Contents

Chapter 1

Span

1.1 Spans of Vectors

Knowing whether two vectors are linearly dependent or independent allows us to accurately describe the span of those two vectors (this expands to include any number of vectors). In the previous chapter, we saw that linear combinations of two linearly dependent vectors can only make vectors that lie on the same line as the two starting vectors. We saw this in 2D, but it also applies to 3D vectors. Consider the two vectors $\mathbf{u} = [2, 4, 3]$ and $\mathbf{v} = [4, 8, 6]$, shown in figure 1.1.

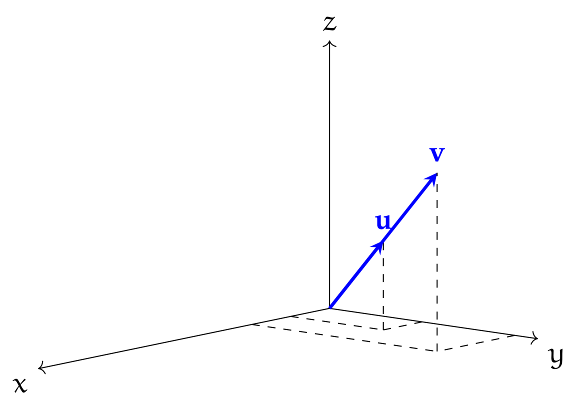


Figure 1.1: 3-dimensional vectors, **u** and **v**

Notice that these two vectors are colinear (that is, they are on the same line), therefore they are linearly dependent and any combination of **u** and **v** will lie on the same line as **u** and **v**. Therefore, we say the *the span of* **u** *and v is a line*. In fact, for any size list of linearly dependent vectors (whether it's one vector or one hundred), the span of that list is a line.

Now that you have a sense of what a span is, it is time for the formal mathematical definition. A vector span is the collection of vectors obtained by scaling and combining the original set of vectors in all possible proportions. Formally, if the set $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ contains vectors from a vector space V, then the span of S is given by:

$$\text{Span}(S) = \{a_1\mathbf{v_1} + a_2\mathbf{v_2} + \cdots + a_n\mathbf{v_n} : a_1, a_2, ..., a_n \in \mathbb{R}\} \tag{1.1}$$

This means that any vector in the $\text{Span}(S)$ can be written as a linear combination of the vectors in S.

1.1.1 Spans of Independent Vectors

What if our list of vectors aren't all linearly dependent on each other? We've seen in 2 dimensions that any two independent vectors can be linearly combined to create any vector in $\mathbb{R}^2$. So, the span is described as a *plane* (in fact, it is the entire xy-plane, which we also call $\mathbb{R}^2$). How does this expand to 3-dimensional vectors? Let's again consider two 3-dimensional vectors: $\mathbf{u} = [2, 4, 3]$ and $\mathbf{v} = [2, 1, 0]$, as shown in figure 1.2.

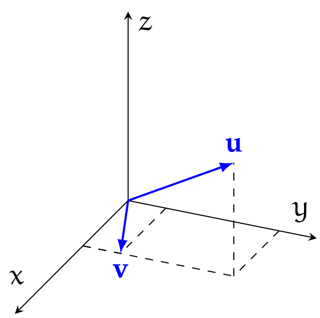


Figure 1.2: Linearly independent 3-dimensional vectors, **u** and **v**

Just like in two dimensions, any two independent vectors in $\mathbb{R}^3$ define a plane (see figure 1.3). This also applies to higher dimensions: the span of any two linearly independent vectors is a plane.

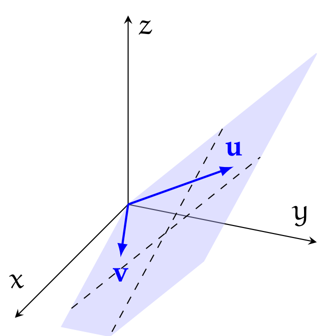


Figure 1.3: Linearly independent 3-dimensional vectors, **u** and **v**, define a plane.

If we have 3 independent vectors, then we can define a *3-dimensional space*. To understand this, first imagine a plane formed by two independent 3-dimensional vectors like in figure 1.3). If a third independent vector is introduced, it must not lie on the plane: if it did, it would be a linear combination of the first two and therefore not independent. This third vector allows us to move off the plane, and therefore all three independent vectors span $\mathbb{R}^3$.

In review, 1 vector or set of dependent vectors span a *line*, 2 vectors or sets of dependent vectors span a *plane*, and 3 vectors or sets of dependent vectors span $\mathbb{R}^3$.

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

Solution: We need to determine the number of *independent vectors*. First, we'll check if $\mathbf{r}$ and $\mathbf{s}$ are independent. They are independent if the only solution to the equation below is $a_1 = a_2 = 0$:

$$a_1 [5, 4, -6] + a_2 [0, -5, -10] = [0, 0, 0]$$

Which we can write as a system of equations:

$$5a_1 + 0a_2 = 0$$

$$4a_1 - 5a_2 = 0$$

$$-6a_2 - 10a_2 = 0$$

From the first equation, we see that $5a_1 = 0$ which implies that $a_1 = 0$. Substituting that into the second equation:

$$4(0) - 5a_2 = 0$$

$$-5a_2 = 0$$

$$a_2 = 0$$

Therefore, vectors $\mathbf{r}$ and $\mathbf{s}$ are independent. Now let's check $\mathbf{r}$ and $\mathbf{t}$:

$$a_1 [5, 4, -6] + a_2 [0, 2, 4,] = [0, 0, 0]$$

Which we can re-write as a system of equations:

$$5a_1 + 0a_2 = 0$$

$$4a_1 + 2a_2 = 0$$

$$-6a_1 + 4a_2 = 0$$

Again, from the first equation, we see that $a_1 = 0$. Substituting into the second:

$$4(0) + 2a_2 = 0$$

$$2a_2 = 0$$

$$a_2 = 0$$

Therefore, $\mathbf{r}$ and $\mathbf{t}$ are also independent. Last, we'll check $\mathbf{s}$ and $\mathbf{t}$ for independence:

$$a_1 [0, -5, -10] + a_2 [0, 1, 2,] = [0, 0, 0]$$

The system of equations:

$$0a_1 + 0a_2 = 0$$
$$-5a_1 + a_2 = 0$$
$$-10a_1 + 2a_2 = 0$$

The first equation doesn't tell us anything, since it would be true no matter what a_1 and a_2 are. We can solve the second equation for a_2 and substitute into the third equation:

$$a_2 = 5a_1$$
$$-10a_1 + 2(5a_1) = 0$$
$$-10a_1 + 10a_1 = 0$$

Which is also true for all a_1. In fact, there are many solutions to $a_1 [0, -5, -10] + a_2 [0, 1, 2,] = [0, 0, 0]$, $a_1 = 1$ and $a_2 = 5$ is an example. Therefore, **s** and **t** are *dependent*. So, we really have 2 independent vectors in the list, and therefore span($\mathbf{r}, \mathbf{s}, \mathbf{t}$) is a plane.

Exercise 1 Determining Span

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

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

Working Space

Answer on Page 61

1.2 Where to Learn More

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

The Wolfram Demonstrations website has a fun, interactive demo where you can enter values for 2D and 3D matrices and see how the area or volume changes. `https://demonstrations.wolfram.com/DeterminantsSeenGeometrically/#more`

If you are curious about the *Expansion of Minors*, see: `https://mathworld.wolfram.com/DeterminantExpansionbyMinors.html`

Chapter 2

Determinants and Inverse Matrices

2.1 Determinants

Checking the independence of multitudes of vectors may take an immense amount of time. What if you had a list of 5, 10, or even 100 vectors? The determinant of a matrix is a scalar value that also indicates whether the columns of a matrix are linearly independent. So, if you put all your vectors together in a matrix and take the determinant of that matrix, the result will tell you if all the vectors are independent or not. For a 2D matrix, the determinant is the area of the parallelogram defined by the column vectors. For a 3D matrix, the determinant is the volume of the parallelepiped (a six-dimensional figure formed by six parallelograms, such as a cube).[1]

Let's plot the parallelogram for this matrix (see figure 2.1):

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

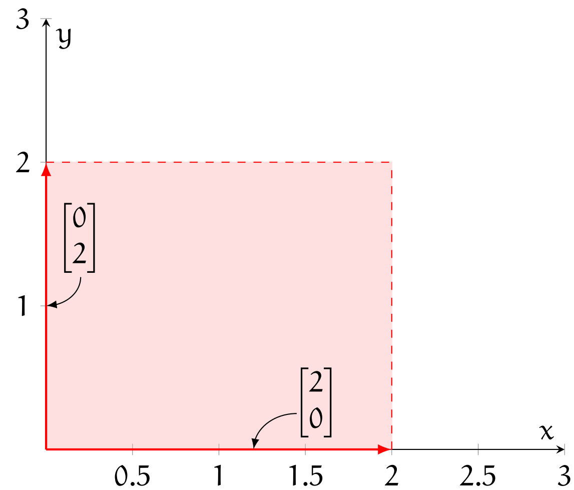


Figure 2.1: A parallelogram constructed from vectors $[2, 0]$ and $[0, 2]$

[1]Note that determinants can only be found for square, $n \times n$ matrices.

