



CONTENTS

1	Lagrange Multipliers	5
1.1	Constrained Optimization	5
1.2	The Picture	5
1.2.1	Tangent Curves Have Parallel Gradients	6
1.3	The Method	7
1.4	The Soda Can, Without Substitution	8
1.5	A Probability Example: The Most Likely Spinner	9
1.6	What λ Means	10
1.7	Lagrange Multipliers with SymPy	10
2	Introduction to Linear Algebra	13
2.1	What is With the Linear?	13
2.2	Linear Combinations	14
2.3	Image Operations	14
2.4	The Numbers Behind Some Image Operations	16
2.5	Applications of Linear Algebra	22
2.6	Let us Observe the Sun!	23
2.7	Images in Python	24
2.8	Exercise	25
3	Vectors and Matrices	27
3.1	Matrices	27
3.1.1	Zero Matrices	28
3.2	Operations of Matrices	29

3.2.1	Adding and Subtracting Matrices	29
3.2.2	Multiplying Matrices	30
3.2.3	Transposing a Matrix	35
3.2.4	Row-Major Order, Column-Major Order, and Computer Memory	36
3.2.5	Properties of Transposes	37
3.3	Types of Matrices	37
3.3.1	Symmetric Matrices	39
3.3.2	Matrices in Python	40
3.3.3	Can We Divide Matrices?	41
4	Linear Combinations of Vectors	43
4.1	Visualizing Linear Combinations	46
4.2	Span in Three Dimensions	50
4.2.1	A Line	50
4.2.2	A Plane	51
4.2.3	All of Space	52
4.3	Summary	53
4.4	Where to Learn More	54
5	Matrices and Systems	55
5.1	Systems of Equations as Matrices	55
5.1.1	Systems of Equations, REF, and RREF	55
5.1.2	Row-Echelon Form	56
5.1.3	Larger RREF Example	60
5.1.4	Free Variables	64
5.1.5	Basis, Dimension, Rank	66
5.1.6	Creating Matrices in Python	69
5.1.7	Conclusion	71
5.2	Applications of Matrices in System Form	72
5.2.1	Trail Mix for Mars	72
5.2.2	Where to Learn More	75
6	Linear Independence and Dependence	77
6.1	Linear Independence	77
6.1.1	Connection to Span	80
6.2	Linearly Dependent Vectors: Geometrically	81
6.3	Three or More Vectors	83
6.4	What is next?	84

7	Subspaces	85
7.1	What is a Subspace?	85
7.2	Basis and Dimension	86
7.3	Nullspace	88
7.3.1	Linear Combinations and Span	89
7.3.2	Example: Nullspace of a 2 by 2 Matrix	89
7.3.3	Example: Nullspace of a 4 by 4 Matrix	90
7.4	Row Space	92
7.4.1	Example: Row Space	93
7.5	Column Space	94
7.5.1	Example: Column Space	94
7.6	Summary	96
A	Answers to Exercises	97
	Index	113

Lagrange Multipliers

Imagine hiking on a mountain, but you must stay on a marked trail. The highest point you can reach is probably not the summit. It is the highest point *on the trail*. Most real optimization problems look like this. A company wants the most profit, but it has a fixed budget. An engineer wants the strongest beam, but it must fit in a given space. A can maker wants to use the least aluminum, but the can must hold 355 cm^3 .

In the Optimization chapter, you solved problems like these by using the constraint to eliminate a variable. That works when the algebra cooperates. This chapter gives a method that works even when it does not, and it comes with a beautiful picture: at the best point, two gradients line up.

1.1 Constrained Optimization

The general problem is:

$$\text{maximize (or minimize) } f(x, y) \quad \text{subject to} \quad g(x, y) = c$$

The function f is the *objective*: the thing you care about. The equation $g(x, y) = c$ is the *constraint*: a curve in the plane that you must stay on.

Here is a simple example to keep in mind. You have 16 m of fence to make a rectangle. If the sides are x and y , then $2x + 2y = 16$, or $x + y = 8$, and you want to maximize the area $f(x, y) = xy$. (You can solve this one by substitution, but it is a good test case for the new method.)

1.2 The Picture

Draw the level curves of $f(x, y) = xy$: the hyperbolas $xy = 4$, $xy = 9$, $xy = 16$, and $xy = 25$ (the Multivariate Functions chapter introduced level curves). Then draw the constraint, the line $x + y = 8$ (figure 1.1).

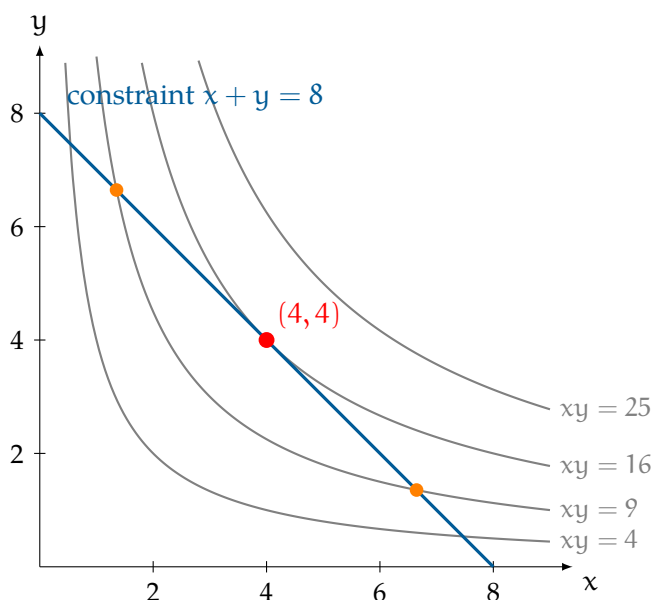


Figure 1.1: Level curves of the area $f = xy$ (gray) and the fence constraint $x + y = 8$ (blue). The line crosses $xy = 9$ at the orange points but only touches $xy = 16$, at $(4, 4)$.

Walk along the blue line from the top left. At first, you cross level curves with bigger and bigger values: $xy = 4$, then $xy = 9$. You are going uphill. At $(4, 4)$, the line just *touches* the level curve $xy = 16$ without crossing it, and after that you start crossing the same curves again in reverse: downhill. The curve $xy = 25$ is out of reach entirely. So the best area is 16 m^2 , at $x = y = 4$.

That is the whole idea: **at a constrained maximum or minimum, the constraint curve is tangent to a level curve of f** . If the path crossed the level curve instead, you could keep walking in one direction and climb higher.

1.2.1 Tangent Curves Have Parallel Gradients

From the Gradients chapter, the gradient ∇f points in the direction of steepest ascent, and it is perpendicular to the level curve through that point (moving along a level curve does not change f , so the gradient has no component along it). The constraint curve is a level curve of g , so ∇g is perpendicular to the constraint.

When the two curves are tangent, they share a tangent line, so their perpendiculars point along the same line. That means the gradients are parallel: one is a multiple of the other.

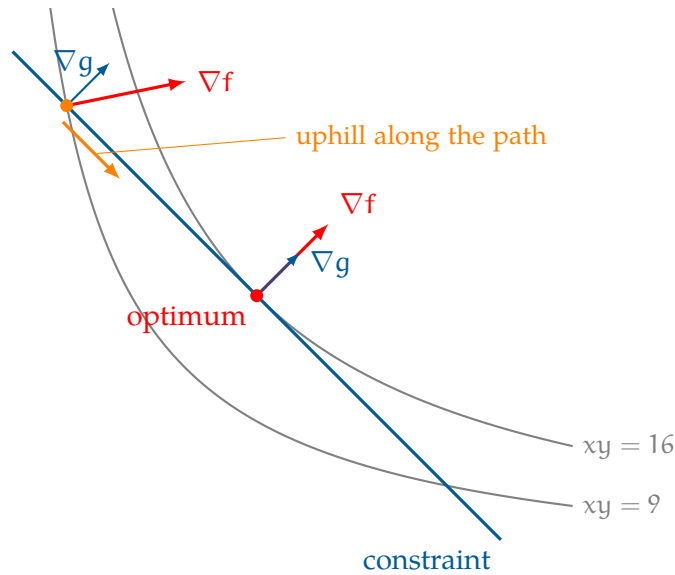


Figure 1.2: At the optimum, ∇f and ∇g point the same way. At the orange point, ∇f tilts away from ∇g . The part of ∇f that runs along the blue path (orange arrow) says you can still increase f by moving along the path.

Figure 1.2 shows the difference. At the orange point, $\nabla f = (6.65, 1.35)$ is not parallel to $\nabla g = (1, 1)$. Part of ∇f points along the fence line, so sliding along the line raises the area. At $(4, 4)$, $\nabla f = (4, 4)$ is exactly $4\nabla g$. No part of it points along the line, so there is nowhere better to slide.

1.3 The Method

Write the parallel-gradient condition with a multiplier λ (the Greek letter lambda), called a *Lagrange multiplier*, after Joseph-Louis Lagrange, who invented the method in 1788.

Lagrange Multipliers

To find the maximum and minimum of $f(x, y)$ subject to $g(x, y) = c$, solve

$$\nabla f = \lambda \nabla g \quad \text{and} \quad g(x, y) = c$$

that is, the three equations

$$f_x = \lambda g_x, \quad f_y = \lambda g_y, \quad g(x, y) = c$$

for x , y , and λ . Then evaluate f at every solution. The largest value is the maximum, and the smallest is the minimum.

The method finds *candidates*, just as setting a derivative to zero did in one variable. Always compare the values of f at all of them.

Example: Maximize $f(x, y) = xy$ subject to $x + y = 8$.

Solution: $\nabla f = (y, x)$ and $\nabla g = (1, 1)$. The equations are

$$y = \lambda, \quad x = \lambda, \quad x + y = 8$$

So $x = y$, and then $x = y = 4$, with $\lambda = 4$. The maximum area is 16, just as the picture showed.

Example: Find the maximum and minimum of $f(x, y) = x^2 + 2y^2$ on the unit circle $x^2 + y^2 = 1$.

Solution: $\nabla f = (2x, 4y)$ and $\nabla g = (2x, 2y)$. The equations are

$$2x = 2\lambda x, \quad 4y = 2\lambda y, \quad x^2 + y^2 = 1$$

The first says $x(1 - \lambda) = 0$, so $x = 0$ or $\lambda = 1$. The second says $y(2 - \lambda) = 0$, so $y = 0$ or $\lambda = 2$.

- If $\lambda = 1$, then $y = 0$, and the constraint gives $x = \pm 1$. There, $f = 1$.
- If $\lambda = 2$, then $x = 0$, and the constraint gives $y = \pm 1$. There, $f = 2$.

($x = y = 0$ is not on the circle.) The maximum is 2, at $(0, \pm 1)$, and the minimum is 1, at $(\pm 1, 0)$. Notice that one set of equations found both.

1.4 The Soda Can, Without Substitution

In the Optimization chapter, you found the can that holds 355 cm^3 using the least aluminum. Here is the same problem with Lagrange multipliers. Minimize the surface area $A(r, h) = 2\pi r^2 + 2\pi rh$ subject to the volume $V(r, h) = \pi r^2 h = 355$. The gradient equations are

$$4\pi r + 2\pi h = \lambda(2\pi rh), \quad 2\pi r = \lambda(\pi r^2)$$

The second gives $\lambda = \frac{2}{r}$. Put that into the first:

$$4\pi r + 2\pi h = \frac{2}{r}(2\pi rh) = 4\pi h \implies 4\pi r = 2\pi h \implies h = 2r$$

Look at what happened. Before using the number 355 at all, we learned the *shape* of the best can: its height equals its diameter. That holds for a can of any size. Then $\pi r^2(2r) = 355$ gives $r \approx 3.84 \text{ cm}$ and $h \approx 7.67 \text{ cm}$.

(Real soda cans are about 12 cm tall and 6.6 cm wide, much taller than the ideal. They are easier to hold, the top and bottom are made of thicker metal than the sides, and a narrow can fits more neatly in a vending machine.)

1.5 A Probability Example: The Most Likely Spinner

A spinner has three colors, but you do not know how big each section is. You spin it 20 times and get red 10 times, green 6 times, and blue 4 times. What are the most believable probabilities p , q , and s for red, green, and blue?

Spins are independent, so the probability of getting this particular sequence of results is the product of the individual chances:

$$P = p^{10} q^6 s^4$$

(For example, two heads in a row with a fair coin has probability $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.) The probabilities must add to 1, so we want to

$$\text{maximize } p^{10} q^6 s^4 \quad \text{subject to } p + q + s = 1$$

This is called finding the *maximum likelihood* estimate: the probabilities that make what you actually saw most likely.

Products of powers are awkward to differentiate, so take the natural log first. Since $\ln$ always increases, P and $\ln P$ are largest at the same point, and

$$\ln P = 10 \ln p + 6 \ln q + 4 \ln s$$

The method works the same way with three variables: set $\nabla(\ln P) = \lambda \nabla(p + q + s)$:

$$\frac{10}{p} = \lambda, \quad \frac{6}{q} = \lambda, \quad \frac{4}{s} = \lambda$$

So $p = \frac{10}{\lambda}$, $q = \frac{6}{\lambda}$, and $s = \frac{4}{\lambda}$. The constraint gives $\frac{20}{\lambda} = 1$, so $\lambda = 20$, and

$$p = \frac{10}{20} = 0.5, \quad q = \frac{6}{20} = 0.3, \quad s = \frac{4}{20} = 0.2$$

The most likely probabilities are just the fractions you observed. That may sound obvious, but now you have a proof, and the same argument works for any number of outcomes and any counts. It is the starting point for much of statistics and machine learning, which you will meet in later volumes.

1.6 What λ Means

The multiplier is more than a bookkeeping device. It tells you how much the best value of f changes when you loosen the constraint: if c increases by a small amount Δc , the optimum changes by about $\lambda \Delta c$.

Check it on the fence. With $x + y = c$, the best rectangle is a square with side $\frac{c}{2}$, so the best area is $\frac{c^2}{4}$. Its derivative is $\frac{c}{2}$, which is 4 when $c = 8$: exactly the λ we found. So one more meter of $x + y$ (two more meters of fence) buys about 4 more square meters of garden. Economists call λ a *shadow price*: what one more unit of a limited resource is worth.

1.7 Lagrange Multipliers with SymPy

The equations can get messy, and a computer algebra system can solve them for you. Here is the circle example in Python with the SymPy library:

```

1 import sympy as sp
2
3 x, y, lam = sp.symbols("x y lambda", real=True)
4 f = x**2 + 2*y**2          # what we optimize
5 g = x**2 + y**2           # the constraint is g = 1
6
7 # grad f = lambda * grad g, together with the constraint
8 equations = [
9     sp.Eq(sp.diff(f, x), lam * sp.diff(g, x)),
10    sp.Eq(sp.diff(f, y), lam * sp.diff(g, y)),
11    sp.Eq(g, 1),
12 ]
13 for sol in sp.solve(equations, [x, y, lam], dict=True):
14     print(sol, " f =", f.subs(sol))

```

```

{lambda: 1, x: -1, y: 0} f = 1
{lambda: 1, x: 1, y: 0} f = 1
{lambda: 2, x: 0, y: -1} f = 2
{lambda: 2, x: 0, y: 1} f = 2

```

All four candidate points appear, with their values of f .

Exercise 1 On a Circle

Find the maximum and minimum of $f(x, y) = x + y$ on the circle $x^2 + y^2 = 8$. Draw a sketch like figure 1.1 to check

Working Space

your answer.

Answer on Page 97

Exercise 2 Fencing by a River

A farmer has 100 m of fence to enclose a rectangular field along a straight river. No fence is needed along the river. Let x be the side parallel to the river and y each side perpendicular to it.

1. Use Lagrange multipliers to find the largest possible area.
2. What is λ , and what does it tell the farmer?

Working Space

Answer on Page 97

Exercise 3 The Biggest Box

A closed box has surface area $2(xy + yz + xz) = 24$. Find the largest possible volume xyz . (Hint: with three variables, there are four equations.)

Working Space

Answer on Page 97

Exercise 4 An Ellipse

Find the maximum and minimum of $f(x, y) = xy$ on the ellipse $x^2 + 4y^2 = 8$.

Working Space

Then check your answer with SymPy.

Answer on Page 97

Exercise 5 The Most Uncertain Spinner

A measure of how unpredictable a spinner with probabilities p , q , and s is, called its *entropy*, is

$$H = -p \ln p - q \ln q - s \ln s$$

Use Lagrange multipliers to find the probabilities that make H as large as possible, subject to $p + q + s = 1$. Does the answer make sense?

Working Space

Answer on Page 98

Introduction to Linear Algebra

Every time you rotate your phone and the screen turns, apply a filter to a photo, or play a video game where the camera swings around a 3D world, a computer is doing linear algebra: millions of multiplications and additions on grids of numbers, every second. Linear algebra is the mathematics of vectors, matrices, and linear transformations. You have met all three already; in this book you will see how they work together.

Let us review what you know.

- **Vectors.** Earlier, you saw how vectors can represent forces, as well as how to add and multiply them to figure out such things as rocket engine force and direction.
- **Matrices.** In the Basic Statistics and Spreadsheets chapter, you learned to use spreadsheets to solve problems numerically, such as how to figure out the number of barrels a cooper has to produce to achieve a certain take-home pay. Spreadsheets are essentially matrices — a row by column structure that contains values.
- **Linear transformations.** When you studied congruence, you met transformations such as rotation, reflection, and translation. Rotations and reflections that keep the origin fixed are linear; translation, as you will see, is not.

2.1 What is With the Linear?

You might be thinking, “Have I not been doing algebra all along?”

You have! You have come a long way in your problem solving journey. You have used algebra to solve simple equations like $7x + 10 = 24$ and quadratic equations like $4x^2 + 9x + 31 = 0$. What distinguishes linear algebra is the focus on linear combinations. Any equation with a power greater than 1, such as a quadratic, is nonlinear. Want to see what we will be covering? Take a look at 3blue1brown’s website with papers and videos on Linear Algebra: <https://www.3blue1brown.com/topics/linear-algebra>

We will first take a look at linear combinations

2.2 Linear Combinations

A *linear combination* uses only two operations: adding things together and multiplying them by numbers (scalars). For example, $3\mathbf{u} + 2\mathbf{v}$ is a linear combination of the vectors $\mathbf{u}$ and $\mathbf{v}$. No squaring, no square roots, no sin of a variable. An operation is *linear* if it plays nicely with both: doing it to a sum gives the sum of the results, and doing it to a scaled input gives a scaled result. You will see that linear combinations allow you to solve many types of problems in science and engineering. Before we get deep into the numbers, let us take a look at a few linear operations you can perform on images. This will give you an intuition for the underlying math. After that, we will take a look at some numbers.

2.3 Image Operations

The simplest image, a bitmap, can be represented by a two-dimensional matrix of values — either 0 for black, or 1 for white. Grayscale images are also represented by a two-dimensional matrix of values, but the values typically range from 0 to 255. 0 is black, 255 is white, and the values in between represent shades of gray.

Color images are more complex. The simplest color image is a three-dimensional matrix of values. You can think of it as three 2D matrices, one to represent red values (R), another for green values (G), and the third for blue values (B). The combination of R, G, and B determines the color you see.

