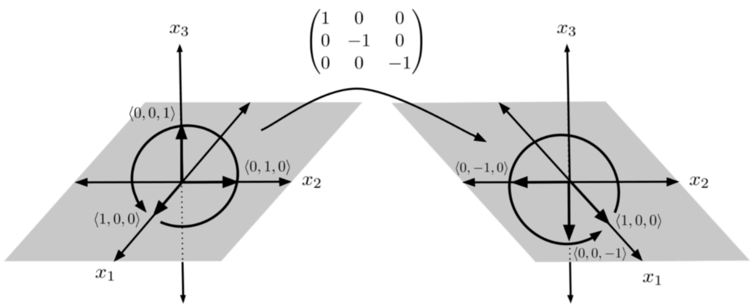
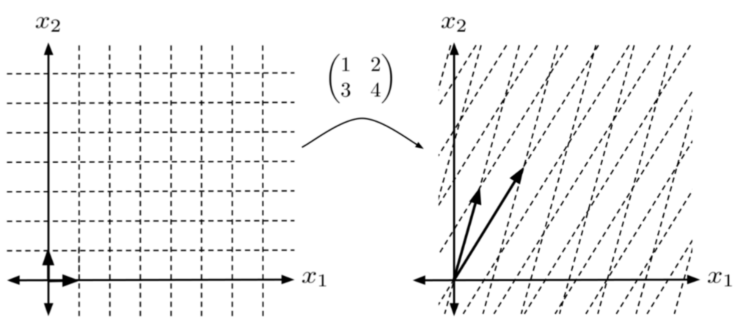
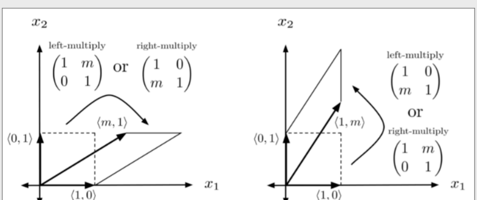
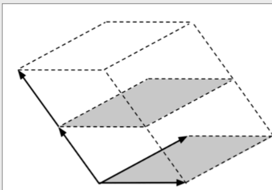


Linear Algebra







Justin Skycak

First edition.

www.justinmath.com

Acknowledgements

A sincere thank you to Sanjana Kulkarni for her thoughtful suggestions and diligent proofreading of this book.

Table of Contents

Part 1
Vectors

1.1 N-Dimensional Space

N-dimensional space consists of points that have N components. For example, $(0,0)$ is the origin in 2-dimensional space, $(0,0,0)$ is the origin in 3-dimensional space, and $(0,0,0,0)$ is the origin in 4-dimensional space.

Similarly, the points $(0,0)$, $(1,0)$, $(0,1)$ are the corners of a triangle in 2-dimensional space, the points $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ are the corners of a tetrahedron in 3-dimensional space, and the points $(0,0,0,0)$, $(1,0,0,0)$, $(0,1,0,0)$, $(0,0,1,0)$, $(0,0,0,1)$ are the corners of a "hypertetrahedron" in 4-dimensional space.

Functions with Multiple Inputs and Outputs

We're used to seeing single variables as inputs and outputs to functions, but functions can really take any number of variables as input and produce any number of variables as output.

For example, the function $f(x,y) = x+y$ takes two input variables, and adds them to produce a single output variable. Thus, it maps points in 2-dimensional space onto points in 1-dimensional space.

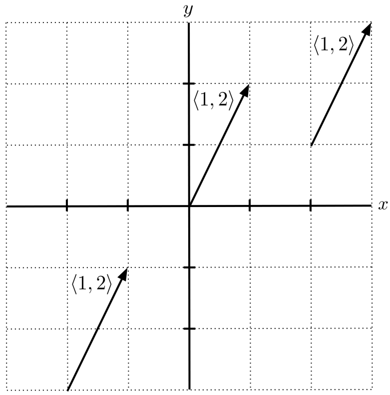
Similarly, the function $f(x,y) = \left(x+y, x-y, xy, \frac{x}{y} \right)$ takes two input variables and produces four output variables: the sum, difference, product, and quotient of the inputs. Thus, it maps points in 2-dimensional space onto points in 4-dimensional space.

Lastly, the function $f(x_1, x_2, \ldots, x_n) = (1x_1, 2x_2, \ldots, mx_m)$ takes N input variables and produces M output variables, which are just the first M input variables multiplied by their indices. Thus, it maps points in N-dimensional space onto points in M-dimensional space.

Vectors

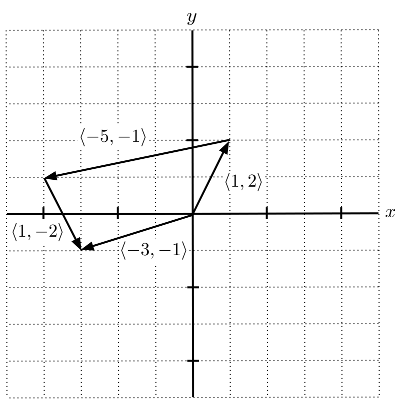
Points in N-dimensional space consist of numbers, but can also be thought of as manipulable entities in their own right, called **vectors**. When we think of points as vectors, we cease to think of them as fixed points in space. Instead, we think of them as displacements through space.

For example, the vector $\langle 1, 2 \rangle$ can represent the displacement from the point $(0,0)$ to the point $(1,2)$ -- but it can also represent the displacement from $(2,1)$ to $(3,3)$, or $(-2,-3)$ to $(-1,-1)$, or any other point (x,y) to $(x+1, y+2)$.



Vectors can be added component-wise, and adding a sequence of vectors together yields a net displacement through all the vectors combined, starting each vector where the previous one ends.

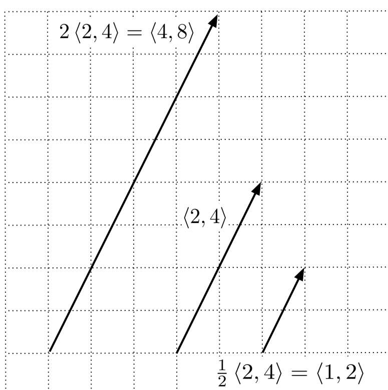
$$\begin{aligned}\langle 1, 2\rangle + \langle -5, -1\rangle + \langle 1, -2\rangle &= \langle 1 - 5 + 1, 2 - 1 - 2\rangle \\ &= \langle -3, -1\rangle\end{aligned}$$



Scalars

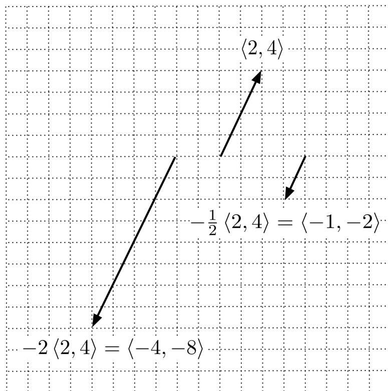
Vectors can also be multiplied by regular numbers called **scalars**. Multiplying a vector by a scalar has the effect of rescaling a vector to become shorter or longer, depending on the magnitude of the scalar.

$$\begin{aligned} 2\langle 2,4\rangle &= \langle 2(2), 2(4)\rangle \\ &= \langle 4,8\rangle \end{aligned} \qquad \begin{aligned} \frac{1}{2}\langle 2,4\rangle &= \left\langle \frac{1}{2}(2), \frac{1}{2}(4)\right\rangle \\ &= \langle 1,2\rangle \end{aligned}$$



If the scalar is negative, then the vector also flips direction, in addition to being rescaled by the magnitude of the scalar.

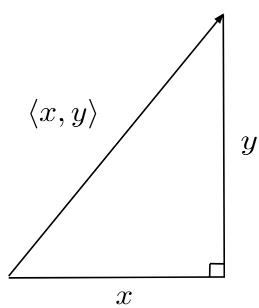
$$\begin{aligned} -2\langle 2,4\rangle &= \langle -2(2), -2(4)\rangle \\ &= \langle -4,-8\rangle \end{aligned} \qquad \begin{aligned} -\frac{1}{2}\langle 2,4\rangle &= \left\langle -\frac{1}{2}(2), -\frac{1}{2}(4)\right\rangle \\ &= \langle -1,-2\rangle \end{aligned}$$



Norm of a Vector

In two dimensions, a vector's length, called its **norm**, can be obtained using the Pythagorean theorem:

$$|\langle x, y \rangle| = \sqrt{x^2 + y^2}$$



In general, the norm of a vector can be computed by extending the Pythagorean theorem to higher dimensions.

$$|\langle x_1, x_2, \ldots, x_n \rangle| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

To see that this definition of the norm is compatible with the idea that scalar multiplication rescales a vector, observe that the norm of a scaled vector is equal to the product of the scalar and the norm of the unscaled vector.

$$
\begin{aligned}
|c\langle x_1, x_2, \ldots, x_n\rangle| &= |\langle cx_1, cx_2, \ldots, cx_n\rangle| \\
&= \sqrt{(cx_1)^2 + (cx_2)^2 + \ldots + (cx_n)^2} \\
&= \sqrt{c^2x_1^2 + c^2x_2^2 + \ldots + c^2x_n^2} \\
&= \sqrt{c^2(x_1^2 + x_2^2 + \ldots + x_n^2)} \\
&= \sqrt{c^2}\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \\
&= c\,|\langle x_1, x_2, \ldots, x_n\rangle|
\end{aligned}
$$

Algebra with Vectors

Lastly, expressions involving multiple vector operations follow the standard rules of arithmetic, and equations involving vector variables follow the standard rules of algebra.

$$
\begin{aligned}
2\left(\langle 3, 2\rangle - x\right) &= \langle 4, 2\rangle \\
2\langle 3, 2\rangle - 2x &= \langle 4, 2\rangle \\
\langle 6, 4\rangle - 2x &= \langle 4, 2\rangle \\
-2x &= \langle 4, 2\rangle - \langle 6, 4\rangle \\
-2x &= \langle -2, -2\rangle \\
x &= -\frac{1}{2}\langle -2, -2\rangle \\
x &= \langle 1, 1\rangle
\end{aligned}
$$

We can also use algebra to solve for unknown components of vectors.

$$\begin{aligned}
\left|2\left\langle 3,2,x\right\rangle\right| &= \left|\left\langle 1,2,3\right\rangle + \left\langle 5,4,1\right\rangle\right| \\
2\left|\left\langle 3,2,x\right\rangle\right| &= \left|\left\langle 6,6,4\right\rangle\right| \\
\left|\left\langle 3,2,x\right\rangle\right| &= \frac{1}{2}\left|\left\langle 6,6,4\right\rangle\right| \\
\left|\left\langle 3,2,x\right\rangle\right| &= \left|\frac{1}{2}\left\langle 6,6,4\right\rangle\right| \\
\left|\left\langle 3,2,x\right\rangle\right| &= \left|\left\langle 3,3,2\right\rangle\right| \\
\sqrt{3^2+2^2+x^2} &= \sqrt{3^2+3^2+2^2} \\
\sqrt{9+4+x^2} &= \sqrt{9+9+4} \\
\sqrt{13+x^2} &= \sqrt{22} \\
13+x^2 &= 22 \\
x^2 &= 9 \\
x &= \pm 3
\end{aligned}$$

Exercises

For each function, list the dimensionalities of the input space and the output space.

1) $f(x,y) = (x,y,x+y)$ 2) $f(x) = x^3$

3) $f(x,y,z) = xyz$

4) $f(x, y) = (xy, x^2y^2, \ldots, x^ny^n)$

5) $f(x_1, x_2, \ldots, x_m) = (x_1 + x_2, x_2 + x_3, \ldots, x_{m-1} + x_m)$

6) $f(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n) = \left(\dfrac{x_1 + x_2 + \ldots + x_m}{m}, \sqrt[n]{y_1 y_2 \ldots y_n} \right)$

Perform the indicated vector operations.

7) $3\langle -1, 2, 1\rangle + \dfrac{1}{2}\langle 4, 4, -2\rangle$ 8) $4\left(\langle 1, 2\rangle - \langle 3, 2\rangle\right)$

9) $|\langle 1, 0, 1, 2, 0\rangle| + |\langle 0, 0, -4, 4, -8\rangle|$

10) $|\langle -1, 1, 0\rangle - \langle -1, 3, -2\rangle|$

11) $|\langle 1, 2\rangle| \left(\langle 1, 1, 1, -1\rangle - \langle 1, -1, 1, 1\rangle\right)$

12) $\dfrac{\langle 12, 4, -16, 0, -5\rangle + \langle 3, 1, -4, 5, -5\rangle}{|\langle 0, 3, -4\rangle|}$

Solve for the unknown variable.

13) $\langle 1, 2\rangle + x = \langle 4, 7\rangle$

14) $\langle 1, 7, -3, 8, 2\rangle = \langle -2, 8, 0, x, 3\rangle - \langle -3, 1, 3, 6, 1\rangle$

15) $|\langle 3, x\rangle| \langle 1, 2, 3\rangle = \langle 5, 10, 15\rangle$

16) $\langle x+1, x-1, 3\rangle + \langle 0,1,2\rangle = \langle -1,-2,5\rangle$

17) $3\left(x + \langle 1,-1,1,-1\rangle\right) = 2x$

18) $|\langle x,x,x\rangle + \langle -1,0,1\rangle| = 2$

1.2 Dot Product and Cross Product

We know how to multiply a vector by a scalar, but what does it mean to multiply a vector by another vector?

The two most common interpretations of vector multiplication are the **dot product**, and for vectors in 3 dimensions, the **cross product**.

Dot Product

The **dot product** is computed as the sum of products of components.

$$\langle x_1, x_2, \ldots, x_n \rangle \cdot \langle y_1, y_2, \ldots, y_n \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

First of all, notice that the dot product of a vector with itself is just the vector's norm, squared.

$$\begin{aligned}
&\langle x_1, x_2, \ldots, x_n \rangle \cdot \langle x_1, x_2, \ldots, x_n \rangle \\
&= x_1^2 + x_2^2 + \ldots + x_n^2 \\
&= \left(\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \right)^2 \\
&= |\langle x_1, x_2, \ldots, x_n \rangle|^2
\end{aligned}$$

Also notice that the dot product can distribute over sums of vectors, just like multiplication.

$$\begin{aligned}
&\langle x_1, x_2, \ldots, x_n \rangle \cdot (\langle y_1, y_2, \ldots, y_n \rangle + \langle z_1, z_2, \ldots, z_n \rangle) \\
&= \langle x_1, x_2, \ldots, x_n \rangle \cdot (\langle y_1 + z_1, y_2 + z_2, \ldots, y_n + z_n \rangle) \\
&= x_1(y_1 + z_1) + x_2(y_2 + z_2) + \ldots + x_n(y_n + z_n) \\
&= x_1 y_1 + x_1 z_1 + x_2 y_2 + x_2 z_2 + \ldots + x_n y_n + x_n z_n \\
&= (x_1 y_1 + x_2 y_2 + \ldots + x_n y_n) + (x_1 z_1 + x_2 z_2 + \ldots + x_n z_n) \\
&= \langle x_1, x_2, \ldots, x_n \rangle \cdot \langle y_1, y_2, \ldots, y_n \rangle + \langle x_1, x_2, \ldots, x_n \rangle \cdot \langle z_1, z_2, \ldots, z_n \rangle
\end{aligned}$$

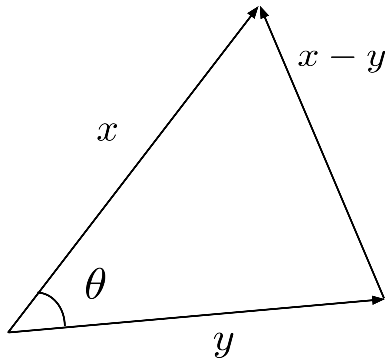
One can also verify that the dot product behaves like multiplication in other ways -- for example for two vectors $x = \langle x_1, x_2, \ldots, x_n \rangle$ and $y = \langle y_1, y_2, \ldots, y_n \rangle$ and any scalar c we have $(cx) \cdot y = x \cdot (cy) = c(x \cdot y)$ and $x \cdot y = y \cdot x$.

Geometric Interpretation of Dot Product

Using the law of cosines on a triangle whose sides are formed by the vectors x, y, and $x - y$, we can find a geometric interpretation of the dot product:

$$x \cdot y = |x||y| \cos \theta,$$

where θ is the angle between the two vectors x and y.



$$\begin{aligned}
|x-y|^2 &= |x|^2 + |y|^2 - 2|x||y|\cos\theta \\
(x-y)\cdot(x-y) &= |x|^2 + |y|^2 - 2|x||y|\cos\theta \\
x\cdot x - x\cdot y - y\cdot x + y\cdot y &= |x|^2 + |y|^2 - 2|x||y|\cos\theta \\
|x|^2 - 2x\cdot y + |y|^2 &= |x|^2 + |y|^2 - 2|x||y|\cos\theta \\
-2x\cdot y &= -2|x||y|\cos\theta \\
x\cdot y &= |x||y|\cos\theta
\end{aligned}$$

One interesting consequence of this formula is that perpendicular vectors have a dot product of zero: the angle between perpendicular vectors is $\frac{\pi}{2}$, and $|x||y|\cos\frac{\pi}{2} = |x||y|(0) = 0$.

Even if the dot product is not zero, we can still use it to compute the angle between the two vectors.

$$\begin{aligned} x \cdot y &= |x||y|\cos\theta \\ \frac{x \cdot y}{|x||y|} &= \cos\theta \\ \arccos\left(\frac{x \cdot y}{|x||y|}\right) &= \theta \end{aligned}$$

Cross Product

For 3-dimensional vectors, we also have another interpretation of vector multiplication called the **cross product**. The cross product is given by

$$\begin{aligned} &\langle x_1, x_2, x_3 \rangle \times \langle y_1, y_2, y_3 \rangle \\ &= \langle x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1 \rangle. \end{aligned}$$

Using the above definition, one can verify that the cross product distributes over sums and satisfies $(cx) \times y = x \times (cy) = c(x \times y)$ for any scalar c.

However, when the two vectors in a cross product are interchanged, the result changes sign: $x \times y = -y \times x$. This is a key difference between the cross product and the dot product.

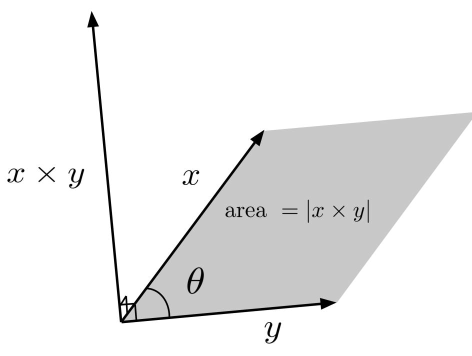
Geometric Interpretation of Cross Product

Like the dot product, the cross product also has a geometric interpretation:

$$|x \times y| = |x||y| \sin \theta$$

This is similar to the geometric interpretation of the dot product, except we have $\sin \theta$ instead of $\cos \theta$, and we are talking about the norm of the vector resulting from the cross product.

As a result, the cross product $x \times y$ represents a vector whose norm is equal to the area enclosed by the parallelogram that has x and y as sides. Moreover, the cross product produces a vector that is perpendicular to the vectors x and y.



To see that the cross product $x \times y$ is perpendicular to x and y, observe that the dot products $(x \times y) \cdot x$ and $(x \times y) \cdot y$ both evaluate to 0.

$$\begin{aligned}&(\langle x_1, x_2, x_3\rangle \times \langle y_1, y_2, y_3\rangle) \cdot \langle x_1, x_2, x_3\rangle \\ &= \langle x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1\rangle \cdot \langle x_1, x_2, x_3\rangle \\ &= x_1(x_2y_3 - x_3y_2) + x_2(x_3y_1 - x_1y_3) + x_3(x_1y_2 - x_2y_1) \\ &= x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_1x_2y_3 + x_1x_3y_2 - x_2x_3y_1 \\ &= 0\end{aligned}$$

$$\begin{aligned}&(\langle x_1, x_2, x_3\rangle \times \langle y_1, y_2, y_3\rangle) \cdot \langle y_1, y_2, y_3\rangle \\ &= \langle x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1\rangle \cdot \langle y_1, y_2, y_3\rangle \\ &= y_1(x_2y_3 - x_3y_2) + y_2(x_3y_1 - x_1y_3) + y_3(x_1y_2 - x_2y_1) \\ &= x_2y_1y_3 - x_3y_1y_2 + x_3y_1y_2 - x_1y_2y_3 + x_1y_2y_3 - x_2y_1y_3 \\ &= 0\end{aligned}$$

To understand why $|x \times y| = |x|\,|y| \sin\theta$, we can begin by squaring both sides of the equation and expressing the right-hand side using the dot product.

$$\begin{aligned}|x \times y|^2 &= |x|^2\,|y|^2 \sin^2\theta \\ &= |x|^2\,|y|^2 \left(1 - \cos^2\theta\right) \\ &= |x|^2\,|y|^2 - |x|^2\,|y|^2 \cos^2\theta \\ &= |x|^2\,|y|^2 - (x \cdot y)^2\end{aligned}$$

Now, we expand out the right hand side using $x = \langle x_1, x_2, x_3\rangle$ and $y = \langle y_1, y_2, y_3\rangle$. We find that some terms cancel, and the remaining terms can be rearranged into the square of the norm of the cross product.

$$\begin{aligned}
&\left(x_1^2+x_2^2+x_3^2\right)\left(y_1^2+y_2^2+y_3^2\right)-(x_1y_1+x_2y_2+x_3y_3)^2 \\
&= \begin{pmatrix} x_1^2y_1^2+x_1^2y_2^2+x_1^2y_3^2 \\ x_2^2y_1^2+x_2^2y_2^2+x_2^2y_3^2 \\ x_3^2y_1^2+x_3^2y_2^2+x_3^2y_3^2 \end{pmatrix} - \begin{pmatrix} x_1^2y_1^2+x_1x_2y_1y_2+x_1x_3y_1y_3 \\ +x_1x_2y_1y_2+x_2^2y_2^2+x_2x_3y_2y_3 \\ +x_1x_3y_1y_3+x_2x_3y_2y_3+x_3^2y_3^2 \end{pmatrix} \\
&= x_2^2y_3^2+x_3^2y_2^2-2x_2x_3y_2y_3+x_3^2y_1^2+x_1^2y_3^2-2x_1x_3y_1y_3 \\
&\quad +x_1^2y_2^2+x_2^2y_1^2-2x_1x_2y_1y_2 \\
&= (x_2y_3-x_3y_2)^2+(x_3y_1-x_1y_3)^2+(x_1y_2-x_2y_1)^2 \\
&= |\langle x\times y\rangle|^2
\end{aligned}$$

Exercises

Evaluate the following vector expressions.

1) $\langle 1,2,4\rangle \cdot \langle -1,-2,3\rangle$

2) $\langle 2,-2,3,3\rangle \cdot \langle -1,1,1,-1\rangle$

3) $2\langle 2,3\rangle \cdot \frac{1}{2}\langle 3,-2\rangle$

4) $\frac{1}{2}\langle 1,4,-2\rangle \cdot (\langle 3,3,1\rangle - \langle 1,3,3\rangle)$

5) $\langle -2,1,0\rangle \times \langle -2,3,0\rangle$

6) $\langle 3,1,5\rangle \times \langle 2,-2,-1\rangle$

7) $\langle 1,1,2\rangle \times (\langle 2,4,6\rangle - \langle 1,4,7\rangle)$

8) $\frac{1}{3}\langle 1,0,2\rangle \times \langle 1,1,1\rangle \times \langle 0,0,3\rangle$

Solve for x.

9) $\langle x,1,2\rangle \cdot \langle 1,2,x\rangle = -4$

10) $\langle x,1\rangle \cdot (\langle 1,1\rangle - \langle 1,x\rangle) = 1$

11) $\langle x,-1,0,1\rangle \cdot \langle x,0,5,7\rangle = 16$

12) $\langle 1,2,3\rangle \times \langle 1,x,3\rangle = \langle -6,0,2\rangle$

13) $\langle 1,1,1\rangle \times (\langle 2,1,-2\rangle + \langle x,1,1\rangle) = \langle -3,2,1\rangle$

14) $\frac{1}{2}\langle 3,1,-2\rangle \times (\langle 0,x,0\rangle - \langle 1,0,2\rangle) = \langle 2,4,5\rangle$

Use the dot product to find the angle θ between the two vectors.

15) $\langle 1,1\rangle$
$\langle 1,2\rangle$

16) $\langle -2,3,1\rangle$
$\langle 0,1,-1\rangle$

17) $\langle 2,0,3,-1\rangle$
$\langle 1,1,0,2\rangle$

18) $\langle 2,0,-1,0,2\rangle$
$\langle -1,2,3,0,0\rangle$

Use the cross product to find the area contained by each parallelogram whose sides are given as vectors.

19) $\langle 2, -1, 0 \rangle$
$\langle 4, 1, 0 \rangle$

20) $\langle 1, 2, 3 \rangle$
$\langle 3, 2, 1 \rangle$

21) $\langle 1, 0, 1 \rangle$
$\langle 1, 1, 1 \rangle$

22) $\langle 3, -3, 1 \rangle$
$\langle -4, 1, -2 \rangle$

1.3 Lines and Planes

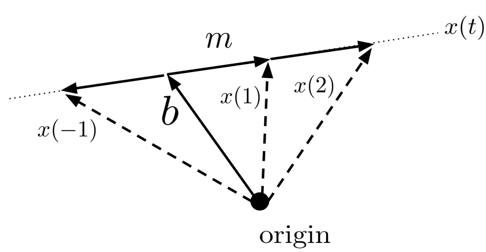
A line starts at an initial point and proceeds straight in a constant direction. Thus, we can write the equation of a line as

$$x(t) = mt + b$$

where

- $b = \langle b_1, b_2, \ldots, b_N \rangle$ is the initial point,
- $m = \langle m_1, m_2, \ldots, m_N \rangle$ is the constant direction in which the line travels, and
- $x(t) = \langle x_1(t), x_2(t), \ldots, x_N(t) \rangle$ is the point reached by traveling t units away from b in the direction of m.

(Though $x(t)$ is actually a vector, we can also refer to it as the point where the vector lands when the vector is placed at the origin.)



Finding the Equation of a Line

For example, to compute the line between the points $(1,2,3,4)$ and $(5,-2,3,7)$ in 4-dimensional space, we can start by computing the direction m as the difference between the two points:

$$\begin{aligned} m &= \langle 5,-2,3,7\rangle - \langle 1,2,3,4\rangle \\ &= \langle 4,-4,0,3\rangle \end{aligned}$$

Taking $b = \langle 1,2,3,4\rangle$ as our initial point, then, we can express the line as

$$x(t) = \langle 4,-4,0,3\rangle\, t + \langle 1,2,3,4\rangle.$$

If we wanted to find another point on the line, we could substitute another value for t, say, $t=-2$.

$$\begin{aligned} x(2) &= \langle 4,-4,0,3\rangle\,(-2) + \langle 1,2,3,4\rangle \\ &= \langle -8,8,0,-6\rangle + \langle 1,2,3,4\rangle \\ &= \langle -7,10,3,-2\rangle \end{aligned}$$

Checking Whether a Point is on a Line

If we wanted to check whether the point $\langle -1,-2,-3,-4\rangle$ is on the line, we could substitute this point for $x(t)$ and try to solve for t.

$$\begin{aligned} \langle -1,-2,-3,-4\rangle &= \langle 4,-4,0,3\rangle\, t + \langle 1,2,3,4\rangle \\ \langle -1,-2,-3,-4\rangle - \langle 1,2,3,4\rangle &= \langle 4,-4,0,3\rangle\, t \\ \langle -2,-4,-6,-8\rangle &= \langle 4,-4,0,3\rangle\, t \end{aligned}$$

Setting first components equal, we find $-2 = 4t$, which implies that $t = -\frac{1}{2}$. But equating second components yields $-4 = -4t$, which implies that $t = 1$. So, there is no solution that matches all pairs of components, and consequently the point $\langle -1,-2,-3,-4\rangle$ is not on the line.

