Linear Algebra and Its Applications

Gilbert Strang

LINEAR ALGEBRA AND ITS APPLICATIONS

GILBERT STRANG
Massachusetts Institute of Technology

ACADEMIC PRESS New York San Francisco London
A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

Library of Congress Cataloging in Publication Data

Strang, Gilbert.
Linear algebra and its applications

Bibliography: p.
Includes index.
1. Algebras, Linear. I. Title.
QA184.S8 512'.5 75-26637
ISBN 0-12-673650-2

AMS (MOS) 1970 Subject Classifications: 15-01, 52-01, 65-01

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

I believe that the teaching of linear algebra has become too abstract. This is a sweeping judgment, and perhaps it is too sweeping to be true. But I feel certain that a text can explain the essentials of linear algebra, and develop the ability to reason mathematically, without ignoring the fact that *this subject is as useful and central and applicable as calculus.* It has a simplicity which is too valuable to be sacrificed.

Of course there are good reasons for the present state of courses in linear algebra: The subject is an excellent introduction to the precision of a mathematical argument, and to the construction of proofs. These virtues I recognize and accept (and hope to preserve); I enjoyed teaching in exactly this way. Nevertheless, once I began to experiment with alternatives at M.I.T., another virtue became equally important: Linear algebra allows and even encourages a very satisfying combination of both elements of mathematics—abstraction and application.

As it is, too many students struggle with the abstraction and never get to see the application. And too many others, especially those who are outside mathematics departments, never take the course. Even our most successful students tend to become adept at abstraction, but inept at any calculation—solving linear equations by Cramer's rule, for example, or understanding eigenvalues only as roots of the characteristic equation. There is a growing desire to make our teaching more useful than that, and more open.

We hope to treat linear algebra in a way which makes sense to a wide variety of students at all levels. This does not imply that we have written a cookbook; the subject deserves better than that. It does imply less concentration on rigor for its own sake, and more on understanding—*we try to explain rather than to deduce.* Some definitions are formal, but others are allowed to come to the surface in the middle of a discussion. In the same way, some proofs are intended to be orderly and precise, but not all. In every case the underlying theory has to be there; it is the core of the subject, but it can be motivated and reinforced by examples.

One specific difficulty in constructing the course is always present, and is hard to postpone: How should it start? Most students come to the first class already knowing something about linear equations. Nevertheless, we are convinced that linear algebra must begin with the fundamental problem of n equations in n unknowns, and that it must teach the simplest and most useful method of solution—Gaussian elimination (not determinants!). Fortunately, even though this method is simple, there are a number of insights that are central to its understanding and new to almost every student. The most important is the equivalence between elimination and matrix factorization; the coefficient matrix is transformed into a product of triangular matrices. This provides a perfect introduction to matrix notation and matrix multiplication.

The other difficulty is to find the right speed. If matrix calculations are already familiar, then *Chapter 1 must not be too slow*; the next chapter is the one which demands hard work. Its goal is a genuine understanding, deeper than elimination can give, of the equation $Ax = b$. I believe that the introduction of four fundamental subspaces—the column space of A; the row space; and their orthogonal complements, the two nullspaces—is an effective way to generate examples of linear dependence and independence, and to illustrate the ideas of basis and dimension and rank. The orthogonality is also a natural extension to n dimensions of the familiar geometry of three-dimensional space. And of course those four subspaces are the key to $Ax = b$.

Chapters 1–5 are really the heart of a course in linear algebra. They contain a large number of applications to physics, engineering, probability and statistics, economics, and biology. (There is also the geometry of a methane molecule, and even an outline of factor analysis in psychology, which is the one application that my colleagues at M.I.T. refuse to teach!) At the same time, you will recognize that this text can certainly not explain every possible application of matrices. It is simply a first course in linear algebra. Our goal is not to develop all the applications, but to prepare for them—and that preparation can only come by understanding the theory.

