Check for updates

Chapter 1
Linear Algebraic Systems

Linear algebra is the core of modern applied mathematics. Its humble origins are to be found in the need to solve "elementary" systems of linear algebraic equations. But its ultimate scope is vast, impinging on all of mathematics, both pure and applied, as well as numerical analysis, statistics, data science, physics, engineering, mathematical biology, financial mathematics, and every other discipline in which mathematical methods are required. A thorough grounding in the methods and theory of linear algebra is an essential prerequisite for understanding and harnessing the power of mathematics throughout its multifaceted applications.

In the first chapter, our focus will be on the most basic method for solving linear algebraic systems, known as *Gaussian Elimination* in honor of one of the all-time mathematical greats, the early nineteenth-century German mathematician Carl Friedrich Gauss, although the method appears in Chinese mathematical texts from around 150 CE, if not earlier, and was also known to Isaac Newton. Gaussian Elimination is quite elementary, but remains one of *the* most important algorithms in applied (as well as theoretical) mathematics. Our initial focus will be on the most important class of systems: those involving the same number of equations as unknowns — although we will eventually develop techniques for handling completely general linear systems. While the former typically have a unique solution, general linear systems may have either no solutions or infinitely many solutions. Since physical models require existence and uniqueness of their solution, the systems arising in applications often (but not always) involve the same number of equations as unknowns. Nevertheless, the ability to confidently handle all types of linear systems is a basic prerequisite for further progress in the subject. In contemporary applications, particularly those arising in numerical solutions of differential equations, in signal and image processing, and in contemporary data analysis, the governing linear systems can be huge, sometimes involving millions of equations in millions of unknowns, challenging even the most powerful supercomputer. So, a systematic and careful development of solution techniques is essential. Section 1.7 discusses some of the practical issues and limitations in computer implementations of the Gaussian Elimination method for large systems arising in applications.

Modern linear algebra relies on the basic concepts of scalar, vector, and matrix, and so we must quickly review the fundamentals of matrix arithmetic. Gaussian Elimination can be profitably reinterpreted as a certain matrix factorization, known as the (permuted) LU decomposition, which provides valuable insight into the solution algorithms. Matrix inverses and determinants are also discussed in brief, primarily for their theoretical properties. As we shall see, formulas relying on the inverse or the determinant are extremely inefficient, and so, except in low-dimensional or highly structured environments, are to be avoided in almost all practical computations. In the theater of applied linear algebra, Gaussian Elimination and matrix factorization are the stars, while inverses and determinants are relegated to the supporting cast.

1.1 Solution of Linear Systems

Gaussian Elimination is a simple, systematic algorithm to solve systems of linear equations. It is the workhorse of linear algebra, and, as such, of absolutely fundamental importance

© Springer International Publishing AG, part of Springer Nature 2018

P. J. Olver, C. Shakiban, *Applied Linear Algebra*, Undergraduate Texts in Mathematics,
https://doi.org/10.1007/978-3-319-91041-3_1

in applied mathematics. In this section, we review the method in the most important case, in which there is the same number of equations as unknowns. The general situation will be deferred until Section 1.8.

To illustrate, consider an elementary system of three linear equations

$$\begin{aligned} x + 2y + z &= 2, \\ 2x + 6y + z &= 7, \\ x + y + 4z &= 3, \end{aligned} \tag{1.1}$$

in three unknowns x, y, z. Linearity[†] refers to the fact that the unknowns only appear to the first power, and there are no product terms like xy or xyz. The basic solution method is to systematically employ the following fundamental operation:

Linear System Operation #1: Add a multiple of one equation to another equation.

Before continuing, you might try to convince yourself that this operation doesn't change the solutions to the system. Our goal is to judiciously apply the operation and so be led to a much simpler linear system that is easy to solve, and, moreover, has the same solutions as the original. Any linear system that is derived from the original system by successive application of such operations will be called an *equivalent system*. By the preceding remark, *equivalent linear systems have the same solutions*.

The systematic feature is that we successively eliminate the variables in our equations in order of appearance. We begin by eliminating the first variable, x, from the second equation. To this end, we subtract twice the first equation from the second, leading to the equivalent system

$$\begin{aligned} x + 2y + z &= 2, \\ 2y - z &= 3, \\ x + y + 4z &= 3. \end{aligned} \tag{1.2}$$

Next, we eliminate x from the third equation by subtracting the first equation from it:

$$\begin{aligned} x + 2y + z &= 2, \\ 2y - z &= 3, \\ -y + 3z &= 1. \end{aligned} \tag{1.3}$$

The equivalent system (1.3) is already simpler than the original (1.1). Notice that the second and third equations do not involve x (by design) and so constitute a system of two linear equations for two unknowns. Moreover, once we have solved this subsystem for y and z, we can substitute the answer into the first equation, and we need only solve a single linear equation for x.

We continue on in this fashion, the next phase being the elimination of the second variable, y, from the third equation by adding $\frac{1}{2}$ the second equation to it. The result is

$$\begin{aligned} x + 2y + z &= 2, \\ 2y - z &= 3, \\ \tfrac{5}{2} z &= \tfrac{5}{2}, \end{aligned} \tag{1.4}$$

which is the simple system we are after. It is in what is called *triangular form*, which means that, while the first equation involves all three variables, the second equation involves only the second and third variables, and the last equation involves only the last variable.

† The "official" definition of linearity will be deferred until Chapter 7.

Any triangular system can be straightforwardly solved by the method of *Back Substitution*. As the name suggests, we work backwards, solving the last equation first, which requires that $z = 1$. We substitute this result back into the penultimate equation, which becomes $2y - 1 = 3$, with solution $y = 2$. We finally substitute these two values for y and z into the first equation, which becomes $x + 5 = 2$, and so the solution to the triangular system (1.4) is

$$x = -3, \qquad y = 2, \qquad z = 1. \tag{1.5}$$

Moreover, since we used only our basic linear system operation to pass from (1.1) to the triangular system (1.4), this is also the solution to the original system of linear equations, as you can check. We note that the system (1.1) has a unique — meaning one and only one — solution, namely (1.5).

And that, barring a few minor complications that can crop up from time to time, is all that there is to the method of Gaussian Elimination! It is extraordinarily simple, but its importance cannot be overemphasized. Before exploring the relevant issues, it will help to reformulate our method in a more convenient matrix notation.

Exercises

1.1.1. Solve the following systems of linear equations by reducing to triangular form and then using Back Substitution.

(a) $\begin{aligned} x - y &= 7, \\ x + 2y &= 3; \end{aligned}$ (b) $\begin{aligned} 6u + v &= 5, \\ 3u - 2v &= 5; \end{aligned}$ (c) $\begin{aligned} p + q - r &= 0, \\ 2p - q + 3r &= 3, \\ -p - q &= 6; \end{aligned}$ (d) $\begin{aligned} 2u - v + 2w &= 2, \\ -u - v + 3w &= 1, \\ 3u - 2w &= 1; \end{aligned}$

(e) $\begin{aligned} 5x_1 + 3x_2 - x_3 &= 9, \\ 3x_1 + 2x_2 - x_3 &= 5, \\ x_1 + x_2 + x_3 &= -1; \end{aligned}$ (f) $\begin{aligned} x + z - 2w &= -3, \\ 2x - y + 2z - w &= -5, \\ -6y - 4z + 2w &= 2, \\ x + 3y + 2z - w &= 1; \end{aligned}$ (g) $\begin{aligned} 3x_1 + x_2 &= 1, \\ x_1 + 3x_2 + x_3 &= 1, \\ x_2 + 3x_3 + x_4 &= 1, \\ x_3 + 3x_4 &= 1. \end{aligned}$

1.1.2. How should the coefficients a, b, and c be chosen so that the system $ax + by + cz = 3$, $ax - y + cz = 1$, $x + by - cz = 2$, has the solution $x = 1$, $y = 2$ and $z = -1$?

♡ 1.1.3. The system $2x = -6$, $-4x + 3y = 3$, $x + 4y - z = 7$, is in *lower triangular form*. (a) Formulate a method of *Forward Substitution* to solve it. (*b*) What happens if you reduce the system to (upper) triangular form using the algorithm in this section? (*c*) Devise an algorithm that uses our linear system operation to reduce a system to lower triangular form and then solve it by Forward Substitution. (*d*) Check your algorithm by applying it to one or two of the systems in Exercise 1.1.1. Are you able to solve them in all cases?

1.2 Matrices and Vectors

A *matrix* is a rectangular array of numbers. Thus,

$$\begin{pmatrix} 1 & 0 & 3 \\ -2 & 4 & 1 \end{pmatrix}, \qquad \begin{pmatrix} \pi & 0 \\ e & \frac{1}{2} \\ -1 & .83 \\ \sqrt{5} & -\frac{4}{7} \end{pmatrix}, \qquad (\,.2 \quad -1.6 \quad .32\,), \qquad \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 3 \\ -2 & 5 \end{pmatrix},$$

are all examples of matrices. We use the notation

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{1.6}$$

for a general matrix of size $m \times n$ (read "m by n"), where m denotes the number of *rows* in A and n denotes the number of *columns*. Thus, the preceding examples of matrices have respective sizes 2×3, 4×2, 1×3, 2×1, and 2×2. A matrix is *square* if $m = n$, i.e., it has the same number of rows as columns. A *column vector* is an $m \times 1$ matrix, while a *row vector* is a $1 \times n$ matrix. As we shall see, column vectors are by far the more important of the two, and the term "vector" without qualification will always mean "column vector". A 1×1 matrix, which has but a single entry, is both a row and a column vector.

The number that lies in the i^{th} row and the j^{th} column of A is called the (i, j) *entry* of A, and is denoted by a_{ij}. The row index always appears first and the column index second.† Two matrices are equal, $A = B$, if and only if they have the same size, say $m \times n$, and *all* their entries are the same: $a_{ij} = b_{ij}$ for $i = 1, \dots, m$ and $j = 1, \dots, n$.

A general linear system of m equations in n unknowns will take the form

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \ \cdots \ + a_{1n} x_n &= b_1, \\ a_{21} x_1 + a_{22} x_2 + \ \cdots \ + a_{2n} x_n &= b_2, \\ \vdots \qquad\quad \vdots \qquad\quad & \ \ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \ \cdots \ + a_{mn} x_n &= b_m. \end{aligned} \tag{1.7}$$

As such, it is composed of three basic ingredients: the $m \times n$ *coefficient matrix* A, with entries a_{ij} as in (1.6), the column vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ containing the *unknowns*, and the column vector $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ containing *right-hand sides*. In our previous example,
$\begin{aligned} x + 2y + z &= 2, \\ 2x + 6y + z &= 7, \\ x + y + 4z &= 3, \end{aligned}$ the coefficient matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix}$ can be filled in, entry by entry, from the coefficients of the variables appearing in the equations; if a variable does not appear in an equation, the corresponding matrix entry is 0. The vector $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ lists the variables, while the entries of $\mathbf{b} = \begin{pmatrix} 2 \\ 7 \\ 3 \end{pmatrix}$ are the right-hand sides of the equations.

† In tensor analysis, [**1**], a sub- and super-script notation is adopted, with a^i_j denoting the (i, j) entry of the matrix A. This has certain advantages, but, to avoid possible confusion with powers, we shall stick with the simpler subscript notation throughout this text.

Remark. We will consistently use bold face lower case letters to denote vectors, and ordinary capital letters to denote general matrices.

Exercises

1.2.1. Let $A = \begin{pmatrix} -2 & 0 & 1 & 3 \\ -1 & 2 & 7 & -5 \\ 6 & -6 & -3 & 4 \end{pmatrix}$. (a) What is the size of A? (b) What is its $(2,3)$ entry? (c) $(3,1)$ entry? (d) 1^{st} row? (e) 2^{nd} column?

1.2.2. Write down examples of (a) a 3×3 matrix; (b) a 2×3 matrix; (c) a matrix with 3 rows and 4 columns; (d) a row vector with 4 entries; (e) a column vector with 3 entries; (f) a matrix that is both a row vector and a column vector.

1.2.3. For which values of x, y, z, w are the matrices $\begin{pmatrix} x+y & x-z \\ y+w & x+2w \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ equal?

1.2.4. For each of the systems in Exercise 1.1.1, write down the coefficient matrix A and the vectors $\mathbf{x}$ and $\mathbf{b}$.

1.2.5. Write out and solve the linear systems corresponding to the indicated matrix, vector of unknowns, and right-hand side. (a) $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -1 \\ -3 \end{pmatrix}$;
(b) $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$; (c) $A = \begin{pmatrix} 3 & 0 & -1 \\ -2 & -1 & 0 \\ 1 & 1 & -3 \end{pmatrix}$,
$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$; (d) $A = \begin{pmatrix} 1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 2 \\ 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 4 \\ 1 \\ 5 \end{pmatrix}$.

Matrix Arithmetic

Matrix arithmetic involves three basic operations: *matrix addition*, *scalar multiplication*, and *matrix multiplication*. First we define *addition* of matrices. You are allowed to add two matrices only if they are of the *same size*, and matrix addition is performed entry by entry. For example,

$$\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & -5 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 1 & 1 \end{pmatrix}.$$

Therefore, if A and B are $m\times n$ matrices, their sum $C = A+B$ is the $m\times n$ matrix whose entries are given by $c_{ij} = a_{ij} + b_{ij}$ for $i = 1,\ldots,m$ and $j = 1,\ldots,n$. When defined, matrix addition is commutative, $A + B = B + A$, and associative, $A + (B + C) = (A + B) + C$, just like ordinary addition.

A *scalar* is a fancy name for an ordinary number — the term merely distinguishes it from a vector or a matrix. For the time being, we will restrict our attention to real scalars and matrices with real entries, but eventually complex scalars and complex matrices must be dealt with. We will consistently identify a scalar $c \in \mathbb{R}$ with the 1×1 matrix $(\,c\,)$ in which it is the sole entry, and so will omit the redundant parentheses in the latter case. *Scalar multiplication* takes a scalar c and an $m\times n$ matrix A and computes the $m\times n$

matrix $B = c\,A$ by multiplying each entry of A by c. For example,

$$3\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ -3 & 0 \end{pmatrix}.$$

In general, $b_{ij} = c\,a_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Basic properties of scalar multiplication are summarized at the end of this section.

Finally, we define *matrix multiplication*. First, the product of a row vector $\mathbf{a}$ and a column vector $\mathbf{x}$ having the *same* number of entries is the *scalar* or 1×1 matrix defined by the following rule:

$$\mathbf{a}\,\mathbf{x} = (\, a_1 \;\, a_2 \; \ldots \; a_n \,) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + a_2 x_2 + \; \cdots \; + a_n x_n = \sum_{k=1}^{n} a_k x_k. \tag{1.8}$$

More generally, if A is an $m \times n$ matrix and B is an $n \times p$ matrix, so that the number of *columns* in A equals the number of *rows* in B, then the matrix product $C = A\,B$ is defined as the $m \times p$ matrix whose (i,j) entry equals the vector product of the i^{th} row of A and the j^{th} column of B. Therefore,

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}. \tag{1.9}$$

Note that our restriction on the sizes of A and B guarantees that the relevant row and column vectors will have the same number of entries, and so their product is defined.

For example, the product of the coefficient matrix A and vector of unknowns $\mathbf{x}$ for our original system (1.1) is given by

$$A\,\mathbf{x} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ 2x + 6y + z \\ x + y + 4z \end{pmatrix}.$$

The result is a column vector whose entries reproduce the left-hand sides of the original linear system! As a result, we can rewrite the system

$$A\,\mathbf{x} = \mathbf{b} \tag{1.10}$$

as an equality between two column vectors. This result is general; a linear system (1.7) consisting of m equations in n unknowns can be written in the matrix form (1.10), where A is the $m \times n$ coefficient matrix (1.6), $\mathbf{x}$ is the $n \times 1$ column vector of unknowns, and $\mathbf{b}$ is the $m \times 1$ column vector containing the right-hand sides. This is one of the principal reasons for the non-evident definition of matrix multiplication. Component-wise multiplication of matrix entries turns out to be almost completely useless in applications.

Now, the bad news. Matrix multiplication is *not* commutative — that is, BA is not necessarily equal to AB. For example, BA may not be defined even when AB is. Even if both are defined, they may be different sized matrices. For example the product $s = \mathbf{r}\,\mathbf{c}$ of a row vector $\mathbf{r}$, a $1 \times n$ matrix, and a column vector $\mathbf{c}$, an $n \times 1$ matrix with the same number of entries, is a 1×1 matrix, or scalar, whereas the reversed product $C = \mathbf{c}\,\mathbf{r}$ is an $n \times n$ matrix. For instance,

$$(\,1 \;\; 2\,)\begin{pmatrix} 3 \\ 0 \end{pmatrix} = 3, \qquad \text{whereas} \qquad \begin{pmatrix} 3 \\ 0 \end{pmatrix}(\,1 \;\; 2\,) = \begin{pmatrix} 3 & 6 \\ 0 & 0 \end{pmatrix}.$$

In computing the latter product, don't forget that we multiply the *rows* of the first matrix by the *columns* of the second, each of which has but a single entry. Moreover, even if the matrix products AB and BA have the same size, which requires both A and B to be square matrices, we may still have $AB \neq BA$. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 5 \\ -4 & 11 \end{pmatrix} \neq \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

On the other hand, matrix multiplication is associative, so $A(BC) = (AB)C$ whenever A has size $m \times n$, B has size $n \times p$, and C has size $p \times q$; the result is a matrix of size $m \times q$. The proof of associativity is a tedious computation based on the definition of matrix multiplication that, for brevity, we omit.† Consequently, the one difference between matrix algebra and ordinary algebra is that you need to be careful not to change the order of multiplicative factors without proper justification.

Since matrix multiplication acts by multiplying rows by columns, one can compute the columns in a matrix product AB by multiplying the matrix A and the individual columns of B. For example, the two columns of the matrix product

$$\begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 8 & 6 \end{pmatrix}$$

are obtained by multiplying the first matrix with the individual columns of the second:

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

In general, if we use $\mathbf{b}_k$ to denote the k^{th} column of B, then

$$AB = A\begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_p \end{pmatrix} = \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \ldots & A\mathbf{b}_p \end{pmatrix}, \tag{1.11}$$

indicating that the k^{th} column of their matrix product is $A\mathbf{b}_k$.

There are two important special matrices. The first is the *zero matrix*, all of whose entries are 0. We use $\mathrm{O}_{m \times n}$ to denote the $m \times n$ zero matrix, often written as just O if the size is clear from the context. The zero matrix is the additive unit, so $A + \mathrm{O} = A = \mathrm{O} + A$ when O has the same size as A. In particular, we will use a bold face $\mathbf{0}$ to denote a column vector with all zero entries, i.e., $\mathrm{O}_{1 \times n}$.

The role of the multiplicative unit is played by the square *identity matrix*

$$\mathrm{I} = \mathrm{I}_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

of size $n \times n$. The entries along the *main diagonal* — which runs from top left to bottom right — are equal to 1, while the *off-diagonal* entries are all 0. As you can check, if A is

† A much simpler — but more abstract proof can be found in Exercise 7.1.45.

Basic Matrix Arithmetic

Matrix Addition:	Commutativity	$A+B=B+A$
	Associativity	$(A+B)+C=A+(B+C)$
	Zero Matrix	$A+\mathrm{O}=A=\mathrm{O}+A$
	Additive Inverse	$A+(-A)=\mathrm{O}, \quad -A=(-1)A$
Scalar Multiplication:	Associativity	$c\,(d\,A)=(c\,d)\,A$
	Distributivity	$c\,(A+B)=(c\,A)+(c\,B)$ $(c+d)\,A=(c\,A)+(d\,A)$
	Unit Scalar	$1\,A=A$
	Zero Scalar	$0\,A=\mathrm{O}$
Matrix Multiplication:	Associativity	$(A\,B)\,C=A\,(B\,C)$
	Distributivity	$A\,(B+C)=A\,B+A\,C,$ $(A+B)\,C=A\,C+B\,C,$
	Compatibility	$c\,(A\,B)=(c\,A)\,B=A\,(c\,B)$
	Identity Matrix	$A\,\mathrm{I}=A=\mathrm{I}\,A$
	Zero Matrix	$A\,\mathrm{O}=\mathrm{O}, \quad \mathrm{O}\,A=\mathrm{O}$

any $m\times n$ matrix, then $\mathrm{I}_m\,A=A=A\,\mathrm{I}_n$. We will sometimes write the preceding equation as just $\mathrm{I}\,A=A=A\,\mathrm{I}$, since each matrix product is well-defined for exactly one size of identity matrix.

The identity matrix is a particular example of a *diagonal matrix*. In general, a square matrix A is diagonal if all its off-diagonal entries are zero: $a_{ij}=0$ for all $i\neq j$. We will sometimes write $D=\operatorname{diag}(c_1,\ldots,c_n)$ for the $n\times n$ diagonal matrix with diagonal entries $d_{ii}=c_i$. Thus, $\operatorname{diag}(1,3,0)$ refers to the diagonal matrix $\begin{pmatrix}1&0&0\\0&3&0\\0&0&0\end{pmatrix}$, while the 4×4 identity matrix can be written as

$$\mathrm{I}_4=\operatorname{diag}(1,1,1,1)=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

Let us conclude this section by summarizing the basic properties of matrix arithmetic. In the accompanying table, A,B,C are matrices; c,d are scalars; O is a zero matrix; and I is an identity matrix. All matrices are assumed to have the correct sizes so that the indicated operations are defined.

Exercises

1.2.6. (*a*) Write down the 5×5 identity and zero matrices. (*b*) Write down their sum and their product. Does the order of multiplication matter?

1.2.7. Consider the matrices $A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 4 & -2 \\ 3 & 0 & 6 \end{pmatrix}$, $B = \begin{pmatrix} -6 & 0 & 3 \\ 4 & 2 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 3 \\ -3 & -4 \\ 1 & 2 \end{pmatrix}$.
Compute the indicated combinations where possible. (*a*) $3A - B$, (*b*) AB, (*c*) BA, (*d*) $(A+B)C$, (*e*) $A+BC$, (*f*) $A+2CB$, (*g*) $BCB-\mathrm{I}$, (*h*) $A^2-3A+\mathrm{I}$, (*i*) $(B-\mathrm{I})(C+\mathrm{I})$.

1.2.8. Which of the following pairs of matrices commute under matrix multiplication?
(*a*) $\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 5 & 0 \end{pmatrix}$, (*b*) $\begin{pmatrix} 3 & -1 \\ 0 & 2 \\ 1 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 2 & -2 \\ 5 & 2 & 4 \end{pmatrix}$, (*c*) $\begin{pmatrix} 3 & 0 & -1 \\ -2 & -1 & 2 \\ 2 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ 2 & 0 & -1 \end{pmatrix}$.

1.2.9. List the diagonal entries of $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$.

1.2.10. Write out the following diagonal matrices: (*a*) $\operatorname{diag}(1, 0, -1)$, (*b*) $\operatorname{diag}(2, -2, 3, -3)$.

1.2.11. *True or false*: (*a*) The sum of two diagonal matrices of the same size is a diagonal matrix. (*b*) The product is also diagonal.

♡ 1.2.12. (*a*) Show that if $D = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ is a 2×2 diagonal matrix with $a \neq b$, then the only matrices that commute (under matrix multiplication) with D are other 2×2 diagonal matrices. (*b*) What if $a = b$? (*c*) Find all matrices that commute with $D = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$, where a, b, c are all different. (*d*) Answer the same question for the case when $a \neq b = c$. (*e*) Prove that a matrix A commutes with an $n \times n$ diagonal matrix D with all *distinct* diagonal entries if and only if A is a diagonal matrix.

1.2.13. Show that the matrix products AB and BA have the same size if and only if A and B are square matrices of the same size.

1.2.14. Find all matrices B that commute (under matrix multiplication) with $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$.

1.2.15. (*a*) Show that, if A, B are commuting square matrices, then $(A + B)^2 = A^2 + 2AB + B^2$. (*b*) Find a pair of 2×2 matrices A, B such that $(A + B)^2 \neq A^2 + 2AB + B^2$.

1.2.16. Show that if the matrices A and B commute, then they necessarily are both square and the same size.

1.2.17. Let A be an $m \times n$ matrix. What are the permissible sizes for the zero matrices appearing in the identities $A\,\mathrm{O} = \mathrm{O}$ and $\mathrm{O}\,A = \mathrm{O}$?

1.2.18. Let A be an $m \times n$ matrix and let c be a scalar. Show that if $cA = \mathrm{O}$, then either $c = 0$ or $A = \mathrm{O}$.

1.2.19. *True or false*: If $AB = \mathrm{O}$ then either $A = \mathrm{O}$ or $B = \mathrm{O}$.

1.2.20. *True or false*: If A, B are square matrices of the same size, then
$$A^2 - B^2 = (A + B)(A - B).$$

1.2.21. Prove that $A\mathbf{v} = \mathbf{0}$ for every vector $\mathbf{v}$ (with the appropriate number of entries) if and only if $A = \mathrm{O}$ is the zero matrix. *Hint*: If you are stuck, first try to find a proof when A is a small matrix, e.g., of size 2×2.

1.2.22. (*a*) Under what conditions is the square A^2 of a matrix defined? (*b*) Show that A and A^2 commute. (*c*) How many matrix multiplications are needed to compute A^n?

1.2.23. Find a nonzero matrix $A \neq \mathrm{O}$ such that $A^2 = \mathrm{O}$.

♢ 1.2.24. Let A have a row all of whose entries are zero. (*a*) Explain why the product AB also has a zero row. (*b*) Find an example where BA does not have a zero row.

1.2.25. (a) Find all solutions $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ to the matrix equation $AX = \mathrm{I}$ when $A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$. (b) Find all solutions to $XA = \mathrm{I}$. Are they the same?

1.2.26. (a) Find all solutions $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ to the matrix equation $AX = B$ when $A = \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$. (b) Find all solutions to $XA = B$. Are they the same?

1.2.27. (a) Find all solutions $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ to the matrix equation $AX = XB$ when $A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}$. (b) Can you find a pair of nonzero matrices $A \neq B$ such that the matrix equation $AX = XB$ has a nonzero solution $X \neq \mathrm{O}$?

1.2.28. Let A be a matrix and c a scalar. Find all solutions to the matrix equation $cA = \mathrm{I}$.

♢ 1.2.29. Let $\mathbf{e}$ be the $1 \times m$ row vector all of whose entries are equal to 1. (a) Show that if A is an $m \times n$ matrix, then the i^{th} entry of the product $\mathbf{v} = \mathbf{e}A$ is the j^{th} *column sum* of A, meaning the sum of all the entries in its j^{th} row. (b) Let W denote the $m \times m$ matrix whose diagonal entries are equal to $\dfrac{1-m}{m}$ and whose off-diagonal entries are all equal to $\dfrac{1}{m}$. Prove that the column sums of $B = WA$ are all zero. (c) Check both results when $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 3 \\ -4 & 5 & -1 \end{pmatrix}$. **Remark.** If the rows of A represent experimental data values, then the entries of $\dfrac{1}{m}\mathbf{e}A$ represent the means or averages of the data values, while $B = WA$ corresponds to data that has been normalized to have mean 0; see Section 8.8.

♡ 1.2.30. The *commutator* of two matrices A, B, is defined to be the matrix

$$C = [\,A, B\,] = AB - BA. \tag{1.12}$$

(a) Explain why $[\,A, B\,]$ is defined if and only if A and B are square matrices of the same size. (b) Show that A and B commute under matrix multiplication if and only if $[\,A, B\,] = \mathrm{O}$. (c) Compute the commutator of the following matrices: (i) $\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ -2 & 0 \end{pmatrix}$; (ii) $\begin{pmatrix} -1 & 3 \\ 3 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 7 \\ 7 & 1 \end{pmatrix}$; (iii) $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$;

(d) Prove that the commutator is (i) *Bilinear*: $[\,cA + dB, C\,] = c[\,A, C\,] + d[\,B, C\,]$ and $[\,A, cB + dC\,] = c[\,A, B\,] + d[\,A, C\,]$ for any scalars c, d; (ii) *Skew-symmetric*: $[\,A, B\,] = -[\,B, A\,]$; (iii) satisfies the the *Jacobi identity*:

$$[\,[\,A, B\,], C\,] + [\,[\,C, A\,], B\,] + [\,[\,B, C\,], A\,] = \mathrm{O},$$

for any square matrices A, B, C of the same size.