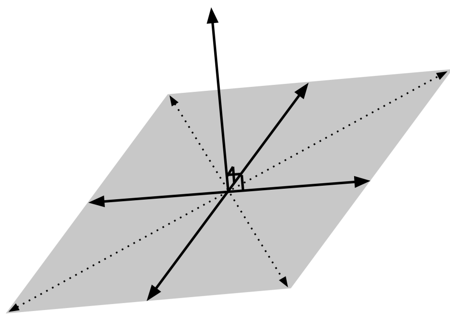
However, we can verify that the point $\langle 9,-6,3,10\rangle$ is on the line using the same method.

$$\begin{aligned} \langle 9,-6,3,10\rangle &= \langle 4,-4,0,3\rangle\, t + \langle 1,2,3,4\rangle \\ \langle 9,-6,3,10\rangle - \langle 1,2,3,4\rangle &= \langle 4,-4,0,3\rangle\, t \\ \langle 8,-8,0,6\rangle &= \langle 4,-4,0,3\rangle\, t \end{aligned}$$

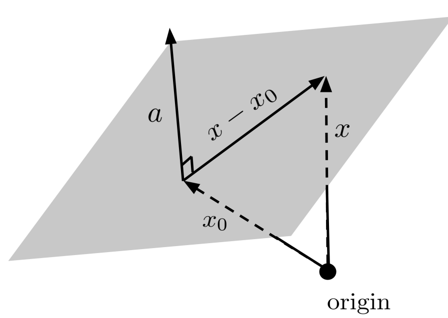
Equating first components yields $8 = 4t$ which is valid for $t = 2$; equating second components yields $-8 = -4t$ which is also valid for $t = 2$; equating third components yields $0 = 0$ which is valid for all choices of t; and equating fourth components yields $6 = 3t$ which is also valid for $t = 2$. Thus the point $\langle 9,-6,3,10\rangle$ is on the line because it is simply $x(t)$ evaluated at $t = 2$.

The Equation of a Plane

Now, let's talk about how to write the equation of a plane in N-dimensional space. A plane can be visualized as a flat sheet that makes a right angle with some particular vector. Thus, a plane just consists of all vectors through some point in the plane, that are perpendicular to a single vector.



If x_0 is a point in the plane, then the vectors in the plane can be written $x - x_0$, where x represents other points on the plane. These vectors are all perpendicular to a single vector, call it a, so their dot product must be zero: $a \cdot (x - x_0) = 0$.



Distributing the dot product, we have $a \cdot x - a \cdot x_0 = 0$, and rearranging we have $a \cdot x = a \cdot x_0$. The right-hand side $a \cdot x_0$ is just a constant, so we can simply call it k. Thus, we have the general equation for a plane:

$$a \cdot x = k$$

Here, a is a vector that is perpendicular to the plane, x are points on the plane, and k is some constant. Writing $a = \langle a_1, a_2, \ldots, a_N \rangle$ and $x = \langle x_1, x_2, \ldots, x_N \rangle$, we can expand out the general equation for a plane into an equation consisting only of scalars:

$$a_1 x_1 + a_2 x_2 + \ldots + a_N x_N = k$$

Finding a Plane Given a Point and Perpendicular Vector

For example, to compute the plane that passes through the point $\langle 5, 3, 1 \rangle$ and has a perpendicular vector of $\langle 1, -1, 2 \rangle$, we can start by setting up the equation with the perpendicular vector substituted.

$$\langle 1, -1, 2 \rangle \cdot x = k$$

To solve for k, we can simply substitute the point $\langle 5, 3, 1 \rangle$ for x and take the dot product.

$$\begin{aligned} \langle 1, -1, 2 \rangle \cdot \langle 5, 3, 1 \rangle &= k \\ 1(5) - 1(3) + 2(1) &= k \\ 4 &= k \end{aligned}$$

Then, we can substitute for k and expand out the dot product in the initial equation.

$$\begin{aligned}\langle 1,-1,2\rangle \cdot x &= 4 \\ \langle 1,-1,2\rangle \cdot \langle x_1,x_2,x_3\rangle &= 4 \\ x_1 - x_2 + 2x_3 &= 4\end{aligned}$$

Now, suppose we have an equation for a plane, and we want to find the perpendicular vector.

$$2x_1 - 5x_2 + 3 = 10 - x_3 + x_4$$

To do this, we can simply organize the equation and convert it to the vector equation of the plane, using the dot product.

$$\begin{aligned}2x_1 - 5x_2 + 3 &= 10 - x_3 + x_4 \\ 2x_1 - 5x_2 + x_3 - x_4 &= 7 \\ \langle 2,-5,1,-1\rangle \cdot \langle x_1,x_2,x_3,x_4\rangle &= 7 \\ \langle 2,-5,1,-1\rangle \cdot x &= 7\end{aligned}$$

The perpendicular vector is just the first vector in the dot product, $\langle 2,-5,1,-1\rangle$.

Finding a Plane Given Three Points

Lastly, suppose that we want to find the equation of the plane that contains the three points $(1,2,3)$, $(5,4,-1)$, and $(-2,3,0)$.

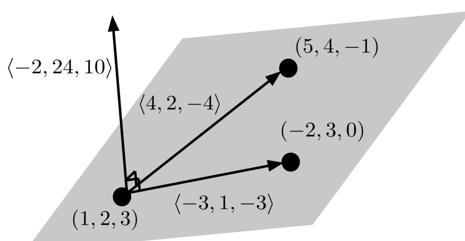
To start off, we can find two vectors within the plane by starting at one of the points, say $(1,2,3)$, and computing the displacement vectors to the other two points:

$$\langle 5,4,-1\rangle - \langle 1,2,3\rangle = \langle 4,2,-4\rangle$$
$$\langle -2,3,0\rangle - \langle 1,2,3\rangle = \langle -3,1,-3\rangle$$

These displacement vectors are within the plane, so if we can find a vector that is perpendicular to these displacement vectors, then we will have a vector that is normal to the plane.

Since these displacement vectors are 3-dimensional, we can compute their cross product, which yields a perpendicular vector.

$$\begin{aligned}&\langle 4,2,-4\rangle \times \langle -3,1,-3\rangle \\ &= \langle 2(-3)-(-4)(1), -4(-3)-4(-3), 4(1)-2(-3)\rangle \\ &= \langle -2,24,10\rangle\end{aligned}$$



Using this vector as the normal vector of the plane, the equation of the plane becomes

$$-2x+24y+10z=k$$

for some constant k. To find the value of k, we can simply substitute one of the points in the plane, say, $(1,2,3)$.

$$\begin{aligned} -2(1)+24(2)+10(3) &= k \\ 76 &= k \end{aligned}$$

Thus, the equation of the plane is

$$-2x+24y+10z=76$$

which can be simplified to

$$-x+12y+5z=38.$$

Looking back, we could have saved some work by using the vector $\langle 2,1,-2\rangle$ instead of the displacement vector $\langle 4,2,-4\rangle$, since $\langle 2,1,-2\rangle$ follows the same direction (it's just half as long).

The normal vector resulting from the cross product would then have been $\langle -1,12,5\rangle$, which is the normal vector in the fully simplified equation of the plane.

(The vector $\langle -1,12,5\rangle$ points in the same direction as the original normal vector $\langle -2,24,10\rangle$; it's just half as long.)

Exercises

Compute the equation of the line that passes through the given points.

1) $(1,2,3)$
$(3,2,1)$

2) $(1,-1,2,-2)$
$(0,1,4,-3)$

3) $(-5,-1,2,1,3)$
$(4,4,2,2,0)$

4) $(1,0,-3,-3,0,1)$
$(3,1,-1,0,2,0)$

Check whether the given point P is on the given line. If so, determine the value of t for which $x(t)=P$.

5) $P(5,10,5)$
$x(t)=\langle 1,4,1\rangle+\langle 2,3,2\rangle\, t$

6) $P(2,7,16,0)$
$x(t)=\langle 2,-1,0,3\rangle+\langle 0,1,2,5\rangle\, t$

7) $P(5,6,7,8,9)$
$x(t)=\langle 1,2,3,4,5\rangle+\langle 5,4,3,2,1\rangle\, t$

8) $P(-2,3,-9,9,-13,12)$
$x(t)=\langle 3,-2,1,-1,2,-3\rangle+\langle -1,1,-2,2,-3,3\rangle\, t$

Write the equation of the plane that contains the given point P and is perpendicular to the given normal vector n.

9) $P(0,0,0)$
$n = \langle 1,2,3 \rangle$

10) $P(1,-1,2)$
$n = \langle 2,-2,1 \rangle$

11) $P(3,0,2,1)$
$n = \langle 4,-2,0,-5 \rangle$

12) $P(1,0,1,0,1)$
$n = \langle 1,-2,-3,2,-1 \rangle$

Write the equation of the plane that contains the three given points.

13) $(0,0,0)$
$(0,1,1)$
$(1,0,1)$

14) $(1,3,-5)$
$(2,-1,2)$
$(1,1,1)$

15) $(2,0,-1)$
$(-2,-3,-4)$
$(1,2,3)$

16) $(5,1,1)$
$(0,3,-2)$
$(0,1,1)$

1.4 Span, Subspaces, and Reduction

The **span** of a set of vectors consists of all vectors that can be made by adding multiples of vectors in the set.

For example, the span of the set $\{\langle 1,0\rangle, \langle 0,1\rangle\}$ is just the entirety of the 2-dimensional plane: any vector $\langle x,y\rangle$ in this plane can be made by adding $x\langle 1,0\rangle + y\langle 0,1\rangle$. For instance, $\langle 5,-2\rangle$ can be written as $5\langle 1,0\rangle - 2\langle 0,1\rangle$.

Similarly, the span of the set $\{\langle 1,1\rangle, \langle -1,1\rangle\}$ is also the entirety of the 2-dimensional plane: any vector $\langle x,y\rangle$ in this plane can be made by adding $\left(\frac{x+y}{2}\right)\langle 1,1\rangle + \left(\frac{x-y}{2}\right)\langle 1,-1\rangle$. This is a little less obvious, but it's true: for example, $\langle 3,-7\rangle$ can be written as $-2\langle 1,1\rangle + 5\langle 1,-1\rangle$.

The span of the set $\{\langle 1,1\rangle, \langle 2,2\rangle\}$, however, is just a 1-dimensional line within the 2-dimensional plane: it contains only vectors of the form $k\langle 1,1\rangle$, where k is a constant. To see why this is, observe what happens when we try to add multiples of the vectors:

$$\begin{aligned} &a\langle 1,1\rangle + b\langle 2,2\rangle \\ &= a\langle 1,1\rangle + 2b\langle 1,1\rangle \\ &= (a+2b)\langle 1,1\rangle \\ &= k\langle 1,1\rangle \end{aligned}$$

Subspaces of Two-Dimensional Space

The span of $\{\langle 1,1\rangle, \langle 2,2\rangle\}$ is a 1-dimensional line within the 2-dimensional plane. So, we say that the span forms a 1-dimensional **subspace** of the 2-dimensional plane.

In 2 dimensions, it turns out that any set of two vectors that are multiples of one another will span a line, and any set of two vectors that are NOT multiples of one another will span the entire space.

For example, the set $\{\langle 2,-3\rangle, \langle -4,6\rangle\}$ spans a line because $\langle -4,6\rangle = -2\langle 2,-3\rangle$. On the other hand, the set $\{\langle 7,4\rangle, \langle 1,-2\rangle\}$ spans the entire space because the vectors cannot be written as multiples of each other.

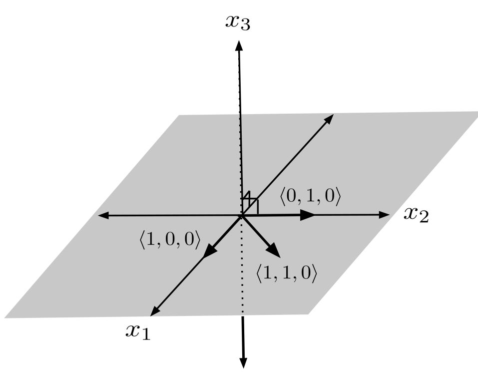
But just because two vectors are multiples, doesn't mean they can't be included in a set that spans the space. For example, the set $\{\langle 1,1\rangle, \langle 2,2\rangle\}$ spans only a line, but if we add include a third vector $\langle 1,2\rangle$ that is not a multiple of the original two vectors, then the set $\{\langle 1,1\rangle, \langle 2,2\rangle, \langle 1,2\rangle\}$ spans the entire plane.

Subspaces of N-Dimensional Space

Now, let's generalize these ideas to N dimensions. It might be tempting to think that in general, a set of vectors will span the entire space provided there are some three vectors that aren't multiples of one another. But this isn't always true.

For example, consider the set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 1,1,0\rangle\}$. None of these vectors are multiples of each other, but there is no way to combine the vectors to reach a point whose third component is not zero.

The issue here is that the third vector is the sum of the first two vectors. As a result, the third vector is redundant -- we can already reach any point using the first two vectors, that we can reach using the third vector. The set, then, has the same span as the set of just the first two vectors. It covers just a plane, a 2-dimensional subspace of 3-dimensional space.



The vectors span the plane $x_3 = 0$. No matter how we combine vectors, the third component will always be 0 -- so we cannot reach any points above or below the plane, by adding multiples of vectors in the set.

Independence

In general, the dimension of the span of a set of vectors is equal to the number of **independent** vectors that remain after we remove the **dependent** vectors. A vector is said to be *dependent* if it can be written as a sum of multiples of other vectors in the set.

The labeling of vectors as independent or dependent depends on the order in which the vectors are considered, but regardless of order, removing all dependent vectors will leave the same number of independent vectors, even if the independent vectors themselves are different for different orders.

For example, in the set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 1,1,0\rangle\}$ we can start by looking at the first vector, $\langle 1,0,0\rangle$. This vector is dependent since it can be produced by subtracting the other two vectors: $\langle 1,0,0\rangle = \langle 1,1,0\rangle - \langle 0,1,0\rangle$. Removing this vector from the set yields the reduced set $\{\langle 0,1,0\rangle, \langle 1,1,0\rangle\}$, which contains two independent vectors and thus cannot be reduced any further.

Since the fully reduced set has two independent vectors, it spans a 2-dimensional plane, and since the original set has the same span as the reduced set, the original set also spans the same 2-dimensional plane.

Alternatively, beginning with the original set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 1,1,0\rangle\}$ we can start by looking at the second vector, $\langle 0,1,0\rangle$. This vector is dependent since it can be produced by subtracting the other two vectors: $\langle 0,1,0\rangle = \langle 1,1,0\rangle - \langle 1,0,0\rangle$.

Removing this vector from the set yields the reduced set $\{\langle 1,0,0\rangle, \langle 1,1,0\rangle\}$, which contains two independent vectors and thus cannot be reduced any further.

Again, since the fully reduced set has two independent vectors, it spans a 2-dimensional plane, and since the original set has the same span as the reduced set, the original set also spans the same 2-dimensional plane.

The last alternative, beginning with the original set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 1,1,0\rangle\}$, is to start by looking at the third vector, $\langle 1,1,0\rangle$. This vector is dependent since it can be produced by adding the other two vectors: $\langle 1,1,0\rangle = \langle 0,1,0\rangle + \langle 1,1,0\rangle$. Removing this vector from the set yields the reduced set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle\}$, which contains two independent vectors and thus cannot be reduced any further.

Again, since the fully reduced set has two independent vectors, it spans a 2-dimensional plane, and since the original set has the same span as the reduced set, the original set also spans the same 2-dimensional plane.

Maximum Number of Independent Vectors

Since the number of independent vectors in a set tells us the dimension of the span of that set, we can make the general conclusion that **a set of N-dimensional vectors can have at most N independent vectors.**

The N-dimensional vectors reside in N-dimensional space, so the largest space they can possibly span is their full N-dimensional space. Consequently, it’s not possible for the set of vectors to contain more than N independent vectors -- otherwise, they would need to span a space of more than N dimensions.

For example, consider the following set of vectors:

$$\{\langle 1,1\rangle, \langle 1,-2\rangle, \langle -3,1\rangle, \langle 2,3\rangle\}$$

Looking at the first two vectors $\langle 1,1\rangle$ and $\langle 1,-2\rangle$, we see that these two vectors are independent since they are not multiples of each other. As a result, the span of the vectors must have a dimension of at least 2.

But the vectors reside in 2-dimensional space, so their span is limited to at most 2 dimensions. Thus, we can conclude that the set of vectors spans exactly 2 dimensions, and that the third and fourth vectors in the set must be dependent, without even needing to check whether they can be written as sums of multiples of the first two vectors.

Now, consider the following set of vectors:

$$\left\{\begin{matrix} \langle 1,-1,1,2\rangle \\ \langle 2,0,4,2\rangle \\ \langle -1,3,3,2\rangle \\ \langle 3,-4,1,4\rangle \end{matrix}\right\}$$

We can tell the first two vectors $\langle 1, -1, 1, 2 \rangle$ and $\langle 2, 0, 4, 2 \rangle$ are independent since they aren't multiples of each other, but it's harder to see whether the remaining two vectors $\langle -1, 3, 3, 2 \rangle$ and $\langle 3, -4, 1, 4 \rangle$ are independent because we also have to make sure they can't be written as sums of multiples of other vectors in the set.

Reduction

To make it easier for us to tell whether these vectors are independent, we can **reduce** the set of vectors to a simpler set with the same span, by adding multiples of vectors from each other.

To begin the process of reduction, we can add multiples of the first vector to the other vectors so that we eliminate the first component from each of the other vectors.

- The second vector has a first component of 2, so we can eliminate it by adding -2 times the first vector.
- The third vector has a first component of -1, so we can eliminate it by adding 1 times the first vector.
- The fourth vector has a first component of 3, so we can eliminate it by adding -3 times the first vector.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 2, 0, 4, 2 \rangle - 2 \langle 1, -1, 1, 2 \rangle \\ \langle -1, 3, 3, 2 \rangle + 1 \langle 1, -1, 1, 2 \rangle \\ \langle 3, -4, 1, 4 \rangle - 3 \langle 1, -1, 1, 2 \rangle \end{matrix} \right\}$$

The resulting set of vectors is shown below.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 2, 2, -2 \rangle \\ \langle 0, 2, 4, 4 \rangle \\ \langle 0, -1, -2, -2 \rangle \end{matrix} \right\}$$

Since we only added multiples of vectors, we haven't changed the span at all. But now all of the first components are zero EXCEPT for the first component in the first vector, so we can see that the first vector cannot be written as a sum of other vectors in the set. All the other vectors have zero in their first component, so every time we add multiples of them, the result will still have zero in the first component.

To check whether the second vector is independent, we can add multiples of the second vector to the remaining vectors to eliminate their second components.

But to make this easier, we can start by rescaling (i.e. multiplying) the second vector to have a second component of 1. Its second component is currently 2, so to convert 2 to 1, we need to multiply by $\frac{1}{2}$.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \frac{1}{2} \langle 0, 2, 2, -2 \rangle \\ \langle 0, 2, 4, 4 \rangle \\ \langle 0, -1, -2, -2 \rangle \end{matrix} \right\}$$

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \langle 0, 2, 4, 4 \rangle \\ \langle 0, -1, -2, -2 \rangle \end{matrix} \right\}$$

Now, we can add multiples of the second vector to the third and fourth vectors so that we eliminate their second components.

- The third vector has a second component of 2, so we can eliminate it by adding -2 times the second vector.
- The fourth vector has a second component of -1, so we can eliminate it by adding 1 times the second vector.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \langle 0, 2, 4, 4 \rangle - 2 \langle 0, 1, 1, -1 \rangle \\ \langle 0, -1, -2, -2 \rangle + 1 \langle 0, 1, 1, -1 \rangle \end{matrix} \right\}$$

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \langle 0, 0, 2, 6 \rangle \\ \langle 0, 0, -1, -3 \rangle \end{matrix} \right\}$$

Clearly, the second vector cannot be written as a sum of multiples including the first vector, since including the first vector would cause the first component to become nonzero. And the second vector cannot be written as a sum of multiples of the third and fourth vectors, either, because no combination of them can produce a nonzero second component. So the second vector must be independent.

To determine whether the third and fourth vectors are independent, we can repeat the usual process once more. First, we'll rescale the third vector by $\frac{1}{2}$ so that its first component is 1.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \frac{1}{2} \langle 0, 0, 2, 6 \rangle \\ \langle 0, 0, -1, -3 \rangle \end{matrix} \right\}$$

The result is shown below.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \langle 0, 0, 1, 3 \rangle \\ \langle 0, 0, -1, -3 \rangle \end{matrix} \right\}$$

The fourth vector has a third component of -1, so we can eliminate it by adding 1 times the third vector.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \langle 0, 0, 1, 3 \rangle \\ \langle 0, 0, -1, -3 \rangle + 1 \langle 0, 0, 1, 3 \rangle \end{matrix} \right\}$$

Our final result is shown below.

$$\left\{ \begin{matrix} \langle 1, -1, 1, 2 \rangle \\ \langle 0, 1, 1, -1 \rangle \\ \langle 0, 0, 1, 3 \rangle \\ \langle 0, 0, 0, 0 \rangle \end{matrix} \right\}$$

We see that the first three vectors are independent, whereas the fourth vector is dependent since it is a multiple of every vector in the set (you can multiply any other vector by 0 to obtain the fourth vector). As a result, our set spans a 3-dimensional subspace of 4-dimensional space.

Exercises

Tell the dimension of the span of the set of vectors.

1) $\{\langle 1,1\rangle, \langle 2,0\rangle\}$

2) $\{\langle 1,1\rangle, \langle 2,2\rangle\}$

3) $\{\langle 1,1\rangle, \langle 2,0\rangle, \langle 3,2\rangle\}$

4) $\{\langle 1,2,3\rangle, \langle 3,2,1\rangle\}$

5)
$$\left\{\begin{matrix} \langle 1,2,3\rangle \\ \langle 3,2,1\rangle \\ \langle 1,0,0\rangle \end{matrix}\right\}$$

6)
$$\left\{\begin{matrix} \langle 1,2,3\rangle \\ \langle 3,2,1\rangle \\ \langle 1,1,1\rangle \end{matrix}\right\}$$

7)
$$\left\{\begin{matrix} \langle 1,1,1\rangle \\ \langle -1,-1,-1\rangle \\ \langle 2,2,2\rangle \end{matrix}\right\}$$

8)
$$\left\{\begin{matrix} \langle 1,0,-1\rangle \\ \langle 1,0,1\rangle \\ \langle 0,1,1\rangle \end{matrix}\right\}$$

9)
$$\left\{\begin{matrix} \langle 1,1,0\rangle \\ \langle 0,1,1\rangle \\ \langle 1,0,1\rangle \\ \langle 0,1,0\rangle \end{matrix}\right\}$$

10)
$$\left\{\begin{matrix} \langle 1,2,1,0\rangle \\ \langle 4,3,3,1\rangle \\ \langle 3,4,3,3\rangle \\ \langle 4,0,0,0\rangle \end{matrix}\right\}$$

11)

$$\left\{ \begin{matrix} \langle 1,2,3,4,5 \rangle \\ \langle 1,2,1,2,1 \rangle \\ \langle -1,2,-1,2,-1 \rangle \\ \langle 1,0,1,0,1 \rangle \end{matrix} \right\}$$

12)

$$\left\{ \begin{matrix} \langle 0,1,2,3,4 \rangle \\ \langle 1,2,3,4,5 \rangle \\ \langle 2,3,4,5,6 \rangle \\ \langle 3,4,5,6,7 \rangle \\ \langle 4,5,6,7,8 \rangle \\ \langle 5,6,7,8,9 \rangle \end{matrix} \right\}$$

13)

$$\left\{ \begin{matrix} \langle 1,-1,2,-2,3 \rangle \\ \langle 1,1,0,0,0 \rangle \\ \langle 0,1,1,0,0 \rangle \\ \langle 0,0,1,1,0 \rangle \\ \langle 0,0,0,1,1 \rangle \\ \langle 0,1,0,1,0 \rangle \\ \langle 1,0,1,0,1 \rangle \end{matrix} \right\}$$

14)

$$\left\{ \begin{matrix} \langle 1,3,3,1,0 \rangle \\ \langle 1,4,6,4,1 \rangle \\ \langle 0,1,3,3,1 \rangle \\ \langle 1,2,1,2,1 \rangle \end{matrix} \right\}$$

1.5 Elimination as Vector Reduction

Recall that systems of linear equations can be solved through **elimination**, multiplying equations by constants and adding equations to each other to cancel variables.

For example, to solve the following linear system

$$\begin{aligned} x + 2y + 3z &= 2 \\ -x - y - 2z &= 1 \\ 2x + 4y + 7z &= 5 \end{aligned}$$

we can start by adding the first equation to the second equation and subtracting two times the first equation from the third equation.

$$\begin{gathered} x + 2y + 3z = 2 \\ (-x - y - 2z = 1) + (x + 2y + 3z = 2) \\ (2x + 4y + 7z = 5) - 2(x + 2y + 3z = 2) \end{gathered}$$

$$\begin{aligned} x + 2y + 3z &= 2 \\ (-x - y - 2z) + (x + 2y + 3z) &= 1 + 2 \\ (2x + 4y + 7z) - 2(x + 2y + 3z) &= 5 - 2(2) \end{aligned}$$

$$\begin{aligned} x + 2y + 3z &= 2 \\ y + z &= 3 \\ z &= 1 \end{aligned}$$

Then, starting with $z = 1$, we can back-substitute to solve for each of the variables:

$z = 1$	$y + z = 3$	$x + 2y + 3z = 2$
	$y + 1 = 3$	$x + 2(2) + 3(1) = 2$
	$y = 2$	$x + 7 = 2$
		$x = -5$

We reach the final solution $x = -5$, $y = 2$, and $z = 1$.

Interpreting Elimination as Vector Reduction

In light of the previous chapter, elimination can also be interpreted as vector reduction.

