Contents

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 for theoretical analysis and computation 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

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

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:

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

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, $\ldots$, 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 assigns label $\vec{A}$ to 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.

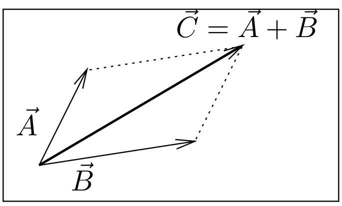

Figure 1. Free vector addition. The diagonal of the parallelogram formed by free vectors $\vec{A}$, $\vec{B}$ is the sum vector $\vec{C} = \vec{A} + \vec{B}$.

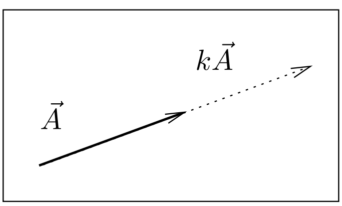

Figure 2. Free vector scalar multiplication. To form $k\vec{A}$, the head of free vector $\vec{A}$ is moved to a new location along the line formed by the head and tail.

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.

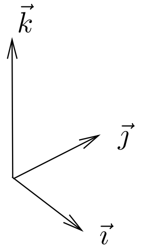

Figure 3. Fundamental triad. The free vectors $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ are 90° apart and of unit length.

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 $|k| > 1$.

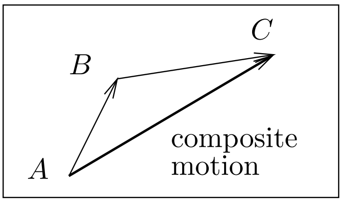

Figure 4. Planar composition of motions. The motion A to C is the composition of two motions or the *sum* of vectors AB and BC.

Comparison of Vector Models

In free vector diagrams it is possible to use free, physics and Gibbs vector models almost interchangeably. In the Gibbs model, the negative of a vector and the zero vector are natural objects, whereas in the other models they can be 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 is realized as a free vector by drawing an arrow 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 $\vec{Y}$.

Subtraction 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.

Head Minus 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. This statement is called 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.

[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 .

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
	$k_1(k_2\vec{X}) = (k_1 k_2)\vec{X}$	distributive III
	$1\vec{X} = \vec{X}$	identity

Definition 5.1 (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 to burst vectors.

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 numerical approximation sequences. Additional applications are coefficients of Fourier series and Taylor series.
>
> A **Matrix space** is a structured system which can organize two-dimensional data sets. Examples are the array of pixels for a digital photograph and robotic mechanical component manipulators represented 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 operations $\boxed{+}$ and $\boxed{\cdot}$ (details postponed). 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.

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 5.1, page 300.

Definition 5.2 (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.

Theorem 5.1 (Subspace Criterion)
Assume abstract 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 5.2 (Subspaces of $\mathcal{R}^n$: The Kernel Theorem)
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 on page 314.

When does Theorem 5.2 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 5.2, 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 5.2.

Theorem 5.3 (Subspaces of $\mathcal{R}^n$: Restriction Equations)
Let V be one of the vector spaces $\mathcal{R}^n$ and let data set S be defined by a system of restriction equations. If the restriction equations are homogeneous linear algebraic equations, then S is a subspace of V.

How to apply Theorem 5.2 and Theorem 5.3. 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* $x_4 = 0$.

By Theorem 5.3, S is a subspace of V, with no further details required.

[2]This key theorem is named the **kernel theorem**, because solutions $\vec{x}$ of $A\vec{x} = \vec{0}$ define the **kernel** of A. It is also named the **Nullspace Theorem**.

To apply Theorem 5.2, the restriction equations have to be re-written as a homogeneous matrix equation $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$:

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

Then Theorem 5.2 applies to conclude that S is a subspace of V.

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 5.4 (Test 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 5.3 (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 5.4 (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

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

Important: The symbols $c_1, \ldots, c_n$ exhaust all possible choices of scalars: expect the **span** to contain infinitely many data packages (called *abstract* vectors) from data set V.

Theorem 5.5 (Span of Vectors is a Subspace)
Let V be an abstract vector space. A subset $S = \mathbf{span}(\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_k)$ is a subspace of V. Proof on page 314.

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.

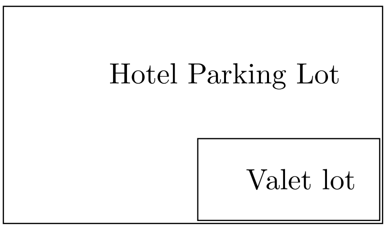

Figure 5. Parking lot analogy. An abstract vector space V and one of its subspaces S can be visualized through the analogy of a parking lot (V) containing a valet lot (S).

Vector Algebra

Definition 5.5 (Norm of a Fixed Vector)
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 $\vec{Y}$:

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

Definition 5.6 (Dot Product or Scalar Product)
The **dot product** $\vec{X} \cdot \vec{Y}$ of two fixed vectors $\vec{X}$ and $\vec{Y}$ is defined by

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

Definition 5.7 (Angle Between Vectors)
Assume $|\vec{X}| > 0$ and $|\vec{Y}| > 0$. Define the angle θ, $0 \le \theta \le \pi$, between vectors $\vec{X}$ and $\vec{Y}$ by:

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

Calculus vector geometry for $n = 3$ derives formula $|\vec{X}||\vec{Y}|\cos\theta = \vec{X} \cdot \vec{Y}$, which produces the above equation by solving for $\cos\theta$, motivation for the definition.

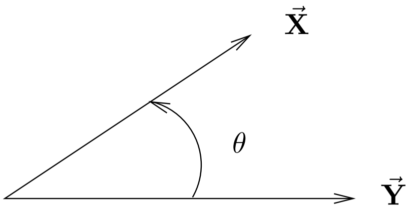

Figure 6. Angle θ between two vectors $\vec{\mathbf{X}}, \vec{\mathbf{Y}}$.

Definition 5.8 (Orthogonal Vectors)
Two n-vectors $\vec{X}$, $\vec{Y}$ are said to be **orthogonal** provided $\vec{X} \cdot \vec{Y} = 0$.

If both vectors are nonzero, then $\cos(\theta) = \dfrac{\vec{X} \cdot \vec{Y}}{|\vec{X}||\vec{Y}|} = 0$, which implies the angle between the vectors is $\theta = 90°$.

Definition 5.9 (Shadow Projection)
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.

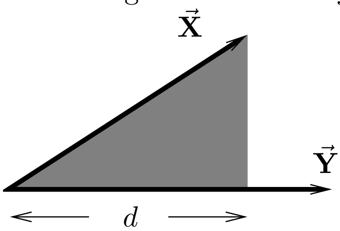

Figure 7. Shadow projection d
Distance d is the length of the shadow formed by vector $\vec{\mathbf{X}}$ onto the direction of vector $\vec{\mathbf{Y}}$.

Definition 5.10 (Vector Projection)
The **vector projection** of $\vec{X}$ onto the line L through the origin in the direction of $\vec{Y}$ is defined by

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

Definition 5.11 (Vector Reflection)
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**:

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

Definition 5.12 (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 might 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 this 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.

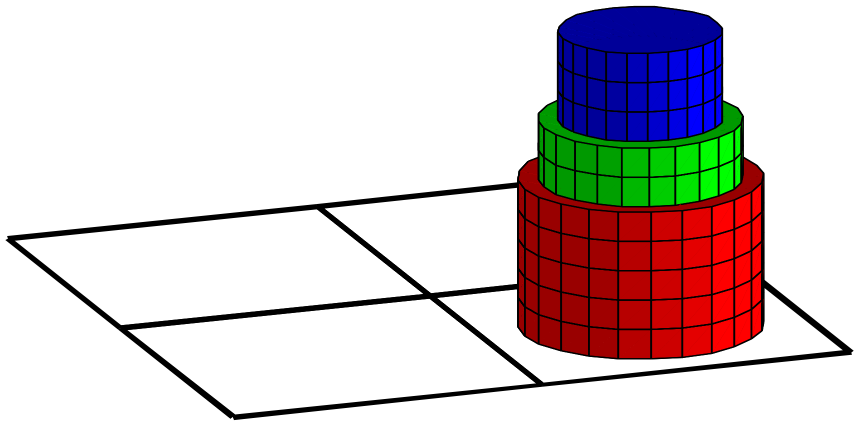

Figure 8. Checkerboard visualization.
Illustrated is a stack of checkers, representing one photodiode site on an image sensor inside a digital camera. There are 5 red, 2 green and 3 blue checkers stacked on one square, representing electron counts.

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 $r = g = b = 255$ is the color **white**.

A matrix of size $m \times n$ is visualized as a checkerboard with mn squares, each square stacked with red, green and blue checkers. Higher resolution image sensors store image data in huge matrices with richer color information, for instance 32-bit and 128-bit color.[3]

Visualization of Matrix Addition and Scalar Multiply

Matrix addition can be visualized through matrices representing color separations.[4] 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 $R + G + B$.

Matrix scalar multiplication 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 ma-

[3] A beginner's digital camera manufactured in the early days of digital photography made low resolution color photos using 24-bit color. The photo is constructed from 240 rows of dots with 320 dots per row. The associated storage matrix A is of size 240×320. The identical small format was used for video clips.

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. Storage format **JPEG** has replaced the early formats on phones.

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

trix 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, Figure 8 page 305, 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 R, G, B in the equation $X = R + 2^8 G + 2^{16} B$.

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 between 0 and 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.

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 page 305.

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.

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, a relation used so often that it deserves a formal statement.

Theorem 5.6 (Matrix Multiply as a Linear Combination of Columns)
Let matrix A have vector columns $\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_n$ and let vector $\vec{X}$ have scalar components $x_1, \ldots, x_n$. Then the definition of matrix multiply implies

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

General matrix product AB. The evaluation of matrix products $A\vec{Y}_1$, $A\vec{Y}_2$, $\ldots$, $A\vec{Y}_k$ is a list of k column vectors which can be packaged into a matrix C. Let B be the matrix which packages the columns $\vec{Y}_1, \ldots, \vec{Y}_k$. *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. More arithmetic errors are made when computing dot products written in the form

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

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

On paper, work can be arranged 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.

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

Visualization of Matrix Multiply

Discussed here is 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

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

Geometric rotation and scaling of planar figures have equations of this form. Adopt below definitions of A, B:

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

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

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

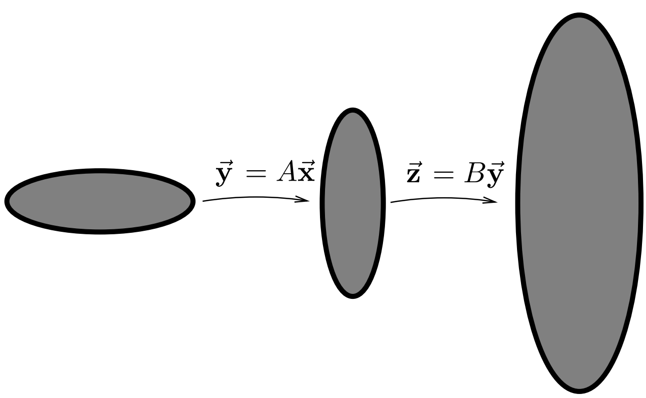

Figure 9. An ellipse is mapped into a rotated and re-scaled ellipse.
The rotation is $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$, which is followed by re-scaling $\vec{\mathbf{z}} = B\vec{\mathbf{y}}$. The composite geometric transformation is $\vec{\mathbf{z}} = BA\vec{\mathbf{x}}$, which maps the ellipse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

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

$$\text{diagonal} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \quad \text{scalar} = \begin{pmatrix} c\,0 \cdots 0 \\ 0\,c \cdots 0 \\ \vdots \\ 0\,0 \cdots c \end{pmatrix}.$$

Matrix Algebra

A matrix can be viewed as a single long array, or fixed vector, therefore the vector space toolkit page 297 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 (page 298).

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. Exercises outline the 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

Definition 5.13 (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, $B = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ is an inverse of $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ because

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

The zero matrix does not have an inverse. To justify, let $A = \mathbf{0}$ and assume square matrix B is a inverse of A. Then relation $\mathbf{0}B = B\mathbf{0} = I$ holds. The zero matrix times any matrix is the zero matrix, which leads to the contradiction $\mathbf{0} = I$.

A given matrix A may not have an inverse.

Definition 5.14 (Inverse Notation A^{-1})
If matrix A has an inverse B, then notation A^{-1} is used for B:

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

Theorem 5.7 (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 $(AB)^{-1} = B^{-1}A^{-1}$.

Proofs on page 315.

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

Theorem 5.8 (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.

Definition 5.15 (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. Proofs are in the exercises.

Proofs and Details

Proof of the Kernel Theorem 5.2: 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:

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

■

Proof of the Span Theorem 5.5: 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. It is impossible, therefore, to **burst** the vectors into components! 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. ■

Proof of the Inverse Theorem 5.7:

(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 $A^TC = CA^T = I$. Start with $AB = BA = I$, take the transpose to get $B^TA^T = A^TB^T = I$. Substitute $C = B^T$, then $CA^T = A^TC = I$, 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 5.2, page 300.

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

Test S Not a Subspace

Test the following restriction equations for $V = \mathcal{R}^3$ and show that the corresponding subset S is not a subspace of V. Use Theorem 5.4 page 301.

33. $x = 1$

34. $x + z = 1$

35. $xz = 2$

36. $xz + y = 1$

37. $xz + y = 0$

38. $xyz = 0$

39. $z \geq 0$

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

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 169 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 200 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. $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$

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

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

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

Matrix Classification

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

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

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

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

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

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

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

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

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

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

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

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 $X = R + 256G + 65536B$, find the red, green and blue color separation matrices R, G, B. Computer assist expected.

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

94. $X = \begin{pmatrix} 514 & 3 \\ 131331 & 66049 \end{pmatrix}$

95. $X = \begin{pmatrix} 513 & 7 \\ 131333 & 66057 \end{pmatrix}$

96. $X = \begin{pmatrix} 257 & 7 \\ 131101 & 66057 \end{pmatrix}$

97. $X = \begin{pmatrix} 257 & 17 \\ 131101 & 265 \end{pmatrix}$

98. $X = \begin{pmatrix} 65537 & 269 \\ 65829 & 261 \end{pmatrix}$

99. $X = \begin{pmatrix} 65538 & 65803 \\ 65833 & 7 \end{pmatrix}$

100. $X = \begin{pmatrix} 259 & 65805 \\ 299 & 5 \end{pmatrix}$

Matrix Properties

Verify the result.

101. 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.

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

103. Let A be a 2×3 matrix and let $\vec{Y}_1$, $\vec{Y}_2$, $\vec{Y}_3$ 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 $AC = \mathbf{0}$.

104. 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.

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

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

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

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

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

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

Matrix Multiply Properties

Verify the result.

111. 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 $L_{ij} = R_{ij}$.

112. 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 $L_{ij} = \sum_{k=1}^n a_{ik}(b_{kj} + c_{kj})$ and $R_{ij} = \sum_{k=1}^n a_{ik}b_{kj} + a_{ik}c_{kj}$. Then $L_{ij} = R_{ij}$.

113. 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 $L_{ij} = a_{ij}$.

114. 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 $L_{ij} = R_{ij}$.

115. 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 $L_{ij} = R_{ij}$.

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

Invertible Matrices

Verify the result.

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

118. The zero matrix is not invertible.

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

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

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

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

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

Symmetric Matrices

Verify the result.

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

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

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

126. 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.

5.2 Matrix Equations

Linear Algebraic Equations

An $m \times n$ system of linear equations

$$(1) \qquad \begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2, \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m, \end{array}$$

can be written as a matrix multiply equation $A\vec{X} = \vec{b}$. Let A be the matrix of coefficients a_{ij}, let $\vec{X}$ be the column vector of variable names $x_1, \ldots, x_n$ and let $\vec{b}$ be the column vector with components $b_1, \ldots, b_n$. Assume equations (1) hold. Then:

$$\begin{aligned} A\vec{X} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \\ &= \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \qquad \text{by equation (1)} \end{aligned}$$

Therefore, equations (1) imply $A\vec{X} = \vec{b}$. Conversely, assume matrix equation $A\vec{X} = \vec{b}$. Reversible steps above give the last vector equality. Vector equality page 304 implies system (1) is satisfied.

A system of linear equations can be represented by its **variable list** $x_1, x_2, \ldots, x_n$ and its **augmented matrix**.

Definition 5.16 (Augmented Matrix)
The augmented matrix of A and $\vec{\mathbf{b}}$ for system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is

$$(2) \qquad \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right) \quad \text{or symbol} \quad \left\langle A \,|\, \vec{b} \right\rangle$$

Vertical Line Notation. The present text uses a vertical line in a matrix display to mean it is an augmented matrix. While symbol $\langle A\,|\,\vec{b}\rangle$ has a vertical bar, the matrix itself has no vertical line as in display (2). Given a matrix C, it certainly has no vertical line. It may be a coefficient matrix in some system $C\vec{\mathbf{x}} = \vec{\mathbf{d}}$, or C could be an augmented matrix for some system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. Computers do not display nor store the vertical line appearing in equation (2). References may not use a vertical line.

Convert Augmented Matrix to Linear Algebraic Equations. Given an augmented $n \times (n+1)$ matrix C and a variable list $x_1, \ldots, x_n$, the conversion back to a linear system of algebraic equations is made by expanding $C\vec{Y} = \mathbf{0}$, where $\vec{Y}$ has components $x_1, \ldots, x_n, -1$. Hand work might contain an exposition like this:

$$
(3) \qquad \begin{array}{c}
\begin{array}{cccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array} \\
\left(\begin{array}{cccc|c}
a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
 & & \vdots & & \\
a_{m1} & a_{m2} & \cdots & a_{mn} & b_n
\end{array}\right)
\end{array}
$$

In (3), a dot product is applied to the first n elements of each row, using the variable list written above the columns. The symbolic answer is set equal to the rightmost column's entry, in order to recover the equations. An example:

$$
\begin{array}{c}
\begin{array}{ccc} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{array} \\
\left(\begin{array}{ccc|c}
1 & 5 & -2 & 7 \\
2 & 0 & -1 & 10 \\
3 & 2 & 4 & 12
\end{array}\right)
\end{array}
\longrightarrow
\left\{\begin{array}{rcr}
x_1 + 5x_2 - 2x_3 & = & 7 \\
2x_1 + 0x_2 - x_3 & = & 10 \\
3x_1 + 2x_2 + 4x_3 & = & 12
\end{array}\right.
$$

Homogeneous System Augmented Matrix. It is usual in homogeneous systems $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ to *omit* the column of zeros and deal directly with A instead of $\langle A\,|\,\vec{0}\rangle$. The convention is justified by arguing that the rightmost column of zeros is unchanged by swap, multiply and combination rules which are defined below. A negative is remembering to insert the column of zeros when using a computation. An example:

$$
\left\{\begin{array}{rcrcrcl}
x_1 & + & 5x_2 & - & 2x_3 & = & 0 \\
2x_1 & + & 0x_2 & - & x_3 & = & 0 \\
3x_1 & + & 2x_2 & + & 4x_3 & = & 0
\end{array}\right.
$$

$$
\text{Use } \begin{pmatrix} 1 & 5 & -2 \\ 2 & 0 & -1 \\ 3 & 2 & 4 \end{pmatrix} \text{ instead of } \left(\begin{array}{ccc|c} 1 & 5 & -2 & 0 \\ 2 & 0 & -1 & 0 \\ 3 & 2 & 4 & 0 \end{array}\right)
$$

Elementary Row Operations

The three operations on equations which produce equivalent systems can be translated directly to row operations on the augmented matrix for the system.

The rules produce **equivalent systems**, that is, the three rules neither create nor destroy solutions.

Swap	Two rows can be interchanged.
Multiply	A row can be multiplied by multiplier $m \neq 0$.
Combination	A multiple of one row can be added to a different row.

Documentation of Row Operations

Throughout the display below, symbol **s** stands for **source**, symbol **t** for **target**, symbol **m** for **multiplier** and symbol **c** for **constant**.

Swap	**swap(s,t)** $\equiv$ swap rows **s** and **t**.
Multiply	**mult(t,m)** $\equiv$ multiply row **t** by **m**$\neq 0$.
Combination	**combo(s,t,c)** $\equiv$ add **c** times row **s** to row **t** $\neq$ **s**.

The standard for documentation is to write the notation next to the target row, which is the row to be changed. For swap operations, the notation is written next to the first row that was swapped, and optionally next to both rows. The notation was developed from early `maple` notation for the corresponding operations `swaprow`, `mulrow` and `addrow`, appearing in the `maple` package `linalg`. For instance, `addrow(A,1,3,-5)` selects matrix A as the target of the combination rule, which is documented in written work as `combo(1,3,-5)`. In written work on paper, symbol A is omitted, because A is the matrix appearing on the previous line of the sequence of steps.

Maple Remarks. Versions of `maple` use packages to perform toolkit operations. A short conversion table appears below.

On paper	**Maple** `with(linalg)`	**Maple** `with(LinearAlgebra)`
`swap(s,t)`	`swaprow(A,s,t)`	`RowOperation(A,[t,s])`
`mult(t,c)`	`mulrow(A,t,c)`	`RowOperation(A,t,c)`
`combo(s,t,c)`	`addrow(A,s,t,c)`	`RowOperation(A,[t,s],c)`

Conversion between packages can be controlled by the following function definitions, which causes the maple code to be the same regardless of which linear algebra package is used.[5]

Maple linalg

```
combo:=(a,s,t,c)->addrow(a,s,t,c);
swap:=(a,s,t)->swaprow(a,s,t);
mult:=(a,t,c)->mulrow(a,t,c);
```

[5]The acronym ASTC is used for the signs of the trigonometric functions in quadrants I through IV. The argument lists for `combo, swap, mult` use the same order, ASTC, memorized in trigonometry as **A***ll* **S***tudents* **T***ake* **C***alculus*.

Maple LinearAlgebra

```
combo:=(a,s,t,c)->RowOperation(a,[t,s],c);
swap:=(a,s,t)->RowOperation(a,[t,s]);
mult:=(a,t,c)->RowOperation(a,t,c);
macro(matrix=Matrix);
```

RREF Test

A linear algebraic equation example of *RREF*:

$$
\begin{array}{rrrrrrrrl}
x_1 & +2x_2 & & +3x_4 & +4x_5 & & +5x_7 & & = 6 \\
 & & x_3 & +7x_4 & +8x_5 & & +9x_7 & & = 10 \\
 & & & & & x_6 & +11x_7 & & = 12 \\
 & & & & & & & x_8 & = 13
\end{array} \tag{4}
$$

The corresponding vector-matrix augmented matrix, no vertical line:

$$
\begin{pmatrix}
1 & 2 & 0 & 3 & 4 & 0 & 5 & 0 & 6 \\
0 & 0 & 1 & 7 & 8 & 0 & 9 & 0 & 10 \\
0 & 0 & 0 & 0 & 0 & 1 & 11 & 0 & 12 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 13 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \tag{5}
$$

Definition 5.17 (Reduced Row-echelon Form or RREF)
The reduced row-echelon form of a matrix, or **rref**, is defined by:

1. Zero rows appear last. Each nonzero row has first element 1, called a **leading one**. The column in which the leading one appears, called a **pivot column**, has all other entries zero.

2. The pivot columns appear as consecutive initial columns of the identity matrix I. Trailing columns of I might be absent.

Matrix (5) is a typical **rref** which satisfies the preceding properties. The initial 4 columns of the 7×7 identity matrix I appear in natural order in matrix (5); the trailing 3 columns of I are absent.

If the **rref** of the augmented matrix has a leading one in the last column, then the corresponding system of equations then has an equation "$0 = 1$" displayed, which signals an **inconsistent** system. Important: the **rref** always exists, even if the corresponding linear algebraic equations are inconsistent.

Elimination Method

The elimination algorithm for equations page **??** has an implementation for matrices. A row is marked **processed** if either (1) the row is all zeros, or else (2) the row contains a leading one and all other entries in that column are zero. Otherwise, the row is called **unprocessed**.

1. Move each unprocessed row of zeros to the last row using **swap** and mark it *processed.*
2. Identify an unprocessed nonzero row having the least number of leading zeros. Apply the **swap** rule to make this row the very first unprocessed row. Apply the **multiply** rule to insure a leading one. Apply the **combination** rule to change to zero all other entries in that column. The number of leading ones (lead variables) has been increased by one and the current column is a column of the identity matrix. Mark the row as *processed*, e.g., box the leading one: $\boxed{1}$.
3. Repeat steps 1–2, until all rows have been processed. Then all leading ones have been defined and the resulting matrix is in reduced row-echelon form.

Computer algebra systems and computer numerical laboratories automate computation of the **reduced row-echelon form** of a matrix A.

Literature calls the algorithm **Gauss-Jordan elimination**. Two examples:

$\mathbf{rref}(\mathbf{0}) = \mathbf{0}$	In **step 2**, all rows of the zero matrix $\mathbf{0}$ are zero. No changes are made to the zero matrix.
$\mathbf{rref}(I) = I$	In **step 2**, each row has a leading one. No changes are made to the identity matrix I.

Visual RREF Test. The habit to mark pivots with a box leads to a visual test for a RREF. An illustration:

$$\left(\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \boxed{1} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Each boxed leading one $\boxed{1}$ appears in a column of the identity matrix. The boxes trail downward, ordered by columns 1, 2, 3 of the identity. There is no 4th pivot, therefore trailing identity column 4 is not used.

Toolkit Sequence

A sequence of swap, multiply and combination steps applied to a system of equations is called a **toolkit sequence**. The viewpoint is that a camera is pointed over the shoulder of an expert who writes the mathematics, and after the completion of each toolkit step, a photo is taken. The ordered sequence of cropped photo frames is a **filmstrip** or a sequence of frames. The **First Frame** displays the original system and the **Last Frame** displays the reduced row echelon system.

The terminology applies to systems $A\vec{x} = \vec{b}$ represented by an augmented matrix $C = \left\langle A \,|\, \vec{b} \right\rangle$. The First Frame is C and the Last Frame is $\mathbf{rref}(C)$.

Documentation of toolkit sequence steps will use this textbook's notation, page 323:

```
swap(s,t), mult(t,m), combo(s,t,c),
```

each written next to the target row `t`. During the sequence, consecutive initial columns of the identity, called **pivot columns**, are created as steps toward the **rref**. The remaining consecutive columns of the identity might not appear. An illustration:

Frame 1: $\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 1 & 4 & -1 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ Original augmented matrix.

Frame 2: $\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ `combo(1,2,-1)` Pivot column 1 completed.

Frame 3: $\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ `swap(2,3)`

Frame 4: $\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ `combo(2,3,-2)`

Frame 5: $\left(\begin{array}{cccc|c} 1 & 0 & -3 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ Pivot column 2 completed by operation `combo(2,1,-2)`. Back-substitution postpones this step.

Frame 6: $\left(\begin{array}{cccc|c} 1 & 0 & -3 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ All leading ones found. `mult(3,-1/2)`

Frame 7: $\left(\begin{array}{cccc|c} 1 & 0 & -3 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ `combo(3,2,-1)` Zero other column 3 entries. Next, finish pivot column 3.

$$\text{Last Frame:} \quad \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

`combo(3,1,3)`
rref found. Column 4 of the identity does not appear! There is no 4th pivot column.

Avoiding fractions. A matrix A with only integer entries can often be put into reduced row-echelon form without introducing fractions. The **multiply** rule introduces fractions, so its use should be limited. It is advised that leading ones be introduced only when convenient, otherwise make the leading coefficient nonzero and positive. Divisions at the end of the computation will produce the **rref**.

Clever use of the **combination** rule can sometimes create a leading one without introducing fractions. Consider the two rows

$$\begin{array}{ccccc} 25 & 0 & 1 & 0 & 5 \\ 7 & 0 & 2 & 0 & 2 \end{array}$$

The second row multiplied by -4 and added to the first row effectively replaces the 25 by -3, whereupon adding the first row twice to the second gives a leading one in the second row. The resulting rows are fraction-free.

$$\begin{array}{ccccc} -3 & 0 & -7 & 0 & -3 \\ 1 & 0 & -12 & 0 & -4 \end{array}$$

Rank and Nullity. What does it mean, if the first column of a **rref** is the zero vector? It means that the corresponding variable x_1 is a **free variable**. In fact, every column that does not contain a leading one corresponds to a free variable in the standard general solution of the system of equations. Symmetrically, each leading one identifies a pivot column and corresponds to a **leading variable**.

The number of leading ones is the **rank** of the matrix, denoted $\mathbf{rank}(A)$. The rank cannot exceed the row dimension nor the column dimension. The column count less the number of leading ones is the **nullity** of the matrix, denoted $\mathbf{nullity}(A)$. It equals the number of free variables.

Regardless of how matrix B arises, augmented or not, we have the relation

$$\text{variable count} = \mathbf{rank}(B) + \mathbf{nullity}(B).$$

If $B = \langle A \,|\, \vec{b} \rangle$ for $A\vec{X} = \vec{b}$, then the variable count n comes from $\vec{X}$ and the column count of B is *one more*, or $n + 1$. Replacing the *variable count* by the *column count* can therefore lead to fundamental errors.

Back-substitution and efficiency. The algorithm implemented in the preceding toolkit sequence is *easy to learn*, because the actual work is organized by creating pivot columns, via swap, combination and multiply. The created pivot columns are initial columns of the identity. You are advised to learn the algorithm in this form, but *please change the algorithm* as you become more efficient at doing the steps. See the examples for illustrations.

Back Substitution. Computer implementations and also hand computation can be made more efficient by changing steps 2 and 3, then adding **step 4**, as outlined below.

1. Move each unprocessed row of zeros to the last row using **swap** and mark it *processed*.

2a. Identify an unprocessed nonzero row having the least number of leading zeros. Apply the **swap** rule to make this row the very first unprocessed row. Apply the **multiply** rule to insure a leading one. Apply the **combination** rule to change to zero all other entries in that column which are **below** the leading one.

3a. Repeat steps 1–2a, until all rows have been processed. The matrix has all leading ones identified, a triangular shape, but it is not generally a RREF.

4. **Back-Substitution**. Identify the last row with a leading one. Apply the **combination** rule to change to zero all other entries in that column which are **above** the leading one. Repeat until all rows have been processed. The resulting matrix is a RREF.

Literature refers to **step 4** as **back-substitution**, a process which is exactly the original elimination algorithm applied to the system created by step 3a with *reversed variable list*.

Inverse Matrix. An efficient method to find the inverse B of a square matrix A, should it happen to exist, is to form the augmented matrix $C = \left\langle A \,|\, I \right\rangle$ and then read off B as the package of the last n columns of $\mathbf{rref}(C)$. This method is based upon the equivalence

$$\mathbf{rref}(\left\langle A \,|\, I \right\rangle) = \left\langle I \,|\, B \right\rangle \quad \text{if and only if} \quad AB = I.$$

The next theorem aids not only in establishing this equivalence but also in the practical matter of testing a candidate solution for the inverse matrix.

Theorem 5.9 (Inverse Test for Matrices)
If A and B are square matrices such that $AB = I$, then also $BA = I$. Therefore, only one of the equalities $AB = I$ or $BA = I$ is required to check an inverse. Proof on page 338.

Theorem 5.10 (Matrix Inverse and the rref)
Let A and B denote square matrices. Then

(a) If $\mathbf{rref}\left(\left\langle A \,|\, I \right\rangle\right) = \left\langle I \,|\, B \right\rangle$, then $AB = BA = I$ and B is the inverse of A.

(b) If $AB = BA = I$, then $\mathbf{rref}\left(\left\langle A \,|\, I \right\rangle\right) = \left\langle I \,|\, B \right\rangle$.

(c) If $\mathbf{rref}\left(\langle A\,|\,I\rangle\right) = \langle C\,|\,B\rangle$, then $C = \mathbf{rref}(A)$. If $C \neq I$, then A is not invertible. If $C = I$, then B is the inverse of A.

(d) Identity $\mathbf{rref}(A) = I$ holds if and only if A has an inverse.

Proof on page 338.

Matrix Inverse: Find A^{-1}

The method will be illustrated for the matrix

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

Define the first frame of the sequence to be $C_1 = \langle A\,|\,I\rangle$, then compute the toolkit sequence to $\mathbf{rref}(C_1)$ as follows.

$$C_1 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \qquad \text{First Frame}$$

$$C_2 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & -1 & 1 \end{array}\right) \qquad \texttt{combo(3,2,-1)}$$

$$C_3 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array}\right) \qquad \texttt{mult(3,1/2)}$$

$$C_4 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array}\right) \qquad \texttt{combo(3,2,1)}$$

$$C_5 = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1/2 & -1/2 \\ 0 & 1 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array}\right) \qquad \texttt{combo(3,1,-1)} \quad \text{Last Frame}$$

The theory implies that the inverse of A is the matrix in the right half of the last frame:

$$A^{-1} = \begin{pmatrix} 1 & 1/2 & -1/2 \\ 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix}$$

Answer Check. Let B equal the matrix of the last display, claimed to be A^{-1}. The **Inverse Test**, Theorem 5.9 page 328, says that only one of $AB = I$ or $BA = I$ needs to be checked. Details:

$$\begin{aligned}
AB &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & -1/2 \\ 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1/2 - 1/2 & -1/2 + 1/2 \\ 0 & 1/2 + 1/2 & 1/2 - 1/2 \\ 0 & 1/2 - 1/2 & 1/2 + 1/2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

Elementary Matrices

Elementary matrices express toolkit operations of swap, combination and multiply as matrix multiply equations.

Typically, toolkit operations produce a finite sequence of k linear algebraic equations, the first is the original system and the last is the reduced row echelon form of the system. We are going to re-write a typical toolkit sequence as matrix multiply equations. Each step is obtained from the previous by left-multiplication by a square matrix E:

$$\begin{array}{rcll}
A\vec{X} &=& \vec{b} & \text{Original system} \\
E_1 A\vec{X} &=& E_1\vec{b} & \text{After one toolkit step} \\
E_2 E_1 A\vec{X} &=& E_2 E_1\vec{b} & \text{After two toolkit steps} \\
E_3 E_2 E_1 A\vec{X} &=& E_3 E_2 E_1\vec{b} & \text{After three toolkit steps}
\end{array} \tag{6}$$

Definition 5.18 (Elementary Matrix)
An elementary matrix E is created from the identity matrix by applying a single toolkit operation, that is, exactly one of the operations combination, multiply or swap.

Elementary Combination Matrix. Create square matrix E by applying the operation `combo(s,t,c)` to the identity matrix. The result equals the identity matrix except for the zero in row t and column s which is replaced by c.

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ Identity matrix.

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{pmatrix}$$ Elementary combination matrix, `combo(2,3,c)`.