Working with images means working with millions of pixels. Fortunately, modern techniques make this a snap. Let us look at some common operations on an image of a rocket. Flipping is a linear operation: each pixel just moves to a new spot.

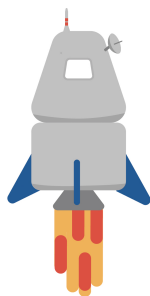


Figure 2.1: Rocket image.

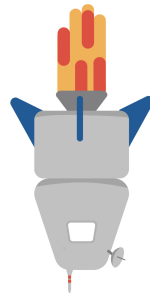


Figure 2.2: The rocket, flipped upside down.

Next, the image is rotated 90 degrees. Rotation is linear too, at any angle. To rotate by an angle that is not a multiple of 90, you will need $\cos \theta$ and $\sin \theta$, but those are just fixed numbers once you pick the angle. The rotation itself only multiplies and adds coordinates.

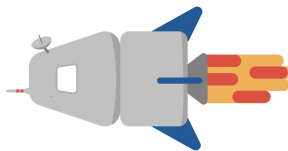


Figure 2.3: Rotating the rocket by 90°

Inversion replaces each red, green, and blue value with 1.0 minus the old value. The white background turns black, and the rocket seems to fly through deep space. Strictly speaking, inversion is not linear: because of the “1.0 minus,” inverting a black pixel (0) gives white (1), not 0. Operations like this, a linear operation plus a shift, are called *affine*. Translation is affine for the same reason.

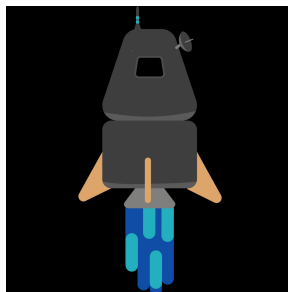


Figure 2.4: Inverting the rocket replaces each pixel with its “opposite” color.

It is possible to redefine the red, green, and blue values in many ways. Visit the NASA website and search for false color images. NASA and other scientists redefine colors to

communicate such things as the amount of vegetation or water in an area, the temperatures of the sun's surface, and so on. Photographers often do this for artistic effect. For example, the image on the left was taken with an infrared camera. (This is near-infrared light, not the thermal infrared that heat cameras see. Healthy plants reflect near-infrared strongly.) The image is further processed to swap channels. For example, the matrix representing red might be swapped with the matrix representing blue. The image on the right shows the image after swapping color values. All these swapping operations are linear.

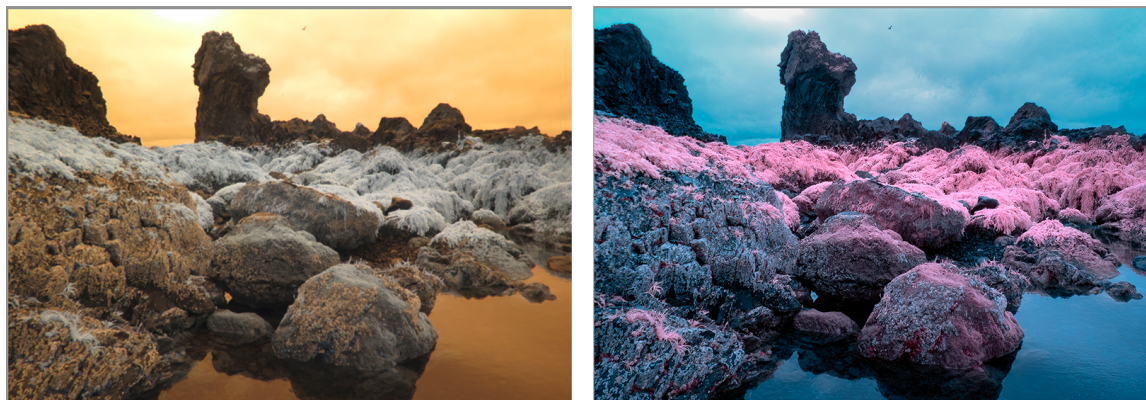


Figure 2.5: Left: image with infrared camera. Right: image shown post swapping color values

2.4 The Numbers Behind Some Image Operations

You will see a few *matrices* in this section. Let us first look at how a spreadsheet can be represented as a matrix. Recall the barrel-making shop example. This is part of that spreadsheet.

	A	B	C	D
1	Barrels Produced (per month)	115	120	125
2	Materials cost (per barrel)	\$45.00	\$45.00	\$45.00
3	Sale price (per barrel)	\$100.00	\$100.00	\$100.00
4	Rent (per month)	\$2,000.00	\$2,000.00	\$2,000.00
5	Pretax Earnings (per month)	\$4,325.00	\$4,600.00	\$4,875.00
6	Taxes (per month)	\$865.00	\$920.00	\$975.00
7	Take home pay (per month)	\$3,460.00	\$3,680.00	\$3,900.00

Figure 2.6: Spreadsheet of the barrel example.

Represented as a matrix, it looks like the following. Note the differences: A matrix con-

tains only values, no labels. This matrix uses floating point values, hence the inclusion of decimal points.

$$\begin{bmatrix} 115.0 & 120.0 & 125.0 \\ 45.0 & 45.0 & 45.0 \\ 100.0 & 100.0 & 100.0 \\ 2000.0 & 2000.0 & 2000.0 \\ 4325.0 & 4600.0 & 4875.0 \\ 865.0 & 920.0 & 975.0 \\ 3460.0 & 3680.0 & 3900.0 \end{bmatrix}$$

A matrix that represents an image contains only pixel values, whereas the barrel-making shop matrix represents seven kinds of variables: barrels produced, materials cost, sales price, rent, pretax earnings, taxes, and take home pay.

The simplest image to create is a bitmap, because that requires a matrix of zeros and ones. This is a matrix for a 10 pixel by 10 pixel image. Why use decimal points when this is obviously a matrix of integers? It turns out that when you use tools like Python to process matrices, you must be conscious of data types. Most of the Python methods we use for image operations expect floating points. A few expect integer types, but you will see how to handle type conversion later, in the section on Python.

$$\begin{bmatrix} 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 1. & 1. & 1. & 1. & 1 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 1. & 1. & 1. & 1. & 1 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \\ 0. & 0. & 0. & 0. & 1. & 0. & 0. & 0. & 1. & 0 \end{bmatrix}$$

When converted to an image, it is very tiny. This is an enlarged version, so you can see the pattern.

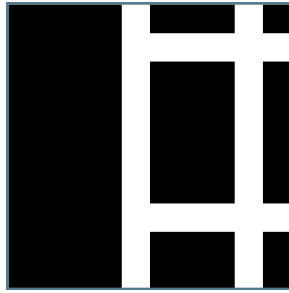


Figure 2.7: Resulting bitmap vectors as a black and white image.

We can invert this image by changing all the values in the matrix so that 0 becomes 1 and 1 becomes 0. (Do not confuse this with the *inverse of a matrix*, which you will meet later and which means something entirely different.)

$$\begin{bmatrix} 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 0. & 0. & 0. & 0. & 0 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 0. & 0. & 0. & 0. & 0 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \\ 1. & 1. & 1. & 1. & 0. & 1. & 1. & 1. & 0. & 1 \end{bmatrix}$$

When converted to an image and enlarged, it looks like this:

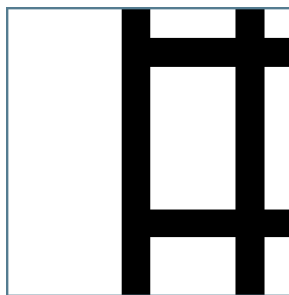


Figure 2.8: The inverted bitmap.

Rotating the original matrix by 90 degrees gives this:

$$\begin{bmatrix}
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.
 \end{bmatrix}$$

This is the resulting enlarged image:

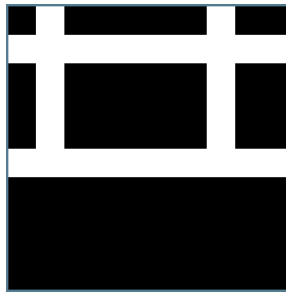


Figure 2.9: The bitmap rotated 90° counterclockwise.

You will transpose many matrices in the upcoming pages. It requires swapping rows for columns.

$$\begin{bmatrix}
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0. \\
 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. \\
 0. & 1. & 0. & 0. & 0. & 0. & 0. & 1. & 0. & 0.
 \end{bmatrix}$$

The resulting image looks like this:

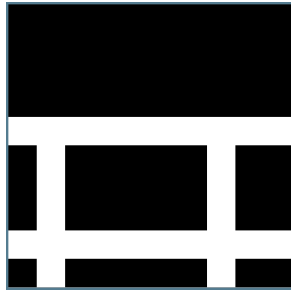


Figure 2.10: Transposed image reversing rows and columns.

What about adding images? That fits the definition of a linear combination. Recall that grayscale images have values from 0 to 255. To make things simple, let us define two matrices with values ranging from 0.0 to 1.0. When we want a grayscale image, it is easy to multiply the matrix by 255.

Let us call this matrix f .

$$\begin{bmatrix} 1.0 & 0.5 & 0.0 \\ 1.0 & 1.0 & 1.0 \\ 0.5 & 0.0 & 0.0 \end{bmatrix}$$

When multiplied by 255 and converted to a grayscale image:

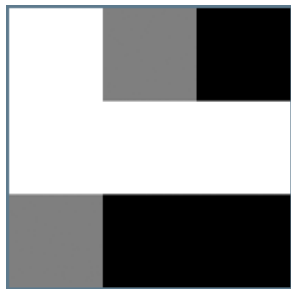


Figure 2.11: Bitmap of matrix f .

Let us call this matrix g :

$$\begin{bmatrix} 0.5 & 0.0 & 0.0 \\ 1.0 & 0.5 & 1.0 \\ 1.0 & 1.0 & 1.0 \end{bmatrix}$$

When multiplied by 255 and converted to a grayscale image:

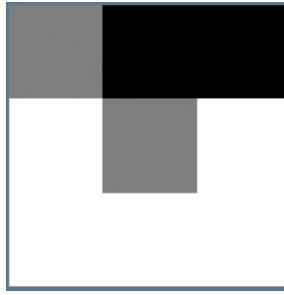


Figure 2.12: Bitmap of matrix g.

When we add f and g we get k:

$$\begin{bmatrix} 1.5 & 0.5 & 0.0 \\ 2.0 & 1.5 & 2.0 \\ 1.5 & 1.0 & 1.0 \end{bmatrix}$$

However, the values in k exceed the range of 0.0 to 1.0, so we normalize by dividing the matrix by 2.0:

$$\begin{bmatrix} 0.75 & 0.25 & 0.00 \\ 1.00 & 0.75 & 1.00 \\ 0.75 & 0.50 & 0.50 \end{bmatrix}$$

When multiplied by 255 and converted to grayscale, we get:

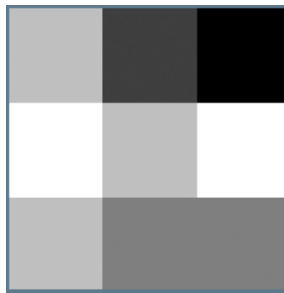


Figure 2.13: Adding f + g represented as a bitmap.

Let us go back to the first small grayscale image (from Figure 2.11):



Figure 2.14: Figure 2.11, repeated.

To keep only the first column, multiply each row of the matrix, entry by entry, by the vector $[1.0 \ 0.0 \ 0.0]$. The 1.0 keeps the values in the first column; the 0.0s knock out the rest, because 0.0 times anything is 0.0.

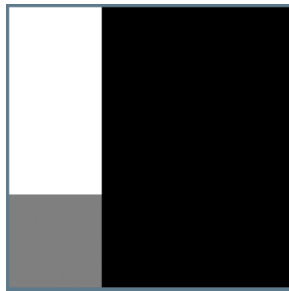


Figure 2.15: Multiplying columns by 0 turns them black, while 1 keeps them the same.

What do you think will happen if you use the vector $[0.0 \ 1.0 \ 0.0]$ or $[0.0 \ 0.0 \ 1.0]$? In a color image, the same trick with the red, green, and blue values isolates one color channel.

Except for inversion, all of these operations are linear: they preserve addition and scalar multiplication.

2.5 Applications of Linear Algebra

So far you have seen how linear operations on matrices can process images by:

- multiplying a matrix using a scalar (e.g., normalize, change the range)
- adding one matrix to another to get a composite image
- multiplying two matrices to perform a transform (e.g., flipping)
- multiplying entry by entry (isolating a column or a channel)

Many areas in engineering and science rely on the matrix operations defined by linear algebra. Besides image processing, linear algebra is used for:

- **Computer Graphics.** When you play a video game or watch the latest CG animation, matrix operations transform objects in the scene to make them appear as if moving, getting closer, and so on. You can represent the vertices of objects as vectors, and then apply a transformation matrix.
- **Data Analysis.** We live in an era in which it is easy to collect so much data that it is difficult to make sense of the data by just looking at it. You can represent the data in matrix form and then find a solution vector. For example, scientists use this technique to figure out the effectiveness of drug treatments on disease.
- **Economics.** Take a look at financial section of any news organization and you will see headlines such as "Economic Data Points to Faster Growth" or "Is the Inflation Battle Won?" Economists can use systems of linear equations to represent economic indicators, such as consumer consumption, government spending, investment rate, and gross national product. By using various methods that you will learn about later, they can get a good idea of the state of the economy.
- **Engineering.** Engineers could not do without linear algebra. For example, the orbital dynamics of space travel relies on it. Engineers must predict and calculate the motion of planetary bodies, satellites, and spacecraft. By solving systems of linear equations engineers can make sure that a spacecraft travels to its destination without running into a satellite or space rock.

2.6 Let us Observe the Sun!

In 2023, India launched the Aditya-L1 spacecraft to study the Sun. Without a thorough understanding of linear algebra (among other things), the engineers would not have accomplished the amazing feat of getting Aditya in a stable orbit around a Lagrange point.

In previous chapters you learned about gravity and its effects. A Lagrange point is a spot where the pulls of two bodies (here, the Earth and the Sun) balance so that a spacecraft can keep pace with the Earth around the Sun. Aditya-L1 went to L1, about 1.5 million km from Earth toward the Sun. With the right trajectory, a spacecraft will orbit around a Lagrange point in a stable position that does not require much energy to maintain. This is called a halo orbit. Because there are no fueling stations in space, a Halo orbit will allow Aditya to maintain position for about 5 years. Pretty good mileage!

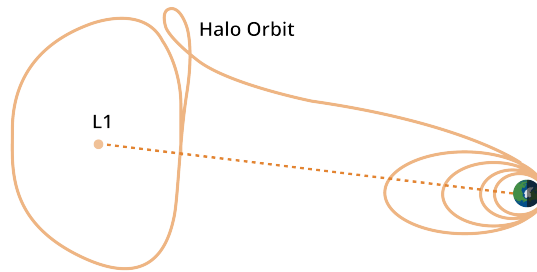


Figure 2.16: The Aditya Orbit.

Aditya's engineers had to calculate a looping maneuver that would precisely inject the Aditya spacecraft into the Halo orbit. They determined the angles and burn times for the thrust engine. If they were wrong in one direction, the spacecraft would fly off to the sun. The other direction would send the spacecraft back in the direction of Earth. Their success is due to a solid understanding of vectors and linear algebra.

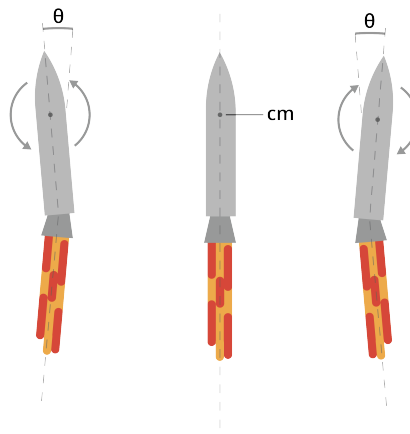


Figure 2.17: A gimbaled engine tilts by a small angle θ to steer the rocket.

2.7 Images in Python

One of the wonderful things about python is the availability of libraries for specialized computation. The Python Imaging Library, PIL, is what you will use to create images, read existing images from disk, and perform operation on images. To create and manipulate arrays, you will use NumPy.

Create a file called `image_creation.py` and enter this code. It makes each of the images in this chapter:

```

1  import numpy as np
2  from PIL import Image
3
4  # A 10 by 10 bitmap: 0 is black, 1 is white.
5  # The decimal points make these floating-point numbers.
6  bitmap = np.array([
7  [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
8  [0., 0., 0., 0., 1., 1., 1., 1., 1., 1.],
9  [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
10 [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
11 [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
12 [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
13 [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
14 [0., 0., 0., 0., 1., 1., 1., 1., 1., 1.],
15 [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.],
16 [0., 0., 0., 0., 1., 0., 0., 0., 1., 0.]])
17
18 def show(array):
19     # Scale 0-1 up to 0-255, convert to bytes, and enlarge 20x so you can see it
20     pixels = (array * 255).astype(np.uint8)
21     Image.fromarray(pixels).resize((200, 200), Image.NEAREST).show()
22
23 show(bitmap)                # the original
24 show(bitmap.transpose())    # rows become columns
25 show(np.rot90(bitmap))      # rotated 90 degrees counterclockwise
26 show(np.flipud(bitmap))     # flipped upside down
27 show(1.0 - bitmap)          # inverted: black becomes white

```

2.8 Exercise

Exercise 6 Linear or Not?

An operation T is linear if $T(a + b) = T(a) + T(b)$ and $T(ca) = cT(a)$ for all inputs a , b and numbers c . Which of these operations on numbers are linear?

1. $T(x) = 3x$
2. $T(x) = 3x + 1$
3. $T(x) = x^2$

Working Space

4. $T(x) = -x$

Answer on Page 98

Exercise 7 Matrix Operations in Python

Using NumPy, create the matrices $f = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Compute $f+g$, the matrix product fg , and the transpose f^T .

Working Space

Answer on Page 98

Vectors and Matrices

A spreadsheet of your class's test scores, a photo on your phone, and the positions of every vertex in a video game character all have something in common: they are grids of numbers. Mathematicians call a grid of numbers a *matrix*. In this chapter, you will learn to add, multiply, and transpose matrices, by hand and with Python.

3.1 Matrices

We have been looking at vectors. We have seen them in physics as a straight line comprised of *x* and *y* components, or represented as a column of numbers. For example, while we may write $\mathbf{v} = [1, 2, 3]$ in line, the vector is really:

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

A matrix can be made of many columns, like the 3×3 matrix shown below:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \\ 3 & 9 & 27 \end{bmatrix}$$

We describe the size and shape of matrices by saying *an* $m \times n$ *matrix*, where *m* is the number of *rows* and *n* is the number of *columns*. A *vector* is simply a one-column matrix. For example, the vector $\mathbf{v}$ above is of size 3×1 . Grids of numbers can also have more than two dimensions: a $3 \times 2 \times 4$ array is 4 stacked 3×2 matrices. (A color image is height $\times$ width $\times$ 3.) In this book, “matrix” means two dimensions.

A matrix is considered *square* when it has the same number of rows and columns. Some important operations work only on square matrices: the determinant, the inverse, and eigenvalues, all of which you will learn in future chapters.

