

## Chapter 1

# Introduction to Vectors

The heart of linear algebra is in two operations—both with vectors. We add vectors to get  $v + w$ . We multiply them by numbers  $c$  and  $d$  to get  $cv$  and  $d w$ . Combining those two operations (adding  $cv$  to  $d w$ ) gives the **linear combination**  $cv + d w$ .

**Linear combination**  $cv + d w = c \begin{bmatrix} 1 \\ 1 \end{bmatrix} + d \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} c + 2d \\ c + 3d \end{bmatrix}$

**Example**  $v + w = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$  is the combination with  $c = d = 1$

Linear combinations are all-important in this subject! Sometimes we want one particular combination, the specific choice  $c = 2$  and  $d = 1$  that produces  $cv + d w = (4, 5)$ . Other times we want *all the combinations* of  $v$  and  $w$  (coming from all  $c$  and  $d$ ).

The vectors  $cv$  lie along a line. When  $w$  is not on that line, **the combinations**  $cv + d w$  **fill the whole two-dimensional plane**. (I have to say “two-dimensional” because linear algebra allows higher-dimensional planes.) Starting from four vectors  $u, v, w, z$  in four-dimensional space, their combinations  $cu + dv + ew + fz$  are likely to fill the space—but not always. The vectors and their combinations could even lie on one line.

Chapter 1 explains these central ideas, on which everything builds. We start with two-dimensional vectors and three-dimensional vectors, which are reasonable to draw. Then we move into higher dimensions. The really impressive feature of linear algebra is how smoothly it takes that step into  $n$ -dimensional space. Your mental picture stays completely correct, even if drawing a ten-dimensional vector is impossible.

This is where the book is going (into  $n$ -dimensional space). The first steps are the operations in Sections 1.1 and 1.2. Then Section 1.3 outlines three fundamental ideas.

1.1 Vector addition  $v + w$  and linear combinations  $cv + d w$ .

1.2 The dot product  $v \cdot w$  of two vectors and the length  $\|v\| = \sqrt{v \cdot v}$ .

1.3 Matrices  $A$ , linear equations  $Ax = b$ , solutions  $x = A^{-1}b$ .

## 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**  $v$ :

**Column vector**  $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$   $v_1 =$  first component  
 $v_2 =$  second component

We write  $v$  as a **column**, not as a row. The main point so far is to have a single letter  $v$  (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  $v$  and  $w$  stay separate from the second components:

**VECTOR ADDITION**  $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$  and  $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$  add to  $v + w = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$ .

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

## 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*  $v$ :

$$\text{Column vector } v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \begin{array}{l} v_1 = \text{first component} \\ v_2 = \text{second component} \end{array}$$

We write  $v$  as a *column*, not as a row. The main point so far is to have a single letter  $v$  (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  $v$  and  $w$  stay separate from the second components:

$$\text{VECTOR ADDITION } v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{and} \quad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \text{add to} \quad v + w = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}.$$

You see the reason. We want to add apples to apples. Subtraction of vectors follows the same idea: *The components of*  $v - w$  *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  $v$  to  $v$ . The other way (the usual way) is to multiply each component by 2:

$$\text{SCALAR MULTIPLICATION } 2v = \begin{bmatrix} 2v_1 \\ 2v_2 \end{bmatrix} \quad \text{and} \quad -v = \begin{bmatrix} -v_1 \\ -v_2 \end{bmatrix}.$$

The components of  $cv$  are  $cv_1$  and  $cv_2$ . The number  $c$  is called a "scalar".

Notice that the sum of  $-v$  and  $v$  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  $v + w$  and  $cv$ —*adding vectors and multiplying by scalars*.

The order of addition makes no difference:  $v + w$  equals  $w + v$ . Check that by algebra: The first component is  $v_1 + w_1$  which equals  $w_1 + v_1$ . Check also by an example:

$$v + w = \begin{bmatrix} 1 \\ 5 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix} \quad w + v = \begin{bmatrix} 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}.$$

### Linear Combinations

Combining addition with scalar multiplication, we now form "*linear combinations*" of  $v$  and  $w$ . Multiply  $v$  by  $c$  and multiply  $w$  by  $d$ ; then add  $cv + dw$ .

**DEFINITION** *The sum of*  $cv$  *and*  $dw$  *is a* *linear combination of*  $v$  *and*  $w$ .

Four special linear combinations are: sum, difference, zero, and a scalar multiple  $cv$ :

$$\begin{array}{ll} 1v + 1w & = \text{sum of vectors in Figure 1.1a} \\ 1v - 1w & = \text{difference of vectors in Figure 1.1b} \\ 0v + 0w & = \text{zero vector} \\ cv + 0w & = \text{vector } cv \text{ in the direction of } v \end{array}$$

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  $v$  and  $w$ , 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  $v$  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—so we have three ways to describe  $v$ :

**Represent vector**  $v$    Two numbers   Arrow from (0, 0)   Point in the plane

We add using the numbers. We visualize  $v + w$  using arrows:

## Linear Combinations

Combining addition with scalar multiplication, we now form “linear combinations” of  $v$  and  $w$ . Multiply  $v$  by  $c$  and multiply  $w$  by  $d$ ; then add  $cv + dw$ .