Remark. The commutator plays a very important role in geometry, symmetry, and quantum mechanics. See Section 10.4 as well as [**54**, **60**, **93**] for further developments.

♢ 1.2.31. The *trace* of a $n \times n$ matrix $A \in \mathcal{M}_{n\times n}$ is defined to be the sum of its diagonal entries: $\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$. (a) Compute the trace of (i) $\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 3 & 2 \\ -1 & 0 & 1 \\ -4 & 3 & -1 \end{pmatrix}$.

(b) Prove that $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$. (c) Prove that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. (d) Prove that the commutator matrix $C = AB - BA$ has zero trace: $\operatorname{tr} C = 0$. (e) Is part (c) valid if A has size $m \times n$ and B has size $n \times m$? (f) Prove that $\operatorname{tr}(ABC) = \operatorname{tr}(CAB) = \operatorname{tr}(BCA)$. On the other hand, find an example where $\operatorname{tr}(ABC) \neq \operatorname{tr}(ACB)$.

♢ 1.2.32. Prove that matrix multiplication is associative: $A(BC) = (AB)C$ when defined.

♢ 1.2.33. Justify the following alternative formula for multiplying a matrix A and a column vector $\mathbf{x}$:

$$A\mathbf{x} = x_1 \mathbf{c}_1 + x_2 \mathbf{c}_2 + \cdots + x_n \mathbf{c}_n, \tag{1.13}$$

where $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are the columns of A and $x_1, \ldots, x_n$ the entries of $\mathbf{x}$.

♡ 1.2.34. The basic definition of matrix multiplication AB tells us to multiply rows of A by columns of B. Remarkably, if you suitably interpret the operation, you can also compute AB by multiplying columns of A by rows of B! Suppose A is an $m \times n$ matrix with columns $\mathbf{c}_1, \ldots, \mathbf{c}_n$. Suppose B is an $n \times p$ matrix with rows $\mathbf{r}_1, \ldots, \mathbf{r}_n$. Then we claim that

$$AB = \mathbf{c}_1 \mathbf{r}_1 + \mathbf{c}_2 \mathbf{r}_2 + \cdots + \mathbf{c}_n \mathbf{r}_n, \tag{1.14}$$

where each summand is a matrix of size $m \times p$. (*a*) Verify that the particular case

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}\begin{pmatrix} 0 & -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \end{pmatrix}\begin{pmatrix} 2 & 3 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & -3 \end{pmatrix} + \begin{pmatrix} 4 & 6 \\ 8 & 12 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 9 \end{pmatrix}$$

agrees with the usual method for computing the matrix product. (*b*) Use this method to compute the matrix products (*i*) $\begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & 0 \end{pmatrix}$, (*ii*) $\begin{pmatrix} 1 & -2 & 0 \\ -3 & -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 5 \\ -3 & 0 \\ 1 & -1 \end{pmatrix}$,

(*iii*) $\begin{pmatrix} 3 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & -5 \end{pmatrix}\begin{pmatrix} 2 & 3 & 0 \\ 3 & -1 & 4 \\ 0 & 4 & 1 \end{pmatrix}$, and verify that you get the same answer as that obtained by the traditional method. (*c*) Explain why (1.13) is a special case of (1.14). (*d*) Prove that (1.14) gives the correct formula for the matrix product.

♡ 1.2.35. *Matrix polynomials*. Let $p(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0$ be a polynomial function. If A is a square matrix, we define the corresponding *matrix polynomial* $p(A) = c_n A^n + c_{n-1} A^{n-1} + \cdots + c_1 A + c_0 \mathrm{I}$; the constant term becomes a scalar multiple of the identity matrix. For instance, if $p(x) = x^2 - 2x + 3$, then $p(A) = A^2 - 2A + 3\,\mathrm{I}$. (*a*) Write out the matrix polynomials $p(A), q(A)$ when $p(x) = x^3 - 3x + 2$, $q(x) = 2x^2 + 1$. (*b*) Evaluate $p(A)$ and $q(A)$ when $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$. (*c*) Show that the matrix product $p(A)\,q(A)$ is the matrix polynomial corresponding to the product polynomial $r(x) = p(x)\,q(x)$. (*d*) *True or false*: If $B = p(A)$ and $C = q(A)$, then $BC = CB$. Check your answer in the particular case of part (*b*).

♡ 1.2.36. A *block matrix* has the form $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ in which A, B, C, D are matrices with respective sizes $i \times k$, $i \times l$, $j \times k$, $j \times l$. (*a*) What is the size of M? (*b*) Write out the block matrix M when $A = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}$.

(*c*) Show that if $N = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$ is a block matrix whose blocks have the same size as those of M, then $M + N = \begin{pmatrix} A+P & B+Q \\ C+R & D+S \end{pmatrix}$, i.e., matrix addition can be done in blocks.

(*d*) Show that if $P = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$ has blocks of a compatible size, the matrix product is $MP = \begin{pmatrix} AX + BZ & AY + BW \\ CX + DZ & CY + DW \end{pmatrix}$. Explain what "compatible" means. (*e*) Write down a compatible block matrix P for the matrix M in part (*b*). Then validate the block matrix product identity of part (*d*) for your chosen matrices.

♡ 1.2.37. The matrix S is said to be a *square root* of the matrix A if $S^2 = A$. (*a*) Show that $S = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$ is a square root of the matrix $A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$. Can you find another square root of A? (*b*) Explain why only square matrices can have a square root. (*c*) Find all real square roots of the 2×2 identity matrix $\mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. (*d*) Does $-\mathrm{I} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ have a real square root?

1.3 Gaussian Elimination — Regular Case

With the basic matrix arithmetic operations in hand, let us now return to our primary task. The goal is to develop a systematic method for solving linear systems of equations. While we could continue to work directly with the equations, matrices provide a convenient alternative that begins by merely shortening the amount of writing, but ultimately leads to profound insight into the structure of linear systems and their solutions.

We begin by replacing the system (1.7) by its matrix constituents. It is convenient to ignore the vector of unknowns, and form the *augmented matrix*

$$M = \left(A \mid \mathbf{b} \right) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right), \tag{1.15}$$

which is an $m \times (n+1)$ matrix obtained by tacking the right-hand side vector onto the original coefficient matrix. The extra vertical line is included just to remind us that the last column of this matrix plays a special role. For example, the augmented matrix for the system (1.1), i.e.,

$$\begin{aligned} x + 2y + z &= 2, \\ 2x + 6y + z &= 7, \\ x + y + 4z &= 3, \end{aligned} \qquad \text{is} \qquad M = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 6 & 1 & 7 \\ 1 & 1 & 4 & 3 \end{array} \right). \tag{1.16}$$

Note that one can immediately recover the equations in the original linear system from the augmented matrix. Since operations on equations also affect their right-hand sides, keeping track of everything is most easily done through the augmented matrix.

For the time being, we will concentrate our efforts on linear systems that have the same number, n, of equations as unknowns. The associated coefficient matrix A is square, of size $n \times n$. The corresponding augmented matrix $M = \left(A \mid \mathbf{b} \right)$ then has size $n \times (n+1)$.

The matrix operation that assumes the role of Linear System Operation #1 is:

Elementary Row Operation #1:

Add a scalar multiple of one row of the augmented matrix to another row.

For example, if we add -2 times the first row of the augmented matrix (1.16) to the second row, the result is the row vector

$$-2\,(\,1 \quad 2 \quad 1 \quad 2\,) \;+\; (\,2 \quad 6 \quad 1 \quad 7\,) = (\,0 \quad 2 \quad -1 \quad 3\,).$$

The result can be recognized as the second row of the modified augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -1 & 3 \\ 1 & 1 & 4 & 3 \end{array} \right) \tag{1.17}$$

that corresponds to the first equivalent system (1.2). When elementary row operation #1 is performed, it is critical that the result replaces the row being added to — *not* the row being multiplied by the scalar. Notice that the elimination of a variable in an equation — in this case, the first variable in the second equation — amounts to making its entry in the coefficient matrix equal to zero.

We shall call the $(1,1)$ entry of the coefficient matrix the *first pivot*. The precise definition of pivot will become clear as we continue; the one key requirement is that *a pivot must always be nonzero*. Eliminating the first variable x from the second and third equations amounts to making all the matrix entries in the column below the pivot equal to zero. We have already done this with the $(2,1)$ entry in (1.17). To make the $(3,1)$ entry equal to zero, we subtract (that is, add -1 times) the first row from the last row. The resulting augmented matrix is

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & -1 & 3 & 1 \end{array}\right),$$

which corresponds to the system (1.3). The *second pivot* is the $(2,2)$ entry of this matrix, which is 2, and is the coefficient of the second variable in the second equation. Again, the pivot must be nonzero. We use the elementary row operation of adding $\frac{1}{2}$ of the second row to the third row to make the entry below the second pivot equal to 0; the result is the augmented matrix

$$N = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & \frac{5}{2} & \frac{5}{2} \end{array}\right)$$

that corresponds to the triangular system (1.4). We write the final augmented matrix as

$$N = \bigl(\, U \mid \mathbf{c} \,\bigr), \qquad \text{where} \qquad U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & \frac{5}{2} \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} 2 \\ 3 \\ \frac{5}{2} \end{pmatrix}.$$

The corresponding linear system has vector form

$$U\mathbf{x} = \mathbf{c}. \tag{1.18}$$

Its coefficient matrix U is *upper triangular*, which means that all its entries below the main diagonal are zero: $u_{ij} = 0$ whenever $i > j$. The three nonzero entries on its diagonal, $1, 2, \frac{5}{2}$, including the last one in the $(3,3)$ slot, are the three pivots. Once the system has been reduced to triangular form (1.18), we can easily solve it by Back Substitution.

The preceding algorithm for solving a linear system of n equations in n unknowns is known as *regular Gaussian Elimination*. A square matrix A will be called *regular*† if the algorithm successfully reduces it to upper triangular form U with all non-zero pivots on the diagonal. In other words, for regular matrices, as the algorithm proceeds, each successive pivot appearing on the diagonal must be nonzero; otherwise, the matrix is not regular. We then use the pivot row to make all the entries lying in the column below the pivot equal to zero through elementary row operations. The solution is found by applying Back Substitution to the resulting triangular system.

---

† Strangely, there is no commonly accepted term to describe this kind of matrix. For lack of a better alternative, we propose to use the adjective "regular" in the sequel.

Gaussian Elimination — Regular Case

```
start
    for j = 1 to n
        if m_jj = 0, stop; print "A is not regular"
        else for i = j + 1 to n
            set l_ij = m_ij / m_jj
            add −l_ij times row j of M to row i of M
        next i
    next j
end
```

Let us state this algorithm in the form of a program, written in a general "pseudocode" that can be easily translated into any specific language, e.g., C++, FORTRAN, JAVA, MAPLE, MATHEMATICA, MATLAB. In accordance with the usual programming convention, the same letter $M = (m_{ij})$ will be used to denote the current augmented matrix at each stage in the computation, keeping in mind that its entries will change as the algorithm progresses. We initialize $M = \left(A \mid \mathbf{b} \right)$. The final output of the program, assuming A is regular, is the augmented matrix $M = \left(U \mid \mathbf{c} \right)$, where U is the upper triangular matrix whose diagonal entries are the pivots, while $\mathbf{c}$ is the resulting vector of right-hand sides in the triangular system $U\mathbf{x} = \mathbf{c}$.

For completeness, let us include the pseudocode program for Back Substitution. The input to this program is the upper triangular matrix U and the right-hand side vector $\mathbf{c}$ that results from the Gaussian Elimination pseudocode program, which produces $M = \left(U \mid \mathbf{c} \right)$. The output of the Back Substitution program is the solution vector $\mathbf{x}$ to the triangular system $U\mathbf{x} = \mathbf{c}$, which is the *same* as the solution to the original linear system $A\mathbf{x} = \mathbf{b}$.

Back Substitution

```
start
    set x_n = c_n / u_nn
    for i = n − 1 to 1 with increment −1
        set x_i = (1/u_ii) ( c_i − Σ_{j=1}^{i+1} u_ij x_j )
    next j
end
```

Exercises

1.3.1. Solve the following linear systems by Gaussian Elimination. (*a*) $\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$,

(b) $\begin{pmatrix} 6 & 1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$, (c) $\begin{pmatrix} 2 & 1 & 2 \\ -1 & 3 & 3 \\ 4 & -3 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 7 \end{pmatrix}$,

(d) $\begin{pmatrix} 5 & 3 & -1 \\ 3 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 9 \\ 5 \\ -1 \end{pmatrix}$, (e) $\begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 3 \\ -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix}$,

(f) $\begin{pmatrix} -1 & 1 & 1 & 0 \\ 2 & -1 & 0 & 1 \\ 1 & 0 & 2 & 3 \\ 0 & 1 & -1 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, (g) $\begin{pmatrix} 2 & -3 & 1 & 1 \\ 1 & -1 & -2 & -1 \\ 3 & -2 & 1 & 2 \\ 1 & 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 5 \\ 3 \end{pmatrix}$.

1.3.2. Write out the augmented matrix for the following linear systems. Then solve the system by first applying elementary row operations of type #1 to place the augmented matrix in upper triangular form, followed by Back Substitution.

(a) $\begin{aligned} x_1 + 7x_2 &= 4, \\ -2x_1 - 9x_2 &= 2. \end{aligned}$ (b) $\begin{aligned} 3z - 5w &= -1, \\ 2z + w &= 8. \end{aligned}$ (c) $\begin{aligned} x - 2y + z &= 0, \\ 2y - 8z &= 8, \\ -4x + 5y + 9z &= -9. \end{aligned}$

(d) $\begin{aligned} p + 4q - 2r &= 1, \\ -2p - 3r &= -7, \\ 3p - 2q + 2r &= -1. \end{aligned}$ (e) $\begin{aligned} x_1 - 2x_3 &= -1, \\ x_2 - x_4 &= 2, \\ -3x_2 + 2x_3 &= 0, \\ -4x_1 + 7x_4 &= -5. \end{aligned}$ (f) $\begin{aligned} -x + 3y - z + w &= -2, \\ x - y + 3z - w &= 0, \\ y - z + 4w &= 7, \\ 4x - y + z &= 5. \end{aligned}$

1.3.3. For each of the following augmented matrices write out the corresponding linear system of equations. Solve the system by applying Gaussian Elimination to the augmented matrix.

(a) $\left(\begin{array}{cc|c} 3 & 2 & 2 \\ -4 & -3 & -1 \end{array}\right)$, (b) $\left(\begin{array}{ccc|c} 1 & 2 & 0 & -3 \\ -1 & 2 & 1 & -6 \\ -2 & 0 & -3 & 1 \end{array}\right)$, (c) $\left(\begin{array}{cccc|c} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 \\ 0 & -1 & 2 & -1 & 1 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right)$.

1.3.4. Which of the following matrices are regular? (a) $\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} 0 & -1 \\ 3 & -2 \end{pmatrix}$,

(c) $\begin{pmatrix} 3 & -2 & 1 \\ -1 & 4 & -3 \\ 3 & -2 & 5 \end{pmatrix}$, (d) $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}$, (e) $\begin{pmatrix} 1 & 3 & -3 & 0 \\ -1 & 0 & -1 & 2 \\ 3 & 3 & -6 & 1 \\ 2 & 3 & -3 & 5 \end{pmatrix}$.

1.3.5. The techniques that are developed for solving linear systems are also applicable to systems with complex coefficients, whose solutions may also be complex. Use Gaussian Elimination to solve the following complex linear systems.

(a) $\begin{aligned} -\,\mathrm{i}\,x_1 + (1 + \mathrm{i}\,)x_2 &= -1, \\ (1 - \mathrm{i}\,)x_1 + x_2 &= -3\,\mathrm{i}\,. \end{aligned}$ (b) $\begin{aligned} \mathrm{i}\,x + (1 - \mathrm{i}\,)z &= 2\,\mathrm{i}\,, \\ 2\,\mathrm{i}\,y + (1 + \mathrm{i}\,)z &= 2, \\ -x + 2\,\mathrm{i}\,y + \mathrm{i}\,z &= 1 - 2\,\mathrm{i}\,. \end{aligned}$

(c) $\begin{aligned} (1 - \mathrm{i}\,)x + 2y &= \mathrm{i}\,, \\ -\,\mathrm{i}\,x + (1 + \mathrm{i}\,)y &= -1. \end{aligned}$ (d) $\begin{aligned} (1 + \mathrm{i}\,)x + \mathrm{i}\,y + (2 + 2\,\mathrm{i}\,)z &= 0, \\ (1 - \mathrm{i}\,)x + 2y + \mathrm{i}\,z &= 0, \\ (3 - 3\,\mathrm{i}\,)x + \mathrm{i}\,y + (3 - 11\,\mathrm{i}\,)z &= 6. \end{aligned}$

1.3.6. (a) Write down an example of a system of 5 linear equations in 5 unknowns with regular diagonal coefficient matrix. (b) Solve your system. (c) Explain why solving a system whose coefficient matrix is diagonal is very easy.

1.3.7. Find the equation of the parabola $y = ax^2 + bx + c$ that goes through the points $(1, 6), (2, 4)$, and $(3, 0)$.

♢ 1.3.8. A linear system is called *homogeneous* if all the right-hand sides are zero, and so takes the matrix form $A\mathbf{x} = \mathbf{0}$. Explain why the solution to a homogeneous system with regular coefficient matrix is $\mathbf{x} = \mathbf{0}$.

1.3.9. Under what conditions do two 2×2 upper triangular matrices commute?

1.3.10. A matrix is called *lower triangular* if all entries above the diagonal are zero. Show that a matrix is both lower and upper triangular if and only if it is a diagonal matrix.

♢ 1.3.11. A square matrix is called *strictly lower* (*upper*) *triangular* if all entries on or above (below) the main diagonal are 0. (*a*) Prove that every square matrix can be uniquely written as a sum $A = L + D + U$, with L strictly lower triangular, D diagonal, and U strictly upper triangular. (*b*) Decompose $A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & -4 & 2 \\ -2 & 0 & 5 \end{pmatrix}$ in this manner.

♢ 1.3.12. A square matrix N is called *nilpotent* if $N^k = \mathrm{O}$ for some $k \geq 1$.
(*a*) Show that $N = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ is nilpotent. (*b*) Show that every strictly upper triangular matrix, as defined in Exercise 1.3.11, is nilpotent. (*c*) Find a nilpotent matrix which is neither lower nor upper triangular.

♢ 1.3.13. A square matrix W is called *unipotent* if $N = W - \mathrm{I}$ is nilpotent, as in Exercise 1.3.12, so $(W - \mathrm{I})^k = \mathrm{O}$ for some $k \geq 1$. (*a*) Show that every lower or upper triangular matrix is unipotent if and only if it is unitriangular, meaning its diagonal entries are all equal to 1. (*b*) Find a unipotent matrix which is neither lower nor upper triangular.

1.3.14. A square matrix P is called *idempotent* if $P^2 = P$. (*a*) Find all 2×2 idempotent upper triangular matrices. (*b*) Find all 2×2 idempotent matrices.

Elementary Matrices

A key observation is that elementary row operations can, in fact, be realized by matrix multiplication. To this end, we introduce the first type of "elementary matrix". (Later we will meet two other types of elementary matrix, corresponding to the other two kinds of elementary row operation.)

Definition 1.1. The *elementary matrix* associated with an elementary row operation for m-rowed matrices is the $m \times m$ matrix obtained by applying the row operation to the $m \times m$ identity matrix I_m.

For example, applying the elementary row operation that adds -2 times the first row to the second row of the 3×3 identity matrix $\mathrm{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ results in the corresponding elementary matrix $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. We claim that, if A is *any* 3–rowed matrix, then multiplying $E_1 A$ has the same effect as the given elementary row operation. For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{pmatrix},$$

which you may recognize as the first elementary row operation we used to solve our illustrative example. If we set

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}, \tag{1.19}$$

then multiplication by E_1 will subtract twice the first row from the second row, multiplication by E_2 will subtract the first row from the third row, and multiplication by E_3 will add $\frac{1}{2}$ the second row to the third row — precisely the row operations used to place our original system in triangular form. Therefore, performing them in the correct order, we conclude that when

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix}, \qquad \text{then} \qquad E_3\, E_2\, E_1\, A = U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & \frac{5}{2} \end{pmatrix}. \tag{1.20}$$

The reader is urged to check this by directly multiplying the indicated matrices. Keep in mind that the associative property of matrix multiplication allows us to compute the above matrix product in any convenient order:

$$E_3\, E_2\, E_1\, A = E_3\,(E_2\,(E_1\, A)) = ((E_3\, E_2)\, E_1)\, A = (E_3\,(E_2\, E_1))\, A = (E_3\, E_2)\,(E_1\, A) = \ \cdots\ ,$$

making sure that the overall left to right order of the matrices is maintained, since the matrix products are usually *not* commutative.

In general, then, an $m \times m$ *elementary matrix E of the first type* will have all 1's on the diagonal, one nonzero entry c in some off-diagonal position (i, j), with $i \neq j$, and all other entries equal to zero. If A is any $m \times n$ matrix, then the matrix product $E\,A$ is equal to the matrix obtained from A by the elementary row operation adding c times row j to row i. (Note that the order of i and j is reversed.)

To undo the operation of adding c times row j to row i, we must perform the inverse row operation that subtracts c (or, equivalently, adds $-c$) times row j from row i. The corresponding *inverse elementary matrix* again has 1's along the diagonal and $-c$ in the (i, j) slot. Let us denote the inverses of the particular elementary matrices (1.19) by L_i, so that, according to our general rule,

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \qquad L_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}. \tag{1.21}$$

Note that the products

$$L_1\, E_1 = L_2\, E_2 = L_3\, E_3 = \mathrm{I} \tag{1.22}$$

yield the 3×3 identity matrix, reflecting the fact that the matrices represent mutually inverse row operations. (A more thorough discussion of matrix inverses will be postponed until Section 1.5.)

The product of the latter three elementary matrices (1.21) is equal to

$$L = L_1\, L_2\, L_3 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -\frac{1}{2} & 1 \end{pmatrix}. \tag{1.23}$$

The matrix L is called a *lower unitriangular* matrix, where "lower triangular" means that all the entries above the main diagonal are 0, while "uni-", which is short for "unipotent" as defined in Exercise 1.3.13, imposes the requirement that all the entries on the diagonal are equal to 1. Observe that the entries of L below the diagonal are the same as the corresponding nonzero entries in the L_i. This is a general fact that holds when the lower triangular elementary matrices are multiplied in the correct order. More generally, the following elementary consequence of the laws of matrix multiplication will be used extensively.

Lemma 1.2. If L and $\widehat{L}$ are lower triangular matrices of the same size, so is their product $L\widehat{L}$. If they are both lower unitriangular, so is their product. Similarly, if $U, \widehat{U}$ are upper (uni)triangular matrices, so is their product $U\widehat{U}$.

The LU Factorization

We have almost arrived at our first important result. Let us compute the product of the matrices L and U in (1.20), (1.23). Using associativity of matrix multiplication, equations (1.22), and the basic property of the identity matrix I, we conclude that

$$\begin{aligned} LU = (L_1 L_2 L_3)(E_3 E_2 E_1 A) &= L_1 L_2 (L_3 E_3) E_2 E_1 A = L_1 L_2 \,\mathrm{I}\, E_2 E_1 A \\ &= L_1 (L_2 E_2) E_1 A = L_1 \,\mathrm{I}\, E_1 A = (L_1 E_1) A = \mathrm{I}\, A = A. \end{aligned}$$

In other words, we have *factored* the coefficient matrix $A = LU$ into a product of a lower unitriangular matrix L and an upper triangular matrix U with the nonzero pivots on its main diagonal. By similar reasoning, the same holds true for any regular square matrix.

Theorem 1.3. A matrix A is regular if and only if it can be factored

$$A = LU, \tag{1.24}$$

where L is a lower unitriangular matrix, having all 1's on the diagonal, and U is upper triangular with nonzero diagonal entries, which are the pivots of A. The nonzero off-diagonal entries l_{ij} for $i > j$ appearing in L prescribe the elementary row operations that bring A into upper triangular form; namely, one subtracts l_{ij} times row j from row i at the appropriate step of the Gaussian Elimination process.

In practice, to find the LU factorization of a square matrix A, one applies the regular Gaussian Elimination algorithm to reduce A to its upper triangular form U. The entries of L can be filled in during the course of the calculation with the negatives of the multiples used in the elementary row operations. If the algorithm fails to be completed, which happens whenever zero appears in any diagonal pivot position, then the original matrix is *not* regular, and does *not* have an LU factorization.

Example 1.4. Let us compute the LU factorization of the matrix $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 5 & 2 \\ 2 & -2 & 0 \end{pmatrix}$.

Applying the Gaussian Elimination algorithm, we begin by adding -2 times the first row to the second row, and then adding -1 times the first row to the third. The result is the matrix $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & -3 & -1 \end{pmatrix}$. The next step adds the second row to the third row, leading to the upper triangular matrix $U = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, whose diagonal entries are the pivots. The corresponding lower triangular matrix is $L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$; its entries lying below the main diagonal are the *negatives* of the multiples we used during the elimination procedure. For instance, the $(2,1)$ entry indicates that we added -2 times the first row to the second row, and so on. The reader might wish to verify the resulting factorization

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 5 & 2 \\ 2 & -2 & 0 \end{pmatrix} = A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Exercises

1.3.15. What elementary row operations do the following matrices represent? What size matrices do they apply to?
(*a*) $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 0 \\ 7 & 1 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}$, (*e*) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

1.3.16. Write down the elementary matrix corresponding to the following row operations on 4×4 matrices: (*a*) Add the third row to the fourth row. (*b*) Subtract the fourth row from the third row. (*c*) Add 3 times the last row to the first row. (*d*) Subtract twice the second row from the fourth row.

1.3.17. Compute the product $L_3 L_2 L_1$ of the elementary matrices (1.21). Compare your answer with (1.23).

1.3.18. Determine the product $E_3 E_2 E_1$ of the elementary matrices in (1.19). Is this the same as the product $E_1 E_2 E_3$? Which is easier to predict?

1.3.19. (*a*) Explain, using their interpretation as elementary row operations, why elementary matrices do not generally commute: $E\widetilde{E} \neq \widetilde{E}E$. (*b*) Which pairs of the elementary matrices listed in (1.19) commute? (*c*) Can you formulate a general rule that tells in advance whether two given elementary matrices commute?

1.3.20. Determine which of the following 3×3 matrices is (*i*) upper triangular, (*ii*) upper unitriangular, (*iii*) lower triangular, and/or (*iv*) lower unitriangular:
(*a*) $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & -2 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 3 & 3 \end{pmatrix}$ (*d*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -4 & 1 \end{pmatrix}$ (*e*) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

1.3.21. Find the LU factorization of the following matrices: (*a*) $\begin{pmatrix} 1 & 3 \\ -1 & 0 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$,
(*c*) $\begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, (*d*) $\begin{pmatrix} 2 & 0 & 3 \\ 1 & 3 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, (*e*) $\begin{pmatrix} -1 & 0 & 0 \\ 2 & -3 & 0 \\ 1 & 3 & 2 \end{pmatrix}$, (*f*) $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & 2 \\ -3 & 1 & 0 \end{pmatrix}$,
(*g*) $\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & -1 & -1 \\ -1 & 3 & 0 & 2 \\ 0 & -1 & 2 & 1 \end{pmatrix}$, (*h*) $\begin{pmatrix} 1 & 1 & -2 & 3 \\ -1 & 2 & 3 & 0 \\ -2 & 1 & 1 & -2 \\ 3 & 0 & 1 & 5 \end{pmatrix}$, (*i*) $\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 4 & 0 & 1 \\ 3 & 0 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix}$.

