Chapter 7:

Linear Algebra:

Matrices, Vectors, Determinants. Linear Systems

E. Kreyszig

7.1. Matrices, Vectors

Linear algebra is a fairly extensive subject that covers vectors and matrices, determinants, systems of linear equations, vector spaces and linear transformations, eigenvalue problems, and other topics.

Matrices, which are rectangular arrays of numbers or functions, and **vectors** are the main tools of linear algebra. Matrices are important because they let us express large amounts of data and functions in an organized and concise form.

$$\begin{bmatrix} 0.3 & 1 & -5 \\ 0 & -0.2 & 16 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

$$\begin{bmatrix} e^{-x} & 2x^2 \\ e^{6x} & 4x \end{bmatrix}, \quad [a_1 \quad a_2 \quad a_3], \quad \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}$$

Example: Linear Systems, a Major Application of Matrices

We are given a system of linear equations, briefly a **linear system**, such as

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

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

This means that we can just use the augmented matrix to do the calculations needed to solve the system.

General Concepts and Notations

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

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

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 $\mathbf{A}$ an $n \times n$ **square matrix**. Then its diagonal containing the entries $a_{11}, a_{22}, \cdots, a_{nn}$ is called the **main diagonal** of $\mathbf{A}$. 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 of any size $m \times n$ is called a **rectangular matrix**; this includes square matrices as a special case.

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 $\mathbf{a}$, $\mathbf{b}, \cdots$ or by its general component in brackets, $\mathbf{a} = [a_j]$, and so on. Our special vectors in (1) suggest that a (general) **row vector** is of the form

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

A **column vector** is of the form

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

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

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

Then

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

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

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

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

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

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×n, namely,

$$
\textbf{(3)} \quad
\begin{aligned}
&\text{(a)} \quad && \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \\
&\text{(b)} \quad && (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \\
&\text{(c)} \quad && \mathbf{A} + \mathbf{0} = \mathbf{A} \\
&\text{(d)} \quad && \mathbf{A} + (-\mathbf{A}) = \mathbf{0}.
\end{aligned}
$$

Matrix addition is

commutative and *associative*

Here $\mathbf{0}$ denotes the **zero matrix** (of size $m \times n$), that is, the $m \times n$ matrix with all entries zero. If $m = 1$ or $n = 1$, this is a vector, called a **zero vector**.

$$
\textbf{(4)} \quad
\begin{aligned}
&\text{(a)} \quad && c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B} \\
&\text{(b)} \quad && (c + k)\mathbf{A} = c\mathbf{A} + k\mathbf{A} \\
&\text{(c)} \quad && c(k\mathbf{A}) = (ck)\mathbf{A} \qquad \text{(written } ck\mathbf{A}\text{)} \\
&\text{(d)} \quad && 1\mathbf{A} = \mathbf{A}.
\end{aligned}
$$

7.2. Matrix Multiplication

DEFINITION

Multiplication of a Matrix by a Matrix

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

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

The condition $r = n$ means that the second factor, **B**, must have as many rows as the first factor has columns, namely n. A diagram of sizes that shows when matrix multiplication is possible is as follows:

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

The entry c_{jk} in (1) is obtained by multiplying each entry in the jth row of $\mathbf{A}$ by the corresponding entry in the kth column of $\mathbf{B}$ 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***. For $n = 3$, this is illustrated by

$$
m = 4 \left\{ \vphantom{\begin{bmatrix} a\\a\\a\\a \end{bmatrix}} \right. \overbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}}^{n=3} \overbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}}^{p=2} = \overbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix}}^{p=2} \left. \vphantom{\begin{bmatrix} a\\a\\a\\a \end{bmatrix}} \right\} m = 4
$$

Notations in a product $\mathbf{AB} = \mathbf{C}$

EXAMPLE 1

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

Matrix Multiplication

$$
\mathbf{AB} = \begin{bmatrix} 3 & 5 & -1 \\ 4 & 0 & 2 \\ -6 & -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 & 1 \\ 5 & 0 & 7 & 8 \\ 9 & -4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 22 & -2 & 43 & 42 \\ 26 & -16 & \boxed{14} & 6 \\ -9 & 4 & -37 & -28 \end{bmatrix}
$$

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 $\mathbf{BA}$ is not defined.

EXAMPLE 2

Multiplication of a Matrix and a Vector

$$\begin{bmatrix} 4 & 2 \\ 1 & 8 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \cdot 3 + 2 \cdot 5 \\ 1 \cdot 3 + 8 \cdot 5 \end{bmatrix} = \begin{bmatrix} 22 \\ 43 \end{bmatrix} \qquad \text{whereas} \qquad \begin{bmatrix} 3 \\ 5 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ 1 & 8 \end{bmatrix} \quad \text{is undefined.}$$

■

EXAMPLE 3

Products of Row and Column Vectors

$$\begin{bmatrix} 3 & 6 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 19 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}\begin{bmatrix} 3 & 6 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 1 \\ 6 & 12 & 2 \\ 12 & 24 & 4 \end{bmatrix}.$$

■

EXAMPLE 4

CAUTION! Matrix Multiplication Is Not Commutative, AB ≠ 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 in matrix products ***the order of factors must always be observed very carefully***. Otherwise matrix multiplication satisfies rules similar to those for numbers, namely.

$$
\text{(2)}\quad
\begin{array}{lll}
 & \text{(a)} & (k\mathbf{A})\mathbf{B} = k(\mathbf{AB}) = \mathbf{A}(k\mathbf{B}) \quad \textit{written } k\mathbf{AB} \textit{ or } \mathbf{A}k\mathbf{B} \\
\text{associative law} & \text{(b)} & \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} \quad \textit{written } \mathbf{ABC} \\
 & \text{(c)} & (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \\
\text{distributive laws} & \text{(d)} & \mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}
\end{array}
$$

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

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

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

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,

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

Columns of $\mathbf{B}$ are then assigned to different processors (individually or several to each processor), which simultaneously compute the columns of the product matrix $\mathbf{Ab}_1$, $\mathbf{Ab}_2$, 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 $\mathbf{AB}$ and then write them as a single matrix, as shown in the first formula on the right.

Transposition

We obtain the transpose of a matrix by writing its rows as columns (or equivalently its columns as rows). This also applies to the transpose of vectors. Thus, a row vector becomes a column vector and vice versa. In addition, for square matrices, we can also "reflect" the elements along the main diagonal, that is, interchange entries that are symmetrically positioned with respect to the main diagonal to obtain the transpose.