This theory is well established. After the vector spaces of Chapter 2, we study projections and inner products in Chapter 3, determinants in Chapter 4, and eigenvalues in Chapter 5. I hope that engineers and others will look especially at Chapter 5, where we concentrate on the uses of diagonalization (including the spectral theorem) and save the Jordan form for an appendix. Each chapter is followed by a set of review exercises, and is so organized that its last section is optional; this applies also to Section 3.4 on pseudoinverses. In a one-semester or a one-quarter course, the instructor must decide whether the positive definite matrices of Chapter 6 or the linear programs of Chapter 8 are the more essential to his class; I believe that Sections 8.1 and 8.4 will allow a brief but worthwhile introduction to linear programming and game theory.

The book also contains the essentials of three quite different courses. One is on numerical linear algebra, which involves all of Chapter 1, the essential facts of Chapters 2 to 6, and then Chapter 7 on computations and Section 8.2 on the simplex method. Another is "linear algebra for statistics," which must treat Chapters 3 and 6 much more completely. And the third possibility is to regard inequalities as coequal with equations, as economists do, and go as quickly as possible from $Ax = b$ to linear programming and duality.

We should like to ask one favor of the mathematician who simply wants to teach basic linear algebra. That is the true purpose of the book, and we hope he will not be put off by the "operation counts," and the other remarks about numerical computation, which arise especially in Chapter 1. From a practical viewpoint these comments are obviously important. Also from a theoretical viewpoint they have a serious purpose—to reinforce a detailed grasp of the elimination sequence, by actually counting the steps. I normally ask the class to make this count during the first or second lecture, with completely unpredictable results. But there is no need to discuss this or any other computer-oriented topic in class; any text ought to supplement as well as summarize the lectures.

In short, a book is needed that will permit the applications to be taught successfully, in combination with the underlying mathematics. That is the book I have tried to write.

For help in writing it, I take this special opportunity to give thanks to Tom Slobko for his encouragement, to Ursula for typing everything with such gentle grace, and to my family who are precious above all. Beyond this there is an earlier debt, which I can never fully repay. It is to my parents, and I now dedicate the book to them, hoping that they will understand how much they gave to it: Thank you both.

GILBERT STRANG

CONTENTS

Preface ix

1 GAUSSIAN ELIMINATION

1.1 Introduction 1
1.2 An Example of Gaussian Elimination 2
1.3 Matrix Notation and Matrix Multiplication 6
1.4 Gaussian Elimination = Triangular Factorization 17
1.5 Row Exchanges, Inverses, and Roundoff Errors 25
1.6 Band Matrices, Symmetric Matrices, and Applications 35
Review Exercises 42

2 THE THEORY OF SIMULTANEOUS LINEAR EQUATIONS

2.1 Vector Spaces and Subspaces 44
2.2 The Solution of m Equations in n Unknowns 49
2.3 Linear Independence, Basis, and Dimension 56
2.4 The Four Fundamental Subspaces 64
2.5 Orthogonality of Vectors and Subspaces 75
2.6 Pairs of Subspaces and Products of Matrices 86
Review Exercises 94

3 ORTHOGONAL PROJECTIONS AND LEAST SQUARES

3.1 Inner Products and Transposes 96
3.2 Projections onto Subspaces and Least Squares Approximations 104

3.3 Orthogonal Bases, Orthogonal Matrices, and Gram–Schmidt Orthogonalization 115
3.4 The Pseudoinverse and the Singular Value Decomposition 130
3.5 Weighted Least Squares 138
Review Exercises 144

4 DETERMINANTS

4.1 Introduction 146
4.2 The Properties of the Determinant 148
4.3 Formulas for the Determinant 154
4.4 Applications of Determinants 162
Review Exercises 169

5 EIGENVALUES AND EIGENVECTORS

5.1 Introduction 171
5.2 The Diagonal Form of a Matrix 181
5.3 Difference Equations and the Powers A^k 186
5.4 Differential Equations and the Exponential e^{At} 196
5.5 The Complex Case: Hermitian and Unitary Matrices 206
5.6 Similarity Transformations and Triangular Forms 221
Review Exercises 229

6 POSITIVE DEFINITE MATRICES

6.1 Minima, Maxima, and Saddle Points 231
6.2 Tests for Positive Definiteness 237
6.3 Semidefinite and Indefinite Matrices; $Ax = \lambda Bx$ 244
6.4 Minimum Principles and Rayleigh's Quotient 251
6.5 The Rayleigh–Ritz Principle and Finite Elements 260