2 by 2 Determinant

The formal definition for calculating the determinant of a 2 by 2 matrix A is:

$$\det(A) = (a \cdot d) - (b \cdot c)$$

where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

For the matrix plotted above, the determinant is $(2 * 2) - (0 * 0)$. You can also see that 4.0 is the area, base (2) times height (2).

You can use the determinant to see what happens to a shape when it goes through a linear transformation. Let's scale the 2 by 2 matrix by 4:

$$\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}$$

Plot it (see figure 2.2):

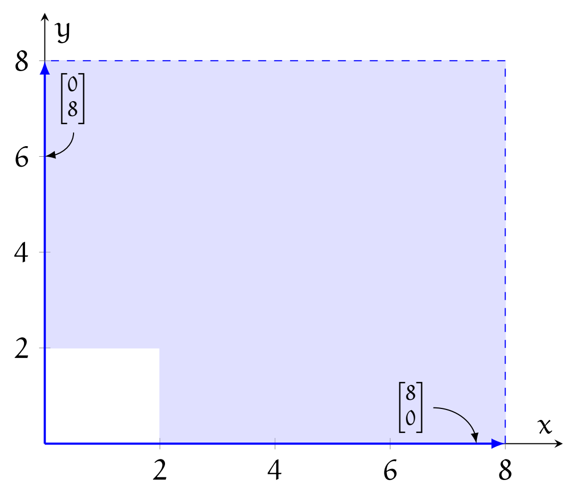


Figure 2.2: Scaling the matrix also scales the parallelogram.

Find the determinant using $(8 * 8) - (0 * 0) = 64$

You can see that scaling the matrix scaled the area by the scaling factor squared (see figure 2.3).

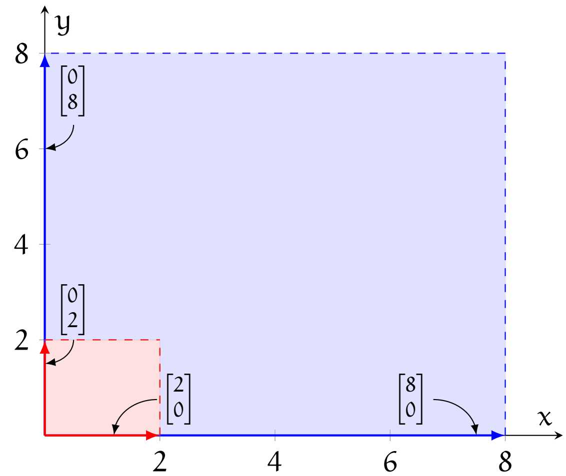


Figure 2.3: Scaling a matrix by a constant c increases the area of the parallelogram by a factor of c^2.

We can show why this is true mathematically. Suppose we have a 2 by 2 matrix A:

$$A = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

Then $\det(A) = wz - xy$. We can scale this matrix by a constant, c:

$$cA = c \cdot \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} cw & cx \\ cy & cz \end{bmatrix}$$

And we can take the determinant:

$$\det(cA) = \det\left(\begin{bmatrix} cw & cx \\ cy & cz \end{bmatrix}\right) = cw(cz) - cx(cy) = c^2(wz - xy) = c^2 \cdot \det(A)$$

Therefore, scaling a 2 by 2 matrix by a factor changes the determinant by that factor squared. What about higher dimensions? If each side of a cube were scaled by a factor of c, then the volume of the cube would change by a factor of c^3 (feel free to confirm this yourself). And if a tesseract (a four-dimensional cube) had each side scaled by a factor of c, then the hypervolume (four-dimensional volume) would be scaled by a factor of c^4. Do you notice a pattern?

In fact, scaling an $n \times n$ matrix by a constant factor, c, changes the determinant of that $n \times n$ matrix by a factor of c^n.

What happens if the columns of a matrix are not independent? Let's plot this matrix (see figure 2.4):

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

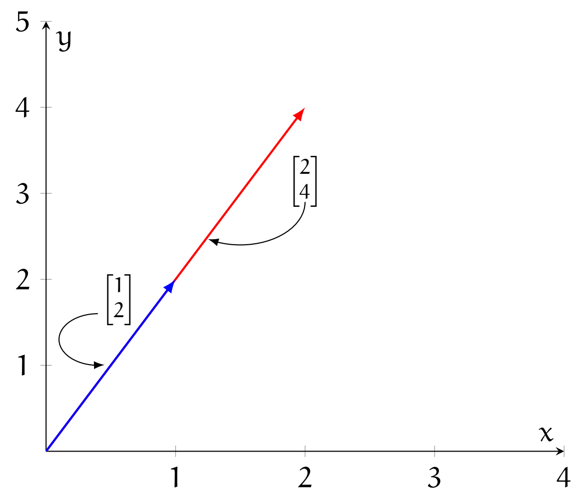


Figure 2.4: The vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ are co-linear, so there is no area between them and the determinant of $\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$ is zero.

One vector overwrites the other. As you can see, the area is 0 because there is no space between the vectors. Therefore, the columns of the matrix are linearly dependent.

Exercise 2 Finding the Determinant

Plot the parallelogram represented by the columns of the matrix. What is the area of this parallelogram?

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

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

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

Working Space

Answer on Page 61

Calculating the determinant for a 2 by 2 matrix is easy. For a larger matrix, finding the determinant is more complex and requires breaking down the matrix into smaller matrices until you reach the 2x2 form. The process is called expansion by minors. For example,

3×3 Determinant

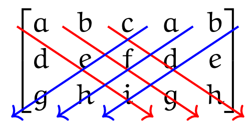
The determinant of a 3 by 3 matrix is found by

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \begin{bmatrix} d & e \\ g & h \end{bmatrix}$$

A trick for this is rewriting the matrix as a 3×5 augmented matrix with the first 2 columns after the third column:

$$B = \begin{bmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{bmatrix}$$

and solving the *down-right diagonals* minus the *down-left diagonals*:



Down-right diagonals (add)   Down-left diagonals (subtract)

$$\det(B) = (a \cdot e \cdot i) + (b \cdot f \cdot g) + (c \cdot d \cdot h) - [(c \cdot e \cdot g) + (a \cdot f \cdot h) + (b \cdot d \cdot i)]$$

Note that this is the same multiplication as above, just formatted differently.

As you can see, this involves a recursive process of breaking a larger matrix into a smaller 2×2 matrix.

For our purposes, we simply want to first check to see if a matrix contains linearly independent rows and columns before using our Python code to solve.

2.2 Determinants in Python

Modify your code so that is uses the $np.linalg.det()$ function. If the determinant is not zero, then you can call the $np.linalg.solve()$ function. Your code should look like this:

```
# Are the rows and columns independent?
# Equivallently, is the determinant 0?
if (np.linalg.det(D) != 0):
    j = np.linalg.solve(D,e)
    print(j)
else:
    print("Rows and columns are dependent.")
```

How does this work below the hood? Let's also write a recursive python function that finds our determinant:

There are two base cases:

- The matrix is of size 1×1
- The matrix is of size 2×2

And further sizes can be simplified into one of the base cases by *cofactor expansion*. The idea behind cofactor expansion is to break a big determinant into smaller ones until we reach cases we already know how to solve. Formally, this is written as

$$|A| = \sum_{j=1}^{n} (-1)^{(i+j)} a_{(ij)} M_{(ij)}$$

We do this as follows:

1. Pick the first row of the matrix.
2. For each entry in that row:
   - Remove the row and column containing that entry.
   - This creates a smaller matrix, called a *minor*, which are submatrices. Recall what we did above for our 3×3 determinant definition.
3. Compute the determinant of that smaller matrix.
4. Multiply the original entry (where the row and column lines originate from), the determinant of its minor, and an alternating sign $+, -, +, \dots$.
5. Sum all of these results together.

This process repeats recursively: each smaller determinant is computed the same way, until we reach one of the base cases.