Elementary Multiply Matrix.
Create square matrix E by applying `mult(t,m)` to the identity matrix. The result equals the identity matrix except the one in row t is replaced by m.

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{Identity matrix.}$$

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & m \end{pmatrix} \qquad \text{Elementary multiply matrix, } \texttt{mult(3,m)}.$$

Elementary Swap Matrix. Create square matrix E by applying `swap(s,t)` to the identity matrix.

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{Identity matrix.}$$

$$E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \text{Elementary swap matrix, } \texttt{swap(1,3)}.$$

If square matrix E represents a combination, multiply or swap rule, then the definition of matrix multiply applied to matrix EB gives the same matrix as obtained by applying the toolkit rule directly to matrix B. The statement is justified by experiment. See the exercises and Theorem 5.11.

Elementary 3×3 matrices (C=Combination, M=Multiply, S=Swap) can be displayed in computer algebra system `maple` as follows.

On Paper	Maple `with(linalg)`	Maple `with(LinearAlgebra)`
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	`B:=diag(1,1,1);`	`B:=IdentityMatrix(3);`
`combo(2,3,c)`	`C:=addrow(B,2,3,c);`	`C:=RowOperation(B,[3,2],c);`
`mult(3,m)`	`M:=mulrow(B,3,m);`	`M:=RowOperation(B,3,m);`
`swap(1,3)`	`S:=swaprow(B,1,3);`	`S:=RowOperation(B,[3,1]);`

A helpful project is to write out several examples of elementary 5 matrices by hand or machine. Such experiments lead to the following observations and theorems, proofs delayed to page 339.

Constructing an Elementary Matrix E

Combination Change a zero in the identity matrix to symbol c.

Multiply Change a one in the identity matrix to symbol $m \neq 0$.

Swap Interchange two rows of the identity matrix.

Constructing E^{-1} from an Elementary Matrix E

Combination	Change multiplier c in E to $-c$.
Multiply	Change diagonal multiplier $m \neq 0$ in E to $1/m$.
Swap	The inverse of E is E itself.

Theorem 5.11 (Matrix Multiply by an Elementary Matrix)
Let B_1 be a given matrix of row dimension n. Select a toolkit operation combination, multiply or swap, then apply it to matrix B_1 to obtain matrix B_2. Apply the identical toolkit operation to the $n \times n$ identity I to obtain elementary matrix E. Then

$$B_2 = EB_1.$$

Theorem 5.12 (Toolkit Sequence Identity)
If C and D are any two frames in a sequence, then corresponding toolkit operations are represented by square elementary matrices E_1, E_2, ..., E_k and the two frames C, D satisfy the matrix multiply equation

$$D = E_k \cdots E_2 E_1 C.$$

Theorem 5.13 (The rref and Elementary Matrices)
Let A be a given matrix of row dimension n. Then there exist $n \times n$ elementary matrices E_1, E_2, ..., E_k representing certain toolkit operations such that

$$\mathbf{rref}(A) = E_k \cdots E_2 E_1 A.$$

Illustration

Consider the following 6-frame toolkit sequence.

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 0 \\ 3 & 6 & 3 \end{pmatrix}$$ Frame 1, original matrix.

$$A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 3 & 6 & 3 \end{pmatrix}$$ Frame 2, combo(1,2,-2).

$$A_3 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 3 & 6 & 3 \end{pmatrix}$$ Frame 3, mult(2,-1/6).

$$A_4 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -6 \end{pmatrix}$$ Frame 4, combo(1,3,-3).

$$A_5 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$ Frame 5, combo(2,3,-6).

$$A_6 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$ Frame 6, combo(2,1,-3). Found **rref**.

The corresponding 3×3 elementary matrices are

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ Frame 2, combo(1,2,-2) applied to I.

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ Frame 3, mult(2,-1/6) applied to I.

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$$ Frame 4, combo(1,3,-3) applied to I.

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{pmatrix}$$ Frame 5, combo(2,3,-6) applied to I.

$$E_5 = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ Frame 6, combo(2,1,-3) applied to I.

Because each frame of the sequence has the succinct form EB, where E is an elementary matrix and B is the previous frame, the complete toolkit sequence can be written as follows.

$A_2 = E_1A_1$	Frame 2, E_1 equals combo(1,2,-2) on I.
$A_3 = E_2A_2$	Frame 3, E_2 equals mult(2,-1/6) on I.
$A_4 = E_3A_3$	Frame 4, E_3 equals combo(1,3,-3) on I.
$A_5 = E_4A_4$	Frame 5, E_4 equals combo(2,3,-6) on I.
$A_6 = E_5A_5$	Frame 6, E_5 equals combo(2,1,-3) on I.
$A_6 = E_5E_4E_3E_2E_1A_1$	Summary, frames 1-6. This relation is $\mathbf{rref}(A_1) = E_5E_4E_3E_2E_1A_1$, which is the result claimed in Theorem 5.13.

The summary is the equation

$$\mathbf{rref}(A_1) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A_1$$

The inverse relationship $A_1 = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1}\,\mathbf{rref}(A_1)$ is formed by the rules for constructing E^{-1} from elementary matrix E, page 331, the result being

$$A_1 = \begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-6&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\3&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&6&1\end{pmatrix}\begin{pmatrix}1&3&0\\0&1&0\\0&0&1\end{pmatrix}\mathbf{rref}(A_1)$$

Examples and Methods

Example 5.1 (Identify a Reduced Row–Echelon Form)
Identify the matrices in reduced row–echelon form using the RREF Test page 324.

$$A = \begin{pmatrix}0&1&3&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix} \quad B = \begin{pmatrix}1&1&3&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

$$C = \begin{pmatrix}2&1&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix} \quad D = \begin{pmatrix}0&1&3&0\\0&0&0&1\\1&0&0&0\\0&0&0&0\end{pmatrix}$$

Solution:

Matrix A. There are two nonzero rows, each with a leading one. The pivot columns are 2, 4 and they are consecutive columns of the 4×4 identity matrix. Yes, it is a RREF.

Matrix B. Same as A but with pivot columns 1, 4. Yes, it is a RREF. Column 2 is not a pivot column. The example shows that a scan for columns of the identity is not enough.

Matrix C. Immediately not a RREF, because the leading nonzero entry in row 1 is not a one.

Matrix D. Not a RREF. Swapping row 3 twice to bring it to row 1 will make it a RREF. This example has pivots in columns 1, 4 but the pivot columns fail to be columns 1, 2 of the identity (they are columns 3, 2).

Visual RREF Test. More experience is needed to use the visual test for RREF, but the effort is rewarded. Details are very brief. The ability to use the visual test is learned by working examples that use the basic RREF test.

Leading ones are boxed:

$$A = \begin{pmatrix}0&\boxed{1}&3&0\\0&0&0&\boxed{1}\\0&0&0&0\\0&0&0&0\end{pmatrix} \quad B = \begin{pmatrix}\boxed{1}&1&3&0\\0&0&0&\boxed{1}\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

$$C = \begin{pmatrix}2&1&1&0\\0&0&0&\boxed{1}\\0&0&0&0\\0&0&0&0\end{pmatrix} \quad D = \begin{pmatrix}0&\boxed{1}&3&0\\0&0&0&\boxed{1}\\\boxed{1}&0&0&0\\0&0&0&0\end{pmatrix}$$

Matrices A, B pass the visual test. Matrices C, D fail the test. Visually, we look for a boxed one starting on row 1. Boxes occupy consecutive rows, marching down and right, to make a triangular diagram. Columns with boxed ones are expected to be consecutive initial columns of identity matrix I.

Example 5.2 (Reduced Row–Echelon Form)
Find the reduced row–echelon form of the coefficient matrix A using the **elimination method**, page 325. Then solve the system.

$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & - & x_3 & + & x_4 & = & 0, \\ x_1 & + & 3x_2 & - & x_3 & + & 2x_4 & = & 0, \\ & & x_2 & & & + & x_4 & = & 0. \end{array}$$

Solution: The coefficient matrix A and its **rref** are given by (details below)

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 1 & 3 & -1 & 2 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{rref}(A) = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Using variable list x_1, x_2, x_2, x_4, the equivalent reduced echelon system is

$$\begin{array}{rcrcrcrcl} x_1 & & & - & x_3 & - & x_4 & = & 0, \\ & & x_2 & & & + & x_4 & = & 0, \\ & & & & & & 0 & = & 0. \end{array}$$

which has lead variables x_1, x_2 and free variables x_3, x_4.

The *last frame algorithm* applies to write the standard general solution. This algorithm assigns invented symbols t_1, t_2 to the free variables, then back-substitution is applied to the lead variables. The solution to the system is

$$\begin{array}{rcl} x_1 & = & t_1 + t_2, \\ x_2 & = & -t_2, \\ x_3 & = & t_1, \\ x_4 & = & t_2, \qquad -\infty < t_1, t_2 < \infty. \end{array}$$

Details of the Elimination Method.

$$\begin{pmatrix} 1^* & 2 & -1 & 1 \\ 1 & 3 & -1 & 2 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

The coefficient matrix A. Leading one identified and marked as 1^*.

$$\begin{pmatrix} \boxed{1} & 2 & -1 & 1 \\ 0 & 1^* & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Apply the **combination** rule to zero the other entries in column 1. Mark the row processed. Identify the next leading one, marked 1^*.

$$\begin{pmatrix} \boxed{1} & 0 & -1 & -1 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Apply the **combination** rule to zero the other entries in column 2. Mark the row processed. The matrix passes the **Visual RREF Test**.

Example 5.3 (Back-Substitution)
Display a toolkit sequence which uses numerical efficiency ideas of back substitution, page 328, in order to find the RREF of the matrix

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

Solution: The answer for the reduced row-echelon form of matrix A is

$$\mathbf{rref}(A) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Back-substitution details appear below.

Meaning of the computation. Finding a RREF is part of solving the homogeneous system $A\vec{X} = \vec{0}$. The Last Frame Algorithm is used to write the general solution. The algorithm requires a toolkit sequence applied to the augmented matrix $\langle A \,|\, \vec{0} \rangle$, ending in the Last Frame, which is the RREF with an added column of zeros.

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 1 & 3 & -1 & 2 \\ 0 & 1 & 0 & 2 \end{pmatrix}$$ The given matrix A. Identify row 1 for the first pivot.

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{pmatrix}$$ `combo(1,2,-1)` applied to introduce zeros below the leading one in row 1.

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$ `combo(2,3,-1)` applied to introduce zeros below the leading one in row 2. The RREF has not yet been found. The matrix is triangular.

$$\begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$ Begin back-substitution: `combo(2,1,-2)` applied to introduce zeros above the leading one in row 2.

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$ Continue back-substitution: `combo(3,2,-1)` and `combo(3,1,1)` applied to introduce zeros above the leading one in row 3.

$$\begin{pmatrix} \boxed{1} & 0 & -1 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$ RREF Visual Test passed. This matrix is the answer.

Example 5.4 (Answer Check a Matrix Inverse)

Display the answer check details for the given matrix A and its proposed inverse B.

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Solution:

Details. We apply the **Inverse Test**, Theorem 5.9, which requires one matrix multiply:

$$AB = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$ Expect $AB = I$.

$$= \begin{pmatrix} 1 & -3+2+1 & 1-2+1 & 1-1 \\ 0 & 1 & -1+1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1-1 & -1+1 & 1 \end{pmatrix}$$ Multiply.

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Simplify. Then $AB = I$. Because of Theorem 5.9, we don't check $BA = I$.

Example 5.5 (Find the Inverse of a Matrix)
Compute the inverse matrix of

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

Solution: The answer:

$$A^{-1} = \begin{pmatrix} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Details. Form the augmented matrix $C = \left\langle A \,|\, I \right\rangle$ and compute its reduced row-echelon form by toolkit steps.

$$\left(\begin{array}{cccc|cccc} 1 & 2 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Augment I onto A.

$$\left(\begin{array}{cccc|cccc} 1 & 2 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

`swap(3,4)`.

$$\left(\begin{array}{cccc|cccc} 1 & 2 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

`combo(2,3,-1)`. Triangular matrix.

$$\left(\begin{array}{cccc|cccc} 1 & 2 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

Back-substitution: `combo(4,2,-1)`.

$$\left(\begin{array}{cccc|cccc} 1 & 2 & -1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

`combo(4,1,-1)`.

$$\left(\begin{array}{cccc|cccc} 1 & 0 & -1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

`combo(2,1,-2)`.

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -3 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

`combo(3,1,1)`. Identity left, inverse right.

Details and Proofs

Proof of Theorem 5.9:

Assume $AB = I$. Let $C = BA - I$. We intend to show $C = \mathbf{0}$, then $BA = C + I = I$, as claimed.

Compute $AC = ABA - A = AI - A = \mathbf{0}$. It follows that the columns $\vec{y}$ of C are solutions of the homogeneous equation $A\vec{y} = \vec{0}$. To complete the proof, we show that the only solution of $A\vec{y} = \vec{0}$ is $\vec{y} = \vec{0}$, because then C has all zero columns, which means C is the zero matrix.

First, $B\vec{u} = \vec{0}$ implies $\vec{u} = I\vec{u} = AB\vec{u} = A\vec{0} = \vec{0}$, hence B has an inverse, and then $B\vec{x} = \vec{y}$ has a unique solution $\vec{x} = B^{-1}\vec{y}$.

Suppose $A\vec{y} = \vec{0}$. Write $\vec{y} = B\vec{x}$. Then $\vec{x} = I\vec{x} = AB\vec{x} = A\vec{y} = \vec{0}$. This implies $\vec{y} = B\vec{x} = B\vec{0} = \vec{0}$. ■

Proof of Theorem 5.10:

Details for (a). Let $C = \left\langle A \,|\, I \right\rangle$ and assume $\mathbf{rref}(C) = \left\langle I \,|\, B \right\rangle$. Solving the $n \times 2n$ system $C\vec{X} = \vec{\mathbf{0}}$ is equivalent to solving the system $A\vec{Y} + I\vec{Z} = \vec{\mathbf{0}}$ with n-vector unknowns $\vec{Y}$ and $\vec{Z}$. This system has exactly the same solutions as $I\vec{Y} + B\vec{Z} = \vec{\mathbf{0}}$, by the equation $\mathbf{rref}(C) = \left\langle I \,|\, B \right\rangle$. The latter is a reduced echelon system with lead variables equal to the components of $\vec{Y}$ and free variables equal to the components of $\vec{Z}$. Multiplying by A gives $A\vec{Y} + AB\vec{Z} = \vec{\mathbf{0}}$, hence $-\vec{Z} + AB\vec{Z} = \vec{\mathbf{0}}$, or equivalently $AB\vec{Z} = \vec{Z}$ for every vector $\vec{Z}$ (because its components are free variables). Letting $\vec{Z}$ be a column of I shows that $AB = I$. Then $AB = BA = I$ by Theorem 5.9, and B is the inverse of A.

Details for (b). Assume $AB = I$. We prove the identity $\mathbf{rref}(\left\langle A \,|\, I \right\rangle) = \left\langle I \,|\, B \right\rangle$. Let the system $A\vec{Y} + I\vec{Z} = \vec{\mathbf{0}}$ have a solution $\vec{Y}$, $\vec{Z}$. Multiply by B to obtain $BA\vec{Y} + B\vec{Z} = \vec{\mathbf{0}}$. Use $BA = I$ to give $\vec{Y} + B\vec{Z} = \vec{\mathbf{0}}$. The latter system therefore has $\vec{Y}$, $\vec{Z}$ as a solution. Conversely, a solution $\vec{Y}$, $\vec{Z}$ of $\vec{Y} + B\vec{Z} = \vec{\mathbf{0}}$ is a solution of the system $A\vec{Y} + I\vec{Z} = \vec{\mathbf{0}}$, because of multiplication by A. Therefore, $A\vec{Y} + I\vec{Z} = \vec{\mathbf{0}}$ and $\vec{Y} + B\vec{Z} = \vec{\mathbf{0}}$ are *equivalent* systems. The latter is in reduced row-echelon form, and therefore $\mathbf{rref}(\left\langle A \,|\, I \right\rangle) = \left\langle I \,|\, B \right\rangle$.

Details for (c). Toolkit steps that compute $\mathbf{rref}(\left\langle A \,|\, I \right\rangle)$ must also compute $\mathbf{rref}(A)$. This fact is learned first by working examples. Elementary matrix formulas can make the proof more transparent: see the Miscellany exercises. Conclusion: $\mathbf{rref}(\left\langle A \,|\, I \right\rangle) = \left\langle C \,|\, B \right\rangle$ implies $C = \mathbf{rref}(A)$.

Let's prove $C \neq I$ implies A is not invertible. Suppose not, then $C \neq I$ and A is invertible. Then (b) implies $\left\langle C \,|\, B \right\rangle = \mathbf{rref}(\left\langle A \,|\, I \right\rangle) = \left\langle I \,|\, B \right\rangle$. Comparing columns, this equation implies $C = I$, a contradiction.

To prove $C = I$ implies B is the inverse of A, apply (a).

Details for (d). Assume A is invertible. We are to prove $\mathbf{rref}(A) = I$. Part (b) says $F = \left\langle A \,|\, I \right\rangle$ satisfies $\mathbf{rref}(F) = \left\langle I \,|\, B \right\rangle$ where B is the inverse of A. Part (c) says $\mathbf{rref}(F) = \left\langle \mathbf{rref}(A) \,|\, \vec{b} \right\rangle$. Comparing matrix columns gives $\mathbf{rref}(A) = I$.

Converse: assume $\mathbf{rref}(A) = I$, to prove A invertible. Let $F = \left\langle A \,|\, I \right\rangle$, then $\mathbf{rref}(F) = \left\langle C \,|\, B \right\rangle$ for some C, B. Part (c) says $C = \mathbf{rref}(A) = I$. Part (a) says B is the inverse of A. This proves A is invertible and completes (d).

Proof of Theorem 5.11: It is possible to organize the proof into three cases, by considering the three possible toolkit operations. We don't do the tedious details. Instead, we refer to the *Elementary Matrix Multiply* exercises page 340, for suitable experiments that provide the intuition needed to develop formal proof details.

Proof of Theorem 5.12: The idea of the proof begins with writing Frame 1 as $C_1 = E_1 C$, using Theorem 5.11. Repeat to write Frame 2 as $C_2 = E_2 C_1 = E_2 E_1 C$. By induction, Frame k is $C_k = E_k C_{k-1} = E_k \cdots E_2 E_1 C$. But Frame k is matrix D in the sequence. ■

Proof of Theorem 5.13: The reduced row-echelon matrix $D = \mathbf{rref}(A)$ paired with $C = A$ imply by Theorem 5.12 that $\mathbf{rref}(A) = D = E_k \cdots E_2 E_1 C = E_k \cdots E_2 E_1 A$. ■

Exercises 5.2

Identify RREF

Mark the matrices which pass the RREF Test, page 324. Explain the failures.

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

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

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

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

Lead and Free Variables

For each matrix A, assume a homogeneous system $A\vec{X} = \vec{\mathbf{0}}$ with variable list $x_1, \ldots, x_n$. List the lead and free variables. Then report the rank and nullity of matrix A.

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

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

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

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

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

10. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

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

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

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

14. $\begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

15. $\begin{pmatrix} 0 & 1 & 0 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

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

Elementary Matrices

Write the 3×3 elementary matrix E and its inverse E^{-1} for each of the following operations, defined on page 323.

17. `combo(1,3,-1)`

18. `combo(2,3,-5)`

19. `combo(3,2,4)`

20. `combo(2,1,4)`

21. `combo(1,2,-1)`

22. `combo(1,2,-`e^2`)`

23. `mult(1,5)`

24. `mult(1,-3)`

25. `mult(2,5)`

26. `mult(2,-2)`

27. `mult(3,4)`

28. `mult(3,5)`

29. `mult(2,-`π`)`

30. `mult(1,`e^2`)`

31. `swap(1,3)`

32. `swap(1,2)`

33. `swap(2,3)`

34. `swap(2,1)`

35. `swap(3,2)`

36. `swap(3,1)`

Elementary Matrix Multiply

For each given matrix B_1, perform the toolkit operation (`combo`, `swap`, `mult`) to obtain the result B_2. Then compute the elementary matrix E for the identical toolkit operation. Finally, verify the matrix multiply equation $B_2 = EB_1$.

37. $\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}$, `mult(2,1/3)`.

38. $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, `mult(1,3)`.

39. $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, `combo(3,2,-1)`.

40. $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, `combo(2,1,-3)`.

41. $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, `swap(2,3)`.

42. $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, `swap(1,2)`.

Inverse Row Operations

Given the final frame B of a sequence starting with matrix A, and the given operations, find matrix A. Do not use matrix multiply.

43. $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,-3)`, `mult(1,-2)`, `swap(2,3)`.

44. $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,3)`, `mult(1,2)`, `swap(3,2)`.

45. $B = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,3)`, `mult(1,4)`, `swap(1,3)`.

46. $B = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,4)`, `mult(1,3)`, `swap(3,2)`.

Elementary Matrix Products

Given the first frame B_1 of a sequence and elementary matrix operations E_1, E_2, E_3, find matrices $F = E_3E_2E_1$ and $B_4 = FB_1$. Hint: Compute $\left\langle B_4|F \right\rangle$ from toolkit operations on $\left\langle B_1|I \right\rangle$.

47. $B_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,-3)`, `mult(1,-2)`.

48. $B_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,3)`, `swap(3,2)`.

49. $B_1 = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `mult(1,4)`, `swap(1,3)`.

50. $B_1 = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations `combo(1,2,-1)`, `combo(2,3,4)`, `mult(1,3)`.

Miscellany

51. Justify with English sentences why all possible 2×2 matrices in reduced row-echelon form must look like

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & * \\ 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

where $*$ denotes an arbitrary number.

52. Display all possible 3×3 matrices in reduced row-echelon form. Besides the zero matrix and the identity matrix, please report five other forms, most containing symbol $*$ representing an arbitrary number.

53. Determine all possible 4×4 matrices in reduced row-echelon form.

54. Display a 6×6 matrix in reduced row-echelon form with rank 4 and only entries of zero and one.

55. Display a 5×5 matrix in reduced row-echelon form with nullity 2 having entries of zero, one and two, but no other entries.

56. Display the rank and nullity of any $n \times n$ elementary matrix.

57. Let $F = \left\langle C|D \right\rangle$ and let E be a square matrix with row dimension matching F. Display the details for the equality

$$EF = \left\langle EC|ED \right\rangle.$$

58. Let $F = \left\langle C|D \right\rangle$ and let E_1, E_2 be $n \times n$ matrices with n equal to the row dimension of F. Display the details for the equality

$$E_2E_1F = \left\langle E_2E_1C|E_2E_1D \right\rangle.$$

59. Assume matrix A is invertible. Display details explaining why $\mathbf{rref}(\left\langle A|I \right\rangle)$ equals the matrix $\left\langle R|E \right\rangle$, where matrix $R = \mathbf{rref}(A)$ and matrix $E = E_k \cdots E_1$. Symbols E_i are elementary matrices in toolkit steps taking matrix A into reduced row-echelon form. Suggestion: Use the preceding exercises.

60. Assume E_1, E_2 are elementary matrices in toolkit steps taking A into reduced row-echelon form. Prove that $A^{-1} = E_2E_1$. In words, A^{-1} is found by doing the same toolkit steps to the identity matrix.

61. Assume matrix A is invertible and $E_1, \ldots, E_k$ are elementary matrices in toolkit steps taking A into reduced row-echelon form. Prove that $A^{-1} = E_k \cdots E_1$.

62. Assume A, B are 2×2 matrices. Assume A is invertible and $\mathbf{rref}(\left\langle A|B \right\rangle) = \left\langle I|D \right\rangle$. Explain why the first column $\vec{x}$ of D is the unique solution of $A\vec{x} = \vec{b}$, where $\vec{b}$ is the first column of B.

63. Assume A, B are $n \times n$ matrices with A invertible. Explain how to solve the matrix equation $AX = B$ for matrix X using the augmented matrix of A, B.

5.3 Determinants and Cramer's Rule

Unique Solution of a 2×2 System

The 2×2 system

$$\begin{array}{rcrcl} ax & + & by & = & e, \\ cx & + & dy & = & f, \end{array} \tag{1}$$

has a unique solution provided $\Delta = ad - bc$ is nonzero, in which case the solution is given by

$$x = \frac{de - bf}{ad - bc}, \quad y = \frac{af - ce}{ad - bc}. \tag{2}$$

This result, called **Cramer's Rule** for 2×2 systems, is first learned in college algebra as a part of determinant theory.

Determinants of Order 2

College algebra introduces matrix notation and determinant notation:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad |A| \text{ or } \det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Evaluation of a 2×2 determinant is by **Sarrus' Rule**:

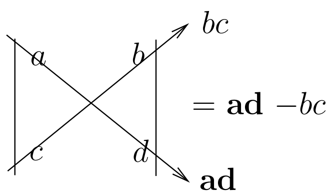

Figure 10. Sarrus' 2×2 rule.
A diagram for $|A| = (ad) - (bc)$.

The boldface product **ad** is the product of the main diagonal entries and the other product bc is from the anti-diagonal.

Cramer's 2×2 rule in determinant notation is

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}. \tag{3}$$

Unique Solution of an $n \times n$ System

Cramer's rule can be generalized to an $n \times n$ system of equations in matrix form $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ or in scalar form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ \vdots & & \vdots & & \cdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n. \end{array} \tag{4}$$

Determinants will be defined shortly; intuition from the 2×2 case and Sarrus' rule should suffice for the moment.

System (4) has a unique solution provided the **determinant of coefficients** $\Delta = \det(A)$ is nonzero, in which case the solution is given by

$$x_1 = \frac{\Delta_1}{\Delta},\ x_2 = \frac{\Delta_2}{\Delta},\ \ldots,\ x_n = \frac{\Delta_n}{\Delta}. \tag{5}$$

The determinant Δ_j equals $\det(B_j)$ where matrix B_j is matrix A modified to have column j equal to $\vec{\mathbf{b}} = (b_1, \ldots, b_n)$. Vector $\vec{\mathbf{b}}$ is the right side of system (4). The result is called **Cramer's Rule** for $n \times n$ systems.

Determinant Notation for Cramer's Rule

The **determinant of coefficients** for system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is denoted by

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{6}$$

The other n determinants in Cramer's rule (5) are given by

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \ldots, \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & b_1 \\ a_{21} & a_{22} & \cdots & b_2 \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n \end{vmatrix}. \tag{7}$$

Determinant Notation Conflicts. The literature is filled with various notations for matrices, vectors and determinants. The expected notation **uses vertical bars only** for determinants and absolute values, e.g., $|A|$ makes sense for a matrix A or a constant A. For clarity, the notation $\det(A)$ may be preferred.

Value of a Determinant. Notation $|A|$ for $\det(A)$ implies that *a determinant is a number*, computed by $|A| = a_{11}a_{22} - a_{12}a_{21}$ when $n = 2$. For $n \geq 3$, $|A|$ is computed by similar but increasingly complicated formulas; see *Sarrus' Rule* page 345 and **Four Determinant Properties** *infra*.

It is false that $|A| = A$ for a 1×1 matrix, because $|A|$ is a number and A is a matrix. The symbol $|c|$ for a *constant* c (not a matrix) is evaluated by algebra rules: $|c| = c$ for $c \geq 0$ and otherwise $|c| = -c$. Overloading of symbols causes equations like $|A| = -1$ for 1×1 matrix $A = (-1)$, whereas $|-1| = 1$ for constant -1.

Sarrus' Rule for 3×3 Matrices

College algebra supplies the following formula for the determinant of a 3×3 matrix A:

$$\begin{aligned} \det(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ &\quad -a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}. \end{aligned} \tag{8}$$

The number $\det(A)$ can be computed by an algorithm similar to the one for 2×2 matrices, as in Figure 11. Important: no further generalizations are possible. *There is no Sarrus' rule for* 4×4 *or larger matrices*!

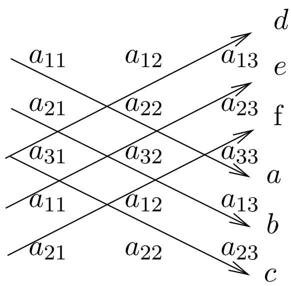

Figure 11. Sarrus' 3×3 rule.
The down arrow sum $a + b + c$ and the up arrow sum $d + e + f$ are subtracted:

$$\det(A) = (a + b + c) - (d + e + f).$$

College Algebra Definition of Determinant

The impractical definition is the formula

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n}. \tag{9}$$

In formula (9), a_{ij} denotes the element in row i and column j of the matrix A. The symbol $\sigma = (\sigma_1, \ldots, \sigma_n)$ stands for a rearrangement of the subscripts $1, 2, \ldots, n$ and S_n is the set of all possible rearrangements. The nonnegative integer parity(σ) is determined by counting the minimum number of pairwise interchanges required to assemble the list of integers $\sigma_1, \ldots, \sigma_n$ into natural order $1, \ldots, n$.

Formula (9) reproduces the definition for 3×3 matrices given in equation (8). We will have no computational use for (9). For computing the value of a determinant, see **four properties** and **cofactor expansion**, *infra*.

Four Determinant Properties

The definition of determinant (9) implies the following four properties:

Triangular	The value of $\det(A)$ for either an upper triangular or a lower triangular matrix A is the product of the diagonal elements: $\det(A) = a_{11}a_{22}\cdots a_{nn}$.
Combination	The value of $\det(A)$ is unchanged by adding a multiple of a row to a different row.
Multiply	If one row of A is multiplied by constant $c \neq 0$ to create matrix B, then $\det(B) = c\det(A)$.
Swap	If B results from A by swapping two rows, then $\det(A) = (-1)\det(B)$.

It is known that these four rules suffice to compute the value of any $n \times n$ determinant. The proof of the four properties is delayed until page 360.

Elementary Matrices and the Four Rules

The rules can be stated in terms of elementary matrices as follows.

Triangular	The value of $\det(A)$ for either an upper triangular or a lower triangular matrix A is the product of the diagonal elements: $\det(A) = a_{11}a_{22}\cdots a_{nn}$. This is a one-arrow Sarrus' rule valid for dimension n.
Combination	If E is an elementary matrix for a combination rule, then $\det(EA) = \det(A)$.
Multiply	If E is an elementary matrix for a multiply rule with multiplier $m \neq 0$, then $\det(EA) = m\det(A)$.
Swap	If E is an elementary matrix for a swap rule, then $\det(EA) = (-1)\det(A)$.

Since $\det(E) = 1$ for a combination rule, $\det(E) = -1$ for a swap rule and $\det(E) = c$ for a multiply rule with multiplier $c \neq 0$, it follows that for any elementary matrix E there is the determinant multiplication rule $\det(EA) = \det(E)\det(A)$.

Theorem 5.14 (Four Rules Compressed)
The Four rules to compute the value of any determinant can be written as two rules.

Triangular Rule	The value of $\|A\|$ for a triangular matrix A is the product of the diagonal elements
Determinant Product Rule	Let E be an elementary matrix, then $\det(EA) = \det(E)\det(A)$.

Additional Determinant Rules

The following rules make for efficient evaluation of certain special determinants. The results are stated for rows, but they also hold for columns, because of Theorem 5.15.

Zero row	If one row of A is zero, then $\det(A) = 0$.
Duplicate rows	If two rows of A are identical, then $\det(A) = 0$.
Dependent rows	If a row of A is a linear combination of the other rows, then $\det(A) = 0$.
RREF $\neq I$	If $\mathbf{rref}(A) \neq I$, then $\det(A) = 0$.
Common factor	The relation $\det(A) = c\det(B)$ holds, provided A and B differ only in one row, say row j, for which $\mathbf{row}(A,j) = c\,\mathbf{row}(B,j)$.
Row linearity	The relation $\det(A) = \det(B) + \det(C)$ holds, provided A, B and C differ only in one row, say row j, for which $\mathbf{row}(A,j) = \mathbf{row}(B,j) + \mathbf{row}(C,j)$.

The proofs of these properties are delayed until page 360.

Determinant of a Transpose

A consequence of (9) is the relation $|A| = \left|A^T\right|$ where A^T means the transpose of A, obtained by swapping rows and columns.

Theorem 5.15 (Determinant of the Transpose)
The relation

$$\det\left(A^T\right) = \det(A) \quad \text{or} \quad \left|A^T\right| = |A|$$

implies that all determinant theory results for rows also apply to columns.

Cofactor Expansion

The special subject of cofactor expansions is used to justify Cramer's rule and to provide an alternative method for computation of determinants. There is no claim that cofactor expansion is efficient, only that it is possible, and different than Sarrus' rule or the use of the four properties.

Background from College Algebra

The cofactor expansion theory is most easily understood from the college algebra topic in dimension 3. Cofactor row expansion computes $|A|$ by one of three possible formulas, recorded below. The **pattern**:

$$|A| = \Sigma(\text{row element} \times \text{checkerboard sign} \times \text{cross-out determinant}).$$

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$|A| = a_{11}(+1)\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(+1)\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$|A| = a_{21}(-1)\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22}(+1)\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + a_{23}(-1)\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

$$|A| = a_{31}(+1)\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{32}(-1)\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}(+1)\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

The formulas expand a 3×3 determinant in terms of 2×2 determinants, along a row of A. The attached signs ± 1 are called the **checkerboard signs**, to be defined shortly. The 2×2 **cross-out determinants** are officially called **minors** of the 3×3 determinant $|A|$. The checkerboard sign multiplied against a minor is called a **cofactor**.

These formulas are generally used when a row has one or two zeros, making it unnecessary to evaluate one or two of the 2×2 determinants in the expansion. To illustrate, row 1 expansion gives

$$\begin{vmatrix} 3 & 0 & 0 \\ 2 & 1 & 7 \\ 5 & 4 & 8 \end{vmatrix} = 3(+1)\begin{vmatrix} 1 & 7 \\ 4 & 8 \end{vmatrix} + 0(-1)\begin{vmatrix} 2 & 7 \\ 5 & 8 \end{vmatrix} + 0(+1)\begin{vmatrix} 2 & 1 \\ 5 & 4 \end{vmatrix} = -60.$$

A clever time–saving choice is always a row which has the most zeros, although success does not depend upon cleverness. What has been said for rows also applies to columns, due to the transpose formula $|A| = |A^T|$.

Minors and Cofactors

The $(n-1) \times (n-1)$ determinant obtained from $\det(A)$ by crossing-out row i and column j is called the (i, j)–minor of A and denoted $\textbf{minor}(A, i, j)$ (M_{ij} is common in literature). The (i, j)–cofactor of $|A|$ is $\textbf{cof}(A, i, j) = (-1)^{i+j}\,\textbf{minor}(A, i, j)$. Multiplicative factor $(-1)^{i+j}$ is called the **checkerboard sign**, because its value can be determined by counting *plus*, *minus*, *plus*, etc., from location $(1, 1)$ to location (i, j) in any checkerboard fashion.

