

PART B

Linear Algebra. Vector Calculus

CHAPTER 7 Linear Algebra: Matrices, Vectors, Determinants. Linear Systems

CHAPTER 8 Linear Algebra: Matrix Eigenvalue Problems

CHAPTER 9 Vector Differential Calculus. Grad, Div, Curl

CHAPTER 10 Vector Integral Calculus. Integral Theorems

Linear algebra in Chaps. 7 and 8 consists of the theory and application of vectors and matrices, mainly related to linear systems of equations, eigenvalue problems, and linear transformations.

Linear algebra is of growing importance in engineering research and teaching because it forms a foundation of numeric methods (see Chaps. 20–22), and its main instruments, matrices, can hold enormous amounts of data—think of a net of millions of telephone connections—in a form readily accessible by the computer.

Linear analysis in Chaps. 9 and 10, usually called **vector calculus**, extends *differentiation* of functions of one variable to functions of several variables—this includes the vector differential operations grad, div, and curl. And it generalizes integration to integrals over curves, surfaces, and solids, with transformations of these integrals into one another, by the basic theorems of Gauss, Green, and Stokes (Chap. 10).

Software suitable for linear algebra (Lapack, Maple, Mathematica, Matlab) can be found in the list at the opening of Part E of the book if needed.

Numeric linear algebra (Chap. 20) *can be studied directly after Chap. 7 or 8* because Chap. 20 is independent of the other chapters in Part E on numerics.

수학의 기초



CHAPTER 7

Linear Algebra: Matrices, Vectors, Determinants. Linear Systems

This is the first of two chapters on **linear algebra**, which concerns mainly systems of linear equations and linear transformations (to be discussed in this chapter) and eigenvalue problems (to follow in Chap. 8).

Systems of linear equations, briefly called **linear systems**, arise in electrical networks, mechanical frameworks, economic models, optimization problems, numerics for differential equations, as we shall see in Chaps. 21–23, and so on.

As main tools, linear algebra uses **matrices** (rectangular arrays of numbers or functions) and **vectors**. Calculations with matrices handle matrices as single objects, denote them by single letters, and calculate with them in a very compact form, almost as with numbers, so that matrix calculations constitute a powerful “mathematical shorthand”.

Calculations with matrices and vectors are defined and explained in Secs. 7.1–7.2. Sections 7.3–7.8 center around linear systems, with a thorough discussion of Gauss elimination, the role of rank, the existence and uniqueness problem for solutions (Sec. 7.5), and matrix inversion. This also includes determinants (Cramer’s rule) in Sec. 7.6 (for quick reference) and Sec. 7.7. Applications are considered throughout this chapter. The last section (Sec. 7.9) on vector spaces, inner product spaces, and linear transformations is more abstract. Eigenvalue problems follow in Chap. 8.

COMMENT. *Numeric linear algebra (Secs. 20.1–20.5) can be studied immediately after this chapter.*

Prerequisite: None.

Sections that may be omitted in a short course: 7.5, 7.9.

References and Answers to Problems: App. 1 Part B, and App. 2.

7.1 Matrices, Vectors: Addition and Scalar Multiplication

In this section and the next one we introduce the basic concepts and rules of matrix and vector algebra. The main application to **linear systems** (systems of linear equations) begins in Sec. 7.3.

A **matrix** is a rectangular array of numbers (or functions) enclosed in brackets. These numbers (or functions) are called the **entries** (or sometimes the *elements*) of the matrix. For example,

$$(1) \quad \begin{bmatrix} 0.3 & 1 & -5 \\ 0 & -0.2 & 16 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \begin{bmatrix} e^{-x} & 2x^2 \\ e^{6x} & 4x \end{bmatrix}, \quad [a_1 \ a_2 \ a_3], \quad \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}$$

수자,
2x^2
1/2

are matrices. The first matrix has two **rows** (horizontal lines of entries) and three **columns** (vertical lines). The second and third matrices are **square matrices**, that is, each has as many rows as columns (3 and 2, respectively). The entries of the second matrix have two indices giving the location of the entry. The first index is the number of the row and the second is the number of the column in which the entry stands. Thus, a_{23} (read *a two three*) is in Row 2 and Column 3, etc. This notation is standard, regardless of whether a matrix is square or not.

Matrices having just a single row or column are called vectors. Thus the fourth matrix in (1) has just one row and is called a **row vector**. The last matrix in (1) has just one column and is called a **column vector**.

We shall see that matrices are practical in various applications for storing and processing data. As a first illustration let us consider two simple but typical examples.

EXAMPLE 1 Linear Systems, a Major Application of Matrices

In a system of linear equations, briefly called a **linear system**, such as

$$\begin{aligned} 4x_1 + 6x_2 + 9x_3 &= 6 \\ 6x_1 &\quad - 2x_3 = 20 \\ 5x_1 - 8x_2 + x_3 &= 10 \end{aligned}$$

the coefficients of the **unknowns** x_1, x_2, x_3 are the entries of the **coefficient matrix**, call it **A**,

$$\mathbf{A} = \begin{bmatrix} 4 & 6 & 9 \\ 6 & 0 & -2 \\ 5 & -8 & 1 \end{bmatrix}, \quad \text{The matrix } \tilde{\mathbf{A}} = \begin{bmatrix} 4 & 6 & 9 & 6 \\ 6 & 0 & -2 & 20 \\ 5 & -8 & 1 & 10 \end{bmatrix}$$

is obtained by augmenting **A** by the right sides of the linear system and is called the **augmented matrix** of the system. In **A** the coefficients of the system are displayed in the pattern of the equations. That is, their position in **A** corresponds to that in the system when written as shown. The same is true for $\tilde{\mathbf{A}}$.

We shall see that the augmented matrix $\tilde{\mathbf{A}}$ contains all the information about the solutions of a system, so that we can solve a system just by calculations on its augmented matrix. We shall discuss this in great detail, beginning in Sec. 7.3. Meanwhile you may verify by substitution that the solution is $x_1 = 3, x_2 = \frac{1}{2}, x_3 = -1$.

The notation x_1, x_2, x_3 for the unknowns is practical but not essential; we could choose x, y, z or some other letters.

EXAMPLE 2 Sales Figures in Matrix Form

Sales figures for three products I, II, III in a store on Monday (M), Tuesday (T), ... may for each week be arranged in a matrix

$$A = \begin{matrix} & \begin{matrix} \text{M} & \text{T} & \text{W} & \text{Th} & \text{F} & \text{S} \end{matrix} \\ \begin{matrix} \text{I} \\ \text{II} \\ \text{III} \end{matrix} & \begin{bmatrix} 400 & 330 & 810 & 0 & 210 & 470 \\ 0 & 120 & 780 & 500 & 500 & 960 \\ 100 & 0 & 0 & 270 & 430 & 780 \end{bmatrix} \end{matrix}$$

9/27
prod

If the company has ten stores, we can set up ten such matrices, one for each store. Then by adding corresponding entries of these matrices we can get a matrix showing the total sales of each product on each day. Can you think of other data for which matrices are feasible? For instance, in transportation or storage problems? Or in recording phone calls, or in listing distances in a network of roads?

General Concepts and Notations

We shall denote matrices by capital boldface letters A, B, C, \dots , or by writing the general entry in brackets; thus $A = [a_{jk}]$, and so on. By an $m \times n$ **matrix** (read m by n matrix) we mean a matrix with m rows and n columns—rows come always first! $m \times n$ is called the **size** of the matrix. Thus an $m \times n$ matrix is of the form

$$(2) \quad A = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

The matrices in (1) are of sizes 2×3 , 3×3 , 2×2 , 1×3 , and 2×1 , respectively.

Each entry in (2) has two subscripts. The first is the *row number* and the second is the *column number*. Thus a_{21} is the entry in Row 2 and Column 1.

If $m = n$, we call A an $n \times n$ **square matrix**. Then its diagonal containing the entries $a_{11}, a_{22}, \dots, a_{nn}$ is called the **main diagonal** of A . Thus the main diagonals of the two square matrices in (1) are a_{11}, a_{22}, a_{33} and $e^{-x}, 4x$, respectively.

Square matrices are particularly important, as we shall see. A matrix that is not square is called a **rectangular matrix**.

Vectors

A **vector** is a matrix with only one row or column. Its entries are called the **components** of the vector. We shall denote vectors by lowercase boldface letters a, b, \dots or by its general component in brackets, $a = [a_j]$, and so on. Our special vectors in (1) suggest that a (general) **row vector** is of the form

$$a = [a_1 \ a_2 \ \cdots \ a_n]. \quad \text{For instance,} \quad a = [-2 \ 5 \ 0.8 \ 0 \ 1].$$

A **column vector** is of the form

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{For instance,} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ -7 \end{bmatrix}.$$

Matrix Addition and Scalar Multiplication

What makes matrices and vectors really useful and particularly suitable for computers is the fact that we can calculate with them almost as easily as with numbers. Indeed, we now introduce rules for addition and for scalar multiplication (multiplication by numbers) that were suggested by practical applications. (Multiplication of matrices by matrices follows in the next section.) We first need the concept of equality.

DEFINITION

Equality of Matrices

Two matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ are **equal**, written $\mathbf{A} = \mathbf{B}$, if and only if they have the same size and the corresponding entries are equal, that is, $a_{11} = b_{11}$, $a_{12} = b_{12}$, and so on. Matrices that are not equal are called **different**. Thus, matrices of different sizes are always different.

EXAMPLE 3

Equality of Matrices

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 4 & 0 \\ 3 & -1 \end{bmatrix}.$$

Then

$$\mathbf{A} = \mathbf{B} \quad \text{if and only if} \quad \begin{array}{ll} a_{11} = 4, & a_{12} = 0, \\ a_{21} = 3, & a_{22} = -1. \end{array}$$

The following matrices are all different. Explain!

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

Size of matrix

DEFINITION

Addition of Matrices

The **sum** of two matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ of the same size is written $\mathbf{A} + \mathbf{B}$ and has the entries $a_{jk} + b_{jk}$ obtained by adding the corresponding entries of \mathbf{A} and \mathbf{B} . Matrices of different sizes cannot be added.

As a special case, the **sum** $\mathbf{a} + \mathbf{b}$ of two row vectors or two column vectors, which must have the same number of components, is obtained by adding the corresponding components.

EXAMPLE 4 Addition of Matrices and Vectors

If $\mathbf{A} = \begin{bmatrix} -4 & 6 & 3 \\ 0 & 1 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 5 & -1 & 0 \\ 3 & 1 & 0 \end{bmatrix}$, then $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1 & 5 & 3 \\ 3 & 2 & 2 \end{bmatrix}$.

\mathbf{A} in Example 3 and our present \mathbf{A} cannot be added. If $\mathbf{a} = [5 \ 7 \ 2]$ and $\mathbf{b} = [-6 \ 2 \ 0]$, then $\mathbf{a} + \mathbf{b} = [-1 \ 9 \ 2]$.

An application of matrix addition was suggested in Example 2. Many others will follow. ■

DEFINITION**Scalar Multiplication (Multiplication by a Number)**

The **product** of any $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ and any **scalar** c (number c) is written $c\mathbf{A}$ and is the $m \times n$ matrix $c\mathbf{A} = [ca_{jk}]$ obtained by multiplying each entry of \mathbf{A} by c .

Here $(-1)\mathbf{A}$ is simply written $-\mathbf{A}$ and is called the **negative** of \mathbf{A} . Similarly, $(-k)\mathbf{A}$ is written $-k\mathbf{A}$. Also, $\mathbf{A} + (-\mathbf{B})$ is written $\mathbf{A} - \mathbf{B}$ and is called the **difference** of \mathbf{A} and \mathbf{B} (which must have the same size!).

EXAMPLE 5 Scalar Multiplication

If $\mathbf{A} = \begin{bmatrix} 2.7 & -1.8 \\ 0 & 0.9 \\ 9.0 & -4.5 \end{bmatrix}$, then $-\mathbf{A} = \begin{bmatrix} -2.7 & 1.8 \\ 0 & -0.9 \\ -9.0 & 4.5 \end{bmatrix}$, $\frac{10}{9}\mathbf{A} = \begin{bmatrix} 3 & -2 \\ 0 & 1 \\ 10 & -5 \end{bmatrix}$, $0\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

If a matrix \mathbf{B} shows the distances between some cities in miles, $1.609\mathbf{B}$ gives these distances in kilometers. ■

Rules for Matrix Addition and Scalar Multiplication. From the familiar laws for the addition of numbers we obtain similar laws for the addition of matrices of the same size $m \times n$, namely,

(3) (a) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
 (b) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$ (written $\mathbf{A} + \mathbf{B} + \mathbf{C}$)
 (c) $\mathbf{A} + \mathbf{0} = \mathbf{A}$
 (d) $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$.

Here $\mathbf{0}$ denotes the **zero matrix** (of size $m \times n$), that is, the $m \times n$ matrix with all entries zero. (The last matrix in Example 5 is a zero matrix.)

Hence matrix addition is *commutative* and *associative* [by (3a) and (3b)].

Similarly, for scalar multiplication we obtain the rules

(4) (a) $c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}$
 (b) $(c + k)\mathbf{A} = c\mathbf{A} + k\mathbf{A}$
 (c) $c(k\mathbf{A}) = (ck)\mathbf{A}$ (written $ck\mathbf{A}$)
 (d) $1\mathbf{A} = \mathbf{A}$.

H.W 2, 8

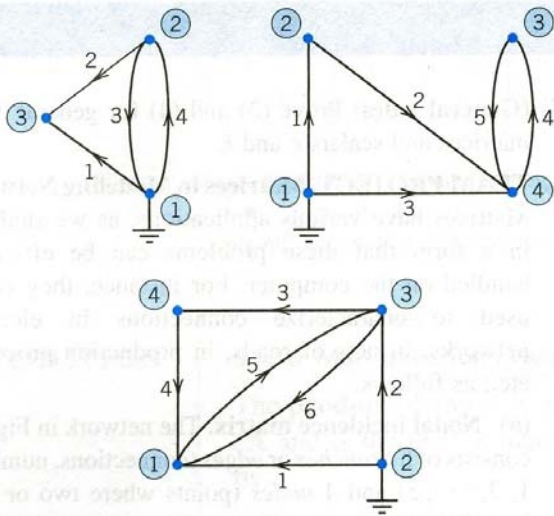


Fig. 153. Networks in Team Project 16(b)

(c) Graph the three networks corresponding to the nodal incidence matrices

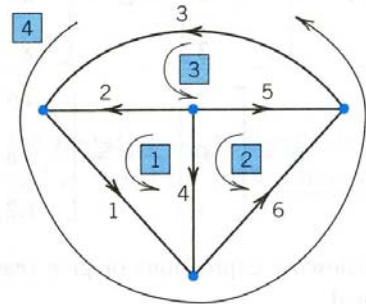
$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & -1 \end{bmatrix}.$$

(d) **Mesh incidence matrix.** A network can also be characterized by the *mesh incidence matrix* $\mathbf{M} = [m_{jk}]$, where

$$m_{jk} = \begin{cases} +1 & \text{if branch } k \text{ is in mesh } \boxed{j} \\ & \text{and has the same orientation} \\ -1 & \text{if branch } k \text{ is in mesh } \boxed{j} \\ & \text{and has the opposite orientation} \\ 0 & \text{if branch } k \text{ is not in mesh } \boxed{j} \end{cases}$$

and a mesh is a loop with no branch in its interior (or in its exterior). Here, the meshes are numbered and directed (oriented) in an arbitrary fashion. Show that in Fig. 154 the matrix \mathbf{M} corresponds to the given figure, where Row 1 corresponds to mesh 1, etc.



$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Fig. 154. Network and matrix \mathbf{M} in Team Project 16(d)

(e) Number the nodes in Fig. 154 from left to right 1, 2, 3 and the low node by 4. Find the corresponding nodal incidence matrix.

7.2 Matrix Multiplication

Matrix multiplication means multiplication of matrices by matrices. This is the last algebraic operation to be defined (except for *transposition*, which is of lesser importance). Now matrices are *added* by *adding* corresponding entries. In *multiplication*, do we *multiply* corresponding entries? The answer is no. Why not? Such an operation would not be of much use in applications. The standard definition of multiplication looks artificial, but will be fully motivated later in this section by the use of matrices in “linear transformations,” by which this multiplication is suggested.

DEFINITION

Multiplication of a Matrix by a Matrix

The **product** $C = AB$ (in this order) of an $m \times n$ matrix $A = [a_{jk}]$ times an $r \times p$ matrix $B = [b_{jk}]$ is defined if and only if $r = n$ and is then the $m \times p$ matrix $C = [c_{jk}]$ with entries

$$(1) \quad c_{jk} = \sum_{l=1}^n a_{jl}b_{lk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} \quad \begin{matrix} j = 1, \dots, m \\ k = 1, \dots, p. \end{matrix}$$

The condition $r = n$ means that the second factor, B , must have as many rows as the first factor has columns, namely n . As a diagram of sizes (denoted as shown):

$$\begin{matrix} A & B & = & C \\ [m \times n] & [n \times r] & = & [m \times r]. \end{matrix}$$

c_{jk} in (1) is obtained by multiplying each entry in the j th row of A by the corresponding entry in the k th column of B and then adding these n products. For instance, $c_{21} = a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1}$, and so on. One calls this briefly a "multiplication of rows into columns." See the illustration in Fig. 155, where $n = 3$.

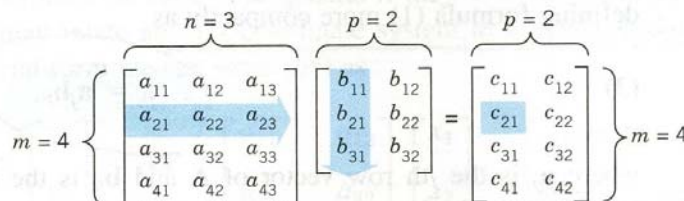


Fig. 155. Notations in a product $AB = C$

EXAMPLE 1 Matrix Multiplication

$$AB = \begin{bmatrix} 3 & 5 & -1 \\ 4 & 0 & 2 \\ -6 & -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 5 & 0 \\ 9 & -4 \end{bmatrix} = \begin{bmatrix} 22 & -2 & 43 & 42 \\ 26 & -16 & 14 & 6 \\ -9 & 4 & -37 & -28 \end{bmatrix}$$

Here $c_{11} = 3 \cdot 2 + 5 \cdot 5 + (-1) \cdot 9 = 22$, and so on. The entry in the box is $c_{23} = 4 \cdot 3 + 0 \cdot 7 + 2 \cdot 1 = 14$. The product BA is not defined.

