Chapter 5

Linear Algebra

Contents

Linear algebra topics specific to **linear algebraic equations** were presented earlier in this text as an extension of college algebra topics, without the aid of vector-matrix notation.

The project before us introduces **specialized vector-matrix notation** in order to extend methods for solving linear algebraic equations. Enrichment includes a full study of rank, nullity, basis and independence from a vector-matrix viewpoint.

Engineering science views linear algebra as an essential language interface between an application and a computer algebra system or a computer numerical laboratory. Without the language interface provided by vectors and matrices, computer assist would be impossibly tedious.

Linear algebra with computer assist is advantageous in the study of **mechanical systems** and **electrical networks**, in which the notation and methods of linear algebra play an important and essential role.

5.1 Vectors and Matrices

The advent of computer algebra systems and computer numerical laboratories has precipitated a common need among engineers and scientists to learn the language of vectors and matrices, which is used heavily in applications.

Fixed Vector Model. A **fixed vector** $\vec{X}$ is a one-dimensional array called a **column vector** or a **row vector**, denoted correspondingly by

$$(1) \qquad \vec{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{or} \quad \vec{X} = \begin{pmatrix} x_1, x_2, \ldots, x_n \end{pmatrix}.$$

The **entries** or **components** $x_1, \ldots, x_n$ are numbers and n is correspondingly called the **column dimension** or the **row dimension** of the vector in (1). The set of all n-vectors (1) is denoted $\mathcal{R}^n$.

Practical matters. A fixed vector is a **package** of application data items. The term **vector** means **data item package** and the collection of all data item packages is the **data set**. Data items are usually numbers. A fixed vector imparts an implicit ordering to the package. To illustrate, a fixed vector might have $n = 6$ components x, y, z, p_x, p_y, p_z, where the first three are space position and the last three are momenta, with respective associated units meters and kilogram-meters per second.

Vector addition and **vector scalar multiplication** are defined by componentwise operations as follows.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad k \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{pmatrix}.$$

The Mailbox Analogy. Fixed vectors can be visualized as in Table 1. Fixed vector entries $x_1, \ldots, x_n$ are numbers written individually onto papers $1, 2, \ldots, n$ deposited into mailboxes with names 1, 2, ..., n.

Table 1. The mailbox analogy. Box i has contents x_i.

x_1	mailbox 1
x_2	mailbox 2
$\vdots$	$\vdots$
x_n	mailbox n

Free Vector Model. In the model, **rigid motions** from geometry are applied to directed line segments. A line segment $\overline{PQ}$ is represented as an **arrow** with head at Q and tail at P. Two such arrows are considered **equivalent** if they can be **rigidly translated** to the same arrow whose tail is at the origin. The arrows are called **free vectors**. They are denoted by the symbol $\overrightarrow{PQ}$, or sometimes $\vec{A} = \overrightarrow{PQ}$, which labels the arrow with tail at P and head at Q.

The parallelogram rule defines **free vector addition**, as in Figure 1. To define **free vector scalar multiplication** $k\vec{A}$, we change the location of the head of vector $\vec{A}$; see Figure 2. If $0 < k < 1$, then the head shrinks to a location along the segment between the head and tail. If $k > 1$, then the head moves in the direction of the arrowhead. If $k < 0$, then the head is reflected along the line and then moved.

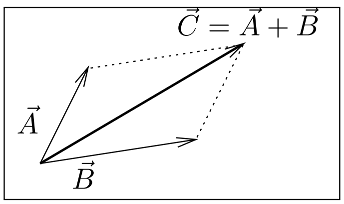

Figure 1. Free vector addition. The diagonal of the parallelogram formed by free vectors $\vec{A}$, $\vec{B}$ is the sum vector $\vec{C} = \vec{A} + \vec{B}$.

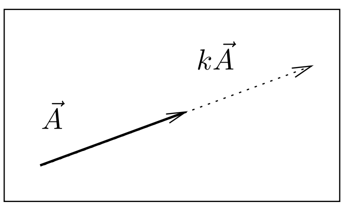

Figure 2. Free vector scalar multiplication. To form $k\vec{A}$, the head of free vector $\vec{A}$ is moved to a new location along the line formed by the head and tail.

Physics Vector Model. This model is also called the **$\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ vector model** and the **orthogonal triad model**. The model arises from the free vector model by inventing symbols $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ for a mutually orthogonal triad of free vectors. Usually, these three vectors represent free vectors of unit length along the coordinate axes, although use in the literature is not restricted to this specialized setting; see Figure 3.

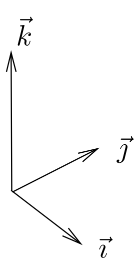

Figure 3. Fundamental triad. The free vectors $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ are 90° apart and of unit length.

The advantage of the model is that any free vector can be represented as $a\vec{\imath} + b\vec{\jmath} + c\vec{k}$ for some constants a, b, c, which gives an immediate connection to the free vector with head at (a, b, c) and tail at $(0, 0, 0)$, as well as to the fixed vector whose components are a, b, c.

Vector addition and scalar multiplication are defined **componentwise**: if $\vec{A} = a_1\vec{\imath} + a_2\vec{\jmath} + a_3\vec{k}$, $\vec{B} = b_1\vec{\imath} + b_2\vec{\jmath} + b_3\vec{k}$ and c is a constant, then

$$\begin{aligned}\vec{A} + \vec{B} &= (a_1 + b_1)\vec{\imath} + (a_2 + b_2)\vec{\jmath} + (a_3 + b_3)\vec{k},\\ c\vec{A} &= (ca_1)\vec{\imath} + (ca_2)\vec{\jmath} + (ca_3)\vec{k}.\end{aligned}$$

Formally, computations involving the **physics model** amount to fixed vector computations and the so-called *equalities* between free vectors and fixed vectors: $\vec{\imath} = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\vec{\jmath} = \begin{pmatrix}0\\1\\0\end{pmatrix}$, $\vec{k} = \begin{pmatrix}0\\0\\1\end{pmatrix}$.

Gibbs Vector Model. The model assigns physical properties to vectors, thus avoiding the pitfalls of free vectors and fixed vectors. Gibbs defines a vector as a **linear motion** that takes a point A into a point B. Visualize this idea as a workman who carries material from A to B: the material is loaded at A, transported along a straight line to B, and then deposited at B. Arrow diagrams arise from this idea by representing a motion from A to B as an arrow with tail at A and head at B.

Vector addition is defined as composition of motions: material is loaded at A and transported to B, then loaded at B and transported to C. Gibbs' idea in the plane is the parallelogram law; see Figure 4.

Vector scalar multiplication is defined so that 1 times a motion is itself, 0 times a motion is no motion and -1 times a motion loads at B and transports to A (the reverse motion). If $k > 0$, then k times a motion from A to B causes the load to be deposited at C instead of B, where k is the ratio of the lengths of segments $\overline{AC}$ and $\overline{AB}$. If $k < 0$, then the definition is applied to the reverse motion from B to A using instead of k the constant $|k|$. Briefly, the load to be deposited along the direction to B is dropped earlier if $0 < |k| < 1$ and later if $|k| > 1$.

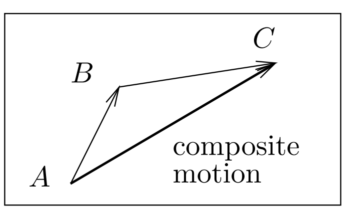

Figure 4. Planar composition of motions. The motion A to C is the composition of two motions or the *sum* of vectors AB and BC.

Comparison of Vector Models. It is possible to use free, physics and Gibbs vector models in free vector diagrams, almost interchangeably. In the Gibbs model, the negative of a vector and the zero vector are natural objects, whereas in the free and physics models they are problematic. To understand the theoretical difficulties, try to answer these questions:

1. What is the zero vector?
2. What is the meaning of the negative of a vector?

Some working rules which connect the free, physics and Gibbs models to the fixed model are the following.

Conversion A fixed vector $\vec{X}$ with components a, b, c converts to a free vector drawn from $(0,0,0)$ to (a,b,c).