EXAMPLE 7 Transposition of Matrices and Vectors

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

$$\begin{bmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 7 \\ 8 & -1 & 5 \\ 1 & -9 & 4 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 3 & 8 & 1 \\ 0 & -1 & -9 \\ 7 & 5 & 4 \end{bmatrix},$$

$$[6 \quad 2 \quad 3]^{\mathsf{T}} = \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix} \cdot \qquad \text{Conversely,} \qquad \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix}^{\mathsf{T}} = [6 \quad 2 \quad 3].$$

DEFINITION

Transposition of Matrices and Vectors

The transpose of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the $n \times m$ matrix $\mathbf{A}^\mathsf{T}$ (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 $\mathbf{A}^\mathsf{T} = [a_{kj}]$, written out

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

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

Transposition gives us a choice in that we can work either with the matrix or its transpose, whichever is more convenient.

Rules for transposition are

$$
\textbf{(10)} \qquad
\begin{aligned}
&\text{(a)} \quad & (\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} &= \mathbf{A} \\
&\text{(b)} \quad & (\mathbf{A} + \mathbf{B})^{\mathsf{T}} &= \mathbf{A}^{\mathsf{T}} + \mathbf{B}^{\mathsf{T}} \\
&\text{(c)} \quad & (c\mathbf{A})^{\mathsf{T}} &= c\mathbf{A}^{\mathsf{T}} \\
&\text{(d)} \quad & (\mathbf{A}\mathbf{B})^{\mathsf{T}} &= \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}.
\end{aligned}
$$

CAUTION! Note that in (10d) the transposed matrices are ***in reversed order*.** We leave the proofs as an exercise in Probs. 9 and 10.

Special Matrices

Symmetric and Skew-Symmetric Matrices:

- **Symmetric** matrices are square matrices whose transpose equals the matrix itself.
- **Skew-symmetric** matrices are square matrices whose transpose equals ***minus*** the matrix.

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

Symmetric Matrix Skew-Symmetric Matrix

EXAMPLE 8 Symmetric and Skew-Symmetric Matrices

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

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.
- **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

$$\underbrace{\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 6 \end{bmatrix}}_{\text{Upper triangular}}, \quad \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 8 & -1 & 0 \\ 7 & 6 & 8 \end{bmatrix}, \begin{bmatrix} 3 & 0 & 0 & 0 \\ 9 & -3 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 9 & 3 & 6 \end{bmatrix}}_{\text{Lower triangular}}.$$

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 $\mathbf{S}$ are equal, say, c, we call $\mathbf{S}$ a **scalar matrix** because multiplication of any square matrix $\mathbf{A}$ of the same size by $\mathbf{S}$ has the same effect as the multiplication by a scalar, that is,

$$\mathbf{AS} = \mathbf{SA} = c\mathbf{A}. \tag{12}$$

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 $\mathbf{I}_n$ or simply by $\mathbf{I}$. For $\mathbf{I}$, formula (12) becomes

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}. \tag{13}$$

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

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

7.3. Linear Systems of Equations

Linear System, Coefficient Matrix, Augmented Matrix

A **linear system of m equations in n unknowns** $x_1, \cdots, x_n$ is a set of equations of the form

$$
\text{(1)} \qquad
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots&\cdots\cdots \\
a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}
$$

The system is called *linear* because each variable x_j appears in the first power only, just as in the equation of a straight line. $a_{11}, \cdots, a_{mn}$ are given numbers, called the **coefficients** of the system. $b_1, \cdots, b_m$ on the right are also given numbers. If all the b_j are zero, then (1) is called a **homogeneous system**. If at least one b_j is not zero, then (1) is called a **nonhomogeneous system**.

A **solution** of (1) is a set of numbers $x_1, \cdots, x_n$ that satisfies all the m equations. A **solution vector** of (1) is a vector $\mathbf{x}$ whose components form a solution of (1). If the system (1) is homogeneous, it always has at least the **trivial solution** $x_1 = 0, \cdots, x_n = 0$.

Matrix Form of the Linear System (1). From the definition of matrix multiplication we see that the m equations of (1) may be written as a single vector equation

$$\mathbf{Ax} = \mathbf{b} \tag{2}$$

where the **coefficient matrix** $\mathbf{A} = [a_{jk}]$ is the $m \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

are column vectors. We assume that the coefficients a_{jk} are not all zero, so that $\mathbf{A}$ is not a zero matrix. Note that $\mathbf{x}$ has n components, whereas $\mathbf{b}$ has m components.

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

is called the **augmented matrix** of the system (1). The dashed vertical line could be omitted, as we shall do later. It is merely a reminder that the last column of $\tilde{\mathbf{A}}$ did not come from matrix **A** but came from vector **b**. Thus, we *augmented* the matrix **A**.

Note that ***the augmented matrix*** $\tilde{\mathbf{A}}$ ***determines the system*** **(1)** ***completely*** because it contains all the given numbers appearing in (1).

EXAMPLE 1 Geometric Interpretation. Existence and Uniqueness of Solutions

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

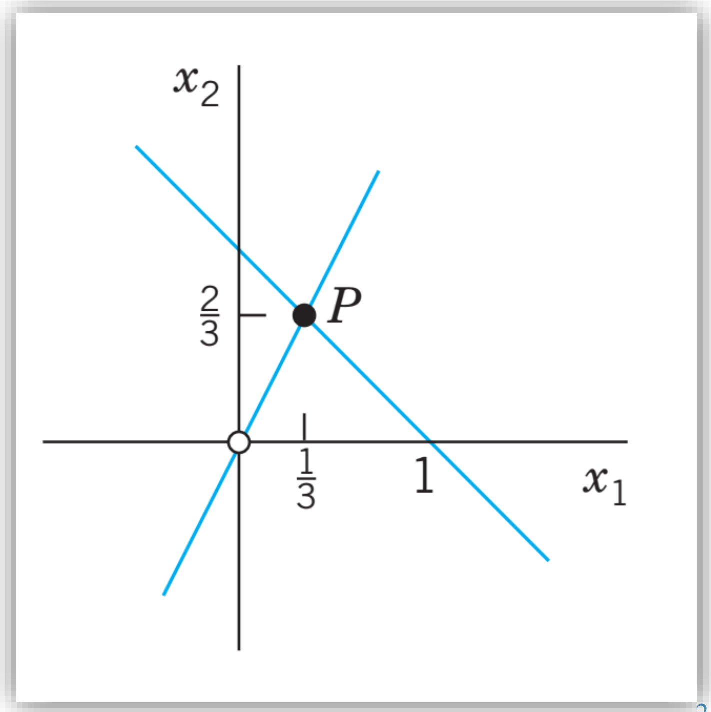
$$a_{11}x_1 + a_{12}x_2 = b_1$$

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

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 (see Fig. 158 on next page):