EXAMPLE 2 Multiplication of a Matrix and a Vector

$$\begin{bmatrix} 4 & 2 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \cdot 3 + 2 \cdot 5 \\ 1 \cdot 3 + 8 \cdot 5 \end{bmatrix} = \begin{bmatrix} 22 \\ 43 \end{bmatrix}$$

whereas

$$\begin{bmatrix} 3 \\ 5 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 1 & 8 \end{bmatrix}$$

is undefined.

EXAMPLE 3 Products of Row and Column Vectors

$$\begin{bmatrix} 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = [19],$$

1 over 1 and

$$\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \begin{bmatrix} 3 & 6 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 1 \\ 6 & 12 & 2 \\ 12 & 24 & 4 \end{bmatrix}$$

matrix

EXAMPLE 4 CAUTION! Matrix Multiplication Is Not Commutative, $AB \neq BA$ in General

This is illustrated by Examples 1 and 2, where one of the two products is not even defined, and by Example 3, where the two products have different sizes. But it also holds for square matrices. For instance,

$$\begin{bmatrix} 1 & 1 \\ 100 & 100 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 100 & 100 \end{bmatrix} = \begin{bmatrix} 99 & 99 \\ -99 & -99 \end{bmatrix}.$$

It is interesting that this also shows that $\mathbf{AB} = \mathbf{0}$ does *not* necessarily imply $\mathbf{BA} = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$. We shall discuss this further in Sec. 7.8, along with reasons when this happens. ■

Our examples show that the *order of factors* in matrix products *must always be observed very carefully*. Otherwise matrix multiplication satisfies rules similar to those for numbers, namely,

- (a) $(k\mathbf{A})\mathbf{B} = k(\mathbf{AB}) = \mathbf{A}(k\mathbf{B})$ written $k\mathbf{AB}$ or \mathbf{AkB}
 (b) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ written \mathbf{ABC}
 (2) (c) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
 (d) $\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$

provided \mathbf{A} , \mathbf{B} , and \mathbf{C} are such that the expressions on the left are defined; here, k is any scalar. (2b) is called the **associative law**. (2c) and (2d) are called the **distributive laws**.

Since matrix multiplication is a multiplication of rows into columns, we can write the defining formula (1) more compactly as

$$(3) \quad c_{jk} = \mathbf{a}_j \mathbf{b}_k, \quad j = 1, \dots, m; \quad k = 1, \dots, p,$$

where \mathbf{a}_j is the j th row vector of \mathbf{A} and \mathbf{b}_k is the k th column vector of \mathbf{B} , so that in agreement with (1),

$$\mathbf{a}_j \mathbf{b}_k = [a_{j1} \quad a_{j2} \quad \cdots \quad a_{jn}] \begin{bmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{bmatrix} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk}.$$

EXAMPLE 5 Product in Terms of Row and Column Vectors

If $\mathbf{A} = [a_{jk}]$ is of size 3×3 and $\mathbf{B} = [b_{jk}]$ is of size 3×4 , then

$$(4) \quad \mathbf{AB} = \begin{bmatrix} \mathbf{a}_1 \mathbf{b}_1 & \mathbf{a}_1 \mathbf{b}_2 & \mathbf{a}_1 \mathbf{b}_3 & \mathbf{a}_1 \mathbf{b}_4 \\ \mathbf{a}_2 \mathbf{b}_1 & \mathbf{a}_2 \mathbf{b}_2 & \mathbf{a}_2 \mathbf{b}_3 & \mathbf{a}_2 \mathbf{b}_4 \\ \mathbf{a}_3 \mathbf{b}_1 & \mathbf{a}_3 \mathbf{b}_2 & \mathbf{a}_3 \mathbf{b}_3 & \mathbf{a}_3 \mathbf{b}_4 \end{bmatrix}.$$

Taking $\mathbf{a}_1 = [3 \quad 5 \quad -1]$, $\mathbf{a}_2 = [4 \quad 0 \quad 2]$, etc., verify (4) for the product in Example 1. ■

Parallel processing of products on the computer is facilitated by a variant of (3) for computing $\mathbf{C} = \mathbf{AB}$, which is used by standard algorithms (such as in Lapack). In this method, \mathbf{A} is used as given, \mathbf{B} is taken in terms of its column vectors, and the product is computed columnwise; thus,

$$(5) \quad \mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p] = [\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \cdots \quad \mathbf{Ab}_p].$$

Columns of **B** are then assigned to different processors (individually or several to each processor), which simultaneously compute the columns of the product matrix **Ab**₁, **Ab**₂, etc.

EXAMPLE 6 Computing Products Columnwise by (5)

To obtain

$$\mathbf{AB} = \begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 & 7 \\ -1 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 11 & 4 & 34 \\ -17 & 8 & -23 \end{bmatrix}$$

from (5), calculate the columns

$$\begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 11 \\ -17 \end{bmatrix}, \quad \begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \quad \begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 34 \\ -23 \end{bmatrix}$$

of **AB** and then write them as a single matrix, as shown in the first formula on the right. ■

Motivation of Multiplication by Linear Transformations

Let us now motivate the “unnatural” matrix multiplication by its use in **linear transformations**. For $n = 2$ variables these transformations are of the form

$$(6^*) \quad \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned}$$

and suffice to explain the idea. (For general n they will be discussed in Sec. 7.9.) For instance, (6*) may relate an x_1x_2 -coordinate system to a y_1y_2 -coordinate system in the plane. In vectorial form we can write (6*) as

$$(6) \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{Ax} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}.$$

Now suppose further that the x_1x_2 -system is related to a w_1w_2 -system by another linear transformation, say,

$$(7) \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{Bw} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} b_{11}w_1 + b_{12}w_2 \\ b_{21}w_1 + b_{22}w_2 \end{bmatrix}.$$

Then the y_1y_2 -system is related to the w_1w_2 -system indirectly via the x_1x_2 -system, and we wish to express this relation directly. Substitution will show that this direct relation is a linear transformation, too, say,

$$(8) \quad \mathbf{y} = \mathbf{Cw} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} c_{11}w_1 + c_{12}w_2 \\ c_{21}w_1 + c_{22}w_2 \end{bmatrix}.$$

Indeed, substituting (7) into (6), we obtain

$$\begin{aligned} y_1 &= a_{11}(b_{11}w_1 + b_{12}w_2) + a_{12}(b_{21}w_1 + b_{22}w_2) \\ &= (a_{11}b_{11} + a_{12}b_{21})w_1 + (a_{11}b_{12} + a_{12}b_{22})w_2 \\ y_2 &= a_{21}(b_{11}w_1 + b_{12}w_2) + a_{22}(b_{21}w_1 + b_{22}w_2) \\ &= (a_{21}b_{11} + a_{22}b_{21})w_1 + (a_{21}b_{12} + a_{22}b_{22})w_2. \end{aligned}$$

Comparing this with (8), we see that

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} & c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} & c_{22} &= a_{21}b_{12} + a_{22}b_{22}. \end{aligned}$$

This proves that $\mathbf{C} = \mathbf{AB}$ with the product defined as in (1). For larger matrix sizes the idea and result are exactly the same. Only the number of variables changes. We then have m variables y and n variables x and p variables w . The matrices \mathbf{A} , \mathbf{B} , and $\mathbf{C} = \mathbf{AB}$ then have sizes $m \times n$, $n \times p$, and $m \times p$, respectively. And the requirement that \mathbf{C} be the product \mathbf{AB} leads to formula (1) in its general form. This motivates matrix multiplication completely.

Transposition

Transposition provides a transition from row vectors to column vectors and conversely. More generally, it gives us a choice to work either with a matrix or with its transpose, whatever will be more practical in a specific situation.

DEFINITION

Transposition of Matrices and Vectors

The transpose of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the $n \times m$ matrix \mathbf{A}^T (read *A transpose*) that has the first row of \mathbf{A} as its first column, the second row of \mathbf{A} as its second column, and so on. Thus the transpose of \mathbf{A} in (2) is $\mathbf{A}^T = [a_{kj}]$, written out

$$(9) \quad \mathbf{A}^T = [a_{kj}] = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdot & \cdot & \cdots & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

As a special case, transposition converts row vectors to column vectors and conversely.

EXAMPLE 7 Transposition of Matrices and Vectors

If $\mathbf{A} = \begin{bmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{bmatrix}$, then $\mathbf{A}^T = \begin{bmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{bmatrix}$.

A little more compactly, we can write

$$\begin{aligned} \begin{bmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{bmatrix}^T &= \begin{bmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{bmatrix}, & \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}^T &= \begin{bmatrix} 3 & 8 \\ 0 & -1 \end{bmatrix} \\ [6 & 2 & 3]^T &= \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix}, & \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix}^T &= [6 & 2 & 3]. \end{aligned}$$

Note that for a square matrix, the transpose is obtained by interchanging entries that are symmetrically positioned with respect to the main diagonal, e.g., a_{12} and a_{21} , and so on. ■

Rules for transposition are

$$(10) \quad \begin{aligned} (a) \quad & (A^T)^T = A \\ (b) \quad & (A + B)^T = A^T + B^T \\ (c) \quad & (cA)^T = cA^T \\ (d) \quad & (AB)^T = B^T A^T. \end{aligned}$$

CAUTION! Note that in (10d) the transposed matrices are *in reversed order*. We leave the proofs to the student. (See Prob. 22.)

Special Matrices

Certain kinds of matrices will occur quite frequently in our work, and we now list the most important ones of them.

Symmetric and Skew-Symmetric Matrices. Transposition gives rise to two useful classes of matrices, as follows. *Symmetric matrices* and *skew-symmetric matrices* are square matrices whose transpose equals the matrix itself or minus the matrix, respectively:

$$(11) \quad A^T = A \quad (\text{thus } a_{kj} = a_{jk}), \quad A^T = -A \quad (\text{thus } a_{kj} = -a_{jk}, \text{ hence } a_{jj} = 0).$$

Symmetric Matrix

Skew-Symmetric Matrix

EXAMPLE 8 Symmetric and Skew-Symmetric Matrices

$$A = \begin{bmatrix} 20 & 120 & 200 \\ 120 & 10 & 150 \\ 200 & 150 & 30 \end{bmatrix} \quad \text{is symmetric, and} \quad B = \begin{bmatrix} 0 & 1 & -3 \\ -1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix} \quad \text{is skew-symmetric.}$$

For instance, if a company has three building supply centers C_1, C_2, C_3 , then A could show costs, say, a_{jj} for handling 1000 bags of cement on center C_j , and a_{jk} ($j \neq k$) the cost of shipping 1000 bags from C_j to C_k . Clearly, $a_{jk} = a_{kj}$ because shipping in the opposite direction will usually cost the same.

Symmetric matrices have several general properties which make them important. This will be seen as we proceed.

Triangular Matrices. Upper triangular matrices are square matrices that can have nonzero entries only on and *above* the main diagonal, whereas any entry below the diagonal must be zero. Similarly, lower triangular matrices can have nonzero entries only on and *below* the main diagonal. Any entry on the main diagonal of a triangular matrix may be zero or not.

EXAMPLE 9 Upper and Lower Triangular Matrices

$$\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 6 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 \\ 8 & -1 & 0 \\ 7 & 6 & 8 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 0 & 0 \\ 9 & -3 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 9 & 3 & 6 \end{bmatrix}$$

Upper triangular

Lower triangular

Diagonal Matrices. These are square matrices that can have nonzero entries only on the main diagonal. Any entry above or below the main diagonal must be zero.

If all the diagonal entries of a diagonal matrix S are equal, say, c , we call S a scalar matrix because multiplication of any square matrix A of the same size by S has the same effect as the multiplication by a scalar, that is,

$$(12) \quad \underline{AS = SA = cA.}$$

In particular, a scalar matrix whose entries on the main diagonal are all 1 is called a unit matrix (or identity matrix) and is denoted by I_n or simply by I . For I , formula (12) becomes

$$(13) \quad \underline{AI = IA = A.}$$

EXAMPLE 10 Diagonal Matrix D. Scalar Matrix S. Unit Matrix I

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Applications of Matrix Multiplication

Matrix multiplication will play a crucial role in connection with linear systems of equations, beginning in the next section. For the time being we mention some other simple applications that need no lengthy explanations.

EXAMPLE 11 Computer Production. Matrix Times Matrix

Supercomp Ltd produces two computer models PC1086 and PC1186. The matrix A shows the cost per computer (in thousands of dollars) and B the production figures for the year 2005 (in multiples of 10000 units.) Find a matrix C that shows the shareholders the cost per quarter (in millions of dollars) for raw material, labor, and miscellaneous.

	PC1086	PC1186		Quarter		
			1	2	3	4
$A =$	1.2	1.6	Raw Components			
	0.3	0.4	Labor			
	0.5	0.6	Miscellaneous			

$B =$	3	8	6	9	PC1086
	6	2	4	3	PC1186

Solution.

		Quarter				
		1	2	3	4	
$C = AB =$	13.2	12.8	13.6	15.6	Raw Components	
	3.3	3.2	3.4	3.9	Labor	
	5.1	5.2	5.4	6.3	Miscellaneous	

Since cost is given in multiples of \$1000 and production in multiples of 10 000 units, the entries of C are multiples of \$10 millions; thus $c_{11} = 13.2$ means \$132 million, etc.

EXAMPLE 12 Weight Watching. Matrix Times Vector

Suppose that in a weight-watching program, a person of 185 lb burns 350 cal/hr in walking (3 mph), 500 in bicycling (13 mph) and 950 in jogging (5.5 mph). Bill, weighing 185 lb, plans to exercise according to the matrix shown. Verify the calculations (W = Walking, B = Bicycling, J = Jogging).

$$\begin{array}{l}
 \text{MON} \\
 \text{WED} \\
 \text{FRI} \\
 \text{SAT}
 \end{array}
 \begin{array}{c}
 \begin{matrix} W & B & J \\ \swarrow & \searrow & \swarrow \\ \text{hr} & & \text{cal/hr} \end{matrix} \\
 \begin{bmatrix} 1.0 & 0 & 0.5 \\ 1.0 & 1.0 & 0.5 \\ 1.5 & 0 & 0.5 \\ 2.0 & 1.5 & 1.0 \end{bmatrix}
 \end{array}
 \begin{array}{c}
 \begin{bmatrix} 350 \\ 500 \\ 950 \end{bmatrix}
 \end{array}
 =
 \begin{array}{c}
 \begin{bmatrix} 825 \\ 1325 \\ 1000 \\ 2400 \end{bmatrix}
 \end{array}
 \begin{array}{l}
 \text{MON} \\
 \text{WED} \\
 \text{FRI} \\
 \text{SAT}
 \end{array}$$

EXAMPLE 13 Markov Process. Powers of a Matrix. Stochastic Matrix

Suppose that the 2004 state of land use in a city of 60 mi² of built-up area is

C: Commercially Used 25% I: Industrially Used 20% R: Residentially Used 55%.

Find the states in 2009, 2014, and 2019, assuming that the transition probabilities for 5-year intervals are given by the matrix **A** and remain practically the same over the time considered.

$$\mathbf{A} = \begin{array}{c}
 \begin{matrix} \text{From C} & \text{From I} & \text{From R} \\ \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix} \\ \text{To C} \\ \text{To I} \\ \text{To R} \end{matrix}
 \end{array}$$

A is a **stochastic matrix**, that is, a square matrix with all entries nonnegative and all column sums equal to 1. Our example concerns a **Markov process**¹, that is, a process for which the probability of entering a certain state depends only on the last state occupied (and the matrix **A**), not on any earlier state.

Solution. From the matrix **A** and the 2004 state we can compute the 2009 state,

$$\begin{array}{l}
 \text{C} \\
 \text{I} \\
 \text{R}
 \end{array}
 \begin{array}{c}
 \begin{matrix} C \rightarrow C & I \rightarrow C & R \rightarrow C \\ \begin{bmatrix} 0.7 \cdot 25 + 0.1 \cdot 20 + 0 \cdot 55 \\ 0.2 \cdot 25 + 0.9 \cdot 20 + 0.2 \cdot 55 \\ 0.1 \cdot 25 + 0 \cdot 20 + 0.8 \cdot 55 \end{bmatrix} \\ = \\ \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix} \\ \begin{bmatrix} 25 \\ 20 \\ 55 \end{bmatrix} \\ = \\ \begin{bmatrix} 19.5 \\ 34.0 \\ 46.5 \end{bmatrix} \end{matrix}
 \end{array}$$

To explain: The 2009 figure for C equals 25% times the probability 0.7 that C goes into C, plus 20% times the probability 0.1 that I goes into C, plus 55% times the probability 0 that R goes into C. Together,

$$25 \cdot 0.7 + 20 \cdot 0.1 + 55 \cdot 0 = 19.5 \text{ [\%]}. \quad \text{Also} \quad 25 \cdot 0.2 + 20 \cdot 0.9 + 55 \cdot 0.2 = 34 \text{ [\%]}.$$

Similarly, the new R is 46.5%. We see that the 2009 state vector is the column vector

$$\mathbf{y} = [19.5 \quad 34.0 \quad 46.5]^T = \mathbf{Ax} = \mathbf{A} [25 \quad 20 \quad 55]^T$$

where the column vector $\mathbf{x} = [25 \quad 20 \quad 55]^T$ is the given 2004 state vector. Note that the sum of the entries of **y** is 100 [%]. Similarly, you may verify that for 2014 and 2019 we get the state vectors

$$\begin{array}{l}
 \mathbf{z} = \mathbf{Ay} = \mathbf{A(Ax)} = \mathbf{A}^2\mathbf{x} = [17.05 \quad 43.80 \quad 39.15]^T \rightarrow 2014 \\
 \mathbf{u} = \mathbf{Az} = \mathbf{A}^2\mathbf{y} = \mathbf{A}^3\mathbf{x} = [16.315 \quad 50.660 \quad 33.025]^T \rightarrow 2019
 \end{array}$$

¹ANDREI ANDREJEVITCH MARKOV (1856–1922), Russian mathematician, known for his work in probability theory.

Answer. In 2009 the commercial area will be 19.5% (11.7 mi²), the industrial 34% (20.4 mi²) and the residential 46.5% (27.9 mi²). For 2014 the corresponding figures are 17.05%, 43.80%, 39.15%. For 2019 they are 16.315%, 50.660%, 33.025%. (In Sec. 8.2 we shall see what happens in the limit, assuming that those probabilities remain the same. In the meantime, can you experiment or guess?)