Addition To add two free vectors, $\vec{Z} = \vec{X} + \vec{Y}$, place the tail of $\vec{Y}$ at the head of $\vec{X}$, then draw vector $\vec{Z}$ to form a triangle, from the tail of $\vec{X}$ to the head of $\vec{Y}$.

Head–Tail A free vector $\vec{X}$ converts to a fixed vector whose components are the componentwise differences between the point at the head and the point at the tail.
To subtract two free vectors, $\vec{Z} = \vec{Y} - \vec{X}$, place the tails of $\vec{X}$ and $\vec{Y}$ together, then draw $\vec{Z}$ between the heads of $\vec{X}$ and $\vec{Y}$, with the heads of $\vec{Z}$ and $\vec{Y}$ together.

The last item can be memorized as the phrase **head minus tail**. We shall reference both statements as the **head minus tail rule**.

Vector Spaces and the Toolkit. Consider any vector model: fixed, free, physics or Gibbs. Let V denote the **data set** of one of these models. The data set consists of packages of data items, called **vectors**.[1] Assume a particular dimension, n for fixed, 2 or 3 for the others. Let k, k_1, k_2 be constants. Let $\vec{X}$, $\vec{Y}$, $\vec{Z}$ represent three vectors in V. The following **toolkit** of eight (8) vector properties can be verified from the definitions.

Closure	The operations $\vec{X}+\vec{Y}$ and $k\vec{X}$ are defined and result in a new data item package [a vector] which is also in V.	
Addition	$\vec{X}+\vec{Y}=\vec{Y}+\vec{X}$	commutative
	$\vec{X}+(\vec{Y}+\vec{Z})=(\vec{Y}+\vec{X})+\vec{Z}$	associative
	Vector $\vec{0}$ is defined and $\vec{0}+\vec{X}=\vec{X}$	zero
	Vector $-\vec{X}$ is defined and $\vec{X}+(-\vec{X})=\vec{0}$	negative
Scalar multiply	$k(\vec{X}+\vec{Y})=k\vec{X}+k\vec{Y}$	distributive I
	$(k_1+k_2)\vec{X}=k_1\vec{X}+k_2\vec{X}$	distributive II
	$k_1(k_2\vec{X})=(k_1k_2)\vec{X}$	distributive III
	$1\vec{X}=\vec{X}$	identity

[1] If you think vectors are arrows, then re-tool your thoughts. Think of vectors as **data item packages**. A technical word, **vector** can also mean a graph, a matrix for a digital photo, a sequence, a signal, an impulse, or a differential equation solution .

Definition 1 (Vector Space)
A data set V equipped with $\boxed{+}$ and $\boxed{\cdot}$ operations satisfying the closure law and the eight toolkit properties is called an **abstract vector space**.

What's a *space*? There is no intended geometrical implication in this term. The usage of **space** originates from phrases like **parking space** and **storage space**. An abstract vector space is a data set for an application, organized as packages of data items, together with $\boxed{+}$ and $\boxed{\cdot}$ operations, which satisfy the eight toolkit manipulation rules. The packaging of individual data items is structured, or organized, by some scheme, which amounts to a *storage space*, hence the term *space*.

What does *abstract* mean? The technical details of the packaging and the organization of the data set are invisible to the toolkit rules. The toolkit acts on the formal packages, which are called **vectors**. Briefly, the toolkit is used **abstractly**, devoid of any details of the storage scheme. Bursting data packages into data items is generally counterproductive for algebraic manipulations. Resist the temptation.

A variety of data sets. The following key examples are a basis for initial intuition about vector spaces.

> **Coordinate space** $\mathcal{R}^n$ is the set of all fixed n-vectors. Sets $\mathcal{R}^n$ are structured packaging systems which organize data sets from calculations, geometrical problems and physical vector diagrams.
>
> **Function spaces** are structured packages of graphs, such as solutions to differential equations.
>
> **Infinite sequence spaces** are suited to organize the coefficients of series expansions, like Fourier series and Taylor series.
>
> **Matrix spaces** are structured systems which can organize two-dimensional data sets. Examples are the array of pixels for a digital photograph and robotic mechanical component locations given by 3×3 or 4×4 matrices.

Subspaces and Data Analysis. Subspaces address the issue of how to do efficient data analysis on a smaller subset S of a data set V. We assume the larger data set V is equipped with $\boxed{+}$ and $\boxed{\cdot}$ and has the 8-property toolkit: it is an abstract vector space by assumption.

Slot racer on a track. To illustrate the idea, consider a problem in planar kinematics and a laboratory data recorder that approximates the x, y, z location of an object in 3-dimensional space. The recorder puts the data set of the kinematics problem into fixed 3-vectors. After the recording, the data analysis begins.

From the beginning, the kinematics problem is planar, and we should have done the data recording using 2-vectors. However, the plane of action may not be nicely aligned with the axes set up by the data recorder, and this spin on the experiment causes the 3-dimensional recording.

The kinematics problem and its algebraic structure are exactly planar, but the geometry for the recorder data may be opaque. For instance, the experiment's acquisition plane might be given approximately by a **homogeneous restriction equation** like

$$x + 2y - 1000z = 0.$$

The **restriction equation** is preserved by $\boxed{+}$ and $\boxed{\cdot}$ (details later). Then data analysis on the smaller planar data set can proceed to use the toolkit at will, knowing that all calculations will be in the plane, hence physically relevant to the original kinematics problem.

Physical data in reality contains errors, preventing the data from exactly satisfying an ideal restriction equation like $x+2y-1000z = 0$. Methods like **least squares** can construct the idealized equations. The physical data is then converted by projection, making a new data set S that exactly satisfies the restriction equation $x + 2y - 1000z = 0$. It is this modified set S, the working data set of the application, that we call a subspace.

Definition 2 (Subspace)
A subset S of an abstract vector space V is called a **subspace** if it is a nonempty vector space under the operations of addition and scalar multiplication inherited from V.

In applications, a subspace S of V is a smaller data set, recorded using the same data packages as V. The smaller set S contains at least the zero vector $\vec{\mathbf{0}}$. Required is that the algebraic operations of addition and scalar multiplication acting on S give answers back in S. Then the entire 8-property toolkit is available for calculations in the smaller data set S. Applied scientists view subspaces as **working sets**, which are actively constructed and rarely discovered without mathematical effort. The construction is guided by the subspace criterion.

Theorem 1 (Subspace Criterion)
Assume vector space V is equipped with addition $(+)$ and scalar multiplication $(\cdot)$. A subset S is a subspace of V provided these checkpoints hold:

- Vector $\vec{\mathbf{0}}$ is in S (S is nonvoid).
- For each pair $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ in S, the vector $\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2$ is in S.
- For each $\vec{\mathbf{v}}$ in S and constant c, the combination $c\vec{\mathbf{v}}$ belongs to S.

Actual use of the subspace criterion is rare, because most applications define a subspace S by a *restriction* on elements of V, normally realized as a set of linear homogeneous equations. Such systems can be re-written as a matrix equation $A\vec{\mathbf{u}} = \vec{\mathbf{0}}$. To illustrate, $x + y + z = 0$ is re-written as a matrix equation as follows:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Theorem 2 (Subspaces and Restriction Equations)
Let V be one of the vector spaces $\mathcal{R}^n$ and let A be an $m \times n$ matrix. Define the data set

$$S = \{\vec{\mathbf{v}} \;:\; \vec{\mathbf{v}} \text{ in } V \text{ and } A\vec{\mathbf{v}} = \vec{\mathbf{0}}\}.$$

Then S is a subspace of V, that is, operations of addition and scalar multiplication applied to data in S give data back in S and the 8-property toolkit applies to S-data.[2]

Proof: Zero is in S because $A\vec{\mathbf{0}} = \vec{\mathbf{0}}$ for any matrix A. To verify the subspace criterion, we verify that, for $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ in S, the vector $\vec{\mathbf{z}} = c_1\vec{\mathbf{x}} + c_2\vec{\mathbf{y}}$ also belongs to S. The details:

$$\begin{aligned} A\vec{\mathbf{z}} &= A(c_1\vec{\mathbf{x}} + c_2\vec{\mathbf{y}}) \\ &= A(c_1\vec{\mathbf{x}}) + A(c_2\vec{\mathbf{y}}) \\ &= c_1A\vec{\mathbf{x}} + c_2A\vec{\mathbf{y}} \\ &= c_1\vec{\mathbf{0}} + c_2\vec{\mathbf{0}} && \text{Because } A\vec{\mathbf{x}} = A\vec{\mathbf{y}} = \vec{\mathbf{0}}, \text{ due to } \vec{\mathbf{x}}, \vec{\mathbf{y}} \text{ in } S. \\ &= \vec{\mathbf{0}} && \text{Therefore, } A\vec{\mathbf{z}} = \vec{\mathbf{0}}, \text{ and } \vec{\mathbf{z}} \text{ is in } S. \end{aligned}$$

The proof is complete.