1.3.22. Given the factorization $A = \begin{pmatrix} 2 & -1 & 0 \\ -6 & 4 & -1 \\ 4 & -6 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -4 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}$,
explain, without computing, which elementary row operations are used to reduce A to upper triangular form. Be careful to state the order in which they should be applied. Then check the correctness of your answer by performing the elimination.

1.3.23. (a) Write down a 4×4 lower unitriangular matrix whose entries below the diagonal are distinct nonzero numbers. (b) Explain which elementary row operation each entry corresponds to. (c) Indicate the order in which the elementary row operations should be performed by labeling the entries $1, 2, 3, \ldots$.

♢ 1.3.24. Let $t_1, t_2, \ldots$ be distinct real numbers. Find the LU factorization of the following *Vandermonde matrices*: (a) $\begin{pmatrix} 1 & 1 \\ t_1 & t_2 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ t_1 & t_2 & t_3 & t_4 \\ t_1^2 & t_2^2 & t_3^2 & t_4^2 \\ t_1^3 & t_2^3 & t_3^3 & t_4^3 \end{pmatrix}$.
Can you spot a pattern? Test your conjecture with the 5×5 Vandermonde matrix.

1.3.25. Write down the explicit requirements on its entries a_{ij} for a square matrix A to be (a) diagonal, (b) upper triangular, (c) upper unitriangular, (d) lower triangular, (e) lower unitriangular.

♢ 1.3.26. (a) Explain why the product of two lower triangular matrices is lower triangular. (b) What can you say concerning the diagonal entries of the product of two lower triangular matrices? (c) Explain why the product of two lower unitriangular matrices is also lower unitriangular.

1.3.27. *True or false*: If A has a zero entry on its main diagonal, it is not regular.

1.3.28. In general, how many elementary row operations does one need to perform in order to reduce a regular $n \times n$ matrix to upper triangular form?

1.3.29. Prove that if A is a regular 2×2 matrix, then its LU factorization is unique. In other words, if $A = LU = \widehat{L}\,\widehat{U}$ where $L, \widehat{L}$ are lower unitriangular and $U, \widehat{U}$ are upper triangular, then $L = \widehat{L}$ and $U = \widehat{U}$. (The general case appears in Proposition 1.30.)

♢ 1.3.30. Prove directly that the matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ does not have an LU factorization.

♢ 1.3.31. Suppose A is regular. (a) Show that the matrix obtained by multiplying each column of A by the sign of its pivot is also regular and, moreover, has all positive pivots. (b) Show that the matrix obtained by multiplying each row of A by the sign of its pivot is also regular and has all positive pivots.
(c) Check these results in the particular case $A = \begin{pmatrix} -2 & 2 & 1 \\ 1 & 0 & 1 \\ 4 & 2 & 3 \end{pmatrix}$.

Forward and Back Substitution

Knowing the LU factorization of a regular matrix A enables us to solve any associated linear system $A\mathbf{x} = \mathbf{b}$ in two easy stages:

(1) First, solve the lower triangular system

$$L\,\mathbf{c} = \mathbf{b} \tag{1.25}$$

for the vector $\mathbf{c}$ by *Forward Substitution*. This is the same as Back Substitution, except one solves the equations for the variables in the direct order — from first to last. Explicitly,

$$c_1 = b_1, \qquad c_i = b_i - \sum_{j=1}^{i-1} l_{ij}\, c_j, \qquad \text{for} \qquad i = 2, 3, \ldots, n, \tag{1.26}$$

noting that the previously computed values of $c_1, \ldots, c_{i-1}$ are used to determine c_i.

(2) Second, solve the resulting upper triangular system

$$U\,\mathbf{x} = \mathbf{c} \tag{1.27}$$

by *Back Substitution*. The values of the unknowns

$$x_n = \frac{c_n}{u_{nn}}, \qquad x_i = \frac{1}{u_{ii}} \left(c_i - \sum_{j=i+1}^{n} u_{ij} x_j \right), \qquad \text{for} \qquad i = n-1, \ldots, 2, 1, \tag{1.28}$$

are successively computed, but now in reverse order. It is worth pointing out that the requirement that each pivot be nonzero, $u_{ii} \neq 0$, is essential here, as otherwise we would not be able to solve for the corresponding variable x_i.

Note that the combined algorithm does indeed solve the original system, since if

$$U\mathbf{x} = \mathbf{c} \qquad \text{and} \qquad L\mathbf{c} = \mathbf{b}, \qquad \text{then} \qquad A\mathbf{x} = LU\mathbf{x} = L\mathbf{c} = \mathbf{b}.$$

Example 1.5. With the LU decomposition

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 5 & 2 \\ 2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

found in Example 1.4, we can readily solve any linear system with the given coefficient matrix by Forward and Back Substitution. For instance, to find the solution to

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 5 & 2 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix},$$

we first solve the lower triangular system

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \qquad \text{or, explicitly,} \qquad \begin{aligned} a \qquad\quad &= 1, \\ 2a + b \qquad &= 2, \\ a - b + c &= 2. \end{aligned}$$

The first equation says $a = 1$; substituting into the second, we find $b = 0$; the final equation yields $c = 1$. We then use Back Substitution to solve the upper triangular system

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \text{which is} \qquad \begin{aligned} 2x + y + z &= 1, \\ 3y \qquad &= 0, \\ -z &= 1. \end{aligned}$$

We find $z = -1$, then $y = 0$, and then $x = 1$, which is indeed the solution.

Thus, once we have found the LU factorization of the coefficient matrix A, the Forward and Back Substitution processes quickly produce the solution to any system $A\mathbf{x} = \mathbf{b}$. Moreover, they can be straightforwardly programmed on a computer. In practice, to solve a system from scratch, it is just a matter of taste whether you work directly with the augmented matrix, or first determine the LU factorization of the coefficient matrix, and then apply Forward and Back Substitution to compute the solution.

Exercises

1.3.32. Given the LU factorizations you calculated in Exercise 1.3.21, solve the associated linear systems $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is the column vector with all entries equal to 1.

1.3.33. In each of the following problems, find the $A = LU$ factorization of the coefficient matrix, and then use Forward and Back Substitution to solve the corresponding linear systems $A\mathbf{x} = \mathbf{b}_j$ for each of the indicated right-hand sides:

(a) $A = \begin{pmatrix} -1 & 3 \\ 3 & 2 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$, $\mathbf{b}_3 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.

(b) $A = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix}$.

(c) $A = \begin{pmatrix} 9 & -2 & -1 \\ -6 & 1 & 1 \\ 2 & -1 & 0 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$.

(d) $A = \begin{pmatrix} 2.0 & .3 & .4 \\ .3 & 4.0 & .5 \\ .4 & .5 & 6.0 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

(e) $A = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & 3 & -1 \\ -1 & 3 & 2 & 2 \\ 0 & -1 & 2 & 1 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$.

(f) $A = \begin{pmatrix} 1 & -2 & 0 & 2 \\ 4 & 1 & -1 & -1 \\ -8 & -1 & 2 & 1 \\ -4 & -1 & 1 & 2 \end{pmatrix}$, $\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} 3 \\ 0 \\ -1 \\ 2 \end{pmatrix}$, $\mathbf{b}_3 = \begin{pmatrix} 2 \\ 3 \\ -2 \\ 1 \end{pmatrix}$.

1.4 Pivoting and Permutations

The method of Gaussian Elimination presented so far applies only to regular matrices. But not every square matrix is regular; a simple class of examples is matrices whose upper left, i.e., $(1,1)$, entry is zero, and so cannot serve as the first pivot. More generally, the algorithm cannot proceed whenever a zero entry appears in the current pivot position on the diagonal. What then to do? The answer requires revisiting the source of the method.

Consider, as a specific example, the linear system

$$\begin{aligned} 2y + z &= 2, \\ 2x + 6y + z &= 7, \\ x + y + 4z &= 3. \end{aligned} \tag{1.29}$$

The augmented coefficient matrix is

$$\left(\begin{array}{ccc|c} 0 & 2 & 1 & 2 \\ 2 & 6 & 1 & 7 \\ 1 & 1 & 4 & 3 \end{array}\right).$$

In this case, the $(1,1)$ entry is 0, and so is not a legitimate pivot. The problem, of course, is that the first variable x does not appear in the first equation, and so we cannot use it to eliminate x in the other two equations. But this "problem" is actually a bonus — we already have an equation with only two variables in it, and so we need to eliminate x from only one of the other two equations. To be systematic, we rewrite the system in a different order,

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

by interchanging the first two equations. In other words, we employ

Linear System Operation #2: Interchange two equations.

Clearly, this operation does not change the solution and so produces an equivalent linear system. In our case, the augmented coefficient matrix,

$$\left(\begin{array}{ccc|c} 2 & 6 & 1 & 7 \\ 0 & 2 & 1 & 2 \\ 1 & 1 & 4 & 3 \end{array}\right),$$

can be obtained from the original by performing the second type of row operation:

Elementary Row Operation #2: Interchange two rows of the matrix.

The new nonzero upper left entry, 2, can now serve as the first pivot, and we may continue to apply elementary row operations of type #1 to reduce our matrix to upper triangular form. For this particular example, we eliminate the remaining nonzero entry in the first column by subtracting $\frac{1}{2}$ the first row from the last:

$$\left(\begin{array}{ccc|c} 2 & 6 & 1 & 7 \\ 0 & 2 & 1 & 2 \\ 0 & -2 & \frac{7}{2} & -\frac{1}{2} \end{array}\right).$$

The $(2,2)$ entry serves as the next pivot. To eliminate the nonzero entry below it, we add the second to the third row:

$$\left(\begin{array}{ccc|c} 2 & 6 & 1 & 7 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & \frac{9}{2} & \frac{3}{2} \end{array}\right).$$

We have now placed the system in upper triangular form, with the three pivots $2, 2,$ and $\frac{9}{2}$ along the diagonal. Back Substitution produces the solution $x = \frac{5}{6},\ \ y = \frac{5}{6},\ \ z = \frac{1}{3}$.

The row interchange that is required when a zero shows up in the diagonal pivot position is known as *pivoting*. Later, in Section 1.7, we will discuss practical reasons for pivoting even when a diagonal entry is nonzero. Let us distinguish the class of matrices that can be reduced to upper triangular form by Gaussian Elimination with pivoting. These matrices will prove to be of fundamental importance throughout linear algebra.

Definition 1.6. A square matrix is called *nonsingular* if it can be reduced to upper triangular form with all non-zero elements on the diagonal — the pivots — by elementary row operations of types 1 and 2.

In contrast, a *singular* square matrix cannot be reduced to such upper triangular form by such row operations, because at some stage in the elimination procedure the diagonal entry and all the entries below it are zero. Every regular matrix is nonsingular, but, as we just saw, not every nonsingular matrix is regular. Uniqueness of solutions is the key defining characteristic of nonsingularity.

Theorem 1.7. A linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution for *every* choice of right-hand side $\mathbf{b}$ if and only if its coefficient matrix A is square and nonsingular.

We are able to prove the "if" part of this theorem, since nonsingularity implies reduction to an equivalent upper triangular form that has the same solutions as the original system.

The unique solution to the system is then found by Back Substitution. The "only if" part will be proved in Section 1.8.

The revised version of the Gaussian Elimination algorithm, valid for all nonsingular coefficient matrices, is implemented by the accompanying pseudocode program. The starting point is the augmented matrix $M = \left(A \mid \mathbf{b} \right)$ representing the linear system $A\mathbf{x} = \mathbf{b}$. After successful termination of the program, the result is an augmented matrix in upper triangular form $M = \left(U \mid \mathbf{c} \right)$ representing the equivalent linear system $U\mathbf{x} = \mathbf{c}$. One then uses Back Substitution to determine the solution $\mathbf{x}$ to the linear system.

Gaussian Elimination — Nonsingular Case

```
start
    for j = 1 to n
        if m_kj = 0 for all k ≥ j, stop; print "A is singular"
        if m_jj = 0 but m_kj ≠ 0 for some k > j, switch rows k and j
        for i = j + 1 to n
            set l_ij = m_ij / m_jj
            add −l_ij times row j to row i of M
        next i
    next j
end
```

Remark. When performing the algorithm using exact arithmetic, when pivoting is required it does not matter which row k one chooses to switch with row j, as long as it lies below and the (k, j) entry is nonzero. When dealing with matters involving numerical precision and round off errors, there are some practical rules of thumb to be followed to maintain accuracy in the intervening computations. These will be discussed in Section 1.7.

Exercises

1.4.1. Determine whether the following matrices are singular or nonsingular:

(a) $\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$, (b) $\begin{pmatrix} -1 & 2 \\ 4 & -8 \end{pmatrix}$, (c) $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 1 & 3 \\ 2 & -2 & 0 \end{pmatrix}$, (d) $\begin{pmatrix} 1 & 1 & 3 \\ 2 & 2 & 2 \\ 3 & -1 & 1 \end{pmatrix}$, (e) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$,

(f) $\begin{pmatrix} -1 & 1 & 0 & -3 \\ 2 & -2 & 4 & 0 \\ 1 & -2 & 2 & -1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, (g) $\begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & -2 \\ 2 & 0 & 2 & 0 \end{pmatrix}$, (h) $\begin{pmatrix} 1 & -2 & 0 & 2 \\ 4 & 1 & -1 & -1 \\ -8 & -1 & 2 & 1 \\ -4 & -1 & 1 & 2 \end{pmatrix}$.

1.4.2. Classify the following matrices as (i) regular, (ii) nonsingular, and/or (iii) singular:

(a) $\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} 3 & -2 & 1 \\ -1 & 4 & 4 \\ 2 & 2 & 5 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}$, (d) $\begin{pmatrix} 1 & 3 & -3 & 0 \\ -1 & 0 & -1 & 2 \\ 3 & -2 & 6 & 1 \\ 2 & -1 & 3 & 5 \end{pmatrix}$.

1.4.3. Find the equation $z = ax + by + c$ for the plane passing through the three points $\mathbf{p}_1 = (0, 2, -1)$, $\mathbf{p}_2 = (-2, 4, 3)$, $\mathbf{p}_3 = (2, -1, -3)$.

1.4.4. Show that a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is (*a*) nonsingular if and only if $a\,d - b\,c \neq 0$, (*b*) regular if and only if $a\,d - b\,c \neq 0$ and $a \neq 0$.

1.4.5. Solve the following systems of equations by Gaussian Elimination:

(a) $\begin{aligned} x_1 - 2x_2 + 2x_3 &= 15, \\ x_1 - 2x_2 + x_3 &= 10, \\ 2x_1 - x_2 - 2x_3 &= -10. \end{aligned}$ (*b*) $\begin{aligned} 2x_1 - x_2 &= 1, \\ -4x_1 + 2x_2 - 3x_3 &= -8, \\ x_1 - 3x_2 + x_3 &= 5. \end{aligned}$ (*c*) $\begin{aligned} x_2 - x_3 &= 4, \\ -2x_1 - 5x_2 &= 2, \\ x_1 + x_3 &= -8. \end{aligned}$

(*d*) $\begin{aligned} x - y + z - w = 0,\ \ -2x + 2y - z + w = 2, \\ -4x + 4y + 3z = 5,\ \ x - 3y + w = 4. \end{aligned}$ (*e*) $\begin{aligned} -3x_2 + 2x_3 = 0,\ \ x_3 - x_4 = 2, \\ x_1 - 2x_3 = -1,\ \ -4x_1 + 7x_4 = -5. \end{aligned}$

1.4.6. *True or false*: A singular matrix cannot be regular.

1.4.7. *True or false*: A square matrix that has a column with all 0 entries is singular. What can you say about a linear system that has such a coefficient matrix?

♢ 1.4.8. Explain why the solution to the homogeneous system $A\,\mathbf{x} = \mathbf{0}$ with nonsingular coefficient matrix is $\mathbf{x} = \mathbf{0}$.

1.4.9. Write out the details of the proof of the "if" part of Theorem 1.7: if A is nonsingular, then the linear system $A\,\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.

Permutations and Permutation Matrices

As with the first type of elementary row operation, row interchanges can be accomplished by multiplication by a second type of elementary matrix, which is found by applying the row operation to the identity matrix of the appropriate size. For instance, interchanging rows 1 and 2 of the 3×3 identity matrix produces the elementary interchange matrix $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The result $P\,A$ of multiplying any 3-rowed matrix A on the left by P is the same as interchanging the first two rows of A. For instance,

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

Multiple row interchanges are accomplished by combining such elementary interchange matrices. Each such combination of row interchanges uniquely corresponds to what is called a permutation matrix.

Definition 1.8. A *permutation matrix* is a matrix obtained from the identity matrix by any combination of row interchanges.

In particular, applying a row interchange to a permutation matrix produces another permutation matrix. The following result is easily established.

Lemma 1.9. A matrix P is a permutation matrix if and only if each row of P contains all 0 entries except for a single 1, and, in addition, each column of P also contains all 0 entries except for a single 1.

In general, if, in the permutation matrix P, a 1 appears in position (i, j), then multiplication by P will move the j^{th} row of A into the i^{th} row of the product $P\,A$.

Example 1.10. There are six different 3×3 permutation matrices, namely

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{1.30}$$

These have the following effects: if A is a matrix with row vectors $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$, then multiplication on the left by each of the six permutation matrices produces, respectively,

$$\begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{r}_2 \\ \mathbf{r}_3 \\ \mathbf{r}_1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{r}_3 \\ \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{r}_2 \\ \mathbf{r}_1 \\ \mathbf{r}_3 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{r}_3 \\ \mathbf{r}_2 \\ \mathbf{r}_1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_3 \\ \mathbf{r}_2 \end{pmatrix}. \tag{1.31}$$

Thus, the first permutation matrix, which is the identity, does nothing — the *identity permutation*. The fourth, fifth, sixth represent row interchanges. The second and third are non-elementary permutations, and can be realized by a pair of successive row interchanges.

In general, any rearrangement of a finite ordered collection of objects is called a *permutation*. Thus, the 6 permutation matrices (1.30) produce the 6 possible permutations (1.31) of the rows of a 3×3 matrix. In general, if a permutation π rearranges the integers $(1, \ldots, n)$ to form $(\pi(1), \ldots, \pi(n))$, then the corresponding permutation matrix $P = P_\pi$ that maps row $\mathbf{r}_i$ to row $\mathbf{r}_{\pi(i)}$ will have 1's in positions $(i, \pi(i))$ for $i = 1, \ldots, n$ and zeros everywhere else. For example, the second permutation matrix in (1.30) corresponds to the permutation with $\pi(1) = 2$, $\pi(2) = 3$, $\pi(3) = 1$. Keep in mind that $\pi(1), \ldots, \pi(n)$ is merely a rearrangement of the integers $1, \ldots, n$, so that $1 \leq \pi(i) \leq n$ and $\pi(i) \neq \pi(j)$ when $i \neq j$.

An elementary combinatorial argument proves that there is a total of

$$n! = n\,(n-1)\,(n-2) \;\cdots\; 3 \cdot 2 \cdot 1 \tag{1.32}$$

different permutations of $(1, \ldots, n)$, and hence the same number of permutation matrices of size $n \times n$. Moreover, the product $P = P_1 P_2$ of any two permutation matrices is also a permutation matrix, and corresponds to the composition of the two permutations, meaning one permutes according to P_2 and then permutes the result according to P_1. An important point is that multiplication of permutation matrices is *noncommutative* — the order in which one permutes makes a difference. Switching the first and second rows, and then switching the second and third rows, *does not* have the same effect as first switching the second and third rows and then switching the first and second rows!

Exercises

1.4.10. Write down the elementary 4×4 permutation matrix (*a*) P_1 that permutes the second and fourth rows, and (*b*) P_2 that permutes the first and fourth rows. (*c*) Do P_1 and P_2 commute? (*d*) Explain what the matrix products $P_1 P_2$ and $P_2 P_1$ do to a 4×4 matrix.

1.4.11. Write down the permutation matrix P such that

(a) $P\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} v \\ w \\ u \end{pmatrix}$, (b) $P\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} d \\ c \\ a \\ b \end{pmatrix}$, (c) $P\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} b \\ a \\ d \\ c \end{pmatrix}$, (d) $P\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_4 \\ x_1 \\ x_3 \\ x_2 \\ x_5 \end{pmatrix}$.

1.4.12. Construct a multiplication table that shows all possible products of the 3×3 permutation matrices (1.30). List all pairs that commute.

1.4.13. Write down all 4×4 permutation matrices that (*a*) fix the third row of a 4×4 matrix A; (*b*) take the third row to the fourth row; (*c*) interchange the second and third rows.

1.4.14. *True or false*: (a) Every elementary permutation matrix satisfies $P^2 = \mathrm{I}$. (*b*) Every permutation matrix satisfies $P^2 = \mathrm{I}$. (*c*) A matrix that satisfies $P^2 = \mathrm{I}$ is necessarily a permutation matrix.

1.4.15. (*a*) Let P and Q be $n \times n$ permutation matrices and $\mathbf{v} \in \mathbb{R}^n$ a vector. Under what conditions does the equation $P\mathbf{v} = Q\mathbf{v}$ imply that $P = Q$? (*b*) Answer the same question when $PA = QA$, where A is an $n \times k$ matrix.

1.4.16. Let P be the 3×3 permutation matrix such that the product PA permutes the first and third rows of the 3×3 matrix A. (a) Write down P. (*b*) *True or false*: The product AP is obtained by permuting the first and third columns of A.
(*c*) Does the same conclusion hold for every permutation matrix: is the effect of PA on the rows of a square matrix A the same as the effect of AP on the columns of A?

♡ 1.4.17. A common notation for a permutation π of the integers $\{1, \ldots, m\}$ is as a $2 \times m$ matrix $\begin{pmatrix} 1 & 2 & 3 & \ldots & m \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(m) \end{pmatrix}$, indicating that π takes i to $\pi(i)$. (*a*) Show that such a permutation corresponds to the permutation matrix with 1's in positions $(\pi(j), j)$ for $j = 1, \ldots, m$. (*b*) Write down the permutation matrices corresponding to the following permutations: (*i*) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, (*ii*) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$, (*iii*) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$, (*iv*) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}$. Which are elementary matrices? (*c*) Write down, using the preceding notation, the permutations corresponding to the following permutation matrices:
(*i*) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, (*ii*) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$, (*iii*) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$, (*iv*) $\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$.

♢ 1.4.18. Justify the statement that there are $n!$ different $n \times n$ permutation matrices.

1.4.19. Consider the following combination of elementary row operations of type #1: (*i*) Add row i to row j. (*ii*) Subtract row j from row i. (*iii*) Add row i to row j again. Prove that the net effect is to interchange -1 times row i with row j. Thus, we can *almost* produce an elementary row operation of type #2 by a combination of elementary row operations of type #1. Lest you be tempted to try, Exercise 1.9.16 proves that one *cannot* produce a bona fide row interchange by a combination of elementary row operations of type #1.

1.4.20. What is the effect of permuting the *columns* of its coefficient matrix on a linear system?

The Permuted LU Factorization

As we now know, every nonsingular matrix A can be reduced to upper triangular form by elementary row operations of types #1 and #2. The row interchanges merely reorder the equations. If one performs all of the required row interchanges in advance, then the elimination algorithm can proceed without requiring any further pivoting. Thus, the matrix obtained by permuting the rows of A in the prescribed manner is regular. In other words, if A is a nonsingular matrix, then there is a permutation matrix P such that the product PA is regular, and hence admits an LU factorization. As a result, we deduce the general *permuted LU factorization*

$$PA = LU, \qquad (1.33)$$

where P is a permutation matrix, L is lower unitriangular, and U is upper triangular with the pivots on the diagonal. For instance, in the preceding example, we permuted the first and second rows, and hence equation (1.33) has the explicit form

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 6 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & \frac{9}{2} \end{pmatrix}.$$

We have now established the following generalization of Theorem 1.3.

Theorem 1.11. Let A be an $n \times n$ matrix. Then the following conditions are equivalent:

(i) A is nonsingular.

(ii) A has n nonzero pivots.

(iii) A admits a permuted LU factorization: $PA = LU$.

A practical method to construct a permuted LU factorization of a given matrix A would proceed as follows. First set up $P = L = \mathrm{I}$ as $n \times n$ identity matrices. The matrix P will keep track of the permutations performed during the Gaussian Elimination process, while the entries of L below the diagonal are gradually replaced by the negatives of the multiples used in the corresponding row operations of type #1. Each time two rows of A are interchanged, the same two rows of P will be interchanged. Moreover, any pair of entries that both lie *below* the diagonal in these same two rows of L must also be interchanged, while entries lying on and above its diagonal need to stay in their place. At a successful conclusion to the procedure, A will have been converted into the upper triangular matrix U, while L and P will assume their final form. Here is an illustrative example.

Example 1.12. Our goal is to produce a permuted LU factorization of the matrix

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -2 & -1 \\ -3 & -5 & 6 & 1 \\ -1 & 2 & 8 & -2 \end{pmatrix}.$$

To begin the procedure, we apply row operations of type #1 to eliminate the entries below the first pivot. The updated matrices[†] are

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 4 & 7 & -2 \end{pmatrix}, \qquad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where L keeps track of the row operations, and we initialize P to be the identity matrix. The $(2,2)$ entry of the new A is zero, and so we interchange its second and third rows, leading to

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 4 & 7 & -2 \end{pmatrix}, \qquad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

[†] Here, we are adopting computer programming conventions, where updates of a matrix are all given the same name.

We interchanged the same two rows of P, while in L we only interchanged the already computed entries in its second and third rows that lie in its first column below the diagonal. We then eliminate the nonzero entry lying below the $(2,2)$ pivot, leading to

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -5 & -6 \end{pmatrix}, \qquad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ -1 & 4 & 0 & 1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A final row interchange places the matrix in upper triangular form:

$$U = A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -5 & -6 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ -1 & 4 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Again, we performed the same row interchange on P, while interchanging only the third and fourth row entries of L that lie below the diagonal. You can verify that

$$PA = \begin{pmatrix} 1 & 2 & -1 & 0 \\ -3 & -5 & 6 & 1 \\ -1 & 2 & 8 & -2 \\ 2 & 4 & -2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ -1 & 4 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -5 & -6 \\ 0 & 0 & 0 & -1 \end{pmatrix} = LU, \tag{1.34}$$

as promised. Thus, by rearranging the equations in the order first, third, fourth, second, as prescribed by P, we obtain an equivalent linear system whose coefficient matrix PA is regular, in accordance with Theorem 1.11.

Once the permuted LU factorization is established, the solution to the original system $A\mathbf{x} = \mathbf{b}$ is obtained by applying the same Forward and Back Substitution algorithm presented above. Explicitly, we first multiply the system $A\mathbf{x} = \mathbf{b}$ by the permutation matrix, leading to

$$PA\mathbf{x} = P\mathbf{b} = \widehat{\mathbf{b}}, \tag{1.35}$$

whose right-hand side $\widehat{\mathbf{b}}$ has been obtained by permuting the entries of $\mathbf{b}$ in the same fashion as the rows of A. We then solve the two triangular systems

$$L\mathbf{c} = \widehat{\mathbf{b}} \qquad \text{and} \qquad U\mathbf{x} = \mathbf{c} \tag{1.36}$$

by, respectively, Forward and Back Substitution, as before.