If we format our matrices as a nested array, we can use python's indexing to check and reduce the matrices. Take a look at this recursive determinant program:

```
def recursive_determinant(matrix):
    """
    Compute the determinant of a square matrix recursively.
    matrix: list of lists, forming a matrix of size n by n
    """
    n = len(matrix)

    # base case 1x1:
    if n == 1:
        return matrix[0][0]
    # base case 2x2
    if n == 2:
        return matrix[0][0] * matrix[1][1] - matrix[0][1] * matrix[1][0]

    det = 0

    for col in range(n):
        # build a minor matrix:
        minor = []
        # for every row in the matrix
        for row in matrix[1:]:
```

```
            # remove column col and store as new_row
            new_row = row[:col]+row[col+1:]
            # apprend new row to the minor matrix
            minor.append(new_row)

        sign = (-1) ** col #exponential for alternating sign
        det += sign * matrix[0][col] * recursive_determinant(minor)
    return det
```

We implemented our known base case, and recursively reduce our array until it fits a known base case. This is exactly how numpy's `np.linalg.det()` works.

2.3 Inverse Matrices

Now that we have talked about determinants, we can talk about Inverse Matrices. The idea of a matrix inverse is a natural generalization of the multiplicative inverse of a real number. For example, since

$$3 \cdot \frac{1}{3} = 1,$$

the number $\frac{1}{3}$ is called the multiplicative inverse of 3.

Similarly, for matrices, we define an inverse in terms of matrix multiplication and the identity matrix. If A is an $n \times n$ matrix, its inverse (when it exists) is denoted by A^{-1} and satisfies

$$AA^{-1} = A^{-1}A = I_n,$$

where I_n is the $n \times n$ identity matrix.

Inverse Matrix

An inverse matrix is a square matrix that satisfies the following property:

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

where A is the original matrix, B is the inverse matrix, and I_n is the identity matrix of size $n \times n$.

When such a matrix B exists, it is denoted by A^{-1}, and is said to be *invertible*. Note that sometimes an inverse matrix may not exist.

It is important to note that not every matrix has an inverse. In particular, only *square matrices* can be invertible, and even among square matrices, an inverse may fail to exist.

2.3.1 Existence of Square Matrices

Existence of the Inverse

A square matrix A has an inverse if and only if its determinant is nonzero:

$$A^{-1} \text{ exists if and only if } \det(A) \neq 0.$$

If

$$\det(A) = 0,$$

then A is called a **singular matrix**, and no inverse exists. Equivalently, an inverse matrix does not exist when the rows or columns of A are *linearly dependent*. As discussed in the chapter on linear dependence, this occurs when one row (or column) is a scalar multiple of another, or when a row (or column) can be written as a linear combination of the others.

Linear dependence implies that the matrix does not contain enough independent information to reverse its action, making an inverse impossible.

2.3.2 Finding an Inverse Matrix

2×2 Inverse

There are several methods for finding the inverse of a matrix. For small matrices, especially 2×2 matrices, there is a direct formula. For larger matrices, a more systematic method using row operations is required.

Inverse of a 2×2 Matrix

A trick for a 2×2 Matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is

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

Note that we multiply by the a new matrix has a and d swapped, and b and c switch signs. Finally, we divide by the determinant of the original A.

This formula is valid only when the determinant of A is nonzero. Recall that the determinant of a 2×2 matrix is

$$\det(A) = ad - bc.$$

If $ad - bc = 0$, then A is singular and has no inverse.

Exercise 3 2×2 Practice

Find the inverse of the matrix of

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

Working Space

Answer on Page 63

$n \times n$ Inverse

For matrices of size 3×3 or greater, we create an augmented matrix.

Recall that we can create an augmented matrix by writing two matrices directly next to each other.

To find the inverse of a square matrix A:

1. Form the augmented matrix $[A \mid I_n]$, where I_n is the identity matrix.
2. Use elementary row operations to transform the left side into I_n.
3. If this is possible, the right side of the augmented matrix becomes A^{-1}.

Symbolically, this process can be written as

$$[A \mid I_n] \longrightarrow [I_n \mid A^{-1}].$$

If row reduction does not result in the identity matrix on the left side, then A does not have an inverse. In this case, the matrix is singular.

Exercise 4 A bigger inverse

Working Space

Find the inverse of the matrix

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

Check your answer using $AA^{-1} = I_3$

Answer on Page 63

Exercise 5 Does a matrix exist?

Working Space

Given the matrix:

$$\begin{bmatrix} 1 & 2 & 4 & 8 \\ 2 & 4 & 8 & 16 \\ 4 & 8 & 16 & 32 \\ 8 & 16 & 32 & 64 \end{bmatrix}$$

Does an inverse matrix exist? Explain why or why not

Answer on Page 64

The augmented matrix row-reduction method not only provides a way to compute inverses, but also offers another criterion for invertibility: a matrix is invertible if and only if it can be row-reduced to the identity matrix.

2.3.3 Relation to $A\vec{x} = \vec{b}$

An inverse matrix is not only for square matrices, but also for systems of equations and our fundamental linear algebra equation $A\vec{x} = \vec{b}$. Let's look at an example.

Given

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

solve for $\vec{x}$ by first finding the inverse.

First, we can interpret our givens as a systems of matrices:

$$A\vec{x} = \vec{b} \quad \Longleftrightarrow \quad \begin{cases} 2x_1 + x_2 = 5 \\ x_1 + x_2 = 3 \end{cases}$$

Checking that an inverse matrix first exists, we have

$$\det(A) = 2(1) - 1(1) = 1 \neq 0$$

So we know an inverse matrix exists! To find A^{-1}

$$A^{-1} = 1 \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

Here is the fun part! Let's do some computational equivalences:

$$\begin{aligned} A\vec{x} &= \vec{b} \\ A^{-1}(A\vec{x}) &= A^{-1}\vec{b} \\ I\vec{x} &= A^{-1}\vec{b} \\ \vec{x} &= A^{-1}\vec{b} \end{aligned} \tag{2.1}$$

Equation (2.1) show that, by only multiplying by the matrix equivalent of 1, the identity matrix I, we can state that $\vec{x} = A^{-1}\vec{b}$. Since we have both A^{-1} and $\vec{b}$, we can find $\vec{x}$:

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

Geometrically, this did a very simple operation. We used the fact that A vectors transforms $\vec{x}$ into $\vec{b}$. Then we noted that A^{-1} undoes that transformation, recovering our original $\vec{x}$ by finding $A^{-1}\vec{b}$. This may be hard to visualize, so don't worry if it is hard to grasp immediately. We will review it again in a following chapter.

2.3.4 Matrix Inverse using NumPy

We can use our Python library NumPy to find inverses a lot easier using the `np.linalg.inv` function to find the inverse of a given array.

```
import numpy as np

A = np.array([[2, 1],
              [1, 1]])

A_inv = np.linalg.inv(A)
print(A_inv)
```

This outputs

```
[[ 1. -1.]
 [-1.  2.]]
```

If given a matrix that is *linearly dependent*, NumPy will raise an exception:

```
raise LinAlgError("Singular matrix")
numpy.linalg.LinAlgError: Singular matrix
```

letting us know that the inverse does not exist.

To improve our code, we check for the determinant of the matrix beforehand. In this code, we introduce a tolerance as computers handle numbers near zero as very small but non-zero digits.

```
import numpy as np

A = np.array([[2, 1],
              [4, 2]], dtype=float)

# BAD PRACTICE:
# A_inv = np.linalg.inv(A)
# print(A_inv)

detA = np.linalg.det(A)

tolerance = 1e-12

if abs(detA) < tolerance:
    print("Matrix is singular or nearly singular. No reliable inverse.")
else:
```

```
    A_inv = np.linalg.inv(A)
    print("Determinant:", detA) # you may see a large floating point number
    print("Inverse:\n", A_inv)
```

Output:

```
Matrix is singular or nearly singular. No reliable inverse.
```

2.4 Summary

In this chapter, we established the determinant of a matrix and inverse matrix.

The determinant of a matrix provides a powerful geometric and algebraic description of how a matrix acts on space.

In two dimensions, the determinant of a 2×2 matrix represents the signed area of the parallelogram formed by the images of the standard basis vectors. In higher dimensions, the determinant represents signed volume.

Further, if a matrix A multiplies a region in space, then $|\det(A)|$ describes the factor by which areas or volumes are scaled, noting the sign and absolute value of the determinant as a scalar.

The inverse of a matrix *reverses the effect of the original matrix*. If A is invertible, then multiplying by A^{-1} restores any vector to its original position:

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

Thus, invertibility, nonzero determinant, linear independence of rows and columns, and reversibility of action are all different ways of describing the same underlying property.

In the next chapter, we will interpret matrices as *functions* that transform vectors.

This idea will allow us to visualize matrix multiplication, understand invertibility deeper, and connect algebraic properties such as determinants to transformations of space.

The existence of an inverse matrix is closely related to whether a matrix represents a reversible transformation.

The next chapter will be very graph and program heavy, as we expand the geometric properties of determinants, inverses, and matrices.