**DEFINITION** The sum of  $cv$  and  $dw$  is a linear combination of  $v$  and  $w$ .

Four special linear combinations are: sum, difference, zero, and a scalar multiple  $cv$ :

$$\begin{aligned} 1v + 1w &= \text{sum of vectors in Figure 1.1a} \\ 1v - 1w &= \text{difference of vectors in Figure 1.1b} \\ 0v + 0w &= \text{zero vector} \\ cv + 0w &= \text{vector } cv \text{ in the direction of } v \end{aligned}$$

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  $v$  and  $w$ , 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  $v$  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—so we have three ways to describe  $v$ :

**Represent vector  $v$**  Two numbers    Arrow from  $(0, 0)$     Point in the plane

We add using the numbers. We visualize  $v + w$  using arrows:

*Vector addition* (head to tail)    *At the end of  $v$ , place the start of  $w$ .*

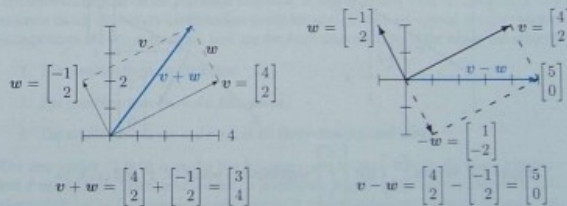


Figure 1.1: Vector addition  $v + w = (3, 4)$  produces the diagonal of a parallelogram. The linear combination on the right is  $v - w = (5, 0)$ .

We travel along  $v$  and then along  $w$ . Or we take the diagonal shortcut along  $v + w$ . We could also go along  $w$  and then  $v$ . In other words,  $w + v$  gives the same answer as  $v + w$ .

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

The zero vector  $\mathbf{0} = (0, 0)$  is too short to draw a decent arrow, but you know that  $v + \mathbf{0} = v$ . For  $2v$  we double the length of the arrow. We reverse  $w$  to get  $-w$ . 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  $v$  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):

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

The vector  $v$  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*.

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

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

16 / 5

## Vectors in Three Dimensions

A vector with two components corresponds to a point in the  $xy$  plane. The components of  $v$  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):

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

The vector  $v$  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*.

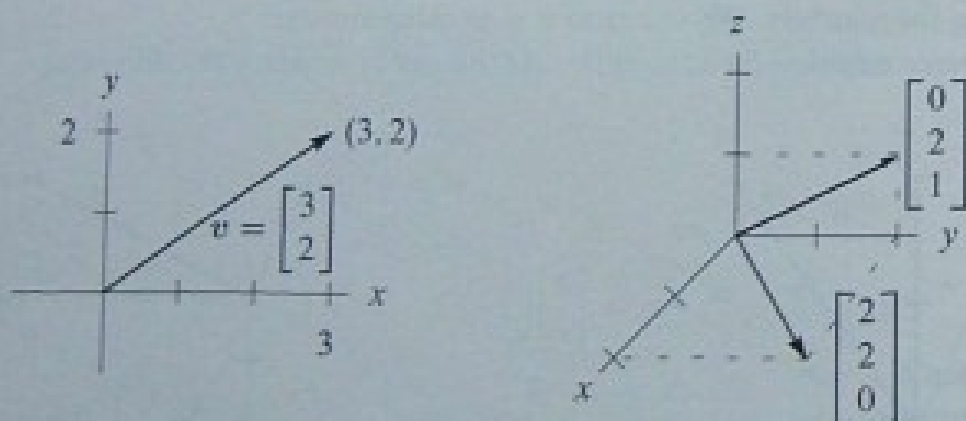


Figure 1.2: Vectors  $\begin{bmatrix} x \\ y \end{bmatrix}$  and  $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$  correspond to points  $(x, y)$  and  $(x, y, z)$ .

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

The reason for the row form (in parentheses) is to save space. But  $v = (1, 1, -1)$  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  $v$ .

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

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

$$\begin{array}{l} \text{Linear combination} \\ \text{Multiply by } 1, 4, -2 \\ \text{Then add} \end{array} \quad \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}.$$

### The Important Questions

For one vector  $u$ , the only linear combinations are the multiples  $cu$ . For two vectors, the combinations are  $cu + dv$ . For three vectors, the combinations are  $cu + dv + ew$ . Will you take the big step from *one* combination to *all* combinations? Every  $c$  and  $d$  and  $e$  are allowed. Suppose the vectors  $u, v, w$  are in three-dimensional space:

1. What is the picture of *all* combinations  $cu$ ?
2. What is the picture of *all* combinations  $cu + dv$ ?
3. What is the picture of *all* combinations  $cu + dv + ew$ ?

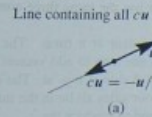
The answers depend on the particular vectors  $u, v$ , and  $w$ . If they were zero vectors (a very extreme case), then every combination would be zero. If they are typical nonzero vectors (components chosen at random), here are the three answers. This is the key to our subject:

1. The combinations  $cu$  fill a *line*.
2. The combinations  $cu + dv$  fill a *plane*.
3. The combinations  $cu + dv + ew$  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  $cu$  is infinitely long (forward and backward). It is the plane of all  $cu + dv$  (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 in the plane in Figure 1.3.*

When we include a third vector  $w$ , the multiples  $ew$  give a third line. Suppose that third line is not in the plane of  $u$  and  $v$ . Then combining all  $ew$  with all  $cu + dv$  fills up the whole three-dimensional space.



