Chapter 1. Introduction to Vectors

1.1 Vectors and Linear Combinations

"You can't add apples and oranges." In a strange way, this is the reason for vectors. We have two separate numbers \( v_1 \) and \( v_2 \). That pair produces a two-dimensional vector \( \mathbf{v} \):

\[
\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}
\]

\( v_1 \) is the first component
\( v_2 \) is the second component

We write \( \mathbf{v} \) as a column, not as a row. The main point so far is to have a single letter \( \mathbf{v} \) (in **boldface italic**) for this pair of numbers \( v_1 \) and \( v_2 \) (in **lightface italic**).

Even if we don't add \( v_1 \) to \( v_2 \), we do **add vectors**. The first components of \( \mathbf{v} \) and \( \mathbf{w} \) stay separate from the second components:

\[
\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}
\]

\( \mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix} \)

You see the reason. We want to add apples to apples. **Subtraction** of vectors follows the same idea: The components of \( \mathbf{v} - \mathbf{w} \) are \( v_1 - w_1 \) and \( v_2 - w_2 \).

The other basic operation is **scalar multiplication**. Vectors can be multiplied by \( 2 \) or by \( -1 \) or by any number \( c \). There are two ways to double a vector. One way is to add \( \mathbf{v} + \mathbf{v} \). The other way (the usual way) is to multiply each component by \( 2 \):

\[
2\mathbf{v} = \begin{bmatrix} 2v_1 \\ 2v_2 \end{bmatrix} \quad \text{and} \quad -\mathbf{v} = \begin{bmatrix} -v_1 \\ -v_2 \end{bmatrix}
\]

The components of \( c\mathbf{v} \) are \( cv_1 \) and \( cv_2 \). The number \( c \) is called a "**scalar**".

Notice that the sum of \( -\mathbf{v} \) and \( \mathbf{v} \) is the zero vector. This is \( \mathbf{0} \), which is not the same as the number zero! The vector \( \mathbf{0} \) has components \( 0 \) and \( 0 \). Forgive me for hammering away at the difference between a vector and its components. Linear algebra is built on these operations \( \mathbf{v} + \mathbf{w} \) and \( c\mathbf{v} \)—adding vectors and multiplying by scalars.

The order of addition makes no difference: \( \mathbf{v} + \mathbf{w} \) equals \( \mathbf{w} + \mathbf{v} \). Check that by algebra:

The first component is \( v_1 + w_1 \) which equals \( w_1 + v_1 \). Check also by an example:

\[
\mathbf{v} + \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 8 \end{bmatrix} \quad \text{and} \quad \mathbf{w} + \mathbf{v} = \begin{bmatrix} 3 \\ 3 \\ 8 \end{bmatrix}
\]

**Linear Combinations**

Combining addition with scalar multiplication, we now form "**linear combinations**" of \( \mathbf{v} \) and \( \mathbf{w} \). Multiply \( \mathbf{v} \) by \( c \) and multiply \( \mathbf{w} \) by \( d \); then add \( c\mathbf{v} + d\mathbf{w} \).

**Definition** The sum of \( c\mathbf{v} \) and \( d\mathbf{w} \) is a linear combination of \( \mathbf{v} \) and \( \mathbf{w} \).

Four special linear combinations are: sum, difference, zero, and a scalar multiple \( c\mathbf{v} \):

\[
\begin{align*}
\mathbf{v} + \mathbf{w} &= \text{sum of vectors in Figure 1.1a} \\
\mathbf{v} - \mathbf{w} &= \text{difference of vectors in Figure 1.1b} \\
0\mathbf{v} + 0\mathbf{w} &= \text{zero vector} \\
c\mathbf{v} + d\mathbf{w} &= \text{vector } c\mathbf{v} \text{ in the direction of } \mathbf{v}
\end{align*}
\]

The zero vector is always a possible combination (its coefficients are zero). Every time we see a "space" of vectors, that zero vector will be included. This big view, taking all the combinations of \( \mathbf{v} \) and \( \mathbf{w} \), is linear algebra at work.

The figures show how you can visualize vectors. For algebra, we just need the components (like \( 4 \) and \( 2 \)). That vector \( \mathbf{v} \) is represented by an arrow. The arrow goes \( v_1 = 4 \) units to the right and \( v_2 = 2 \) units up. It ends at the point whose \( x \), \( y \) coordinates are \( 4 \), \( 2 \).

This point is another representation of the vector—we have three ways to describe \( \mathbf{v} \): **Represent vector \( \mathbf{v} \)**

- Two numbers
- Arrow from \((0, 0)\)
- Point in the plane

We add using the numbers. We visualize \( \mathbf{v} + \mathbf{w} \) using arrows:

**Vector addition (head to tail)** At the end of \( \mathbf{v} \), place the start of \( \mathbf{w} \).

\[
\begin{align*}
\mathbf{w} &= \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} \\
\mathbf{v} &= \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \\
\mathbf{v} + \mathbf{w} &= \begin{bmatrix} 5 \\ 2 \\ 6 \end{bmatrix} \\
\mathbf{w} - \mathbf{v} &= \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}
\end{align*}
\]

Figure 1.1: Vector addition \( \mathbf{v} + \mathbf{w} = (3, 4) \) produces the diagonal of a parallelogram.

The linear combination on the right is \( \mathbf{v} - \mathbf{w} = (5, 0) \).

We travel along \( \mathbf{v} \) and then along \( \mathbf{w} \). Or we take the diagonal shortcut along \( \mathbf{v} + \mathbf{w} \). We could also go along \( \mathbf{w} \) and then \( \mathbf{v} \). In other words, \( \mathbf{w} + \mathbf{v} \) gives the same answer as \( \mathbf{v} + \mathbf{w} \).
Chapter 1. Introduction to Vectors

These are different ways along the parallelogram (in this example it is a rectangle). The sum is the diagonal vector \( \mathbf{v} + \mathbf{w} \).

The zero vector \( \mathbf{0} = (0, 0) \) is too short to draw a decent arrow, but you know that \( \mathbf{v} + \mathbf{0} = \mathbf{v} \). For 2\( \mathbf{v} \) we double the length of the arrow. We reverse \( \mathbf{w} \) to get \(-\mathbf{w}\). This reversing gives the subtraction on the right side of Figure 1.1.

Vectors in Three Dimensions

A vector with two components corresponds to a point in the \( xy \) plane. The components of \( \mathbf{v} \) are the coordinates of the point: \( x = v_1 \) and \( y = v_2 \). The arrow ends at this point \((v_1, v_2)\), when it starts from \((0, 0)\). Now we allow vectors to have three components \((v_1, v_2, v_3)\).

The \( xy \) plane is replaced by three-dimensional space. Here are typical vectors (still column vectors but with three components):

\[
\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v} + \mathbf{w} = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}.
\]

The vector \( \mathbf{v} \) corresponds to an arrow in 3-space. Usually the arrow starts at the "origin", where the \( xyz \) axes meet and the coordinates are \((0, 0, 0)\). The arrow ends at the point with coordinates \(v_1, v_2, v_3\). There is a perfect match between the column vector and the arrow from the origin and the point where the arrow ends.

Figure 1.2: Vectors \( \begin{bmatrix} x \\ y \end{bmatrix} \) and \( \begin{bmatrix} x \\ y \\ z \end{bmatrix} \) correspond to points \((x, y)\) and \((x, y, z)\).

\[\text{From now on } \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \text{ is also written as } \mathbf{v} = (1, 1, -1).\]

1.1. Vectors and Linear Combinations

The reason for the row form (in parentheses) is to save space. But \( \mathbf{v} = (1, 1, -1) \) is not a row vector! It is in actuality a column vector, just temporarily lying down. The row vector \([1 \ 1 \ -1]\) is absolutely different, even though it has the same three components.

That row vector is the "transpose" of the column \( \mathbf{v} \). In three dimensions, \( \mathbf{v} + \mathbf{w} \) is still found a component at a time. The sum has components \(v_1 + w_1\), \(v_2 + w_2\), and \(v_3 + w_3\). You see how to add vectors in 4 or 5 or \( n \) dimensions. When \( \mathbf{w} \) starts at the end of \( \mathbf{v} \), the third side is \( \mathbf{w} \). The other way around the parallelogram is \( \mathbf{w} + \mathbf{v} \). Question: Do the four sides all lie in the same plane? Yes: And the sum \( \mathbf{w} + \mathbf{v} = \mathbf{v} + \mathbf{w} \) goes completely around to produce the \( \mathbf{v} \) vector.

A typical linear combination of three vectors in three dimensions is \( u + 4v - 2w \):

\[
\begin{align*}
\text{Linear combination} & \\
\text{Multiply by } 1, 4, -2 & \\
\text{Then add} & \\
\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} & + 4 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.
\end{align*}
\]

The Important Questions

For one vector \( \mathbf{u} \), the only linear combinations are the multiples \( c \mathbf{u} \). For two vectors, the combinations are \( \mathbf{u} + d \mathbf{v} \). For three vectors, the combinations are \( \mathbf{u} + d \mathbf{v} + e \mathbf{w} \). Will you take the big step from one combination to all combinations? Every \( c \) and \( d \) and \( e \) are allowed. Suppose the vectors \( \mathbf{u}, \mathbf{v}, \mathbf{w} \) are in three-dimensional space:

1. What is the picture of all combinations \( c \mathbf{u} \)?
2. What is the picture of all combinations \( c \mathbf{u} + d \mathbf{v} \)?
3. What is the picture of all combinations \( c \mathbf{u} + d \mathbf{v} + e \mathbf{w} \)?

The answers depend on the particular vectors \( \mathbf{u}, \mathbf{v}, \) and \( \mathbf{w} \). If they were zero vectors (a very extreme case), then every combination would be zero. If they are typical nonzero vectors (components chosen at random), here are the three answers. This is the key to our subject:

1. The combinations \( c \mathbf{u} \) fill a line.
2. The combinations \( c \mathbf{u} + d \mathbf{v} \) fill a plane.
3. The combinations \( c \mathbf{u} + d \mathbf{v} + e \mathbf{w} \) fill three-dimensional space.

The zero vector \((0, 0, 0)\) is on the line because \( c \) can be zero. It is on the plane because \( c \) and \( d \) can be zero. The line of vectors \( c \mathbf{u} \) is infinitely long (forward and backward). It is the plane of all \( c \mathbf{u} + d \mathbf{v} \) (combining two vectors in three-dimensional space) that I especially ask you to think about.

Adding all \( cu \) on one line to all \( dv \) on the other line fills the plane in Figure 1.3.

When we include a third vector \( \mathbf{w} \), the multiples \( c \mathbf{w} \) give a third line. Suppose that third line is not in the plane of \( \mathbf{u} \) and \( \mathbf{v} \). Then combining all \( c \mathbf{w} \) with all \( c \mathbf{u} + d \mathbf{v} \) fills up the whole three-dimensional space.
Chapter 1. Introduction to Vectors

1.1. Vectors and Linear Combinations

Another description of this plane through \((0,0,0)\) is to know that \(v = (1,-1,1)\) is perpendicular to the plane. Section 1.2 will confirm that 90° angle by using dot products: \(v \cdot n = 0\) and \(w \cdot n = 0\).

1.1 A For \(v = (1,0)\) and \(w = (0,1)\), describe all points \(cv + dw\) with (1) whole numbers \(c\) and (2) nonnegative \(c \geq 0\). Then add all vectors \(dw\) and describe all \(cv + dw\).

Solution

1. The vectors \(cv = (c,0)\) with whole numbers \(c\) are equally spaced points along the \(x\) axis (the direction of \(v\)). They include \((-2,0), (-1,0), (0,0), (1,0), (2,0)\).

2. The vectors \(cv\) with \(c \geq 0\) fill a half-line. It is the positive \(x\) axis. This half-line starts at \((0,0)\) where \(c = 0\). It includes \((x,0)\) but not \((-x,0)\).

3. Adding all vectors \(dw = (0,d)\) puts a vertical line through those points \(cv\). We have infinitely many parallel lines from \((whole number\ c, any number\ d)\).

4. Adding all vectors \(dw\) puts a vertical line through every \(cv\) on the half-line. Now we have a half-plane. It is the right half of the \(x y\) plane (any \(x \geq 0\), any height \(y\)).

