n-Dimensional Euclidean Space and Matrices

Version: April, 2008

Definition of n space. As was learned in Math 1b, a point in Euclidean three space can be thought of in any of three ways:

(i) as the set of *triples* (x, y, z) where x, y, and z are real numbers;

(ii) as the set of *points* in space;

(iii) as the set of *directed line segments* in space, based at the origin.

The first of these points of view is easily extended from 3 to any number of dimensions. We define $\mathbb{R}^n$, where n is a positive integer (possibly greater than 3), as the set of all ordered n-tuples $(x_1, x_2, \ldots, x_n)$, where the x_i are real numbers. For instance, $(1, \sqrt{5}, 2, 4) \in \mathbb{R}^4$.

The set $\mathbb{R}^n$ is known as ***Euclidean n-space***, and we may think of its elements $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ as ***vectors*** or ***n-vectors***. By setting $n = 1, 2$, or 3, we recover the line, the plane, and three-dimensional space respectively.

Addition and Scalar Multiplication. We begin our study of Euclidean n-space by introducing algebraic operations analogous to those learned for $\mathbb{R}^2$ and $\mathbb{R}^3$. Addition and scalar multiplication are defined as follows:

(i) $(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$; and

(ii) for any real number α,

$$\alpha(a_1, a_2, \ldots, a_n) = (\alpha a_1, \alpha a_2, \ldots, \alpha a_n).$$

The geometric significance of these operations for $\mathbb{R}^2$ and $\mathbb{R}^3$ is discussed in "Calculus", Chapter 13.

Informal Discussion. What is the point of going from $\mathbb{R}, \mathbb{R}^2$ and $\mathbb{R}^3$, which seem comfortable and "real world", to $\mathbb{R}^n$? Our world is, after all, *three-dimensional*, not n-dimensional!

First, it *is true* that the bulk of multivariable calculus is about $\mathbb{R}^2$ and $\mathbb{R}^3$. However, many of the ideas work in $\mathbb{R}^n$ with little extra effort, so why not do it? Second, the ideas really are *useful*! For instance, if you are studying a chemical reaction involving 5 chemicals, you will probably want to store and manipulate their concentrations as a 5-tuple; that is, an element of $\mathbb{R}^5$. The laws governing chemical reaction rates also demand we do *calculus* in this 5-dimensional space.

The Standard Basis of $\mathbb{R}^n$. The n vectors in $\mathbb{R}^n$ defined by

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 0, 1)$$

are called the ***standard basis vectors*** of $\mathbb{R}^n$, and they generalize the three mutually orthogonal unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ of $\mathbb{R}^3$. Any vector $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ can be written in terms of the $\mathbf{e}_i$'s as

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \ldots + a_n\mathbf{e}_n.$$

Dot Product and Length. Recall that for two vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ in $\mathbb{R}^3$, we defined the ***dot product*** or ***inner product*** $\mathbf{a} \cdot \mathbf{b}$ to be the real number

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

This definition easily extends to $\mathbb{R}^n$; specifically, for $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, we define

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + \ldots + a_nb_n.$$

We also define the ***length*** or ***norm*** of a vector $\mathbf{a}$ by the formula

$$\text{length of } \mathbf{a} = \|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}.$$

The algebraic properties 1–5 (page 669 in "Calculus", Chapter 13) for the dot product in three space are still valid for the dot product in $\mathbb{R}^n$. Each can be proven by a simple computation.

Vector Spaces and Inner Product Spaces. The notion of a ***vector space*** focusses on having a set of objects called ***vectors*** that one can add and multiply by scalars, where these operations obey the familiar rules of vector addition. Thus, we refer to $\mathbb{R}^n$ as an example of a vector space (also called a linear space). For example, the space of all continuous functions f defined on the interval $[0, 1]$ is a vector space. The student is familiar with how to add functions and how to multipy them by scalars and they obey similar algebraic properties as the addition of vectors (such as, for example, $\alpha(f + g) = \alpha f + \alpha g$).