To illustrate how to create the smaller cross-out determinant, denoted by the symbol $\textbf{minor}(A, i, j)$, consider this example:

$$\textbf{minor}\left(\begin{pmatrix} 3 & 0 & 0 \\ 2 & 1 & 7 \\ 5 & 4 & 8 \end{pmatrix}, 2, 3\right) = \begin{vmatrix} 3 & 0 & \cancel{0} \\ \cancel{2} & \cancel{1} & \cancel{7} \\ 5 & 4 & \cancel{8} \end{vmatrix} = \begin{vmatrix} 3 & 0 \\ 5 & 4 \end{vmatrix}$$

cross-out row=2 and column=3, red strikeouts removed

Expansion of Determinants by Cofactors

The formulas are

$$\text{(10)} \qquad \det(A) = \sum_{j=1}^{n} a_{kj}\,\mathbf{cof}(A,k,j), \quad \det(A) = \sum_{i=1}^{n} a_{i\ell}\,\mathbf{cof}(A,i,\ell),$$

where $1 \le k \le n$, $1 \le \ell \le n$. The first expansion in (10) is called a **cofactor row expansion** and the second is called a **cofactor column expansion**. The value $\mathbf{cof}(A,i,j)$ is the cofactor of element a_{ij} in $\det(A)$, that is, the checkerboard sign times the minor of a_{ij}. The proof of expansion (10) is delayed until page 361.

The Adjugate Matrix

The **adjugate** of an $n \times n$ matrix A, denoted $\mathbf{adj}(A)$, is the transpose of the matrix of cofactors:

$$\mathbf{adj}(A) = \begin{pmatrix} \mathbf{cof}(A,1,1) & \mathbf{cof}(A,1,2) & \cdots & \mathbf{cof}(A,1,n) \\ \mathbf{cof}(A,2,1) & \mathbf{cof}(A,2,2) & \cdots & \mathbf{cof}(A,2,n) \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{cof}(A,n,1) & \mathbf{cof}(A,n,2) & \cdots & \mathbf{cof}(A,n,n) \end{pmatrix}^T.$$

A cofactor $\mathbf{cof}(A,i,j)$ is the checkerboard sign $(-1)^{i+j}$ times the corresponding cross-out determinant $\mathbf{minor}(A,i,j)$. In the 2×2 case,

$$\mathbf{adj}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

In words: *swap the diagonal elements and change the sign of the off–diagonal elements.*

The Inverse Matrix

The adjugate appears in the inverse matrix formula for a 2×2 matrix:

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

This formula is verified by direct matrix multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The $n \times n$ matrix identity $A \cdot \mathbf{adj}(A) = \det(A)\,I$ implies

$$A^{-1} = \frac{1}{\det(A)}\begin{pmatrix} \mathbf{cof}(A,1,1) & \mathbf{cof}(A,1,2) & \cdots & \mathbf{cof}(A,1,n) \\ \mathbf{cof}(A,2,1) & \mathbf{cof}(A,2,2) & \cdots & \mathbf{cof}(A,2,n) \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{cof}(A,n,1) & \mathbf{cof}(A,n,2) & \cdots & \mathbf{cof}(A,n,n) \end{pmatrix}^T$$

Theorem 5.16 (Fundamental Adjugate Identity)

$$A \cdot \textbf{adj}(A) = \textbf{adj}(A) \cdot A = \det(A)\, I$$

The proof is delayed until page 362.

Determinants of Elementary Matrices

An elementary matrix E is the result of applying a combination, multiply or swap rule to the identity matrix. This definition implies that an elementary matrix is the identity matrix with a minor change applied, to wit:

Combination	Change an off-diagonal zero of I to c.
Multiply	Change a diagonal one of I to multiplier $m \neq 0$.
Swap	Swap two rows of I.

Theorem 5.17 (Determinants and Elementary Matrices)
Let E be an $n \times n$ elementary matrix. Then

Combination	$\det(E) = 1$
Multiply	$\det(E) = m$ for multiplier m.
Swap	$\det(E) = -1$
Product	$\det(EX) = \det(E)\det(X)$ for all $n \times n$ matrices X.

Theorem 5.18 (Determinants and Invertible Matrices)
Let A be a given invertible matrix. Then

$$\det(A) = \frac{(-1)^s}{m_1 m_2 \cdots m_r}$$

where s is the number of swap rules applied and m_1, m_2, ..., m_r are the nonzero multipliers used in multiply rules when A is reduced to $\mathbf{rref}(A)$.

Determinant Product Rule

The determinant rules of combination, multiply and swap imply that $\det(EX) = \det(E)\det(X)$ for elementary matrices E and square matrices X. We show that a more general relationship holds.

Theorem 5.19 (Product Rule for Determinants)
Let A and B be given $n \times n$ matrices. Then

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

Proof:

Used in the proof is the equivalence of invertibility of a square matrix C with $\det(C) \neq 0$ and $\mathbf{rref}(C) = I$.

Assume one of A or B has zero determinant. Then $\det(A)\det(B) = 0$. If $\det(B) = 0$, then $B\vec{\mathbf{x}} = \vec{\mathbf{0}}$ has infinitely many solutions, in particular a nonzero solution $\vec{\mathbf{x}}$. Multiply $B\vec{\mathbf{x}} = \vec{\mathbf{0}}$ by A, then $AB\vec{\mathbf{x}} = \vec{\mathbf{0}}$ which implies AB is not invertible. Then the identity $\det(AB) = \det(A)\det(B)$ holds, because both sides are zero. If $\det(B) \neq 0$ but $\det(A) = 0$, then there is a nonzero $\vec{\mathbf{y}}$ with $A\vec{\mathbf{y}} = \vec{\mathbf{0}}$. Because B has an inverse, then $\vec{\mathbf{x}} = B^{-1}\vec{\mathbf{y}}$ is defined and nonzero. Then $AB\vec{\mathbf{x}} = A\vec{\mathbf{y}} = \vec{\mathbf{0}}$, with $\vec{\mathbf{x}} \neq \vec{\mathbf{0}}$, which implies $\mathbf{rref}(AB) \neq I$ and $|AB| = 0$. Therefore, both sides of $\det(AB) = \det(A)\det(B)$ are zero and the identity holds.

Assume A, B are invertible. Then $C = AB$ is invertible. In particular, $\mathbf{rref}(A^{-1}) = \mathbf{rref}(B^{-1}) = I$. Write $I = \mathbf{rref}(A^{-1}) = E_1E_2\cdots E_kA^{-1}$ and $I = \mathbf{rref}(B^{-1}) = F_1F_2\cdots F_mB^{-1}$ for elementary matrices E_i, F_j. Then

$$AB = E_1E_2\cdots E_kF_1F_2\cdots F_m. \tag{11}$$

The theorem follows from repeated application of identity $\det(EX) = \det(E)\det(X)$ to relation (11), because

$$\det(A) = \det(E_1)\cdots\det(E_k), \quad \det(B) = \det(F_1)\cdots\det(F_m).$$

Cramer's Rule and the Determinant Product Formula

The equation $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ in the 3×3 case is used routinely to produce the three matrix multiply equations

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix},$$
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 1 & x_1 & 0 \\ 0 & x_2 & 0 \\ 0 & x_3 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix},$$
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 1 & 0 & x_1 \\ 0 & 1 & x_2 \\ 0 & 0 & x_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}.$$

The determinant of the second matrix on the left in the first equation evaluates to x_1. Similarly, in the other equations, the determinant of the second matrix evaluates to x_2, x_3, respectively. Therefore, **the determinant product theorem** applied to these three equations, followed by dividing by $\det(A)$, derives Cramer's Rule:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{|A|}, x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{|A|}, x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{|A|}.$$

Examples

Example 5.6 (Four Properties)
Apply the four properties of a determinant to justify the formula

$$\begin{vmatrix} 12 & 6 & 0 \\ 11 & 5 & 1 \\ 10 & 2 & 2 \end{vmatrix} = 24.$$

Solution: The details:

$\begin{vmatrix} 12 & 6 & 0 \\ 11 & 5 & 1 \\ 10 & 2 & 2 \end{vmatrix}$	Given.
$= \begin{vmatrix} 12 & 6 & 0 \\ -1 & -1 & 1 \\ -2 & -4 & 2 \end{vmatrix}$	Combination rule twice: `combo(1,2,-1)`, `combo(1,3,-1)`.
$= 6\begin{vmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ -2 & -4 & 2 \end{vmatrix}$	Multiply rule: factor out 6 from row 1.
$= 6\begin{vmatrix} 0 & -1 & 2 \\ -1 & -1 & 1 \\ 0 & -3 & 2 \end{vmatrix}$	Combination rule twice: `combo(1,3,1)`, `combo(2,1,2)`.
$= 6(-1)\begin{vmatrix} -1 & -1 & 1 \\ 0 & -1 & 2 \\ 0 & -3 & 2 \end{vmatrix}$	Swap rule: `swap(1,2)`.
$= 6(-1)^2\begin{vmatrix} 1 & 1 & -1 \\ 0 & -1 & 2 \\ 0 & -3 & 2 \end{vmatrix}$	Multiply rule: factor out (-1) from row 1.
$= 6\begin{vmatrix} 1 & 1 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & -4 \end{vmatrix}$	Combination rule: `combo(2,3,-3)`.
$= 6(1)(-1)(-4)$	Triangular rule.
$= 24$	Formula verified.

Example 5.7 (Determinant of an Elementary Matrix)
Compute the determinants of the following elementary matrices.

$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

Solution: The matrices correspond to toolkit operations:

```
swap(1,2), combo(3,1,c), mult(3,10).
```

Therefore, the determinant values are $-1, 1, 10$, by Theorem 5.17.

Example 5.8 (Additional Determinant Rules)
Compute the determinants by applying the additional determinant rules, page 347.

$$\begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 & 10 \\ 0 & 1 & 0 \\ 1 & 1 & 10 \end{vmatrix}, \quad \begin{vmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 6 & 4 & 2 \end{vmatrix}.$$

Solution: Answer: $0, 0, 0$. A row of zeros implies determinant zero, for the 2×2. Row 3 equal to the sum of rows 1 and 2 implies determinant zero, for the 3×3. Row 4 equals twice row 1 implies determinant zero, for the 4×4.

Example 5.9 (Adjugate and Inverse)
Compute the adjugate matrix $\mathbf{adj}(A)$ and the inverse matrix $B = \frac{\mathbf{adj}(A)}{|A|}$, given

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

Solution: The adjugate matrix is the transpose of the matrix of cofactors. **A common mistake** is to compute instead the transpose matrix, a tragic over-simplification, considering the effort required: the matrix of cofactors requires the evaluation of 16 determinants of size 3×3.

For example, the effort for one 3×3 cofactor (=(checkerboard sign)(3×3 minor determinant)) is about 30 seconds:

$$\mathbf{cof}(A, 1, 2) = (-1)^{1+2}\,\mathbf{minor}(A, 1, 2) = -\begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 4 & 2 \end{vmatrix} = 0.$$

Reported here is the answer for the adjugate matrix, an effort on paper of about 8 minutes.

$$\begin{aligned} \mathbf{adj}(A) &= \text{transpose of} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & -1 & 0 & 2 \\ 0 & 0 & -1 & 2 \\ -1 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 2 & 2 & 1 \end{pmatrix} \end{aligned}$$

The determinant of A is already known, because of the formula $A\,\mathbf{adj}(A) = |A|I$. For instance, the $(1,1)$-position in matrix $|A|I$ has value $|A|$, which from the left side of $A\,\mathbf{adj}(A) = |A|I$ equals the dot product of row 1 of A and column 1 of $\mathbf{adj}(A)$. Then $|A| = -1$.

The inverse matrix B is the adjugate matrix $\textbf{adj}(A)$ divided by the determinant $|A| = -1$:

$$B = \frac{\textbf{adj}(A)}{|A|} = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & -2 & -2 & -1 \end{pmatrix}.$$

Answer Check. The inverse answer can be checked by matrix multiply, using the equation $A\,\textbf{adj}(A) = |A|I$, or the equation $AB = I$. For example,

$$AB = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & -2 & -2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Alternate solution without determinants. Define $C = \big\langle A|I \big\rangle$ and compute with toolkit steps $\mathbf{rref}(C) = \big\langle I|B \big\rangle$. Toolkit steps can evaluate $|A|$, and B is the inverse of A. Report $\textbf{adj}(A) = |A|B$.

Example 5.10 (Cofactor Expansion Method)

Justify by cofactor expansion the identity

$$\begin{vmatrix} 10 & 5 & 0 & 0 \\ 11 & 5 & a & 0 \\ 10 & 2 & b & 0 \\ 15 & 8 & 4 & 2 \end{vmatrix} = 10(6a - b).$$

Solution: The plan is to choose the row or column with most zeros, then expand by cofactors. The greatest advantage is column 4, effectively reducing the determinant to 3×3. The resulting 3×3 is treated by a hybrid method in the next example. Here, we will expand it by cofactors, again choosing a column or row with most zeros. The details:

$\begin{vmatrix} 10 & 5 & 0 & 0 \\ 11 & 5 & a & 0 \\ 10 & 2 & b & 0 \\ 15 & 8 & 4 & 2 \end{vmatrix}$	Given 4×4 determinant with symbols a, b.
$= 2(-1)^{4+4} \begin{vmatrix} 10 & 5 & 0 \\ 11 & 5 & a \\ 10 & 2 & b \end{vmatrix}$	Cofactor expansion on column 4. Three zero terms are not written. See $\boxed{\mathbf{1}}$ below.
$= \begin{matrix} 2\left(a(-1)^{2+3}\begin{vmatrix} 10 & 5 \\ 10 & 2 \end{vmatrix}\right) + \\ 2\left(b(-1)^{3+3}\begin{vmatrix} 10 & 5 \\ 11 & 5 \end{vmatrix}\right) \end{matrix}$	Expand by cofactors on column 3. The zero term is not written. See $\boxed{\mathbf{2}}$ below.
$= \begin{matrix} 2\left(a(-1)^{2+3}(-30)\right) + \\ 2\left(b(-1)^{3+3}(-5)\right) \end{matrix}$	Expand 2×2 determinants by Sarrus' rule.
$= 60a - 10b$	Final answer with symbols a, b.

$\boxed{1}$ The factor 2 is from element $4,4$. The factor $(-1)^{4+4}$ is the checkerboard sign of element $4,4$. The 3×3 determinant is the **minor** obtained by cross-out of row 4, column 4.

$\boxed{2}$ For example, $2\left(a(-1)^{2+3}\begin{vmatrix} 10 & 5 \\ 10 & 2 \end{vmatrix}\right)$ is decoded as follows. Factor 2 is from the 4×4 cofactor expansion. Inside the parentheses, factor a is from the 3×3 determinant element in row 2, column 3. Factor $(-1)^{2+3}$ is the checkerboard sign of that row and column. Factor $\begin{vmatrix} 10 & 5 \\ 10 & 2 \end{vmatrix}$ is the minor determinant obtained by cross–out of row 2 and column 3.

Example 5.11 (Hybrid Method)
Justify by cofactor expansion and the four properties the identity

$$\begin{vmatrix} 10 & 5 & 0 \\ 11 & 5 & a \\ 10 & 2 & b \end{vmatrix} = 5(6a - b).$$

Solution: The details:

$$\begin{array}{ll}
\begin{vmatrix} 10 & 5 & 0 \\ 11 & 5 & a \\ 10 & 2 & b \end{vmatrix} & \text{Given.} \\
= \begin{vmatrix} 10 & 5 & 0 \\ 1 & 0 & a \\ 0 & -3 & b \end{vmatrix} & \text{Combination: subtract row 1 from the other rows.} \\
= \begin{vmatrix} 0 & 5 & -10a \\ 1 & 0 & a \\ 0 & -3 & b \end{vmatrix} & \text{Combination: add } -10 \text{ times row 2 to row 1.} \\
= (1)(-1)\begin{vmatrix} 5 & -10a \\ -3 & b \end{vmatrix} & \text{Cofactor expansion on column 1.} \\
= (1)(-1)(5b - 30a) & \text{Sarrus' rule for } n = 2. \\
= 5(6a - b). & \text{Formula verified.}
\end{array}$$

Example 5.12 (Determinant Product Rule)
Let A, B be 4×4 matrices. Let E_1, E_2, E_3 be elementary matrices of the same size corresponding to toolkit operations

```
combo(1,3,-2), mult(3,-5), swap(2,4).
```

Find $|A|$, given $|B| = 3$ and the equation

$$A^3B^2 = E_3E_2E_1B.$$

Solution: The idea is to use the determinant product rule $|CD| = |C||D|$ repeatedly, on the given equation, to obtain the scalar equation

$$|A|^3|B|^2 = |E_3||E_2||E_1||B|.$$

Determinant values for elementary matrices are completely determined by the given toolkit operation: $|E_1| = 1, |E_2| = -5, |E_3| = -1$. Then the scalar equation above reduces, because of $|B| = 3$, to the algebraic equation

$$|A|^3(3)^2 = (-1)(-5)(1)(3).$$

Solving for symbol $|A|$ gives the answer $|A| = \sqrt[3]{15/9} = 1.1856$.

Example 5.13 (Cramer's Rule)
Solve by Cramer's rule the system of equations

$$\begin{array}{rcrcrcrcr} 2x_1 & + & 3x_2 & + & x_3 & - & x_4 & = & 1, \\ x_1 & + & x_2 & - & & & x_4 & = & -1, \\ & & 3x_2 & + & x_3 & + & x_4 & = & 3, \\ x_1 & + & & & x_3 & - & x_4 & = & 0, \end{array}$$

verifying $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$.

Solution: Form the four determinants $\Delta_1, \ldots, \Delta_4$ from the determinant of coefficients Δ as follows:

$$\Delta = \begin{vmatrix} 2 & 3 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 3 & 1 & 1 \\ 1 & 0 & 1 & -1 \end{vmatrix},$$

$$\Delta_1 = \begin{vmatrix} \mathbf{1} & 3 & 1 & -1 \\ \mathbf{-1} & 1 & 0 & -1 \\ \mathbf{3} & 3 & 1 & 1 \\ \mathbf{0} & 0 & 1 & -1 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} 2 & \mathbf{1} & 1 & -1 \\ 1 & \mathbf{-1} & 0 & -1 \\ 0 & \mathbf{3} & 1 & 1 \\ 1 & \mathbf{0} & 1 & -1 \end{vmatrix},$$

$$\Delta_3 = \begin{vmatrix} 2 & 3 & \mathbf{1} & -1 \\ 1 & 1 & \mathbf{-1} & -1 \\ 0 & 3 & \mathbf{3} & 1 \\ 1 & 0 & \mathbf{0} & -1 \end{vmatrix}, \quad \Delta_4 = \begin{vmatrix} 2 & 3 & 1 & \mathbf{1} \\ 1 & 1 & 0 & \mathbf{-1} \\ 0 & 3 & 1 & \mathbf{3} \\ 1 & 0 & 1 & \mathbf{0} \end{vmatrix}.$$

Five repetitions of the methods used in the previous examples give the answers $\Delta = -2$, $\Delta_1 = -2$, $\Delta_2 = 0$, $\Delta_3 = -2$, $\Delta_4 = -4$, therefore Cramer's rule implies the solution $x_i = \Delta_i/\Delta$, $1 \le i \le 4$. Then $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$.

Answer Check. The details of the computation above can be checked in computer algebra system `maple` as follows.

```
A:=Matrix([[2,  3,  1, -1], [1,  1,  0, -1],
           [0,  3,  1,  1], [1,  0,  1, -1]]);
B1:=Matrix([[ 1,  3,  1, -1], [-1,  1,  0, -1],
            [ 3,  3,  1,  1], [ 0,  0,  1, -1]]);
Delta:= linalg[det](A); Delta1:=linalg[det](B1);
x[1]:=Delta1/Delta;
```

The Cayley-Hamilton Theorem

Presented here is an adjoint formula $F^{-1} = \textbf{adj}(F)/\det(F)$ derivation for the celebrated Cayley-Hamilton formula

$$(-A)^n + p_{n-1}(-A)^{n-1} + \cdots + p_0 I = 0.$$

The $n \times n$ matrix A is given and I is the identity matrix. The coefficients p_k in (14) are determined by the **characteristic polynomial** of matrix A, which is defined by the determinant expansion formula

$$|A - \lambda I| = (-\lambda)^n + p_{n-1}(-\lambda)^{n-1} + \cdots + p_0(-\lambda)^0. \tag{12}$$

The **characteristic equation of** A is $|A - \lambda I| = 0$, explicitly

$$(-\lambda)^n + p_{n-1}(-\lambda)^{n-1} + \cdots + p_0(-\lambda)^0 = 0. \tag{13}$$

Theorem 5.20 (Cayley-Hamilton)
A square matrix A satisfies its own characteristic equation. In detail, given characteristic equation $(-\lambda)^n + p_{n-1}(-\lambda)^{n-1} + \cdots + p_0(-\lambda)^0 = 0$, then replace λ on the left by A and zero on the right side by the zero matrix $\mathbf{0}$ to obtain

$$(-A)^n + p_{n-1}(-A)^{n-1} + \cdots + p_0 I = \mathbf{0}. \tag{14}$$

Proof of (14): Define $x = -\lambda$, $F = A + xI$ and $G = \textbf{adj}(F)$. A cofactor of $\det(F)$ is a polynomial in x of degree at most $n-1$. Therefore, there are $n \times n$ constant matrices $C_0, \ldots, C_{n-1}$ such that

$$\textbf{adj}(F) = x^{n-1}C_{n-1} + \cdots + xC_1 + C_0.$$

The adjugate identity $\det(F)I = \textbf{adj}(F)\,F$ is valid for any square matrix F, even if $\det(F)$ is zero. Relation (13) implies $\det(F) = x^n + p_{n-1}x^{n-1} + \cdots + p_0$. Expand the matrix product $\textbf{adj}(F)F$ in powers of x as follows:

$$\begin{aligned} \textbf{adj}(F)F &= \left(\sum_{j=0}^{n-1} x^j C_j\right)(A + xI) \\ &= C_0 A + \sum_{i=1}^{n-1} x^i (C_i A + C_{i-1}) + x^n C_{n-1}. \end{aligned}$$

Match coefficients of powers of x on each side of $\det(F)I = \textbf{adj}(F)F$ to give the relations

$$\left\{ \begin{array}{rcl} p_0 I &=& C_0 A, \\ p_1 I &=& C_1 A + C_0, \\ p_2 I &=& C_2 A + C_1, \\ &\vdots& \\ I &=& C_{n-1}. \end{array} \right. \tag{15}$$

To complete the proof of the Cayley-Hamilton identity (14), multiply the equations in (15) by I, $(-A)$, $(-A)^2$, $(-A)^3$, ..., $(-A)^n$, respectively. Then add all the equations. The left side matches the left side of (14). The right side is a telescoping sum which adds to the zero matrix. ■

An Applied Definition of Determinant

To be developed here is another way to look at formula (9), which emphasizes the column and row structure of a determinant. The definition, which agrees with (9), leads to a short proof of the four properties, which are used to find the value of any determinant.

Permutation Matrices

A matrix P obtained from the identity matrix I by swapping rows is called a **permutation matrix**. There are $n!$ permutation matrices. To illustrate, the 3×3 permutation matrices are

$$\begin{pmatrix} 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{pmatrix}, \begin{pmatrix} 1\,0\,0 \\ 0\,0\,1 \\ 0\,1\,0 \end{pmatrix}, \begin{pmatrix} 0\,1\,0 \\ 1\,0\,0 \\ 0\,0\,1 \end{pmatrix}, \begin{pmatrix} 0\,1\,0 \\ 0\,0\,1 \\ 1\,0\,0 \end{pmatrix}, \begin{pmatrix} 0\,0\,1 \\ 1\,0\,0 \\ 0\,1\,0 \end{pmatrix}, \begin{pmatrix} 0\,0\,1 \\ 0\,1\,0 \\ 1\,0\,0 \end{pmatrix}.$$

Define for a permutation matrix P the determinant by

$$\det(P) = (-1)^k$$

where k is the least number of row swaps required to convert P to the identity. The number k satisfies $r = k+2m$, where r is *any* count of row swaps that changes P to the identity, and m is some integer. Therefore, $\det(P) = (-1)^k = (-1)^r$. In the illustration, the corresponding determinants are 1, -1, -1, 1, 1, -1, as computed from $\det(P) = (-1)^r$, where r row swaps change P into I.

It can be verified that $\det(P)$ agrees with the value reported by formula (9). Each σ in (9) corresponds to a permutation matrix P with rows arranged in the order specified by σ. The summation in (9) for $A = P$ has exactly one nonzero term.

Sampled Product

Let A be an $n \times n$ matrix and P an $n \times n$ permutation matrix. The matrix P has ones in exactly n locations. Sampled product $A.P$ multiplies entries from the matrix A, selected by the location of the ones in P.

Definition 5.19 (Sampled Product $A.P$)
Let $\vec{A}_1, \ldots, \vec{A}_n$ be the rows of A and let $\vec{P}_1, \ldots, \vec{P}_n$ be the rows of P. Let the rows of P be rows $\sigma_1,\ldots,\sigma_n$ of identity matrix I. Define via the normal dot product $(\cdot)$ the **sampled product**

$$(16) \qquad \begin{aligned} A.P &= (A_1 \cdot P_1)(A_2 \cdot P_2)\cdots(A_n \cdot P_n) \\ &= a_{1\sigma_1} \cdots a_{n\sigma_n}. \end{aligned}$$

Equation (16) implies that $A.P$ is a linear function of the rows of A. Replace rows by columns and repeat definition (16) to show $A.P$ is a linear function of the columns of A with value $a_{\sigma_1 1} \cdots a_{\sigma_n n}$.

Sampled-Product Determinant Formula

An alternative definition of determinant is

$$\det(A) = \textstyle\sum_P \det(P)\, A.P, \tag{17}$$

where the summation extends over all possible permutation matrices P. The definition emphasizes the explicit linear dependence of the determinant upon the rows of A (or the columns of A). A tedious but otherwise routine justification shows that the college algebra definition of determinant (9) and the sampled product definition of determinant (17) give the same value.

Three Properties that Define a Determinant

Write the determinant $\det(A)$ in terms of the rows $A_1, \ldots, A_n$ of the matrix A as follows:

$$D_1(A_1, \ldots, A_n) = \sum_P \det(P)\, A.P.$$

Already known is that $D_1(A_1, \ldots, A_n)$ is a function D that satisfies the following **three properties**:

Linearity	D is linear in each argument $A_1, \ldots, A_n$.
Swap	D changes sign if two arguments are swapped. Equivalently, $D = 0$ if two arguments are equal.
Identity	$D = 1$ when $A = I$.

The equivalence reported in **swap** is obtained by expansion, e.g., for $n = 2$, $A_1 = A_2$ implies $D(A_1, A_2) = -D(A_1, A_2)$ and hence $D = 0$. Similarly, $D(A_1 + A_2, A_1 + A_2) = 0$ implies by linearity that $D(A_1, A_2) = -D(A_2, A_1)$, which is the swap property for $n = 2$.

It is less obvious that *the three properties uniquely define the determinant*:

Theorem 5.21 (Uniqueness)
If $D(A_1, \ldots, A_n)$ satisfies the properties of **linearity**, **swap** and **identity**, then $D(A_1, \ldots, A_n) = \det(A)$.

Proof: The rows of the identity matrix I are denoted $E_1, \ldots, E_n$, so that for $1 \le j \le n$ we may write the expansion

$$A_j = a_{j1}E_1 + a_{j2}E_2 + \cdots + a_{jn}E_n. \tag{18}$$

We illustrate the proof for the case $n = 2$:

$$\begin{aligned}
D(A_1, A_2) &= D(a_{11}E_1 + a_{12}E_2, A_2) && \text{By (18).}\\
&= a_{11}D(E_1, A_2) + a_{12}D(E_2, A_2) && \text{By linearity.}\\
&= a_{11}a_{22}D(E_1, E_2) + a_{11}a_{21}D(E_1, E_1) && \text{Repeat for } A_2.\\
&\quad + a_{12}a_{21}D(E_2, E_1) + a_{12}a_{22}D(E_2, E_2)
\end{aligned}$$

The swap and identity properties give $D(E_1, E_1) = D(E_2, E_2) = 0$ and $1 = D(E_1, E_2) = -D(E_2, E_1)$. Therefore, $D(A_1, A_2) = a_{11}a_{22} - a_{12}a_{21}$ and this implies that $D(A_1, A_2) = \det(A)$.

The proof for general n depends upon the identity

$$\begin{aligned} D(E_{\sigma_1}, \ldots, E_{\sigma_n}) &= (-1)^{\text{parity}(\sigma)} D(E_1, \ldots, E_n) \\ &= (-1)^{\text{parity}(\sigma)} \end{aligned}$$

where $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a rearrangement of the integers 1, ..., n. This identity is implied by the **swap** and **identity** properties. Then, as in the case $n = 2$, **linearity** implies that

$$\begin{aligned} D(A_1, \ldots, A_n) &= \textstyle\sum_\sigma a_{1\sigma_1} \cdots a_{n\sigma_n} D(E_{\sigma_1}, \ldots, E_{\sigma_n}) \\ &= \textstyle\sum_\sigma (-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n} \\ &= \det(A). \end{aligned}$$

Proofs and Details

Verification of the Four Properties:

The details will use the sampled product $A.P$ defined on page 358 and the sampled product determinant formula (17) page 359. This is done only for clarity of proof, because it is possible to use the clumsier college algebra definition of determinant (9) page 345.

Triangular. If A is $n \times n$ triangular, then in (17) appears only one nonzero term, due to zero factors in the product $A.P$. The term that appears corresponds to P=identity, therefore $A.P$ is the product of the diagonal elements of A. Since $\det(P) = \det(I) = 1$, the result follows. A similar proof can be constructed from college algebra determinant definition (9), using intuition from Sarrus' rule.

Swap. Let elementary swap matrix Q be obtained from I by swapping rows i and j. Let $B = QA$, then B equals matrix A with rows i and j swapped. To be shown: $\det(A) = -\det(B)$. By definition, $B.P = QA.P$. With effort, it is possible to show that $QA.P = P.QA = PQ.A = A.PQ$ and $\det(PQ) = -\det(P)$. Matrices PQ over all possible P duplicates the list of all permutation matrices. Then definition (17) implies the result.

Combination. Let matrix B be obtained from matrix A by adding to row j the row vector k times row i ($i \neq j$). Then $\mathbf{row}(B, j) = \mathbf{row}(A, j) + k\,\mathbf{row}(A, i)$ and $B.P = (B_1 \cdot P) \cdots (B_n \cdot P) = A.P + k\,C.P$, where C is the matrix obtained from A by *replacing* $\mathbf{row}(A, j)$ with $\mathbf{row}(A, i)$.

Matrix C has equal rows $\mathbf{row}(C, i) = \mathbf{row}(C, j) = \mathbf{row}(A, i)$. By the swap rule applied to rows i and j, $|C| = -|C|$, or $|C| = 0$. Add on P across $B.P = A.P + k\,C.P$ to obtain $|B| = |A| + k|C|$. Then $|B| = |A|$.

Multiply. Let matrices A and B have the same rows, except $\mathbf{row}(B, i) = c\,\mathbf{row}(A, i)$ for some index i. Then $B.P = c\,A.P$. Add on P across this equation to obtain $|B| = c|A|$.

Verification of the Additional Rules:

Zero row. Apply the common factor rule with $c = 2$, possible since the row has all zero entries. Then $|A| = 2|A|$, which implies $|A| = 0$.

Duplicate rows. The swap rule applies to the two duplicate rows to give $|A| = -|A|$, which implies $|A| = 0$.

Dependent rows. The determinant is unchanged by adding a linear combination of rows of A to a different row, the result a matrix B. Then $|A| = |B|$. Select the combination to create a row of zeros in B. Then $|B| = 0$ from **zero row**, implying $|A| = 0$.

RREF $\neq I$. Each step in a toolkit sequence to the RREF gives $|A| = |EB| = |E||B|$ where E is an elementary matrix and B is one frame closer to $rref(A)$. At some point $B = \mathbf{rref}(A)$, then $B \neq I$ means B has a row of zeros. Therefore, $|B| = 0$, which implies $|A| = |E||B| = 0$.

Common factor and row linearity. The sampled product $A.P$ is a linear function of each row, therefore the same is true of $|A|$ by the sampled product determinant formula (17) page 359.

Derivation of cofactor expansion (10): The column expansion formula is derived from the row expansion formula applied to the transpose. We consider only the derivation of the row expansion formula (10) for $k = 1$, because the case for general k is the same except for notation. The plan is to establish equality of the two sides of (10) for $k = 1$, which in terms of $\mathbf{minor}(A, 1, j) = (-1)^{1+j}\,\mathbf{cof}(A, 1, j)$ is the equality

$$\det(A) = \sum_{j=1}^{n} a_{1j}(-1)^{1+j}\,\mathbf{minor}(A, 1, j). \tag{19}$$

The details require expansion of $\mathbf{minor}(A, 1, j)$ in (19) via the definition of determinant $\det(A) = \sum_\sigma (-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n}$. A typical term on the right in (19) after expansion looks like

$$a_{1j}\,(-1)^{1+j}(-1)^{\text{parity}(\alpha)} a_{2\alpha_2} \cdots a_{n\alpha_n}.$$

Here, α is a rearrangement of the set of $n-1$ elements consisting of $1, \ldots, j-1, j+1, \ldots, n$. Define $\sigma = (j, \alpha_2, \ldots, \alpha_n)$, which is a rearrangement of symbols $1, \ldots, n$. After parity(α) interchanges, α is changed into $(1, \ldots, j-1, j+1, \ldots, n)$ and therefore these same interchanges transform σ into $(j, 1, \ldots, j-1, j+1, \ldots, n)$. An additional $j-1$ interchanges will transform σ into natural order $(1, \ldots, n)$. This establishes, because of $(-1)^{j-1} = (-1)^{j+1}$, the identity

$$\begin{aligned}(-1)^{\text{parity}(\sigma)} &= (-1)^{j-1+\text{parity}(\alpha)} \\ &= (-1)^{j+1+\text{parity}(\alpha)}.\end{aligned}$$

Collecting formulas gives

$$(-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n} = a_{1j}\,(-1)^{1+j}(-1)^{\text{parity}(\alpha)} a_{2\alpha_2} \cdots a_{n\alpha_n}.$$

Adding across this formula over all α and j gives a sum on the right which matches the right side of (19). Some additional thought reveals that the terms on the left add exactly to $\det(A)$, hence (19) is proved.

