplication.

3. Given

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & -1 & 2 \\ 3 & 1 & 2 \end{bmatrix} \text{ and } b = \begin{bmatrix} 2 \\ 11 \\ 9 \end{bmatrix}$$

(a) Find the determinate of A.

(b) Solve the system $Ax = b$ using Gaussian elimination or Gauss-Jordan elimination. Specify which you are using.

(c) Find the inverse of A.

(d) How would you prove that your solution to problem 3c is actually the inverse of A?

4. Create a matrix for which an inverse does not exist and explain why an inverse does not exist.

5. Create a system for each of the following descriptions and explain why the system fits the description.

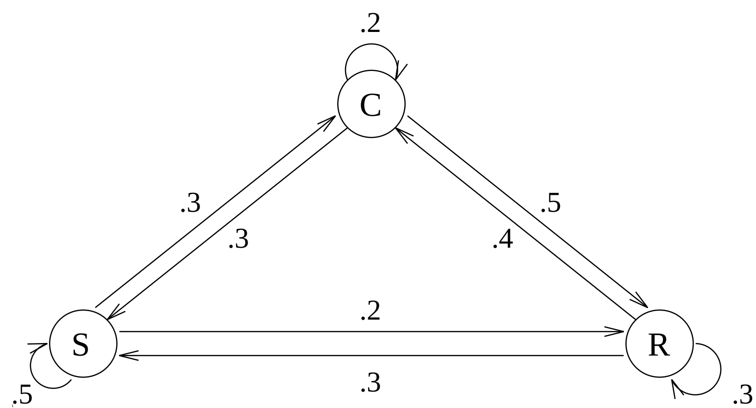
   (a) Consistent and underdetermined

   (b) Consistent and uniquely determined

   (c) Inconsistent

Matrix Test 2A

1. An amateur meteorologist predicts the weather based solely on today's weather. The picture below depicts the probability that the weather will either be sunny, cloudy, or rainy based on the weather today.



   (a) Construct and label the transition matrix that corresponds to this picture. Name the matrix A.

   (b) If the weather is rainy today, what is the probability that it will be cloudy tomorrow?

   (c) If it is sunny today, what is the probability that it will be cloudy the day after tomorrow (ie., 2 days from now)?

   (d) Matrix A is the transition matrix for one day. Find the transition matrix for two days.

   (e) Find the transition matrix for three days.

   (f) Find, to three decimal places, the matrix to which A appears to converges after many days.

   (g) Explain the meaning of your solution to problem 1f.

2. Using the following data and the normal equation, find the "best fit" straight line, accurate to one decimal place, to the points.

x	y
-2	4
0	1
1	0
2	3

3. Using the following data and the normal equation, find the "best fit" parabola, accurate to one decimal place, to the points.

x	y
-2	-4
0	5
1	0
2	-12

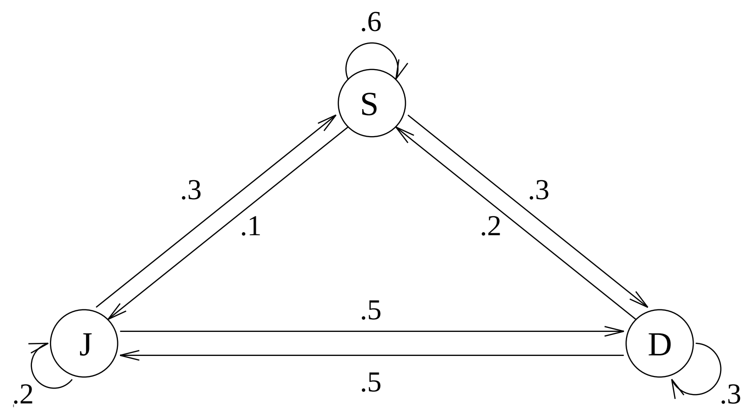
4. (a) Use the power method to find the dominant eigenpair of the matrix A from problem 1.

   (b) Use the power method to find the dominant eigenpair of the matrix $\begin{bmatrix} -6 & 2 \\ -3 & 1 \end{bmatrix}$.

   (c) Use the characteristic equation to find both of the eigenpairs of the matrix $\begin{bmatrix} -2 & -9 \\ -2 & 1 \end{bmatrix}$.

Matrix Test 2B

1. A very mathematical young lady decides what type of clothes that she will wear tomorrow based on what she wears today. The picture below depicts the probability that she will wear jeans, slacks, or a dress tomorrow based on what she wears today.



   (a) Construct and label the transition matrix that corresponds to this picture. Name the matrix A.

   (b) If she wears jeans today, what is the probability that she will wear a dress tomorrow?

   (c) If she wears a dress today, what is the probability that she will wear slacks the day after tomorrow (ie., 2 days from now)?

   (d) Matrix A is the transition matrix for one day. Find the transition matrix for two days.

   (e) Find the transition matrix for three days.

   (f) Find, to two decimal places, the matrix to which A appears to converge after many days.

(g) Explain the meaning of your solution to problem 1f.

2. Using the following data and the normal equation, find the "best fit" straight line, accurate to one decimal place, to the points.

x	y
-1	-8
1	1
2	7
4	19

3. Using the following data and the normal equation, find the "best fit" parabola, accurate to one decimal place, to the points.

x	y
-2	10
1	2
2	4
3	7

4. (a) Use the power method to find the dominant eigenpair of the matrix A from problem 1.

   (b) Use the power method to find the dominant eigenpair of the matrix
   $$\begin{bmatrix} -8 & 3 \\ -6 & 3 \end{bmatrix}.$$

   (c) Use the characteristic equation to find both of the eigenpairs of the matrix
   $$\begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}.$$

Solutions

12.1 Solutions to Introduction - Problems from page 8

1. (a) $B = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 3 & 6 & 9 & 12 & 15 \\ 4 & 8 & 12 & 16 & 20 \end{bmatrix}$

   (b) $B^T = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \\ 4 & 8 & 12 & 16 \\ 5 & 10 & 15 & 20 \end{bmatrix}$

   (c) No. One reason is that $B \neq B^T$. Another reason is that B is not square. Note: Only one of the reasons is necessary.

2. (a) matrices are easy

   (b) friends

   (c) calculator

   (d) The answers will vary.

3. (a)

	Won	Lost
Lions	5	8
Tigers	9	4
Bears	7	6

   Note: On all solutions of this type (including this and the next 2 problems), rows could be switched and/or columns may be switched. The only requirement is that each of these numbers lines up with both of its labels.

Therefore,

	Lost	Won
Lions	8	5
Bears	6	7
Tigers	4	9

is also an appropriate solution.

(b)

	L	T	B
Won	5	9	7
Lost	8	4	6

Note: Make sure that this matrix corresponds to the solution that the student gave on part (a) if the columns or rows were not as we ordered them in part (a).

4. (a)

	S	L	F	T
Won	6	8	9	5
Lost	8	7	4	9
Tied	1	0	2	1

(b)

	W	L	T
Snakes	6	8	1
Lizards	8	7	0
Frogs	9	4	2
Toads	5	9	1

Please refer to the notes on the previous problem.

5. (a)

$$\begin{array}{c} \\ \text{Amit} \\ \text{Perry} \\ \text{Don} \\ \text{Heather} \\ \text{Shelly} \end{array} \begin{array}{c} \begin{array}{ccc} \text{GPA} & \text{PSAT} & \text{SAT} \end{array} \\ \left[\begin{array}{ccc} 3.48 & 160 & 1580 \\ 3.65 & 121 & 1320 \\ 2.76 & 102 & 840 \\ 3.80 & 99 & 1260 \\ 3.01 & 83 & 980 \end{array} \right] \end{array}$$

(b)

$$\begin{array}{c} \\ \text{GPA} \\ \text{PSAT} \\ \text{SAT} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{Amit} & \text{Perry} & \text{Don} & \text{Heather} & \text{Shelly} \end{array} \\ \left[\begin{array}{ccccc} 3.48 & 3.65 & 2.76 & 3.80 & 3.01 \\ 160 & 121 & 102 & 99 & 83 \\ 1580 & 1320 & 840 & 1260 & 980 \end{array} \right] \end{array}$$

Please refer to the notes on problem 4.

Computer Program

We have written sample programs in QBasic and Pascal. These programs are not meant to be examples of perfect programming style, but they do work. The goal of the programming assignments is to help students express what they have learned about matrices. Programming makes students think about each step of a problem.

QBasic:

```
REM This program asks for a matrix to be input.
REM It prints the matrix and its transpose.
REM This program uses no commands that are specific to matrices.
CLS
PRINT ''This program prints your matrix and its transpose.''
INPUT ''Enter the dimensions of the matrix separated by a comma.'',m,n
DIM a(m,n)
PRINT
```

```
PRINT ``Please press enter after each element of the matrix. ''
PRINT ``Enter all the elements of one row before the next row. ''

REM This loop reads the matrix
FOR i=1 to m
   FOR j=1 to n
      INPUT a(i,j)
   NEXT j
NEXT i

REM This loop prints the matrix
PRINT
PRINT ``This is the matrix that you entered: ''
FOR i=1 to m
   FOR j=1 to n
      PRINT a(i,j),
   NEXT j
   PRINT
NEXT i

REM This loop prints the transpose of the matrix
PRINT
PRINT ``This is the transpose of your matrix: ''
FOR j=1 to n
   FOR i=1 to m
      PRINT a(i,j),
```

```
   NEXT i
   PRINT
NEXT j
END
```