First, we can interpret the linear system itself as a set of vectors, consisting of the coefficients and constants.

$$\begin{Bmatrix} x & +2y & +3z & = 2 \\ -x & -y & -2z & = 1 \\ 2x & +4y & +7z & = 5 \end{Bmatrix} \longrightarrow \begin{Bmatrix} \langle 1, 2, 3, 2 \rangle \\ \langle -1, -1, -2, 1 \rangle \\ \langle 2, 4, 7, 5 \rangle \end{Bmatrix}$$

Then, to reduce the set of vectors, we can add the first equation to the second equation, and subtract two of the first equation from the third equation.

$$\left\{ \begin{matrix} \langle 1,2,3,2 \rangle \\ \langle -1,-1,-2,1 \rangle + \langle 1,2,3,2 \rangle \\ \langle 2,4,7,5 \rangle - 2 \langle 1,2,3,2 \rangle \end{matrix} \right\}$$

$$\left\{ \begin{matrix} \langle 1,2,3,2 \rangle \\ \langle 0,1,1,3 \rangle \\ \langle 0,0,1,1 \rangle \end{matrix} \right\}$$

Then, we can convert the set of vectors back into equations that can be solved in the same way via back-substitution.

$$\left\{ \begin{matrix} \langle 1,2,3,2 \rangle \\ \langle 0,1,1,3 \rangle \\ \langle 0,0,1,1 \rangle \end{matrix} \right\} \longrightarrow \left\{ \begin{matrix} x & +2y & +3z & = 2 \\ & y & +z & = 3 \\ & & z & = 1 \end{matrix} \right\}$$

The big geometric insight here is that **the space of linear equations is actually a vector space**.

This occurs because we're allowed to add/subtract multiples of the equations. The particular linear equations in our system span a subspace of this vector space, and reducing the vectors allows us to simplify the system while maintaining the original span.

$$
\begin{array}{ccccc}
\begin{aligned} x + 2y + 3z &= 2 \\ -x - y - 2z &= 1 \\ 2x + 4y + 7z &= 5 \end{aligned} & & \longleftrightarrow & & \left\{ \begin{matrix} \langle 1, 2, 3, 2 \rangle \\ \langle -1, -1, -2, 1 \rangle \\ \langle 2, 4, 7, 5 \rangle \end{matrix} \right\} \\
\downarrow & \searrow & & \swarrow & \downarrow \\
\text{elimination} & & \boxed{\begin{matrix}\text{spanning} \\ \text{subspace}\end{matrix}} & & \text{reduction} \\
\downarrow & \nearrow & & \nwarrow & \downarrow \\
\begin{aligned} x + 2y + 3z &= 2 \\ y + z &= 3 \\ z &= 1 \end{aligned} & & \longleftrightarrow & & \left\{ \begin{matrix} \langle 1, 2, 3, 2 \rangle \\ \langle 0, 1, 1, 3 \rangle \\ \langle 0, 0, 1, 1 \rangle \end{matrix} \right\}
\end{array}
$$

Thinking of linear equations in terms of vectors can sometimes yield additional insight. For example, notice that for a system of linear equations to have a single solution, the vectors must be reducible to the following form:

$$
\left\{ \begin{matrix} \langle 1, \square, \square, \cdots, \square, \square \rangle \\ \langle 0, 1, \square, \cdots, \square, \square \rangle \\ \vdots \\ \langle 0, 0, 0, \cdots, 1, \square \rangle \end{matrix} \right\}
$$

In other words, the vectors must span all components except the last. For a system of linear equations in n variables, the vectors consist of $n+1$ components: the first n components correspond to variable coefficients, and the last component corresponds to the constant.

As a result, **for a system of linear equations in n variables to have a single solution, at least n equations are required.**

Exercises

Solve the following systems.

1)
$$\begin{aligned} 2x + 3y &= 1 \\ x + y &= 1 \end{aligned}$$

2)
$$\begin{aligned} x - 2y &= -1 \\ 2x - y &= 7 \end{aligned}$$

3)
$$\begin{aligned} x + 2y - z &= 2 \\ 2x + y + z &= 7 \\ x + y - 2z &= -3 \end{aligned}$$

4)
$$\begin{aligned} 3x - 4y + z &= 8 \\ 3x + y &= 5 \\ y + z &= -3 \end{aligned}$$

5)
$$\begin{aligned} x + 2z &= 3 \\ x + y &= 6 \\ y - z &= 4 \end{aligned}$$

6)
$$\begin{aligned} -x + 5y &= -7 \\ 3x + 3z &= 15 \\ x + y + z &= 4 \end{aligned}$$

7)
$$\begin{aligned} w + x - y - z &= 4 \\ 2w - x &= 1 \\ x + y + z &= 4 \\ x + 2z &= 5 \end{aligned}$$

8)
$$\begin{aligned} 3w + 2x + y &= -5 \\ 3x + 2y + z &= -2 \\ w + 2y + 3z &= 1 \\ w - x - y - z &= -1 \end{aligned}$$

Part 2
Volume

2.1 N-Dimensional Volume Formula

N-dimensional volume generalizes the idea of the space occupied by an object:

- 1-dimensional volume refers to the space occupied by a 1-dimensional object, such as the length of a line segment.

- 2-dimensional volume refers to the space occupied by a 2-dimensional object, such as the area of a square.

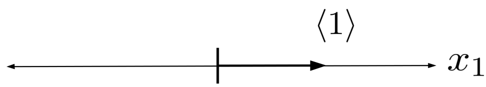
- 3-dimensional volume is what we normally mean by the word "volume" -- the amount of space occupied by a 3-dimensional object, such as the volume of a cube.

Continuing this pattern, we can infer that 4-dimensional volume refers to the space occupied by a 4-dimensional object. It's harder to come up with an example, though, since it's difficult to visualize shapes in 4 and higher dimensions.

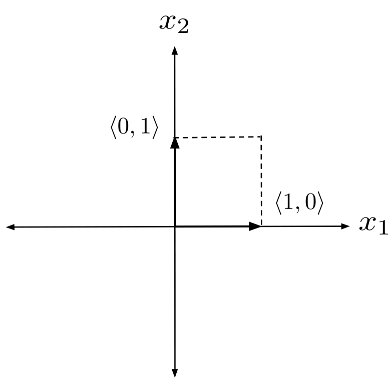
Volume Enclosed by N-Dimensional Vectors

However, it becomes easier if we think about N-dimensional volume as being enclosed by N-dimensional vectors.

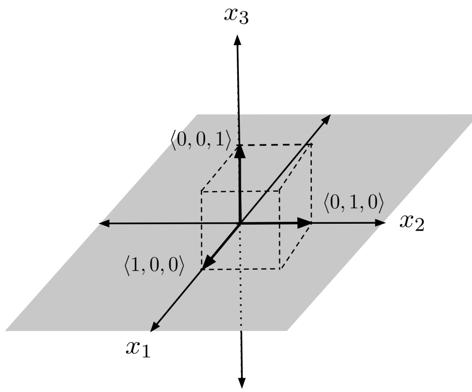
The length of a unit line segment can be interpreted as the space enclosed by the 1-dimensional unit vector $\langle 1 \rangle$.



The area of a unit square can be interpreted as the space enclosed by the two 2-dimensional vectors $\langle 1, 0 \rangle$ and $\langle 0, 1 \rangle$.



The volume of a unit cube can be interpreted as the space enclosed by the three 3-dimensional unit vectors: $\langle 1, 0, 0 \rangle$, $\langle 0, 1, 0 \rangle$, and $\langle 0, 0, 1 \rangle$.



Continuing this pattern, the volume of a unit 4-dimensional cube can be interpreted as the space enclosed by the four 4-dimensional unit vectors: $\langle 1,0,0,0\rangle$, $\langle 0,1,0,0\rangle$, $\langle 0,0,1,0\rangle$, and $\langle 0,0,0,1\rangle$.

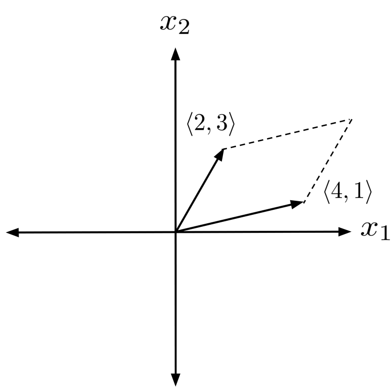
This is harder to draw, but it gives us a general way to think about volume in N dimensions. The volume of a unit N-dimensional cube can be interpreted as the space enclosed by the N N-dimensional unit vectors:

$$\begin{gathered}\langle 1,0,0,\cdots,0,0\rangle\\ \langle 0,1,0,\cdots,0,0\rangle\\ \vdots\\ \langle 0,0,0,\cdots,1,0\rangle\\ \langle 0,0,0,\cdots,0,1\rangle\end{gathered}$$

Volume of a Parallelogram

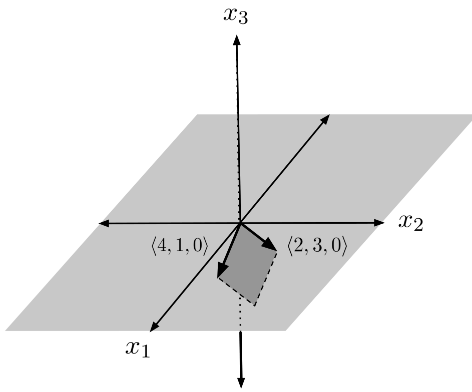
Given an object whose sides are perpendicular unit vectors, it's easy to see that the volume of the object is 1, since the distance of the object in each perpendicular dimension is 1.

But it's difficult when the vectors are not perpendicular. For example, how would we compute the area of the parallelogram enclosed by the vectors $\langle 4, 1 \rangle$ and $\langle 2, 3 \rangle$?



Remember, the area of a parallelogram enclosed by two 3-dimensional vectors is just the magnitude of their cross product.

Although the vectors $\langle 4, 1 \rangle$ and $\langle 2, 3 \rangle$ are 2-dimensional, we can interpret them as the 3-dimensional vectors $\langle 4, 1, 0 \rangle$ and $\langle 2, 3, 0 \rangle$ in the $x_1 x_2$ plane.



Taking the magnitude of the cross product, we find that the area of the parallelogram is 10.

$$|\langle 4, 1, 0 \rangle \times \langle 2, 3, 0 \rangle| = |\langle 0, 0, 10 \rangle| = 10$$

Volume of a Parallelepiped

We can also use the cross product as a starting point to find the volume of a parallelepiped enclosed by three 3-dimensional vectors x, y, and z.

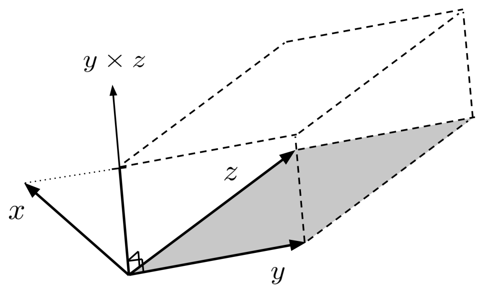
The cross product of two of the vectors, say $y \times z$, gives a vector whose magnitude is the area of a face of the parallelepiped, and which points in the direction perpendicular to the face.

The height of the parallelepiped, then, is the length of the remaining vector x in the direction of $y \times z$.

Thus, the volume can be obtained using the dot product:

$$x \cdot (y \times z)$$

This result is known as the **triple product**.



That being said, the triple product may come out negative, depending on the order of the vectors in the cross product. So, to compute the volume, we have to take the absolute value of the final result. A simple example for the unit cube is shown below.

$$\begin{aligned}
&|\langle 1,0,0\rangle \cdot (\langle 0,0,1\rangle \times \langle 0,1,0\rangle)| \\
&= |\langle 1,0,0\rangle \cdot \langle -1,0,0\rangle| \\
&= |-1| \\
&= 1
\end{aligned}$$

N-Dimensional Volume Formula

Now that we know general methods to compute volume enclosed by 2-dimensional and 3-dimensional vectors, how do we extend this to 4-dimensional vectors?

If we rewrite the 3-dimensional volume formula in terms of the 2-dimensional volume formula, a pattern jumps out at us.

To start, let's write down the 2-dimensional volume formula for two vectors $x_1 = \langle x_{11}, x_{12} \rangle$ and $x_2 = \langle x_{21}, x_{22} \rangle$.

$$\begin{aligned} V \begin{pmatrix} \langle x_{11}, x_{12} \rangle \\ \langle x_{21}, x_{22} \rangle \end{pmatrix} &= |\langle x_{11}, x_{12}, 0 \rangle \times \langle x_{21}, x_{22}, 0 \rangle| \\ &= |\langle 0, 0, x_{11}x_{22} - x_{12}x_{21} \rangle| \\ &= |x_{11}x_{22} - x_{12}x_{21}| \end{aligned}$$

However, in order to make the pattern clear, we will leave off the absolute value sign, thereby permitting "signed" volume.

$$V \begin{pmatrix} \langle x_{11}, x_{12} \rangle \\ \langle x_{21}, x_{22} \rangle \end{pmatrix} = x_{11}x_{22} - x_{12}x_{21}$$

In 2 dimensions, the volume is traced out from the first vector $x_1 = \langle x_{11}, x_{12} \rangle$ to the second vector $x_2 = \langle x_{21}, x_{22} \rangle$, and the sign of the volume just tells us whether the tracing occurs counterclockwise (positive) or clockwise (negative). Intuitively, this convention matches that which is used for tracing out positive or negative angles in the unit circle.

Now, let's write down the 3-dimensional volume formula for 3 vectors $x_1 = \langle x_{11}, x_{12}, x_{13} \rangle$, $x_2 = \langle x_{21}, x_{22}, x_{23} \rangle$, and $x_3 = \langle x_{31}, x_{32}, x_{33} \rangle$, again leaving off the absolute value sign and thereby permitting "signed" volume. (The meaning of "signed" volume in 3 dimensions will be addressed later.)

$$\begin{aligned} V \begin{pmatrix} \langle x_{11}, x_{12}, x_{13} \rangle \\ \langle x_{21}, x_{22}, x_{23} \rangle \\ \langle x_{31}, x_{32}, x_{33} \rangle \end{pmatrix} &= \langle x_{11}, x_{12}, x_{13} \rangle \cdot (\langle x_{21}, x_{22}, x_{23} \rangle \times \langle x_{31}, x_{32}, x_{33} \rangle) \\ &= \langle x_{11}, x_{12}, x_{13} \rangle \cdot (\langle x_{22}x_{33} - x_{23}x_{32}, x_{23}x_{31} - x_{21}x_{33}, x_{21}x_{32} - x_{22}x_{31} \rangle) \\ &= \langle x_{11}, x_{12}, x_{13} \rangle \cdot \left\langle V \begin{pmatrix} \langle x_{22}, x_{23} \rangle \\ \langle x_{32}, x_{33} \rangle \end{pmatrix}, V \begin{pmatrix} \langle x_{23}, x_{21} \rangle \\ \langle x_{33}, x_{31} \rangle \end{pmatrix}, V \begin{pmatrix} \langle x_{21}, x_{22} \rangle \\ \langle x_{31}, x_{32} \rangle \end{pmatrix} \right\rangle \end{aligned}$$

To ease notation, we define another volume function V_k that also computes volumes of vectors, but first re-indexes the vectors by removing the k^th component and moving entries after the k^th component to the beginning of the vector. This way, the 3-dimensional volume formula can be simplified to

$$V \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \cdot \left\langle V_1 \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}, V_2 \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}, V_3 \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} \right\rangle .$$

Using this form, we can guess at an N-dimensional volume formula:

$$V \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} = x_1 \cdot \left\langle V_1 \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, V_2 \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, V_N \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle$$

Sanity Checks

Let’s test out this formula on a simple case: the 4-dimensional unit cube, which is enclosed by the vectors $x_1 = \langle 1,0,0,0 \rangle$, $x_2 = \langle 0,1,0,0 \rangle$, $x_3 = \langle 0,0,1,0 \rangle$, and $x_4 = \langle 0,0,0,1 \rangle$.

$$V \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_1 \cdot \left\langle V_1 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}, V_2 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}, V_3 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}, V_4 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix} \right\rangle$$

$$= \langle 1,0,0,0 \rangle \cdot \left\langle V_1 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}, V_2 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}, V_3 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}, V_4 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix} \right\rangle$$

$$= V_1 \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$= V_1 \begin{pmatrix} \langle 0,1,0,0 \rangle \\ \langle 0,0,1,0 \rangle \\ \langle 0,0,0,1 \rangle \end{pmatrix}$$

$$= V \begin{pmatrix} \langle 1,0,0 \rangle \\ \langle 0,1,0 \rangle \\ \langle 0,0,1 \rangle \end{pmatrix}$$

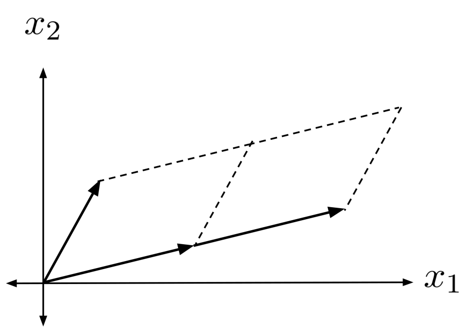
$$= \langle 1,0,0 \rangle \cdot (\langle 0,1,0 \rangle \times \langle 0,0,1 \rangle)$$

$$= \langle 1,0,0 \rangle \cdot \langle 1,0,0 \rangle$$

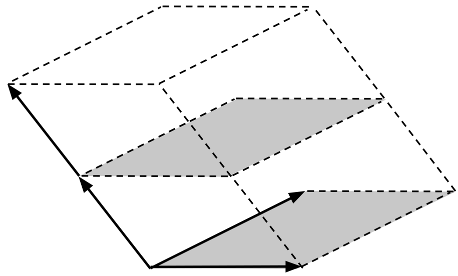
$$= 1$$

Now we're reaching the point where it's hard to actually "see" what's happening. But the math shows us a pattern, the pattern matches our intuition on a simple case, and given that 3-dimensional volume is the sum of multiples of 2-dimensional volumes, it seems plausible that N-dimensional volume could be the sum of multiples of (N-1)-dimensional volumes.

Moreover, the volume formula matches our intuition when we rescale a vector. Intuitively, rescaling a vector should have the effect of rescaling the volume: for example, if we double the length of one of the sides of a parallelogram, the area should double.



Likewise, if we double the length of one of the sides of a parallelepiped, then the volume should double.



More generally, if we rescale a single vector by a factor r, then the volume should also be rescaled by a factor r. Indeed, this is the case with our formula.

$$\begin{aligned}
V\begin{pmatrix} rx_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} &= rx_1 \cdot \left\langle V_1\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, V_2\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, V_N\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle \\
&= r\left(x_1 \cdot \left\langle V_1\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, V_2\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, V_N\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle \right) \\
&= rV\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}
\end{aligned}$$

Of course, we have only shown that rescaling the first vector rescales the volume. However, we can use this fact to show that rescaling the second vector also rescales the volume.

$$V\begin{pmatrix} x_1 \\ rx_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} = x_1 \cdot \left\langle V_1\begin{pmatrix} rx_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, V_2\begin{pmatrix} rx_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, V_N\begin{pmatrix} rx_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle$$

$$= x_1 \cdot \left\langle rV_1\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, rV_2\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, rV_N\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle$$

$$= x_1 \cdot r\left\langle V_1\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, V_2\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, V_N\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle$$

$$= r\left(x_1 \cdot \left\langle V_1\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, V_2\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}, \ldots, V_N\begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \right\rangle \right)$$

$$= rV\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix}$$

We could keep on going, using similar arguments to show that rescaling the 3rd, 4th, so on, and Nth vector have the same effect of rescaling the volume.

Final Remarks

Unfortunately, this volume formula is unintuitive and unwieldy for volume computations in high-dimensional space.

Soon, though, we will introduce a more intuitive concept called **shearing**, which will lead us to an easier, more intuitive process of computing high-dimensional volumes.

The volume formula presented in this chapter is still noteworthy, though, because although shearing will provide us with a *process* for computing volume, it won't give us a *formula* for computing volume.

Exercises

Compute the N-dimensional unsigned volume $|V(x_1, x_2, \ldots, x_n)|$ enclosed by the vectors $x_1, x_2, \ldots, x_n$.

1)

$$x_1 = \langle 1, 0 \rangle$$
$$x_2 = \langle 0, 2 \rangle$$

2)

$$x_1 = \langle 1, 1 \rangle$$
$$x_2 = \langle 1, 2 \rangle$$

3)

$$x_1 = \langle 3, 4 \rangle$$
$$x_2 = \langle -1, 2 \rangle$$

4)

$$x_1 = \langle 7, -3 \rangle$$
$$x_2 = \langle 2, -6 \rangle$$

5)

$$x_1 = \langle 1, 2, 3 \rangle$$
$$x_2 = \langle 3, 2, 1 \rangle$$
$$x_3 = \langle 1, 3, 2 \rangle$$

6)

$$x_1 = \langle 2, 0, -1 \rangle$$
$$x_2 = \langle 3, 4, 2 \rangle$$
$$x_3 = \langle 0, -1, 2 \rangle$$

7)

$$x_1 = \langle 1, 5, 7 \rangle$$
$$x_2 = \langle 3, 2, 1 \rangle$$
$$x_3 = \langle 0, 1, 2 \rangle$$

8)

$$x_1 = \langle 4, 0, -5 \rangle$$
$$x_2 = \langle -9, 1, 7 \rangle$$
$$x_3 = \langle 2, -5, 4 \rangle$$

9)

$$x_1 = \langle 1, 3, 2, 0 \rangle$$
$$x_2 = \langle 4, 3, 0, 1 \rangle$$
$$x_3 = \langle 0, 0, 2, 1 \rangle$$
$$x_4 = \langle 3, 0, 4, 2 \rangle$$

10)

$$x_1 = \langle 0, -2, 3, 1 \rangle$$
$$x_2 = \langle -1, -1, 2, 2 \rangle$$
$$x_3 = \langle 1, 0, 1, 0 \rangle$$
$$x_4 = \langle 0, 1, 1, 1 \rangle$$

2.2 Volume as the Determinant of a Square Linear System

We have seen that the space of linear equations is actually a vector space, and that the linear equations in any particular system span a subspace of this vector space.

Linear Systems as Vector Equations

However, there is also another way to interpret linear systems in terms of vectors: **a linear system can be interpreted as a single vector equation stating that some multiples of particular vectors add up to another particular vector.**

For example, we can write the system below as a vector equation by interpreting each side of the equation as a vector:

$$\begin{aligned} x + 2y + 3z &= 2 \\ -x - y - 2z &= 1 \\ 2x + 4y + 7z &= 5 \end{aligned}$$

$$\begin{pmatrix} x + 2y + 3z \\ -x - y - 2z \\ 2x + 4y + 7z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} x \\ -x \\ 2x \end{pmatrix} + \begin{pmatrix} 2y \\ -y \\ 4y \end{pmatrix} + \begin{pmatrix} 3z \\ -2z \\ 7z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} x + \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} y + \begin{pmatrix} 3 \\ -2 \\ 7 \end{pmatrix} z = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$$

This equation states that some multiples x, y, and z of the coefficient vectors $\langle 1, -1, 2\rangle$, $\langle 2, -1, 4\rangle$, and $\langle 3, -2, 7\rangle$ sum to the constant vector $\langle 2, 1, 5\rangle$.

You might recall that we solved this system earlier using reduction, and we found that the solution was $x = -5$, $y = 2$, and $z = 1$. Now, we see that these are simply the multiples of the coefficient vectors that sum to the constant vector.

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} (-5) + \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} (2) + \begin{pmatrix} 3 \\ -2 \\ 7 \end{pmatrix} (1) = \begin{pmatrix} -5 \\ 5 \\ -10 \end{pmatrix} + \begin{pmatrix} 4 \\ -2 \\ 8 \end{pmatrix} + \begin{pmatrix} 3 \\ -2 \\ 7 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$$

For the linear system to have a solution, there must be some multiples of the coefficient vectors that add to the constant vector. In other words, for the linear system to have a solution, the constant vector must be in the span of the coefficient vectors.

Thinking about linear systems in terms of coefficient vectors can provide useful intuition. For example, we can tell that the linear system below has a solution because its coefficient vectors span the full 2-dimensional plane.

$$\begin{array}{ccccccc} x & + & y & + & 2z & = & 5 \\ x & + & 2y & + & 2z & = & -3 \end{array} \quad \rightarrow \quad x\begin{pmatrix}1\\1\end{pmatrix} + y\begin{pmatrix}1\\2\end{pmatrix} + z\begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}5\\-3\end{pmatrix}$$

Moreover, since there are 3 coefficient vectors spanning a 2-dimensional plane, there must be a dependent vector, so there must be infinitely many solutions.

For example, $\langle 2, 2 \rangle$ is a multiple of $\langle 1, 1 \rangle$, so in any solution we can increase z by some amount and decrease x by twice that amount to yield another solution. Thus since $x = 13, y = -8, z = 0$ is a solution, so is $x = 11, y = -8, z = 1$, and $x = 9, y = -8, z = 2$, and so on.

The Determinant

When there are exactly N coefficient vectors that form an N-dimensional parallelepiped, we can also extend this intuition to relate to the volume of the coefficient vectors. Such linear systems are called **square** linear systems because they consist of N rows of equations and N columns of variables. In a square system, the volume of the coefficient vectors is called the **determinant**, because it *determines* much about the solutions of the system.

When the determinant is nonzero, there is exactly one solution. When the determinant is nonzero, the N coefficient vectors form a parallelepiped that extends some nonzero amount in all N dimensions, and consequently the coefficient vectors span the full N-dimensional space, guaranteeing a solution.

Moreover, the solution must be unique. For N vectors to span N dimensions, the vectors must be independent -- meaning that no vector can be written in terms of the others, and thus guaranteeing that there is only one solution.

For example, the following linear system has a nonzero determinant, and a single solution $x = 0, y = 1, z = 1$.

$$\begin{array}{ccccccc} x & + & y & + & 2z & = & 3 \\ x & + & y & + & z & = & 2 \\ x & + & 2y & + & 3z & = & 5 \end{array} \quad \rightarrow \quad x \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + z \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$$

$$\text{determinant} = \det \begin{pmatrix} \langle 1,1,1 \rangle \\ \langle 1,1,2 \rangle \\ \langle 2,1,3 \rangle \end{pmatrix} = V \begin{pmatrix} \langle 1,1,1 \rangle \\ \langle 1,1,2 \rangle \\ \langle 2,1,3 \rangle \end{pmatrix} = 1$$

On the other hand, **when the determinant is zero, there are either no solutions or infinitely many solutions.** When the determinant is zero, the coefficient vectors form a parallelepiped that is flat in at least one dimension, and consequently the coefficient vectors span only a smaller subspace of N-dimensional space, which may or may not contain the constant vector.

If the subspace *does not* contain the constant vector, then there is no solution.

If the subspace *does* contain the constant vector, then there is a solution, and moreover, since a set of N vectors spanning fewer than N dimensions must contain at least one dependent vector, there must be infinitely many solutions.