Chapter 3

Cross Product

3.1 The Cross Product

Similar to the dot product, we can also take the *cross product* of two vectors. While the dot product measures how closely the two vectors *align* and creates a *scalar*, the cross product measures how much two vectors *differ in direction*, producing a *vector* that is perpendicular to both vectors. The resulting vector's magnitude is equal to the area of the parallelogram spanned by the two input vectors. Before defining the cross product, recall the determinant of a 3×3 matrix. The determinant measures the *signed* volume scaling factor of the linear transformation represented by the matrix. If the sign is negative, orientation is reversed, as if flipping over a piece of paper.

The cross product formula often uses the notation of a 3×3 determinant as a compact way to organize its components. However, these were harder to solve without using the expansion by minors. It is important to understand that, in the case of the cross product, the "determinant" we write includes *unit vectors* in the first row; this is a memory device, not an actual determinant in the strict algebraic sense. However, understanding determinants is important before approaching the cross product.

Given two vectors in $\mathbb{R}^3$,

$$\vec{\mathbf{a}} = \langle a_x, a_y, a_z \rangle, \quad \vec{\mathbf{b}} = \langle b_x, b_y, b_z \rangle,$$

their *cross product* is defined as

$$\vec{\mathbf{a}} \times \vec{\mathbf{b}} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}$$

Expanding this yields:

$$\vec{\mathbf{a}} \times \vec{\mathbf{b}} = (a_y b_z - a_z b_y)\,\mathbf{i} - (a_x b_z - a_z b_x)\,\mathbf{j} + (a_x b_y - a_y b_x)\,\mathbf{k}.$$

where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are unit vectors.

3.1.1 Geometric Interpretation

The cross product $\vec{\mathbf{a}} \times \vec{\mathbf{b}}$ has the following properties:

1. It is perpendicular to both $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, making it *orthogonal* as well (orthogonality can be thought of as 3D perpendicularity).
2. Orthogonality implies $\vec{\mathbf{a}} \times \vec{\mathbf{b}} = \vec{\mathbf{c}}$, $\vec{\mathbf{a}} \cdot \vec{\mathbf{c}} = 0$ and $\vec{\mathbf{b}} \cdot \vec{\mathbf{c}} = 0$.
3. Its magnitude is
$$\|\vec{\mathbf{a}} \times \vec{\mathbf{b}}\| = \|\vec{\mathbf{a}}\| \, \|\vec{\mathbf{b}}\| \sin\theta,$$
where θ is the angle between $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.
4. The magnitude equals the area of the parallelogram spanned by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$.
5. Two parallel vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ result in the zero vector, $\vec{\mathbf{0}}$.

FIXME diagram here

Exercise 6 Using the cross product

Find the cross product of the vectors $\vec{v_1} = \langle 4, 5, 6 \rangle$ and $\vec{v_2} = \langle 3, 7, -8 \rangle$

Working Space

Answer on Page 64

Chapter 4

Matrices as Transformations

Recall that a *function*, informally, is a rule that takes an input and produces an output.

For example, we know the common function $f(x) = x^3$ takes the input and cubes it, such that the result is $x \times x \times x$. Inputting $x = 3$ outputs $f(3) = 27$.

In Linear Algebra, we can think of matrices as a type of function. Recall our systems of matrices equation,

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

We can rearrange this to look closer to function notation:

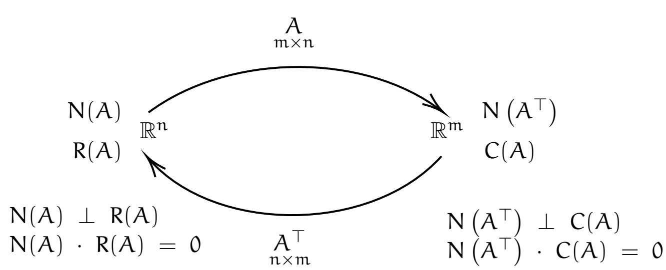
$$\vec{b} = A\vec{x}$$

We can change our input vector $\vec{x}$, which directly affects the output variable $\vec{b}$.

Recall our subspace diagram; the matrix A, which is size $m \times n$, takes the input vectors $\vec{x} \in \mathbb{R}^n$, and transforms them to the output $\vec{b} \in \mathbb{R}^m$.

The transformation T is said to map from the reals of n to the reals of m, such that: $T : \mathbb{R}^n \to \mathbb{R}^m$

- $\mathbb{R}^n$ is called the **domain**
- $\mathbb{R}^m$ is called the **codomain**



Let's call the transformation T, such that

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

We can rewrite our transformation with matrix A

$$T(\vec{x}) = A\vec{x}$$

A matrix transformation is completely determined by where it sends the *standard basis vectors*. Each column of A is the image of a basis vector under the transformation, and every output vector is a linear combination of these columns.

Consequently, the column space of A represents the full set of possible outputs of the transformation.

4.0.1 The Identity Transformation

For this chapter, we will restrict our transformations to $\mathbb{R}^2 \to \mathbb{R}^2$, such that our matrix is 2×2. This will help calculations be simpler and also will allow us to make easy to follow transformation diagrams.

Let's take the simplest transformation; the *identity transformation*. Take the matrix:

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

Every vector in the subspace of $\mathbb{R}^n$ **stays where it is**.

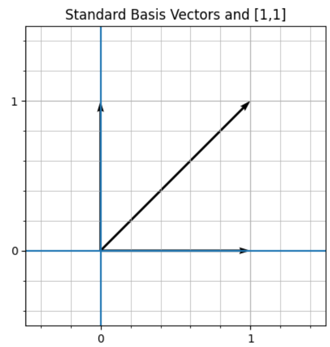
- e_1 stays e_1
- e_2 stays e_2

Every vector stays where it originally started in this transformation. In a way, no transformation is truly applied.

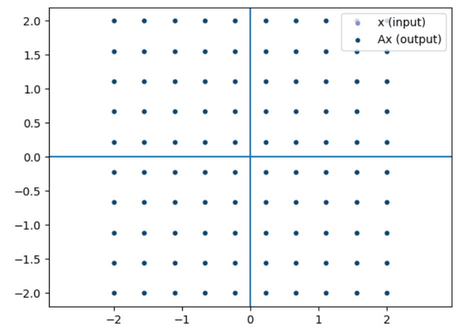
Under the identity transformation, nothing moves. We still get something very important out of this: the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ can be written as a combination of the basis vectors:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1\vec{e}_1 + 1\vec{e}_2$$

When a linear transformation is applied, the way a vector is built from the basis vectors does not change, but *rather the direction of the basis vectors*.



(a) Identity vectors before transformation



(b) Identity vectors after transformation (no change)

Figure 4.1: An identity transformation acting on basis vectors.

A linear transformation acts on $\vec{x}$ by acting on each basis vector individually:

$$A\vec{x} = x_1 A\vec{e}_1 + x_2 A\vec{e}_2$$

such that any coefficient x_1 and x_2 are unchanged, while e_1 and e_2 are transformed.

Our chapter will also take a look at transforming a simple house. Here is the house before any transformations (or, with the identity transformation applied):

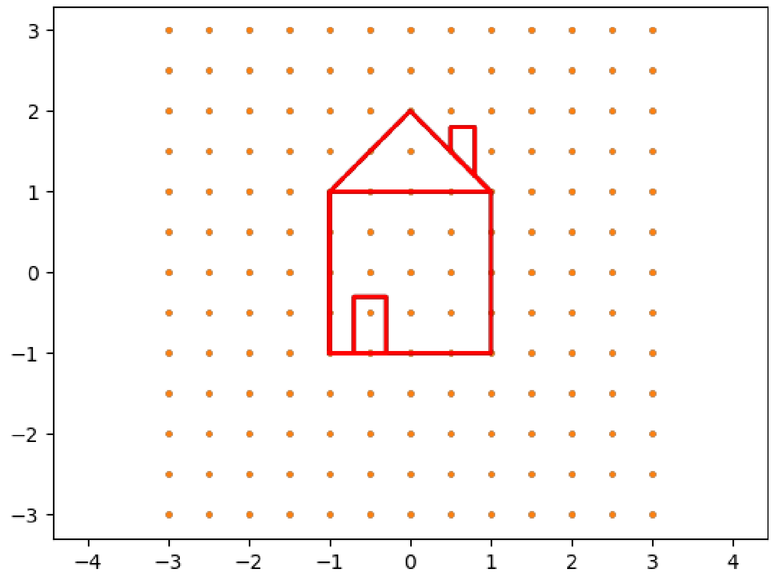


Figure 4.2: A simple house before any transformations.

Linear transformations preserve linear combinations while *altering the directions that define the coordinate system*.