When does Theorem 2 apply? Briefly, the kernel theorem hypothesis requires V to be a space of fixed vectors and S a subset defined by homogeneous restriction equations. A vector space of functions, used as data sets in differential equations, does not satisfy the hypothesis of Theorem 2, because V is not one of the spaces $\mathcal{R}^n$. This is why a subspace check for a function space uses the basic subspace criterion, and not Theorem 2.

How to apply Theorem 2. We illustrate with V the vector space $\mathcal{R}^4$ of all fixed 4-vectors with components x_1, x_2, x_3, x_4. Let S be the subset of V defined by the *restriction equation* $x_4 = 0$.

The matrix equation $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ of the theorem can be taken to be

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

[2] This key theorem is sometimes called the **kernel theorem**, because solutions $\vec{x}$ of $A\vec{x} = \vec{0}$ define the **kernel** of A.

The theorem applies to conclude that S is a subspace of V. The reader should be delighted that the application of the theorem involves no abstract proof details. The only details are conversion of a system of equations to matrix form.

When is S not a subspace? The following test enumerates three common conditions for which S fails to pass the subspace test. It is justified from the subspace criterion.

Theorem 3 (Testing S not a Subspace)
Let V be an abstract vector space and assume S is a subset of V. Then S is not a subspace of V provided one of the following holds.

(1) The vector 0 is not in S.

(2) Some $\vec{\mathbf{x}}$ and $-\vec{\mathbf{x}}$ are not both in S.

(3) Vector $\vec{\mathbf{x}} + \vec{\mathbf{y}}$ is not in S for some $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ in S.

Linear Combinations and Closure.

Definition 3 (Linear Combination)
A **linear combination** of vectors $\vec{\mathbf{v}}_1,\ldots,\vec{\mathbf{v}}_k$ is defined to be a sum

$$\vec{\mathbf{x}} = c_1\vec{\mathbf{v}}_1 + \cdots + c_k\vec{\mathbf{v}}_k,$$

where $c_1,\ldots,c_k$ are constants.

The **closure** property for a subspace S can be stated as *linear combinations of vectors in S are again in S*. Therefore, according to the subspace criterion, S is a subspace of V provided $\vec{\mathbf{0}}$ is in S and S is closed under the operations $+$ and $\cdot$ inherited from the larger data set V.

Definition 4 (Span)
Let vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k$ be given in a vector space V. The subset S of V consisting of all linear combinations $\vec{\mathbf{v}} = c_1\vec{\mathbf{v}}_1 + \cdots + c_k\vec{\mathbf{v}}_k$ is called the **span** of the vectors $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k$ and written

$$S = \mathbf{span}(\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k).$$

Theorem 4 (A Span of Vectors is a Subspace)
A subset $S = \mathbf{span}(\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k)$ is a subspace of V.

Proof: Details will be supplied for $k = 3$, because the text of the proof can be easily edited to give the details for general k. The vector space V is an abstract vector space, and we do not assume that the vectors are fixed vectors. Let $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ be given vectors in V and let

$$S = \mathbf{span}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3) = \{\vec{\mathbf{v}} \ : \ \vec{\mathbf{v}} = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3\}.$$

The subspace criterion will be applied to prove that S is a subspace of V.

(1) We show $\vec{\mathbf{0}}$ is in S. Choose $c_1 = c_2 = c_3 = 0$, then $\vec{\mathbf{v}} = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3 = \vec{\mathbf{0}}$. Therefore, $\vec{\mathbf{0}}$ is in S.

(2) Assume $\vec{\mathbf{v}} = a_1\vec{\mathbf{v}}_1 + a_2\vec{\mathbf{v}}_2 + a_3\vec{\mathbf{v}}_3$ and $\vec{\mathbf{w}} = b_1\vec{\mathbf{v}}_1 + b_2\vec{\mathbf{v}}_2 + b_3\vec{\mathbf{v}}_3$ are in S. We show that $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ is in S, by adding the equations:

$$\begin{aligned} \vec{\mathbf{v}} + \vec{\mathbf{w}} &= a_1\vec{\mathbf{v}}_1 + a_2\vec{\mathbf{v}}_2 + a_3\vec{\mathbf{v}}_3 + b_1\vec{\mathbf{v}}_1 + b_2\vec{\mathbf{v}}_2 + b_3\vec{\mathbf{v}}_3 \\ &= (a_1 + b_1)\vec{\mathbf{v}}_1 + (a_2 + b_2)\vec{\mathbf{v}}_2 + (a_3 + b_3)\vec{\mathbf{v}}_3 \\ &= c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3 \end{aligned}$$

where the constants are defined by $c_1 = a_1 + b_1$, $c_2 = a_2 + b_2$, $c_3 = a_3 + b_3$. Then $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ is in S.

(3) Assume $\vec{\mathbf{v}} = a_1\vec{\mathbf{v}}_1 + a_2\vec{\mathbf{v}}_2 + a_3\vec{\mathbf{v}}_3$ and c is a constant. We show $c\vec{\mathbf{v}}$ is in S. Multiply the equation for $\vec{\mathbf{v}}$ by c to obtain

$$\begin{aligned} c\vec{\mathbf{v}} &= ca_1\vec{\mathbf{v}}_1 + ca_2\vec{\mathbf{v}}_2 + ca_3\vec{\mathbf{v}}_3 \\ &= c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3 \end{aligned}$$

where the constants are defined by $c_1 = ca_1$, $c_2 = ca_2$, $c_3 = ca_3$. Then $c\vec{\mathbf{v}}$ is in S.

The proof is complete.

The Parking Lot Analogy. A useful visualization for *vector space* and *subspace* is a parking lot with valet parking. The large lot represents the **storage space** of the larger data set associated with a vector space V. The parking lot rules, such as *display your ticket*, *park between the lines*, correspond to the toolkit of 8 vector space rules. The valet parking lot S, which is a smaller roped-off area within the larger lot V, is also storage space, subject to the same rules as the larger lot. The smaller data set S corresponds to a subspace of V. Just as additional restrictions apply to the valet lot, a subspace S is generally defined by equations, relations or restrictions on the data items of V.

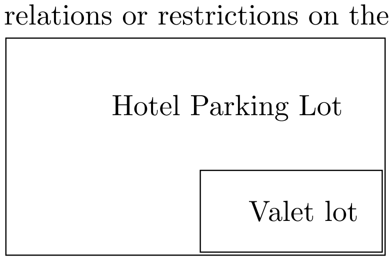

Figure 5. Parking lot analogy. An abstract vector space V and one of its subspaces S can be visualized through the analogy of a parking lot (V) containing a valet lot (S).