For example, the following linear system has a zero determinant and no solutions.

$$\begin{array}{ccccccc} x & + & y & + & 2z & = & 3 \\ x & + & y & + & 2z & = & 4 \\ x & + & 2y & + & 3z & = & 5 \end{array} \quad \rightarrow \quad x\begin{pmatrix}1\\1\\1\end{pmatrix} + y\begin{pmatrix}1\\1\\2\end{pmatrix} + z\begin{pmatrix}2\\2\\3\end{pmatrix} = \begin{pmatrix}3\\4\\5\end{pmatrix}$$

$$\text{determinant} = \det\begin{pmatrix}\langle 1,1,1\rangle\\ \langle 1,1,2\rangle\\ \langle 2,2,3\rangle\end{pmatrix} = V\begin{pmatrix}\langle 1,1,1\rangle\\ \langle 1,1,2\rangle\\ \langle 2,2,3\rangle\end{pmatrix} = 0$$

On the other hand, the following linear system has a zero determinant, and infinitely many solutions.

$$\begin{array}{ccccccc} x & + & y & + & 2z & = & 3 \\ x & + & y & + & 2z & = & 3 \\ x & + & 2y & + & 3z & = & 5 \end{array} \quad \rightarrow \quad x\begin{pmatrix}1\\1\\1\end{pmatrix} + y\begin{pmatrix}1\\1\\2\end{pmatrix} + z\begin{pmatrix}2\\2\\3\end{pmatrix} = \begin{pmatrix}3\\3\\5\end{pmatrix}$$

$$\text{determinant} = \det\begin{pmatrix}\langle 1,1,1\rangle\\ \langle 1,1,2\rangle\\ \langle 2,2,3\rangle\end{pmatrix} = V\begin{pmatrix}\langle 1,1,1\rangle\\ \langle 1,1,2\rangle\\ \langle 2,2,3\rangle\end{pmatrix} = 0$$

One solution, for example, is $x = 0, y = 1, z = 1$. But since $\langle 2,2,3\rangle$ is the sum of $\langle 1,1,1\rangle$ and $\langle 1,1,2\rangle$, we can obtain another solution by increasing x and y by some amount, and decreasing z by that same amount. For example, another solution is $x = 1, y = 2, z = 0$, and yet another solution is $x = 2, y = 3, z = -1$.

Later, we will see that the determinant plays a fundamental role in understanding transformations of vectors, which are called

matrices. For now, though, we will just get in the habit of writing volume using the determinant operator $\det$ in place of V.

Exercises

Determine whether the system has A) exactly one solution, or B) no solutions or infinitely many solutions.

1)

$$\begin{aligned} x + y &= 1 \\ 2x + y &= 3 \end{aligned}$$

2)

$$\begin{aligned} 2x + 4y &= 0 \\ -x - 2y &= 3 \end{aligned}$$

3)

$$\begin{aligned} x - y + z &= 10 \\ 2x + y + z &= 12 \\ 4x - y + 3z &= 13 \end{aligned}$$

4)

$$\begin{aligned} x + y + z &= 27 \\ x + 2y &= 31 \\ y + 2z &= 42 \end{aligned}$$

5)

$$\begin{aligned} x - y &= 53 \\ x + y + z &= 11 \\ x + z &= 16 \end{aligned}$$

6)

$$\begin{aligned} 7x + 3y - 2z &= 11 \\ x - y - z &= 12 \\ 8x - 2y - 3z &= -7 \end{aligned}$$

7)

$$\begin{aligned} 2x + 2y - z &= -27 \\ x + y + z &= 11 \\ 7x + 5y &= 83 \end{aligned}$$

8)

$$\begin{aligned} x + y - 3z &= 12 \\ x - y + 3z &= -12 \\ y - 3z &= 4 \end{aligned}$$

9)

$$\begin{aligned} w + x + y + z &= 42 \\ w - x - y - z &= 27 \\ x + z &= -11 \\ w + 2x + z &= -14 \end{aligned}$$

10)

$$\begin{aligned} w + x + y - z &= 12 \\ 3w + 2x + y &= -7 \\ x + 2y + 3z &= 11 \\ w + 2x + 3y + 2z &= 14 \end{aligned}$$

2.3 Shearing, Cramer’s Rule, and Volume by Reduction

Not only can a nonzero determinant tell us that a linear system has exactly one solution -- the nonzero determinant can also help us quickly find that solution through a process known as **Cramer’s rule.**

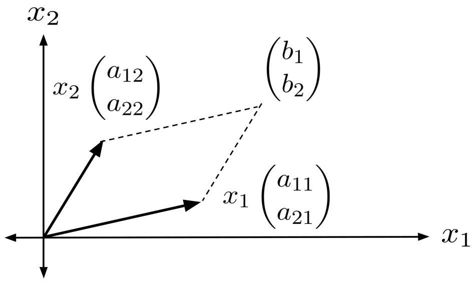
Shearing

The key bit of intuition surrounding Cramer’s rule is the idea that moving one of the sides of a parallelepiped in a parallel direction does not change the volume of the parallelepiped. This kind of transformation is known as **shearing**, and the intuition can be most easily illustrated in 2 dimensions.

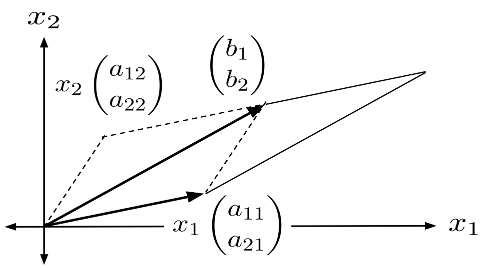
Suppose we have the following 2-dimensional linear system:

$$\begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \end{array} \quad \rightarrow \quad x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

The three vectors in this system -- two coefficient vectors and one constant vector -- can be represented visually as the vertices of a parallelogram.



Notice that shearing the parallelogram does not change its area.



Cramer’s Rule in Two Dimensions

Equating the volumes of the original parallelogram and the sheared parallelogram, we have

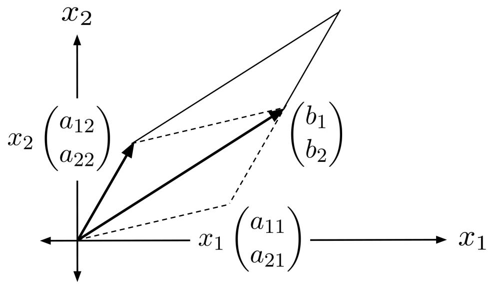
$$\det\left(x_1\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, x_2\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) = \det\left(x_1\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right).$$

Using the fact that scaling a vector results in the volume being scaled by the same amount, we can simplify and solve for x_2.

$$x_1 x_2 \det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) = x_1 \det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right)$$

$$x_2 = \frac{\det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\right)}{\det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right)}$$

This is the solution for x_2 in the original system! We can use the same method to solve for x_1, too.



$$\det\left(x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) = \det\left(\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right)$$

$$x_1 x_2 \det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) = x_2 \det\left(\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right)$$

$$x_1 = \frac{\det\left(\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right)}{\det\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right)}$$

This method is known as **Cramer's rule**. We illustrate it below on a concrete example, which would otherwise be annoying to solve by reduction because its solutions are fractional.

$$\begin{matrix} 2x_1 & - & 3x_2 & = & 7 \\ 5x_1 & + & x_2 & = & 4 \end{matrix} \quad \rightarrow \quad x_1 \begin{pmatrix} 2 \\ 5 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}$$

Using Cramer's rule, however, the solutions are much easier to compute.

$$x_1 = \frac{\det\left(\begin{pmatrix} 7 \\ 4 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix}\right)}{\det\left(\begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix}\right)} = \frac{19}{17}$$

$$x_2 = \frac{\det\left(\begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 7 \\ 4 \end{pmatrix}\right)}{\det\left(\begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix}\right)} = \frac{-27}{17}$$

Cramer's Rule in N Dimensions

To generalize Cramer's rule to N dimensions, we first come up with a compact notation for writing systems of linear equations.

$$\begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \end{array}$$

$$\downarrow$$

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} x_2 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\downarrow$$

$$a_1 x_1 + a_2 x_2 = b$$

In this notation, we reduce the linear system to a single equation in terms of the vectors $a_1 = \langle a_{11}, a_{21} \rangle$, $a_2 = \langle a_{12}, a_{22} \rangle$, and $b = \langle b_1, b_2 \rangle$. The solutions from Cramer's rule can then be written as follows:

$$x_1 = \frac{\det(b, a_2)}{\det(a_1, a_2)} \qquad x_2 = \frac{\det(a_1, b)}{\det(a_1, a_2)}$$

The pattern is clear: x_i is given by a fraction whose denominator is the determinant of the coefficient vectors, and whose numerator is the same except that the ith coefficient vector a_i is replaced with the constant vector b.

For an N-dimensional square linear system, then, the solutions to

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b$$

are given by

$$x_i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det(a_1, a_2, \ldots, a_n)}.$$

Volume by Reduction

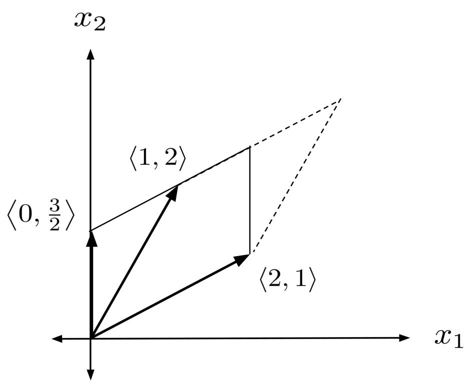
Now, let's take a step back and talk more about the elegance of shearing. We have seen that through Cramer's rule, shearing can be used to express the solution of a linear system using ratios of volumes. Now, we will see that **shearing can also be used to compute volumes themselves, without having to use the volume formula.**

In a set of vectors, shearing simply amounts to adding one vector to another vector. Consequently, reducing a set of vectors preserves the volume of the parallelepiped formed by those vectors, provided that we don't rescale any of the vectors themselves (otherwise, the volume would be rescaled as well).

The volume of a reduced set of vectors is much easier to compute: we can simply multiply the diagonal, because the diagonal entries are the parallelepiped's lengths in each dimension.

As a simple example, we can use shearing to compute the volume enclosed by the vectors $\langle 2,1 \rangle$ and $\langle 1,2 \rangle$.

$$\det \begin{pmatrix} \langle 2,1 \rangle \\ \langle 1,2 \rangle \end{pmatrix} = \det \begin{pmatrix} \langle 2,1 \rangle \\ \langle 1,2 \rangle - \frac{1}{2} \langle 2,1 \rangle \end{pmatrix} = \det \begin{pmatrix} \langle 2,1 \rangle \\ \langle 0, \frac{3}{2} \rangle \end{pmatrix} = 2 \cdot \frac{3}{2} = 3$$



When computing the volume of 4 or more vectors, it is much faster to use shearing instead of the volume formula. Below is an example of a 4-dimensional volume calculation using shearing.

$$\begin{aligned}
\det \begin{pmatrix} \langle 1,2,3,4 \rangle \\ \langle 1,3,2,4 \rangle \\ \langle 4,3,2,1 \rangle \\ \langle 3,2,1,4 \rangle \end{pmatrix} &= \det \begin{pmatrix} \langle 1,2,3,4 \rangle \\ \langle 1,3,2,4 \rangle - \langle 1,2,3,4 \rangle \\ \langle 4,3,2,1 \rangle - 4\langle 1,2,3,4 \rangle \\ \langle 3,2,1,4 \rangle - 3\langle 1,2,3,4 \rangle \end{pmatrix} \\
&= \det \begin{pmatrix} \langle 1,2,3,4 \rangle \\ \langle 0,1,-1,0 \rangle \\ \langle 0,-5,-10,-15 \rangle + 5\langle 0,1,-1,0 \rangle \\ \langle 0,-4,-8,-8 \rangle + 4\langle 0,1,-1,0 \rangle \end{pmatrix} \\
&= \det \begin{pmatrix} \langle 1,2,3,4 \rangle \\ \langle 0,1,-1,0 \rangle \\ \langle 0,0,-15,-15 \rangle \\ \langle 0,0,-12,-8 \rangle - \frac{12}{15}\langle 0,0,-15,-15 \rangle \end{pmatrix} \\
&= \det \begin{pmatrix} \langle 1,2,3,4 \rangle \\ \langle 0,1,-1,0 \rangle \\ \langle 0,0,-15,-15 \rangle \\ \langle 0,0,0,4 \rangle \end{pmatrix} \\
&= (1)(1)(-15)(4) \\
&= -60
\end{aligned}$$

Exercises

State whether the linear system has A) exactly one solution, or B) no solutions or infinitely many solutions. If there is A) exactly one solution, then use Cramer's rule to find it.

In your calculations of determinants in higher than 3 dimensions, be sure to use the technique of shearing -- it will save lots of time!

1)

$$\begin{aligned} x + 2y &= 4 \\ 2x - 3y &= 3 \end{aligned}$$

2)

$$\begin{aligned} 3x + y &= -1 \\ x - 2y &= 4 \end{aligned}$$

3)

$$\begin{aligned} 5x + 8y &= 2 \\ 3x - 7y &= 5 \end{aligned}$$

4)

$$\begin{aligned} 3x + 7y &= 4 \\ 10x - 6y &= -3 \end{aligned}$$

5)

$$\begin{aligned} x + 2y + 3z &= 8 \\ 2x + y - z &= 1 \\ x + y - z &= 2 \end{aligned}$$

6)

$$\begin{aligned} 3x - y + z &= 1 \\ 6x + y + z &= 2 \\ 3y - z &= 3 \end{aligned}$$

7)

$$\begin{aligned} x + 2y - z &= 1 \\ 3x - 3y + z &= 7 \\ -x + 7y + z &= -3 \end{aligned}$$

8)

$$\begin{aligned} 4x - 3y + 2z &= 1 \\ 2x + 3y + 4z &= 1 \\ x - y + z &= 1 \end{aligned}$$

9)

$$\begin{aligned} w + x + y + z &= 3 \\ w + y + z &= 2 \\ x + y + z &= -1 \\ 3w - 4z &= 7 \end{aligned}$$

10)

$$\begin{aligned} 4w + 4x - y + z &= 4 \\ 2w - 4x + y + z &= 3 \\ w + 4x - y &= 3 \\ w + 3x + y &= 1 \end{aligned}$$

11)

$$\begin{aligned} u + 3w + 4x - y + z &= 1 \\ u - 2w - 4x + y + z &= -1 \\ 4u + 4x - y &= 0 \\ w + 2x - y &= 1 \\ 3w - x &= 2 \end{aligned}$$

12)

$$\begin{aligned} 3u - 2w + x + y &= 0 \\ u - y - z &= 1 \\ w + x + y - 3z &= 3 \\ u + 2x + 3z &= -2 \\ 2u - x + y - z &= -2 \end{aligned}$$

2.4 Higher-Order Variation of Parameters

Until this point, we have been working exclusively with linear systems. However, solving linear systems can sometimes be a necessary component of solving nonlinear systems.

Second-Order Variation of Parameters

For example, recall the **variation of parameters** method for solving a second-order differential equation of the form

$$y'' + a_1(x)y' + a_0(x)y = f(x).$$

Variation of parameters proceeds by first guessing a solution of the form

$$y_f(x) = u_1(x)y_1(x) + u_2(x)y_2(x),$$

where y_1 and y_2 are the two zero solutions of the differential equation

$$y'' + a_1(x)y' + a_0(x)y = 0,$$

and $u_1(x)$ and $u_2(x)$ are some unknown multiplier functions that we solve for by setting up a system of equations.

To set up the first equation in our system, we force

$$y'_f(x) = u_1(x)y'_1(x) + u_2(x)y'_2(x)$$

and equate it to the true derivative of y_f:

$$\begin{aligned} y'_f &= u_1y'_1 + u_2y'_2 \\ (u_1y_1 + u_2y_2)' &= u_1y'_1 + u_2y'_2 \\ u'_1y_1 + u_1y'_1 + u'_2y_2 + u_2y'_2 &= u_1y'_1 + u_2y'_2 \\ u'_1y_1 + u'_2y_2 &= 0 \end{aligned}$$

The second equation comes from substituting our guess for y_f into the differential equation and simplifying, using the fact that y_1 and y_2 are the zero solutions.

$$\begin{aligned} &f = y''_f + a_1y'_f + a_0y_f \\ &f = (u_1y'_1 + u_2y'_2)' + a_1(u_1y'_1 + u_2y'_2) + a_0(u_1y_1 + u_2y_2) \\ &f = (u'_1y'_1 + u_1y''_1 + u'_2y'_2 + u_2y''_2) + a_1(u_1y'_1 + u_2y'_2) + a_0(u_1y_1 + u_2y_2) \\ &f = (u_1)(y''_1 + a_1y'_1 + a_0y_1) + (u_2)(y''_2 + a_1y'_2 + a_0y_2) + u'_1y'_1 + u'_2y'_2 \\ &f = (u_1)(0) + (u_2)(0) + u'_1y'_1 + u'_2y'_2 \\ &f = u'_1y'_1 + u'_2y'_2 \end{aligned}$$

This results in a square system of 2 equations.

$$\begin{cases} u'_1y_1 + u'_2y_2 = 0 \\ u'_1y'_1 + u'_2y'_2 = f \end{cases}$$

In 2 dimensions, we can easily solve for u_1' and u_2' using elimination, obtaining the result below.

$$u_1' = -\frac{y_2 f}{y_1 y_2' - y_2 y_1'}$$
$$u_2' = \frac{y_1 f}{y_1 y_2' - y_2 y_1'}$$

Then, we can simply integrate and substitute these back into our particular solution.

$$\begin{aligned} y_f &= u_1 y_1 + u_2 y_y \\ &= -y_1 \int \frac{y_2 f}{y_1 y_2' - y_2 y_1'} \, dx + y_2 \int \frac{y_1 f}{y_1 y_2' - y_2 y_1'} \, dx \end{aligned}$$

Higher-Order Variation of Parameters

When we wish to use variation of parameters to find the particular solution of an Nth order differential equation

$$y^n + a_1(x) y^{n-1} + \ldots + a_n(x) y = f(x)$$

we guess a solution of the form

$$y_f(x) = u_1(x) y_1(x) + u_2(x) y_2(x) + \ldots + u_n(x) y_n(x)$$

and force

$$\begin{array}{rclcccccc} y_f'(x) & = & u_1(x)y_1'(x) & + & u_2(x)y_2'(x) & + & \ldots & + & u_n(x)y_n'(x) \\ y_f''(x) & = & u_1(x)y_1''(x) & + & u_2(x)y_2''(x) & + & \ldots & + & u_n(x)y_n''(x) \\ & \vdots & & & & & & & \\ y_f^{(n-1)}(x) & = & u_1(x)y_1^{(n-1)}(x) & + & u_2(x)y_2^{(n-1)}(x) & + & \ldots & + & u_n(x)y_n^{(n-1)}(x) \end{array}.$$

By equating each derivative with the true derivative of y_f up to order N, we can set up a system of equations.

$$\begin{array}{ccccccccc} u_1'y_1 & + & u_2'y_2 & + & \ldots & + & u_n'y_n & = & 0 \\ u_1'y_1' & + & u_2'y_2' & + & \ldots & + & u_n'y_n' & = & 0 \\ & & & & & & & \vdots & \\ u_1'y_1^{(n-2)} & + & u_2'y_2^{(n-2)} & + & \ldots & + & u_n'y_n^{(n-2)} & = & 0 \\ u_1'y_1^{(n-1)} & + & u_2'y_2^{(n-1)} & + & \ldots & + & u_n'y_n^{(n-1)} & = & f \end{array}$$

This system is difficult to solve by elimination. But now we can use Cramer's rule! First, let's write our system more compactly, using the notation

$$y_i^{(0:n-1)} = \left\langle y_i, y_i', \ldots, y_i^{(n-1)} \right\rangle.$$

The system becomes

$$u_1'y_1^{(0:n-1)} + u_2'y_2^{(0:n-1)} + \ldots + u_n'y_n^{(0:n-1)} = \langle 0, \ldots, 0, f \rangle.$$

According to Cramer's rule, each u_i' is given by

$$u_i' = \frac{\det\left(y_1^{(0:n-1)}, \ldots, y_{i-1}^{(0:n-1)}, \langle 0, \ldots, 0, f \rangle, y_{i+1}^{(0:n-1)}, \ldots, y_n^{(0:n-1)}\right)}{\det\left(y_1^{(0:n-1)}, y_2^{(0:n-1)}, \ldots, y_n^{(0:n-1)}\right)}.$$

The denominator of this fraction is also known as the **Wronskian**, denoted

$$W(y_1, y_2, \ldots, y_n) = \det\left(y_1^{(0:n-1)}, y_2^{(0:n-1)}, \ldots, y_n^{(0:n-1)}\right).$$

If we define $W_{i,f}$ as

$$W_{i,f}(y_1, y_2, \ldots, y_n) = \det\left(y_1^{(0:n-1)}, \ldots, y_{i-1}^{(0:n-1)}, \langle 0, \ldots, 0, f\rangle, y_{i+1}^{(0:n-1)}, \ldots, y_n^{(0:n-1)}\right)$$

then we have

$$u_i' = \frac{W_{i,f}\left(y_1, y_2, \ldots, y_n\right)}{W\left(y_1, y_2, \ldots, y_n\right)}.$$

Finally, we can write the particular solution to the differential equation by integrating and substituting into our initial guess.

$$y_f = y_1 \int \frac{W_{1,f}\left(y_1, y_2, \ldots, y_n\right)}{W\left(y_1, y_2, \ldots, y_n\right)}\, dx + \ldots + y_n \int \frac{W_{n,f}\left(y_1, y_2, \ldots, y_n\right)}{W\left(y_1, y_2, \ldots, y_n\right)}\, dx$$

Demonstration

Let's illustrate this method on a simple example. To make it easier to find the zero solutions, we'll choose an example with constant coefficients, but remember that this method works even when the coefficients are functions themselves.

$$y''' - 2y'' - y' + 2y = e^{3x}$$

We start off by finding the zero solutions, i.e. those that satisfy the equation whose right-hand side is zero.

$$y''' - 2y'' - y' + 2y = 0$$

We do this by finding the roots of the characteristic polynomial $p(r) = r^3 - 2r^2 - r + 2$. We can find the roots via factoring by grouping.

$$\begin{aligned}
0 &= r^3 - 2r^2 - r + 2 \\
0 &= r^2(r-2) - 1(r-2) \\
0 &= (r^2 - 1)(r-2) \\
0 &= (r+1)(r-1)(r-2) \\
r &= -1, 1, 2
\end{aligned}$$

These roots correspond to the following zero solutions:

$$y_{-1} = C_{-1}e^{-x} \qquad y_1 = C_1 e^{x} \qquad y_2 = C_2 e^{2x}$$

It remains to find the particular solution. To use variation of parameters, we need three independent zero solutions, so we'll choose the simplest ones from above: e^{-x}, e^{x}, e^{2x}.

Substituting these into the variation of parameters formula, we have a particular solution of the form

$$\begin{aligned} y_f = e^{-x} \int \frac{W_{1,f}\left(e^{-x}, e^{x}, e^{2x}\right)}{W\left(e^{-x}, e^{x}, e^{2x}\right)}\, dx + e^{x} \int \frac{W_{2,f}\left(e^{-x}, e^{x}, e^{2x}\right)}{W\left(e^{-x}, e^{x}, e^{2x}\right)}\, dx \\ + e^{2x} \int \frac{W_{3,f}\left(e^{-x}, e^{x}, e^{2x}\right)}{W\left(e^{-x}, e^{x}, e^{2x}\right)}\, dx \end{aligned}$$

with $f(x) = e^{3x}$. Now, it remains to do the computations. First, we compute the standard Wronskian in the denominator.

$$\begin{aligned} W\left(e^{-x}, e^{x}, e^{2x}\right) &= \det\left(\begin{pmatrix} e^{-x} \\ \left(e^{-x}\right)' \\ \left(e^{-x}\right)'' \end{pmatrix}, \begin{pmatrix} e^{x} \\ \left(e^{x}\right)' \\ \left(e^{x}\right)'' \end{pmatrix}, \begin{pmatrix} e^{2x} \\ \left(e^{2x}\right)' \\ \left(e^{2x}\right)'' \end{pmatrix}\right) \\ &= \det\left(\begin{pmatrix} e^{-x} \\ -e^{-x} \\ e^{-x} \end{pmatrix}, \begin{pmatrix} e^{x} \\ e^{x} \\ e^{x} \end{pmatrix}, \begin{pmatrix} e^{2x} \\ 2e^{2x} \\ 4e^{2x} \end{pmatrix}\right) \\ &= \left(e^{-x}\right)\left(e^{x}\right)\left(e^{2x}\right) \det\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}\right) \\ &= \left(e^{-x}\right)\left(e^{x}\right)\left(e^{2x}\right)(6) \\ &= 6e^{2x} \end{aligned}$$

Next, we compute the modified Wronskians in the numerators.

$$\begin{aligned}
W_{1,f}\left(e^{-x}, e^{x}, e^{2x}\right) &= \det\left(\begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}, \begin{pmatrix} e^{x} \\ (e^{x})' \\ (e^{x})'' \end{pmatrix}, \begin{pmatrix} e^{2x} \\ \left(e^{2x}\right)' \\ \left(e^{2x}\right)'' \end{pmatrix}\right) \\
&= \det\left(\begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}, \begin{pmatrix} e^{x} \\ e^{x} \\ e^{x} \end{pmatrix}, \begin{pmatrix} e^{2x} \\ 2e^{2x} \\ 4e^{2x} \end{pmatrix}\right) \\
&= \left(e^{3x}\right)\left(e^{x}\right)\left(e^{2x}\right)\det\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}\right) \\
&= \left(e^{3x}\right)\left(e^{x}\right)\left(e^{2x}\right)(1) \\
&= e^{6x}
\end{aligned}$$

$$\begin{aligned}
W_{2,f}\left(e^{-x}, e^{x}, e^{2x}\right) &= \det\left(\begin{pmatrix} e^{-x} \\ (e^{-x})' \\ (e^{-x})'' \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}, \begin{pmatrix} e^{2x} \\ \left(e^{2x}\right)' \\ \left(e^{2x}\right)'' \end{pmatrix}\right) \\
&= \det\left(\begin{pmatrix} e^{-x} \\ -e^{-x} \\ e^{-x} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}, \begin{pmatrix} e^{2x} \\ 2e^{2x} \\ 4e^{2x} \end{pmatrix}\right) \\
&= \left(e^{-x}\right)\left(e^{3x}\right)\left(e^{2x}\right)\det\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}\right) \\
&= \left(e^{-x}\right)\left(e^{3x}\right)\left(e^{2x}\right)(-3) \\
&= -3e^{4x}
\end{aligned}$$

$$\begin{aligned}
W_{3,f}\left(e^{-x}, e^{x}, e^{2x}\right) &= \det\left(\begin{pmatrix} e^{-x} \\ (e^{-x})' \\ (e^{-x})'' \end{pmatrix}, \begin{pmatrix} e^{x} \\ (e^{x})' \\ (e^{x})'' \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}\right) \\
&= \det\left(\begin{pmatrix} e^{-x} \\ -e^{-x} \\ e^{-x} \end{pmatrix}, \begin{pmatrix} e^{x} \\ e^{x} \\ e^{x} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ e^{3x} \end{pmatrix}\right) \\
&= \left(e^{-x}\right)\left(e^{x}\right)\left(e^{3x}\right)\det\left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) \\
&= \left(e^{-x}\right)\left(e^{x}\right)\left(e^{3x}\right)(2) \\
&= 2e^{3x}
\end{aligned}$$

Lastly, we substitute back into the formula for the particular solution, and simplify.

$$\begin{aligned}
y_f &= e^{-x}\int \frac{W_{1,f}\left(e^{-x}, e^{x}, e^{2x}\right)}{W\left(e^{-x}, e^{x}, e^{2x}\right)}\,dx + e^{x}\int \frac{W_{2,f}\left(e^{-x}, e^{x}, e^{2x}\right)}{W\left(e^{-x}, e^{x}, e^{2x}\right)}\,dx \\
&\quad + e^{2x}\int \frac{W_{3,f}\left(e^{-x}, e^{x}, e^{2x}\right)}{W\left(e^{-x}, e^{x}, e^{2x}\right)}\,dx \\
&= e^{-x}\int \frac{e^{6x}}{6e^{2x}}\,dx + e^{x}\int \frac{-3e^{4x}}{6e^{2x}}\,dx + e^{2x}\int \frac{2e^{3x}}{6e^{2x}}\,dx \\
&= e^{-x}\int \frac{1}{6}e^{4x}\,dx + e^{x}\int -\frac{1}{2}e^{2x}\,dx + e^{2x}\int \frac{1}{3}e^{x}\,dx \\
&= e^{-x}\left(\frac{1}{24}e^{4x}\right) + e^{x}\left(-\frac{1}{4}e^{2x}\right) + e^{2x}\left(\frac{1}{3}e^{x}\right) \\
&= \frac{1}{24}e^{3x} + -\frac{1}{4}e^{3x} + \frac{1}{3}e^{3x} \\
&= \frac{1}{8}e^{3x}
\end{aligned}$$

The full solution to the differential equation, then, is

$$y = C_{-1}e^{-x} + C_1 e^x + C_2 e^{2x} + \frac{1}{8}e^{3x}.$$

Exercises

Solve the following differential equations using variation of parameters.