Pascal:

```
Program intro(input, output);
{This program asks for a matrix to be input.
 It prints the matrix and its transpose.
 This program uses no commands that are specific to matrices.}

uses crt; {Necessary for some Pascal compilers}

type
   matrix=array[1..10,1..10] of real;

var
   m,n: integer; {dimensions of the matrices}
   a: matrix;  {matrix}

procedure readmatrix(var a:matrix; m,n:integer);
   var
      i,j: integer; {counters}

   begin {read}
```

```
   for i:=1 to m do
      begin {do}
      for j:=1 to n do
         read(a[i,j]);
         readln;
      end; {do}
   end; {read}

procedure writematrix(a:matrix; m,n:integer);
   var
      i,j: integer; {counters}

   begin {write}
   for i:=1 to m do
      begin {each line}
      writeln;
      for j:=1 to n do
         write(a[i,j]:6:2);
      end; {each line}
   writeln;
   end; {write}

procedure writetranspose(a:matrix; m,n:integer);
   var
      i,j: integer; {counters}
```

```
   begin {write}
   for j:=1 to n do
      begin {each line}
      writeln;
      for i:=1 to m do
         write(a[i,j]:6:2);
      end; {each line}
   writeln;
   end; {write}

begin{main program}
   clrscr;
   writeln('Enter the dimensions of the matrix ');
   writeln('separated by a space. Then hit return.');
   read(m,n);
   writeln('Enter your matrix.');
   writeln('Enter each element followed by a return.');
   writeln('Enter the first row before you go to the next row.');
   readmatrix(a,m,n);
   writeln('The matrix that you entered is :');
   writematrix(a,m,n);
   writeln('The transpose of your matrix is :');
   writetranspose(a,m,n);
   writeln('Press return to leave the program');
   readln;
end. {main program}
```

12.2 Solutions to Addition - Problems from page 15

1. (a) $A + C = \begin{bmatrix} 8 & 13 & 8 & 8 \\ 16 & 7 & 6 & 7 \\ 4 & 13 & 13 & 4 \end{bmatrix}$

   (b) $D + E = \begin{bmatrix} 7 & 11 & 5 \\ 18 & 13 & 2 \\ 6 & 6 & 7 \\ 7 & 13 & 6 \end{bmatrix}$

   (c) $F - D = \begin{bmatrix} 2 & -4 & -2 \\ -9 & -1 & 2 \\ 7 & 4 & 1 \\ -3 & -5 & 7 \end{bmatrix}$

   (d) $F + B =$not possible since the dimensions do not match.

   (e) $B - (A + C) = \begin{bmatrix} 0 & -6 & -4 & -8 \\ -7 & -1 & -4 & -2 \\ -3 & -9 & -6 & -2 \end{bmatrix}$

   (f) $D - (E + F) = \begin{bmatrix} -2 & -1 & 0 \\ 0 & -7 & -3 \\ -13 & -8 & -4 \\ 3 & -2 & -12 \end{bmatrix}$

   (g) $B + C - B = \begin{bmatrix} 1 & 9 & 0 & 2 \\ 7 & 4 & 6 & 5 \\ 3 & 8 & 7 & 1 \end{bmatrix}$

   (h) $A - D =$not possible since the dimensions do not match.

(i) $A + D^T = \begin{bmatrix} 14 & 13 & 8 & 13 \\ 15 & 8 & 2 & 8 \\ 4 & 6 & 10 & 4 \end{bmatrix}$

(j) $D + E - B^T = \begin{bmatrix} -1 & 2 & 4 \\ 11 & 7 & -2 \\ 2 & 4 & 0 \\ 7 & 8 & 4 \end{bmatrix}$

2. $\begin{bmatrix} -2 & 0 & 8 & -7 & 9 \\ -\frac{1}{2} & -5 & 6 & -4 & -1 \\ 2 & -10 & -3 & -13 & 7 \end{bmatrix}$

3. (a) $A + B =$

	W	L
Lions	12	13
Tigers	15	10
Bears	11	14

(b) Row three tells us that, in two years, the bears won 11 games and lost 14 games.

(c) $C - (A + B) =$

	W	L
Lions	8	6
Tigers	7	7
Bears	5	9

4. (a)

	FT	FG	T
Home	6	18	0
Visitor	5	16	3

(b) Home won 85 to 84.

5. Yes. This can be proven using the facts learned in the Questions and Answers in Chapters 1 and 2. The matrix $(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$.

6. Yes. The matrices A and B must both be square because they are symmetric. Therefore, if they have the same dimensions, then $A - B$ must also be square. A generic element of $A - B$ is $a_{ij} - b_{ij}$. Since A and B are symmetric, this element is the same as $a_{ji} - b_{ji}$ which is also a generic element for $(A - B)^T$ because $(A - B)^T = A^T - B^T$. Therefore, $A - B$ is symmetric.

Computer Program

QBasic:

```
REM This program adds and subtracts matrices
REM It uses no commands that are specific to matrices
CLS
PRINT ''This program will add or subtract matrices. ''
INPUT ''Enter the dimensions of the matrices separated by a comma.'',m,n
DIM a(m,n)
DIM b(m,n)

REM This loop reads the first matrix
PRINT ''Please press enter after each element of the matrix. ''
PRINT ''Enter all the elements of one row before the next row. ''
FOR i=1 to m
   FOR j=1 to n
      INPUT a(i,j)
   NEXT j
```

```
NEXT i

100 REM This allows the user to choose what to do next
INPUT ``Do you wish to 1) add 2) subtract or 3) end '',s
IF s=3 then goto 500

200 PRINT ``Enter the next matrix. ''
FOR i=1 to m
   FOR j=1 to n
      INPUT b(i,j)
   NEXT j
NEXT i
If s=1 then GOSUB 1000 ELSE IF s=2 THEN GOSUB 2000 ELSE IF s=3
             THEN GOTO 500 ELSE GOTO 100
REM   The previous line should be part of the line above it
PRINT ``The resulting matrix is: ''
FOR i=1 to m
   FOR j=1 to n
      PRINT a(i,j),
   NEXT j
   PRINT
NEXT i
PRINT
GOTO 100
500 PRINT ``Your final solution is: ''
FOR i=1 to m
```

```
   FOR j=1 to n
      PRINT a(i,j),
   NEXT j
   PRINT
NEXT i
END

1000 REM This adds matrices
FOR i=1 to m
   FOR i=1 to n
      a(i,j)=a(i,j)+b(i,j)
   NEXT j
NEXT i
RETURN

2000 REM This subtracts matrices
FOR i=1 to m
   FOR i=1 to n
      a(i,j)=a(i,j)-b(i,j)
   NEXT j
NEXT i
RETURN
```

Pascal:

```
Program add(input, output);
```

```
{This program adds and subtracts matrices.}

uses crt; {Necessary for some Pascal compilers}

type
   matrix=array[1..10,1..10] of real;

var
   choice : integer;  {for choosing operation}
   m,n: integer; {dimensions of the matrices}
   a,b: matrix;  {matrix}

procedure readmatrix(var a:matrix; m,n:integer);
   var
      i,j: integer; {counters}

   begin {read}
   for i:=1 to m do
      begin {do}
      for j:=1 to n do
         read(a[i,j]);
         readln;
      end {do}
   end; {read}

procedure writematrix(a:matrix; m,n:integer);
```

```
   var
      i,j: integer; {counters}

   begin {write}
   for i:=1 to m do
      begin {each line}
      writeln;
      for j:=1 to n do
         write(a[i,j]:6:2);
      end; {each line}
   writeln;
   end; {write}

procedure addmatrix(var a:matrix; b: matrix);
   var
      i,j: integer; {counters}

   begin {addmatrix}
   for i:=1 to m do
      for j:=1 to n do
         a[i,j]:=a[i,j]+b[i,j];
   end;{addmatrix}

procedure submatrix(var a:matrix; b: matrix);
   var
      i,j: integer; {counters}
```

```
   begin {submatrix}
   for i:=1 to m do
      for j:=1 to n do
         a[i,j]:=a[i,j]-b[i,j];
   end;{submatrix}

procedure menu{(var choice:char)};
   begin {menu}
   writeln;
   writeln('Do you wish to :');
   writeln('  1. Add');
   writeln('  2. Subtract');
   writeln('  3. End the program');
   readln(choice);
   end;{menu}

procedure operation(var a : matrix);
   begin {operation}
   writeln('Enter your second matrix.');
   writeln('Enter the elements of each row separated');
   writeln('by a space. Hit return at the end of each row.');
   readmatrix(b,m,n);
   if choice = 1 then
      addmatrix(a,b)
   else if choice = 2 then
```

```
      submatrix(a,b)
   else if choice = 3 then
      choice := 3
   else if ((choice <> 1) and (choice <> 2) and (choice <> 3)) then
      begin {else if}
      writeln('That was not a choice.');
      choice := 4;
      end;{else if}
   end;{operation}

begin{main program}
   clrscr;
   choice := 0;
   writeln('Enter the dimensions of the matrix ');
   writeln('separated by a space. Then hit return.');
   read(m,n);
   writeln('Enter your matrix.');
   writeln('Enter each element followed by a return.');
   writeln('Enter the first row before you go to the next row.');
   readmatrix(a,m,n);
   menu;
   while choice <> 3 do
      begin {while}
      operation(a);
      writeln('The resulting matrix is :');
      writematrix(a,m,n);
```

```
    menu
  end; {while}
  writeln('The final resulting matrix is :');
  writematrix(a,m,n);
  writeln('Press return to leave the program');
  readln;
end. {main program}
```

12.3 Solutions to Multiplication - Problems from page 30

1. (a) Symmetric

   (b) Multiply the matrix by the scalar 1.6

   (c) $\begin{bmatrix} 0 & 2464 & 2576 & 2192 \\ 2464 & 0 & 4464 & 4240 \\ 2576 & 4464 & 0 & 384 \\ 2192 & 4240 & 384 & 0 \end{bmatrix}$

2. (a) $4C = \begin{bmatrix} 20 & 12 & 24 \end{bmatrix}$

   (b) $AD = \begin{bmatrix} 94 \\ 66 \\ 52 \end{bmatrix}$

   (c) DA - The dimensions are wrong (3 by 1 multiplied by 3 by 3). The inside dimensions do not agree.