This same space of functions also provides an example of an ***inner product space***, that is, a vector space in which one has a dot product that satisfies the properties 1–5 (page 669 in "Calculus", Chapter 13. Namely, we define the inner product of f and g to be

$$f \cdot g = \int_0^1 f(x)g(x)\,dx.$$

Notice that this is analogous to the definition of the inner product of two vectors in $\mathbb{R}^n$, with the sum replaced by an integral.

The Cauchy-Schwarz Inequality. Using the algebraic properties of the dot product, we will prove algebraically the following inequality, which is normally proved by a geometric argument in $\mathbb{R}^3$.

Cauchy-Schwarz Inequality

For vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$, we have

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|.$$

If either $\mathbf{a}$ or $\mathbf{b}$ is zero, the inequality reduces to $0 \leq 0$, so we can assume that both are non-zero. Then

$$0 \leq \left\| \frac{\mathbf{a}}{\|\mathbf{a}\|} - \frac{\mathbf{b}}{\|\mathbf{b}\|} \right\|^2 = 1 - \frac{2\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|} + 1;$$

that is,

$$0 \leq -\frac{2\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|}.$$

Rearranging, we get

$$\mathbf{a} \cdot \mathbf{b} \leq \|\mathbf{a}\| \, \|\mathbf{b}\|.$$

The same argument using a plus sign instead of minus gives

$$-\mathbf{a} \cdot \mathbf{b} \leq \|\mathbf{a}\| \, \|\mathbf{b}\|.$$

The two together give the stated inequality, since $|\mathbf{a} \cdot \mathbf{b}| = \pm \mathbf{a} \cdot \mathbf{b}$ depending on the sign of $\mathbf{a} \cdot \mathbf{b}$.

If the vectors $\mathbf{a}$ and $\mathbf{b}$ in the Cauchy-Schwarz inequality are both nonzero, the inequality shows that the quotient $\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|}$ lies between -1 and 1. Any number in the interval $[-1, 1]$ is the cosine of a unique angle θ between 0 and π; by analogy with the 2 and 3-dimensional cases, we call

$$\theta = \cos^{-1} \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|}$$

the ***angle between the vectors*** $\mathbf{a}$ and $\mathbf{b}$.

Informal Discussion. Notice that the reasoning above is opposite to the geometric reasoning normally done in $\mathbb{R}^3$. That is because, in $\mathbb{R}^n$, we have no geometry to begin with, so we start with algebraic definitions and define the geometric notions in terms of them.

The Triangle Inequality. We can derive the triangle inequality from the Cauchy-Schwarz inequality: $\mathbf{a}\cdot\mathbf{b} \le |\mathbf{a}\cdot\mathbf{b}| \le \|\mathbf{a}\|\,\|\mathbf{b}\|$, so that

$$\|\mathbf{a}+\mathbf{b}\|^2 = \|\mathbf{a}\|^2 + 2\mathbf{a}\cdot\mathbf{b} + \|\mathbf{b}\|^2 \le \|\mathbf{a}\|^2 + 2\|\mathbf{a}\|\|\mathbf{b}\| + \|\mathbf{b}\|^2.$$

Hence, we get $\|\mathbf{a}+\mathbf{b}\|^2 \le (\|\mathbf{a}\|+\|\mathbf{b}\|)^2$; taking square roots gives the following result.

Triangle Inequality

Let vectors $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{R}^n$. Then

$$\|\mathbf{a}+\mathbf{b}\| \le \|\mathbf{a}\| + \|\mathbf{b}\|.$$

If the Cauchy-Schwarz and triangle inequalities are written in terms of components, they look like this:

$$\left|\sum_{i=1}^n a_i b_i\right| \le \left(\sum_{i=1}^n a_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n b_i^2\right)^{\frac{1}{2}};$$

and

$$\left(\sum_{i=1}^n (a_i+b_i)^2\right)^{\frac{1}{2}} \le \left(\sum_{i=1}^n a_i^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^n b_i^2\right)^{\frac{1}{2}}.$$

