Lyryx with Open Texts

LINEAR ALGEBRA
with Applications

Open Edition

Adapted for

Emory University

Math 221

Linear Algebra

Sections **1 & 2**

Lectured and adapted by

Le Chen

April 15, 2021

le.chen@emory.edu

Course page

http://math.emory.edu/~lchen41/teaching/2021_Spring_Math221

ADAPTABLE | ACCESSIBLE | AFFORDABLE

by W. Keith Nicholson

Contents

2.2 Matrix-Vector Multiplication

Up to now we have used matrices to solve systems of linear equations by manipulating the rows of the augmented matrix. In this section we introduce a different way of describing linear systems that makes more use of the coefficient matrix of the system and leads to a useful way of "multiplying" matrices.

Vectors

It is a well-known fact in analytic geometry that two points in the plane with coordinates (a_1, a_2) and (b_1, b_2) are equal if and only if $a_1 = b_1$ and $a_2 = b_2$. Moreover, a similar condition applies to points (a_1, a_2, a_3) in space. We extend this idea as follows.

An ordered sequence $(a_1, a_2, \ldots, a_n)$ of real numbers is called an **ordered n-tuple**. The word "ordered" here reflects our insistence that two ordered n-tuples are equal if and only if corresponding entries are the same. In other words,

$$(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n) \quad \text{if and only if} \quad a_1 = b_1, a_2 = b_2, \ldots, \text{ and } a_n = b_n.$$

Thus the ordered 2-tuples and 3-tuples are just the ordered pairs and triples familiar from geometry.

> **Definition 2.4 The set $\mathbb{R}^n$ of ordered n-tuples of real numbers**
>
> *Let $\mathbb{R}$ denote the set of all real numbers. The set of all ordered n-tuples from $\mathbb{R}$ has a special notation:*
>
> $$\mathbb{R}^n \text{ denotes the set of all ordered } n\text{-tuples of real numbers.}$$

There are two commonly used ways to denote the n-tuples in $\mathbb{R}^n$: As rows $(r_1, r_2, \ldots, r_n)$ or columns $\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$; the notation we use depends on the context. In any event they are called **vectors** or n-**vectors** and will be denoted using bold type such as $\mathbf{x}$ or $\mathbf{v}$. For example, an $m \times n$ matrix A will be written as a row of columns:

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \text{ where } \mathbf{a}_j \text{ denotes column } j \text{ of } A \text{ for each } j.$$

If $\mathbf{x}$ and $\mathbf{y}$ are two n-vectors in $\mathbb{R}^n$, it is clear that their matrix sum $\mathbf{x} + \mathbf{y}$ is also in $\mathbb{R}^n$ as is the scalar multiple $k\mathbf{x}$ for any real number k. We express this observation by saying that $\mathbb{R}^n$ is **closed** under addition and scalar multiplication. In particular, all the basic properties in Theorem 2.1.1 are true of these n-vectors. These properties are fundamental and will be used frequently below without comment. As for matrices in general, the $n \times 1$ zero matrix is called the **zero n-vector** in $\mathbb{R}^n$ and, if $\mathbf{x}$ is an n-vector, the n-vector $-\mathbf{x}$ is called the **negative x**.

Of course, we have already encountered these n-vectors in Section 1.3 as the solutions to systems of linear equations with n variables. In particular we defined the notion of a linear combination

of vectors and showed that a linear combination of solutions to a homogeneous system is again a solution. Clearly, a linear combination of n-vectors in $\mathbb{R}^n$ is again in $\mathbb{R}^n$, a fact that we will be using.

Matrix-Vector Multiplication

Given a system of linear equations, the left sides of the equations depend only on the coefficient matrix A and the column $\mathbf{x}$ of variables, and not on the constants. This observation leads to a fundamental idea in linear algebra: We view the left sides of the equations as the "product" $A\mathbf{x}$ of the matrix A and the vector $\mathbf{x}$. This simple change of perspective leads to a completely new way of viewing linear systems—one that is very useful and will occupy our attention throughout this book.

To motivate the definition of the "product" $A\mathbf{x}$, consider first the following system of two equations in three variables:

$$\begin{aligned} ax_1 + bx_2 + cx_3 &= b_1 \\ a'x_1 + b'x_2 + c'x_3 &= b_1 \end{aligned} \tag{2.2}$$

and let $A = \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ denote the coefficient matrix, the variable matrix, and the constant matrix, respectively. The system (2.2) can be expressed as a single vector equation

$$\begin{bmatrix} ax_1 + bx_2 + cx_3 \\ a'x_1 + b'x_2 + c'x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

which in turn can be written as follows:

$$x_1 \begin{bmatrix} a \\ a' \end{bmatrix} + x_2 \begin{bmatrix} b \\ b' \end{bmatrix} + x_3 \begin{bmatrix} c \\ c' \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Now observe that the vectors appearing on the left side are just the columns

$$\mathbf{a}_1 = \begin{bmatrix} a \\ a' \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} b \\ b' \end{bmatrix}, \quad \text{and } \mathbf{a}_3 = \begin{bmatrix} c \\ c' \end{bmatrix}$$

of the coefficient matrix A. Hence the system (2.2) takes the form

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b} \tag{2.3}$$

This shows that the system (2.2) has a solution if and only if the constant matrix $\mathbf{b}$ is a linear combination[3] of the columns of A, and that in this case the entries of the solution are the coefficients x_1, x_2, and x_3 in this linear combination.

Moreover, this holds in general. If A is any $m \times n$ matrix, it is often convenient to view A as a row of columns. That is, if $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ are the columns of A, we write

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

and say that $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ is *given in terms of its columns*.

---

[3]Linear combinations were introduced in Section 1.3 to describe the solutions of homogeneous systems of linear equations. They will be used extensively in what follows.

Now consider any system of linear equations with $m \times n$ coefficient matrix A. If $\mathbf{b}$ is the constant matrix of the system, and if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is the matrix of variables then, exactly as above, the system can be written as a single vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{2.4}$$

Example 2.2.1

Write the system $\begin{cases} 3x_1 + 2x_2 - 4x_3 = 0 \\ x_1 - 3x_2 + x_3 = 3 \\ x_2 - 5x_3 = -1 \end{cases}$ in the form given in (2.4).

Solution.

$$x_1 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -4 \\ 1 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}$$

As mentioned above, we view the left side of (2.4) as the *product* of the matrix A and the vector $\mathbf{x}$. This basic idea is formalized in the following definition:

Definition 2.5 Matrix-Vector Multiplication

*Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ be an $m \times n$ matrix, written in terms of its columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is any n-vector, the **product** $A\mathbf{x}$ is defined to be the m-vector given by:*

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

In other words, if A is $m \times n$ and $\mathbf{x}$ is an n-vector, the product $A\mathbf{x}$ is the linear combination of the columns of A where the coefficients are the entries of $\mathbf{x}$ (in order).