   (d) BC - The dimensions are wrong (3 by 3 multiplied by 1 by 3). The inside dimensions do not agree.

   (e) $3CB = \begin{bmatrix} 342 & 213 & 180 \end{bmatrix}$

(f) $C(A+B) = \begin{bmatrix} 160 & 161 & 136 \end{bmatrix}$

(g) $AB = \begin{bmatrix} 124 & 43 & 72 \\ 106 & 75 & 52 \\ 119 & 91 & 58 \end{bmatrix}$

(h) $BA = \begin{bmatrix} 64 & 112 & 72 \\ 49 & 93 & 58 \\ 59 & 154 & 100 \end{bmatrix}$

(i) $CAD = 980$

(j) DBC - The dimensions are wrong (3 by 1 x 3 by 3 x 1 by 3). The inside dimensions do not agree on either multiplication.

(k) $AD + (CB)^T = \begin{bmatrix} 208 \\ 137 \\ 112 \end{bmatrix}$

(l) $DC = \begin{bmatrix} 20 & 12 & 24 \\ 40 & 24 & 48 \\ 5 & 3 & 6 \end{bmatrix}$

(m) $CD = 50$

3. (a) $P = \begin{matrix} \text{T} \\ \text{Q} \\ \text{H} \end{matrix} \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix}$. Notice that this is a column vector.

(b) GP

(c) $GP = \begin{bmatrix} 78 \\ 84 \\ 74 \\ 67 \\ 84 \end{bmatrix}$

4. Yes. The matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, the vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and the scalar $c = c$.

Because the order of multiplication does not matter for real numbers,

$$\begin{aligned}
c\,(Ax) &\stackrel{?}{=} A\,(cx) \\
c\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &\stackrel{?}{=} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\left(c\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \\
c\left(\begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}\right) &\stackrel{?}{=} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\left(\begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix}\right) \\
\begin{bmatrix} ca_{11}x_1 + ca_{12}x_2 \\ ca_{21}x_1 + ca_{22}x_2 \end{bmatrix} &\stackrel{?}{=} \begin{bmatrix} a_{11}cx_1 + a_{12}cx_2 \\ a_{21}cx_1 + a_{22}cx_2 \end{bmatrix} \\
\begin{bmatrix} ca_{11}x_1 + ca_{12}x_2 \\ ca_{21}x_1 + ca_{22}x_2 \end{bmatrix} &\stackrel{?}{=} \begin{bmatrix} ca_{11}x_1 + ca_{12}x_2 \\ ca_{21}x_1 + ca_{22}x_2 \end{bmatrix}
\end{aligned}$$

Make sure that you tell the students that $c(Ax) = A(cx)$ for all matrices that have the correct dimensions for Ax to be multiplied.

5. $\begin{bmatrix} a \\ b \\ c \end{bmatrix}\left(\begin{bmatrix} d & e & f \end{bmatrix}\begin{bmatrix} g \\ h \\ k \end{bmatrix}\right)\begin{bmatrix} l & p & q \end{bmatrix}$ This grouping requires only 15 simple multiplications to find T.

6. Yes. $A^2 = \begin{bmatrix} 22 & 27 \\ 18 & 31 \end{bmatrix}$, so

$$\begin{aligned} 3A^2 - 2A &= 3\begin{bmatrix} 22 & 27 \\ 18 & 31 \end{bmatrix} - 2\begin{bmatrix} 4 & 3 \\ 2 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 66 & 81 \\ 54 & 93 \end{bmatrix} - \begin{bmatrix} 8 & 6 \\ 4 & 10 \end{bmatrix} \\ &= \begin{bmatrix} 58 & 75 \\ 50 & 83 \end{bmatrix} \end{aligned}$$

Computer Program

QBasic:

```
REM This program adds, subtracts and multiplies matrices
REM It uses no commands that are specific to matrices
CLS
DIM a(10,10)
DIM b(10,10)
DIM c(10,10)
PRINT ‘‘This program will add, subtract or multiply matrices whose’’
PRINT ‘‘dimensions are less than 10. ’’
PRINT
PRINT ‘‘Enter the dimensions of the first matrix’’
INPUT ‘‘ separated by a comma. ’’,m,n
REM This loop reads the first matrix
PRINT
PRINT ‘‘Please press enter after each element of the matrix. ’’
PRINT ‘‘Enter all the elements of one row before the next row. ’’
```

```
FOR i=1 to m
   FOR j=1 to n
      INPUT a(i,j)
   NEXT j
NEXT i

100 REM This allows the user to choose what to do next.
PRINT
INPUT ''Do you wish to 1)add 2) subtract 3) multiply or 4) end '',s
IF s=4 THEN GOTO 500
200 PRINT
INPUT ''Enter the dimensions of the next matrix. '',m2,n2
REM The next statements check to make sure the dimensions are correct
REM for the operation
IF (((s=1 or s=2) AND (m<<>m2 OR n<>n2) OR (s=3 AND n<>n2)) THEN
   PRINT ''The matrix dimensions are not correct''
   GOTO 200
END IF
PRINT
PRINT ''Enter the next matrix. ''
FOR i=1 to m2
   FOR j=1 to n2
      INPUT b(i,j)
   NEXT j
NEXT i
IF s=1 THEN GOSUB 1000 ELSE IF S=2 THEN GOSUB 2000 ELSE IF S=3
```

```
            THEN GOSUB 3000 ELSE IF S=4 THEN GOTO ELSE GOTO 100
REM  The previous line should be part of the one above it
PRINT
PRINT ``The resulting matrix is: ''
FOR i=1 to m
   FOR j=1 to n
      PRINT a(i,j),
   NEXT j
   PRINT
NEXT i
PRINT
GOTO 100
500 PRINT
PRINT ``Your final solution is: ''
FOR i=1 to m
   FOR j=1 to n
      PRINT a(i,j),
   NEXT j
   PRINT
NEXT i
END

1000 REM This adds matrices
FOR i=1 to m
   FOR j=1 to n
      a(i,j)=a(i,j) + b(i,j)
```

```
   NEXT j
NEXT i
RETURN

2000 REM This subtracts matrices
FOR i=1 to m
   FOR j=1 to n
      a(i,j)=a(i,j) - b(i,j)
   NEXT j
NEXT i
RETURN

3000 REM This multiplies matrices
FOR i=1 to m
   FOR j=1 to n
      c(i,j)= 0
      FOR k=1 to n
         c(i,j)=c(i,j) + a(i,k) $*$ b(k,j)
      NEXT k
   NEXT j
NEXT i
n=n2
FOR i=1 to m
   FOR j=1 to n
      a(i,j) = c(i,j)
   NEXT j
```

```
NEXT i
RETURN
```

Pascal:

```
Program mult(input, output);
{This program adds, subtracts, and multiplies matrices.}

uses crt;  {Necessary for some Pascal compilers}

type
   matrix=array[1..10,1..10] of real;

var
   choice : integer;  {for choosing operation}
   m,n,m2,n2 : integer; {dimensions of the matrices}
   a,b,c : matrix;  {matrix}

procedure readmatrix(var a:matrix; m,n:integer);
   var
      i,j: integer; {counters}

   begin {read}
   for i:=1 to m do
      begin {do}
      for j:=1 to n do
```

```
      read(a[i,j]);
    readln;
    end {do}
  end; {read}

procedure writematrix(a:matrix; m,n:integer);
  var
    i,j: integer; {counters}

  begin {write}
  for i:=1 to m do
    begin {each line}
    writeln;
    for j:=1 to n do
      write(a[i,j]:6:2);
    end; {each line}
  writeln;
  end; {write}

procedure addmatrix(var a:matrix; b: matrix);
  var
    i,j: integer; {counters}

  begin {addmatrix}
  for i:=1 to m do
    for j:=1 to n do
```

```
        a[i,j]:=a[i,j]+b[i,j];
   end;{addmatrix}

procedure submatrix(var a:matrix; b: matrix);
   var
      i,j: integer; {counters}

   begin {submatrix}
   for i:=1 to m do
      for j:=1 to n do
         a[i,j]:=a[i,j]-b[i,j];
   end;{submatrix}

procedure multmatrix(var a:matrix; b:matrix; m:integer;
                  var n:integer; m2,n2:integer);
   {The above row should be attached to the first row}
   var
      i,j,k: integer; {counters}
      c: matrix; {temporary matrix}

   begin {multmatrix}
   for i:=1 to m do
      for j:=1 to n2 do
         begin {inner product}
         c[i,j]:=0;
         for k:=1 to n do
```

```
          c[i,j]:=c[i,j]+a[i,k]*b[k,j];
        end; {inner product}
   n:=n2;
   for i:=1 to m do
      for j:=1 to n do a[i,j]:=c[i,j];
   end;{multmatrix}

procedure menu{(var choice:char)};
   begin {menu}
   writeln;
   writeln('Do you wish to :');
   writeln('  1. Add');
   writeln('  2. Subtract');
   writeln('  3. Multiply');
   writeln('  4. End the program');
   readln(choice);
   end;{menu}

procedure operation(var a : matrix);
   begin {operation}
   writeln('Enter the dimensions of the second matrix ');
   writeln('separated by a space. Then hit return.');
   read(m2,n2);
   if (((m<>m2) or (n<>n2)) and ((choice=1) or (choice=2))) then
      begin {if}
      writeln('The dimensions are not correct for that operation.');
```

```
   choice := 5;
   end {if m}
else if ((choice = 3) and (m2 <> n)) then
   begin {else if choice}
   writeln('The dimensions are not correct for multiplication.');
   choice := 5;
   end{elseif choice}
else if ((choice<>1) and (choice<>2) and (choice<>3)
      and (choice<>4)) then {this line should be attached
                             to the previous line}
   begin {else if}
   writeln('That was not a choice');
   choice := 5;
   end;{else if}
if (choice<>5) then
   begin {if}
   writeln('Enter your matrix.);
   writeln('Enter the elements of each row separated');
   writeln('by a space. Hit return at the end of each row.');
   readmatrix(b,m2,n2);
   end; {if}
if choice = 1 then
   addmatrix(a,b)
else if choice = 2 then
   submatrix(a,b)
else if choice = 3 then
```

```
      multmatrix(a,b,m,n,m2,n2);
   end;{operation}


begin{main program}
   clrscr;
   choice := 0;
   writeln('Enter the dimensions of the first matrix ');
   writeln('separated by a space. Then hit return.');
   read(m,n);
   writeln('Enter your matrix.);
   writeln('Enter each element followed by a return.');
   writeln('Enter the first row before you go to the next row.');
   readmatrix(a,m,n);
   menu;
   while choice <> 4 do
      begin {while}
      operation(a);
      writeln('The resulting matrix is :');
      writematrix(a,m,n);
      menu
   end; {while}
   writeln('The final resulting matrix is :');
   writematrix(a,m,n);
   writeln('Press return to leave the program');
   readln;
end. {main program}
```

12.4 Solutions to Systems of Equations - Problems from page 51

1. No, we do not know that A has an inverse.

2. I. This is because $AA^{-1} = I$, $A^{-1}A = I$, $I^T = I$ and $II = I$.

3. On these problems, we have provided the solution and one of the paths that leads to that solution. There are different paths, but only one solution.