1.1 C Find two equations for the unknowns \(c\) and \(d\) so that the linear combination \(cv + dw\) equals the vector \(b\):

\[
\begin{bmatrix}
2 \\ -1
\end{bmatrix}
\]

Solution

In applying mathematics, many problems have two parts:

1. Modeling part. Express the problem by a set of equations.

2. Computational part. Solve those equations by a fast and accurate algorithm.

Here we are only asked for the first part (the equations). Chapter 2 is devoted to the second part (the algorithm). Our example fits into a fundamental model for linear algebra:

\[
\begin{bmatrix}
c_1 \\ c_2 \\ \vdots \\ c_n
\end{bmatrix}
\]

For \(n = 2\) we could find a formula for the \(c's\). The "elimination method" in Chapter 2 succeeds far beyond \(n = 100\). For \(n\) greater than 1 million, see Chapter 9. Here \(n = 2\).

Vector equation

\[
\begin{bmatrix}
2 \\ -1
\end{bmatrix}
\]

\[
\begin{bmatrix}
2 \\ -1
\end{bmatrix} + \begin{bmatrix}
-1 \\ 2
\end{bmatrix} = \begin{bmatrix}
1 \\ 0
\end{bmatrix}
\]

The required equations for \(c\) and \(d\) just come from the two components separately:

Two scalar equations

\[
2c - d = 1
\]

\[
-c + 2d = 0
\]

You could think of those as two lines that cross at the solution \(c = \frac{2}{3}, d = \frac{1}{3}\).
Chapter 1. Introduction to Vectors

Problem Set 1.1

Problems 1–9 are about addition of vectors and linear combinations.

1. Describe geometrically (line, plane, or all of \( \mathbb{R}^3 \)) all linear combinations of

   (a) \[
   \begin{bmatrix}
   1 \\
   2 \\
   3 \\
   \end{bmatrix}
   \text{ and }
   \begin{bmatrix}
   2 \\
   6 \\
   9 \\
   \end{bmatrix}
   \]

   (b) \[
   \begin{bmatrix}
   1 \\
   0 \\
   0 \\
   \end{bmatrix}
   \text{ and }
   \begin{bmatrix}
   0 \\
   2 \\
   3 \\
   \end{bmatrix}
   \]

   (c) \[
   \begin{bmatrix}
   2 \\
   0 \\
   2 \\
   \end{bmatrix}
   \text{ and }
   \begin{bmatrix}
   2 \\
   0 \\
   2 \\
   \end{bmatrix}
   \]

2. Draw \( v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \) and \( w = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \) and \( v + w \) and \( v - w \) in a single \( xy \) plane.

3. If \( v + w = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix} \) and \( v - w = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} \) compute and draw \( v \) and \( w \).

4. From \( v = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \) and \( w = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \), find the components of \( 3w + w \) and \( cv + dw \).

5. Compute \( u + v + w \) and \( 2u + 2v + w \). How do you know \( u, v, w \) lie in a plane?

   In a plane

   \[
   u = \begin{bmatrix}
   1 \\
   2 \\
   3 \\
   \end{bmatrix},
   \quad v = \begin{bmatrix}
   -3 \\
   1 \\
   -2 \\
   \end{bmatrix},
   \quad w = \begin{bmatrix}
   2 \\
   -3 \\
   -1 \\
   \end{bmatrix}.
   \]

6. Every combination of \( v = (1, -2, 1) \) and \( w = (0, 1, -1) \) has components that add to \( \_ \). Find \( c \) and \( d \) so that \( cv + dw = (3, 3, -6) \).

7. In the \( xy \) plane mark all nine of these linear combinations:

   \[
   c \begin{bmatrix}
   2 \\
   1 \\
   0 \\
   \end{bmatrix} + d \begin{bmatrix}
   1 \\
   0 \\
   1 \\
   \end{bmatrix}
   \quad \text{with} \quad c = 0, 1, 2 \quad \text{and} \quad d = 0, 1, 2.
   \]

8. The parallelogram in Figure 1.1 has diagonal \( v + w \). What is its other diagonal? What is the sum of the two diagonals? Draw that vector sum.

9. If there are three corners of a parallelogram are \( (1, 1), (4, 2), \) and \( (1, 3) \), what are all of the possible fourth corners? Draw two of them.

Problems 10–14 are about special vectors on cubes and clocks in Figure 1.4.

10. Which point of the cube is \( i + j \)? Which point is the vector sum of \( i = (1, 0, 0) \) and \( j = (0, 1, 0) \)? Describe all points \( (x, y, z) \) in the cube.

11. Four corners of the cube are \( (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1) \). What are the other four corners? Find the coordinates of the center point of the cube. The center points of the six faces are \( \_ \).

12. How many corners does a cube have in 4 dimensions? How many 3D faces? How many edges? A typical corner is \( (0, 0, 1, 0) \). A typical edge goes to \( (0, 1, 0, 0) \).

13. (a) What is the sum \( v + w \) of the twelve vectors that go from the center of a clock to the hours 1:00, 2:00, ..., 12:00?

(b) If the 2:00 vector is removed, why do the 11 remaining vectors add to 8:00?

(c) What are the components of that 2:00 vector \( v = (\cos \theta, \sin \theta) \)?

14. Suppose the twelve vectors start from 6:00 at the bottom instead of \( (0, 0, 0) \) at the center. The vector to 12:00 is doubled to \( (0, 2) \). Add the new twelve vectors.

Problems 15–19 go further with linear combinations of \( v \) and \( w \) (Figure 1.5a).

15. Figure 1.5a shows \( \frac{1}{2}v + \frac{1}{2}w \). Mark the points \( \frac{3}{2}v + \frac{1}{2}w \) and \( \frac{1}{2}v + \frac{1}{2}w \) and \( v + w \).

16. Mark the point \( -v + 2w \) and any other combination \( cv + dw \) with \( c + d = 1 \). Draw the line of all combinations that have \( c + d = 1 \).

17. Locate \( \frac{1}{2}v + \frac{3}{2}w \) and \( \frac{3}{2}v + \frac{1}{2}w \). The combinations \( cv + dw \) fill out what line?

18. Restricted by \( 0 \leq c \leq 1 \) and \( 0 \leq d \leq 1 \), shade in all combinations \( cv + dw \).

19. Restricted only by \( c \geq 0 \) and \( d \geq 0 \) draw the "cone" of all combinations \( cv + dw \).

Figure 1.5: Problems 15–19 in a plane

Problems 20–25 in 3-dimensional space
1.2 Lengths and Dot Products

The first section backed off from multiplying vectors. Now we go forward to define the "dot product" of \( \mathbf{v} \) and \( \mathbf{w} \). This multiplication involves the separate products \( v_1w_1 \) and \( v_2w_2 \), but it doesn't stop there. Those two numbers are added to produce the single number \( \mathbf{v} \cdot \mathbf{w} \). This is the geometry section (lengths and angles).

**Definition** The dot product or inner product of \( \mathbf{v} = (v_1, v_2) \) and \( \mathbf{w} = (w_1, w_2) \) is the number \( \mathbf{v} \cdot \mathbf{w} \).

\[
\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2.
\]  

**Example 1** The vectors \( \mathbf{v} = (4, 2) \) and \( \mathbf{w} = (-1, 2) \) have a zero dot product:

Dot product is zero

Perpendicular vectors

\[
\begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -4 + 4 = 0.
\]

In mathematics, zero is always a special number. For dot products, it means that these two vectors are perpendicular. The angle between them is 90°. When we drew them in Figure 1.1, we saw a rectangle (not just any parallelogram). The clearest example of perpendicular vectors is \( i = (1, 0) \) along the x axis and \( j = (0, 1) \) up the y axis. Again the dot product is \( i \cdot j = 0 + 0 = 0 \). Those vectors \( i \) and \( j \) form a right angle.

The dot product of \( \mathbf{v} = (1, 2) \) and \( \mathbf{w} = (3, 1) \) is 5. Soon \( \mathbf{v} \cdot \mathbf{w} \) will reveal the angle between \( \mathbf{v} \) and \( \mathbf{w} \) (not 90°). Please check that \( \mathbf{w} \cdot \mathbf{v} \) is also 5.

The dot product \( \mathbf{v} \cdot \mathbf{w} \) equals \( \mathbf{w} \cdot \mathbf{v} \). The order of \( \mathbf{v} \) and \( \mathbf{w} \) makes no difference.

**Example 2** Put a weight of 4 at the point \( x = -1 \) (left of zero) and 1 weight of 2 at the point \( x = 2 \) (right of zero). The x axis will balance on the center point (like a see-saw). The weights balance because the dot product is \((4)(-1) + (2)(2) = 0\).

This example is typical of engineering and science. The vector of weights is \( (w_1, w_2) = (4, 2) \). The vector of distances from the center is \( (u_1, u_2) = (-1, 2) \). The weights times the distances, \( w_1u_1 \) and \( w_2u_2 \), give the "moments". The equation for the see-saw to balance is \( w_1u_1 + w_2u_2 = 0 \).

**Example 3** Dot products enter in economics and business. We have three goods to buy and sell. Their prices are \( (p_1, p_2, p_3) \) for each unit—this is the "price vector" \( \mathbf{p} \). The quantities we buy or sell are \( (q_1, q_2, q_3) \)—positive when we sell, negative when we buy. Selling \( q_1 \) units at the price \( p_1 \) brings in \( q_1p_1 \). The total income (quantities \( q \) times prices \( p \)) is the dot product \( \mathbf{q} \cdot \mathbf{p} \) in three dimensions:

\[
\text{Income} = (q_1, q_2, q_3) \cdot (p_1, p_2, p_3) = q_1p_1 + q_2p_2 + q_3p_3 = \text{dot product}.
\]

A zero dot product means that "the books balance". Total sales equal total purchases if \( \mathbf{q} \cdot \mathbf{p} = 0 \). Then \( \mathbf{p} \) is perpendicular to \( \mathbf{q} \) (in three-dimensional space). A supermarket with thousands of goods goes quickly into high dimensions.
1.2. Lengths and Dot Products

An important case is the dot product of a vector with itself. In this case $\mathbf{v}$ equals $\mathbf{w}$. When the vector is $\mathbf{v} = (1, 2, 3)$, the dot product with itself is $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2 = 14$.

**Dot product** $\mathbf{v} \cdot \mathbf{v}$

**Length squared**

$$\|\mathbf{v}\|^2 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 + 4 + 9 = 14.$$  

Instead of a 90° angle between vectors we have $0°$. The answer is not zero because $\mathbf{v}$ is not perpendicular to itself. The dot product $\mathbf{v} \cdot \mathbf{v}$ gives the length of $\mathbf{v}$ squared.

**Definition** The length $\|\mathbf{v}\|$ of a vector $\mathbf{v}$ is the square root of $\mathbf{v} \cdot \mathbf{v}$:

$$\text{Length} = \|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$  

In two dimensions the length is $\sqrt{1^2 + 2^2}$. In three dimensions it is $\sqrt{1^2 + 2^2 + 3^2}$. By the calculation above, the length of $\mathbf{v} = (1, 2, 3)$ is $\|\mathbf{v}\| = \sqrt{14}$.

When $\mathbf{v} = \sqrt{\mathbf{v} \cdot \mathbf{v}}$ is just the ordinate length of the array that represents the vector.

In two dimensions, the arrow is in a plane. If the components are 1 and 2, the arrow is the third side of a right triangle (Figure 1.6). The Pythagoras formula $a^2 + b^2 = c^2$, which connects the three sides, is $1^2 + 2^2 = \|\mathbf{v}\|^2$.

For the length of $\mathbf{v} = (1, 2, 3)$, we used the right triangle twice. The vector $(1, 2, 0)$ in the base has length $\sqrt{5}$. This base vector is perpendicular to $(0, 0, 3)$ that goes straight up. So the diagonal of the box has length $\|\mathbf{v}\| = \sqrt{5 + 9} = \sqrt{14}$.

The length of a four-dimensional vector would be $\sqrt{1^2 + 2^2 + 3^2 + 3^2}$. Thus the vector $(1, 1, 1, 1)$ has length $\sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2$. This is the diagonal through a unit cube in four-dimensional space. The diagonal in $n$ dimensions has length $\sqrt{n}$.

The word "unit" is always indicating that some measurement equals "one". The unit price is the price for one item. A unit cube has sides of length one. A unit circle is a circle with radius one. Now we define the idea of a "unit vector".

**Definition** A unit vector $\mathbf{u}$ is a vector whose length equals one. Then $\mathbf{u} \cdot \mathbf{u} = 1$.

An example in four dimensions is $\mathbf{u} = \left( \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right)$. Then $\mathbf{u} \cdot \mathbf{u}$ is $\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$.