Example 1.12 (continued). Suppose we wish to solve the linear system

$$\begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -2 & -1 \\ -3 & -5 & 6 & 1 \\ -1 & 2 & 8 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 0 \end{pmatrix}.$$

In view of the $PA = LU$ factorization established in (1.34), we need only solve the two auxiliary lower and upper triangular systems (1.36). The lower triangular system is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ -1 & 4 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \\ -1 \end{pmatrix};$$

whose right-hand side was obtained by applying the permutation matrix P to the right-hand side of the original system. Its solution, namely $a = 1$, $b = 6$, $c = -23$, $d = -3$, is obtained through Forward Substitution. The resulting upper triangular system is

$$\begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -5 & -6 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ -23 \\ -3 \end{pmatrix}.$$

Its solution, $w = 3$, $z = 1$, $y = 0$, $x = 2$, which is also the solution to the original system, is easily obtained by Back Substitution.

Exercises

1.4.21. For each of the listed matrices A and vectors $\mathbf{b}$, find a permuted LU factorization of the matrix, and use your factorization to solve the system $A\mathbf{x} = \mathbf{b}$. (a) $\begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$,
(b) $\begin{pmatrix} 0 & 0 & -4 \\ 1 & 2 & 3 \\ 0 & 1 & 7 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, (c) $\begin{pmatrix} 0 & 1 & -3 \\ 0 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, (d) $\begin{pmatrix} 1 & 2 & -1 & 0 \\ 3 & 6 & 2 & -1 \\ 1 & 1 & -7 & 2 \\ 1 & -1 & 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 3 \end{pmatrix}$,
(e) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 1 & 4 & -1 & 2 \\ 7 & -1 & 2 & 3 \end{pmatrix}$, $\begin{pmatrix} -1 \\ -4 \\ 0 \\ 5 \end{pmatrix}$, (f) $\begin{pmatrix} 0 & 0 & 2 & 3 & 4 \\ 0 & 1 & -7 & 2 & 3 \\ 1 & 4 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 7 & 3 \end{pmatrix}$, $\begin{pmatrix} -3 \\ -2 \\ 0 \\ 0 \\ -7 \end{pmatrix}$.

1.4.22. For each of the following linear systems find a permuted LU factorization of the coefficient matrix and then use it to solve the system by Forward and Back Substitution.

(a) $\begin{aligned} 4x_1 - 4x_2 + 2x_3 &= 1, \\ -3x_1 + 3x_2 + x_3 &= 3, \\ -3x_1 + x_2 - 2x_3 &= -5. \end{aligned}$ (b) $\begin{aligned} y - z + w &= 0, \\ y + z &= 1, \\ x - y + z - 3w &= 2, \\ x + 2y - z + w &= 4. \end{aligned}$ (c) $\begin{aligned} x - y + 2z + w &= 0, \\ -x + y - 3z &= 1, \\ x - y + 4z - 3w &= 2, \\ x + 2y - z + w &= 4. \end{aligned}$

♢ 1.4.23. (a) Explain why

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 3 \\ 2 & -1 & 1 \\ 2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 3 \\ 2 & -1 & 1 \\ 2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 3 \\ 2 & -1 & 1 \\ 2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix},$$

are all legitimate permuted LU factorizations of the same matrix. List the elementary row operations that are being used in each case.

(b) Use each of the factorizations to solve the linear system $\begin{pmatrix} 0 & 1 & 3 \\ 2 & -1 & 1 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -5 \\ -1 \\ 0 \end{pmatrix}$.
Do you always obtain the same result? Explain why or why not.

1.4.24. (a) Find three different permuted LU factorizations of the matrix $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 1 & 1 & 3 \end{pmatrix}$.
(b) How many different permuted LU factorizations does A have?

1.4.25. What is the maximal number of permuted LU factorizations a regular 3×3 matrix can have? Give an example of such a matrix.

1.4.26. *True or false*: The pivots of a nonsingular matrix are uniquely defined.

♠ 1.4.27. (a) Write a pseudocode program implementing the algorithm for finding the permuted LU factorization of a matrix. (b) Program your algorithm and test it on the examples in Exercise 1.4.21.

1.5 Matrix Inverses

The inverse of a matrix is analogous to the reciprocal $a^{-1} = 1/a$ of a nonzero scalar $a \neq 0$. We already encountered the inverses of matrices corresponding to elementary row operations. In this section, we will study inverses of general square matrices. We begin with the formal definition.

Definition 1.13. Let A be a square matrix of size $n \times n$. An $n \times n$ matrix X is called the *inverse* of A if it satisfies

$$X A = \mathrm{I} = A X, \tag{1.37}$$

where $\mathrm{I} = \mathrm{I}_n$ is the $n \times n$ identity matrix. The inverse of A is commonly denoted by A^{-1}.

Remark. Noncommutativity of matrix multiplication requires that we impose both conditions in (1.37) in order to properly define an inverse to the matrix A. The first condition, $X A = \mathrm{I}$, says that X is a *left inverse*, while the second, $A X = \mathrm{I}$, requires that X also be a *right inverse*. Rectangular matrices might have either a left inverse or a right inverse, but, as we shall see, *only* square matrices have both, and so only square matrices can have full-fledged inverses. However, not every square matrix has an inverse. Indeed, not every scalar has an inverse: $0^{-1} = 1/0$ is not defined, since the equation $0\,x = 1$ has no solution.

Example 1.14. Since

$$\begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 4 & -5 \\ 1 & 1 & -1 \\ 4 & 6 & -7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 4 & -5 \\ 1 & 1 & -1 \\ 4 & 6 & -7 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix},$$

we conclude that when $A = \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$, then $A^{-1} = \begin{pmatrix} 3 & 4 & -5 \\ 1 & 1 & -1 \\ 4 & 6 & -7 \end{pmatrix}$. Observe that there is no obvious way to anticipate the entries of A^{-1} from the entries of A.

Example 1.15. Let us compute the inverse $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$, when it exists, of a general 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The right inverse condition

$$A X = \begin{pmatrix} a x + b z & a y + b w \\ c x + d z & c y + d w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}$$

holds if and only if x, y, z, w satisfy the linear system

$$\begin{aligned} a x + b z &= 1, & a y + b w &= 0, \\ c x + d z &= 0, & c y + d w &= 1. \end{aligned}$$

Solving by Gaussian Elimination (or directly), we find

$$x = \frac{d}{a\,d - b\,c}, \qquad y = -\frac{b}{a\,d - b\,c}, \qquad z = -\frac{c}{a\,d - b\,c}, \qquad w = \frac{a}{a\,d - b\,c},$$

provided the common denominator $a\,d - b\,c \neq 0$ does not vanish. Therefore, the matrix

$$X = \frac{1}{a\,d - b\,c}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

forms a right inverse to A. However, a short computation shows that it also defines a left inverse:

$$X\,A = \begin{pmatrix} x\,a + y\,c & x\,b + y\,d \\ z\,a + w\,c & z\,b + w\,d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I},$$

and hence $X = A^{-1}$ is the inverse of A.

The denominator appearing in the preceding formulas has a special name; it is called the *determinant* of the 2×2 matrix A, and denoted by

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,d - b\,c. \tag{1.38}$$

Thus, the determinant of a 2×2 matrix is the product of the diagonal entries minus the product of the off-diagonal entries. (Determinants of larger square matrices will be discussed in Section 1.9.) Thus, the 2×2 matrix A is invertible, with

$$A^{-1} = \frac{1}{a\,d - b\,c}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{1.39}$$

if and only if $\det A \neq 0$. For example, if $A = \begin{pmatrix} 1 & 3 \\ -2 & -4 \end{pmatrix}$, then $\det A = 2 \neq 0$. We conclude that A has an inverse, which, by (1.39), is $A^{-1} = \frac{1}{2}\begin{pmatrix} -4 & -3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -\frac{3}{2} \\ 1 & \frac{1}{2} \end{pmatrix}$.

Example 1.16. We already learned how to find the inverse of an elementary matrix of type #1: we just negate the one nonzero off-diagonal entry. For example, if

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \qquad \text{then} \qquad E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

This is because the inverse of the elementary row operation that adds twice the first row to the third row is the operation of subtracting twice the first row from the third row.

Example 1.17. Let $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ denote the elementary matrix that has the effect of interchanging rows 1 and 2 of a 3–rowed matrix. Then $P^2 = \mathrm{I}$, since performing the interchange twice returns us to where we began. This implies that $P^{-1} = P$ is its own inverse. Indeed, the same result holds for all elementary permutation matrices that correspond to row operations of type #2. However, it is not true for more general permutation matrices.

The following fundamental result will be established later in this chapter.

Theorem 1.18. A square matrix has an inverse if and only if it is nonsingular.

Consequently, an $n \times n$ matrix will have an inverse if and only if it can be reduced to upper triangular form, with n nonzero pivots on the diagonal, by a combination of elementary row operations. Indeed, "invertible" is often used as a synonym for "nonsingular". All other matrices are singular and do not have an inverse as defined above. Before attempting to prove Theorem 1.18, we need first to become familiar with some elementary properties of matrix inverses.

Lemma 1.19. The inverse of a square matrix, if it exists, is unique.

Proof: Suppose both X and Y satisfy (1.37), so

$$X A = \mathrm{I} = A X \qquad \text{and} \qquad Y A = \mathrm{I} = A Y.$$

Then, by associativity,

$$X = X \mathrm{I} = X(A Y) = (X A) Y = \mathrm{I} Y = Y. \qquad \textit{Q.E.D.}$$

Inverting a matrix twice brings us back to where we started.

Lemma 1.20. If A is an invertible matrix, then A^{-1} is also invertible and $(A^{-1})^{-1} = A$.

Proof: The matrix inverse equations $A^{-1} A = \mathrm{I} = A A^{-1}$ are sufficient to prove that A is the inverse of A^{-1}. *Q.E.D.*

Lemma 1.21. If A and B are invertible matrices of the same size, then their product, $A B$, is invertible, and

$$(A B)^{-1} = B^{-1} A^{-1}. \tag{1.40}$$

Note that the order of the factors is reversed under inversion.

Proof: Let $X = B^{-1} A^{-1}$. Then, by associativity,

$$X\,(A B) = B^{-1} A^{-1} A B = B^{-1} \mathrm{I} B = B^{-1} B = \mathrm{I},$$
$$(A B)\,X = A B B^{-1} A^{-1} = A \mathrm{I} A^{-1} = A A^{-1} = \mathrm{I}.$$

Thus X is both a left and a right inverse for the product matrix $A B$. *Q.E.D.*

Example 1.22. One verifies, directly, that the inverse of $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ is $A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$, while the inverse of $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is $B^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Therefore, the inverse of their product $C = A B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix}$ is given by $C^{-1} = B^{-1} A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}$.

We can straightforwardly generalize the preceding result. The inverse of a k-fold product of invertible matrices is the product of their inverses, *in the reverse order*:

$$(A_1 A_2 \cdots A_{k-1} A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \cdots A_2^{-1} A_1^{-1}. \tag{1.41}$$

Warning. In general, $(A + B)^{-1} \neq A^{-1} + B^{-1}$. Indeed, this equation is not even true for scalars (1×1 matrices)!

Exercises

1.5.1. Verify by direct multiplication that the following matrices are inverses, i.e., both conditions in (1.37) hold: (*a*) $A = \begin{pmatrix} 2 & 3 \\ -1 & -1 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} -1 & -3 \\ 1 & 2 \end{pmatrix}$; (*b*) $A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$; (*c*) $A = \begin{pmatrix} -1 & 3 & 2 \\ 2 & 2 & -1 \\ -2 & 1 & 3 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} -1 & 1 & 1 \\ \frac{4}{7} & -\frac{1}{7} & -\frac{3}{7} \\ -\frac{6}{7} & \frac{5}{7} & \frac{8}{7} \end{pmatrix}$.

1.5.2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & -1 & -8 \end{pmatrix}$. Find the right inverse of A by setting up and solving the linear system $AX = \mathrm{I}$. Verify that the resulting matrix X is also a left inverse.

1.5.3. Write down the inverse of each of the following elementary matrices: (*a*) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 0 \\ 5 & 1 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}$, (*e*) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 6 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, (*f*) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$.

1.5.4. Show that the inverse of $L = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{pmatrix}$ is $L^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ -b & 0 & 1 \end{pmatrix}$. However, the inverse of $M = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}$ is *not* $\begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ -b & -c & 1 \end{pmatrix}$. What is M^{-1}?

1.5.5. Explain why a matrix with a row of all zeros does not have an inverse.

1.5.6. (a) Write down the inverse of the matrices $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}$. (*b*) Write down the product matrix $C = AB$ and its inverse C^{-1} using the inverse product formula.

1.5.7. (a) Find the inverse of the *rotation matrix* $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, where $\theta \in \mathbb{R}$.
(*b*) Use your result to solve the system $x = a\cos\theta - b\sin\theta$, $y = a\sin\theta + b\cos\theta$, for a and b in terms of x and y. (*c*) Prove that, for all $a \in \mathbb{R}$ and $0 < \theta < \pi$, the matrix $R_\theta - a\,\mathrm{I}$ has an inverse.

1.5.8. (a) Write down the inverses of each of the 3×3 permutation matrices (1.30). (*b*) Which ones are their own inverses, $P^{-1} = P$? (*c*) Can you find a non-elementary permutation matrix P that is its own inverse: $P^{-1} = P$?

1.5.9. Find the inverse of the following permutation matrices:
(a) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$, (*b*) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$.

1.5.10. Explain how to write down the inverse permutation using the notation of Exercise 1.4.17. Apply your method to the examples in Exercise 1.5.9, and check the result by verifying that it produces the inverse permutation matrix.

1.5.11. Find all real 2×2 matrices that are their own inverses: $A^{-1} = A$.

1.5.12. Show that if a square matrix A satisfies $A^2 - 3A + \mathrm{I} = \mathrm{O}$, then $A^{-1} = 3\,\mathrm{I} - A$.

1.5.13. Prove that if $c \neq 0$ is any nonzero scalar and A is an invertible matrix, then the scalar product matrix cA is invertible, and $(cA)^{-1} = \dfrac{1}{c}\,A^{-1}$.

1.5.14. Show that $A = \begin{pmatrix} 0 & a & 0 & 0 & 0 \\ b & 0 & c & 0 & 0 \\ 0 & d & 0 & e & 0 \\ 0 & 0 & f & 0 & g \\ 0 & 0 & 0 & h & 0 \end{pmatrix}$ is not invertible for any value of the entries.

1.5.15. Show that if A is a nonsingular matrix, so is everyy power A^n.

1.5.16. Prove that a diagonal matrix $D = \operatorname{diag}(d_1, \ldots, d_n)$ is invertible if and only if all its diagonal entries are nonzero, in which case $D^{-1} = \operatorname{diag}(1/d_1, \ldots, 1/d_n)$.

1.5.17. Prove that if U is a nonsingular upper triangular matrix, then the diagonal entries of U^{-1} are the reciprocals of the diagonal entries of U.

♢ 1.5.18. (a) Let U be a $m \times n$ matrix and V an $n \times m$ matrix, such that the $m \times m$ matrix $\mathrm{I}_m + UV$ is invertible. Prove that $\mathrm{I}_n + VU$ is also invertible, and is given by

$$(\mathrm{I}_n + VU)^{-1} = \mathrm{I}_n - V(\mathrm{I}_m + UV)^{-1}U.$$

(b) The *Sherman–Morrison–Woodbury formula* generalizes this identity to

$$(A + VBU)^{-1} = A^{-1} - A^{-1}V(B^{-1} + UA^{-1}V)^{-1}UA^{-1}. \tag{1.42}$$

Explain what assumptions must be made on the matrices A, B, U, V for (1.42) to be valid.

♢ 1.5.19. Two matrices A and B are said to be *similar*, written $A \sim B$, if there exists an invertible matrix S such that $B = S^{-1}AS$. Prove: (a) $A \sim A$. (b) If $A \sim B$, then $B \sim A$. (c) If $A \sim B$ and $B \sim C$, then $A \sim C$.

♡ 1.5.20. (a) A block matrix $D = \begin{pmatrix} A & \mathrm{O} \\ \mathrm{O} & B \end{pmatrix}$ is called *block diagonal* if A and B are square matrices, not necessarily of the same size, while the O's are zero matrices of the appropriate sizes. Prove that D has an inverse if and only if both A and B do, and $D^{-1} = \begin{pmatrix} A^{-1} & \mathrm{O} \\ \mathrm{O} & B^{-1} \end{pmatrix}$. (b) Find the inverse of $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ and $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 2 & 5 \end{pmatrix}$ by using this method.

1.5.21. (a) Show that $B = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$ is a left inverse of $A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$. (b) Show that A does not have a right inverse. (c) Can you find any other left inverses of A?

1.5.22. Prove that the rectangular matrix $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 2 & 0 \end{pmatrix}$ has a right inverse, but no left inverse.

1.5.23. (a) Are there any nonzero real scalars that satisfy $(a+b)^{-1} = a^{-1} + b^{-1}$?
(b) Are there any nonsingular real 2×2 matrices that satisfy $(A+B)^{-1} = A^{-1} + B^{-1}$?

Gauss–Jordan Elimination

The principal algorithm used to compute the inverse of a nonsingular matrix is known as *Gauss–Jordan Elimination*, in honor of Gauss and Wilhelm Jordan, a nineteenth-century German engineer. A key fact is that, given that A is square, we need to solve only the right inverse equation

$$AX = \mathrm{I} \tag{1.43}$$

in order to compute $X = A^{-1}$. The left inverse equation in (1.37), namely $X A = \mathrm{I}$, will then follow as an automatic consequence. In other words, for square matrices, a right inverse is automatically a left inverse, and conversely! A proof will appear below.

The reader may well ask, then, why use both left and right inverse conditions in the original definition? There are several good reasons. First of all, a non-square matrix may satisfy one of the two conditions — having either a left inverse or a right inverse — but can never satisfy both. Moreover, even when we restrict our attention to square matrices, starting with only one of the conditions makes the logical development of the subject considerably more difficult, and not really worth the extra effort. Once we have established the basic properties of the inverse of a square matrix, we can then safely discard the superfluous left inverse condition. Finally, when we generalize the notion of an inverse to linear operators in Chapter 7, then, in contrast to the case of square matrices, we *cannot* dispense with either of the conditions.

Let us write out the individual columns of the right inverse equation (1.43). The j^{th} column of the $n \times n$ identity matrix I is the vector $\mathbf{e}_j$ that has a 1 in the j^{th} slot and 0's elsewhere, so

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \ldots \qquad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{1.44}$$

According to (1.11), the j^{th} column of the matrix product $A X$ is equal to $A \mathbf{x}_j$, where $\mathbf{x}_j$ denotes the j^{th} column of the inverse matrix X. Therefore, the single matrix equation (1.43) is equivalent to n linear systems

$$A \mathbf{x}_1 = \mathbf{e}_1, \qquad A \mathbf{x}_2 = \mathbf{e}_2, \qquad \ldots \qquad A \mathbf{x}_n = \mathbf{e}_n, \tag{1.45}$$

all having the same coefficient matrix. As such, to solve them we should form the n augmented matrices $M_1 = \left(A \mid \mathbf{e}_1 \right), \ldots, M_n = \left(A \mid \mathbf{e}_n \right)$, and then apply our Gaussian Elimination algorithm to each. But this would be a waste of effort. Since the coefficient matrix is the same, we will end up performing *identical* row operations on each augmented matrix. Clearly, it will be more efficient to combine them into one large augmented matrix $M = \left(A \mid \mathbf{e}_1 \ \ldots \ \mathbf{e}_n \right) = \left(A \mid \mathrm{I} \right)$, of size $n \times (2n)$, in which the right-hand sides $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of our systems are placed into n different columns, which we then recognize as reassembling the columns of an $n \times n$ identity matrix. We may then simultaneously apply our elementary row operations to reduce, if possible, the large augmented matrix so that its first n columns are in upper triangular form.

Example 1.23. For example, to find the inverse of the matrix $A = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix}$, we form the large augmented matrix

$$\left(\begin{array}{ccc|ccc} 0 & 2 & 1 & 1 & 0 & 0 \\ 2 & 6 & 1 & 0 & 1 & 0 \\ 1 & 1 & 4 & 0 & 0 & 1 \end{array}\right).$$

Applying the same sequence of elementary row operations as in Section 1.4, we first interchange the rows

$$\left(\begin{array}{ccc|ccc} 2 & 6 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 4 & 0 & 0 & 1 \end{array}\right),$$

and then eliminate the nonzero entries below the first pivot,

$$\left(\begin{array}{ccc|ccc} 2 & 6 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & -2 & \frac{7}{2} & 0 & -\frac{1}{2} & 1 \end{array}\right).$$

Next we eliminate the entry below the second pivot:

$$\left(\begin{array}{ccc|ccc} 2 & 6 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{9}{2} & 1 & -\frac{1}{2} & 1 \end{array}\right).$$

At this stage, we have reduced our augmented matrix to the form $\left(U \mid C \right)$, where U is upper triangular. This is equivalent to reducing the original n linear systems $A\mathbf{x}_i = \mathbf{e}_i$ to n upper triangular systems $U\mathbf{x}_i = \mathbf{c}_i$. We can therefore perform n back substitutions to produce the solutions $\mathbf{x}_i$, which would form the individual columns of the inverse matrix $X = (\mathbf{x}_1 \ \ldots \ \mathbf{x}_n)$. In the more common version of the Gauss–Jordan scheme, one instead continues to employ elementary row operations to fully reduce the augmented matrix. The goal is to produce an augmented matrix $\left(\mathrm{I} \mid X \right)$ in which the left-hand $n \times n$ matrix has become the identity, while the right-hand matrix is the desired solution $X = A^{-1}$. Indeed, $\left(\mathrm{I} \mid X \right)$ represents the n trivial linear systems $\mathrm{I}\,\mathbf{x} = \mathbf{x}_i$ whose solutions $\mathbf{x} = \mathbf{x}_i$ are the columns of the inverse matrix X.

Now, the identity matrix has 0's below the diagonal, just like U. It also has 1's along the diagonal, whereas U has the pivots (which are all nonzero) along the diagonal. Thus, the next phase in the reduction process is to make all the diagonal entries of U equal to 1. To proceed, we need to introduce the last, and least, of our linear systems operations.

Linear System Operation #3: Multiply an equation by a nonzero constant.

This operation clearly does not affect the solution, and so yields an equivalent linear system. The corresponding elementary row operation is:

Elementary Row Operation #3: Multiply a row of the matrix by a nonzero scalar.

Dividing the rows of the upper triangular augmented matrix $\left(U \mid C \right)$ by the diagonal pivots of U will produce a matrix of the form $\left(V \mid B \right)$, where V is *upper unitriangular*, meaning it has all 1's along the diagonal. In our particular example, the result of these three elementary row operations of type #3 is

$$\left(\begin{array}{ccc|ccc} 1 & 3 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & \frac{2}{9} & -\frac{1}{9} & \frac{2}{9} \end{array}\right),$$

where we multiplied the first and second rows by $\frac{1}{2}$ and the third row by $\frac{2}{9}$.

We are now over halfway towards our goal. We need only make the entries above the diagonal of the left-hand matrix equal to zero. This can be done by elementary row operations of type #1, but now we work backwards. First, we eliminate the nonzero entries in the third column lying above the $(3,3)$ entry by subtracting one half the third row from the second and also from the first:

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -\frac{1}{9} & \frac{5}{9} & -\frac{1}{9} \\ 0 & 1 & 0 & \frac{7}{18} & \frac{1}{18} & -\frac{1}{9} \\ 0 & 0 & 1 & \frac{2}{9} & -\frac{1}{9} & \frac{2}{9} \end{array}\right).$$

Finally, we subtract 3 times the second row from the first to eliminate the remaining nonzero off-diagonal entry, thereby completing the Gauss–Jordan procedure:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{23}{18} & \frac{7}{18} & \frac{2}{9} \\ 0 & 1 & 0 & \frac{7}{18} & \frac{1}{18} & -\frac{1}{9} \\ 0 & 0 & 1 & \frac{2}{9} & -\frac{1}{9} & \frac{2}{9} \end{array}\right).$$

The left-hand matrix is the identity, and therefore the final right-hand matrix is our desired inverse:

$$A^{-1} = \begin{pmatrix} -\frac{23}{18} & \frac{7}{18} & \frac{2}{9} \\ \frac{7}{18} & \frac{1}{18} & -\frac{1}{9} \\ \frac{2}{9} & -\frac{1}{9} & \frac{2}{9} \end{pmatrix}. \tag{1.46}$$

The reader may wish to verify that the final result does satisfy both inverse conditions $A\,A^{-1} = \mathrm{I} = A^{-1}A$.

We are now able to complete the proofs of the basic results on inverse matrices. First, we need to determine the elementary matrix corresponding to an elementary row operation of type #3. Again, this is obtained by performing the row operation in question on the identity matrix. Thus, the elementary matrix that multiplies row i by the nonzero scalar c is the diagonal matrix having c in the i^{th} diagonal position, and 1's elsewhere along the diagonal. The inverse elementary matrix is the diagonal matrix with $1/c$ in the i^{th} diagonal position and 1's elsewhere on the main diagonal; it corresponds to the inverse operation that divides row i by c. For example, the elementary matrix that multiplies the second row of a 3-rowed matrix by 5 is $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; its inverse is $E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

In summary:

Lemma 1.24. Every elementary matrix is nonsingular, and its inverse is also an elementary matrix of the same type.

The Gauss–Jordan method tells us how to reduce any nonsingular square matrix A to the identity matrix by a sequence of elementary row operations. Let $E_1, E_2, \ldots, E_N$ be the corresponding elementary matrices. The elimination procedure that reduces A to I amounts to multiplying A by a succession of elementary matrices:

$$E_N\,E_{N-1}\;\cdots\;E_2\,E_1\,A = \mathrm{I}\,. \tag{1.47}$$

We claim that the product matrix

$$X = E_N\,E_{N-1}\;\cdots\;E_2\,E_1 \tag{1.48}$$

is the inverse of A. Indeed, formula (1.47) says that $X\,A = \mathrm{I}$, and so X is a left inverse. Furthermore, each elementary matrix has an inverse, and so by (1.41), X itself is invertible, with

$$X^{-1} = E_1^{-1}\,E_2^{-1}\;\cdots\;E_{N-1}^{-1}\,E_N^{-1}. \tag{1.49}$$

Therefore, multiplying formula (1.47), namely $X\,A = \mathrm{I}$, on the left by X^{-1} leads to $A = X^{-1}$. Lemma 1.20 implies $X = A^{-1}$, as claimed, completing the proof of Theorem 1.18. Finally, equating $A = X^{-1}$ to the product (1.49), and invoking Lemma 1.24, we have established the following result.

Proposition 1.25. Every nonsingular matrix can be written as the product of elementary matrices.

Example 1.26. The 2×2 matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix}$ is converted into the identity matrix by first interchanging its rows, $\begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix}$, then scaling the second row by -1, $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, and, finally, subtracting 3 times the second row from the first to obtain $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}$. The corresponding elementary matrices are

$$E_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}.$$

Therefore, by (1.48),

$$A^{-1} = E_3 E_2 E_1 = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix},$$

while

$$A = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix}.$$

As an application, let us prove that the inverse of a nonsingular triangular matrix is also triangular. Specifically:

Proposition 1.27. If L is a lower triangular matrix with all nonzero entries on the main diagonal, then L is nonsingular and its inverse L^{-1} is also lower triangular. In particular, if L is lower unitriangular, so is L^{-1}. A similar result holds for upper triangular matrices.

Proof: It suffices to note that if L has all nonzero diagonal entries, one can reduce L to the identity by elementary row operations of types #1 and #3, whose associated elementary matrices are all lower triangular. Lemma 1.2 implies that the product (1.48) is then also lower triangular. If L is unitriangular, then all the pivots are equal to 1. Thus, no elementary row operations of type #3 are required, and so L can be reduced to the identity matrix by elementary row operations of type #1 alone. Therefore, its inverse is a product of lower unitriangular matrices, and hence is itself lower unitriangular. A similar argument applies in the upper triangular case. *Q.E.D.*

Exercises

1.5.24. (*a*) Write down the elementary matrix that multiplies the third row of a 4×4 matrix by 7. (*b*) Write down its inverse.