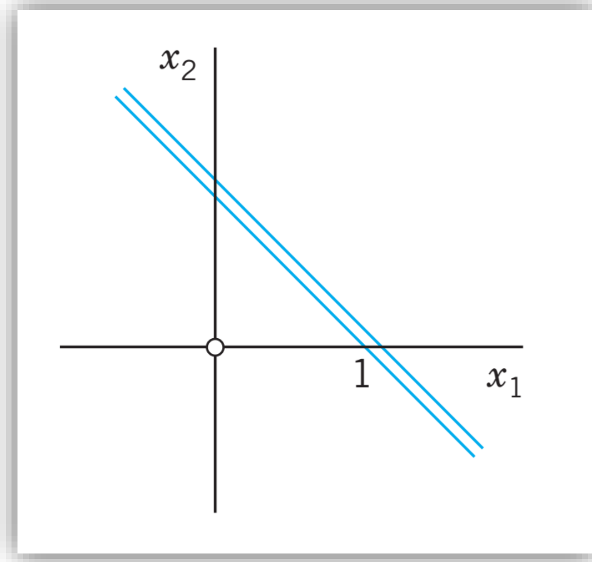
(*a*) Precisely one solution if the lines intersect

$$x_1 + x_2 = 1$$
$$2x_1 - x_2 = 0$$



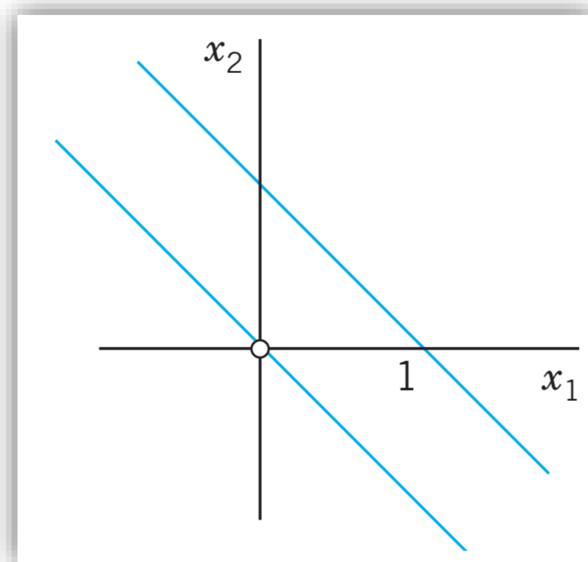
(*b*) Infinitely many solutions if the lines coincide

$$x_1 + x_2 = 1$$
$$2x_1 + 2x_2 = 2$$

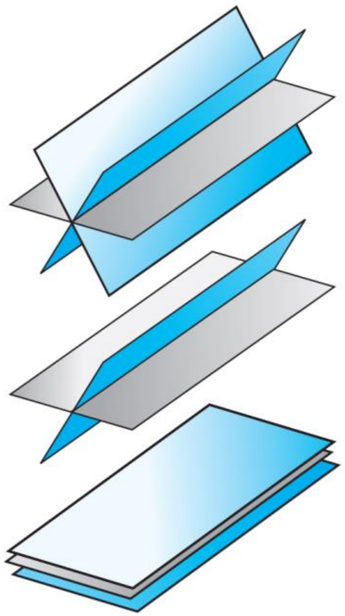


(*c*) No solution if the lines are parallel

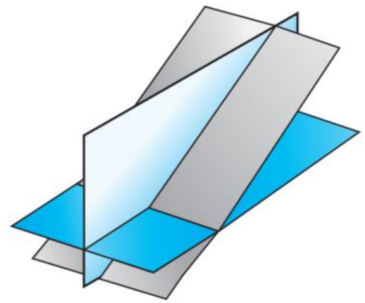
$$x_1 + x_2 = 1$$
$$x_1 + x_2 = 0$$



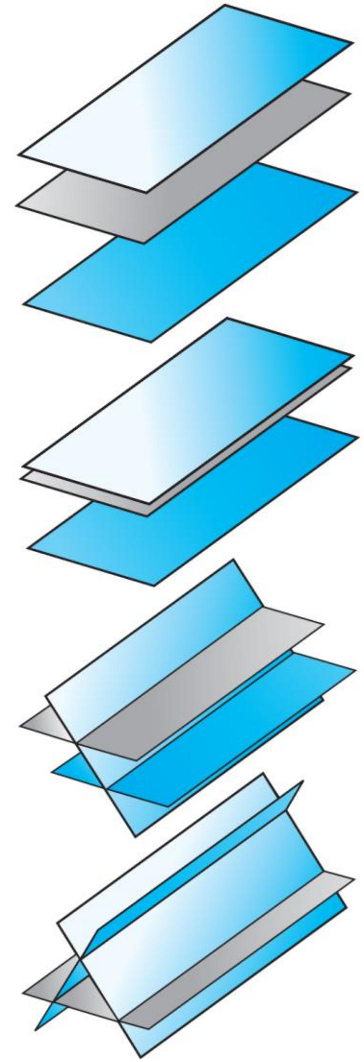
Three equations in three unknowns interpreted as planes in space



Infinitely many solutions



Unique solution



No solution

Gauss Elimination and Back Substitution

The Gauss elimination method can be motivated as follows. Consider a linear system that is in ***triangular form*** (in full, *upper* triangular form) such as

$$\boxed{\begin{aligned} 2x_1 + 5x_2 &= 2 \\ 13x_2 &= -26 \end{aligned}}$$

(*Triangular* means that all the nonzero entries of the corresponding coefficient matrix lie above the diagonal and form an upside-down 90° triangle.) Then we can solve the system by **back substitution**, that is, we solve the last equation for the variable, $x_2 = -26/13 = -2$, and then work backward, substituting $x_2 = -2$ into the first equation and solving 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

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ -4x_1 + 3x_2 &= -30. \end{aligned}$$

Its augmented matrix is

$$\begin{bmatrix} 2 & 5 & 2 \\ -4 & 3 & -30 \end{bmatrix}.$$

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ -4x_1 + 3x_2 &= -30. \end{aligned} \qquad \text{Its augmented matrix is} \qquad \begin{bmatrix} 2 & 5 & 2 \\ -4 & 3 & -30 \end{bmatrix}.$$

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,

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

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.

EXAMPLE 2 Gauss Elimination.

Solve the linear system

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 20x_1 + 10x_2 \qquad &= 80. \end{aligned}$$

Solution by Gauss Elimination.