7 COMPUTATIONS WITH MATRICES

7.1 Introduction 265
7.2 The Norm and Condition Number of a Matrix 266
7.3 The Computation of Eigenvalues 273
7.4 Iterative Methods for $Ax = b$ 283

8 LINEAR PROGRAMMING AND GAME THEORY

8.1 Linear Inequalities 291
8.2 The Simplex Method 297

8.3 The Theory of Duality 309
8.4 Game Theory and the Minimax Theorem 319

APPENDIX A LINEAR TRANSFORMATIONS, MATRICES, AND CHANGE OF BASIS 327

APPENDIX B THE JORDAN FORM 335

References 341

Solutions to Exercises 343

Index 369

1

GAUSSIAN ELIMINATION

INTRODUCTION ■ 1.1

The central problem of linear algebra is the solution of simultaneous linear equations. The most important case, and the simplest, is when the number of unknowns equals the number of equations. Therefore we begin with this problem: *n equations in n unknowns.*

Two ways of solving simultaneous equations are proposed, almost in a sort of competition, from high school texts on. The first is the method of ***elimination***: Multiples of the first equation in the system are subtracted from the other equations, in such a way as to remove the first unknown from those equations. This leaves a smaller system, of $n - 1$ equations in $n - 1$ unknowns. The process is repeated over and over until there remains only one equation and one unknown, which can be solved immediately. Then it is not hard to go backward, and find all the other unknowns in reverse order; we shall work out an example in a moment. A second and more sophisticated way introduces the idea of ***determinants***. There is an exact formula, called Cramer's rule, which gives the solution (the correct values of the unknowns) as a ratio of two n by n determinants. It is not always obvious from the examples that are worked in a textbook ($n = 3$ or $n = 4$ is about the upper limit on the patience of a reasonable human being) which way is better.

In fact, the more sophisticated formula involving determinants is a disaster, and elimination is the algorithm that is constantly used to solve large systems of simultaneous equations. Our first goal is to understand this algorithm. It is generally called *Gaussian elimination.*

The algorithm is deceptively simple, and in some form it may already be familiar to the reader. But there are four aspects that lie deeper than the simple mechanics of elimination, and which—together with the algorithm itself—we want to explain in this chapter. They are:

(1) The interpretation of the elimination method as a factorization of the coefficient matrix. We shall introduce *matrix notation* for the system of simultaneous equations, writing the n unknowns as a vector x and the n equations in the matrix shorthand $Ax = b$. Then *elimination amounts to factoring A into a product LU of a lower triangular matrix L and an upper triangular matrix U.* This is a basic and very useful observation.

Of course, we have to introduce matrices and vectors in a systematic way, as well as the rules for their multiplication. We also define the transpose A^T and the inverse A^{-1} of a matrix A.

(2) In most cases the elimination method works without any difficulties or modifications. In some exceptional cases it breaks down—either because the equations were originally written in the wrong order, which is easily fixed by exchanging them, or else because the equations $Ax = b$ fail to have a unique solution. In the latter case there may be no solution, or infinitely many. We want to understand how, at the time of breakdown, the elimination process identifies each of these possibilities.

(3) It is essential to have a rough count of the *number of arithmetic operations* required to solve a system by elimination. In many practical problems the decision of how many unknowns to introduce—balancing extra accuracy in a mathematical model against extra expense in computing—is governed by this operation count.

(4) We also want to see, intuitively, how sensitive to *roundoff error* the solution x might be. Some problems are sensitive; others are not. Once the source of difficulty becomes clear, it is easy to guess how to try to control it. Without control, a computer could carry out millions of operations, rounding each result to a fixed number of digits, and produce a totally useless "solution."

The final result of this chapter will be an elimination algorithm which is about as efficient as possible. It is essentially the algorithm that is in constant use in a tremendous variety of applications. And at the same time, understanding it in terms of matrices—the coefficient matrix, the matrices that carry out an elimination step or an exchange of rows, and the final triangular factors L and U—is an essential foundation for the theory.