We divided $\mathbf{u} = (1, 1, 1, 1)$ by its length $\|\mathbf{u}\| = 2$ to get this unit vector.

**Example 4** The standard unit vectors along the $x$ and $y$ axes are written $\mathbf{i}$ and $\mathbf{j}$. In the $xy$ plane, the unit vector that makes an angle "theta" with the $x$ axis is $(\cos \theta, \sin \theta)$:

Unit vectors $\mathbf{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$.

When $\theta = 0$, the horizontal vector $\mathbf{u}$ is $\mathbf{i}$. When $\theta = 90°$ (or $\frac{\pi}{2}$ radians), the vertical vector is $\mathbf{j}$. At any angle, the components $\cos \theta$ and $\sin \theta$ produce $\mathbf{u} \cdot \mathbf{u} = 1$ because $\cos^2 \theta + \sin^2 \theta = 1$. These vectors reach out to the unit circle in Figure 1.7. Thus $\cos \theta$ and $\sin \theta$ are simply the coordinates of that point at angle $\theta$ on the unit circle.

Since $(2, 2, 1)$ has length 3, the vector $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ has length 1. Check that $\mathbf{u} \cdot \mathbf{u} = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = 1$. For a unit vector, divide any nonzero $\mathbf{v}$ by its length $\|\mathbf{v}\|$.

**Unit vector** $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ is a unit vector in the same direction as $\mathbf{v}$.

The coordinate vectors $\mathbf{i}$ and $\mathbf{j}$. The unit vector $\mathbf{u}$ = $(\cos \theta, \sin \theta)$ divides $\mathbf{v} = (1, 1)$ by its length $\|\mathbf{v}\| = \sqrt{2}$. The unit vector $\mathbf{u} = (\cos \theta, \sin \theta)$ is at angle $\theta$. 

Figure 1.6: The length $\sqrt{\mathbf{v} \cdot \mathbf{v}}$ of two-dimensional and three-dimensional vectors.

Figure 1.7: The coordinate vectors $\mathbf{i}$ and $\mathbf{j}$. The unit vector $\mathbf{u}$ at angle $45°$ (left) divides $\mathbf{v} = (1, 1)$ by its length $\|\mathbf{v}\| = \sqrt{2}$. The unit vector $\mathbf{u} = (\cos \theta, \sin \theta)$ is at angle $\theta$. 

The Angle Between Two Vectors

We stated that perpendicular vectors have \( \mathbf{v} \cdot \mathbf{w} = 0 \). The dot product is zero when the angle is 90°. To explain this, we have to connect angles to dot products. Then we show how \( \mathbf{v} \cdot \mathbf{w} \) finds the angle between any two nonzero vectors \( \mathbf{v} \) and \( \mathbf{w} \).

### Right angles

**The dot product is \( \mathbf{v} \cdot \mathbf{w} = 0 \) when \( \mathbf{v} \) is perpendicular to \( \mathbf{w} \).**

**Proof** When \( \mathbf{v} \) and \( \mathbf{w} \) are perpendicular, they form two sides of a right triangle. The third side is \( \mathbf{v} - \mathbf{w} \) (the hypotenuse going across in Figure 1.8). The Pythagorean Law for the sides of a right triangle is \( a^2 + b^2 = c^2 \).

**Perpendicular vectors:**

\[
\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{v} - \mathbf{w}\|^2
\]

Writing out the formulas for those lengths in two dimensions, this equation is

\[
(v_1^2 + v_2^2) + (w_1^2 + w_2^2) = (v_1 - w_1)^2 + (v_2 - w_2)^2.
\]

The right side begins with \( v_1^2 - 2v_1w_1 + w_1^2 \). Then \( v_1^2 \) and \( w_1^2 \) are on both sides of the equation and they cancel, leaving \(-2v_1w_1\). Also \( v_2^2 \) and \( w_2^2 \) cancel, leaving \(-2v_2w_2\). (In three dimensions there would be \(-2v_3w_3\).) Now divide by \(-2\):

\[
0 = -2v_1w_1 - 2v_2w_2 \quad \text{which leads to} \quad v_1w_1 + v_2w_2 = 0.
\]

**Conclusion** Right angles produce \( \mathbf{v} \cdot \mathbf{w} = 0 \). The dot product is zero when the angle is \( \theta = 90^\circ \). Then \( \cos \theta = 0 \). The zero vector \( \mathbf{v} = \theta = 0 \) is perpendicular to every vector \( \mathbf{w} \) because \( \mathbf{0} \cdot \mathbf{w} \) is always zero.

Now suppose \( \mathbf{v} \cdot \mathbf{w} \) is not zero. It may be positive, it may be negative. The sign of \( \mathbf{v} \cdot \mathbf{w} \) immediately tells whether we are below or above a right angle. The angle is less than 90° when \( \mathbf{v} \cdot \mathbf{w} \) is positive. The angle is above 90° when \( \mathbf{v} \cdot \mathbf{w} \) is negative. The right side of Figure 1.8 shows a typical vector \( \mathbf{v} = (3, 1) \). The angle with \( \mathbf{w} = (1, 3) \) is less than 90° because \( \mathbf{v} \cdot \mathbf{w} = 6 \) is positive.

![Diagram of perpendicular vectors](image)

**Figure 1.8:** Perpendicular vectors have \( \mathbf{v} \cdot \mathbf{w} = 0 \). Then \( \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{v} - \mathbf{w}\|^2 \).

### 1.2. Lengths and Dot Products

The borderline is where vectors are perpendicular to \( \mathbf{v} \). On that dividing line between plus and minus, \( (1, -3) \) is perpendicular to \( (3, 1) \). The dot product is zero.

The dot product reveals the exact angle \( \theta \). This is not necessary for linear algebra—you could stop here! Once we have matrices, we won’t come back to \( \theta \). But while we are on the subject of angles, this is the place for the formula.

Start with unit vectors \( \mathbf{u} \) and \( \mathbf{U} \). The sign of \( \mathbf{u} \cdot \mathbf{U} \) tells whether \( \theta < 90^\circ \) or \( \theta > 90^\circ \). Because the vectors have length 1, we learn more than that. **The dot product \( \mathbf{u} \cdot \mathbf{U} \) is the cosine of \( \theta \).** This is true in any number of dimensions.

**Unit vectors \( \mathbf{u} \) and \( \mathbf{U} \) at angle \( \theta \) have**

\[ \mathbf{u} \cdot \mathbf{U} = \cos \theta. \]

**Certainly** \( |\mathbf{u} \cdot \mathbf{U}| \leq 1 \).

Remember that \( \cos \theta \) is never greater than 1. It is never less than \(-1\). **The dot product of unit vectors is between \(-1\) and \(1\).**

Figure 1.9 shows this clearly when the vectors are \( \mathbf{u} = (\cos \theta, \sin \theta) \) and \( \mathbf{i} = (1, 0) \). The dot product is \( \mathbf{u} \cdot \mathbf{i} = \cos \theta \). That is the cosine of the angle between them.

After rotation through any angle \( \alpha \), these are still unit vectors. The vector \( \mathbf{i} \) rotates to \( (\cos \alpha, \sin \alpha) \). The vector \( \mathbf{u} \) rotates to \( (\cos \beta, \sin \beta) \) with \( \beta = \alpha + \theta \). Their dot product is \( \cos \alpha \cos \beta + \sin \alpha \sin \beta \). From trigonometry this is the same as \( \cos(\beta - \alpha) \). But \( \beta - \alpha \) is the angle \( \theta \), so the dot product is \( \cos \theta \).

![Diagram of dot product](image)

**Figure 1.9:** The dot product of unit vectors is the cosine of the angle \( \theta \).

**Problem 24** proves \( |\mathbf{u} \cdot \mathbf{U}| \leq 1 \) directly, without mentioning angles. The inequality and the cosine formula \( \mathbf{u} \cdot \mathbf{U} = \cos \theta \) are always true for unit vectors.

**What if \( \mathbf{u} \) and \( \mathbf{w} \) are not unit vectors?** Divide by their lengths to get \( \mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} \) and \( \mathbf{U} = \frac{\mathbf{w}}{\|\mathbf{w}\|} \). Then the dot product of those unit vectors \( \mathbf{u} \) and \( \mathbf{U} \) gives \( \cos \theta \).

**Cosine Formula** If \( \mathbf{v} \) and \( \mathbf{w} \) are nonzero vectors then

\[ \mathbf{v} \cdot \mathbf{w} = \cos \theta. \]
Chapter 1. Introduction to Vectors

Whatever the angle, this dot product of \( \|v\| \) with \( w/\|w\| \) never exceeds one. That is the “Schwartz inequality” \( |v \cdot w| \leq \|v\| \|w\| \) for dot products—or more correctly the Cauchy-Schwarz-Bunyakovsky inequality. It was found in France and Germany and Russia (and maybe elsewhere)—it is the most important inequality in mathematics.

Since \( |\cos \theta| \) never exceeds 1, the cosine formula gives two great inequalities:

\[
\text{Schwarz Inequality} \quad |v \cdot w| \leq \|v\| \|w\|
\]

\[
\text{Triangle Inequality} \quad \|v + w\| \leq \|v\| + \|w\|
\]

**Example 5** Find \( \cos \theta \) for \( v = \left[ \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right] \) and \( w = \left[ \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right] \) and check both inequalities.

**Solution** The dot product is \( v \cdot w = 4 \). Both \( v \) and \( w \) have length \( \sqrt{3} \). The cosine is \( 4/3 \).

\[
\cos \theta = \frac{v \cdot w}{\|v\| \|w\|} = \frac{4}{\sqrt{3}} = \frac{4}{3}.
\]

The angle is below \( 90^\circ \) because \( v \cdot w = 4 \) is positive. By the Schwarz inequality, \( v \cdot w = 4 \) is less than \( \|v\| \|w\| = 5 \). Side 3 = \( \|v + w\| \) is less than side 1 + side 2, by the triangle inequality. For \( v + w = (3, 3) \) that says \( \sqrt{18} < \sqrt{5} + \sqrt{5} \). Square this to get 18 < 20.

**Example 6** The dot product of \( v = (a, b) \) and \( w = (b, a) \) is \( 2ab \). Both lengths are \( \sqrt{a^2 + b^2} \). The Schwarz inequality in this case says that \( 2ab \leq a^2 + b^2 \).

This is more famous if we write \( x = a^2 \) and \( y = b^2 \). The “geometric mean” \( \sqrt{xy} \) is not larger than the “arithmetic mean” = average \( \frac{1}{2}(x + y) \).

\[
\text{Geometric mean} \quad ab \leq \frac{a^2 + b^2}{2} \quad \text{becomes} \quad \sqrt{xy} \leq \frac{x + y}{2}.
\]

Example 5 had \( a = 2 \) and \( b = 1 \). So \( x = 4 \) and \( y = 1 \). The geometric mean \( \sqrt{xy} = 2 \) is below the arithmetic mean \( \frac{1}{2}(1 + 4) = 2.5 \).

**Notes on Computing**

Write the components of \( v \) as \( v(1), \ldots, v(N) \) and similarly for \( w \). In FORTRAN, the sum \( v + w \) requires a loop to add components separately. The dot product also uses a loop \( w \) add the separate \( v(j)w(j) \). Here are VPLUSW and VDOTW:

FORTRAN

```
DO 10 J = 1,N
10 VPLUSW(J) = V(J) + W(J)
```

MATLAB and also PYTHON work directly with whole vectors, not their components. No loop is needed. When \( v \) and \( w \) have been defined, \( v + w \) is immediately understood.

1.2. Lengths and Dot Products

Input \( v \) and \( w \) as rows—the prime \(^t\) transposes them to columns. \( 2v + 3w \) uses \( * \) for multiplication by 2 and 3. The result will be printed unless the line ends in a semicolon.

MATLAB

```
   \text{MATLAB} \quad v = [2 \ 3 \ 4]^{\prime}; \quad w = [1 \ 1 \ 1]^{\prime}; \quad u = 2v + 3w
```

The dot product \( v \cdot w \) is usually seen as \textit{a row times a column} (with no dot):

```
   \quad \text{Instead of} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \text{we more often see} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \text{or} \quad v^\prime \cdot w
```

The length of \( v \) is known to MATLAB as norm(\( v \)). We could define it ourselves as \( \sqrt{v^\prime \cdot v} \), using the square root function—also known. The cosine we have to define ourselves! The angle (in radians) comes from the \textit{arc cosine} (acos) function:

```
   \quad \text{Cosine formula} \quad \cos \theta = v^\prime \cdot w / (\text{norm}(v) \cdot \text{norm}(w))
   \quad \text{Angle formula} \quad \theta = \text{acos} (\cos \theta)
```

An M-file would create a new function \( \text{cosine}(v, w) \) for future use. The M-files created especially for this book are listed at the end. R and PYTHON are open source software.