5/20/14, 22

PROBLEM SET 7.2

1-14 MULTIPLICATION, ADDITION, AND TRANSPOSITION OF MATRICES AND VECTORS

Let

$$A = \begin{bmatrix} 6 & -2 & -2 \\ 10 & -3 & 1 \\ -10 & 5 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 9 & 4 & -4 \\ 4 & 7 & 0 \\ -4 & 0 & 11 \end{bmatrix},$$

$$C = \begin{bmatrix} 3 & 1 \\ 0 & -2 \\ 4 & 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{b} = [3 \quad 0 \quad 8].$$

Calculate the following products and sums or give reasons why they are not defined. (Show all intermediate results.)

- $A\mathbf{a}$, $A\mathbf{b}$, $A\mathbf{b}^T$, $A\mathbf{B}$
- $A\mathbf{b}^T + B\mathbf{b}^T$, $(A + B)\mathbf{b}^T$, $\mathbf{b}A$, $B - B^T$
- AB , BA , AA^T , $A^T A$
- A^2 , B^2 , $(A^T)^2$, $(A^2)^T$
- $\mathbf{a}^T A$, $\mathbf{b}A$, $5B(3\mathbf{a} + 2\mathbf{b}^T)$, $15B\mathbf{a} + 10B\mathbf{b}^T$
- $A^T \mathbf{b}$, $\mathbf{b}^T B$, $(3A - 2B)^T \mathbf{a}$, $\mathbf{a}^T (3A - 2B)$
- $\mathbf{a}\mathbf{b}$, $\mathbf{b}\mathbf{a}$, $(\mathbf{a}\mathbf{b})A$, $\mathbf{a}(\mathbf{b}A)$
- $\mathbf{a}\mathbf{b} - \mathbf{b}\mathbf{a}$, $-(4\mathbf{b})(7\mathbf{a})$, $-28\mathbf{b}\mathbf{a}$, $5\mathbf{a}\mathbf{b}B$
- $(A + B)^2$, $A^2 + AB + BA + B^2$, $A^2 + 2AB + B^2$
- $(A + B)(A - B)$, $A^2 - AB + BA - B^2$, $A^2 - B^2$
- $A^2 B$, A^3 , $(AB)^2$, $A^2 B^2$
- B^3 , BC , $(BC)^2$, $(BC)(BC)^T$
- $\mathbf{a}^T A \mathbf{a}$, $\mathbf{a}^T (A + A^T) \mathbf{a}$, $\mathbf{b} B \mathbf{b}^T$, $\mathbf{b} (B - B^T) \mathbf{b}^T$
- $\mathbf{a}^T C C^T \mathbf{a}$, $\mathbf{a}^T C^2 \mathbf{a}$, $\mathbf{b} C^T C \mathbf{b}^T$, $\mathbf{b} C C^T \mathbf{b}^T$
- (General rules) Prove (2) for 2×2 matrices $A = [a_{jk}]$, $B = [b_{jk}]$, $C = [c_{jk}]$ and a general scalar.
- (Commutativity) Find all 2×2 matrices $A = [a_{jk}]$ that commute with $B = [b_{jk}]$, where $b_{jk} = j + k$.
- (Product) Write AB in Probs. 1-14 in terms of row and column vectors.
- (Product) Calculate AB in Prob. 1 columnwise. (See Example 6.)
- TEAM PROJECT. **Symmetric and Skew-Symmetric Matrices.** These matrices occur quite frequently in applications, so it is worthwhile to study some of their most important properties.
 - Verify the claims in (11) that $a_{kj} = a_{jk}$ for a symmetric matrix, and $a_{kj} = -a_{jk}$ for a skew-symmetric matrix. Give examples.

(b) Show that for every square matrix C the matrix $C + C^T$ is symmetric and $C - C^T$ is skew-symmetric. Write C in the form $C = S + T$, where S is symmetric and T is skew-symmetric and find S and T in terms of C . Represent A and B in Probs. 1-14 in this form.

(c) A linear combination of matrices A, B, C, \dots, M of the same size is an expression of the form

$$(14) \quad aA + bB + cC + \dots + mM,$$

where a, \dots, m are any scalars. Show that if these matrices are square and symmetric, so is (14); similarly, if they are skew-symmetric, so is (14).

(d) Show that AB with symmetric A and B is symmetric if and only if A and B commute, that is, $AB = BA$.

(e) Under what condition is the product of skew-symmetric matrices skew-symmetric?

20. (Idempotent and nilpotent matrices) By definition, A is idempotent if $A^2 = A$, and B is nilpotent if $B^m = \mathbf{0}$ for some positive integer m . Give examples (different from $\mathbf{0}$ or I). Also give examples such that $A^2 = I$ (the unit matrix).

21. (Triangular matrices) Let U_1, U_2 be upper triangular and L_1, L_2 lower triangular. Which of the following are triangular? Give examples. How can you save half of your work by transposition?

$$U_1 + U_2, U_1 U_2, U_1^2, U_1 + L_1, U_1 L_1, L_1 + L_2,$$

$$L_1 L_2, L_1^2$$

22. (Transposition of products) Prove (10a)-(10c). Illustrate the basic formula (10d) by examples of your own. Then prove it.

APPLICATIONS

23. (Markov process) If the transition matrix A has the entries $a_{11} = 0.5$, $a_{12} = 0.3$, $a_{21} = 0.5$, $a_{22} = 0.7$ and the initial state is $[1 \quad 1]^T$, what will the next three states be?

24. (Concert subscription) In a community of 300 000 adults, subscribers to a concert series tend to renew their subscription with probability 90% and persons presently not subscribing will subscribe for the next season with probability 0.1%. If the present number of subscribers is 2000, can one predict an increase, decrease, or no change over each of the next three seasons?

25. **CAS Experiment. Markov Process.** Write a program for a Markov process. Use it to calculate further steps in Example 13 of the text. Experiment with other stochastic 3×3 matrices, also using different starting values.
26. **(Production)** In a production process, let N mean “no trouble” and T “trouble.” Let the transition probabilities from one day to the next be 0.9 for $N \rightarrow N$, hence 0.1 for $N \rightarrow T$, and 0.5 for $T \rightarrow N$, hence 0.5 for $T \rightarrow T$. If today there is no trouble, what is the probability of N two days after today? Three days after today?
27. **(Profit vector)** Two factory outlets F_1 and F_2 in New York and Los Angeles sell sofas (S), chairs (C), and tables (T) with a profit of \$110, \$45, and \$80, respectively. Let the sales in a certain week be given by the matrix

$$A = \begin{matrix} & \begin{matrix} S & C & T \end{matrix} \\ \begin{matrix} F_1 \\ F_2 \end{matrix} & \begin{bmatrix} 600 & 400 & 100 \\ 300 & 820 & 205 \end{bmatrix} \end{matrix}$$

Introduce a “profit vector” \mathbf{p} such that the components of $\mathbf{v} = A\mathbf{p}$ give the total profits of F_1 and F_2 .

28. **TEAM PROJECT. Special Linear Transformations.** **Rotations** have various applications. We show in this project how they can be handled by matrices.
- (a) **Rotation in the plane.** Show that the linear transformation $\mathbf{y} = A\mathbf{x}$ with matrix

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

is a counterclockwise rotation of the Cartesian x_1x_2 -coordinate system in the plane about the origin, where θ is the angle of rotation.

- (b) **Rotation through $n\theta$.** Show that in (a)

$$A^n = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}.$$

Is this plausible? Explain this in words.

- (c) **Addition formulas for cosine and sine.** By geometry we should have

$$\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} = \begin{bmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{bmatrix}.$$

Derive from this the addition formulas (6) in App. A3.1.

- (d) **Computer graphics.** To visualize a three-dimensional object with plane faces (e.g., a cube), we may store the position vectors of the vertices with respect to a suitable $x_1x_2x_3$ -coordinate system (and a list of the connecting edges) and then obtain a two-dimensional image on a video screen by projecting the object onto a coordinate plane, for instance, onto the x_1x_2 -plane by setting $x_3 = 0$. To change the appearance of the image, we can impose a linear transformation on the position vectors stored. Show that a diagonal matrix \mathbf{D} with main diagonal entries 3, 1, $\frac{1}{2}$ gives from an $\mathbf{x} = [x_j]$ the new position vector $\mathbf{y} = \mathbf{D}\mathbf{x}$, where $y_1 = 3x_1$ (stretch in the x_1 -direction by a factor 3), $y_2 = x_2$ (unchanged), $y_3 = \frac{1}{2}x_3$ (contraction in the x_3 -direction). What effect would a scalar matrix have?

- (e) **Rotations in space.** Explain $\mathbf{y} = A\mathbf{x}$ geometrically when A is one of the three matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix},$$

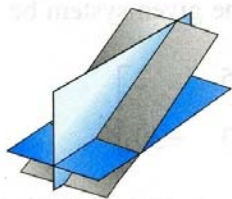
$$\begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix}, \quad \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

What effect would these transformations have in situations such as that described in (d)?

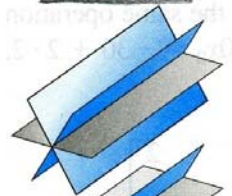
7.3 Linear Systems of Equations. Gauss Elimination

The most important use of matrices occurs in the solution of *systems of linear equations*, briefly called **linear systems**. Such systems model various problems, for instance, in frameworks, electrical networks, traffic flow, economics, statistics, and many others. In this section we show an important solution method, the *Gauss elimination*. General properties of solutions will be discussed in the next sections.

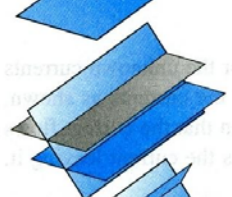
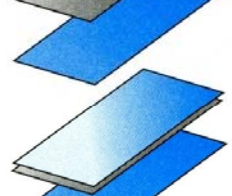
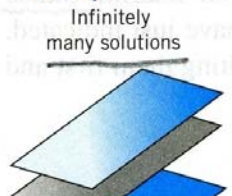
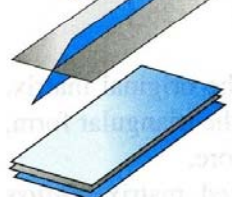
EXAMPLE 1 Geometric Interpretation. Existence and Uniqueness of Solutions



Unique solution



Infinitely many solutions



No solution

If $m = n = 2$, we have two equations in two unknowns x_1, x_2

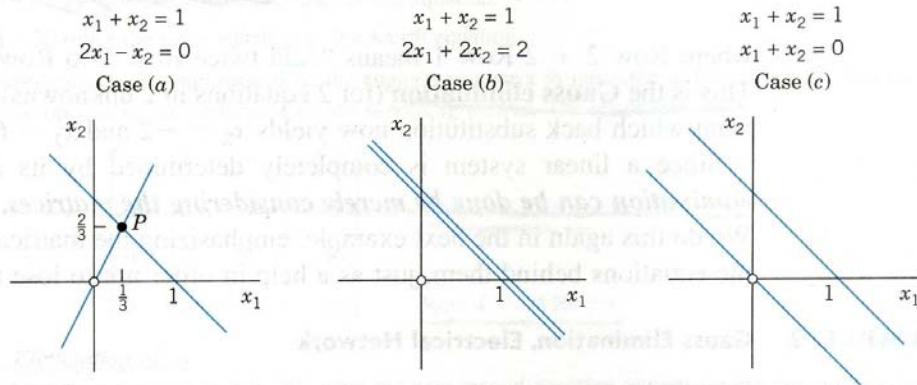
$$a_{11}x_1 + a_{12}x_2 = b_1$$

$$a_{21}x_1 + a_{22}x_2 = b_2.$$

If we interpret x_1, x_2 as coordinates in the x_1x_2 -plane, then each of the two equations represents a straight line, and (x_1, x_2) is a solution if and only if the point P with coordinates x_1, x_2 lies on both lines. Hence there are three possible cases:

- (a) Precisely one solution if the lines intersect.
- (b) Infinitely many solutions if the lines coincide.
- (c) No solution if the lines are parallel

For instance,



If the system is homogenous, Case (c) cannot happen, because then those two straight lines pass through the origin, whose coordinates $0, 0$ constitute the trivial solution. If you wish, consider three equations in three unknowns as representations of three planes in space and discuss the various possible cases in a similar fashion. See Fig. 156. ■

Our simple example illustrates that a system (1) may perhaps have no solution. This poses the following problem. Does a given system (1) have a solution? Under what conditions does it have precisely one solution? If it has more than one solution, how can we characterize the set of all solutions? How can we actually obtain the solutions? Perhaps the last question is the most immediate one from a practical viewpoint. We shall answer it first and discuss the other questions in Sec. 7.5.

Gauss Elimination and Back Substitution

This is a standard elimination method for solving linear systems that proceeds systematically irrespective of particular features of the coefficients. It is a method of great practical importance and is reasonable with respect to computing time and storage demand (two aspects we shall consider in Sec. 20.1 in the chapter on numeric linear algebra). We begin by motivating the method. If a system is in “**triangular form**,” say,

$$2x_1 + 5x_2 = 2$$

$$13x_2 = -26$$

we can solve it by “**back substitution**,” that is, solve the last equation for the variable, $x_2 = -26/13 = -2$, and then work backward, substituting $x_2 = -2$ into the first equation

Fig. 156. Three equations in three unknowns interpreted as planes in space

and solve it for x_1 , obtaining $x_1 = \frac{1}{2}(2 - 5x_2) = \frac{1}{2}(2 - 5 \cdot (-2)) = 6$. This gives us the idea of first reducing a general system to triangular form. For instance, let the given system be

$$\begin{array}{l} \text{X2 3un} \rightarrow \\ \text{(2)ri 2122} \end{array} \quad \begin{array}{r} 2x_1 + 5x_2 = 2 \\ -4x_1 + 3x_2 = -30. \end{array} \quad \text{Its augmented matrix is } \begin{bmatrix} 2 & 5 & 2 \\ -4 & 3 & -30 \end{bmatrix}.$$

We leave the first equation as it is. We eliminate x_1 from the second equation, to get a triangular system. For this we add twice the first equation to the second, and we do the same operation on the *rows* of the augmented matrix. This gives $-4x_1 + 4x_1 + 3x_2 + 10x_2 = -30 + 2 \cdot 2$, that is,

$$\begin{array}{r} 2x_1 + 5x_2 = 2 \\ 13x_2 = -26 \end{array} \quad \text{Row 2} + 2 \text{ Row 1} \quad \begin{bmatrix} 2 & 5 & 2 \\ 0 & 13 & -26 \end{bmatrix}$$

where **Row 2 + 2 Row 1** means "Add twice Row 1 to Row 2" in the original matrix. This is the **Gauss elimination** (for 2 equations in 2 unknowns) giving the triangular form, from which back substitution now yields $x_2 = -2$ and $x_1 = 6$, as before.

Since a linear system is completely determined by its augmented matrix, **Gauss elimination can be done by merely considering the matrices**, as we have just indicated. We do this again in the next example, emphasizing the matrices by writing them first and the equations behind them, just as a help in order not to lose track.

EXAMPLE 2 Gauss Elimination. Electrical Network

Solve the linear system

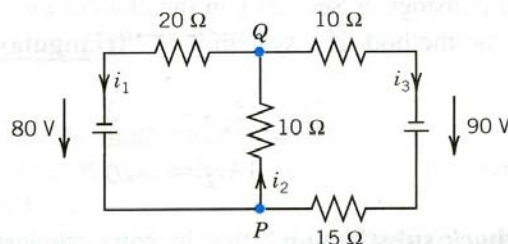
$$\begin{array}{r} x_1 - x_2 + x_3 = 0 \\ -x_1 + x_2 - x_3 = 0 \\ 10x_2 + 25x_3 = 90 \\ 20x_1 + 10x_2 = 80. \end{array}$$

Derivation from the circuit in Fig. 157 (Optional). This is the system for the unknown currents $x_1 = i_1$, $x_2 = i_2$, $x_3 = i_3$ in the electrical network in Fig. 157. To obtain it, we label the currents as shown, choosing directions arbitrarily; if a current will come out negative, this will simply mean that the current flows against the direction of our arrow. The current entering each battery will be the same as the current leaving it. The equations for the currents result from Kirchhoff's laws:

Kirchhoff's current law (KCL). At any point of a circuit, the sum of the inflowing currents equals the sum of the outflowing currents.

Kirchhoff's voltage law (KVL). In any closed loop, the sum of all voltage drops equals the impressed electromotive force.

Node P gives the first equation, node Q the second, the right loop the third, and the left loop the fourth, as indicated in the figure.



$$\text{Node } P: \quad i_1 - i_2 + i_3 = 0$$

$$\text{Node } Q: \quad -i_1 + i_2 - i_3 = 0$$

$$\text{Right loop:} \quad 10i_2 + 25i_3 = 90$$

$$\text{Left loop:} \quad 20i_1 + 10i_2 = 80$$

Fig. 157. Network in Example 2 and equations relating the currents

Solution by Gauss Elimination. This system could be solved rather quickly by noticing its particular form. But this is not the point. The point is that the Gauss elimination is systematic and will work in general, also for large systems. We apply it to our system and then do back substitution. As indicated let us write the augmented matrix of the system first and then the system itself:

	Augmented Matrix \tilde{A}		Equations
Pivot 1 →	$\left[\begin{array}{ccc c} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{array} \right]$	Pivot 1 →	$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 20x_1 + 10x_2 &= 80. \end{aligned}$
Eliminate →		Eliminate →	

Step 1. Elimination of x_1

Call the first row of A the **pivot row** and the first equation the **pivot equation**. Call the coefficient 1 of its x_1 -term the **pivot** in this step. Use this equation to eliminate x_1 (get rid of x_1) in the other equations. For this, do:

- Add 1 times the pivot equation to the second equation.
- Add -20 times the pivot equation to the fourth equation.

This corresponds to **row operations** on the augmented matrix as indicated in **BLUE** behind the **new matrix** in (3). So the operations are performed on the **preceding matrix**. The result is

	$\left[\begin{array}{ccc c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 30 & -20 & 80 \end{array} \right]$		$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 0 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 30x_2 - 20x_3 &= 80. \end{aligned}$
(3)		Row 2 + Row 1	
		Row 4 - 20 Row 1	

Step 2. Elimination of x_2

The first equation remains as it is. We want the new second equation to serve as the next pivot equation. But since it has no x_2 -term (in fact, it is $0 = 0$), we must first change the order of the equations and the corresponding rows of the new matrix. We put $0 = 0$ at the end and move the third equation and the fourth equation one place up. This is called **partial pivoting** (as opposed to the rarely used **total pivoting**, in which also the order of the unknowns is changed). It gives

	$\left[\begin{array}{ccc c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 30 & -20 & 80 \\ 0 & 0 & 0 & 0 \end{array} \right]$		$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 30x_2 - 20x_3 &= 80 \\ 0 &= 0 \end{aligned}$
Pivot 10 →		Pivot 10 →	
Eliminate 30 →		Eliminate 30x ₂ →	

To eliminate x_2 , do:

- Add -3 times the pivot equation to the third equation.