Note that if A is an $m \times n$ matrix, the product $A\mathbf{x}$ is only defined if $\mathbf{x}$ is an n-vector and then the vector $A\mathbf{x}$ is an m-vector because this is true of each column $\mathbf{a}_j$ of A. But in this case the *system* of linear equations with coefficient matrix A and constant vector $\mathbf{b}$ takes the form of a *single* matrix equation

$$A\mathbf{x} = \mathbf{b}$$

The following theorem combines Definition 2.5 and equation (2.4) and summarizes the above discussion. Recall that a system of linear equations is said to be *consistent* if it has at least one solution.

Theorem 2.2.1

1. *Every system of linear equations has the form $A\mathbf{x} = \mathbf{b}$ where A is the coefficient matrix, $\mathbf{b}$ is the constant matrix, and $\mathbf{x}$ is the matrix of variables.*

2. *The system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is a linear combination of the columns of A.*

3. *If $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ are the columns of A and if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then $\mathbf{x}$ is a solution to the linear system $A\mathbf{x} = \mathbf{b}$ if and only if x_1, x_2, ..., x_n are a solution of the vector equation*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

A system of linear equations in the form $A\mathbf{x} = \mathbf{b}$ as in (1) of Theorem 2.2.1 is said to be written in **matrix form**. This is a useful way to view linear systems as we shall see.

Theorem 2.2.1 transforms the problem of solving the linear system $A\mathbf{x} = \mathbf{b}$ into the problem of expressing the constant matrix B as a linear combination of the columns of the coefficient matrix A. Such a change in perspective is very useful because one approach or the other may be better in a particular situation; the importance of the theorem is that there is a choice.

Example 2.2.2

If $A = \begin{bmatrix} 2 & -1 & 3 & 5 \\ 0 & 2 & -3 & 1 \\ -3 & 4 & 1 & 2 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ -2 \end{bmatrix}$, compute $A\mathbf{x}$.

Solution. By Definition 2.5: $A\mathbf{x} = 2\begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix} + 1\begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix} + 0\begin{bmatrix} 3 \\ -3 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7 \\ 0 \\ -6 \end{bmatrix}$.

Example 2.2.3

Given columns $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$ in $\mathbb{R}^3$, write $2\mathbf{a}_1 - 3\mathbf{a}_2 + 5\mathbf{a}_3 + \mathbf{a}_4$ in the form $A\mathbf{x}$ where A is a matrix and $\mathbf{x}$ is a vector.

Solution. Here the column of coefficients is $\mathbf{x} = \begin{bmatrix} 2 \\ -3 \\ 5 \\ 1 \end{bmatrix}$. Hence Definition 2.5 gives

$$A\mathbf{x} = 2\mathbf{a}_1 - 3\mathbf{a}_2 + 5\mathbf{a}_3 + \mathbf{a}_4$$

where $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix}$ is the matrix with $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$ as its columns.

Example 2.2.4

Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix}$ be the 3×4 matrix given in terms of its columns $\mathbf{a}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$,

$\mathbf{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 3 \\ -1 \\ -3 \end{bmatrix}$, and $\mathbf{a}_4 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$. In each case below, either express $\mathbf{b}$ as a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$, or show that it is not such a linear combination. Explain what your answer means for the corresponding system $A\mathbf{x} = \mathbf{b}$ of linear equations.

$\quad$ a. $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\qquad\qquad\qquad\qquad\qquad$ b. $\mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$

Solution. By Theorem 2.2.1, $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$ if and only if the system $A\mathbf{x} = \mathbf{b}$ is consistent (that is, it has a solution). So in each case we carry the augmented matrix $[A|\mathbf{b}]$ of the system $A\mathbf{x} = \mathbf{b}$ to reduced form.

$\quad$ a. Here $\left[\begin{array}{cccc|c} 2 & 1 & 3 & 3 & 1 \\ 0 & 1 & -1 & 1 & 2 \\ -1 & 1 & -3 & 0 & 3 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right]$, so the system $A\mathbf{x} = \mathbf{b}$ has no

$\quad$ solution in this case. Hence $\mathbf{b}$ is *not* a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$.

$\quad$ b. Now $\left[\begin{array}{cccc|c} 2 & 1 & 3 & 3 & 4 \\ 0 & 1 & -1 & 1 & 2 \\ -1 & 1 & -3 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$, so the system $A\mathbf{x} = \mathbf{b}$ is

$\quad$ consistent.

Thus $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$ in this case. In fact the general solution is $x_1 = 1 - 2s - t$, $x_2 = 2 + s - t$, $x_3 = s$, and $x_4 = t$ where s and t are arbitrary parameters.

Hence $x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + x_3 \mathbf{a}_3 + x_4 \mathbf{a}_4 = \mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$ for *any* choice of s and t. If we take $s = 0$ and

$t = 0$, this becomes $\mathbf{a}_1 + 2\mathbf{a}_2 = \mathbf{b}$, whereas taking $s = 1 = t$ gives $-2\mathbf{a}_1 + 2\mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4 = \mathbf{b}$.

Example 2.2.5

Taking A to be the zero matrix, we have $0\mathbf{x} = \mathbf{0}$ for all vectors $\mathbf{x}$ by Definition 2.5 because every column of the zero matrix is zero. Similarly, $A\mathbf{0} = \mathbf{0}$ for all matrices A because every entry of the zero vector is zero.

> **Example 2.2.6**
>
> If $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, show that $I\mathbf{x} = \mathbf{x}$ for any vector $\mathbf{x}$ in $\mathbb{R}^3$.
>
> **Solution.** If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ then Definition 2.5 gives
>
> $$I\mathbf{x} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}$$

The matrix I in Example 2.2.6 is called the 3×3 **identity matrix**, and we will encounter such matrices again in Example 2.2.11 below. Before proceeding, we develop some algebraic properties of matrix-vector multiplication that are used extensively throughout linear algebra.

> **Theorem 2.2.2**
>
> Let A and B be $m \times n$ matrices, and let $\mathbf{x}$ and $\mathbf{y}$ be n-vectors in $\mathbb{R}^n$. Then:
>
> 1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$.
>
> 2. $A(a\mathbf{x}) = a(A\mathbf{x}) = (aA)\mathbf{x}$ for all scalars a.
>
> 3. $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$.