Derivation of Cramer's Rule: The cofactor column expansion theory implies that the Cramer's rule solution of $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is given by

$$x_j = \frac{\Delta_j}{\Delta} = \frac{1}{\Delta}\sum_{k=1}^{n} b_k\,\mathbf{cof}(A, k, j). \tag{20}$$

We will verify that $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. Let $E_1, \ldots, E_n$ be the rows of the identity matrix. The question reduces to showing that $E_p A\vec{\mathbf{x}} = b_p$. The details will use the fact

$$\sum_{j=1}^{n} a_{pj}\,\mathbf{cof}(A,k,j) = \begin{cases} \det(A) & \text{for } k = p, \\ 0 & \text{for } k \ne p, \end{cases} \tag{21}$$

Equation (21) follows by cofactor row expansion, because the sum on the left is $\det(B)$ where B is matrix A with row k replaced by row p. If B has two equal rows, then $\det(B) = 0$; otherwise, $B = A$ and $\det(B) = \det(A)$.

$$\begin{aligned}
E_p A\vec{\mathbf{x}} &= \sum_{j=1}^{n} a_{pj} x_j \\
&= \frac{1}{\Delta}\sum_{j=1}^{n} a_{pj} \sum_{k=1}^{n} b_k\,\mathbf{cof}(A,k,j) && \text{Apply formula (20).} \\
&= \frac{1}{\Delta}\sum_{k=1}^{n} b_k \left(\sum_{j=1}^{n} a_{pj}\,\mathbf{cof}(A,k,j)\right) && \text{Switch order of summation.} \\
&= b_p && \text{Apply (21).}
\end{aligned}$$

Derivation of $A \cdot \mathbf{adj}(A) = \det(A) I$: The proof uses formula (21). Consider column k of $\mathbf{adj}(A)$, denoted $\vec{X}$, multiplied against matrix A, which gives

$$A\vec{X} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\,\mathbf{cof}(A,k,j) \\ \sum_{j=1}^{n} a_{2j}\,\mathbf{cof}(A,k,j) \\ \vdots \\ \sum_{j=1}^{n} a_{nj}\,\mathbf{cof}(A,k,j) \end{pmatrix}.$$

By formula (21),

$$\sum_{j=1}^{n} a_{ij}\,\mathbf{cof}(A,k,j) = \begin{cases} \det(A) & i = k, \\ 0 & i \ne k. \end{cases}$$

Therefore, $A\vec{X}$ is $\det(A)$ times column k of the identity I. ■

Exercises 5.3

Determinant Notation

Write formulae for x and y as quotients of 2×2 determinants. Do not evaluate the determinants!

1. $\begin{pmatrix} 1 & -1 \\ 2 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -10 \\ 3 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10 \\ -6 \end{pmatrix}$

3. $\begin{pmatrix} 0 & -1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 10 \end{pmatrix}$

4. $\begin{pmatrix} 0 & -3 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

Sarrus' 2×2 rule

Evaluate $\det(A)$.

5. $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

6. $A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$

7. $A = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$

8. $A = \begin{pmatrix} 5a & 1 \\ -1 & 2a \end{pmatrix}$

Sarrus' rule 3×3

Evaluate $\det(A)$.

9. $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$

10. $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

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

12. $A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 2 & -1 \\ 1 & 1 & -1 \end{pmatrix}$

Inverse of a 2×2 Matrix

Define matrix A and its adjugate C:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad C = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

13. Verify $AC = |A| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

14. Display the details of the argument that $|A| \neq 0$ implies A^{-1} exists and $A^{-1} = \dfrac{C}{|A|}$.

15. Show that A^{-1} exists implies $|A| \neq 0$. Suggestion: Assume not, then $AB = BA = I$ for some matrix B and also $|A| = 0$. Find a contradiction using $AC = |A|I$ from Exercise 13.

16. Calculate the inverse of $\begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}$ using the formula developed in these exercises.

Unique Solution of a 2×2 System

Solve $A\vec{\mathbf{X}} = \vec{\mathbf{b}}$ for $\vec{\mathbf{X}}$ using Cramer's rule for 2×2 systems.

17. $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$, $\vec{\mathbf{b}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

18. $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$, $\vec{\mathbf{b}} = \begin{pmatrix} 5 \\ -5 \end{pmatrix}$

19. $A = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$, $\vec{\mathbf{b}} = \begin{pmatrix} -4 \\ 4 \end{pmatrix}$

20. $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $\vec{\mathbf{b}} = \begin{pmatrix} -10 \\ 10 \end{pmatrix}$

Definition of Determinant

21. Let A be 3×3 with zero first row. Use the college algebra definition of determinant to show that $\det(A) = 0$.

22. Let A be 3×3 with equal first and second row. Use the college algebra definition of determinant to show that $\det(A) = 0$.

23. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Use the college algebra definition of determinant to verify that $|A| = ad - bc$.

24. Let $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Use the college algebra definition of determinant to verify that the determinant of A equals

$$\begin{array}{l} a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} \\ +a_{31}a_{12}a_{23} - a_{11}a_{32}a_{23} \\ -a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} \end{array}$$

Four Properties

Evaluate $\det(A)$ using the four properties for determinants, page 345.

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

26. $A = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

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

28. $A = \begin{pmatrix} 2 & 4 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

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

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

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

32. $A = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{pmatrix}$

Elementary Matrices and the Four Rules

Find $\det(A)$.

33. A is 3×3 and obtained from the identity matrix I by three row swaps.

34. A is 7×7, obtained from I by swapping rows 1 and 2, then rows 4 and 1, then rows 1 and 3.

35. A is obtained from the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ by swapping rows 1 and 3, then two row combinations.

36. A is obtained from the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ by two row combinations, then multiply row 2 by -5.

More Determinant Rules

Cite the determinant rule that verifies $\det(A) = 0$. **Never** expand $\det(A)$! See page 347.

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

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

39. $A = \begin{pmatrix} 4 & -8 & -8 \\ 2 & -4 & -4 \\ 1 & 1 & -3 \end{pmatrix}$

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

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

42. $A = \begin{pmatrix} -1 & 5 & 4 \\ 2 & -4 & -2 \\ 1 & 1 & 2 \end{pmatrix}$

Cofactor Expansion and College Algebra

Evaluate the determinant with an efficient cofactor expansion.

43. $\begin{vmatrix} 2 & 5 & 1 \\ 2 & 0 & -4 \\ 1 & 0 & 0 \end{vmatrix}$

44. $\begin{vmatrix} 2 & 5 & 1 \\ 2 & 0 & -4 \\ 1 & 0 & 1 \end{vmatrix}$

45. $\begin{vmatrix} 2 & 5 & 0 & 0 \\ 2 & 1 & 4 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{vmatrix}$

46. $\begin{vmatrix} 0 & 2 & 0 & 1 \\ 2 & 3 & 2 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{vmatrix}$

47. $\begin{vmatrix} 2 & 5 & 1 & -1 & 1 \\ 0 & -1 & -4 & 1 & -1 \\ 1 & 2 & 3 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{vmatrix}$

48. $\begin{vmatrix} 2 & 0 & 1 & -1 & 1 \\ 0 & -1 & -4 & 1 & -1 \\ 1 & 2 & 3 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 \\ 1 & 2 & 0 & 1 & 1 \end{vmatrix}$

Minors and Cofactors

Write out and then evaluate the minor and cofactor of each element cited for the matrix $A = \begin{pmatrix} 2 & 5 & y \\ x & -1 & -4 \\ 1 & 2 & z \end{pmatrix}$

49. Row 1 and column 3.

50. Row 2 and column 1.

51. Row 3 and column 2.

52. Row 3 and column 1.

Cofactor Expansion

Use cofactors to evaluate the determinant.

53. $\begin{vmatrix} 2 & 7 & 1 \\ -1 & 0 & -4 \\ 1 & 0 & 3 \end{vmatrix}$

54. $\begin{vmatrix} 2 & 7 & 7 \\ -1 & 1 & 0 \\ 1 & 2 & 0 \end{vmatrix}$

55. $\begin{vmatrix} 0 & 2 & 7 & 7 \\ 0 & -1 & 1 & 0 \\ 3 & 1 & 2 & 0 \\ 0 & -1 & 1 & 0 \end{vmatrix}$

56. $\begin{vmatrix} 0 & 2 & 7 & 7 \\ 0 & -1 & y & 0 \\ x & 1 & 2 & 0 \\ 0 & -1 & 1 & 0 \end{vmatrix}$

57. $\begin{vmatrix} 0 & 2 & 7 & 7 & 3 \\ 0 & -1 & 0 & 0 & 1 \\ x & 1 & 2 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 \end{vmatrix}$

58. $\begin{vmatrix} 0 & 2 & 7 & 7 & 3 \\ 0 & -1 & 2 & 0 & 1 \\ x & 1 & 2 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 \end{vmatrix}$

Adjugate and Inverse Matrix

Find the adjugate of A and the inverse B of A. Check the answers via the formulas $A\,\mathbf{adj}(A) = \det(A)I$ and $AB = I$.

59. $A = \begin{pmatrix} 2 & 7 \\ -1 & 0 \end{pmatrix}$

60. $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$

61. $A = \begin{pmatrix} 5 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & 0 & 3 \end{pmatrix}$

62. $A = \begin{pmatrix} 5 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & 0 & 3 \end{pmatrix}$

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

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

Transpose and Inverse

65. Verify that $A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ satisfies $A^T = A^{-1}$.

66. Find all 2×2 matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $\det(A) = 1$ and $A^T = A^{-1}$.

67. Find all 3×3 diagonal matrices A such that $A^T = A^{-1}$.

68. Find all 3×3 upper triangular matrices A such that $A^T = A^{-1}$.

69. Find all $n \times n$ diagonal matrices A such that $A^T = A^{-1}$.

70. Determine the $n \times n$ triangular matrices A such that $\det(A) = 1$ and $A^T = \mathbf{adj}(A)$.

Elementary Matrices

Find the determinant of A from the given equation.

71. Let $A = 5E_2E_1$ be 3×3, where E_1 multiplies row 3 of the identity by -7 and E_2 swaps rows 3 and 1 of the identity. Hint: $A = (5I)E_2E_1$.

72. Let $A = 2E_2E_1$ be 5×5, where E_1 multiplies row 3 of the identity by -2 and E_2 swaps rows 3 and 5 of the identity.

73. Let $A = E_2E_1B$ be 4×4, where E_1 multiplies row 2 of the identity by 3 and E_2 is a combination. Find $|A|$ in terms of $|B|$.

74. Let $A = 3E_2E_1B$ be 3×3, where E_1 multiplies row 2 of the identity by 3 and E_2 is a combination. Find $|A|$ in terms of $|B|$.

75. Let $A = 4E_2E_1B$ be 3×3, where E_1 multiplies row 1 of the identity by 2, E_2 is a combination and $|B| = -1$.

76. Let $A = 2E_3E_2E_1B^3$ be 3×3, where E_1 multiplies row 2 of the identity by -1, E_2 and E_3 are swaps and $|B| = -2$.

Determinants and the Toolkit

Display the toolkit steps for $\mathbf{rref}(A)$. Using only the steps, report:

- The determinant of the elementary matrix E for each step.
- The determinant of A.

77. $A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 0 & 2 \\ 1 & 0 & 4 \end{pmatrix}$

78. $A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$

79. $A = \begin{pmatrix} 2 & 3 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 0 & 2 & 1 \end{pmatrix}$

80. $A = \begin{pmatrix} 2 & 3 & 1 & 2 \\ 0 & 3 & 0 & 0 \\ 2 & 6 & 1 & 2 \\ 1 & 0 & 2 & 1 \end{pmatrix}$

Determinant Product Rule

Apply the product rule $\det(AB) = \det(A)\det(B)$.

81. Let $\det(A) = 5$ and $\det(B) = -2$. Find $\det(A^2B^3)$.

82. Let $\det(A) = 4$ and $A(B - 2A) = 0$. Find $\det(B)$.

83. Let $A = E_1E_2E_3$ where E_1, E_2 are elementary swap matrices and E_3 is an elementary combination matrix. Find $\det(A)$.

84. Assume $\det(AB+A) = 0$ and $\det(A) \neq 0$. Show that $\det(B + I) = 0$.

Cramer's 2×2 Rule

Assume

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}.$$

85. Derive the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & 0 \\ y & 1 \end{pmatrix} = \begin{pmatrix} e & b \\ f & d \end{pmatrix}.$$

86. Derive the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix} = \begin{pmatrix} a & e \\ c & f \end{pmatrix}.$$

87. Use the determinant product rule to derive the Cramer's Rule formula

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

88. Derive, using the determinant product rule, the Cramer's Rule formula

$$y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

Cramer's 3×3 Rule

Let A be the coefficient matrix in the equation

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

89. Derive the formula

$$A \begin{pmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}$$

90. Derive the formula

$$A \begin{pmatrix} 1 & 0 & x_1 \\ 0 & 1 & x_2 \\ 0 & 0 & x_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}$$

91. Derive, using the determinant product rule, the Cramer's Rule formula

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

92. Use the determinant product rule to derive the Cramer's Rule formula

$$x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

Cayley-Hamilton Theorem

93. Let $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$. Expand $|A - rI|$ to compute the characteristic polynomial of A. Answer: $r^2 - 4r + 5$.

94. Let $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$. Apply the Cayley-Hamiltion theorem to justify the equation

$$A^2 - 4A + 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

95. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Expand $|A - rI|$ by Sarrus' Rule to obtain $r^2 - (a+b)r + (ad - bc)$.

96. The result of the previous exercise is often written as $(-r)^2 + \mathbf{trace}(A)(-r) + |A|$ where $\mathbf{trace}(A) = a + d =$ sum of the diagonal elements. Display the details.

97. Let $\lambda^2 - 2\lambda + 1 = 0$ be the characteristic equation of a matrix A. Find a formula for A^2 in terms of A and I.

98. Let A be an $n \times n$ triangular matrix with all diagonal entries zero. Prove that $A^n = 0$.

99. Find all 2×2 matrices A such that $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, discovered from values of $\mathbf{trace}(A)$ and $|A|$.

100. Find four 2×2 matrices A such that $A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Applied Definition of Determinant

Miscellany for permutation matrices and the sampled product page 358

$$\begin{aligned} A.P &= (A_1 \cdot P_1)(A_2 \cdot P_2) \cdots (A_n \cdot P_n) \\ &= a_{1\sigma_1} \cdots a_{n\sigma_n}. \end{aligned}$$

101. Compute the sampled product of $\begin{pmatrix} 5 & 3 & 1 \\ 0 & 5 & 7 \\ 1 & 9 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

102. Compute the sampled product of $\begin{pmatrix} 5 & 3 & 3 \\ 0 & 2 & 7 \\ 1 & 9 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

103. Determine the permutation matrices P required to evaluate $\det(A)$ when A is 2×2.

104. Determine the permutation matrices P required to evaluate $\det(A)$ when A is 4×4.

Three Properties

Reference: Page 359, three properties that define a determinant

105. Assume $n = 3$. Prove that the three properties imply $D = 0$ when two rows are identical.

106. Assume $n = 3$. Prove that the three properties imply $D = 0$ when a row is zero.

5.4 Vector Spaces, Independence, Basis

The technical topics of independence, dependence and span apply to the study of Euclidean spaces $\mathcal{R}^2$, $\mathcal{R}^3$, ..., $\mathcal{R}^n$ and also to the continuous function space $C(E)$, the space of differentiable functions $C^1(E)$ and its generalization $C^n(E)$, and to general abstract vector spaces.

Basis and General Solution: Algebraic Equations

The term **basis** was introduced on page **??** for systems of linear algebraic equations. To review, a basis is obtained from the vector general solution $\vec{x}$ of matrix equation $A\vec{x} = \vec{0}$ by computing the partial derivatives ∂_{t_1}, ∂_{t_2}, ... of $\vec{x}$, where t_1, t_2, ... is the list of invented symbols assigned to the free variables identified in **rref**(A). The partial derivatives are **Strang's special solutions**[6] to the homogeneous equation $A\vec{x} = \vec{0}$. Solution $\vec{v}_i$ is also found by letting $t_i = 1$ with all other invented symbols zero, $1 \le i \le k$. Knowing the special solutions enables reconstruction of the general solution: multiply by constants and add.

> The general solution of $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ is the sum of constants times Strang's special solutions (they are a **basis**).

Deeper properties have been isolated for the list of Strang's special solutions, the partial derivatives $\partial_{t_1}\vec{x}$, $\partial_{t_2}\vec{x}$, ... The most important properties are **span** and **independence**.

Span, Independence and Basis

Definition 5.20 (Span of a Set of Vectors)
A list of vectors $\vec{v}_1$, ..., $\vec{v}_k$ is said to **span** an abstract vector space V (page 301), written

$$V = \mathbf{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k),$$

provided V consists of exactly the set of all linear combinations

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

for all choices of constants $c_1, \ldots, c_k$.

The notion originates with the general solution $\vec{v}$ of a homogeneous matrix system $A\vec{v} = \vec{0}$, where the invented symbols t_1, ..., t_k are the constants c_1, ..., c_k and the vector partial derivative list $\partial_{t_1}\vec{v}$, ..., $\partial_{t_k}\vec{v}$ is the list of vectors $\vec{v}_1$, ..., $\vec{v}_k$.

[6]The nomenclature is due to Gilbert Strang [?], with **Strang's special solutions** an appropriate reference.

Definition 5.21 (Independence of Vectors)
Vectors $\vec{v}_1$, ..., $\vec{v}_k$ in an abstract vector space V are said to be **Independent** or **Linearly independent** provided each linear combination $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is represented by a unique set of constants c_1, ..., c_k. The unique constants are called the **weights** of vector $\vec{v}$ relative to $\vec{v}_1$, ..., $\vec{v}_k$.

See pages 377 and 382 for independence tests.

Unique representation of linear combinations has an algebraic equivalent:

Linear Independence of Vectors $\vec{v}_1$, ..., $\vec{v}_k$

If two linear combinations are equal,

$$a_1\vec{v}_1 + \cdots + a_k\vec{v}_k = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k,$$

then the coefficients match

$$a_1 = b_1,\ a_2 = b_2, \ldots, a_k = b_k.$$

Definition 5.22 (Basis)
A **basis** of an abstract vector space V is defined to be a list of independent vectors $\vec{v}_1$, ..., $\vec{v}_k$ which spans V. A basis is tested by two checkpoints:

1. The list of vectors $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_k$ is independent.
2. The vectors span V, written $V = \mathbf{span}(\vec{v}_1, \ldots, \vec{v}_k)$.

A basis expresses the **general solution** of a linear problem with *the fewest possible terms.*

Theorem 5.22 (Independence of Strang's Special Solutions)
Assume matrix equation $A\vec{x} = \vec{0}$ with scalar general solution $x_1, x_2, \ldots, x_n$ using invented symbols $t_1, t_2, \ldots, t_k$. Define k **special solutions** by partial differentiation:

$$\vec{v}_1 = \partial_{t_1}\vec{x}, \quad \vec{v}_2 = \partial_{t_2}\vec{x}, \quad \ldots, \quad \vec{v}_k = \partial_{t_k}\vec{x}$$

Then:

1. Each solution $\vec{x}$ of $A\vec{x} = \vec{0}$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$.
2. The vectors $\vec{v}_1, \ldots, \vec{v}_k$ are independent.

Briefly: Strang's special solutions are independent and they form a basis for the set of solutions to $A\vec{x} = \vec{0}$. See also the Kernel Theorem 5.2 page 300.

Proof on page 393

Vector Space $\mathcal{R}^n$

The vector space $\mathcal{R}^n$ of n-element fixed column vectors (or row vectors) is from the view of applications a *storage system for organization of numerical data sets* that is equipped with an algebraic toolkit. The scheme induces a *data structure* onto the numerical data set. In particular, whether needed or not, there are pre-defined operations of addition (+) and scalar multiplication ($\cdot$) which apply to fixed vectors. The two operations on fixed vectors satisfy the *closure law* and in addition obey the *eight algebraic vector space properties.* The vector space $V = \mathcal{R}^n$ is viewed as a **data set** consisting of data item packages.

Algebraic Toolkit

The **toolkit** for an abstract vector space V is the following set of eight algebraic properties. Set V is a data set. Elements of V are data packages called **vectors**, denoted $\vec{\mathbf{X}}$ and $\vec{\mathbf{Y}}$ in the toolkit.

Closure	The operations $\vec{X}+\vec{Y}$ and $k\vec{X}$ are defined and result in a new vector which is also in the set V.	
Addition	$\vec{X}+\vec{Y}=\vec{Y}+\vec{X}$	commutative
	$\vec{X}+(\vec{Y}+\vec{Z})=(\vec{X}+\vec{Y})+\vec{Z}$	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

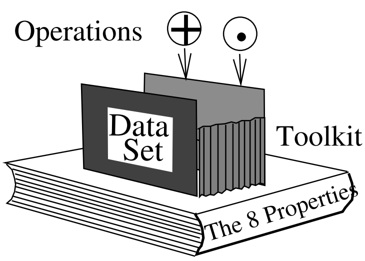

Figure 12. A Data Storage System.
A vector space is a data set of data item packages plus a storage system which organizes the data. A toolkit is provided consisting of operations + and $\cdot$ plus 8 algebraic vector space properties.

Fixed Vectors and the Toolkit

Scalar multiplication of fixed vectors is commonly used for re-scaling, especially to unit systems fps, cgs and mks. For instance, a numerical data set of lengths recorded in meters (mks) is re-scaled to centimeters (cgs) using scale factor $k = 100$.

Addition and subtraction of fixed vectors is used in a variety of calculations, which includes averages, difference quotients and calculus operations like integration.

Planar Plot Vector Toolkit Example

The data set for a plot problem consists of plot points in $\mathcal{R}^2$ which are the **dots** for the connect-the-dots graphic. Assume the function $y(x)$ to be plotted comes from differential equation $y' = f(x, y)$. Euler's numerical method applies to compute the sequence of dots in the graphic. In this algorithm, the next dot is represented as $\vec{v}_2 = \vec{v}_1 + \vec{E}(\vec{v}_1)$ where symbol $\vec{v}_1$ is the previous dot. Symbol $\vec{E}(\vec{v}_1)$ is the Euler increment. Definitions:

$$\vec{v}_1 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad \vec{E}(\vec{v}_1) = h \begin{pmatrix} 1 \\ f(x_0, y_0) \end{pmatrix},$$
$$\vec{v}_2 = \vec{v}_1 + \vec{E}(\vec{v}_1) = \begin{pmatrix} x_0 + h \\ y_0 + hf(x_0, y_0) \end{pmatrix}.$$

Step size $h = 0.05$ is a common instance. The Euler increment $\vec{E}(\vec{v}_1)$ is defined as scalar multiplication by h against an $\mathcal{R}^2$-vector which contains an evaluation of f at the previous dot $\vec{v}_1$.

Summary. The **dots** for the graphic of $y(x)$ form a data set in the vector space $\mathcal{R}^2$. The dots are obtained by algorithm rules, which are easily expressed by vector addition $(+)$ and scalar multiplication $(\cdot)$. The 8 properties of the toolkit were used in a limited way.

Digital Photographs

A digital photo has many **pixels** arranged in a two dimensional array. Structure can be assigned to the photo by storing the pixel digital color data in a matrix A of size $n \times m$. Each entry of A is an integer which encodes the color information at a specific pixel location.

The set V of all $n \times m$ matrices is a vector space under the usual rules for matrix addition and scalar multiplication. Initially, V is just a storage system for photos. However, the algebraic toolkit for V (page 371) is a convenient way to express operations on photos. An illustration: reconstruction of a photo from RGB (Red, Green, Blue) separation photos.

Let $A = (a_{ij})$ be an $n \times m$ matrix of color data for a photo. One way to encode each entry of A is to define $a_{ij} = r_{ij} + g_{ij}x + b_{ij}x^2$ where x is some convenient base. The integers r_{ij}, g_{ij}, b_{ij} represent the amount of red, green and blue present in the pixel with data a_{ij}. Then $A = R + Gx + Bx^2$ where $R = [r_{ij}]$, $G = [g_{ij}]$, $B = [b_{ij}]$ are $n \times m$ matrices that represent the color separation photos. Construction of matrices R, G, B can be done from A by decoding integer a_{ij} into respective matrix entries. It is done with modular arithmetic. Matrices R, xG and x^2B correspond to three monochromatic photos, which can be realized as color transparencies. The transparencies placed on a standard overhead projector will reconstruct the original photograph.

Printing machinery from many years ago employed separation negatives and multiple printing runs in primary ink colors to make book photos. The advent of digital printers and simpler inexpensive technologies has made the separation process nearly obsolete. To document the historical events, we quote Sam Wang[7]:

> I encountered many difficulties when I first began making gum prints: it was not clear which paper to use; my exposing light (a sun lamp) was highly inadequate; plus a myriad of other problems. I was also using panchromatic film, making in–camera separations, holding RGB filters in front of the camera lens for three exposures onto 3 separate pieces of black and white film. I also made color separation negatives from color transparencies by enlarging in the darkroom. Both of these methods were not only tedious but often produced negatives very difficult to print — densities and contrasts that were hard to control and working in the dark with panchromatic film was definitely not fun. The fact that I got a few halfway decent prints is something of a small miracle, and represents hundreds of hours of frustrating work! Digital negatives by comparison greatly simplify the process. Nowadays (2004) I use color images from digital cameras as well as scans from slides, and the negatives print much more predictably.

Function Spaces

The default storage system used for applications involving ordinary or partial differential equations is a *function space*. The data item packages for differential equations are their solutions, which are *functions*, or in an applied context, a graphic defined on a certain graph window. They are **not** column vectors of numbers.

Functions and Column Vectors

An alternative view, adopted by researchers in numerical solutions of differential equations, is that a solution is a table of numbers, consisting of pairs of x and y values.

It is possible to think of the function as being a fixed vector. The viewpoint is that a function is a **graph** and a graph is determined by so many **dots**, which are practically obtained by **sampling** the function $y(x)$ at a reasonably dense set of x-values. The approximation is

$$y \approx \begin{pmatrix} y(x_1) \\ y(x_2) \\ \vdots \\ y(x_n) \end{pmatrix}$$

where $x_1, \ldots, x_n$ are the **samples** and $y(x_1), \ldots, y(x_n)$ are the **sampled values** of function y.

The trouble with the approximation is that two different functions may need different sampling rates to properly represent their graphic. The result is that

[7] Sam Wang lectured on photography and art with computer at Clemson University in South Carolina. **Reference**: *A Gallery of Tri-Color Prints*, by Sam Wang

the two functions might need data storage of different dimensions, e.g., f needs its sampled values in $\mathcal{R}^{200}$ and g needs its sampled values in $\mathcal{R}^{400}$. The absence of a universal fixed vector storage system for sampled functions explains the appeal of a system like the set of all functions.

Infinitely Long Column Vectors

Is there a way around the lack of a universal numerical data storage system for sampled functions? Is it possible to *develop a theory of column vectors with infinitely many components*? It may help you to think of any function f as an infinitely long column vector, with one entry $f(x)$ for each possible sample x, e.g.,

$$\vec{f} = \begin{pmatrix} \vdots \\ f(x) \\ \vdots \end{pmatrix} \qquad \text{level } x$$

It is not clear how to order or address the entries of such a column vector: at algebraic stages it hinders. Can computers store infinitely long column vectors? The safest path through the algebra is to deal exactly with functions and function notation. Still, there is something attractive about the change from sampled approximations to a single column vector with infinite extent:

$$\vec{f} \approx \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} \to \begin{pmatrix} \vdots \\ f(x) \\ \vdots \end{pmatrix} \qquad \text{level } x$$

The thinking behind the *level* x annotation is that x stands for one of the infinite possibilities for an invented sample. Alternatively, with a rich set of invented samples $x_1, \ldots, x_n$, value $f(x)$ equals approximately $f(x_j)$, where x is closest to some sample x_j.

The Vector Space V of all Functions on a Set E

The rules for function addition and scalar multiplication come from college algebra and pre-calculus backgrounds:

$$(f+g)(x) = f(x) + g(x), \quad (cf)(x) = c \cdot f(x).$$

These rules can be motivated and remembered by the notation of infinitely long column vectors, where level x is an arbitrary **sample**:

$$c_1\vec{f} + c_2\vec{g} = c_1 \begin{pmatrix} \vdots \\ f(x) \\ \vdots \end{pmatrix} + c_2 \begin{pmatrix} \vdots \\ g(x) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ c_1 f(x) + c_2 g(x) \\ \vdots \end{pmatrix}$$

The rules define **addition** and **scalar multiplication** of functions. The closure law for a vector space holds. Routine tedious justifications show that V, under the above rules for addition and scalar multiplication, has the required 8-property toolkit to make it a vector space:

Closure	The operations $f + g$ and kf are defined and result in a new function which is also in the set V of all functions on the set E.	
Addition	$f + g = g + f$	commutative
	$f + (g + h) = (f + g) + h$	associative
	The zero function 0 is defined and $0 + f = f$	zero
	The function $-f$ is defined and $f + (-f) = 0$	negative
Scalar multiply	$k(f + g) = kf + kg$	distributive I
	$(k_1 + k_2)f = k_1 f + k_2 f$	distributive II
	$k_1(k_2 f) = (k_1 k_2)f$	distributive III
	$1f = f$	identity

Important subspaces of the vector space V of all functions appear in applied literature as the storage systems for solutions to differential equations and solutions of related models.

Vector Space $C(E)$

Let $E = \{x : a < x < b\}$ be an open interval on the real line, a, b possibly infinite. The set $C(E)$ is defined to be the subset S of the set V of all functions on E obtained by restricting the function to be continuous. Because sums and scalar multiples of continuous functions are continuous, then $S = C(E)$ is a subspace of V and a vector space in its own right. The definition applies to any nonvoid subset E of $\mathcal{R}^1$.

Vector Space $C^1(E)$

The set $C^1(E)$ is the subset of the vector space $C(E)$ of all continuous functions on open interval E obtained by restricting the function to be continuously differentiable. Because sums and scalar multiples of continuously differentiable functions are continuously differentiable, then $C^1(E)$ is a subspace of $C(E)$ and a vector space in its own right.

Vector Space $C^k(E)$

The set $C^k(E)$ is the subset of the vector space $C(E)$ of all continuous functions on open interval E obtained by restricting the function to be k times continuously differentiable. Because sums and scalar multiples of k times continuously differentiable functions are k times continuously differentiable, then $C^k(E)$ is a subspace of $C(E)$ and a vector space in its own right.

Solution Space of a Differential Equation

The differential equation $y'' - y = 0$ has general solution $y = c_1 e^x + c_2 e^{-x}$, which means that the set S of all solutions of the differential equation consists of all possible linear combinations of the two functions e^x and e^{-x}. Briefly,

$$S = \textbf{span}\left(e^x, e^{-x}\right).$$

The functions e^x, e^{-x} are in $C^2(E)$ for any interval E on the x-axis. Therefore, S is a subspace of $C^2(E)$ and a vector space in its own right.

More generally, every homogeneous linear differential equation, of any order, has a solution set S which is a vector space in its own right.

Invented Vector Spaces

The number of different vector spaces used as data storage systems in scientific literature is finite, but growing with new discoveries. There is really no limit to the number of different vector spaces possible, because creative individuals are able to invent new ones.

Here is an example of how creation begets new vector spaces. Consider the problem $y' = 2y + f(x)$ and the task of storing data for the plotting of an initial value problem with initial condition $y(x_0) = y_0$. The data set V suitable for plotting consists of column vectors

$$\vec{v} = \begin{pmatrix} x_0 \\ y_0 \\ f \end{pmatrix}.$$

A plot command takes such a data item, computes the solution

$$y(x) = y_0 e^{2x} + e^{2x} \int_0^x e^{-2t} f(t)dt$$

and then plots it in a window of fixed size with center at (x_0, y_0). The column vectors are not numerical vectors in $\mathcal{R}^3$, but some **hybrid** of vectors in $\mathcal{R}^2$ and the space of continuous functions $C(E)$ where E is the real line.

It is relatively easy to come up with definitions of vector addition and scalar multiplication on V. The closure law holds and the eight vector space properties can be routinely verified. Therefore, V is an abstract vector space, unlike any found in this text. To reiterate:

> An abstract vector space is a set V and two operations of $\boxed{+}$ and $\boxed{\cdot}$ such that the closure law holds and the eight algebraic vector space properties are satisfied.

The paycheck for having recognized a vector space setting in an application is clarity of exposition and economy of effort in details. Algebraic details in $\mathcal{R}^2$ often transfer unchanged to an abstract vector space setting, line for line, to obtain the details in the more abstract setting.

Independence and Dependence

Independence is defined in Definition 5.21 page 370:

Vectors $\vec{v}_1, \ldots, \vec{v}_k$ are called **independent** provided each linear combination $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is represented by a **unique** set of constants $c_1, \ldots, c_k$.

Independence means **unique representation of linear combinations** of $\vec{v}_1$, $\ldots$, $\vec{v}_k$, which is the statement

$$a_1\vec{v}_1 + \cdots + a_k\vec{v}_k = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$$

implies the coefficients match:

$$\begin{cases} a_1 & = & b_1 \\ a_2 & = & b_2 \\ & \vdots & \\ a_k & = & b_k \end{cases}$$

The subject of *independence* applies to coordinate spaces $\mathcal{R}^n$, function spaces and in particular solution spaces of differential equations, digital photos, sequences of Fourier coefficients or Taylor coefficients, and general abstract vector spaces. Introduced here are definitions for low dimensions, the geometrical meaning of independence, geometric tests for independence and basic algebraic tests for independence.

The motivation for the study of independence is the theory of general solutions, which are expressions representing *all possible solutions* of a linear problem. Independence is a central issue for discovery of *the shortest possible expression* for a general solution.

Definition 5.23 (Dependence)
Vectors $\vec{v}_1$, $\ldots$, $\vec{v}_k$ are called **dependent** provided they are not independent. This means that some linear combination $\vec{v} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$ can be represented in a second way as $\vec{v} = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$ where for at least one index j, $a_j \neq b_j$.

Publications and proofs routinely use a brief abstract definition of independence which is a consequence of Theorem 5.23 below. See Definition 5.24 page 381 for the abstract definition normally used in mathematical proofs and technical publications.

Theorem 5.23 (Unique Representation of the Zero Vector)
Vectors $\vec{v}_1$, $\ldots$, $\vec{v}_k$ are independent in vector space V if and only if the system of equations

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$$

has unique solution $c_1 = \cdots = c_k = 0$. Proof on page 394.

Theorem 5.24 (Subsets of Independent Sets)
Any nonvoid subset of an independent set is also independent.