Vector Algebra. The **norm** or **length** of a fixed vector $\vec{X}$ with components $x_1, \ldots, x_n$ is given by the formula

$$|\vec{X}| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

This measurement can be used to quantify the numerical error between two data sets stored in vectors $\vec{X}$ and $\vec{Y}$:

$$\textbf{norm-error} = |\vec{X} - \vec{Y}|.$$

The **dot product** $\vec{X} \cdot \vec{Y}$ of two fixed vectors $\vec{X}$ and $\vec{Y}$ is defined by

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n.$$

If $n = 3$, then $|\vec{X}||\vec{Y}| \cos\theta = \vec{X} \cdot \vec{Y}$ where θ is the **angle between** $\vec{X}$ and $\vec{Y}$. In analogy, two n-vectors are said to be **orthogonal** provided $\vec{X} \cdot \vec{Y} = 0$. It is usual to require that $|\vec{X}| > 0$ and $|\vec{Y}| > 0$ when talking about the angle θ between vectors, in which case we *define* θ to be the angle θ $(0 \le \theta \le \pi)$ satisfying

$$\cos\theta = \frac{\vec{X} \cdot \vec{Y}}{|\vec{X}||\vec{Y}|}.$$

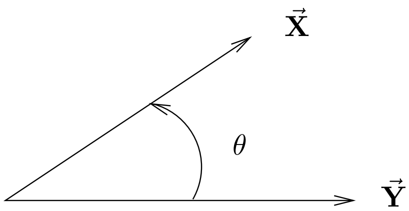

Figure 6. Angle θ between two vectors $\vec{\mathbf{X}}$, $\vec{\mathbf{Y}}$.

The **shadow projection** of vector $\vec{X}$ onto the direction of vector $\vec{Y}$ is the number d defined by

$$d = \frac{\vec{X} \cdot \vec{Y}}{|\vec{Y}|}.$$

The triangle determined by $\vec{X}$ and $(d/|\vec{Y}|)\vec{Y}$ is a right triangle.

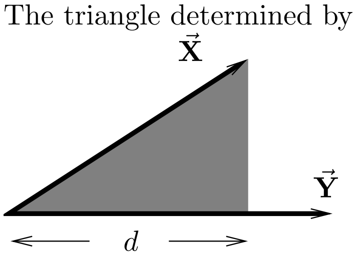

Figure 7. Shadow projection d of vector $\vec{\mathbf{X}}$ onto the direction of vector $\vec{\mathbf{Y}}$.

The **vector projection** of $\vec{X}$ onto the line L through the origin in the direction of $\vec{Y}$ is defined by

$$\mathbf{proj}_{\vec{Y}}(\vec{X}) = d\frac{\vec{Y}}{|\vec{Y}|} = \frac{\vec{X} \cdot \vec{Y}}{\vec{Y} \cdot \vec{Y}}\vec{Y}.$$

The **vector reflection** of vector $\vec{X}$ in the line L through the origin having the direction of vector $\vec{Y}$ is defined to be the vector

$$\mathbf{refl}_{\vec{Y}}(\vec{X}) = 2\,\mathbf{proj}_{\vec{Y}}(\vec{X}) - \vec{X} = 2\frac{\vec{X} \cdot \vec{Y}}{\vec{Y} \cdot \vec{Y}}\vec{Y} - \vec{X}.$$

It is the formal analog of the complex conjugate map $a + ib \to a - ib$ with the x-axis replaced by line L.

Matrices are Vector Packages. A **matrix** A is a package of so many fixed vectors, considered together, and written as a 2-dimensional array

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

The packaging can be in terms of **column vectors** or **row vectors**:

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{pmatrix} \cdots \begin{pmatrix} a_{1m} \\ a_{2m} \\ \cdots \\ a_{nm} \end{pmatrix} \quad \text{or} \quad \begin{cases} (a_{11}, a_{12}, \ldots, a_{1n}) \\ (a_{21}, a_{22}, \ldots, a_{2n}) \\ \vdots \\ (a_{m1}, a_{m2}, \ldots, a_{mn}) \end{cases}.$$

Equality of matrices. Two matrices A and B are said to be **equal** provided they have identical row and column dimensions and corresponding entries are equal. Equivalently, A and B are equal if they have identical columns, or identical rows.

Mailbox analogy. A matrix A can be visualized as a rectangular collection of so many mailboxes labeled (i, j) with contents a_{ij}, where the row index is i and the column index is j; see Table 2.

Table 2. The mailbox analogy for matrices.
A matrix A is visualized as a block of mailboxes, each located by row index i and column index j. The box at (i, j) contains data a_{ij}.

a_{11}	a_{12}	$\cdots$	a_{1n}
a_{21}	a_{22}	$\cdots$	a_{2n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_{m1}	a_{m2}	$\cdots$	a_{mn}

Computer Storage. Computer programs store matrices as a long single array. Array contents are fetched by computing the index into the long array followed by retrieval of the numeric content a_{ij}. From a computer viewpoint, vectors and matrices are the same objects.

For instance, a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be stored by stacking its rows into a column vector, the mathematical equivalent being the one-to-one and onto mapping

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longleftrightarrow \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

This mapping uniquely associates the 2×2 matrix A with a vector in $\mathcal{R}^4$. Similarly, a matrix of size $m \times n$ is associated with a column vector in $\mathcal{R}^k$, where $k = mn$.

Matrix Addition and Scalar Multiplication. Addition of two matrices is defined by applying fixed vector addition on corresponding columns. Similarly, an organization by rows leads to a second definition of matrix addition, which is exactly the same:

$$\begin{pmatrix} a_{11} \cdots a_{1n} \\ a_{21} \cdots a_{2n} \\ \vdots \\ a_{m1} \cdots a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} \cdots b_{1n} \\ b_{21} \cdots b_{2n} \\ \vdots \\ b_{m1} \cdots b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots a_{1n}+b_{1n} \\ a_{21}+b_{21} & \cdots a_{2n}+b_{2n} \\ & \vdots \\ a_{m1}+b_{m1} & \cdots a_{mn}+b_{mn} \end{pmatrix}.$$

Scalar multiplication of matrices is defined by applying scalar multiplication to the columns or rows:

$$k \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ka_{11} & \cdots & ka_{1n} \\ ka_{21} & \cdots & ka_{2n} \\ & \vdots & \\ ka_{m1} & \cdots & ka_{mn} \end{pmatrix}.$$

Both operations on matrices are motivated by considering a matrix to be a long single array or *fixed vector*, to which the standard fixed vector definitions are applied. The operation of addition is properly defined exactly when the two matrices have the same row and column dimensions.

Digital Photographs. A digital camera stores image sensor data as a matrix A of numbers corresponding to the color and intensity of tiny sensor sites called **pixels** or **dots**. The pixel position in the print is given by row and column location in the matrix A.

A visualization of the image sensor is a checkerboard. Each square is stacked with a certain number of checkers, the count proportional to the number of electrons knocked loose by light falling on the photodiode site.[3]

In 24-bit color, a pixel could be represented in matrix A by a coded integer $a = r + (2^8)g + (2^{16})b$. Symbols r, g, b are integers between 0 and 255 which represent the intensity of colors red, green and blue, respectively. For example, $r = g = b = 0$ is the color **black** while

[3]Some digital cameras have three image sensors, one for each of colors red, green and blue (RGB). Other digital cameras integrate the three image sensors into one array, interpolating color-filtered sites to obtain the color data.

$r = g = b = 255$ is the color **white**. Grander schemes exist, e.g., 32-bit and 128-bit color.[4]

Matrix addition can be visualized through matrices representing color separations.[5] When three monochrome transparencies of colors red, green and blue (RGB) are projected simultaneously by a projector, the colors add to make a full color screen projection. The three transparencies can be associated with matrices R, G, B which contain pixel data for the monochrome images. Then the projected image is associated with the matrix sum $R + G + B$.

Scalar multiplication of matrices has a similar visualization. The pixel information in a monochrome image (red, green or blue) is coded for intensity. The associated matrix A of pixel data when multiplied by a scalar k gives a new matrix kA of pixel data with the intensity of each pixel adjusted by factor k. The photographic effect is to adjust the range of intensities. In the checkerboard visualization of an image sensor, factor k increases or decreases the checker stack height at each square.