Proof. We prove (3); the other verifications are similar and are left as exercises. Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ be given in terms of their columns. Since adding two matrices is the same as adding their columns, we have

$$A + B = \begin{bmatrix} \mathbf{a}_1 + \mathbf{b}_1 & \mathbf{a}_2 + \mathbf{b}_2 & \cdots & \mathbf{a}_n + \mathbf{b}_n \end{bmatrix}$$

If we write $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ Definition 2.5 gives

$$\begin{aligned} (A + B)\mathbf{x} &= x_1(\mathbf{a}_1 + \mathbf{b}_1) + x_2(\mathbf{a}_2 + \mathbf{b}_2) + \cdots + x_n(\mathbf{a}_n + \mathbf{b}_n) \\ &= (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n) + (x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n) \\ &= A\mathbf{x} + B\mathbf{x} \end{aligned}$$

☐

Theorem 2.2.2 allows matrix-vector computations to be carried out much as in ordinary arithmetic. For example, for any $m \times n$ matrices A and B and any n-vectors $\mathbf{x}$ and $\mathbf{y}$, we have:

$$A(2\mathbf{x} - 5\mathbf{y}) = 2A\mathbf{x} - 5A\mathbf{y} \quad \text{and} \quad (3A - 7B)\mathbf{x} = 3A\mathbf{x} - 7B\mathbf{x}$$

We will use such manipulations throughout the book, often without mention.

Linear Equations

Theorem 2.2.2 also gives a useful way to describe the solutions to a system

$$Ax = b$$

of linear equations. There is a related system

$$Ax = 0$$

called the **associated homogeneous system**, obtained from the original system $Ax = b$ by re-placing all the constants by zeros. Suppose x_1 is a solution to $Ax = b$ and x_0 is a solution to $Ax = 0$ (that is $Ax_1 = b$ and $Ax_0 = 0$). Then $x_1 + x_0$ is another solution to $Ax = b$. Indeed, Theorem 2.2.2 gives

$$A(x_1 + x_0) = Ax_1 + Ax_0 = b + 0 = b$$

This observation has a useful converse.

Theorem 2.2.3

Suppose x_1 is any particular solution to the system $Ax = b$ of linear equations. Then every solution x_2 to $Ax = b$ has the form

$$x_2 = x_0 + x_1$$

for some solution x_0 of the associated homogeneous system $Ax = 0$.

Proof. Suppose x_2 is also a solution to $Ax = b$, so that $Ax_2 = b$. Write $x_0 = x_2 - x_1$. Then $x_2 = x_0 + x_1$ and, using Theorem 2.2.2, we compute

$$Ax_0 = A(x_2 - x_1) = Ax_2 - Ax_1 = b - b = 0$$

Hence x_0 is a solution to the associated homogeneous system $Ax = 0$. □

Note that gaussian elimination provides one such representation.

Example 2.2.7

Express every solution to the following system as the sum of a specific solution plus a solution to the associated homogeneous system.

$$\begin{aligned}
x_1 - x_2 - x_3 + 3x_4 &= 2 \\
2x_1 - x_2 - 3x_3 + 4x_4 &= 6 \\
x_1 \phantom{{}- x_2} - 2x_3 + x_4 &= 4
\end{aligned}$$

Solution. Gaussian elimination gives $x_1 = 4 + 2s - t$, $x_2 = 2 + s + 2t$, $x_3 = s$, and $x_4 = t$ where s and t are arbitrary parameters. Hence the general solution can be written

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 + 2s - t \\ 2 + s + 2t \\ s \\ t \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \left(s \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right)$$

Thus $\mathbf{x}_1 = \begin{bmatrix} 4 \\ 2 \\ 0 \\ 0 \end{bmatrix}$ is a particular solution (where $s = 0 = t$), and $\mathbf{x}_0 = s \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$

gives *all* solutions to the associated homogeneous system. (To see why this is so, carry out the gaussian elimination again but with all the constants set equal to zero.)

The following useful result is included with no proof.

Theorem 2.2.4

Let $A\mathbf{x} = \mathbf{b}$ be a system of equations with augmented matrix $\begin{bmatrix} A & | & b \end{bmatrix}$. Write $\operatorname{rank} A = r$.

1. $\operatorname{rank} \begin{bmatrix} A & | & b \end{bmatrix}$ is either r or $r + 1$.

2. The system is consistent if and only if $\operatorname{rank} \begin{bmatrix} A & | & b \end{bmatrix} = r$.

3. The system is inconsistent if and only if $\operatorname{rank} \begin{bmatrix} A & | & b \end{bmatrix} = r + 1$.

The Dot Product

Definition 2.5 is not always the easiest way to compute a matrix-vector product $A\mathbf{x}$ because it requires that the columns of A be explicitly identified. There is another way to find such a product which uses the matrix A as a whole with no reference to its columns, and hence is useful in practice. The method depends on the following notion.

Definition 2.6 Dot Product in $\mathbb{R}^n$

If $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ are two ordered n-tuples, their **dot product** is defined to be the number

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

obtained by multiplying corresponding entries and adding the results.

To see how this relates to matrix products, let A denote a 3×4 matrix and let $\mathbf{x}$ be a 4-vector.

Writing

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

in the notation of Section 2.1, we compute

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} + x_4 \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 \end{bmatrix}$$

From this we see that each entry of $A\mathbf{x}$ is the dot product of the corresponding row of A with $\mathbf{x}$. This computation goes through in general, and we record the result in Theorem 2.2.5.

Theorem 2.2.5: Dot Product Rule

Let A be an $m \times n$ matrix and let $\mathbf{x}$ be an n-vector. Then each entry of the vector $A\mathbf{x}$ is the dot product of the corresponding row of A with $\mathbf{x}$.

row i
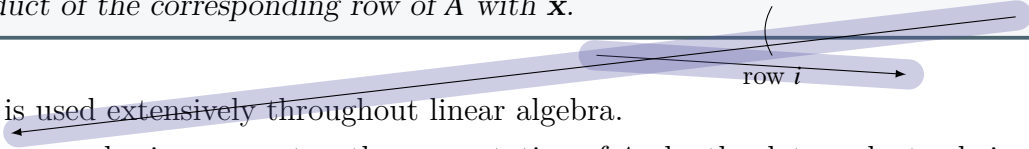
This result is used extensively throughout linear algebra.

If A is $m \times n$ and $\mathbf{x}$ is an n-vector, the computation of $A\mathbf{x}$ by the dot product rule is simpler than using Definition 2.5 because the computation can be carried out directly with no explicit reference to the columns of A (as in Definition 2.5). The first entry of $A\mathbf{x}$ is the dot product of row 1 of A with $\mathbf{x}$. In hand calculations this is computed by going *across* row one of A, going *down* the column $\mathbf{x}$, multiplying corresponding entries, and adding the results. The other entries of $A\mathbf{x}$ are computed in the same way using the other rows of A with the column $\mathbf{x}$.