1.5.25. Find the inverse of each of the following matrices, if possible, by applying the Gauss–Jordan Method.

(*a*) $\begin{pmatrix} 1 & -2 \\ 3 & -3 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$, (*c*) $\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, (*e*) $\begin{pmatrix} 1 & 0 & -2 \\ 3 & -1 & 0 \\ -2 & 1 & -3 \end{pmatrix}$,

(*f*) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 5 & 5 \\ 2 & 1 & 2 \end{pmatrix}$, (*g*) $\begin{pmatrix} 2 & 1 & 2 \\ 4 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix}$, (*h*) $\begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & -2 & -5 \end{pmatrix}$, (*i*) $\begin{pmatrix} 1 & -2 & 1 & 1 \\ 2 & -3 & 3 & 0 \\ 3 & -7 & 2 & 4 \\ 0 & 2 & 1 & 1 \end{pmatrix}$.

1.5.26. Write each of the matrices in Exercise 1.5.25 as a product of elementary matrices.

1.5.27. Express $A = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$ as a product of elementary matrices.

1.5.28. Use the Gauss–Jordan Method to find the inverse of the following complex matrices:
(*a*) $\begin{pmatrix} \mathrm{i} & 1 \\ 1 & \mathrm{i} \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 1-\mathrm{i} \\ 1+\mathrm{i} & 1 \end{pmatrix}$, (*c*) $\begin{pmatrix} 0 & 1 & -\mathrm{i} \\ \mathrm{i} & 0 & -1 \\ -1 & \mathrm{i} & 1 \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 0 & \mathrm{i} \\ \mathrm{i} & -1 & 1+\mathrm{i} \\ -3\,\mathrm{i} & 1-\mathrm{i} & 1+\mathrm{i} \end{pmatrix}$.

1.5.29. Can two nonsingular linear systems have the same solution and yet not be equivalent?

♡ 1.5.30. (*a*) Suppose $\widetilde{A}$ is obtained from A by applying an elementary row operation. Let $C = A B$, where B is any matrix of the appropriate size. Explain why $\widetilde{C} = \widetilde{A}\, B$ can be obtained by applying the same elementary row operation to C. (*b*) Illustrate by adding -2 times the first row to the third row of $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -3 & 2 \\ 0 & 1 & -4 \end{pmatrix}$ and then multiplying the result on the right by $B = \begin{pmatrix} 1 & -2 \\ 3 & 0 \\ -1 & 1 \end{pmatrix}$. Check that the resulting matrix is the same as first multiplying $A B$ and then applying the same row operation to the product matrix.

Solving Linear Systems with the Inverse

The primary motivation for introducing the matrix inverse is that it provides a compact formula for the solution to any linear system with an invertible coefficient matrix.

Theorem 1.28. If the matrix A is nonsingular, then $\mathbf{x} = A^{-1}\,\mathbf{b}$ is the unique solution to the linear system $A\,\mathbf{x} = \mathbf{b}$.

Proof: We merely multiply the system by A^{-1}, which yields $\mathbf{x} = A^{-1} A\,\mathbf{x} = A^{-1}\mathbf{b}$. Moreover, $A\,\mathbf{x} = A\,A^{-1}\mathbf{b} = \mathbf{b}$, proving that $\mathbf{x} = A^{-1}\mathbf{b}$ is indeed the solution. *Q.E.D.*

For example, let us return to the linear system (1.29). Since we computed the inverse of its coefficient matrix in (1.46), a "direct" way to solve the system is to multiply the right-hand side by the inverse matrix:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{23}{18} & \frac{7}{18} & \frac{2}{9} \\ \frac{7}{18} & \frac{1}{18} & -\frac{1}{9} \\ \frac{2}{9} & -\frac{1}{9} & \frac{2}{9} \end{pmatrix} \begin{pmatrix} 2 \\ 7 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{5}{6} \\ \frac{5}{6} \\ \frac{1}{3} \end{pmatrix},$$

reproducing our earlier solution.

However, while æsthetically appealing, the solution method based on the inverse matrix is hopelessly inefficient as compared to direct Gaussian Elimination, and, despite what you may have been told, *should not be used in practical computations*. (A complete justification of this dictum will be provided in Section 1.7.) On the other hand, the inverse does play a useful role in theoretical developments, as well as providing insight into the design of practical algorithms. But the principal message of applied linear algebra is that LU decomposition and Gaussian Elimination are fundamental; matrix inverses are to be avoided in all but the most elementary computations.

Remark. The reader may have learned a version of the Gauss–Jordan algorithm for solving a single linear system that replaces the Back Substitution step by a complete reduction of the coefficient matrix to the identity. In other words, to solve $A\mathbf{x} = \mathbf{b}$, we start with the augmented matrix $M = \left(A \mid \mathbf{b} \right)$ and use all three types of elementary row operations to produce (assuming nonsingularity) the fully reduced form $\left(\mathrm{I} \mid \mathbf{d} \right)$, representing the trivially soluble, equivalent system $\mathbf{x} = \mathbf{d}$, which is the solution to the original system. However, Back Substitution is more efficient, and it remains the method of choice in practical computations.

Exercises

1.5.31. Solve the following systems of linear equations by computing the inverses of their coefficient matrices.

(a) $\begin{aligned} x + 2y &= 1, \\ x - 2y &= -2. \end{aligned}$ (b) $\begin{aligned} 3u - 2v &= 2, \\ u + 5v &= 12. \end{aligned}$ (c) $\begin{aligned} x - y + 3z &= 3, \\ x - 2y + 3z &= -2, \\ x - 2y + z &= 2. \end{aligned}$ (d) $\begin{aligned} y + 5z &= 3, \\ x - y + 3z &= -1, \\ -2x + 3y &= 5. \end{aligned}$

(e) $\begin{aligned} x + 4y - z &= 3, \\ 2x + 7y - 2z &= 5, \\ -x - 5y + 2z &= -7. \end{aligned}$ (f) $\begin{aligned} x + y &= 4, \\ 2x + 3y - w &= 11, \\ -y - z + w &= -7, \\ z - w &= 6. \end{aligned}$ (g) $\begin{aligned} x - 2y + z + 2u &= -2, \\ x - y + z - u &= 3, \\ 2x - y + z + u &= 3, \\ -x + 3y - 2z - u &= 2. \end{aligned}$

1.5.32. For each of the nonsingular matrices in Exercise 1.5.25, use your computed inverse to solve the associated linear system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is the column vector of the appropriate size that has all 1's as its entries.

The LDV Factorization

The second phase of the Gauss–Jordan process leads to a slightly more detailed version of the LU factorization. Let D denote the diagonal matrix having the same diagonal entries as U; in other words, D contains the pivots on its diagonal and zeros everywhere else. Let V be the upper unitriangular matrix obtained from U by dividing each row by its pivot, so that V has all 1's on the diagonal. We already encountered V during the course of the Gauss–Jordan procedure. It is easily seen that $U = DV$, which implies the following result.

Theorem 1.29. A matrix A is regular if and only if it admits a factorization

$$A = LDV, \tag{1.50}$$

where L is a lower unitriangular matrix, D is a diagonal matrix with nonzero diagonal entries, and V is an upper unitriangular matrix.

For the matrix appearing in Example 1.4, we have $U = DV$, where

$$U = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad V = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This leads to the factorization

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 5 & 2 \\ 2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = LDV.$$

Proposition 1.30. If $A = LU$ is regular, then the factors L and U are uniquely determined. The same holds for the $A = LDV$ factorization.

Proof: Suppose $LU = \widetilde{L}\,\widetilde{U}$. Since the diagonal entries of all four matrices are non-zero, Proposition 1.27 implies that they are invertible. Therefore,

$$\widetilde{L}^{-1}L = \widetilde{L}^{-1}L\,U\,U^{-1} = \widetilde{L}^{-1}\widetilde{L}\,\widetilde{U}\,U^{-1} = \widetilde{U}\,U^{-1}. \tag{1.51}$$

The left-hand side of the matrix equation (1.51) is the product of two lower unitriangular matrices, and so, by Lemma 1.2, is itself lower unitriangular. The right-hand side is the product of two upper triangular matrices, and hence is upper triangular. But the only way a lower unitriangular matrix can equal an upper triangular matrix is if they both equal the diagonal identity matrix. Therefore, $\widetilde{L}^{-1}L = \mathrm{I} = \widetilde{U}\,U^{-1}$, and so $\widetilde{L} = L$ and $\widetilde{U} = U$, proving the first result. The LDV version is an immediate consequence. *Q.E.D.*

As you may have guessed, the more general cases requiring one or more row interchanges lead to a permuted LDV factorization in the following form.

Theorem 1.31. A matrix A is nonsingular if and only if there is a permutation matrix P such that

$$PA = LDV, \tag{1.52}$$

where L is a lower unitriangular matrix, D is a diagonal matrix with nonzero diagonal entries, and V is a upper unitriangular matrix.

Uniqueness does not hold for the more general permuted factorizations (1.33), (1.52), since there may be several permutation matrices that place a matrix in regular form; an explicit example can be found in Exercise 1.4.23. Moreover, in contrast to regular Gaussian Elimination, here the pivots, i.e., the diagonal entries of U, are no longer uniquely defined, but depend on the particular combination of row interchanges employed during the course of the computation.

Exercises

1.5.33. Produce the LDV or a permuted LDV factorization of the following matrices:

(*a*) $\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}$, (*b*) $\begin{pmatrix} 0 & 4 \\ -7 & 2 \end{pmatrix}$, (*c*) $\begin{pmatrix} 2 & 1 & 2 \\ 2 & 4 & -1 \\ 0 & -2 & 1 \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 1 & 5 \\ 1 & 1 & -2 \\ 2 & -1 & 3 \end{pmatrix}$,

(*e*) $\begin{pmatrix} 2 & -3 & 2 \\ 1 & -1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$, (*f*) $\begin{pmatrix} 1 & -1 & 1 & 2 \\ 1 & -4 & 1 & 5 \\ 1 & 2 & -1 & -1 \\ 3 & 1 & 1 & 6 \end{pmatrix}$, (*g*) $\begin{pmatrix} 1 & 0 & 2 & -3 \\ 2 & -2 & 0 & 1 \\ 1 & -2 & -2 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$.

1.5.34. Using the LDV factorization for the matrices you found in parts (*a–g*) of Exercise 1.5.33, solve the corresponding linear systems $A\,\mathbf{x} = \mathbf{b}$, for the indicated vector $\mathbf{b}$.

(*a*) $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, (*b*) $\begin{pmatrix} -1 \\ -2 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$, (*d*) $\begin{pmatrix} -1 \\ 4 \\ -1 \end{pmatrix}$, (*e*) $\begin{pmatrix} -1 \\ -2 \\ 5 \end{pmatrix}$, (*f*) $\begin{pmatrix} 2 \\ -9 \\ 3 \\ 4 \end{pmatrix}$, (*g*) $\begin{pmatrix} 6 \\ -4 \\ 0 \\ -3 \end{pmatrix}$.

1.6 Transposes and Symmetric Matrices

Another basic operation on matrices is to interchange their rows and columns. If A is an $m \times n$ matrix, then its *transpose*, denoted by A^T, is the $n \times m$ matrix whose (i,j) entry equals the (j,i) entry of A; thus

$$B = A^T \qquad \text{means that} \qquad b_{ij} = a_{ji}.$$

For example, if

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad \text{then} \qquad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Observe that the rows of A become the columns of A^T and vice versa. In particular, the transpose of a row vector is a column vector, while the transpose of a column vector is a row vector; if $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, then $\mathbf{v}^T = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$. The transpose of a scalar, considered as a 1×1 matrix, is itself: $c^T = c$.

Remark. Most vectors appearing in applied mathematics are column vectors. To conserve vertical space in this text, we will often use the transpose notation, e.g., $\mathbf{v} = (\, v_1, v_2, v_3 \,)^T$, as a compact way of writing the column vector $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$.

In the square case, transposition can be viewed as "reflecting" the matrix entries across the main diagonal. For example,

$$\begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 5 \\ -2 & -4 & 8 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 0 & -4 \\ -1 & 5 & 8 \end{pmatrix}.$$

In particular, the transpose of a lower triangular matrix is upper triangular and vice-versa.

Transposing twice returns you to where you started:

$$(A^T)^T = A. \tag{1.53}$$

Unlike inversion, transposition *is* compatible with matrix addition and scalar multiplication:

$$(A+B)^T = A^T + B^T, \qquad (c\,A)^T = c\,A^T. \tag{1.54}$$

Transposition is also compatible with matrix multiplication, but with a twist. Like the inverse, the transpose *reverses* the order of multiplication:

$$(A\,B)^T = B^T A^T. \tag{1.55}$$

Indeed, if A has size $m \times n$ and B has size $n \times p$, so they can be multiplied, then A^T has size $n \times m$ and B^T has size $p \times n$, and so, in general, one has no choice but to multiply $B^T A^T$ in that order. Formula (1.55) is a straightforward consequence of the basic laws of matrix multiplication. More generally,

$$(A_1 A_2 \,\cdots\, A_{k-1} A_k)^T = A_k^T A_{k-1}^T \,\cdots\, A_2^T A_1^T.$$

An important special case is the product of a row vector $\mathbf{v}^T$ and a column vector $\mathbf{w}$ with the same number of entries. In this case,

$$\mathbf{v}^T\mathbf{w} = (\mathbf{v}^T\mathbf{w})^T = \mathbf{w}^T\mathbf{v}, \tag{1.56}$$

because their product is a scalar and so, as noted above, equals its own transpose.

Lemma 1.32. If A is a nonsingular matrix, so is A^T, and its inverse is denoted by

$$A^{-T} = (A^T)^{-1} = (A^{-1})^T. \tag{1.57}$$

Thus, transposing a matrix and then inverting yields the same result as first inverting and then transposing.

Proof: Let $X = (A^{-1})^T$. Then, according to (1.55),

$$X\,A^T = (A^{-1})^T A^T = (A\,A^{-1})^T = \mathrm{I}^T = \mathrm{I}.$$

The proof that $A^T X = \mathrm{I}$ is similar, and so we conclude that $X = (A^T)^{-1}$. *Q.E.D.*

Exercises

1.6.1. Write down the transpose of the following matrices: (*a*) $\begin{pmatrix} 1 \\ 5 \end{pmatrix}$, (*b*) $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, (*c*) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, (*d*) $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 2 \end{pmatrix}$, (*e*) $\begin{pmatrix} 1 & 2 & -3 \end{pmatrix}$, (*f*) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, (*g*) $\begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ 1 & 1 & 5 \end{pmatrix}$.

1.6.2. Let $A = \begin{pmatrix} 3 & -1 & -1 \\ 1 & 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 2 \\ 2 & 0 \\ -3 & 4 \end{pmatrix}$. Compute A^T and B^T. Then compute $(A\,B)^T$ and $(B\,A)^T$ without first computing $A\,B$ or $B\,A$.

1.6.3. Show that $(A\,B)^T = A^T B^T$ if and only if A and B are square commuting matrices.

♢ 1.6.4. Prove formula (1.55).

1.6.5. Find a formula for the transposed product $(A\,B\,C)^T$ in terms of A^T, B^T and C^T.

1.6.6. *True or false*: Every square matrix A commutes with its transpose A^T.

♢ 1.6.7. A square matrix is called *normal* if it commutes with its transpose: $A^T A = A\,A^T$. Find all normal 2×2 matrices.

1.6.8. (*a*) Prove that the inverse transpose operation (1.57) respects matrix multiplication: $(A\,B)^{-T} = A^{-T} B^{-T}$. (*b*) Verify this identity for $A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

1.6.9. Prove that if A is an invertible matrix, then $A\,A^T$ and $A^T A$ are also invertible.

1.6.10. If $\mathbf{v}, \mathbf{w}$ are column vectors with the same number of entries, does $\mathbf{v}\,\mathbf{w}^T = \mathbf{w}\,\mathbf{v}^T$?

1.6.11. Is there a matrix analogue of formula (1.56), namely $A^T B = B^T A$?

♢ 1.6.12. (*a*) Let A be an $m \times n$ matrix. Let $\mathbf{e}_j$ denote the $1 \times n$ column vector with a single 1 in the j^{th} entry, as in (1.44). Explain why the product $A\,\mathbf{e}_j$ equals the j^{th} column of A. (*b*) Similarly, let $\widehat{\mathbf{e}}_i$ be the $1 \times m$ column vector with a single 1 in the i^{th} entry. Explain why the triple product $\widehat{\mathbf{e}}_i^T A\,\mathbf{e}_j = a_{ij}$ equals the (i,j) entry of the matrix A.

♢ 1.6.13. Let A and B be $m \times n$ matrices. (a) Suppose that $\mathbf{v}^T A \mathbf{w} = \mathbf{v}^T B \mathbf{w}$ for all vectors $\mathbf{v}, \mathbf{w}$. Prove that $A = B$. (b) Give an example of two matrices such that $\mathbf{v}^T A \mathbf{v} = \mathbf{v}^T B \mathbf{v}$ for all vectors $\mathbf{v}$, but $A \neq B$.

♢ 1.6.14. (a) Explain why the inverse of a permutation matrix equals its transpose: $P^{-1} = P^T$. (b) If $A^{-1} = A^T$, is A necessarily a permutation matrix?

♢ 1.6.15. Let A be a square matrix and P a permutation matrix of the same size. (a) Explain why the product $A P^T$ has the effect of applying the permutation defined by P to the columns of A. (b) Explain the effect of multiplying $P A P^T$. *Hint*: Try this on some 3×3 examples first.

♡ 1.6.16. Let $\mathbf{v}, \mathbf{w}$ be $n \times 1$ column vectors. (a) Prove that in most cases the inverse of the $n \times n$ matrix $A = \mathrm{I} - \mathbf{v}\,\mathbf{w}^T$ has the form $A^{-1} = \mathrm{I} - c\,\mathbf{v}\,\mathbf{w}^T$ for some scalar c. Find all $\mathbf{v}$, $\mathbf{w}$ for which such a result is valid. (b) Illustrate the method when $\mathbf{v} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$. (c) What happens when the method fails?

Factorization of Symmetric Matrices

A particularly important class of square matrices consists of those that are unchanged by the transpose operation.

Definition 1.33. A matrix is called *symmetric* if it equals its own transpose: $A = A^T$.

Thus, A is symmetric if and only if it is square and its entries satisfy $a_{ji} = a_{ij}$ for all i, j. In other words, entries lying in "mirror image" positions relative to the main diagonal must be equal. For example, the most general symmetric 3×3 matrix has the form

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}.$$

Note that all diagonal matrices, including the identity, are symmetric. A lower or upper triangular matrix is symmetric if and only if it is, in fact, a diagonal matrix.

The $L D V$ factorization of a nonsingular matrix takes a particularly simple form if the matrix also happens to be symmetric. This result will form the foundation of some significant later developments.

Theorem 1.34. A symmetric matrix A is regular if and only if it can be factored as

$$A = L D L^T, \tag{1.58}$$

where L is a lower unitriangular matrix and D is a diagonal matrix with nonzero diagonal entries.

Proof: We already know, according to Theorem 1.29, that we can factor

$$A = L D V. \tag{1.59}$$

We take the transpose of both sides of this equation:

$$A^T = (L D V)^T = V^T D^T L^T = V^T D L^T, \tag{1.60}$$

since diagonal matrices are automatically symmetric: $D^T = D$. Note that V^T is lower unitriangular, and L^T is upper unitriangular. Therefore (1.60) is the $L D V$ factorization of A^T.

In particular, if A is symmetric, then

$$L D V = A = A^T = V^T D L^T.$$

Uniqueness of the $L D V$ factorization implies that

$$L = V^T \qquad \text{and} \qquad V = L^T$$

(which are two versions of the same equation). Replacing V by L^T in (1.59) establishes the factorization (1.58). *Q.E.D.*

Remark. If $A = L D L^T$, then A is necessarily symmetric. Indeed,

$$A^T = (L D L^T)^T = (L^T)^T D^T L^T = L D L^T = A.$$

However, not every symmetric matrix has an $L D L^T$ factorization. A simple example is the irregular but nonsingular 2×2 matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Example 1.35. The problem is to find the $L D L^T$ factorization of the particular symmetric matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix}$. This requires performing the usual Gaussian Elimination algorithm. Subtracting twice the first row from the second and also the first row from the third produces the matrix $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}$. We then add one half of the second row of the latter matrix to its third row, resulting in the upper triangular form

$$U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & \frac{5}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \frac{5}{2} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} = D V,$$

which we further factor by dividing each row of U by its pivot. On the other hand, the lower unitriangular matrix associated with the preceding row operations is $L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -\frac{1}{2} & 1 \end{pmatrix}$, which, as guaranteed by Theorem 1.34, is the transpose of $V = L^T$. Therefore, the desired $A = L U = L D L^T$ factorizations of this particular symmetric matrix are

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & \frac{5}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \frac{5}{2} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

Example 1.36. Let us look at a general 2×2 symmetric matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Regularity requires that the first pivot be $a \neq 0$. A single row operation will place A in upper triangular form $U = \begin{pmatrix} a & b \\ 0 & \dfrac{a c - b^2}{a} \end{pmatrix}$, and so A is regular provided $a c - b^2 \neq 0$

also. The associated lower triangular matrix is $L = \begin{pmatrix} 1 & 0 \\ \frac{b}{a} & 1 \end{pmatrix}$. Thus, $A = L U$, as you can check. Finally, $D = \begin{pmatrix} a & 0 \\ 0 & \frac{a c - b^2}{a} \end{pmatrix}$ is just the diagonal part of U, and hence $U = D L^T$, so that the $L D L^T$ factorization is explicitly given by

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{b}{a} & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \frac{a c - b^2}{a} \end{pmatrix} \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix}. \tag{1.61}$$

Exercises

1.6.17. Find all values of a, b, and c for which the following matrices are symmetric:
(a) $\begin{pmatrix} 3 & a \\ 2a - 1 & a - 2 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & a & 2 \\ -1 & b & c \\ b & 3 & 0 \end{pmatrix}$, (c) $\begin{pmatrix} 3 & a + 2b - 2c & -4 \\ 6 & 7 & b - c \\ -a + b + c & 4 & b + 3c \end{pmatrix}$.

1.6.18. List all symmetric (a) 3×3 permutation matrices, (b) 4×4 permutation matrices.

1.6.19. *True or false*: If A is symmetric, then A^2 is symmetric.

♢ 1.6.20. *True or false*: If A is a nonsingular symmetric matrix, then A^{-1} is also symmetric.

♢ 1.6.21. *True or false*: If A and B are symmetric $n \times n$ matrices, so is $A B$.

1.6.22. (a) Show that every diagonal matrix is symmetric. (b) Show that an upper (lower) triangular matrix is symmetric if and only if it is diagonal.

1.6.23. Let A be a symmetric matrix. (a) Show that A^n is symmetric for every nonnegative integer n. (b) Show that $2A^2 - 3A + \mathrm{I}$ is symmetric. (c) Show that every matrix polynomial $p(A)$ of A, cf. Exercise 1.2.35, is a symmetric matrix.

1.6.24. Show that if A is any matrix, then $K = A^T A$ and $L = A A^T$ are both well-defined, symmetric matrices.

1.6.25. Find the $L D L^T$ factorization of the following symmetric matrices:
(a) $\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} -2 & 3 \\ 3 & -1 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 3 & 2 \\ -1 & 2 & 0 \end{pmatrix}$, (d) $\begin{pmatrix} 1 & -1 & 0 & 3 \\ -1 & 2 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}$.

1.6.26. Find the $L D L^T$ factorization of the matrices

$$M_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \text{and} \quad M_4 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

♢ 1.6.27. Prove that the 3×3 matrix $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & -1 \\ 1 & -1 & 3 \end{pmatrix}$ cannot be factored as $A = L D L^T$.

♡ 1.6.28. *Skew-symmetric matrices*: An $n \times n$ matrix J is called *skew-symmetric* if $J^T = -J$. (a) Show that every diagonal entry of a skew-symmetric matrix is zero. (b) Write down an example of a nonsingular skew-symmetric matrix. (c) Can you find a regular skew-symmetric matrix? (d) Show that if J is a nonsingular skew-symmetric matrix, then J^{-1} is also skew-symmetric. Verify this fact for the matrix you wrote down in part (b). (e) Show that if J and K are skew-symmetric, then so are J^T, $J + K$, and $J - K$. What about $J K$? (f) Prove that if J is a skew-symmetric matrix, then $\mathbf{v}^T J \mathbf{v} = 0$ for *all* vectors $\mathbf{v} \in \mathbb{R}^n$.

1.6.29. (a) Prove that every square matrix can be expressed as the sum, $A = S + J$, of a symmetric matrix $S = S^T$ and a skew-symmetric matrix $J = -J^T$.

(b) Write $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ as the sum of symmetric and skew-symmetric matrices.

♢ 1.6.30. Suppose $A = LU$ is a regular matrix. Write down the LU factorization of A^T. Prove that A^T is also regular, and its pivots are the *same* as the pivots of A.

1.7 Practical Linear Algebra

For pedagogical and practical reasons, the examples and exercises we have chosen to illustrate the algorithms are all based on relatively small matrices. When dealing with matrices of moderate size, the differences between the various approaches to solving linear systems (Gauss, Gauss–Jordan, matrix inverse, and so on) are relatively unimportant, particularly if one has a decent computer or even hand calculator to do the tedious parts. However, real-world applied mathematics deals with much larger linear systems, and the design of efficient algorithms is a must. For example, numerical solution schemes for ordinary differential equations will typically lead to matrices with thousands of entries, while numerical schemes for partial differential equations arising in fluid and solid mechanics, weather prediction, image and video processing, quantum mechanics, molecular dynamics, chemical processes, etc., will often require dealing with matrices with more than a million entries. It is not hard for such systems to tax even the most sophisticated supercomputer. Thus, it is essential that we understand the computational details of competing methods in order to compare their efficiency, and thereby gain some experience with the issues underlying the design of high performance numerical algorithms.

The most basic question is this: how many arithmetic operations† — in numerical applications these are almost always performed in floating point with various precision levels — are required to complete an algorithm? The number will directly influence the time spent running the algorithm on a computer. We shall keep track of additions and multiplications separately, since the latter typically take longer to process.‡ But we shall not distinguish between addition and subtraction, nor between multiplication and division, since these typically have the same complexity. We shall also assume that the matrices and vectors we deal with are *generic*, with few, if any, zero entries. Modifications of the basic algorithms for *sparse matrices*, meaning those that have lots of zero entries, are an important topic of research, since these include many of the large matrices that appear in applications to differential equations. We refer the interested reader to more advanced treatments of numerical linear algebra, such as [**21**, **40**, **66**, **89**], for such developments.

First, when multiplying an $n \times n$ matrix A and an $n \times 1$ column vector $\mathbf{b}$, each entry of the product $A\mathbf{b}$ requires n multiplications of the form $a_{ij}\, b_j$ and $n-1$ additions to sum the resulting products. Since there are n entries, this means a total of n^2 multiplications

† For simplicity, we will count only the basic arithmetic operations. But it is worth noting that other issues, such as the number of storage and retrieval operations, may also play a role in estimating the computational complexity of a numerical algorithm.

‡ At least, in traditional computer architectures. New algorithms and new methods for performing basic arithmetic operations on a computer, particularly in high precision arithmetic, make this discussion trickier. For simplicity, we will stay with the "classical" version here.

and $n(n-1) = n^2 - n$ additions. Thus, for a matrix of size $n = 100$, one needs about 10,000 distinct multiplications and a similar number of additions. If $n = 1{,}000{,}000 = 10^6$, then $n^2 = 10^{12}$, which is phenomenally large, and the total time required to perform the computation becomes a significant issue†.