Color Separation Illustration. Consider the coded matrix

$$\vec{\mathbf{X}} = \begin{pmatrix} 514 & 3 \\ 131843 & 197125 \end{pmatrix}.$$

We will determine the monochromatic pixel data $\vec{\mathbf{R}}$, $\vec{\mathbf{G}}$, $\vec{\mathbf{B}}$ in the equation $\vec{\mathbf{X}} = \vec{\mathbf{R}} + 2^8\vec{\mathbf{G}} + 2^{16}\vec{\mathbf{B}}$.

First we decode the scalar equation $x = r + 2^8 g + 2^{16} b$ by these algebraic steps, which use the modulus function $\mathbf{mod}(x, m)$, defined to be the remainder after division of x by m. We assume r, g, b are integers in the range 0 to 255.

$y = \mathbf{mod}(x, 2^{16})$	The remainder should be $y = r + 2^8 g$.
$r = \mathbf{mod}(y, 2^8)$	Because $y = r + 2^8 g$, the remainder equals r.
$g = (y - r)/2^8$	Divide $y - r = 2^8 g$ by 2^8 to obtain g.
$b = (x - y)/2^{16}$	Because $x - y = x - r - 2^8 g$ has remainder b.
$r + 2^8 g + 2^{16} b$	Answer check. This should equal x.

[4] A typical beginner's digital camera makes low resolution color photos using 24-bit color. The photo is constructed of 240 rows of dots with 320 dots per row. The associated storage matrix A is of size 240×320. The identical small format is used for video clips at up to 30 frames per second in video-capable digital cameras.

The storage format **BMP** stores data as bytes, in groups of three b, g, r, starting at the lower left corner of the photo. Therefore, 240×320 photos have $230,400$ data bytes. The storage format **JPEG** reduces file size by compression and quality loss.

[5] James Clerk Maxwell is credited with the idea of color separation.

Computer algebra systems can provide an answer for matrices R, G, B by duplicating the scalar steps. Below is a `maple` implementation that gives the answers

$$R = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}, G = \begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 2 & 3 \end{pmatrix}.$$

```
with(LinearAlgebra:-Modular):
X:=Matrix([[514,3],[131843,197125]]);
Y:=Mod(2^16,X,integer); # y=mod(x,65536)
R:=Mod(2^8,Y,integer);  # r=mod(y,256)
G:=(Y-R)/2^8;           # g=(y-r)/256
B:=(X-Y)/2^16;          # b=(x-y)/65536
X-(R+G*2^8+B*2^16);     # answer check
```

The result can be visualized through a checkerboard of 4 squares. The second square has 5 red, 2 green and 3 blue checkers stacked, representing the color $x = (5) + 2^8(2) + 2^{16}(3)$ - see Figure 8. A matrix of size $m \times n$ is visualized as a checkerboard with mn squares, each square stacked with red, green and blue checkers.

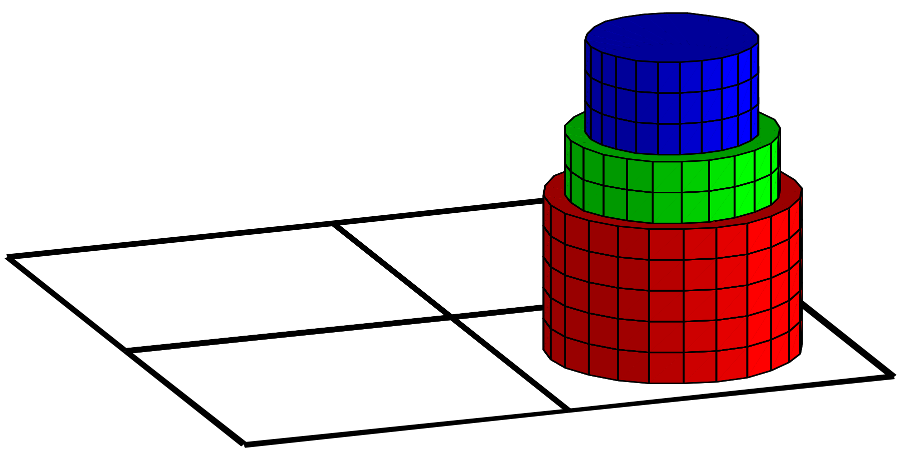

Figure 8. Checkerboard visualization.
Illustrated is a stack of checkers, representing one photodiode site on an image sensor inside a digital camera. There are 5 red, 2 green and 3 blue checkers stacked on one square. The checkers represent the number of electrons knocked loose by photons falling on each RGB-filtered site.

Matrix Multiply. College algebra texts cite the definition of matrix multiplication as *the product AB equals a matrix C given by the relations*

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}, \quad 1 \le i \le m, \ 1 \le j \le k.$$

Below, we motivate the definition of matrix multiplication from an applied point of view, based upon familiarity with the dot product. Microcode implementations in vector supercomputers make use of a similar viewpoint.

Matrix multiplication as a dot product extension. To illustrate the basic idea by example, let

$$A = \begin{pmatrix} -1 & 2 & 1 \\ 3 & 0 & -3 \\ 4 & -2 & 5 \end{pmatrix}, \quad \vec{X} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

The product equation $A\vec{X}$ is displayed as the *dotless juxtaposition*

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

which represents an *unevaluated request* to **gang** the dot product operation onto the rows of the matrix on the left:

$$(-1\,2\,1) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = 3, \quad (3\,0\,-3) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = -3, \quad (4\,-2\,5) \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = 21.$$

The *evaluated request* produces a column vector containing the dot product answers, called the **product of a matrix and a vector** (no mention of dot product), written as

$$\begin{pmatrix} -1 & 2 & 1 \\ 3 & 0 & -3 \\ 4 & -2 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ 21 \end{pmatrix}.$$

The general scheme which gangs the dot product operation onto the matrix rows can be written as

$$\begin{pmatrix} \cdots & \text{row 1} & \cdots \\ \cdots & \text{row 2} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & \text{row m} & \cdots \end{pmatrix} \vec{X} = \begin{pmatrix} (\text{row 1}) \cdot \vec{X} \\ (\text{row 2}) \cdot \vec{X} \\ \vdots \\ (\text{row m}) \cdot \vec{X} \end{pmatrix}.$$

The product is properly defined only in case the number of matrix columns equals the number of entries in $\vec{X}$, so that the dot products on the right are defined.

Matrix multiply as a linear combination of columns. The identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} a \\ c \end{pmatrix} + x_2 \begin{pmatrix} b \\ d \end{pmatrix}$$

implies that $A\vec{\mathbf{x}}$ is a linear combination of the columns of A, where A is the 2×2 matrix on the left.

This result holds in general. Assume $A = \mathbf{aug}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n)$ and $\vec{X}$ has components x_1, ..., x_n. Then the definition of matrix multiply implies

$$A\vec{X} = x_1\vec{\mathbf{v}}_1 + x_2\vec{\mathbf{v}}_2 + \cdots + x_n\vec{\mathbf{v}}_n.$$

This relation is used so often, that we record it as a formal result.

Theorem 5 (Linear Combination of Columns)
The product of a matrix A and a vector $\vec{\mathbf{x}}$ satisfies

$$A\vec{\mathbf{x}} = x_1\,\mathbf{col}(A,1) + \cdots + x_n\,\mathbf{col}(A,n)$$

where $\mathbf{col}(A,i)$ denotes column i of matrix A, and symbols x_1, ..., x_n are the components of vector $\vec{\mathbf{x}}$.

General matrix product AB. The evaluation of matrix products $A\vec{Y}_1$, $A\vec{Y}_2$, ..., $A\vec{Y}_k$ is a list of k column vectors which can be packaged into a matrix C. Let B be the matrix which packages the columns $\vec{Y}_1$, ..., $\vec{Y}_k$. *Define* $C = AB$ by the dot product definition

$$c_{ij} = \mathbf{row}(A,i) \cdot \mathbf{col}(B,j).$$

This definition makes sense provided the column dimension of A matches the row dimension of B. It is consistent with the earlier definition from college algebra and the definition of $A\vec{Y}$, therefore it may be taken as *the basic definition for a matrix product.*

How to multiply matrices on paper. Most persons make arithmetic errors when computing dot products

$$\begin{pmatrix} -7 & 3 & 5 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix} = -9,$$

because alignment of corresponding entries must be done mentally. It is visually easier when the entries are aligned.