In general, compute entry i of $A\mathbf{x}$ as follows (see the diagram):

$$A\mathbf{x} \begin{bmatrix} \\ \\ \end{bmatrix} \begin{bmatrix} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \end{bmatrix}$$

Go *across* row i of A and *down* column $\mathbf{x}$, multiply corresponding entries, and add the results.

entry i

As an illustration, we rework Example 2.2.2 using the dot product rule instead of Definition 2.5.

Example 2.2.8

If $A = \begin{bmatrix} 2 & -1 & 3 & 5 \\ 0 & 2 & -3 & 1 \\ -3 & 4 & 1 & 2 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ -2 \end{bmatrix}$, compute $A\mathbf{x}$.

Solution. The entries of $A\mathbf{x}$ are the dot products of the rows of A with $\mathbf{x}$:

$$A\mathbf{x} = \begin{bmatrix} 2 & -1 & 3 & 5 \\ 0 & 2 & -3 & 1 \\ -3 & 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 2 & + & (-1)1 & + & 3 \cdot 0 & + & 5(-2) \\ 0 \cdot 2 & + & 2 \cdot 1 & + & (-3)0 & + & 1(-2) \\ (-3)2 & + & 4 \cdot 1 & + & 1 \cdot 0 & + & 2(-2) \end{bmatrix} = \begin{bmatrix} -7 \\ 0 \\ -6 \end{bmatrix}$$

Of course, this agrees with the outcome in Example 2.2.2.

Example 2.2.9

Write the following system of linear equations in the form $A\mathbf{x} = \mathbf{b}$.

$$\begin{aligned} 5x_1 - x_2 + 2x_3 + x_4 - 3x_5 &= 8 \\ x_1 + x_2 + 3x_3 - 5x_4 + 2x_5 &= -2 \\ -x_1 + x_2 - 2x_3 \qquad\quad - 3x_5 &= 0 \end{aligned}$$

Solution. Write $A = \begin{bmatrix} 5 & -1 & 2 & 1 & -3 \\ 1 & 1 & 3 & -5 & 2 \\ -1 & 1 & -2 & 0 & -3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 8 \\ -2 \\ 0 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$. Then the

dot product rule gives $A\mathbf{x} = \begin{bmatrix} 5x_1 - x_2 + 2x_3 + x_4 - 3x_5 \\ x_1 + x_2 + 3x_3 - 5x_4 + 2x_5 \\ -x_1 + x_2 - 2x_3 \qquad - 3x_5 \end{bmatrix}$, so the entries of $A\mathbf{x}$ are the left

sides of the equations in the linear system. Hence the system becomes $A\mathbf{x} = \mathbf{b}$ because matrices are equal if and only corresponding entries are equal.

Example 2.2.10

If A is the zero $m \times n$ matrix, then $A\mathbf{x} = \mathbf{0}$ for each n-vector $\mathbf{x}$.

Solution. For each k, entry k of $A\mathbf{x}$ is the dot product of row k of A with $\mathbf{x}$, and this is zero because row k of A consists of zeros.

Definition 2.7 The Identity Matrix

*For each $n > 2$, the **identity matrix** I_n is the $n \times n$ matrix with 1s on the main diagonal (upper left to lower right), and zeros elsewhere.*

The first few identity matrices are

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \cdots$$

In Example 2.2.6 we showed that $I_3\mathbf{x} = \mathbf{x}$ for each 3-vector $\mathbf{x}$ using Definition 2.5. The following result shows that this holds in general, and is the reason for the name.

Example 2.2.11

For each $n \geq 2$ we have $I_n\mathbf{x} = \mathbf{x}$ for each n-vector $\mathbf{x}$ in $\mathbb{R}^n$.

Solution. We verify the case $n = 4$. Given the 4-vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ the dot product rule

gives

$$I_4\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1+0+0+0 \\ 0+x_2+0+0 \\ 0+0+x_3+0 \\ 0+0+0+x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \mathbf{x}$$

In general, $I_n\mathbf{x} = \mathbf{x}$ because entry k of $I_n\mathbf{x}$ is the dot product of row k of I_n with $\mathbf{x}$, and row k of I_n has 1 in position k and zeros elsewhere.

Example 2.2.12

Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ be any $m \times n$ matrix with columns $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$. If $\mathbf{e}_j$ denotes column j of the $n \times n$ identity matrix I_n, then $A\mathbf{e}_j = \mathbf{a}_j$ for each $j = 1, 2, \ldots, n$.

Solution. Write $\mathbf{e}_j = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$ where $t_j = 1$, but $t_i = 0$ for all $i \neq j$. Then Theorem 2.2.5 gives

$$A\mathbf{e}_j = t_1\mathbf{a}_1 + \cdots + t_j\mathbf{a}_j + \cdots + t_n\mathbf{a}_n = 0 + \cdots + \mathbf{a}_j + \cdots + 0 = \mathbf{a}_j$$

Example 2.2.12 will be referred to later; for now we use it to prove:

Theorem 2.2.6

Let A and B be $m \times n$ matrices. If $A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$, then $A = B$.

Proof. Write $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ and $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ and in terms of their columns. It is enough to show that $\mathbf{a}_k = \mathbf{b}_k$ holds for all k. But we are assuming that $A\mathbf{e}_k = B\mathbf{e}_k$, which gives $\mathbf{a}_k = \mathbf{b}_k$ by Example 2.2.12. □

We have introduced matrix-vector multiplication as a new way to think about systems of linear equations. But it has several other uses as well. It turns out that many geometric operations can be described using matrix multiplication, and we now investigate how this happens. As a bonus, this description provides a geometric "picture" of a matrix by revealing the effect on a vector when it is multiplied by A. This "geometric view" of matrices is a fundamental tool in understanding them.
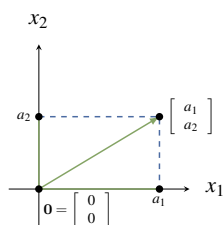
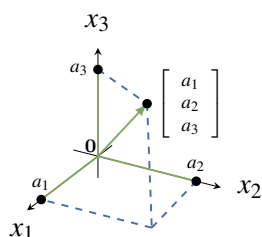## Transformations

Figure 2.2.1

Figure 2.2.2

The set $\mathbb{R}^2$ has a geometrical interpretation as the euclidean plane where a vector $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ in $\mathbb{R}^2$ represents the point $(a_1,\ a_2)$ in the plane (see Figure 2.2.1). In this way we regard $\mathbb{R}^2$ as the set of all points in the plane. Accordingly, we will refer to vectors in $\mathbb{R}^2$ as points, and denote their coordinates as a column rather than a row. To enhance this geometrical interpretation of the vector $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$, it is denoted graphically by an arrow from the origin $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to the vector as in Figure 2.2.1.