(a)



(b)

Figure 1.3: (a) Line through  $u$ . (b) The plane containing the lines through  $u$  and  $v$ .

This is the typical situation! **Line**, then **plane**, then **space**. But other possibilities exist. When  $w$  happens to be  $cu + dv$ , the third vector is in the plane of the first two. The combinations of  $u, v, w$  will not go outside that  $uv$  plane. We do not get the full three-dimensional space. Please think about the special cases in Problem 1.

#### REVIEW OF THE KEY IDEAS

1. A vector  $v$  in two-dimensional space has two components  $v_1$  and  $v_2$ .
2.  $v + w = (v_1 + w_1, v_2 + w_2)$  and  $cv = (cv_1, cv_2)$  are found a component at a time.
3. A linear combination of three vectors  $u$  and  $v$  and  $w$  is  $cu + dv + ew$ .
4. Take *all* linear combinations of  $u$ , or  $u$  and  $v$ , or  $u, v, w$ . In three dimensions, those combinations typically fill a line, then a plane, and the whole space  $\mathbf{R}^3$ .

#### WORKED EXAMPLES

**1.1 A** The linear combinations of  $v = (1, 1, 0)$  and  $w = (0, 1, 1)$  fill a plane. Describe that plane. Find a vector that is *not* a combination of  $v$  and  $w$ .

**Solution** The combinations  $cv + dw$  fill a plane in  $\mathbf{R}^3$ . The vectors in that plane allow any  $c$  and  $d$ . The plane of Figure 1.3 fills in between the “ $u$ -line” and the “ $v$ -line”.

$$\text{Combinations } cv + dw = c \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c \\ c + d \\ d \end{bmatrix} \text{ fill a plane.}$$

Four particular vectors in that plane are  $(0, 0, 0)$  and  $(2, 3, 1)$  and  $(5, 7, 2)$  and  $(\pi, 2\pi, \pi)$ . The second component  $c + d$  is always the sum of the first and third components. The vector  $(1, 2, 3)$  is *not* in the plane, because  $2 \neq 1 + 3$ .

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

**1.1 B** For  $v = (1, 0)$  and  $w = (0, 1)$ , describe all points  $cv$  with (1) whole numbers  $c$  (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  $(\pi, 0)$  but not  $(-\pi, 0)$ .
- (1') 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  $d$ ).
- (2') 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  $xy$  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  $b$ :

$$u = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \quad v = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

## 1.1. Vectors and Linear Combinations

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

**1.1 B** For  $\mathbf{v} = (1, 0)$  and  $\mathbf{w} = (0, 1)$ , describe all points  $c\mathbf{v}$  with (1) whole numbers  $c$  (2) nonnegative  $c \geq 0$ . Then add all vectors  $d\mathbf{w}$  and describe all  $c\mathbf{v} + d\mathbf{w}$ .

**Solution**

- (1) The vectors  $c\mathbf{v} = (c, 0)$  with whole numbers  $c$  are equally spaced points along the  $x$  axis (the direction of  $\mathbf{v}$ ). They include  $(-2, 0)$ ,  $(-1, 0)$ ,  $(0, 0)$ ,  $(1, 0)$ ,  $(2, 0)$ .
- (2) The vectors  $c\mathbf{v}$  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  $(\pi, 0)$  but not  $(-\pi, 0)$ .
- (1') Adding all vectors  $d\mathbf{w} = (0, d)$  puts a vertical line through those points  $c\mathbf{v}$ . We have infinitely many *parallel lines* from (whole number  $c$ , any number  $d$ ).
- (2') Adding all vectors  $d\mathbf{w}$  puts a vertical line through every  $c\mathbf{v}$  on the half-line. Now we have a *half-plane*. It is the right half of the  $xy$  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  $c\mathbf{v} + d\mathbf{w}$  equals the vector  $\mathbf{b}$ :

$$\mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \mathbf{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

**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:

$$\text{Find } c_1, \dots, c_n \text{ so that } c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n = \mathbf{b}.$$

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  $n = 2$ :

$$\text{Vector equation} \quad c \begin{bmatrix} 2 \\ -1 \end{bmatrix} + d \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The required equations for  $c$  and  $d$  just come from the two components separately:

$$\begin{aligned} \text{Two scalar equations} \quad & 2c - d = 1 \\ & -c + 2d = 0 \end{aligned}$$

You could think of those as two lines that cross at the solution  $c = \frac{2}{3}$ ,  $d = \frac{1}{3}$ .

## 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

$$(a) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ and } \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \quad (b) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \quad (c) \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

2 Draw  $v = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$  and  $w = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$  and  $v + w$  and  $v - w$  in a single  $xy$  plane.