Augmented Matrix $\tilde{\mathbf{A}}$

$$\begin{array}{l} \text{Pivot 1} \rightarrow \\ \\ \text{Eliminate} \rightarrow \\ \\ \end{array} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{array}\right]$$

Equations

$$\begin{array}{l} \text{Pivot 1} \rightarrow \\ \\ \text{Eliminate} \rightarrow \\ \\ \end{array} \begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 20x_1 + 10x_2 &= 80. \end{aligned}$$

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

$$\text{(3)} \quad \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 30 & -20 & 80 \end{array}\right] \quad \begin{array}{l} \\ \text{Row 2 + Row 1} \\ \\ \text{Row 4 − 20 Row 1} \end{array} \quad \begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 0 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 30x_2 - 20x_3 &= 80. \end{aligned}$$

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 the order of the unknowns is also changed). It gives

$$\begin{array}{l} \\ \text{Pivot 10} \rightarrow \\ \text{Eliminate 30} \rightarrow \\ \\ \end{array} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & \boxed{30} & -20 & 80 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \begin{array}{rl} & x_1 - x_2 + x_3 = 0 \\ \text{Pivot 10} \rightarrow & 10x_2 + 25x_3 = 90 \\ \text{Eliminate } 30x_2 \rightarrow & \boxed{30x_2} - 20x_3 = 80 \\ & 0 = 0. \end{array}$$

To eliminate x_2, do:

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

The result is

$$(4) \qquad \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & -95 & -190 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \begin{array}{rr} & x_1 - x_2 + x_3 = 0 \\ & 10x_2 + 25x_3 = 90 \\ \text{Row 3} - 3 \text{ Row 2} & -95x_3 = -190 \\ & 0 = 0. \end{array}$$

$$
\begin{aligned}
x_1 - \quad x_2 + \quad x_3 &= \quad 0 \\
10x_2 + 25x_3 &= \quad 90 \\
-\ 95x_3 &= -190 \\
0 &= \quad 0.
\end{aligned}
$$

$$\Downarrow$$

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:

$$
\begin{aligned}
-\ 95x_3 &= -190 & x_3 &= i_3 = 2\ [\mathrm{A}] \\
10x_2 + 25x_3 &= \quad 90 & x_2 &= \tfrac{1}{10}(90 - 25x_3) = i_2 = 4\ [\mathrm{A}] \\
x_1 - \quad x_2 + \quad x_3 &= \quad 0 & x_1 &= x_2 - x_3 = i_1 = 2\ [\mathrm{A}]
\end{aligned}
$$

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

Unique solution

Elementary Row Operations. Row-Equivalent Systems

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*

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. This justifies Gauss elimination and establishes the following result.

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,
- **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),
- **Inconsistent** if it has no solutions at all.

Gauss Elimination:
The Three Possible Cases of Systems

EXAMPLE 3 Gauss Elimination if Infinitely Many Solutions Exist

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

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

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

$$-0.6/3.0 = -0.2 \text{ times the first equation to the second equation,}$$

$$-1.2/3.0 = -0.4 \text{ times the first equation to the third equation.}$$

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

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

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

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

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 $t_1, t_2, \cdots$. 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). ■

Infinitely many solutions

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

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

Step 1. Elimination of x_1

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

Step 2. Elimination of x_2 from the third equation gives

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & 0 & 0 & 12 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} - 6\text{ Row 2} \end{array} \qquad \begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ -\tfrac{1}{3}x_2 + \tfrac{1}{3}x_3 &= -2 \\ 0 &= 12. \end{aligned}$$

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

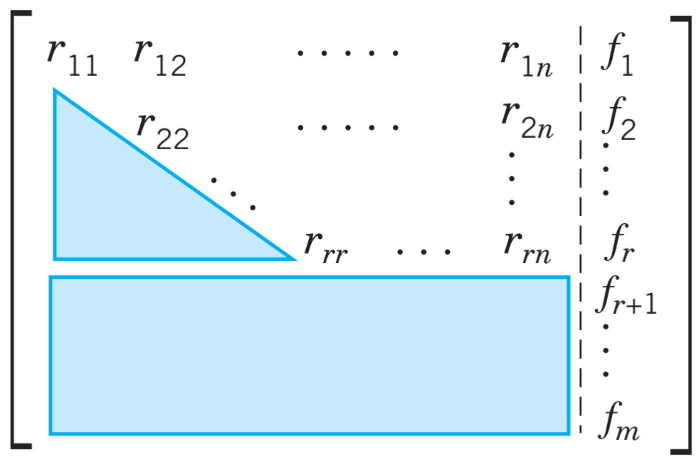
No solution

Row Echelon Form and Information From It

The original system of m equations in n unknowns has augmented matrix $[\mathbf{A}\,|\,\mathbf{b}]$. This is to be row reduced to matrix $[\mathbf{R}\,|\,\mathbf{f}]$. The two systems $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Rx} = \mathbf{f}$ are equivalent: if either one has a solution, so does the other, and the solutions are identical.

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

(9)



.

Here, $r \leqq m$, $r_{11} \neq 0$, and all entries in the blue triangle and blue rectangle are zero.

The number of nonzero rows, r, in the row-reduced coefficient matrix $\mathbf{R}$ is called the **rank of R** and also the **rank of A**. Here is the method for determining whether $\mathbf{Ax} = \mathbf{b}$ has solutions and what they are:

(a) No solution. If r is less than m (meaning that $\mathbf{R}$ actually has at least one row of all 0s) *and* at least one of the numbers $f_{r+1}, f_{r+2}, \cdots, f_m$ is not zero, then the system $\mathbf{Rx} = \mathbf{f}$ is inconsistent: No solution is possible. Therefore the system $\mathbf{Ax} = \mathbf{b}$ is inconsistent as well. See Example 4, where $r = 2 < m = 3$ and $f_{r+1} = f_3 = 12$.

If the system is consistent (either $r = m$, or $r < m$ *and* all the numbers $f_{r+1}, f_{r+2}, \cdots, f_m$ are zero), then there are solutions.

(b) Unique solution. If the system is consistent and $r = n$, there is exactly one solution, which can be found by back substitution. See Example 2, where $r = n = 3$ and $m = 4$.

(c) Infinitely many solutions. To obtain any of these solutions, choose values of $x_{r+1}, \cdots, x_n$ arbitrarily. Then solve the rth equation for x_r (in terms of those arbitrary values), then the $(r - 1)$st equation for x_{r-1}, and so on up the line. See Example 3.

7.4. Linear Independence. Rank of a Matrix. Vector Space

Linear Independence and Dependence of Vectors

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

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

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

$$\textbf{(1)} \qquad c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \cdots + c_m\mathbf{a}_{(m)} = \mathbf{0}.$$

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 $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(m)}$ 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**.

vectors **linearly dependent**. This means that we can express at least one of the vectors as a linear combination of the other vectors. For instance, if (1) holds with, say, $c_1 \neq 0$, we can solve (1) for $\mathbf{a}_{(1)}$:

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

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

EXAMPLE 1 Linear Independence and Dependence

The three vectors

$$\begin{aligned} \mathbf{a}_{(1)} &= [\;3 \quad 0 \quad 2 \quad 2\;] \\ \mathbf{a}_{(2)} &= [-6 \quad 42 \quad 24 \quad 54] \\ \mathbf{a}_{(3)} &= [\;21 \quad -21 \quad 0 \quad -15] \end{aligned}$$

are linearly dependent because

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

Although this is easily checked by vector arithmetic (do it!), it is not so easy to discover.

Rank of a Matrix

DEFINITION

The **rank** of a matrix **A** is the maximum number of linearly independent row vectors of **A**. It is denoted by rank **A**.