On paper, persons often arrange their work for a matrix times a vector as below, so that the entries align. The boldface transcription above the columns is temporary, erased after the dot product step.

$$\begin{array}{c} \begin{matrix} \mathbf{-1} & \mathbf{3} & \mathbf{-5} \end{matrix} \\ \begin{pmatrix} -7 & 3 & 5 \\ -5 & -2 & 3 \\ 1 & -3 & -7 \end{pmatrix} \end{array} \cdot \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix} = \begin{pmatrix} -9 \\ -16 \\ 25 \end{pmatrix}$$

Visualization of matrix multiply. We give a key example of how to interpret 2×2 matrix multiply as a geometric operation.

Let's begin by inspecting a 2×2 system $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$ for its geometric meaning. Consider the system

$$(2) \qquad \left|\begin{array}{lcl} y_1 &=& ax_1 + bx_2 \\ y_2 &=& cx_1 + dx_2 \end{array}\right| \quad \text{or} \quad \vec{\mathbf{y}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \vec{\mathbf{x}}$$

Geometric rotation and scaling of planar figures have equations of this form.

$$(3) \qquad \begin{array}{cc} \textbf{Rotation by angle } \theta & \textbf{Scale by factor } k \\ \hline A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} & B = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \end{array}$$

The geometric effect of mapping points $\vec{\mathbf{x}}$ on an ellipse by the equation $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$ is to rotate the ellipse. If we choose $\theta = \pi/2$, then it is a rotation by 90 degrees. The mapping $\vec{\mathbf{z}} = B\vec{\mathbf{y}}$ re-scales the axes by factor k. If we choose $k = 2$, then the geometric result is to double the dimensions of the rotated ellipse. The resulting geometric transformation of $\vec{\mathbf{x}}$ into $\vec{\mathbf{z}}$ has algebraic realization

$$\vec{\mathbf{z}} = B\vec{\mathbf{y}} = BA\vec{\mathbf{x}},$$

which means the composite transformation of rotation followed by scaling is represented by system (2), with coefficient matrix

$$BA = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \cos\pi/2 & \sin\pi/2 \\ -\sin\pi/2 & \cos\pi/2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$

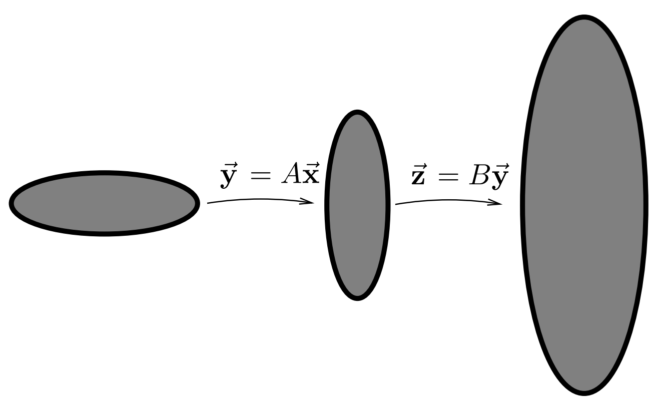

Figure 9. An ellipse is mapped into a rotated and re-scaled ellipse. The rotation is $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$, which is followed by re-scaling $\vec{\mathbf{z}} = B\vec{\mathbf{y}}$. The composite geometric transformation is $\vec{\mathbf{z}} = BA\vec{\mathbf{x}}$, which maps the ellipse into a rotated and re-scaled ellipse.

Special Matrices. The **zero matrix**, denoted $\mathbf{0}$, is the $m \times n$ matrix all of whose entries are zero. The **identity matrix**, denoted I, is the $n \times n$ matrix with ones on the diagonal and zeros elsewhere: $a_{ij} = 1$ for $i = j$ and $a_{ij} = 0$ for $i \neq j$.

$$\mathbf{0} = \begin{pmatrix} 0\,0 \cdots 0 \\ 0\,0 \cdots 0 \\ \vdots \\ 0\,0 \cdots 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1\,0 \cdots 0 \\ 0\,1 \cdots 0 \\ \vdots \\ 0\,0 \cdots 1 \end{pmatrix}.$$

The identity I is a package of column vectors called the **standard unit vectors** of size n. Literature may write the columns of I as $\vec{e}_1, \ldots, \vec{e}_n$ or as $\mathbf{col}(I, 1), \ldots, \mathbf{col}(I, n)$.

The **negative** of a matrix A is $(-1)A$, which multiplies each entry of A by the factor (-1):

$$-A = \begin{pmatrix} -a_{11} \cdots -a_{1n} \\ -a_{21} \cdots -a_{2n} \\ \vdots \\ -a_{m1} \cdots -a_{mn} \end{pmatrix}.$$

Square Matrices. An $n \times n$ matrix A is said to be **square**. The entries a_{kk}, $k = 1, \ldots, n$ of a square matrix make up its **diagonal**. A square matrix A is **lower triangular** if $a_{ij} = 0$ for $i > j$, and **upper triangular** if $a_{ij} = 0$ for $i < j$; it is **triangular** if it is either upper or lower triangular. Therefore, an upper triangular matrix has all zeros below the diagonal and a lower triangular matrix has all zeros above the diagonal. A square matrix A is a **diagonal matrix** if $a_{ij} = 0$ for $i \neq j$, that is, the off-diagonal elements are zero. A square matrix A is a **scalar matrix** if $A = cI$ for some constant c.

$$\begin{array}{ll} \text{upper triangular} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, & \text{lower triangular} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \\ \text{diagonal} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, & \text{scalar} = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & c \end{pmatrix}. \end{array}$$

Matrix Algebra. A matrix can be viewed as a single long array, or fixed vector, therefore the toolkit for fixed vectors applies to matrices.

Let A, B, C be matrices of the same row and column dimensions and let k_1, k_2, k be constants. Then

Closure	The operations $A+B$ and kA are defined and result in a new matrix of the same dimensions.	
Addition rules	$A+B=B+A$	commutative
	$A+(B+C)=(A+B)+C$	associative
	Matrix $\mathbf{0}$ is defined and $\mathbf{0}+A=A$	zero
	Matrix $-A$ is defined and $A+(-A)=\mathbf{0}$	negative
Scalar multiply rules	$k(A+B)=kA+kB$	distributive I
	$(k_1+k_2)A=k_1A+k_2B$	distributive II
	$k_1(k_2A)=(k_1k_2)A$	distributive III
	$1\,A=A$	identity

These rules collectively establish that the set of all $m\times n$ matrices is an abstract vector space.

The operation of matrix multiplication gives rise to some new matrix rules, which are in common use, but do not qualify as vector space rules. The rules are proved by expansion of each side of the equation. Techniques are sketched in the exercises, which carry out the steps of each proof.

Associative	$A(BC)=(AB)C$, provided products BC and AB are defined.
Distributive	$A(B+C)=AB+AC$, provided products AB and AC are defined.
Right Identity	$AI=A$, provided AI is defined.
Left Identity	$IA=A$, provided IA is defined.

Transpose. Swapping rows and columns of a matrix A results in a new matrix B whose entries are given by $b_{ij}=a_{ji}$. The matrix B is denoted A^T (pronounced "A-*transpose*"). The transpose has the following properties. See the exercises for proofs.

$(A^T)^T=A$	Identity
$(A+B)^T=A^T+B^T$	Sum
$(AB)^T=B^TA^T$	Product
$(kA)^T=kA^T$	Scalar

Inverse Matrix. A square matrix B is said to be an **inverse** of a square matrix A provided $AB=BA=I$. The symbol I is the identity matrix of matching dimension. To illustrate,

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

which implies that $B = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ is an inverse of $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

A given matrix A may not have an inverse, for example, $\mathbf{0}$ times any square matrix B is $\mathbf{0}$, which prohibits a relation $\mathbf{0}B = B\mathbf{0} = I$. When A does have an inverse B, then the notation A^{-1} is used for B, hence $AA^{-1} = A^{-1}A = I$. The following properties of inverses will be proved on page 310.