Similarly we identify $\mathbb{R}^3$ with 3-dimensional space by writing a point $(a_1,\ a_2,\ a_3)$ as the vector $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ in $\mathbb{R}^3$, again represented by an arrow[4] from the origin to the point as in Figure 2.2.2. In this way the terms "point" and "vector" mean the same thing in the plane or in space.

We begin by describing a particular geometrical transformation of the plane $\mathbb{R}^2$.
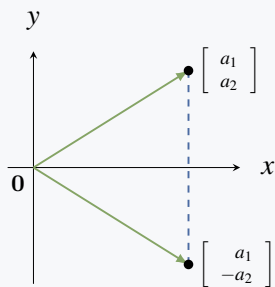
Example 2.2.13

Figure 2.2.3

Consider the transformation of $\mathbb{R}^2$ given by *reflection* in the x axis. This operation carries the vector $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ to its reflection $\begin{bmatrix} a_1 \\ -a_2 \end{bmatrix}$ as in Figure 2.2.3. Now observe that

$$\begin{bmatrix} a_1 \\ -a_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

so reflecting $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ in the x axis can be achieved by

---

[4]This "arrow" representation of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ will be used extensively in Chapter 4.

$$\text{multiplying by the matrix } \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

If we write $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, Example 2.2.13 shows that reflection in the x axis carries each vector $\mathbf{x}$ in $\mathbb{R}^2$ to the vector $A\mathbf{x}$ in $\mathbb{R}^2$. It is thus an example of a function

$$T : \mathbb{R}^2 \to \mathbb{R}^2 \quad \text{where} \quad T(\mathbf{x}) = A\mathbf{x} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^2$$

As such it is a generalization of the familiar functions $f : \mathbb{R} \to \mathbb{R}$ that carry a *number* x to another real *number* $f(x)$.
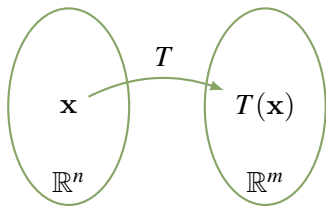

Figure 2.2.4

More generally, functions $T : \mathbb{R}^n \to \mathbb{R}^m$ are called **transformations** from $\mathbb{R}^n$ to $\mathbb{R}^m$. Such a transformation T is a rule that assigns to every vector $\mathbf{x}$ in $\mathbb{R}^n$ a uniquely determined vector $T(\mathbf{x})$ in $\mathbb{R}^m$ called the **image** of $\mathbf{x}$ under T. We denote this state of affairs by writing

$$T : \mathbb{R}^n \to \mathbb{R}^m \quad \text{or} \quad \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$$

The transformation T can be visualized as in Figure 2.2.4.

To describe a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ we must specify the vector $T(\mathbf{x})$ in $\mathbb{R}^m$ for every $\mathbf{x}$ in $\mathbb{R}^n$. This is referred to as **defining** T, or as specifying the **action** of T. Saying that the action *defines* the transformation means that we regard two transformations $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ as **equal** if they have the **same action**; more formally

$$S = T \quad \text{if and only if} \quad S(\mathbf{x}) = T(\mathbf{x}) \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n.$$

Again, this what we mean by $f = g$ where f, $g : \mathbb{R} \to \mathbb{R}$ are ordinary functions.

Functions $f : \mathbb{R} \to \mathbb{R}$ are often described by a formula, examples being $f(x) = x^2 + 1$ and $f(x) = \sin x$. The same is true of transformations; here is an example.

Example 2.2.14

The formula $T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_4 \end{bmatrix}$ defines a transformation $\mathbb{R}^4 \to \mathbb{R}^3$.

Example 2.2.13 suggests that matrix multiplication is an important way of defining transformations $\mathbb{R}^n \to \mathbb{R}^m$. If A is any $m \times n$ matrix, multiplication by A gives a transformation

$$T_A : \mathbb{R}^n \to \mathbb{R}^m \quad \text{defined by} \quad T_A(\mathbf{x}) = A\mathbf{x} \text{ for every } \mathbf{x} \text{ in } \mathbb{R}^n$$

Definition 2.8 Matrix Transformation T_A

T_A is called the **matrix transformation induced** by A.

Thus Example 2.2.13 shows that reflection in the x axis is the matrix transformation $\mathbb{R}^2 \to \mathbb{R}^2$ induced by the matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Also, the transformation $R : \mathbb{R}^4 \to \mathbb{R}^3$ in Example 2.2.13 is the matrix transformation induced by the matrix

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{because} \quad \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_4 \end{bmatrix}$$

Example 2.2.15

Let $R_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$ denote counterclockwise rotation about the origin through $\frac{\pi}{2}$ radians (that is, 90°)[5]. Show that $R_{\frac{\pi}{2}}$ is induced by the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.
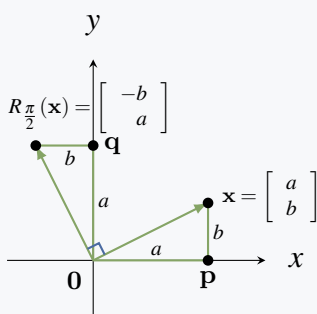
Solution.

Figure 2.2.5

The effect of $R_{\frac{\pi}{2}}$ is to rotate the vector $\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}$ counterclockwise through $\frac{\pi}{2}$ to produce the vector $R_{\frac{\pi}{2}}(\mathbf{x})$ shown in Figure 2.2.5. Since triangles $\mathbf{0}p\mathbf{x}$ and $\mathbf{0}q R_{\frac{\pi}{2}}(\mathbf{x})$ are identical, we obtain $R_{\frac{\pi}{2}}(\mathbf{x}) = \begin{bmatrix} -b \\ a \end{bmatrix}$. But

$$\begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}, \text{ so we obtain } R_{\frac{\pi}{2}}(\mathbf{x}) = A\mathbf{x}$$

for all $\mathbf{x}$ in $\mathbb{R}^2$ where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. In other words, $R_{\frac{\pi}{2}}$ is the matrix transformation induced by A.

If A is the $m \times n$ zero matrix, then A induces the transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m \quad \text{given by} \quad T(\mathbf{x}) = A\mathbf{x} = \mathbf{0} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

This is called the **zero transformation**, and is denoted $T = 0$.