Subsets of dependent sets can be either independent or dependent.

Proof on page 394.

Independence of $1, x^2, x^4$ is decided by Theorem 5.24, because it is known that powers $1, x, x^2, x^3, x^4$ form an independent set.

Independence Test: Abstract Vector Space

Theorem 5.23 provides a simple independence / dependence test.[8]

> Form the system of equations
>
> $$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}.$$
>
> Solve for the constants $c_1, \ldots, c_k$.
>
> **Independence** is proved if $c_1, \ldots, c_k$ are all zero.
>
> **Dependence** is proved if a **nonzero** solution $c_1, \ldots, c_k$ exists. This means $c_j \neq 0$ for at least one index j.

Example 5.14 (Independence of Fixed Vectors in $\mathcal{R}^2$)

Test $\mathcal{R}^2$ vectors $\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ for independence.

Details:
The two column vectors are tested for independence by forming the system of equations $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ and solving for the weights c_1, c_2. Then:

$$c_1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Write the vector equation as a homogeneous system $A\vec{c} = \vec{0}$:

$$\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The system has $\mathbf{rref}(A) = I$, details omitted. Then $c_1 = c_2 = 0$, which verifies independence of the two vectors.

Theorem 5.29 page 382 provides a shorter independence test for two vectors: $\vec{v}_1 \neq (\text{constant})\vec{v}_2$.

[8]The *test* is used in publications and mathematical proofs, often without citing the definition of independence. See Definition 5.24 page 381.

Example 5.15 (Independence of Fixed Vectors in $\mathcal{R}^3$)

Test $\mathcal{R}^3$ vectors $\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ for independence.

Details: The two column vectors are tested for independence by forming the system of equations $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ and solving for the weights c_1, c_2:

$$c_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Write the vector equation as a homogeneous system $A\vec{c} = \vec{0}$:

$$\begin{pmatrix} -1 & 2 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The 3×2 coefficient matrix A has reduced row echelon form

$$\mathbf{rref}(A) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The original homogeneous system is then equivalent to $c_1 = 0$, $c_2 = 0$. This proves the two vectors are **independent** by the independence test page 378.

See the Rank Test page 383 and the Determinant Test page 383 for additional column vector independence tests. Determinants are defined only for square matrices, therefore it is an error to use the Determinant Test on non-square Example 5.15. Determinant shortcuts for non-square problems exist [**?**], but they are not discussed here.

Geometric Independence and Dependence for Two Vectors

Two vectors $\vec{v}_1$, $\vec{v}_2$ in $\mathcal{R}^2$ or $\mathcal{R}^3$ are defined to be **geometrically independent** provided neither is the zero vector and one is not a scalar multiple of the other. Graphically, this means $\vec{v}_1$ and $\vec{v}_2$ form the edges of a non-degenerate parallelogram: Figure 13. Free vector arguments use the parallelogram rule for adding and subtracting vectors: Figure 14.

> Two vectors in $\mathcal{R}^2$ or $\mathcal{R}^3$ are geometrically independent if and only if they form the edges of a parallelogram of positive area.

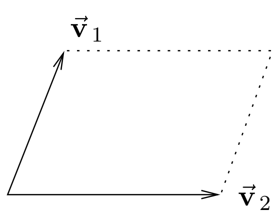

Figure 13. Geometric Independence.
Two nonzero nonparallel vectors $\vec{v}_1$, $\vec{v}_2$ form the edges of a parallelogram. A vector $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2$ lies interior to the parallelogram if and only if the scaling constants satisfy $0 < c_1 < 1$, $0 < c_2 < 1$.

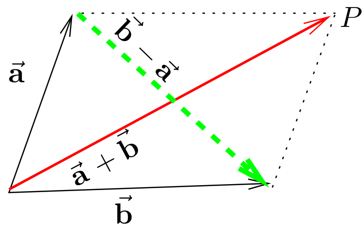

Figure 14. Parallelogram Rule. Given nonzero vectors $\vec{a}$, $\vec{b}$. Red sum vector $\vec{a} + \vec{b}$ has head at vertex P and tail at the joined tails of $\vec{a}$, $\vec{b}$. Green difference vector $\vec{b} - \vec{a}$ connects the head of $\vec{a}$ to the head of $\vec{b}$, according to the **head minus tail rule** on page 297.

Geometric Dependence of Two Vectors

Vectors $\vec{v}_1$, $\vec{v}_2$ in $\mathcal{R}^2$ or $\mathcal{R}^3$ are defined to be **geometrically dependent** provided they are **not geometrically independent**. This means *the two vectors do not form a parallelogram of positive area*: one of $\vec{v}_1$, $\vec{v}_2$ is the zero vector or else $\vec{v}_1$ and $\vec{v}_2$ lie along the same line.

> Two vectors in $\mathcal{R}^2$ or $\mathcal{R}^3$ are geometrically dependent if and only if one is the zero vector or else they are parallel vectors.

Geometric Independence for Three Fixed Vectors

Three vectors in $\mathcal{R}^3$ are said to be **geometrically independent** provided none of them are the zero vector and they form the edges of a non-degenerate parallelepiped of positive volume. Such vectors are called a **triad**. In the special case of all pairs orthogonal (the vectors are 90° apart) they are called an **orthogonal triad**.

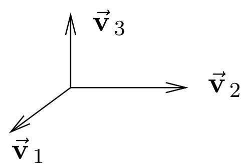

Figure 15. Geometric independence of three vectors. Vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ form the edges of a non-degenerate parallelepiped. A vector $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3$ is located interior to the parallelepiped provided $0 < c_1, c_2, c_3 < 1$.

> Three vectors in $\mathcal{R}^3$ are **geometrically independent** if and only they form the edges of a parallelepiped of positive volume.

Geometric Dependence of Three Fixed Vectors

Given vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, they are **dependent** if and only if they are **not independent**. The three subcases that occur can be analyzed geometrically using Theorem 5.24 page 378:

A nonvoid subset of an independent set is independent.

1. There is a dependent subset of one vector. This vector is the zero vector.

2. There is a dependent subset of two nonzero vectors. Then two of them lie along the same line.

2. There is a dependent subset of three nonzero vectors. Then one of them is in the plane of the other two, because the three cannot form a parallelepiped of positive volume.

> Three vectors in $\mathcal{R}^3$ are **geometrically dependent** if and only if one of them is in the span of the other two. The span is geometrically a point, line or plane.

Theorem 5.25 (Geometric Independence $\equiv$ Algebraic Independence)
The definitions of geometric independence and algebraic independence are equivalent. Proof on page 395.

Independence in an Abstract Vector Space

Linear algebra literature uses a purely algebraic definition of independence, which is equivalent to the **independence test** page 378. The definition and its consequences are recorded here for reference.

Definition 5.24 (Independence in an Abstract Vector Space)
Let $\vec{v}_1$, ..., $\vec{v}_k$ be a finite set of vectors in an abstract vector space V. The set is called **independent** if and only if the vector equation

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$$

has unique solution $c_1 = \cdots = c_k = 0$.

The set of vectors is called **dependent** if and only if the set is not independent. This means that the equation in unknowns c_1, ..., c_k has a solution with at least one constant c_j nonzero.

Theorem 5.26 (Unique Representation)
Let $\vec{v}_1$, ..., $\vec{v}_k$ be independent vectors in an abstract vector space V. If scalars a_1, ..., a_k and b_1, ..., b_k satisfy the relation

$$a_1\vec{v}_1 + \cdots + a_k\vec{v}_k = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$$

then the coefficients must match:

$$\begin{cases} a_1 &= b_1, \\ a_2 &= b_2, \\ \vdots & \\ a_k &= b_k. \end{cases}$$

Proof on page 395.
The result is often used to derive scalar equations from vector equations, e.g., the *Method of Undetermined Coefficients* in differential equations, page ??.

Theorem 5.27 (Zero Vector and Dependent Sets)
An independent set in an abstract vector space V cannot contain the zero vector. Equivalently, a set containing the zero vector is dependent. Proof on page 395

Theorem 5.28 (Linear Combination and Independence)
Let $\vec{v}_1, \ldots, \vec{v}_k$ be given vectors in abstract vector space V. Then:

1. Assume $\vec{v}_1, \ldots, \vec{v}_k$ is an independent set. Suppose $\vec{v}$ from V is not a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$. Then $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ is an independent set.

2. If vector $\vec{v}$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$, then $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ is a dependent set.

Proof on page 395

Theorem 5.29 (Independence of Two Vectors)
Two vectors in an abstract vector space V are independent if and only if neither is the zero vector and one is not a constant multiple of the other. Proof on page 396.

Independence and Dependence Tests for Fixed Vectors

Recorded here are a number of useful algebraic tests to determine independence or dependence of a finite list of fixed vectors.

Rank Test

In the vector space $\mathcal{R}^n$, the key to detection of independence is **zero free variables**, or nullity zero, or equivalently, maximal rank. The test is justified from the formula $\mathbf{nullity}(A) + \mathbf{rank}(A) = k$, where k is the column dimension of A.

Theorem 5.30 (Rank-Nullity Test for Three Vectors)
Let $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ be 3 column vectors in $\mathcal{R}^n$ and let their $n \times 3$ augmented matrix be

$$A = \left\langle \vec{v}_1 | \vec{v}_2 | \vec{v}_3 \right\rangle.$$

The vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ are independent if $\mathbf{rank}(A) = 3$ and dependent if $\mathbf{rank}(A) < 3$. The conditions are equivalent to $\mathbf{nullity}(A) = 0$ and $\mathbf{nullity}(A) > 0$, respectively. Proof on page 396.

Theorem 5.31 (Rank-Nullity Test)
Let $\vec{v}_1, \ldots, \vec{v}_k$ be k column vectors in $\mathcal{R}^n$ and let A be their $n \times k$ augmented matrix. The vectors are independent if $\mathbf{rank}(A) = k$ and dependent if $\mathbf{rank}(A) < k$. The conditions are equivalent to $\mathbf{nullity}(A) = 0$ and $\mathbf{nullity}(A) > 0$, respectively. Proof on page 396.

Determinant Test

In the unusual case when system $A\vec{c} = \vec{0}$ arising in the independence test is square (A is $n \times n$), then $\det(A) = 0$ detects dependence, and $\det(A) \neq 0$ detects independence. The reasoning applies formula $A^{-1} = \mathbf{adj}(A)/\det(A)$, valid exactly when $\det(A) \neq 0$.

Theorem 5.32 (Determinant Test)
Let $\vec{v}_1, \ldots, \vec{v}_n$ be n column vectors in $\mathcal{R}^n$ and let A be the $n \times n$ augmented matrix of these vectors. The vectors are independent if $\det(A) \neq 0$ and dependent if $\det(A) = 0$. Proof on page 396.

Orthogonal Vector Test

In some applications the vectors being tested are known to satisfy **orthogonality conditions**. The dot product conditions for three vectors:

$$\begin{array}{lll} \vec{v}_1 \cdot \vec{v}_1 > 0, & \vec{v}_2 \cdot \vec{v}_2 > 0, & \vec{v}_3 \cdot \vec{v}_3 > 0, \\ \vec{v}_1 \cdot \vec{v}_2 = 0, & \vec{v}_2 \cdot \vec{v}_3 = 0, & \vec{v}_3 \cdot \vec{v}_1 = 0. \end{array} \tag{1}$$

The conditions mean that the vectors are nonzero and pairwise 90° apart. The set of vectors is said to be **pairwise orthogonal**, or briefly, **orthogonal**. The orthogonality conditions for a list of k vectors are written

$$\vec{v}_i \cdot \vec{v}_i > 0, \quad \vec{v}_i \cdot \vec{v}_j = 0, \quad 1 \leq i, j \leq k, \quad i \neq j. \tag{2}$$

Theorem 5.33 (Orthogonal Vector Test)
A set of nonzero pairwise orthogonal vectors $\vec{v}_1, \ldots, \vec{v}_k$ is linearly independent. Proof on page 397.

Independence Tests for Functions

It is not obvious how to solve for $c_1, \ldots, c_k$ in the algebraic independence test page 378, when the vectors $\vec{v}_1, \ldots, \vec{v}_k$ are not fixed vectors. If V is a set of functions, then the methods from linear algebraic equations do not directly apply. This algebraic problem causes development of special tools just for functions, called the **sampling test** and **Wronskian test**. Neither test is an equivalence. Such tests only apply to conclude independence. No results here are equipped to test dependence of a list of functions.

Sampling Test for Functions

Let f_1, f_2, f_3 be three functions defined on a domain D. Let V be the vector space of all functions $\vec{f}$ on D with the usual scalar multiplication and addition rules learned in college algebra.[9] Addressed here is the question of how to test independence and dependence of $\vec{f_1}$, $\vec{f_2}$, $\vec{f_3}$ in V. The vector relation

$$c_1\vec{f_1} + c_2\vec{f_2} + c_3\vec{f_3} = \vec{0}$$

means

$$c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = 0, \quad x \text{ in } D.$$

An idea how to solve for c_1, c_2, c_3 arises by **sampling**, which means 3 relations are obtained by **inventing** 3 values for x, say x_1, x_2, x_3. The equations arising are

$$\begin{array}{ccccccc} c_1 f_1(x_1) & + & c_2 f_2(x_1) & + & c_3 f_3(x_1) & = & 0, \\ c_1 f_1(x_2) & + & c_2 f_2(x_2) & + & c_3 f_3(x_2) & = & 0, \\ c_1 f_1(x_3) & + & c_2 f_2(x_3) & + & c_3 f_3(x_3) & = & 0. \end{array}$$

This system of 3 equations in 3 unknowns can be written in matrix form $A\vec{c} = \vec{0}$, where the coefficient matrix A and vector $\vec{c}$ of unknowns c_1, c_2, c_3 are defined by

$$A = \begin{pmatrix} f_1(x_1) & f_2(x_1) & f_3(x_1) \\ f_1(x_2) & f_2(x_2) & f_3(x_2) \\ f_1(x_3) & f_2(x_3) & f_3(x_3) \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

The matrix A is called the **sampling matrix** for f_1, f_2, f_3 with **samples** x_1, x_2, x_3. Important: you must invent the values for the samples.

The system $A\vec{c} = \vec{0}$ has unique solution $\vec{c} = \vec{0}$, proving $\vec{f_1}$, $\vec{f_2}$, $\vec{f_3}$ independent, provided $\det(A) \neq 0$.

Definition 5.25 (Sampling Matrix)
Let functions $f_1, \ldots, f_k$ be given. Let k samples $x_1, \ldots, x_k$ be given. The **Sampling Matrix** A is defined by:

$$A = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \cdots & \vdots \\ f_1(x_k) & f_2(x_k) & \cdots & f_k(x_k) \end{pmatrix}.$$

Theorem 5.34 (Sampling Test for Functions)
The functions $f_1, \ldots, f_k$ are linearly independent on an x-set D provided there is a sampling matrix A constructed from invented samples $x_1, \ldots, x_k$ in D such that $\det(A) \neq 0$.

The converse is false. An independent list of functions may have $\det(A) = 0$ for a given sampling matrix. Proof on page 397.

[9]Symbol $\vec{f}$ is the vector package for function f. Symbol $f(x)$ is a number, a function value. Symbol f is a graph, equivalently the domain D plus equation $y = f(x)$. Vector $\vec{f}$ is the package of equation $y = f(x)$ and the domain.

Wronskian Test for Functions

The test will be explained first for two functions f_1, f_2. Independence of f_1, f_2, as in the sampling test, is decided by solving for constants c_1, c_2 in the equation

$$c_1 f_1(x) + c_2 f_2(x) = 0, \quad \text{for all } x.$$

J. M. Wronski[10] suggested to solve for the constants by differentiation of this equation, obtaining a pair of equations

$$\begin{array}{rcrcll} c_1 f_1(x) & + & c_2 f_2(x) & = & 0, & \\ c_1 f_1'(x) & + & c_2 f_2'(x) & = & 0, & \text{for all } x. \end{array}$$

This is a system of equations $A\vec{c} = \vec{0}$ with coefficient matrix A and variable list vector $\vec{c}$ given by

$$A = \begin{pmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The **Wronskian Test** is simply $\det(A) \neq 0$ implies $\vec{c} = \vec{0}$, similar to the sampling test:

$$\begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix} \neq 0 \quad \text{for some } x \text{ implies } f_1,\ f_2 \text{ independent.}$$

Interesting about Wronski's idea is that it requires the invention of just one sample x such that the determinant is non-vanishing, in order to establish independence of the two functions.

Definition 5.26 (Wronskian Matrix)
Given functions $f_1, \ldots, f_n$ each differentiable $n-1$ times on an interval $a < x < b$, the **Wronskian determinant** is defined by the relation

$$W(f_1, \ldots, f_n)(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \cdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}.$$

Theorem 5.35 (Wronskian Test)
Let functions $f_1, \ldots, f_n$ be differentiable $n-1$ times on interval $a < x < b$. Assume the Wronskian determinant $W(f_1, \ldots, f_n)(x_0)$ is nonzero for some x_0 in (a, b). Then $f_1, \ldots, f_n$ are independent functions in the vector space V of all functions on (a, b).

The converse is false. Independent functions may have Wronskian determinant identically zero on (a, b).

Proof on page 397.

[10] J. M. Wronski (1776-1853). Born Józef Maria Hoëné in Poland, he resided his final 40 years in France using the name Wronski.

Euler Solution Atom Test

The test originates in linear differential equations. It applies in a variety of situations outside that scope, providing basic intuition about independence of functions.

Definition 5.27 (Euler Solution Atom)
The infinite set of Euler solution atoms is a set of functions on $-\infty < x < \infty$ indexed by three variables a, b, n:

> **Index set**: real a, real $b > 0$, integer $n = 0, 1, 2, \ldots$
> **Distinct functions**: $x^n e^{ax}$, $x^n e^{ax}\cos(bx)$, $x^n e^{ax}\sin(bx)$

A **base atom** is one of $e^{ax}, e^{ax}\cos(bx), e^{ax}\sin(bx)$. An **Euler solution atom** is a base atom times x^n, index set as above.

Theorem 5.36 (Independence of Euler Solution Atoms)
A finite list of distinct Euler solution atoms is independent on any interval E in $-\infty < x < \infty$.

Outline of the proof on page 398. See also Example 5.22, page 391.

Application: Vandermonde Determinant

Choosing the functions in the *sampling test* to be 1, x, x^2 with invented samples x_1, x_2, x_3 gives the sampling matrix

$$V(x_1, x_2, x_3) = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix}.$$

The sampling matrix is called a **Vandermonde matrix**. Using the polynomial basis $f_1(x) = 1$, $f_2(x) = x$, $\ldots$, $f_k(x) = x^{k-1}$ and invented samples x_1, $\ldots$, x_k gives the $k \times k$ Vandermonde matrix

$$V(x_1, \ldots, x_k) = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{k-1} \\ 1 & x_2 & \cdots & x_2^{k-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_k & \cdots & x_k^{k-1} \end{pmatrix}.$$

The most often used Vandermonde determinant identities are

$$\begin{vmatrix} 1 & a \\ 1 & b \end{vmatrix} = b - a,$$

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (c-b)(c-a)(b-a),$$

$$\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix} = (d-c)(d-b)(d-a)(c-b)(c-a)(b-a).$$

Theorem 5.37 (Vandermonde Determinant Identity)
The Vandermonde matrix has a nonzero determinant for distinct samples:

$$\det(V(x_1,\ldots,x_k)) = \prod_{i<j}(x_j - x_i).$$

Proof on page 398.

Examples

Example 5.16 (Vector General Solution)
Find the vector general solution $\vec{\mathbf{u}}$ of $A\vec{\mathbf{u}} = \vec{\mathbf{0}}$, given matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Solution: The solution divides into two distinct sections: $\boxed{\mathbf{1}}$ and $\boxed{\mathbf{2}}$.

$\boxed{\mathbf{1}}$: Find the scalar general solution of the system $A\vec{x} = \vec{0}$.

The toolkit: combination, swap and multiply. Then we use the last frame algorithm. The usual shortcut applies to compute $\mathbf{rref}(A)$. We skip the augmented matrix $\left\langle A|\vec{0}\right\rangle$, knowing that the last column of zeros is unchanged by the toolkit. The details:

$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ First frame.

$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ `combo(1,2,-2)`.

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ `combo(2,1,-2)`. Last frame, this is $\mathbf{rref}(A)$.

$\begin{vmatrix} x_1 & = & 0, \\ x_2 & = & 0, \\ 0 & = & 0. \end{vmatrix}$ Translate to scalar equations.

$\begin{vmatrix} x_1 &=& 0, \\ x_2 &=& 0, \\ x_3 &=& t_1. \end{vmatrix}$	Scalar general solution, obtained from the last frame algorithm: x_1, x_2=lead, x_3=free.

$\boxed{\mathbf{2}}$: Find the vector general solution of the system $A\vec{x} = \vec{0}$.

The plan is to use the answer from $\boxed{\mathbf{1}}$ and partial differentiation to display the vector general solution $\vec{x}$.

$\begin{vmatrix} x_1 &=& 0, \\ x_2 &=& 0, \\ x_3 &=& t_1. \end{vmatrix}$	Scalar general solution, from $\boxed{\mathbf{1}}$.
$\partial_{t_1}\vec{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$	**Strang's special solution** is the partial derivative on symbol t_1. Only one, because of only one invented symbol.
$\vec{x} = t_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$	The vector general solution. It is the sum of terms, an invented symbol times the corresponding **special solution** (partial on that symbol). See also Example 5.19.

Example 5.17 (Independence)

Assume $\vec{v}_1, \vec{v}_2$ are independent vectors in abstract vector space V. Display the details which verify the independence of the vectors $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$.

Solution: The algebraic independence test page 378 will be applied. Form the equation

$$c_1(\vec{v}_1 + 3\vec{v}_2) + c_2(\vec{v}_1 - 2\vec{v}_2) = \vec{0}$$

and somehow solve for c_1, c_2. The plan is to re-write this equation in terms of $\vec{v}_1, \vec{v}_2$, then use the algebraic independence page 378 on vectors $\vec{v}_1, \vec{v}_2$ to obtain scalar equations for c_1, c_2. The equation re-arrangement:

$$(c_1 + c_2)\vec{v}_1 + (3c_1 - 2c_2)\vec{v}_2 = \vec{0}.$$

The independence test applied to a relation $a\vec{v}_1 + b\vec{v}_2 = \vec{0}$ implies scalar equations $a = 0$, $b = 0$. The re-arranged equation has $a = c_1 + c_2$, $b = 3c_1 - 2c_2$. Therefore, independence strips away the vectors from the re-arranged equation, leaving a system of scalar equations in symbols c_1, c_2:

$$\begin{array}{rcrcrl} c_1 &+& c_2 &=& 0, & \quad\text{The equation } a = 0, \\ 3c_1 &-& 2c_2 &=& = 0, & \quad\text{The equation } b = 0. \end{array}$$

These equations have only the zero solution $c_1 = c_2 = 0$, because the coefficient matrix $\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$ is invertible (nonzero determinant). The vectors $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$ are independent by the independence test page 378.

Example 5.18 (Span)

Let $\vec{v}_1, \vec{v}_2$ be two vectors in an abstract vector space V. Define two subspaces

$$S_1 = \mathbf{span}(\vec{v}_1, \vec{v}_2), \quad S_2 = \mathbf{span}(\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2).$$

(a) Display the technical details which show that the two subspaces are equal: $S_1 = S_2$.

(b) Use the result of (a) to prove that independence of $\vec{v}_1, \vec{v}_2$ implies independence of $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$.

Solution:

Details for (a). Sets S_1, S_2 are known to be subspaces of V by the **span theorem** page 301. To show $S_1 = S_2$, we will show each set is a subset of the other, that is, $S_2 \subset S_1$ and $S_1 \subset S_2$.

Show $S_2 \subset S_1$. By definition of **span** page 301, both vectors $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$ belong to the set S_1. Therefore, the span of these two vectors is also in subspace S_1, hence $S_2 \subset S_1$.

Show $S_1 \subset S_2$. Write $\vec{v}_1$ as a linear combination of $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$ in $\boxed{\mathbf{1}}$, $\boxed{\mathbf{2}}$ steps below. This will prove $\vec{v}_1$ belongs to S_2.

$\boxed{\mathbf{1}}$ $5\vec{v}_1 = 2(\vec{v}_1 + 3\vec{v}_2) + 3(\vec{v}_1 - 2\vec{v}_2)$. Eliminate $\vec{v}_2$ with a combination.

$\boxed{\mathbf{2}}$ $\vec{v}_1 = \frac{2}{5}(\vec{v}_1 + 3\vec{v}_2) + \frac{3}{5}(\vec{v}_1 - 2\vec{v}_2)$. Divide by 5.

Similarly, $\vec{v}_2$ belongs to S_2. Therefore, the span of $\vec{v}_1, \vec{v}_2$ belongs to S_2, or $S_1 \subset S_2$, as claimed.

Details for (b). Independence of $\vec{v}_1, \vec{v}_2$ implies $\dim(S_1) = 2$. Therefore, $\dim(S_2) = 2$. If $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$ fail to be independent, then they are dependent and span S_2. Then $\dim(S_2) \le 1$, a contradiction to $\dim(S_2) = 2$. This proves that $\vec{v}_1 + 3\vec{v}_2, \vec{v}_1 - 2\vec{v}_2$ are independent.

Example 5.19 (Independence, Span and Basis)

A 5×5 linear system $A\vec{x} = \vec{0}$ has scalar general solution

$$\begin{aligned} x_1 &= t_1 + 2t_2, \\ x_2 &= t_1, \\ x_3 &= t_2, \\ x_4 &= 4t_2 + t_3, \\ x_5 &= t_3. \end{aligned}$$

Find a basis for the solution space.

Solution: The answer is the set of **Strang's special solutions** obtained by taking partial derivatives on the symbols t_1, t_2, t_3. Details below.

$$\vec{X}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{X}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 4 \\ 0 \end{pmatrix}, \quad \vec{X}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Span. The vector general solution is expressed as the sum $\vec{x} = t_1\vec{X}_1 + t_2\vec{X}_2 + t_3\vec{X}_3$, which implies that the solution space is **span**$(\vec{X}_1, \vec{X}_2, \vec{X}_3)$.

Independence follows from Theorem 5.22, proof on page 393. Let's repeat the proof for the three special solutions $\vec{X}_1, \vec{X}_2, \vec{X}_3$, using the independence test in Theorem 5.23, which is the basis for Definition 5.24, page 381. Form the equation $c_1\vec{X}_1+c_2\vec{X}_2+c_3\vec{X}_3 = \vec{0}$ and solve for c_1, c_2, c_3. The left side of the equation is a vector solution $\vec{x}$ with invented symbols replaced by $t_1 = c_1, t_2 = c_2, t_3 = c_3$. The equation says that $\vec{x} = \vec{0}$, which in scalar form means $x_1 = x_2 = x_3 = x_4 = x_5 = 0$. The scalar general solution has lead variables x_1, x_4 and free variables x_2, x_3, x_5. The free variable equations are:

$$\begin{aligned} x_2 &= t_1, \\ x_3 &= t_2, \\ x_5 &= t_3. \end{aligned}$$

Because $x_2 = x_3 = x_5 = 0$, then $t_1 = t_2 = t_3 = 0$, which implies $c_1 = c_2 = c_3 = 0$. This proves independence of $\vec{X}_1, \vec{X}_2, \vec{X}_3$.

Special Solution Details. Take the partial derivative of the scalar general solution on symbol t_1 to create special solution $\vec{X}_1$. The others are found the same way, by partial derivatives on t_2, t_3. For symbol t_1:

$$\vec{X}_1 = \partial_{t_1}\vec{x} = \begin{pmatrix} \partial_{t_1}x_1 \\ \partial_{t_1}x_2 \\ \partial_{t_1}x_3 \\ \partial_{t_1}x_4 \\ \partial_{t_1}x_5 \end{pmatrix} = \begin{pmatrix} \partial_{t_1}(t_1+2t_2) \\ \partial_{t_1}(t_1) \\ \partial_{t_1}(t_2) \\ \partial_{t_1}(4t_2+t_3) \\ \partial_{t_1}(t_3) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Example 5.20 (Rank Test and Determinant Test)

Apply both the rank test and the determinant test to decide independence or dependence of the vectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \vec{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}.$$

Solution: Answer: The vectors are dependent.

Details for the Rank Test. Form the augmented matrix A of the four vectors and then compute the rank of A. If the rank is 4, then the rank test implies they are independent, otherwise dependent.

$$\begin{aligned} A &= \left\langle \vec{v}_1 | \vec{v}_2 | \vec{v}_3 | \vec{v}_4 \right\rangle \\ &= \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \end{pmatrix}. \end{aligned}$$

How to determine that the rank is not 4? Use rank of A equals the rank of A^T. Equivalently, the row rank equals the column rank. Then a row of zeros implies a dependent set of rows, which implies the row rank is not 4 (the rank is actually 2). Also, columns one and two of A are identical, they are dependent columns, therefore the column rank is not 4.

Details for the Determinant Test. The test uses the square matrix A defined above. The question of independence reduces to testing $|A|$ nonzero. If nonzero, then the columns of A are independent, which implies the four given vectors are independent. Otherwise, $|A| = 0$, which implies the columns of A are dependent, so the given four vectors are dependent.

All depends upon A being square: there is no determinant theory for non-square matrices.

Immediately$|A| = 0$, because A has a row of zeros. Alternatively, $|A| = 0$ because A has duplicate columns. Then the columns of A are dependent, which means dependence of the given four vectors.

Example 5.21 (Sampling Test and Wronskian Test)

Let $V = C(-\infty, \infty)$ and define vectors $\vec{v}_1 = x^2$, $\vec{v}_2 = x^{7/3}$, $\vec{v}_3 = x^5$.[11] Apply the sampling test and the Wronskian test to establish independence of the three vectors in V.

Solution: The vectors are not fixed vectors (column vectors in some $\mathcal{R}^n$), therefore the rank test and determinant test cannot apply. The Euler solution atom test does not apply: the functions are not atoms.

Sampling Test Details. A bad sample choice is $x = 0$, because it will produce a row of zeros, hence a zero determinant, leading to no test. Choose samples $x = 1, 2, 3$ for lack of insight, and then see if it works. The **sample matrix**:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 4 & (\sqrt[3]{2})^7 & 32 \\ 9 & (\sqrt[3]{3})^7 & 243 \end{pmatrix}.$$

Because $|A| \approx 132$ is nonzero, then the given vectors are independent by the sampling test.

Wronskian Test Details. Choose the sample x after finding the Wronskian matrix $W(x)$ for all x. Start with row vector $\begin{pmatrix} x^2, x^{7/3}, x^5 \end{pmatrix}$ and differentiate twice to compute the rows of the Wronskian matrix:

$$W(x) = \begin{pmatrix} x^2 & x^{7/3} & x^5 \\ 2x & \frac{7}{3}x^{4/3} & 5x^4 \\ 2 & \frac{28}{9}x^{1/3} & 20x^3 \end{pmatrix}.$$

The sample $x = 0$ won't work, because $|W(0)|$ has a row of zeros. Choose $x = 1$, then

$$W(1) = \begin{pmatrix} 1 & 1 & 1 \\ 2 & \frac{7}{3} & 5 \\ 2 & \frac{28}{9} & 20 \end{pmatrix}.$$

The determinant $|W(1)| = 8/3$ is nonzero, which implies the three functions are independent by the Wronskian test.

Example 5.22 (Solution Space of a Differential Equation)

A fifth order linear differential equation has general solution

$$y(x) = c_1 + c_2 x + c_3 e^x + c_4 e^{-x} + c_5 e^{2x}.$$

Write the solution space S in vector space $C^5(-\infty, \infty)$ as the span of basis vectors.

[11] Equation $\vec{v}_1 = x^2$ is an abuse of notation which defines vector package $\vec{v}_1$ in V with domain $(-\infty, \infty)$ and equation $y = x^2$. It is used without apology.

Solution: The answer is

$$S = \mathbf{span}\left(1, x, e^x, e^{-x}, e^{2x}\right).$$

Details. A general solution is an expression for all solutions (no solutions skipped) in terms of arbitrary constants, in this case, the constants c_1 to c_5. We think of the constants as the invented symbols $t_1, t_2, \ldots$ in a matrix equation general solution. Then the expected basis vectors should be the partial derivatives on the symbols:

$$\begin{array}{rcl}
\partial_{c_1} y(x) &=& 1, \\
\partial_{c_2} y(x) &=& x, \\
\partial_{c_3} y(x) &=& e^x, \\
\partial_{c_4} y(x) &=& e^{-x}, \\
\partial_{c_5} y(x) &=& e^{2x}.
\end{array}$$

The five vectors so obtained already span the space S. All that remains is to prove they are independent. The easiest method to apply in this case is the Wronskian test.

Independence Details. Let $W(x)$ be the Wronskian of the five solutions above. Then row one is the list $1, x, e^x, e^{-x}, e^{2x}$ and the other four rows are successive derivatives of the first row.

$$W(x) = \left| \begin{array}{ccccr}
1 & x & e^x & e^{-x} & e^{2x} \\
0 & 1 & e^x & -e^{-x} & 2e^{2x} \\
0 & 0 & e^x & e^{-x} & 4e^{2x} \\
0 & 0 & e^x & -e^{-x} & 8e^{2x} \\
0 & 0 & e^x & e^{-x} & 16e^{2x}
\end{array} \right|.$$

The cofactor rule applied twice in succession to column 1 gives

$$W(x) = \left| \begin{array}{ccr}
e^x & e^{-x} & 4e^{2x} \\
e^x & -e^{-x} & 8e^{2x} \\
e^x & e^{-x} & 16e^{2x}
\end{array} \right|.$$

Choose sample $x = 0$ to simplify the work:

$$W(0) = \left| \begin{array}{rrr}
1 & 1 & 4 \\
1 & -1 & 8 \\
1 & 1 & 16
\end{array} \right| = -24.$$

Then the determinant $|W(0)| = -24$ is nonzero, which implies independence of the functions in row one of $W(x)$, by the Wronskian test.

A Faster Independence Test. Generally, the Wronskian test is not used. Instead, apply the Euler solution atom test Theorem 5.36 page 386, which establishes independence without proof details.[12]

The details of the Euler solution atom test are brief: (1) The list $1, x, e^x, e^{-x}, e^{2x}$ is a finite set of distinct Euler solution atoms. (2) The test concludes that the set $1, x, e^x, e^{-x}, e^{2x}$ is independent.

Example 5.23 (Extracting a Basis from a List)

Let V be the vector space of all polynomials. Define subspace

$$S = \mathbf{span}(x + 1, 2x - 1, 3x + 4, x^2).$$

Find a basis for S selected from the list $x + 1, 2x - 1, 3x + 4, x^2$.