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

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

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

$$\text{In a plane} \quad u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad v = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix}, \quad w = \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}.$$

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:

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

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 three corners of a parallelogram are  $(1, 1)$ ,  $(4, 2)$ , and  $(1, 3)$ , what are all three 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  $j = (0, 1, 0)$  and  $k = (0, 0, 1)$ ? 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  $(0, 1, 0, 0)$ .



## 1.2 Lengths and Dot Products

The first section backed off from multiplying vectors. Now we go forward to define the “dot product” of  $v$  and  $w$ . This multiplication involves the separate products  $v_1 w_1$  and  $v_2 w_2$ , but it doesn’t stop there. Those two numbers are added to produce the single number  $v \cdot w$ . *This is the geometry section (lengths and angles).*

**DEFINITION** The *dot product* or *inner product* of  $v = (v_1, v_2)$  and  $w = (w_1, w_2)$  is the number  $v \cdot w$ :

$$v \cdot w = v_1 w_1 + v_2 w_2. \quad (1)$$

**Example 1** The vectors  $v = (4, 2)$  and  $w = (-1, 2)$  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  $90^\circ$ . 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  $v = (1, 2)$  and  $w = (3, 1)$  is 5. Soon  $v \cdot w$  will reveal the angle between  $v$  and  $w$  (not  $90^\circ$ ). Please check that  $w \cdot v$  is also 5.

*The dot product  $w \cdot v$  equals  $v \cdot w$ . The order of  $v$  and  $w$  makes no difference.*

**Example 2** Put a weight of 4 at the point  $x = -1$  (left of zero) and a 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  $(v_1, v_2) = (-1, 2)$ . The weights times the distances,  $w_1 v_1$  and  $w_2 v_2$ , give the “moments”. The equation for the see-saw to balance is  $w_1 v_1 + w_2 v_2 = 0$ .

**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”  $p$ . 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  $q_1 p_1$ .* The total income (quantities  $q$  times prices  $p$ ) is *the dot product  $q \cdot p$  in three dimensions*:

$$\text{Income} = (q_1, q_2, q_3) \cdot (p_1, p_2, p_3) = q_1 p_1 + q_2 p_2 + q_3 p_3 = \text{dot product.}$$

A zero dot product means that “the books balance”. Total sales equal total purchases if  $q \cdot p = 0$ . Then  $p$  is perpendicular to  $q$  (in three-dimensional space). A supermarket with thousands of goods goes quickly into high dimensions.

Small note: Spreadsheets have become essential in management. They compute linear combinations and dot products. What you see on the screen is a matrix.

**Main point** To compute  $v \cdot w$ , multiply each  $v_i$  times  $w_i$ . Then add  $\sum v_i w_i$ .

24 / 5

## Lengths and Unit Vectors

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

$$\begin{array}{l} \text{Dot product } v \cdot v \\ \text{Length squared} \end{array} \quad \|v\|^2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 + 4 + 9 = 14.$$

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

**DEFINITION** The *length*  $\|v\|$  of a vector  $v$  is the square root of  $v \cdot v$ :

Length = norm( $v$ )

$$\text{length} = \|v\| = \sqrt{v \cdot v}.$$

In two dimensions the length is  $\sqrt{v_1^2 + v_2^2}$ . In three dimensions it is  $\sqrt{v_1^2 + v_2^2 + v_3^2}$ . By the calculation above, the length of  $v = (1, 2, 3)$  is  $\|v\| = \sqrt{14}$ .

Here  $\|v\| = \sqrt{v \cdot v}$  is just the ordinary length of the arrow 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  $1^2 + 2^2 = \|v\|^2$ .

For the length of  $v = (1, 2, 3)$ , we used the right triangle formula 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  $\|v\| = \sqrt{5 + 9} = \sqrt{14}$ .

The length of a four-dimensional vector would be  $\sqrt{v_1^2 + v_2^2 + v_3^2 + v_4^2}$ . 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*  $u$  is a vector whose length equals one. Then  $u \cdot u = 1$ .

An example in four dimensions is  $u = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ . Then  $u \cdot u$  is  $\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$ . We divided  $v = (1, 1, 1, 1)$  by its length  $\|v\| = 2$  to get this unit vector.