Another important example is the **identity transformation**

$$1_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n \quad \text{given by} \quad 1_{\mathbb{R}^n}(\mathbf{x}) = \mathbf{x} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

That is, the action of $1_{\mathbb{R}^n}$ on $\mathbf{x}$ is to do nothing to it. If I_n denotes the $n \times n$ identity matrix, we showed in Example 2.2.11 that $I_n \mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Hence $1_{\mathbb{R}^n}(\mathbf{x}) = I_n \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$; that is, the identity matrix I_n induces the identity transformation.

Here are two more examples of matrix transformations with a clear geometric description.

---

[5] *Radian measure* for angles is based on the fact that 360° equals 2π radians. Hence π radians $= 180°$ and $\frac{\pi}{2}$ radians $= 90°$.

Example 2.2.16

If $a > 0$, the matrix transformation $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax \\ y \end{bmatrix}$ induced by the matrix $A = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}$ is called an **x-expansion** of $\mathbb{R}^2$ if $a > 1$, and an **x-compression** if $0 < a < 1$. The reason for the names is clear in the diagram below. Similarly, if $b > 0$ the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix}$ gives rise to **y-expansions** and **y-compressions**.
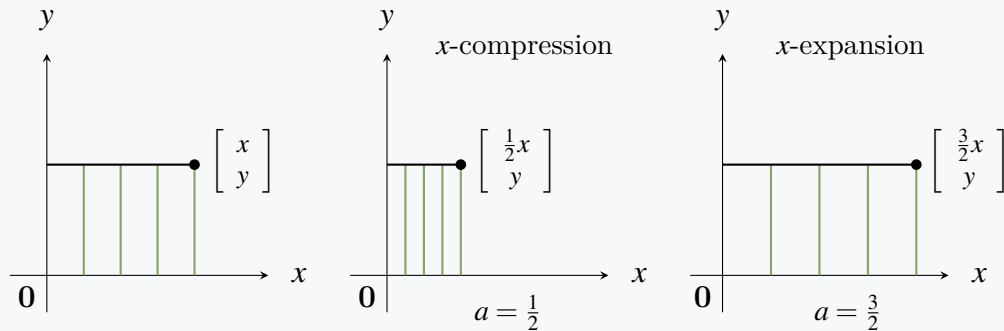
Example 2.2.17

If a is a number, the matrix transformation $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + ay \\ y \end{bmatrix}$ induced by the matrix $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ is called an **x-shear** of $\mathbb{R}^2$ (**positive** if $a > 0$ and **negative** if $a < 0$). Its effect is illustrated below when $a = \frac{1}{4}$ and $a = -\frac{1}{4}$.
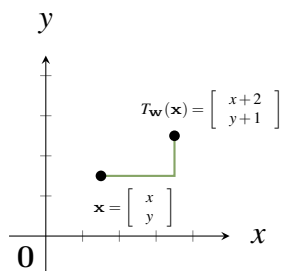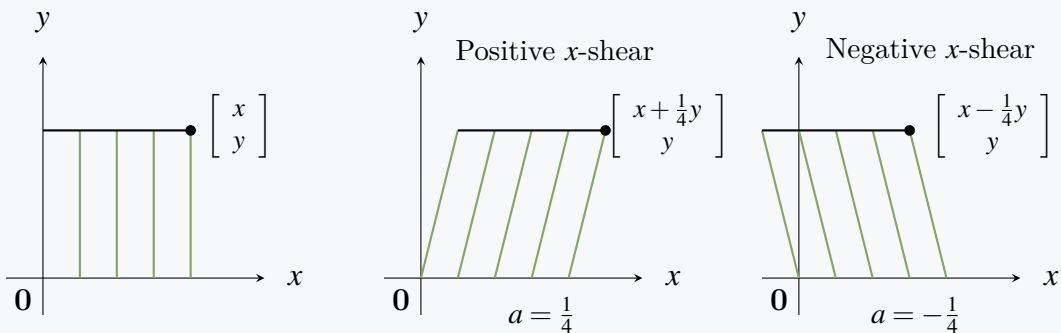
Figure 2.2.6

We hasten to note that there are important geometric transformations that are *not* matrix transformations. For example, if $\mathbf{w}$ is a fixed column in $\mathbb{R}^n$, define the transformation $T_{\mathbf{w}} : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T_{\mathbf{w}}(\mathbf{x}) = \mathbf{x} + \mathbf{w} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

Then $T_{\mathbf{w}}$ is called **translation** by $\mathbf{w}$. In particular, if $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ in

$$\mathbb{R}^2, \text{ the effect of } T_{\mathbf{w}} \text{ on } \begin{bmatrix} x \\ y \end{bmatrix} \text{ is to translate it two units to the right}$$

and one unit up (see Figure 2.2.6).

The translation $T_{\mathbf{w}}$ is not a matrix transformation unless $\mathbf{w} = \mathbf{0}$. Indeed, if $T_{\mathbf{w}}$ were induced by a matrix A, then $A\mathbf{x} = T_{\mathbf{w}}(\mathbf{x}) = \mathbf{x} + \mathbf{w}$ would hold for every $\mathbf{x}$ in $\mathbb{R}^n$. In particular, taking $\mathbf{x} = \mathbf{0}$ gives $\mathbf{w} = A\mathbf{0} = \mathbf{0}$.

Exercises for 2.2

Exercise 2.2.1 In each case find a system of equations that is equivalent to the given vector equation. (Do not solve the system.)

a. $x_1 \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ -3 \end{bmatrix}$

b. $x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 2 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 2 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 2 \\ 0 \end{bmatrix}$

---

b. $\begin{aligned} x_1 - 3x_2 - 3x_3 + 3x_4 &= 5 \\ 8x_2 \qquad\quad + 2x_4 &= 1 \\ x_1 + 2x_2 + 2x_3 \qquad &= 2 \\ x_2 + 2x_3 - 5x_4 &= 0 \end{aligned}$

Exercise 2.2.2 In each case find a vector equation that is equivalent to the given system of equations. (Do not solve the equation.)

a. $\begin{aligned} x_1 - \ x_2 + 3x_3 &= \ 5 \\ -3x_1 + \ x_2 + \ x_3 &= -6 \\ 5x_1 - 8x_2 \qquad &= \ 9 \end{aligned}$

b. $\begin{aligned} x_1 - 2x_2 - \ x_3 + \ x_4 &= \ 5 \\ -x_1 \qquad\quad + \ x_3 - 2x_4 &= -3 \\ 2x_1 - 2x_2 + 7x_3 \qquad &= \ 8 \\ 3x_1 - 4x_2 + 9x_3 - 2x_4 &= 12 \end{aligned}$