[12]The proof of the Euler solution atom test, only outlined but not proved in this textbook, involves determinant evaluations similar to this example. An essential result used in the proof is *subsets of independent sets are independent.*

Solution: One possible answer: $x+1, 2x-1, x^2$.

The vectors $x+1, 2x-1$ are independent, because one is not a scalar multiple of the other (they are lines with slopes $1, 2$); see Theorem 5.29.

The list $x+1, 2x-1, 3x+4$ of three vectors is dependent. In detail, using Theorem 5.28, we first will show $\textbf{span}(x+1, 2x-1) = \textbf{span}(1, x)$, using these two stages:

$\boxed{1}$ $3x = (x+1) + (2x-1)$

$\boxed{2}$ $-3 = -2(x+1) + (2x-1)$

Divide $\boxed{1}$ by 3 and $\boxed{2}$ by -3 to show $\textbf{span}(x+1, 2x-1) = \textbf{span}(1, x)$. Then $3x+4$ is in $\textbf{span}(1, x) = \textbf{span}(x+1, 2x-1)$. Therefore, the list $x+1, 2x-1, 3x+4$ of three vectors is dependent. Skip $3x+4$ and go on to add x^2 to the list. Vector x^2 is not in $\textbf{span}(x+1, 2x-1) = \textbf{span}(1, x)$, because Euler solution atoms $1, x, x^2$ are independent, Theorem 5.36 page 386. The final independent set is $x+1, 2x-1, x^2$, and this is a basis for S. **Important**: a basis is not unique, for instance $1, x, x^2$ is also a basis for S. To extract a basis from the list means the expected answer is the list $x+1, 2x-1, 3x+4, x^2$ with dependent vectors removed. Many correct answers are possible.

Details and Proofs

Proof of Theorem 5.22, Independence of Special Solutions:

1. **To prove**: each solution $\vec{x}$ is a linear combination of $\vec{v_1}, \ldots, \vec{v}_k$. The general solution of $A\vec{x} = \vec{0}$ is written in scalar form by the last frame algorithm page **??**, using invented symbols $t_1, \ldots, t_k$. Special solution $\vec{v}_i = \partial_{t_i}\vec{x}$ $(1 \le i \le k)$ can also be defined as the vector obtained from the scalar general solution with $t_i = 1$ and all other $t_1, \ldots, t_k$ set to zero. The vector general solution is a re-write of the scalar equations in vector form

$$\vec{x} = t_1\vec{v}_1 + \cdots + t_k\vec{v}_k \tag{3}$$

Therefore, each solution is a linear combination of the special solutions.

2. **To prove**: the vectors $\vec{v_1}, \ldots, \vec{v}_k$ are independent. Suppose a given solution $\vec{x}$ can be written in two ways as a linear combination of the special solutions:

$$\vec{x} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k, \quad \vec{x} = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$$

Subtract the two equations and collect on $\vec{v_1}, \ldots, \vec{v}_k$:

$$(a_1 - b_1)\vec{v}_1 + \cdots + (a_k - b_k)\vec{v}_k = \vec{0}$$

Define $c_i = a_i - b_i$, $1 \le i \le k$, then rewrite the preceding equation as

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$$

The left side of this equation is a solution of $A\vec{x} = \vec{0}$ in the form (3) produced by the last frame algorithm. Values $c_1, \ldots, c_k$ are values assigned to the invented symbols $t_1, \ldots, t_k$. Because this solution equals $\vec{0}$, then the corresponding scalar solution $x_1, \ldots, x_n$ of $A\vec{x} = \vec{0}$ is zero: $x_i = 0$ for $1 \le i \le n$. Variables x_i are divided into free variables and lead variables. The free variables in the last frame algorithm are set equal to $t_1, \ldots, t_k$. The lead variables are determined in terms of the free variables. Because all $x_i = 0$, then **all the free variables are zero**: $t_1 = \cdots = t_k = 0$, equivalently $c_1 = \cdots = c_k = 0$.

Equation $c_i = a_i - b_i$ and $c_i = 0$ implies $a_i = b_i$ for $1 \le i \le k$. This proves that a given solution cannot be represented in two different ways: vectors $\vec{v}_1, \ldots, \vec{v}_k$ are independent. ■

Proof of Theorem 5.23, Unique Representation of the Zero Vector: The proof will be given for the characteristic case $k = 3$, because details for general k can be written from this proof, by minor editing of the text.

Assume vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ are independent and $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$. Then $a_1\vec{v}_1 + x_2\vec{v}_2 + a_3\vec{v}_3 = b_1\vec{v}_1 + b_2\vec{v}_2 + b_3\vec{v}_3$ where we define $a_1 = c_1$, $a_2 = c_2$, $a_3 = c_3$ and $b_1 = b_2 = b_3 = 0$. By independence, the coefficients match. By the definition of the symbols, this implies the equations $c_1 = a_1 = b_1 = 0$, $c_2 = a_2 = b_2 = 0$, $c_3 = a_3 = b_3 = 0$. Then $c_1 = c_2 = c_3 = 0$.

Conversely, assume $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$ implies $c_1 = c_2 = c_3 = 0$. If

$$a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 = b_1\vec{v}_1 + b_2\vec{v}_2 + b_3\vec{v}_3,$$

then subtract the right side from the left to obtain

$$(a_1 - b_1)\vec{v}_1 + (a_2 - b_2)\vec{v}_2 + (a_3 - b_3)\vec{v}_3 = \vec{0}.$$

This equation is equivalent to

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$$

where the symbols c_1, c_2, c_3 are defined by $c_1 = a_1 - b_1$, $c_2 = a_2 - b_2$, $c_3 = a_3 - b_3$. The theorem's condition implies that $c_1 = c_2 = c_3 = 0$, which in turn implies $a_1 = b_1$, $a_2 = b_2$, $a_3 = b_3$. ■

Proof of Theorem 5.24, Subsets of Independent Sets are Independent: The idea will be communicated for a set of three independent vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$. Let the subset to be tested consist of the two vectors $\vec{v}_1, \vec{v}_2$. To be applied: the algebraic independence test page 378. Form the vector equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$$

and solve for the constants c_1, c_2. If $c_1 = c_2 = 0$ is the only solution, then $\vec{v}_1, \vec{v}_2$ is a an independent set.

Define $c_3 = 0$. Because $c_3\vec{v}_3 = \vec{0}$, the term $c_3\vec{v}_3$ can be added into the previous vector equation to obtain the new vector equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}.$$

Independence of the three vectors implies $c_1 = c_2 = c_3 = 0$, which in turn implies $c_1 = c_2 = 0$, completing the proof that $\vec{v}_1, \vec{v}_2$ are independent.

The proof for an arbitrary independent set $\vec{v}_1, \ldots, \vec{v}_k$ is similar. By renumbering, we can assume the subset to be tested for independence is $\vec{v}_1, \ldots, \vec{v}_m$ for some index $m \le k$. The proof amounts to adapting the proof for $k = 3$ and $m = 2$, given above. The details are omitted.

Because a single nonzero vector is an independent subset of any list of vectors, then a subset of a dependent set can be independent. If the subset of the dependent set is the whole set, then the subset is dependent. In conclusion, subsets of dependent sets can be either independent or dependent.

Proof of Theorem 5.25: The ideas below for $\mathcal{R}^2$ can be applied to supply details for $\mathcal{R}^3$, the $n = 3$ case omitted.

Assume vectors $\vec{v}_1$, $\vec{v}_2$ are geometrically independent: they are nonzero and nonparallel. To apply the independence test page 378, let's solve for c_1, c_2 in the equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}.$$

Suppose $c_1 \neq 0$. Divide by c_1 to obtain $\vec{v}_1 = -(c_2/c_1)\vec{v}_2$. This equality says $\vec{v}_1$, $\vec{v}_2$ are parallel, so we conclude $c_1 = 0$. Replace $c_1 = 0$, then $0\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$, which implies $c_2\vec{v}_2 = \vec{0}$. Because $\vec{v}_2 \neq \vec{0}$, then $c_2 = 0$. This proves weights $c_1 = c_2 = 0$. By the independence test page 378, vectors $\vec{v}_1, \vec{v}_2$ are algebraically independent.

Assume vectors $\vec{v}_1, \vec{v}_2$ are algebraically independent. To show they are geometrically independent requires: (1) they are nonzero, (2) they are not parallel. If (1) fails, then one of the vectors is zero, say $\vec{v}_1$. The independence test page 378 detects dependence, because $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ holds with $c_1 = 1, c_2 = 0$ (not both weights are zero). Similarly if $\vec{v}_2$ is zero. If (1) holds but (2) fails, then the vectors are nonzero and parallel, meaning $\vec{v}_1 = c\vec{v}_2$ for some scalar c. Let $c_1 = 1, c_2 = -c$ in the independence test page 378 to conclude dependence instead of independence. Therefore, (1) and (2) hold, meaning the vectors are geometrically independent. ■

Proof of Theorem 5.26, Unique Representation Abstract Space:
Assume independence of $\vec{v}_1, \ldots, \vec{v}_k$. Suppose there are two equal linear combinations

$$a_1\vec{v}_1 + \cdots + a_k\vec{v}_k = b_1\vec{v}_1 + \cdots + b_k\vec{v}_k$$

Subtract:

$$(a_1 - b_1)\vec{v}_1 + \cdots + (a_k - b_k)\vec{v}_k = \vec{0}$$

Definition 5.24 page 381 says all the weights are zero: $a_j - b_j = 0$ for $1 \leq j \leq k$. Therefore, the coefficients must match: $a_j = b_j$ for $1 \leq j \leq k$. ■

Proof of Theorem 5.27, Zero Vector Abstract Space: Let $\vec{v}_1, \ldots, \vec{v}_k$ be an independent set in abstract vector space V. Suppose $\vec{0}$ is in the set. Assume $\vec{v}_1 = \vec{0}$ by renumbering the list. Then:

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$$

holds with $c_1 = 1$ and all other weights zero. Applying the independence test page 378 proves the set is *dependent*. ■

Proof of Theorem 5.28, Linear Combination and Independence:
1. Let $\vec{v}_1, \ldots, \vec{v}_k$ be a set of independent vectors in abstract vector space V. Assume $\vec{v}$ is not a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$. Independence test page 378 will be applied to set $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$. Form the equation

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v} = \vec{0}$$

and solve for the coefficients. If $c_{k+1} \neq 0$, then divide by it and solve for vector $\vec{v}$ as a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$, a contradiction. Therefore, $c_{k+1} = 0$. Term $c_{k+1}\vec{v}$ is the zero vector, therefore the equation becomes

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$$

Independence implies $c_1 = \cdots = c_k = 0$. Then all weights are zero, proving independence of $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ by the test page 378.

2. Suppose vector $\vec{v}$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$. Then for some constants $c_1, \ldots, c_k$:

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

Define $c_{k+1} = -1$. Then

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v} = \vec{0}$$

holds for weights $c_1, \ldots, c_{k+1}$ not all zero. Apply the independence test page 378 to prove the set is *dependent*. ■

Proof of Theorem 5.29, Independence Two Vectors Abstract Space: Let $\vec{v}_1, \vec{v}_2$ be two vectors in abstract vector space V.

If they are independent, then Theorem 5.27 implies neither can be the zero vector. If a vector is be a multiple of the other, then $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ holds with either $c_1 = 1$ or $c_2 = 1$ (not both weights zero). Applying the independence test page 378 proves the set is *dependent*, a contradiction. Conclude that neither is a constant multiple of the other.

Assume neither is the zero vector and one is not a constant multiple of the other. Let's apply the independence test page 378. Form the system of equations

$$c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$$

and solve for c_1, c_2. If $c_1 = 0$, then $c_2\vec{v}_2 = \vec{0}$, which implies $c_2 = 0$ because $\vec{v}_2 \neq \vec{0}$. Then $c_1 = c_2 = 0$ and independence is proved by the test on page 378. Otherwise, $c_1 \neq 0$ and division results in

$$\vec{v}_1 = -\frac{c_2}{c_1}\vec{v}_2$$

which implies one vector is a constant multiple of the other, a contradiction. Conclusion: $c_1 = c_2 = 0$ and the two vectors are proved independent by the independence test page 378. ■

Proofs of Theorems 5.30, 5.31, Rank-Nullity Test: The proof will be given for $k = 3$, because a small change in the text of this proof is a proof for general k.

Suppose $\mathbf{rank}(A) = 3$. Then there are 3 leading ones in $\mathbf{rref}(A)$ and zero free variables. Therefore, $A\vec{c} = \vec{0}$ has unique solution $\vec{c} = \vec{0}$.

To be applied: the algebraic independence test page 378. Form the vector equation

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$$

and solve for the constants c_1, c_2, c_3. The vector equation says that a linear combination of the columns of matrix A is the zero vector, or equivalently, $A\vec{c} = \vec{0}$ where $\vec{c}$ has components c_1, c_2, c_3. Therefore, $\mathbf{rank}(A) = 3$ implies $\vec{c} = \vec{0}$, or equivalently, $c_1 = c_2 = c_3 = 0$. This proves that the 3 vectors are linearly independent by the test page 378.

If $\mathbf{rank}(A) < 3$, then there exists at least one free variable. Then the equation $A\vec{c} = \vec{0}$ has at least one nonzero solution $\vec{c}$. This proves that the vectors are dependent by the test page 378. ■

Proof of Theorem 5.32, Determinant Test: The proof details will be done for $n = 3$, because minor edits to this text will give the details for general n.

The algebraic independence test page 378 for vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ in $\mathcal{R}^3$ requires solving the system of linear algebraic equations

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$$

for constants c_1, c_2, c_3. The left side of the equation is a linear combination of the columns of the augmented matrix $A = \left\langle \vec{v}_1 | \vec{v}_2 | \vec{v}_3 \right\rangle$, and therefore the system can be represented as the matrix equation $A\vec{c} = \vec{0}$. If $\det(A) \neq 0$, then A^{-1} exists. Multiply the equation $A\vec{c} = \vec{0}$ by the inverse matrix to give

$$\begin{aligned} A\vec{c} &= \vec{0} \\ A^{-1}A\vec{c} &= A^{-1}\vec{0} \\ I\vec{c} &= A^{-1}\vec{0} \\ \vec{c} &= \vec{0}. \end{aligned}$$

Then $\vec{c} = \vec{0}$, or equivalently, $c_1 = c_2 = c_3 = 0$. The vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ are proved independent by the independence test page 378.

Conversely, if the vectors are independent and $A = \left\langle \vec{v}_1 | \vec{v}_2 | \vec{v}_3 \right\rangle$ is the augmented matrix of these vectors, then the system $A\vec{c} = \vec{0}$ has unique solution $\vec{c} = \vec{0}$ by the independence test page 378. The unique solution case for a homogeneous system $A\vec{c} = \vec{0}$ means no free variables or $\mathbf{rref}(A) = I$. Then A has a inverse. Because A^{-1} exists, then $\det(A) \neq 0$. ■

Proof of Theorem 5.33, Orthogonal Vector Test: The proof will be given for $k = 3$, because the details are easily supplied for k vectors, by editing the text in the proof. To be applied: the algebraic independence test page 378. Form the system of equations

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$$

and solve for the constants c_1, c_2, c_3. Constant c_1 is isolated by taking the dot product of the above equation with vector $\vec{v}_1$, to obtain the scalar equation

$$c_1\vec{v}_1 \cdot \vec{v}_1 + c_2\vec{v}_1 \cdot \vec{v}_2 + c_3\vec{v}_1 \cdot \vec{v}_3 = \vec{v}_1 \cdot \vec{0}.$$

The orthogonality relations $\vec{v}_1 \cdot \vec{v}_2 = 0$, $\vec{v}_2 \cdot \vec{v}_3 = 0$, $\vec{v}_3 \cdot \vec{v}_1 = 0$ reduce the scalar equation to

$$c_1\vec{v}_1 \cdot \vec{v}_1 + c_2(0) + c_3(0) = 0.$$

Because $\vec{v}_1 \cdot \vec{v}_1 > 0$, then $c_1 = 0$. Symmetrically, vector $\vec{v}_2$ replacing $\vec{v}_1$, the scalar equation becomes

$$c_1(0) + c_2\vec{v}_2 \cdot \vec{v}_2 + c_3(0) = 0.$$

Again, $c_2 = 0$. The argument for $c_3 = 0$ is similar. The conclusion: $c_1 = c_2 = c_3 = 0$. The three vectors are proved independent. ■

Proof of Theorem 5.34, Sampling Test: Let A be the sampling matrix of Definition 5.25. Let vector $\vec{c}$ have components $c_1, \ldots, c_k$. The algebraic independence test page 378 will be applied. Form the vector equation

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0},$$

to be solved for $c_1, \ldots, c_k$. Substitute samples $x_1, \ldots, x_k$ into the vector equation and re-write as $A\vec{c} = \vec{0}$. Because $\det(A) \neq 0$, then equation $A\vec{c} = \vec{0}$ has unique solution $\vec{c} = \vec{0}$. Then all the weights are zero, proving that vectors $\vec{v}_1, \ldots, \vec{v}_k$ are independent. ■

Proof of Theorem 5.35, Wronskian Test: To be applied: the algebraic independence test page 378. Form the equation

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0, \quad \text{for all } x,$$

and solve for the constants $c_1, \ldots, c_n$. The functions are proved independent provided all the constants are zero. The idea of the proof, attributed to Wronski, is to differentiate the above equation $n-1$ times, then substitute $x = x_0$ to obtain a homogeneous $n \times n$ system $A\vec{c} = \vec{0}$ for the components $c_1, \ldots, c_n$ of the vector $\vec{c}$. Because $|A| = W(f_1, \ldots, f_n)(x_0) \neq 0$, the inverse matrix $A^{-1} = \textbf{adj}(A)/|A|$ exists. Multiply $A\vec{c} = \vec{0}$ on the left by A^{-1} to obtain $\vec{c} = \vec{0}$, completing the proof.

Proof of Theorem 5.36, Euler Solution Atom Test: An outline of the proof will be given, the excuse being that the details are long and uninteresting.[13] Unpleasantness includes complex numbers, real and imaginary parts of functions and the use of several support theorems.

$\boxed{1}$ The powers $1, x, \ldots, x^k$ are independent: Wronskian test Theorem 5.35.

$\boxed{2}$ Exponential e^x is independent of the powers $1, x, \ldots, x^k$. An easy argument uses Maclaurin series for the exponential. The same is true for e^{ax} with $a \neq 0$. Value a can be complex.

$\boxed{3}$ A list of distinct exponentials $e^{a_i x}$, $i = 1, \ldots, k$ with nonzero exponents is linearly independent. Details use the Wronskian test Theorem 5.35, Vandermonde matrices and determinants Theorem 5.37. Values a_i are allowed complex.

$\boxed{4}$ Powers $1, x, \ldots, x^k$ times e^{ax} $(a \neq 0)$ are independent. The result uses the algebraic independence test page 378 and $\boxed{1}$. Symbol a is allowed complex.

$\boxed{5}$ Powers $1, x, \ldots, x^p$ times a list of distinct complex exponentials $e^{a_i x}$, $i = 1, \ldots, q$ makes a list of pq distinct functions. This list of functions is independent. The details use the algebraic independence test page 378, double mathematical induction on p, q and $\boxed{1}$-$\boxed{4}$.

$\boxed{6}$ Restrict the values a_i in $\boxed{5}$ to be of the form $A + iB$ with $B > 0$. The real and imaginary parts of the list of functions in $\boxed{5}$ makes a set of $2pq$ distinct functions, all of which are Euler solution atoms. The set is independent.

The proof concludes by arguing that any finite set of distinct Euler solution atoms is a subset of an independent set described in $\boxed{6}$. Because *subsets of independent sets are independent*, Theorem 5.24, the proof ends. ■

Proof of Theorem 5.37, Vandermonde Determinant Identity: Let's prove the identity for the case $k = 3$, which simplifies notation. Assume distinct samples x_1, x_2, x_3. To be proved:

$$\det(V(x_1, x_2, x_3)) = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1).$$

The proof uses a recursion:

$$\det(V(x_1, x_2, x_3)) = \det(V(x_2, x_3))(x_3 - x_1)(x_2 - x_1).$$

Expansion of $\det(V(x_2, x_3)) = \begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix} = x_3 - x_2$ by Sarrus' Rule gives the claimed $n = 3$ identity:

$$\det(V(x_1, x_2, x_3)) = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1).$$

Recursion proof. Define matrix $A = V(x, x_2, x_3)$ (x_1 replaced by x). Cofactor expansion along row one of $\det(A)$ gives a quadratic in variable x:

$$\det(A) = (1)\,\textbf{cof}(A, 1, 1) + (x)\,\textbf{cof}(A, 1, 2) + (x^2)\,\textbf{cof}(A, 1, 3).$$

[13] Writing details for this is not preferred to eating shattered glass.

Because a determinant with duplicate rows has zero value, then quadratic equation $\det(A) = 0$ has roots $x = x_2$ and $x = x_3$. The factor and root theorems of college algebra apply: for some constant c,

$$\det(A) = c(x_3 - x)(x_2 - x).$$

Constant c is the coefficient of x^2 in $\det(A)$, therefore

$$c = \textbf{cof}(A, 1, 3) = (-1)^{1+3}\,\textbf{minor}(A, 1, 3) = \det(V(x_2, x_3)).$$

Then

$$\det(A) = \det(V(x_2, x_3))(x_3 - x)(x_2 - x).$$

Upon substitution of $x = x_1$, this equation becomes the claimed recursion

$$\det(V(x_1, x_2, x_3)) = \det(V(x_2, x_3))(x_3 - x_1)(x_2 - x_1).$$

Mathematical Induction. The $k \times k$ case first proves by cofactor expansion the recursion

$$\det(V(x_1, x_2, \ldots, x_k)) = \det(V(x_2, \ldots, x_k)\prod_{j=2}^{k}(x_j - x_1). \tag{4}$$

Identity (4) provides the induction step used to prove Theorem 5.37 by induction. To understand the derivation of identity (4), which also requires mathematical induction, experiment with special case $k = 4$:

$$\det(V(x_1, x_2, x_3, x_4)) = \det(V(x_2, x_3, x_4))\prod_{j=2}^{4}(x_j - x_1).$$

Exercises 5.4

Scalar and Vector General Solution

Given the scalar general solution of $A\vec{x} = \vec{0}$, find the vector general solution

$$\vec{x} = t_1\vec{u}_1 + t_2\vec{u}_2 + \cdots$$

where symbols t_1, t_2, ... denote arbitrary constants and $\vec{u}_1$, $\vec{u}_2$, ... are fixed vectors.

1. $x_1 = 2t_1$, $x_2 = t_1 - t_2$, $x_3 = t_2$

2. $x_1 = t_1 + 3t_2$, $x_2 = t_1$, $x_3 = 4t_2$, $x_4 = t_2$

3. $x_1 = t_1$, $x_2 = t_2$, $x_3 = 2t_1 + 3t_2$

4. $x_1 = 2t_1 + 3t_2 + t_3$, $x_2 = t_1$, $x_3 = t_2$, $x_4 = t_3$

Vector General Solution

Find the vector general solution $\vec{x}$ of $A\vec{x} = \vec{0}$.

5. $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

6. $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

7. $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

8. $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

9. $A = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 2 & -1 & 0 \end{pmatrix}$

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

Dimension

11. Give four examples in $\mathcal{R}^3$ of $S = \textbf{span}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ (3 vectors required) which have respectively dimensions $0, 1, 2, 3$.

12. Give an example in $\mathcal{R}^3$ of 2-dimensional subspaces S_1, S_2 with only the zero vector in common.

13. Let $S = \textbf{span}(\vec{v}_1, \vec{v}_2)$ in abstract vector space V. Explain why $\dim(S) \le 2$.

14. Let $S = \textbf{span}(\vec{v}_1, \dots, \vec{v}_k)$ in abstract vector space V. Explain why $\dim(S) \le k$.

15. Let S be a subspace of $\mathcal{R}^3$ with basis $\vec{v}_1, \vec{v}_2$. Define $\vec{v}_3$ to be the **cross product** of $\vec{v}_1, \vec{v}_2$. What is $\dim(\textbf{span}(\vec{v}_2, \vec{v}_3))$?

16. Let S_1, S_2 be subspaces of $\mathcal{R}^4$ such that $\dim(S_1) = \dim(S_2) = 2$. Assume S_1, S_2 have only the zero vector in common. Prove or give a counter-example: the span of the union of S_1, S_2 equals $\mathcal{R}^4$.

Independence in Abstract Spaces

17. Assume linear combinations of vectors $\vec{v}_1$, $\vec{v}_2$ are uniquely determined, that is, $a_1\vec{v}_1 + a_2\vec{v}_2 = b_1\vec{v}_1 + b_2\vec{v}_2$ implies $a_1 = b_1$, $a_2 = b_2$. **Prove** this result: If $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$, then $c_1 = c_2 = 0$.

18. Assume the zero linear combination of vectors $\vec{v}_1$, $\vec{v}_2$ is uniquely determined, that is, $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ implies $c_1 = c_2 = 0$. **Prove** this result: If $a_1\vec{v}_1 + a_2\vec{v}_2 = b_1\vec{v}_1 + b_2\vec{v}_2$, then $a_1 = b_1$, $a_2 = b_2$.

19. Prove that two **nonzero** vectors $\vec{v}_1$, $\vec{v}_2$ in an abstract vector space V are independent if and only if each of $\vec{v}_1$, $\vec{v}_2$ is not a constant multiple of the other.

20. Let $\vec{v}_1$ be a vector in an abstract vector space V. Prove that the one-element set $\vec{v}_1$ is independent if and only if $\vec{v}_1$ is not the zero vector.

21. Let V be an abstract vector space and assume $\vec{v}_1$, $\vec{v}_2$ are independent vectors in V. Define $\vec{u}_1 = \vec{v}_1 + \vec{v}_2$, $\vec{u}_2 = \vec{v}_1 + 2\vec{v}_2$. Prove that $\vec{u}_1$, $\vec{u}_2$ are independent in V.
Advice: Fixed vectors not assumed! Bursting the vector packages is impossible, there are no components.

22. Let V be an abstract vector space and assume $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ are independent vectors in V. Define $\vec{u}_1 = \vec{v}_1 + \vec{v}_2$, $\vec{u}_2 = \vec{v}_1 + 4\vec{v}_2$, $\vec{u}_3 = \vec{v}_3 - \vec{v}_1$. Prove that $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ are independent in V.

23. Let S be a finite set of independent vectors in an abstract vector space V. Prove that none of the vectors can be the zero vector.

24. Let S be a finite set of independent vectors in an abstract vector space V. Prove that no vector in the list can be a linear combination of the other vectors.

The Spaces $\mathcal{R}^n$

25. (Scalar Multiply) Let $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ have components measured in centimeters. Report constants c_1, c_2, c_3 for re-scaled data $c_1\vec{x}$, $c_2\vec{x}$, $c_3\vec{x}$ in units of kilometers, meters and millimeters.

26. (Matrix Multiply) Let $\vec{u} = \left(x_1, x_2, x_3, p_1, p_2, p_3 \right)^T$ have position x-units in kilometers and momentum p-units in kilogram-centimeters per millisecond. Determine a matrix M such that the vector $\vec{y} = M\vec{u}$ has SI units of meters and kilogram-meters per second.

27. Let $\vec{v}_1$, $\vec{v}_2$ be two independent vectors in $\mathcal{R}^n$. Assume $c_1\vec{v}_1 + c_2\vec{v}_2$ lies strictly interior to the parallelogram determined by $\vec{v}_1$, $\vec{v}_2$. Give geometric details explaining why $0 < c_1 < 1$ and $0 < c_2 < 1$.

28. Prove the 4 scalar multiply toolkit properties for fixed vectors in $\mathcal{R}^3$.

29. Define

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, -\vec{v} = \begin{pmatrix} -v_1 \\ -v_2 \\ -v_3 \end{pmatrix}.$$

Prove the 4 addition toolkit properties for fixed vectors in $\mathcal{R}^3$.

30. Use the 8 property toolkit in $\mathcal{R}^3$ to prove that zero times a vector is the zero vector.

31. Let A be an invertible 3×3 matrix. Let $\vec{v}_1, \vec{v}_2, \vec{v}_3$ be a basis for $\mathcal{R}^3$. Prove that $A\vec{v}_1, A\vec{v}_2, A\vec{v}_3$ is a basis for $\mathcal{R}^3$.

32. Let A be an invertible 3×3 matrix. Let $\vec{v}_1, \vec{v}_2, \vec{v}_3$ be dependent in $\mathcal{R}^3$. Prove that $A\vec{v}_1, A\vec{v}_2, A\vec{v}_3$ is a dependent set in $\mathcal{R}^3$.

Digital Photographs

Let V be the vector space of all 2×3 matrices. A matrix in V is a 6-pixel digital photo, a sub-section of a larger photo.

Let $B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, ..., $B_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Each B_j **lights up** one pixel in the 2×3 sub-photo.

33. Prove that $B_1, \ldots, B_6$ are independent and span V: they are a **basis** for V.

34. Let $A = 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 4\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Assume a black and white image and 0 means black. Describe photo A, from the checkerboard analogy.

Digital RGB Photos

Define red, green and blue monochrome matrices R, G, B by

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 5 & 8 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 5 \end{pmatrix}.$$

35. Define base $x = 16$. Compute $A = R + xG + x^2B$.

36. Define base $x = 32$. Compute $A = R + xG + x^2B$.

Polynomial Spaces

Let V be the vector space of all cubic or less polynomials $p(x) = c_0 + c_1x + c_2x^2 + c_3x^3$.

37. Find a subspace S of V, $\dim(S) = 2$, which contains the vector $1 + x$.

38. Let S be the subset of V spanned by x, x^2 and x^3. Prove that S is a subspace of V which does not contain the polynomial $1 + x$.

39. Define set S by the conditions $p(0) = 0, p(1) = 0$. Find a basis for S.

40. Define set S by the condition $p(0) = \int_0^1 p(x)dx$. Find a basis for S.

The Space $C(E)$

Define $\vec{f}$ to be the vector package with domain $E = \{x : -2 \le x \le 2\}$ and equation $y = |x|$. Similarly, $\vec{g}$ is defined by equation $y = x$.

41. Show independence of $\vec{f}, \vec{g}$.

42. Find the dimension of $\mathbf{span}(\vec{f}, \vec{g})$.

43. Let $h(x) = 0$ on $-1 \le x \le 0$, $h(x) = -x$ on $0 \le x \le 1$. Show that $\vec{h}$ is in $C(E)$.

44. Let $h(x) = -1$ on $-2 \le x \le 0$, $h(x) = 1$ on $0 \le x \le 2$. Show that $\vec{h}$ is not in $C(E)$.

45. Let $h(x) = 0$ on $-2 \le x \le 0$, $h(x) = -x$ on $0 \le x \le 2$. Show that $\vec{h}$ is in $\mathbf{span}(\vec{f}, \vec{g})$.

46. Let $h(x) = \tan(\pi x/2)$ on $-2 < x < 2$, $h(2) = h(-2) = 0$. Explain why $\vec{h}$ is not in $C(E)$

The Space $C^1(E)$

Define $\vec{f}$ to be the vector package with domain $E = \{x : -1 \le x \le 1\}$ and equation $y = x|x|$. Similarly, $\vec{g}$ is defined by equation $y = x^2$.

47. Verify that $\vec{f}$ is in $C^1(E)$, but its derivative is not.

48. Show that $\vec{f}, \vec{g}$ are independent in $C^1(E)$.

The Space $C^k(E)$

49. Compute the first three derivatives of $y(x) = e^{-x^2}$ at $x = 0$.

50. Justify that $y(x) = e^{-x^2}$ belongs to $C^k(0,1)$ for all $k \ge 1$.

51. Prove that the span of a finite list of distinct Euler solution atoms (page 386) is a subspace of $C^k(E)$ for any interval E.

52. Prove that $y(x) = |x|$ is in $C^k(0,1)$ but not in $C^1(-1,1)$.

Solution Space

A differential equations solver finds general solution $y = c_1 + c_2 x + c_3 e^x + c_4 e^{-x}$. Use vector space $V = C^4(E)$ where E is the whole real line.

53. Write the solution set S as the span of four vectors in V.

54. Find a basis for the solution space S of the differential equation. Verify independence using the sampling test or Wronskian test.

55. Find a differential equation $y'' + a_1 y' + a_0 y = 0$ which has solution $y = c_1 + c_2 x$.

56. Find a differential equation $y'''' + a_3 y''' + a_2 y'' + a_1 y' + a_0 y = 0$ which has solution $y = c_1 + c_2 x + c_3 e^x + c_4 e^{-x}$.

Algebraic Independence Test for Two Vectors

Solve for c_1, c_2 in the independence test for two vectors, showing all details.

57. $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

58. $\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

Dependence of two vectors

Solve for c_1, c_2 not both zero in the independence test for two vectors, showing all details for dependency of the two vectors.

59. $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$

60. $\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix}$

Independence Test for Three Vectors

Solve for the constants c_1, c_2, c_3 in the relation $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$. Report dependent of independent vectors. If dependent, then display a dependency relation.

61. $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$

62. $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Independence in an Abstract Vector Space

In vector space V, report independence or a dependency relation for the given vectors.

63. Space $V = C(-\infty, \infty)$, $\vec{v}_1 = 1 + x$, $\vec{v}_2 = 2 + x$, $\vec{v}_3 = 3 + x^2$.

64. Space $V = C(-\infty, \infty)$, $\vec{v}_1 = x^{3/5}$, $\vec{v}_2 = x^2$, $\vec{v}_3 = 2x^2 + 3x^{3/5}$

65. Space V is all 3×3 matrices. Let $\vec{v}_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 2 & 5 & 0 \\ 0 & 2 & 5 \\ 0 & 3 & 5 \end{pmatrix}$.

66. Space V is all 2×2 matrices. Let $\vec{v}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix}$.

Rank Test

Compute the rank of the augmented matrix to determine independence or dependence of the given vectors.

67. $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$

68. $\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$

Determinant Test

Evaluate the determinant of the augmented matrix to determine independence or dependence of the given vectors.

69. $\begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 0 \end{pmatrix}$

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

Sampling Test for Functions

Invent samples to verify independence.

71. $\cosh(x), \sinh(x)$

72. $x^{7/3}, x\sin(x)$

73. $1, x, \sin(x)$

74. $1, \cos^2(x), \sin(x)$

Sampling Test and Dependence

For three functions f_1, f_2, f_3 to be dependent, constants c_1, c_2, c_3 must be found such that

$$c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = 0.$$

The trick is that c_1, c_2, c_3 are not all zero and the relation holds **for all** x. The sampling test method can discover the constants, but it is **unable to prove dependence**!