1) $y''' - y'' - 4y' + 4y = e^{2x}$

2) $y''' - y'' - 5y' - 3y = e^x$

3) $y''' + y'' - y' - y = \cos x$

4) $y''' - 2y'' - 9y' + 18y = \sin x$

5) $y''' - 2y'' + y' - 2 = \cos x$

6) $y''' - y'' + 4y' - 4 = \sin x$

Part 3
Matrices

3.1 Linear Systems as Transformations of Vectors by Matrices

Let's create a compact notation for expressing systems of linear equations like the one shown below.

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

Matrices of Column Vectors

We're familiar with a slightly condensed version using coefficient vectors.

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{mn} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

However, we can condense this even further by putting the coefficient vectors in a vector themselves and taking the dot product with the vector of variables.

$$\left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{mn} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right) \cdot \langle x_1, x_2, \cdots, x_n \rangle = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

To save space, the vector of variables can be written as a column vector as well.

$$\left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{mn} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Finally, to simplify the notation, we can remove the vector braces around the individual coefficient vectors and remove the dot product symbol.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

The array containing the coefficients is called a **matrix**. It's really just a vector of sub-vectors, written without braces on the individual sub-vectors.

Looking back, it makes sense to define a matrix multiplying a vector as follows:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \to \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{mn} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n$$

Matrices of Row Vectors

Keeping this form of matrix notation and multiplication in mind, let's start from scratch and proceed to condense a system of linear equations in a different way. We'll get an interesting result.

Again, we will start with the system below.

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

This time, however, we will begin by writing each equation as a dot product.

$$\begin{aligned} \langle a_{11}, a_{12}, \cdots, a_{1n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle &= b_1 \\ \langle a_{21}, a_{22}, \cdots, a_{2n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle &= b_2 \\ &\vdots \\ \langle a_{m1}, a_{m2}, \cdots, a_{mn} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle &= b_m \end{aligned}$$

Then, we will write the system as a single vector equation by interpreting each side of the equation as a vector.

$$\begin{pmatrix} \langle a_{11}, a_{12}, \cdots, a_{1n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \\ \langle a_{21}, a_{22}, \cdots, a_{2n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \\ \vdots \\ \langle a_{m1}, a_{m2}, \cdots, a_{mn} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Each component of the left-hand-side vector includes a dot product with the vector of variables, so we can factor out the vector of variables.

$$\begin{pmatrix} \langle a_{11}, a_{12}, \cdots, a_{1n} \rangle \\ \langle a_{21}, a_{22}, \cdots, a_{2n} \rangle \\ \vdots \\ \langle a_{m1}, a_{m2}, \cdots, a_{mn} \rangle \end{pmatrix} \langle x_1, x_2, \cdots, x_n \rangle = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Again, to save space, the vector of variables can be written as a column vector.

$$\begin{pmatrix} \langle a_{11}, a_{12}, \cdots, a_{1n} \rangle \\ \langle a_{21}, a_{22}, \cdots, a_{2n} \rangle \\ \vdots \\ \langle a_{m1}, a_{m2}, \cdots, a_{mn} \rangle \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Finally, to simplify the notation, we can again remove the vector braces around the individual coefficient vectors.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Again, there is a matrix! And again, the matrix just represents a vector of sub-vectors, written without braces on the individual sub-vectors.

But this time, looking back, it makes sense to define a matrix multiplying a vector by a different rule.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \to \begin{pmatrix} \langle a_{11}, a_{12}, \cdots, a_{1n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \\ \langle a_{21}, a_{22}, \cdots, a_{2n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \\ \vdots \\ \langle a_{m1}, a_{m2}, \cdots, a_{mn} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \end{pmatrix}$$

Matrix Multiplication

Which rule is correct? It turns out, they both are. Before we do an example, though, let's recap.

We're stumbling upon the following structure:

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

$$\downarrow$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

The array on the left-hand side is called a matrix, and we have two ways to compute the product of a matrix and a vector -- one which involves interpreting the columns of the matrix as individual vectors, and another which involves interpreting the rows of the matrix as individual vectors.

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{mn} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n$$

$$\nearrow$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\searrow$$

$$\begin{pmatrix} \langle a_{11}, a_{12}, \cdots, a_{1n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \\ \langle a_{21}, a_{22}, \cdots, a_{2n} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \\ \vdots \\ \langle a_{m1}, a_{m2}, \cdots, a_{mn} \rangle \cdot \langle x_1, x_2, \cdots, x_n \rangle \end{pmatrix}$$

To verify that both methods of computation indeed yield the same result, we can try out a simple example using the two different methods to compute the product of a 2-by-2 matrix and a 2-dimensional vector.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} (5) + \begin{pmatrix} 2 \\ 4 \end{pmatrix} (6) = \begin{pmatrix} 5 \\ 15 \end{pmatrix} + \begin{pmatrix} 12 \\ 24 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} \langle 1, 2 \rangle \cdot \langle 5, 6 \rangle \\ \langle 3, 4 \rangle \cdot \langle 5, 6 \rangle \end{pmatrix} = \begin{pmatrix} 5 + 12 \\ 15 + 24 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}$$

Geometric Intuition

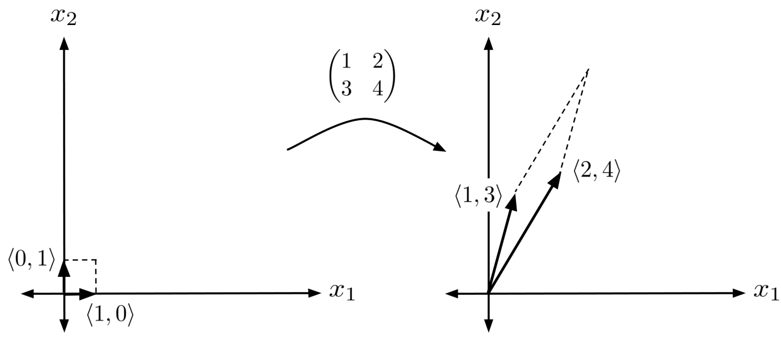
Lastly, let's build some geometric intuition. Geometrically, a matrix represents a transformation of a vector space, and we can visualize this transformation by thinking about what the matrix does to the N-dimensional unit cube.

For example, to see what the matrix from the example does to the unit square, we can multiply the vertices $(1, 0)$ and $(0, 1)$ of the unit square by the matrix.

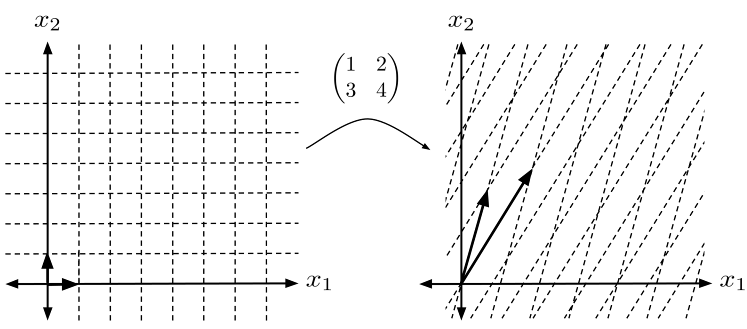
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} (1) + \begin{pmatrix} 2 \\ 4 \end{pmatrix} (0) = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

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

We see that the matrix moves the vertices of the unit square from $(1, 0)$ and $(0, 1)$, to $(1, 2)$ and $(3, 4)$. Notice that these are just the columns of the matrix!



But it’s not just the unit square that is transformed in this way. The *entire space* undergoes this transformation as well.



And it’s more than simple stretching -- the space is actually flipped over, since the original bottom vertex $(1, 0)$ is now the top vertex $(1, 3)$, and the original top vertex $(0, 1)$ is now the bottom vertex $(2, 4)$.

Exercises

Convert the following linear systems to matrix form.

1)
$$\begin{aligned} 3x_1 - 2x_2 &= 7 \\ 5x_1 + 4x_2 &= 6 \end{aligned}$$

2)
$$\begin{aligned} x_1 - 8x_2 &= 3 \\ x_1 + x_2 &= -2 \end{aligned}$$

3)
$$\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 5x_1 &= 8 \end{aligned}$$

4)
$$\begin{aligned} 8x_2 &= -7 \\ 3x_1 - x_2 &= 5 \end{aligned}$$

5)
$$\begin{aligned} 2x_1 + 3x_2 - 4x_3 &= 5 \\ 7x_1 - 2x_2 + 3x_3 &= 2 \\ 9x_2 + 5x_2 + 4x_3 &= 1 \end{aligned}$$

6)
$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 2x_1 - 5x_2 + x_3 &= -2 \\ x_1 + 4x_2 + 2x_3 &= 3 \end{aligned}$$

7)
$$\begin{aligned} x_1 - x_3 &= 4 \\ x_2 + 3x_3 &= 7 \\ x_1 + x_2 + x_3 &= -5 \end{aligned}$$

8)
$$\begin{aligned} x_2 + x_3 &= 6 \\ x_1 + x_3 &= 5 \\ x_1 + x_2 &= 4 \end{aligned}$$

9)
$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 7 \\ x_2 + x_3 - 7x_4 &= 5 \\ x_1 + 8x_3 &= 11 \\ 4x_2 + x_4 &= 3 \end{aligned}$$

10)
$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ x_2 - 2x_3 + 3x_4 &= 0 \\ x_1 - 2x_3 + 3x_4 &= 1 \\ x_1 - 2x_2 + 3x_4 &= 1 \end{aligned}$$

Compute the product of the given vector and matrix by A) interpreting the columns of the matrix as individual vectors, and B) interpreting the rows of the matrix as individual vectors. Verify that the results are the same.

11)

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

12)

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

13)

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

14)

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

15)

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

16)

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

17)

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

18)

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

19)

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

20)

$$\begin{pmatrix} 7 & 0 & -1 & -1 \\ 2 & 3 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 4 & 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ -1 \\ 2 \end{pmatrix}$$

3.2 Matrix Multiplication

We have seen how to multiply a vector by a matrix. Now, we will see how to multiply a matrix by another matrix.

Whereas multiplying a vector by a matrix corresponds to a linear transformation of that vector, multiplying a matrix by another matrix corresponds to a composition of linear transformations.

General Procedure

The procedure for matrix multiplication is quite familiar: we simply multiply each column vector in the right matrix by the left matrix.

Really, we’re just trying to figure out where the points $(1, 0)$ and $(0, 1)$ map to after being transformed once by the right matrix and then again by the left matrix. We already know that the right matrix maps those points to its columns, so all we have to do is map those columns according to the left matrix.

An example is shown below.

$$\begin{aligned}\begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix} &= \left(\begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}1\\3\end{pmatrix} \quad \begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}2\\4\end{pmatrix}\right)\\ &= \left(\begin{pmatrix}23\\31\end{pmatrix}\begin{pmatrix}34\\46\end{pmatrix}\right)\\ &= \begin{pmatrix}23 & 34\\31 & 46\end{pmatrix}\end{aligned}$$

We can verify that multiplying a vector by this new matrix gives the same result as multiplying the vector first by the original right matrix, and then by the original left matrix.

$$\begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}5 & 6\\7 & 8\end{pmatrix}\begin{pmatrix}5\\11\end{pmatrix} = \begin{pmatrix}91\\123\end{pmatrix}$$

$$\begin{pmatrix}23 & 34\\31 & 46\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}91\\123\end{pmatrix}$$

Case of Rectangular Matrices

Matrix multiplication isn't limited to just square matrices. The matrices can be rectangular, too.

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} &= \left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \right) \\
&= \left(\begin{pmatrix} 22 \\ 49 \\ 76 \\ 103 \end{pmatrix} \begin{pmatrix} 28 \\ 64 \\ 100 \\ 136 \end{pmatrix} \right) \\
&= \begin{pmatrix} 22 & 28 \\ 49 & 64 \\ 76 & 100 \\ 103 & 136 \end{pmatrix}
\end{aligned}$$

But notice that if we switch the above example around, it no longer makes sense to multiply the matrices, because we are unable to multiply each column vector in the right matrix by the left matrix. There are fewer columns in the left matrix than there are entries in each column of the right matrix.

$$\begin{aligned}
\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} &= \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 7 \\ 10 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \\ 8 \\ 11 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \\ 9 \\ 12 \end{pmatrix} \right) \\
&= (? \quad ? \quad ?)
\end{aligned}$$

Criterion for Multiplication

The trick to telling whether matrix multiplication is defined in a particular case is to check whether the width of the left matrix matches the height of the right matrix.

Matrix dimensions are usually written as $\text{height} \times \text{width}$, so matrix multiplication is defined whenever the inner dimensions match up.

For example, in the multiplication

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \\ 76 & 100 \\ 103 & 136 \end{pmatrix}$$

the left matrix has dimensions 4×3 and the right matrix has dimensions 3×2.

Writing these dimensions in the order of multiplication, we see that the inner dimensions do indeed match up: they're 3 and 3.

$$(4 \times 3) \times (3 \times 2)$$

Moreover, the outer dimensions give the dimensions of the resulting product: 4×2.

On the other hand, the matrices in the multiplication

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

have dimensions $(3 \times 2) \times (4 \times 3)$. The inner dimensions don't match up: they're 2 and 4. Therefore, the matrix multiplication is not defined.

Notice the implications for square matrices: multiplication is defined for square matrices only when they both have the same dimensions, say $N \times N$, and multiplication remains defined even if we switch the order of the square matrices, because the dimensions of the product stay the same:

$$(N \times N) \times (N \times N)$$

Moreover, the output is itself a square matrix of the same dimension, $N \times N$.

Non-Commutativity

Even for square matrices, though, **matrix multiplication is generally not commutative** -- if we switch the order of two matrices in a product, we tend to get a different result.

For example, switching the two matrices in the most recent example yields a different result:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

Even simple matrices generally do not commute:

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

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

The reason matrices tend not to commute is that left-multiplication and right-multiplication have different interpretations: left-multiplication sums combinations of row vectors, whereas right-multiplication sums combinations of column vectors.

Applying some operation to the rows of a matrix is generally not the same as applying that operation to the columns of a matrix.

Left-multiplication of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ **by** $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$

$$\begin{aligned} &\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &\quad = \begin{pmatrix} Aa + Bc & Ab + Bd \\ Ca + Dc & Cb + Dd \end{pmatrix} \\ &\quad = \begin{pmatrix} A \langle a, b \rangle + B \langle c, d \rangle \\ C \langle a, b \rangle + D \langle c, d \rangle \end{pmatrix} \end{aligned}$$

Right-multiplication of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ **by** $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$

$$\begin{aligned} &\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ &\quad = \begin{pmatrix} aA + bC & aB + bD \\ cA + dC & cB + dD \end{pmatrix} \\ &\quad = \begin{pmatrix} A \begin{pmatrix} a \\ c \end{pmatrix} + C \begin{pmatrix} b \\ d \end{pmatrix} & B \begin{pmatrix} a \\ c \end{pmatrix} + D \begin{pmatrix} b \\ d \end{pmatrix} \end{pmatrix} \end{aligned}$$

Diagonal Matrices

That being said, there are some instances in which matrices do commute.

For example, **diagonal** matrices commute with each other. (A diagonal matrix consists of zero everywhere except the diagonal running from the top-left entry to the bottom-right entry.)

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix}$$

Diagonal matrices commute with each other because the diagonal components end up being multiplied independently as scalars rather than vectors, and scalar multiplication does in fact commute.

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} Aa & 0 \\ 0 & Bb \end{pmatrix} = \begin{pmatrix} aA & 0 \\ 0 & bB \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

Be aware, though, that **antidiagonal** matrices generally do not commute with each other. (An antidiagonal matrix is like a diagonal matrix, but with the diagonal running from top-right to bottom-left.)

$$\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 4 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 4 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix}$$

Exercises

Compute the product of the given matrices, if possible, using A) the left-multiplication interpretation, and B) the right-multiplication interpretation.

Otherwise, if it is not possible to compute the product, then state the dimensions that A) the left matrix would need to have for the multiplication to be defined, or B) that the right matrix would need to have for the multiplication to be defined.

1)

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

2)

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

3)

$$\begin{pmatrix} -4 & 1 \\ 2 & 3 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 0 & 4 \\ 7 & -3 \end{pmatrix}$$

4)

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

5)

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

6)

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

7)

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

8)

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

9)

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

10)

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

3.3 Rescaling, Shearing, and the Determinant

The key insight in this chapter is that **every square matrix can be decomposed into a product of rescalings and shears**. Before we elaborate on that, though, let's discuss what rescalings and shears are, in terms of matrices.

Rescaling Matrices

Rescaling matrices are matrices that rescale the dimensions of space, with each dimension potentially being rescaled by a different amount. That is to say, the dimensions of space maintain their original direction, but their lengths are multiplied by some factors.

For example, in 2-dimensional space, the rescaling of $\langle 1, 0 \rangle$ into $\langle 2, 0 \rangle$ and $\langle 0, 1 \rangle$ into $\left\langle 0, \frac{1}{3} \right\rangle$ is given by left- or right-multiplication by the following rescaling matrix:

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

We could also have negative rescalings, or even zero rescalings that collapse a vector's length down to 0. For example, the rescaling matrix that rescales $\langle 1,0 \rangle$ into $\langle -5,0 \rangle$ and $\langle 0,1 \rangle$ into $\langle 0,0 \rangle$ is given by

$$\begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix}.$$

We can extend to higher dimensions as well. In 3-dimensional space, the rescaling matrix that rescales $\langle 1,0,0 \rangle$ to $\langle 2,0,0 \rangle$, and $\langle 0,1,0 \rangle$ to $\langle 0,-3,0 \rangle$, and $\langle 0,0,1 \rangle$ to $\left\langle 0,0,\frac{4}{5} \right\rangle$ is given by

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & \frac{4}{5} \end{pmatrix}.$$

Do you notice a pattern? **Rescaling matrices are just diagonal matrices!**

There is a fast trick for multiplying rescaling matrices: just multiply the diagonal entries independently. Consequently, the product of two rescaling matrices is itself always a rescaling matrix as well.

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & \frac{4}{5} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} (2)(4) & 0 & 0 \\ 0 & (-3)\left(\frac{1}{2}\right) & 0 \\ 0 & 0 & \left(\frac{4}{5}\right)(5) \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -\frac{3}{2} & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

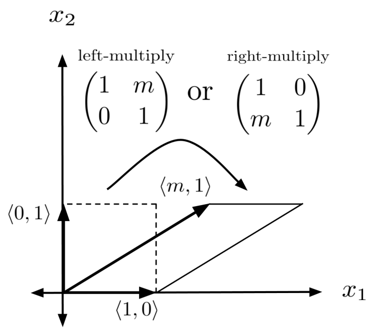
Likewise, there is also a fast trick to compute the determinant of a rescaling matrix. Since the vectors in a rescaling matrix form a rectangular prism, and the volume of that prism is obtained by multiplying the side lengths, the determinant of a rescaling matrix is simply the product of the rescalings, i.e. the product of the diagonal.

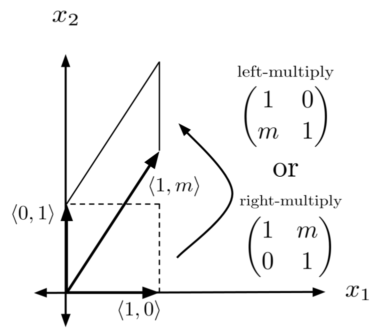
$$\det \begin{pmatrix} 8 & 0 & 0 \\ 0 & -\frac{3}{2} & 0 \\ 0 & 0 & 4 \end{pmatrix} = (8)\left(-\frac{3}{2}\right)(4) = -48$$

Shearing Matrices

Now let's talk about **shearing matrices**. Recall that shearing involves moving one of the sides of a parallelepiped in a parallel direction, and does not change the volume of the parallelepiped. We have also seen that in a set of vectors, shearing simply amounts to adding a multiple of some vector to a different vector.

Since a matrix is defined by its transformation of the unit cube, we can consider just the shears of the unit cube. In 2 dimensions, for example, a shear of the unit cube would either consist of vectors $\langle 1, 0 \rangle$ and $\langle m, 1 \rangle$, or $\langle 1, m \rangle$ and $\langle 0, 1 \rangle$, where m is the multiple of the vector that is added.





Likewise, in 3 dimensions, a shear of the unit cube could consist of vectors $\langle 1,0,0\rangle$, $\langle 0,1,0\rangle$, and $\langle m,n,1\rangle$; or $\langle 1,0,0\rangle$, $\langle m,1,n\rangle$, and $\langle 0,0,1\rangle$; or $\langle 1,0,0\rangle$, $\langle 0,1,0\rangle$, and $\langle m,n,1\rangle$, where m and n are the multiples of the vectors that are added. These correspond to the following matrices, respectively:

<u>Left-multiply</u>: $\begin{pmatrix}1 & 0 & m\\0 & 1 & n\\0 & 0 & 1\end{pmatrix}$ $\begin{pmatrix}1 & m & 0\\0 & 1 & 0\\0 & n & 1\end{pmatrix}$ $\begin{pmatrix}1 & 0 & 0\\m & 1 & 0\\n & 0 & 1\end{pmatrix}$

OR

<u>Right-multiply</u>: $\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\m & n & 1\end{pmatrix}$ $\begin{pmatrix}1 & 0 & 0\\m & 1 & n\\0 & 0 & 1\end{pmatrix}$ $\begin{pmatrix}1 & m & n\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}$

Do you notice a pattern? **Shear matrices consist of a diagonal of 1s, with all other entries zero except for possibly a single row or column.**

Unfortunately, there is no easy trick for multiplying shear matrices, other than just adding multiples of one row/column to another row/column. The result of multiplying two shear matrices might not even maintain a diagonal of 1s, for example:

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} = \begin{pmatrix} 1 + ab & a \\ b & 1 \end{pmatrix}$$

However, there is one property that is conserved in the result of multiplying shear matrices: the determinant of a product of shear matrices has to remain 1. This is because shear matrices don't change the volume of any parallelepiped within a vector space.

Decomposing into Rescalings and Shears

Now let's move onto the main idea of this chapter: **every square matrix can be decomposed into a product of rescalings and shears**. We'll illustrate the process through a couple of examples.

The process of decomposing a matrix into a product of rescalings and shears is very familiar -- it mainly consists of reducing the row or column vectors while keeping track of our multipliers in rescaling and shear vectors.

The only catch is that we need to keep track of the process *in reverse*, which means we have to flip the sign of the multipliers that we put in shear matrices, and take the reciprocal of the multipliers that we put in rescaling matrices.

For example, to decompose the matrix below into a product of rescalings and shears, we start by adding -3 times the first row to the second row, which means we put 3 in our left shear matrix to represent the reverse operation. Then, we multiply the bottom row by $-\frac{1}{2}$, which means we put -2 in our left rescaling matrix to represent the reverse operation.

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

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

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

Here is another example, which might initially seem tricky because the first row has a 0 as its first entry. However, we can create a 1 as the first entry by adding $-\frac{5}{4}$ of the second row. Then, all that remains is to rescale the second row.

$$\begin{pmatrix} 0 & 3 \\ -\frac{4}{5} & 0 \end{pmatrix}$$

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

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

Note that sometimes we may need to rescale by 0 to introduce a 1 into a row of 0s, such as in the top row of the matrix below.

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

Likewise, to introduce a 1 into a *column* of 0s, we can *right*-multiply by a rescaling matrix having a 0 entry on the diagonal.

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

Below is a final example of decomposing a 3×3 matrix into rescalings and shears.

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

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

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

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

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

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

Determinant of a Product

One important consequence of decomposing square matrices into rescalings and shears is that, **for two square matrices A and B, we have**

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

To understand why this is, imagine writing A and B each as a product of rescalings and shears.

Since the shears have no effect on volume, they can be removed from the product AB without changing $\det(AB)$.

Then, we are left with the rescaling matrices for A and B, which give the determinants for A and B, respectively.

Meaning of Negative Determinant

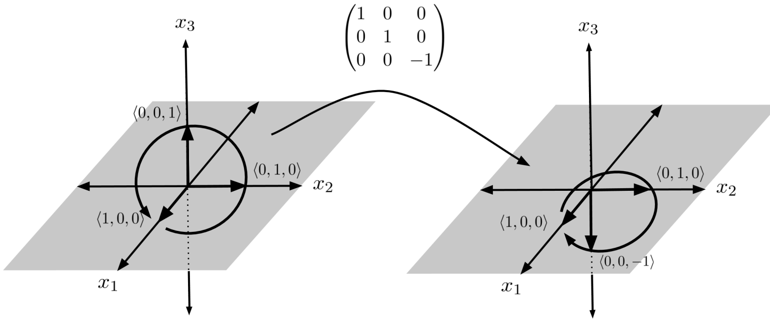
Another consequence of decomposing square matrices into rescalings and shears is that it makes clear the meaning of negative determinant.

Since shears don't change the determinant, a negative in a determinant must come from the rescalings -- meaning that the total number of negative entries in the diagonals of all the rescaling matrices must be odd.