Let us next look at the (regular) Gaussian Elimination algorithm, referring back to our pseudocode program for the notational details. First, we count how many arithmetic operations are based on the j^{th} pivot m_{jj}. For each of the $n-j$ rows lying below it, we must perform one division to compute the factor $l_{ij} = m_{ij}/m_{jj}$ used in the elementary row operation. The entries in the column below the pivot will be set to zero automatically, and so we need only compute the updated entries lying strictly below and to the right of the pivot. There are $(n-j)^2$ such entries in the coefficient matrix and an additional $n-j$ entries in the last column of the augmented matrix. Let us concentrate on the former for the moment. For each of these, we replace m_{ik} by $m_{ik} - l_{ij}\,m_{jk}$, and so must perform one multiplication and one addition. For the j^{th} pivot, there is a total of $(n-j)(n-j+1)$ multiplications — including the initial $n-j$ divisions needed to produce the l_{ij} — and $(n-j)^2$ additions needed to update the coefficient matrix. Therefore, to reduce a regular $n \times n$ matrix to upper triangular form requires a total‡ of

$$\begin{aligned} &\sum_{j=1}^{n} (n-j)(n-j+1) = \frac{n^3-n}{3} \qquad && \text{multiplications, and} \\ &\sum_{j=1}^{n} (n-j)^2 = \frac{2n^3-3n^2+n}{6} && \text{additions.} \end{aligned} \tag{1.62}$$

Thus, when n is large, both involve approximately $\frac{1}{3}n^3$ operations.

We should also be keeping track of the number of operations on the right-hand side of the system. No pivots appear there, and so there are

$$\sum_{j=1}^{n} (n-j) = \frac{n^2-n}{2} \tag{1.63}$$

multiplications and the same number of additions required to produce the right-hand side in the resulting triangular system $U\mathbf{x} = \mathbf{c}$. For large n, this count is considerably smaller than the coefficient matrix totals (1.62). We note that the Forward Substitution equations (1.26) require precisely the same number of arithmetic operations to solve $L\mathbf{c} = \mathbf{b}$ for the right-hand side of the upper triangular system. Indeed, the j^{th} equation

$$c_j = b_j - \sum_{k=1}^{j-1} l_{jk}\,c_k$$

requires $j-1$ multiplications and the same number of additions, giving a total of

$$\sum_{j=1}^{n} (j-1) = \frac{n^2-n}{2}$$

operations of each type. Therefore, to reduce a linear system to upper triangular form, it makes no difference in computational efficiency whether one works directly with the

† See Exercise 1.7.8 for more sophisticated computational algorithms that can be employed to (slightly) speed up multiplication of large matrices.

‡ In Exercise 1.7.4, the reader is asked to prove these summation formulaes by induction.

augmented matrix or employs Forward Substitution after the LU factorization of the coefficient matrix has been established.

The Back Substitution phase of the algorithm can be similarly analyzed. To find the value of

$$x_j = \frac{1}{u_{jj}} \left(c_j - \sum_{k=j+1}^{n} u_{jk}\, x_k \right)$$

once we have computed $x_{j+1}, \ldots, x_n$, requires $n-j+1$ multiplications/divisions and $n-j$ additions. Therefore, Back Substitution requires

$$\begin{aligned} \sum_{j=1}^{n} (n-j+1) &= \frac{n^2+n}{2} \qquad \text{multiplications, along with} \\ \sum_{j=1}^{n} (n-j) &= \frac{n^2-n}{2} \qquad \text{additions.} \end{aligned} \tag{1.64}$$

For n large, both of these are approximately equal to $\frac{1}{2}n^2$. Comparing the counts, we conclude that the bulk of the computational effort goes into the reduction of the coefficient matrix to upper triangular form.

Combining the two counts (1.63–64), we discover that, once we have computed the $A = LU$ decomposition of the coefficient matrix, the Forward and Back Substitution process requires n^2 multiplications and $n^2 - n$ additions to solve a linear system $A\mathbf{x} = \mathbf{b}$. This is exactly the *same* as the number of multiplications and additions needed to compute the product $A^{-1}\mathbf{b}$. Thus, even if we happen to know the inverse of A, it is still *just as efficient* to use Forward and Back Substitution to compute the solution!

On the other hand, the computation of A^{-1} is decidedly more inefficient. There are two possible strategies. First, we can solve the n linear systems (1.45), namely

$$A\mathbf{x} = \mathbf{e}_i, \qquad i = 1, \ldots, n, \tag{1.65}$$

for the individual columns of A^{-1}. This requires first computing the LU decomposition, which uses about $\frac{1}{3}n^3$ multiplications and a similar number of additions, followed by applying Forward and Back Substitution to each of the systems, using $n \cdot n^2 = n^3$ multiplications and $n(n^2 - n) \approx n^3$ additions, for a grand total of about $\frac{4}{3}n^3$ operations of each type in order to compute A^{-1}. Gauss–Jordan Elimination fares no better (in fact, slightly worse), also requiring about the same number, $\frac{4}{3}n^3$, of each type of arithmetic operation. Both algorithms can be made more efficient by exploiting the fact that there are lots of zeros on the right-hand sides of the systems (1.65). Designing the algorithm to avoid adding or subtracting a preordained 0, or multiplying or dividing by a preordained ± 1, reduces the total number of operations required to compute A^{-1} to exactly n^3 multiplications and $n(n-1)^2 \approx n^3$ additions. (Details are relegated to the exercises.) And don't forget that we still need to multiply $A^{-1}\mathbf{b}$ to solve the original system. As a result, solving a linear system with the inverse matrix requires approximately *three* times as many arithmetic operations, and so would take three times as long to complete, as the more elementary Gaussian Elimination and Back Substitution algorithm. This justifies our earlier contention that matrix inversion is inefficient, and, except in very special situations, should never be used for solving linear systems in practice.

Exercises

1.7.1. Solve the following linear systems by (i) Gaussian Elimination with Back Substitution; (ii) the Gauss–Jordan algorithm to convert the augmented matrix to the fully reduced form $\left(\mathrm{I} \mid \mathbf{x} \right)$ with solution $\mathbf{x}$; (iii) computing the inverse of the coefficient matrix, and then multiplying it by the right-hand side. Keep track of the number of arithmetic operations you need to perform to complete each computation, and discuss their relative efficiency.

(a) $\begin{aligned} x - 2y &= 4 \\ 3x + y &= -7, \end{aligned}$ (b) $\begin{aligned} 2x - 4y + 6z &= 6, \\ 3x - 3y + 4z &= -1, \\ -4x + 3y - 4z &= 5, \end{aligned}$ (c) $\begin{aligned} x - 3y &= 1, \\ 3x - 7y + 5z &= -1, \\ -2x + 6y - 5z &= 0. \end{aligned}$

1.7.2. (a) Let A be an $n \times n$ matrix. Which is faster to compute, A^2 or A^{-1}? Justify your answer. (b) What about A^3 versus A^{-1}? (c) How many operations are needed to compute A^k? *Hint*: When $k > 3$, you can get away with less than $k-1$ matrix multiplications!

1.7.3. Which is faster: Back Substitution or multiplying a matrix by a vector? How much faster?

♢ 1.7.4. Use induction to prove the summation formulas (1.62), (1.63) and (1.64).

♡ 1.7.5. Let A be a general $n \times n$ matrix. Determine the exact number of arithmetic operations needed to compute A^{-1} using (a) Gaussian Elimination to factor $PA = LU$ and then Forward and Back Substitution to solve the n linear systems (1.65); (b) the Gauss–Jordan method. Make sure your totals do not count adding or subtracting a known 0, or multiplying or dividing by a known ± 1.

1.7.6. Count the number of arithmetic operations needed to solve a system the "old-fashioned" way, by using elementary row operations of all three types, in the same order as the Gauss–Jordan scheme, to fully reduce the augmented matrix $M = \left(A \mid \mathbf{b} \right)$ to the form $\left(\mathrm{I} \mid \mathbf{d} \right)$, with $\mathbf{x} = \mathbf{d}$ being the solution.

1.7.7. An alternative solution strategy, also called *Gauss–Jordan* in some texts, is, once a pivot is in position, to use elementary row operations of type #1 to eliminate all entries both above and below it, thereby reducing the augmented matrix to diagonal form $\left(D \mid \mathbf{c} \right)$ where $D = \operatorname{diag}(d_1, \ldots, d_n)$ is a diagonal matrix containing the pivots. The solutions $x_i = c_i / d_i$ are then obtained by simple division. Is this strategy more efficient, less efficient, or the same as Gaussian Elimination with Back Substitution? Justify your answer with an exact operations count.

♡ 1.7.8. Here, we describe a remarkable algorithm for matrix multiplication discovered by Strassen, [**82**]. Let $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$, and $C = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} = AB$ be block matrices of size $n = 2m$, where all blocks are of size $m \times m$. (a) Let $D_1 = (A_1 + A_4)(B_1 + B_4)$, $D_2 = (A_1 - A_3)(B_1 + B_2)$, $D_3 = (A_2 - A_4)(B_3 + B_4)$, $D_4 = (A_1 + A_2) B_4$, $D_5 = (A_3 + A_4) B_1$, $D_6 = A_4(B_1 - B_3)$, $D_7 = A_1(B_2 - B_4)$. Show that $C_1 = D_1 + D_3 - D_4 - D_6$, $C_2 = D_4 + D_7$, $C_3 = D_5 - D_6$, $C_4 = D_1 - D_2 - D_5 + D_7$. (b) How many arithmetic operations are required when A and B are 2×2 matrices? How does this compare with the usual method of multiplying 2×2 matrices?
(c) In the general case, suppose we use standard matrix multiplication for the matrix products in $D_1, \ldots, D_7$. Prove that Strassen's Method is faster than the direct algorithm for computing AB by a factor of $\approx \frac{7}{8}$. (d) When A and B have size $n \times n$ with $n = 2^r$, we can recursively apply Strassen's Method to multiply the $2^{r-1} \times 2^{r-1}$ blocks A_i, B_i. Prove that the resulting algorithm requires a total of $7^r = n^{\log_2 7} = n^{2.80735}$ multiplications

and $6(7^{r-1} - 4^{r-1}) \le 7^r = n^{\log_2 7}$ additions/subtractions, versus n^3 multiplications and $n^3 - n^2 \approx n^3$ additions for the ordinary matrix multiplication algorithm. How much faster is Strassen's Method when $n = 2^{10}$? 2^{25}? 2^{100}? (*e*) How might you proceed if the size of the matrices does not happen to be a power of 2? Further developments of these ideas can be found in [**11**, **40**].

Tridiagonal Matrices

Of course, in special cases, the actual arithmetic operation count might be considerably reduced, particularly if A is a sparse matrix with many zero entries. A number of specialized techniques have been designed to handle sparse linear systems. A particularly important class consists of the *tridiagonal matrices*

$$A = \begin{pmatrix} q_1 & r_1 & & & & & \\ p_1 & q_2 & r_2 & & & & \\ & p_2 & q_3 & r_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & p_{n-2} & q_{n-1} & r_{n-1} \\ & & & & p_{n-1} & q_n \end{pmatrix} \qquad (1.66)$$

with all entries zero except for those on the main diagonal, namely $a_{ii} = q_i$, the *subdiagonal*, meaning the $n-1$ entries $a_{i+1,i} = p_i$ immediately below the main diagonal, and the *superdiagonal*, meaning the entries $a_{i,i+1} = r_i$ immediately above the main diagonal. (Blanks are used to indicate 0 entries.) Such matrices arise in the numerical solution of ordinary differential equations and the spline fitting of curves for interpolation and computer graphics. If $A = LU$ is regular, it turns out that the factors are lower and upper *bidiagonal matrices*, of the form

$$L = \begin{pmatrix} 1 & & & & & \\ l_1 & 1 & & & & \\ & l_2 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & l_{n-2} & 1 & \\ & & & & l_{n-1} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} d_1 & u_1 & & & & \\ & d_2 & u_2 & & & \\ & & d_3 & u_3 & & \\ & & & \ddots & \ddots & \\ & & & & d_{n-1} & u_{n-1} \\ & & & & & d_n \end{pmatrix}. \qquad (1.67)$$

Multiplying out LU and equating the result to A leads to the equations

$$\begin{array}{lll} d_1 = q_1, & u_1 = r_1, & l_1 d_1 = p_1, \\ l_1 u_1 + d_2 = q_2, & u_2 = r_2, & l_2 d_2 = p_2, \\ \quad\vdots & \quad\vdots & \quad\vdots \\ l_{j-1} u_{j-1} + d_j = q_j, & u_j = r_j, & l_j d_j = p_j, \\ \quad\vdots & \quad\vdots & \quad\vdots \\ l_{n-2} u_{n-2} + d_{n-1} = q_{n-1}, & u_{n-1} = r_{n-1}, & l_{n-1} d_{n-1} = p_{n-1}, \\ l_{n-1} u_{n-1} + d_n = q_n. & & \end{array} \qquad (1.68)$$

These elementary algebraic equations can be successively solved for the entries of L and U in the following order: $d_1, u_1, l_1, d_2, u_2, l_2, d_3, u_3 \; \ldots$. The original matrix A is regular provided none of the diagonal entries $d_1, d_2, \ldots$ are zero, which allows the recursive procedure to successfully proceed to termination.

Once the LU factors are in place, we can apply Forward and Back Substitution to solve the tridiagonal linear system $A\mathbf{x} = \mathbf{b}$. We first solve the lower triangular system $L\mathbf{c} = \mathbf{b}$ by Forward Substitution, which leads to the recursive equations

$$c_1 = b_1, \qquad c_2 = b_2 - l_1 c_1, \qquad \ldots \qquad c_n = b_n - l_{n-1} c_{n-1}. \tag{1.69}$$

We then solve the upper triangular system $U\mathbf{x} = \mathbf{c}$ by Back Substitution, again recursively:

$$x_n = \frac{c_n}{d_n}, \qquad x_{n-1} = \frac{c_{n-1} - u_{n-1} x_n}{d_{n-1}}, \qquad \ldots \qquad x_1 = \frac{c_1 - u_1 x_2}{d_1}. \tag{1.70}$$

As you can check, there are a total of $5n - 4$ multiplications/divisions and $3n - 3$ additions/subtractions required to solve a general tridiagonal system of n linear equations — a striking improvement over the general case.

Example 1.37. Consider the $n \times n$ tridiagonal matrix

$$A = \begin{pmatrix} 4 & 1 & & & & & \\ 1 & 4 & 1 & & & & \\ & 1 & 4 & 1 & & & \\ & & 1 & 4 & 1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 1 & 4 & 1 \\ & & & & & 1 & 4 \end{pmatrix}$$

in which the diagonal entries are all $q_i = 4$, while the entries immediately above and below the main diagonal are all $p_i = r_i = 1$. According to (1.68), the tridiagonal factorization (1.67) has $u_1 = u_2 = \cdots = u_{n-1} = 1$, while

$$d_1 = 4, \qquad l_j = 1/d_j, \qquad d_{j+1} = 4 - l_j, \qquad j = 1, 2, \ldots, n-1.$$

The computed values are

j	1	2	3	4	5	6	7
d_j	4.0	3.75	3.733333	3.732143	3.732057	3.732051	3.732051
l_j	.25	.266666	.267857	.267942	.267948	.267949	.267949

These converge rapidly to

$$d_j \longrightarrow 2 + \sqrt{3} = 3.732050\ldots, \qquad l_j \longrightarrow 2 - \sqrt{3} = .267949\ldots,$$

which makes the factorization for large n almost trivial. The numbers $2 \pm \sqrt{3}$ are the roots of the quadratic equation $x^2 - 4x + 1 = 0$, and are characterized as the fixed points of the nonlinear iterative system $d_{j+1} = 4 - 1/d_j$.

Exercises

1.7.9. For each of the following tridiagonal systems find the LU factorization of the coefficient matrix, and then solve the system. (a) $\begin{pmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & -2 & 3 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 4 \\ -1 \\ -6 \end{pmatrix}$,
(b) $\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 0 & -1 & 4 & 1 \\ 0 & 0 & -5 & 6 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 6 \\ 7 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & 2 & 0 & 0 \\ -1 & -3 & 0 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & -1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ -2 \\ -3 \\ 1 \end{pmatrix}$.

1.7.10. *True or false*: (a) The product of two tridiagonal matrices is tridiagonal. (b) The inverse of a tridiagonal matrix is tridiagonal.

1.7.11. (a) Find the LU factorization of the $n \times n$ tridiagonal matrix A_n with all 2's along the diagonal and all -1's along the sub- and super-diagonals for $n = 3, 4$, and 5. (b) Use your factorizations to solve the system $A_n \mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (1, 1, 1, \ldots, 1)^T$. (c) Can you write down the LU factorization of A_n for general n? Do the entries in the factors approach a limit as n gets larger and larger? (d) Can you find the solution to the system $A_n \mathbf{x} = \mathbf{b} = (1, 1, 1, \ldots, 1)^T$ for general n?

♠ 1.7.12. Answer Exercise 1.7.11 if the super-diagonal entries of A_n are changed to $+1$.

♠ 1.7.13. Find the LU factorizations of $\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}$, $\begin{pmatrix} 4 & 1 & 0 & 1 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 1 & 0 & 1 & 4 \end{pmatrix}$, $\begin{pmatrix} 4 & 1 & 0 & 0 & 1 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 1 & 0 & 0 & 1 & 4 \end{pmatrix}$.

Do you see a pattern? Try the 6×6 version. The following exercise should now be clear.

♡ 1.7.14. A *tricirculant matrix* $C = \begin{pmatrix} q_1 & r_1 & & & & p_1 \\ p_2 & q_2 & r_2 & & & \\ & p_3 & q_3 & r_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & p_{n-1} & q_{n-1} & r_{n-1} \\ r_n & & & & p_n & q_n \end{pmatrix}$ is tridiagonal except for its $(1, n)$ and $(n, 1)$ entries. Tricirculant matrices arise in the numerical solution of periodic boundary value problems and in spline interpolation.
(a) Prove that if $C = LU$ is regular, its factors have the form

$$\begin{pmatrix} 1 & & & & & & \\ l_1 & 1 & & & & & \\ & l_2 & 1 & & & & \\ & & l_3 & 1 & & & \\ & & & \ddots & \ddots & & \\ & & & & l_{n-2} & 1 & \\ m_1 & m_2 & m_3 & \ldots & m_{n-2} & l_{n-1} & 1 \end{pmatrix}, \quad \begin{pmatrix} d_1 & u_1 & & & & v_1 \\ & d_2 & u_2 & & & v_2 \\ & & d_3 & u_3 & & v_3 \\ & & & \ddots & \ddots & \vdots \\ & & & d_{n-2} & u_{n-2} & v_{n-2} \\ & & & & d_{n-1} & u_{n-1} \\ & & & & & d_n \end{pmatrix}.$$

(b) Compute the LU factorization of the $n \times n$ tricirculant matrix

$$C_n = \begin{pmatrix} 1 & -1 & & & & -1 \\ -1 & 2 & -1 & & & \\ & -1 & 3 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & n-1 & -1 \\ -1 & & & & -1 & 1 \end{pmatrix}$$ for $n = 3, 5$, and 6. What goes wrong when $n = 4$?

♡ 1.7.15. A matrix A is said to have *bandwidth* k if all entries that are more than k slots away from the main diagonal are zero: $a_{ij} = 0$ whenever $|\, i - j \,| > k$. (a) Show that a tridiagonal matrix has band width 1. (*b*) Write down an example of a 6×6 matrix of band width 2 and one of band width 3. (*c*) Prove that the L and U factors of a regular banded matrix have the same band width. (*d*) Find the $L U$ factorization of the matrices you wrote down in part (*b*). (*e*) Use the factorization to solve the system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is the column vector with all entries equal to 1. (*f*) How many arithmetic operations are needed to solve $A\mathbf{x} = \mathbf{b}$ if A is banded? (*g*) Prove or give a counterexample: the inverse of a banded matrix is banded.

Pivoting Strategies

Let us now investigate the practical side of pivoting. As we know, in the irregular situations when a zero shows up in a diagonal pivot position, a row interchange is required to proceed with the elimination algorithm. But even when a nonzero pivot element is in place, there may be good numerical reasons for exchanging rows in order to install a more desirable element in the pivot position. Here is a simple example:

$$.01\, x + 1.6\, y = 32.1, \qquad\qquad x + .6\, y = 22. \tag{1.71}$$

The exact solution to the system is easily found:

$$x = 10, \qquad\qquad y = 20.$$

Suppose we are working with a very primitive calculator that only retains 3 digits of accuracy. (Of course, this is not a very realistic situation, but the example could be suitably modified to produce similar difficulties no matter how many digits of accuracy our computer is capable of retaining.) The augmented matrix is

$$\left(\begin{array}{cc|c} .01 & 1.6 & 32.1 \\ 1 & .6 & 22 \end{array}\right).$$

Choosing the $(1,1)$ entry as our pivot, and subtracting 100 times the first row from the second produces the upper triangular form

$$\left(\begin{array}{cc|c} .01 & 1.6 & 32.1 \\ 0 & -159.4 & -3188 \end{array}\right).$$

Since our calculator has only three–place accuracy, it will round the entries in the second row, producing the augmented coefficient matrix

$$\left(\begin{array}{cc|c} .01 & 1.6 & 32.1 \\ 0 & -159.0 & -3190 \end{array}\right).$$

The solution by Back Substitution gives

$$\begin{aligned} & y = 3190/159 = 20.0628\ldots \ \simeq\ 20.1, \qquad\qquad \text{and then} \\ & x = 100\,(32.1 - 1.6\, y) = 100\,(32.1 - 32.16) \simeq 100\,(32.1 - 32.2) = -10. \end{aligned}$$

The relatively small error in y has produced a very large error in x — not even its sign is correct!

The problem is that the first pivot, .01, is much smaller than the other element, 1, that appears in the column below it. Interchanging the two rows before performing the row

Gaussian Elimination With Partial Pivoting

```
start
    for i = 1 to n
        set r(i) = i
    next i
    for j = 1 to n
        if m_{r(i),j} = 0 for all i ≥ j, stop; print "A is singular"
        choose i > j such that m_{r(i),j} is maximal
        interchange r(i) ⟷ r(j)
        for i = j + 1 to n
            set l_{r(i)j} = m_{r(i)j} / m_{r(j)j}
            for k = j + 1 to n + 1
                set m_{r(i)k} = m_{r(i)k} − l_{r(i)j} m_{r(j)k}
            next k
        next i
    next j
end
```

operation would resolve the difficulty — even with such an inaccurate calculator! After the interchange, we have

$$\left(\begin{array}{cc|c} 1 & .6 & 22 \\ .01 & 1.6 & 32.1 \end{array}\right),$$

which results in the rounded-off upper triangular form

$$\left(\begin{array}{cc|c} 1 & .6 & 22 \\ 0 & 1.594 & 31.88 \end{array}\right) \simeq \left(\begin{array}{cc|c} 1 & .6 & 22 \\ 0 & 1.59 & 31.9 \end{array}\right).$$

The solution by Back Substitution now gives a respectable answer:

$$y = 31.9/1.59 = 20.0628\ldots \simeq 20.1, \qquad x = 22 - .6\,y = 22 - 12.06 \simeq 22 - 12.1 = 9.9.$$

The general strategy, known as *Partial Pivoting*, says that at each stage, we should use the largest (in absolute value) legitimate (i.e., in the pivot column on or below the diagonal) element as the pivot, even if the diagonal element is nonzero. Partial Pivoting can help suppress the undesirable effects of round-off errors during the computation.

In a computer implementation of pivoting, there is no need to waste processor time physically exchanging the row entries in memory. Rather, one introduces a separate array of pointers that serve to indicate which original row is currently in which permuted position. More concretely, one initializes n row pointers $r(1) = 1, \ldots, r(n) = n$. Interchanging row i and row j of the coefficient or augmented matrix is then accomplished by merely interchanging $r(i)$ and $r(j)$. Thus, to access a matrix element that is currently in row i of the augmented matrix, one merely retrieves the element that is in row $r(i)$ in the computer's memory. An explicit implementation of this strategy is provided in the accompanying pseudocode program.

Partial pivoting will solve most problems, although there can still be difficulties. For instance, it does not accurately solve the system

$$10\,x + 1600\,y = 32100, \qquad x + .6\,y = 22,$$

obtained by multiplying the first equation in (1.71) by 1000. The tip-off is that, while the entries in the column containing the pivot are smaller, those in its row are much larger. The solution to this difficulty is *Full Pivoting*, in which one also performs column interchanges — preferably with a column pointer — to move the largest legitimate element into the pivot position. In practice, a column interchange amounts to reordering the variables in the system, which, as long as one keeps proper track of the order, also doesn't change the solutions. Thus, switching the order of x, y leads to the augmented matrix

$$\left(\begin{array}{cc|c} 1600 & 10 & 32100 \\ .6 & 1 & 22 \end{array}\right),$$

in which the first column now refers to y and the second to x. Now Gaussian Elimination will produce a reasonably accurate solution to the system.

Finally, there are some matrices that are hard to handle even with sophisticated pivoting strategies. Such *ill-conditioned* matrices are typically characterized by being "almost" singular. A famous example of an ill-conditioned matrix is the $n \times n$ *Hilbert matrix*

$$H_n = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \cdots & \frac{1}{n+2} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \cdots & \frac{1}{n+3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \frac{1}{n+3} & \cdots & \frac{1}{2n-1} \end{pmatrix}. \tag{1.72}$$

Later, in Proposition 3.40, we will prove that H_n is nonsingular for all n. However, the solution of a linear system whose coefficient matrix is a Hilbert matrix H_n, even for moderately large n, is a very challenging problem, even using high precision computer arithmetic†. This is because the larger n is, the closer H_n is, in a sense, to being singular. A full discussion of the so-called condition number of a matrix can be found in Section 8.7.

The reader is urged to try the following computer experiment. Fix a moderately large value of n, say 20. Choose a column vector $\mathbf{x}$ with n entries chosen at random. Compute $\mathbf{b} = H_n \mathbf{x}$ directly. Then try to solve the system $H_n \mathbf{x} = \mathbf{b}$ by Gaussian Elimination, and compare the result with the original vector $\mathbf{x}$. If you obtain an accurate solution with $n = 20$, try $n = 50$ or 100. This will give you a good indicator of the degree of arithmetic precision used by your computer hardware, and the accuracy of the numerical solution algorithm(s) in your software.

---

† In computer algebra systems such as MAPLE and MATHEMATICA, one can use exact rational arithmetic to perform the computations. Then the important issues are time and computational efficiency. Incidentally, there is an explicit formula for the inverse of a Hilbert matrix, which appears in Exercise 1.7.23.

Exercises

1.7.16. (a) Find the exact solution to the linear system $\begin{pmatrix} .1 & 2.7 \\ 1.0 & .5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10. \\ -6.0 \end{pmatrix}$. (b) Solve the system using Gaussian Elimination with 2-digit rounding. (c) Solve the system using Partial Pivoting and 2-digit rounding. (d) Compare your answers and discuss.

1.7.17. (a) Find the exact solution to the linear system $x - 5y - z = 1$, $\frac{1}{6}x - \frac{5}{6}y + z = 0$, $2x - y = 3$. (b) Solve the system using Gaussian Elimination with 4-digit rounding. (c) Solve the system using Partial Pivoting and 4-digit rounding. Compare your answers.

1.7.18. Answer Exercise 1.7.17 for the system

$$x + 4y - 3z = -3, \qquad 25x + 97y - 35z = 39, \qquad 35x - 22y + 33z = -15.$$

1.7.19. Employ 2 digit arithmetic with rounding to compute an approximate solution of the linear system $0.2x + 2y - 3z = 6$, $5x + 43y + 27z = 58$, $3x + 23y - 42z = -87$, using the following methods: (a) Regular Gaussian Elimination with Back Substitution; (b) Gaussian Elimination with Partial Pivoting; (c) Gaussian Elimination with Full Pivoting. (d) Compare your answers and discuss their accuracy.