---

b. $x_1 \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 0 \\ -2 \\ -4 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 7 \\ 9 \end{bmatrix} +$

$$x_4 \begin{bmatrix} 1 \\ -2 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ 8 \\ 12 \end{bmatrix}$$

Exercise 2.2.3 In each case compute $A\mathbf{x}$ using: (i) Definition 2.5. (ii) Theorem 2.2.5.

a. $A = \begin{bmatrix} 3 & -2 & 0 \\ 5 & -4 & 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

b. $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 5 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

c. $A = \begin{bmatrix} -2 & 0 & 5 & 4 \\ 1 & 2 & 0 & 3 \\ -5 & 6 & -7 & 8 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

d. $A = \begin{bmatrix} 3 & -4 & 1 & 6 \\ 0 & 2 & 1 & 5 \\ -8 & 7 & -3 & 0 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

---

b. $A\mathbf{x} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} +$

$$x_2 \begin{bmatrix} 2 \\ -4 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ -4x_2 + 5x_3 \end{bmatrix}$$

d. $A\mathbf{x} = \begin{bmatrix} 3 & -4 & 1 & 6 \\ 0 & 2 & 1 & 5 \\ -8 & 7 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$

$$= x_1 \begin{bmatrix} 3 \\ 0 \\ -8 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} +$$

$$x_4 \begin{bmatrix} 6 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 3x_1 - 4x_2 + x_3 + 6x_4 \\ 2x_2 + x_3 + 5x_4 \\ -8x_1 + 7x_2 - 3x_3 \end{bmatrix}$$

Exercise 2.2.4 Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix}$ be the 3×4 matrix given in terms of its columns

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \ \mathbf{a}_2 = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \ \mathbf{a}_3 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \text{ and } \mathbf{a}_4 =$$

$$\begin{bmatrix} 0 \\ -3 \\ 5 \end{bmatrix}. \text{ In each case either express } \mathbf{b} \text{ as a linear}$$

combination of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and $\mathbf{a}_4$, or show that it is not such a linear combination. Explain what your answer means for the corresponding system $A\mathbf{x} = \mathbf{b}$ of linear equations.

a) $\mathbf{b} = \begin{bmatrix} 0 \\ 3 \\ 5 \end{bmatrix}$ b) $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$

---

b. To solve $A\mathbf{x} = \mathbf{b}$ the reduction is

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

$$\downarrow$$

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

So the general solution is $\begin{bmatrix} 1+s+3t \\ 1-s-t \\ s \\ t \end{bmatrix}$.

Hence $(1+s+3t)\mathbf{a}_1 + (1-s-t)\mathbf{a}_2 + s\mathbf{a}_3 + t\mathbf{a}_4 = \mathbf{b}$ for any choice of s and t. If $s = t = 0$, we get $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{b}$; if $s = 1$ and $t = 0$, we have $2\mathbf{a}_1 + \mathbf{a}_3 = \mathbf{b}$.

Exercise 2.2.5 In each case, express every solution of the system as a sum of a specific solution plus a solution of the associated homogeneous system.

a) $\begin{aligned} x+y+ z&=2 \\ 2x+y\quad\ &=3 \\ x-y-3z&=0 \end{aligned}$ b) $\begin{aligned} x- y-4z&=-4 \\ x+2y+5z&=2 \\ x+ y+2z&=0 \end{aligned}$

c) $\begin{aligned} x_1+x_2- x_3\quad\ -5x_5&=2 \\ x_2+ x_3\quad\ -4x_5&=-1 \\ x_2+ x_3+x_4- x_5&=-1 \\ 2x_1\quad\ -4x_3+x_4+ x_5&=6 \end{aligned}$

d) $\begin{aligned} 2x_1+x_2- x_3- x_4&=-1 \\ 3x_1+x_2+ x_3-2x_4&=-2 \\ -x_1-x_2+2x_3+ x_4&=2 \\ -2x_1-x_2\quad\ +2x_4&=3 \end{aligned}$

---

b. $\begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$

d. $\begin{bmatrix} 3 \\ -9 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 4 \\ 1 \\ 1 \end{bmatrix}$

Exercise 2.2.6 If $\mathbf{x}_0$ and $\mathbf{x}_1$ are solutions to the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$, use Theorem 2.2.2 to show that $s\mathbf{x}_0 + t\mathbf{x}_1$ is also a solution for any scalars s and t (called a **linear combination** of $\mathbf{x}_0$ and $\mathbf{x}_1$).

We have $A\mathbf{x}_0 = \mathbf{0}$ and $A\mathbf{x}_1 = \mathbf{0}$ and so $A(s\mathbf{x}_0 + t\mathbf{x}_1) = s(A\mathbf{x}_0) + t(A\mathbf{x}_1) = s \cdot \mathbf{0} + t \cdot \mathbf{0} = \mathbf{0}$.

Exercise 2.2.7 Assume that $A \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \mathbf{0} =$ $A \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$. Show that $\mathbf{x}_0 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$ is a solution to $A\mathbf{x} = \mathbf{b}$. Find a two-parameter family of solutions to $A\mathbf{x} = \mathbf{b}$.

Exercise 2.2.8 In each case write the system in the form $A\mathbf{x} = \mathbf{b}$, use the gaussian algorithm to solve the system, and express the solution as a particular solution plus a linear combination of basic solutions to the associated homogeneous system $A\mathbf{x} = \mathbf{0}$.

a. $\begin{aligned} x_1- 2x_2+ x_3+ 4x_4- x_5&=8 \\ -2x_1+ 4x_2+ x_3- 2x_4- 4x_5&=-1 \\ 3x_1- 6x_2+8x_3+ 4x_4-13x_5&=1 \\ 8x_1-16x_2+7x_3+12x_4- 6x_5&=11 \end{aligned}$

b.
$$x_1 - 2x_2 + x_3 + 2x_4 + 3x_5 = -4$$
$$-3x_1 + 6x_2 - 2x_3 - 3x_4 - 11x_5 = 11$$
$$-2x_1 + 4x_2 - x_3 + x_4 - 8x_5 = 7$$
$$-x_1 + 2x_2 + 3x_4 - 5x_5 = 3$$

―――――――――

b. $\mathbf{x} = \begin{bmatrix} -3 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \left(s \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -5 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right).$

Exercise 2.2.9 Given vectors $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$,

$\mathbf{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{a}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$, find a vector $\mathbf{b}$ that

is *not* a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. Justify
your answer. [*Hint*: Part (2) of Theorem 2.2.1.]