Exercise 8 Matrix Dimensions 1

Write the dimensions of the following matrices:

Working Space

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

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

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

Answer on Page 98

Exercise 9 Matrix Dimensions 2

Create a matrix with the indicated dimensions.

Working Space

1. 1×3

2. 2×4

3. 4×3

Answer on Page 99

3.1.1 Zero Matrices

Recall that we can represent a generic zero vector as $\mathbf{0}$ or $\vec{0}$ (you may see both), which indicates a vector of any number of dimensions filled with zeros. Just like vectors, there are *zero matrices*, which can be any number of dimensions, all filled with zeros. In two dimensions, zero matrices are denoted as $0_{m \times n}$, where the subscript is the dimension of the matrix. The subscript can be expanded to denote any number of dimensions.

3.2 Operations of Matrices

3.2.1 Adding and Subtracting Matrices

Matrices that are the same dimension can be added and subtracted. Just like vectors, to add matrices you add the elements in the same position:

$$\begin{bmatrix} -2 & -1 \\ 2 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 2 \\ -1 & -4 \end{bmatrix} = \begin{bmatrix} -2+5 & -1+2 \\ 2+(-1) & 4+(-4) \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}$$

And to subtract matrices, you subtract the elements in the same position:

$$\begin{bmatrix} -2 & -1 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 2 \\ -1 & -4 \end{bmatrix} = \begin{bmatrix} -2-5 & -1-2 \\ 2-(-1) & 4-(-4) \end{bmatrix} = \begin{bmatrix} -7 & -3 \\ 3 & 8 \end{bmatrix}$$

Formally, for 2-dimensional matrices, we can say that:

Adding and Subtracting Matrices

For two $m \times n$ matrices, the sum of the matrices is the matrix of the sums of the elements in analogous positions:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mn} \end{bmatrix} + \begin{bmatrix} y_{11} & y_{12} & y_{13} & \cdots & y_{1n} \\ y_{21} & y_{22} & y_{23} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y_{m1} & y_{m2} & y_{m3} & \cdots & y_{mn} \end{bmatrix} = \begin{bmatrix} x_{11} + y_{11} & x_{12} + y_{12} & x_{13} + y_{13} & \cdots & x_{1n} + y_{1n} \\ x_{21} + y_{21} & x_{22} + y_{22} & x_{23} + y_{23} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{m1} + y_{m1} & x_{m2} + y_{m2} & x_{m3} + y_{m3} & \cdots & x_{mn} + y_{mn} \end{bmatrix}$$

To subtract matrices, simply add the negative of the second matrix (that is, $A - B = A + -B$). Additionally, matrix addition is commutative ($A + B = B + A$).

Matrices of different dimensions cannot be added or subtracted.

Exercise 10 Adding and Subtracting Matrices

Find $A + B$, $A - B$, and $B - A$.

Working Space

1. $A = \begin{bmatrix} 0 & 4 & 0 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 3 & -2 & 5 \end{bmatrix}$

2. $A = \begin{bmatrix} 4 & -4 & -2 \\ 1 & -3 & 5 \\ -5 & 3 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 0 & -1 \\ -5 & -3 & -2 \\ -5 & 3 & -4 \end{bmatrix}$.

3. $A = \begin{bmatrix} -2 & -1 & -5 & -1 \\ 5 & -4 & 4 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} -5 & -2 & 3 & -5 \\ 0 & 5 & -4 & -3 \end{bmatrix}$.

Answer on Page 99

3.2.2 Multiplying Matrices

Matrix multiplication has dimensional limits. We cannot multiply any two matrices; the first matrix must have the same number of columns as the second has number of rows. Let us examine the origin of the dimension limits on matrix multiplication. We begin with a review of the vector dot product.

Recall that in order to find the dot product of two vectors, they must be the same length (that is, the same number of dimensions). The result is always a scalar: one number. You can review finding the dot product of vectors and practice the dimension limits on the vector dot product in the next exercise.

Exercise 11 Vector Dot Product Review

Find all possible pairs of vectors that can be used to find a dot product, then find the dot products.

Working Space

1. $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

2. $\mathbf{b} = \begin{bmatrix} -3 \\ 3 \\ 5 \\ -5 \end{bmatrix}$

$$3. \mathbf{c} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$4. \mathbf{d} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$$

$$5. \mathbf{e} = \begin{bmatrix} 1 \\ -5 \\ 3 \\ 1 \end{bmatrix}$$

$$6. \mathbf{f} = \begin{bmatrix} 4 \\ 1 \\ -3 \end{bmatrix}$$

Answer on Page 99

To multiply two matrices, it is helpful to think of the rows of the first matrix and the columns of the second matrix as vectors. Let us see how this shakes out for two 2×2 matrices in Figure 3.1:

$$\begin{array}{c} \mathbf{a}_1 \rightarrow \\ \mathbf{a}_2 \rightarrow \end{array} \begin{bmatrix} 5 & 4 \\ -5 & 1 \end{bmatrix} \times \begin{array}{c} \mathbf{b}_1 \downarrow \quad \mathbf{b}_2 \downarrow \\ \begin{bmatrix} -1 & -2 \\ -5 & -4 \end{bmatrix} \end{array} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 \end{bmatrix}$$

$$\mathbf{A} \quad \times \quad \mathbf{B}$$

Figure 3.1: Each entry in $\mathbf{C}$, c_{ij} , is the dot product of the i^{th} row of $\mathbf{A}$, $\mathbf{a}_i$, and the j^{th} column of $\mathbf{B}$, $\mathbf{b}_j$.

Let us look at this more concretely. For two-dimensional matrices, it can be helpful to move your left index finger across the row and right index finger down the column, as shown in Figure 3.2.

$$\begin{array}{c}
 \begin{bmatrix} 5 & 4 \\ -5 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -2 \\ 5 & -4 \end{bmatrix} = \begin{bmatrix} 5 \times -1 + 4 \times -4 \\ -5 \times 1 + 1 \times -4 \end{bmatrix} \\
 \begin{bmatrix} 5 & 4 \\ -5 & 1 \end{bmatrix} \times \begin{bmatrix} -1 & -2 \\ -5 & -4 \end{bmatrix} = \begin{bmatrix} -21 & 5 \times -2 + 4 \times -4 \\ -5 \times -1 + 1 \times -5 \end{bmatrix} \\
 \begin{bmatrix} 5 & 4 \\ -5 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -2 \\ 5 & -4 \end{bmatrix} = \begin{bmatrix} -21 & -26 \\ -5 \times -1 + 1 \times -5 \end{bmatrix} \\
 \begin{bmatrix} 5 & 4 \\ -5 & 1 \end{bmatrix} \times \begin{bmatrix} -1 & -2 \\ -5 & -4 \end{bmatrix} = \begin{bmatrix} -21 & -26 \\ 0 & -5 \times -2 + 1 \times -4 \end{bmatrix} \\
 \begin{bmatrix} 5 & 4 \\ -5 & 1 \end{bmatrix} \times \begin{bmatrix} -1 & -2 \\ -5 & -4 \end{bmatrix} = \begin{bmatrix} -21 & -26 \\ 0 & 6 \end{bmatrix}
 \end{array}$$

Figure 3.2: You can use your fingers to trace across matrix A and down matrix B to find $A \cdot B$.

Since each entry in the product matrix is the dot product between a row of the first matrix and a column of the second matrix, the first matrix must have the same number of elements in each row as the second has in each column. Another way to say this is that the number of columns of the first matrix must match the number of rows in the second matrix.

Matrix Multiplication

For two-dimensional matrices, the inner dimensions must match in order to carry out matrix multiplication. That is, if we want to find $A \times B$, and A has dimensions $m \times n$, then B must have dimensions $n \times p$, where m , n , and p are integers. The resulting matrix will have dimensions $m \times p$ (m and p may be equal or unequal). Note that we use $\times$ to differentiate from the dot product, represented by $\cdot$.

Exercise 12 Multiplying Matrices 1

Multiply the matrices.

Working Space

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

$$2. \begin{bmatrix} 1 \\ 5 \\ -5 \\ 4 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 & 5 & 1 \end{bmatrix}$$

$$3. \begin{bmatrix} -1 & 4 & -4 \\ 5 & -3 & 5 \\ -1 & -4 & 4 \\ -4 & 1 & 4 \end{bmatrix} \times \begin{bmatrix} -3 & 5 & 1 \\ -3 & 0 & -3 \\ 0 & 3 & 0 \end{bmatrix}$$

Answer on Page 100

Exercise 13 Multiplying Matrices 2

Find $A \times B$ and $B \times A$.

Working Space

$$1. A = \begin{bmatrix} -2 \\ 2 \\ 1 \\ -2 \end{bmatrix} \text{ and } B = \begin{bmatrix} -4 & 3 & -5 & -2 \end{bmatrix}$$

$$2. A = \begin{bmatrix} -4 & -2 \\ 2 & 5 \\ -3 & -4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & -2 & -4 \\ 1 & -4 & 0 \end{bmatrix}$$

$$3. A = \begin{bmatrix} 2 & 0 & 1 & 4 \\ -4 & 0 & -5 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & -3 \\ -4 & -1 \\ -2 & 3 \\ -5 & 1 \end{bmatrix}$$

Answer on Page 100

What have you noticed about the results of $A \times B$ as compared to $B \times A$? You should have noticed that the product matrices are *different dimensions*. This leads us to the next unusual property of matrix multiplication: it is *non-commutative*. That is, the *order* in which you multiply matrices affects the result. This is very different from scalar values!

As you saw in the second matrix multiplication exercise, A is a 2×4 matrix and B is a 4×2 matrix, then AB is a 2×2 matrix, while BA is a 4×4 matrix. It is obvious, then, that $A \cdot B \neq B \cdot A$. What if A and B are square matrices?

Exercise 14 Are Matrices Commutative?

Find $A \cdot B$ and $B \cdot A$ if $A = \begin{bmatrix} -3 & 5 \\ -1 & 0 \end{bmatrix}$ and

$$B = \begin{bmatrix} -1 & 1 \\ 4 & -3 \end{bmatrix}.$$

Working Space

Answer on Page 100

As you can see, even if A and B are square, matrix multiplication is still not commutative.

Non-Commutation of Matrix Multiplication

In general, for two matrices A and B ,

$$A \times B \neq B \times A$$

Some pairs do happen to commute (for example, any matrix and the identity matrix), but you can never assume it.

Properties of the Zero Matrix

Just like the number 0, the zero matrix, $\mathbf{0}$ has unique mathematical properties:

Properties of the Zero Matrix

For a matrix, A , and a zero matrix, $\mathbf{0}$

1. $A + \mathbf{0} = A$
2. $A + (-A) = \mathbf{0}$
3. $\mathbf{0} \cdot A = \mathbf{0}$

The Identity Matrix

There is another special matrix, called the *identity matrix*, usually denoted with I . An identity matrix is all zeroes except for a diagonal line of ones. A 3×3 identity matrix is shown below:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

All identity matrices are square (that is, they have the same number of rows as they do columns). The identity matrix has the special property that whenever a vector or matrix is multiplied by I , it does not change. Let us look at some examples:

Example: If $\mathbf{x} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, what is $I\mathbf{x}$? (Take I to be a 2×2 identity matrix.)

Solution:

$$I\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \cdot (2) + 0 \cdot (-3) \\ 0 \cdot (2) + 1 \cdot (-3) \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Example: If $B = \begin{bmatrix} -2 & 5 \\ 3 & -4 \end{bmatrix}$, what is $I \times B$?

Solution:

$$I \times B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 & 5 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-2) + 0 \cdot (5) & 1 \cdot (5) + 0 \cdot (-4) \\ 0 \cdot (-2) + 1 \cdot (5) & 0 \cdot (5) + 1 \cdot (-4) \end{bmatrix} = \begin{bmatrix} -2 & 5 \\ 3 & -4 \end{bmatrix}$$

Properties of the Identity Matrix

An $n \times n$ identity matrix, I , does not change any vectors or matrices it multiplies. That is:

1. $I \times \mathbf{x} = \mathbf{x}$
2. $I \times B = B$

where $\mathbf{x}$ is an $n \times 1$ vector and B is an $n \times p$ matrix (p may be, but is not necessarily, equal to n).

3.2.3 Transposing a Matrix

The **transpose** of a matrix is an operation that flips a matrix over its diagonal. In other words, rows become columns and columns become rows. If a matrix A has entries a_{ij} , then its transpose, written A^T , sometimes $A^\top$, has entries a_{ji} .

Transpose of a Matrix

If

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

then the **Transpose of A** is

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

That is, entry a_{ij} becomes a_{ji} .

Example: Let

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

Then the transpose of A is

$$A^T = \begin{bmatrix} 2 & 0 \\ -1 & 3 \\ 4 & 5 \end{bmatrix}.$$

Notice how the first *row* of A becomes the first *column* of A^T , and so on.

3.2.4 Row-Major Order, Column-Major Order, and Computer Memory

For big matrices, we let computers do the work. Computer memory is one long row of numbered slots, so a matrix has to be laid out in a line. In **row-major order** (used by C, C++, and NumPy by default), the first row is stored, then the second row, and so on. In **column-major order** (used by Fortran, MATLAB, and R), the columns are stored one after another. Here is a neat trick: the transpose of a row-major matrix, read in column-major order, is the same list of numbers. So a library can “transpose” a huge matrix without moving any data, just by changing how it reads it. NumPy’s `.T` does exactly this.



Figure 3.3: Row-major and column-major ordering of matrix elements.

3.2.5 Properties of Transposes

The transpose operation behaves nicely with matrix addition and multiplication:

- (1) $(A^T)^T = A$
- (2) $(A + B)^T = A^T + B^T$
- (3) $(cA)^T = cA^T$ for any scalar c
- (4) $(AB)^T = B^T A^T$, and $(ABC)^T = C^T B^T A^T$ (order reverses)

Property 4 often surprises students: the transpose of a product reverses the order. (It has to, for the dimensions to work. If A is 2×3 and B is 3×4 , then $A^T B^T$ would be 3×2 times 4×3 , which is not even allowed.)

Transposing a matrix is used throughout linear algebra, especially when:

- converting rows into column vectors (and vice versa),
- defining dot products using matrix multiplication,
- forming symmetric matrices ($A = A^T$),
- working with orthogonal matrices ($Q^T Q = I$),

3.3 Types of Matrices

Matrices can be described by their shape:

Row Matrix A matrix of size $1 \times n$ with 1 row and n columns.

Column Matrix A matrix of size $n \times 1$ with n rows and 1 column, typically used to represent vectors.

Square Matrix A matrix of size $n \times n$, containing the same number of rows and columns

Rectangular Matrix A matrix of size $m \times n$, that has an unequal number of rows and columns.

They can also be described by their unique numerical properties. Special matrices that come in handy for certain types of computations. These are a few of the most common special matrices:

Zero Matrix A matrix that only contains entries that are zero.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The matrix above could also be

$$0_{2 \times 2}$$

Diagonal Matrix A square matrix whose entries off the main diagonal are all zero.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Identity Matrix A diagonal matrix with entries of 1 along the diagonal.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Symmetric Matrix A matrix that equals its own transpose.

$$A = \begin{bmatrix} 20 & 40 & 60 \\ 40 & 50 & 80 \\ 60 & 80 & 100 \end{bmatrix}$$

Triangular Matrix This is a special square matrix that can be upper triangular or lower triangular. In an upper triangular matrix, every entry below the main diagonal is zero. In a lower triangular matrix, every entry above the main diagonal is zero. A matrix that is both upper and lower triangular becomes a diagonal matrix.

Upper Triangular Matrix

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

Lower Triangular matrix

$$\begin{bmatrix} 7 & 0 & 0 \\ 2 & 5 & 0 \\ 4 & 3 & 9 \end{bmatrix}$$

3.3.1 Symmetric Matrices

If you want to find out if a square matrix is symmetric, you need to transpose it. If the transpose is equal to the original matrix, then the matrix is symmetric.

To transpose a matrix, flip it over its diagonal so that the rows and columns are switched, like this:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

After transposing:

$$A^T = \begin{bmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{1,2} & a_{2,2} & a_{3,2} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{bmatrix}$$

Note that A^T means the transpose of A .

Let us see how this works for the following square matrix, A .

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

Switch the rows and columns to get the transpose:

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

Notice that $A = A^T$, so the matrix is symmetric.

What about matrix B ?

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 7 & 8 & 9 \end{bmatrix}$$

Switch the rows and columns to get the transpose:

$$B^T = \begin{bmatrix} 1 & 2 & 7 \\ 2 & 4 & 8 \\ 3 & 5 & 9 \end{bmatrix}$$

Note that $B \neq B^T$, so B is not symmetric.

Exercise 15 **Matrix Transposition**

Find the transpose of this matrix. Is it symmetric?

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Working Space

Answer on Page 101

3.3.2 Matrices in Python

NumPy does all of this for you. The @ operator multiplies matrices; + and - work entry by entry; and .T transposes. Create vectors-matrices.py:

```

1 import numpy as np
2
3 A = np.array([[5, 1, 3, -2],
4               [1, -1, 8, 4],
5               [6, 2, 1, 3]])
6 b = np.array([1, 2, 3, -8])
7
8 print(A.shape)    # (3, 4): 3 rows, 4 columns
9 print(A @ b)      # [ 32 -9 -11]
10 print(A.T)        # the 4 x 3 transpose

```

Careful: $A * B$ in NumPy multiplies entry by entry (like the column trick in the last chapter). It is *not* matrix multiplication.

Exercise 16 **Check Your Work**

Use NumPy to check your answers to the Multiplying Matrices 2 exercise. Then compute $A \times b$ by hand for the A and b above and confirm that you get

Working Space

$[32, -9, -11]$.

[Answer on Page 101](#)

3.3.3 Can We Divide Matrices?

Matrices cannot be divided. Suppose we have a matrix, A , a vector $\mathbf{x}$, and another vector $\mathbf{b}$ such that:

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

Now, if we know A and $\mathbf{x}$, it is easy to find $\mathbf{b}$. What if, on the other hand, we know A and $\mathbf{b}$ and want to find $\mathbf{x}$? We might be tempted to do something like this:

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

That would work if $\mathbf{x}$, $\mathbf{b}$, and A were numbers, but it means nothing for matrices. However, there is an analogy we can make. Instead of trying to divide by A , we can multiply by its *inverse*:

Inverse Matrices

Given a matrix A , and vectors $\mathbf{b}$ and $\mathbf{x}$, if

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

Then,

$$\mathbf{x} = A^{-1} \times \mathbf{b}$$

A^{-1} is called the *inverse matrix*. But, the inverse *does not always exist*! We will explore inverse matrices and how to find them in the next chapter.

Linear Combinations of Vectors

A smoothie shop sells two blends. A cup of the berry blend has 2 g of protein and 30 g of sugar; a cup of the green blend has 4 g of protein and 10 g of sugar. Write each as a vector, $[2, 30]$ and $[4, 10]$. If you mix 2 cups of berry with 1 cup of green, you get $2[2, 30] + 1[4, 10] = [8, 70]$. Could you mix them to get exactly 10 g of protein and 50 g of sugar? Questions like this are what this chapter is about.