**REVIEW OF THE KEY IDEAS**

1. The dot product \( v \cdot w \) multiplies each component \( v_j \) by \( w_j \) and adds all \( v_jw_j \).
2. The length \( \|v\| \) of \( v \) is the square root of \( v \cdot v \).
3. \( u = v/\|v\| \) is a \textit{unit vector}. Its length is 1.
4. The dot product is \( v \cdot w = 0 \) when vectors \( v \) and \( w \) are perpendicular.
5. The cosine of \( \theta \) (the angle between any nonzero \( v \) and \( w \)) never exceeds \( 1 \):

\[
\cos \theta = \frac{v \cdot w}{\|v\| \|w\|} \quad \text{Schwarz inequality} \quad |v \cdot w| \leq \|v\| \|w\|.
\]

Problem 21 will produce the \textit{triangle inequality} \( \|v + w\| \leq \|v\| + \|w\| \).

**WORKED EXAMPLES**

1.2 A. For the vectors \( v = (3, 4) \) and \( w = (4, 3) \), test the Schwarz inequality on \( v \cdot w \) and the triangle inequality on \( \|v + w\| \). Find \( \cos \theta \) for the angle between \( v \) and \( w \). When will we have equality \( |v \cdot w| = \|v\| \|w\| \) and \( \|v + w\| = \|v\| + \|w\| \)?
1.2. Lengths and Dot Products

Problem Set 1.2

1. Calculate the dot products $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{u} \cdot \mathbf{w}$ and $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$ and $\mathbf{w} \cdot \mathbf{v}$:

$$\mathbf{u} = \begin{pmatrix} -6 \\ 3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 8 \\ 6 \end{pmatrix}$$

2. Compute the lengths $||\mathbf{u}||$ and $||\mathbf{v}||$ and $||\mathbf{w}||$ of those vectors. Check the Schwarz inequalities $||\mathbf{u} \cdot \mathbf{v}|| \leq ||\mathbf{u}|| ||\mathbf{v}||$ and $||\mathbf{v} \cdot \mathbf{w}|| \leq ||\mathbf{v}|| ||\mathbf{w}||$.

3. Find unit vectors in the directions of $\mathbf{v}$ and $\mathbf{w}$ in Problem 1, and the cosine of the angle between $\mathbf{u}$ and $\mathbf{v}$ refuse. Choose vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ that make $0^\circ$, $90^\circ$, and $180^\circ$ angles with $\mathbf{v}$.

4. For all unit vectors $\mathbf{v}$ and $\mathbf{w}$, find the dot products (actual numbers) of

(a) $\mathbf{v} \cdot -\mathbf{v}$
(b) $\mathbf{v} \cdot \mathbf{v} - \mathbf{w}$
(c) $\mathbf{v} - \mathbf{w}$

5. Find unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ in the directions of $\mathbf{v} = (3, 1)$ and $\mathbf{w} = (2, 1, 2).$

6. Find unit vectors $\mathbf{U}_1$ and $\mathbf{U}_2$ that are perpendicular to $\mathbf{u}_1$ and $\mathbf{u}_2$.

7. (a) Describe every vector $\mathbf{w} = (w_1, w_2)$ that is perpendicular to $\mathbf{v} = (2, -1)$.
(b) The vectors that are perpendicular to $\mathbf{V} = (1, 1, 1)$ lie on a _______.
(c) The vectors that are perpendicular to $(1, 1, 1)$ and $(1, 2, 3)$ lie on a _______.

8. Find the angle $\theta$ (from its cosine) between these pairs of vectors:

(a) $\mathbf{v} = \left[ \frac{1}{\sqrt{3}} \right]$ and $\mathbf{w} = \left[ \frac{1}{0} \right]$ (b) $\mathbf{v} = \left[ \frac{2}{2} \right]$ and $\mathbf{w} = \left[ \frac{2}{2} \right]$
(c) $\mathbf{v} = \left[ \frac{1}{\sqrt{3}} \right]$ and $\mathbf{w} = \left[ \frac{-1}{1} \right]$ (d) $\mathbf{v} = \left[ \frac{1}{3} \right]$ and $\mathbf{w} = \left[ \frac{-1}{-2} \right]$

9. True or false (give a reason or a counterexample if false).

(a) If $\mathbf{u}$ is perpendicular (in three dimensions) to $\mathbf{v}$ and $\mathbf{w}$, those vectors $\mathbf{v}$ and $\mathbf{w}$ are parallel.
(b) If $\mathbf{u}$ is perpendicular to $\mathbf{v}$ and $\mathbf{w}$, then $\mathbf{u}$ is perpendicular to $\mathbf{v} + \mathbf{w}$.
(c) If $\mathbf{u}$ and $\mathbf{v}$ are perpendicular unit vectors then $||\mathbf{u} \cdot \mathbf{v}|| = 1/2$.

10. The slopes of the arrows from $(0,0)$ to $(v_1, v_2)$ and $(w_1, w_2)$ are $v_2/v_1$ and $w_2/w_1$. Suppose the product $v_1 w_2 - v_2 w_1$ of those slopes is $-1$. Show that $\mathbf{v} \cdot \mathbf{w} = 0$ and the vectors are perpendicular.

11. Draw arrows from $(0,0)$ to the points $\mathbf{v} = (1, 2)$ and $\mathbf{w} = (-2, 1)$. Multiply their slopes. That answer is a signal that $\mathbf{v} \cdot \mathbf{w} = 0$ and the arrows are _______.

11. If $\mathbf{v} \cdot \mathbf{w}$ is negative, what does this say about the angle between $\mathbf{v}$ and $\mathbf{w}$? Draw a 3-dimensional vector $\mathbf{v}$ (an arrow), and show where to find all $\mathbf{w}$'s with $\mathbf{v} \cdot \mathbf{w} < 0$. 

Chapter 1. Introduction to Vectors

12. With \( v = (1, 1) \) and \( w = (1, 5) \) choose a number \( c \) so that \( w - cv \) is perpendicular to \( v \). Then find the formula that gives this number \( c \) for any nonzero \( v \) and \( w \). (Note: \( cv \) is the "projection" of \( w \) onto \( v \).)

13. Find two vectors \( u \) and \( w \) that are perpendicular to \((1, 0, 1)\) and to each other.

14. Find nonzero vectors \( u \), \( v \), \( w \) that are perpendicular to \((1, 1, 1)\) and to each other.

15. The geometric mean of \( x = 2 \) and \( y = 8 \) is \( \sqrt{xy} = 4 \). The arithmetic mean is larger: \( \frac{1}{2}(x + y) = \). This would come in Example 6 from the Schwarz inequality for \( v = (\sqrt{2}, \sqrt{8}) \) and \( w = (\sqrt{8}, \sqrt{2}) \). Find cos \( \theta \) for this \( v \) and \( w \).

16. How long is the vector \( v = (1, 1, \ldots, 1) \) in 9 dimensions? Find a unit vector \( u \) in the same direction as \( v \) and a unit vector \( w \) that is perpendicular to \( v \).

17. What are the cosines of the angles \( \alpha, \beta, \theta \) between the vector \((1, 0, -1)\) and the unit vectors \( \mathbf{i}, \mathbf{j}, \mathbf{k} \) along the axes? Check the formula \( \cos^2 \alpha + \cos^2 \beta + \cos^2 \theta = 1 \).

Problems 18–31 lead to the main facts about lengths and angles in triangles.

18. The parallelogram with sides \( v = (4, 2) \) and \( w = (-1, 2) \) is a rectangle. Check the Pythagoras formula \( a^2 + b^2 = c^2 \) which is for \textit{right triangles only}:

\[
\text{length of } v + \text{length of } w = \text{length of } v + w.
\]

19. (Rules for dot products) These equations are simple but useful:

(1) \( v \cdot w = w \cdot v \)

(2) \( u \cdot (v + w) = u \cdot v + u \cdot w \)

(3) \( c(v \cdot w) = (cv) \cdot w = c(v \cdot w) \)

Use (2) with \( u = v + w \) to prove \( \|v + w\|^2 = v \cdot v + 2v \cdot w + w \cdot w \).

20. The "Law of Cosines" comes from \( (v - w) \cdot (v - w) = v \cdot v - 2v \cdot w + w \cdot w \). Use this to prove:

\[
\text{Cosine Law: } \|v - w\|^2 = \|v\|^2 - 2v \cdot w + \|w\|^2.
\]

If \( \theta < 90^\circ \) show that \( \|v\|^2 + \|w\|^2 \) is larger than \( \|v - w\|^2 \) (the third side).

21. The \textit{triangle inequality} says: (length of \( v + w \) \( \leq \) (length of \( v \) + (length of \( w \)).

Problem 19 found \( \|v + w\|^2 = \|v\|^2 + 2v \cdot w + \|w\|^2 \). Use the Schwarz inequality \( v \cdot w \leq \|v\| \|w\| \) to show that \( \text{side 1} \) can not exceed \( \|v\| \|w\| \).

\[
\text{Triangle Inequality: } \|v + w\|^2 \leq (\|v\| + \|w\|)^2 \text{ or } \|v + w\| \leq \|v\| + \|w\|.
\]

22. The Schwarz inequality \( v \cdot w \leq \|v\| \|w\| \) by algebra instead of trigonometry:

(a) Multiply out both sides of \((v_1 w_1 + v_2 w_2)^2 \leq (v_1^2 + v_2^2)(w_1^2 + w_2^2) \).

(b) Show that the difference between those two sides equals \((v_1 w_2 - v_2 w_1)^2 \). This cannot be negative since it is a square—so the inequality is true.

23. The figure shows that \( \cos \alpha = v_1 /\|v\| \) and \( \sin \alpha = v_2 /\|v\| \). Similarly \( \cos \beta = w_1 /\|w\| \) and \( \sin \beta = w_2 /\|w\| \). The angle \( \theta \) is \( \beta - \alpha \). Substitute into the trigonometric formula \( \cos(\beta - \alpha) = \cos \beta \cos \alpha + \sin \beta \sin \alpha \) to find \( \cos \theta = v \cdot w /\|v\|\|w\| \).

24. One-line proof of the Schwarz inequality \( \|v \cdot U\| \leq 1 \) for unit vectors:

\[
\|v \cdot U\| \leq \|v\| \|U_1\| + \|U_2\| \|U_2\| \leq \frac{v_1^2 + U_1^2}{2} + \frac{v_2^2 + U_2^2}{2} = 1 + \frac{1}{2} = 1.
\]

Put \((U_1, U_2) = (6, 8) \) and \((U_1, U_2) = (8, 6) \) in that whole line and find cos \( \theta \).

25. Why is \( \cos \theta \) never greater than 1 in the first place?

26. If \( v = (1, 2) \) draw all vectors \( w = (x, y) \) in the xy plane with \( v \cdot w = x + 2y = 5 \). Which is the shortest \( w \)?

27. (Recommended) If \( \|v\| = 5 \) and \( \|w\| = 3 \), what are the smallest and largest values of \( \|v - w\| \)? What are the smallest and largest values of \( v \cdot w \)?

Challenge Problems

28. Can three vectors in the xy plane have \( u \cdot v < 0 \) and \( u \cdot w < 0 \) and \( u \cdot w < 0 \)? I don’t know how many vectors in xyz space can have all negative dot products. (Four of those vectors in the plane would certainly be impossible . . .)

29. Pick any numbers that add to \( x + y + z = 0 \). Find the angle between your vector \( v = (x, y, z) \) and the vector \( w = (z, x, y) \). Challenge question: Explain why \( v \cdot w /\|v\|\|w\| \) is always \( -1 \).