In the next example, we will apply a non-identity matrix and observe how transforming the basis vectors reshapes the entire space. Each figure will show a grid of points in $\mathbb{R}^2$ alongside the resulting transformed points after applying the matrix transformation. We encourage you to play with the python script `transforms.py` to see how different matrices affect the transformation. Also, we will limit our view to a square region from $(-2, -2)$ to $(2, 2)$ and only use 2×2 matrices for simplicity.

4.0.2 Scaling and Shearing

4.1 Scaling

Now, let's try scaling up and down the basis vectors. Consider the matrix:

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

This matrix will scale both the e_1 and e_2 vectors by a factor of 2, effectively doubling their lengths.

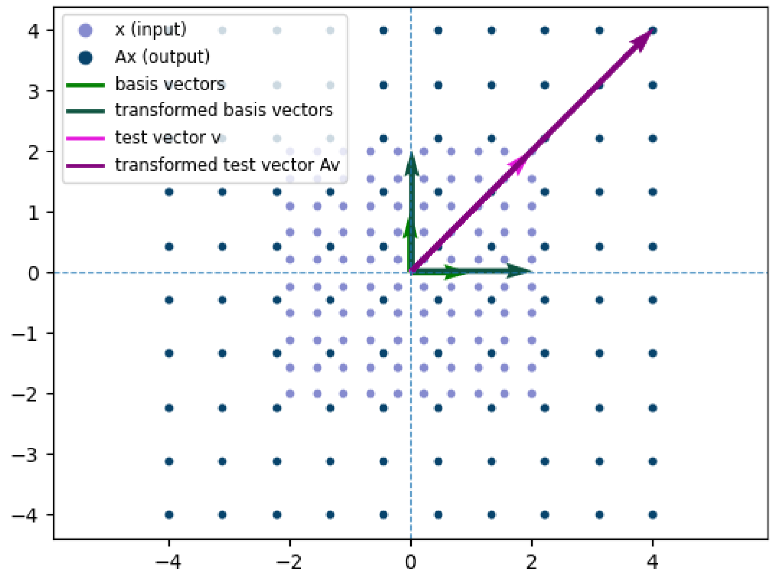


Figure 4.3: The effect of the scaling transformation $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ on a grid of points.

What happens to our house? Well, each part of the house is transformed linearly by 2, a linear *scaling*.

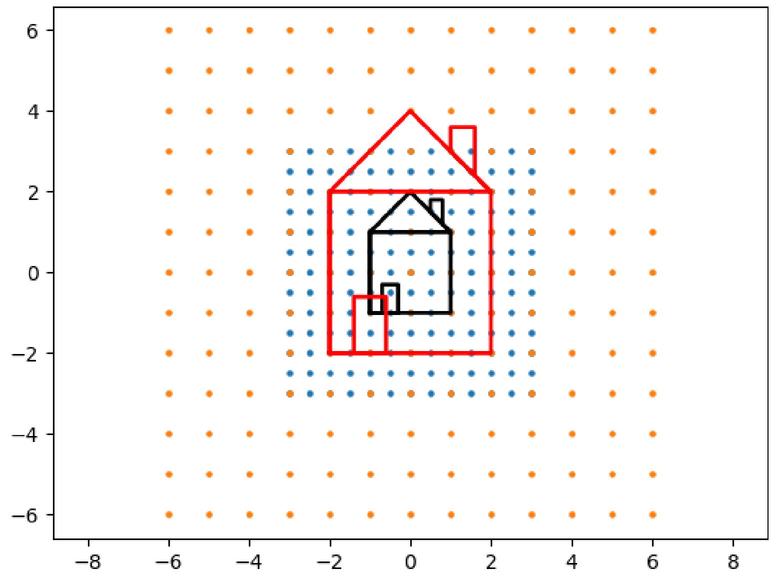


Figure 4.4: The effect of the scaling transformation $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ on a simple house.

This tells us that the transformation A is a scaling transformation, which stretches the space by a factor of 2 in both directions. What do you think happens if we scale by a factor of $\frac{1}{2}$ instead? Try it out!

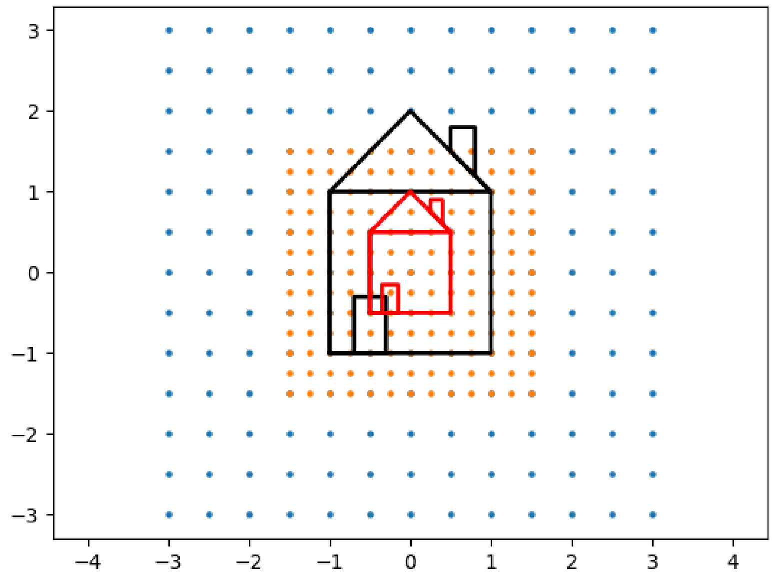


Figure 4.5: The effect of the scaling transformation $A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$ on a simple house.

This, as expected, shrinks the house by a factor of $\frac{1}{2}$ in both directions.

Now, let's try scaling up *only one* of the basis vectors. Consider the matrix:

$$A_{\text{rightward shear}} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

This matrix will scale the e_1 vector by a factor of 1 (no change) and the e_2 vector by a factor of 2 (doubling its length).

Applying this transformation to the basis vectors produces this result:

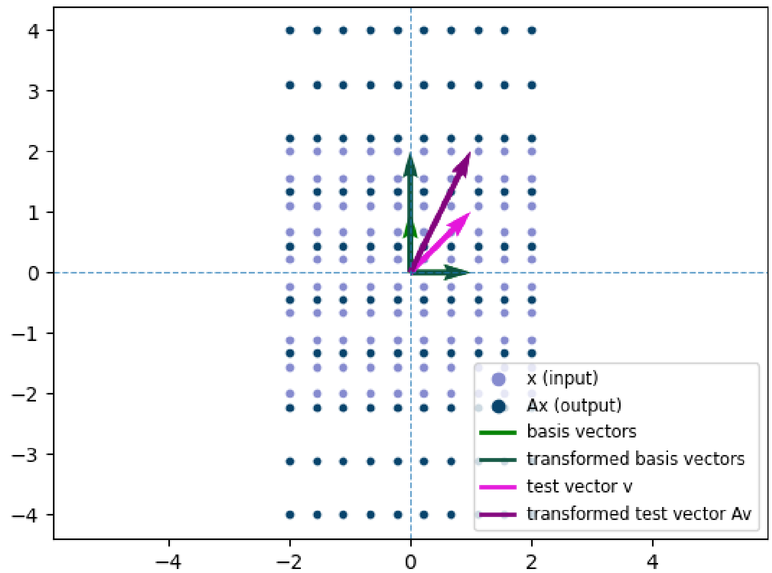


Figure 4.6: The effect of the scaling transformation $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ on the standard basis vectors.

Applying this transformation to our house:

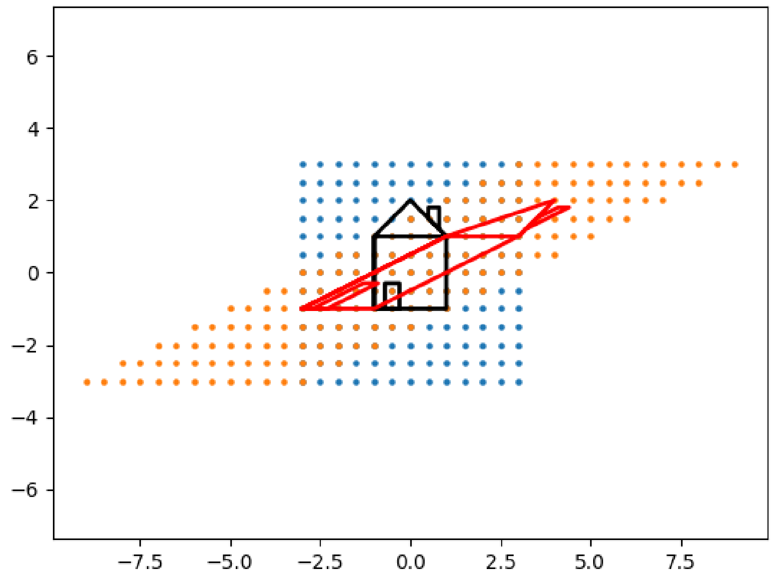


Figure 4.7: The effect of the shearing transformation $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ on a simple house.