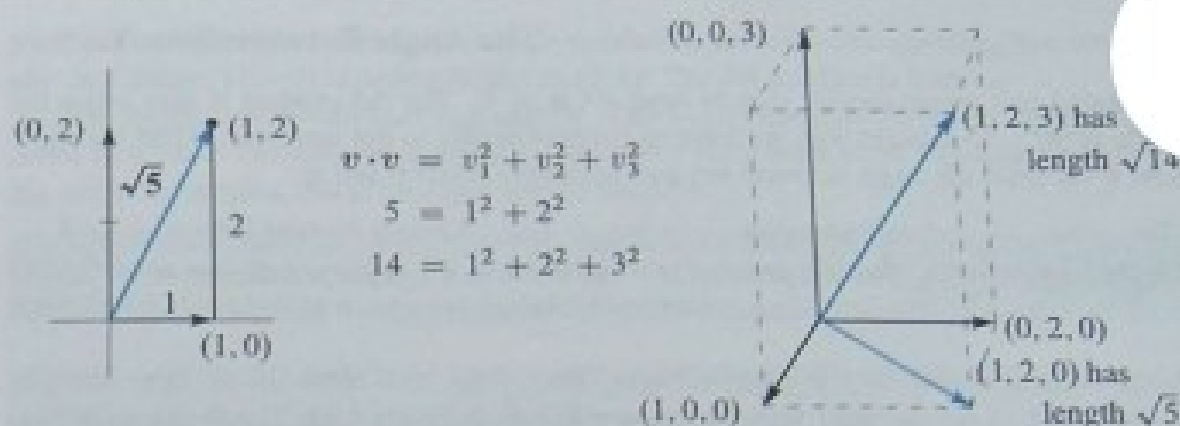


Figure 1.6: The length  $\sqrt{v \cdot v}$  of two-dimensional and three-dimensional vectors.

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

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

When  $\theta = 0$ , the horizontal vector  $u$  is  $i$ . When  $\theta = 90^\circ$  (or  $\frac{\pi}{2}$  radians), the vertical vector is  $j$ . At any angle, the components  $\cos \theta$  and  $\sin \theta$  produce  $u \cdot u = 1$  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  $(\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$  has length 1. Check that  $u \cdot u = \frac{4}{9} + \frac{4}{9} + \frac{1}{9} = 1$ . For a unit vector, divide any nonzero  $v$  by its length  $\|v\|$ .

**Unit vector**  $u = v/\|v\|$  is a unit vector in the same direction as  $v$ .

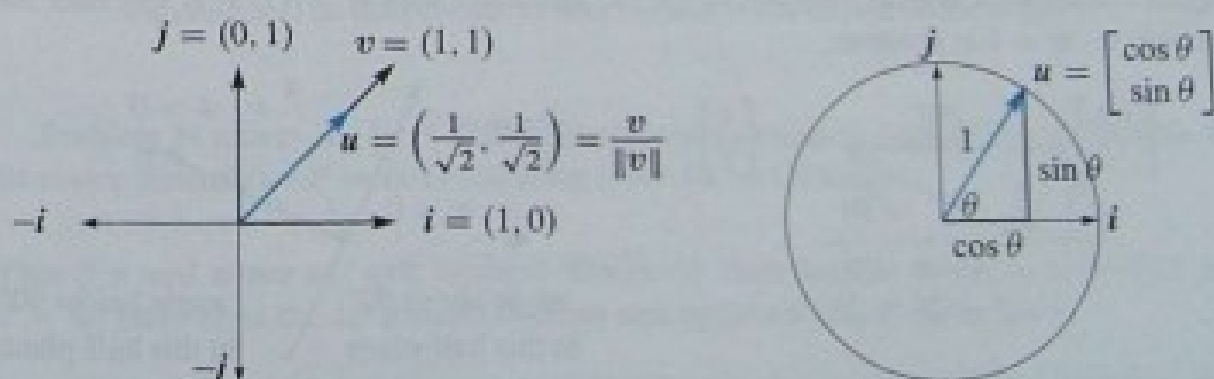


Figure 1.7: The coordinate vectors  $i$  and  $j$ . The unit vector  $u$  at angle  $45^\circ$  (left) divides  $v = (1, 1)$  by its length  $\|v\| = \sqrt{2}$ . The unit vector  $u = (\cos \theta, \sin \theta)$  is at angle  $\theta$ .

## The Angle Between Two Vectors

We stated that perpendicular vectors have  $v \cdot w = 0$ . The dot product is zero when the angle is  $90^\circ$ . To explain this, we have to connect angles to dot products. To see how  $v \cdot w$  finds the angle between any two nonzero vectors  $v$  and  $w$ .

26 / 58

### Right angles

*The dot product is  $v \cdot w = 0$  when  $v$  is perpendicular to  $w$ .*

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

$$\text{Perpendicular vectors} \quad \|v\|^2 + \|w\|^2 = \|v - w\|^2 \quad (2)$$

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

$$\text{Pythagoras} \quad (v_1^2 + v_2^2) + (w_1^2 + w_2^2) = (v_1 - w_1)^2 + (v_2 - w_2)^2. \quad (3)$$

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  $-2$ :

$$0 = -2v_1w_1 - 2v_2w_2 \quad \text{which leads to} \quad v_1w_1 + v_2w_2 = 0. \quad (4)$$

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

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

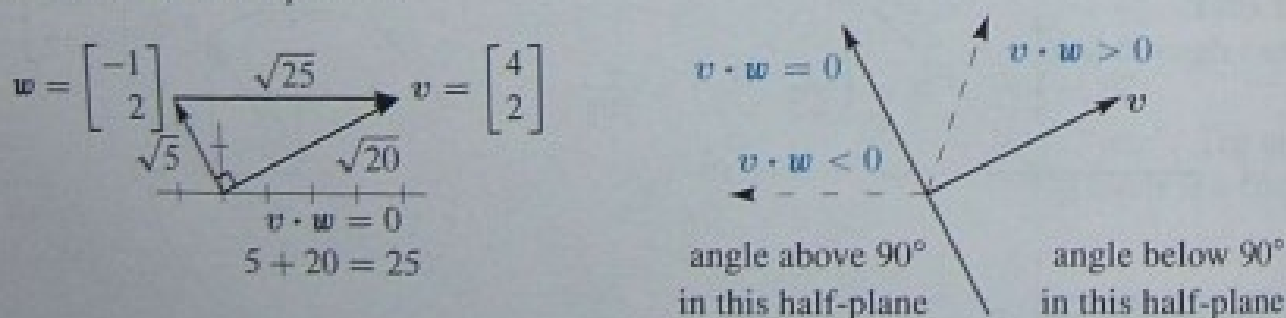


Figure 1.8: Perpendicular vectors have  $v \cdot w = 0$ . Then  $\|v\|^2 + \|w\|^2 = \|v - w\|^2$ .