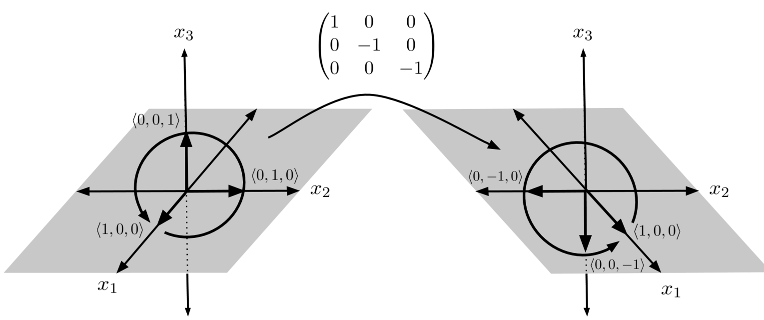
There is geometric intuition for negative determinants as well, having to do with the orientation of space.

The orientation of space can be thought of as a "curl" proceeding from x_1 to x_2, and then to x_3, and so on, until x_n, and then back to x_1. For example, for the unit cube in 3 dimensions, the curl is counterclockwise (when viewed opposite the origin).

However, applying a matrix with a single negative rescaling and thus a determinant of -1, one of the sides of the unit cube is flipped in the opposite direction. This causes the curl to reverse its orientation from counterclockwise to clockwise.



On the other hand, applying a matrix with two negative rescalings and thus a determinant of 1, two of the sides of the unit cube are flipped in the opposite direction, and the curl maintains its counterclockwise orientation.



Exercises

Decompose the following matrices into products of rescalings and shears.

1)

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

2)

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

3)

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

4)

$$\begin{pmatrix} 2 & 1 \\ 6 & 4 \end{pmatrix}$$

5)

$$\begin{pmatrix} 4 & 4 \\ 6 & 3 \end{pmatrix}$$

6)

$$\begin{pmatrix} 4 & 2 \\ 2 & 0 \end{pmatrix}$$

7)

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

8)

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

9)

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

10)

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

Compute $\det(X)$ for the matrix X in the equations below, given that $\det(A) = 1$, $\det(B) = 2$, and $\det(C) = 3$.

11) $X = ABC$

12) $X = A^3B^2C$

13) $AX = B$

14) $BX = C$

15) $AB^2X = AC$

16) $-XC^2 = BA$

17) $X^2A = B^2$

18) $AXBX = C$

3.4 Inverse Matrices

In this chapter, we introduce the idea of the **inverse** of a matrix, which undoes the transformation of that matrix.

Verifying an Inverse Matrix

For example, it's straightforward that the inverse of the rescaling matrix

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

is obtained as the rescaling matrix that rescales each dimension by the inverse amount.

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

We can verify that by multiplying the matrix by its inverse, and observing that the inverse takes the matrix back to the unit square.

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

Procedure for Finding the Inverse

But when we consider a more general matrix like the one below, it's less straightforward how to find the inverse.

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

We could try inverting each of the components separately, like we did with the diagonal of the rescaling matrix, but the resulting matrix doesn't take the original matrix back to the unit square -- so it can't be the inverse.

$$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} & 4 \\ \frac{13}{12} & \frac{5}{3} \end{pmatrix}$$

Here is another idea: since we want to end up with the unit square, let's left-multiply our matrix by other matrices representing row operations until we get to the unit square.

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

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

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

Then, let's take all the matrices we multiplied by, and find their product. That will be our inverse matrix.

$$\begin{aligned} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \end{aligned}$$

We can verify that indeed, this is the correct inverse matrix.

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

Left and Right Inverses

Based on the fact that we computed the inverse by left-multiplying, we should only expect the inverse to work for left-multiplication.

Interestingly, it works for right-multiplication as well!

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

This result is general to any inverse matrix -- regardless of whether we multiply a matrix by its inverse on the left or right, the result will be the unit cube.

To see why, we'll need to do a bit of simple algebra. To ease notation, we'll denote the unit cube matrix by I, which stands for the **identity matrix** and comes from the fact that $AI = IA = A$ for any matrix A.

Since $A^{-1}A = I$ for an inverse matrix obtained by left-multiplication, and since matrix multiplication is associative, we have

$$A = AI = A(A^{-1}A) = (AA^{-1})A.$$

But if we left-multiply A by a matrix and maintain a result of A, that matrix must be the identity! That is, if $A = (AA^{-1})A$, then we must have $AA^{-1} = I$. Hence, left and right inverses are one and the same.

Non-Invertible Matrices

Now, let's try to find the inverse of the matrix below. Something weird will happen.

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

This is simply a rescaling matrix with the rescaling quantities 2 and 0 on the diagonal. With rescaling matrices, we're used to finding the inverse by inverting the diagonal entries. We can invert 2 and get $\frac{1}{2}$, but we can't invert 0 -- the fraction $\frac{1}{0}$ is undefined.

It turns out, this matrix has no inverse. In general, any matrix having a 0 rescaling has no inverse, because once a vector is rescaled by a factor of 0, it's impossible to recover the original length of the vector -- as far as we know, it could be any length, because 0 times any number results in 0.

Criterion for Invertibility

By the same token, any matrix whose rescalings are all nonzero has an inverse. Once a vector is rescaled by a factor of $r \neq 0$, we can recover the original length of the vector by simply rescaling again by $\frac{1}{r}$.

Since the determinant of a matrix is the product of its rescalings, we can put all this together into an elegant statement: **a matrix is invertible if and only if its determinant is nonzero.**

This statement gives another perspective on why a linear system with nonzero determinant has exactly 1 solution, whereas a linear system with zero determinant has none or infinitely many solutions.

Any linear system can be written as a matrix equation $Ax = b$, and if $\det(A) \neq 0$, then A^{-1} exists, resulting in a single solution given by $x = A^{-1}b$.

On the other hand, if $\det(A) = 0$, then A contains some zero rescaling, and thus if there is any solution at all, then there must be infinitely many solutions because multiplication by zero gives the same result for infinitely many numbers.

Faster Method for Computing Inverses

Lastly, let's end by discussing a faster method to compute inverse matrices, based on the technique of reduction.

We already know how to use reduction to keep track of coefficients when solving linear systems by elimination -- but we'll introduce a more compact **augmented matrix** notation that will allow us to compute inverse matrices.

To solve the linear system below, we first convert it to an augmented matrix.

$$\begin{cases} 2x_1 + x_2 = 4 \\ x_1 + x_2 = 3 \end{cases} \to \left(\begin{array}{cc|c} 2 & 1 & 4 \\ 1 & 1 & 3 \end{array}\right)$$

Then, we perform row operations on the augmented matrix until we have reduced the left-hand side to the identity matrix.

$$\left(\begin{array}{cc|c} 2 & 1 & 4 \\ 1 & 1 & 3 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & 1 & 3 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right)$$

Finally, the solutions are displayed on the right-hand side: $x_1 = 1$ and $x_2 = 2$.

This process is familiar -- we're just left-multiplying by matrices corresponding to row operations until we get to the identity matrix, at which point we have effectively multiplied the original left-hand side matrix by its inverse.

Since we perform those same operations on the right-hand side vector, we are effectively multiplying the vector by the inverse matrix as well, which yields the solution.

If we want to find the actual inverse matrix, rather than just using it to solve the system, we can modify this process slightly by replacing the original right-hand side vector with the identity matrix.

Then, once the left-hand side matrix is taken to the identity matrix, the right-hand side identity matrix will be taken to the inverse matrix.

To find the actual inverse matrix in the previous example, we replace the right-hand side with the identity matrix and perform the same row operations to reduce the left-hand side.

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

Thus, we have the inverse matrix:

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

Formula for the Inverse of a 2-by-2 Matrix

There is a nice general formula for the inverse of a 2×2 matrix, which is given below.

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

It is recommended to memorize this formula to ease manipulations with 2×2 matrices, since the whole point of doing examples with 2×2 matrices is to ensure that they are relatively simple and fast.

Exercises

For each given matrix A, compute $\det(A)$ to tell whether A is invertible. If it is, then compute A^{-1}, and verify that $A^{-1}A = I$ and $AA^{-1} = I$.

1)
$$\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

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

3)
$$\begin{pmatrix} 0 & 1 \\ 5 & 7 \end{pmatrix}$$

4)
$$\begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}$$

5)

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

6)

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

7)

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 7 \\ 5 & 2 & -4 \end{pmatrix}$$

8)

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

9)

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

10)

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

For each equation $Ax = b$, tell whether A^{-1} exists. If it does, then compute the solution $x = A^{-1}b$.

11)

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} x = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

12)

$$\begin{pmatrix} -1 & 5 \\ 1 & 1 \end{pmatrix} x = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

13)

$$\begin{pmatrix} 3 & -7 \\ -9 & 21 \end{pmatrix} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

14)

$$\begin{pmatrix} 2 & 7 \\ 5 & -1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

15)

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

16)

$$\begin{pmatrix} 7 & 3 & 4 \\ 1 & 2 & 3 \\ 4 & -3 & -5 \end{pmatrix} x = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

17)

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

18)

$$\begin{pmatrix} 3 & 1 & 1 \\ -1 & 1 & 2 \\ 4 & 3 & 5 \end{pmatrix} x = \begin{pmatrix} -3 \\ 3 \\ 4 \end{pmatrix}$$

19)

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

20)

$$\begin{pmatrix} 0 & 7 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 2 & 3 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix}$$

Part 4
Eigenspace

4.1 Eigenvalues, Eigenvectors, and Diagonalization

Suppose we want to compute a matrix raised to a large power, i.e. multiplied by itself many times.

$$\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}^{999} = \underbrace{\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \cdots \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}}_{\text{999 copies}}$$

Of course, we could perform this computation using sheer brute force, multiplying out each of the 999 matrices -- but this would take a while.

On the other hand, we could go about the multiplications in a more clever way -- for example, if the matrix is A, then we could compute $AA = A^2$, $A^2A^2 = A^4$, and so on until we get to $A^{256}A^{256} = A^{512}$, and then compute

$$A^{999} = A^{512}A^{256}A^{128}A^{64}A^{32}A^4A^2A.$$

However, this would still require us to compute 14 multiplications, which -- although it is much better than the original 999 -- is still an annoyingly large amount of work, especially once the numbers inside the matrices become large.

Inverse Shearings and Rescalings

Fortunately, there is an even better way. First, notice that there is a way to express this matrix as a particular product of shearings and rescalings shown below.

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

The two shearings surrounding the rescaling are special in that they are inverses of each other:

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

As a result, if we multiply 999 copies of the *decomposed* matrix, we see that all of the shears cancel except the very first and the very last, leaving us with a product of 999 rescaling matrices in between.

$$\underbrace{\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \cdots \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}}_{\text{999 copies}}$$

$$\underbrace{\left[\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\right] \left[\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\right] \cdots \left[\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\right]}_{\text{999 copies}}$$

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}\underbrace{\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}}_{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}}\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}\underbrace{\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}}_{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}}\cdots\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}$$

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\underbrace{\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}\cdots\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix}}_{\text{999 copies}}\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}$$

But rescaling matrices are easy to multiply -- we can just multiply the diagonal entries separately! This leaves us with only 3 remaining matrix multiplications, which isn't too much work to do by hand.

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}2^{999} & 0\\0 & 3^{999}\end{pmatrix}\begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}$$

$$\begin{pmatrix}2^{999} & 3^{999}-2^{999}\\0 & 3^{999}\end{pmatrix}$$

Diagonalized Form

In order to reproduce this trick on other matrices, we need to come up with a general method for expressing a matrix A in the **diagonalized** form

$$A = PDP^{-1}$$

where D is a diagonal rescaling matrix and the surrounding matrices P and P^{-1} are inverses of each other.

In order to solve for P and D, it helps to right-multiply both sides of the equation by P so that

$$AP = PD.$$

Then, we can express P in terms of its column vectors v_i and D in terms of its diagonal entries λ_i, and multiply.

$$A \begin{pmatrix} | & | & \\ v_1 & v_2 & \cdots \\ | & | & \end{pmatrix} = \begin{pmatrix} | & | & \\ v_1 & v_2 & \cdots \\ | & | & \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix}$$

$$\begin{pmatrix} | & | & \\ Av_1 & Av_2 & \cdots \\ | & | & \end{pmatrix} = \begin{pmatrix} | & | & \\ \lambda_1 v_1 & \lambda_2 v_2 & \cdots \\ | & | & \end{pmatrix}$$

We see that the problem amounts to finding pairs of vectors v and scalars λ such that

$$Av = \lambda v.$$

Eigenvectors and Eigenvalues

Such vectors v are called **eigenvectors** of the matrix A, and the scalars λ that the eigenvectors are paired with are called **eigenvalues**.

Essentially, the eigenvectors of a matrix are those vectors that the matrix simply rescales, and the factor by which an eigenvector is rescaled is called its eigenvalue.

There is one important constraint: **the eigenvectors must be nonzero and independent**, since we need to be able to compute the inverse of the matrix that has them as columns.

In order to solve for the eigenvector and eigenvalue pairs, we rearrange the equation once more, introducing the identity matrix I so that we may factor out the eigenvector v.

$$\begin{aligned} Av &= \lambda v \\ Av - \lambda v &= 0 \\ Av - \lambda I v &= 0 \\ (A - \lambda I)v &= 0 \end{aligned}$$

Since we're assuming v is not the zero vector, the last equation tells us that some combination of not-all-zero multiples of columns of $A - \lambda I$ makes the zero vector. Consequently, the columns of $A - \lambda I$ must be dependent, and thus

$$\det(A - \lambda I) = 0.$$

Finally, we have an equation that we can use to solve for λ. Then, for each solution that we find for the eigenvalue λ, we can simply substitute back into $(A - \lambda I)v = 0$ to solve for the corresponding eigenvector v.

Demonstration of Diagonalization

Let's work an example. Say we want to diagonalize the matrix below.

$$\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix}$$

We start by solving the equation $\det(A - \lambda I) = 0$ for the eigenvalues λ.

$$\begin{aligned}
\det\left(\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) &= 0 \\
\det\left(\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) &= 0 \\
\det\begin{pmatrix} -10-\lambda & -6 \\ 18 & 11-\lambda \end{pmatrix} &= 0 \\
(-10-\lambda)(11-\lambda) - (-6)(18) &= 0 \\
-110 - \lambda + \lambda^2 + 108 &= 0 \\
\lambda^2 - \lambda - 2 &= 0 \\
(\lambda+1)(\lambda-2) &= 0 \\
\lambda &= -1, 2
\end{aligned}$$

Now that we have the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 2$, we solve the equation $(A - \lambda I)v = 0$ for corresponding eigenvectors v_1 and v_2.

$$\begin{aligned} (A - \lambda_1)\, v_1 &= 0 \\ \left(\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} - (-1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -9 & -6 \\ 18 & 12 \end{pmatrix} v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

At this point, one option is to write v_1 in terms of its components, say $v_1 = \langle s, t \rangle$, and simplify the matrix equation into a linear system in s and t.

$$\begin{pmatrix} -9 & -6 \\ 18 & 12 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} -9s & - & 6t & = & 0 \\ 18s & + & 12t & = & 0 \end{cases}$$

We can simplify the system by dividing the top equation by -3 and the bottom equation by 6. This reveals that the two equations are really just the same equation.

$$\begin{cases} 3s & + & 2t & = & 0 \\ 3s & + & 2t & = & 0 \end{cases}$$

As a result, they can be reduced down to a single equation, and we can easily solve for s in terms of t.

$$\begin{aligned} 3s + 2t &= 0 \\ 3s &= -2t \\ s &= -\frac{2}{3}t \end{aligned}$$

Substituting back into v_1, we have

$$\begin{aligned} v_1 &= \begin{pmatrix} s \\ t \end{pmatrix} \\ &= \begin{pmatrix} -\frac{2}{3}t \\ t \end{pmatrix} \\ &= \begin{pmatrix} -\frac{2}{3} \\ 1 \end{pmatrix} t \end{aligned}.$$

In other words, the eigenvector v_1 can be chosen as any multiple of the vector $\left\langle -\frac{2}{3}, 1 \right\rangle$. Intuitively, this makes sense: if $Av = \lambda v$, then any multiple cv of v will have the same property:

$$A(cv) = cAv = c\lambda v = \lambda(cv)$$

We only need to choose a single vector for v_1. For the sake of simplicity, we will choose v_1 to be the least multiple of $\left\langle -\frac{2}{3}, 1 \right\rangle$ that has whole number coefficients, and a positive first component. We multiply the vector by -3 to reach

$$\begin{aligned} v_1 &= \begin{pmatrix} -\frac{2}{3} \\ 1 \end{pmatrix} (-3) \\ &= \begin{pmatrix} 2 \\ -3 \end{pmatrix} \end{aligned} .$$

Thus, we have our first eigenvalue-eigenvector pair!

$$\lambda_1 = -1, \quad v_1 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

Solving for an eigenvector might seem like a bit of work, but once you go through the process several times, you'll notice a faster method: we can simply multiply by a diagonal matrix.

$$\begin{aligned} (A - \lambda_1)\, v_1 &= 0 \\ \left(\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} - (-1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -9 & -6 \\ 18 & 12 \end{pmatrix} v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -\frac{1}{3} & 0 \\ 0 & \frac{1}{6} \end{pmatrix} \begin{pmatrix} -9 & -6 \\ 18 & 12 \end{pmatrix} v_1 &= \begin{pmatrix} -\frac{1}{3} & 0 \\ 0 & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix} v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ v_1 &= \begin{pmatrix} 2 \\ -3 \end{pmatrix} t \end{aligned}$$

The diagonal matrix represents the operations we did the long way on the system of equations: dividing the top equation by -3 and the bottom equation by 6.

Then, we just have to choose v_1 as a vector whose dot product with $\langle 3, 2 \rangle$ is equal to 0. The simplest choice is $\langle 2, -3 \rangle$, and to keep the solution general, we introduce a parameter t to mean that v_1 is any nonzero multiple of $\langle 2, -3 \rangle$.

For the purposes of diagonalization, we just need one particular such vector, so we will choose the simplest case, $t = 1$ (and we will implicitly assume such choice when solving for other eigenvectors).

Using this method, we reach the same eigenvalue-eigenvector pair.

$$\lambda_1 = -1, \quad v_1 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

Next we repeat the same process to find the second eigenvalue-eigenvector pair, this time starting with our second eigenvalue $\lambda_2 = 2$.

$$\begin{aligned}
(A - \lambda_2)\, v_2 &= 0 \\
\left(\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) v_2 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -12 & -6 \\ 18 & 9 \end{pmatrix} v_2 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -\frac{1}{6} & 0 \\ 0 & \frac{1}{9} \end{pmatrix} \begin{pmatrix} -12 & -6 \\ 18 & 9 \end{pmatrix} v_2 &= \begin{pmatrix} -\frac{1}{6} & 0 \\ 0 & \frac{1}{9} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} v_2 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
v_2 &= \begin{pmatrix} -1 \\ 2 \end{pmatrix}
\end{aligned}$$

Now that we have our eigenvalues and eigenvectors, we can substitute them into our diagonalization.

$$A = PDP^{-1}$$

$$\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix}^{-1}$$

$$\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}^{-1}$$

$$\begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$$

More Complicated Case

In this example, the eigenvalues came out to nice integer values. As we'll see in the next example, eigenvalues and eigenvectors might be messy, involving roots or even complex numbers.

The next example will also be on a 3×3 matrix, to illustrate that the method of diagonalization is the same even for higher-dimensional matrices.

To diagonalize the matrix

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

we begin by computing the eigenvalues:

$$\begin{aligned}
\det\left(\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) &= 0 \\
\det \begin{pmatrix} 1-\lambda & 1 & 0 \\ -1 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} &= 0 \\
\det \begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda+\frac{1}{1-\lambda} & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix} &= 0 \\
(1-\lambda)\left(1-\lambda+\frac{1}{1-\lambda}\right)(1-\lambda) &= 0 \\
(1-\lambda)^3 + (1-\lambda) &= 0 \\
(1-\lambda)\left((1-\lambda)^2+1\right) &= 0 \\
\lambda = 1 \text{ or } (1-\lambda)^2 &= -1 \\
\lambda &= 1, 1 \pm i
\end{aligned}$$

Then, we solve for the eigenvectors corresponding to the eigenvectors $\lambda_1 = 1$, $\lambda_2 = 1+i$, and $\lambda_3 = 1-i$.

$$\begin{aligned}
(A - \lambda_1)\, v_1 &= 0 \\
\left(\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) v_1 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} v_1 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
v_1 &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
(A - \lambda_2)\, v_2 &= 0 \\
\left(\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - (1+i) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) v_2 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -i & 1 & 0 \\ -1 & -i & 0 \\ 0 & 0 & -i \end{pmatrix} v_2 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} -i & 1 & 0 \\ -1 & -i & 0 \\ 0 & 0 & -i \end{pmatrix} v_2 &= \begin{pmatrix} i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 1 & i & 0 \\ 1 & i & 0 \\ 0 & 0 & 1 \end{pmatrix} v_2 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
v_2 &= \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
(A-\lambda_3)\,v_3 &= 0 \\
\left(\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - (1-i) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) v_3 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} i & 1 & 0 \\ -1 & i & 0 \\ 0 & 0 & i \end{pmatrix} v_3 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -i \end{pmatrix} \begin{pmatrix} i & 1 & 0 \\ -1 & i & 0 \\ 0 & 0 & i \end{pmatrix} v_3 &= \begin{pmatrix} -i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -i \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 1 & -i & 0 \\ 1 & -i & 0 \\ 0 & 0 & 1 \end{pmatrix} v_3 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
v_3 &= \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}
\end{aligned}$$

Collecting our eigenvalues and eigenvectors, we have

$$\begin{aligned}
\lambda_1 = 1, \;\; v_1 &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
\lambda_2 = 1+i, \;\; v_2 &= \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \\
\lambda_3 = 1-i, \;\; v_3 &= \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}.
\end{aligned}$$

We substitute the eigenvalues and eigenvectors into our diagonalization.

$$A = PDP^{-1}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix}^{-1}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & i & -i \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+i & 0 \\ 0 & 0 & 1-i \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & i & -i \\ 1 & 0 & 0 \end{pmatrix}^{-1}$$

Then we compute P^{-1}.

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & i & -i & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & -i & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right)$$

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

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{i}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{i}{2} & 0 \end{array}\right)$$

Finally, we're done!

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & i & -i \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+i & 0 \\ 0 & 0 & 1-i \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & -\frac{i}{2} & 0 \\ \frac{1}{2} & \frac{i}{2} & 0 \end{pmatrix}$$

Eigenvalues with Multiple Eigenvectors

When diagonalizing some matrices such as the one below, we may end up with a single repeated eigenvalue, which corresponds to multiple independent eigenvectors.

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

This matrix consists of two distinct eigenvalues, one of which is repeated.

$$\begin{aligned}
\det\left(\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) &= 0 \\
\det\begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 1 & 2-\lambda \end{pmatrix} &= 0 \\
\det\begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} &= 0 \\
(2-\lambda)^2(1-\lambda) &= 0 \\
\lambda &= 2, 2, 1
\end{aligned}$$

When we solve for the eigenvector corresponding to the eigenvalue $\lambda = 2$, we find that the solution consists of combinations of *two* independent vectors.

$$\begin{aligned} \left(\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) v &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} v &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ v &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} s + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} t \end{aligned}$$

We shall use the simplest cases, $s = 1, t = 0$ and $s = 0, t = 1$, to choose two eigenvectors corresponding to the eigenvalue $\lambda = 2$. Thus, we have two eigenvalue-eigenvector pairs!

$$\lambda_1 = 2, \;\; v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\lambda_2 = 2, \;\; v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

We solve for the third eigenvector, corresponding to the eigenvalue $\lambda_3 = 1$, as usual.

$$\begin{aligned}
(A - \lambda_3)\, v_3 &= 0 \\
\left(\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) v_3 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} v_3 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
v_3 &= \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}
\end{aligned}$$

Then, we can invert the eigenvector matrix and diagonalize.

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

Non-Diagonalizable Matrices

Other times, though, we may not find enough independent eigenvectors to create the matrix P.

In such cases, A simply cannot be diagonalized (though we will later learn a different method to exponentiate such matrices without too much more work).

For an example of a non-diagonalizable matrix, consider the matrix below:

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

We are able to solve for the eigenvalues of this matrix:

$$\begin{aligned} \det\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) &= 0 \\ \det\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} &= 0 \\ (1-\lambda)^2 &= 0 \\ \lambda &= 1, 1 \end{aligned}$$

However, when we attempt to solve for the eigenvectors, we reach a problem: there is only one independent vector that satisfies $(A - \lambda I)v = 0$.

$$\begin{aligned} \left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) v &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} v &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ v &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

We need two pairs of eigenvalues and eigenvectors to diagonalize the matrix, but we have a repeated eigenvalue and only one independent eigenvector corresponding to that eigenvalue.

Thus, we simply do not have enough independent eigenvectors to diagonalize the matrix.

Exercises

Diagonalize the given matrices A, if possible. If diagonalization is possible, then use the diagonalization to compute a formula for A^n. Check your formula on the case $n = 2$.

1)
$$\begin{pmatrix} 2 & 0 \\ -1 & 4 \end{pmatrix}$$

2)
$$\begin{pmatrix} 6 & 6 \\ 0 & 2 \end{pmatrix}$$

3)
$$\begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix}$$

4)
$$\begin{pmatrix} 7 & -3 \\ -2 & 8 \end{pmatrix}$$

5)
$$\begin{pmatrix} -3 & -4 \\ -4 & 3 \end{pmatrix}$$

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

7)
$$\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}$$

8)
$$\begin{pmatrix} -4 & 0 \\ 5 & -4 \end{pmatrix}$$

9)

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

10)

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

11)

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

12)

$$\begin{pmatrix} 0 & 6 & 4 \\ -5 & 11 & 6 \\ 6 & -9 & -4 \end{pmatrix}$$

13)

$$\begin{pmatrix} 4 & 2 & -2 \\ -12 & -10 & 8 \\ -9 & -9 & 7 \end{pmatrix}$$

14)

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

4.2 Recursive Sequence Formulas via Diagonalization

In this chapter, we introduce an interesting application of matrix diagonalization: constructing closed-form expressions for recursive sequences.

Recursive Sequences

A recursive sequence is defined according to one or more initial terms and an update rule for obtaining the next term after some number of previous terms.

For example, the sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ given below are arithmetic and geometric sequences given in recursive form.

$$\begin{aligned} a_0 &= 3 \\ a_{n+1} &= a_n + 2 \end{aligned} \qquad \Bigg| \qquad \begin{aligned} b_0 &= 3 \\ b_{n+1} &= b_n \times 2 \end{aligned}$$

For both of these sequences, it is straightforward to write a closed-form expression for the Nth term:

$$\begin{aligned} a_n &= 3 + \underbrace{2 + 2 + \cdots + 2}_{\text{n copies}} \\ &= 3 + 2n \end{aligned} \qquad \Bigg| \qquad \begin{aligned} b_n &= 3 \times \underbrace{2 \times 2 \times \cdots \times 2}_{\text{n copies}} \\ &= 3 \times 2^n \end{aligned}$$

For other sequences, however, this is not so straightforward. For example, consider the **Fibonacci sequence**, whose first term is 0, whose second term is 1, and whose successive terms are obtained by adding the previous two terms together.