A *linear combination* of vectors is a sum of vectors, each multiplied by a number. For example, $3 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 5 \end{bmatrix}$. The set of all the vectors you can make this way is called the **span** of the vectors. Formally:

Linear Combination of Vectors

A linear combination of a list of n vectors, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ takes the form:

$$a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_n \mathbf{v}_n$$

where $a_1, a_2, \dots, a_n \in \mathbb{R}$. You may see the list of vectors as $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, both are acceptable variables.

Example: Find a linear combination of $\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$ that gives the vector $\begin{bmatrix} 17 \\ -4 \\ 2 \end{bmatrix}$.

Solution: We are looking for a_1 and a_2 such that:

$$a_1 \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17 \\ -4 \\ 2 \end{bmatrix}$$

Looking at each dimension separately, we get the system of equations:

$$2a_1 + 1a_2 = 17$$

$$1a_1 - 2a_2 = -4$$

$$-3a_1 + 4a_2 = 2$$

If we can solve this system of equations, we will find a_1 and a_2 . Let us multiply the first equation by 2 and add it to the second equation:

$$2(2a_1 + a_2) + (a_1 - 2a_2) = 2(17) + (-4)$$

$$4a_1 + 2a_2 + a_1 - 2a_2 = 34 - 4$$

$$5a_1 = 30$$

$$a_1 = 6$$

Now we can take a_1 and substitute it back into any equation in our system to find a_2 . Let us use the third equation:

$$-3(6) + 4a_2 = 2$$

$$-18 + 4a_2 = 2$$

$$4a_2 = 20$$

$$a_2 = 5$$

We found a_1 from the first two equations combined and a_2 from the third, so we must check that both values also satisfy the first two equations: $2(6) + 5 = 17$ and $6 - 2(5) = -4$. They do. Always check every equation: three equations in two unknowns might have no solution at all, as you will see shortly.

Therefore,

$$6 \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17 \\ -4 \\ 2 \end{bmatrix}.$$

Exercise 17 Linear Combinations

Find a linear combination of the first two vectors that yields the third vector.

Working Space

1. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \end{bmatrix}$
2. $\begin{bmatrix} 9 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \end{bmatrix}$
3. $\begin{bmatrix} 7 \\ -2 \end{bmatrix}, \begin{bmatrix} -8 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \end{bmatrix}$

Answer on Page 101

Sometimes, a set of vectors cannot be combined to make a specific vector. Take the pair of vectors we have looked at before: $\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$. Can we find a combination to make vector $\begin{bmatrix} 17 \\ -4 \\ 5 \end{bmatrix}$? Let us try. We define a_1 and a_2 such that:

$$a_1 \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17 \\ -4 \\ 5 \end{bmatrix}$$

Which creates the system of equations:

$$2a_1 + a_2 = 17$$

$$a_1 - 2a_2 = -4$$

$$-3a_1 + 4a_2 = 5$$

We have two variables (a_1 and a_2) and three equations. Let us use the first two to find a_1 and a_2 , then check our answers by substituting our solutions into the third equation. First, we will multiply the second equation by -2 and add that to the first equation:

$$2a_1 + a_2 + (-2)(a_1 - 2a_2) = 17 + (-2)(-4)$$

$$2a_1 + a_2 - 2a_1 + 4a_2 = 17 + 8$$

$$5a_2 = 25$$

$$a_2 = 5$$

Substituting for a_2 back into the first equation and solving for a_1 :

$$2a_1 + 5 = 17$$

$$2a_1 = 12$$

$$a_1 = 6$$

Now, let us check if $a_1 = 6$, $a_2 = 5$ is a solution to the third equation:

$$-3(6) + 4(5) = 5$$

$$-18 + 20 = 2 \neq 5$$

Therefore, there is no linear combination of the vectors $\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$ that yields $\begin{bmatrix} 17 \\ -4 \\ 5 \end{bmatrix}$.

Linear Combinations as systems.

We are aiming to know if $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 = \mathbf{w}$. We can rewrite this as a matrix multiplication:

$$[v_1 \ v_2] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \mathbf{w}$$

Then, to check whether w is a linear combination of v_1 and v_2 , solve the system of equations. If a solution exists, then yes; if not, then no. Note that the a_n values are only scalar numbers, while the v_n 's remain vectors.

We will talk more about systems of equations in Section 5.1.

4.1 Visualizing Linear Combinations

First, let us look at what vectors can be made from linear combinations of the 2-dimensional unit vectors $\mathbf{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suppose we are looking for a linear combination of $\mathbf{i}$ and $\mathbf{j}$ to create the vector $\begin{bmatrix} 3 \\ -4 \end{bmatrix}$. We can find such a linear combination:

$$3\mathbf{i} + (-4)\mathbf{j} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

In fact, with $\mathbf{i}$ and $\mathbf{j}$, we can create any 2-dimensional vector. To prove this, consider a generic vector, $\mathbf{z} = \begin{bmatrix} a \\ b \end{bmatrix}$, where $a, b \in \mathbb{R}$. We are looking for a linear combination of $\mathbf{i}$ and $\mathbf{j}$ such that:

$$c_1\mathbf{i} + c_2\mathbf{j} = \begin{bmatrix} a \\ b \end{bmatrix}$$

The above equation yields the system of equations:

$$c_1(1) + c_2(0) = a$$

$$c_1(0) + c_2(1) = b$$

And the solution to this system of equations is:

$$c_1 = a$$

$$c_2 = b$$

Therefore, using $\mathbf{i}$ and $\mathbf{j}$, we can construct any vector in $\mathbb{R}^2$ (that is, any vector in the xy -plane). What about combinations of other vectors?

Let us consider linear combinations of two vectors: $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. The vectors are shown in figure 4.1.

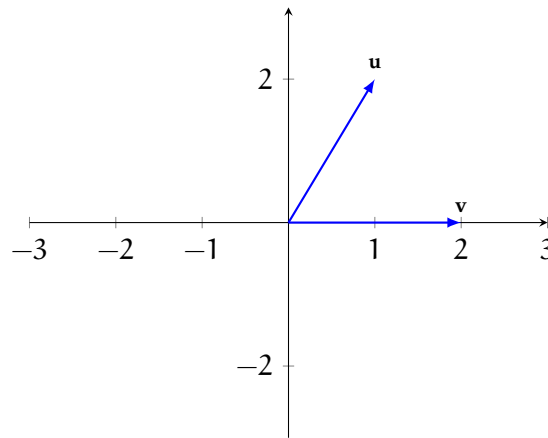


Figure 4.1: The vectors $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

Suppose we want to construct the vector $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$. Since only $\mathbf{u}$ has value in the y -dimension, we can start by adding $\mathbf{u}$ vectors to reach $y = 4$ (see figure 4.2a). Next, we can use $\mathbf{v}$ vectors to reach $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$ (see figure 4.2b).

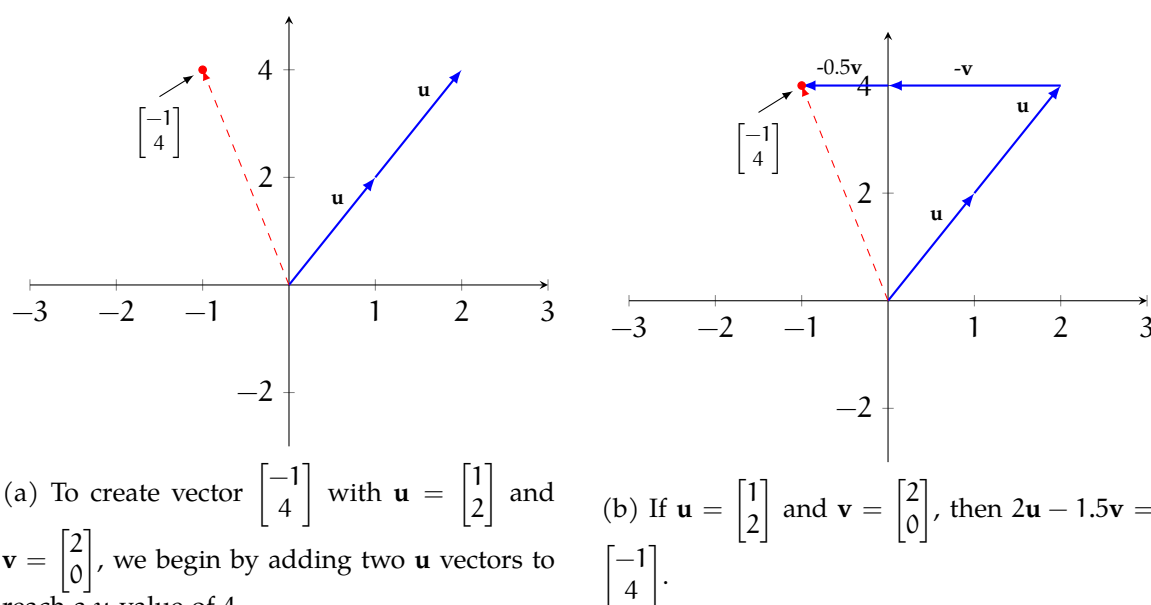
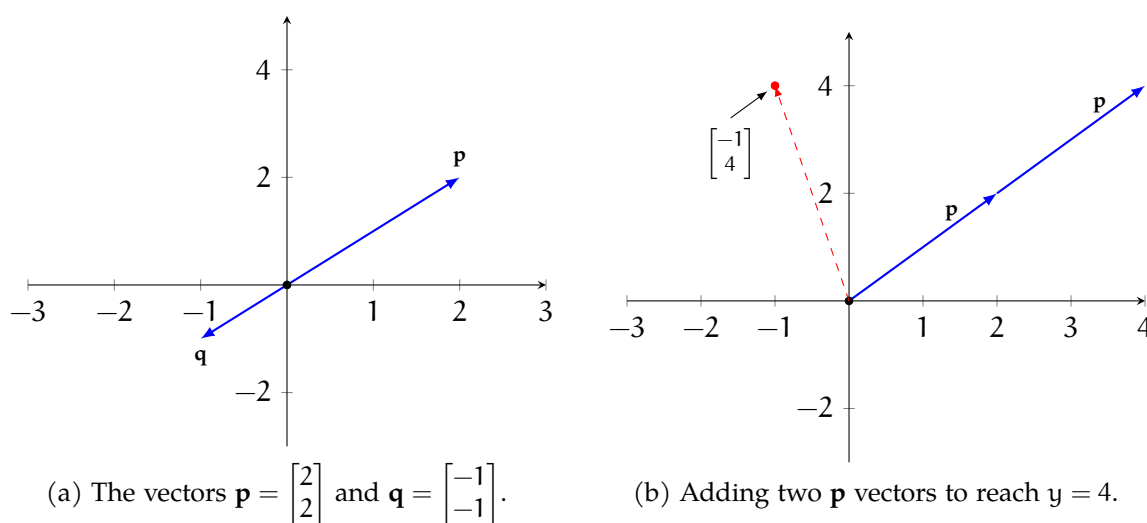


Figure 4.2: Using linear combinations of $\mathbf{u}$ and $\mathbf{v}$ to construct $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$.

Using this method, we can imagine reaching any point in $\mathbb{R}^2$: we add or subtract as many $\mathbf{u}$ vectors as needed to reach the appropriate y-value, then add or subtract as many $\mathbf{v}$ vectors to reach the appropriate x-value. The vectors are *not multiples* of each other, so we can say that they span all of $\mathbb{R}^2$.

Let us look at another pair of vectors: $\mathbf{p} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $\mathbf{q} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ (see figure 4.3a). Again, let us try to use $\mathbf{p}$ and $\mathbf{q}$ to construct the vector $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$. We begin by using $\mathbf{p}$ to reach the y-value of 4 (see figure 4.3b).


 Figure 4.3: Visualizing combinations of $\mathbf{p}$ and $\mathbf{q}$.

But now we run into a problem: no matter how many multiples of $\mathbf{q}$ vectors we add or subtract, we just move along the $\mathbf{p}$ vector and never reach our goal of $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$ (see figure 4.4). Notice that $\mathbf{p}$ and $\mathbf{q}$ lie on the same line, (for a better visualization, refer back to 4.3a). In other words, the vectors are *scalar multiples* of each other: $\begin{bmatrix} 2 \\ 2 \end{bmatrix} = -2 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix}$. When two vectors lie on the same line, so that one is a scalar multiple of the other, we call them *linearly dependent*.

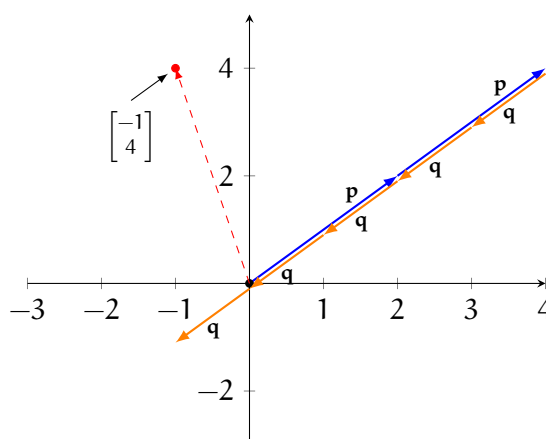


Figure 4.4: There is no linear combination of $\mathbf{p} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $\mathbf{q} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ that yields the vector $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$.

Back to the smoothies. To get 10 g of protein and 50 g of sugar, solve $a_1[2, 30] + a_2[4, 10] =$

[10, 50]: $2a_1 + 4a_2 = 10$ and $30a_1 + 10a_2 = 50$. The solution is $a_1 = 1$ and $a_2 = 2$: one cup of berry and two cups of green. Because $[2, 30]$ and $[4, 10]$ are not multiples of each other, they span $\mathbb{R}^2$, and you can hit any target (as long as you allow negative cups, which the shop does not).

Exercise 18 Span or Not?

Do the vectors span $\mathbb{R}^2$? If not, describe the set of vectors they do span.

Working Space

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

2. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Answer on Page 103

4.2 Span in Three Dimensions

Imagine flying a drone whose controller has one stick that moves it forward and back. You can reach any spot on one straight line, and nowhere else. Add a second stick for left and right, and you can reach any point at that altitude: a whole plane. Add a third for up and down, and the whole sky is yours. Each stick is a vector, and the places you can reach are their span.

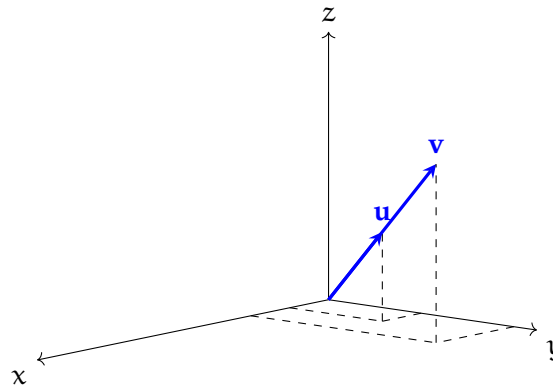
Here is the formal definition. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a set of vectors, then the span of S is the set of all their linear combinations:

$$\text{Span}(S) = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n : a_1, a_2, \dots, a_n \in \mathbb{R}\} \quad (4.1)$$

A vector is in $\text{Span}(S)$ exactly when it can be written as a linear combination of the vectors in S .

4.2.1 A Line

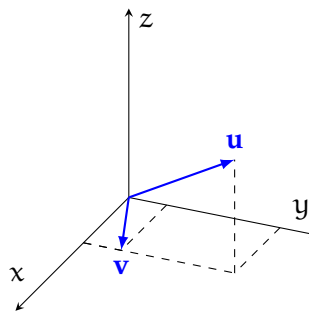
In two dimensions, two vectors that are multiples of each other span only a line. The same is true in three dimensions. Consider $\mathbf{u} = [2, 4, 3]$ and $\mathbf{v} = [4, 8, 6]$, shown in figure 4.5.

Figure 4.5: 3-dimensional vectors, $\mathbf{u}$ and $\mathbf{v}$

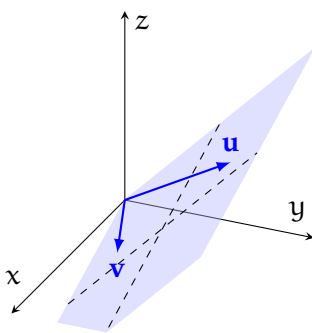
Since $\mathbf{v} = 2\mathbf{u}$, the two vectors lie on the same line through the origin, and every combination $a_1\mathbf{u} + a_2\mathbf{v} = (a_1 + 2a_2)\mathbf{u}$ lies on that line too. The span of $\mathbf{u}$ and $\mathbf{v}$ is a line. In fact, if every vector in a list lies on one line through the origin, the span of the list is that line, no matter how many vectors are in the list.

4.2.2 A Plane

Now take two 3-dimensional vectors that are *not* multiples of each other: $\mathbf{u} = [2, 4, 3]$ and $\mathbf{v} = [2, 1, 0]$ (figure 4.6).

Figure 4.6: Two 3-dimensional vectors, $\mathbf{u}$ and $\mathbf{v}$, that are not multiples of each other

In two dimensions, two such vectors spanned the whole plane $\mathbb{R}^2$. In three dimensions, they span a plane through the origin, a flat sheet that contains both vectors (figure 4.7). You can slide along the sheet in any direction, but you can never leave it.

Figure 4.7: The vectors $\mathbf{u}$ and $\mathbf{v}$ span a plane.

4.2.3 All of Space

To reach every point in $\mathbb{R}^3$, add a third vector that is *not* on that plane. A vector is on the plane exactly when it is a linear combination of the first two. A third vector that is not a combination of the first two lets you move off the sheet, and together the three vectors span all of $\mathbb{R}^3$.

Think of walking around a city. Moving forward and backward on one street covers a line. Adding a sideways direction covers a flat field, which is a plane. Adding a third direction that is not in that field lets you move up and down, so you cover all of $\mathbb{R}^3$.

So, to describe the span of a list of vectors in $\mathbb{R}^3$:

1. If all the vectors are multiples of one nonzero vector, the span is a *line*.
2. Otherwise, pick two that are not multiples of each other; they span a plane. If every other vector in the list is a linear combination of those two, the span is that *plane*.
3. If some vector is not a linear combination of those two, the span is all of $\mathbb{R}^3$.

Step 3 is the question this chapter started with: can a vector be written as a linear combination of others? You answer it by solving a system of equations.

Example: Do the vectors $\mathbf{r} = [5, 4, -6]$, $\mathbf{s} = [0, -5, -10]$, and $\mathbf{t} = [0, 2, 4]$ span a line, a plane, or $\mathbb{R}^3$?

Solution: $\mathbf{r}$ and $\mathbf{s}$ are not multiples of each other ($\mathbf{s}$ has a 0 where $\mathbf{r}$ has a 5), so they span a plane. Is $\mathbf{t}$ on that plane? We look for α_1 and α_2 with

$$\alpha_1 [5, 4, -6] + \alpha_2 [0, -5, -10] = [0, 2, 4]$$

which gives the system

$$5\alpha_1 = 0$$

$$4a_1 - 5a_2 = 2$$

$$-6a_1 - 10a_2 = 4$$

The first equation gives $a_1 = 0$. The second then gives $-5a_2 = 2$, so $a_2 = -\frac{2}{5}$. Check the third: $-6(0) - 10(-\frac{2}{5}) = 4$. It works, so $\mathbf{t} = -\frac{2}{5}\mathbf{s}$ lies on the plane, and $\text{Span}(\mathbf{r}, \mathbf{s}, \mathbf{t})$ is a plane.