The borderline is where vectors are perpendicular to  $v$ . 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're on the subject of angles, this is the place for the formula.

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

**Unit vectors  $u$  and  $U$  at angle  $\theta$  have  $u \cdot U = \cos \theta$ . Certainly  $|u \cdot U| \leq 1$ .**

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  $u = (\cos \theta, \sin \theta)$  and  $i = (1, 0)$ . The dot product is  $u \cdot i = \cos \theta$ . That is the cosine of the angle between them.

After rotation through any angle  $\alpha$ , these are still unit vectors. The vector  $i = (1, 0)$  rotates to  $(\cos \alpha, \sin \alpha)$ . The vector  $u$  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  $\cos \theta$ .

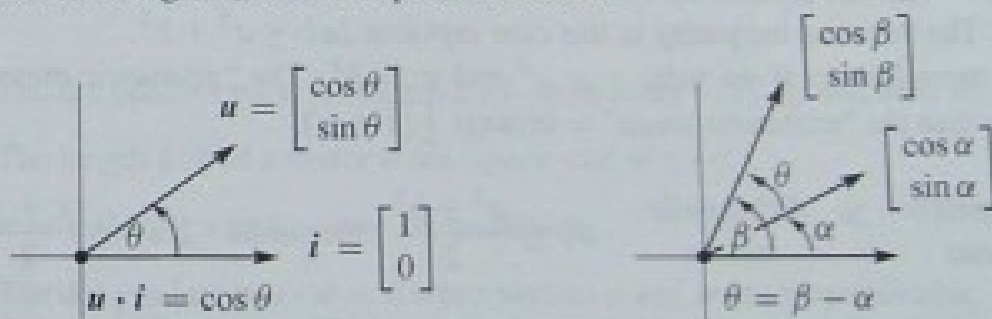


Figure 1.9: The dot product of unit vectors is the cosine of the angle  $\theta$ .

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

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

**COSINE FORMULA** If  $v$  and  $w$  are nonzero vectors then  $\frac{v \cdot w}{\|v\| \|w\|} = \cos \theta$ .

Whatever the angle, this dot product of  $v/\|v\|$  with  $w/\|w\|$  never exceeds 1. This is the “*Schwarz inequality*”  $|v \cdot w| \leq \|v\| \|w\|$  for dot products—or more generally the Cauchy-Schwarz-Buniakowsky inequality. It was found in France 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:

**SCHWARZ INEQUALITY**  $|v \cdot w| \leq \|v\| \|w\|$

**TRIANGLE INEQUALITY**  $\|v + w\| \leq \|v\| + \|w\|$

**Example 5** Find  $\cos \theta$  for  $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  and  $w = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  and check both inequalities.

**Solution** The dot product is  $v \cdot w = 4$ . Both  $v$  and  $w$  have length  $\sqrt{5}$ . The cosine is  $4/5$ .

$$\cos \theta = \frac{v \cdot w}{\|v\| \|w\|} = \frac{4}{\sqrt{5}\sqrt{5}} = \frac{4}{5}.$$

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  $\frac{1}{2}(x + y)$ .