75. Functions $1, x, 1+x$ are dependent. Insert $x = 1, 2, -1$ and solve for c_1, c_2, c_3, to discover a dependency relation.

76. Functions $1, \cos^2(x), \sin^2(x)$ are dependent. Cleverly choose 3 values of x, insert them, then solve for c_1, c_2, c_3, to discover a dependency relation.

Vandermonde Determinant

77. Let $V = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix}$. Verify by direct computation the formula

$$|V| = x_2 - x_1.$$

78. Let $V = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix}$. Verify by direct computation the formula

$$|V| = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1).$$

Wronskian Test for Functions

Apply the Wronskian Test to verify independence.

79. $\cos(x), \sin(x)$.

80. $\cos(x), \sin(x), \sin(2x)$.

81. $x, x^{5/3}$.

82. $\cosh(x), \sinh(x)$.

Wronskian Test: Theory

83. The functions x^2 and $x|x|$ are continuously differentiable and have zero Wronskian. Verify that they **fail to be dependent** on $-1 < x < 1$.

84. The Wronskian Test can verify the independence of the powers $1, x, \ldots, x^k$. Show the determinant details.

Extracting a Basis

Given a list of vectors in space $V = \mathcal{R}^4$, extract a largest independent subset.

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

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

Extracting a Basis

Given a list of vectors in space $V = C(-\infty, \infty)$, extract a largest independent subset.

87. $x, x\cos^2(x), x\sin^2(x), e^x, x + e^x$

88. $1, 2 + x, \frac{x}{1+x^2}, \frac{x^2}{1+x^2}$

Euler Solution Atom

Identify the Euler solution atoms in the given list. Strictly apply the definition: e^x is an atom but $2e^x$ is not.

89. $1, 2 + x, e^{2.15x}, e^{x^2}, \frac{x}{1+x^2}$

90. $2, x^3, e^{x/\pi}, e^{2x+1}, \ln|1 + x|$

Euler Solution Atom Test

Establish independence of set S_1.
Suggestion: First establish an identity $\textbf{span}(S_1) = \textbf{span}(S_2)$, where S_2 is an invented list of distinct atoms. The Test implies S_2 is independent. Extract a largest independent subset of S_1, using independence of S_2.

91. Set S_1 is the list $2, 1 + x^2, 4 + 5e^x, \pi e^{2x+\pi}, 10x\cos(x)$.

92. Set S_1 is the list $1 + x^2, 1 - x^2, 2\cos(3x), \cos(3x) + \sin(3x)$.

5.5 Basis, Dimension and Rank

The topics of basis, dimension and rank apply to the study of Euclidean spaces, continuous function spaces, spaces of differentiable functions and general abstract vector spaces.

Definition 5.28 (Basis)
A **basis** for a vector space V is defined to be an independent set of vectors such that each vector in V is a linear combination of finitely many vectors in the basis. The independent vectors are said to **span** V, with notation

$$V = \mathbf{span}(\text{the set of basis vectors}).$$

If the set of independent vectors is finite, then V is called **finite dimensional**. An important example is $\mathcal{R}^n$. Otherwise, V is said to be **infinite dimensional**. A Fourier series example: the space V spanned by $\sin(nx)$ on $-\pi \le x \le \pi$, $n = 1, 2, 3, \ldots$ is infinite dimensional.

Theorem 5.38 (Size of a Basis)
If vector space V has two bases $\vec{v}_1, \ldots, \vec{v}_p$ and $\vec{u}_1, \ldots, \vec{u}_q$, then $p = q$. Proof on page 422.

Definition 5.29 (Dimension)
The **dimension** of a finite-dimensional vector space V is defined to be the number of vectors in a basis.
Because of Theorem 5.38, the term *dimension* is well-defined.

Theorem 5.39 (Basis of a Finite-Dimensional Vector Space)
Let V be an n-dimensional vector space and $L = \{\vec{v}_1, \ldots, \vec{v}_p\}$ a list of vectors in V, not assumed linearly independent. Then:

1. If $p = n$ and L is an independent set, then L is a basis for V.

2. If $p = n$ and $\mathbf{span}(L) = V$, then L is a basis for V.

3. Always V has a basis containing a given independent subset of L.

4. If $\mathbf{span}(L) = V$, then L contains a basis for V.

Proof on page 422.

Euclidean Spaces

The space $\mathcal{R}^n$ has a **standard basis** consisting of the columns of the $n \times n$ identity matrix:

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

The determinant test implies they are independent. They span $\mathcal{R}^n$ due to the formula

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \quad \cdots \quad + c_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Definition 5.29 implies the columns of the identity matrix form a basis of $\mathcal{R}^n$ of dimension n.

Theorem 5.40 (Basis and Dimension in $\mathcal{R}^n$)
Any basis of $\mathcal{R}^n$ has exactly n independent vectors. Further, any list of $n+1$ or more vectors in $\mathcal{R}^n$ is dependent.
Proof on page 423.

Polynomial Spaces

The vector space of all polynomials $p(x) = p_0 + p_1 x + p_2 x^2$ has dimension 3, justified by producing a basis 1, x, x^2. Formally, the basis elements are obtained from the expression $p(x)$ by partial differentiation on the symbols p_0, p_1, p_2.

Illustration. The subspace $S = \textbf{span}(1-x, 1+x, x)$ is the set of combinations $c_1(1-x) + c_2(1+x) + c_3 x$. Partial differentiation on symbols c_1, c_2, c_3 produces the list of vectors $1-x, 1+x, x$. While they span S, they fail to be independent. Extract a **largest independent subset** from this list to find a basis for S, for example $1-x, 1+x$. Basis size 2 verifies that S has **dimension** 2: see Theorem 5.38 and Definition 5.29.

Differential Equations

The equation $y'' + y = 0$ has general solution $y = c_1 \cos x + c_2 \sin x$. Therefore, the formal partial derivatives ∂_{c_1}, ∂_{c_2} applied to the general solution y give a basis $\cos x$, $\sin x$. The solution space of $y'' + y = 0$ has dimension 2.

Similarly, $y''' = 0$ has a solution basis 1, x, x^2 and therefore its solution space has dimension 3. Generally, an nth order linear homogeneous scalar differential equation has solution space V of dimension n, and an $n \times n$ linear homogeneous system $\frac{d}{dx}\vec{y} = A\vec{y}$ has solution space V of dimension n. There is a general procedure for finding a basis for a differential equation:

> Let a linear differential equation have general solution expressed in terms of arbitrary constants c_1, c_2, ..., then a basis is found by taking the partial derivatives ∂_{c_1}, ∂_{c_2}, ...

Largest Subset of Independent Fixed Vectors

Let vectors $\vec{v}_1, \ldots, \vec{v}_k$ be given in $\mathcal{R}^n$. Then the subset

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

of $\mathcal{R}^n$ consisting of all linear combinations $\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ is a subspace of $\mathcal{R}^n$ by Theorem 5.5. The subset S is identical to the set of all linear combinations of the columns of the augmented matrix A of $\vec{v}_1, \ldots, \vec{v}_k$.

Because matrix multiply is a linear combination of columns, that is,

$$A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = c_1\vec{v}_1 + \quad \cdots \quad + c_k\vec{v}_k,$$

then S is also equals the **image** of the matrix, $S = \mathbf{Image}(A)$.

Definition 5.30 (Image of a Matrix)

$$\mathbf{Image}(A) = \{A\vec{c} \;:\; \text{vector } \vec{c} \text{ arbitrary}\}.$$

Discussed here are efficient methods for finding a basis for any subspace S given as the span of a finite list L of vectors: $S = \mathbf{span}(L)$. The methods apply in particular when the list L consists of the columns of a matrix. Equivalently, the methods find a **largest subset of independent vectors** L_1 from the vectors in set L. This largest subset L_1 is independent and spans S, which makes it a basis for S.

Iterative Method for a Largest Independent Subset

A largest independent subset of vectors $\vec{v}_1, \ldots, \vec{v}_k$ in an abstract vector space V is identified as $\vec{v}_{i_1}, \ldots, \vec{v}_{i_p}$ for some distinct subscripts $i_1 < \cdots < i_p$. Described here is how to find such subscripts. A set containing only the zero vector is dependent, therefore let's assume at least one nonzero vector is listed. Let i_1 be the first subscript such that $\vec{v}_{i_1} \neq \vec{0}$. Define i_2 to be the *first* subscript greater than i_1 such that $\vec{v}_{i_2}$ is not a scalar multiple of $\vec{v}_{i_1}$. The process terminates if there is no such $i_2 > i_1$. Otherwise, proceed in a similar way to define i_3, i_4, ..., i_p. At each stage q we let $S = \{\vec{v}_{i_1}, \ldots, \vec{v}_{i_q}\}$ and select another vector $\vec{v}_{i_{q+1}}$ from $\vec{v}_1, \ldots, \vec{v}_k$ which is not in $\mathbf{span}(S)$. Then

$$\dim(\mathbf{span}(S)) < \dim(\mathbf{span}(S \cup \{\vec{v}_{i_{q+1}}\})).$$

Why does it work? Because each vector added which increases the dimension cannot be a linear combination of the preceding vectors, in short, the list of vectors at each stage is independent. See Example 5.24.

Pivot Theorem Method

Definition 5.31 (Pivot Column of Matrix A)
A column j of A is called a **pivot column** provided $\mathbf{rref}(A)$ has a leading one in column j. The leading ones in $\mathbf{rref}(A)$ belong to consecutive initial columns of the identity matrix I; the **matching columns in** A are the pivot columns of A.

Theorem 5.41 (Pivot Theorem: Independent Columns of A)

1. The pivot columns of a matrix A are linearly independent.

2. A non-pivot column is a linear combination of the pivot columns.

Proof on page 423.

Example 5.24 (Largest Independent Subset)
Find a largest independent subset from the five vectors

$$\vec{v}_1 = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \vec{v}_2 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \vec{v}_3 = \begin{pmatrix}1\\1\\0\\1\end{pmatrix}, \vec{v}_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \vec{v}_5 = \begin{pmatrix}1\\1\\0\\2\end{pmatrix}.$$

Solution:

The **Iterative Method** applies. A visual inspection shows that we should skip the zero vector $\vec{v}_1$ and add $\vec{v}_2, \vec{v}_3$ to the proposed largest independent set. Here, we use the fact that two nonzero vectors are independent if one is not a scalar multiple of the other. Because $\vec{v}_4 = \vec{v}_2 + \vec{v}_3$, we also skip $\vec{v}_4$. Formally, this dependence relation can be computed from toolkit steps on augmented matrix $B = \left\langle \vec{v}_2|\vec{v}_3|\vec{v}_4 \right\rangle$. Similarly, $\vec{v}_5 = 2\vec{v}_4 + \vec{v}_2$, causing a skip of $\vec{v}_5$. A largest independent subset is $\vec{v}_2, \vec{v}_3$.

The **Pivot Theorem** applies. This method has a computer implementation. Form the augmented matrix A of the five vectors and then compute $\mathbf{rref}(A)$.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{rref}(A) = \begin{pmatrix} 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then columns 2, 3 of matrix A are the pivot columns of A. We report $\vec{v}_2, \vec{v}_3$ as a largest independent subset, namely the pivot columns of A. **Beware**: The wrong answer is column 2, 3 of $\mathbf{rref}(A)$, because $\mathbf{rref}(A)$ columns are not in the original list of vectors! Example 5.24 is complete.

The Pivot Theorem can be restated as a method, called the **pivot method**, for finding a largest independent subset.

Theorem 5.42 (Pivot Method)
Let A be the augmented matrix of fixed vectors $\vec{v}_1, \ldots, \vec{v}_k$. Let the leading ones in $\mathbf{rref}(A)$ occur in columns $i_1, \ldots, i_p$. Then a largest independent subset of the k vectors $\vec{v}_1, \ldots, \vec{v}_k$ is the set of pivot columns of A, that is, the vectors

$$\vec{v}_{i_1}, \vec{v}_{i_2}, \ldots, \vec{v}_{i_p}.$$

Proof on page 424.

Rank and Nullity

Definition 5.32 (Rank of a Matrix)
The **rank** of an $m \times n$ matrix A, symbol $\mathbf{rank}(A)$, equals the number of leading ones in $\mathbf{rref}(A)$. Alternatively, the rank is the number of pivot columns of A.

Definition 5.33 (Nullity of a Matrix)
The **nullity** of an $m \times n$ matrix A is the number of free variables in the system $\mathbf{rref}(A)\vec{u} = \vec{0}$, or equivalently, the number of columns of A less the rank of A. The nullity equals the number of non-pivot columns of A in the Pivot Theorem.

The variable count in $\vec{u}$ equals the column dimension of A, which leads to the main result for rank and nullity.

Theorem 5.43 (Rank-Nullity Theorem)

$$\mathbf{rank}(A) + \mathbf{nullity}(A) = \text{column dimension of } A.$$

Proof on page 424.

In terms of homogeneous system $A\vec{u} = \vec{0}$, the rank of A is the number of leading variables and the nullity of A is the number of free variables, reliably computed from the system $\mathbf{rref}(A)\vec{x} = \vec{0}$.

Theorem 5.44 (Basis for $Ax = 0$)
Assume

$$k = \mathbf{nullity}(A) = \dim \left\{ \vec{x} : A\vec{x} = \vec{0} \right\} > 0.$$

Then the solution set of $A\vec{x} = \vec{0}$ can be expressed as

$$\vec{x} = t_1 \vec{X}_1 + \cdots + t_k \vec{X}_k \tag{1}$$

where $\vec{X}_1, \ldots, \vec{X}_k$ are special linearly independent solutions of $A\vec{x} = \vec{0}$ and $t_1, \ldots, t_k$ are arbitrary scalars (free variable invented symbols).

Proof on page 424.

Theorem 5.45 (Row Rank Equals Column Rank)
The number of independent rows of a matrix A equals the number of independent columns of A. Equivalently, $\mathbf{rank}(A) = \mathbf{rank}(A^T)$.

Proof on page 424.

Nullspace, Column Space and Row Space

Definition 5.34 (Kernel and Nullspace)
The **kernel** or **nullspace** of an $m \times n$ matrix A is the vector space of all solutions $\vec{x}$ to the homogeneous system $A\vec{x} = \vec{0}$. In symbols,

$$\mathbf{kernel}(A) \quad = \quad \mathbf{nullspace}(A) \quad = \quad \{\vec{x} \ : \ A\vec{x} = \vec{0}\}.$$

Definition 5.35 (Column Space)
The **column space** of $m \times n$ matrix A is the vector space consisting of all vectors $\vec{y} = A\vec{x}$, where $\vec{x}$ is arbitrary in $\mathcal{R}^n$.

In literature, the column space is also called the **image** of A, or the **range** of A, or the span of the columns of A. Because $A\vec{x}$ can be written as a linear combination of the columns $\vec{v}_1, \ldots, \vec{v}_n$ of A, the column space is the set of all linear combinations

$$\vec{y} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n.$$

In symbols,

$$\begin{aligned} \mathbf{colspace}(A) &= \{\vec{y} \ : \ \vec{y} = A\vec{x} \text{ for some } \vec{x}\} \\ &= \mathbf{Image}(A) \\ &= \mathbf{Range}(A) \\ &= \mathbf{span}(\vec{v}_1, \ldots, \vec{v}_n). \end{aligned}$$

Definition 5.36 (Row Space)
The **row space** of $m \times n$ matrix A is the vector space consisting of vectors $\vec{w} = A^T\vec{y}$, where $\vec{y}$ is arbitrary in $\mathcal{R}^m$. Technically, the row space of A is the column space of A^T. This vector space is viewed as the set of all linear combinations of rows of A. In symbols,

$$\begin{aligned} \mathbf{rowspace}(A) &= \mathbf{colspace}\left(A^T\right) \\ &= \{\vec{w} \ : \ \vec{w} = A^T\vec{y} \text{ for some } \vec{y}\} \\ &= \mathbf{Image}\left(A^T\right) \\ &= \mathbf{Range}\left(A^T\right). \end{aligned}$$

The row space of A and the null space of A live in $\mathcal{R}^n$, but the column space of A lives in $\mathcal{R}^m$, a different dimension. The correct bases are obtained as follows. If an alternative basis for **rowspace**(A) is suitable (rows of A not reported), then bases for **rowspace**(A), **colspace**(A), **nullspace**(A) can all be found by calculating just **rref**(A).

Null Space. Compute **rref**(A). Write out the general solution $\vec{x}$ to $A\vec{x} = \vec{0}$, where the free variables are assigned invented symbols $t_1, \ldots, t_k$. Report the basis for **nullspace**(A) as the list of partial derivatives $\partial_{t_1}\vec{x}, \ldots, \partial_{t_k}\vec{x}$, which are **special solutions** of $A\vec{x} = \vec{0}$.

Column Space. Compute $\mathbf{rref}(A)$. Identify the lead variable columns $i_1, \ldots, i_k$. Report the basis for $\mathbf{colspace}(A)$ as the list of columns $i_1, \ldots, i_k$ of A. These are the **pivot columns** of A.

Row Space. Compute $\mathbf{rref}\left(A^T\right)$. Identify the lead variable columns $i_1, \ldots, i_k$. Report the basis for $\mathbf{rowspace}(A)$ as the list of rows $i_1, \ldots, i_k$ of A.

Alternatively, compute $\mathbf{rref}(A)$, then $\mathbf{rowspace}(A)$ has a basis consisting of the list of nonzero rows of $\mathbf{rref}(A)$. The two bases obtained by these methods are different, but equivalent.

Due to the identity $\mathbf{nullity}(A) + \mathbf{rank}(A) = n$, where n is the column dimension of A, the following results hold. Notation: $\dim(V)$ is the dimension of vector space V, which equals the number of elements in a basis for V. Subspaces $\mathbf{nullspace}(A) = \mathbf{kernel}(A)$ and $\mathbf{colspace}(A) = \mathbf{Image}(A)$ have dual naming conventions in the literature.

Theorem 5.46 (Dimension Identities)

(a) $\dim(\mathbf{nullspace}(A)) = \dim(\mathbf{kernel}(A)) = \mathbf{nullity}(A)$

(b) $\dim(\mathbf{colspace}(A)) = \dim(\mathbf{Image}(A)) = \mathbf{rank}(A)$

(c) $\dim(\mathbf{rowspace}(A)) = \dim(\mathbf{Image}\left(A^T\right) = \mathbf{rank}(A)$

(d) $\dim(\mathbf{kernel}(A)) + \dim(\mathbf{Image}(A)) =$ column dimension of A

(e) $\dim(\mathbf{kernel}(A)) + \dim(\mathbf{kernel}\left(A^T\right)) =$ column dimension of A

Proof on page 425.

Equivalent Bases

Assume $\vec{v}_1, \ldots, \vec{v}_k$ are independent vectors in an abstract vector space V and let $S = \mathbf{span}(\vec{v}_1, \ldots, \vec{v}_n)$. Let $\vec{u}_1, \ldots, \vec{u}_\ell$ be another set of independent vectors in V.

Studied here is the question of whether or not $\vec{u}_1, \ldots, \vec{u}_\ell$ is a basis for S. First of all, all the vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ have to be in S, otherwise this set cannot possibly span S. Secondly, to be a basis, the vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ must be independent. Two bases for S must have the same number of elements, by Theorem 5.38. Therefore, $k = \ell$ is necessary for a possible second basis of S. These remarks establish:

Theorem 5.47 (Equivalent Bases of a Subspace S)

Let $\vec{v}_1, \ldots, \vec{v}_k$ be independent vectors in an abstract vector space V. Let S be the subspace of V consisting of all linear combinations of $\vec{v}_1, \ldots, \vec{v}_k$.

A set of vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ in V is an equivalent basis for S if and only

(1) Each of $\vec{u}_1, \ldots, \vec{u}_\ell$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$.

(2) The set $\vec{u}_1, \ldots, \vec{u}_\ell$ is independent.

(3) The sets are the same size, $k = \ell$.

Proof on page 425.

Equivalent Basis Test in $\mathcal{R}^n$

Assume given two sets of fixed vectors $\vec{v}_1, \ldots, \vec{v}_k$ and $\vec{u}_1, \ldots, \vec{u}_\ell$, in the same space $\mathcal{R}^n$. A test is developed for equivalence of bases, in a form suited for use in computer algebra systems and numerical laboratories.

Theorem 5.48 (Equivalence Test for Bases in $\mathcal{R}^n$)
Define augmented matrices

$$B = \left\langle \vec{v}_1 | \cdots | \vec{v}_k \right\rangle, \quad C = \left\langle \vec{u}_1 | \cdots | \vec{u}_\ell \right\rangle, \quad W = \left\langle B | C \right\rangle.$$

The relation

$$k = \ell = \mathbf{rank}(B) = \mathbf{rank}(C) = \mathbf{rank}(W)$$

implies

1. $\vec{v}_1, \ldots, \vec{v}_k$ is an independent set.
2. $\vec{u}_1, \ldots, \vec{u}_\ell$ is an independent set.
3. $\mathbf{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \mathbf{span}\{\vec{u}_1, \ldots, \vec{u}_\ell\}$

In particular, $\mathbf{colspace}(B) = \mathbf{colspace}(C)$ and each set of vectors is an equivalent basis for this vector space.

Proof on page 426.

Examples

Example 5.25 (Basis and Dimension)
Let S be the solution space in $V = \mathcal{R}^4$ of the system of equations

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & & & = & 0, \\ 2x_1 & + & 5x_2 & & & = & 0, \\ & & & & x_4 & = & 0, \\ & & & & 0 & = & 0. \end{array} \tag{2}$$

Find a basis for S, then report the dimension of S.

Solution: The solution divides into two distinct sections: $\boxed{\mathbf{1}}$ and $\boxed{\mathbf{2}}$.

$\boxed{\mathbf{1}}$: Find the scalar general solution of system (2).

The toolkit: matrix combination, swap and multiply on the coefficient matrix. The last frame algorithm finds the general solution. The details:

$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ First frame.

$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ `combo(1,2,-2)`.

$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ `combo(2,1,-2)`. Last frame, this is the **rref**.

$\left|\begin{array}{rcl} x_1 &=& 0, \\ x_2 &=& 0, \\ x_4 &=& 0, \\ 0 &=& 0. \end{array}\right|$ Translate to scalar equations.

$\left|\begin{array}{rcl} x_1 &=& 0, \\ x_2 &=& 0, \\ x_3 &=& t_1, \\ x_4 &=& 0. \end{array}\right|$ Scalar general solution, obtained from the last frame algorithm: x_1, x_2, x_4=lead, x_3=free.

$\boxed{\mathbf{2}}$: Find the vector general solution of the system (2).

The plan is to use the answer from $\boxed{\mathbf{1}}$ and partial differentiation to display the vector general solution $\vec{x}$.

$\partial_{t_1}\vec{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ The **special solution** is the partial on symbol t_1. Only one, because there is only one invented symbol.

$\vec{x} = t_1 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ The vector general solution.

Therefore, solution space $S = \mathbf{span}(\vec{X}_1)$, where $\vec{X}_1$ is the special solution obtained above. Because the spanning set is independent with one element, then $\dim(S) = 1$.

Example 5.26 (Euclidean Spaces)

Let A be an $m \times n$ matrix with columns $\vec{v}_1, \ldots, \vec{v}_n$ and let $\vec{b}$ be a vector in $\mathcal{R}^m$. Write a mathematical proof for each of the following facts.

$\boxed{\mathbf{1}}$. If the equation $A\vec{x} = \vec{b}$ has a solution $\vec{x}$, then $\vec{b}$ belongs to the span of the columns of A.

$\boxed{\mathbf{2}}$. If $\vec{b}$ belongs to the span of the columns of A, then the equation $A\vec{x} = \vec{b}$ has a solution $\vec{x}$.

$\boxed{\mathbf{3}}$. If $A\vec{x} = \vec{b}$ has a solution $\vec{x}$, then $\vec{b}, \vec{v}_1, \ldots, \vec{v}_n$ is a dependent set.

Solution:
$\boxed{\mathbf{1}}$: Let equation $A\vec{x} = \vec{b}$ have a solution $\vec{x}$. Write the equation backwards, then express the matrix product as a linear combination of the columns of A:

$$\vec{b} = A\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n.$$

This proves $\vec{b}$ is a linear combination of the columns of A.

$\boxed{\mathbf{1}}$: Let $\vec{b}$ be a linear combination of the columns of A. We show $A\vec{x} = \vec{b}$ has a solution $\vec{x}$. By hypothesis, there are constants $x_1, \ldots, x_n$ such that

$$\vec{b} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n.$$

Let $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Because $A\vec{x}$ can be written as a linear combination of the columns of A, then $A\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n = \vec{b}$, which proves that $A\vec{x} = \vec{b}$ has a solution $\vec{x}$.

$\boxed{\mathbf{3}}$: Assume $A\vec{x} = \vec{b}$ has a solution $\vec{x}$. Write the equation backwards, then express the matrix product as a linear combination of the columns of A:

$$\vec{b} = A\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n.$$

Define $c_0 = -1$, $c_1 = x_1, \ldots, c_n = x_n$. Then

$$c_0\vec{b} + c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}.$$

The definition of dependence implies that vectors $\vec{b}, \vec{v}_1, \ldots, \vec{v}_n$ are dependent.

The details for $\boxed{\mathbf{1}}$, $\boxed{\mathbf{2}}$, $\boxed{\mathbf{3}}$ are complete.

Example 5.27 (Sequence Spaces)

Let V be the vector space of all real sequences $\{x_n\}_{n=1}^{\infty}$ with componentwise addition and scalar multiplication. Let S be the subset of V defined by the equation $x_1 = 0$. Show that S is an infinite-dimensional subspace of V.

Solution: The space V is the abstraction of addition and scalar multiplication of Taylor series

$$f(t) = \sum_{n=1}^{\infty} x_n t^{n-1}.$$

The subspace S corresponds to all Taylor series which satisfy $f(0) = 0$. We assume it is known, or easily verified, that the larger set V is a vector space.

The **subspace criterion** applies to prove that S is a subspace of V. The omitted details are constructed from a similar set of details for the $\mathcal{R}^3$ subspace defined by the linear algebraic restriction equation $x_1 = 0$.

The remainder of the proof establishes $\dim(S) = \infty$. These details produce a list L with $\mathbf{span}(L) \subset S$. It is false that $S = \mathbf{span}(L)$, even though $\mathbf{span}(L)$ is a subspace by the **span theorem**. Further details are delayed to after $\dim(S) = \infty$ is established.

A standard method to find a basis L for S computes the partial derivatives on the symbols used to define S. The symbols are $x_2, x_3, \ldots$. We abuse notation and think of the sequences as column vectors with infinitely many components:

$$\{x_n\}_{n=1}^{\infty} \longrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix}.$$

Then S is the subset of all infinitely long column vectors with $x_1 = 0$. Take partial derivatives on $x_2, x_3, \ldots$ to obtain the candidate basis vectors:

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \ldots$$

The list is infinite. Any finite subset of this list is independent. The intuition:

$$c_1 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$$

implies

$$\begin{pmatrix} 0 \\ c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$$

and therefore $c_1 = c_2 = c_3 = 0$.

The list L of infinite sequences is correspondingly

$$\begin{array}{c} 0, 1, 0, 0, 0, \ldots \\ 0, 0, 1, 0, 0, \ldots \\ 0, 0, 0, 1, 0, \ldots \\ \vdots \end{array}$$

and there are infinitely many.

The details are finished by the method of contradiction. Suppose a true hypothesis and false conclusion. Then S has finite dimension n. Let Z be the span of a list L_1 of $n+1$ vectors from the above list. A proof can be constructed, based upon the above ideas, for independence of L_1, and then $\dim(Z) = n+1$. Because Z a subset of S, then $\dim(Z) \le \dim(S) = n$, a contradiction to $\dim(Z) = n+1$. Therefore, S cannot have finite dimension.

Conclusion: S is an infinite dimensional subspace of V.

Complaints. The preceding details do not prove Z is an independent set. The notation with infinitely many components is certainly not standard notation, therefore the reader is advised not to use it to present proof details. But it is excellent for intuition, and that is why you see it presented here, instead of more abstract details.

Is L a basis of S? The answer is **NO**.
Subspace $W = \mathbf{span}(L)$ is contained in subspace S. Then L is a basis for W. But L is not a basis for S. For example, the Taylor series for $f(t) = e^t - 1$ corresponds to a sequence in S with $x_k > 0$ for $k > 1$, and this sequence cannot be written as a finite linear combination of vectors selected from L.

Example 5.28 (Polynomial Spaces)

Let V be the vector space of all polynomials $p(x)$. Find a basis and hence the dimension of the subspace S defined by these conditions:

1. Polynomial $p(x)$ has degree no larger than two.

2. The equation $p(0) = \int_0^1 xp(x)dx$ is satisfied.

Solution: The answer is a list of independent polynomials in S: $\frac{2}{3} + x$, $\quad \frac{1}{2} + x^2$. Then $\dim(S) = 2 =$ number of basis elements.

Details. Start by inventing symbols for the coefficients of $p(x)$, for example $p(x) = a_1 + a_2x + a_3x^2$ because of requirement **1**. Insert the $p(x)$ expression into requirement **2**, in order to find a relation for the three symbols a_1, a_2, a_3.

$p(0) = \int_0^1 xp(x)dx$	Requirement **2**.
$a_1 = \int_0^1 (a_1x + a_2x^2 + a_3x^3)dx$	Insert for $p(x)$ the expression $a_1 + a_2x + a_3x^2$. Then evaluate $p(0) = a_1$.
$a_1 = \dfrac{a_1}{2} + \dfrac{a_2}{3} + \dfrac{a_3}{4}$	Evaluate integral.

Rearrangement of the last equation gives the linear equation $a_1 - \dfrac{2}{3}a_2 - \dfrac{1}{2}a_3 = 0$ in unknowns a_1, a_2, a_3. This linear system is in reduced echelon form. It has general solution

$$\begin{array}{rcl} a_1 & = & \frac{2}{3}t_1 + \frac{1}{2}t_2, \\ a_2 & = & t_1, \\ a_3 & = & t_2, \end{array} \tag{3}$$

with basis of solutions

$$\vec{v}_1 = \begin{pmatrix} \frac{2}{3} \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix}.$$

Translation to the corresponding polynomials, via the correspondence

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \longrightarrow p(x) = a_1 + a_2x + a_3x^2$$

gives the two polynomials

$$p_1(x) = \frac{2}{3} + x, \quad p_2(x) = \frac{1}{2} + x^2.$$

Why are these polynomials a basis for S?

A sophisticated answer is that the correspondence used to find the two polynomials is a one-to-one linear map from $W = \mathbf{span}(\vec{v}_1, \vec{v}_2)$ onto S, mapping $\vec{v}_1 \to p_1$ and $\vec{v}_2 \to p_2$.

A computational method will justify independence and span for the polynomials p_1, p_2. Start with $p(x) = a_1 + a_2x + a_3x^2$ in S. Equation $a_1 = \frac{2}{3}a_2 + \frac{1}{2}a_3$ holds because p belongs to S. Define $t_1 = a_2$, $t_2 = a_3$ (idea from equation (3)). Then all three equations in (3) are satisfied. Expand:

$$\begin{aligned} t_1p_1 + t_2p_2 &= a_2\left(\frac{2}{3} + x\right) + a_3\left(\frac{1}{2} + x^2\right) \\ &= \frac{2a_2}{3} + \frac{a_3}{2} + a_2x + a_3x^2 \\ &= a_1 + a_2x + a_3x^2. \end{aligned}$$

This computation proves that each polynomial in S is also in $\mathbf{span}(p_1, p_2)$, written as $S \subset \mathbf{span}(p_1, p_2)$. Because p_1, p_2 are already in S, then $\mathbf{span}(p_1, p_2) \subset S$. Then $S = \mathbf{span}(p_1, p_2)$, proving list p_1, p_2 spans S. The two polynomials p_1, p_2 are independent, because one is not a scalar multiple of the other. Then p_1, p_2 is independent and spans, which proves that p_1, p_2 is a basis for S and $\dim(S) = 2$.

Example 5.29 (Differential Equations)

A given homogeneous 5th order linear differential equation has general solution $y(x) = c_1 + c_2x + c_3x^2 + c_4\cos x + c_5\sin x$. Find a basis for the solution space S, a subspace of vector space $V = C^5(-\infty, \infty)$.

Solution: The answer is a list of 5 independent solutions in S: 1, x, x^2, $\cos x$, $\sin x$.

Details. The general solution expression implies S is the span of the reported list. We explain how to find the list. In the case of linear algebraic equations, we would take partial derivatives on the invented symbols to determine the list of special solutions, which is the basis. Here, we imagine c_1 to c_5 to be the invented symbols and take partial derivatives to determine a list of special vectors which span S. Let y abbreviate $y(x) = c_1 + c_2x + c_3x^2 + c_4\cos x + c_5\sin x$.

$$\partial_{c_1}y = 1, \quad \partial_{c_2}y = x, \quad \partial_{c_3}y = x^2, \quad \partial_{c_4}y = \cos x, \quad \partial_{c_5}y = \sin x.$$

The five answers are Euler solution atoms (defined on page 386). They are independent by Theorem 5.36, page 386. The general solution expression implies are solutions and they span S. They are a basis for S, dimension five.

Alternative Independence Test. The Wronskian test applies with sample $x = 0$. The Wronskian matrix is formed by rows which are successive derivatives of the list in row 1:

$$W(x) = \begin{pmatrix} 1 & x & x^2 & \cos x & \sin x \\ 0 & 1 & 2x & -\sin x & \cos x \\ 0 & 0 & 2 & -\cos x & -\sin x \\ 0 & 0 & 0 & \sin x & -\cos x \\ 0 & 0 & 0 & \cos x & \sin x \end{pmatrix}.$$

The determinant of $W(x)$ for sample $x = 0$ is $|W(0)| = 2$. The Wronskian test page 385 implies the list in row 1 of $W(x)$ is independent.

Example 5.30 (Largest Independent Subset)
Let $V = C(-\infty, \infty)$ and consider this list of vectors in V:

$$1,\ x + x^2,\ 2 + x,\ 1 + x^2,\ e^x,\ x + e^x.$$

Find a largest independent subset of this list.

Solution: One answer of the many possible answers is the list

$$1, x + x^2, 2 + x, e^x.$$

Details. Start with the nonzero vectors $1, x + x^2$. They are not scalar multiples of each other, hence they are independent. The initial independent subset is $1, x + x^2$. Vector $1 + x$ cannot be expressed as a combination of 1 and $x + x^2$, because such a relation

$$2 + x = c_1(1) + c_2(x + x^2)$$

requires both c_1 and c_2 nonzero, in which case we reach the impossibility that a linear polynomial equals a quadratic polynomial. The vector is added to the list to extend the initial independent subset to $1, x + x^2, 2 + x$. The different growth rate at $x = \infty$ of exponential term e^x explains why the independent subset is extended to $1, x+x^2, 2+x, e^x$.