   (a) $x = \begin{bmatrix} 1\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$

$$\left[\begin{array}{cc|c} 3 & 5 & 2 \\ 2 & 4 & 1 \end{array}\right] \begin{array}{c}\text{Original} \\ \text{Augmented Matrix}\end{array} \Rightarrow \left[\begin{array}{cc|c} 1 & 1\frac{2}{3} & \frac{2}{3} \\ 2 & 4 & 1 \end{array}\right] r1 \div 3$$

$$\left[\begin{array}{cc|c} 1 & 1\frac{2}{3} & \frac{2}{3} \\ 0 & \frac{2}{3} & -\frac{1}{3} \end{array}\right] -2 * r1 + r2 \Rightarrow \left[\begin{array}{cc|c} 1 & 1\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & -\frac{1}{2} \end{array}\right] r2 \div \frac{2}{3}$$

$$\left[\begin{array}{cc|c} 1 & 0 & 1\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \end{array}\right] -\frac{5}{3} * r2 + r1 \Rightarrow x = \begin{bmatrix} 1\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

   (b) $x = \begin{bmatrix} 21 \\ -5 \end{bmatrix}$

$$\left[\begin{array}{cc|c} 2 & 9 & -3 \\ 1 & 3 & 6 \end{array}\right] \begin{array}{c}\text{Original} \\ \text{Augmented Matrix}\end{array} \Rightarrow \left[\begin{array}{cc|c} 1 & 4.5 & -1.5 \\ 1 & 3 & 6 \end{array}\right] r1 \div 2$$

$$\left[\begin{array}{cc|c} 1 & 4.5 & -1.5 \\ 0 & -1.5 & 7.5 \end{array}\right] -1 * r1 + r2 \Rightarrow \left[\begin{array}{cc|c} 1 & 4.5 & -1.5 \\ 0 & 1 & -5 \end{array}\right] r2 \div (-1.5)$$

$$x_2 = -5 \Rightarrow x_1 + 4.5(-5) = -1.5 \Rightarrow x_1 = 21$$

(c) $x = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 6 & 3 & 3 \\ 3 & 8 & 5 & 4 \end{array}\right] \begin{array}{c} \text{Original} \\ \text{Augmented} \\ \text{Matrix} \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 2 & 2 & -2 \end{array}\right] \begin{array}{l} \\ -2r1 + r2 \\ -3r1 + r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & .5 & -.5 \\ 0 & 2 & 2 & -2 \end{array}\right] r2 \div 2 \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & .5 & -.5 \\ 0 & 0 & 1 & -1 \end{array}\right] \begin{array}{l} \\ \\ -2 * r2 + r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \begin{array}{l} -1 * r3 + r1 \\ -.5 * r3 + r2 \\ \\ \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \begin{array}{l} -2 * r2 + r1 \\ \\ \\ \end{array}$$

(d) $x = \begin{bmatrix} 3 \\ -3 \\ 6 \end{bmatrix}$

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 3 \\ 1 & 3 & 3 & 12 \\ 3 & 3 & 1 & 6 \end{array}\right] \begin{array}{c} \text{Original} \\ \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 1.5 & 0.5 & 1.5 \\ 1 & 3 & 3 & 12 \\ 3 & 3 & 1 & 6 \end{array}\right] \begin{array}{l} r1 \div 2 \\ \\ \\ \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 1.5 & 0.5 & 1.5 \\ 0 & 1.5 & 2.5 & 10.5 \\ 0 & -1.5 & -0.5 & 1.5 \end{array}\right] \begin{array}{l} \\ -1 * r1 + r2 \\ -3 * r1 + r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 1.5 & 0.5 & 1.5 \\ 0 & 1 & 1\frac{2}{3} & 7 \\ 0 & -1.5 & -0.5 & 1.5 \end{array}\right] \quad r2 \div 1.5$$

$$\left[\begin{array}{ccc|c} 1 & 1.5 & 0.5 & 1.5 \\ 0 & 1 & 1\frac{2}{3} & 7 \\ 0 & 0 & 2 & 12 \end{array}\right] \quad 1.5 * r2 + r3$$

$$\left[\begin{array}{ccc|c} 1 & 1.5 & 0.5 & 1.5 \\ 0 & 1 & 1\frac{2}{3} & 7 \\ 0 & 0 & 1 & 6 \end{array}\right] \quad r3 \div 2$$

$x_3 = 6$

$x_2 + 1\frac{2}{3}(6) = 7 \Rightarrow -3$

$x_1 + 1.5(-3) + 0.5(6) = 1.5 \Rightarrow 3$

4. (a) $\begin{bmatrix} \frac{4}{5} & \frac{-3}{5} \\ -1 & 1 \end{bmatrix}$ Use the formula that we found for 2 by 2 matrices.

(b) $\begin{bmatrix} -1\frac{3}{5} & \frac{3}{5} \\ 1\frac{2}{5} & \frac{-2}{5} \end{bmatrix}$ Use the formula that we found for 2 by 2 matrices.

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

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 1 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \quad \text{Original Augmented Matrix}$$

$$\left[\begin{array}{ccc|ccc} 2 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \begin{array}{c} \text{Switch r1 and r2} \\ \text{so that 0 is not} \\ \text{a pivot} \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right] r1 \div 2$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -0.5 & 0 & -1.5 & 1 \end{array}\right] -3 * r1 + r3$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 & -2 \end{array}\right] r3 \div (-.5)$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 & -2 \end{array}\right] -0.5 * r3 + r1$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 & -2 \end{array}\right] -1 * r2 + r1$$

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

5. (a) Live

   (b) Love

   (c) Learn

(d) Answers will vary

6. $Ax = b$. Therefore, if A^{-1} exists (which it does for the stated problems), $AA^{-1}x = A^{-1}b$. Therefore, $x = A^{-1}b$.

(a) $\begin{bmatrix} -1 \\ 5 \end{bmatrix}$

(b) $\begin{bmatrix} \frac{-2}{5} \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 4\frac{3}{5} \\ -3\frac{2}{5} \end{bmatrix}$

(d) $\begin{bmatrix} -2\frac{4}{5} \\ 2\frac{1}{5} \end{bmatrix}$

(e) $\begin{bmatrix} -6 \\ 3 \\ 10 \end{bmatrix}$

(f) $\begin{bmatrix} -2 \\ -2 \\ 9 \end{bmatrix}$

(g) $\begin{bmatrix} 0 \\ 5 \\ -12 \end{bmatrix}$

7. $A = (A^{-1})^{-1}$ since $AA^{-1} = I$ and $A^{-1}A = I$, A and A^{-1} are inverses of each other. Therefore, the inverse of A^{-1} is A.

8. (a) $\begin{bmatrix} \frac{1}{a} & 0 & 0 \\ 0 & \frac{1}{b} & 0 \\ 0 & 0 & \frac{1}{c} \end{bmatrix}$

   (b) If A is a diagonal matrix consisting of the elements a_{ii}, then A^{-1} is a diagonal matrix consisting of the elements $\frac{1}{a_{ii}}$.

12.5 Solutions to Determinants - Problems from page 64

1. (a) -10 $\quad -6-4=-10$

   (b) 36 $\quad 30-(-6)=36$

   (c) -211 $\quad (18+0-168)-(40+0+21)=-211$

   (d) 153 $\quad (0+75+14)-(0-84+20)=153$

   (e) -2 $\quad (0+0+0)-(27+0-25)=-2$

   (f) -44 $\quad (-1)(-2)\begin{vmatrix} -4 & 2 & 1 \\ 5 & 0 & -2 \\ 2 & -1 & 0 \end{vmatrix} + (-1)(2)\begin{vmatrix} 3 & 7 & 6 \\ -4 & 2 & 1 \\ 2 & -1 & 0 \end{vmatrix}$ is one method.