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

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

```
FORTRAN      DO 10 J = 1,N          DO 10 J = 1,N
              10 VPLUSW(J) = v(J) + w(J)    10 VDOTW = VDOTW + 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.

Input  $v$  and  $w$  as rows—the prime  $'$  transposes them to columns.  $2v + 3w$  uses multiplication by 2 and 3. The result will be printed unless the line ends in a semicolon.

**MATLAB**  $v = [2 \ 3 \ 4]'$  ;  $w = [1 \ 1 \ 1]'$  ;  $u = 2 * v + 3 * w$

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

Instead of  $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}$  we more often see  $[1 \ 2] \begin{bmatrix} 3 \\ 4 \end{bmatrix}$  or  $v' * w$ .

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

**Cosine formula**

$$\text{cosine} = v' * w / (\text{norm}(v) * \text{norm}(w))$$

**Angle formula**

$$\text{angle} = \text{acos}(\text{cosine})$$

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  $v_i$  by  $w_i$  and adds all  $v_i w_i$ .
2. The length  $\|v\|$  of a vector is the square root of  $v \cdot v$ .
3.  $u = v / \|v\|$  is a *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 *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  $\|v + w\| = \|v\| + \|w\|$ ?

**Solution** The dot product is  $v \cdot w = (3)(4) + (4)(3) = 24$ . The length  $\|v\| = \sqrt{9 + 16} = 5$  and also  $\|w\| = 5$ . The sum  $v + w = (7, 7)$  has length

**Schwarz inequality**  $|v \cdot w| \leq \|v\| \|w\|$  is  $24 < 25$ .

**Triangle inequality**  $\|v + w\| \leq \|v\| + \|w\|$  is  $7\sqrt{2} < 5 + 5$ .

**Cosine of angle**  $\cos \theta = \frac{24}{25}$  This angle from  $v = (3, 4)$  to  $w = (4, 3)$

Suppose one vector is a multiple of the other as in  $w = cv$ . Then the angle is  $0^\circ$  or  $180^\circ$ . In this case  $|\cos \theta| = 1$  and  $|v \cdot w|$  equals  $\|v\| \|w\|$ . If the angle is  $0^\circ$ , as in  $w = 2v$ , then  $\|v + w\| = \|v\| + \|w\|$ . The triangle is completely flat.

**1.2 B** Find a unit vector  $u$  in the direction of  $v = (3, 4)$ . Find a unit vector  $U$  that is perpendicular to  $u$ . How many possibilities for  $U$ ?

**Solution** For a unit vector  $u$ , divide  $v$  by its length  $\|v\| = 5$ . For a perpendicular vector  $V$  we can choose  $(-4, 3)$  since the dot product  $v \cdot V$  is  $(3)(-4) + (4)(3) = 0$ . For a unit vector  $U$ , divide  $V$  by its length  $\|V\|$ :

$$u = \frac{v}{\|v\|} = \left( \frac{3}{5}, \frac{4}{5} \right) \quad U = \frac{V}{\|V\|} = \left( -\frac{4}{5}, \frac{3}{5} \right) \quad u \cdot U = 0$$

The only other perpendicular unit vector would be  $-U = \left( \frac{4}{5}, -\frac{3}{5} \right)$ .

**1.2 C** Find a vector  $x = (c, d)$  that has dot products  $x \cdot r = 1$  and  $x \cdot s = 0$  with the given vectors  $r = (2, -1)$  and  $s = (-1, 2)$ .

How is this question related to Example 1.1 C, which solved  $cv + dw = b = (1, 0)$ ?

**Solution** Those two dot products give linear equations for  $c$  and  $d$ . Then  $x = (c, d)$ .

$$\begin{array}{lll} x \cdot r = 1 & 2c - d = 1 & \text{The same equations as} \\ x \cdot s = 0 & -c + 2d = 0 & \text{in Worked Example 1.1 C} \end{array}$$

The second equation makes  $x$  perpendicular to  $s = (-1, 2)$ . So I can see the geometry: Go in the perpendicular direction  $(2, 1)$ . When you reach  $x = \frac{1}{3}(2, 1)$ , the dot product with  $r = (2, -1)$  has the required value  $x \cdot r = 1$ .

*Comment on  $n$  equations for  $x = (x_1, \dots, x_n)$  in  $n$ -dimensional space*

Section 1.1 would start with column vectors  $v_1, \dots, v_n$ . The goal is to combine them to produce a required vector  $x_1 v_1 + \dots + x_n v_n = b$ . This section would start from vectors  $r_1, \dots, r_n$ . Now the goal is to find  $x$  with the required dot products  $x \cdot r_j = b_j$ .

Soon the  $v$ 's will be the columns of a matrix  $A$ , and the  $r$ 's will be the rows of  $A$ . Then the (one and only) problem will be to solve  $Ax = b$ .



## Problem Set 1.2

- 1 Calculate the dot products  $u \cdot v$  and  $u \cdot w$  and  $u \cdot (v + w)$  and  $w \cdot v$ :

$$u = \begin{bmatrix} -.6 \\ .8 \end{bmatrix} \quad v = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad w = \begin{bmatrix} 8 \\ 6 \end{bmatrix}.$$

- 2 Compute the lengths  $\|u\|$  and  $\|v\|$  and  $\|w\|$  of those vectors. Check the Schwarz inequalities  $|u \cdot v| \leq \|u\| \|v\|$  and  $|v \cdot w| \leq \|v\| \|w\|$ .
- 3 Find unit vectors in the directions of  $v$  and  $w$  in Problem 1, and the cosine of the angle  $\theta$ . Choose vectors  $a, b, c$  that make  $0^\circ, 90^\circ,$  and  $180^\circ$  angles with  $w$ .
- 4 For any *unit* vectors  $v$  and  $w$ , find the dot products (actual numbers) of
- (a)  $v$  and  $-v$       (b)  $v + w$  and  $v - w$       (c)  $v - 2w$  and  $v + 2w$
- 5 Find unit vectors  $u_1$  and  $u_2$  in the directions of  $v = (3, 1)$  and  $w = (2, 1, 2)$ . Find unit vectors  $U_1$  and  $U_2$  that are perpendicular to  $u_1$  and  $u_2$ .
- 6 (a) Describe every vector  $w = (w_1, w_2)$  that is perpendicular to  $v = (2, -1)$ .  
 (b) The vectors that are perpendicular to  $V = (1, 1, 1)$  lie on a \_\_\_\_\_.  
 (c) The vectors that are perpendicular to  $(1, 1, 1)$  and  $(1, 2, 3)$  lie on a \_\_\_\_\_.
- 7 Find the angle  $\theta$  (from its cosine) between these pairs of vectors:

$$\begin{array}{ll} \text{(a) } v = \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix} \text{ and } w = \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \text{(b) } v = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \text{ and } w = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} \\ \text{(c) } v = \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix} \text{ and } w = \begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix} & \text{(d) } v = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ and } w = \begin{bmatrix} -1 \\ -2 \end{bmatrix}. \end{array}$$