Be careful: checking only that no two vectors are multiples of each other is *not* enough. Three vectors can be pairwise non-multiples and still lie on one plane, such as $[1, 0, 0]$, $[0, 1, 0]$, and $[1, 1, 0]$.

Exercise 19 Determining Span

Geometrically describe (line, plane, or $\mathbb{R}^3$) the span of the list of vectors.

Working Space

1. $[1, 2, 4]$ and $[-2, -4, -8]$
2. $[2, 0, 0]$ and $[0, 1, 3]$
3. $[3, 0, 0]$ and $[0, 3, 3]$ and $[3, 3, 2]$

Answer on Page 103

Exercise 20 Span of Three Vectors

Do the vectors $[1, 1, 0]$, $[0, 1, 1]$, and $[1, 0, 1]$ span a line, a plane, or $\mathbb{R}^3$? Explain your answer.

Working Space

Answer on Page 104

4.3 Summary

- A linear combination is a sum of scalar multiples of vectors.

- Asking whether $\mathbf{w}$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ is the same as asking whether a system of equations has a solution.
- The span of a set of vectors is everything you can build from them. Two vectors in $\mathbb{R}^2$ that are not multiples of each other span the whole plane; two that are multiples span only a line.
- In $\mathbb{R}^3$, a span is a line, a plane, or all of space. A vector adds a new dimension to the span only if it is not a linear combination of the vectors already there.

In the next chapter, we will be looking more at linear combinations, span, and independence in the form of systems of equations. Future chapters will look even further at these concepts, linear dependence and independence, and their applications to various subspaces.

4.4 Where to Learn More

Watch this video on *Linear Combinations and Vector Spans* from Khan Academy: <http://rb.gy/g1snk>

Matrices and Systems

5.1 Systems of Equations as Matrices

5.1.1 Systems of Equations, REF, and RREF

You have solved systems of two equations in two unknowns since algebra class. But real problems, like balancing a chemical reaction, routing traffic, or fitting a curve through data, can have hundreds of equations and hundreds of unknowns. Matrices let us write any system compactly and solve it with one systematic recipe that a computer can follow. Let us start small:

$$\begin{cases} 1x_1 + 2x_2 = -1 \\ 2x_1 + 0x_2 = 4 \end{cases}$$

We can write this with matrices. Write the coefficients as a matrix, the x_n 's as a column vector, and the constants as a column vector called $\vec{b}$, the right-hand-side vector. Here is the coefficient matrix which we will call A :

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}$$

Then we can write the column vector of x values, which we will call $\vec{x}$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

And we can write the $\vec{b}$ as the following column vector:

$$\vec{b} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

Now, if we multiply A and $\vec{x}$,

$$A\vec{x} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 + 2x_2 \\ 2x_1 + 0x_2 \end{bmatrix}$$

We get exactly the left-hand sides of our original system of equations! Therefore, when we set this equal to the right-hand-side vector $\vec{b}$:

$$\vec{b} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

we get the matrix equation

$$A\vec{x} = \vec{b} \quad (5.1)$$

This is a very important equation for linear algebra as a whole. Notice that the sizes of the matrices must follow the rules of matrix multiplication:

$$\begin{matrix} A & \cdot & \vec{x} & = & \vec{b} \\ m \times n & & n \times 1 & & m \times 1 \end{matrix}$$

Exercise 21 Matrix Equation

Rewrite the systems of equations in matrix form and identify A , $\vec{x}$, and $\vec{b}$.

$$\begin{cases} -4x_1 + 9x_2 - 8x_3 = 5 \\ -1x_1 + 0x_2 + 6x_3 = 7 \end{cases}$$

Working Space

Answer on Page 104

5.1.2 Row-Echelon Form

How do we solve for $\vec{x}$? The most common way is to form an augmented matrix and reduce it to row-echelon form.

The Augmented Matrix

What is the augmented matrix? It is a way of representing the A matrix and $\vec{b}$ vector. We write the A matrix, separated by a vertical line from the $\vec{b}$ vector:

$$[A \mid \vec{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right]$$

The left block contains all the coefficients of the system of equations, and the right block contains the constants. Then, to solve for $\vec{x}$, we perform *elementary row operations* to reach *row-echelon form*. Let us continue with the matrix vector set we were experimenting with:

$$A\vec{x} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 + 2x_2 \\ 2x_1 + 0x_2 \end{bmatrix}$$

This can be written as an augmented matrix:

$$[A \mid \vec{b}] = \left[\begin{array}{cc|c} 1 & 2 & -1 \\ 2 & 0 & 4 \end{array} \right]$$

Elementary Row Operations

The Row-Echelon Form of a matrix is an *equivalent* form of a matrix obtained through Gaussian and Gauss-Jordan elimination. The rules of Gaussian elimination are simple. Only three *elementary row operations* (EROs) are allowed:

- Swapping two rows
- Multiply a row by a nonzero scalar
- Add a multiple of one row to another row

Why do these operations preserve the given linear system? Swapping two rows only changes the order in which the equations are written and therefore has no effect on the solution set. Multiplying a row by a nonzero scalar corresponds to multiplying both sides of an equation by the same scalar, which does not alter its solutions. Finally, if two equations are satisfied by the same solution, then any linear combination of those equations is also satisfied by that solution; in particular, adding a multiple of one equation to another preserves the solution set. Before and after each operation, the augmented matrix contains exactly the same information as the system of equations and therefore

represents the same solution set.

With these ERO's, we are trying to achieve a matrix with the following properties:

- Leading entries (not necessarily 1) move to the right as you go down
- Rows of zeros, if exist, are at the bottom
- Each pivot has zeros in all positions below it in the same column.

Let us walk through the steps to simplify $\text{ref}(A|\vec{b})$:

$$\begin{bmatrix} 1 & 2 & | & -1 \\ 2 & 0 & | & 4 \end{bmatrix}$$
$$R_2 \rightarrow R_2 - 2R_1 \implies \begin{bmatrix} 1 & 2 & | & -1 \\ 0 & -4 & | & 6 \end{bmatrix}$$

As you can see, we set Row 2 to be Row 2 minus 2 times Row 1. This gives us a zero in the first column of Row 2. Technically, this matrix is now in Row-Echelon Form (REF). However, we can continue applying elementary row operations to obtain an even more structured form called the *Reduced Row-Echelon Form*.

Reduced Row Echelon Form

A matrix is in *Reduced Row-Echelon Form* (RREF) if it satisfies all of the conditions for Row-Echelon Form, and in addition:

- Each leading entry is equal to 1
- Each leading entry is the only nonzero entry in its column

In other words, not only do pivots move to the right as we move down the rows, but each pivot column contains zeros both above and below the pivot, and each pivot is normalized to 1. These additional conditions make RREF especially useful, since the solutions to the system can often be read directly from the matrix.

Continuing our example, we start with the matrix in REF:

$$\begin{bmatrix} 1 & 2 & | & -1 \\ 0 & -4 & | & 6 \end{bmatrix}$$

First, make the leading entry in Row 2 equal to 1 by dividing the row by -4 :

$$R_2 \rightarrow -\frac{1}{4}R_2 \implies \left[\begin{array}{cc|c} 1 & 2 & -1 \\ 0 & 1 & -\frac{3}{2} \end{array} \right]$$

Next, eliminate the entry above the pivot in column 2:

$$R_1 \rightarrow R_1 - 2R_2 \implies \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -\frac{3}{2} \end{array} \right]$$

The matrix is now in Reduced Row-Echelon Form.

Writing the Solution Vector

Recall that the augmented matrix represents a system of linear equations. After reducing the matrix to Reduced Row-Echelon Form, we obtain

$$\left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -\frac{3}{2} \end{array} \right]$$

Each row now corresponds directly to an equation in the variables x_1 and x_2 :

$$\begin{aligned} x_1 &= 2 \\ x_2 &= -\frac{3}{2} \end{aligned}$$

Thus, the solution to the system can be written as the vector

$$\vec{x} = \begin{bmatrix} 2 \\ -\frac{3}{2} \end{bmatrix}$$

Since every variable corresponds to a pivot column, the system has a *unique solution*.

We have successfully solved the system of equations using matrix methods:

$$\begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -\frac{3}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

This is a fundamental technique in linear algebra, and you will use it often in this module and beyond!

5.1.3 Larger RREF Example

Let us solve a larger system by finding $\vec{x}$ in the matrix equation

$$A\vec{x} = \vec{b}.$$

Write the System and Identify A , $\vec{x}$, and $\vec{b}$

Consider the following system of equations:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 4 \\ 2x_1 + x_2 + 3x_3 + x_4 = 9 \\ x_1 + 3x_2 + 2x_3 + 2x_4 = 10 \\ 3x_1 + x_2 + 2x_3 + 4x_4 = 13 \end{cases}$$

The coefficient matrix, variable vector, and constant vector are:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 1 \\ 1 & 3 & 2 & 2 \\ 3 & 1 & 2 & 4 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 4 \\ 9 \\ 10 \\ 13 \end{bmatrix}.$$

So the matrix equation is:

$$A\vec{x} = \vec{b}.$$

Form the Augmented Matrix

$$\left[A | \vec{b} \right] = \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 2 & 1 & 3 & 1 & 9 \\ 1 & 3 & 2 & 2 & 10 \\ 3 & 1 & 2 & 4 & 13 \end{array} \right].$$

Row-Reduce to REF

Use the pivot in Row 1 to clear entries below it in column 1:

$$R_2 \rightarrow R_2 - 2R_1$$

$$R_3 \rightarrow R_3 - R_1$$

$$R_4 \rightarrow R_4 - 3R_1$$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 2 & 1 & 1 & 6 \\ 0 & -2 & -1 & 1 & 1 \end{array} \right].$$

Now use the pivot in Row 2 (column 2) to clear entries below it in column 2:

$$R_3 \rightarrow R_3 + 2R_2$$

$$R_4 \rightarrow R_4 - 2R_2$$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 0 & 3 & -1 & 8 \\ 0 & 0 & -3 & 3 & -1 \end{array} \right].$$

Next use the pivot in Row 3 (column 3) to clear the entry below it in column 3:

$$R_4 \rightarrow R_4 + R_3$$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 0 & 3 & -1 & 8 \\ 0 & 0 & 0 & 2 & 7 \end{array} \right].$$

At this point, the matrix is in Row-Echelon Form (REF).

Continue to RREF

Make the pivot in Row 4 equal to 1:

$$R_4 \rightarrow \frac{1}{2}R_4 \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 0 & 3 & -1 & 8 \\ 0 & 0 & 0 & 1 & \frac{7}{2} \end{array} \right].$$

Clear the entries above the pivot in column 4:

$$R_1 \rightarrow R_1 - R_4$$

$$R_2 \rightarrow R_2 + R_4$$

$$R_3 \rightarrow R_3 + R_4$$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & \frac{1}{2} \\ 0 & -1 & 1 & 0 & \frac{9}{2} \\ 0 & 0 & 3 & 0 & \frac{23}{2} \\ 0 & 0 & 0 & 1 & \frac{7}{2} \end{array} \right].$$

Make the pivot in Row 3 equal to 1:

$$R_3 \rightarrow \frac{1}{3}R_3 \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & \frac{1}{2} \\ 0 & -1 & 1 & 0 & \frac{9}{2} \\ 0 & 0 & 1 & 0 & \frac{23}{6} \\ 0 & 0 & 0 & 1 & \frac{7}{2} \end{array} \right].$$

Clear the entries above the pivot in column 3:

$$R_1 \rightarrow R_1 - R_3$$

$$R_2 \rightarrow R_2 - R_3$$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & -\frac{10}{3} \\ 0 & -1 & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 1 & 0 & \frac{23}{6} \\ 0 & 0 & 0 & 1 & \frac{7}{2} \end{array} \right].$$

Make the pivot in Row 2 equal to 1:

$$R_2 \rightarrow -R_2 \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & -\frac{10}{3} \\ 0 & 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & 0 & \frac{23}{6} \\ 0 & 0 & 0 & 1 & \frac{7}{2} \end{array} \right].$$

Clear the entry above the pivot in column 2:

$$R_1 \rightarrow R_1 - R_2$$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -\frac{8}{3} \\ 0 & 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & 0 & \frac{23}{6} \\ 0 & 0 & 0 & 1 & \frac{7}{2} \end{array} \right].$$

This matrix is now in Reduced Row-Echelon Form (RREF).

Write the Solution Vector

Reading each row as an equation gives:

$$\begin{aligned} x_1 &= -\frac{8}{3} \\ x_2 &= -\frac{2}{3} \\ x_3 &= \frac{23}{6} \\ x_4 &= \frac{7}{2} \end{aligned}$$

Thus the solution vector is:

$$\vec{x} = \begin{bmatrix} -\frac{8}{3} \\ -\frac{2}{3} \\ \frac{23}{6} \\ \frac{7}{2} \end{bmatrix}.$$

You can verify the solution by confirming that $A\vec{x} = \vec{b}$ using matrix multiplication.

Note that solution sets must be one of the following:

- No solution (for example, two parallel lines that never meet),
- Exactly one unique solution (for example, two lines or planes that cross at one point),
- or Infinitely many solutions (for example, two equations describe the same line)

5.1.4 Free Variables

In the reduced row-echelon form (RREF) of a matrix, pivots determine which variables are constrained by the system.

Zero rows and redundant equations

A row of all zeros in RREF represents a redundant equation. That row corresponds to the equation

$$0x_1 + 0x_2 + \cdots + 0x_n = 0,$$

which is always true and therefore imposes no restriction on the variables. Such a row arises when one of the original equations is a linear combination of the others.

Every row of zeros is a row without a pivot. Consequently, if a matrix has at least one zero row in RREF:

- The number of pivots is less than the number of rows,
- The rank of the matrix is less than the number of rows,
- The rows of the matrix are linearly dependent.

Free variables

Pivots are associated with *columns*, not rows. A variable is called a *free variable* if its column does not contain a pivot.

Thus:

- A *row* without a pivot (a zero row) indicates a redundant equation,
- A *column* without a pivot corresponds to a free variable.

A free variable can take any value. When the system is consistent and has a free variable, it has infinitely many solutions.

Infinitely Many Solutions

Example: Consider the augmented matrix in RREF:

$$\begin{bmatrix} 1 & 0 & 0 & 8 \\ 0 & 0 & 1 & -6 \end{bmatrix}$$

Labeling the variable columns (and ignoring the constants c column), we write:

$$\begin{bmatrix} x_1 & x_2 & x_3 & c \\ 1 & 0 & 0 & 8 \\ 0 & 0 & 1 & -6 \end{bmatrix}$$

The pivot positions are in the columns corresponding to x_1 and x_3 . Since the column corresponding to x_2 contains no pivot, x_2 is a free variable.

The system of equations represented by this matrix is:

$$\begin{aligned} x_1 &= 8, \\ x_3 &= -6. \end{aligned}$$

The variable x_2 does not appear in any equation and may take any real value. Since x_1 and x_3 are bound, we can say the solution is :

$$x_1 = 8, \quad x_3 = -6, \quad x_2 \in \mathbb{R}$$

written as a vector this is: $\vec{x} = \begin{bmatrix} 8 \\ x_2 \\ -6 \end{bmatrix}$ or in set builder notation: $\left\{ \begin{pmatrix} 8 \\ x_2 \\ -6 \end{pmatrix} : x_2 \in \mathbb{R} \right\}$

Thus, the presence of a column without a pivot leads to a free variable and infinitely many solutions.

Inconsistent sets

Let us look at another example: **Example** The augmented matrix

$$\begin{bmatrix} 1 & 2 & 4 & 12 \\ 12 & 24 & 48 & 9 \\ 3 & 6 & 12 & -4 \end{bmatrix}$$

simplifies to

$$\begin{bmatrix} 1 & 2 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We can interpret this as

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 0, \\ 0x_1 + 0x_2 + 0x_3 &= 1. \end{aligned}$$

Note that the equation $0 = 1$ is never true, so the system is *inconsistent*, and there is *no solution*.

A row that is all zeros except for a nonzero entry in the last column signifies an inconsistent system.

5.1.5 Basis, Dimension, Rank

Let us define a few terms. These definitions are very basic, and will come back in the future:

Rank

Rank is the number of pivots in the RREF state of a matrix of size $m \times n$. In other words, the number of nonzero rows in any row-echelon form of a matrix A is denoted $\text{rank}(A)$. This has the constraints of:

- $\text{rank}(A) \leq m$
- $\text{rank}(A) \leq n$

Dimension

The **dimension** of the solution set of a system is the number of free variables. We calculate this with

$$\text{dimension} = \text{number of variables} - \text{rank}$$

or equivalently, for matrix A of size $m \times n$

$$\text{dimension} = n - \text{rank}(A)$$

Basis

A **basis** is a minimal set of direction vectors that describes all solutions. Specifically, this consists of $n - \text{rank}(A)$ vectors, one for each free variable

In summary,

$$\begin{array}{ll} \text{rank} & = \text{number of pivots} \\ \text{number of free variables} & = n - \text{rank} \\ \text{dimension of solution set} & = n - \text{rank} \\ \text{number of basis vectors} & = n - \text{rank} \end{array}$$

Exercises

Exercise 22 Augmented Matrix 1

Simplify the matrix:

$$\begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & -2 & 6 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

to RREF state and determine the type of solution set.

Working Space

Answer on Page 104

Exercise 23 Augmented Matrices 2

Put the augmented matrix

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

Working Space

into RREF and determine its solution.

Answer on Page 105

Exercise 24 **Augmented Matrices 3**

Determine the solution for the following set:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Working Space

Answer on Page 106

Exercise 25 **Rank 1**

Find the rank of

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

Working Space

Answer on Page 106

Exercise 26 **Rank 2**

Let A be a 4×6 matrix with exactly 3 pivots in RREF.

Working Space

- What is $\text{rank}(A)$?

- How many free variables are there?

Answer on Page 106

Exercise 27 Free Variables

The RREF of a system has the form:

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

1. Identify the pivot variables
2. Identify the free variables
3. State the dimension of the solution set

Working Space

Answer on Page 106

5.1.6 Creating Matrices in Python

NumPy stores matrices as 2D arrays. Create `matrices_creation.py`:

```

1  import numpy as np
2
3  A = np.array([[5, 1, 3],
4               [1, -1, 8],
5               [6, 2, 1]])
6  print(A.T)                # not equal to A: not symmetric
7
8  B = np.array([[5, 1, 6],
9               [1, -1, 2],
10              [6, 2, 1]])
11 print(np.array_equal(B, B.T)) # True: B is symmetric
12
13 J = np.array([[5, 1, 3, 0],
```

```

14         [1, -1, 8, 11],
15         [6, 2, 1, -7]])
16 print(J.T.shape)           # (4, 3)

```

When you run it, you should see:

```

1  [[ 5  1  6]
2   [ 1 -1  2]
3   [ 3  8  1]]
4  True
5  (4, 3)

```

$A \neq A^T$, so A is not symmetric, but B is. You can transpose any matrix, but only a square matrix can be symmetric: transposing the 3×4 matrix J gives a 4×3 matrix.