1.2 ■ AN EXAMPLE OF GAUSSIAN ELIMINATION

The way to understand this subject is by example. We begin in three dimensions with the system

$$\begin{aligned} 2u + v + w &= 1 \\ 4u + v &= -2 \\ -2u + 2v + w &= 7. \end{aligned} \tag{1}$$

The problem is to find the unknown values of u, v, and w, and we shall apply Gaussian elimination. (Gauss is recognized as the greatest of all mathematicians, but certainly not because of this invention, which probably took him ten minutes. Ironically, however, it is the most frequently used of all the ideas that bear his name.) The method starts by *subtracting multiples of the first equation from the others, so as to eliminate u from the last two equations*. This requires that we

(a) subtract 2 times the first equation from the second;
(b) subtract -1 times the first equation from the third.

The result is an equivalent system of equations

$$\begin{aligned} 2u + v + w &= 1 \\ -1v - 2w &= -4 \\ 3v + 2w &= 8. \end{aligned} \tag{2}$$

The coefficient 2, which multiplied the first unknown u in the first equation, is known as the *pivot* in this first elimination step.

At the second stage of elimination, we ignore the first equation. The other two equations involve only the two unknowns v and w, and the same elimination procedure can be applied to them. *The pivot for this stage is* -1, and a multiple of this second equation will be subtracted from the remaining equations (in this case there is only the third one remaining) so as to eliminate the second unknown v. This means that we

(c) subtract -3 times the second equation from the third.

The elimination process is now complete, at least in the "forward" direction, and leaves the simplified system

$$\begin{aligned} 2u + v + w &= 1 \\ -1v - 2w &= -4 \\ -4w &= -4. \end{aligned} \tag{3}$$

There is an obvious order in which to solve this system. The last equation gives $w = 1$; substituting into the second equation, we find $v = 2$; then the first equation gives $u = -1$. This simple process is called ***back-substitution***.

It is easy to understand how the elimination idea can be extended to n equations in n unknowns, no matter how large the system may be. At the first stage, we use multiples of the first equation to annihilate all coefficients below the first pivot. Next, the second column is cleared out below the second pivot; and so on. Finally, the last equation contains only the last unknown. Back-substitution yields the answer in the opposite order, beginning with the last unknown, then solving for the next to last, and eventually for the first.

EXERCISE 1.2.1 Apply elimination and back-substitution to solve

$$\begin{aligned} 2u - 3v &= 3 \\ 4u - 5v + w &= 7 \\ 2u - v - 3w &= 5. \end{aligned}$$

What are the pivots? List the three operations in which a multiple of one row is subtracted from another.

EXERCISE 1.2.2 Solve the system

$$\begin{aligned} 2u - v &= 0 \\ -u + 2v - w &= 0 \\ - v + 2w - z &= 0 \\ - w + 2z &= 5. \end{aligned}$$

We want to ask two questions. They may seem a little premature—after all, we have barely got the algorithm working—but their answers will shed more light on the method itself. The first question is whether this elimination procedure always leads to the solution. *Under what circumstances could the process break down?* The answer is: If none of the pivots are zero, there is only one solution to the problem and it is found by forward elimination and back-substitution. But if any of the pivots happen to be zero, the elimination technique cannot proceed.

If the first pivot were zero, for example, the elimination of u from the other equations would be impossible. The same is true at every intermediate stage. Notice that an intermediate pivot may become zero during the elimination process (as in Exercise 1.2.3 below) even though in the original system the coefficient in that place was not zero. Roughly speaking, *we do not know whether the pivots are nonzero until we try*, by actually going through the elimination process.

In most cases this problem of a zero pivot can be cured, and elimination can proceed to find the unique solution to the problem. In other cases, a breakdown is unavoidable since the equations have either no solution or infinitely many.

We postpone to a later section the analysis of breakdown.

The second question is very practical, in fact it is financial. *How many separate arithmetical operations does elimination require for a system of n equations in n unknowns?* If n is large, a computer is going to take our place in carrying out the elimination (you may have such a program available, or be able to write one) but since all the steps are known in advance, we should be able to predict how long the computer will take. For the moment we ignore the right-hand sides of the equations, and count only the operations on the left. These operations are of two kinds. One is a division by the pivot in order to find out what multiple (say l) of the pivotal equation is to be subtracted from an equation below it. Then when we actually do this subtraction of one equation from another, we continually meet a "multiply–subtract" combination; the terms in the pivotal equation are multiplied by l, and then subtracted from the equation beneath it.