- 8 True or false (give a reason if true or a counterexample if false):
- (a) If  $u$  is perpendicular (in three dimensions) to  $v$  and  $w$ , those vectors  $v$  and  $w$  are parallel.
- (b) If  $u$  is perpendicular to  $v$  and  $w$ , then  $u$  is perpendicular to  $v + 2w$ .
- (c) If  $u$  and  $v$  are perpendicular unit vectors then  $\|u - v\| = \sqrt{2}$ .
- 9 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  $v_2w_2/v_1w_1$  of those slopes is  $-1$ . Show that  $v \cdot w = 0$  and the vectors are perpendicular.
- 10 Draw arrows from  $(0, 0)$  to the points  $v = (1, 2)$  and  $w = (-2, 1)$ . Multiply their slopes. That answer is a signal that  $v \cdot w = 0$  and the arrows are \_\_\_\_\_.
- 11 If  $v \cdot w$  is negative, what does this say about the angle between  $v$  and  $w$ ? Draw 3-dimensional vector  $v$  (an arrow), and show where to find all  $w$ 's with  $v \cdot w < 0$ .

### 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  $w$ :

$$\text{First example} \quad u = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \quad w = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Their linear combinations in three-dimensional space are  $cu + dv + ew$ :

$$\text{Combinations} \quad c \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + e \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c \\ d - c \\ e - d \end{bmatrix}. \quad (1)$$

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:

$$\text{Same combination} \quad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} c \\ d \\ e \end{bmatrix} = \begin{bmatrix} c \\ d - c \\ e - d \end{bmatrix}. \quad (2)$$

is now  $A$  times  $x$

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:

$$\text{Matrix times vector} \quad Ax = \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} c \\ d \\ e \end{bmatrix} = cu + dv + ew. \quad (3)$$

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  $Ax$ . With new letters we see

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b. \quad (4)$$

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  $b$ ):

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

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  $x_2 - x_1$  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, columns 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  $x_1u + x_2v + x_3w$  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  $x_1, x_2, x_3$ :

	$x_1 = b_1$		
$Ax = b$	$-x_1 + x_2 = b_2$	<b>Solution</b>	$x_2 = b_1 + b_2$
	$-x_2 + x_3 = b_3$		$x_3 = b_1 + b_2 + b_3$

(6)

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.

Look at two specific choices 0, 0, 0 and 1, 3, 5 of the right sides  $b_1, b_2, b_3$ :

$$b = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ gives } x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \text{ gives } x = \begin{bmatrix} 1 \\ 1+3 \\ 1+3+5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix}.$$

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!

$$Ax = b \text{ is solved by } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 + b_2 \\ b_1 + b_2 + b_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad (7)$$

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  $x = Sb$ :

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad Sb = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

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  $A^{-1}$ :

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

*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  $A$ .*

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

### Problem Set 1.3

- 1 Find the linear combination  $2s_1 + 3s_2 + 4s_3 = b$ . Then write  $b$  as a matrix-vector multiplication  $Sx$ . Compute the dot products (row of  $S$ )  $\cdot x$ :

$$s_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad s_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad s_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ go into the columns of } S.$$

- 2 Solve these equations  $Sy = b$  with  $s_1, s_2, s_3$  in the columns of  $S$ :

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix}.$$

The sum of the first  $n$  odd numbers is \_\_\_\_\_.

- 3 Solve these three equations for  $y_1, y_2, y_3$  in terms of  $B_1, B_2, B_3$ :

$$Sy = B \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}.$$

Write the solution  $y$  as a matrix  $A = S^{-1}$  times the vector  $B$ . Are the columns of  $S$  independent or dependent?

- 4 Find a combination  $x_1w_1 + x_2w_2 + x_3w_3$  that gives the zero vector:

$$w_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad w_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad w_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}.$$

Those vectors are (independent) (dependent). The three vectors lie in a \_\_\_\_\_. The matrix  $W$  with those columns is *not invertible*.

- 5 The rows of that matrix  $W$  produce three vectors (I write them as columns):

$$r_1 = \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} \quad r_2 = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} \quad r_3 = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}.$$

Linear algebra says that these vectors must also lie in a plane. There must be many combinations with  $y_1r_1 + y_2r_2 + y_3r_3 = 0$ . Find two sets of  $y$ 's.

- 6 Which values of  $c$  give dependent columns (combination equals zero)?

$$\begin{bmatrix} 1 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 1 & c \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & c \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} c & c & c \\ 2 & 1 & 5 \\ 3 & 3 & 6 \end{bmatrix}$$