Creating Special Matrices in Python

NumPy has shortcuts for the special matrices. `np.diag` works both ways: given a matrix, it extracts the diagonal; given a list, it builds a diagonal matrix. (`np.diag(W, 1)` extracts the diagonal just above the main one, and `np.diag(W, -1)` the one just below.)

```

1  import numpy as np
2
3  print(np.zeros((2, 3)))      # a 2 x 3 zero matrix
4  print(np.eye(3))           # a 3 x 3 identity matrix
5
6  W = np.array([[1, 2, 3, 4],
7               [5, 6, 7, 8],
8               [-8, -7, -6, -5],
9               [-4, -3, -2, -1]])
10 print(np.diag(W))           # the main diagonal
11 print(np.diag([1, 2, 3]))   # a diagonal matrix
12 print(np.triu(W))           # upper triangle
13 print(np.tril(W))           # lower triangle

```

You should see:

```

1  [[0.  0.  0.]
2   [0.  0.  0.]]
3  [[1.  0.  0.]
4   [0.  1.  0.]
5   [0.  0.  1.]]
6  [ 1  6 -6 -1]
7  [[1  0  0]

```

```

8  [0 2 0]
9  [0 0 3]]
10 [[ 1  2  3  4]
11  [ 0  6  7  8]
12  [ 0  0 -6 -5]
13  [ 0  0  0 -1]]
14 [[ 1  0  0  0]
15  [ 5  6  0  0]
16  [-8 -7 -6  0]
17  [-4 -3 -2 -1]]

```

Solving Systems in Python

Row reduction by hand is great for understanding, but for real problems you let the computer do it. `np.linalg.solve` solves $A\vec{x} = \vec{b}$ when there is exactly one solution. Here is the 4×4 system from earlier in the chapter:

```

1  import numpy as np
2
3  A = np.array([[1, 1, 1, 1],
4               [2, 1, 3, 1],
5               [1, 3, 2, 2],
6               [3, 1, 2, 4]])
7  b = np.array([4, 9, 10, 13])
8
9  x = np.linalg.solve(A, b)
10 print(x)
11 print(A @ x)    # check: should equal b

```

You should see:

```

1  [-2.66666667 -0.66666667  3.83333333  3.5      ]
2  [ 4.   9.  10.  13.]

```

That is $-\frac{8}{3}$, $-\frac{2}{3}$, $\frac{23}{6}$, $\frac{7}{2}$, just as we found by hand.

5.1.7 Conclusion

In this module, we have explored the fundamental concepts of matrices, including their types, properties, and how they can be used to represent and solve systems of linear equations. We have also learned how to perform row operations to convert matrices into Row-Echelon Form (REF) and Reduced Row-Echelon Form (RREF), which are essential

techniques in linear algebra. In Chapter 7, we will go deeper into vector spaces and subspaces, building on RREF.

5.2 Applications of Matrices in System Form

In the chapter on linear combinations, we saw that we can linearly combine vectors to create other vectors. Consider 3 vectors:

$$\mathbf{x} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

We can write a linear combination of these vectors:

$$c\mathbf{x} + d\mathbf{y} + e\mathbf{z}$$

Which we can expand to show the vectors:

$$c \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + d \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} + e \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -c - d - 2e \\ 2c + 2d + e \\ d \end{bmatrix}$$

We can also represent this combination with a matrix where each column is one of the vectors:

$$\begin{bmatrix} -1 & -1 & -2 \\ 2 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} c \\ d \\ e \end{bmatrix} = \begin{bmatrix} -c - d - 2e \\ 2c + 2d + e \\ d \end{bmatrix}$$

5.2.1 Trail Mix for Mars

Let us look at an applied problem. Three astronauts, Pat, Kai, and River, are getting ready for a trip to Mars. NASA food service is preparing trail mix for the voyage, tailored to each astronaut's taste. The chef needs to submit a budget based on the cost of the trail mix for each astronaut. The mix is made up of raisins, almonds, and chocolate.

Pat prefers a raisins:almonds:chocolate ratio of 6:10:4, Kai likes 2:3:15, and River wants 14:1:5. The chef can buy a kg of raisins for \$7.50, a kg of almonds for \$14.75, and a kg of chocolate for \$22.25. Assuming each astronaut will get 20 kg of trail mix, which astronaut will cost more to feed?

First, set up a matrix to represent the raisins:almonds:chocolate ratios. (Conveniently,

these ratios already add to 20.)

$$\text{MixRatios} = \begin{bmatrix} 6 & 10 & 4 \\ 2 & 3 & 15 \\ 14 & 1 & 5 \end{bmatrix}$$

Use a vector to represent the cost of each item:

$$\text{IngredientCost} = \begin{bmatrix} 7.50 \\ 14.75 \\ 22.25 \end{bmatrix}$$

To find the cost of trail mix for each astronaut, we simply find the dot product between the mix ratios and the ingredient costs to get:

$$\text{Pat} = \$281.50$$

$$\text{Kai} = \$393.00$$

$$\text{River} = \$231.00$$

So Kai will cost the most to feed.

Exercise 28 Vector Matrix Multiplication

Multiply the array A with the vector v . Compute this by hand, and make sure to show your computations.

$$A = \begin{bmatrix} 1 & -2 & 3 & 5 \\ -4 & 2 & 7 & 1 \\ 3 & 3 & -9 & 1 \end{bmatrix}$$

$$v = \begin{bmatrix} 2 \\ 2 \\ 6 \\ -1 \end{bmatrix}$$

Working Space

Answer on Page 107

Exercise 29 Using Vector Matrix Multiplication

A college professor offers three different methods of determining a student's final grade. In method A, the student's grade is 20% based on attendance, 50% homework, 15% midterm, and 15% final. This professor knows many students can learn the material without attending every class, so with method B the student's grade is 50% homework, 20% midterm, and 30% final. Last, the professor knows some students do not do the homework but still show they understand the material by doing well on the tests. With method C, a student's grade is 40% midterm and 60% final. The professor uses whatever method results in the highest grade to determine each student's final grade.

Suppose Suzy has attended 65% of classes, has an average homework grade of 30%, earned a 80% on the midterm, and earned a 75% on the final. What final grade will her professor post?

Working Space

Answer on Page 107

Vector-Matrix Multiplication in Python

Real problems use matrices far too big to multiply by hand. Check your answers to the exercises above with NumPy's @ operator, as in the Vectors and Matrices chapter:

```
1 import numpy as np
2
3 mix = np.array([[6, 10, 4], [2, 3, 15], [14, 1, 5]])
4 cost = np.array([7.50, 14.75, 22.25])
5 print(mix @ cost)      # [281.5 393. 231. ]
```

5.2.2 Where to Learn More

Watch this video from Khan Academy about matrix-vector products: <https://rb.gy/frga5>

Linear Independence and Dependence

Remember the smoothie shop? Suppose it adds a third drink, the “combo,” which is just one cup of berry blend mixed with one cup of green. The combo looks new on the menu, but it adds nothing you could not already make from the other two. Mathematicians would say the three drinks are *linearly dependent*. This chapter makes that idea precise.

6.1 Linear Independence

Linear Independence

A set of vectors is **linearly independent** if no vector in the set can be written as a linear combination of the others.

Mathematically, this means that

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = \mathbf{0}$$

is only satisfied by

$$c_1 = c_2 = \cdots = c_k = 0$$

Linear Dependence

A set of vectors $\{v_1, \dots, v_k\}$ is **linearly dependent** if there exist scalars $c_1, \dots, c_k$, not all zero, such that

$$c_1v_1 + \cdots + c_kv_k = \mathbf{0}.$$

For two vectors, this reduces to one being a scalar multiple of the other.

When vectors are arranged as columns of a matrix, linear independence means that no column can be written as a linear combination of the others.

We say that each vector contributes a new *direction* not captured by any previous ones, meaning that no vector can be written as a linear combination of the others. Each vector adds something genuinely new to the set, so there is no redundancy.

Think about giving directions on a city grid. If one route goes 2 blocks east and 3 blocks north, then a route of 4 blocks east and 6 blocks north is the same direction, just longer. It adds nothing new. Those two routes are linearly dependent. A route that goes 3 blocks east and 1 block north adds a new direction, so it is independent of the first route.

Exercise 30 Testing Linear Independence

Determine whether each pair of vectors is linearly independent or linearly dependent.

Working Space

1. $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ -2 \end{bmatrix}$

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

3. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$

4. $\begin{bmatrix} -2 \\ 4 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix}$

Answer on Page 108

Any set of vectors that includes the zero vector is automatically linearly dependent, since

$$1 \cdot \vec{0} = \vec{0}$$

gives a nontrivial solution.

Exercise 31 Zero Vector and Dependence

Determine whether each set of vectors is linearly independent.

Working Space

1. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$

3. $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Answer on Page 108

Geometrically, in $\mathbb{R}^2$, two vectors are independent when they do not lie on the same line through the origin. The example below uses perpendicular vectors, but independence does not require that.

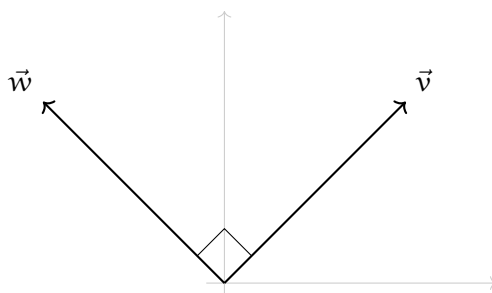


Figure 6.1: An example of independent vectors in $\mathbb{R}^2$.

In $\mathbb{R}^3$, two independent vectors lie in a plane through the origin, while three independent vectors do not lie in the same plane through the origin.

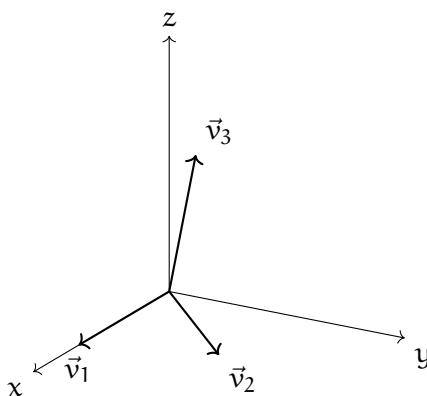


Figure 6.2: An example of linearly independent vectors in $\mathbb{R}^3$.

Exercise 32 **Values Causing Dependence**

Find all values of k for which the vectors are linearly dependent.

Working Space

1. $\begin{bmatrix} 1 \\ k \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$

2. $\begin{bmatrix} k \\ 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2k \\ 2 \\ 4 \end{bmatrix}$

Answer on Page 108

6.1.1 Connection to Span

The *span* of a set of vectors is the collection of all vectors that can be written as linear combinations of those vectors. In other words, the span consists of all vectors that are reachable using the given set.

We now use this idea to better understand linear independence.

Consider the vector space $\mathbb{R}^3$. Let

$$V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Any vector in $\mathbb{R}^3$ can be written as a linear combination of these three vectors. Therefore,

$$\text{span}(V) = \mathbb{R}^3.$$

Moreover, no vector in V can be written as a linear combination of the other two. Each vector contributes a new direction to the span. For this reason, the vectors in V are *linearly independent*.

Equivalently, a set of vectors is linearly independent if every vector in their span has a *unique* representation as a linear combination of those vectors.

This example illustrates an important idea: a linearly independent set contains no redundant vectors, and when such a set spans a space, the set is called a basis for the space.

A set B is a basis if its elements are linearly independent and every element of V is a linear combination of elements of B . In other words, a basis is a linearly independent spanning set. A vector space can have several bases; however all the bases have the same number of elements, called the dimension of the vector space. We will dive into this concept more in the subspaces chapter, so do not worry if this is a bit of a jump!

6.2 Linearly Dependent Vectors: Geometrically

Two vectors are linearly dependent if one is a multiple of the other. Mathematically,

Linearly dependent vectors in $\mathbb{R}^n$

Two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are linearly dependent if one is a scalar multiple of the other: $\mathbf{v} = a\mathbf{u}$ or $\mathbf{u} = a\mathbf{v}$ for some $a \in \mathbb{R}$.

Graphically, two linearly dependent vectors in $\mathbb{R}^2$ lie on the same line through the origin (or one of them is the zero vector).

If two vectors are linearly dependent, then linear combinations of them can only produce vectors lying on that same line. If they are *not* linearly dependent, they are called linearly *independent*, and their linear combinations can produce every vector in $\mathbb{R}^2$.

Example: Which of the following 3 vectors are linearly dependent, if any? $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$,

$$\mathbf{v} = \begin{bmatrix} -3 \\ 4 \\ -1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 6 \\ -8 \\ 2 \end{bmatrix}.$$

Solution: Two vectors are linearly dependent if one is a scalar multiple of the other. Let us compare $\mathbf{u}$ and $\mathbf{v}$. Since the first component of $\mathbf{u}$ is 1 and the first component of $\mathbf{v}$ is -3, let us multiply $\mathbf{u}$ by -3 to see if we get $\mathbf{v}$:

$$-3\mathbf{u} = -3 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \\ -9 \end{bmatrix} \neq \mathbf{v}$$

Therefore, $\mathbf{u}$ and $\mathbf{v}$ are *not* linearly dependent. Now let us examine $\mathbf{v}$ and $\mathbf{w}$. Again, we will use the first components: the first component of $\mathbf{w}$ is 6, so let us see if multiplying $\mathbf{v}$

by -2 yields $\mathbf{w}$:

$$-2\mathbf{v} = -2 \begin{bmatrix} -3 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 6 \\ -8 \\ 2 \end{bmatrix} = \mathbf{w}$$

Therefore, $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent. Since $\mathbf{w} = -2\mathbf{v}$, if $\mathbf{u}$ and $\mathbf{w}$ were dependent, then $\mathbf{u}$ and $\mathbf{v}$ would be too. They are not, so $\mathbf{u}$ and $\mathbf{w}$ are not linearly dependent either.

As a *set*, though, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent, because it contains the dependent pair $\mathbf{v}$ and $\mathbf{w}$: $0\mathbf{u} + 2\mathbf{v} + 1\mathbf{w} = \mathbf{0}$.

Exercise 33 Linear Dependence

Identify which, if any, of the following vectors are linearly dependent:

Working Space

1. $\mathbf{a} = \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix}$

2. $\mathbf{b} = \begin{bmatrix} -4 \\ 5 \\ -3 \end{bmatrix}$

3. $\mathbf{c} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

4. $\mathbf{d} = \begin{bmatrix} 1 \\ -\frac{1}{4} \\ -1 \end{bmatrix}$

5. $\mathbf{e} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$

6. $\mathbf{f} = \begin{bmatrix} -6 \\ \frac{3}{2} \\ 6 \end{bmatrix}$

Answer on Page 109

6.3 Three or More Vectors

With three or more vectors, checking pairs is not enough. The vectors $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ are not multiples of each other, yet the third is the sum of the first two: they are dependent, like the smoothie combo.

To test any set, put the vectors in the columns of a matrix A and solve $A\vec{c} = \mathbf{0}$ by row reducing. If every column gets a pivot, the only solution is $\vec{c} = \mathbf{0}$ and the vectors are independent. If some column has no pivot, there is a free variable, so there are nonzero solutions, and the vectors are dependent.

Example: For the three vectors above,

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - R_1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Column 3 has no pivot, so the vectors are dependent.

One shortcut: in $\mathbb{R}^n$, any set of more than n vectors is dependent, because the matrix has more columns than rows, so some column must miss a pivot. Three vectors in the plane are always dependent.

Exercise 34 Independent or Not?

Are these vectors linearly independent?

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Working Space

Answer on Page 109

6.4 What is next?

In this chapter, we defined linear independence and linear dependence, which describe whether a set of vectors contains redundancy. Linearly independent vectors introduce new directions, while dependent vectors can be written as linear combinations of others.

These ideas lead naturally to the study of subspaces, which describe collections of vectors that are closed under addition and scalar multiplication. In the next chapter, we formalize how sets of vectors generate larger structures and how linear independence helps describe them efficiently.

Subspaces

Picture the room you are sitting in as $\mathbb{R}^3$, with the origin in one corner. Some sets of vectors in the room are special: a flat, endless sheet through the corner, or a straight, endless wire through it. If you start on the sheet and add or stretch vectors that lie on it, you never leave the sheet. These self-contained sets are called *subspaces*, and they describe exactly what the solutions of $A\vec{x} = \vec{0}$ look like (Section 5.1).

Quickly, let us review some vocabulary. The zero vector, denoted as $\vec{0}$, is the vector where all components are zero. It will be a column vector of appropriate size $n \times 1$.

If $A\vec{x} = \vec{0}$, then the system is called **homogeneous**. If $A\vec{x} = \vec{b}$ with $\vec{b} \neq \vec{0}$, then the system is called **non-homogeneous**.

7.1 What is a Subspace?

A subspace V of $\mathbb{R}^n$ is a subset V of $\mathbb{R}^n$ satisfying the following properties:

- The zero vector $\vec{0}$ is in the subspace.
- If $\vec{u}$ and $\vec{v}$ are in the subspace, then their sum $\vec{u} + \vec{v}$ is also in the subspace.
- If $\vec{u}$ is in the subspace and c is a scalar, then the scalar multiple $c\vec{u}$ is also in the subspace.

In short, a subspace is exactly the set of all linear combinations of some collection of vectors. Additionally, every subspace is a span.

Subspace Span

If V is a subspace of $\mathbb{R}^n$, then there exists a set of vectors

$$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subseteq \mathbb{R}^n$$

such that

$$V = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$$

That is, every vector in V can be written as a linear combination of the vectors $\vec{v}_1, \dots, \vec{v}_k$.

Picture a flat sheet of paper that passes through the origin. If you stay on the sheet and

combine vectors that lie on it, you never leave the sheet. A subspace works the same way.

Because V is closed under:

- vector addition, and
- scalar multiplication,

any linear combination of vectors in V must also lie in V .

So if $\vec{u}, \vec{v} \in V$ and $a, b \in \mathbb{R}$, then

$$a\vec{u} + b\vec{v} \in V$$

The converse is also true: any span is a subspace.

$$\text{span}(V) \text{ is a subspace of } \mathbb{R}^n$$

For example, the vectors $\vec{v}_1 = [1, 0]$ and $\vec{v}_2 = [0, 1]$ span all of $\mathbb{R}^2$, since every vector in $\mathbb{R}^2$ is a linear combination of them.

- the span of one nonzero vector $\vec{v}_1$ is a line through $\vec{0}$
- the span of two vectors $\vec{v}_1, \vec{v}_2$ is a plane through $\vec{0}$ (or a line, if the two vectors lie on the same line)
- the span of three vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ is a line, a plane, or all of $\mathbb{R}^3$, depending on how the vectors point

We need to review and reinforce some vocabulary that will come up often during this section. Let us look at Basis and Dimension.

7.2 Basis and Dimension

A basis of a subspace V is a set of linearly independent vectors that span V .

Basis

A set of vectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ in a subspace V is called a **basis** for V if

- the vectors span V , and
- the vectors are linearly independent.

For example, a basis for $\mathbb{R}^3$ is given by the standard unit vectors:

$$V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Note that the span of V is all of $\mathbb{R}^3$, and the vectors in V are linearly independent.