4.2 Reflections

How would we do reflections across a given axis? Let's think about what happens to each point. Reflecting individual points is the same as reflecting the basis vectors, since every point is a linear combination of the basis vectors.

4.2.1 Reflection across the x-axis

For a reflection across the x-axis, the y-values of each point change sign, while the x-values stay the same.

$$(x, y) \mapsto (x, -y)$$

This means that e_1 stays e_1, while e_2 becomes $-e_2$. This allows us to rewrite a 2×2 matrix for this transformation:

$$A_{x-axis} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Below is a pink test vector $\begin{bmatrix}1\\1\end{bmatrix}$ and its transformation in purple, which is $\begin{bmatrix}1\\-1\end{bmatrix}$.

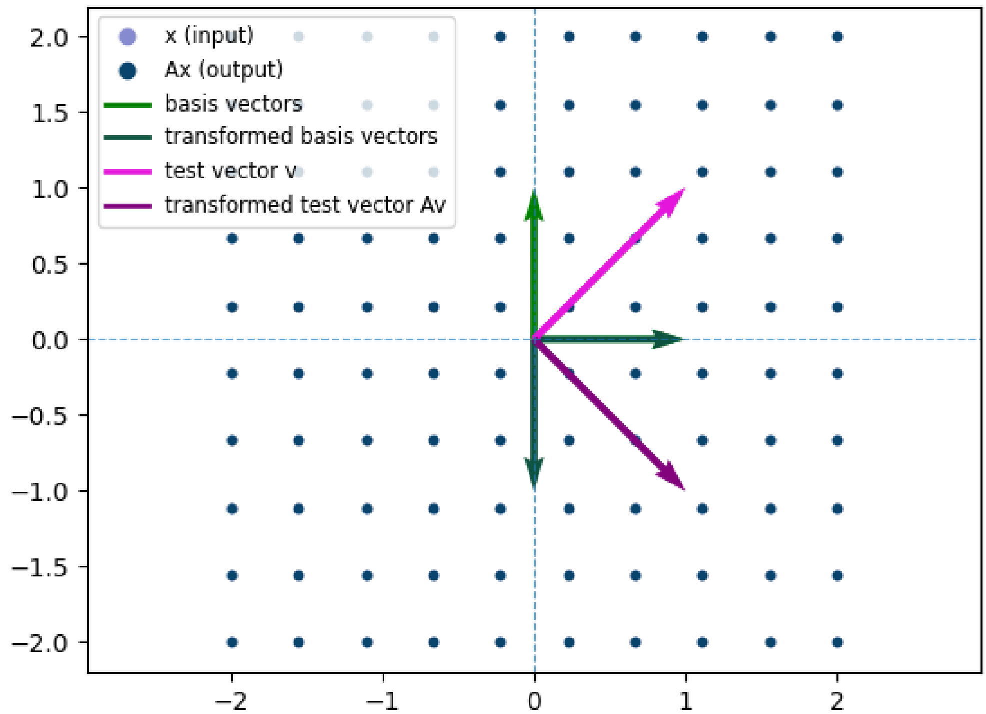


Figure 4.8: The effect of reflecting across the x-axis. Added is the vector $\begin{bmatrix}1\\1\end{bmatrix}$ and its transformation in purple.

Also, we can reflect our house across the x-axis. We will see a difference in where the chimney, roof, and door are located.

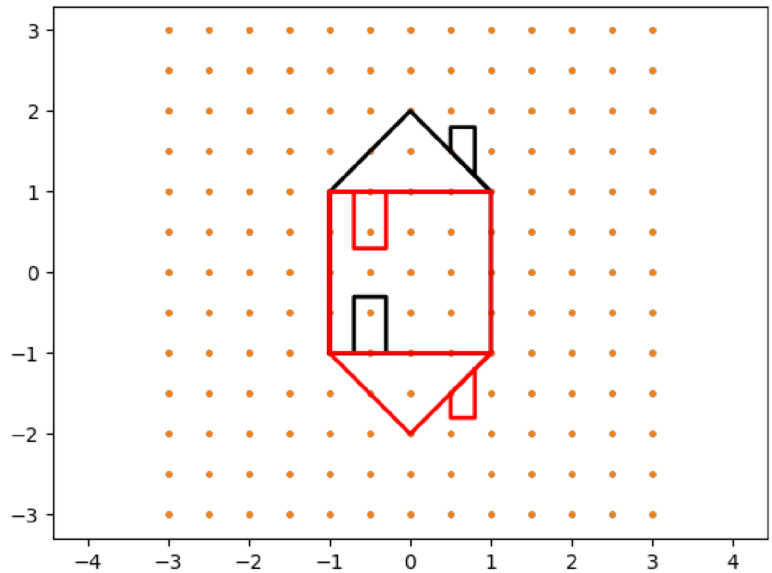


Figure 4.9: The effect of reflecting a house across the x-axis. Red house is the transformed house, while the black house is the original.

4.2.2 Reflection across the y-axis

What if we want to reflect across the y-axis instead? In this case, the x-values of each point change sign, while the y-values stay the same. As a function of points, this is

$$(x, y) \mapsto (-x, y)$$

Our basis vector e_1 becomes $-e_1$, while e_2 stays e_2. This allows us to rewrite a matrix for this transformation:

$$A_{y-axis} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

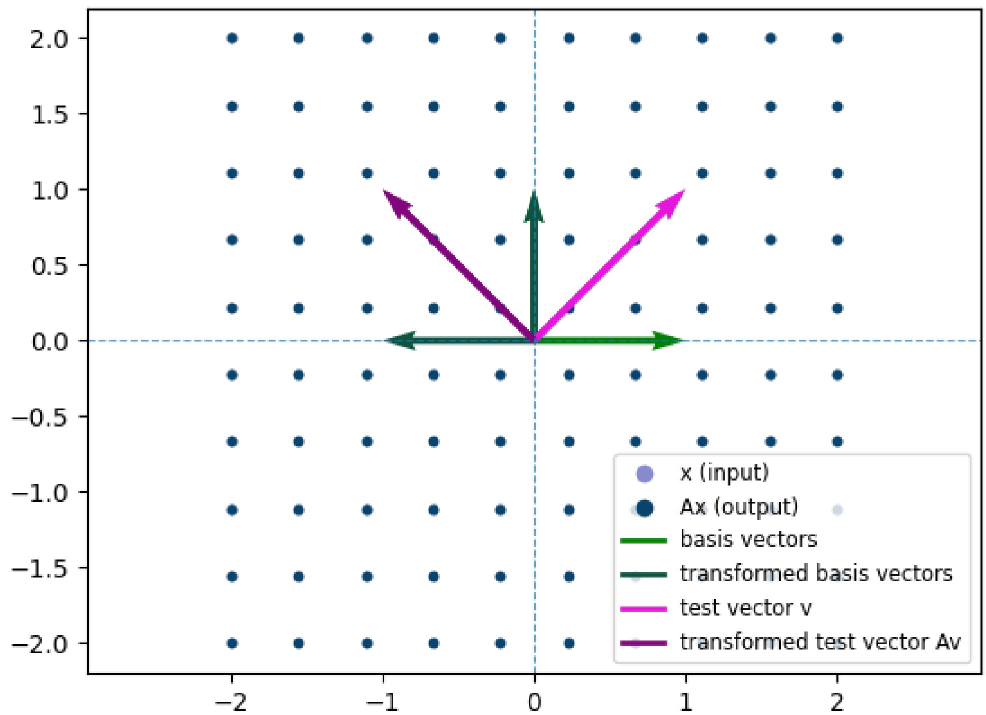


Figure 4.10: The effect of reflecting across the y-axis. Added is the vector $\begin{bmatrix}1\\1\end{bmatrix}$ and its transformation in purple.

Notice that it is simlar to the effect of a selfie, where the left and right sides are flipped. This may be more noticable in the house transformation, where the door and chimney are on the opposite sides of the house.