   (g) -510 $\quad (+1)(-3)\begin{vmatrix} 9 & 3 & 1 \\ -2 & 6 & -4 \\ 2 & -1 & 4 \end{vmatrix}$

   (h) 151 $\quad (-1)(2)\begin{vmatrix} 3 & 1 & 0 \\ -5 & 6 & 2 \\ 8 & -3 & 1 \end{vmatrix} + (+1)(-1)\begin{vmatrix} 7 & 3 & 1 \\ -2 & -5 & 6 \\ 0 & 8 & -3 \end{vmatrix}$

(i) -402 $\quad 2\begin{vmatrix} 1 & 2 & 0 & 4 \\ 6 & 9 & 3 & 7 \\ 3 & 5 & 0 & 1 \\ 6 & 4 & 0 & 3 \end{vmatrix} = 2\left(3\begin{vmatrix} 1 & 2 & 4 \\ 3 & 5 & 1 \\ 6 & 4 & 3 \end{vmatrix}\right)$

2. (a) $x_1 = 1 \quad x_2 = -1 \quad x_3 = 2$

$$A = \text{Denominator} = \begin{vmatrix} 2 & 3 & -5 \\ -4 & -1 & 3 \\ 3 & -2 & 1 \end{vmatrix} = -6$$

$$B_1 = \text{Numerator of } \mathrm{x}_1 = \begin{vmatrix} -11 & 3 & -5 \\ 3 & -1 & 3 \\ 7 & -2 & 1 \end{vmatrix} = -6$$

$$B_2 = \text{Numerator of } \mathrm{x}_2 = \begin{vmatrix} 2 & -11 & -5 \\ -4 & 3 & 3 \\ 3 & 7 & 1 \end{vmatrix} = 6$$

$$B_3 = \text{Numerator of } \mathrm{x}_3 = \begin{vmatrix} 2 & 3 & -11 \\ -4 & -1 & 3 \\ 3 & -2 & 7 \end{vmatrix} = -12$$

(b) $x_1 = 2 \quad x_2 = 1 \quad x_3 = -1$

$$A = \text{Denominator} = \begin{vmatrix} 1 & -5 & 7 \\ 0 & 9 & 2 \\ 1 & 3 & -1 \end{vmatrix} = -88$$

$$B_1 = \text{Numerator of } \mathrm{x}_1 = \begin{vmatrix} -10 & -5 & 7 \\ 7 & 9 & 2 \\ 6 & 3 & -1 \end{vmatrix} = -176$$

$$B_2 = \text{Numerator of } x_2 = \begin{vmatrix} 1 & -10 & 7 \\ 0 & 7 & 2 \\ 1 & 6 & -1 \end{vmatrix} = -88$$

$$B_3 = \text{Numerator of } x_3 = \begin{vmatrix} 1 & -5 & -10 \\ 0 & 9 & 7 \\ 1 & 3 & 6 \end{vmatrix} = 88$$

3. $A^{-1} = \begin{bmatrix} \frac{1}{5} & \frac{-1}{25} & \frac{8}{25} \\ \frac{1}{5} & \frac{4}{25} & \frac{-7}{25} \\ \frac{-1}{5} & \frac{11}{25} & \frac{-13}{25} \end{bmatrix}$

4. (a) True

$$\begin{aligned}
A &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \\
\det(A) &= ad - bc \qquad \det(B) = eh - fg \\
\det(A)\det(B) &= adeh - adfg - bceh + bcfg \\
AB &= \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \\
\det(AB) &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\
&= acef + adeh + bcfg + bdgh - acef - adfg - bceh - bdgh \\
&= adeh + bcfg - adfg - bceh
\end{aligned}$$

(b) True

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$\det(A) = ad - bc$

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

$$\det(A^{-1}) = \frac{ad-bc}{(ad-bc)^2}$$

$$= \frac{1}{ad-bc} \text{ because } ad-bc \neq 0 \text{ since } A^{-1} \text{ exists.}$$

(c) False

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \qquad A+B = \begin{bmatrix} 2 & 2 \\ 3 & 4 \end{bmatrix}$$

$$\det(A) = -1 \qquad \det(B) = 2 \qquad \det(A+B) = 2 \qquad -1+2 \neq 2$$

This is only one example. The students' answers will vary.

(d) True

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

$$\det(A) = ad - cb \qquad \det(A^T) = ad - bc$$

These are equal because $cb = bc$ in scalar multiplication

5. Use a generic matrix and expansion by minors.

$$A = \begin{bmatrix} a & b & c & d \\ 0 & e & f & g \\ 0 & 0 & h & k \\ 0 & 0 & 0 & p \end{bmatrix}$$

$$\det(A) = a \begin{vmatrix} e & f & g \\ 0 & h & k \\ 0 & 0 & p \end{vmatrix} = ae \begin{vmatrix} h & k \\ 0 & p \end{vmatrix} = aeh \begin{vmatrix} p \end{vmatrix} = aehp$$

12.6 Solutions to Consistent and Inconsistent - Problems from page 74

1. (a) Inconsistent

   The last line of the augmented matrix after EROs states that $0x_1 + 0x_2 \neq -0.5$.

   These are parallel lines. No real numbers solve this system.

   (b) Consistent and underdetermined

   The last line of the augmented matrix after EROs states that $0x_1+0x_2 = 0$.

   There are an infinite number of solutions along a line because both equations express the same line.

   (c) Consistent and uniquely determined

   A solution can be obtained.

   This represents a single point where two lines intersect.

   (d) Consistent and underdetermined

   The last line of the augmented matrix after EROs states that $0x_1 + 0x_2 + 0x_3 = 0$.

   This represents an infinite number of solutions along a line where two planes intersect. The third plane is a linear combination of the other two planes, so it gives us no further information.

   (e) Consistent and uniquely determined

   A solution can be obtained.

   This represents a single point where three planes intersect.

   (f) Consistent and underdetermined

The last two lines of the augmented matrix after EROs states that $0x_1 + 0x_2 + 0x_3 = 0$.

This system has an infinite number of solutions in a plane. All three equations represent the same plane.

(g) Inconsistent

The last line of the augmented matrix after EROs states that $0x_1 + 0x_2 + 0x_3 \neq -0.5$.

This represents parallel planes. No real numbers solve this system because no real numbered solution lies in the intersection of these three planes. This tells us that at least two of the planes must be parallel to one another. Note that they do not all three have to be parallel. Even if two planes intersect, if the third plane is parallel to one of the first two planes, there will be no point that lies in all three planes.

2. Those that are consistent and uniquely determined - c and e.

12.7 Solutions to First Review from page 75

This review is quite long, but it provides a thorough review of the information covered. You might want to give the students a few days to work on this assignment. If they can work all these problems, then the test should not be difficult.

1. 2 by 3

2. $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix}$

3. (a)

$$\begin{array}{c} \\ \text{Test} \\ \text{HW} \end{array} \begin{array}{c} \begin{array}{ccc} \text{K} & \text{J} & \text{Y} \end{array} \\ \begin{bmatrix} 94 & 75 & 70 \\ 99 & 80 & 90 \end{bmatrix} \end{array}$$

(b)

$$\begin{array}{c} \\ \text{Keith} \\ \text{Juan} \\ \text{Yolanda} \end{array} \begin{array}{c} \begin{array}{cc} \text{T} & \text{HW} \end{array} \\ \begin{bmatrix} 94 & 99 \\ 75 & 80 \\ 70 & 90 \end{bmatrix} \end{array}$$

4. (a) $A + B = \begin{bmatrix} 13 & 1 & 15 \\ 8 & -1 & 6 \\ 7 & 14 & -1 \end{bmatrix}$

(b) $A^T + B = \begin{bmatrix} 13 & -2 & 16 \\ 11 & -1 & 3 \\ 6 & 17 & -1 \end{bmatrix}$

5. Yes because $B = B^T$

6. (a) $W1 + W2 = \begin{array}{c} \\ \text{Fresh.} \\ \text{Soph.} \\ \text{Jr.} \\ \text{Sr.} \end{array} \begin{array}{c} \begin{array}{cc} \text{N} & \text{P} \end{array} \\ \begin{bmatrix} 700 & 650 \\ 500 & 600 \\ 650 & 500 \\ 450 & 450 \end{bmatrix} \end{array}$

(b) Freshmen

(c) $T - (W1 + W2) = \begin{array}{c} \\ \text{Fresh.} \\ \text{Soph.} \\ \text{Jr.} \\ \text{Sr.} \end{array} \begin{array}{c} \begin{array}{cc} \text{N} & \text{P} \end{array} \\ \begin{bmatrix} 300 & 200 \\ 200 & 150 \\ 250 & 200 \\ 150 & 150 \end{bmatrix} \end{array}$

(d) 2900

(e) $\begin{bmatrix} 1000 & 850 \\ 700 & 750 \\ 900 & 700 \\ 600 & 600 \end{bmatrix} \begin{bmatrix} 0.3 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 470 \\ 360 \\ 410 \\ 300 \end{bmatrix}$ These solutions are given in dollars because we represented 30 cents as 0.3.

(f) \$1540

7. (a) No because the inside dimensions of 2 by 3 multiplied by 2 by 3 do not match.

(b) $AB^T = \begin{bmatrix} 54 & 87 \\ -5 & 15 \end{bmatrix}$, $BA^T = \begin{bmatrix} 54 & -5 \\ 87 & 15 \end{bmatrix}$, $BB^T = \begin{bmatrix} 65 & 0 \\ 0 & 90 \end{bmatrix}$. Remember that the students were only asked for 2 of these.