The fact that the proof of the Cauchy-Schwarz inequality depends only on the properties of vectors and of the inner product means that the proof works for any inner product space and not just for $\mathbb{R}^n$. For example, for the space of continuous functions on $[0,1]$, the Cauchy-Schwarz inequality reads

$$\left|\int_0^1 f(x)g(x)\,dx\right| \le \sqrt{\int_0^1 [f(x)]^2\,dx}\sqrt{\int_0^1 [g(x)]^2\,dx}$$

and the triangle inequality reads

$$\sqrt{\int_0^1 [f(x)+g(x)]^2\,dx} \le \sqrt{\int_0^1 [f(x)]^2\,dx} + \sqrt{\int_0^1 [g(x)]^2\,dx}.$$

Example 1. Verify the Cauchy-Schwarz and triangle inequalities for $\mathbf{a} = (1, 2, 0, -1)$ and $\mathbf{b} = (-1, 1, 1, 0)$.

Solution.

$$\begin{aligned}
\|\mathbf{a}\| &= \sqrt{1^2 + 2^2 + 0^2 + (-1)^2} = \sqrt{6} \\
\|\mathbf{b}\| &= \sqrt{(-1)^2 + 1^2 + 1^2 + 0^2} = \sqrt{3} \\
\mathbf{a} \cdot \mathbf{b} &= 1(-1) + 2 \cdot 1 + 0 \cdot 1 + (-1)0 = 1 \\
\mathbf{a} + \mathbf{b} &= (0, 3, 1, -1) \\
\|\mathbf{a} + \mathbf{b}\| &= \sqrt{0^2 + 3^2 + 1^2 + (-1)^2} = \sqrt{11}.
\end{aligned}$$

We compute $\mathbf{a} \cdot \mathbf{b} = 1$ and $\|\mathbf{a}\| \, \|\mathbf{b}\| = \sqrt{6}\sqrt{3} \approx 4.24$, which verifies the Cauchy-Schwarz inequality. Similarly, we can check the triangle inequality:

$$\|\mathbf{a} + \mathbf{b}\| = \sqrt{11} \approx 3.32,$$

while

$$\|\mathbf{a}\| \; + \; \|\mathbf{b}\| = 2.45 + 1.73 \approx \sqrt{6} + \sqrt{3} \approx 4.18,$$

which is larger. ♦

Distance. By analogy with $\mathbb{R}^3$, we can define the notion of distance in $\mathbb{R}^n$; namely, if $\mathbf{a}$ and $\mathbf{b}$ are points in $\mathbb{R}^n$, the ***distance between* a *and* b** is defined to be $\|\mathbf{a} - \mathbf{b}\|$, or the length of the vector $\mathbf{a} - \mathbf{b}$. There is *no cross product* defined on $\mathbb{R}^n$ except for $n = 3$, because it is only in $\mathbb{R}^3$ that there is a unique direction perpendicular to two (linearly independent) vectors. It is only the dot product that has been generalized.

> **Informal Discussion.** By focussing on *different* ways of thinking about the cross product, it turns out that there *is* a generalization of the cross product to $\mathbb{R}^n$. This is part of a calculus one learns in more advanced courses that was discovered by Cartan around 1920. This calculus is closely related to how one generalizes the main integral theorems of Green, Gauss and Stokes' to $\mathbb{R}^n$.

Matrices. Generalizing 2×2 and 3×3 matrices, we can consider $m \times n$ matrices, that is, arrays of mn numbers:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{n2} & \cdots & a_{mn} \end{bmatrix}.$$

Note that an $m \times n$ matrix has m rows and n columns. The entry a_{ij} goes in the ith row and the jth column. We shall also write A as $[a_{ij}]$. A ***square matrix*** is one for which $m = n$.

Addition and Scalar Multiplication. We define addition and multiplication by a scalar componentwise, as we did for vectors. Given two $m \times n$ matrices A and B, we can add them to obtain a new $m \times n$ matrix $C = A + B$, whose ijth entry c_{ij} is the sum of a_{ij} and b_{ij}. It is clear that $A + B = B + A$. Similarly, the difference $D = A - B$ is the matrix whose entries are $d_{ij} = a_{ij} - b_{ij}$.

Example 2.