Theorem 6 (Inverses)
Let A, B, C denote square matrices. Then

(a) A matrix has at most one inverse, that is, if $AB = BA = I$ and $AC = CA = I$, then $B = C$.

(b) If A has an inverse, then so does A^{-1} and $(A^{-1})^{-1} = A$.

(c) If A has an inverse, then $(A^{-1})^T = (A^T)^{-1}$.

(d) If A and B have inverses , then $(AB)^{-1} = B^{-1}A^{-1}$.

Left to be discussed is how to find the inverse A^{-1}. For a 2×2 matrix, there is an easily justified formula.

Theorem 7 (Inverse of a 2×2)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The formula is commonly committed to memory, because of repeated use. In words, the theorem says:

> Swap the diagonal entries, change signs on the off-diagonal entries, then divide by the determinant $ad - bc$.

There is a generalization of this formula to $n \times n$ matrices, which is equivalent to the formulas in **Cramer's rule**. It will be derived during the study of determinants; the statement is paraphrased as follows:

$$A^{-1} = \frac{\text{adjugate matrix of } A}{\text{determinant of } A}.$$

A general and efficient method for computing inverses, based upon **rref** methods, will be presented in the next section. The method can be implemented on hand calculators, computer algebra systems and computer numerical laboratories.

Symmetric Matrix. A matrix A is said to be **symmetric** if $A^T = A$, which implies that the row and column dimensions of A are the same and

$a_{ij} = a_{ji}$. If A is symmetric and invertible, then its inverse is symmetric. If B is any matrix, not necessarily square, then $A = B^T B$ is symmetric. See the exercises for proofs.

Proof of Theorem 6:

(a) If $AB = BA = I$ and $AC = CA = I$, then $B = BI = BAC = IC = C$.
(b) Let $B = A^{-1}$. Given $AB = BA = I$, then by definition A is an inverse of B, but by (a) it is the only one, so $(A^{-1})^{-1} = B^{-1} = A$.
(c) Let $B = A^{-1}$. We show $B^T = (A^T)^{-1}$ or equivalently $C = B^T$ satisfies $A^T C = CA^T = I$. Start with $AB = BA = I$, take the transpose to get $B^T A^T = A^T B^T = I$. Substitute $C = B^T$, then $CA^T = A^T C = I$, which was to be proved.
(d) The formula is proved by showing that $C = B^{-1}A^{-1}$ satisfies $(AB)C = C(AB) = I$. The left side is $(AB)C = ABB^{-1}A^{-1} = I$ and the right side $C(AB) = B^{-1}A^{-1}AB = I$, proving LHS = RHS.

Exercises 5.1

Fixed vectors. Perform the indicated operation(s).

1. $\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \end{pmatrix}$

2. $\begin{pmatrix} 2 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 \\ -3 \end{pmatrix}$

3. $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$

4. $\begin{pmatrix} 2 \\ -2 \\ 9 \end{pmatrix} - \begin{pmatrix} 1 \\ -3 \\ 7 \end{pmatrix}$

5. $2\begin{pmatrix} 1 \\ -1 \end{pmatrix} + 3\begin{pmatrix} -2 \\ 1 \end{pmatrix}$

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

7. $5\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + 3\begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$

8. $3\begin{pmatrix} 2 \\ -2 \\ 9 \end{pmatrix} - 5\begin{pmatrix} 1 \\ -3 \\ 7 \end{pmatrix}$

9. $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$

10. $\begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$

Parallelogram Rule. Determine the resultant vector in two ways: (a) the parallelogram rule, and (b) fixed vector addition.

11. $\begin{pmatrix} 2 \\ -2 \end{pmatrix} + \begin{pmatrix} 1 \\ -3 \end{pmatrix}$

12. $(2\vec{\imath} - 2\vec{\jmath}) + (\vec{\imath} - 3\vec{\jmath})$

13. $\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}$

14. $(2\vec{\imath} - 2\vec{\jmath} + 3\vec{k}) + (\vec{\imath} - 3\vec{\jmath} - \vec{k})$

Toolkit. Let V be the data set of all fixed 2-vectors, $V = \mathcal{R}^2$. Define addition and scalar multiplication componentwise. Verify the following toolkit rules by direct computation.

15. (Commutative)
$\vec{X} + \vec{Y} = \vec{Y} + \vec{X}$

16. (Associative)
$\vec{X} + (\vec{Y} + \vec{Z}) = (\vec{Y} + \vec{X}) + \vec{Z}$

17. (Zero)
Vector $\vec{0}$ is defined and $\vec{0} + \vec{X} = \vec{X}$

18. (Negative)
Vector $-\vec{X}$ is defined and
$\vec{X} + (-\vec{X}) = \vec{0}$

19. (Distributive I)
$k(\vec{X} + \vec{Y}) = k\vec{X} + k\vec{Y}$

20. (Distributive II)
$(k_1 + k_2)\vec{X} = k_1\vec{X} + k_2\vec{X}$

21. (Distributive III)
$k_1(k_2\vec{X}) = (k_1k_2)\vec{X}$

22. (Identity)
$1\vec{X} = \vec{X}$

Subspaces. Verify that the given restriction equation defines a subspace S of $V = \mathcal{R}^3$. Use Theorem 2, page 296.

23. $z = 0$

24. $y = 0$

25. $x + z = 0$

26. $2x + y + z = 0$

27. $x = 2y + 3z$

28. $x = 0$, $z = x$

29. $z = 0$, $x + y = 0$

30. $x = 3z - y$, $2x = z$

31. $x + y + z = 0$, $x + y = 0$

32. $x + y - z = 0$, $x - z = y$

Not a Subspace. Test the following restiction equations for $V = \mathcal{R}^3$ and show that the corresponding subset S is not a subspace of V.

33. $x = 1$

34. $x + z = 1$

35. $xz = 2$

36. $xz + y = 1$

37. $xz + y = 0$

38. $xyz = 0$

39. $z \ge 0$

40. $x \ge 0$ and $y \ge 0$

41. Octant I

42. The interior of the unit sphere

Dot Product. Find the dot product of $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.

43. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$.

44. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

45. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$.

46. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$.

47. $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ are in $\mathcal{R}^{169}$, $\vec{\mathbf{a}}$ has all components 1 and $\vec{\mathbf{b}}$ has all components -1, except four, which all equal 5.

48. $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ are in $\mathcal{R}^{200}$, $\vec{\mathbf{a}}$ has all components -1 and $\vec{\mathbf{b}}$ has all components -1 except three, which are zero.

Length of a Vector. Find the length of the vector $\vec{\mathbf{v}}$.

49. $\vec{\mathbf{v}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

50. $\vec{\mathbf{v}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

51. $\vec{\mathbf{v}} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$.

52. $\vec{\mathbf{v}} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}$.

Shadow Projection. Find the shadow projection $d = \vec{\mathbf{a}} \cdot \vec{\mathbf{b}}/|\vec{\mathbf{b}}|$.

53. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$.

54. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

55. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$.

56. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$.

Projections and Reflections. Let L denote a line through the origin with unit direction $\vec{\mathbf{u}}$.

The **projection** of vector $\vec{\mathbf{x}}$ onto L is $P(\vec{\mathbf{x}}) = d\vec{\mathbf{u}}$, where $d = \vec{\mathbf{x}} \cdot \vec{\mathbf{u}}$ is the shadow projection.

The **reflection** of vector $\vec{\mathbf{x}}$ across L is $R(\vec{\mathbf{x}}) = 2d\vec{\mathbf{u}} - \vec{\mathbf{x}}$ (a generalized complex conjugate).

57. Let $\vec{\mathbf{u}}$ be the direction of the x-axis in the plane. Establish that $P(\vec{\mathbf{x}})$ and $R(\vec{\mathbf{x}})$ are sides of a right triangle and P duplicates the complex conjugate operation $z \to \overline{z}$. Include a figure.