8. $AB = \begin{bmatrix} 48 & 121 & 101 \\ -49 & -20 & 5 \\ 55 & 90 & 121 \end{bmatrix}$

9. $x = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$

$$\left[\begin{array}{ccc|c} 1 & 3 & -1 & -2 \\ 2 & 3 & 1 & 2 \\ 3 & 6 & 1 & 3 \end{array}\right] \text{Original Matrix} \Longrightarrow \left[\begin{array}{ccc|c} 1 & 3 & -1 & -2 \\ 0 & -3 & 3 & 6 \\ 0 & -3 & 4 & 9 \end{array}\right] \begin{array}{c} \\ -2*r1+r2 \\ -3*r1+r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & -1 & -2 \\ 0 & 1 & -1 & -2 \\ 0 & -3 & 4 & 9 \end{array}\right] \; r2 \div (-3) \; \Longrightarrow \left[\begin{array}{ccc|c} 1 & 3 & -1 & -2 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right] \begin{array}{c} \\ \\ 3*r2+r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \begin{array}{c} r3+r1 \\ r3+r2 \\ \\ \end{array} \Longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \begin{array}{c} -3*r2+r1 \\ \\ \\ \end{array}$$

10. $x = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$

$$\left[\begin{array}{ccc|c} 4 & 2 & -1 & -8 \\ 3 & -1 & 2 & -3 \\ 1 & 0 & 5 & 8 \end{array}\right] \begin{array}{c} \text{Original} \\ \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 0.5 & -0.25 & -2 \\ 3 & -1 & 2 & -3 \\ 1 & 0 & 5 & 8 \end{array}\right] \begin{array}{c} r1 \div 4 \\ \\ \\ \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 0.5 & -0.25 & -2 \\ 0 & -2.5 & 2.75 & 3 \\ 0 & -0.5 & 5.25 & 10 \end{array}\right] \begin{array}{c} \\ -3*r1+r2 \\ -1*r1+r3 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 0.5 & -0.25 & -2 \\ 0 & 1 & -1.1 & -1.2 \\ 0 & -0.5 & 5.25 & 10 \end{array}\right] \quad r2 \div (-2.5)$$

$$\left[\begin{array}{ccc|c} 1 & 0.5 & -0.25 & -2 \\ 0 & 1 & -1.1 & -1.2 \\ 0 & 0 & 4.7 & 9.4 \end{array}\right] \quad 0.5 * r2 + r3$$

$$\left[\begin{array}{ccc|c} 1 & 0.5 & -0.25 & -2 \\ 0 & 1 & -1.1 & -1.2 \\ 0 & 0 & 1 & 2 \end{array}\right] \quad r3 \div 4.7$$

$$x_3 = 2$$

$$x_2 - 1.1(2) = -1.2 \Rightarrow x_2 = 1$$

$$x_1 + 0.5(1) - 0.25(2) = -2 \Rightarrow x_1 = -2$$

11. $\begin{bmatrix} \frac{1}{5} & 0 & -\frac{4}{5} \\ \frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} & \frac{3}{5} \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 1 & 8 & 4 & 1 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{c} \text{Original} \\ \text{Augmented} \\ \text{Matrix} \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 8 & 4 & 1 & 0 & 0 \\ 0 & -15 & -5 & -2 & 1 & 0 \\ 0 & 10 & 5 & 1 & 0 & 1 \end{array}\right] \quad \begin{array}{c} \\ -2 * r1 + r2 \\ r1 + r3 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 8 & 4 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{2}{15} & -\frac{1}{15} & 0 \\ 0 & 10 & 5 & 1 & 0 & 1 \end{array}\right] \quad r2 \div (-15)$$

$$\left[\begin{array}{ccc|ccc} 1 & 8 & 4 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{2}{15} & -\frac{1}{15} & 0 \\ 0 & 0 & 1\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} & 1 \end{array}\right] -10*r2+r3$$

$$\left[\begin{array}{ccc|ccc} 1 & 8 & 4 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{2}{15} & -\frac{1}{15} & 0 \\ 0 & 0 & 1 & -\frac{1}{5} & \frac{2}{5} & \frac{3}{5} \end{array}\right] r3 \div 1\frac{2}{3}$$

$$\left[\begin{array}{ccc|ccc} 1 & 8 & 0 & 1\frac{4}{5} & -1\frac{3}{5} & -2\frac{2}{5} \\ 0 & 1 & 0 & \frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & -\frac{1}{5} & \frac{2}{5} & \frac{3}{5} \end{array}\right] \begin{array}{l} -4*r3+r1 \\ -\frac{1}{3}*r3+r2 \\ \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{5} & 0 & -\frac{4}{5} \\ 0 & 1 & 0 & \frac{1}{5} & -\frac{1}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & -\frac{1}{5} & \frac{2}{5} & \frac{3}{5} \end{array}\right] -8*r2+r1$$

12. No. If the matrix is inconsistent or underdetermined, an inverse will not exist because there is not a matrix such that $AA^{-1} = A^{-1}A = I$.

13. (a) -2 $\quad (-8) - (-6) = -2$

(b) 140 $\quad (0 + 216 - 12) - (0 - 8 + 72) = 140$

(c) 940 $\quad 2\begin{vmatrix} -5 & 4 & 4 \\ 2 & 1 & 6 \\ 3 & 7 & 2 \end{vmatrix} + \begin{vmatrix} 7 & 9 & -6 \\ -5 & 4 & 4 \\ 3 & 7 & 2 \end{vmatrix} = 940$

(d) -120 $\quad 2\begin{vmatrix} 6 & 6 & 0 & 1 \\ -9 & -8 & 4 & 7 \\ 0 & 3 & 0 & 0 \\ 1 & 10 & 0 & 1 \end{vmatrix} = 2(3)\begin{vmatrix} 6 & 0 & 1 \\ -9 & 4 & 7 \\ 1 & 0 & 1 \end{vmatrix} = -120$

14. (a) Consistent and underdetermined because row 3 is a combination of the other two rows (because row 3 after elimination becomes all zeros).

    (b) Consistent and uniquely determined because there is a single solution to the system. The matrix A is invertible.

    (c) Inconsistent because after elimination, the last row is an unsolvable equation.

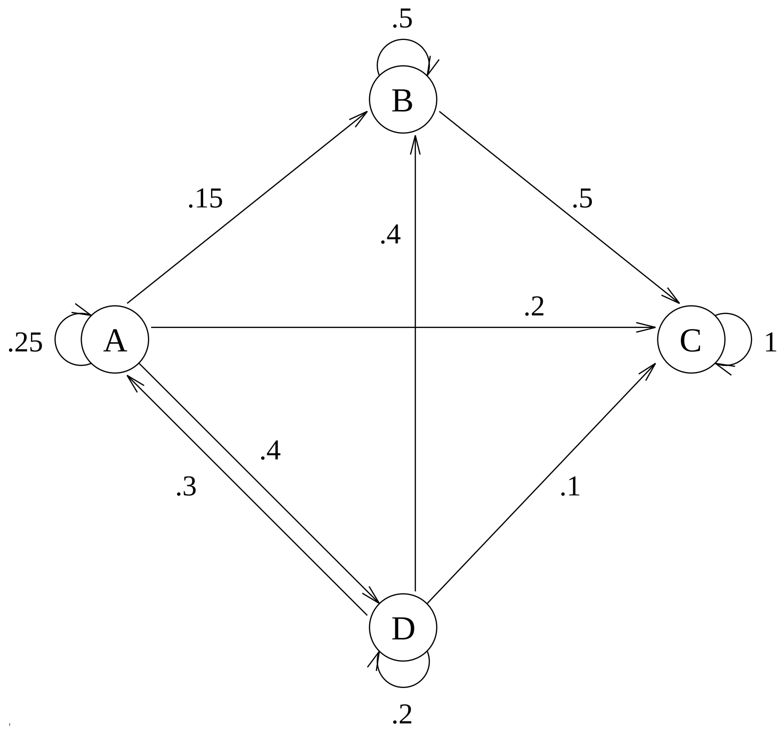
    (d) Consistent and uniquely determined because there is a single solution to the system. The matrix A is invertible.

12.8 Solutions to Markov Chains - Problems from page 90

Many of these problems require tedious multiplications. You might want to consider letting the students use calculators on this and the following chapters.

1.
$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \left[\begin{array}{ccc} .7 & .2 & .1 \\ .3 & .4 & .3 \\ 0 & .8 & .2 \end{array}\right] \end{array}$$

2. (a)



   (b) Information that flows into C never leaves. Since matter has a way to get to C from A and B, and because matter can never leave C once it flows into C, eventually, C should absorb all the matter in the system.

3. (a) Yes. Each element of the transition matrix is a probability. The elements of each row sum to 1. The matrix has a row and a column for each state.

   (b) No. $.6 + .1 + .2 = .9 \neq 1$

   (c) No. $.25 + .15 + .3 + .4 = 1.1 \neq 1$

4. (a) No. The probability of leaving B (including the chance of immediately returning to B) is less than 1.

   (b) Yes.

5. $$\begin{array}{r} \\ \text{Sandwich} \\ \text{Hamburger} \\ \text{Pizza} \end{array} \begin{array}{c} \begin{array}{ccc} \text{S} & \text{H} & \text{P} \end{array} \\ \begin{bmatrix} .2 & .2 & .6 \\ .5 & .2 & .3 \\ .4 & .3 & .3 \end{bmatrix} \end{array}$$

The students might arrange the rows and, therefore the columns, in a different order.

6. (a) 0.3

(b) 0.28

(c) 0.336

(d) $S_2 = S^2 = \begin{bmatrix} .36 & .48 & .16 \\ .32 & .40 & .28 \\ .32 & .36 & .32 \end{bmatrix}$

(e) $S_3 = S^3 = S^2 S = \begin{bmatrix} .328 & .396 & .276 \\ .336 & .424 & .240 \\ .336 & .432 & .232 \end{bmatrix}$

(f) $S_4 = S^4 s^3 S = \begin{bmatrix} .3344 & .4224 & .2432 \\ .3328 & .4144 & .2528 \\ .3328 & .4128 & .2544 \end{bmatrix}$

(g) $\begin{bmatrix} .33 & .42 & .25 \\ .33 & .42 & .25 \\ .33 & .42 & .25 \end{bmatrix}$

7. (a) $$A = \begin{array}{r} \\ \text{Full} \\ \text{Partial} \\ \text{None} \end{array} \begin{array}{c} \begin{array}{ccc} \text{F} & \text{P} & \text{N} \end{array} \\ \begin{bmatrix} .4 & .35 & .25 \\ .45 & .4 & .15 \\ .8 & .15 & .05 \end{bmatrix} \end{array}$$

The students might arrange the rows and, therefore the columns, in a different order.