The result is

	$\left[\begin{array}{ccc c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & -95 & -190 \\ 0 & 0 & 0 & 0 \end{array} \right]$		$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ -95x_3 &= -190 \\ 0 &= 0 \end{aligned}$
(4)		Row 3 - 3 Row 2	

Back Substitution. Determination of x_3, x_2, x_1 (in this order)

Working backward from the last to the first equation of this "triangular" system (4), we can now readily find x_3 , then x_2 , and then x_1 :

$-95x_3 = -190$	$x_3 = i_3 = 2 [A]$
$10x_2 + 25x_3 = 90$	$x_2 = \frac{1}{10}(90 - 25x_3) = i_2 = 4 [A]$
$x_1 - x_2 + x_3 = 0$	$x_1 = x_2 - x_3 = i_1 = 2 [A]$

where A stands for "amperes." This is the answer to our problem. The solution is unique. ■

Elementary Row Operations. Row-Equivalent Systems

Example 2 illustrates the operations of the Gauss elimination. These are the first two of three operations, which are called

Elementary Row Operations for Matrices:

Interchange of two rows

Addition of a constant multiple of one row to another row.

*Multiplication of a row by a **nonzero** constant c .*

CAUTION! These operations are for rows, *not for columns!* They correspond to the following

Elementary Operations for Equations:

Interchange of two equations

Addition of a constant multiple of one equation to another equation

*Multiplication of an equation by a **nonzero** constant c .*

Clearly, the interchange of two equations does not alter the solution set. Neither does that addition because we can undo it by a corresponding subtraction. Similarly for that multiplication, which we can undo by multiplying the new equation by $1/c$ (since $c \neq 0$), producing the original equation.

We now call a linear system S_1 **row-equivalent** to a linear system S_2 if S_1 can be obtained from S_2 by (finitely many!) row operations. Thus we have proved the following result, which also justifies the Gauss elimination.

THEOREM 1

Row-Equivalent Systems

Row-equivalent linear systems have the same set of solutions.

Because of this theorem, systems having the same solution sets are often called *equivalent systems*. But note well that we are dealing with **row operations**. No column operations on the augmented matrix are permitted in this context because they would generally alter the solution set.

A linear system (1) is called **overdetermined** if it has more equations than unknowns, as in Example 2, **determined** if $m = n$, as in Example 1, and **underdetermined** if it has fewer equations than unknowns.

Furthermore, a system (1) is called **consistent** if it has at least one solution (thus, one solution or infinitely many solutions), but **inconsistent** if it has no solutions at all, as $x_1 + x_2 = 1$, $x_1 + x_2 = 0$ in Example 1.

Gauss Elimination: The Three Possible Cases of Systems

The Gauss elimination can take care of linear systems with a unique solution (see Example 2), with infinitely many solutions (Example 3, below), and without solutions (inconsistent systems; see Example 4).

EXAMPLE 3 Gauss Elimination if Infinitely Many Solutions Exist

Solve the following linear systems of three equations in four unknowns whose augmented matrix is

$$(5) \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array} \right]. \quad \text{Thus, } \begin{cases} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 = 8.0 \\ 0.6x_1 + 1.5x_2 + 1.5x_3 - 5.4x_4 = 2.7 \\ 1.2x_1 - 0.3x_2 - 0.3x_3 + 2.4x_4 = 2.1. \end{cases}$$

Solution. As in the previous example, we circle pivots and box terms of equations and corresponding entries to be eliminated. We indicate the operations in terms of equations and operate on both equations and matrices.

Step 1. Elimination of x_1 from the second and third equations by adding

$$\begin{aligned} -0.6/3.0 &= -0.2 \text{ times the first equation to the second equation,} \\ -1.2/3.0 &= -0.4 \text{ times the first equation to the third equation.} \end{aligned}$$

This gives the following, in which the pivot of the next step is circled.

$$(6) \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & -1.1 & -1.1 & 4.4 & -1.1 \end{array} \right] \quad \begin{array}{l} \text{Row 2} - 0.2 \text{ Row 1} \\ \text{Row 3} - 0.4 \text{ Row 1} \end{array} \quad \begin{cases} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 = 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 = 1.1 \\ -1.1x_2 - 1.1x_3 + 4.4x_4 = -1.1 \end{cases}$$

Step 2. Elimination of x_2 from the third equation of (6) by adding

$$1.1/1.1 = 1 \text{ times the second equation to the third equation.}$$

This gives

$$(7) \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \quad \begin{array}{l} \text{Row 3} + \text{Row 2} \end{array} \quad \begin{cases} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 = 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 = 1.1 \\ 0 = 0 \end{cases}$$

dependent / freedom!

Back Substitution. From the second equation, $x_2 = 1 - x_3 + 4x_4$. From this and the first equation, $x_1 = 2 - x_4$. Since x_3 and x_4 remain arbitrary, we have infinitely many solutions. If we choose a value of x_3 and a value of x_4 , then the corresponding values of x_1 and x_2 are uniquely determined.

On Notation. If unknowns remain arbitrary, it is also customary to denote them by other letters t_1, t_2, \dots . In this example we may thus write $x_1 = 2 - x_4 = 2 - t_2$, $x_2 = 1 - x_3 + 4x_4 = 1 - t_1 + 4t_2$, $x_3 = t_1$ (first arbitrary unknown), $x_4 = t_2$ (second arbitrary unknown).

EXAMPLE 4 Gauss Elimination if no Solution Exists

What will happen if we apply the Gauss elimination to a linear system that has no solution? The answer is that in this case the method will show this fact by producing a contradiction. For instance, consider

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 4 & 6 \end{array} \right] \quad \begin{cases} 3x_1 + 2x_2 + x_3 = 3 \\ 2x_1 + x_2 + x_3 = 0 \\ 6x_1 + 2x_2 + 4x_3 = 6. \end{cases}$$

Step 1. Elimination of x_1 from the second and third equations by adding

$$\begin{aligned} -\frac{2}{3} \text{ times the first equation to the second equation,} \\ -\frac{6}{3} = -2 \text{ times the first equation to the third equation.} \end{aligned}$$

This gives

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & -2 & 2 & 0 \end{array} \right] \quad \begin{array}{l} \text{Row 2} - \frac{2}{3} \text{ Row 1} \\ \text{Row 3} - 2 \text{ Row 1} \end{array} \quad \begin{cases} 3x_1 + 2x_2 + x_3 = 3 \\ -\frac{1}{3}x_2 + \frac{1}{3}x_3 = -2 \\ -2x_2 + 2x_3 = 0. \end{cases}$$

Step 2. Elimination of x_2 from the third equation gives

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & 0 & 0 & 12 \end{array} \right] \quad \text{Row 3} - 6 \text{ Row 2}$$

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ -\frac{1}{3}x_2 + \frac{1}{3}x_3 &= -2 \\ 0 &= 12. \end{aligned}$$

The false statement $0 = 12$ shows that the system has no solution. ■

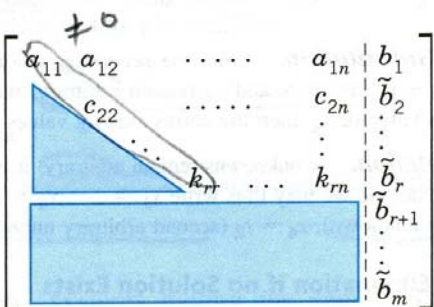
Row Echelon Form and Information From It

At the end of the Gauss elimination the form of the coefficient matrix, the augmented matrix, and the system itself are called the **row echelon form**. In it, rows of zeros, if present, are the last rows, and in each nonzero row the leftmost nonzero entry is farther to the right than in the previous row. For instance, in Example 4 the coefficient matrix and its augmented in row echelon form are

$$\left[\begin{array}{ccc} 3 & 2 & 1 \\ 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 \end{array} \right] \quad \text{and} \quad \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & 0 & 0 & 12 \end{array} \right].$$

Note that we do not require that the leftmost nonzero entries be 1 since this would have no theoretic or numeric advantage. (The so-called *reduced echelon form*, in which those entries are 1, will be discussed in Sec. 7.8.)

At the end of the Gauss elimination (before the back substitution) the row echelon form of the augmented matrix will be

(8) 

Here, $r \leq m$ and $a_{11} \neq 0, c_{22} \neq 0, \dots, k_{rr} \neq 0$, and all the entries in the blue triangle as well as in the blue rectangle are zero. From this we see that with respect to solutions of the system with augmented matrix (8) (and thus with respect to the originally given system) there are three possible cases:

(a) **Exactly one solution** if $r = n$ and $\tilde{b}_{r+1}, \dots, \tilde{b}_m$, if present, are zero. To get the solution, solve the n th equation corresponding to (8) (which is $k_{nn}x_n = \tilde{b}_n$) for x_n , then the $(n - 1)$ st equation for x_{n-1} , and so on up the line. See Example 2, where $r = n = 3$ and $m = 4$.

(b) **Infinitely many solutions** if $r < n$ and $\tilde{b}_{r+1}, \dots, \tilde{b}_m$, if present, are zero. To obtain any of these solutions, choose values of x_{r+1}, \dots, x_n arbitrarily. Then solve the r th equation for x_r , then the $(r - 1)$ st equation for x_{r-1} , and so on up the line. See Example 3.

(c) **No solution** if $r < m$ and one of the entries $\tilde{b}_{r+1}, \dots, \tilde{b}_m$ is not zero. See Example 4, where $r = 2 < m = 3$ and $\tilde{b}_{r+1} = \tilde{b}_3 = 12$.

home work (4, 22)

20. (Wheatstone bridge) Show that if $R_x/R_3 = R_1/R_2$ in the figure, then $I = 0$. (R_0 is the resistance of the instrument by which I is measured.) This bridge is a method for determining R_x . R_1, R_2, R_3 are known. R_3 is variable. To get R_x , make $I = 0$ by varying R_3 . Then calculate $R_x = R_3 R_1 / R_2$.

21. (Traffic flow) Methods of electrical circuit analysis have applications to other fields. For instance, applying the analog of Kirchhoff's current law, find the traffic flow (cars per hour) in the net of one-way streets (in the directions indicated by the arrows) shown in the figure. Is the solution unique?

22. (Models of markets) Determine the equilibrium solution ($D_1 = S_1, D_2 = S_2$) of the two-commodity market with linear model ($D, S, P =$ demand, supply, price; index 1 = first commodity, index 2 = second commodity)

$$D_1 = 60 - 2P_1 - P_2, \quad S_1 = 4P_1 - 2P_2 + 14$$

$$D_2 = 4P_1 - P_2 + 10, \quad S_2 = 5P_2 - 2.$$

23. (Equivalence relation) By definition, an *equivalence relation* on a set is a relation satisfying three conditions (named as indicated):

- Each element A of the set is equivalent to itself ("Reflexivity").
- If A is equivalent to B , then B is equivalent to A ("Symmetry").
- If A is equivalent to B and B is equivalent to C , then A is equivalent to C ("Transitivity").

Show that row equivalence of matrices satisfies these three conditions. *Hint.* Show that for each of the three elementary row operations these conditions hold.

24. **PROJECT. Elementary Matrices.** The idea is that elementary operations can be accomplished by matrix multiplication. If A is an $m \times n$ matrix on which we want to do an elementary operation, then there is a matrix E such that EA is the new matrix after the operation. Such an E is called an **elementary matrix**. This idea can be helpful, for instance, in the design of algorithms. (Computationally, it is generally preferable

to do row operations *directly*, rather than by multiplication by E .)

(a) Show that the following are elementary matrices, for interchanging Rows 2 and 3, for adding -5 times the first row to the third, and for multiplying the fourth row by 8.

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -5 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

Apply E_1, E_2, E_3 to a vector and to a 4×3 matrix of your choice. Find $B = E_3 E_2 E_1 A$, where $A = [a_{jk}]$ is the general 4×2 matrix. Is B equal to $C = E_1 E_2 E_3 A$?

(b) Conclude that E_1, E_2, E_3 are obtained by doing the corresponding elementary operations on the 4×4 unit matrix. Prove that if M is obtained from A by an elementary row operation, then

$$M = EA,$$

where E is obtained from the $n \times n$ unit matrix I_n by the same row operation.

25. **CAS PROJECT. Gauss Elimination and Back Substitution.** Write a program for Gauss elimination and back substitution (a) that does not include pivoting, (b) that does include pivoting. Apply the programs to Probs. 13–16 and to some larger systems of your choice.

7.4 Linear Independence. Rank of a Matrix. Vector Space

In the last section we explained the Gauss elimination with back substitution, the most important numeric solution method for linear systems of equations. It appeared that such a system may have a unique solution or infinitely many solutions, or it may be inconsistent, that is, have no solution at all. Hence we are confronted with the questions of **existence and uniqueness of solutions**. We shall answer these questions in the next section. As the

key concept for this (and other questions) we introduce the **rank of a matrix**. To define rank, we first need the following concepts, which are of general importance.

Linear Independence and Dependence of Vectors

Given any set of m vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(m)}$ (with the same number of components), a **linear combination** of these vectors is an expression of the form

$$c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \dots + c_m\mathbf{a}_{(m)}$$

where c_1, c_2, \dots, c_m are any scalars. Now consider the equation

$$(1) \quad c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \dots + c_m\mathbf{a}_{(m)} = \mathbf{0}.$$

Clearly, this vector equation (1) holds if we choose all c_j 's zero, because then it becomes $\mathbf{0} = \mathbf{0}$. If this is the only m -tuple of scalars for which (1) holds, then our vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(m)}$ are said to form a **linearly independent set** or, more briefly, we call them **linearly independent**. Otherwise, if (1) also holds with scalars not all zero, we call these vectors **linearly dependent**, because then we can express (at least) one of them as a linear combination of the others. For instance, if (1) holds with, say, $c_1 \neq 0$, we can solve (1) for $\mathbf{a}_{(1)}$:

$$\mathbf{a}_{(1)} = k_2\mathbf{a}_{(2)} + \dots + k_m\mathbf{a}_{(m)} \quad \text{where } k_j = -c_j/c_1.$$

(Some k_j 's may be zero. Or even all of them, namely, if $\mathbf{a}_{(1)} = \mathbf{0}$.)

Why is this important? Well, in the case of linear dependence we can get rid of some of the vectors until we arrive at a linearly independent set that is optimal to work with because it is smallest possible in the sense that it consists only of the "really essential" vectors, which can no longer be expressed linearly in terms of each other. This motivates the idea of a "basis" used in various contexts, notably later in our present section.

EXAMPLE 1 Linear Independence and Dependence

The three vectors

$$\mathbf{a}_{(1)} = [3 \quad 0 \quad 2 \quad 2]$$

$$\mathbf{a}_{(2)} = [-6 \quad 42 \quad 24 \quad 54]$$

$$\mathbf{a}_{(3)} = [21 \quad -21 \quad 0 \quad -15]$$

are linearly dependent because

$$6\mathbf{a}_{(1)} - \frac{1}{2}\mathbf{a}_{(2)} - \mathbf{a}_{(3)} = \mathbf{0}.$$

Although this is easily checked (do it!), it is not so easy to discover. However, a systematic method for finding out about linear independence and dependence follows below.

The first two of the three vectors are linearly independent because $c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} = \mathbf{0}$ implies $c_2 = 0$ (from the second components) and then $c_1 = 0$ (from any other component of $\mathbf{a}_{(1)}$). ■

Rank of a Matrix

DEFINITION

The **rank** of a matrix \mathbf{A} is the maximum number of linearly independent row vectors of \mathbf{A} . It is denoted by $\text{rank } \mathbf{A}$.

Our further discussion will show that the rank of a matrix is an important key concept for understanding general properties of matrices and linear systems of equations.

EXAMPLE 2 Rank

The matrix

$$(2) \quad A = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix}$$

has rank 2, because Example 1 shows that the first two row vectors are linearly independent, whereas all three row vectors are linearly dependent.

Note further that $\text{rank } A = 0$ if and only if $A = \mathbf{0}$. This follows directly from the definition. ■

We call a matrix A_1 **row-equivalent** to a matrix A_2 if A_1 can be obtained from A_2 by (finitely many!) elementary row operations.

Now the maximum number of linearly independent row vectors of a matrix does not change if we change the order of rows or multiply a row by a nonzero c or take a linear combination by adding a multiple of a row to another row. This proves that rank is **invariant** under elementary row operations:

THEOREM 1

Row-Equivalent Matrices

Row-equivalent matrices have the same rank.

Hence we can determine the rank of a matrix by reduction to row-echelon form (Sec. 7.3) and then see the rank directly.

EXAMPLE 3 Determination of Rank

For the matrix in Example 2 we obtain successively

$$A = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix} \quad (\text{given})$$

$$\begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & -21 & -14 & -29 \end{bmatrix} \quad \begin{array}{l} \text{Row 2} + 2 \text{ Row 1} \\ \text{Row 3} - 7 \text{ Row 1} \end{array}$$

$$\begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{Row 3} + \frac{1}{2} \text{ Row 2}$$

Since rank is defined in terms of two vectors, we immediately have the useful

THEOREM 2

Linear Independence and Dependence of Vectors

p vectors with n components each are linearly independent if the matrix with these vectors as row vectors has rank p , but they are linearly dependent if that rank is less than p .

Further important properties will result from the basic

THEOREM 3

Rank in Terms of Column Vectors

The rank r of a matrix A equals the maximum number of linearly independent column vectors of A .

Hence A and its transpose A^T have the same rank.