58. Let $\vec{\mathbf{u}}$ be any direction in the plane. Establish that $P(\vec{\mathbf{x}})$ and $R(\vec{\mathbf{x}})$ are sides of a right triangle. Draw a suitable figure, which includes $\vec{\mathbf{x}}$.

59. Let $\vec{\mathbf{u}}$ be the direction of $2\vec{\imath} + \vec{\jmath}$. Define $\vec{\mathbf{x}} = 4\vec{\imath} + 3\vec{\jmath}$. Compute the vectors $P(\vec{\mathbf{x}})$ and $R(\vec{\mathbf{x}})$.

60. Let $\vec{\mathbf{u}}$ be the direction of $\vec{\imath} + 2\vec{\jmath}$. Define $\vec{\mathbf{x}} = 3\vec{\imath} + 5\vec{\jmath}$. Compute the vectors $P(\vec{\mathbf{x}})$ and $R(\vec{\mathbf{x}})$.

Angle. Find the angle θ between the given vectors.

61. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$.

62. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

63. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$.

64. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$.

65. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$.

66. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \end{pmatrix}$.

67. $\vec{\mathbf{a}} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$.

68. $\vec{\mathbf{a}} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$ and $\vec{\mathbf{b}} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$.

Matrix Multiply. Find the given matrix product or else explain why it does not exist.

69. $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

70. $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

71. $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

72. $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

73. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$

74. $\begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$

75. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$

76. $\begin{pmatrix} 1 & 2 & 1 \\ 1 & -2 & 0 \\ 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

77. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

78. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

79. $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix}$

80. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

81. $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$

82. $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$

83. $\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

84. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$

Matrix Classification. Classify as square, non-square, upper triangular, lower triangular, scalar, diagonal, symmetric, non-symmetric. Cite as many terms as apply.

85. $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

86. $\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$

87. $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$

88. $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$

89. $\begin{pmatrix} 1 & 3 & 4 \\ 5 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

90. $\begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

91. $\begin{pmatrix} 1 & 3 & 4 \\ 3 & 2 & 0 \\ 4 & 0 & 3 \end{pmatrix}$

92. $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

93. $\begin{pmatrix} i & 0 \\ 0 & 2i \end{pmatrix}$

94. $\begin{pmatrix} i & 3 \\ 3 & 2i \end{pmatrix}$

Digital Photographs. Assume integer 24-bit color encoding $x = r + (256)g + (65536)b$, which means r units **red**, g units **green** and b units **blue**. Given matrix $\vec{\mathbf{X}} = \vec{\mathbf{R}} + 256\vec{\mathbf{G}} + 65536\vec{\mathbf{B}}$, find the red, green and blue color separation matrices $\vec{\mathbf{R}}$, $\vec{\mathbf{G}}$, $\vec{\mathbf{B}}$.

95. $\vec{\mathbf{X}} = \begin{pmatrix} 514 & 3 \\ 131843 & 197125 \end{pmatrix}$

96. $\vec{\mathbf{X}} = \begin{pmatrix} 514 & 3 \\ 131331 & 66049 \end{pmatrix}$

97. $\vec{\mathbf{X}} = \begin{pmatrix} 513 & 7 \\ 131333 & 66057 \end{pmatrix}$

98. $\vec{\mathbf{X}} = \begin{pmatrix} 257 & 7 \\ 131101 & 66057 \end{pmatrix}$

99. $\vec{\mathbf{X}} = \begin{pmatrix} 257 & 17 \\ 131101 & 265 \end{pmatrix}$

100. $\vec{\mathbf{X}} = \begin{pmatrix} 65537 & 269 \\ 65829 & 261 \end{pmatrix}$

101. $\vec{\mathbf{X}} = \begin{pmatrix} 65538 & 65803 \\ 65833 & 7 \end{pmatrix}$

102. $\vec{\mathbf{X}} = \begin{pmatrix} 259 & 65805 \\ 299 & 5 \end{pmatrix}$

Matrix Properties. Verify the result.

103. Let C be an $m \times n$ matrix. Let $\vec{X}$ be column i of the $n \times n$ identity I. Define $\vec{Y} = C\vec{X}$. Verify that $\vec{Y}$ is column i of C.

104. Let A and C be an $m \times n$ matrices such that $AC = \mathbf{0}$. Verify that each column $\vec{Y}$ of C satisfies $A\vec{Y} = \vec{\mathbf{0}}$.

105. Let A be a 2×3 matrix and let $\vec{Y}_1$, $\vec{Y}_2$, $\vec{Y}_n$ be column vectors packaged into a 3×3 matrix C. Assume each column vector $\vec{Y}_i$ satisfies the equation $A\vec{Y}_i = \vec{\mathbf{0}}$, $1 \le i \le 3$. Show that $AC = \mathbf{0}$.

106. Let A be an $m \times n$ matrix and let $\vec{Y}_1, \ldots, \vec{Y}_n$ be column vectors packaged into an $n \times n$ matrix C. Assume each column vector $\vec{Y}_i$ satisfies the equation $A\vec{Y}_i = \vec{\mathbf{0}}$, $1 \le i \le n$. Show that $AC = \mathbf{0}$.

Triangular Matrices. Verify the result.

107. The product of two upper triangular 2×2 matrices is upper triangular.

108. The product of two lower triangular 2×2 matrices is lower triangular.

109. The product of two triangular 2×2 matrices is not necessarily triangular.

110. The product of two upper triangular $n \times n$ matrices is upper triangular.

111. The product of two lower triangular $n \times n$ matrices is lower triangular.

112. The only 3×3 matrices which are both upper and lower triangular are the 3×3 diagonal matrices.

Matrix Multiply Properties. Verify the result.

113. The associative law $A(BC) = (AB)C$ holds for matrix multiplication.
Sketch: Expand $L = A(BC)$ entry L_{ij} according to matrix multiply rules. Expand $R = (AB)C$ entry R_{ij} the same way. Show $L_{ij} = R_{ij}$.

114. The distributive law $A(B+C) = AB + AC$ holds for matrices.
Sketch: Expand $L = A(B+C)$ entry L_{ij} according to matrix multiply rules. Expand $R = AB + AC$ entry R_{ij} the same way. Show $L_{ij} = a_{ik}(b_{kj} + c_{kj})$ and $R_{ij} = a_{ik}b_{kj} + a_{ik}c_{kj}$. Then $L_{ij} = R_{ij}$.

115. For any matrix A the transpose formula $(A^T)^T = A$ holds.
Sketch: Expand $L = (A^T)^T$ entry L_{ij} according to matrix transpose rules. Then $L_{ij} = a_{ij}$.

116. For matrices A, B the transpose formula $(A+B)^T = A^T + B^T$ holds.
Sketch: Expand $L = (A+B)^T$ entry L_{ij} according to matrix transpose rules. Repeat for entry R_{ij} of $R = A^T + B^T$. Show $L_{ij} = R_{ij}$.

117. For matrices A, B the transpose formula $(AB)^T = B^T A^T$ holds.
Sketch: Expand $L = (AB)^T$ entry L_{ij} according to matrix multiply and transpose rules. Repeat for entry R_{ij} of $R = B^T A^T$. Show $L_{ij} = R_{ij}$.

118. For a matrix A and constant k, the transpose formula $(kA)^T = kA^T$ holds.

Invertible Matrices. Verify the result.

119. There are infinitely many 2×2 matrices A, B such that $AB = 0$.

120. The zero matrix is not invertible.

121. The matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ is not invertible.

122. The matrix $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ is invertible.

123. The matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ satisfy

$$AB = BA = (ad - bc)I.$$

124. If $AB = 0$, then one of A or B is not invertible.

Symmetric Matrices. Verify the result.

125. The product of two symmetric $n \times n$ matrices A, B such that $AB = BA$ is symmetric.

126. The product of two symmetric 2×2 matrices may not be symmetric.

127. If A is symmetric, then so is A^{-1}.
Sketch: Let $B = A^{-1}$. Compute B^T using transpose rules.

128. If B is an $m \times n$ matrix and $A = B^T B$, then A is $n \times n$ symmetric.
Sketch: Compute A^T using transpose rules.