(b) 0.4

(c) 0.05

(d) 0.18

(e) $A^2 = \begin{matrix} F \\ P \\ N \end{matrix} \begin{bmatrix} .5175 & .3175 & .165 \\ .48 & .34 & .18 \\ .4275 & .3475 & .225 \end{bmatrix}$

(f) $A^3 = \begin{matrix} F \\ P \\ N \end{matrix} \begin{bmatrix} .481875 & .332875 & .18525 \\ .489 & .331 & .18 \\ .507375 & .322375 & .17025 \end{bmatrix}$

(g) 0.18

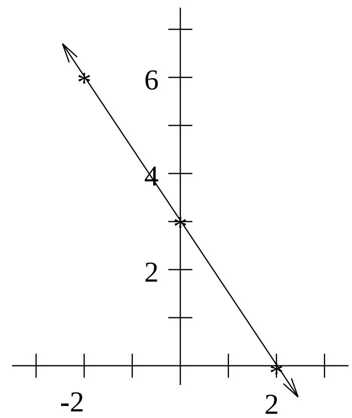
(h) $\begin{matrix} F \\ P \\ N \end{matrix} \begin{bmatrix} .49 & .33 & .18 \\ .49 & .33 & .18 \\ .49 & .33 & .18 \end{bmatrix}$

(i) If we are looking far enough into the future (a few weeks or longer), it doesn't matter what kind of assignment we have today. We have a 49% chance of having a full assignment, a 33% chance of having a partial assignment and an 18% chance of not having an assignment.

12.9 Solutions to Least Squares - Problems from page 110

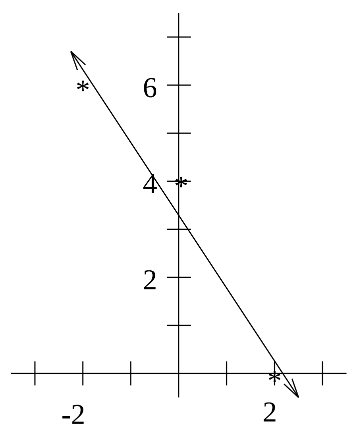
1. (a) $y = 3 - 1.5x$

Sum of squared errors = 0



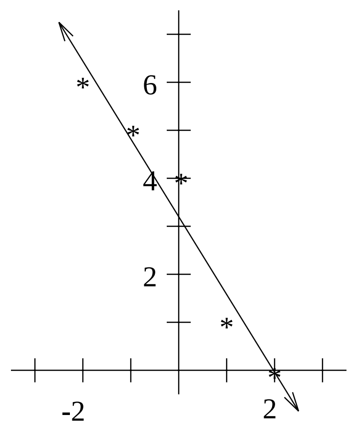
(b) $y = 3.3 - 1.5x$

Sum of squared errors = 0.67



(a) $y = 3.2 - 1.6x$

Sum of squared errors $= 1.2$



2. $y = 5.8 - 1.2x$

3. $y = 0.1 + 3x$

4. $y = 0.13 - 1.29x + 3.31x^2$

   Sum of squared errors $= 0.12$

5. $y = 3.23 - 1.90x + 0.79x^2$

6. Linear model: $5.5811 + 1.9054x$

   Parabolic model: $5.6200 + 1.8903x - 0.0140x^2$

7. $X = \begin{bmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \end{bmatrix} \qquad y = \begin{bmatrix} 44 \\ 11 \\ 3 \\ 1 \\ -91 \end{bmatrix}$

8. Answers will vary.

12.10 Solutions to Eigenpairs - Problems from page 130

Note: If your students have never seen the video footage of the Tacoma Bridge disaster, it is well worth your time to find it and show it to them.

1. $x = \begin{bmatrix} .33 \\ .42 \\ .25 \end{bmatrix} \qquad \lambda = 1$

2. $x = \begin{bmatrix} .49 \\ .33 \\ .18 \end{bmatrix} \qquad \lambda = 1$

3. $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \lambda = 8$

4. $x = \begin{bmatrix} 2.5 \\ 1 \end{bmatrix} \qquad \lambda = 4$

5. $x = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \qquad \lambda = 4$

   $x = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \lambda = -2$

6. $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \lambda = -2$

   $x = \begin{bmatrix} 3.5 \\ 1 \end{bmatrix} \qquad \lambda = 5$

7. $x = \begin{bmatrix} 3.5 \\ 1 \end{bmatrix} \qquad \lambda = 5$

$x = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \lambda = -4$

8. The solutions can be any non-zero, constant multiples of $\begin{bmatrix} -3 \\ 2 \\ -1 \\ 1 \end{bmatrix}$. Examples are: $\begin{bmatrix} -9 \\ 6 \\ -3 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -6 \\ 4 \\ -2 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} 3 \\ -2 \\ 1 \\ -1 \end{bmatrix}$.

12.11 Solutions to Second Review from page 140

1. (a) $A = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} .6 & .3 & .1 \\ .2 & .6 & .2 \\ .2 & .1 & .7 \end{bmatrix} \end{array}$

   (b) .1

   (c) .19

   (d) $A^2 = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} .44 & .37 & .19 \\ .28 & .44 & .28 \\ .28 & .19 & .53 \end{bmatrix} \end{array}$

(e) $A^3 = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} .376 & .373 & .251 \\ .312 & .376 & .312 \\ .312 & .251 & .437 \end{bmatrix} \end{array}$

(f) $A^n = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} .\overline{3} & .\overline{3} & .\overline{3} \\ .\overline{3} & .\overline{3} & .\overline{3} \\ .\overline{3} & .\overline{3} & .\overline{3} \end{bmatrix} \end{array}$, where n is large

(g) This means that after many years, a third of the teachers in the district will be at each school, A, B, and C, regardless of how many teachers are in each school now.

2. (a) $2x + 1.5$

   (b) $3x - 1.5$

3. (a) $1.9x^2 + 2.8x - 1.1$

   (b) $3x^2 - 2.2x + 1.1$

4. (a) $\lambda = 1 \qquad x = \begin{bmatrix} .\overline{3} \\ .\overline{3} \\ .\overline{3} \end{bmatrix}$

   (b) $\lambda = 7 \qquad x = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

   (c) $\lambda = 1 \qquad x = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\qquad \lambda = -2 \qquad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

12.12 Solutions to Matrix Test 1A

This test and the one which begins with a question beginning with the name Gail are intended to be similar enough that you could use the tests in the same class if you do not want to give all the students the same test.

Note: The first name on this test is Heather.

1. (a)
$$\begin{array}{c} \\ \text{Heather} \\ \text{Reena} \\ \text{Edwin} \\ \text{Mark} \end{array} \begin{array}{c} \begin{array}{ccc} \text{H} & \text{T} & \text{S} \end{array} \\ \left[\begin{array}{ccc} 3 & 4 & 8 \\ 5 & 1 & 6 \\ 1 & 0 & 9 \\ 2 & 3 & 7 \end{array}\right] \end{array}$$

(b)
$$\begin{array}{c} \\ \text{HW} \\ \text{TV} \\ \text{Sleep} \end{array} \begin{array}{c} \begin{array}{cccc} \text{H} & \text{R} & \text{E} & \text{M} \end{array} \\ \left[\begin{array}{cccc} 3 & 5 & 1 & 2 \\ 4 & 1 & 0 & 3 \\ 8 & 6 & 9 & 7 \end{array}\right] \end{array}$$

(c)
$$\begin{array}{c} \\ \text{Heather} \\ \text{Reena} \\ \text{Edwin} \\ \text{Mark} \end{array} \begin{array}{c} \begin{array}{cccc} \text{H} & \text{T} & \text{S} & \text{C} \end{array} \\ \left[\begin{array}{cccc} 3 & 4 & 8 & 0 \\ 5 & 1 & 6 & 0 \\ 1 & 0 & 9 & 5 \\ 2 & 3 & 7 & 0 \end{array}\right] \end{array}$$

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

3. $\begin{bmatrix} -13 & -32 \\ 52 & -6 \end{bmatrix}$

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

5. Augment the matrix with the identity matrix like this: $\left[\begin{array}{ccc|ccc} 3 & 4 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ -2 & 2 & 3 & 0 & 0 & 1 \end{array}\right]$.

   Then perform Gauss-Jordan elimination on the augmented matrix.

   The inverse is $\begin{bmatrix} -2 & 5 & -1 \\ 2.5 & -5.5 & 1 \\ -3 & 7 & -1 \end{bmatrix}$.

   Note that the students were only asked for the steps OR the inverse.

6. Answers will vary. Matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is a generic 2 by 2 symmetric matrix.

   The matrix is symmetric because $A = A^T$.

7. $x = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$

   Gauss-Jordan elimination uses elementary row operations until the matrix to the left of the bar is the identity matrix. Gaussian elimination uses elementary row operations until the matrix to the left of the bar is an upper triangular matrix. Then it uses back-substitution to find the values of x.

8. 116

9. (a) Consistent and uniquely determined because there is a single solution.

   (b) Consistent and underdetermined because the last equation does not give you any extra information.

Note that the students could also have used the appearance of the last row for reason for their solutions.

12.13 Solutions to Matrix Test 1B

This test and the one with question one beginning with the name Heather are intended to be similar enough that you could use the tests in the same class if you do not want to give all the students the same test.

Note: The first name on this test is Gail.

1. (a)

	W	St	T	Sl
Gail	3	2	1	8
Kerry	2	3	2	7
Brad	5	4	0	6

(b)

	G	K	B
Work	3	2	5
Study	2	3	4
TV	1	2	0
Sleep	8	7	6

(c)

	W	St	T	Sl
Gail	3	2	1	8
Kerry	2	3	2	7
Brad	5	4	0	6
Adam	12	0	0	7

2. $\begin{bmatrix} 4 & 9 \\ 9 & 3 \end{bmatrix}$

3. $\begin{bmatrix} 14 & -33 \\ 21 & 28 \end{bmatrix}$

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

5. Augment the matrix with the identity matrix like this: $\left[\begin{array}{rrr|rrr} 1 & 2 & 1 & 1 & 0 & 0 \\ 3 & -1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 1 & 0 & 0 & 1 \end{array}\right]$.