EXAMPLE 2 Rank

The matrix

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

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 rank $\mathbf{A} = 0$ if and only if $\mathbf{A} = \mathbf{0}$. This follows directly from the definition.

We call a matrix $\mathbf{A}_1$ **row-equivalent** to a matrix $\mathbf{A}_2$ if $\mathbf{A}_1$ can be obtained from $\mathbf{A}_2$ 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 shows 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 reducing the matrix to row-echelon form, as was done in Sec. 7.3. Once the matrix is in row-echelon form, we count the number of nonzero rows, which is precisely the rank of the matrix.

EXAMPLE 3 Determination of Rank

For the matrix in Example 2 we obtain successively

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

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

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

The last matrix is in row-echelon form and has two nonzero rows. Hence rank $\mathbf{A} = 2$, as before. ■

THEOREM 2

Linear Independence and Dependence of Vectors

Consider p vectors that each have n components. Then these vectors are linearly independent if the matrix formed, with these vectors as row vectors, has rank p. However, these vectors are linearly dependent if that matrix has rank less than p.

THEOREM 3

Rank in Terms of Column Vectors

The rank r of a matrix $\mathbf{A}$ equals the maximum number of linearly independent ***column*** *vectors of $\mathbf{A}$.*

Hence $\mathbf{A}$ and its transpose $\mathbf{A}^T$ have the same rank.

Proof in the book

THEOREM 4

Linear Dependence of Vectors

Consider p vectors each having n components. If $n < p$, then these vectors *are linearly dependent.*

Proof in the book

Vector Space

Consider a nonempty set V of vectors where each vector has the same number of components. If, for any two vectors $\mathbf{a}$ and $\mathbf{b}$ in V, we have that all their linear combinations $\alpha\mathbf{a} + \beta\mathbf{b}$ (α, β any real numbers) are also elements of V, and if, furthermore, $\mathbf{a}$ and $\mathbf{b}$ satisfy the laws (3a), (3c), (3d), and (4) in Sec. 7.1, as well as any vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ in V satisfy (3b) then V is a vector space. Note that here we wrote laws (3) and (4) of Sec. 7.1 in lowercase letters $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, which is our notation for vectors.

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;

A linearly independent set in V consisting of a maximum possible number of vectors in V is called a **basis** for V. In other words, any largest possible set of independent vectors in V forms basis for V. That means, if we add one or more vector to that set, the set will be linearly dependent. Thus, the number of vectors of a basis for V equals dim V.

The set of all linear combinations of given vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(p)}$ with the same number of components is called the **span** of these vectors. Obviously, a span is a vector space. If in addition, the given vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(p)}$ are linearly independent, then they form a basis for that vector space.

This then leads to another equivalent definition of basis. A set of vectors is a **basis** for a vector space V if (1) the vectors in the set are linearly independent, and if (2) any vector in V can be expressed as a linear combination of the vectors in the set. If (2) holds, we also say that the set of vectors **spans** the vector space V.

EXAMPLE 5 Vector Space, Dimension, Basis

The span of the three vectors in Example 1 is a vector space of dimension 2. A basis of this vector space consists of any two of those three vectors, for instance, $\mathbf{a}_{(1)}$, $\mathbf{a}_{(2)}$, or $\mathbf{a}_{(1)}$, $\mathbf{a}_{(3)}$, etc.

THEOREM 5

Vector Space R^n

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

THEOREM 6

Row Space and Column Space

The row space and the column space of a matrix $\mathbf{A}$ *have the same dimension, equal to rank* $\mathbf{A}$.

Finally, for a given matrix **A** the solution set of the homogeneous system $\mathbf{Ax} = 0$ 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)
$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = \text{Number of columns of } \mathbf{A}.$$

7.5. Solutions of Linear Systems: Existence, Uniqueness

Rank, as just defined, gives complete information about existence, uniqueness, and general structure of the solution set of linear systems as follows.

A linear system of equations in n unknowns has a

- Unique solution if the coefficient matrix and the augmented matrix have the same rank n
- Infinitely many solutions if that common rank is less than n
- No solution if those two matrices have different rank

THEOREM 1

Fundamental Theorem for Linear Systems

(a) Existence. *A linear system of m equations in n unknowns $x_1, \cdots, x_n$*

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

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

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

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

Proof in the book

In Example 2 there is a unique solution since rank $\widetilde{\mathbf{A}} = \text{rank } \mathbf{A} = n = 3$ (as can be seen from the last matrix in the example). In Example 3 we have rank $\widetilde{\mathbf{A}} = \text{rank } \mathbf{A} = 2 < n = 4$ and can choose x_3 and x_4 arbitrarily. In Example 4 there is no solution because rank $\mathbf{A} = 2 < \text{rank } \widetilde{\mathbf{A}} = 3$.

Homogeneous Linear System

Recall from Sec. 7.3 that a linear system (1) is called **homogeneous** if all the b_j's are zero, and **nonhomogeneous** if one or several b_j's are not zero.

THEOREM 2

Homogeneous Linear System

A homogeneous linear system

(4)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot& \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

always has the **trivial solution** $x_1 = 0, \cdots, x_n = 0$. *Nontrivial solutions exist if and only if rank* $\mathbf{A} < n$. *If* rank $\mathbf{A} = r < n$, *these solutions, together with* $\mathbf{x} = \mathbf{0}$, *form a vector space* (see Sec. 7.4) *of dimension* $n - r$ called the **solution space** of (4).

In particular, if $\mathbf{x}_{(1)}$ *and* $\mathbf{x}_{(2)}$ *are solution vectors of* (4), *then* $\mathbf{x} = c_1\mathbf{x}_{(1)} + c_2\mathbf{x}_{(2)}$ *with any scalars* c_1 *and* c_2 *is a solution vector of* (4). (This ***does not hold*** for nonhomogeneous systems. Also, the term *solution space is used for homogeneous* systems only.)

Proof in the book

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

$$\text{rank}\,\mathbf{A} + \text{nullity}\,\mathbf{A} = n \tag{5}$$

where n is the number of unknowns (number of columns of $\mathbf{A}$).

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

THEOREM 3

Homogeneous Linear System with Fewer Equations Than Unknowns

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

Nonhomogeneous Linear System

THEOREM 4

Nonhomogeneous Linear System

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

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h \tag{6}$$

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). ■

7.6. For Reference: Second- and Third-Order Determinants

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

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

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

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

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

is

(3) $$\boxed{x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{D} = \frac{b_1a_{22} - a_{12}b_2}{D}} \qquad \boxed{x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{D} = \frac{a_{11}b_2 - b_1a_{21}}{D}}$$

with D as in (1), provided $D \neq 0$.

EXAMPLE 1 Cramer's Rule for Two Equations

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

Third-Order Determinants

A **determinant of third order** can be defined by

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

$$\Downarrow$$

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

Cramer's Rule for Linear Systems of Three Equations

$$\boxed{\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}} \tag{5}$$

is

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

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).