(a) $\begin{bmatrix} 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 3 \\ 0 & 0 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 8 \end{bmatrix}.$

(b) $\begin{bmatrix} 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix}.$

(c) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$ $\blacklozenge$

Given a scalar λ and an $m \times n$ matrix A, we can multiply A by λ to obtain a new $m \times n$ matrix $\lambda A = C$, whose ijth entry c_{ij} is the product λa_{ij}.

Example 3.

$$3 \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 5 \\ 1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -3 & 6 \\ 0 & 3 & 15 \\ 3 & 0 & 9 \end{bmatrix}. \quad \blacklozenge$$

Matrix Multiplication. Next we turn to the most important operation, that of ***matrix multiplication***. If $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then $AB = C$ is the $m \times p$ matrix whose ijth entry is

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

which is the dot product of the ith row of A and the jth column of B.

One way to motivate matrix multiplication is via substitution of one set of linear equations into another. Suppose, for example, one has the linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= y_1 \\ a_{21}x_1 + a_{22}x_2 &= y_2 \end{aligned}$$

and we want to substitute (or change variables) the expressions

$$\begin{aligned} x_1 &= b_{11}u_1 + b_{12}u_2 \\ x_2 &= b_{21}u_1 + b_{22}u_2 \end{aligned}$$

Then one gets the equations (as is readily checked)

$$\begin{aligned} c_{11}u_1 + c_{12}u_2 &= y_1 \\ c_{21}u_1 + c_{22}u_2 &= y_2 \end{aligned}$$

where the coefficient matrix of c_{ij} is given by the product of the matrices a_{ij} and b_{ij}. Similar considerations can be used to motivate the multiplication of arbitrary matrices (as long as the matrices can indeed be multiplied.)

Example 4. Let

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 0 & 4 & 3 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 3 \\ 3 & 1 & 0 \end{bmatrix}. \quad \blacklozenge$$

This example shows that matrix multiplication is ***not*** commutative. That is, in general,

$$AB \neq BA.$$

Note that for AB to be defined, the number of columns of A must equal the number of rows of B.

Example 5. Let

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 3 & 1 & 5 \\ 3 & 4 & 5 \end{bmatrix},$$

but BA is not defined. ♦

Example 6. Let

$$A = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 2 & 1 & 2 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 4 & 4 & 2 & 4 \\ 2 & 2 & 1 & 2 \\ 6 & 6 & 3 & 6 \end{bmatrix} \quad \text{and} \quad BA = [13]. \quad \blacklozenge$$

If we have three matrices A, B, and C such that the products AB and BC are defined, then the products $(AB)C$ and $A(BC)$ will be defined and equal (that is, *matrix multiplication is associative*). It is legitimate, therefore, to drop the parenthesis and denote the product by ABC.

Example 7. Let

$$A = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \quad B = [1 \quad 1], \quad \text{and} \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Then

$$A(BC) = \begin{bmatrix} 3 \\ 5 \end{bmatrix} [3] = \begin{bmatrix} 9 \\ 15 \end{bmatrix}.$$

Also,

$$AB(C) = \begin{bmatrix} 3 & 3 \\ 5 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 15 \end{bmatrix}$$

as well. ♦

It is convenient when working with matrices to represent the vector $\mathbf{a} = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$ by the $n \times 1$ ***column matrix*** $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, which we also denote by $\mathbf{a}$.

> **Informal Discussion.** Less frequently, one also encounters the $1 \times n$ ***row matrix*** $[a_1 \ldots a_n]$, which is called the ***transpose*** of $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ and is denoted by $\mathbf{a}^T$. Note that an element of $\mathbb{R}^n$ (e. g. (4, 1, 3.2, 6)) is written with parentheses and commas, while a row matrix (e.g. $[7 \quad 1 \quad 3 \quad 2]$) is written with square brackets and no commas. It is also convenient at this point to use the letters $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$ instead of $\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$ to denote vectors and their components.