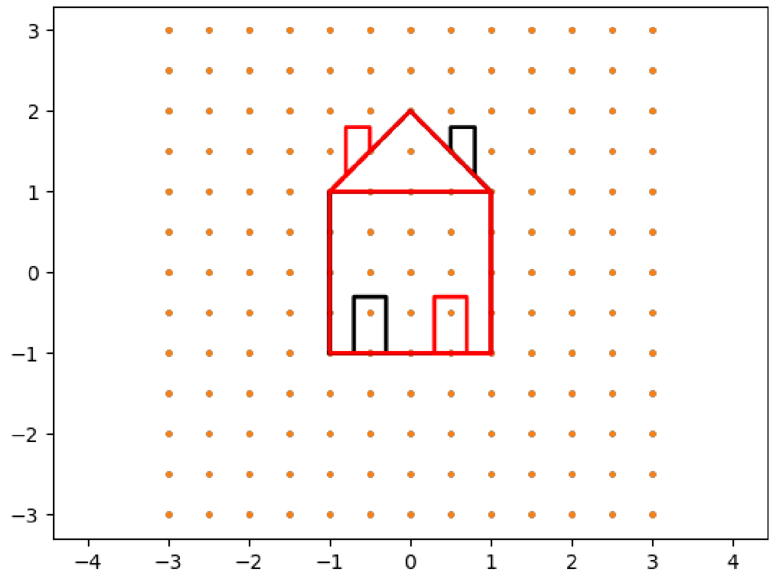


Figure 4.11: The effect of reflecting a house across the y-axis.

Note that the red house is still upright, but the left and right sides are flipped. This makes sense, since the y-values of the points are unchanged, while the x-values flip signs.

4.2.3 Reflection across the lines $y = x$ and $y = -x$

As we have seen, there are two standard lines important for inverse functions: $y = x$ and $y = -x$. We can also reflect across these lines.

The line $y = x$ is the line where the x and y values are equal. When we reflect across this line, the x and y values swap places. As a function of points, this is

$$(x, y) \mapsto (y, x)$$

Since each point swaps places, the basis vectors also swap places. e_1 becomes e_2, while e_2 becomes e_1. This allows us to rewrite a matrix for this transformation:

$$A_{y=x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Let's look at our house picture, reflected across the line $y = x$. Notice that the house is flipped across the line, such that the roof is now on the bottom and the door is on top.

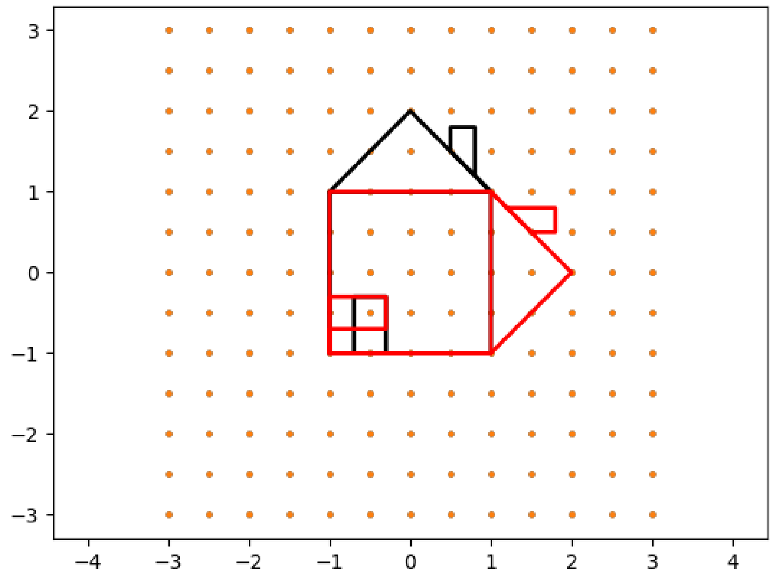


Figure 4.12: The effect of reflecting a house across the line $y = x$.

The line $y = -x$ is the line where the x and y values are equal but with opposite signs. When we reflect across this line, the x and y values swap places and flip signs. As a function of points, this is

$$(x, y) \mapsto (-y, -x)$$

Not only do the x and y values swap places, but they also flip signs. This means that e_1 becomes $-e_2$, while e_2 becomes $-e_1$. This allows us to rewrite a matrix for this transformation, where

$$A_{y=-x} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

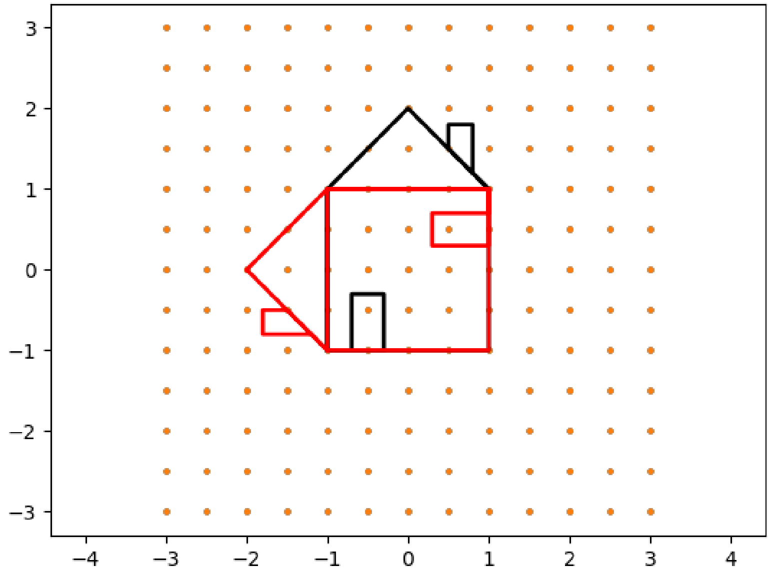


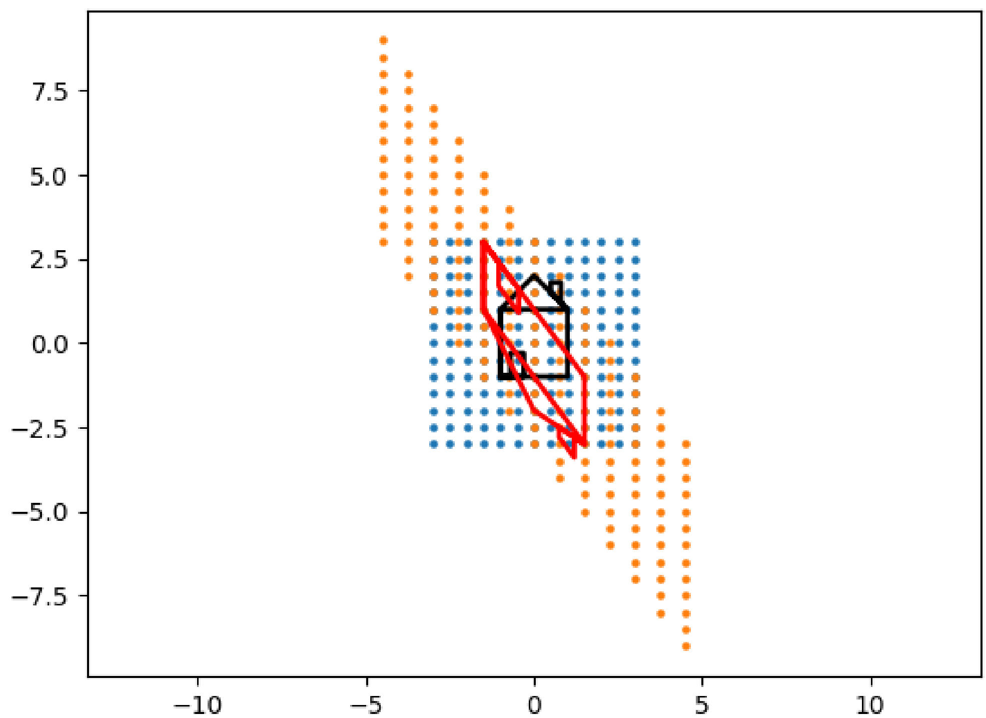
Figure 4.13: The effect of reflecting a house across the line $y = -x$.

Exercise 7 Guess the transformation matrix

Work

Given the following transformation of the house, find the matrix that represents this transformation.

Hint: the grid is 3×3 unit direction with 13 points in each row/column.



Answer on Page 65

4.3 Rotations

Those past four transformations (4.9, 4.11, 4.12, 4.13) all had something similar

What if we want to rotate our house? We can also do this with a matrix transformation. It may look similar to the reflections, but it is not the same. Let's track our basis vectors to see where they end up:

$$\begin{bmatrix}1\\0\end{bmatrix} \mapsto \begin{bmatrix}0\\1\end{bmatrix}, \qquad \begin{bmatrix}0\\1\end{bmatrix} \mapsto \begin{bmatrix}-1\\0\end{bmatrix}$$

From this, we can write the matrix for a 90° rotation clockwise:

$$A_{90^\circ}^{cw} = \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}$$

What about a 90° rotation counterclockwise? We can track our basis vectors again:

$$\begin{bmatrix}1\\0\end{bmatrix} \mapsto \begin{bmatrix}0\\-1\end{bmatrix}, \qquad \begin{bmatrix}0\\1\end{bmatrix} \mapsto \begin{bmatrix}1\\0\end{bmatrix}$$

And so the matrix for a 90° rotation counterclockwise is:

$$A_{90^\circ}^{ccw} = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$$

However, we have learned trigonometric functions that result in these values. We know that at 90°, $\cos(90^\circ) = 0$ and $\sin(90^\circ) = 1$. So we could say:

$$A_{90^\circ}^{cw} = \begin{bmatrix}\cos(90^\circ) & \sin(90^\circ)\\-\sin(90^\circ) & \cos(90^\circ)\end{bmatrix}$$

Note that this aligns with our *polar* coordinate system where cosine is the x-value and sine is the y-value.