1.7.20. Solve the following systems by hand, using pointers instead of physically interchanging the rows: (a) $\begin{pmatrix} 0 & 1 & -2 \\ 1 & -1 & 1 \\ 3 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, (b) $\begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & 0 & -2 & 1 \\ 1 & 0 & 2 & 0 \\ -1 & 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$,

(c) $\begin{pmatrix} 3 & -1 & 2 & -1 \\ 6 & -2 & 4 & 3 \\ 3 & 1 & 0 & -2 \\ -1 & 3 & -2 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$, (d) $\begin{pmatrix} 0 & -1 & 5 & -1 \\ 1 & -2 & 0 & 1 \\ 2 & -3 & -3 & -1 \\ 2 & 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 3 \\ 0 \end{pmatrix}$.

1.7.21. Solve the following systems using Partial Pivoting and pointers:

(a) $\begin{pmatrix} 1 & 5 \\ 2 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & 2 & -1 \\ 4 & -2 & 1 \\ 3 & 5 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$,

(c) $\begin{pmatrix} 1 & -3 & 6 & -1 \\ 2 & -5 & 0 & 1 \\ -1 & -6 & 4 & -2 \\ 3 & 0 & 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}$, (d) $\begin{pmatrix} .01 & 4 & 2 \\ 2 & -802 & 3 \\ 7 & .03 & 250 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 122 \end{pmatrix}$.

1.7.22. Use Full Pivoting with pointers to solve the systems in Exercise 1.7.21.

♠ 1.7.23. Let H_n be the $n \times n$ Hilbert matrix (1.72), and $K_n = H_n^{-1}$ its inverse. It can be proved, [**40**; p. 513], that the (i, j) entry of K_n is

$$(-1)^{i+j}(i+j-1)\binom{n+i-1}{n-j}\binom{n+j-1}{n-i}\binom{i+j-2}{i-1}^2,$$

where $\binom{n}{k} = \dfrac{n!}{k!\,(n-k)!}$ is the standard binomial coefficient. (**Warning.** Proving this formula is a nontrivial combinatorial challenge.) (a) Write down the inverse of the Hilbert matrices H_3, H_4, H_5 using the formula or the Gauss–Jordan Method with exact rational arithmetic. Check your results by multiplying the matrix by its inverse.
(b) Recompute the inverses on your computer using floating point arithmetic and compare with the exact answers. (c) Try using floating point arithmetic to find K_{10} and K_{20}. Test the answer by multiplying the Hilbert matrix by its computed inverse.

♠ 1.7.24. (a) Write out a pseudo-code algorithm, using both row and column pointers, for Gaussian Elimination with Full Pivoting. (b) Implement your code on a computer, and try it on the systems in Exercise 1.7.21.

1.8 General Linear Systems

So far, we have treated only linear systems involving the same number of equations as unknowns, and then only those with nonsingular coefficient matrices. These are precisely the systems that always have a unique solution. We now turn to the problem of solving a general linear system of m equations in n unknowns. The cases not treated as yet are non-square systems, with $m \neq n$, as well as square systems with singular coefficient matrices. The basic idea underlying the Gaussian Elimination algorithm for nonsingular systems can be straightforwardly adapted to these cases, too. One systematically applies the same two types of elementary row operation to reduce the coefficient matrix to a simplified form that generalizes the upper triangular form we aimed for in the nonsingular situation.

Definition 1.38. An $m \times n$ matrix U is said to be in *row echelon form* if it has the following "staircase" structure:

$$U = \begin{pmatrix} \circledast & * & \ldots & * & * & \ldots & * & * & \ldots & \ldots & * & * & * & \ldots & * \\ 0 & 0 & \ldots & 0 & \circledast & \ldots & * & * & \ldots & \ldots & * & * & * & \ldots & * \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & \circledast & \ldots & \ldots & * & * & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & & \ddots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & \ldots & 0 & \circledast & * & \ldots & * \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & \ldots & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

The entries indicated by $\circledast$ are the *pivots*, and must be nonzero. The first r rows of U each contain exactly one pivot, but not all columns are required to include a pivot entry. The entries below the "staircase", indicated by the solid line, are all zero, while the non-pivot entries above the staircase, indicated by stars, can be anything. The last $m - r$ rows are identically zero, and do not contain any pivots. There may, in exceptional situations, be one or more all zero initial columns. Here is an explicit example of a matrix in row echelon form:

$$\begin{pmatrix} 3 & 1 & 0 & 4 & 5 & -7 \\ 0 & -1 & -2 & 1 & 8 & 0 \\ 0 & 0 & 0 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The three pivots are the first nonzero entries in the three nonzero rows, namely, $3, -1, 2$.

Slightly more generally, U may have several initial columns consisting of all zeros. An example is the row echelon matrix

$$\begin{pmatrix} 0 & 0 & 3 & 5 & -2 & 0 \\ 0 & 0 & 0 & 0 & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which also has three pivots. The latter matrix corresponds to a linear system in which the first two variables do not appear in any of the equations. Thus, such row echelon forms almost never appear in applications.

Proposition 1.39. Every matrix can be reduced to row echelon form by a sequence of elementary row operations of types #1 and #2.

In matrix language, Proposition 1.39 implies that if A is any $m \times n$ matrix, then there exists an $m \times m$ permutation matrix P and an $m \times m$ lower unitriangular matrix L such that

$$PA = LU, \tag{1.73}$$

where U is an $m \times n$ row echelon matrix. The factorization (1.73) is not unique. Observe that P and L are square matrices of the same size, while A and U are rectangular, also of the same size. As with a square matrix, the entries of L below the diagonal correspond to the row operations of type #1, while P keeps track of row interchanges. As before, one can keep track of row interchanges with a row pointer.

A constructive proof of this result is based on the general Gaussian Elimination algorithm, which proceeds as follows. Starting on the left of the matrix, one searches for the first column that is not identically zero. Any of the nonzero entries in that column may serve as the pivot. Partial pivoting indicates that it is probably best to choose the largest one, although this is not essential for the algorithm to proceed. One places the chosen pivot in the first row of the matrix via a row interchange, if necessary. The entries below the pivot are made equal to zero by the appropriate elementary row operations of type #1. One then proceeds iteratively, performing the same reduction algorithm on the submatrix consisting of all entries strictly to the right and below the pivot. The algorithm terminates when either there is a nonzero pivot in the last row, or all of the rows lying below the last pivot are identically zero, and so no more pivots can be found.

Example 1.40. The easiest way to learn the general Gaussian Elimination algorithm is to follow through an illustrative example. Consider the linear system

$$\begin{aligned} x + 3y + 2z - u &= a, \\ 2x + 6y + z + 4u + 3v &= b, \\ -x - 3y - 3z + 3u + v &= c, \\ 3x + 9y + 8z - 7u + 2v &= d, \end{aligned} \tag{1.74}$$

of 4 equations in 5 unknowns, where a, b, c, d are given numbers[†]. The coefficient matrix is

$$A = \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 2 & 6 & 1 & 4 & 3 \\ -1 & -3 & -3 & 3 & 1 \\ 3 & 9 & 8 & -7 & 2 \end{pmatrix}. \tag{1.75}$$

To solve the system, we introduce the augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & 0 & a \\ 2 & 6 & 1 & 4 & 3 & b \\ -1 & -3 & -3 & 3 & 1 & c \\ 3 & 9 & 8 & -7 & 2 & d \end{array}\right),$$

obtained by appending the right-hand side of the system. The upper left entry is nonzero, and so can serve as the first pivot. We eliminate the entries below it by elementary row

[†] It will be convenient to work with the right-hand side in general form, although the reader may prefer, at least initially, to assign numerical values to a, b, c, d.

operations, resulting in

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & 0 & a \\ 0 & 0 & -3 & 6 & 3 & b-2a \\ 0 & 0 & -1 & 2 & 1 & c+a \\ 0 & 0 & 2 & -4 & 2 & d-3a \end{array}\right).$$

Now, the second column contains no suitable nonzero entry to serve as the second pivot. (The top entry already lies in a row containing a pivot, and so cannot be used.) Therefore, we move on to the third column, choosing the $(2,3)$ entry, -3, as our second pivot. Again, we eliminate the entries below it, leading to

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & 0 & a \\ 0 & 0 & -3 & 6 & 3 & b-2a \\ 0 & 0 & 0 & 0 & 0 & c-\frac{1}{3}b+\frac{5}{3}a \\ 0 & 0 & 0 & 0 & 4 & d+\frac{2}{3}b-\frac{13}{3}a \end{array}\right).$$

The fourth column has no pivot candidates, and so the final pivot is the 4 in the fifth column. We interchange the last two rows in order to place the coefficient matrix in row echelon form:

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & 0 & a \\ 0 & 0 & -3 & 6 & 3 & b-2a \\ 0 & 0 & 0 & 0 & 4 & d+\frac{2}{3}b-\frac{13}{3}a \\ 0 & 0 & 0 & 0 & 0 & c-\frac{1}{3}b+\frac{5}{3}a \end{array}\right). \tag{1.76}$$

There are three pivots, $1, -3$, and 4, sitting in positions $(1,1)$, $(2,3)$, and $(3,5)$. Note the staircase form, with the pivots on the steps and everything below the staircase being zero. Recalling the row operations used to construct the solution (and keeping in mind that the row interchange that appears at the end also affects the entries of L), we find the factorization (1.73) takes the explicit form

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 2 & 6 & 1 & 4 & 3 \\ -1 & -3 & -3 & 3 & 1 \\ 3 & 9 & 8 & -7 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & -\frac{2}{3} & 1 & 0 \\ -1 & \frac{1}{3} & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 0 & -3 & 6 & 3 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We shall return to find the solution to our linear system after a brief theoretical interlude.

Warning. In the augmented matrix, pivots can *never* appear in the last column, representing the right-hand side of the system. Thus, even if $c-\frac{1}{3}b+\frac{5}{3}a \neq 0$, that entry does not qualify as a pivot.

We now introduce the most important numerical quantity associated with a matrix.

Definition 1.41. The *rank* of a matrix is the number of pivots.

For instance, the rank of the matrix (1.75) equals 3, since its reduced row echelon form, i.e., the first five columns of (1.76), has three pivots. Since there is at most one pivot per row and one pivot per column, the rank of an $m \times n$ matrix is bounded by both m and n, and so

$$0 \le r = \operatorname{rank} A \le \min\{m,n\}. \tag{1.77}$$

The only $m \times n$ matrix of rank 0 is the zero matrix O — which is the only matrix without any pivots.

Proposition 1.42. A square matrix of size $n \times n$ is nonsingular if and only if its rank is equal to n.

Indeed, the only way an $n \times n$ matrix can end up having n pivots is if its reduced row echelon form is upper triangular with nonzero diagonal entries. But a matrix that reduces to such triangular form is, by definition, nonsingular.

Interestingly, the rank of a matrix *does not depend* on which elementary row operations are performed along the way to row echelon form. Indeed, performing a different sequence of row operations — say using Partial Pivoting versus no pivoting — can produce a completely different reduced form. The remarkable result is that all such row echelon forms end up having exactly the same number of pivots, and this number is the rank of the matrix. A formal proof of this fact will appear in Chapter 2; see Theorem 2.49.

Once the coefficient matrix has been reduced to row echelon form $\left(U \mid \mathbf{c} \right)$, the solution to the equivalent linear system $U\mathbf{x} = \mathbf{c}$ proceeds as follows. The first step is to see whether there are any equations that do not have a solution. Suppose one of the rows in the echelon form U is identically zero, but the corresponding entry in the last column $\mathbf{c}$ of the augmented matrix is nonzero. What linear equation would this represent? Well, the coefficients of all the variables are zero, and so the equation is of the form

$$0 = c_i, \tag{1.78}$$

where i is the row's index. If $c_i \neq 0$, then the equation cannot be satisfied — it is *inconsistent*. The reduced system does not have a solution. Since the reduced system was obtained by elementary row operations, the original linear system is *incompatible*, meaning it also has no solutions. *Note*: It takes only one inconsistency to render the entire system incompatible. On the other hand, if $c_i = 0$, so the entire row in the augmented matrix is zero, then (1.78) is merely $0 = 0$, and is trivially satisfied. Such all-zero rows do not affect the solvability of the system.

In our example, the last row in the echelon form (1.76) is all zero, and hence the last entry in the final column must also vanish in order that the system be compatible. Therefore, the linear system (1.74) will have a solution if and only if the right-hand sides a, b, c, d satisfy the linear constraint

$$\tfrac{5}{3}\, a - \tfrac{1}{3}\, b + c = 0. \tag{1.79}$$

In general, if the system is incompatible, there is nothing else to do. Otherwise, every all zero row in the row echelon form of the coefficient matrix also has a zero entry in the last column of the augmented matrix; the system is *compatible* and admits one or more solutions. (If there are no all-zero rows in the coefficient matrix, meaning that every row contains a pivot, then the system is automatically compatible.) To find the solution(s), we split the variables in the system into two classes.

Definition 1.43. In a linear system $U\mathbf{x} = \mathbf{c}$ in row echelon form, the variables corresponding to columns containing a pivot are called *basic variables*, while the variables corresponding to the columns without a pivot are called *free variables*.

The solution to the system then proceeds by an adaptation of the Back Substitution procedure. Working in reverse order, each nonzero equation is solved for the basic variable associated with its pivot. The result is substituted into the preceding equations before they in turn are solved. The solution then specifies all the basic variables as certain combinations of the remaining free variables. As their name indicates, the free variables, if any, are allowed to take on any values whatsoever, and so serve to parameterize the general solution to the system.

Example 1.44. Let us illustrate the solution procedure with our particular system (1.74). The values $a = 0$, $b = 3$, $c = 1$, $d = 1$, satisfy the consistency constraint (1.79), and the corresponding reduced augmented matrix (1.76) is

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & -1 & 0 & 0 \\ 0 & 0 & -3 & 6 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The pivots are found in columns $1, 3, 5$, and so the corresponding variables, x, z, v, are basic; the other variables, y, u, corresponding to the non-pivot columns $2, 4$, are free. Our task is to solve the reduced system

$$\begin{aligned} x + 3\,y + 2\,z - u \qquad\quad &= 0, \\ -3\,z + 6\,u + 3\,v &= 3, \\ 4\,v &= 3, \\ 0 &= 0, \end{aligned}$$

for the basic variables x, z, v in terms of the free variables y, u. As before, this is done in the reverse order, by successively substituting the resulting values in the preceding equation. The result is the general solution

$$v = \tfrac{3}{4}, \qquad z = -1 + 2\,u + v = -\tfrac{1}{4} + 2\,u, \qquad x = -3\,y - 2\,z + u = \tfrac{1}{2} - 3\,y - 3\,u.$$

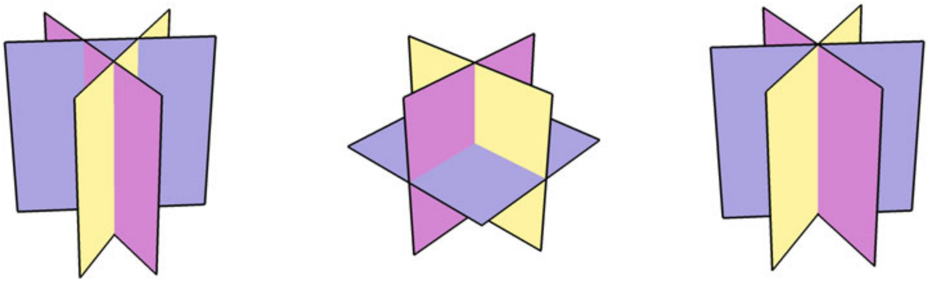
The free variables y, u remain completely arbitrary; any assigned values will produce a solution to the original system. For instance, if $y = -1, u = \pi$, then $x = \frac{7}{2} - 3\pi$, $z = -\frac{1}{4} + 2\pi$, $v = \frac{3}{4}$. But keep in mind that this is merely one of an infinite number of valid solutions.

In general, if the $m\times n$ coefficient matrix of a system of m linear equations in n unknowns has rank r, there are $m-r$ all-zero rows in the row echelon form, and these $m-r$ equations must have zero right-hand side in order that the system be compatible and have a solution. Moreover, there is a total of$r$ basic variables and $n - r$ free variables, and so the general solution depends upon $n - r$ parameters.

Summarizing the preceding discussion, we have learned that there are only three possible outcomes for the solution to a system of linear equations.

Theorem 1.45. A system $A\,\mathbf{x} = \mathbf{b}$ of m linear equations in n unknowns has either (i) exactly one solution, (ii) infinitely many solutions, or (iii) no solution.

Case (iii) occurs if the system is incompatible, producing a zero row in the echelon form that has a nonzero right-hand side. Case (ii) occurs if the system is compatible and there are one or more free variables, and so the rank of the coefficient matrix is strictly less than the number of columns: $r < n$. Case (i) occurs for nonsingular square coefficient matrices, and, more generally, for compatible systems for which $r = n$, implying there are no free variables. Since $r \leq m$, this case can arise only if the coefficient matrix has at least as many rows as columns, i.e., the linear system has at least as many equations as unknowns. A linear system can *never* have a finite number — other than 0 or 1 — of solutions. As a consequence, any linear system that admits two or more solutions automatically has infinitely many!



No Solution Unique Solution Infinitely Many Solutions

Figure 1.1. Intersecting Planes.

Warning. This property requires linearity, and is *not* valid for nonlinear systems. For instance, the real quadratic equation $x^2 + x - 2 = 0$ has exactly two real solutions: $x = 1$ and $x = -2$.

Example 1.46. Consider the linear system

$$y + 4z = a, \qquad 3x - y + 2z = b, \qquad x + y + 6z = c,$$

consisting of three equations in three unknowns. The augmented coefficient matrix is

$$\left(\begin{array}{ccc|c} 0 & 1 & 4 & a \\ 3 & -1 & 2 & b \\ 1 & 1 & 6 & c \end{array}\right).$$

Interchanging the first two rows, and then eliminating the elements below the first pivot leads to

$$\left(\begin{array}{ccc|c} 3 & -1 & 2 & b \\ 0 & 1 & 4 & a \\ 0 & \frac{4}{3} & \frac{16}{3} & c - \frac{1}{3}b \end{array}\right).$$

The second pivot is in the $(2,2)$ position, but after eliminating the entry below it, we find the row echelon form to be

$$\left(\begin{array}{ccc|c} 3 & -1 & 2 & b \\ 0 & 1 & 4 & a \\ 0 & 0 & 0 & c - \frac{1}{3}b - \frac{4}{3}a \end{array}\right).$$

Since there is a row of all zeros, the original coefficient matrix is singular, and its rank is only 2.

The consistency condition follows from this last row in the reduced echelon form, which requires

$$\tfrac{4}{3}a + \tfrac{1}{3}b - c = 0.$$

If this is not satisfied, the system has no solutions; otherwise, it has infinitely many. The free variable is z, since there is no pivot in the third column. The general solution is

$$y = a - 4z, \qquad x = \tfrac{1}{3}b + \tfrac{1}{3}y - \tfrac{2}{3}z = \tfrac{1}{3}a + \tfrac{1}{3}b - 2z,$$

where z is arbitrary.

Geometrically, Theorem 1.45 is telling us about the possible configurations of linear subsets (lines, planes, etc.) of an n-dimensional space. For example, a single linear equation $a\,x + b\,y + c\,z = d$, with $(a, b, c) \neq \mathbf{0}$, defines a plane P in three-dimensional space. The solutions to a system of three linear equations in three unknowns belong to all three planes; that is, they lie in their *intersection* $P_1 \cap P_2 \cap P_3$. Generically, three planes intersect in a single common point; this is case (i) of the theorem, which occurs if and only if the coefficient matrix is nonsingular. The case of infinitely many solutions occurs when the three planes intersect in a common line, or, even more degenerately, when they all coincide. On the other hand, parallel planes, or planes intersecting in parallel lines, have no common point of intersection, and this occurs when the system is incompatible and has no solutions. There are no other possibilities: the total number of points in the intersection is either 0, 1, or ∞. Some sample geometric configurations appear in Figure 1.1.

Exercises

1.8.1. Which of the following systems has (i) a unique solution? (ii) infinitely many solutions? (iii) no solution? In each case, find all solutions: (a) $\begin{aligned} x - 2y &= 1, \\ 3x + 2y &= -3. \end{aligned}$

(b) $\begin{aligned} 2x + y + 3z &= 1, \\ x + 4y - 2z &= -3. \end{aligned}$ (c) $\begin{aligned} x + y - 2z &= -3, \\ 2x - y + 3z &= 7, \\ x - 2y + 5z &= 1. \end{aligned}$ (d) $\begin{aligned} x - 2y + z &= 6, \\ 2x + y - 3z &= -3, \\ x - 3y + 3z &= 10. \end{aligned}$

(e) $\begin{aligned} x - 2y + 2z - w &= 3, \\ 3x + y + 6z + 11w &= 16, \\ 2x - y + 4z + w &= 9. \end{aligned}$ (f) $\begin{aligned} 3x - 2y + z &= 4, \\ x + 3y - 4z &= -3, \\ 2x - 3y + 5z &= 7, \\ x - 8y + 9z &= 10. \end{aligned}$ (g) $\begin{aligned} x + 2y + 17z - 5w &= 50, \\ 9x - 16y + 10z - 8w &= 24, \\ 2x - 5y - 4z &= -13, \\ 6x - 12y + z - 4w &= -1. \end{aligned}$

1.8.2. Determine if the following systems are compatible and, if so, find the general solution:

(a) $\begin{aligned} 6x_1 + 3x_2 &= 12, \\ 4x_1 + 2x_2 &= 9. \end{aligned}$ (b) $\begin{aligned} 8x_1 + 12x_2 &= 16, \\ 6x_1 + 9x_2 &= 13. \end{aligned}$ (c) $\begin{aligned} x_1 + 2x_2 &= 1, \\ 2x_1 + 5x_2 &= 2, \\ 3x_1 + 6x_2 &= 3. \end{aligned}$ (d) $\begin{aligned} 2x_1 - 6x_2 + 4x_3 &= 2, \\ -x_1 + 3x_2 - 2x_3 &= -1. \end{aligned}$

(e) $\begin{aligned} 2x_1 + 2x_2 + 3x_3 &= 1, \\ x_2 + 2x_3 &= 3, \\ 4x_1 + 5x_2 + 7x_3 &= 15. \end{aligned}$ (f) $\begin{aligned} x_1 + x_2 + x_3 + 9x_4 &= 8, \\ x_2 + 2x_3 + 8x_4 &= 7, \\ -3x_1 + x_3 - 7x_4 &= 9. \end{aligned}$ (g) $\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 1, \\ 2x_1 + 4x_2 + 6x_3 + 5x_4 &= 0, \\ 3x_1 + 4x_2 + x_3 + x_4 &= 0, \\ 4x_1 + 6x_2 + 4x_3 - x_4 &= 0. \end{aligned}$

1.8.3. Graph the following planes and determine whether they have a common intersection:

$$x + y + z = 1, \qquad x + y = 1, \qquad x + z = 1.$$

1.8.4. Let $A = \left(\begin{array}{ccc|c} a & 0 & b & 2 \\ a & 2 & a & b \\ b & 2 & a & a \end{array}\right)$ be the augmented matrix for a linear system. For which values of a and b does the system have (i) a unique solution? (ii) infinitely many solutions? (iii) no solution?

1.8.5. Determine the general (complex) solution to the following systems:

(a) $\begin{aligned} 2x + (1+\mathrm{i})y - 2\,\mathrm{i}\,z &= 2\,\mathrm{i}, \\ (1-\mathrm{i})x + y - 2\,\mathrm{i}\,z &= 0. \end{aligned}$ (b) $\begin{aligned} x + 2\,\mathrm{i}\,y + (2 - 4\,\mathrm{i})z &= 5 + 5\,\mathrm{i}, \\ (-1+\mathrm{i})x + 2y + (4 + 2\,\mathrm{i})z &= 0, \\ (1-\mathrm{i})x + (1 + 4\,\mathrm{i})y - 5\,\mathrm{i}\,z &= 10 + 5\,\mathrm{i}. \end{aligned}$

$$(c)\ \begin{aligned} x_1 + \mathrm{i}\,x_2 + x_3 &= 1 + 4\,\mathrm{i}\,, \\ -x_1 + x_2 - \mathrm{i}\,x_3 &= -1, \\ \mathrm{i}\,x_1 - x_2 - x_3 &= -1 - 2\,\mathrm{i}\,. \end{aligned} \qquad (d)\ \begin{aligned} (2+\mathrm{i}\,)x + \mathrm{i}\,y + (2+2\,\mathrm{i}\,)z + (1+12\,\mathrm{i}\,)w &= 0, \\ (1-\mathrm{i}\,)x + y + (2-\mathrm{i}\,)z + (8+2\,\mathrm{i}\,)w &= 0, \\ (3+2\,\mathrm{i}\,)x + \mathrm{i}\,y + (3+3\,\mathrm{i}\,)z + 19\,\mathrm{i}\,w &= 0. \end{aligned}$$

1.8.6. For which values of b and c does the system $x_1 + x_2 + b\,x_3 = 1$, $b\,x_1 + 3\,x_2 - x_3 = -2$, $3\,x_1 + 4\,x_2 + x_3 = c$, have (*a*) no solution? (*b*) exactly one solution? (*c*) infinitely many solutions?

1.8.7. Determine the rank of the following matrices: (*a*) $\begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$, (*b*) $\begin{pmatrix} 2 & 1 & 3 \\ -2 & -1 & -3 \end{pmatrix}$,

(*c*) $\begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 2 \\ -1 & 1 & 0 \end{pmatrix}$, (*d*) $\begin{pmatrix} 2 & -1 & 0 \\ 2 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$, (*e*) $\begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}$, (*f*) $(\,0 \quad -1 \quad 2 \quad 5\,)$,

(*g*) $\begin{pmatrix} 0 & -3 \\ 4 & -1 \\ 1 & 2 \\ -1 & -5 \end{pmatrix}$, (*h*) $\begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 1 & -1 & 0 \\ 1 & 2 & -3 & -1 \\ 4 & -1 & 3 & 2 \\ 0 & 3 & -5 & -2 \end{pmatrix}$, (*i*) $\begin{pmatrix} 0 & 0 & 0 & 3 & 1 \\ 1 & 2 & -3 & 1 & -2 \\ 2 & 4 & -2 & 1 & -2 \end{pmatrix}$.

1.8.8. Write out a $P A = L U$ factorization for each of the matrices in Exercise 1.8.7.

1.8.9. Construct a system of three linear equations in three unknowns that has (*a*) one and only one solution; (*b*) more than one solution; (*c*) no solution.

1.8.10. Find a coefficient matrix A such that the associated linear system $A\,\mathbf{x} = \mathbf{b}$ has (a) infinitely many solutions for every $\mathbf{b}$; (*b*) 0 or ∞ solutions, depending on $\mathbf{b}$; (*c*) 0 or 1 solution depending on $\mathbf{b}$; (*d*) exactly 1 solution for all $\mathbf{b}$.

1.8.11. Give an example of a *nonlinear* system of two equations in two unknowns that has (*a*) no solution; (*b*) exactly two solutions; (*c*) exactly three solutions; (*d*) infinitely many solutions.

1.8.12. What does it mean if a linear system has a coefficient matrix with a column of all 0's?

1.8.13. *True or false*: One can find an $m \times n$ matrix of rank r for every $0 \le r \le \min\{m, n\}$.

1.8.14. *True or false*: Every $m \times n$ matrix has (*a*) exactly m pivots; (*b*) at least one pivot.

♥ 1.8.15. (*a*) Prove that the product $A = \mathbf{v}\,\mathbf{w}^T$ of a nonzero $m \times 1$ column vector $\mathbf{v}$ and a nonzero $1 \times n$ row vector $\mathbf{w}^T$ is an $m \times n$ matrix of rank $r = 1$. (*b*) Compute the following rank one products: (*i*) $\begin{pmatrix} 1 \\ 3 \end{pmatrix} (\,-1 \quad 2\,)$, (*ii*) $\begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} (\,-2 \quad 1\,)$, (*iii*) $\begin{pmatrix} 2 \\ -3 \end{pmatrix} (\,1 \quad 3 \quad -1\,)$.
(*c*) Prove that every rank one matrix can be written in the form $A = \mathbf{v}\,\mathbf{w}^T$.