Suppose we agree to call each division, and each multiplication–subtraction, a single operation. At the beginning, when the first equation has length n, *it takes n operations for every zero we achieve in the first column*—one to find the multiple l, and the others to find the new entries along the row. There are $n - 1$ rows underneath the first one, and therefore $n - 1$ zeros to be produced below the first pivot, so *the first stage of elimination needs $n(n - 1) = n^2 - n$ operations.* After that stage, the first column is set. Now notice that later stages are faster because the equations are becoming progressively shorter; at the second stage we are working with only $n - 1$ equations in $n - 1$ unknowns. When the elimination is down to k equations, only $k(k - 1) = k^2 - k$ operations are needed to clear out the column below the pivot—by the same reasoning that applied to the first stage, when k equaled n. Altogether, therefore, *the total number of arithmetical operations on the left side of the equations is*

$$P = (n^2 - n) + \cdots + (k^2 - k) + \cdots + (1^2 - 1).$$

(Notice that there was no work to do at the last stage, $1^2 - 1 = 0$, when we are down to one equation in one unknown.) This sum P is known to equal

$$P = \sum_1^n k^2 - \sum_1^n k = \frac{1}{3}n\left(n + \frac{1}{2}\right)(n + 1) - \frac{1}{2}n(n + 1) = \frac{n^3 - n}{3}.$$

(Here calculus is actually useful as a check. The integral of x^2 from 0 to n is $n^3/3$, and the integral of x is $n^2/2$. These are exactly the leading terms in the two sums.) If n is at all large, *a very good estimate for the number of operations is $P \approx n^3/3$.*

Back-substitution is considerably faster. The last unknown is found from one operation (a division by the last pivot), the second to last unknown requires two (a multiplication–subtraction and then a division), and so on. The kth step involves only k operations. Therefore, back-substitution requires altogether

$$Q = \sum_{k=1}^n k = \frac{1}{2}n(n + 1) \approx \frac{n^2}{2} \quad \text{operations.}$$

A few years ago, almost every mathematician would have guessed that these numbers were essentially optimal, in other words that a general system of order n could not be solved with much fewer than $n^3/3$ multiplications. (There were even theorems to demonstrate it, but they did not allow for all possible methods.) Astonishingly, that guess has been proved wrong, and *there now exists a method that requires only $Cn^{\log_2 7}$ operations*! Fortunately for elimination, the constant C is by comparison so large, and so many more additions are required, and the computer programming is so awkward, that the new method is largely of theoretical interest. It seems to be completely unknown whether the exponent can be made any smaller.

EXERCISE 1.2.3 Apply elimination to the system

$$\begin{aligned} u + v + w &= -2 \\ 3u + 3v - w &= 6 \\ u - v + w &= -1. \end{aligned}$$

When a zero pivot arises, exchange that equation for the one below it, and proceed. What coefficient of v in the third equation, in place of the present -1, would make it impossible to proceed—and force the elimination method to break down?

EXERCISE 1.2.4 For a system of two equations like

$$\begin{aligned} au + bv &= 0 \\ cu + dv &= 1 \end{aligned}$$

list explicitly the $P = 2$ individual operations that are applied to the left side.

EXERCISE 1.2.5 With reasonable assumptions on computer speed and cost, how large a system can be solved for \$1, and for \$1000? Use $n^3/3$ as the operation count, and you might pay \$1000 an hour for a computer that could average a million operations a second.

EXERCISE 1.2.6 (very optional) Normally the multiplication of two complex numbers

$$(a + ib)(c + id) = (ac - bd) + i(bc + ad)$$

involves the four separate multiplications ac, bd, bc, ad. Ignoring i, can you compute the quantities $ac - bd$ and $bc + ad$ with only three multiplications? (You may do additions, such as forming $a + b$ before multiplying, without any penalty.)