Exercise 2.2.10 In each case either show that the
statement is true, or give an example showing that
it is false.

a. $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

b. If $A\mathbf{x}$ has a zero entry, then A has a row of
zeros.

c. If $A\mathbf{x} = \mathbf{0}$ where $\mathbf{x} \neq \mathbf{0}$, then $A = 0$.

d. Every linear combination of vectors in $\mathbb{R}^n$ can
be written in the form $A\mathbf{x}$.

e. If $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$ in terms of its columns,
and if $\mathbf{b} = 3\mathbf{a}_1 - 2\mathbf{a}_2$, then the system $A\mathbf{x} = \mathbf{b}$
has a solution.

f. If $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$ in terms of its columns,
and if the system $A\mathbf{x} = \mathbf{b}$ has a solution, then
$\mathbf{b} = s\mathbf{a}_1 + t\mathbf{a}_2$ for some s, t.

g. If A is $m \times n$ and $m < n$, then $A\mathbf{x} = \mathbf{b}$ has a
solution for every column $\mathbf{b}$.

h. If $A\mathbf{x} = \mathbf{b}$ has a solution for some column $\mathbf{b}$,
then it has a solution for every column $\mathbf{b}$.

i. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions to $A\mathbf{x} = \mathbf{b}$, then
$\mathbf{x}_1 - \mathbf{x}_2$ is a solution to $A\mathbf{x} = \mathbf{0}$.

j. Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$ in terms of its
columns. If $\mathbf{a}_3 = s\mathbf{a}_1 + t\mathbf{a}_2$, then $A\mathbf{x} = \mathbf{0}$, where
$$\mathbf{x} = \begin{bmatrix} s \\ t \\ -1 \end{bmatrix}.$$

―――――――――

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

d. True. The linear combination $x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n$
equals $A\mathbf{x}$ where $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ by The-
orem 2.2.1.

f. False. If $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & 0 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$,

then

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \neq s \begin{bmatrix} 1 \\ 2 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ for any } s \text{ and } t.$$

h. False. If $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$, there is a so-

lution for $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ but not for $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Exercise 2.2.11 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a transfor-
mation. In each case show that T is induced by a
matrix and find the matrix.

a. T is a reflection in the y axis.

b. T is a reflection in the line $y = x$.

c. T is a reflection in the line $y = -x$.

d. T is a clockwise rotation through $\frac{\pi}{2}$.

―――――――――

b. Here $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.

d. Here $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.

Exercise 2.2.12 The **projection** $P : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $P \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$. Show that P is induced by a matrix and find the matrix.

Exercise 2.2.13 Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a transformation. In each case show that T is induced by a matrix and find the matrix.

a. T is a reflection in the $x - y$ plane.

b. T is a reflection in the $y - z$ plane.

_____

b. Here

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

so the matrix is $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Exercise 2.2.14 Fix $a > 0$ in $\mathbb{R}$, and define $T_a : \mathbb{R}^4 \to \mathbb{R}^4$ by $T_a(\mathbf{x}) = a\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^4$. Show that T is induced by a matrix and find the matrix. [T is called a **dilation** if $a > 1$ and a **contraction** if $a < 1$.]

Exercise 2.2.15 Let A be $m \times n$ and let $\mathbf{x}$ be in $\mathbb{R}^n$. If A has a row of zeros, show that $A\mathbf{x}$ has a zero entry.

Exercise 2.2.16 If a vector $\mathbf{b}$ is a linear combination of the columns of A, show that the system $A\mathbf{x} = \mathbf{b}$ is consistent (that is, it has at least one solution.)

_____

Write $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ in terms of its columns. If $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$ where the x_i are scalars, then $A\mathbf{x} = \mathbf{b}$ by Theorem 2.2.1 where $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$. That is, $\mathbf{x}$ is a solution to the system $A\mathbf{x} = \mathbf{b}$.

Exercise 2.2.17 If a system $A\mathbf{x} = \mathbf{b}$ is inconsistent (no solution), show that $\mathbf{b}$ is not a linear combination of the columns of A.

Exercise 2.2.18 Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$.

a. Show that $\mathbf{x}_1 + \mathbf{x}_2$ is a solution to $A\mathbf{x} = \mathbf{0}$.

b. Show that $t\mathbf{x}_1$ is a solution to $A\mathbf{x} = \mathbf{0}$ for any scalar t.

_____

b. By Theorem 2.2.3, $A(t\mathbf{x}_1) = t(A\mathbf{x}_1) = t \cdot \mathbf{0} = \mathbf{0}$; that is, $t\mathbf{x}_1$ is a solution to $A\mathbf{x} = \mathbf{0}$.

Exercise 2.2.19 Suppose $\mathbf{x}_1$ is a solution to the system $A\mathbf{x} = \mathbf{b}$. If $\mathbf{x}_0$ is any nontrivial solution to the associated homogeneous system $A\mathbf{x} = \mathbf{0}$, show that $\mathbf{x}_1 + t\mathbf{x}_0$, t a scalar, is an infinite one parameter family of solutions to $A\mathbf{x} = \mathbf{b}$. [*Hint*: Example 2.1.7 Section 2.1.]

Exercise 2.2.20 Let A and B be matrices of the same size. If $\mathbf{x}$ is a solution to both the system $A\mathbf{x} = \mathbf{0}$ and the system $B\mathbf{x} = \mathbf{0}$, show that $\mathbf{x}$ is a solution to the system $(A + B)\mathbf{x} = \mathbf{0}$.

Exercise 2.2.21 If A is $m \times n$ and $A\mathbf{x} = \mathbf{0}$ for every $\mathbf{x}$ in $\mathbb{R}^n$, show that $A = 0$ is the zero matrix. [*Hint*: Consider $A\mathbf{e}_j$ where $\mathbf{e}_j$ is the jth column of I_n; that is, $\mathbf{e}_j$ is the vector in $\mathbb{R}^n$ with 1 as entry j and every other entry 0.]

Exercise 2.2.22 Prove part (1) of Theorem 2.2.2.

_____

If A is $m \times n$ and $\mathbf{x}$ and $\mathbf{y}$ are n-vectors, we must show that $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$. Denote the columns of A by $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$, and write $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$. Then $\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 & x_2 + y_2 & \cdots & x_n + y_n \end{bmatrix}^T$, so Definition 2.1 and Theorem 2.1.1 give $A(\mathbf{x} + \mathbf{y}) = (x_1 + y_1)\mathbf{a}_1 + (x_2 + y_2)\mathbf{a}_2 + \cdots + (x_n + y_n)\mathbf{a}_n = (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n) + (y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + \cdots + y_n\mathbf{a}_n) = A\mathbf{x} + A\mathbf{y}$.

Exercise 2.2.23 Prove part (2) of Theorem 2.2.2.