30. How could you prove \( \sqrt{xy} \leq \frac{1}{2}(x + y) \) (geometric mean \( \leq \) arithmetic mean) \?

31. Find four perpendicular unit vectors with all components equal to \( \frac{1}{2} \) or \( -\frac{1}{2} \).

32. Using \( v = \text{randn}(3, 1) \) in MATLAB, create a random unit vector \( u = v /\|v\| \). Using \( V = \text{randn}(3, 30) \) create 30 more random unit vectors \( U \). What is the average size of the dot products \( \|u \cdot U\| \)? In calculus, the average \( \int_0^\pi \cos \theta d\theta / \pi = 2 / \pi \).
1.3 Matrices

This section is based on two carefully chosen examples. They both start with three vectors. I will take their combinations using matrices. The three vectors in the first example are \( u, v, \) and \( w \):

First example

\[
\begin{bmatrix}
1 \\
-1 \\
0
\end{bmatrix}
\quad
\begin{bmatrix}
0 \\
0 \\
1
\end{bmatrix}
\quad
\begin{bmatrix}
0 \\
1 \\
1
\end{bmatrix}
\]

Their linear combinations in three-dimensional space are \( cu + dv - ew \):

\[
\begin{bmatrix}
1 \\
0 \\
-1
\end{bmatrix}
+ \begin{bmatrix}
0 \\
1 \\
0
\end{bmatrix}
- \begin{bmatrix}
0 \\
0 \\
1
\end{bmatrix}
= \begin{bmatrix}
c \\
d \\
e - d
\end{bmatrix}.
\]

Now something important: Rewrite that combination using a matrix. The vectors \( u, v, w \) go into the columns of the matrix \( A \). That matrix "multiplies" a vector.

\[
\begin{align*}
\text{Combination} & \quad \text{is now } A \text{ times } x \\
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
-1 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
c \\
d \\
e - d
\end{bmatrix}
& = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
-1 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
\end{align*}
\]

The numbers \( c, d, e \) are the components of a vector \( x \). The matrix \( A \) times the vector \( x \) is the same as the combination \( cu + dv - ew \) of the three columns:

\[
\begin{bmatrix}
1 \\
0 \\
-1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= \begin{bmatrix}
x_1 - x_3 \\
x_2 - x_1 \\
x_3 - x_2
\end{bmatrix}
\]

This is more than a definition of \( Ax \), because the rewriting brings a crucial change in viewpoint. At first, the numbers \( c, d, e \) were multiplying the vectors. Now the matrix is multiplying those numbers. The matrix \( A \) acts on the vector \( x \). The result \( Ax \) is a combination \( b \) of the columns of \( A \).

To see that action, I will write \( x_1, x_2, x_3 \) instead of \( c, d, e \). I will write \( b_1, b_2, b_3 \) for the components of \( b \). With new letters we see

\[
\begin{bmatrix}
1 \\
0 \\
-1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix} = b.
\]

The input is \( x \) and the output is \( b = Ax \). This \( A \) is a "difference matrix" because \( b \) contains differences of the input vector \( x \). The top difference is \( x_1 - x_0 = x_1 - 0 \).

Here is an example to show differences of numbers (squares in \( x \), odd numbers in \( b \)):

\[
x = \begin{bmatrix}
1 \\
4 \\
9
\end{bmatrix} = \text{squares} \quad Ax = \begin{bmatrix}
1 - 0 \\
4 - 1 \\
9 - 4
\end{bmatrix} = \begin{bmatrix}
1 \\
3 \\
5
\end{bmatrix} = b.
\]

That pattern would continue for a 4 by 4 difference matrix. The next square would be \( x_4 = 16 \). The next difference would be \( x_4 - x_3 = 16 - 9 = 7 \) (this is the next odd number). The matrix finds all the differences at once.

**Important Note.** You may already have learned about multiplying \( Ax \), a matrix times a vector. Probably it was explained differently, using the rows instead of the columns. The usual way takes the dot product of each row with \( x \):

\[
\text{Dot products with rows} \quad Ax = \begin{bmatrix}
1 & 0 & 0 \\
-1 & 1 & 0 \\
0 & -1 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= \begin{bmatrix}
(1, 0, 0) \cdot (x_1, x_2, x_3) \\
(-1, 1, 0) \cdot (x_1, x_2, x_3) \\
(0, -1, 1) \cdot (x_1, x_2, x_3)
\end{bmatrix}
\]

Those dot products are the same \( x_1 \) and \( x_3 - x_1 \) and \( x_3 - x_2 \) that we wrote in equation (4). The new way is to work with \( Ax \) a column at a time. Linear combinations are the key to linear algebra, and the output \( Ax \) is a linear combination of the columns of \( A \).

With numbers, you can multiply \( Ax \) either way (I admit to using rows). With letters, combinations are the good way. Chapter 2 will repeat these rules of matrix multiplication, and explain the underlying ideas. There we will multiply matrices both ways.

### Linear Equations

One more change in viewpoint is crucial. Up to now, the numbers \( x_1, x_2, x_3 \) were known (called \( c, d, e \) at first). The right hand side \( b \) was not known. We found that vector of differences by multiplying \( Ax \). Now we think of \( b \) as known and we look for \( x \).

**Old question:** Compute the linear combination \( x_1 u + x_2 v + x_3 w \) to find \( b \).

**New question:** Which combination of \( u, v, w \) produces a particular vector \( b \)?

This is the inverse problem—to find the input \( x \) that gives the desired output \( b \) = \( Ax \). You have seen this before, as a system of linear equations for \( x_1, x_2, x_3 \). The right hand sides of the equations are \( b_1, b_2, b_3 \). We can solve that system to find \( x_1, x_2, x_3 \):

\[
Ax = b \quad \begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix} \quad \text{Solution} \quad \begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\]

Let me admit right away—most linear systems are not so easy to solve. In this example, the first equation decided \( x_1 = b_1 \). Then the second equation produced \( x_2 = b_1 + b_2 \). The equations could be solved in order (top to bottom) because the matrix \( A \) was selected to be lower triangular.
Chapter 1. Introduction to Vectors

Look at two specific choices 0, 0, 0 and 1, 3, 5 of the right sides \( b_1, b_2, b_3 \):\[ b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ gives } x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \text{ gives } x = \begin{bmatrix} 1 + 3 + 5 \\ 4 \\ 9 \end{bmatrix}. \]
The first solution (all zeros) is more important than it looks. In words: if the output is \( b = 0 \), then the input must be \( x = 0 \). That statement is true for this matrix \( A \). It is not true for all matrices. Our second example will show (for a different matrix \( C \)) how we can have \( Cx = 0 \) when \( C \neq 0 \) and \( x \neq 0 \).

This matrix \( A \) is "invertible". From \( b \) we can recover \( x \).

### The Inverse Matrix

Let me repeat the solution \( x \) in equation (6). A sum matrix will appear!

\[ Ax = b \text{ solved by } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 + b_3 \\ b_1 + b_2 + b_3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} b_1 \begin{bmatrix} b_2 \\ b_3 \end{bmatrix}. \quad (7) \]

If the differences of the \( x \)'s are the \( b \)'s, the sums of the \( b \)'s are the \( x \)'s. That was true for the odd numbers \( b = (1, 3, 5) \) and the squares \( x = (1, 4, 9) \). It is true for all vectors.

The sum matrix \( S \) in equation (7) is the inverse of the difference matrix \( A \).

**Example:** The differences of \( x = (1, 2, 3) \) are \( b = (1, 1, 1) \). So \( b = Ax \) and \( x = Sb \):

\[ Ax = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ and } Sb = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \]

Equation (7) for the solution vector \( x = (x_1, x_2, x_3) \) tells us two important facts:

1. For every \( b \) there is one solution to \( Ax = b \).
2. A matrix \( S \) produces \( x = Sb \).

The next chapters ask about other equations \( Ax = b \). Is there a solution? How is it computed? In linear algebra, the notation for the "inverse matrix" is \( A^{-1} \):

\[ Ax = b \text{ is solved by } x = A^{-1}b = Sb. \]

**Note on calculus.** Let me connect these special matrices \( A \) and \( S \) to calculus. The vector \( x \) changes to a function \( x(t) \). The differences \( Ax \) become the derivative \( dx/dt = b(t) \). In the inverse direction, the sum \( Sb \) becomes the integral of \( b(t) \). The Fundamental Theorem of Calculus says that integration \( S \) is the inverse of differentiation \( A \):

\[ Ax = b \text{ and } x = Sb \quad \frac{dx}{dt} = b \text{ and } x(t) = \int_0^t b. \quad (8) \]

### 1.3. Matrices

The derivative of distance traveled \( x(t) \) is the velocity \( b(t) \). The integral of \( b(t) \) is the distance \( x(t) \). Instead of adding \( 4C \), I measured the distance from \( x(0) = 0 \). In the same way, the differences started at \( x_0 = 0 \). This zero start makes the pattern complete, when we write \( x_1 - x_0 \) for the first component of \( Ax \) (we just wrote \( x_1 \)).

Notice another analogy with calculus. The differences of squares \( 0, 1, 4, 9 \) are odd numbers \( 1, 3, 5 \). The derivative of \( x(t) = t^2 \) is \( 2t \). A perfect analogy would have produced the even numbers \( b = 2, 4, 6 \) at times \( t = 1, 2, 3 \). But differences are not the same as derivatives, and our matrix \( A \) produces not \( 2t \) but \( 2t - 1 \) (these one-sided "backward differences" are centered at \( t = 1/2 \)):

\[ x(t) - x(t - 1) = t^2 - (t - 1)^2 = t^2 - (t^2 - 2t + 1) = 2t - 1. \quad (9) \]

The Problem Set will follow up to show that "forward differences" produce \( 2t + 1 \). A better choice (not always seen in calculus courses) is a **centered difference** that uses \( x(t + 1) - x(t - 1) \). Divide \( \Delta x \) by the distance \( \Delta t \) from \( t \) to \( t + 1 \), which is 2:

\[ \text{Centered difference of } x(t) = t^2 \Rightarrow \frac{(t + 1)^2 - (t - 1)^2}{2} = 2t \text{ exactly.} \quad (10) \]

Centered differences are great. Centered is best. Our second example is not invertible.

### Cyclic Differences

This example keeps the same columns \( u \) and \( v \) but changes \( w \) to a new vector \( w^* \):

**Second example** \( u = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \quad w^* = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \).

Now the linear combinations of \( u, v, w^* \) lead to a cyclic difference matrix \( C \):

**Cyclic** \( Cx = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} x_1 \quad \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} x_2 \quad \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} x_3 = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = b. \quad (11) \)

This matrix \( C \) is not triangular. It is not so simple to solve for \( x \) when we are given \( b \). Actually it is impossible to find \( x \) when \( Cx = 0 \), because the three equations either have **infinitely many solutions** or else **no solution**:

\[ Cx = 0 \quad \text{Infinitely many } x \quad \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} \text{ is solved by all vectors } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} c \\ c \\ c \end{bmatrix}. \quad (12) \]
Every constant vector \((c, c, c)\) has zero differences when we go cyclically. This undetermined constant \(c\) is like the \(+C\) that we add to integrals. The cyclic differences have \(x_1 - x_3\) in the first component, instead of starting from \(x_0 = 0\).

The other very likely possibility for \(C x = b\) is no solution at all:

\[
C x = b \quad \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{Left sides add to 0} \\
\begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_1 - x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \quad \text{Right sides add to 9} \\
\text{No solution} \\
x_1, x_2, x_3
\] (13)

Look at this example geometrically. No combination of \(u, v, w\), and \(w^*\) will produce the vector \(b = (1, 3, 5)\). The combinations don’t fill the whole three-dimensional space. The right sides must have \(b_1 + b_2 + b_3 = 0\) to allow a solution to \(C x = b\), because the left sides \(x_1 - x_3, x_2 - x_1, \text{ and } x_3 - x_2\) always add to zero.

Put that in different words. All linear combinations \(x_1 u + x_2 v + x_3 w^* = b\) lie on the plane given by \(b_1 + b_2 + b_3 = 0\). This subject is suddenly connecting algebra with geometry. Linear combinations can fill all of space, or only a plane. We need a picture to show the crucial difference between \(u, v, w\) (the first example) and \(u, v, w^*\).

Figure 1.10: Independent vectors \(u, v, w\). Dependent vectors \(u, v, w^*\) in a plane.