7.7. Determinants. Cramer's Rule

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

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

For $n = 1$, this determinant is defined by

$$D = a_{11}. \tag{2}$$

For $n \geqq 2$ by

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

or

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

Here,

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

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 jth row and the kth column.

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

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

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

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

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

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**

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

■

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

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 no longer gives an elementary row operation.)

Proof in the book

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

EXAMPLE 4 Evaluation of Determinants by Reduction to Triangular Form

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

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

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -0 & 47.25 \end{vmatrix} \begin{array}{l} \\ \\ \\ \text{Row 4} + 4.75\ \text{Row 3} \end{array}$$

$$= 2 \cdot 5 \cdot 2.4 \cdot 47.25 = 1134.$$

THEOREM 2

Proof in the book

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.*

THEOREM 3

Rank in Terms of Determinants

Consider an $m \times n$ *matrix* $\mathbf{A} = [a_{jk}]$:

(1) $\mathbf{A}$ *has rank* $r \geqq 1$ *if and only if* $\mathbf{A}$ *has an* $r \times r$ *submatrix with a nonzero determinant.*

(2) *The determinant of any square submatrix with more than r rows, contained in* $\mathbf{A}$ *(if such a matrix exists!) has a value equal to zero.*

Furthermore, if $m = n$, *we have:*

(3) *An* $n \times n$ *square matrix* $\mathbf{A}$ *has rank n if and only if*

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

Proof in the book

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* $x_1, \cdots, x_n$

(6)
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

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

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

where D_k *is the determinant obtained from D by replacing in D the kth column by the column with the entries* $b_1, \cdots, b_n$.

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

Proof in the book

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

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{1}$$

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

If $\mathbf{A}$ has an inverse, then $\mathbf{A}$ is called a **nonsingular matrix**. If $\mathbf{A}$ has no inverse, then $\mathbf{A}$ is called a **singular matrix**.

If $\mathbf{A}$ *has an inverse, the inverse is unique.*

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

$$\mathbf{B} = \mathbf{IB} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{CI} = \mathbf{C}.$$

We prove next that $\mathbf{A}$ has an inverse (is nonsingular) if and only if it has maximum possible rank n.

THEOREM 1

Existence of the Inverse

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

Proof in the book

Determination of the Inverse by the Gauss–Jordan Method

Using $\mathbf{A}$, we form n linear systems

$$\mathbf{A}\mathbf{x}_{(1)} = \mathbf{e}_{(1)}, \quad \cdots, \quad \mathbf{A}\mathbf{x}_{(n)} = \mathbf{e}_{(n)}$$

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

EXAMPLE 1 Finding the Inverse of a Matrix by Gauss–Jordan Elimination

Determine the inverse $\mathbf{A}^{-1}$ of

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

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.

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

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

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

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

$$\left[\begin{array}{rrr|rrr} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} -\text{Row 1} \\ 0.5 \text{ Row 2} \\ -0.2 \text{ Row 3} \end{array}$$

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

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} \text{Row 1 + Row 2} \\ \\ \\ \end{array}$$

The last three columns constitute $\mathbf{A}^{-1}$. 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 $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Similarly, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. ■

Formulas for Inverses

THEOREM 2

Inverse of a Matrix by Determinants

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

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

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

In particular, the inverse of

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

Proof in the book

EXAMPLE 2 Inverse of a 2 × 2 Matrix by Determinants

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

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),

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

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

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

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

EXAMPLE 4 Inverse of a Diagonal Matrix

$$\mathbf{A} = \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \Longrightarrow \quad \mathbf{A}^{-1} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

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

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

Hence for more than two factors,

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

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

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

Unusual Properties of Matrix Multiplication. Cancellation Laws

[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 $\mathbf{A} \neq \mathbf{0}$).

THEOREM 3

Cancellation Laws

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

(a) *If* rank $\mathbf{A} = n$ *and* $\mathbf{AB} = \mathbf{AC}$, *then* $\mathbf{B} = \mathbf{C}$.

(b) *If* 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* rank $\mathbf{A} < n$ *and* rank $\mathbf{B} < n$.

(c) *If* $\mathbf{A}$ *is singular, so are* $\mathbf{BA}$ *and* $\mathbf{AB}$.

Proof in the book

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 $\boxed{\det \mathbf{AB} = \det \mathbf{BA}}$, although $\boxed{\mathbf{AB} \neq \mathbf{BA}}$ 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* $\mathbf{A}$ *and* $\mathbf{B}$,

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det \mathbf{A} \det \mathbf{B}. \tag{10}$$

Proof in the book

7.9. Vector Spaces, Inner Product Spaces, Linear Transformations

We can generalize this idea by taking *all* vectors with n real numbers as components and obtain the very important ***real n-dimensional vector space*** R^n. The vectors are known as "real vectors." Thus, each vector in R^n is an ordered n-tuple of real numbers.

Now we can consider special values for n. For $n = 2$, we obtain R^2, the vector space of all ordered pairs, which correspond to the **vectors in the plane**. For $n = 3$, we obtain R^3, the vector space of all ordered triples, which are the **vectors in 3-space**. These vectors have wide applications in mechanics, geometry, and calculus and are basic to the engineer and physicist.

DEFINITION

Real Vector Space

A nonempty set V of elements $\mathbf{a}$, $\mathbf{b}$, $\cdots$ is called a **real vector space** (or *real linear space*), and these elements are called **vectors** (regardless of their nature, which will come out from the context or will be left arbitrary) if, in V, there are defined two algebraic operations (called *vector addition* and *scalar multiplication*) as follows.

I. Vector addition associates with every pair of vectors $\mathbf{a}$ and $\mathbf{b}$ of V a unique vector of V, called the *sum* of $\mathbf{a}$ and $\mathbf{b}$ and denoted by $\mathbf{a} + \mathbf{b}$, such that the following axioms are satisfied.

I.1 *Commutativity.* For any two vectors $\mathbf{a}$ and $\mathbf{b}$ of V,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

I.2 *Associativity.* For any three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ of V,

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \qquad (\text{written } \mathbf{a} + \mathbf{b} + \mathbf{c}).$$

I.3 There is a unique vector in V, called the *zero vector* and denoted by $\mathbf{0}$, such that for every $\mathbf{a}$ in V,

$$\mathbf{a} + \mathbf{0} = \mathbf{a}.$$

I.4 For every $\mathbf{a}$ in V there is a unique vector in V that is denoted by $-\mathbf{a}$ and is such that

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}.$$

II. Scalar multiplication. The real numbers are called **scalars**. Scalar multiplication associates with every $\mathbf{a}$ in V and every scalar c a unique vector of V, called the *product* of c and $\mathbf{a}$ and denoted by $c\mathbf{a}$ (or $\mathbf{a}c$) such that the following axioms are satisfied.