EXERCISE 1.2.7 Use elimination to solve

$$\begin{aligned} u + v + w &= 6 \\ u + 2v + 2w &= 11 \\ 2u + 3v - 4w &= 3. \end{aligned}$$

1.3 ■ MATRIX NOTATION AND MATRIX MULTIPLICATION

So far, with our 3 by 3 example, we have been able to write out all the equations in full. We could even list in detail each of the elimination steps, subtracting a multiple of one row from another, which put the system of equations into a simpler form. For a large system, however, this way of keeping track of the elimination would be hopeless; a much more concise record is needed. We shall now introduce matrix notation to describe the original system of equations, and matrix multiplication to describe the operations that make it simpler.

Notice that in our example

$$\begin{aligned} 2u + v + w &= 1 \\ 4u + v \qquad &= -2 \\ -2u + 2v + w &= 7 \end{aligned}$$

three different types of quantities appear. There are the unknowns u, v, w; there are the right sides 1, -2, 7; and finally, there is a set of nine numerical coefficients on the left side (one of which happens to be zero). For the column of numbers on the right side—the *inhomogeneous terms* in the equations—we introduce the vector notation

$$b = \begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix}.$$

This is a three-dimensional column vector. To represent it geometrically, we can take its three components as the coordinates of a point in three-dimensional space. Then every point in the space is matched with a three-dimensional vector (which we may visualize as an arrow, or a directed line segment, which starts at the origin and ends at the point).

The basic operations are the addition of two such vectors and the multiplication of a vector by a scalar. Geometrically, $2b$ is a vector in the same direction as b but twice as long; $-2b$ goes in the opposite direction; and $b + c$ is found by placing the starting point of the vector c at the end point of b. Algebraically, this just means that vector operations are carried out *component by component*:

$$2b = 2\begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ 14 \end{bmatrix}, \qquad -2b = \begin{bmatrix} -2 \\ 4 \\ -14 \end{bmatrix},$$

$$b + c = \begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix} + \begin{bmatrix} 1 \\ -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ -6 \\ 3 \end{bmatrix}.$$

Two vectors can be added only if they have the same dimension, that is, the same number of components.

The three unknowns in the equation are also represented by a vector:

$$\text{the unknown is} \quad x = \begin{bmatrix} u \\ v \\ w \end{bmatrix}; \qquad \text{the solution is} \quad x = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Again these are three-dimensional column vectors. For the array of nine coefficients, we introduce a "matrix" with three rows and three columns. This is the

coefficient matrix

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

Notice that, because the number of equations in our example agrees with the number of unknowns, A is a *square matrix* (of order three). More generally, we might have n equations in n unknowns—with a square coefficient matrix of order n. Or still more generally, we might have m equations and n unknowns. In this case the coefficient matrix will be rectangular, with m rows and n columns—in other words, it will be an "m by n matrix."

Matrices are added to each other, or multiplied by numerical constants, exactly as vectors are—one component at a time. In fact we may regard vectors as special cases of matrices; they are matrices with only one column. As before, two matrices can be added only if they have the same shape:

$$\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 1 \\ 1 & 6 \end{bmatrix}, \qquad 2\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 6 & 0 \\ 0 & 8 \end{bmatrix}.$$

Multiplication of a Matrix and a Vector

Now we put this notation to use. We propose to rewrite the system (1) of three equations in three unknowns in the simplified matrix form $Ax = b$. Written out in full, this form is

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 7 \end{bmatrix}.$$

The right side is clear enough; it is the column vector of inhomogeneous terms. The left side consists of the vector x, premultiplied by the matrix A. Obviously, this multiplication will be defined *exactly so as to reproduce the original system* (1). Therefore, the first component of the product Ax must come from "multiplying" the first row of A into the column vector x:

$$\begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2u + v + w \end{bmatrix}. \tag{4}$$

This equals the first component of b; $2u + v + w = 1$ is the first equation in our system. The second component of the product Ax is determined by the second row of A—it is $4u + v$—and the third component $-2u + 2v + w$ comes from the third row. Thus the matrix equation $Ax = b$ is precisely equivalent to the three simultaneous equations with which we started.