**Independence and Dependence**

Figure 1.10 shows those column vectors, first of the matrix \(A\) and then of \(C\). The first two columns \(u\) and \(v\) are the same in both pictures. If we only look at the combinations of those two vectors, we get a two-dimensional plane. The key question is whether the third vector is in that plane:

**Independence** \(w\) is not in the plane of \(u\) and \(v\).

**Dependence** \(w^*\) is in the plane of \(u\) and \(v\).

The important point is that the new vector \(w^*\) is a linear combination of \(u\) and \(v\):

\[
u + v + w^* = 0 \\
w^* = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = -u - v. \tag{14}
\]

1.3 Matrices

All three vectors \(u, v, w^*\) have components adding to zero. Then all their combinations will have \(b_1 + b_2 + b_3 = 0\) (as we saw above, by adding the three equations). This is the equation for the plane containing all combinations of \(u\) and \(v\). By including \(w^*\) we get no new vectors because \(w^*\) is already on that plane.

The original \(w = (0, 0, 1)\) is not on the plane: \(0 + 0 + 1 \neq 0\). The combinations of \(u, v, w\) fill the whole three-dimensional space. We know this already, because the solution \(x = 5b\) in equation (6) gave the right combination to produce any \(b\).

The two matrices \(A\) and \(C\), with third columns \(w\) and \(w^*\), allowed me to mention two key words of linear algebra: independence and dependence. The first half of the course will develop these ideas much further—I am happy if you see them early in the two examples:

**\(u, v, w\)** are independent. No combination except \(0u + 0v + 0w = 0\) gives \(b = 0\).

**\(u, v, w^*\)** are dependent. Other combinations (specifically \(u + v + w^*\)) give \(b = 0\).

You can picture this in three dimensions. The three vectors lie in a plane or they don’t. Chapter 2 has \(n\) vectors in \(n\)-dimensional space. Independence or dependence is the key point. The vectors go into the columns of an \(n \times n\) matrix.

Independent columns: \(Ax = 0\) has one solution. \(A\) is an invertible matrix.

Dependent columns: \(Ax = 0\) has many solutions. \(A\) is a singular matrix.

Eventually we will have \(n\) vectors in \(m\)-dimensional space. The matrix \(A\) with those \(n\) columns is now rectangular (\(m\) by \(n\)). Understanding \(Ax = w\) is the problem of Chapter 3.

### REVIEW OF THE KEY IDEAS

1. **Matrix times vector**: \(Ax = \text{combination of the columns of } A\).
2. The solution to \(Ax = b\) is \(x = A^{-1}b\), when \(A\) is an invertible matrix.
3. The difference matrix \(A\) is inverted by the sum matrix \(S = A^{-1}\).
4. The cyclic matrix \(C\) has no inverse. Its three columns lie in the same plane.

5. This section is looking ahead to key ideas, not fully explained yet.

### WORKED EXAMPLES

1.3 A

Charge the southwest entry \(a_{23}\) of \(A\) (row 3, column 1) to \(a_{31} = 1\):

\[
Ax = b \quad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}
\]

Find the solution \(x\) for any \(b\). From \(x = A^{-1}b\) read off the inverse matrix \(A^{-1}\).
Chapter 1. Introduction to Vectors

1.3 Matrices

1.3.1 Matrices

Solution Solve the (linear triangular) system $Ax = b$ from top to bottom:

First $x_1 = b_1$

Then $x_2 = b_1 + b_2$

This says that $x = A^{-1}b = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

Then $x_3 = b_2 + b_3$

This is good practice to see the columns of the inverse matrix multiplying $b_1, b_2,$ and $b_3$.

The first column of $A^{-1}$ is the solution for $b = (1, 0, 0)$. The second column is the solution for $b = (0, 1, 0)$. The third column $x = A^{-1}b$ is the solution for $b = (0, 0, 1)$. The three columns of $A$ are still independent. They don’t lie in a plane. The combinations of those three columns, using the right weights $x_1, x_2, x_3$, can produce any three-dimensional vector $b = (b_1, b_2, b_3)$. Those weights come from $x = A^{-1}b$.

1.3.2 This $E$ is an elimination matrix. $E$ has a subtraction, $E^{-1}$ has an addition.

$E = \begin{bmatrix} 1 & 0 \\ -t & 1 \end{bmatrix}$

The first equation is $x_1 = b_1$. The second equation is $x_2 - t x_1 = b_2$. The inverse will add $t x_1 = b_2$, because the elimination matrix subtracted $t x_1$:

$x = E^{-1}b = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

1.3.3 Change $C$ from a cyclic difference to a centered difference producing $x_3 - x_1$:

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

Show that $Cx = b$ can only be solved when $b_1 + b_3 = 0$. That is a plane of vectors $b$ in three-dimensional space. Each column of $C$ is in the plane. The matrix has no inverse. So this plane contains all combinations of those columns (which are all the vectors $Cx$).

Solution The first component of $b = Cx$ is $x_2$, and the last component of $b$ is $-x_2$.

So we always have $b_1 + b_3 = 0$, for every choice of $x_2$.

If you draw the column vectors in $C$, the first and third columns fall on the same line. In fact (column 1) $= -(column 3)$. So the three columns will lie in a plane, and $C$ is not an invertible matrix. We cannot solve $Cx = b$ unless $b_1 + b_3 = 0$.

I included the zero row so you could see that this matrix produces "centered differences".

Row $i$ of $C$ is $x_{i+1}$ (right of center) minus $x_{i-1}$ (left of center). Here is the 4 by 4 centered difference matrix:

$Cx = b = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_3 - x_1 \\ x_4 - x_2 \\ x_5 - x_3 \\ x_6 - x_4 \end{bmatrix} = \begin{bmatrix} b_2 \\ b_3 \\ b_4 \end{bmatrix}$

Surprisingly this matrix is now invertible! The first and last rows give $x_2$ and $x_3$. Then the middle rows give $x_1$ and $x_4$. It is possible to write down the inverse matrix $C^{-1}$. But 5 by 5 will be singular (not invertible) again...
Chapter 2

Solving Linear Equations

2.1 Vectors and Linear Equations

The central problem of linear algebra is to solve a system of equations. Those equations are linear, which means that the unknowns are only multiplied by numbers—we never see times. Our first linear system is certainly not big. But you will see how far it leads:

Two equations

\[ \begin{align*}
3x + 2y &= 11 \\
2x - 2y &= 1
\end{align*} \]

(1)

We begin a row at a time. The first equation \( x - 2y = 1 \) produces a straight line in the \( xy \) plane. The point \( x = 1, y = 0 \) is on the line because it solves that equation. The point \( x = 3, y = 1 \) is also on the line because \( 3 - 2 = 1 \). If we choose \( x = 101 \) we find \( y = 50 \).

The slope of this particular line is \( \frac{1}{2} \), because \( y \) increases by 1 when \( x \) changes by 2. But slopes are important in calculus and this is linear algebra!

![Figure 2.1: Row picture: The point (3, 1) where the lines meet is the solution.](image)

Figure 2.1 shows that line \( x - 2y = 1 \). The second line in this "row picture" comes from the second equation \( 3x + 2y = 11 \). You can't miss the intersection point where the
two lines meet. The point \( x = 3, y = 1 \) lies on both lines. That point solves both equations at once. This is the solution to our system of linear equations.

**ROWS**
The row picture shows two lines meeting at a single point (the solution).

Turn now to the column picture. I want to recognize the same linear system as a "vector equation". Instead of numbers we need to see vectors. If you separate the original system into its columns instead of its rows, you get a vector equation:

\[
\begin{align*}
\text{Combination equals } b & \quad \text{ Turn this into a vector equation: } \\
\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + x \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \\ 2 \end{bmatrix} = b.
\end{align*}
\]

(2)

This has two column vectors on the left side. The problem is to find the combination of those vectors that equals the vector on the right. We are multiplying the first column by \( x \) and the second column by \( y \), and adding. With the right choices \( x = 3 \) and \( y = 1 \) (the same numbers as before), this produces \( 3(\text{column } 1) + 1(\text{column } 2) = b \).

**COLUMNS**
The column picture combines the column vectors on the left side to produce the vector \( b \) on the right side.

![Figure 2.2: Column picture: A combination of columns produces the right side (1,11).](image)

Figure 2.2 is the "column picture" of two equations in two unknowns. The first part shows the two separate columns, and that first column multiplied by \( 3 \). This multiplication by a scalar (a number) is one of the two basic operations in linear algebra:

**Scalar multiplication**

\[
\begin{bmatrix} 1 \\ 3 \end{bmatrix} \times 3 = \begin{bmatrix} 3 \\ 9 \end{bmatrix}.
\]

---

### 2.1. Vectors and Linear Equations

If the components of a vector \( u \), \( v \), and \( c \) have components \( u_1, v_1, \) and \( c_1, v_2, \) then \( cv \) has components \( cu_1, cv_2 \).

The other basic operation is vector addition. We add the first components and the second components separately. The vector sum is \((1,11)\) as desired:

\[
\begin{bmatrix} 3 \\ 9 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.
\]

The right side of Figure 2.2 shows this addition. The sum along the diagonal is the vector \( b = (1,11) \) on the right side of the linear equations.

To repeat: The left side of the vector equation is a linear combination of the columns.

The problem is to find the right coefficients \( x = 3 \) and \( y = 1 \). We are combining scalar multiplication and vector addition into one step. That step is crucially important, because it contains both of the basic operations:

**Linear combination**

\[
3 \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 11 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.
\]

Of course the solution \( x = 3, y = 1 \) is the same as in the row picture. I don't know which picture you prefer? I suspect that the two intersecting lines are more familiar at first. You may like the row picture better, but only for one day. My own preference is to combine column vectors. It is a lot easier to see a combination of four vectors in four-dimensional space, than to visualize how four hyperplanes might possibly meet at a point. (Even one hyperplane is hard enough...)

The coefficient matrix on the left side of the equations is the 2 by 2 matrix \( A \):

\[
\begin{bmatrix} 1 & 3 \\ -2 & 2 \end{bmatrix}.
\]

This is very typical of linear algebra, to look at a matrix by rows and by columns. Its rows give the row picture and its columns give the column picture. Same numbers, different pictures, same equations. We write those equations as a matrix problem \( Ax = b \):

\[
\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.
\]

The row picture deals with the two rows of \( A \). The column picture combines the columns. The numbers \( x = 3 \) and \( y = 1 \) go into \( x \). Here is matrix-vector multiplication:

\[
\begin{bmatrix} 1 \\ 3 \\ -2 \\ 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}.
\]

**Looking ahead**

This chapter is going to solve \( n \) equations in \( n \) unknowns (for any \( n \)). I am not going at top speed, because smaller systems allow examples and pictures and a complete understanding. You are free to go faster, as long as matrix multiplication and inversion become clear. Those two ideas will be the keys to invertible matrices.

I can list four steps to understanding elimination using matrices.
1. Elimination goes from \( A \) to a triangular \( U \) by a sequence of matrix steps \( E_{ij} \).

2. The inverse matrices \( E_{ij}^{-1} \) in reverse order bring \( U \) back to the original \( A \).

3. In matrix language that reverse order is \( A = LU = (\text{lower triangle}) (\text{upper triangle}) \).

4. Elimination succeeds if \( A \) is invertible. (It may need row exchanges.)

The most-used algorithm in computational science takes those steps (MATLAB calls it lu).

But linear algebra goes beyond square invertible matrices! For \( m \) by \( n \) matrices, \( Ax = 0 \) may have many solutions. Those solutions will go into a vector space. The rank of \( A \) leads to the dimension of that vector space.

All this comes in Chapter 3, and I don’t want to hurry. But I must get there.

Three Equations in Three Unknowns

The three unknowns are \( x \), \( y \), \( z \). We have three linear equations:

\[
Ax = b
\]

\[
\begin{align*}
x + 2y + 3z &= 6 \\
2x + 5y + 2z &= 4 \\
6x - 3y + z &= 2
\end{align*}
\]

We look for numbers \( x \), \( y \), \( z \) that solve all these equations at once. Those desired numbers might or might not exist. For this system, they do exist. When the number of unknowns matches the number of equations, there is usually one solution. Before solving the problem, we visualize it both ways:

**ROW** The row picture shows three planes meeting at a single point.

**COLUMN** The column picture combines three columns to produce \((6, 4, 2)\).

In the row picture, each equation produces a plane in three-dimensional space. The first plane in Figure 2.3 comes from the first equation \( x + 2y + 3z = 6 \). That plane crosses the \( x \) and \( y \) axes at the points \((6, 0, 0)\) and \((0, 3, 0)\) and \((0, 0, 2)\). Those three points solve the equation and they determine the whole plane.