♦ 1.8.16. (*a*) Let A be an $m \times n$ matrix and let $M = \big(\, A \mid \mathbf{b} \,\big)$ be the augmented matrix for the linear system $A\,\mathbf{x} = \mathbf{b}$. Show that either (*i*) $\operatorname{rank} A = \operatorname{rank} M$, or (*ii*) $\operatorname{rank} A = \operatorname{rank} M - 1$. (*b*) Prove that the system is compatible if and only if case (*i*) holds.

1.8.17. Find the rank of the matrix $\begin{pmatrix} a & a\,r & \ldots & a\,r^{n-1} \\ a\,r^{n} & a\,r^{n+1} & \ldots & a\,r^{2n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a\,r^{(n-1)n} & a\,r^{(n-1)n+1} & \ldots & a\,r^{n^2-1} \end{pmatrix}$ when $a, r \ne 0$.

1.8.18. Find the rank of the $n \times n$ matrix $\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ n+1 & n+2 & n+3 & \ldots & 2n \\ 2n+1 & 2n+2 & 2n+3 & \ldots & 3n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n^2-n+1 & n^2-n+2 & \ldots & \ldots & n^2 \end{pmatrix}$.

1.8.19. Find two matrices A, B such that $\operatorname{rank} AB \neq \operatorname{rank} BA$.

♢ 1.8.20. Let A be an $m \times n$ matrix of rank r. (a) Suppose $C = (A \; B)$ is an $m \times k$ matrix, $k > n$, whose first n columns are the same as the columns of A. Prove that $\operatorname{rank} C \geq \operatorname{rank} A$. Give an example with $\operatorname{rank} C = \operatorname{rank} A$; with $\operatorname{rank} C > \operatorname{rank} A$. (b) Let $E = \begin{pmatrix} A \\ D \end{pmatrix}$ be a $j \times n$ matrix, $j > m$, whose first m rows are the same as those of A. Prove that $\operatorname{rank} E \geq \operatorname{rank} A$. Give an example with $\operatorname{rank} E = \operatorname{rank} A$; with $\operatorname{rank} E > \operatorname{rank} A$.

♢ 1.8.21. Let A be a singular square matrix. Prove that there exist elementary matrices $E_1, \ldots, E_N$ such that $A = E_1 E_2 \cdots E_N Z$, where Z is a matrix with at least one all-zero row.

Homogeneous Systems

A linear system with all 0's on the right-hand side is called *homogeneous*. Conversely, if at least one of the right-hand sides is nonzero, the system is called *inhomogeneous*.

In matrix notation, a homogeneous system takes the form

$$A\mathbf{x} = \mathbf{0}, \tag{1.80}$$

where the zero vector $\mathbf{0}$ indicates that every entry on the right-hand side is zero. Homogeneous systems are always compatible, since $\mathbf{x} = \mathbf{0}$ is a solution, known as the *trivial solution*. If a homogeneous system has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$, then Theorem 1.45 assures us that it must have infinitely many solutions. This will occur if and only if the reduced system has one or more free variables.

Theorem 1.47. A homogeneous linear system $A\mathbf{x} = \mathbf{0}$ of m equations in n unknowns has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$ if and only if the rank of A is $r < n$. If $m < n$, the system *always* has a nontrivial solution. If $m = n$, the system has a nontrivial solution if and only if A is singular.

Thus, homogeneous systems with fewer equations than unknowns always have infinitely many solutions. Indeed, the coefficient matrix of such a system has more columns than rows, and so at least one column cannot contain a pivot, meaning that there is at least one free variable in the general solution formula.

Example 1.48. Consider the homogeneous linear system

$$\begin{aligned} 2x_1 + x_2 + 5x_4 &= 0, \\ 4x_1 + 2x_2 - x_3 + 8x_4 &= 0, \\ -2x_1 - x_2 + 3x_3 - 4x_4 &= 0, \end{aligned}$$

with coefficient matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 5 \\ 4 & 2 & -1 & 8 \\ -2 & -1 & 3 & -4 \end{pmatrix}.$$

Since there are only three equations in four unknowns, we already know that the system has infinitely many solutions, including the trivial solution $x_1 = x_2 = x_3 = x_4 = 0$.

When solving a homogeneous system, the final column of the augmented matrix consists of all zeros. As such, it will never be altered by row operations, and so it is a waste of effort to carry it along during the process. We therefore perform the Gaussian Elimination algorithm directly on the coefficient matrix A. Working with the $(1,1)$ entry as the first pivot, we first obtain

$$\begin{pmatrix} 2 & 1 & 0 & 5 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 3 & 1 \end{pmatrix}.$$

The $(2,3)$ entry is the second pivot, and we apply one final row operation to place the matrix in row echelon form

$$\begin{pmatrix} 2 & 1 & 0 & 5 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & -5 \end{pmatrix}.$$

This corresponds to the reduced homogeneous system

$$2x_1 + x_2 + 5x_4 = 0, \qquad -x_3 - 2x_4 = 0, \qquad -5x_4 = 0.$$

Since there are three pivots in the final row echelon form, the rank of the coefficient matrix A is 3. There is one free variable, namely x_2. Using Back Substitution, we easily obtain the general solution

$$x_1 = -\tfrac{1}{2}t, \qquad x_2 = t, \qquad x_3 = x_4 = 0,$$

which depends upon a single free parameter $t = x_2$.

Example 1.49. Consider the homogeneous linear system

$$\begin{aligned} 2x - y + 3z &= 0, \\ -4x + 2y - 6z &= 0, \\ 2x - y + z &= 0, \\ 6x - 3y + 3z &= 0, \end{aligned} \qquad \text{with coefficient matrix} \qquad A = \begin{pmatrix} 2 & -1 & 3 \\ -4 & 2 & -6 \\ 2 & -1 & 1 \\ 6 & -3 & 3 \end{pmatrix}.$$

The system admits the trivial solution $x = y = z = 0$, but in this case we need to complete the elimination algorithm before we know for sure whether there are other solutions. After the first stage in the reduction process, the coefficient matrix becomes $\begin{pmatrix} 2 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & -6 \end{pmatrix}$. To continue, we need to interchange the second and third rows to place a nonzero entry in the final pivot position; after that, the reduction to the row echelon form $\begin{pmatrix} 2 & -1 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is immediate. Thus, the system reduces to the equations

$$2x - y + 3z = 0, \qquad -2z = 0, \qquad 0 = 0, \qquad 0 = 0.$$

The third and fourth equations are trivially compatible, as they must be in the homogeneous case. The rank of the coefficient matrix is equal to two, which is less than the number of columns, and so, even though the system has more equations than unknowns, it has infinitely many solutions. These can be written in terms of the free variable y; the general solution is $x = \frac{1}{2}y, \ z = 0$, where y is arbitrary.

Exercises

1.8.22. Solve the following homogeneous linear systems.

$$(a)\ \begin{aligned} x + y - 2z &= 0,\\ -x + 4y - 3z &= 0. \end{aligned} \qquad (b)\ \begin{aligned} 2x + 3y - z &= 0,\\ -4x + 3y - 5z &= 0,\\ x - 3y + 3z &= 0. \end{aligned} \qquad (c)\ \begin{aligned} -x + y - 4z &= 0,\\ -2x + 2y - 6z &= 0,\\ x + 3y + 3z &= 0. \end{aligned}$$

$$(d)\ \begin{aligned} x + 2y - 2z + w &= 0,\\ -3x + z - 2w &= 0. \end{aligned} \qquad (e)\ \begin{aligned} -x + 3y - 2z + w &= 0,\\ -2x + 5y + z - 2w &= 0,\\ 3x - 8y + z - 4w &= 0. \end{aligned} \qquad (f)\ \begin{aligned} -y + z &= 0,\\ 2x - 3w &= 0,\\ x + y - 2w &= 0,\\ y - 3z + w &= 0. \end{aligned}$$

1.8.23. Find all solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$ for the coefficient matrix

$$(a)\ \begin{pmatrix} 3 & -1\\ -9 & 3 \end{pmatrix}, \quad (b)\ \begin{pmatrix} 2 & -1 & 4\\ 3 & 1 & 2 \end{pmatrix}, \quad (c)\ \begin{pmatrix} 1 & -2 & 3 & -3\\ 2 & 1 & 4 & 0 \end{pmatrix}, \quad (d)\ \begin{pmatrix} 1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9 \end{pmatrix},$$

$$(e)\ \begin{pmatrix} 0 & 2 & -1\\ -2 & 0 & 3\\ 1 & 3 & 0 \end{pmatrix}, \quad (f)\ \begin{pmatrix} 1 & -2\\ 1 & -1\\ 2 & -1\\ 1 & 0 \end{pmatrix}, \quad (g)\ \begin{pmatrix} 1 & 2 & 0\\ -1 & -3 & 2\\ 4 & 7 & 2\\ -1 & 1 & 6 \end{pmatrix}, \quad (h)\ \begin{pmatrix} 0 & 0 & 3 & -3\\ 1 & -1 & 0 & 3\\ 2 & -2 & 1 & 5\\ -1 & 1 & 1 & -4 \end{pmatrix}.$$

1.8.24. Let U be an upper triangular matrix. Show that the homogeneous system $U\mathbf{x} = \mathbf{0}$ admits a nontrivial solution if and only if U has at least one 0 on its diagonal.

1.8.25. Find the solution to the homogeneous system $2x_1 + x_2 - 2x_3 = 0$, $2x_1 - x_2 - 2x_3 = 0$. Then solve the inhomogeneous version where the right-hand sides are changed to a, b, respectively. What do you observe?

1.8.26. Answer Exercise 1.8.25 for the system $2x_1 + x_2 + x_3 - x_4 = 0$, $2x_1 - 2x_2 - x_3 + 3x_4 = 0$.

1.8.27. Find all values of k for which the following homogeneous systems of linear equations have a non-trivial solution:

$$(a)\ \begin{aligned} x + ky &= 0,\\ kx + 4y &= 0, \end{aligned} \qquad (b)\ \begin{aligned} x_1 + kx_2 + 4x_3 &= 0,\\ kx_1 + x_2 + 2x_3 &= 0,\\ 2x_1 + kx_2 + 8x_3 &= 0. \end{aligned} \qquad (c)\ \begin{aligned} x + ky + 2z &= 0,\\ 3x - ky - 2z &= 0,\\ (k+1)x - 2y - 4z &= 0,\\ kx + 3y + 6z &= 0. \end{aligned}$$

1.9 Determinants

You may be surprised that, so far, we have not mentioned determinants — a topic that typically assumes a central role in many treatments of basic linear algebra. Determinants can be useful in low-dimensional and highly structured problems, and have many fascinating properties. They also prominently feature in theoretical developments of the subject. But, like matrix inverses, they are almost completely irrelevant when it comes to large scale applications and practical computations. Indeed, for most matrices, the best way to compute a determinant is (surprise) Gaussian Elimination! Consequently, from a computational standpoint, the determinant adds no new information concerning the linear system and its solutions. However, for completeness and in preparation for certain later developments (particularly computing eigenvalues of small matrices), you should be familiar with the basic facts and properties of determinants, as summarized in this final section.

The determinant of a square matrix[†] A is a scalar, written $\det A$, that will distinguish between singular and nonsingular matrices. We already encountered in (1.38) the determinant of a 2×2 matrix[‡]: $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,d - b\,c$. The key fact is that the determinant is nonzero if and only if the matrix has an inverse, or, equivalently, is nonsingular. Our goal is to find an analogous quantity for general square matrices.

There are many different ways to define determinants. The difficulty is that the actual formula is very unwieldy — see (1.87) below — and not well motivated. We prefer an axiomatic approach that explains how our three elementary row operations affect the determinant.

Theorem 1.50. Associated with every square matrix, there exists a uniquely defined scalar quantity, known as its *determinant*, that obeys the following axioms:

(*i*) Adding a multiple of one row to another does not change the determinant.

(*ii*) Interchanging two rows changes the sign of the determinant.

(*iii*) Multiplying a row by any scalar (including zero) multiplies the determinant by the same scalar.

(*iv*) The determinant of an upper triangular matrix U is equal to the product of its diagonal entries: $\det U = u_{11} u_{22} \cdots u_{nn}$.

In particular, axiom (*iv*) implies that the determinant of the identity matrix is

$$\det \mathrm{I} = 1. \tag{1.81}$$

Checking that all four of these axioms hold in the 2×2 case is an elementary exercise.

The proof of Theorem 1.50 is based on the following results. Suppose, in particular, we multiply a row of the matrix A by the zero scalar. The resulting matrix has a row of all zeros, and, by axiom (*iii*), has zero determinant. Since any matrix with a zero row can be obtained in this fashion, we conclude:

Lemma 1.51. Any matrix with one or more all-zero rows has zero determinant.

Using these properties, one is able to compute the determinant of any square matrix by Gaussian Elimination, which is, in fact, the fastest and most practical computational method in all but the simplest situations.

Theorem 1.52. If $A = L U$ is a regular matrix, then

$$\det A = \det U = u_{11} u_{22} \cdots u_{nn} \tag{1.82}$$

equals the product of the pivots. More generally, if A is nonsingular, and requires k row interchanges to arrive at its permuted factorization $P A = L U$, then

$$\det A = \det P \, \det U = (-1)^k \, u_{11} u_{22} \cdots u_{nn}. \tag{1.83}$$

Finally, A is singular if and only if

$$\det A = 0. \tag{1.84}$$

† Non-square matrices do not have determinants.

‡ Some authors use vertical lines to indicate the determinant: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Proof: In the regular case, we need only elementary row operations of type #1 to reduce A to upper triangular form U, and axiom (i) says these do not change the determinant. Therefore, $\det A = \det U$, the formula for the latter being given by axiom (iv). The nonsingular case follows in a similar fashion. By axiom (ii), each row interchange changes the sign of the determinant, and so $\det A$ equals $\det U$ if there has been an even number of interchanges, but equals $-\det U$ if there has been an odd number. For the same reason, the determinant of the permutation matrix P equals $+1$ if there has been an even number of row interchanges, and -1 for an odd number. Finally, if A is singular, then we can reduce it to a matrix with at least one row of zeros by elementary row operations of types #1 and #2. Lemma 1.51 implies that the resulting matrix has zero determinant, and so $\det A = 0$, also. *Q.E.D.*

Remark. If we then apply Gauss–Jordan elimination to reduce the upper triangular matrix U to the identity matrix I, and use axiom (ii) when each row is divided by its pivot, we find that axiom (iv) follows from the simpler formula (1.81), which could thus replace it in Theorem 1.50.

Example 1.53. Let us compute the determinant of the 4×4 matrix

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 2 & 1 & -3 & 4 \\ 0 & 2 & -2 & 3 \\ 1 & 1 & -4 & -2 \end{pmatrix}.$$

We perform our usual Gaussian Elimination algorithm, successively leading to the matrices

$$A \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 2 & -2 & 3 \\ 0 & 1 & -3 & -4 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & -2 & -4 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -2 & -4 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

where we used a single row interchange to obtain the final upper triangular form. Owing to the row interchange, the determinant of the original matrix is -1 times the product of the pivots:

$$\det A = -1 \cdot (1 \cdot 1 \cdot (-2) \cdot 3) = 6.$$

In particular, this tells us that A is nonsingular. But, of course, this was already evident, since we successfully reduced the matrix to upper triangular form with 4 nonzero pivots.

There is a variety of other approaches to evaluating determinants. However, except for very small (2×2 or 3×3) matrices or other special situations, the most efficient algorithm for computing the determinant of a matrix is to apply Gaussian Elimination, with pivoting if necessary, and then invoke the relevant formula from Theorem 1.52. In particular, the determinantal criterion (1.84) for singular matrices, while of theoretical interest, is unnecessary in practice, since we will have already detected whether the matrix is singular during the course of the elimination procedure by observing that it has fewer than the full number of pivots.

Let us finish by stating a few of the basic properties of determinants. Proofs are outlined in the exercises.

Proposition 1.54. The determinant of the product of two square matrices of the same size is the product of their determinants:

$$\det(A\,B) = \det A \;\det B. \tag{1.85}$$

Therefore, even though matrix multiplication is not commutative, and so $A\,B \neq B\,A$ in general, both matrix products have the same determinant:

$$\det(A\,B) = \det A \;\det B = \det B \;\det A = \det(B\,A),$$

because ordinary (scalar) multiplication *is* commutative. In particular, setting $B = A^{-1}$ and using axiom (*iv*), we find that the determinant of the inverse matrix is the reciprocal of the matrix's determinant.

Proposition 1.55. If A is a nonsingular matrix, then

$$\det A^{-1} = \frac{1}{\det A}. \tag{1.86}$$

Finally, for later reference, we end with the general formula for the determinant of an $n \times n$ matrix A with entries a_{ij}:

$$\det A = \sum_{\pi} (\operatorname{sign}\pi)\, a_{\pi(1),1}\, a_{\pi(2),2} \;\cdots\; a_{\pi(n),n}. \tag{1.87}$$

The sum is over all possible permutations π of the rows of A. The *sign* of the permutation, written $\operatorname{sign}\pi$, equals the determinant of the corresponding permutation matrix P, so $\operatorname{sign}\pi = \det P = +1$ if the permutation is composed of an even number of row interchanges and -1 if composed of an odd number. For example, the six terms in the well-known formula

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}\,a_{22}\,a_{33} + a_{31}\,a_{12}\,a_{23} + a_{21}a_{32}\,a_{13} - \\ \quad - a_{11}\,a_{32}\,a_{23} - a_{21}\,a_{12}\,a_{33} - a_{31}\,a_{22}\,a_{13} \end{array} \tag{1.88}$$

for a 3×3 determinant correspond to the six possible permutations (1.31) of a 3-rowed matrix. A proof that the formula (1.87) satisfies the defining properties of the determinant listed in Theorem 1.50 is tedious, but not hard. The reader might wish to try out the 3×3 case to be convinced that it works.

The explicit formula (1.87) proves that the determinant function is well-defined, and formally completes the proof of Theorem 1.50. One consequence of this formula is that the determinant is unaffected by the transpose operation.

Proposition 1.56. Transposing a matrix does not change its determinant:

$$\det A^T = \det A. \tag{1.89}$$

Remark. Proposition 1.56 has the interesting consequence that one can equally well use "elementary column operations" to compute determinants. We will not develop this approach in any detail here, since it does not help us to solve linear equations.

However, the explicit determinant formula (1.87) is not used in practice. Since there are $n!$ different permutations of the n rows, the determinantal sum (1.87) contains $n!$ distinct terms, which, as soon as n is of moderate size, renders it completely useless for practical computations. For instance, the determinant of a 10×10 matrix contains $10! = 3{,}628{,}800$

terms, while a 100×100 determinant would require summing 9.3326×10^{157} terms, each of which is a product of 100 matrix entries! The most efficient way to compute determinants is still our mainstay — Gaussian Elimination, coupled with the fact that the determinant is $\pm$ the product of the pivots! On this note, we conclude our brief introduction.

Exercises

1.9.1. Use Gaussian Elimination to find the determinant of the following matrices:

$$(a)\ \begin{pmatrix} 2 & -1 \\ -4 & 3 \end{pmatrix},\quad (b)\ \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix},\quad (c)\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 8 & 10 \end{pmatrix},\quad (d)\ \begin{pmatrix} 0 & 1 & -1 \\ -2 & 1 & 3 \\ 2 & 7 & -8 \end{pmatrix},$$

$$(e)\ \begin{pmatrix} 5 & -1 & 0 & 2 \\ 0 & 3 & -1 & 5 \\ 0 & 0 & -4 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix},\quad (f)\ \begin{pmatrix} 1 & -2 & 1 & 4 \\ 2 & -4 & 0 & 0 \\ 3 & -4 & 2 & 5 \\ 0 & 2 & -4 & -9 \end{pmatrix},\quad (g)\ \begin{pmatrix} 1 & -2 & 1 & 4 & -5 \\ 1 & 1 & -2 & 3 & -3 \\ 2 & -1 & -1 & 2 & 2 \\ 5 & -1 & 0 & 5 & 5 \\ 2 & 2 & 0 & 4 & -1 \end{pmatrix}.$$

1.9.2. Verify the determinant product formula (1.85) when

$$A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -1 & 1 \\ 4 & -2 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -3 & -2 \\ 2 & 0 & 1 \end{pmatrix}.$$

1.9.3. (*a*) Give an example of a non-diagonal 2×2 matrix for which $A^2 = \mathrm{I}$. (*b*) In general, if $A^2 = \mathrm{I}$, show that $\det A = \pm 1$. (*c*) If $A^2 = A$, what can you say about $\det A$?

1.9.4. *True or false*: If true, explain why. If false, give an explicit counterexample. (*a*) If $\det A \neq 0$ then A^{-1} exists. (*b*) $\det(2A) = 2 \det A$. (*c*) $\det(A + B) = \det A + \det B$. (*d*) $\det A^{-T} = \dfrac{1}{\det A}$. (*e*) $\det(AB^{-1}) = \dfrac{\det A}{\det B}$. (*f*) $\det[(A+B)(A-B)] = \det(A^2 - B^2)$. (*g*) If A is an $n \times n$ matrix with $\det A = 0$, then $\operatorname{rank} A < n$. (*h*) If $\det A = 1$ and $AB = \mathrm{O}$, then $B = \mathrm{O}$.

1.9.5. Prove that the similar matrices $B = S^{-1} A S$ have the same determinant: $\det A = \det B$.

1.9.6. Prove that if A is a $n \times n$ matrix and c is a scalar, then $\det(c\,A) = c^n \det A$.

1.9.7. Prove that the determinant of a lower triangular matrix is the product of its diagonal entries.

1.9.8. (*a*) Show that if A has size $n \times n$, then $\det(-A) = (-1)^n \det A$. (*b*) Prove that, for n odd, any $n \times n$ skew-symmetric matrix $A = -A^T$ is singular. (*c*) Find a nonsingular skew-symmetric matrix.

♢ 1.9.9. Prove directly that the 2×2 determinant formula (1.38) satisfies the four determinant axioms listed in Theorem 1.50.

♢ 1.9.10. In this exercise, we prove the determinantal product formula (1.85). (*a*) Prove that if E is any elementary matrix (of the appropriate size), then $\det(EB) = \det E \ \det B$. (*b*) Use induction to prove that if $A = E_1 E_2 \cdots E_N$ is a product of elementary matrices, then $\det(AB) = \det A \ \det B$. Explain why this proves the product formula whenever A is a nonsingular matrix. (*c*) Prove that if Z is a matrix with a zero row, then ZB also has a zero row, and so $\det(ZB) = 0 = \det Z \ \det B$. (*d*) Use Exercise 1.8.21 to complete the proof of the product formula.

1.9.11. Prove (1.86).

♢ 1.9.12. Prove (1.89). *Hint*: Use Exercise 1.6.30 in the regular case. Then extend to the nonsingular case. Finally, explain why the result also holds for singular matrices.

1.9.13. Write out the formula for a 4×4 determinant. It should contain $24 = 4!$ terms.

◇ 1.9.14. Show that (1.87) satisfies all four determinant axioms, and hence is the correct formula for a determinant.

◇ 1.9.15. Prove that axiom (*iv*) in Theorem 1.50 can be proved as a consequence of the first three axioms and the property $\det \mathrm{I} = 1$.

◇ 1.9.16. Prove that one cannot produce an elementary row operation of type #2 by a combination of elementary row operations of type #1.

♡ 1.9.17. Show that (*a*) if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is regular, then its pivots are a and $\dfrac{\det A}{a}$; (*b*) if $A = \begin{pmatrix} a & b & e \\ c & d & f \\ g & h & j \end{pmatrix}$ is regular, then its pivots are a, $\dfrac{ad - bc}{a}$, and $\dfrac{\det A}{ad - bc}$. (*c*) Can you generalize this observation to regular $n \times n$ matrices?

♡ 1.9.18. In this exercise, we justify the use of "elementary column operations" to compute determinants. Prove that (*a*) adding a scalar multiple of one column to another does not change the determinant; (*b*) multiplying a column by a scalar multiplies the determinant by the same scalar; (*c*) interchanging two columns changes the sign of the determinant. (*d*) Explain how to use elementary column operations to reduce a matrix to lower triangular form and thereby compute its determinant.

◇ 1.9.19. Find the determinant of the Vandermonde matrices listed in Exercise 1.3.24. Can you guess the general $n \times n$ formula?

♡ 1.9.20. *Cramer's Rule.* (a) Show that the nonsingular system $ax + by = p$, $cx + dy = q$ has the solution given by the determinantal ratios

$$x = \frac{1}{\Delta} \det \begin{pmatrix} p & b \\ q & d \end{pmatrix}, \qquad y = \frac{1}{\Delta} \det \begin{pmatrix} a & p \\ c & q \end{pmatrix}, \qquad \text{where} \qquad \Delta = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \qquad (1.90)$$

(*b*) Use Cramer's Rule (1.90) to solve the systems (*i*) $\begin{aligned} x + 3y &= 13, \\ 4x + 2y &= 0, \end{aligned}$ (*ii*) $\begin{aligned} x - 2y &= 4, \\ 3x + 6y &= -2. \end{aligned}$

(*c*) Prove that the solution to $\begin{aligned} ax + by + cz &= p, \\ dx + ey + fz &= q, \\ gx + hy + jz &= r, \end{aligned}$ with $\Delta = \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} \neq 0$ is

$$x = \frac{1}{\Delta} \det \begin{pmatrix} p & b & c \\ q & e & f \\ r & h & j \end{pmatrix}, \qquad y = \frac{1}{\Delta} \det \begin{pmatrix} a & p & c \\ d & q & f \\ g & r & j \end{pmatrix}, \qquad z = \frac{1}{\Delta} \det \begin{pmatrix} a & b & p \\ d & e & q \\ g & h & r \end{pmatrix}. \qquad (1.91)$$

(*d*) Use Cramer's Rule (1.91) to solve (*i*) $\begin{aligned} x + 4y &= 3, \\ 4x + 2y + z &= 2, \\ -x + y - z &= 0, \end{aligned}$ (*ii*) $\begin{aligned} 3x + 2y - z &= 1, \\ x - 3y + 2z &= 2, \\ 2x - y + z &= 3. \end{aligned}$

(*e*) Can you see the pattern that will generalize to n equations in n unknowns?

Remark. Although elegant, Cramer's rule is not a very practical solution method.

◇ 1.9.21. (a) Show that if $D = \begin{pmatrix} A & \mathrm{O} \\ \mathrm{O} & B \end{pmatrix}$ is a block diagonal matrix, where A and B are square matrices, then $\det D = \det A \, \det B$. (*b*) Prove that the same holds for a block upper triangular matrix $\det \begin{pmatrix} A & C \\ \mathrm{O} & B \end{pmatrix} = \det A \, \det B$. (*c*) Use this method to compute the determinant of the following matrices:

(*i*) $\begin{pmatrix} 3 & 2 & -2 \\ 0 & 4 & -5 \\ 0 & 3 & 7 \end{pmatrix}$, (*ii*) $\begin{pmatrix} 1 & 2 & -2 & 5 \\ -3 & 1 & 0 & -5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 2 & -2 \end{pmatrix}$, (*iii*) $\begin{pmatrix} 1 & 2 & 0 & 4 \\ -3 & 1 & 4 & -1 \\ 0 & 3 & 1 & 8 \\ 0 & 0 & 0 & -3 \end{pmatrix}$, (*iv*) $\begin{pmatrix} 5 & -1 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 2 & 4 & 4 & -2 \\ 3 & -2 & 9 & -5 \end{pmatrix}$.