II.1 *Distributivity.* For every scalar c and vectors $\mathbf{a}$ and $\mathbf{b}$ in V,

$$c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}.$$

II.2 *Distributivity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$(c + k)\mathbf{a} = c\mathbf{a} + k\mathbf{a}.$$

II.3 *Associativity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$c(k\mathbf{a}) = (ck)\mathbf{a} \qquad \text{(written } ck\mathbf{a}\text{)}.$$

II.4 For every $\mathbf{a}$ in V,

$$1\mathbf{a} = \mathbf{a}.$$

If, in the above definition, we take complex numbers as scalars instead of real numbers, we obtain the axiomatic definition of a **complex vector space**.

Inner Product Spaces

If **a** and **b** are vectors in R^n, regarded as column vectors, we can form the product $\mathbf{a}^\mathsf{T}\mathbf{b}$. This is a 1×1 matrix, which we can identify with its single entry, that is, with a number. This product is called the **inner product** or **dot product** of **a** and **b**. Other notations for it are $(\mathbf{a}, \mathbf{b})$ and $\mathbf{a} \bullet \mathbf{b}$. Thus

$$\mathbf{a}^\mathsf{T}\mathbf{b} = (\mathbf{a}, \mathbf{b}) = \mathbf{a} \bullet \mathbf{b} = [a_1 \cdots a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_{i=1}^{n} a_l b_l = a_1 b_1 + \cdots + a_n b_n.$$

DEFINITION

Real Inner Product Space

A real vector space V is called a **real inner product space** (or *real pre-Hilbert*[4] *space*) if it has the following property. With every pair of vectors **a** and **b** in V there is associated a real number, which is denoted by $(\mathbf{a}, \mathbf{b})$ and is called the **inner product** of **a** and **b**, such that the following axioms are satisfied.

I. For all scalars q_1 and q_2 and all vectors **a**, **b**, **c** in V,

$$(q_1\mathbf{a} + q_2\mathbf{b}, \mathbf{c}) = q_1(\mathbf{a}, \mathbf{c}) + q_2(\mathbf{b}, \mathbf{c}) \qquad (\textit{Linearity}).$$

II. For all vectors **a** and **b** in V,

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a}) \qquad \textit{(Symmetry)}.$$

III. For every **a** in V,

$$\left.\begin{array}{r} (\mathbf{a}, \mathbf{a}) \geqq 0, \\ (\mathbf{a}, \mathbf{a}) = 0 \quad \text{if and only if} \quad \mathbf{a} = \mathbf{0} \end{array}\right\} \quad \textit{(Positive-definiteness)}.$$

Vectors whose inner product is zero are called **orthogonal**.
The *length* or **norm** of a vector in V is defined by

$$\|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})} \quad (\geqq 0). \tag{2}$$

A vector of norm 1 is called a **unit vector**.

From these axioms and from (2) one can derive the basic inequality

$$|(\mathbf{a}, \mathbf{b})| \leqq \|\mathbf{a}\| \, \|\mathbf{b}\| \qquad \textit{(Cauchy–Schwarz}^{5} \textit{ inequality)}. \tag{3}$$

From this follows

$$(4) \qquad \|\mathbf{a} + \mathbf{b}\| \leqq \|\mathbf{a}\| + \|\mathbf{b}\| \qquad \textit{(Triangle inequality)}.$$

A simple direct calculation gives

$$(5) \qquad \|\mathbf{a} + \mathbf{b}\|^2 + \|\mathbf{a} - \mathbf{b}\|^2 = 2(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2) \qquad \textit{(Parallelogram equality)}.$$

EXAMPLE 3 *n*-Dimensional Euclidean Space

R^n with the inner product

$$(6) \qquad (\mathbf{a}, \mathbf{b}) = \mathbf{a}^\mathsf{T}\mathbf{b} = a_1 b_1 + \cdots + a_n b_n$$

(where both **a** and **b** are *column* vectors) is called the ***n*-dimensional Euclidean space** and is denoted by E^n or again simply by R^n. Axioms I–III hold, as direct calculation shows. Equation (2) gives the "**Euclidean norm**"

$$(7) \qquad \|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})} = \sqrt{\mathbf{a}^\mathsf{T}\mathbf{a}} = \sqrt{a_1^2 + \cdots + a_n^2}. \qquad \blacksquare$$

EXAMPLE 4 An Inner Product for Functions. Function Space

The set of all real-valued continuous functions $f(x), g(x), \cdots$ on a given interval $\alpha \leqq x \leqq \beta$ is a real vector space under the usual addition of functions and multiplication by scalars (real numbers). On this "**function space**" we can define an inner product by the integral

$$(f, g) = \int_{\alpha}^{\beta} f(x)\, g(x)\, dx. \tag{8}$$

Axioms I–III can be verified by direct calculation. Equation (2) gives the norm

$$\|f\| = \sqrt{(f, f)} = \sqrt{\int_{\alpha}^{\beta} f(x)^2\, dx}. \tag{9}$$

Linear Transformations

Let X and Y be any vector spaces. To each vector $\mathbf{x}$ in X we assign a unique vector $\mathbf{y}$ in Y. Then we say that a **mapping** (or **transformation** or **operator**) of X into Y is given. Such a mapping is denoted by a capital letter, say F. The vector $\mathbf{y}$ in Y assigned to a vector $\mathbf{x}$ in X is called the **image** of $\mathbf{x}$ under F and is denoted by $F(\mathbf{x})$ [or $F\mathbf{x}$, without parentheses].

F is called a **linear mapping** or **linear transformation** if, for all vectors $\mathbf{v}$ and $\mathbf{x}$ in X and scalars c,

$$(10) \qquad \begin{aligned} F(\mathbf{v} + \mathbf{x}) &= F(\mathbf{v}) + F(\mathbf{x}) \\ F(c\mathbf{x}) &= cF(\mathbf{x}). \end{aligned}$$

Linear Transformation of Space R^n into Space R^m

From now on we let $X = R^n$ and $Y = R^m$. Then any real $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ gives a transformation of R^n into R^m,

$$(11) \qquad \mathbf{y} = \mathbf{A}\mathbf{x}.$$

Since $\mathbf{A}(\mathbf{u} + \mathbf{x}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{x}$ and $\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x}$, this transformation is linear.

We show that, conversely, every linear transformation F of R^n into R^m can be given in terms of an $m \times n$ matrix $\mathbf{A}$, after a basis for R^n and a basis for R^m have been chosen. This can be proved as follows.