   Then perform Gauss-Jordan elimination on the augmented matrix.

   The inverse is $\begin{bmatrix} -7 & 1 & 5 \\ -1 & 0 & 1 \\ 10 & -1 & -7 \end{bmatrix}$.

   Note that the students were only asked for the steps OR the inverse.

6. Answers will vary. Matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is a generic 2 by 2 symmetric matrix.

   The matrix is symmetric because $A = A^T$.

7. $x = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$

   Gauss-Jordan elimination uses elementary row operations until the matrix to the left of the bar is the identity matrix. Gaussian elimination uses elementary row operations until the matrix to the left of the bar is an upper triangular matrix. Then it uses back-substitution to find the values of x.

8. 22

9. (a) Inconsistent because the last equation reduces to an contradictory equation. Note that the students could also have used the appearance of the last row for reason for their solutions.
   (b) Consistent and uniquely determined because there is a single solution.

12.14 Solutions to Matrix Test 1C

This is the test that begins with a question about Canine Cabins. This test is more abstract and more difficult than the other two tests. If your students are advanced, you might consider using this test. Otherwise, these problems can serve as good extra credit problems or enrichment problems. This test will also take longer to grade because many of the questions have answers that may vary, so you will have to check to make sure the students' answers are correct without the help of an answer key.

1. (a) $$Walls = \begin{array}{c} \\ \text{Small} \\ \text{Large} \end{array} \begin{array}{c} \begin{array}{ccc} \text{W} & \text{L} & \text{P} \end{array} \\ \begin{bmatrix} 25 & 1 & 0 \\ 70 & 1 & 0 \end{bmatrix} \end{array} \qquad Roof\ and\ Floor = \begin{array}{c} \\ \text{Small} \\ \text{Large} \end{array} \begin{array}{c} \begin{array}{ccc} \text{W} & \text{L} & \text{P} \end{array} \\ \begin{bmatrix} 15 & 1 & 0 \\ 30 & 1 & 0 \end{bmatrix} \end{array}$$

   (b) $$Walls + R\&F = \begin{array}{c} \\ \text{Small} \\ \text{Large} \end{array} \begin{array}{c} \begin{array}{ccc} \text{W} & \text{L} & \text{P} \end{array} \\ \begin{bmatrix} 40 & 2 & 0 \\ 100 & 2 & 0 \end{bmatrix} \end{array}$$

   (c) $$Complete = \begin{array}{c} \\ \text{Small} \\ \text{Large} \end{array} \begin{array}{c} \begin{array}{ccc} \text{W} & \text{L} & \text{P} \end{array} \\ \begin{bmatrix} 40 & 3.5 & 1 \\ 100 & 4 & 1.5 \end{bmatrix} \end{array}$$

(d) $Attach = Complete - (Walls + R\&F) = \begin{array}{c|ccc} & W & L & P \\ \hline Small & 0 & 1.5 & 1 \\ Large & 0 & 2 & 1.5 \end{array}$

(e) $Cost\ per\ house = Complete * Cost = \begin{array}{c|c} & \$/house \\ \hline Small & 49 \\ Large & 86 \end{array}$ where

$$Cost = \begin{array}{c|c} & \$/unit \\ \hline W & .5 \\ L & 6 \\ P & 8 \end{array}$$

(f) $Materials\ needed = Ordered * Complete = \begin{array}{ccc} Wood & Labor & Paint \\ \hline 3200 & 185 & 60 \end{array}$ where

$$Ordered = \begin{array}{cc} Small & Large \\ \hline 30 & 20 \end{array}$$

(g) $Money\ needed = Ordered * Complete * Cost = \$3190$

2. AB^TC, A^TC^TB, BA^TC^T, CAB^T, C^TBA^T, B^TCA

3. (a) $\det(A) = (-2 + 12 + 3) - (-3 + 12 + 2) = 2$

(b) $x = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$

Gauss-Jordan elimination uses elementary row operations until the matrix to the left of the bar is the identity matrix. Gaussian elimination uses elementary row operations until the matrix to the left of the bar is an upper triangular matrix. Then it uses back substitution to find the values of x.

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

(d) Show that $AA^{-1} = A^{-1}A = I$.

4. The answers may vary.

5. The answers may vary.

12.15 Solutions to Matrix Test 2A

Note: The first question on this test is about a meteorologist.

1. (a) $A = \begin{array}{c} \\ S \\ C \\ R \end{array} \begin{array}{c} \begin{array}{ccc} S & C & R \end{array} \\ \begin{bmatrix} .5 & .3 & .2 \\ .3 & .2 & .5 \\ .3 & .4 & .3 \end{bmatrix} \end{array}$

(b) .4

(c) .29

(d) $A^2 = \begin{array}{c} \\ S \\ C \\ R \end{array} \begin{array}{c} \begin{array}{ccc} S & C & R \end{array} \\ \begin{bmatrix} .4 & .29 & .31 \\ .36 & .33 & .31 \\ .36 & .29 & .35 \end{bmatrix} \end{array}$

(e) $A^3 = \begin{array}{c} \\ S \\ C \\ R \end{array} \begin{array}{c} \begin{array}{ccc} S & C & R \end{array} \\ \begin{bmatrix} .38 & .302 & .318 \\ .372 & .298 & .33 \\ .372 & .306 & .322 \end{bmatrix} \end{array}$

(f) $$A^n = \begin{array}{c} \\ S \\ C \\ R \end{array} \begin{array}{c} \begin{array}{ccc} S & C & R \end{array} \\ \left[\begin{array}{ccc} .375 & .302 & .323 \\ .375 & .302 & .323 \\ .375 & .302 & .323 \end{array}\right] \end{array}$$, where n is large.

(g) The weather will be sunny 37.5% of the time, cloudy 30.2% of the time, and rainy 32.3% of the time regardless of today's weather.

2. $-.5x + 2.1$

3. $-3.3x^2 - 2.0x + 5.1$

4. (a) $\lambda = 1 \qquad x = \begin{bmatrix} .375 \\ .302 \\ .323 \end{bmatrix}$

   (b) $\lambda = -5 \qquad x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

   (c) $\lambda = 4 \qquad x = \begin{bmatrix} \frac{-3}{2} \\ 1 \end{bmatrix}$ and $\lambda = -5 \qquad x = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

12.16 Solutions to Matrix Test 2B

Note: The first question on this test is about the clothes of a young lady.

1. (a) $$A = \begin{array}{c} \\ J \\ S \\ D \end{array} \begin{array}{c} \begin{array}{ccc} J & S & D \end{array} \\ \left[\begin{array}{ccc} .2 & .3 & .5 \\ .1 & .6 & .3 \\ .5 & .2 & .3 \end{array}\right] \end{array}$$

   (b) .5

   (c) .33

(d) $$A^2 = \begin{array}{c} \\ J \\ S \\ D \end{array}\begin{array}{c} \begin{array}{ccc} J & S & D \end{array} \\ \left[\begin{array}{ccc} .32 & .34 & .34 \\ .23 & .45 & .32 \\ .27 & .33 & .4 \end{array}\right] \end{array}$$

(e) $$A^3 = \begin{array}{c} \\ J \\ S \\ D \end{array}\begin{array}{c} \begin{array}{ccc} J & S & D \end{array} \\ \left[\begin{array}{ccc} .268 & .368 & .364 \\ .251 & .403 & .346 \\ .287 & .359 & .354 \end{array}\right] \end{array}$$

(f) $$A^n = \begin{array}{c} \\ J \\ S \\ D \end{array}\begin{array}{c} \begin{array}{ccc} J & S & D \end{array} \\ \left[\begin{array}{ccc} .27 & .38 & .35 \\ .27 & .38 & .35 \\ .27 & .38 & .35 \end{array}\right] \end{array}$$, where n is large.

(g) She will wear jeans 27% of the time, slacks 38% of the time, and dresses 35% of the time regardless of what she is wearing today.

2. $5.4x - 3.3$

3. $x^2 - 1.6x + 2.8$

4. (a) $\lambda = 1 \qquad x = \begin{bmatrix} .27 \\ .38 \\ .35 \end{bmatrix}$

   (b) $\lambda = -6 \qquad x = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$

   (c) $\lambda = 0 \qquad x = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and $\qquad \lambda = 5 \qquad x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Bibliography

Braun, M. (1983). *Differential equations and their applications.* New York: Springer-Verlag.

Houston Independent School District. (1992). *Houston Independent School District scope and sequence grades 9-12.* Houston, TX: Houston Independent School District.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Nering, Evar D. (1974). *Elementary Linear Algebra.* Philadelphia, PA: W.B. Saunders Company.

Skemp, R. (1987). *The psychology of learning mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Index

adding matrices, 11
augmented matrix, 35

back-substitution, 42
by, 4

characteristic equation, 120
cofactor, 56
consistent, 68
counterexample, 7
Cramer's rule, 60

determinant, 54
diagonal matrix, 41
dimensions, 4
dominant eigenvector, 116
dot product, 20

eigenpair, 114
eigenvalue, 114
eigenvector, 114
element, 1, 3
elementary row operations, 34
expansion by minors, 56

Gauss-Jordan elimination, 36
Gaussian elimination, 42

identity matrix, 27, 37
inconsistent, 68, 70
independent, 71
inner product, 20
inverse, 37

least squares, 96, 98
lower triangular, 47

main diagonal, 6
Markov chain, 81, 82
matrices, 3
matrix, 3
matrix product, 23
minor, 56

normal equation, 99

ordered pair, 1

pivot, 40
power method, 116, 117

Rayleigh quotient, 117

scalar, 18
scalar multiplication, 18
scalar product, 19

square matrix, 6

state, 80

subtracting matrices, 12

symmetric, 6

transition matrix, 81

transpose, 5

upper triangular, 41

vector, 1

column, 1

row, 1

zero matrix, 13