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

The recursive rule for the Fibonacci sequence is as follows:

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_{n+2} &= a_n + a_{n+1} \end{aligned}$$

Finding a Closed-Form Expression

Notice that we can express the recursive update rule using matrices.

$$\begin{pmatrix} a_{n+1} \\ a_{n+2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$$

Repeatedly multiplying by this matrix, we can write a closed-form expression for the Nth and (N+1)st terms.

$$\begin{aligned} \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

We can simplify this expression even further by diagonalizing the matrix. First, we solve for the eigenvalues.

$$\begin{aligned}
\det\left(\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) &= 0 \\
\det\begin{pmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} &= 0 \\
-\lambda(1-\lambda) - (1)(1) &= 0 \\
\lambda^2 - \lambda - 1 &= 0 \\
\lambda &= \frac{1 \pm \sqrt{5}}{2}
\end{aligned}$$

Now, we find the eigenvectors v_1 and v_2 that correspond to the eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

$$\begin{aligned}
\left(\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} - \left(\frac{1+\sqrt{5}}{2}\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 2 & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} v_1 &= \begin{pmatrix} 2 & 0 \\ 0 & \frac{1+\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 1+\sqrt{5} & -2 \\ 1+\sqrt{5} & -2 \end{pmatrix} v_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
v_1 &= \begin{pmatrix} 2 \\ 1+\sqrt{5} \end{pmatrix}
\end{aligned}$$

$$\left(\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} - \left(\frac{1-\sqrt{5}}{2} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) v_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \frac{-1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} v_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \frac{-1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} v_2 = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -1+\sqrt{5} & 2 \\ -1+\sqrt{5} & 2 \end{pmatrix} v_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$v_2 = \begin{pmatrix} 2 \\ 1-\sqrt{5} \end{pmatrix}$$

Finally, we can diagonalize the matrix.

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix}^{-1}$$

Substituting the diagonalized matrix into the original formula, we are able to simplify so much that we find a closed-form, non-matrix formula for the Nth term of the sequence.

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \left[\begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix}^{-1} \right]^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \left[\frac{1}{-4\sqrt{5}} \begin{pmatrix} 1-\sqrt{5} & -2 \\ -1-\sqrt{5} & 2 \end{pmatrix} \right] \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \frac{1}{2\sqrt{5}} \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n \\ -\left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix}$$

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right] \\ \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \right] \end{pmatrix}$$

$$a_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right]$$

This formula might be a little surprising -- the Fibonacci sequence consists only of whole numbers, yet $\sqrt{5}$ appears often in the formula!

However, the formula is indeed correct. We verify the formula for a_0, a_1, and a_2 -- and it will work on all the other terms as well.

$$\begin{aligned} a_0 &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^0 - \left(\frac{1-\sqrt{5}}{2}\right)^0\right] \\ &= \frac{1}{\sqrt{5}}[1-1] \\ &= 0 \end{aligned}$$

$$\begin{aligned} a_1 &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^1 - \left(\frac{1-\sqrt{5}}{2}\right)^1\right] \\ &= \frac{1}{\sqrt{5}}\left[\sqrt{5}\right] \\ &= 1 \end{aligned}$$

$$\begin{aligned} a_2 &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^2 - \left(\frac{1-\sqrt{5}}{2}\right)^2\right] \\ &= \frac{1}{\sqrt{5}}\left[\frac{1+2\sqrt{5}+5}{4} - \frac{1-2\sqrt{5}+5}{4}\right] \\ &= \frac{1}{\sqrt{5}}\left[\sqrt{5}\right] \\ &= 1 \end{aligned}$$

Case when Approximation is Required

This same method applies for any recursive sequence, though we may need to diagonalize a higher-dimensional matrix and numerically approximate the eigenvalues.

For example, consider the following spin-off of the Fibonacci sequence:

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_2 &= 1 \\ a_{n+3} &= 2a_n + a_{n+2} \end{aligned}$$

First, we express the recursive update rule using matrices, and write a closed-form expression involving an exponentiated matrix multiplying the first few terms.

$$\begin{aligned} \begin{pmatrix} a_{n+1} \\ a_{n+2} \\ a_{n+3} \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n+1} \\ a_{n+2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 1 \end{pmatrix}^n \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

We omit the steps of diagonalizing the matrix since they should be routine by now -- but it is worthwhile to discuss the method of approximating the eigenvalues.

When solving for the eigenvalues, we reach the following equation:

$$\lambda^3 - \lambda^2 - 2 = 0$$

This cubic cannot be factored manually -- not even using synthetic division -- since it has no rational roots.

Hence, we turn to a graphing utility to approximate a root $\lambda_1 \approx 1.696$. Then, we can perform synthetic division with that root to factor the polynomial into

$$(\lambda - 1.696)(\lambda^2 + 0.696\lambda + 1.180) = 0.$$

We can use the quadratic equation to solve

$$\lambda^2 + 0.696\lambda + 1.180 = 0$$

for the other two roots, which we find as $\lambda_2 \approx -0.348 - 1.029i$ and $\lambda_3 \approx -0.348 + 1.029i$.

Then, with a bit of grunt work, we can use these approximations to solve for the eigenvectors, substitute the diagonalization into the original equation, and multiply to find the formula for the Nth term.

The result, with each term rounded to 3 decimal places, is

$$\begin{aligned} a_n &\approx (-0.162 + 0.164i)(-0.348 - 1.029i)^n \\ &\quad - (0.162 + 0.164i)(-0.348 + 1.029i)^n \\ &\quad + 0.324(1.696)^n \end{aligned}.$$

Lastly, we can verify that the first several terms match up with the actual sequence $0, 1, 1, 1, 3, 5, 7, 13$.

Our estimates are slightly off due to compounded rounding error, but they could be made more accurate by using greater precision in the decimals that occur in the formula for a_n.

$$\begin{array}{cccc} a_0 = 0 & a_1 \approx 1.000 & a_2 \approx 1.001 & a_3 \approx 1.001 \\ a_4 \approx 3.003 & a_5 \approx 5.006 & a_6 \approx 7.011 & a_7 \approx 13.022 \end{array}$$

Exercises

Use diagonalization to compute a closed-form expression for the recursive sequence a_n.

1)

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_{n+2} &= 2a_n + a_{n+1} \end{aligned}$$

2)

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_{n+2} &= a_n + 2a_{n+1} \end{aligned}$$

3)

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_2 &= 1 \\ a_{n+3} &= a_n + a_{n+2} \end{aligned}$$

4)

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_2 &= 1 \\ a_{n+3} &= a_n + a_{n+1} + a_{n+2} \end{aligned}$$

4.3 Generalized Eigenvectors and Jordan Form

As we saw previously, not every matrix is diagonalizable, even when allowing complex eigenvalues/eigenvectors. The matrix below was given as an example of a non-diagonalizable matrix.

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

Patterns in Powers

However, notice that there's a pattern in the powers of this matrix.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^3 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^4 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$$

Leveraging this pattern, we can still write a formula for the Nth power of this matrix.

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

When we conduct this same experiment with a 3×3 matrix of similar form, a more general pattern pops up.

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

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

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^4 = \begin{pmatrix} 1 & 4 & 6 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

The pattern is that the numbers are all just binomial coefficients taken from Pascal's triangle! Writing this pattern more generally for a $k \times k$ square matrix, we have

$$\begin{pmatrix} 1 & 1 & 0 & \cdots & \cdots & 0 \\ & 1 & 1 & \cdots & \cdots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & \binom{n}{1} & \binom{n}{2} & \cdots & \cdots & \binom{n}{k-1} \\ & 1 & \binom{n}{1} & \cdots & \cdots & \binom{n}{k-2} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & 1 & \binom{n}{1} \\ & & & & & 1 \end{pmatrix}.$$

If we replace the diagonal with another number, say λ, then similar experimentation reveals the following formula:

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ & \lambda & 1 & \cdots & \cdots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ & \lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \cdots & \binom{n}{k-2}\lambda^{n-k+2} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ & & & & & \lambda^n \end{pmatrix}$$

Jordan Form

These matrices consisting of a diagonal λ directly below an off-diagonal of 1s are called **Jordan blocks**, and a matrix consisting of Jordan blocks is called a **Jordan matrix**.

For example, the matrix below consists of two Jordan blocks. (Note that blank entries correspond to 0.)

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

The big question, then, is: which matrices A can be expressed as

$$A = PJP^{-1}$$

where J is a Jordan matrix?

The answer is quite satisfying: all of them! Thus, **Jordan form provides a guaranteed backup plan for exponentiating matrices that are non-diagonalizable.**

Procedure for Finding a Jordan Form

So, how do we construct the matrices P and J? Let's start out like we did with diagonalization, right-multiplying both sides of the equation by P.

$$AP = PJ$$

To keep things simple but interesting enough to generalize our results, let's assume the following two-block Jordan matrix.

$$J = \begin{pmatrix} \lambda_1 & 1 & & & \\ 0 & \lambda_1 & & & \\ & & \lambda_2 & 1 & 0 \\ & & 0 & \lambda_2 & 1 \\ & & 0 & 0 & \lambda_2 \end{pmatrix}$$

Then we have

$$A \begin{pmatrix} | & | & | & | & | \\ v_{11} & v_{12} & v_{21} & v_{22} & v_{23} \\ | & | & | & | & | \end{pmatrix} = \begin{pmatrix} | & | & | & | & | \\ v_{11} & v_{12} & v_{21} & v_{22} & v_{23} \\ | & | & | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & & & \\ 0 & \lambda_1 & & & \\ & & \lambda_2 & 1 & 0 \\ & & 0 & \lambda_2 & 1 \\ & & 0 & 0 & \lambda_2 \end{pmatrix}$$

$$\begin{pmatrix} | & | & | & | & | \\ Av_{11} & Av_{12} & Av_{21} & Av_{22} & Av_{23} \\ | & | & | & | & | \end{pmatrix} = \begin{pmatrix} | & | & | & | & | \\ \lambda_1 v_{11} & v_{11} + \lambda_1 v_{12} & \lambda_2 v_{21} & v_{21} + \lambda_2 v_{22} & v_{22} + \lambda_2 v_{23} \\ | & | & | & | & | \end{pmatrix}.$$

First of all, since

$$A v_{11} = \lambda_1 v_{11}$$
$$A v_{21} = \lambda_2 v_{21}$$

we see that v_{11} and v_{21} are eigenvectors corresponding to the eigenvalues λ_1 and λ_2, respectively.

This makes intuitive sense because these columns mark the start of the Jordan blocks and thus don't have a 1 above them -- these columns are perfect diagonals.

Before we go on, notice that we can rearrange the above equations as follows:

$$(A - \lambda_1 I)\, v_{11} = 0$$
$$(A - \lambda_2 I)\, v_{21} = 0$$

This will be helpful shortly. Now, we move into the more novel cases, beginning by equating the second columns.

$$A v_{12} = v_{11} + \lambda_1 v_{12}$$

By rearranging the equation, we come up with an equation similar to those we found for the eigenvectors v_{11} and v_{21}.

$$\begin{aligned} Av_{12} &= v_{11} + \lambda_1 v_{12} \\ (A - \lambda_1 I)\, v_{12} &= v_{11} \\ (A - \lambda_1 I)^2\, v_{12} &= (A - \lambda_1 I)\, v_{11} \\ &= Av_{11} - \lambda_1 I v_{11} \\ &= \lambda_1 v_{11} - \lambda_1 v_{11} \\ &= 0 \end{aligned}$$

We call v_{12} a **generalized eigenvector** of order 2 for the eigenvalue λ_1 because it solves the equation $(A - \lambda_1 I)^2 v = 0$, whereas normal eigenvectors (i.e. generalized eigenvectors of order 1) for the eigenvalue λ_1 solve the equation $(A - \lambda_1 I)v = 0$.

By the same reasoning, we conclude that v_{22} is a generalized eigenvector of order 2 for λ_2, and v_{23} is a generalized eigenvector of order 3 for λ_2.

Demonstration

To conclude this chapter, we walk through an example of exponentiating the non-diagonalizable matrix below by converting it to Jordan form.

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

First, we compute the eigenvalues. In the row manipulations, we use the symbol $\square$ to denote matrix entries that change but have no consequence when computing the determinant.

$$\det \begin{pmatrix} 1-\lambda & 1 & -1 & 1 & -1 \\ 0 & -1-\lambda & 0 & 0 & 1 \\ 0 & 1 & -1-\lambda & 0 & 0 \\ 0 & 3 & -2 & 1-\lambda & -3 \\ 0 & 0 & 0 & 0 & -1-\lambda \end{pmatrix} = 0$$

$$\det \begin{pmatrix} 1-\lambda & 1 & -1 & 1 & -1 \\ 0 & -1-\lambda & 0 & 0 & 1 \\ 0 & 0 & -1-\lambda & 0 & \square \\ 0 & 0 & -2 & 1-\lambda & \square \\ 0 & 0 & 0 & 0 & -1-\lambda \end{pmatrix} = 0$$

$$\det \begin{pmatrix} 1-\lambda & 1 & -1 & 1 & -1 \\ 0 & -1-\lambda & 0 & 0 & 1 \\ 0 & 0 & -1-\lambda & 0 & \square \\ 0 & 0 & 0 & 1-\lambda & \square \\ 0 & 0 & 0 & 0 & -1-\lambda \end{pmatrix} = 0$$

$$(1-\lambda)^2(-1-\lambda)^3 = 0$$

$$\lambda = 1, 1, -1, -1, -1$$

Now that we have the eigenvalues $\lambda_1 = 1$ repeated twice and $\lambda_2 = -1$ repeated three times, we solve for the first and second-order generalized eigenvectors for λ_1, and the first, second, and third-order generalized eigenvectors for λ_2.

First, we solve for the first-order generalized eigenvector v_{11} of $\lambda_1 = 1$.

$$(A - \lambda_1 I)v_{11} = 0$$

$$(A - (1)I)v_{11} = 0$$

$$(A - I)v_{11} = 0$$

$$\begin{pmatrix} 0 & 1 & -1 & 1 & -1 \\ 0 & -2 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -2 & 0 & -3 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix} v_{11} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & -1 & 1 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{11} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{11} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{11} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{11} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$v_{11} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Next, we solve for the second-order generalized eigenvector v_{12} of $\lambda_1 = 1$, which is independent of the first-order generalized eigenvector v_{11}.

$$(A - \lambda_1 I)^2 v_{12} = 0$$

$$\begin{pmatrix} 0 & 1 & -1 & 1 & -1 \\ 0 & -2 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -2 & 0 & -3 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}^2 v_{12} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & -4 \\ 0 & -4 & 4 & 0 & 1 \\ 0 & -8 & 4 & 0 & 9 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} v_{12} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & -4 & 4 & 0 & 0 \\ 0 & -8 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{12} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{12} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{12} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$v_{12} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Before we go on, let's take inventory of what we have, filling in part of our Jordan form expression.

$$\begin{pmatrix} 1 & 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 3 & -2 & 1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} | & | & | & | & | \\ v_{11} & v_{12} & v_{21} & v_{22} & v_{23} \\ | & | & | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 & & & \\ 0 & \lambda_1 & & & \\ & & \lambda_2 & 1 & 0 \\ & & 0 & \lambda_2 & 1 \\ & & 0 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} | & | & | & | & | \\ v_{11} & v_{12} & v_{21} & v_{22} & v_{23} \\ | & | & | & | & | \end{pmatrix}^{-1}$$

$$\begin{pmatrix} 1 & 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 3 & -2 & 1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & | & | & | \\ 0 & 0 & | & | & | \\ 0 & 0 & v_{21} & v_{22} & v_{23} \\ 0 & 1 & | & | & | \\ 0 & 0 & | & | & | \end{pmatrix} \begin{pmatrix} 1 & 1 & & & \\ 0 & 1 & & & \\ & & \lambda_2 & 1 & 0 \\ & & 0 & \lambda_2 & 1 \\ & & 0 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & | & | & | \\ 0 & 0 & | & | & | \\ 0 & 0 & v_{21} & v_{22} & v_{23} \\ 0 & 1 & | & | & | \\ 0 & 0 & | & | & | \end{pmatrix}^{-1}$$

Continuing, we solve for the first-order generalized eigenvector v_{21} of $\lambda_2 = -1$.

$$(A - \lambda_2 I)v_{21} = 0$$

$$(A - (-1)I)v_{21} = 0$$

$$(A + I)v_{21} = 0$$

$$\begin{pmatrix} 2 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & -2 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{21} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{21} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{21} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{21} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$v_{21} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Then, we solve for the second-order generalized eigenvector v_{22} of $\lambda_2 = -1$, which is independent of the first-order generalized eigenvector v_{21}.

$$(A - \lambda_2 I)^2 v_{22} = 0$$

$$\begin{pmatrix} 2 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & -2 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}^2 v_{22} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 4 & -4 & 4 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 4 & -4 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{22} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{22} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} v_{22} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$v_{22} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Lastly, we solve for the third-order generalized eigenvector v_{23} of $\lambda_2 = -1$, which is independent of the first and second-order generalized eigenvectors v_{21} and v_{22}.

$$(A - \lambda_2 I)^3 v_{23} = 0$$

$$\begin{pmatrix} 2 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 3 & -2 & 2 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}^3 v_{23} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 8 & 12 & -12 & 12 & -12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & -8 & 8 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{23} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{23} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} v_{23} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$v_{23} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Finally, we can fill in the rest of our Jordan form expression.

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

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

Exponentiating our matrix, we have

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

To exponentiate the middle matrix, it suffices to exponentiate the two blocks separately. The first block is simple and familiar:

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

For the second block, we make use of the general formula

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ & \lambda & 1 & \cdots & \cdots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ & \lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \cdots & \binom{n}{k-2}\lambda^{n-k+2} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ & & & & & \lambda^n \end{pmatrix}$$

and find that

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}^n = \begin{pmatrix} (-1)^n & \binom{n}{1}(-1)^{n-1} & \binom{n}{2}(-1)^{n-2} \\ 0 & (-1)^n & \binom{n}{1}(-1)^{n-1} \\ 0 & 0 & (-1)^n \end{pmatrix}$$
$$= \begin{pmatrix} (-1)^n & n(-1)^{n-1} & \frac{n(n-1)}{2}(-1)^{n-2} \\ 0 & (-1)^n & n(-1)^{n-1} \\ 0 & 0 & (-1)^n \end{pmatrix}.$$

Thus, we have

$$\begin{pmatrix} 1 & 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 3 & -2 & 1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}^n = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n & & & \\ 0 & 1 & & & \\ & & (-1)^n & n(-1)^{n-1} & \frac{n(n-1)}{2}(-1)^{n-2} \\ & & 0 & (-1)^n & n(-1)^{n-1} \\ & & 0 & 0 & (-1)^n \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Multiplying out and simplifying, we reach

$$\begin{pmatrix} 1 & 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 3 & -2 & 1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}^n = \begin{pmatrix} 1 & n & -n & n & -n \\ 0 & (-1)^n & 0 & 0 & n(-1)^{n+1} \\ 0 & n(-1)^{n+1} & (-1)^n & 0 & \frac{n(n-1)}{2}(-1)^n \\ 0 & (n+1)(-1)^{n+1}+1 & -1+(-1)^n & 1 & \frac{n(n+1)}{2}(-1)^n+(-1)^n-1 \\ 0 & 0 & 0 & 0 & (-1)^n \end{pmatrix}.$$

Lastly, let's verify this formula on the case $n = 2$.

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

$$\begin{pmatrix} 1 & 2 & -2 & 2 & -2 \\ 0 & (-1)^2 & 0 & 0 & 2(-1)^{2+1} \\ 0 & 2(-1)^{2+1} & (-1)^2 & 0 & \frac{2(2-1)}{2}(-1)^2 \\ 0 & (2+1)(-1)^{2+1}+1 & -1+(-1)^2 & 1 & \frac{2(2+1)}{2}(-1)^2+(-1)^2-1 \\ 0 & 0 & 0 & 0 & (-1)^2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -2 & 2 & -2 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & -2 & 1 & 0 & 1 \\ 0 & -2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

It checks out!

Exercises

For each matrix A, express $A = PJP^{-1}$ where J is a Jordan matrix, and use this Jordan expression to compute A^n. Check your formula on the case $n = 2$.

1)

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

2)

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

3)

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

4)

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

5)

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

6)

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

7)

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

8)

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

9)

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

10)

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

4.4 Matrix Exponential and Systems of Linear Differential Equations

In this chapter, we will learn how to solve systems of linear differential equations. These systems take the form shown below.

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k \\ &\vdots \\ \frac{dx_k}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{kk}x_k \end{aligned}$$

Converting to Matrix Form

We can write the system in matrix form:

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_k}{dt} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$$

Defining

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}$$

the system can be written more compactly as

$$\frac{dx}{dt} = Ax.$$

This bears resemblance to a familiar differential equation $\frac{dx}{dt} = ax$, where x and a are both scalars. We know that the solution to such a system is given by $x = e^{at}x(0)$.

We infer, then, that the solution to the matrix differential equation is given by

$$x = e^{At}x(0).$$

But what does it mean to exponentiate a matrix? How should we compute e^{At}?

Matrix Exponential

Recall that e^t can be written as the power series

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}.$$

Consequently, e^{at} can be written as the power series

$$e^{at} = \sum_{n=0}^{\infty} \frac{(at)^n}{n!} = \sum_{n=0}^{\infty} \frac{a^n t^n}{n!}.$$

Extending this to the matrix exponential e^{At}, then, we have

$$e^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!} = \sum_{n=0}^{\infty} \frac{A^n t^n}{n!}.$$

Writing A in the form

$$A = PJP^{-1}$$

where J is a Jordan matrix, we have

$$\begin{aligned} e^{At} &= \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(PJP^{-1})^n t^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{PJ^n P^{-1} t^n}{n!} \\ &= P \left[\sum_{n=0}^{\infty} \frac{J^n t^n}{n!} \right] P^{-1} \\ &= Pe^{Jt}P^{-1} \end{aligned}.$$

Thus, computing the exponential of a matrix reduces to the problem of computing the exponential of the corresponding Jordan matrix. As such, we need only investigate how to compute exponentials of Jordan blocks.

First, we consider the simplest case: a perfectly diagonal block.

$$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_k \end{pmatrix}$$

In this case, we have

$$e^{Dt} = \sum_{n=0}^{\infty} \frac{D^n t^n}{n!}$$

$$e^{Dt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_k \end{pmatrix}^n$$

$$e^{Dt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} \lambda_1^n & & & \\ & \lambda_2^n & & \\ & & \ddots & \\ & & & \lambda_k^n \end{pmatrix}$$

$$e^{Dt} = \begin{pmatrix} \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda_1^n & & & \\ & \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda_2^n & & \\ & & \ddots & \\ & & & \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda_k^n \end{pmatrix}$$

$$e^{Dt} = \begin{pmatrix} \sum\limits_{n=0}^{\infty} \frac{(\lambda_1 t)^n}{n!} & & & \\ & \sum\limits_{n=0}^{\infty} \frac{(\lambda_2 t)^n}{n!} & & \\ & & \ddots & \\ & & & \sum\limits_{n=0}^{\infty} \frac{(\lambda_k t)^n}{n!} \end{pmatrix}$$

$$e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_k t} \end{pmatrix}.$$

Second, we consider the more involved case: a block with an off-diagonal of 1s.

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ & \lambda & 1 & \cdots & \cdots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}$$

In this case, we have

$$e^{Jt} = \sum_{n=0}^{\infty} \frac{J^n t^n}{n!}$$

$$e^{Jt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ & \lambda & 1 & \cdots & \cdots & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}^n .$$

Taking the convention $\binom{n}{c} = 0$ when $c > n$, we can write

$$e^{Jt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ & \lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \cdots & \binom{n}{k-2}\lambda^{n-k+2} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ & & & & & \lambda^n \end{pmatrix}$$

$$e^{Jt} = \begin{pmatrix} \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda^n & \sum\limits_{n=1}^{\infty} \frac{t^n}{n!}\binom{n}{1}\lambda^{n-1} & \sum\limits_{n=2}^{\infty} \frac{t^n}{n!}\binom{n}{2}\lambda^{n-2} & \cdots & \cdots & \sum\limits_{n=k-1}^{\infty} \frac{t^n}{n!}\binom{n}{k-1}\lambda^{n-k+1} \\ & \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda^n & \sum\limits_{n=1}^{\infty} \frac{t^n}{n!}\binom{n}{1}\lambda^{n-1} & \cdots & \cdots & \sum\limits_{n=k-2}^{\infty} \frac{t^n}{n!}\binom{n}{k-2}\lambda^{n-k+2} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda^n & \sum\limits_{n=1}^{\infty} \frac{t^n}{n!}\binom{n}{1}\lambda^{n-1} \\ & & & & & \sum\limits_{n=0}^{\infty} \frac{t^n}{n!}\lambda^n \end{pmatrix} .$$

Notice that the entries in the matrix take the form

$$\sum_{n=c}^{\infty} \frac{t^n}{n!} \binom{n}{c} \lambda^{n-c}$$

where c is the column index of the matrix. We can simplify these expressions as follows:

$$\begin{aligned}
&\sum_{n=c}^{\infty} \frac{t^n}{n!} \binom{n}{c} \lambda^{n-c} \\
&= \sum_{n=c}^{\infty} \frac{t^n}{n!} \frac{n!}{c!(n-c)!} \lambda^{n-c} \\
&= \sum_{n=c}^{\infty} \frac{t^n \lambda^{n-c}}{c!(n-c)!} \\
&= \sum_{n=0}^{\infty} \frac{t^{n+c} \lambda^n}{c! n!} \\
&= \frac{t^c}{c!} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \\
&= \frac{t^c}{c!} e^{\lambda t}
\end{aligned}$$

Thus,

$$e^{Jt} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} & \cdots & \cdots & \frac{t^{k-1}}{(k-1)!}e^{\lambda t} \\ & e^{\lambda t} & te^{\lambda t} & \cdots & \cdots & \frac{t^{k-2}}{(k-2)!}e^{\lambda t} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & e^{\lambda t} & te^{\lambda t} \\ & & & & & e^{\lambda t} \end{pmatrix}$$

$$e^{Jt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \cdots & \frac{t^{k-1}}{(k-1)!} \\ & 1 & t & \cdots & \cdots & \frac{t^{k-2}}{(k-2)!} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & 1 & t \\ & & & & & 1 \end{pmatrix}.$$

Demonstration

Now, we're ready to run through an example. We shall solve the system below.

$$\begin{array}{ccc|c} \frac{dx_1}{dt} & = & x_1 + x_2 - x_3 + x_4 - x_5 & x_1(0) = 1 \\ \frac{dx_2}{dt} & = & -x_2 + x_5 & x_2(0) = 0 \\ \frac{dx_3}{dt} & = & x_2 - x_3 & x_3(0) = 1 \\ \frac{dx_4}{dt} & = & 3x_2 - 2x_3 + x_4 - 3x_5 & x_4(0) = 0 \\ \frac{dx_5}{dt} & = & -x_5 & x_5(0) = 1 \end{array}$$

First, we convert the system to matrix form.

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \\ \frac{dx_4}{dt} \\ \frac{dx_5}{dt} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 3 & -2 & 1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

$$\frac{dx}{dt} = \begin{pmatrix} 1 & 1 & -1 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 3 & -2 & 1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} x$$

Next, we write the matrix in the form PJP^{-1} where J is a Jordan matrix. We did this with the same matrix in the previous chapter, so we will just assume our previous result.