Why is $x + e^x$ eliminated from the list? First, assemble two facts:

1. Vector x belongs to $\mathbf{span}(1, x + x^2, 2 + x)$.
2. Vectors x and e^x belong to $\mathbf{span}(1, x + x^2, 2 + x, e^x)$.

Then $x + e^x$ is in the span of the preceding vectors in the independent subset $1, x + x^2, 1 + x, e^x$. The final independent subset has been found.

Remark on the method. The pivot theorem does not directly apply to this example, because the vector space V is not a space $\mathcal{R}^n$ of fixed vectors. The pivot theorem can be used by reducing the original problem to an equivalent problem in some $\mathcal{R}^n$. This method is explored later, keyword **isomorphism**.

Example 5.31 (Pivot Theorem Method)
Extract a largest independent subset from the columns of the matrix

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

Solution: The answer is columns 2,3.

Details. The quickest solution is to observe that column 5 equals column 3 minus column 2, but columns 2,3 are nonzero and not scalar multiples of one another, therefore they are independent. Zero columns do not add to an independent subset of columns, therefore a largest independent subset of columns is obtained from columns 2,3.

A solution with computer implementation computes the pivot columns of A to be columns 2,3, and then we report a largest independent set of columns of A to be the pivot columns 2, 3.

The pivot columns of A are computed from the $\mathbf{rref}(A)$, which is found on paper using the toolkit *combo, swap, multiply.* It is a one-step process with computer assist: enter the matrix A and then write a command line for $\mathbf{rref}(A)$. The answer:

$$\mathbf{rref}(A) = \begin{pmatrix} 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then the pivot columns are columns $2, 3$ of matrix A. This is a largest independent subset of the columns of A.

Sample Code, Computer Algebra System `maple`:

```
A:=Matrix([[0,1,2,0,1],[0,1,1,0,0],[0,2,1,0,-1],
           [0,0,1,0,1],[0,0,1,0,1]]);
LinearAlgebra[ReducedRowEchelonForm](A);
```

Example 5.32 (Nullspace, Row Space, Column Space)
Compute the nullspace, column space and row space of the matrix

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

Solution: The answers appear below.

Details. The first computation is $\mathbf{rref}(A)$, which provides one answer for each of the three subspaces. The steps:

$\begin{pmatrix} 0&1&0&1 \\ 1&1&1&0 \\ 0&0&0&0 \\ 1&1&1&0 \end{pmatrix}$ Given matrix A.

$\begin{pmatrix} 1&1&1&0 \\ 0&1&0&1 \\ 0&0&0&0 \\ 1&1&1&0 \end{pmatrix}$ `swap(1,2)`

$\begin{pmatrix} 1&1&1&0 \\ 0&1&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$ `combo(1,4,-1)`

$\begin{pmatrix} 1&0&1&-1 \\ 0&1&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$ Begin back-substitution: `combo(2,1,-1)`. Found $\mathbf{rref}(A)$.

The last frame algorithm is applied to find the general solution of $A\vec{x} = \vec{0}$, using the scalar form of the last frame:

$$\begin{array}{ccccccccc} x_1 & & & + & x_3 & - & x_4 & = & 0, \\ & & x_2 & & & + & x_4 & = & 0, \\ & & & & & & 0 & = & 0, \\ & & & & & & 0 & = & 0. \end{array}$$

The lead variables are x_1, x_2 and the free variables x_3, x_4. Using invented symbols t_1, t_2 gives the general solution

$$\begin{aligned} x_1 &= -t_1 + t_2, \\ x_2 &= -t_2, \\ x_3 &= t_1, \\ x_4 &= t_2. \end{aligned}$$

Nullspace. The partial derivatives on the invented symbols, the *special solutions*, form a basis for the nullspace of A:

$$\mathbf{nullspace}(A) = \mathbf{kernel}(A) = \mathbf{span}\left(\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}\right).$$

Column Space. The column space of A is the span of the pivot columns of A, which according to the computed **rref** are columns 1, 2 of A. Then

$$\mathbf{colspace}(A) = \mathbf{span}(\text{pivot columns of } A) = \mathbf{span}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}\right).$$

Row space. One answer is the set of nonzero rows of $\mathbf{rref}(A)$. This gives the first answer

$$\mathbf{rowspace}(A) = \mathbf{span}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right).$$

A second answer is the set of pivot columns of A^T, columns 1,2 of A^T, found from

$$A^T = \begin{pmatrix} 0\,1\,0\,1 \\ 1\,1\,0\,1 \\ 0\,1\,0\,1 \\ 1\,0\,0\,0 \end{pmatrix}, \quad \mathbf{rref}(A^T) = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,1 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \end{pmatrix}.$$

Then the second answer for the row space is

$$\mathbf{rowspace}(A) = \mathbf{span}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right).$$

Example 5.33 (Fundamental Subspaces)

Compute the nullspace and column space for both A and A^T, given

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

The 4 computed subspaces are known as *Gilbert Strang's Four Fundamental Subspaces.*

Solution: Let $N_1 = \textbf{nullspace}(A) = \textbf{span}(\text{Strang's special solutions for } A)$ and $C_1 = \textbf{colspace}(A) = \textbf{span}(\text{pivot columns of } A)$. Both N_1 and C_1 were computed in the previous example:

$$N_1 = \textbf{span}\left(\begin{pmatrix}-1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\\1\end{pmatrix}\right), \quad C_1 = \textbf{span}\left(\begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\end{pmatrix}\right).$$

Define

$$\begin{aligned} N_2 &= \textbf{nullspace}(A^T) \\ &= \textbf{span}(\text{Strang's special solutions of } A^T), \\ C_2 &= \textbf{colspace}(A^T) \\ &= \textbf{span}(\text{pivot columns of } A^T). \end{aligned}$$

The computation of C_2 was completed in the previous example, which also computed

$$A^T = \begin{pmatrix}0&1&0&1\\1&1&0&1\\0&1&0&1\\1&0&0&0\end{pmatrix}, \quad \textbf{rref}(A^T) = \begin{pmatrix}1&0&0&0\\0&1&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix}.$$

Then the general solution for $A^T\vec{x} = \vec{0}$ is

$$x_1 = 0, x_2 = -t_2, x_3 = t_1, x_4 = t_2, \quad \text{Strang's special solutions} = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\0\\1\end{pmatrix}.$$

The newly found answer for N_2 plus the transcribed answer for C_2, taken from the previous example, give the equations

$$N_2 = \textbf{span}\left(\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\0\\1\end{pmatrix}\right), \quad C_2 = \textbf{span}\left(\begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right).$$

Example 5.34 (Equivalent Bases)
Let

$$\vec{v}_1 = \begin{pmatrix}0\\1\\\frac{3}{2}\end{pmatrix}, \vec{v}_2 = \begin{pmatrix}1\\0\\-\frac{1}{2}\end{pmatrix}, \vec{u}_1 = \begin{pmatrix}1\\3\\4\end{pmatrix}, \vec{u}_2 = \begin{pmatrix}3\\1\\0\end{pmatrix}.$$

Verify that $\{\vec{v}_1, \vec{v}_2\}$ and $\{\vec{u}_1, \vec{u}_2\}$ are equivalent bases for a subspace S.

Solution:

Define $B = \begin{pmatrix}0&1\\1&0\\\frac{3}{2}&-\frac{1}{2}\end{pmatrix}$, $C = \begin{pmatrix}1&3\\3&1\\4&0\end{pmatrix}$, $W = \begin{pmatrix}0&1&1&3\\1&0&3&1\\\frac{3}{2}&-\frac{1}{2}&4&0\end{pmatrix}$. Compute the **rank** of each matrix to be 2. Apply the theorem.

Maple Illustration.

```
v1:=<0,1,3/2>;v2:=<1,0,-1/2>; # Basis v1,v2
u1:=<1,3,4>;u2:=<3,1,0>;
B:=<v1|v2>; C:=<u1|u2>; W:=<B|C>;
# Test: ranks of B, C, W must equal 2
linalg[rank](B),linalg[rank](C),linalg[rank](W);
```

Example 5.35 (Equivalent Bases: False Test)
Does $\mathbf{rref}(B) = \mathbf{rref}(C)$ imply that each column of C is a linear combination of the columns of B? The answer is **no**. Supply a counter-example.

Solution: Define $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Then $\mathbf{rref}(B) = \mathbf{rref}(C) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, but column 2 of C is not a linear combination of the columns of B. This means $S_1 = \mathbf{colspace}(B)$ is not equal to $S_2 = \mathbf{colspace}(C)$. Geometrically, S_1 and S_2 are planes in $\mathcal{R}^3$ which intersect only along the line L through the two points $(0,0,0)$ and $(1,0,1)$.

What went wrong? The culprit is the toolkit operation `swap`.

Details and Proofs

Proof of Theorem 5.38, Size of a Basis: The proof proceeds by the formal method of contradiction. Assume the hypotheses are true and the conclusion is false. Then $p \neq q$. Without loss of generality, let the larger basis be listed first, $p > q$.

Because $\vec{u}_1, \ldots, \vec{u}_q$ is a basis of the vector space V, then there are coefficients $\{a_{ij}\}$ such that

$$\begin{array}{rcccccc} \vec{v}_1 & = & a_{11}\vec{u}_1 & + & \cdots & + & a_{1q}\vec{u}_q, \\ \vec{v}_2 & = & a_{21}\vec{u}_1 & + & \cdots & + & a_{2q}\vec{u}_q, \\ & \vdots & & & & & \\ \vec{v}_p & = & a_{p1}\vec{u}_1 & + & \cdots & + & a_{pq}\vec{u}_q. \end{array}$$

Let $A = [a_{ij}]$ be the $p \times q$ matrix of coefficients. Because $p > q$, then $\mathbf{rref}(A^T)$ has at most q leading variables and at least $p - q > 0$ free variables.

Then the $q \times p$ homogeneous system $A^T\vec{x} = \vec{0}$ has infinitely many solutions. Let $\vec{x}$ be a nonzero solution of $A^T\vec{x} = \vec{0}$.

The equation $A^T\vec{x} = \vec{0}$ means $\sum_{i=1}^p a_{ij}x_i = 0$ for $1 \leq j \leq p$, giving the dependence relation

$$\begin{array}{rcl} \sum_{i=1}^p x_i\vec{v}_i & = & \sum_{i=1}^p x_i \sum_{j=1}^q a_{ij}\vec{u}_j \\ & = & \sum_{j=1}^q \sum_{i=1}^p a_{ij}x_i\vec{u}_j \\ & = & \sum_{j=1}^q (0)\vec{u}_j \\ & = & \vec{0} \end{array}$$

The independence of $\vec{v}_1, \ldots, \vec{v}_p$ is contradicted. Arrival of the contradiction implies $p = q$. ■

Proof of Theorem 5.39, Basis of a finite dimensional vector space:

$\boxed{1}$ Let $S = \mathbf{span}(L)$, a subspace of V. By independence, $\dim(S) = n$. By hypothesis, $\dim(V) = n$. Suppose $\vec{v}$ is in V. Let L equal the list of $n + 1$ elements $\vec{v}_1, \ldots, \vec{v}_n, \vec{v}$. Then L is contained in V. Space V has dimension n, which means that no independent subset exists of size larger than n. So list L is not an independent set, which implies that $\vec{v}$ is in $S = \mathbf{span}(L)$. Therefore $S = V$ and L is a basis for V.

$\boxed{2}$ It suffices to prove under the given hypotheses that L is an independent set. If not, then $V = \mathbf{span}(L)$ is spanned by less than n independent vectors. This implies the dimension of V is less than n. A contradiction is reached, therefore L is an independent set.

$\boxed{3}$ Choose any independent subset of L, call it $\vec{w}_1, \ldots, \vec{w}_q$. If $q = n$, then we are done, by $\boxed{1}$. Otherwise, the span of these q vectors is a subspace of V not equal to V. Choose a vector $\vec{v}_{q+1}$ not in the subspace. Then $\vec{w}_1, \ldots, \vec{w}_{q+1}$ is an independent set in V. Repeat the construction until the number of constructed vectors equals n. Then the constructed list is a basis for V.

$\boxed{4}$ Assume $S = \mathbf{span}(L) = V$. If L contains fewer than n independent vectors, then V would have a basis of fewer than n elements, a violation of $\dim(V) = n$. Therefore, L contains n independent vectors. It cannot have more than n, without violating $\dim(V) = n$. Therefore, L contains exactly n independent vectors, which form a basis for V.

Proof of Theorem 5.40, Basis and Dimension in $\mathcal{R}^n$: The first result is due to the fact that all bases contain the same identical number of vectors. Because the columns of the $n \times n$ identity are independent and span $\mathcal{R}^n$, then all bases must contain n vectors, exactly.

A list of $n+1$ vectors $\vec{v}_1, \ldots, \vec{v}_{n+1}$ generates a subspace $S = \mathbf{span}(\vec{v}_1, \ldots, \vec{v}_{n+1})$. Because S is contained in $\mathcal{R}^n$, then S has a basis of n elements or less. Therefore, the list of $n+1$ vectors is dependent. ■

Proof of Theorem 5.41, The Pivot Theorem:

$\boxed{1}$: To prove: the pivot columns of A are independent. Let $\vec{v}_1, \ldots, \vec{v}_k$ be the vectors columns of matrix A. Let $i_1, \ldots, i_p$ be the pivot columns of A.

To apply the independence test, form the system of equations

$$c_1 \vec{v}_{i_1} + \cdots + c_p \vec{v}_{i_p} = \vec{0}$$

and solve for the constants $c_1, \ldots, c_p$, independence confirmed if they are all zero. The tool used to solve for the constants is the elementary matrix formula

$$A = M\,\mathbf{rref}(A), \quad M = E_1 E_2 \cdots E_r,$$

where $E_1, \ldots, E_r$ denote certain elementary matrices. Each elementary matrix is the inverse of a swap, multiply or combination operation applied to A, in order to reduce A to $\mathbf{rref}(A)$. Because elementary matrices are invertible, then M is invertible. The equation $A = \left\langle \vec{v}_1 | \cdots | \vec{v}_k \right\rangle$ implies the pivot columns of A satisfy the equation

$$\vec{v}_{i_q} = M \vec{e}_q, \quad q = 1, \ldots, p,$$

where $\vec{e}_1 = \mathbf{col}(I, 1), \ldots, \vec{e}_p = \mathbf{col}(I, p)$ are the consecutive columns of the identity matrix which occupy the columns of the leading ones in $\mathbf{rref}(A)$. Then

$$\begin{aligned} \vec{0} &= c_1 \vec{v}_{i_1} + \cdots + c_p \vec{v}_{i_p} \\ &= M(c_1 \vec{e}_1 + \cdots + c_p \vec{e}_p) \end{aligned}$$

implies by invertibility of M that

$$c_1\vec{e}_1 + \cdots + c_p\vec{e}_p = \vec{0}.$$

Distinct columns of the identity matrix are independent (subsets of independent sets are independent), therefore $c_1 = \cdots = c_p = 0$. The independence of the pivot columns of A is established. The proof of $\boxed{\mathbf{1}}$ is complete.

$\boxed{\mathbf{2}}$: To prove: a non-pivot column of A is a linear combination of the pivot columns of A. Let column j of A be non-pivot. Let's express this column as a linear combination of the pivot columns of A.

Consider the homogeneous system $A\vec{x} = \vec{0}$ and its equivalent system $\mathbf{rref}(A)\vec{x} = \vec{0}$. The pivot column subscripts determine the leading variables and the remaining column subscripts determine the free variables. Then column j matches a free variable x_j. Define $x_j = 1$. Define all other free variables to be zero. The lead variables are now determined and the resulting nonzero vector $\vec{x}$ satisfies the homogeneous equation $\mathbf{rref}(A)\vec{x} = \vec{0}$, and hence also $A\vec{x} = \vec{0}$. Translating this equation into a linear combination of columns implies

$$\left(\sum_{\text{pivot subscripts } i} x_i\vec{v}_i\right) + \vec{v}_j = \vec{0}$$

which in turn implies that column j of A is a linear combination of the pivot columns of A. The proof of $\boxed{\mathbf{2}}$ is complete.

Proof of Theorem 5.42, The Pivot Method: According to the Pivot Theorem 5.41, the fixed vectors are independent. An attempt to add another column of A to these chosen columns results in a non-pivot column being added. The Pivot Theorem applies: the column added is dependent on the pivot columns. Therefore, the set of pivot columns of A forms a largest independent subset of the columns of A. ■

Proof of Theorem 5.43, The Rank-Nullity Theorem: The rank of A is the number of leading ones in $\mathbf{rref}(A)$. The nullity of A is the number of non-pivot columns in A. The sum of the rank and nullity is the number of variables, which is the column dimension n of A. Then the rank + nullity = n, as claimed. ■

Proof of Theorem 5.44, Basis for $A\vec{x} = \vec{0}$: The system $\mathbf{rref}(A)\vec{x} = \vec{0}$ has exactly the same solution set as $A\vec{x} = \vec{0}$. This system has a standard general solution $\vec{x}$ expressed in terms of invented symbols $t_1, \ldots, t_k$. Define $\vec{X}_j = \partial_{t_j}\vec{x}$, $j = 1, \ldots, k$. Then (1) holds. It remains to prove independence, which means we are to solve for $c_1, \ldots, c_k$ in the system

$$c_1\vec{X}_1 + \cdots + c_k\vec{X}_k = \vec{0}.$$

The left side is a solution $\vec{x}$ of $A\vec{x} = \vec{0}$ in which the invented symbols have been assigned values $c_1, \ldots, c_k$. The right side implies each component of $\vec{x}$ is zero. Because the standard general solution assigns invented symbols to free variables, the relation above implies that each free variable is zero. But free variables have already been assigned values $c_1, \ldots, c_k$. Therefore, $c_1 = \cdots = c_k = 0$. ■

Proof Theorem 5.45, Row Rank equals Column Rank: Let S be the set of all linear combinations of columns of A. Then $S = \mathbf{span}(\text{columns of } A) = \mathbf{Image}(A)$. The non-pivot columns of A are linear combinations of pivot columns of A. Therefore, any linear combination of columns of A is a linear combination of the $p = \mathbf{rank}(A)$ linearly independent pivot columns. By definition, the pivot columns form a **basis** for the vector space S, and $p = \mathbf{rank}(A) = \dim(S)$.

The **span** R **of the rows of** A is defined to be the set of all linear combinations of the columns of A^T.

Let $q = \mathbf{rank}(A^T) = \dim(R)$. It will be shown that $p = q$, which proves the theorem.

Let $\mathbf{rref}(A) = E_1 \cdots E_k A$ where $E_1, \ldots, E_k$ are elementary swap, multiply and combination matrices. The invertible matrix $M = E_1 \cdots E_k$ satisfies the equation $\mathbf{rref}(A) = MA$. Then:

$$\mathbf{rref}(A)^T = A^T M^T$$

Matrix $\mathbf{rref}(A)^T$ has its first p columns independent and its remaining columns are zero. Each nonzero column of $\mathbf{rref}(A)^T$ is expressed on the right as a linear combination of the columns of A^T. Therefore, R contains p independent vectors. The number $q = \dim(R)$ is the vector count in any basis for R. This implies $p \le q$.

The preceding display can be solved for A^T, because M^T is invertible, giving

$$A^T = \mathbf{rref}(A)^T (M^T)^{-1}.$$

Then every column of A^T is a linear combination of the p nonzero columns of $\mathbf{rref}(A)^T$. This implies a basis for R contains at most p elements, i.e., $q \le p$.

Combining $p \le q$ with $q \le p$ proves $p = q$. ■

Proof of Theorem 5.46, Dimension Identities:

(a) $\dim(\mathbf{nullspace}(A)) = \dim(\mathbf{kernel}(A)) = \mathbf{nullity}(A)$

The nullspace is the kernel, defined as the set of solutions to $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$. This set has basis *Strang's Special Solutions*, the number of which matches the number of free variables. That number is the nullity of A.

(b) $\dim(\mathbf{colspace}(A)) = \dim(\mathbf{Image}(A)) = \mathbf{rank}(A)$

The column space has as a basis the pivot columns of A. The number of pivot columns is the rank of A.

(c) $\dim(\mathbf{rowspace}(A)) = \dim(\mathbf{Image}\left(A^T\right) = \mathbf{rank}(A)$

The row space has a basis given by the pivot columns of A^T. The number of columns is the number of independent rows of A, or the row rank of A, which by Theorem 5.45 equals the rank of A.

(d) $\dim(\mathbf{kernel}(A)) + \dim(\mathbf{Image}(A)) =$ column dimension of A

This identity restates the Rank-Nullity Theorem 5.43.

(e) $\dim(\mathbf{kernel}(A)) + \dim(\mathbf{kernel}\left(A^T\right)) =$ column dimension of A

Apply part (d) to A^T. If $\dim(\mathbf{kernel}(A)) = \dim(\mathbf{Image}(A^T))$ then identity (e) follows. Let $r = \dim(\mathbf{kernel}(A))$ and $s = \dim(\mathbf{Image}(A^T))$. We must show $r = s$. Already known is $r = \mathbf{nullity}(A)$, which equals the number of Strang's Special Solutions. Number s is the number of independent columns in A^T, which equals the row rank of A. Theorem 5.45 applies: s equals the row rank of A, which is the rank of A, which is r. Then $r = s$, as claimed. ■

Proof of Theorem 5.47, Equivalent Bases: Vectors $\vec{w}_1, \ldots, \vec{w}_k$ are a basis for S provided they are independent and span S. The three items from the theorem:

(1) Each of $\vec{u}_1, \ldots, \vec{u}_\ell$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_k$.

(2) The set $\vec{u}_1, \ldots, \vec{u}_\ell$ is independent.

(3) The sets are the same size, $k = \ell$.

Sufficiency. Assume given vectors $\vec{v}_1, \ldots, \vec{v}_k$ which form a basis for S. Assume vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ are also a basis for S. Then these vectors are independent and span S. The spanning condition $S = \textbf{span}(\vec{u}_1, \ldots, \vec{u}_k)$ implies (1). Independence implies (2). Theorem 5.38 applies: the two bases have the same size: $k = \ell$, which proves (3) holds.

Necessity. Assume that vectors $\vec{v}_1, \ldots, \vec{v}_k$ form a basis for S. Assume given vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ in S satisfying (1), (2), (3). We prove $\vec{u}_1, \ldots, \vec{u}_\ell$ is a basis for S. Item (2) implies the vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ are independent and (1) implies they span S, because $\vec{v}_1, \ldots, \vec{v}_k$ span S. The definition of basis applies: vectors $\vec{u}_1, \ldots, \vec{u}_\ell$ form a basis for S. ■

Proof of Theorem 5.48, Equivalence test for bases in $\mathcal{R}^n$:
Because $\textbf{rank}(B) = k$, then the first k columns of W are independent. If some column of C is independent of the columns of B, then W would have $k+1$ independent columns, which violates $k = \textbf{rank}(W)$. Therefore, the columns of C are linear combinations of the columns of B. The vector space $\mathcal{U} = \textbf{colspace}(C)$ is therefore a subspace of the vector space $\mathcal{V} = \textbf{colspace}(B)$. Because each vector space has dimension k, then $\mathcal{U} = \mathcal{V}$. ■

Exercises 5.5

Basis and Dimension

Compute a basis and the report the dimension of the subspace S.

1. In $\mathcal{R}^3$, S is the solution space of

$$\left|\begin{array}{ccccccc} x_1 & & & + & x_3 & = & 0, \\ & & x_2 & + & x_3 & = & 0. \end{array}\right|$$

2. In $\mathcal{R}^4$, S is the solution space of

$$\left|\begin{array}{ccc} x_1 + 2x_2 + x_3 & & = & 0, \\ & x_4 & = & 0. \end{array}\right|$$

3. In $\mathcal{R}^2$, $S = \textbf{span}(\vec{v}_1, \vec{v}_2)$. Vectors $\vec{v}_1, \vec{v}_2$ are columns of an invertible matrix.

4. Set $S = \textbf{span}(\vec{v}_1, \vec{v}_2)$, in $\mathcal{R}^4$. The vectors are columns in a 4×4 invertible matrix.

5. Set $S = \textbf{span}(\sin^2 x, \cos^2 x, 1)$, in the vector space V of continuous functions.

6. Set $S = \textbf{span}(x, x-1, x+2)$, in the vector space V of all polynomials.

7. Set $S = \textbf{span}(\sin x, \cos x)$, the solution space of $y'' + y = 0$.

8. Set $S = \textbf{span}\left(e^{2x}, e^{3x}\right)$, the solution space of $y'' - 5y' + 6y = 0$.

Euclidean Spaces

9. Let A be 3×2. Why is it impossible for the columns of A to be a basis for $\mathcal{R}^3$?

10. Let A be $m \times n$. What condition on indices m, n implies it is impossible for the columns of A to be a basis for $\mathcal{R}^m$?

11. Find a pairwise orthogonal basis for $\mathcal{R}^3$ which contains $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

12. Display a basis for $\mathcal{R}^4$ which contains the independent columns of $\begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

13. Let S be a subspace of $\mathcal{R}^{10}$ of dimension 5. Insert a basis for S into an $m \times n$ augmented matrix A. What are m and n?

14. Suppose A and B are 3×3 matrices and let $C = AB$. Assume the columns of A are not a basis for $\mathcal{R}^3$. Is there a matrix B so that the columns of C form a basis for $\mathcal{R}^3$?

15. The term **Hyperplane** is used for an equation like $x_4 = 0$, which in $\mathcal{R}^4$ defines a subspace S of dimension 3. Find a basis for S.

16. Find a 3-dimensional subspace S of $\mathcal{R}^4$ which has no basis consisting of columns of the identity matrix.

Polynomial Spaces

Symbol V is the vector space of all polynomials $p(x)$. Given subspace S of V, find a basis for S and $\dim(S)$.

17. The subset S of $\mathbf{span}(1, x, x^2)$ is defined by $\frac{dp}{dx}(1) = 0$.

18. The subset S of $\mathbf{span}(1, x, x^2, x^3)$ is defined by $p(0) = \frac{dp}{dx}(1) = 0$.

19. The subset S of $\mathbf{span}(1, x, x^2)$ is defined by $\int_0^1 p(x)dx = 0$.

20. The subset S of $\mathbf{span}(1, x, x^2, x^3)$ is defined by $\int_0^1 xp(x)dx = 0$.

Differential Equations

Find a basis for solution subspace S. Assume the general solution of the 4th order linear differential equation is

$$y(x) = c_1 + c_2 x + c_3 e^x + c_4 e^{-x}.$$

21. Subspace S_1 is defined by $y(0) = \frac{dy}{dx}(0) = 0$.

22. Subspace S_2 is defined by $y(1) = 0$.

23. Subspace S_3 is defined by $y(0) = \int_0^1 y(x)dx$.

24. Subspace S_4 is defined by $y(1) = 0, \int_0^1 y(x)dx = 0$.

Largest Subset of Independent Vectors

Find a largest independent subset of the given vectors.

25. The columns of $\begin{pmatrix} 0&0&1&1 \\ 0&0&1&1 \\ 0&1&1&0 \\ 0&1&2&1 \end{pmatrix}$.

26. The columns of $\begin{pmatrix} 3&1&2&0&5 \\ 2&1&1&0&4 \\ 3&2&1&0&7 \\ 1&0&1&0&1 \\ 3&2&1&0&7 \end{pmatrix}$.

27. The polynomials $x, 1+x, 1-x, x^2$.

28. The continuous functions x, e^x, $x+e^x$, e^{2x}.

Pivot Theorem Method

Extract a largest independent set from the columns of the given matrix A. The answer is a list of independent columns of A, called the pivot columns of A.

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

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

31. $\begin{pmatrix} 0&2&1&0&1 \\ 1&5&2&0&3 \\ 1&3&1&0&2 \\ 0&2&1&0&3 \\ 0&2&1&0&1 \end{pmatrix}$

32. $\begin{pmatrix} 0&0&2&1&0&1 \\ 0&1&5&2&0&3 \\ 0&1&3&1&0&2 \\ 0&2&4&1&0&3 \\ 0&0&2&1&0&1 \\ 0&2&4&1&0&3 \end{pmatrix}$

Row and Column Rank

Justify by direct computation that $\mathbf{rank}(A) = \mathbf{rank}\left(A^T\right)$, which means that the row rank equals the column rank.

33. $A = \begin{pmatrix} 1&0&1 \\ 0&1&1 \\ 0&0&0 \end{pmatrix}$

34. $A = \begin{pmatrix} 1&0&1&0 \\ 0&1&1&1 \\ 0&0&0&0 \\ 0&1&1&1 \end{pmatrix}$

Nullspace or Kernel

Find a basis for the nullspace of A, which is also called the kernel of A.

35. $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

36. $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$

Row Space

Find a basis for the row space of A. There are two possible answers: (1) The nonzero rows of $\mathbf{rref}(A)$, (2) The pivot columns of A^T. Answers (1) and (2) can differ wildly.

37. $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

38. $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$

Column Space

Find a basis for the column space of A, in terms of the columns of A. Normally, we report the pivot columns of A.

39. $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

40. $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$

Dimension Identities

Let A be an $m \times n$ matrix of rank r. Prove the following dimension identities in Theorem 5.46.

41. $\dim(\mathbf{nullspace}(A)) = n - r$

42. $\dim(\mathbf{colspace}(A)) = r$

43. $\dim(\mathbf{rowspace}(A)) = r$

44. The dimensions of $\mathbf{nullspace}(A)$ and $\mathbf{colspace}(A)$ add to n.

Orthogonal Complement $S^\perp$

Let S be a subspace of vector space $V = \mathcal{R}^n$. Define the **Orthogonal complement** by

(4) $\quad S^\perp = \{\vec{x} \;:\; \vec{x}^T\vec{y} = 0, \;\vec{y} \text{ in } S\}.$

45. Let $V = \mathcal{R}^3$ and let S be the xy-plane. Compute $S^\perp$. Answer: The z-axis.

46. Prove that $S^\perp$ is a subspace, using the **Subspace Criterion**.

47. Prove that the orthogonal complement of $S^\perp$ is S. In symbols, $\left(S^\perp\right)^\perp = S$.

48. Prove that

$$V = \{\vec{x} + \vec{y} \;:\; \vec{x} \in S, \vec{y} \in S^\perp\}.$$

This relation is called the **Direct Sum** of S and $S^\perp$.

Fundamental Theorem of Linear Algebra

Let A be an $m \times n$ matrix.

49. Write a short proof:
Lemma. Any solution of $A\vec{x} = \vec{0}$ is orthogonal to every row of A.

50. Find the dimension of the kernel and image for both A and A^T. The four answers use symbols $m, n, \mathbf{rank}(A)$. The main tool is the rank-nullity theorem.

51. Prove
$\mathbf{kernel}(A) = \mathbf{Image}\left(A^T\right)^\perp$. Use Exercise 49.

52. Prove
$\mathbf{kernel}\left(A^T\right) = \mathbf{Image}\,(A)^\perp$.

Fundamental Subspaces

The kernel and image of both A and A^T are called *The Four Fundamental Subspaces* by Gilbert Strang. Let A denote an $n \times m$ matrix.

53. Prove using Exercise 51:
$\mathbf{kernel}(A) = \mathbf{rowspace}(A)^\perp$

54. Establish these four identities.
$\mathbf{kernel}(A) = \mathbf{Image}\left(A^T\right)^\perp$
$\mathbf{kernel}\left(A^T\right) = \mathbf{Image}\,(A)^\perp$
$\mathbf{Image}\,(A) = \mathbf{kernel}(A^T)^\perp$
$\mathbf{Image}\left(A^T\right) = \mathbf{kernel}(A)^\perp$

Notation. *kernel* is null space, *image* is column space, symbol $\perp$ is orthogonal complement: see equation (4).

Equivalent Bases

Test the given subspaces for equality.

55. $S_1 = \textbf{span}\left(\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right)$,

$S_2 = \textbf{span}\left(\begin{pmatrix}3\\3\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right)$

56. $S_3 = \textbf{span}\left(\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\end{pmatrix}\right)$,

$S_4 = \textbf{span}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right)$

57. $S_5 = \textbf{span}\left(\begin{pmatrix}1\\0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\1\end{pmatrix}\right)$,

$S_6 = \textbf{span}\left(\begin{pmatrix}1\\0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}\right)$

58. $S_7 = \textbf{span}\left(\begin{pmatrix}2\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\1\end{pmatrix}\right)$,

$S_8 = \textbf{span}\left(\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}, \begin{pmatrix}3\\3\\2\\2\end{pmatrix}\right)$

PDF Sources

Text, Solutions and Corrections

Author: Grant B. Gustafson, University of Utah, Salt Lake City 84112.
Paperback Textbook: There are 12 chapters on differential equations and linear algebra, book format 7 x 10 inches, 1077 pages. Copies of the textbook are available in two volumes at **Amazon** Kindle Direct Publishing for Amazon's cost of printing and shipping. No author profit. Volume I chapters 1-7, ISBN 9798705491124, 661 pages. Volume II chapters 8-12, ISBN 9798711123651, 479 pages. Both paperbacks have extra pages of backmatter: background topics Chapter A, the whole book index and the bibliography.

Textbook PDF with Solution Manual: Packaged as one PDF (13 MB) with hyperlink navigation to displayed equations and theorems. The header in an exercise set has a blue hyperlink to the same section in the solutions. The header of the exercise section within a solution Appendix has a red hyperlink to the textbook exercises. Solutions are organized by chapter, e.g., Appendix 5 for Chapter 5. Odd-numbered exercises have a solution. A few even-numbered exercises have hints and answers. Computer code can be mouse-copied directly from the PDF. Free to use or download, no restrictions for educational use.

Sources at Utah:

https://math.utah.edu/g̃ustafso/indexUtahBookGG.html

Sources for a Local Folder No Internet: The same PDF can be downloaded to a tablet, computer or phone to be viewed locally. After download, no internet is required. Best for computer or tablet using a PDF viewer (Adobe Reader, Evince) or web browser with PDF support (Chrome, FireFox). Smart phones can be used in landscape mode.

Sources at GitHub and GitLab Projects: Utah sources are duplicated at

`https://github.com/ggustaf/github.io` and mirror
`https://gitlab.com/ggustaf/answers`.

Communication: To contribute a solution or correction, ask a question or request an answer, click the link below, then create a GitHub issue and post. Contributions and corrections are credited, privacy respected.

`https://github.com/ggustaf/github.io/issues`