The vector \((x, y, z) = (0, 0, 0)\) does not solve \( x + 2y + 3z = 6 \). Therefore that plane does not contain the origin. The plane \( x + 2y + 3z = 0 \) does pass through the origin, and it is parallel to \( x + 2y + 3z = 6 \). When the right side increases to 6, the parallel plane moves away from the origin.

The second plane is given by the second equation \( 2x + 5y + 2z = 4 \). It intersects the first plane in a line \( L \). The usual result of two equations in three unknowns is a line \( L \) of solutions. (Not if the equations were \( x + 2y + 3z = 6 \) and \( x + 2y + 3z = 0 \).)

The third equation gives a third plane. It cuts the line \( L \) at a single point. That point lies on all three planes and it solves all three equations. It is harder to draw this triple intersection point than to imagine it. The three planes meet at the solution (which we haven’t found yet). The column form will now show immediately why \( z = 2 \).

\[
\begin{align*}
2x + 5y + 2z &= 4 \\
\text{plane } x + 2y + 3z &= 6 \\
3d \text{ plane } 6x - 3y + z &= 2
\end{align*}
\]

\((0, 0, 0)\) is not on these planes.

Figure 2.3: **Row picture:** Two planes meet at a line, three planes at a point.

The column picture starts with the vector form of the equations \( Ax = b \):

\[
\begin{align*}
\text{Combine columns } x &\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} + y &\begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} + z &\begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix} = &\begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}
\end{align*}
\]

The unknowns are the coefficients \( x \), \( y \), \( z \). We want to multiply the three column vectors by the correct numbers \( x \), \( y \), \( z \) to produce \( b = (6, 4, 2) \).

\[
\begin{pmatrix}
1 \\
2 \\
6
\end{pmatrix} = \text{column 1}
\]

\[
\begin{pmatrix}
3 \\
2 \\
1
\end{pmatrix} = \text{column 2}
\]

\[
2 \text{ times column 3 is } b = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}
\]

\[
\begin{pmatrix}
2 \\
5 \\
-3
\end{pmatrix} = \text{column 3}
\]

Figure 2.4: **Column picture:** \((x, y, z) = (0, 0, 2)\) because \(2(3, 2, 1) = (6, 4, 2) = b\).

Figure 2.4 shows this column picture. Linear combinations of those columns can produce any vector \( b \). The combination that produces \( b = (6, 4, 2) \) is just 2 times the third column. The coefficients we need are \( x = 0 \), \( y = 0 \), and \( z = 2 \).
The three planes in the row picture meet at that same solution point \((0, 0, 2)\):

Correct combination
\[
\begin{bmatrix}
1 \\
2 \\
6
\end{bmatrix} + 0 \begin{bmatrix}
2 \\
5 \\
-3
\end{bmatrix} + 2 \begin{bmatrix}
3 \\
2 \\
1
\end{bmatrix} = \begin{bmatrix}
6 \\
4 \\
2
\end{bmatrix}.
\]

\((x, y, z) = (0, 0, 2)\)

**The Matrix Form of the Equations**

We have three rows in the row picture and three columns in the column picture (plus the right side). The three rows and three columns contain nine numbers. These nine numbers fill a 3 by 3 matrix \(A\):

\[
A = \begin{bmatrix}
1 & 2 & 3 \\
2 & 5 & 2 \\
6 & -3 & 1
\end{bmatrix}.
\]

The capital letter \(A\) stands for all nine coefficients (in this square array). The letter \(b\) denotes the column vector with components \(6, 4, 2\). The unknown \(x\) is also a column vector, with components \(x, y, z\). The Greek letter \(\lambda\) is used because it is a variable. By rows the equations were \((3)\), by columns they were \((4)\), and by matrices they are \((5)\):

\[
\text{Matrix equation } Ax = b
\]

\[
\begin{bmatrix}
1 & 2 & 3 \\
2 & 5 & 2 \\
6 & -3 & 1
\end{bmatrix} \begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \begin{bmatrix}
6 \\
4 \\
2
\end{bmatrix}.
\]

Basic question: What does it mean to “multiply \(A\) times \(x\)?” We can multiply by rows or by columns. Either way, \(Ax = b\) must be a correct representation of the three equations. You do the same nine multiplications either way.

Multiplication by rows

\[
Ax = \begin{bmatrix}
\text{(row 1)} \cdot x \\
\text{(row 2)} \cdot x \\
\text{(row 3)} \cdot x
\end{bmatrix}.
\]

Multiplication by columns

\[
Ax = (\text{column 1}) + (\text{column 2}) + (\text{column 3}).
\]

When we substitute the solution \((x, y, z) = (0, 0, 2)\), the multiplication \(Ax\) produces \(b\):

\[
\begin{bmatrix}
1 & 2 & 3 \\
2 & 5 & 2 \\
6 & -3 & 1
\end{bmatrix} \begin{bmatrix}
0 \\
0 \\
2
\end{bmatrix} = 2 \text{ times column } 3 = \begin{bmatrix}
6 \\
4 \\
2
\end{bmatrix}.
\]

The dot product from the first row is \((1, 2, 3) \cdot (0, 0, 2) = 6\). The other rows give dot products 4 and 2. This book sees \(Ax\) as a combination of the columns of \(A\).

---

2.1. Vectors and Linear Equations

**Example 1**

Here are 3 by 3 matrices \(A\) and \(I = \text{identity}\), with three 1’s and six 0’s:

\[
Ax = \begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 0 \\
1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
4 \\
5 \\
6
\end{bmatrix} = \begin{bmatrix}
4 \\
4 \\
4
\end{bmatrix}
\]

\[
Ix = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
4 \\
5 \\
6
\end{bmatrix} = \begin{bmatrix}
4 \\
5 \\
6
\end{bmatrix}
\]

If you are a row person, the dot product of \((1, 0, 0)\) with \((4, 5, 6)\) is 4. If you are a column person, the linear combination \(Ax = Ix\) is 4 times the first column \((1, 1, 1)\). In that matrix \(A\), the second and third columns are zero vectors.

The other matrix \(I\) is special. It has ones on the “main diagonal”. Whatever vector this matrix multiplies, that vector is not changed. This is like multiplication by 1, but for matrices and vectors. The exceptional matrix in this example is the 3 by 3 identity matrix:

\[
I = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

a way yields the multiplication \(Ix = x\).

**Matrix Notation**

The first row of a 2 by 2 matrix contains \(a_{11}\) and \(a_{12}\). The second row contains \(a_{21}\) and \(a_{22}\). The first index gives the row number, so that \(a_{ij}\) is an entry in row \(i\). The second index \(j\) gives the column number. But those subscripts are not very convenient on a keyboard! Instead of \(a_{ij}\) we type \(A(i, j)\). The entry \(a_{57} = A(5, 7)\) would be in row 5, column 7.

\[
A = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix} = \begin{bmatrix}
A(1, 1) & A(1, 2) \\
A(2, 1) & A(2, 2)
\end{bmatrix}
\]

For an \(m\) by \(n\) matrix, the row index \(i\) goes from 1 to \(m\). The column index \(j\) stops at \(n\). There are \(mn\) entries \(a_{ij} = A(i, j)\). A square matrix of order \(n\) has \(n^2\) entries.

**Multiplication in MATLAB**

I want to express \(A\) and \(x\) and their product \(Ax\) using MATLAB commands. This is a first step in learning that language. I begin by defining the matrix \(A\) and the vector \(x\). This vector is a 3 by 1 matrix, with three rows and one column. Enter matrices a row at a time, and use a semicolon to signal the end of a row:

\[
A = [1 \ 2 \ 3; \ 2 \ 5 \ 2; \ 6 \ -3 \ 1]
\]

\[
x = [0; \ 0; \ 2]
\]

Here are three ways to multiply \(Ax\) in MATLAB. In reality, \(A \cdot x\) is the good way to do it. MATLAB is a high level language, and it works with matrices:

\[
\text{Matrix multiplication } b = A \cdot x
\]
2.1. Vectors and Linear Equations

**WORKED EXAMPLES**

2.1 A Describe the column picture of these three equations \( Ax = b \). Solve by careful inspection of the columns (instead of elimination):

\[
\begin{align*}
    z + 3y + 2z &= -3 \\
    2x + 2y + 2z &= -2 \\
    3z + 5y + 6z &= -5
\end{align*}
\]

which is

\[
\begin{bmatrix}
    1 & 3 & 2 \\
    2 & 2 & 2 \\
    3 & 5 & 6
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix} =
\begin{bmatrix}
    -3 \\
    -2 \\
    -5
\end{bmatrix}
\]

Solution

The column picture asks for a linear combination that produces \( b \) from the three columns of \( A \). In this example \( b \) is **not the second column**. So the solution is \( x = 0, y = -1, z = 0 \). To show that \((0, -1, 0)\) is the only solution we have to know that "\( A \) is invertible" and "the columns are independent" and "the determinant isn’t zero."

Those words are not yet defined but the test comes from elimination: We need (and for this matrix we find) a full set of three nonzero pivots.

Suppose the right side changes to \( b = (4, 4, 8) \) is sum of the first two columns. Then the good combination has \( x = 1, y = 1, z = 0 \). The solution becomes \( x = (1, 1, 0) \).

2.1 B This system has no solution. The planes in the row picture don’t meet at a point.

No combination of the three columns produces \( b \). How to show this?

\[
\begin{align*}
    x + 3y + 5z &= 4 \\
    x + 2y - 3z &= 5 \\
    2x + 5y + 2z &= 8
\end{align*}
\]

(1) Multiply the equations by \( 1, 1, -1 \) and add to get \( 0 = 1 \). No solution. Are any two of the planes parallel? What are the equations of planes parallel to \( x - 3y + 5z = 4 \)?

(2) Take the dot product of each column of \( A \) (and also \( b \)) with \( y = (1, 1, -1) \).

How do those dot products show that the system \( Ax = b \) has no solution?

(3) Find three right side vectors \( b^- \) and \( b^{**} \) and \( b^{***} \) that do allow solutions.

Solution

(1) Multiplying the equations by \( 1, 1, -1 \) and adding gives \( 0 = 1 \):

\[
\begin{align*}
    x + 3y + 5z &= 4 \\
    x + 2y - 3z &= 5 \\
    -[2x + 5y + 2z] &= 8
\end{align*}
\]

\(0x + 0y + 0z = 1\) No Solution

The planes don’t meet at a point, even though no two planes are parallel. For a plane parallel to \( x + 3y + 5z = 4 \), change the "4". The parallel plane \( x + 3y + 5z = 0 \) goes through the origin \((0,0,0)\). And the equation multiplied by any nonzero constant still gives the same plane, as in \( 2x + 6y + 10z = 8 \).
(2) The dot product of each column of $A$ with $y = (1, 1, -1)$ is zero. On the right side,

$$y \cdot b = ((1, 1, -1) \cdot (4, 5, 8) = 1$$

is not zero. So a solution is impossible.

(3) There is a solution when $b$ is a combination of the columns. These three choices of $b$

have solutions $x^* = (1, 0, 0)$ and $x^{**} = (1, 1, 1)$ and $x^{***} = (0, 0, 0)$:

$$b^* = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \quad b^{**} = \begin{bmatrix} 0 \\ 9 \\ 0 \end{bmatrix} \quad b^{***} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Problem Set 2.1

Problems 1-8 are about the row and column pictures of $Ax = b$.

1. With $A = I$ (the identity matrix) draw the planes in the row picture. Three sides of a box meet at the solution $x = (x, y, z) = (2, 3, 4)$,

$$1x + 0y + 0z = 2 \\
0x + 1y + 0z = 3 \\
0x + 0y + 1z = 4$$

or draw the vectors in the column picture. Two times column 1 plus three times column 2 plus four times column 3 equals the right side $b$.

2. If the equations in Problem 1 are multiplied by 2, 3, 4, they become $DX = B$:

$$2x + 0y + 0z = 4 \\
0x + 3y + 0z = 9 \\
0x + 0y + 4z = 16$$

Why is the row picture the same? Is the solution $X$ the same as $x$? What is changed in the column picture—the columns or the right combination to give $B$?

3. If equation 1 is added to equation 2, which of these are changed: the planes in the row picture, the vectors in the column picture, the coefficient matrix, the solution?

The new equations in Problem 1 would be $x = 2$, $x + y = 5$, $z = 4$.

4. Find a point with $z = 2$ on the intersection line of the planes $x + y + 3z = 6$ and $x - y + z = 4$. Find the point with $z = 0$. Find a third point halfway between.