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

We know that the solution to the system is given by $x = Pe^{Jt}P^{-1}x(0)$, which we will be able to multiply once we compute e^{Jt}. We break up the computation of e^{Jt} across the two blocks within J.

$$J_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad J_2 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

Applying our formula from earlier, we have

$$e^{J_1 t} = e^t \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \qquad e^{J_2 t} = e^{-t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Putting this together, we have

$$e^{Jt} = \begin{pmatrix} e^t & te^t & & & \\ 0 & e^t & & & \\ & & e^{-t} & te^{-t} & \frac{t^2}{2}e^{-t} \\ & & 0 & e^{-t} & te^{-t} \\ & & 0 & 0 & e^{-t} \end{pmatrix}.$$

Finally, we compute the solution.

$$x = Pe^{Jt}P^{-1}x(0)$$

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^t & te^t & & & \\ 0 & e^t & & & \\ & & e^{-t} & te^{-t} & \frac{t^2}{2}e^{-t} \\ & & 0 & e^{-t} & te^{-t} \\ & & 0 & 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^t & te^t & & & \\ 0 & e^t & & & \\ & & e^{-t} & te^{-t} & \frac{t^2}{2}e^{-t} \\ & & 0 & e^{-t} & te^{-t} \\ & & 0 & 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & -1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^t & te^t & & & \\ 0 & e^t & & & \\ & & e^{-t} & te^{-t} & \frac{t^2}{2}e^{-t} \\ & & 0 & e^{-t} & te^{-t} \\ & & 0 & 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 2 \\ -1 \\ 1 \end{pmatrix}$$

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} (1-2t)e^t \\ -2e^t \\ \left(\frac{t^2}{2} - t + 2\right)e^{-t} \\ (t-1)e^{-t} \\ e^{-t} \end{pmatrix}$$

$$x = \begin{pmatrix} (1-2t)\,e^t \\ te^{-t} \\ \left(\frac{t^2}{2} + 1\right)e^{-t} \\ \left(\frac{t^2}{2} - t + 2\right)e^{-t} - 2e^t \\ e^{-t} \end{pmatrix}$$

Exercises

Solve each system of linear differential equations by converting it to a matrix equation $\frac{dx}{dt} = Ax$ and computing the solution $x = e^{At}x(0)$.

1)

$$\begin{array}{rcl|l} \frac{dx_1}{dt} & = & x_1 + x_2 & x_1(0) = 0 \\ \frac{dx_2}{dt} & = & -x_2 & x_2(0) = 1 \end{array}$$

2)

$$\begin{array}{rcc|l} \frac{dx_1}{dt} & = & x_1 + x_2 & x_1(0) = 1 \\ \frac{dx_2}{dt} & = & x_1 - x_2 & x_2(0) = 1 \end{array}$$

3)

$$\begin{array}{rcc|l} \frac{dx_1}{dt} & = & x_2 & x_1(0) = 0 \\ \frac{dx_2}{dt} & = & x_1 + x_3 & x_2(0) = 0 \\ \frac{dx_3}{dt} & = & x_3 & x_3(0) = 1 \end{array}$$

4)

$$\begin{array}{rcc|l} \frac{dx_1}{dt} & = & x_1 + x_3 & x_1(0) = 0 \\ \frac{dx_2}{dt} & = & x_1 + x_2 & x_2(0) = 1 \\ \frac{dx_3}{dt} & = & x_2 + x_3 & x_3(0) = 1 \end{array}$$

5)

$$\begin{array}{rcc|l} \frac{dx_1}{dt} & = & x_2 & x_1(0) = 1 \\ \frac{dx_2}{dt} & = & x_3 & x_2(0) = 0 \\ \frac{dx_3}{dt} & = & x_4 & x_3(0) = 0 \\ \frac{dx_4}{dt} & = & x_1 & x_3(0) = 0 \end{array}$$

6)

$$\begin{array}{rcc|l} \frac{dx_1}{dt} & = & x_2 - x_3 & x_1(0) = 1 \\ \frac{dx_2}{dt} & = & x_3 - x_4 & x_2(0) = 0 \\ \frac{dx_3}{dt} & = & x_4 - x_1 & x_3(0) = 0 \\ \frac{dx_4}{dt} & = & x_1 - x_2 & x_3(0) = 0 \end{array}$$

Solutions

to Exercises

Part 1

Chapter 1.1

1) input dimension: 2
output dimension: 3

2) input dimension: 1
output dimension: 1

3) input dimension: 3
output dimension: 1

4) input dimension: 2
output dimension: n

5) input dimension: m
output dimension: $m-1$

6) input dimension: $m+n$
output dimension: 2

7) $\langle -1, 8, 2\rangle$

8) $\langle -8, 0\rangle$

9) $5\sqrt{6}$

10) $2\sqrt{2}$

11) $\langle 0, 2\sqrt{5}, 0, -2\sqrt{5}\rangle$

12) $\langle 3, 1, -4, 1, -2\rangle$

13) $x = \langle 3, 5\rangle$

14) $x = 14$

15) $x = \pm 4$

16) $x = -2$

17) $x = \langle -1, 1, -1, 1\rangle$

18) $x = \pm\sqrt{\frac{2}{3}}$

Chapter 1.2

1) 7

2) -4

3) 0

4) 3

5) $\langle 0, 0, -4 \rangle$

6) $\langle 9, 13, -8 \rangle$

7) $\langle -1, 3, -1 \rangle$

8) $\langle 1, 2, 0 \rangle$

9) $x = -2$

10) $x = 0$

11) $x = \pm 3$

12) $x = 4$

13) $x = -1$

14) $x = 3$

15) $\arccos\left(\frac{3}{\sqrt{10}}\right)$
$\approx 18.43^\circ$

16) $\arccos\left(\frac{1}{\sqrt{7}}\right)$
$\approx 67.79^\circ$

17) $\frac{\pi}{2} = 90^\circ$

18) $\arccos\left(-\frac{5}{3\sqrt{14}}\right)$
$\approx 116.45^\circ$

19) 6

20) $4\sqrt{6}$

21) $\sqrt{2}$

22) $\sqrt{110}$

Chapter 1.3

1) answers may vary; one correct answer is
$x(t) = \langle 1, 2, 3 \rangle + \langle 2, 0, -2 \rangle\, t$

2) answers may vary; one correct answer is
$x(t) = \langle 1, -1, 2, -2 \rangle + \langle -1, 2, 2, -1 \rangle\, t$

3) answers may vary; one correct answer is $x(t) = \langle -5, -1, 2, 1, 3 \rangle + \langle 9, 5, 0, 1, -3 \rangle\, t$

4) answers may vary; one correct answer is $x(t) = \langle 1, 0, -3, -3, 0, 1 \rangle + \langle 2, 1, 2, 3, 2, -1 \rangle\, t$

5) yes; $t = 2$

6) no

7) no

8) yes; $t = 5$

9) $x_1 + 2x_2 + 3x_3 = 0$

10) $2x_1 - 2x_2 + x_3 = 6$

11) $4x_1 - x_2 - 5x_4 = 7$

12) $x_1 - 2x_2 - 3x_3 + 2x_4 - x_5 = -3$

13) $x_1 + x_2 - x_3 = 0$

14) $5x_1 + 3x_2 + x_3 = 9$

15) $6x_1 - 19x_2 + 11x_3 = 1$

16) $3x_2 + 2x_3 = 5$

Chapter 1.4

1) 2

2) 1

3) 2

4) 2

5) 3

6) 2

7) 1

8) 3

9) 3

10) 4

11) 3

12) 2

13) 5

14) 3

Chapter 1.5

1)
$$x = 2$$
$$y = -1$$

2)
$$x = 5$$
$$y = 3$$

3)
$$x = 1$$
$$y = 2$$
$$z = 3$$

4)
$$x = 2$$
$$y = -1$$
$$z = -2$$

5)
$$x = 1$$
$$y = 5$$
$$z = 1$$

6)
$$x = 2$$
$$y = -1$$
$$z = 3$$

7)
$$w = 2$$
$$x = 3$$
$$y = 0$$
$$z = 1$$

8)
$$w = -1$$
$$x = 0$$
$$y = -2$$
$$z = 2$$

Part 2

Chapter 2.1

1) 2
2) 1
3) 10
4) 36
5) 12
6) 23
7) 6
8) 59
9) 36
10) 3

Chapter 2.2

1) A) exactly one solution
2) B) no solutions or infinitely many solutions
3) B) no solutions or infinitely many solutions
4) A) exactly one solution
5) A) exactly one solution
6) A) exactly one solution
7) A) exactly one solution
8) B) no solutions or infinitely many solutions
9) A) exactly one solution
10) B) no solutions or infinitely many solutions

Chapter 2.3

1) A) exactly one solution

$x = \frac{18}{7}$

$y = \frac{5}{7}$

2) A) exactly one solution

$x = \frac{2}{7}$

$y = -\frac{13}{7}$

3) A) exactly one solution

$x = \frac{54}{59}$

$y = -\frac{19}{59}$

4) A) exactly one solution

$x = \frac{3}{88}$

$y = \frac{49}{88}$

5) A) exactly one solution

$x = -1$

$y = \frac{18}{5}$

$z = \frac{3}{5}$

6) B) no solutions or infinitely many solutions

7) A) exactly one solution

$x = \frac{35}{18}$

$y = -\frac{2}{9}$

$z = \frac{1}{2}$

8) A) exactly one solution

$x = -\frac{5}{6}$

$y = -\frac{2}{3}$

$z = \frac{7}{6}$

9) A) exactly one solution

$w = 4$

$x = -\frac{13}{4}$

$y = 1$

$z = \frac{5}{4}$

10) B) no solutions or infinitely many solution

11) A) exactly one solution

$u = \frac{1}{12}$

$w = \frac{8}{15}$

$x = -\frac{2}{5}$

$y = -\frac{19}{15}$

$z = -\frac{7}{20}$

12) A) exactly one solution

$u = -\frac{11}{15}$

$w = -\frac{13}{15}$

$x = \frac{16}{15}$

$y = -\frac{3}{5}$

$z = -\frac{17}{15}$

Chapter 2.4

1) $y = C_{-2}e^{-2x} + C_1 e^x + \left(C_2 + \frac{1}{4}x\right) e^{2x}$

2) $y = \left(C_{-1,0} + C_{-1,1}x\right) e^{-x} + C_3 e^{3x} - \frac{1}{8}e^x$

3) $y = C_1 e^x + \left(C_{-1,0} + C_{-1,1}x\right) e^{-x} - \frac{1}{4}\left(\sin x + \cos x\right)$

4) $$y = C_{-3}e^{-3x} + C_2e^{2x} + C_3e^{3x} + \frac{1}{25}\sin x + \frac{1}{50}\cos x$$

5) $$y = C_0 + (C_{1,0} + C_{1,1}x)\, e^{x} + 2x + \frac{1}{2}\cos x$$

6) $$y = \left[C_1 \sin\left(\frac{\sqrt{15}}{2}x\right) + C_2 \cos\left(\frac{\sqrt{15}}{2}x\right)\right] e^{\frac{1}{2}x} + x + \frac{1}{10}\sin x - \frac{3}{10}\cos x + C_3$$

Part 3

Chapter 3.1

1)

$$\begin{pmatrix} 3 & -2 \\ 5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

2)

$$\begin{pmatrix} 1 & -8 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

3)

$$\begin{pmatrix} 2 & 3 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$$

4)

$$\begin{pmatrix} 0 & 8 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -7 \\ 5 \end{pmatrix}$$

5)

$$\begin{pmatrix} 2 & 3 & -4 \\ 7 & -2 & 3 \\ 9 & 5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$$

6)

$$\begin{pmatrix} 1 & -1 & 1 \\ 2 & -5 & 1 \\ 1 & 4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix}$$

7)

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ -5 \end{pmatrix}$$

8)

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix}$$

9)

$$\begin{pmatrix} 1 & 1 & 1 & -1 \\ 0 & 1 & 1 & -7 \\ 1 & 0 & 8 & 0 \\ 0 & 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 11 \\ 3 \end{pmatrix}$$

10)

$$\begin{pmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -2 & 3 \\ 1 & 0 & -2 & 3 \\ 1 & -2 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

11)

$$\begin{pmatrix} 17 \\ 13 \end{pmatrix}$$

12)

$$\begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

13)

$$\begin{pmatrix} -19 \\ 1 \end{pmatrix}$$

14)

$$\begin{pmatrix} 0 \\ -12 \end{pmatrix}$$

15)

$$\begin{pmatrix} 11 \\ 1 \\ 5 \end{pmatrix}$$

16)

$$\begin{pmatrix} 16 \\ 10 \\ -3 \end{pmatrix}$$

17)
$$\begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}$$

18)
$$\begin{pmatrix} 4 \\ 28 \\ 9 \end{pmatrix}$$

19)
$$\begin{pmatrix} 14 \\ 6 \\ -1 \\ -13 \end{pmatrix}$$

20)
$$\begin{pmatrix} 20 \\ 6 \\ 5 \\ 19 \end{pmatrix}$$

Chapter 3.2

1)
$$\begin{pmatrix} 6 & 4 \\ 2 & -7 \end{pmatrix}$$

2)
$$\begin{pmatrix} 12 & 4 & 4 \\ 6 & 13 & 1 \\ 13 & 8 & 4 \\ 23 & 26 & 6 \end{pmatrix}$$

3) A) $N \times 3$
B) $2 \times N$

4)
$$\begin{pmatrix} 4 & 5 & 10 \\ 4 & 1 & 4 \end{pmatrix}$$

5)
$$\begin{pmatrix} 17 & 20 \end{pmatrix}$$

6) A) $N \times 1$
B) $2 \times N$

7)
$$\begin{pmatrix} 10 \end{pmatrix}$$

8)
$$\begin{pmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

9)

$$\begin{pmatrix} 12 & -5 \\ 15 & -10 \\ 11 & 3 \end{pmatrix}$$

10)

$$\begin{pmatrix} 24 & 8 & 17 & -5 \\ 3 & 1 & 10 & 2 \end{pmatrix}$$

Chapter 3.3

1) answers may vary; one correct answer is

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

2) answers may vary; one correct answer is

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

3) answers may vary; one correct answer is

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

4) answers may vary; one correct answer is

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

5) answers may vary; one correct answer is

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

6) answers may vary; one correct answer is

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

7) answers may vary; one correct answer is

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

8) answers may vary; one correct answer is

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

9) answers may vary; one correct answer is

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -4 & 0 \\ 0 & 1 & 0 \\ 0 & 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{8} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

10) answers may vary; one correct answer is

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{8}{3} & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -\frac{3}{14} & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{3}{7} & 1 \end{pmatrix} \begin{pmatrix} -3 & 0 & 0 \\ 0 & -\frac{28}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

11) $\det(X) = 6$

12) $\det(X) = 12$

13) $\det(X) = 2$

14) $\det(X) = \dfrac{3}{2}$

15) $\det(X) = \dfrac{3}{4}$

16) $\det(X) = -\dfrac{2}{9}$

17) $\det(X) = \pm 2$

18) $\det(X) = \pm\sqrt{\frac{3}{2}}$

Chapter 3.4

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

2)
$$\begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{pmatrix}$$

3)
$$\begin{pmatrix} -\frac{7}{5} & \frac{1}{5} \\ 1 & 0 \end{pmatrix}$$

4) not invertible

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

6)
$$\begin{pmatrix} -4 & 1 & -3 \\ -4 & 1 & -2 \\ 5 & -1 & 3 \end{pmatrix}$$

7) not invertible

8)
$$\frac{1}{24}\begin{pmatrix} 6 & -3 & -6 \\ 4 & 2 & 12 \\ -2 & 5 & -6 \end{pmatrix}$$

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

10)

$$\frac{1}{25}\begin{pmatrix} -11 & -17 & 16 & -1 \\ 14 & 8 & -9 & -1 \\ 5 & 10 & -5 & 5 \\ -26 & -22 & 31 & -16 \end{pmatrix}$$

11)

$$x = \frac{1}{5}\begin{pmatrix} 9 \\ 4 \end{pmatrix}$$

12)

$$x = \frac{1}{6}\begin{pmatrix} -7 \\ 1 \end{pmatrix}$$

13) no inverse

14)

$$x = \frac{1}{37}\begin{pmatrix} 31 \\ 7 \end{pmatrix}$$

15)

$$x = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}$$

16) no inverse

17)

$$x = \frac{1}{19}\begin{pmatrix} 12 \\ 16 \\ -5 \end{pmatrix}$$

18)

$$x = \frac{1}{3}\begin{pmatrix} 1 \\ -34 \\ 22 \end{pmatrix}$$

19)

$$x = \frac{1}{10}\begin{pmatrix} 45 \\ -9 \\ -41 \\ 29 \end{pmatrix}$$

20)

$$x = \frac{1}{19}\begin{pmatrix} -31 \\ -2 \\ 33 \\ -34 \end{pmatrix}$$

Part 4

Chapter 4.1

1) answers may vary; one correct answer is

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix}$$

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

2) answers may vary; one correct answer is

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

$$A^n = 2^{n-1} \begin{pmatrix} 2(3)^n & 3(3^n - 1) \\ 0 & 2 \end{pmatrix}$$

3) not diagonalizable

4) answers may vary; one correct answer is

$$A = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{2}{5} & -\frac{3}{5} \end{pmatrix}$$

$$A^n = 5^{n-1} \begin{pmatrix} 3 + 2^{n-1} & 3(1 - 2^n) \\ 2(1 - 2^n) & 2 + 3(2)^n \end{pmatrix}$$

5) answers may vary; one correct answer is

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix}$$

$$A^n = \frac{1}{5} \begin{pmatrix} 4(-5)^n + 5^n & 2\left[(-5)^n - 5^n\right] \\ 2\left[(-5)^n - 5^n\right] & (-5)^n + 4(5)^n \end{pmatrix}$$

6) answers may vary; one correct answer is

$$A = \begin{pmatrix} -\sqrt{3} & \sqrt{3} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2-\sqrt{3} & 0 \\ 0 & 2+\sqrt{3} \end{pmatrix} \begin{pmatrix} -\frac{\sqrt{3}}{6} & \frac{1}{2} \\ \frac{\sqrt{3}}{6} & \frac{1}{2} \end{pmatrix}$$

$$A^n = \frac{1}{6} \begin{pmatrix} 3\left[(2+\sqrt{3})^n + (2-\sqrt{3})^n\right] & 3\sqrt{3}\left[(2+\sqrt{3})^n - (2-\sqrt{3})^n\right] \\ \sqrt{3}\left[(2+\sqrt{3})^n - (2-\sqrt{3})^n\right] & 3\left[(2+\sqrt{3})^n + (2-\sqrt{3})^n\right] \end{pmatrix}$$

7) answers may vary; one correct answer is

$$A = \begin{pmatrix} -1-\sqrt{3} & -1+\sqrt{3} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -\sqrt{3} & 0 \\ 0 & \sqrt{3} \end{pmatrix} \begin{pmatrix} -\frac{\sqrt{3}}{6} & \frac{3-\sqrt{3}}{6} \\ \frac{\sqrt{3}}{6} & \frac{3+\sqrt{3}}{6} \end{pmatrix}$$

$$A^n = \frac{\sqrt{3^{n-1}}}{2} \begin{pmatrix} -1+\sqrt{3}+(1+\sqrt{3})(-1)^n & 2-2(-1)^n \\ 1-(-1)^n & 1+\sqrt{3}+(\sqrt{3}-1)(-1)^n \end{pmatrix}$$

8) not diagonalizable

9) answers may vary; one correct answer is

$$A = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{1}{2} & -\frac{i}{2} \end{pmatrix}$$

$$A^n = \frac{1}{2} \begin{pmatrix} (1+i)^n + (1-i)^n & i\left[(1-i)^n - (1+i)^n\right] \\ i\left[(1+i)^n - (1-i)^n\right] & (1+i)^n - (1-i)^n \end{pmatrix}$$

10) answers may vary; one correct answer is

$$A = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1-2i & 0 \\ 0 & 1+2i \end{pmatrix} \begin{pmatrix} -\frac{i}{2} & \frac{1}{2} \\ \frac{i}{2} & \frac{1}{2} \end{pmatrix}$$

$$A^n = \frac{1}{2} \begin{pmatrix} (1+2i)^n + (1-2i)^n & i\left[(1-2i)^n - (1+2i)^n\right] \\ i\left[(1+2i)^n - (1-2i)^n\right] & (1+2i)^n + (1-2i)^n \end{pmatrix}$$

11) answers may vary; one correct answer is

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

$$A^n = \frac{1}{2} \begin{pmatrix} 2 & 0 & 2(2^n - 1) \\ 3\left[1 - (-1)^n\right] & 2(-1)^n & 3\left[(-1)^n - 1\right] \\ 0 & 0 & -2^{n+1} \end{pmatrix}$$

12) not diagonalizable

13) answers may vary; one correct answer is

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

$$A^n = \begin{pmatrix} 3(2)^n - 2 & 2(2^n - 1) & 2(1 - 2^n) \\ 3\left[(-2)^n - 2^n\right] & 3(-2)^n - 2^{n+1} & (-2)^{n+1} + 2^{n+1} \\ 3\left[(-2)^n - 1\right] & 3\left[(-2)^n - 1\right] & 3 + (-2)^{n+1} \end{pmatrix}$$

14) not diagonalizable

Chapter 4.2

1) $a_n = \dfrac{2^n - (-1)^n}{3}$

2) $a_n = \frac{1}{2\sqrt{2}}\left[\left(1+\sqrt{2}\right)^n - \left(1-\sqrt{2}\right)^n\right]$

3) $$\begin{aligned} a_n \approx &(-0.209+0.184i)(-0.233-0.793i)^n \\ &-(0.209+0.184i)(-0.233+0.793i)^n \\ &+0.417(1.466)^n \end{aligned}$$

4) $$\begin{aligned} a_n \approx &(-0.168+0.198i)(-0.420-0.606i)^n \\ &-(0.168+0.198i)(-0.420+0.606i)^n \\ &+0.336(1.839)^n \end{aligned}$$

Chapter 4.3

1) answers may vary; one correct answer is

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

$$A^n = \begin{pmatrix} n+1 & n \\ -n & 1-n \end{pmatrix}$$

2) answers may vary; one correct answer is

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

$$A^n = \begin{pmatrix} 2n+1 & -2n \\ 2n & 1-2n \end{pmatrix}$$

3) answers may vary; one correct answer is

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

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

4) answers may vary; one correct answer is

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

$$A^n = (-1)^n \begin{pmatrix} 1-n & n \\ -n & n+1 \end{pmatrix}$$

5) answers may vary; one correct answer is

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

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

6) answers may vary; one correct answer is

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

$$A^n = \begin{pmatrix} (-1)^n(1-2n) & (-1)^n n & (-1)^n n \\ -4(-1)^n n & (-1)^n(2n+1) & (2n+1)(-1)^n - 3^n \\ 0 & 0 & 3^n \end{pmatrix}$$

7) answers may vary; one correct answer is

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

$$A^n = \frac{1}{2} \begin{pmatrix} 2 + 6n - 2n^2 & n^2 - 3n & 2n \\ 12n - 4n^2 & 2n^2 - 6n + 2 & 4n \\ -4n & 2n & 2 \end{pmatrix}$$

8) answers may vary; one correct answer is

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

$$A^n = 2^{n-3} \begin{pmatrix} 8 - 4n & -4n & 4n \\ 5n - n^2 & 8 + n - n^2 & n^2 - n \\ n - n^2 & -n^2 - 3n & n^2 + 3n + 8 \end{pmatrix}$$

9) answers may vary; one correct answer is

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

$$A^n = \frac{2^{n-4}}{3} \begin{pmatrix} 48 & 0 & 0 & 0 \\ n^3 + 3n^2 + 20n & 6n^2 - 6n + 48 & 24n & 6n - 6n^2 \\ 6n^2 + 18n & 24n & 48 & -24n \\ n^3 + 3n^2 - 4n & 6n^2 - 6n & 24n & 48 + 6n - 6n^2 \end{pmatrix}$$

10) answers may vary; one correct answer is

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

$$A^n = \begin{pmatrix} -1+2(-1)^n & 2(-1)^n - 2 & 2-2(-1)^n & 1-(-1)^n \\ -\left[1+2(-1)^n\right]n + 1 - (-1)^n & -2\left[1+(-1)^n\right]n + 2 - (-1)^n & 2\left[1+(-1)^n\right]n + (-1)^n - 1 & \left[1+(-1)^n\right]n + (-1)^n - 1 \\ -n & -2n & 2n+1 & n \\ -4(-1)^n n & -4(-1)^n n & 4(-1)^n n & (-1)^n(2n+1) \end{pmatrix}$$

Chapter 4.4

1)

$$x = \begin{pmatrix} \frac{1}{2}e^{x} - \frac{1}{2}e^{-x} \\ e^{-x} \end{pmatrix}$$

2)

$$x = \begin{pmatrix} \left(\frac{1+\sqrt{2}}{2}\right) e^{\sqrt{2}t} + \left(\frac{1-\sqrt{2}}{2}\right) e^{-\sqrt{2}t} \\ \frac{1}{2}e^{\sqrt{2}t} + \frac{1}{2}e^{-\sqrt{2}t} \end{pmatrix}$$

3)

$$x = \begin{pmatrix} \frac{1}{4}(2t-1)e^{t} + \frac{1}{4}e^{-t} \\ \frac{1}{4}(2t+1)e^{t} - \frac{1}{4}e^{-t} \\ e^{t} \end{pmatrix}$$

4)

$$x = \begin{pmatrix} \frac{2}{3}e^{2t} - \frac{2}{3}e^{t/2}\cos\left(\frac{\sqrt{3}}{2}t\right) \\ \frac{2}{3}e^{2t} + \frac{1}{3}e^{t/2}\left[\cos\left(\frac{\sqrt{3}}{2}t\right) - \sqrt{3}\sin\left(\frac{\sqrt{3}}{2}t\right)\right] \\ \frac{2}{3}e^{2t} + \frac{1}{3}e^{t/2}\left[\cos\left(\frac{\sqrt{3}}{2}t\right) + \sqrt{3}\sin\left(\frac{\sqrt{3}}{2}t\right)\right] \end{pmatrix}$$

5)

$$x = \begin{pmatrix} \frac{1}{4}\left[e^t + e^{-t} + 2\cos t\right] \\ \frac{1}{4}\left[e^t - e^{-t} - 2\sin t\right] \\ \frac{1}{4}\left[e^t + e^{-t} - 2\cos t\right] \\ \frac{1}{4}\left[e^t - e^{-t} + 2\sin t\right] \end{pmatrix}$$

6)

$$x = \begin{pmatrix} \frac{1}{4}\left[1 + e^{-2t} + 2e^t\cos t\right] \\ \frac{1}{4}\left[1 - e^{-2t} - 2e^t\sin t\right] \\ \frac{1}{4}\left[1 + e^{-2t} - 2e^t\cos t\right] \\ \frac{1}{4}\left[1 - e^{-2t} + 2e^t\sin t\right] \end{pmatrix}$$