Linear Transformations. If $A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}$ is a $m \times n$ matrix, and $\mathbf{x} = (x_1, \ldots, x_n)$ is in $\mathbb{R}^n$, we can form the matrix product

$$A\mathbf{x} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

The resulting $m \times 1$ column matrix can be considered as a vector $(y_1, \ldots, y_m)$ in $\mathbb{R}^m$. In this way, the matrix A determines a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $\mathbf{y} = A\mathbf{x}$. This function, sometimes indicated by the notation $\mathbf{x} \mapsto A\mathbf{x}$ to emphasize that $\mathbf{x}$ is transformed into $A\mathbf{x}$, is called a ***linear transformation***, since it has the following ***linearity properties***

$$\begin{aligned} A(\mathbf{x} + \mathbf{y}) &= A\mathbf{x} + A\mathbf{y} \quad \text{for } \mathbf{x} \text{ and } \mathbf{y} \text{ in } \mathbb{R}^n \\ A(\alpha\mathbf{x}) &= \alpha(A\mathbf{x}) \quad \text{for } \mathbf{x} \text{ in } \mathbb{R}^n \text{ and } \alpha \text{ in } \mathbb{R}. \end{aligned}$$

This is related to the general notion of a ***linear transformation***; that is, a map between vector spaces that satisfies these same properties.

One can show that any linear transformation $T(x)$ of the vector space $\mathbb{R}^n$ to $\mathbb{R}^m$ actually is of the form $T(x) = A\mathbf{x}$ above for some matrix A. In view of this, one speaks of A as *the matrix of the linear transformation.* As in calculus, an important operation on trasformations is their composition. The main link with matrix multiplication is this: The matrix of a composition of two linear transformations is the product of the matrices of the two transformations. Again, one can use this to motivate the definition of matrix multiplication and it is basic in the study of linear algebra.

Informal Discussion. We will be using abstract linear algebra very sparingly. If you have had a course in linear algebra it will only enrich your vector calculus experience. If you have not had such a course there is no need to worry — we will provide everything that you will need for vector calculus.

Example 8. If

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

then the function $\mathbf{x} \mapsto A\mathbf{x}$ from $\mathbb{R}^3$ to $\mathbb{R}^4$ is defined by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 3 \\ -1 & 0 & 1 \\ 2 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_3 \\ -x_1 + x_3 \\ 2x_1 + x_2 + 2x_3 \\ -x_1 + 2x_2 + x_3 \end{bmatrix}. \quad \blacklozenge$$

Example 9. The following illustrates what happens to a specific point when mapped by the 4×3 matrix of Example 8:

$$A\mathbf{e}_2 = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 0 & 1 \\ 2 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \text{2nd column of } A. \quad \blacklozenge$$

Inverses of Matrices. An $n \times n$ matrix is said to be ***invertible*** if there is an $n \times n$ matrix B such that

$$AB = BA = I_n,$$

where

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

is the $n \times n$ identity matrix. The matrix I_n has the property that $I_nC = CI_n = C$ for any $n \times n$ matrix C. We denote B by A^{-1} and call A^{-1} the ***inverse*** of A. The inverse, when it exists, is unique.

Example 10. If

$$A = \begin{bmatrix} 2 & 4 & 0 \\ 0 & 2 & 1 \\ 3 & 0 & 2 \end{bmatrix}, \quad \text{then} \quad A^{-1} = \frac{1}{20} \begin{bmatrix} 4 & -8 & 4 \\ 3 & 4 & -2 \\ -6 & 12 & 4 \end{bmatrix},$$

since $AA^{-1} = I_3 = A^{-1}A$, as may be checked by matrix multiplication. ♦

If A is invertible, the equation $A\mathbf{x} = \mathbf{b}$ can be solved for the vector $\mathbf{x}$ by multiplying both sides by A^{-1} to obtain $\mathbf{x} = A^{-1}\mathbf{b}$.

Solving Systems of Equations. Finding inverses of matrices is more or less equivalent to solving systems of linear equations. The fundamental fact alluded to above is that

> Write a system of n linear equations in n unknowns in the form $A\mathbf{x} = \mathbf{b}$. It has a unique solution for any $\mathbf{b}$ if and only if the matrix A has an inverse and in this case the solution is given by $\mathbf{x} = A^{-1}\mathbf{b}$.