PROOF

In this proof we write simply "rows" and "columns" for row and column vectors. Let A be an $m \times n$ matrix of rank $A = r$. Then by definition of rank, A has r linearly independent rows which we denote by $v_{(1)}, \dots, v_{(r)}$ (regardless of their position in A), and all the rows $a_{(1)}, \dots, a_{(m)}$ of A are linear combinations of those, say,

$$\begin{aligned} a_{(1)} &= c_{11}v_{(1)} + c_{12}v_{(2)} + \dots + c_{1r}v_{(r)} \\ a_{(2)} &= c_{21}v_{(1)} + c_{22}v_{(2)} + \dots + c_{2r}v_{(r)} \\ &\vdots \\ &\vdots \\ a_{(m)} &= c_{m1}v_{(1)} + c_{m2}v_{(2)} + \dots + c_{mr}v_{(r)}. \end{aligned} \tag{3}$$

These are vector equations for rows. To switch to columns, we write (3) in terms of components as n such systems, with $k = 1, \dots, n$,

$$\begin{aligned} a_{1k} &= c_{11}v_{1k} + c_{12}v_{2k} + \dots + c_{1r}v_{rk} \\ a_{2k} &= c_{21}v_{1k} + c_{22}v_{2k} + \dots + c_{2r}v_{rk} \\ &\vdots \\ &\vdots \\ a_{mk} &= c_{m1}v_{1k} + c_{m2}v_{2k} + \dots + c_{mr}v_{rk} \end{aligned} \tag{4}$$

and collect components in columns. Indeed, we can write (4) as

$$\begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} = v_{1k} \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + v_{2k} \begin{bmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{bmatrix} + \dots + v_{rk} \begin{bmatrix} c_{1r} \\ c_{2r} \\ \vdots \\ c_{mr} \end{bmatrix} \tag{5}$$

** column m 의 row 보다 큰 rank를 뜻하는 것*

(3) r의 column

where $k = 1, \dots, n$. Now the vector on the left is the k th column vector of A . We see that each of these n columns is a linear combination of the same r columns on the right. Hence A cannot have more linearly independent columns than rows, whose number is rank $A = r$. Now rows of A are columns of the transpose A^T . For A^T our conclusion is that A^T cannot have more linearly independent columns than rows, so that A cannot have more linearly independent rows than columns. Together, the number of linearly independent columns of A must be r , the rank of A . This completes the proof. ■

EXAMPLE 4

Illustration of Theorem 3

The matrix in (2) has rank 2. From Example 3 we see that the first two row vectors are linearly independent and by "working backward" we can verify that Row 3 = 6 Row 1 - 1/2 Row 2. Similarly, the first two columns are linearly independent, and by reducing the last matrix in Example 3 by columns we find that

$$\text{Column 3} = \frac{2}{3} \text{Column 1} + \frac{2}{3} \text{Column 2} \quad \text{and} \quad \text{Column 4} = \frac{2}{3} \text{Column 1} + \frac{29}{21} \text{Column 2}. \quad \blacksquare$$

Combining Theorems 2 and 3 we obtain

THEOREM 4

Linear Dependence of Vectors

p vectors with $n < p$ components are always linearly dependent.

PROOF

The matrix \mathbf{A} with those p vectors as row vectors has p rows and $n < p$ columns; hence by Theorem 3 it has rank $\mathbf{A} \leq n < p$, which implies linear dependence by Theorem 2. ■

Vector Space

The following related concepts are of general interest in linear algebra. In the present context they provide a clarification of essential properties of matrices and their role in connection with linear systems.

A **vector space** is a (nonempty) set V of vectors such that with any two vectors \mathbf{a} and \mathbf{b} in V all their linear combinations $\alpha \mathbf{a} + \beta \mathbf{b}$ (α, β any real numbers) are elements of V , and these vectors satisfy the laws (3) and (4) in Sec. 7.1 (written in lowercase letters $\mathbf{a}, \mathbf{b}, \mathbf{u}, \dots$, which is our notation for vectors). (This definition is presently sufficient. General vector spaces will be discussed in Sec. 7.9.)

The maximum number of linearly independent vectors in V is called the **dimension** of V and is denoted by $\dim V$. Here we assume the dimension to be finite; infinite dimension will be defined in Sec. 7.9.

A linearly independent set in V consisting of a maximum possible number of vectors in V is called a **basis** for V . Thus the number of vectors of a basis for V equals $\dim V$.

The set of all linear combinations of given vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(p)}$ with the same number of components is called the **span** of these vectors. Obviously, a span is a vector space.

By a **subspace** of a vector space V we mean a nonempty subset of V (including V itself) that forms itself a vector space with respect to the two algebraic operations (addition and scalar multiplication) defined for the vectors of V .

EXAMPLE 5

Vector Space, Dimension, Basis

The span of the three vectors in Example 1 is a vector space of dimension 2, and a basis is $\mathbf{a}_{(1)}, \mathbf{a}_{(2)}$, for instance, or $\mathbf{a}_{(1)}, \mathbf{a}_{(3)}$, etc. ■

We further note the simple

THEOREM 5

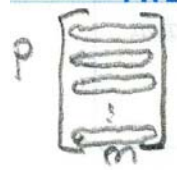
Vector Space \mathbb{R}^n

The vector space \mathbb{R}^n consisting of all vectors with n components (n real numbers) has dimension n .

PROOF

A basis of n vectors is $\mathbf{a}_{(1)} = [1 \ 0 \ \dots \ 0]$, $\mathbf{a}_{(2)} = [0 \ 1 \ 0 \ \dots \ 0], \dots$, $\mathbf{a}_{(n)} = [0 \ \dots \ 0 \ 1]$. ■

In the case of a matrix \mathbf{A} we call the span of the row vectors the **row space** of \mathbf{A} and the span of the column vectors the **column space** of \mathbf{A} .



(associative
and
distributive)

a basis
= dim V

Now, Theorem 3 shows that a matrix A has as many linearly independent rows as columns. By the definition of dimension, their number is the dimension of the row space or the column space of **A**. This proves

THEOREM 6

Row Space and Column Space

The row space and the column space of a matrix **A** have the same dimension, equal to rank **A**.

Finally, for a given matrix **A** the solution set of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ is a vector space, called the null space of **A**, and its dimension is called the nullity of **A**. In the next section we motivate and prove the basic relation

(6) $\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = \text{Number of columns of } \mathbf{A}.$

rank: matrix, dimension: Vector space home work 10, 22

PROBLEM SET 7.4

1-12 RANK, ROW SPACE, COLUMN SPACE

Find the rank and a basis for the row space and for the column space. *Hint.* Row-reduce the matrix and its transpose. (You may omit obvious factors from the vectors of these bases.)

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

2. $\begin{bmatrix} 8 & 2 & 5 \\ 16 & 6 & 29 \\ 4 & 0 & -7 \end{bmatrix}$

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

4. $\begin{bmatrix} a & b & c \\ b & a & c \end{bmatrix}$

5. $\begin{bmatrix} 0 & 3 & 4 \\ -3 & 0 & -5 \\ -4 & 5 & 0 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 1 & a \\ 1 & a & 1 \\ a & 1 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 8 & 0 & 4 \\ 0 & 2 & 0 \\ 4 & 0 & 2 \\ 0 & 4 & 0 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -2 & 3 & -4 \\ 2 & -3 & 4 & -1 \\ 3 & -4 & 1 & -2 \\ 4 & -1 & 2 & -3 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 5 & 8 & -37 \\ 3 & 8 & 7 & 0 \\ 0 & -37 & 0 & 37 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}$

11. $\begin{bmatrix} 2 & 4 & 8 & 16 \\ 16 & 8 & 4 & 2 \\ 4 & 8 & 16 & 2 \\ 2 & 16 & 8 & 4 \end{bmatrix}$

12. $\begin{bmatrix} 0 & 0 & -7 & 1 \\ 0 & 0 & 5 & 0 \\ -7 & 5 & 0 & 2 \\ 1 & 0 & 2 & 0 \end{bmatrix}$

13-20 LINEAR INDEPENDENCE

Are the following sets of vectors linearly independent? (Show the details.)

13. $[3 \ -2 \ 0 \ 4], [5 \ 0 \ 0 \ 1], [-6 \ 1 \ 0 \ 1], [2 \ 0 \ 0 \ 3]$

14. $[1 \ 1 \ 0], [1 \ 0 \ 0], [1 \ 1 \ 1]$

15. $[6 \ 0 \ 3 \ 1 \ 4 \ 2], [0 \ -1 \ 2 \ 7 \ 0 \ 5], [12 \ 3 \ 0 \ -19 \ 8 \ -11]$

16. $[3 \ 4 \ 7], [2 \ 0 \ 3], [8 \ 2 \ 3], [5 \ 5 \ 6]$

17. $[0.2 \ 1.2 \ 5.3 \ 2.8 \ 1.6], [4.3 \ 3.4 \ 0.9 \ 2.0 \ -4.3]$

is **consistent**, that is, has solutions, if and only if the coefficient matrix \mathbf{A} and the augmented matrix $\tilde{\mathbf{A}}$ have the same rank. Here,

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{A}} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right]$$

(b) **Uniqueness.** The system (1) has precisely one solution if and only if this common rank r of \mathbf{A} and $\tilde{\mathbf{A}}$ equals n .

(c) **Infinitely many solutions.** If this common rank r is less than n , the system (1) has infinitely many solutions. All of these solutions are obtained by determining r suitable unknowns (whose submatrix of coefficients must have rank r) in terms of the remaining $n - r$ unknowns, to which arbitrary values can be assigned. (See Example 3 in Sec. 7.3.)

(d) **Gauss elimination (Sec. 7.3).** If solutions exist, they can all be obtained by the Gauss elimination. (This method will automatically reveal whether or not solutions exist; see Sec. 7.3.)

rank = r

\times , $m-r$ 자유도: 인의의 수

PROOF (a) We can write the system (1) in vector form $\mathbf{Ax} = \mathbf{b}$ or in terms of column vectors $\mathbf{c}_{(1)}, \dots, \mathbf{c}_{(n)}$ of \mathbf{A} :

$$(2) \quad \mathbf{c}_{(1)}x_1 + \mathbf{c}_{(2)}x_2 + \cdots + \mathbf{c}_{(n)}x_n = \mathbf{b}.$$

rank r 은
→ 특정 column 수
아래다

$\tilde{\mathbf{A}}$ is obtained by augmenting \mathbf{A} by a single column \mathbf{b} . Hence, by Theorem 3 in Sec. 7.4, rank $\tilde{\mathbf{A}}$ equals rank \mathbf{A} or rank $\mathbf{A} + 1$. Now if (1) has a solution \mathbf{x} , then (2) shows that \mathbf{b} must be a linear combination of those column vectors, so that $\tilde{\mathbf{A}}$ and \mathbf{A} have the same maximum number of linearly independent column vectors and thus the same rank.

Conversely, if rank $\tilde{\mathbf{A}} = \text{rank } \mathbf{A}$, then \mathbf{b} must be a linear combination of the column vectors of \mathbf{A} , say,

$$(2^*) \quad \mathbf{b} = \alpha_1 \mathbf{c}_{(1)} + \cdots + \alpha_n \mathbf{c}_{(n)}$$

$\Rightarrow \alpha$ 의 해를 구하면
의미

since otherwise rank $\tilde{\mathbf{A}} = \text{rank } \mathbf{A} + 1$. But (2*) means that (1) has a solution, namely, $x_1 = \alpha_1, \dots, x_n = \alpha_n$, as can be seen by comparing (2*) and (2).

(b) If rank $\mathbf{A} = n$, the n column vectors in (2) are linearly independent by Theorem 3 in Sec. 7.4. We claim that then the representation (2) of \mathbf{b} is unique because otherwise

$$\mathbf{c}_{(1)}x_1 + \cdots + \mathbf{c}_{(n)}x_n = \mathbf{c}_{(1)}\tilde{x}_1 + \cdots + \mathbf{c}_{(n)}\tilde{x}_n.$$

unique 하기
안된다면

This would imply (take all terms to the left, with a minus sign)

$$(x_1 - \tilde{x}_1)\mathbf{c}_{(1)} + \cdots + (x_n - \tilde{x}_n)\mathbf{c}_{(n)} = \mathbf{0}$$

and $x_1 - \tilde{x}_1 = 0, \dots, x_n - \tilde{x}_n = 0$ by linear independence. But this means that the scalars x_1, \dots, x_n in (2) are uniquely determined, that is, the solution of (1) is unique.

해가 있으면
b는 A의 column
으로 선형결합
 $\Rightarrow A$ 와 \tilde{A} rank는
같다.

PROOF

The first proposition can be seen directly from the system. It agrees with the fact that $\mathbf{b} = \mathbf{0}$ implies that $\text{rank } \tilde{\mathbf{A}} = \text{rank } \mathbf{A}$, so that a homogeneous system is always **consistent**. If $\text{rank } \mathbf{A} = n$, the trivial solution is the unique solution according to (b) in Theorem 1. If $\text{rank } \mathbf{A} < n$, there are nontrivial solutions according to (c) in Theorem 1. The solutions form a vector space because if $\mathbf{x}_{(1)}$ and $\mathbf{x}_{(2)}$ are any of them, then $\mathbf{A}\mathbf{x}_{(1)} = \mathbf{0}$, $\mathbf{A}\mathbf{x}_{(2)} = \mathbf{0}$, and this implies $\mathbf{A}(\mathbf{x}_{(1)} + \mathbf{x}_{(2)}) = \mathbf{A}\mathbf{x}_{(1)} + \mathbf{A}\mathbf{x}_{(2)} = \mathbf{0}$ as well as $\mathbf{A}(c\mathbf{x}_{(1)}) = c\mathbf{A}\mathbf{x}_{(1)} = \mathbf{0}$, where c is arbitrary. If $\text{rank } \mathbf{A} = r < n$, Theorem 1 (c) implies that we can choose $n - r$ suitable unknowns, call them x_{r+1}, \dots, x_n , in an arbitrary fashion, and every solution is obtained in this way. Hence a basis for the solution space, briefly called a **basis of solutions** of (4), is $\mathbf{y}_{(1)}, \dots, \mathbf{y}_{(n-r)}$, where the basis vector $\mathbf{y}_{(j)}$ is obtained by choosing $x_{r+j} = 1$ and the other x_{r+1}, \dots, x_n zero; the corresponding first r components of this solution vector are then determined. Thus the solution space of (4) has dimension $n - r$. This proves Theorem 2. ■

The solution space of (4) is also called the **null space** of \mathbf{A} because $\mathbf{A}\mathbf{x} = \mathbf{0}$ for every \mathbf{x} in the solution space of (4). Its dimension is called the **nullity** of \mathbf{A} . Hence Theorem 2 states that

(5)

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = n$$

where n is the number of unknowns (number of columns of \mathbf{A}). Furthermore, by the definition of rank we have $\text{rank } \mathbf{A} \leq m$ in (4). Hence if $m < n$, then $\text{rank } \mathbf{A} < n$. By Theorem 2 this gives the practically important

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \tilde{\mathbf{A}}$
 trivial only
 nontrivial
 $n-r$ 중
 인의 위치를
 보면 \rightarrow
 $n-r$ 개의
 basis의
 vector space
 수 있다.

THEOREM 3

Homogeneous Linear System with Fewer Equations Than Unknowns

A homogeneous linear system with fewer equations than unknowns has always nontrivial solutions.

Nonhomogeneous Linear Systems

The characterization of all solutions of the linear system (1) is now quite simple, as follows.

THEOREM 4

Nonhomogeneous Linear System

If a nonhomogeneous linear system (1) is **consistent**, then all of its solutions are obtained as

(6)
$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$$

where \mathbf{x}_0 is any (fixed) solution of (1) and \mathbf{x}_h runs through all the solutions of the corresponding homogeneous system (4).

PROOF

The difference $\mathbf{x}_h = \mathbf{x} - \mathbf{x}_0$ of any two solutions of (1) is a solution of (4) because $\mathbf{A}\mathbf{x}_h = \mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}$. Since \mathbf{x} is any solution of (1), we get all the solutions of (1) if in (6) we take any solution \mathbf{x}_0 of (1) and let \mathbf{x}_h vary throughout the solution space of (4). ■

\rightarrow 해의 개수

7.6 For Reference: Second- and Third-Order Determinants

We explain these determinants separately from the general theory in Sec. 7.7 because they will be sufficient for many of our examples and problems. Since this section is for reference, *go on to the next section, consulting this material only when needed.*

A **determinant of second order** is denoted and defined by

$$(1) \quad D = \det \mathbf{A} \triangleq \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \triangleq a_{11}a_{22} - a_{12}a_{21}.$$

So here we have *bars* (whereas a matrix has *brackets*).

Cramer's rule for solving linear systems of two equations in two unknowns

$$(2) \quad \begin{aligned} (a) \quad & a_{11}x_1 + a_{12}x_2 = b_1 \\ (b) \quad & a_{21}x_1 + a_{22}x_2 = b_2 \end{aligned}$$

is

$$(3) \quad x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{D} = \frac{b_1a_{22} - a_{12}b_2}{D},$$

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{D} = \frac{a_{11}b_2 - b_1a_{21}}{D}.$$

with D as in (1), provided

$$D \neq 0.$$

The value $D = 0$ appears for inconsistent nonhomogeneous systems and for homogeneous systems with nontrivial solutions.

PROOF

We prove (3). To eliminate x_2 , multiply (2a) by a_{22} and (2b) by $-a_{12}$ and add,

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - a_{12}b_2.$$

Similarly, to eliminate x_1 , multiply (2a) by $-a_{21}$ and (2b) by a_{11} and add,

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - b_1a_{21}.$$

Assuming that $D = a_{11}a_{22} - a_{12}a_{21} \neq 0$, dividing, and writing the right sides of these two equations as determinants, we obtain (3).

EXAMPLE 1 Cramer's Rule for Two Equations

If
$$\begin{cases} 4x_1 + 3x_2 = 12 \\ 2x_1 + 5x_2 = -8 \end{cases}$$
 then
$$x_1 = \frac{\begin{vmatrix} 12 & 3 \\ -8 & 5 \end{vmatrix}}{\begin{vmatrix} 4 & 3 \\ 2 & 5 \end{vmatrix}} = \frac{84}{14} = 6, \quad x_2 = \frac{\begin{vmatrix} 4 & 12 \\ 2 & -8 \end{vmatrix}}{\begin{vmatrix} 4 & 3 \\ 2 & 5 \end{vmatrix}} = \frac{-56}{14} = -4. \quad \blacksquare$$

Third-Order Determinants

A determinant of third order can be defined by

$$(4) \quad D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

Note the following. The signs on the right are + - +. Each of the three terms on the right is an entry in the first column of D times its **minor**, that is, the second-order determinant obtained from D by deleting the row and column of that entry; thus, for a_{11} delete the first row and first column, and so on.