Let $\mathbf{e}_{(1)}, \cdots, \mathbf{e}_{(n)}$ be any basis for R^n. Then every $\mathbf{x}$ in R^n has a unique representation

$$\mathbf{x} = x_1\mathbf{e}_{(1)} + \cdots + x_n\mathbf{e}_{(n)}.$$

Since F is linear, this representation implies for the image $F(\mathbf{x})$:

$$F(\mathbf{x}) = F(x_1\mathbf{e}_{(1)} + \cdots + x_n\mathbf{e}_{(n)}) = x_1F(\mathbf{e}_{(1)}) + \cdots + x_nF(\mathbf{e}_{(n)}).$$

Hence F is uniquely determined by the images of the vectors of a basis for R^n. We now choose for R^n the "**standard basis**"

$$\text{(12)} \qquad \mathbf{e}_{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \mathbf{e}_{(n)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

where $\mathbf{e}_{(j)}$ has its jth component equal to 1 and all others 0. We show that we can now determine an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ such that for every $\mathbf{x}$ in R^n and image $\mathbf{y} = F(\mathbf{x})$ in R^m,

$$\mathbf{y} = F(\mathbf{x}) = \mathbf{A}\mathbf{x}.$$

Indeed, from the image $\mathbf{y}^{(1)} = F(\mathbf{e}_{(1)})$ of $\mathbf{e}_{(1)}$ we get the condition

$$\mathbf{y}^{(1)} = \begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \\ \vdots \\ y_m^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

from which we can determine the first column of $\mathbf{A}$, namely $a_{11} = y_1^{(1)}$, $a_{21} = y_2^{(1)}$, $\cdots$, $a_{m1} = y_m^{(1)}$. Similarly, from the image of $\mathbf{e}_{(2)}$ we get the second column of $\mathbf{A}$, and so on. This completes the proof. ■

In three-dimensional Euclidean space E^3 the standard basis is usually written $\mathbf{e}_{(1)} = \mathbf{i}$, $\mathbf{e}_{(2)} = \mathbf{j}$, $\mathbf{e}_{(3)} = \mathbf{k}$. Thus,

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{13}$$

EXAMPLE 5 Linear Transformations

Interpreted as transformations of Cartesian coordinates in the plane, the matrices

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

(a) (b) (c) (d)

represent

(a) a reflection in the line $x_2 = x_1$,

(b) a reflection in the x_1-axis,

(c) a reflection in the origin, and

(d) a stretch (when a > 1, or a contraction when 0 < a < 1) in the x_1-direction.

EXAMPLE 6 Linear Transformations

Our discussion preceding Example 5 is simpler than it may look at first sight. To see this, find **A** representing the linear transformation that maps (x_1, x_2) onto $(2x_1 - 5x_2, 3x_1 + 4x_2)$.

Solution. Obviously, the transformation is

$$y_1 = 2x_1 - 5x_2$$
$$y_2 = 3x_1 + 4x_2.$$

From this we can directly see that the matrix is

$$\mathbf{A} = \begin{bmatrix} 2 & -5 \\ 3 & 4 \end{bmatrix}. \qquad \text{Check:} \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - 5x_2 \\ 3x_1 + 4x_2 \end{bmatrix}.$$

If **A** in (11) is square, $n \times n$, then (11) maps R^n into R^n. If this **A** is nonsingular, so that $\mathbf{A}^{-1}$ exists (see Sec. 7.8), then multiplication of (11) by $\mathbf{A}^{-1}$ from the left and use of $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ gives the **inverse transformation**

(14) $$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}.$$

It maps every $\mathbf{y} = \mathbf{y}_0$ onto that $\mathbf{x}$, which by (11) is mapped onto $\mathbf{y}_0$. *The inverse of a linear transformation is itself linear,* because it is given by a matrix, as (14) shows.

Composition of Linear Transformations

The last operation we want to discuss is composition of linear transformations. Let X, Y, W be general vector spaces. As before, let F be a linear transformation from X to Y. Let G be a linear transformation from W to X. Then we denote, by H, the **composition** of F and G, that is,

$$H = F \circ G = FG = F(G),$$

which means *we take transformation G and then apply transformation F to it* (***in that order!***, *i.e. you go from left to right*).

Now, to give this a more concrete meaning, if we let $\mathbf{w}$ be a vector in W, then $G(\mathbf{w})$ is a vector in X and $F(G(\mathbf{w}))$ is a vector in Y. Thus, H maps W to Y, and we can write

$$H(\mathbf{w}) = (F \circ G)(\mathbf{w}) = (FG)(\mathbf{w}) = F(G(\mathbf{w})), \tag{15}$$

EXAMPLE 7 The Composition of Linear Transformations Is Linear

To show that H is indeed linear we must show that (10) holds. We have, for two vectors $\mathbf{w}_1$, $\mathbf{w}_2$ in W,

$$\begin{aligned} H(\mathbf{w}_1 + \mathbf{w}_2) &= (F \circ G)(\mathbf{w}_1 + \mathbf{w}_2) \\ &= F(G(\mathbf{w}_1 + \mathbf{w}_2)) \\ &= F(G(\mathbf{w}_1) + G(\mathbf{w}_2)) && \text{(by linearity of } G) \\ &= F(G(\mathbf{w}_1)) + F(G(\mathbf{w}_2)) && \text{(by linearity of } F) \\ &= (F \circ G)(\mathbf{w}_1) + (F \circ G)(\mathbf{w}_2) && \text{(by (15))} \\ &= H(\mathbf{w}_1) + H(\mathbf{w}_2) && \text{(by definition of } H). \end{aligned}$$

Similarly, $H(c\mathbf{w}_2) = (F \circ G)(c\mathbf{w}_2) = F(G(c\mathbf{w}_2)) = F(c(G(\mathbf{w}_2))$

$$= cF(G(\mathbf{w}_2)) = c(F \circ G)(\mathbf{w}_2) = cH(\mathbf{w}_2).$$

Next we want to relate composition of linear transformations to matrix multiplication.

To do so we let $X = R^n$, $Y = R^m$, and $W = R^p$. This choice of particular vector spaces allows us to represent the linear transformations as matrices and form matrix equations, as was done in (11). Thus F can be represented by a general real $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ and G by an $n \times p$ matrix $\mathbf{B} = [b_{jk}]$. Then we can write for F, with column vectors $\mathbf{x}$ with n entries, and resulting vector $\mathbf{y}$, with m entries

(16) $$\mathbf{y} = \mathbf{A}\mathbf{x}$$

and similarly for G, with column vector $\mathbf{w}$ with p entries,

(17) $$\mathbf{x} = \mathbf{B}\mathbf{w}.$$

Substituting (17) into (16) gives

(18) $$\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{B}\mathbf{w}) = (\mathbf{A}\mathbf{B})\mathbf{w} = \mathbf{A}\mathbf{B}\mathbf{w} = \mathbf{C}\mathbf{w} \qquad \text{where } \mathbf{C} = \mathbf{A}\mathbf{B}.$$

This is (15) in a matrix setting, this is, ***we can define the composition of linear transformations in the Euclidean spaces as multiplication by matrices***. Hence, the real $m \times p$ matrix $\mathbf{C}$ represents a linear transformation H which maps R^p to R^n with vector $\mathbf{w}$, a column vector with p entries.