The basic methods for solving linear equations learned in Math 1b can be brought to bear in this context. These include Cramer's Rule and row reduction.

Example 11. Consider the system of equations

$$\begin{aligned} 3x + 2y &= u \\ -x + y &= v \end{aligned}$$

which can be readily solved by row reduction to give

$$\begin{aligned} x &= \frac{u}{5} - \frac{2v}{5} \\ y &= \frac{u}{5} + \frac{3v}{5} \end{aligned}$$

If we write the system of equations as

$$\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

and the solution as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

then we see that

$$\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{3}{5} \end{bmatrix}$$

which can also be checked by matrix multiplication.

If one encounters an $n \times n$ system that *does not have a solution*, then one can conclude that the matrix of coefficients does not have an inverse. One can also conclude that the matrix of coefficients does not have an inverse if *there are multiple solutions.*

Determinants. One way to approach determinants is to take the definition of the determinant of a 2×2 and a 3×3 matrix as a starting point. This can then be generalized to $n \times n$ determinants. We illustrate here how to write the determinant of a 4×4 matrix in terms of the determinants of 3×3 matrices:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix}$$
$$+ a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}$$

(note that the signs alternate $+, -, +, -, \ldots$). Continuing this way, one defines 5×5 determinants in terms of 4×4 determinants, and so on.

The basic properties of 3×3 determinants remain valid for $n \times n$ determinants. In particular, we note the fact that if A is an $n \times n$ matrix and B is the matrix formed by adding a scalar multiple of the kth row (or column) of A to the lth row (or, respectively, column) of A, then the determinant of A is equal to the determinant of B. (This fact is used in Example 11 below.)

A basic theorem of linear algebra states that

> An $n \times n$ matrix A is invertible if and only if the determinant of A is not zero.

Another basic property is that $\det(AB) = (\det A)(\det B)$. In this text we leave these assertions unproved.

Example 12. Let

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

Find $\det A$. Does A have an inverse?

Solution. Adding $(-1)\times$ first column to the third column and then expanding by minors of the first row, we get

$$\det A = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & -1 & 2 \end{vmatrix} = 1 \begin{vmatrix} 1 & 0 & 1 \\ 1 & -2 & 1 \\ 1 & -1 & 2 \end{vmatrix}.$$

Adding $(-1)\times$ first column to the third column of this 3×3 determinant gives

$$\det A = \begin{vmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 1 & -1 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 0 \\ -1 & 1 \end{vmatrix} = -2.$$

Thus, $\det A = -2 \neq 0$, and so A has an inverse. To actually find the inverse (if this were asked), one can solve the relevant system of linear equations (by, for example, row reduction) and proceed as in the previous example. ♦

Exercises.

Perform the calculations indicated in Exercises 1 – 4.

1. $(1, 4, 5, 6, 7) + (1, 2, 3, 4, 5) =$
2. $(1, 2, 3, \ldots, n) + (0, 1, 2, \ldots, n-1) =$
3. $2(1, 2, 3, 4, 5) \cdot (5, 4, 3, 2, 1) =$
4. $4(5, 4, 3, 2) \cdot 6(8, 4, 1, 7) =$

Verify the Cauchy-Schwarz inequality and the triangle inequality for the vectors given in Exercises 5 – 8.

5. $\mathbf{a} = (2, 0, -1), \mathbf{b} = (4, 0, -2)$
6. $\mathbf{a} = (1, 0, 2, 6), \mathbf{b} = (3, 8, 4, 1)$
7. $\mathbf{i} + \mathbf{j} + \mathbf{k}, \mathbf{i} + \mathbf{k}$
8. $2\mathbf{i} + \mathbf{j}, 3\mathbf{i} + 4\mathbf{j}$

Perform the calculations indicated in Exercises 9 – 12.

9. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} + \begin{bmatrix} 4 & 7 & 3 \\ 8 & 2 & 1 \\ 0 & 6 & 6 \end{bmatrix} =$
10. $\begin{bmatrix} 0 & 6 & 3 \\ 2 & 9 & 8 \\ 1 & 3 & 3 \\ 2 & 7 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 1 & 7 \\ 6 & 6 & 6 \\ 4 & 4 & 4 \\ 9 & 8 & 1 \end{bmatrix} =$
11. $\begin{bmatrix} 2 & 3 & 4 \\ 7 & 7 & 7 \\ 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 2 \\ 3 & 2 & 3 \end{bmatrix} =$
12. $6 \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} =$

In Exercises 13 – 20, find the matrix product or explain why it is not defined.

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

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

15. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 3 & 2 & 1 \end{bmatrix}$

18. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

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

20. $\left(\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & -1 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

In Exercises 21 – 24, for the given A, (a) define the mapping $\mathbf{x} \mapsto A\mathbf{x}$ as was done in Example 8, and (b) calculate $A\mathbf{a}$.

21. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

22. $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 4 \\ 9 \\ 8 \end{bmatrix}$

23. $A = \begin{bmatrix} 4 & 5 \\ 9 & 0 \\ 1 & 1 \\ 7 & 3 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$

24. $A = \begin{bmatrix} 4 & 4 & 4 & 0 \\ 3 & 5 & 5 & 7 \end{bmatrix}$, $\quad \mathbf{a} = \begin{bmatrix} 7 \\ 9 \\ 0 \\ 1 \end{bmatrix}$

In Exercises 25 – 28, determine whether the given matrix has an inverse.

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

26. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

27. $\begin{bmatrix} 0 & 4 & 5 \\ 7 & 0 & 6 \\ 8 & 9 & 0 \end{bmatrix}$

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

29. Let $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Find a matrix B such that $AB = I$. Check that $BA = I$.

30. Verify that the inverse of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

31. Assuming $\det(AB) = (\det A)(\det B)$, verify that $(\det A)(\det A^{-1}) = 1$ and conclude that if A has an inverse, then $\det A \neq 0$.

32. Show that, if A, B and C are $n \times n$ matrices such that $AB = I_n$ and $CA = I_n$, then $B = C$.

33. Define the ***transpose*** A^T of an $m \times n$ matrix A as follows, A^T is the $n \times m$ matrix whose ijth entry is the jith entry of A. For instance

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \qquad \begin{bmatrix} a & b \\ d & e \\ g & h \end{bmatrix}^T = \begin{bmatrix} a & d & g \\ b & e & h \end{bmatrix}.$$

One can see directly that, if A is a 2×2 or 3×3 matrix, then $\det(A^T) = \det A$. Use this to prove the same fact for 4×4 matrices.

34. Let B be the $m \times 1$ column matrix $\begin{bmatrix} \frac{1}{m} \\ \frac{1}{m} \\ \vdots \\ \frac{1}{m} \end{bmatrix}$. If $A = [a_1 \dots a_m]$ is any row matrix, what is AB?

35. Prove that if A is a 4×4 matrix, then
    (a) if B is a matrix obtained from a 4×4 matrix A by multiplying any row or column by a scalar λ, then $\det B = \lambda \det A$; and
    (b) $\det(\lambda A) = \lambda^4 \det A$.

In Exercises 36 – 38, A, B, and C denote $n \times n$ matrices.

36. Is $\det(A + B) = \det A + \det B$? Give a proof or counterexample.

37. Does $(A + B)(A - B) = A^2 - B^2$?

38. Assuming the law $\det(AB) = (\det A)(\det B)$, prove that $\det(ABC) = (\det A)(\det B)(\det C)$.

39. In the Mathematics 1b final you were asked to solve the system of equations

$$\begin{aligned} x + y + 2z &= 1 \\ 3x + y + 4z &= 3 \\ 2x + y &= 4 \end{aligned}$$

Relate your solution to the inversion problem for the matrix of coefficients. Find the inverse if it exists.

40. (a) Try to invert the matrix

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

by solving the appropriate system of equations by row reduction. Find the inverse if it exists or show that an inverse does not exist.

(b) Try to invert the matrix

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

by solving the appropriate system of equations by row reduction. Find the inverse if it exists or show that an inverse does not exist.