If we write out the minors in (4), we obtain

$$(4^*) \quad D = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{21}a_{13}a_{32} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}.$$

Cramer's Rule for Linear Systems of Three Equations

$$(5) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

is

$$(6) \quad x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad x_3 = \frac{D_3}{D} \quad (D \neq 0)$$

with the determinant D of the system given by (4) and

$$D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

Note that D_1, D_2, D_3 are obtained by replacing Columns 1, 2, 3, respectively, by the column of the right sides of (5).

Cramer's rule (6) can be derived by eliminations similar to those for (3), but it also follows from the general case (Theorem 4) in the next section.

7.7 Determinants. Cramer's Rule

Determinants were originally introduced for solving linear systems. Although impractical in computations, they have important engineering applications in eigenvalue problems (Sec. 8.1), differential equations, vector algebra (Sec. 9.3), and so on. They can be introduced in several equivalent ways. Our definition is particularly practical in connection with linear systems.

A determinant of order n is a scalar associated with an $n \times n$ (hence square!) matrix $A = [a_{jk}]$, which is written

$$(1) \quad D = \det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

and is defined for $n = 1$ by

$$(2) \quad D = a_{11}$$

and for $n \geq 2$ by

$$(3a) \quad D \triangleq a_{j1}C_{j1} + a_{j2}C_{j2} + \cdots + a_{jn}C_{jn} \quad (j = 1, 2, \dots, \text{or } n)$$

or

$$(3b) \quad D \triangleq a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk} \quad (k = 1, 2, \dots, \text{or } n)$$

Here,

Cofactor $C_{jk} = (-1)^{j+k} M_{jk}$ — minor

and M_{jk} is a determinant of order $n - 1$, namely, the determinant of the submatrix of A obtained from A by omitting the row and column of the entry a_{jk} , that is, the j th row and the k th column.

In this way, D is defined in terms of n determinants of order $n - 1$, each of which is, in turn, defined in terms of $n - 1$ determinants of order $n - 2$, and so on; we finally arrive at second-order determinants, in which those submatrices consist of single entries whose determinant is defined to be the entry itself.

From the definition it follows that we may expand D by any row or column, that is, choose in (3) the entries in any row or column, similarly when expanding the C_{jk} 's in (3), and so on.

This definition is unambiguous, that is, yields the same value for D no matter which columns or rows we choose in expanding. A proof is given in App. 4.

definition!!

row

column

Terms used in connection with determinants are taken from matrices. In D we have n^2 entries a_{jk} , also n rows and n columns, and a **main diagonal** on which $a_{11}, a_{22}, \dots, a_{nn}$ stand. Two terms are new:

M_{jk} is called the **minor of a_{jk} in D** , and C_{jk} the **cofactor of a_{jk} in D** .

For later use we note that (3) may also be written in terms of minors

$$(4a) \quad D = \sum_{k=1}^n (-1)^{j+k} a_{jk} M_{jk} \quad (j = 1, 2, \dots, \text{or } n)$$

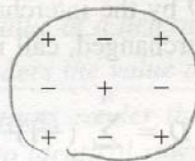
$$(4b) \quad D = \sum_{j=1}^n (-1)^{j+k} a_{jk} M_{jk} \quad (k = 1, 2, \dots, \text{or } n).$$

EXAMPLE 1 Minors and Cofactors of a Third-Order Determinant

In (4) of the previous section the minors and cofactors of the entries in the first column can be seen directly. For the entries in the second row the minors are

$$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \quad M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

and the cofactors are $C_{21} = -M_{21}$, $C_{22} = +M_{22}$, and $C_{23} = -M_{23}$. Similarly for the third row—write these down yourself. And verify that the signs in C_{jk} form a **checkerboard pattern**



EXAMPLE 2 Expansions of a Third-Order Determinant

$$D = \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 1 \begin{vmatrix} 6 & 4 \\ 0 & 2 \end{vmatrix} - 3 \begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} + 0 \begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix} = 1(12 - 0) - 3(4 + 4) + 0(0 + 6) = -12.$$

This is the expansion by the first row. The expansion by the third column is

$$D = 0 \begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix} - 4 \begin{vmatrix} 1 & 3 \\ -1 & 0 \end{vmatrix} + 2 \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0 - 12 + 0 = -12.$$

Verify that the other four expansions also give the value -12 .

EXAMPLE 3 Determinant of a Triangular Matrix

$$\begin{vmatrix} -3 & 0 & 0 \\ 6 & 4 & 0 \\ -1 & 2 & 5 \end{vmatrix} = -3 \begin{vmatrix} 4 & 0 \\ 2 & 5 \end{vmatrix} = -3 \cdot 4 \cdot 5 = -60.$$

Inspired by this, can you formulate a little theorem on determinants of triangular matrices? Of diagonal matrices?

General Properties of Determinants

To obtain the value of a determinant (1), we can first simplify it systematically by elementary row operations, similar to those for matrices in Sec. 7.3, as follows.

THEOREM 1

Behavior of an n th-Order Determinant under Elementary Row Operations

- (a) Interchange of two rows multiplies the value of the determinant by -1 .
- (b) Addition of a multiple of a row to another row does not alter the value of the determinant.
- (c) Multiplication of a row by a nonzero constant c multiplies the value of the determinant by c . (This holds also when $c = 0$, but gives no longer an elementary row operation.)

PROOF (a) By induction. The statement holds for $n = 2$ because

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \quad \text{but} \quad \begin{vmatrix} c & d \\ a & b \end{vmatrix} = bc - ad.$$

We now make the induction hypothesis that (a) holds for determinants of order $n - 1 \geq 2$ and show that it then holds for determinants of order n . Let D be of order n . Let E be obtained from D by the interchange of two rows. Expand D and E by a row that is not one of those interchanged, call it the j th row. Then by (4a),

*j*th row; not interchanged →

$$(5) \quad D = \sum_{k=1}^n (-1)^{j+k} a_{jk} M_{jk}, \quad E = \sum_{k=1}^n (-1)^{j+k} a_{jk} N_{jk}$$

where N_{jk} is obtained from the minor M_{jk} of a_{jk} in D by the interchange of those two rows which have been interchanged in D (and which N_{jk} must both contain because we expand by another row!). Now these minors are of order $n - 1$. Hence the induction hypothesis applies and gives $N_{jk} = -M_{jk}$. Thus $E = -D$ by (5).

(b) Add c times Row i to Row j . Let \tilde{D} be the new determinant. Its entries in Row j are $a_{jk} + ca_{ik}$. If we expand \tilde{D} by this Row j , we see that we can write it as $\tilde{D} = D_1 + cD_2$, where $D_1 = D$ has in Row j the a_{jk} , whereas D_2 has in that Row j the a_{ik} from the addition. Hence D_2 has a_{ik} in both Row i and Row j . Interchanging these two rows gives D_2 back, but on the other hand it gives $-D_2$ by (a). Together $D_2 = -D_2 = 0$, so that $\tilde{D} = D_1 = D$.

(c) Expand the determinant by the row that has been multiplied.

CAUTION! $\det(cA) = c^n \det A$ (not $c \det A$). Explain why. ■

EXAMPLE 4 Evaluation of Determinants by Reduction to Triangular Form

Because of Theorem 1 we may evaluate determinants by reduction to triangular form, as in the Gauss elimination for a matrix. For instance (with the blue explanations always referring to the preceding determinant)

$$D = \begin{vmatrix} 2 & 0 & -4 & 6 \\ 4 & 5 & 1 & 0 \\ 0 & 2 & 6 & -1 \\ -3 & 8 & 9 & 1 \end{vmatrix}$$

$$\begin{aligned}
 &= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 2 & 6 & -1 \\ 0 & 8 & 3 & 10 \end{vmatrix} \begin{array}{l} \text{Row 2} - 2 \text{ Row 1} \\ \text{Row 4} + 1.5 \text{ Row 1} \end{array} \\
 &= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -11.4 & 29.2 \end{vmatrix} \begin{array}{l} \text{Row 3} - 0.4 \text{ Row 2} \\ \text{Row 4} - 1.6 \text{ Row 2} \end{array} \\
 &= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -0 & 47.25 \end{vmatrix} \begin{array}{l} \text{Row 4} + 4.75 \text{ Row 3} \end{array} \\
 &= 2 \cdot 5 \cdot 2.4 \cdot 47.25 = 1134.
 \end{aligned}$$

THEOREM 2

Further Properties of n th-Order Determinants

- (a)–(c) in Theorem 1 hold also for columns.
- (d) Transposition leaves the value of a determinant unaltered.
- (e) A zero row or column renders the value of a determinant zero.
- (f) Proportional rows or columns render the value of a determinant zero. In particular, a determinant with two identical rows or columns has the value zero.

PROOF

(a)–(e) follow directly from the fact that a determinant can be expanded by any row column. In (d), transposition is defined as for matrices, that is, the j th row becomes the j th column of the transpose.

(f) If Row $j = c$ times Row i , then $D = cD_1$, where D_1 has Row $j =$ Row i . Hence an interchange of these rows reproduces D_1 , but it also gives $-D_1$ by Theorem 1(a). Hence $D_1 = 0$ and $D = cD_1 = 0$. Similarly for columns.

It is quite remarkable that the important concept of the rank of a matrix \mathbf{A} , which is the maximum number of linearly independent row or column vectors of \mathbf{A} (see Sec. 7.4), can be related to determinants. Here we may assume that rank $\mathbf{A} > 0$ because the only matrices with rank 0 are the zero matrices (see Sec. 7.4).

THEOREM 3

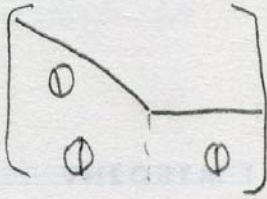
Rank in Terms of Determinants

An $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ has rank $r \geq 1$ if and only if \mathbf{A} has an $r \times r$ submatrix with nonzero determinant, whereas every square submatrix with more than r rows that \mathbf{A} has (or does not have!) has determinant equal to zero.

In particular, if \mathbf{A} is square, $n \times n$, it has rank n if and only if

$$\det \mathbf{A} \neq 0.$$

PROOF



The key idea is that elementary row operations (Sec. 7.3) alter neither rank (by Theorem 1 in Sec. 7.4) nor the property of a determinant being nonzero (by Theorem 1 in this section). The echelon form \hat{A} of A (see Sec. 7.3) has r nonzero row vectors (which are the first r row vectors) if and only if $\text{rank } A = r$. Let \hat{R} be the $r \times r$ submatrix in the left upper corner of \hat{A} (so that the entries of \hat{R} are in both the first r rows and r columns of \hat{A}). Now \hat{R} is triangular, with all diagonal entries r_{jj} nonzero. Thus, $\det \hat{R} = r_{11} \cdots r_{rr} \neq 0$. Also $\det R \neq 0$ for the corresponding $r \times r$ submatrix R of A because \hat{R} results from R by elementary row operations. Similarly, $\det S = 0$ for any square submatrix S of $r + 1$ or more rows perhaps contained in A because the corresponding submatrix \hat{S} of \hat{A} must contain a row of zeros (otherwise we would have $\text{rank } A \geq r + 1$), so that $\det \hat{S} = 0$ by Theorem 2. This proves the theorem for an $m \times n$ matrix.

In particular, if A is square, $n \times n$, then $\text{rank } A = n$ if and only if A contains an $n \times n$ submatrix with nonzero determinant. But the only such submatrix can be A itself, hence $\det A \neq 0$. ■

Cramer's Rule

Theorem 3 opens the way to the classical solution formula for linear systems known as **Cramer's rule**², which gives solutions as quotients of determinants. **Cramer's rule is not practical in computations** (for which the methods in Secs. 7.3 and 20.1–20.3 are suitable), but is of **theoretical interest** in differential equations (Secs. 2.10, 3.3) and other theories that have engineering applications.

THEOREM 4

Cramer's Theorem (Solution of Linear Systems by Determinants)

(a) If a linear system of n equations in the same number of unknowns x_1, \dots, x_n

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$\dots \dots \dots$$

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n$$

(6)

has a nonzero coefficient determinant $D = \det A$, the system has precisely one solution. This solution is given by the formulas

$$(7) \quad x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad \dots, \quad x_n = \frac{D_n}{D} \quad \text{(Cramer's rule)}$$

where D_k is the determinant obtained from D by replacing in D the k th column by the column with the entries b_1, \dots, b_n .

(b) Hence if the system (6) is homogeneous and $D \neq 0$, it has only the trivial solution $x_1 = 0, x_2 = 0, \dots, x_n = 0$. If $D = 0$, the homogeneous system also has nontrivial solutions.

²GABRIEL CRAMER (1704–1752), Swiss mathematician.

PROOF The augmented matrix $\tilde{\mathbf{A}}$ of the system (6) is of size $n \times (n + 1)$. Hence its rank can be at most n . Now if

$$(8) \quad D = \det \mathbf{A} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \neq 0,$$

then $\text{rank } \mathbf{A} = n$ by Theorem 3. Thus $\text{rank } \tilde{\mathbf{A}} = \text{rank } \mathbf{A}$. Hence, by the Fundamental Theorem in Sec. 7.5, the system (6) has a unique solution.

Let us now prove (7). Expanding D by its k th column, we obtain

$$(9) \quad D = a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk},$$

where C_{ik} is the cofactor of entry a_{ik} in D . If we replace the entries in the k th column of D by any other numbers, we obtain a new determinant, say, \hat{D} . Clearly, its expansion by the k th column will be of the form (9), with a_{1k}, \dots, a_{nk} replaced by those new numbers and the cofactors C_{ik} as before. In particular, if we choose as new numbers the entries a_{1l}, \dots, a_{nl} of the l th column of D (where $l \neq k$), we have a new determinant \hat{D} which has twice the column $[a_{1l} \ \cdots \ a_{nl}]^T$, once as its l th column, and once as its k th column because of the replacement. Hence $\hat{D} = 0$ by Theorem 2(f). If we now expand \hat{D} by the column that has been replaced (the k th column), we thus obtain

$$(10) \quad a_{1l}C_{1k} + a_{2l}C_{2k} + \cdots + a_{nl}C_{nk} = 0 \quad (l \neq k).$$

We now multiply the first equation in (6) by C_{1k} on both sides, the second by C_{2k}, \dots , the last by C_{nk} , and add the resulting equations. This gives

$$(11) \quad C_{1k}(a_{11}x_1 + \cdots + a_{1n}x_n) + \cdots + C_{nk}(a_{n1}x_1 + \cdots + a_{nn}x_n) = b_1C_{1k} + \cdots + b_nC_{nk}.$$

Collecting terms with the same x_j , we can write the left side as

$$x_1(a_{11}C_{1k} + a_{21}C_{2k} + \cdots + a_{n1}C_{nk}) + \cdots + x_n(a_{1n}C_{1k} + a_{2n}C_{2k} + \cdots + a_{nn}C_{nk}).$$

From this we see that x_k is multiplied by

$$a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk} = D$$

Equation (9) shows that this equals D . Similarly, x_l is multiplied by

$$a_{1l}C_{1k} + a_{2l}C_{2k} + \cdots + a_{nl}C_{nk}.$$

Equation (10) shows that this is zero when $l \neq k$. Accordingly, the left side of (11) equals simply $x_k D$, so that (11) becomes

$$x_k D = b_1 C_{1k} + b_2 C_{2k} + \cdots + b_n C_{nk}.$$

Handwritten notes:
 k-th column expansion
 k 번째 열의 모든 원소 곱하기
 l 번째 열의 모든 원소 곱하기

Handwritten note:
 D_k : k column 이
 l 번째 열의 모든 원소 곱하기

Now the right side of this is (D_k) as defined in the theorem, expanded by its k th column, so that division by D gives (7). This proves Cramer's rule.

If (6) is homogeneous and $D \neq 0$, then each D_k has a column of zeros, so that $D_k = 0$ by Theorem 2(e), and (7) gives the trivial solution.

Finally, if (6) is homogeneous and $D = 0$, then $\text{rank } A < n$ by Theorem 3, so that nontrivial solutions exist by Theorem 2 in Sec. 7.5. ■

Illustrations of Theorem 4 for $n = 2$ and 3 are given in Sec. 7.6, and an important application of the present formulas will follow in the next section.

H.W. 4, 20

PROBLEM SET 7.7

- (Second-order determinant)** Expand a general second-order determinant in four possible ways and show that the results agree.
- (Minors, cofactors)** Complete the list of minors and cofactors in Example 1.
- (Third-order determinant)** Do the task indicated in Example 2. Also evaluate D by reduction to triangular form.
- (Scalar multiplication)** Show that $\det(kA) = k^n \det A$ (not $k \det A$), where A is any $n \times n$ matrix. Give an example.

$$15. \begin{vmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{vmatrix} \quad 16. \begin{vmatrix} 0 & -2 & 1 & 0 \\ 2 & 0 & -2 & 4 \\ -1 & 2 & 0 & 1 \\ 0 & -4 & -1 & 0 \end{vmatrix}$$

- (Expansion numerically impractical)** Show that the computation of an n th-order determinant by expansion involves $n!$ multiplications, which if a multiplication takes 10^{-9} sec would take these times:

n	10	15	20	25
Time	0.004 sec	22 min	77 years	$0.5 \cdot 10^9$ years

5-16 EVALUATION OF DETERMINANTS

Evaluate, showing the details of your work.

- $\begin{vmatrix} 13 & 8 \\ -2 & 7 \end{vmatrix}$
- $\begin{vmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{vmatrix}$
- $\begin{vmatrix} \cos \alpha & \sin \alpha \\ \sin \beta & \cos \beta \end{vmatrix}$
- $\begin{vmatrix} 14 & 2 & 5 \\ 2 & 0 & 8 \\ 5 & 8 & -2 \end{vmatrix}$
- $\begin{vmatrix} 70.4 & 0.3 & 0.8 \\ 0 & 0.5 & 2.6 \\ 0 & 0 & -1.9 \end{vmatrix}$
- $\begin{vmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{vmatrix}$
- $\begin{vmatrix} 0 & 3 & -1 \\ -3 & 0 & -4 \\ 1 & 4 & 0 \end{vmatrix}$
- $\begin{vmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{vmatrix}$
- $\begin{vmatrix} u & v & w \\ w & u & v \\ v & w & u \end{vmatrix}$
- $\begin{vmatrix} 1 & -2 & 0 & 0 \\ 4 & 3 & 5 & 0 \\ 0 & 2 & 7 & 5 \\ 0 & 0 & 2 & 4 \end{vmatrix}$

18-20 CRAMER'S RULE

Solve by Cramer's rule and check by Gauss elimination and back substitution. (Show details.)