For a subspace V , there are often many possible bases. For example, another basis for $\mathbb{R}^3$ is given by the vectors:

$$W = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Put simply, a basis is a smallest possible set of vectors that can be used to build every vector in the space, with no redundancy. No vector in the basis can be written as a linear combination of the others.

A standard basis is a special type of basis that is often used for $\mathbb{R}^n$.

Standard Basis

The **standard basis** for $\mathbb{R}^n$ is the set of vectors

$$\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\},$$

where $\vec{e}_i$ is the vector in $\mathbb{R}^n$ with a 1 in the i -th position and 0 in all other positions.

An example of this is the standard basis for $\mathbb{R}^2$:

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Recall that we often write these as $\hat{i}$ and $\hat{j}$:

$$\{\hat{i}, \hat{j}\} \quad c_1 \hat{i} + c_2 \hat{j}$$

The standard basis consists of the vectors that point along the coordinate axes. Each standard basis vector measures exactly one coordinate.

The number of vectors in the basis is called the dimension of the subspace.

Dimension

The **dimension** of a vector space V is the number of vectors in any basis for V . If V has a basis consisting of k vectors, then we say that V has dimension k , and write

$$\dim(V) = k.$$

In particular, the dimension of $\mathbb{R}^n$ is n , since the standard basis contains exactly n vectors.

Properties of Dimension:

- A basis of $\mathbb{R}^n$ contains exactly n vectors.
- If W is a linear subspace of V , then $\dim(W) \leq \dim(V)$.
- If V is a finite-dimensional vector space and W is a linear subspace of V with $\dim(W) = \dim(V)$, then $W = V$.
- The space $\mathbb{R}^n$ has the standard basis $\{e_1, \dots, e_n\}$, where e_i is the i -th column of the corresponding identity matrix. Therefore, $\mathbb{R}^n$ has dimension n .

7.3 Nullspace

We now examine a fundamental example of a subspace: the **nullspace** of a matrix. Recall that a subspace is a subset of $\mathbb{R}^n$ that is closed under vector addition and scalar multiplication and contains the zero vector.

The nullspace of a matrix A , denoted $\text{Null}(A)$, is the set of all vectors $\vec{x}$ such that

$$A\vec{x} = \vec{0}.$$

That is, the nullspace is precisely the solution set of the homogeneous system $A\vec{x} = \vec{0}$.

Nullspace

The nullspace of a matrix A is defined as

$$\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}.$$

The nullspace is the set of all vectors $\vec{x}$ that A sends to the zero vector ($\vec{0}$).

Because the equation $A\vec{x} = \vec{0}$ is homogeneous, its solution set always contains the zero

vector. Moreover, if $\vec{x}_1$ and $\vec{x}_2$ are solutions, then any linear combination of the form

$$a\vec{x}_1 + b\vec{x}_2$$

is also a solution. For this reason, the nullspace of a matrix is always a subspace of $\mathbb{R}^n$.

7.3.1 Linear Combinations and Span

Recall that a **linear combination** of vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is any vector of the form

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n,$$

where $a_1, a_2, \dots, a_n \in \mathbb{R}$ are scalars.

The set of all possible linear combinations of a collection of vectors is called their **span**. If a subspace can be written as the span of one or more vectors, those vectors describe all possible directions within the subspace.

7.3.2 Example: Nullspace of a 2 by 2 Matrix

Consider the matrix

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

To find the nullspace of A , we solve the homogeneous system $A\vec{x} = \vec{0}$:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This corresponds to the system of equations

$$x_1 + 2x_2 = 0, \quad 2x_1 + 4x_2 = 0.$$

Solving for x_1 in terms of x_2 gives

$$x_1 = -2x_2.$$

Notice that x_1 is determined by x_2 . This means the columns of A are *linearly dependent*: the second column is 2 times the first. Thus, every vector in the nullspace can be written in the form

$$\vec{x} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Therefore, the nullspace of A is

$$\text{Null}(A) = \text{span} \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}.$$

Geometrically, this nullspace is a line through the origin in $\mathbb{R}^2$, which is a one-dimensional subspace of $\mathbb{R}^2$. The nullspace contains all scalar multiples of the vector $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$. This tells us that the solutions to the homogeneous system $A\vec{x} = \vec{0}$ form a line in $\mathbb{R}^2$. All vectors along this line produce the zero vector when substituted into the equation $A\vec{x} = \vec{0}$.

7.3.3 Example: Nullspace of a 4 by 4 Matrix

Find a basis for the nullspace of the matrix

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 4 & -2 & 0 \\ 0 & 1 & 1 & 3 \\ 0 & 2 & 2 & 6 \end{bmatrix}.$$

First, find the RREF form of A . Notice that $R_2 = 2R_1$ and $R_4 = 2R_3$ in A , so the rows are dependent and the RREF will have zero rows. We get that

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & -3 & -6 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Our $\vec{x}$ must have one entry for each column of A , so it lives in $\mathbb{R}^4$. This gives us $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

So finding $A\vec{x} = \vec{0}$ provides the systems of equations:

$$\begin{cases} x_1 - 3x_3 - 6x_4 = 0 \\ x_2 + x_3 + 3x_4 = 0 \end{cases}$$

We can rewrite this as:

$$\begin{cases} x_1 = 3x_3 + 6x_4 \\ x_2 = -x_3 - 3x_4 \end{cases} \implies x_3 \begin{bmatrix} 3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 6 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

So a basis for the nullspace is formed by

$$\text{Null}(A) = \left\{ \begin{bmatrix} 3 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Exercise 35 Finding a Basis for a Nullspace

Working Space

Consider the matrix

$$A = \begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 2 & -1 & 1 & 5 & 1 \\ 3 & -1 & 1 & 7 & 0 \\ 0 & 1 & -1 & -1 & -3 \end{bmatrix}$$

When put in reduced row echelon form,

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & -1 & -1 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

a Find a basis for the nullspace of A .

b Find a vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$ in the nullspace of A such that $x_1 = -3$, $x_3 = 1$, and $x_4 = 2$.

Answer on Page 109

Key Idea: Nullspace and Linear Independence

Solving for the *nullspace* of a matrix tells you whether a set of vectors is **linearly independent or linearly dependent**.

- If the nullspace contains *only* the zero vector, then the vectors are **linearly independent**.
- If the nullspace contains a *nonzero* vector, then the vectors are **linearly dependent**.

Practical test (via RREF):

- If every column has a *pivot*, the columns are linearly independent.
- If one or more columns lack a pivot, the columns are linearly dependent.

Equivalently, the equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the columns of A are linearly dependent.

7.4 Row Space

The row space of a matrix A , denoted $\text{Row}(A)$, is the subspace of $\mathbb{R}^n$ spanned by the row vectors of A . Each row vector can be viewed as a vector in $\mathbb{R}^n$, and the row space consists of all linear combinations of these row vectors.

Row Space

The row space of a matrix A is defined as

$$\text{Row}(A) = \text{span}\{\text{row}_1, \text{row}_2, \dots, \text{row}_m\},$$

where row_i represents the i -th row of the matrix A . The row space consists of all directions in $\mathbb{R}^n$ that can be built from the rows of the matrix. Row operations replace rows with combinations of rows, so they never change the row space. That means the nonzero rows of $\text{RREF}(A)$ form a convenient basis for $\text{Row}(A)$.

Another way to think about it is that the row space represents all possible linear combinations of the equations represented by the rows of the matrix A . Each row is “tested” on by $\vec{x}$ to produce a component of the output vector $A\vec{x}$. Equivalently, the row space is the set of all vectors that can be formed by adding and scaling the rows of A .

When we compute:

$$A\vec{x}$$

each row of A gets dotted with $\vec{x}$ to produce a component of the output vector.

If each of the rows of A are denoted as $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_m$, then $A\vec{x}$,

$$\begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vec{r}_2 \cdot \vec{x} \\ \vdots \\ \vec{r}_m \cdot \vec{x} \end{bmatrix}$$

The row space, then, is a subspace of $\mathbb{R}^n$ that captures all possible linear combinations of the rows of A . The row space consists of all possible tests on $\vec{x}$ that can be performed by the rows of A .

$$A = \begin{bmatrix} \text{---}\vec{r}_1\text{---} \\ \text{---}\vec{r}_2\text{---} \\ \vdots \\ \text{---}\vec{r}_m\text{---} \end{bmatrix} \quad \text{with } \vec{r}_i \in \mathbb{R}^n$$

Why does the row space live in $\mathbb{R}^n$?

- $\vec{x}$ has n components and lives in $\mathbb{R}^n$, and
- each row $\vec{r}_i$ has n components and lives in $\mathbb{R}^n$

7.4.1 Example: Row Space

Consider the matrix

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

The row vectors of A are

$$\vec{r}_1 = [1 \quad 2 \quad -1], \quad \vec{r}_2 = [0 \quad 1 \quad 3]$$

dotting them with an $\vec{x}$ of appropriate size gives

$$A\vec{x} = \begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vec{r}_2 \cdot \vec{x} \end{bmatrix}$$

The row space of A is the span of its row vectors:

$$\text{Row}(A) = \text{span} \{ [1 \quad 2 \quad -1], [0 \quad 1 \quad 3] \}$$

7.5 Column Space

The column space of a matrix A , denoted $\text{Col}(A)$, is the subspace of $\mathbb{R}^m$ spanned by the column vectors of A . Each column of A can be viewed as a vector in $\mathbb{R}^m$, and the column space consists of all linear combinations of these column vectors.

Column Space

The **column space** of a matrix A is defined as

$$\text{Col}(A) = \text{span}\{\vec{c}_1, \vec{c}_2, \dots, \vec{c}_n\},$$

where $\vec{c}_i$ represents the i -th column of the matrix A . The column space consists of all vectors in $\mathbb{R}^m$ that can be formed as linear combinations of the columns of A .

Equivalently, the column space can be described as the set of all possible outputs of the matrix:

$$\text{Col}(A) = \{A\vec{x} : \vec{x} \in \mathbb{R}^n\}.$$

The columns of the *original* matrix A that correspond to pivot columns in $\text{RREF}(A)$ form a basis for the column space of A . We will discuss this idea further in a future chapter, but it corresponds to the fact that the pivot columns are linearly independent and form a basis for the column space.

7.5.1 Example: Column Space

Consider the matrix

$$A = \begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 7 & 3 & 9 \\ 1 & 5 & 3 & 1 \\ 1 & 2 & 0 & 8 \end{bmatrix}.$$

After applying elementary row operations, we find that the RREF of A is

$$\text{RREF}(A) = \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Column 3 has no pivot, so it depends on the pivot columns. In fact, column 3 equals -2 times column 1 plus column 2. The pivot columns are columns 1, 2, and 4. Thus, a basis

for the column space of A is given by the original columns 1, 2, and 4:

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 7 \\ 5 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 9 \\ 1 \\ 8 \end{bmatrix}.$$

This is a three-dimensional subspace of $\mathbb{R}^4$, since the rank is 3. It is not all of $\mathbb{R}^4$.

Exercise 36 Nullspace, Rowspace, and Columnspace

Working Space

Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & -2 & 6 \\ 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{bmatrix}.$$

- Compute the reduced row echelon form of A .
- Identify the pivot columns and free columns of A . What does this tell you about the solutions?
- Find a basis for the **nullspace** $N(A)$. What is the dimension of $N(A)$?
- Find a basis for the **row space** $R(A)$. What is the dimension of $R(A)$?
- Find a basis for the **column space** $C(A)$. State clearly where the vectors for the basis come from.

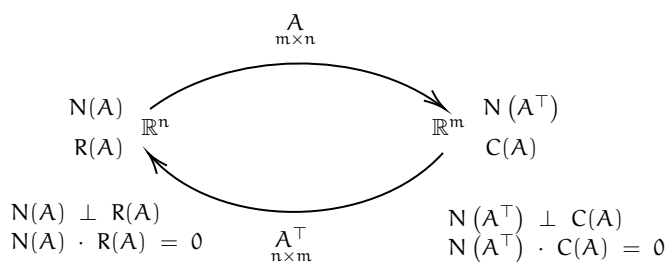
Answer on Page 110

7.6 Summary

In this chapter, we have established a few different common subspaces: the Nullspace, Row space, and Column space. Each of these subspaces captures different structural information about a matrix and the system of equations it represents. The nullspace describes all input vectors that are sent to the zero vector, the column space contains all possible outputs of the matrix, and the row space encodes the independent constraints imposed by the system.

We have now set the stage for a new way of viewing matrices. Rather than thinking of a matrix solely as a collection of numbers or equations, we can begin to interpret a matrix as a transformation – a rule that takes input vectors and maps them to output vectors. In the next few chapters, we will build on our understanding of subspaces to study how matrices act on vectors and reshape space, providing a unifying viewpoint for many of the ideas introduced so far. We will look at the geometric implications of matrices and how they *transform* space. Ideas like this sit behind many computer graphics programs, such as the ones that move objects in video games.

Before ending this chapter, take a look at this graphic. It shows the relationship between the subspaces we talked about, plus a new one: $\text{Null}(A^T)$, the nullspace of the transpose (sometimes called the *left nullspace*). If matrix A has size $m \times n$, then A^T has size $n \times m$. Note which subspaces live in $\mathbb{R}^n$ and which live in $\mathbb{R}^m$.



Instead of asking what vectors solve $A\vec{x} = \vec{0}$, we will begin asking:

What does the matrix A do to an arbitrary vector?

The big take away: A matrix of size $m \times n$ can be viewed as a rule that takes vectors in $\mathbb{R}^n$ as inputs and produces vectors in $\mathbb{R}^m$ as outputs.

Answers to Exercises

Answer to Exercise 1 (on page 10)

$\nabla f = (1, 1)$ and $\nabla g = (2x, 2y)$, so $1 = 2\lambda x$ and $1 = 2\lambda y$. Thus $x = y$, and $2x^2 = 8$ gives $x = y = \pm 2$. The maximum is $f(2, 2) = 4$ and the minimum is $f(-2, -2) = -4$. In a sketch, the level lines $x + y = c$ are parallel lines, and the two extreme ones just touch the circle at $(2, 2)$ and $(-2, -2)$.

Answer to Exercise 2 (on page 11)

1. Maximize xy subject to $x + 2y = 100$. The equations are $y = \lambda$ and $x = 2\lambda$, so $x = 2y$. Then $4y = 100$, so $y = 25$ m, $x = 50$ m, and the area is 1250 m².
2. $\lambda = 25$. Each extra meter of fence adds about 25 m² of field. (Check: with F meters of fence, the best area is $\frac{F^2}{8}$, and its derivative at $F = 100$ is $\frac{100}{4} = 25$.)

Answer to Exercise 3 (on page 11)

The equations are $yz = 2\lambda(y + z)$, $xz = 2\lambda(x + z)$, and $xy = 2\lambda(x + y)$. Multiply the first by x and the second by y : $xyz = 2\lambda(xy + xz)$ and $xyz = 2\lambda(xy + yz)$, so $xz = yz$, and $x = y$. In the same way, $y = z$. So the box is a cube, $6x^2 = 24$ gives $x = 2$, and the largest volume is 8 .

Answer to Exercise 4 (on page 11)

$y = 2\lambda x$ and $x = 8\lambda y$. Substituting, $y = 16\lambda^2 y$, so $\lambda = \pm \frac{1}{4}$ (the case $y = 0$ forces $x = 0$, which is not on the ellipse). With $\lambda = \frac{1}{4}$, $x = 2y$, and $4y^2 + 4y^2 = 8$ gives $y = \pm 1$: the points $(2, 1)$ and $(-2, -1)$, where $f = 2$. With $\lambda = -\frac{1}{4}$, $x = -2y$: the points $(-2, 1)$ and $(2, -1)$, where $f = -2$. The maximum is 2 and the minimum is -2 . In the SymPy program, use $f = x*y$ and $g = x**2 + 4*y**2$, with the constraint `sp.Eq(g, 8)`.

Answer to Exercise 5 (on page 12)

$\frac{\partial H}{\partial p} = -\ln p - 1$, and similarly for q and s . So $-\ln p - 1 = \lambda$, $-\ln q - 1 = \lambda$, and $-\ln s - 1 = \lambda$, which means $\ln p = \ln q = \ln s$, so $p = q = s = \frac{1}{3}$. The most unpredictable spinner gives every color the same chance, which makes sense: any imbalance would make one color a safer guess.

Answer to Exercise 6 (on page 25)

1. Linear: $3(a + b) = 3a + 3b$ and $3(ca) = c(3a)$.
2. Not linear: $T(1 + 1) = 7$, but $T(1) + T(1) = 8$. (It is affine.)
3. Not linear: $T(1 + 1) = 4$, but $T(1) + T(1) = 2$.
4. Linear. (This is a reflection.)

A quick test: a linear operation always sends 0 to 0. That alone rules out $3x + 1$.