5. The first of these equations plus the second equals the third:

$$x + y + z = 2$$

$$x + 2y + z = 3$$

$$2x + 3y + 2z = 5$$

The first two planes meet along a line. The third plane contains that line, because

if $x, y, z$ satisfy the first two equations they also _______. The equations have

indefinitely many solutions (the whole line L). Find three solutions on L.

2.1. Vectors and Linear Equations

6. Move the third plane in Problem 5 to a parallel plane $2x + 3y + 2z = 9$. Now the three equations have no solution—why not? The first two planes meet along the line $L$, but the third plane doesn’t ______ that line.

7. In Problem 5 the columns are $(1, 1, 2)$ and $(1, 2, 3)$ and $(1, 1, 2)$. This is a "singular case" because the third column is ______. Find two combinations of the columns that give $b = (2, 3, 5)$. This is only possible for $b = (4, 6, c)$ if $c = ______$.

8. Normally 4 "planes" in 4-dimensional space meet at a ______. Normally 4 column vectors in 4-dimensional space can combine to produce $b$. What combination of $(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)$ produces $b = (3, 3, 3, 2)$? What 4 equations for $x, y, z, t$ are you solving?

Problems 9-14 are about multiplying matrices and vectors.

9. Compute each $Ax$ by dot products of the rows with the column vector:

(a) \[
\begin{bmatrix}
1 & 2 & 4 \\
-2 & 3 & 1 \\
-4 & 1 & 2
\end{bmatrix}
\] 

(b) \[
\begin{bmatrix}
2 & 1 & 0 & 0 \\
1 & 2 & 1 & 0 \\
0 & 1 & 2 & 1 \\
0 & 0 & 1 & 2
\end{bmatrix}
\]

10. Compute each $Ax$ in Problem 9 as a combination of the columns:

9(a) becomes $Ax = 2 \begin{bmatrix} -1 \\ -2 \\ 2 \\ -4 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 3 \\ 1 \\ 2 \\ 2 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$

How many separate multiplications for $Ax$, when the matrix is "3 by 3"?

11. Find the two components of $Ax$ by rows or by columns:

\[
\begin{bmatrix}
2 & 3 \\
5 & 1 \\
6 & 12 \\
1 & 2 & 4
\end{bmatrix}
\] 

and \[
\begin{bmatrix}
4 & 2 \\
6 & 1 \\
12 & -1 \\
1 & 2 & 1
\end{bmatrix}
\]

12. Multiply $A$ times $x$ to find three components of $Ax$:

\[
\begin{bmatrix}
0 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 0
\end{bmatrix} \quad \begin{bmatrix}
2 & 1 & 3 \\
1 & 2 & 3 \\
3 & 3 & 6
\end{bmatrix} \quad \begin{bmatrix}
-1 \\
-1 \\
-1
\end{bmatrix}
\]

13. (a) A matrix with $m$ rows and $n$ columns multiplies a vector with ______ components to produce a vector with ______ components.

(b) The planes from the $m$ equations $Ax = b$ are in ______-dimensional space. The combination of the columns of $A$ is in ______-dimensional space.
Chapter 2. Solving Linear Equations

14 Write $2x + 3y + z + 5t = 8$ as a matrix $A$ (how many rows?) multiplying the column vector $x = (x, y, z, t)$ to produce $b$. The solutions $x$ fill a plane or "hyperplane" in 4-dimensional space. The plane is 3-dimensional with no 4D volume.

Problems 15–22 ask for matrices that act in special ways on vectors.

15 (a) What is the 2 by 2 identity matrix? $I$ times $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ equals $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$.
15 (b) What is the 2 by 2 exchange matrix? $P$ times $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ equals $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

16 (a) What 2 by 2 matrix $R$ rotates every vector by 90°? $R$ times $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ is $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$.
16 (b) What 2 by 2 matrix $R^2$ rotates every vector by 180°?

17 Find the matrix $P$ that multiplies $(x, y, z)$ to give $(y, z, x)$. Find the matrix $Q$ that multiplies $(y, x, z)$ to bring back $(x, y, z)$.

18 What 2 by 2 matrix $E$ subtracts the first component from the second component? What 3 by 3 matrix does the same?

$$E \begin{bmatrix} 3 \\ 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

19 What 3 by 3 matrix $E$ multiplies $(x, y, z)$ to give $(x, y, z-x)$? What matrix $E^{-1}$ multiplies $(x, y, z)$ to give $(x, y, z-x)$? If you multiply $(3, 4, 5)$ by $E$ and then multiply by $E^{-1}$, the two results are (___) and (___).

20 What 2 by 2 matrix $P_1$ projects the vector $(x, y)$ onto the $x$ axis to produce $(x, 0)$? What matrix $P_2$ projects onto the $y$ axis to produce $(0, y)$? If you multiply $(5, 7)$ by $P_1$ and then multiply by $P_2$, you get (___) and (___).

21 What 2 by 2 matrix $R$ rotates every vector through 45°? The vector $(1, 0)$ goes to $(\sqrt{2}/2, \sqrt{2}/2)$. The vector $(0, 1)$ goes to $(-\sqrt{2}/2, \sqrt{2}/2)$. Those determine the matrix. Draw these particular vectors in the $xy$ plane and find $R$.

22 Write the dot product of $(1, 4, 5)$ and $(x, y, z)$ as a matrix multiplication $Ax$. The matrix $A$ has one row. The solutions to $Ax = 0$ lie on a perpendicular to the vector (___). The columns of $A$ are in a ___-dimensional space.

23 In MATLAB notation, write the commands that define this matrix $A$ and the column vectors $x$ and $b$. What command would test whether or not $Ax = b$?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} 5 \\ -2 \end{bmatrix}, \quad b = \begin{bmatrix} 7 \end{bmatrix}$$

24 The MATLAB commands $A = \text{eye}(3)$ and $v = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}$ produce the 3 by 3 identity matrix and the column vector $(3, 4, 5)$. What are the outputs of $Av$ and $v'v$? (Computer not needed!) If you ask for $v'Av$, what happens?

25 If you multiply the 4 by 4 all-ones matrix $A = \text{ones}(4)$ and the column $v = \text{ones}(4, 1)$, what is $Avv$? (Computer not needed.) If you multiply $B = \text{eye}(4) + \text{ones}(4)$ times $w = \text{zeros}(4, 1) + 2*\text{ones}(4, 1)$, what is $Bww$?

Questions 26–28 review the row and column pictures in 2, 3, and 4 dimensions.

26 Draw the row and column pictures for the equations $x - 2y = 0, x + y = 6$.

27 For two linear equations in three unknowns $x, y, z$, the row picture will show (2 or 3) (lines or planes) in (2 or 3)-dimensional space. The column picture is in (2 or 3)-dimensional space. The solutions normally lie on a ___.

28 For four linear equations in two unknowns $x$ and $y$, the row picture shows four possible solutions. The column picture is in ___-dimensional space. The equations have no solution unless the vector on the right side is a combination of ___.

29 Start with the vector $v_0 = (1, 0)$. Multiply again and again by the same "Markov matrix" $A = (0.8, 0.3, 0.2, 0.7)$. The next three vectors are $u_1, u_2, u_3$:

$$u_1 = \begin{bmatrix} 0.8 \\ 0.3 \\ 0.1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0.2 \\ 0.7 \\ 0 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0.2 \\ 0.7 \\ 0 \end{bmatrix}$$

What property do you notice for all vectors $u_0, u_1, u_2, u_3$?

Challenge Problems

30 Continue Problem 29 from $v_0 = (1, 0)$ to $v_7$, and also from $v_0 = (0, 1)$ to $v_7$. What do you notice about $v_7$ and $v_5$? Here are two MATLAB codes, with while and for. They plot $v_0$ to $u_7$ and $v_0$ to $v_7$. You can use other languages:

```
for $j = 1:7$
  $u = A*v; x = [x; u];$
  end
plot(x, x)
```

The $u$’s and $v$’s are approaching a steady state vector $s$. Guess that vector and check that $As = s$. If you start with $x$, you stay with $x$.

31 Invent a 3 by 3 magic matrix $M_5$ with entries $1, 2, \ldots, 9$. All rows and columns and diagonals add to 15. The first row could be $8, 3, 4$. What is $M_5$ times $(1, 1, 1)$? What is $M_5$ times $(1, 1, 1)$ if a 4 by 4 magic matrix has entries $1, \ldots, 16$?

32 Suppose $u$ and $v$ are the first two columns of a 3 by 3 matrix $A$. Which third column $w$ would make this matrix singular? Describe a typical column picture of $Ax = b$ in that singular case, and a typical row picture (for a random $b$).
Chapter 2. Solving Linear Equations

Multiplying by \( A \) is a “linear transformation”. Those important words mean:

If \( w \) is a combination of \( u \) and \( v \), then \( Aw \) is the same combination of \( Au \) and \( Av \).

It is this “linearity” \( Aw = cAu + dAv \) that gives us the name linear algebra.

Problem: If \( u = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \) and \( v = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \) then \( Au \) and \( Av \) are the columns of \( A \).

Combine \( w = cu + dv \). If \( w = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \) how is \( Aw \) connected to \( Au \) and \( Av \)?

Start from the four equations: \( -x_1 + x_2 + 2x_3 - x_4 = i \) (for \( i = 1, 2, 3, 4 \) with \( x_0 = x_5 = 0 \)). Write those equations in their matrix form \( Ax = b \). Can you solve them for \( x_1, x_2, x_3, x_4, x_5 \)?

35 A 9 by 9 Sudoku matrix \( S \) has the numbers 1, \ldots, 9 in every row and column, and in every 3 by 3 block. For the all-ones vector \( x = (1, \ldots, 1) \), what is \( Sx \)?

A better question is: Which row exchanges will produce another Sudoku matrix?

Also, which exchanges of block rows give another Sudoku matrix?

Section 2.7 will look at all possible permutations (reorderings) of the rows. I can see next 3 rows, and of the last 3 rows. And 6 block permutations of the block rows?

2.2 The Idea of Elimination

This chapter explains a systematic way to solve linear equations. The method is called “elimination”, and you can see it immediately in our 2 by 2 example. Before elimination, \( x \) and \( y \) appear in both equations. After elimination, the first unknown \( x \) has disappeared from the second equation \( 8y \neq 8 \):

Before \( \begin{array}{c}
\begin{align*}
x - 2y &= 1 \\
x + 2y &= 11
\end{align*}
\end{array} \)

After \( \begin{array}{c}
\begin{align*}
x - 2y &= 1 \\
8y &= 8
\end{align*}
\end{array} \) (multiply equation 1 by 3)

(3, 1) (3x, 3y)

(subtract to eliminate 3x)

The new equation \( 8y = 8 \) instantly gives \( y = 1 \). Substituting \( y = 1 \) back into the first equation leaves \( x - 2 = 1 \). Therefore \( x = 3 \) and the solution \( (x, y) = (3, 1) \) is complete.

Elimination produces an upper triangular system—this is the goal. The nonzero coefficients 1, -2, 8 form a triangle. That system is solved from the bottom upwards first \( y = 1 \) and then \( x = 3 \). This quick process is called back substitution. It is used for upper triangular systems of any size, after elimination gives a triangle.

Important point: The original equations have the same solution \( x = 3 \) and \( y = 1 \). Figure 2.5 shows each system as a pair of lines, intersecting at the solution point (3, 1). After elimination, the lines still meet at the same point. Every step worked with correct equations.

How did we get from the first pair of lines to the second pair? We subtracted 3 times the first equation from the second equation. The step that eliminates \( x \) from equation 2 is the fundamental operation in this chapter. We use it so often that we look at it closely:

To eliminate \( x \): Subtract a multiple of equation 1 from equation 2.

Three times \( x - 2y = 1 \) gives \( 3x - 6y = 3 \). When this is subtracted from \( 3x + 2y = 11 \), the right side becomes 8. The main point is that \( 3x \) cancels \( 3x \). What remains on the left side is \( 2y - (-6y) = 8y \), and \( x \) is eliminated. The system became triangular.

Ask yourself how that multiplier \( \ell = 3 \) was found. The first equation contains \( 1x \). So the first pivot was 1 (the coefficient of \( x \). The second equation contains \( 3x \), so the multiplier was 3. Then subtraction \( 3x - 3x \) produced the zero and the triangle.

Figure 2.5: Eliminating \( x \) makes the second line horizontal. Then \( 8y = 8 \) gives \( y = 1 \).