- $2x - 5y = 23$
 $4x + 6y = -2$
- $3y + 4z = 14.8$
 $4x + 2y - z = -6.3$
 $x - y + 5z = 13.5$
- $w + 2x - 3z = 30$
 $4x - 5y + 2z = 13$
 $2w + 8x - 4y + z = 42$
 $3w + y - 5z = 35$

21-23 RANK BY DETERMINANTS

Find the rank by Theorem 3 (which is not a very practical way) and check by row reduction. (Show details.)

21.
$$\begin{bmatrix} 8 & 4 \\ -2 & -1 \\ 6 & 3 \end{bmatrix}$$

22.
$$\begin{bmatrix} 2 & 1 & 0 \\ 13 & -13 & 12 \\ -3 & 5 & -4 \end{bmatrix}$$

23.
$$\begin{bmatrix} 0.4 & 0 & -2.4 & 3.0 \\ 1.2 & 0.6 & 0 & 0.3 \\ 0 & 1.2 & 1.2 & 0 \end{bmatrix}$$

24. TEAM PROJECT. Geometrical Applications: Curves and Surfaces Through Given Points. The idea is to get an equation from the vanishing of the determinant of a homogeneous linear system as the condition for a nontrivial solution in Cramer’s theorem. We explain the trick for obtaining such a system for the case of a line L through two given points $P_1: (x_1, y_1)$ and $P_2: (x_2, y_2)$. The unknown line is $ax + by = -c$, say. We write it as $ax + by + c \cdot 1 = 0$. To get a nontrivial solution a, b, c , the determinant of the “coefficients” $x, y, 1$ must be zero. The system is

$$ax + by + c \cdot 1 = 0 \quad (\text{Line } L)$$

$$(12) \quad ax_1 + by_1 + c \cdot 1 = 0 \quad (P_1 \text{ on } L)$$

$$ax_2 + by_2 + c \cdot 1 = 0 \quad (P_2 \text{ on } L).$$

(a) **Line through two points.** Derive from $D = 0$ in (12) the familiar formula

$$\frac{x - x_1}{x_1 - x_2} = \frac{y - y_1}{y_1 - y_2}$$

(b) **Plane.** Find the analog of (12) for a plane through three given points. Apply it when the points are $(1, 1, 1)$, $(3, 2, 6)$, $(5, 0, 5)$.

(c) **Circle.** Find a similar formula for a circle in the plane through three given points. Find and sketch the circle through $(2, 6)$, $(6, 4)$, $(7, 1)$.

(d) **Sphere.** Find the analog of the formula in (c) for a sphere through four given points. Find the sphere through $(0, 0, 5)$, $(4, 0, 1)$, $(0, 4, 1)$, $(0, 0, -3)$ by this formula or by inspection.

(e) **General conic section.** Find a formula for a general conic section (the vanishing of a determinant of 6th order). Try it out for a quadratic parabola and for a more general conic section of your own choice.

25. WRITING PROJECT. General Properties of Determinants. Illustrate each statement in Theorems 1 and 2 with an example of your choice.

26. CAS EXPERIMENT. Determinant of Zeros and Ones. Find the value of the determinant of the $n \times n$ matrix A_n with main diagonal entries all 0 and all others 1. Try to find a formula for this. Try to prove it by induction. Interpret A_3 and A_4 as “incidence matrices” (as in Problem Set 7.1 but without the minuses) of a triangle and a tetrahedron, respectively; similarly for an “ n -simplex”, having n vertices and $n(n - 1)/2$ edges (and spanning R^{n-1} , $n = 5, 6, \dots$).

7.8 Inverse of a Matrix. Gauss–Jordan Elimination

In this section we consider square matrices exclusively.

The **inverse** of an $n \times n$ matrix $A = [a_{jk}]$ is denoted by A^{-1} and is an $n \times n$ matrix such that

(1)

$$AA^{-1} = A^{-1}A = I$$

where I is the $n \times n$ unit matrix (see Sec. 7.2).

If A has an inverse, then A is called a **nonsingular matrix**. If A has no inverse, then A is called a **singular matrix**.

If A has an inverse, the inverse is unique.

Indeed, if both B and C are inverses of A , then $AB = I$ and $CA = I$, so that we obtain the uniqueness from

Uniqueness

$$B = IB = (CA)B = C(AB) = CI = C.$$

We prove next that \mathbf{A} has an inverse (is nonsingular) if and only if it has maximum possible rank n . The proof will also show that $\mathbf{Ax} = \mathbf{b}$ implies $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ provided \mathbf{A}^{-1} exists, and will thus give a motivation for the inverse as well as a relation to linear systems. (But this will not give a good method of solving $\mathbf{Ax} = \mathbf{b}$ numerically because the Gauss elimination in Sec. 7.3 requires fewer computations.)

THEOREM 1**Existence of the Inverse**

The inverse \mathbf{A}^{-1} of an $n \times n$ matrix \mathbf{A} exists if and only if rank $\mathbf{A} = n$, thus (by Theorem 3, Sec. 7.7) if and only if $\det \mathbf{A} \neq 0$. Hence \mathbf{A} is nonsingular if rank $\mathbf{A} = n$, and is singular if rank $\mathbf{A} < n$.

PROOF Let \mathbf{A} be a given $n \times n$ matrix and consider the linear system

$$(2) \quad \mathbf{Ax} = \mathbf{b}.$$

If the inverse \mathbf{A}^{-1} exists, then multiplication from the left on both sides and use of (1) gives

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

This shows that (2) has a unique solution \mathbf{x} . Hence \mathbf{A} must have rank n by the Fundamental Theorem in Sec. 7.5.

Conversely, let rank $\mathbf{A} = n$. Then by the same theorem, the system (2) has a unique solution \mathbf{x} for any \mathbf{b} . Now the back substitution following the Gauss elimination (Sec. 7.3) shows that the components x_j of \mathbf{x} are linear combinations of those of \mathbf{b} . Hence we can write

$$(3) \quad \mathbf{x} = \mathbf{Bb}$$

with \mathbf{B} to be determined. Substitution into (2) gives

$$\mathbf{Ax} = \mathbf{A}(\mathbf{Bb}) = (\mathbf{AB})\mathbf{b} = \mathbf{Cb} = \mathbf{b} \quad (\mathbf{C} = \mathbf{AB})$$

for any \mathbf{b} . Hence $\mathbf{C} = \mathbf{AB} = \mathbf{I}$, the unit matrix. Similarly, if we substitute (2) into (3) we get

$$\mathbf{x} = \mathbf{Bb} = \mathbf{B}(\mathbf{Ax}) = (\mathbf{BA})\mathbf{x}$$

for any \mathbf{x} (and $\mathbf{b} = \mathbf{Ax}$). Hence $\mathbf{BA} = \mathbf{I}$. Together, $\mathbf{B} = \mathbf{A}^{-1}$ exists. ■

³WILHELM JORDAN (1842–1899), German mathematician and geodesist. [See *American Mathematical Monthly* 94 (1987), 130–142.]

We do not recommend it as a method for solving systems of linear equations, since the number of operations in addition to those of the Gauss elimination is larger than that for back substitution, which the Gauss–Jordan elimination avoids. See also Sec. 20.1.

rank of $\mathbf{A} = n$
 0) 1) 2) unique solutions
 Back Substitution

(2) 3) 4)

Determination of the Inverse by the Gauss–Jordan Method

For the practical determination of the inverse A^{-1} of a nonsingular $n \times n$ matrix A we can use the Gauss elimination (Sec. 7.3), actually a variant of it, called the **Gauss–Jordan elimination**³ (footnote of p. 316). The idea of the method is as follows.

Using A , we form n linear systems

$$Ax_{(1)} = e_{(1)}, \quad \dots, \quad Ax_{(n)} = e_{(n)}$$

where $e_{(1)}, \dots, e_{(n)}$ are the columns of the $n \times n$ unit matrix I ; thus, $e_{(1)} = [1 \ 0 \ \dots \ 0]^T$, $e_{(2)} = [0 \ 1 \ 0 \ \dots \ 0]^T$, etc. These are n vector equations in the unknown vectors $x_{(1)}, \dots, x_{(n)}$. We combine them into a single matrix equation $AX = I$, with the unknown matrix X having the columns $x_{(1)}, \dots, x_{(n)}$. Correspondingly, we combine the n augmented matrices $[A \ e_{(1)}], \dots, [A \ e_{(n)}]$ into one $n \times 2n$ “augmented matrix” $\tilde{A} = [A \ I]$. Now multiplication of $AX = I$ by A^{-1} from the left gives $X = A^{-1}I = A^{-1}$. Hence, to solve $AX = I$ for X , we can apply the Gauss elimination to $\tilde{A} = [A \ I]$. This gives a matrix of the form $[U \ H]$ with upper triangular U because the Gauss elimination triangularizes systems. The Gauss–Jordan method reduces U by further elementary row operations to diagonal form, in fact to the unit matrix I . This is done by eliminating the entries of U above the main diagonal and making the diagonal entries all 1 by multiplication (see the example below). Of course, the method operates on the entire matrix $[U \ H]$, transforming H into some matrix K , hence the entire $[U \ H]$ to $[I \ K]$. This is the “augmented matrix” of $IX = K$. Now $IX = X = A^{-1}$, as shown before. By comparison, $K = A^{-1}$, so that we can read A^{-1} directly from $[I \ K]$.

The following example illustrates the practical details of the method.

$AX = I$
 X : matrix
 $\tilde{A} = [A : I]$
 Gaussian E.
 \downarrow
 $[U : H]$
 Jordan
 \downarrow
 $[I : K]$
 $\uparrow A^{-1}$

EXAMPLE 1 Inverse of a Matrix. Gauss–Jordan Elimination

Determine the inverse A^{-1} of

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

Solution. We apply the Gauss elimination (Sec. 7.3) to the following $n \times 2n = 3 \times 6$ matrix, where BLUE always refers to the previous matrix.

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

$$\left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array} \right] \begin{array}{l} \text{Row 2} + 3 \text{ Row 1} \\ \text{Row 3} - \text{Row 1} \end{array}$$

$$\left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 0 & -5 & -4 & -1 & 1 \end{array} \right] \text{Row 3} - \text{Row 2}$$

II IH

This is $[U \ H]$ as produced by the Gauss elimination. Now follow the additional Gauss–Jordan steps, reducing U to I , that is, to diagonal form with entries 1 on the main diagonal.

$$\begin{array}{l} \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array} \right] \begin{array}{l} \text{— Row 1} \\ 0.5 \text{ Row 2} \\ -0.2 \text{ Row 3} \end{array} \\ \\ \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0.6 & 0.4 & -0.4 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array} \right] \begin{array}{l} \text{Row 1} + 2 \text{ Row 3} \\ \text{Row 2} - 3.5 \text{ Row 3} \end{array} \\ \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array} \right] \begin{array}{l} \text{Row 1} + \text{Row 2} \end{array} \end{array}$$

The last three columns constitute A^{-1} . Check:

$$\begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix} \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hence $AA^{-1} = I$. Similarly, $A^{-1}A = I$.

Useful Formulas for Inverses

The explicit formula (4) in the following theorem is often useful in theoretical studies (as opposed to *computing* inverses). In fact, the special case $n = 2$ occurs quite frequently in geometrical and other applications.

THEOREM 2

Inverse of a Matrix

The inverse of a nonsingular $n \times n$ matrix $A = [a_{jk}]$ is given by

$$(4) \quad A^{-1} = \frac{1}{\det A} [C_{jk}]^T = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix},$$

where C_{jk} is the cofactor of a_{jk} in $\det A$ (see Sec. 7.7). (**CAUTION!** Note well that in A^{-1} , the cofactor C_{jk} occupies the same place as a_{kj} (not a_{jk}) does in A .)

In particular, the inverse of

$$(4^*) \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{is} \quad A^{-1} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

PROOF We denote the right side of (4) by **B** and show that **BA** = **I**. We first write

$$(5) \quad \mathbf{BA} = \mathbf{G} = [g_{kl}]$$

and then show that **G** = **I**. Now by the definition of matrix multiplication and because of the form of **B** in (4), we obtain (**CAUTION!** C_{sk} , not C_{ks})

$$(6) \quad g_{kl} = \sum_{s=1}^n \frac{C_{sk}}{\det \mathbf{A}} a_{sl} = \frac{1}{\det \mathbf{A}} (a_{1l}C_{1k} + \cdots + a_{nl}C_{nk}).$$

$l = k \Rightarrow \det A$
 $D = \det A$

Now (9) and (10) in Sec. 7.7 show that the sum (\cdots) on the right is $D = \det \mathbf{A}$ when $l = k$, and is zero when $l \neq k$. Hence

$$g_{kk} = \frac{1}{\det \mathbf{A}} \det \mathbf{A} = 1,$$

$$g_{kl} = 0 \quad (l \neq k),$$

In particular, for $n = 2$ we have in (4) in the first row $C_{11} = a_{22}$, $C_{21} = -a_{12}$ and in the second row $C_{12} = -a_{21}$, $C_{22} = a_{11}$. This gives (4*). ■

EXAMPLE 2 Inverse of a 2×2 Matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{A}^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{bmatrix}$$

EXAMPLE 3 Further Illustration of Theorem 2

Using (4), find the inverse of

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

Solution. We obtain $\det \mathbf{A} = -1(-7) - 1 \cdot 13 + 2 \cdot 8 = 10$, and in (4),

$$C_{11} = \begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = -7, \quad C_{21} = -\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 2, \quad C_{31} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3,$$

$$C_{12} = -\begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} = -13, \quad C_{22} = \begin{vmatrix} -1 & 2 \\ -1 & 4 \end{vmatrix} = -2, \quad C_{32} = -\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} = 7,$$

$$C_{13} = \begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix} = 8, \quad C_{23} = -\begin{vmatrix} -1 & 1 \\ -1 & 3 \end{vmatrix} = 2, \quad C_{33} = \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = -2,$$

so that by (4), in agreement with Example 1,

$$\mathbf{A}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

Diagonal matrices $A = [a_{jk}]$, $a_{jk} = 0$ when $j \neq k$, have an inverse if and only if all $a_{jj} \neq 0$. Then A^{-1} is diagonal, too, with entries $1/a_{11}, \dots, 1/a_{nn}$.

PROOF For a diagonal matrix we have in (4)

$$\frac{C_{11}}{D} = \frac{a_{22} \cdots a_{nn}}{a_{11} a_{22} \cdots a_{nn}} = \frac{1}{a_{11}}, \quad \text{etc.}$$

EXAMPLE 4 Inverse of a Diagonal Matrix

Let

$$A = \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then the inverse is

$$A^{-1} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Products can be inverted by taking the inverse of each factor and multiplying these inverses *in reverse order*,

$$(7) \quad (AC)^{-1} = C^{-1}A^{-1}.$$

Hence for more than two factors,

$$(8) \quad (AC \cdots PQ)^{-1} = Q^{-1}P^{-1} \cdots C^{-1}A^{-1}.$$

PROOF The idea is to start from (1) for AC instead of A , that is, $AC(AC)^{-1} = I$, and multiply it on both sides from the left, first by A^{-1} , which because of $A^{-1}A = I$ gives

$$\begin{aligned} A^{-1}AC(AC)^{-1} &= C(AC)^{-1} \\ &= A^{-1}I = A^{-1} \end{aligned}$$

and then multiplying this on both sides from the left, this time by C^{-1} and by using $C^{-1}C = I$,

$$C^{-1}C(AC)^{-1} = (AC)^{-1} = C^{-1}A^{-1}.$$

This proves (7), and from it, (8) follows by induction.

We also note that *the inverse of the inverse is the given matrix*, as you may prove,

$$(9) \quad (A^{-1})^{-1} = A.$$

$$AC(AC)^{-1} = I$$

양변에 A^{-1} 곱함

$$C(AC)^{-1} = A^{-1}$$

양변에 C^{-1} 곱함

Unusual Properties of Matrix Multiplication. Cancellation Laws

Section 7.2 contains warnings that some properties of matrix multiplication deviate from those for numbers, and we are now able to explain the restricted validity of the so-called cancellation laws [2.] and [3.] below, using rank and inverse, concepts that were not yet available in Sec. 7.2. The deviations from the usual are of great practical importance and must be carefully observed. They are as follows.

[1.] Matrix multiplication is not commutative, that is, in general we have

$$\mathbf{AB} \neq \mathbf{BA}.$$

[2.] $\mathbf{AB} = \mathbf{0}$ does not generally imply $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$ (or $\mathbf{BA} = \mathbf{0}$); for example,

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

[3.] $\mathbf{AC} = \mathbf{AD}$ does not generally imply $\mathbf{C} = \mathbf{D}$ (even when $\mathbf{A} \neq \mathbf{0}$).

Complete answers to [2.] and [3.] are contained in the following theorem.

THEOREM 3

Cancellation Laws

Let \mathbf{A} , \mathbf{B} , \mathbf{C} be $n \times n$ matrices. Then:

- (a) If $\text{rank } \mathbf{A} = n$ and $\mathbf{AB} = \mathbf{AC}$, then $\mathbf{B} = \mathbf{C}$. full rank
- (b) If $\text{rank } \mathbf{A} = n$, then $\mathbf{AB} = \mathbf{0}$ implies $\mathbf{B} = \mathbf{0}$. Hence if $\mathbf{AB} = \mathbf{0}$, but $\mathbf{A} \neq \mathbf{0}$ as well as $\mathbf{B} \neq \mathbf{0}$, then $\text{rank } \mathbf{A} < n$ and $\text{rank } \mathbf{B} < n$.
- (c) If \mathbf{A} is singular, so are \mathbf{BA} and \mathbf{AB} .

→ no inverse matrix

PROOF

(a) The inverse of \mathbf{A} exists by Theorem 1. Multiplication by \mathbf{A}^{-1} from the left gives $\mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{AC}$, hence $\mathbf{B} = \mathbf{C}$.

(b) Let $\text{rank } \mathbf{A} = n$. Then \mathbf{A}^{-1} exists, and $\mathbf{AB} = \mathbf{0}$ implies $\mathbf{A}^{-1}\mathbf{AB} = \mathbf{B} = \mathbf{0}$. Similarly when $\text{rank } \mathbf{B} = n$. This implies the second statement in (b).

(c₁) $\text{rank } \mathbf{A} < n$ by Theorem 1. Hence $\mathbf{Ax} = \mathbf{0}$ has nontrivial solutions by Theorem 2 in Sec. 7.5. Multiplication by \mathbf{B} shows that these solutions are also solutions of $\mathbf{BAx} = \mathbf{0}$, so that $\text{rank } (\mathbf{BA}) < n$ by Theorem 2 in Sec. 7.5 and \mathbf{BA} is singular by Theorem 1.