Answer to Exercise 7 (on page 26)

```
1 import numpy as np
2
3 f = np.array([[1, 2], [3, 4]])
4 g = np.array([[0, 1], [1, 0]])
5 print(f + g)
6 print(f @ g)
7 print(f.T)
```

The output is:

```
1 [[1 3]
2  [4 4]]
3 [[2 1]
4  [4 3]]
5 [[1 3]
6  [2 4]]
```

Answer to Exercise 8 (on page 27)

1. 3×5
2. 1×2
3. 4×3

Answer to Exercise 9 (on page 28)

1. The matrix should have 1 row and 3 columns. For example,

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

2. The matrix should have 2 rows and 4 columns. For example,

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

3. The matrix should have 4 rows and 3 columns. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}$$

Answer to Exercise 10 (on page 29)

$$1. \mathbf{A} + \mathbf{B} = \begin{bmatrix} -2 & 7 & -2 & 10 \end{bmatrix}, \mathbf{A} - \mathbf{B} = \begin{bmatrix} 2 & 1 & 2 & 0 \end{bmatrix}, \mathbf{B} - \mathbf{A} = \begin{bmatrix} -2 & -1 & -2 & 0 \end{bmatrix}$$

$$2. \mathbf{A} + \mathbf{B} = \begin{bmatrix} 9 & -4 & -3 \\ -4 & -6 & 3 \\ -10 & 6 & -4 \end{bmatrix}, \mathbf{A} - \mathbf{B} = \begin{bmatrix} -1 & -4 & -1 \\ 6 & 0 & 7 \\ 0 & 0 & 4 \end{bmatrix}, \mathbf{B} - \mathbf{A} = \begin{bmatrix} 1 & 4 & 1 \\ -6 & 0 & -7 \\ 0 & 0 & -4 \end{bmatrix}.$$

$$3. \mathbf{A} + \mathbf{B} = \begin{bmatrix} -7 & -3 & -2 & -6 \\ 5 & 1 & 0 & 0 \end{bmatrix}, \mathbf{A} - \mathbf{B} = \begin{bmatrix} 3 & 1 & -8 & 4 \\ 5 & -9 & 8 & 6 \end{bmatrix}, \mathbf{B} - \mathbf{A} = \begin{bmatrix} -3 & -1 & 8 & -4 \\ -5 & 9 & -8 & -6 \end{bmatrix}.$$

Answer to Exercise 11 (on page 30)

It is possible to compute $\mathbf{a} \cdot \mathbf{d}$, $\mathbf{b} \cdot \mathbf{e}$, and $\mathbf{c} \cdot \mathbf{f}$:

$$1. \mathbf{a} \cdot \mathbf{d} = 1(-5) + 2(-1) = -5 + (-2) = -7$$

$$2. \mathbf{b} \cdot \mathbf{e} = -3(1) + 3(-5) + 5(3) + -5(1) = -3 + (-15) + 15 - 5 = -8$$

$$3. \mathbf{c} \cdot \mathbf{f} = 1(4) + 2(1) + -1(-3) = 4 + 2 + 3 = 9$$

Answer to Exercise 12 (on page 32)

$$1. [16 \quad -1 \quad 13]$$

$$2. \begin{bmatrix} 0 & 5 & 1 \\ 0 & 25 & 5 \\ 0 & -25 & -5 \\ 0 & 20 & 4 \\ 0 & 5 & 1 \end{bmatrix}$$

$$3. \begin{bmatrix} -9 & -17 & -13 \\ -6 & 40 & 14 \\ 15 & 7 & 11 \\ 9 & -8 & -7 \end{bmatrix}$$

Answer to Exercise 13 (on page 33)

$$1. A \times B = \begin{bmatrix} 8 & -6 & 10 & 4 \\ -8 & 6 & -10 & -4 \\ -4 & 3 & -5 & -2 \\ 8 & -6 & 10 & 4 \end{bmatrix} \text{ and } B \times A = [13]$$

$$2. A \times B = \begin{bmatrix} -2 & 16 & 16 \\ 5 & -24 & -8 \\ -4 & 22 & 12 \end{bmatrix} \text{ and } B \times A = \begin{bmatrix} 8 & 6 \\ -12 & -22 \end{bmatrix}$$

$$3. A \times B = \begin{bmatrix} -22 & 1 \\ 15 & -4 \end{bmatrix} \text{ and } B \times A = \begin{bmatrix} 12 & 0 & 15 & 3 \\ -4 & 0 & 1 & -15 \\ -16 & 0 & -17 & -11 \\ -14 & 0 & -10 & -21 \end{bmatrix}$$

Answer to Exercise 14 (on page 34)

$$A \times B = \begin{bmatrix} -3 & 5 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 \\ 4 & -3 \end{bmatrix} = \begin{bmatrix} -3(-1) + 5(4) & -3(1) + 5(-3) \\ -1(-1) + 0(4) & -1(1) + 0(-3) \end{bmatrix} = \begin{bmatrix} 23 & -18 \\ 1 & -1 \end{bmatrix}$$

$$B \times A = \begin{bmatrix} -1 & 1 \\ 4 & -3 \end{bmatrix} \cdot \begin{bmatrix} -3 & 5 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1(-3) + 1(-1) & -1(5) + 1(0) \\ 4(-3) + -3(-1) & 4(5) + -3(0) \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ -9 & 20 \end{bmatrix}$$

Answer to Exercise 15 (on page 40)

Yes, it is symmetric:

$$A^T = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Answer to Exercise 16 (on page 40)

Row 1: $5(1) + 1(2) + 3(3) + (-2)(-8) = 5 + 2 + 9 + 16 = 32$. Row 2: $1 - 2 + 24 - 32 = -9$.
Row 3: $6 + 4 + 3 - 24 = -11$.

Answer to Exercise 17 (on page 44)

1. We are looking for a_1 and a_2 such that:

$$a_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + a_2 \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

Which creates the system of equations:

$$a_1 - 3a_2 = 4$$

$$2a_1 + a_2 = 5$$

We can multiply the first equation by -2 and add it to the second to solve for a_2 :

$$-2(a_1 - 3a_2) + 2a_1 + a_2 = -2(4) + 5$$

$$6a_2 + a_2 = -8 + 5$$

$$7a_2 = -3$$

$$a_2 = -\frac{3}{7}$$

Substituting a_2 back into an equation and solving for a_1 :

$$a_1 - 3\left(-\frac{3}{7}\right) = 4$$

$$a_1 + \frac{9}{7} = 4$$

$$a_1 = \frac{19}{7}$$

Therefore, $\frac{19}{7} \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{3}{7} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$.

2. We are looking for a_1 and a_2 such that:

$$a_1 \begin{bmatrix} 9 \\ 4 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

Which creates the system of equations:

$$9a_1 = -5$$

$$4a_1 + a_2 = 3$$

We can find a_1 from the first equation:

$$a_1 = -\frac{5}{9}$$

Substituting for a_1 back into the second equation and solving for a_2 :

$$4\left(-\frac{5}{9}\right) + a_2 = 3$$

$$a_2 - \frac{20}{9} = 3$$

$$a_2 = \frac{47}{9}$$

Therefore, $-\frac{5}{9} \begin{bmatrix} 9 \\ 4 \end{bmatrix} + \frac{47}{9} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$.

3. We are looking for a_1 and a_2 such that:

$$a_1 \begin{bmatrix} 7 \\ -2 \end{bmatrix} + a_2 \begin{bmatrix} -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

Which yields the system of equations:

$$7a_1 - 8a_2 = 6$$

$$-2a_1 + 4a_2 = -2$$

Doubling the second equation and adding it to the first:

$$7a_1 - 8a_2 + 2(-2a_1 + 4a_2) = 6 + 2(-2)$$

$$7a_1 - 8a_2 - 4a_1 + 8a_2 = 6 - 4$$

$$3a_1 = 2$$

$$a_1 = \frac{2}{3}$$

Substituting for a_1 back into the second equation and solving for a_2 :

$$-2\left(\frac{2}{3}\right) + 4a_2 = -2$$

$$-\frac{4}{3} + 4a_2 = -2$$

$$4a_2 = -\frac{2}{3}$$

$$a_2 = -\frac{1}{6}$$

Therefore, $\frac{2}{3} \begin{bmatrix} 7 \\ -2 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$.

Answer to Exercise 18 (on page 50)

1. No. $\begin{bmatrix} -6 \\ -2 \end{bmatrix} = -2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, so they are linearly dependent. They span only the line through the origin in the direction of $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$: the line $y = \frac{x}{3}$.
2. Yes. They are not multiples of each other. (For any $\begin{bmatrix} a \\ b \end{bmatrix}$, take $\frac{a+b}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{a-b}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.)

Answer to Exercise 19 (on page 53)

1. The second vector is -2 times the first, so the span is a *line*.
2. Neither vector is a multiple of the other, so the span is a *plane*.
3. The first two are not multiples of each other, so they span a plane. Is $[3, 3, 2]$ a linear combination of them? Solving $a_1[3, 0, 0] + a_2[0, 3, 3] = [3, 3, 2]$ gives $3a_1 = 3$, $3a_2 = 3$, and $3a_2 = 2$. The last two equations disagree, so there is no solution. The third vector is off the plane, and the span is $\mathbb{R}^3$.

Answer to Exercise 20 (on page 53)

The first two are not multiples of each other, so they span a plane. Solve $a_1[1, 1, 0] + a_2[0, 1, 1] = [1, 0, 1]$: the first entry gives $a_1 = 1$, the third gives $a_2 = 1$, and then the second entry would be $a_1 + a_2 = 2$, not 0. There is no solution, so $[1, 0, 1]$ is off the plane, and the three vectors span $\mathbb{R}^3$. (In the next chapter, row reduction will give you a faster way to answer questions like this.)

Answer to Exercise 21 (on page 56)

$$A = \begin{bmatrix} -4 & 9 & -8 \\ -1 & 0 & 6 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$$A\vec{x} = \vec{b} \implies \begin{bmatrix} -4 & 9 & -8 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

Answer to Exercise 22 (on page 67)

Notice immediately that $R_2 = 2R_1$, which automatically means redundancy.

$$\begin{aligned} & \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & -2 & 6 \\ 1 & 1 & 0 & 2 \end{bmatrix} \xrightarrow{\substack{R_2 \leftarrow R_2 - 2R_1 \\ R_3 \leftarrow R_3 - R_1}} \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix} \\ & \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \leftarrow -R_2} \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ & \xrightarrow{R_1 \leftarrow R_1 - 2R_2} \boxed{\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}} \end{aligned}$$

The RREF is:

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

This gives us the system:

$$\begin{aligned} x_1 + x_3 &= 1, \\ x_2 - x_3 &= 1. \end{aligned}$$

So,

$$x_1 = 1 - x_3, \quad x_2 = 1 + x_3$$

making x_3 free. We can write this as: $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $x_3 \in \mathbb{R}$ Equivalently, as a set:

$$\left\{ \begin{pmatrix} 1 - x_3 \\ 1 + x_3 \\ x_3 \end{pmatrix} : x_3 \in \mathbb{R} \right\}$$

Since x_3 can be anything, there are infinitely many solutions. The row of zeros shows that the system has a free variable, which is another way to see this.

Answer to Exercise 23 (on page 67)

Start with $R_3 \leftarrow R_3 - 2R_1$ to clear the first entry in row 3:

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

Next, $R_3 \leftarrow R_3 - 3R_2$ gives the row echelon form:

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & -8 \end{bmatrix}$$

Then $R_3 \leftarrow -R_3$, $R_2 \leftarrow R_2 + R_3$, and $R_1 \leftarrow R_1 - 2R_3$ give

$$\begin{bmatrix} 1 & -1 & 0 & -16 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & 8 \end{bmatrix}$$

Finally, $R_1 \leftarrow R_1 + R_2$ gives the RREF:

$$\begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & 8 \end{bmatrix}$$

This gives a unique solution:

$$x_1 = -5, \quad x_2 = 11, \quad x_3 = 8$$

or as a vector $\vec{x} = \begin{bmatrix} -5 \\ 11 \\ 8 \end{bmatrix}$

Answer to Exercise 24 (on page 68)

Notice that one row of the matrix states $[0 \ 0 \ 0 \ 1]$, which implies $0x_1 + 0x_2 + 0x_3 = 1$, which is not possible. Therefore the matrix represents an inconsistent system and there is no solution.

Answer to Exercise 25 (on page 68)

The RREF of this matrix is

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \xrightarrow[\substack{R_2 \leftarrow R_2 - 2R_1 \\ R_3 \leftarrow R_3 - 3R_1}]{} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that there are two rows of zeroes, and the first column has a 1. This implies two dependent rows ($R_2 = 2R_1$, $R_3 = 3R_1$), and with only 1 pivot column, $\text{rank}(A) = 1$.

Answer to Exercise 26 (on page 68)

By definition in this chapter:

$$\text{rank}(A) = \text{number of pivots} = 3$$

A 4×6 matrix has $n = 6$ columns, so

$$\# \text{ free variables} = n - \text{rank}(A) = 6 - 3 = 3.$$

So, free variables = 3

Answer to Exercise 27 (on page 69)

There are $n = 4$ variables (x_1, x_2, x_3, x_4) .

1. Pivot columns are columns 1, 2, 4, so the pivot columns are x_1, x_2, x_4 .
2. Free variables: the only non-pivot column is column 3, so the free variable is x_3
3. The dimension of the solution set is the number of free variables: 1. Equivalently, $\text{dimension} = n - \text{rank} = 4 - 3 = 1$

Answer to Exercise 28 (on page 73)

$$Av = \begin{bmatrix} 11 \\ 37 \\ -43 \end{bmatrix}$$

Answer to Exercise 29 (on page 74)

The different methods can be represented in a matrix:

$$\begin{bmatrix} 0.20 & 0.50 & 0.15 & 0.15 \\ 0 & 0.50 & 0.20 & 0.30 \\ 0 & 0 & 0.4 & 0.6 \end{bmatrix}$$

And Suzy's individual grades can be represented by a vector:

$$\begin{bmatrix} 65 \\ 30 \\ 80 \\ 75 \end{bmatrix}$$

To see the results of the three different methods, we can multiply the matrix and the vector:

$$\begin{bmatrix} 0.20 & 0.50 & 0.15 & 0.15 \\ 0 & 0.50 & 0.20 & 0.30 \\ 0 & 0 & 0.4 & 0.6 \end{bmatrix} \cdot \begin{bmatrix} 65 \\ 30 \\ 80 \\ 75 \end{bmatrix} = \begin{bmatrix} 0.2(65) + 0.5(30) + 0.15(80) + 0.15(75) \\ 0(65) + 0.5(30) + 0.2(80) + 0.3(75) \\ 0(65) + 0(30) + 0.4(80) + 0.6(75) \end{bmatrix}$$

Which yields:

$$\begin{bmatrix} 51.25 \\ 53.5 \\ 77 \end{bmatrix}$$

Since method C yields the highest grade, the professor will post a final grade of 77.

Answer to Exercise 30 (on page 78)

Two vectors are linearly dependent if one is a scalar multiple of the other.

1. Dependent, since $\begin{bmatrix} 4 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.
2. Dependent, since $\begin{bmatrix} -6 \\ -2 \end{bmatrix} = -2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.
3. Independent, since neither vector is a scalar multiple of the other.
4. Dependent, since $\begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} -2 \\ 4 \\ 6 \end{bmatrix}$.

Answer to Exercise 31 (on page 78)

Any set containing the zero vector is linearly dependent, since a nontrivial linear combination can equal the zero vector.

1. Dependent.
2. Dependent.
3. Dependent.

Answer to Exercise 32 (on page 80)

1. The vectors are dependent when $\begin{bmatrix} 1 \\ k \end{bmatrix}$ is a scalar multiple of $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$. This occurs when $k = 2$.
2. The vectors are dependent for all values of k , since

$$\begin{bmatrix} 2k \\ 2 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} k \\ 1 \\ 2 \end{bmatrix}.$$

Answer to Exercise 33 (on page 82)

We see that $\frac{\mathbf{a}}{-4} = -\frac{1}{4} \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{4} \\ -1 \end{bmatrix} = \mathbf{d}$. Additionally, $\frac{3}{2}\mathbf{a} = \frac{3}{2} \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -6 \\ \frac{3}{2} \\ 6 \end{bmatrix} = \mathbf{f}$.

Therefore, vectors $\mathbf{a}$, $\mathbf{d}$, and $\mathbf{f}$ are linearly dependent.

We also see that $\frac{1}{2}\mathbf{c} = \frac{1}{2} \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \mathbf{e}$. Therefore, vectors $\mathbf{c}$ and $\mathbf{e}$ are linearly dependent. Vector $\mathbf{b}$ is not linearly dependent to any of the other vectors.

Answer to Exercise 34 (on page 83)

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - 2R_1} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & -4 & 1 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 + 4R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

Every column has a pivot, so the vectors are independent.

Answer to Exercise 35 (on page 91)

a Solving $A\vec{x} = \vec{0}$ gives us two equations:

$$\begin{cases} x_1 + 2x_4 - x_5 = 0 \\ x_2 - x_3 - x_4 - 3x_5 = 0 \end{cases} \implies \begin{cases} x_1 = -2x_4 + x_5 \\ x_2 = x_3 + x_4 + 3x_5 \end{cases}$$

So x_3 , x_4 , and x_5 are free variables, and every vector in the nullspace has the form

$$\vec{x} = \begin{bmatrix} -2x_4 + x_5 \\ x_3 + x_4 + 3x_5 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

or, equivalently

$$\vec{x} = x_3 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

So a basis can be formed by the set

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

b We are given a vector $\vec{x} = \begin{bmatrix} -3 \\ x_2 \\ 1 \\ 2 \\ x_5 \end{bmatrix}$.

Recall that we have $x_1 = -2x_4 + x_5$, so substituting the known values gives $-3 = -2(2) + x_5 \implies x_5 = 1$.

We can then solve $x_2 = x_3 + x_4 + 3x_5$ for x_2 :

$$x_2 = 1 + 2 + 3(1) = 6$$

These two variables give us the $\vec{x}$

$$\vec{x} = \begin{bmatrix} -3 \\ 6 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

Answer to Exercise 36 (on page 95)

a

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

b Columns 1, 2, and 3 are pivot columns (containing a pivot). Column 4 is a non-pivot or free column. This tells you that there are 3 pivot variables and 1 free variable. The nullspace will be 1 dimensional.

c We have the system of equations:

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

Letting x_4 be free, $\vec{x} = x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$. The $N(A) = \left\{ \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ with a dimension of 1.

- d The row space is the subspace spanned by the rows of A . A convenient basis for the row space is given by the nonzero rows of $\text{rref}(A)$. Thus a basis is

$$\{[1 \ 0 \ 0 \ 1], [0 \ 1 \ 0 \ 1], [0 \ 0 \ 1 \ 0]\}.$$

These three rows are linearly independent, and they span the row space because row reduction did not create any new row space, just simplified the set. So $\dim \text{Row}(A) = 3$.

- e For the column space, we must take columns from the *original* matrix A . The usual rule is: pivot columns of the original A form a basis for $C(A)$. Since the pivot columns are 1, 2, 3, a basis is the set of the first three columns of A

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 2 \\ 4 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

and the column space is the span of each original pivot column:

$$C(A) = \text{span}\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$$



INDEX

- identity matrix, 34
- affine, 15
- affine transformation, 15
- augmented matrix, 57
- basis, 67, 81, 87
- bitmap, 14
- coefficient matrix, 55
- column space, 94
- column-major order, 36
- constrained optimization, 5
- constraint, 5
- diagonal matrix, 38
- dimension, 67, 87
- dividing matrices, 41
- dot product, 30
 - application, 73
- elementary row operations, 58
- entropy, 12
- EROs, *see* elementary row operations
- false color images, 15
- free variable, 64
- Halo orbit, 23
- homogeneous, 85
- inconsistent set, 66
- independence
 - in $\mathbb{R}^3$, 79
- inverse matrix, 41
- Lagrange multiplier, 7
- Lagrange multipliers, 5
 - meaning of lambda, 10
- Lagrange point, 23
- line
 - span of vectors, 50
- linear, 14
- linear algebra, 13
- linear combination, 14, 43
 - and span, 89
 - as a system of equations, 46
 - example, 43
 - matrix form, 72
 - no solution, 45
- linear combination of vectors, 43
- linear combinations, 13, 43
 - graphing vectors of, 46
- linear dependence, 81, 89
- linear image operations, 14
- linear independence, 77
 - testing with RREF, 83
- linear operation, 14
- linearly dependent, 49
- matrices, 16, 27
 - normalization, 21
 - properties of, 38
 - shapes of, 37

- vector in form of, 27
- matrix, 27
- matrix addition, 29
- matrix equation, 56
- matrix multiplication, 30
- matrix transpose, 35
- matrix-vector multiplication, 74
- maximum likelihood, 9

- no solution, 66
- non-commutative, 33
- nonlinear operations, 13
- nullspace, 88
 - and linear independence, 92
- NumPy, 24
 - linalg.solve, 71
 - matrices, 40

- objective, 5

- pivot, 64
- plane, 51
- Python Imaging Library, 24

- rank, 66
- reduced row-echelon form, 58, 90
- row space, 92
- row-echelon form, 56
- row-major order, 36

- scalar multiple, 77
- shadow price, 10
- span, 43, 50
 - connection to independence, 80
 - definition, 50
 - example, 52
 - plane, 51
- spanning, 48
- standard basis, 87
- standard basis vectors, 46
- subspace, 85
- symmetric matrices, 39
- system of equations, 55

- trail mix example, 72
- triangular matrix, 38

- unique solution, 59

- vector, 27
- vector space, 85

- weighted grades, 74

- zero matrix, 28
- zero row, 64
- zero vector, 85
 - dependence, 78