(c₂) \mathbf{A}^T is singular by Theorem 2(d) in Sec. 7.7. Hence $\mathbf{B}^T\mathbf{A}^T$ is singular by part (c₁), and is equal to $(\mathbf{AB})^T$ by (10d) in Sec. 7.2. Hence \mathbf{AB} is singular by Theorem 2(d) in Sec. 7.7. ■

Determinants of Matrix Products

The determinant of a matrix product \mathbf{AB} or \mathbf{BA} can be written as the product of the determinants of the factors, and it is interesting that $\det \mathbf{AB} = \det \mathbf{BA}$, although $\mathbf{AB} \neq \mathbf{BA}$ in general. The corresponding formula (10) is needed occasionally and can be obtained by Gauss–Jordan elimination (see Example 1) and from the theorem just proved.

THEOREM 4

Determinant of a Product of Matrices

For any $n \times n$ matrices A and B ,

(10) $\det(AB) = \det(BA) = \det A \det B.$

PROOF

If A or B is singular, so are AB and BA by Theorem 3(c), and (10) reduces to $0 = 0$ by Theorem 3 in Sec. 7.7.

Now let A and B be nonsingular. Then we can reduce A to a diagonal matrix $\hat{A} = [a_{jk}]$ by Gauss–Jordan steps. Under these operations, $\det A$ retains its value, by Theorem 1 in Sec. 7.7, (a) and (b) [not (c)] except perhaps for a sign reversal in row interchanging when pivoting. But the same operations reduce AB to $\hat{A}B$ with the same effect on $\det(AB)$. Hence it remains to prove (10) for $\hat{A}B$; written out,

*A는 G-J 과정을
→ 같은 행렬식인
A를 만든다*

$$\hat{A}B = \begin{bmatrix} \hat{a}_{11} & 0 & \cdots & 0 \\ 0 & \hat{a}_{22} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \hat{a}_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \ddots & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

*$\hat{a}_{11} \hat{a}_{22} \cdots \hat{a}_{nn} | B |$
 $= | \hat{A} B | = | A | | B |$*

$$= \begin{bmatrix} \hat{a}_{11}b_{11} & \hat{a}_{11}b_{12} & \cdots & \hat{a}_{11}b_{1n} \\ \hat{a}_{22}b_{21} & \hat{a}_{22}b_{22} & \cdots & \hat{a}_{22}b_{2n} \\ & & \ddots & \\ \hat{a}_{nn}b_{n1} & \hat{a}_{nn}b_{n2} & \cdots & \hat{a}_{nn}b_{nn} \end{bmatrix}$$

We now take the determinant $\det(\hat{A}B)$. On the right we can take out a factor \hat{a}_{11} from the first row, \hat{a}_{22} from the second, \cdots , \hat{a}_{nn} from the n th. But this product $\hat{a}_{11} \hat{a}_{22} \cdots \hat{a}_{nn}$ equals $\det \hat{A}$ because \hat{A} is diagonal. The remaining determinant is $\det B$. This proves (10) for $\det(AB)$, and the proof for $\det(BA)$ follows by the same idea. ■

H.W. 8, 18

This completes our discussion of linear systems (Secs. 7.3–7.8). Section 7.9 on vector spaces and linear transformations is optional. **Numeric methods** are discussed in Secs. 20.1–20.4, which are independent of other sections on numerics.

chap 7. H.W. 2, 14, 34, 41 10A 4B limit

PROBLEM SET 7.8

1–12 INVERSE

Find the inverse by Gauss–Jordan [or by (4*) if $n = 2$] or state that it does not exist. Check by using (1).

1. $\begin{bmatrix} 1.20 & 4.64 \\ 0.50 & 3.60 \end{bmatrix}$

2. $\begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix}$

3. $\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{bmatrix}$

4. $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$

5. $\begin{bmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix}$
7. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 4 & 1 \end{bmatrix}$
9. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
11. $\begin{bmatrix} 1 & 2 & 5 \\ 0 & -1 & 2 \\ 2 & 4 & 10 \end{bmatrix}$
6. $\begin{bmatrix} 29 & -11 & 10 \\ -160 & 61 & -55 \\ 55 & -21 & 19 \end{bmatrix}$
8. $\begin{bmatrix} 1 & 2 & 5 \\ 0 & -1 & 2 \\ 2 & 4 & 11 \end{bmatrix}$
10. $\begin{bmatrix} 0 & 8 & 0 \\ 0 & 0 & 4 \\ 2 & 0 & 0 \end{bmatrix}$
12. $\begin{bmatrix} 1 & 2 & -9 \\ -2 & -4 & 19 \\ 0 & -1 & 2 \end{bmatrix}$

13. (**Triangular matrix**) Is the inverse of a triangular matrix always triangular (as in Prob. 7)? Give reason.
14. (**Rotation**) Give an application of the matrix in Prob. 3 that makes the form of its inverse obvious.
15. (**Inverse of the square**) Verify $(A^2)^{-1} = (A^{-1})^2$ for A in Prob. 5.
16. Prove the formula in Prob. 15.
17. (**Inverse of the transpose**) Verify $(A^T)^{-1} = (A^{-1})^T$ for A in Prob. 5.
18. Prove the formula in Prob. 17.
19. (**Inverse of the inverse**) Prove that $(A^{-1})^{-1} = A$.
20. (**Row interchange**) Same question as in Prob. 14 for the matrix in Prob. 9.

21–23 **EXPLICIT FORMULA (4) FOR THE INVERSE**

Formula (4) is generally not very practical. To understand its use, apply it:

21. To Prob. 9. 22. To Prob. 4. 23. To Prob. 7.

7.9 Vector Spaces, Inner Product Spaces, Linear Transformations *Optional*

In Sec. 7.4 we have seen that *special vector spaces* arise quite naturally in connection with matrices and linear systems, that their elements, called *vectors*, satisfy rules quite similar to those for numbers [(3) and (4) in Sec. 7.1], and that they are often obtained as *spans* (sets of linear combinations) of finitely many given vectors. Each such vector has n real numbers as its *components*. Look this up before going on.

Now if we take *all* vectors with n real numbers as components (“*real vectors*”), we obtain the very important **real n -dimensional vector space R^n** . This is a standard name and notation. Thus, each vector in R^n is an ordered n -tuple of real numbers.

Particular cases are R^2 , the space of all ordered pairs (“**vectors in the plane**”) and R^3 , the space of all ordered triples (“**vectors in 3-space**”). These vectors have wide applications in mechanics, geometry, and calculus that are basic to the engineer and physicist.

Similarly, if we take all ordered n -tuples of *complex numbers* as vectors and complex numbers as scalars, we obtain the **complex vector space C^n** , which we shall consider in Sec. 8.5.

This is not all. There are other sets of practical interest (sets of matrices, functions, transformations, etc.) for which addition and scalar multiplication can be defined in a natural way so that they form a “vector space”. This suggests to create from the “*concrete model*” R^n the “*abstract concept*” of a “*real vector space*” V by taking the basic properties (3) and (4) in Sec. 7.1 as axioms. These axioms guarantee that one obtains a useful and applicable theory of those more general situations. Note that each axiom expresses a simple property of R^n or, as a matter of fact, of R^3 . Selecting good axioms needs experience and is a process of trial and error that often extends over a long period of time.

13. (Different bases) Find three bases for R^2 .
14. (Uniqueness) Show that the representation $\mathbf{v} = c_1\mathbf{a}_{(1)} + \cdots + c_n\mathbf{a}_{(n)}$ of any given vector in an n -dimensional vector space V in terms of a given basis $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(n)}$ for V is unique.

15–20 LINEAR TRANSFORMATIONS

Find the inverse transformation. (Show the details of your work.)

15. $y_1 = x_1 - 2x_2$ 16. $y_1 = 5x_1 - x_2$
 $y_2 = 4x_1 - 3x_2$ $y_2 = 3x_1 - x_2$
17. $y_1 = 3x_1 - x_2$ 18. $y_1 = 0.25x_1 - 0.1x_3$
 $y_2 = -5x_1 + 2x_2$ $y_2 = x_2 - 0.8x_3$
 $y_3 = 0.2x_3$
19. $y_1 = 2x_1 - 3x_2$
 $y_2 = -10x_1 + 16x_2 + x_3$
 $y_3 = -7x_1 + 11x_2 + x_3$

20. $y_1 = x_1 + x_2 - 2x_3$
 $y_2 = x_1 + x_2 + 2x_3$
 $y_3 = -2x_1 + 2x_2 + 4x_3$

21–26 INNER PRODUCT. ORTHOGONALITY

Find the Euclidean norm of the vectors

21. $[4 \ 2 \ -6]^T$
 22. $[0 \ -3 \ 3 \ 0 \ 5 \ 1]^T$
 23. $[16 \ -32 \ 0]^T$
 24. $[\frac{1}{4} \ \frac{3}{8} \ \frac{1}{2} \ 2]^T$
 25. $[0 \ 1 \ 0 \ 0 \ -1 \ 1 \ -1]^T$
 26. $[\frac{2}{3} \ -\frac{2}{3} \ \frac{1}{3}]^T$
27. (Orthogonality) Show that the vectors in Probs. 21 and 23 are orthogonal.
28. Find all vectors \mathbf{v} in R^3 orthogonal to $[2 \ 0 \ 1]^T$.
29. (Unit vectors) Find all unit vectors orthogonal to $[4 \ -3]^T$. Make a sketch.
30. (Triangle inequality) Verify (4) for the vectors in Probs. 21 and 23.

CHAPTER 7 REVIEW QUESTIONS AND PROBLEMS

1. What properties of matrix multiplication differ from those of the multiplication of numbers? What about division of matrices?
2. Let \mathbf{A} be a 50×50 matrix and \mathbf{B} a 50×20 matrix. Are the following expressions defined or not? $\mathbf{A} + \mathbf{B}$, \mathbf{A}^2 , \mathbf{B}^2 , \mathbf{AB} , \mathbf{BA} , \mathbf{AA}^T , $\mathbf{B}^T\mathbf{A}$, $\mathbf{B}^T\mathbf{B}$, \mathbf{BB}^T , $\mathbf{B}^T\mathbf{AB}$. (Give reasons.)
3. How is matrix multiplication motivated?
4. Are there any linear systems without solutions? With one solution? With more than one solution? Give simple examples.
5. How can you give the rank of a matrix in terms of row vectors? Of column vectors? Of determinants?
6. What is the role of rank in connection with solving linear systems?
7. What is the row space of a matrix? The column space? The null space?
8. What is the idea of Gauss elimination and back substitution?
9. What is the inverse of a matrix? When does it exist? How would you determine it?
10. What is Cramer's rule? When would you apply it?
11. $9x - 3y = 15$
 $5x + 4y = 48$
12. $-2x - 4y + 7z = -6$
 $x + 2y + 16z = 3$
13. $3x + 5y - 8z = 18$ 14. $5x - 10y = 2$
 $x + 2y - 3z = 6$ $3x + y = 13$
 $-x + 6y = 6$
15. $-8x + 2z = 1$ 16. $2y + z = -1$
 $6y + 4z = 3$ $2x + 3y - z = -12$
 $12x + 2y = 2$ $5x - 4y + 3z = 32$
17. $3x + 7y = 0$ 18. $-x + 4y - 2z = 1$
 $5x - 4y = 47$ $3x + 4y + 6z = 1$
 $6x + 9y = 15$ $x - 2y + 2z = -\frac{1}{2}$
19. $7x + 9y - 14z = 36$
 $-x - 3y + 2z = -12$
 $2x + y - 4z = 4$

11–19 LINEAR SYSTEMS

Find all solutions or indicate that no solution exists. (Show the details of your work.)

20–30 CALCULATIONS WITH MATRICES AND VECTORS

Calculate the following expressions (showing the details of your work) or indicate why they do not exist, when

$$A = \begin{bmatrix} 9 & 2 & 8 \\ 2 & 18 & 10 \\ 8 & 10 & 15 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 2 & 6 \\ -2 & 0 & -3 \\ -6 & 3 & 0 \end{bmatrix},$$

$$a = \begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}.$$

- 20. AB, BA
- 21. $A - A^T$
- 22. $A^2 + B^2$
- 23. $\det A, \det B, \det AB$
- 24. $AA^T, A^T A$
- 25. $0.2BB^T$
- 26. $Aa, a^T A, a^T Aa$
- 27. $a^T b, b^T a, ab^T$
- 28. $b^T Bb$
- 29. $a^T B, B^T a$
- 30. $0.1(A + A^T)(B - B^T)$

31–36 RANK

Determine the ranks of the coefficient matrix and the augmented matrix and state how many solutions the linear system will have.

- 31. In Prob. 13
- 32. In Prob. 12
- 33. In Prob. 17
- 34. In Prob. 14
- 35. In Prob. 19
- 36. In Prob. 18

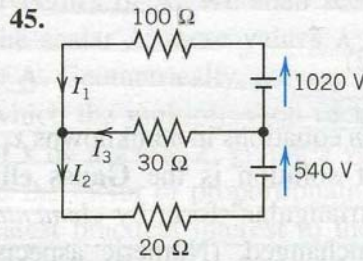
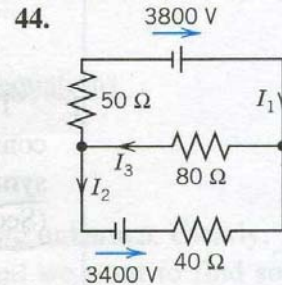
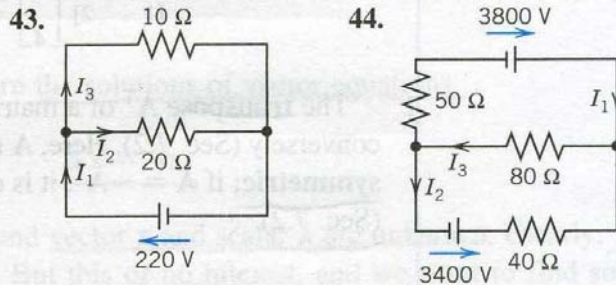
37–42 INVERSE

Find the inverse or state why it does not exist. (Show details.)

- 37. Of the coefficient matrix in Prob. 11
- 38. Of the coefficient matrix in Prob. 15
- 39. Of the coefficient matrix in Prob. 16
- 40. Of the coefficient matrix in Prob. 18
- 41. Of the augmented matrix in Prob. 14
- 42. Of the diagonal matrix with entries 3, -1, 5

43–45 NETWORKS

Find the currents in the following networks.



SUMMARY OF CHAPTER 7

Linear Algebra: Matrices, Vectors, Determinants Linear Systems of Equations

An $m \times n$ **matrix** $A = [a_{jk}]$ is a rectangular array of numbers or functions (“entries”, “elements”) arranged in m horizontal **rows** and n vertical **columns**. If $m = n$, the matrix is called **square**. A $1 \times n$ matrix is called a **row vector** and an $m \times 1$ matrix a **column vector** (Sec. 7.1).

The **sum** $A + B$ of matrices of the same **size** (i.e., both $m \times n$) is obtained by adding corresponding entries. The **product** of A by a scalar c is obtained by multiplying each a_{jk} by c (Sec. 7.1).

The **product** $C = AB$ of an $m \times n$ matrix A by an $r \times p$ matrix $B = [b_{jk}]$ is defined only when $r = n$, and is the $m \times p$ matrix $C = [c_{jk}]$ with entries

$$(1) \quad c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} \quad (\text{row } j \text{ of } A \text{ times column } k \text{ of } B).$$

This multiplication is motivated by the composition of **linear transformations** (Secs. 7.2, 7.9). It is associative, but is **not commutative**: if **AB** is defined, **BA** may not be defined, but even if **BA** is defined, **AB** \neq **BA** in general. Also **AB = 0** may not imply **A = 0** or **B = 0** or **BA = 0** (Secs. 7.2, 7.8). Illustrations:

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

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

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = [11], \quad \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}.$$

The **transpose** \mathbf{A}^T of a matrix $\mathbf{A} = [a_{jk}]$ is $\mathbf{A}^T = [a_{kj}]$; rows become columns and conversely (Sec. 7.2). Here, \mathbf{A} need not be square. If it is and $\mathbf{A} = \mathbf{A}^T$, then \mathbf{A} is called **symmetric**; if $\mathbf{A} = -\mathbf{A}^T$, it is called **skew-symmetric**. For a product, $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$ (Sec. 7.2).

A main application of matrices concerns **linear systems of equations**

$$(2) \quad \mathbf{Ax} = \mathbf{b} \quad (\text{Sec. 7.3})$$

(m equations in n unknowns x_1, \dots, x_n ; \mathbf{A} and \mathbf{b} given). The most important method of solution is the **Gauss elimination** (Sec. 7.3), which reduces the system to "triangular" form by *elementary row operations*, which leave the set of solutions unchanged. (Numeric aspects and variants, such as *Doolittle's* and *Cholesky's methods*, are discussed in Secs. 20.1 and 20.2)

Cramer's rule (Secs. 7.6, 7.7) represents the unknowns in a system (2) of n equations in n unknowns as quotients of determinants; for numeric work it is impractical. **Determinants** (Sec. 7.7) have decreased in importance, but will retain their place in eigenvalue problems, elementary geometry, etc.

The **inverse** \mathbf{A}^{-1} of a square matrix satisfies $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. It exists if and only if $\det \mathbf{A} \neq 0$. It can be computed by the *Gauss-Jordan elimination* (Sec. 7.8).

The **rank** r of a matrix \mathbf{A} is the maximum number of linearly independent rows or columns of \mathbf{A} or, equivalently, the number of rows of the largest square submatrix of \mathbf{A} with nonzero determinant (Secs. 7.4, 7.7).

The system (2) has solutions if and only if $\text{rank } \mathbf{A} = \text{rank } [\mathbf{A} \ \mathbf{b}]$, where $[\mathbf{A} \ \mathbf{b}]$ is the **augmented matrix** (Fundamental Theorem, Sec. 7.5).

The **homogeneous system**

$$(3) \quad \mathbf{Ax} = \mathbf{0}$$

has solutions $\mathbf{x} \neq \mathbf{0}$ ("nontrivial solutions") if and only if $\text{rank } \mathbf{A} < n$, in the case $m = n$ equivalently if and only if $\det \mathbf{A} = 0$ (Secs. 7.6, 7.7).

Vector spaces, inner product spaces, and linear transformations are discussed in Sec. 7.9. See also Sec. 7.4.