For a counterclockwise rotation of 90°, we have:

$$A_{90^\circ}^{ccw} = \begin{bmatrix}\cos(90^\circ) & -\sin(90^\circ)\\\sin(90^\circ) & \cos(90^\circ)\end{bmatrix}$$

But, we can generalize this to any angle θ instead of just 90°. For a clockwise rotation by θ, we have:

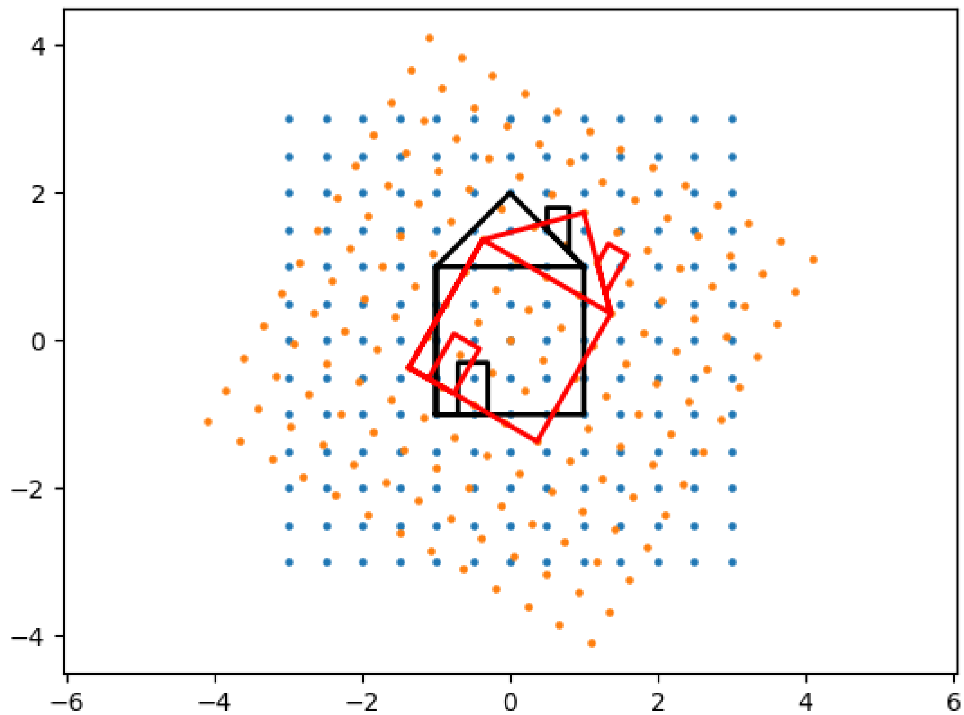
$$A_{\theta}^{cw} = \begin{bmatrix}\cos(\theta) & \sin(\theta)\\-\sin(\theta) & \cos(\theta)\end{bmatrix}$$

And for a counterclockwise rotation by θ, we have:

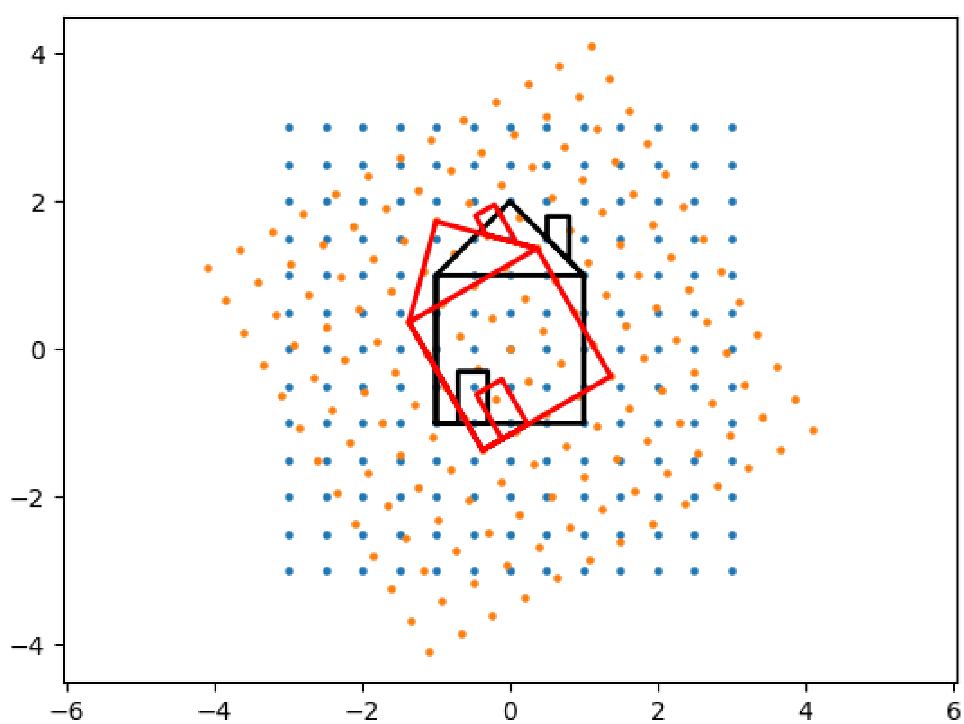
$$A_{\theta}^{ccw} = \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\\sin(\theta) & \cos(\theta)\end{bmatrix}$$

Note how similar they are! The only difference is the sign of the sine values. Sine is an odd function, which means that $\sin(-\theta) = -\sin(\theta)$, so the sign of the sine values determines the direction of rotation.

Let's see what happens when we apply a 30° clockwise and counterclockwise rotation to our house. We can calculate the matrix for this transformation using the cosine and sine values for 30°:



(a) 30 degree clockwise rotation



(b) 30 degree counterclockwise rotation

Figure 4.14: The effect of 30 degree rotations on our house

4.4 What if I want to undo my transform?

We have talked about how to apply a plethora transformations, but what if we want to undo a transformation? For example, if we have a house that is reflected across the rotated 30° clockwise (4.14a), how do we get it back to its original position?

This question brings us back to the application of those *inverse matrices* that we learned about in the previous chapters. If we have a transformation represented by a matrix A, then the inverse transformation is represented by the inverse matrix A^{-1}. Applying A^{-1} to the transformed house will return it to its original position.

Additionally, given a transformed basis vector set $\vec{e_1}'$ and $\vec{e_2}'$, we can apply the inverse transformation to find the original basis vectors. This is because the inverse transformation will reverse the effects of the original transformation, allowing us to recover the original basis vectors from their transformed counterparts.

If you recall, for our 2×2 matrices, the inverse is simple! For a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse is

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

Any larger matrices can get convoluted, so we use gaussian elimination to find the inverse or our numpy function `np.linalg.inv ()`.

The inverse transformation of our 30° clockwise rotation is going to be

$$\begin{aligned} A^{-1}_{30^\circ\text{cw}} &= \frac{1}{\cos(30^\circ)\cos(30^\circ) - (-\sin(30^\circ))\sin(30^\circ)} \begin{bmatrix} \cos(30^\circ) & \sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix} \\ &= \frac{1}{\cos^2(30^\circ) + \sin^2(30^\circ)} \begin{bmatrix} \cos(30^\circ) & \sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix} \\ &= \frac{1}{1} \begin{bmatrix} \cos(30^\circ) & \sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix} \\ &= \begin{bmatrix} \cos(30^\circ) & -\sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix} \end{aligned}$$

Notice that this is the same as the matrix for a 30° counterclockwise rotation! This makes sense, since a 30° counterclockwise rotation would undo a 30° clockwise rotation.

Note that if the determinant is 0, then the associated linear transformation is not invertible. In this case, the transformation collapses space into a lower dimension (for example, a plane into a line or a line into a point), causing information to be lost. As a result, the transformation cannot be reversed, and therefore no inverse matrix exists.

Common examples of this are:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{(projection matrix (covered in a future chapter))}$$
$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{(eliminates one direction)}$$
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{(duplicate / linearly dependent rows)}$$

Exercise 8 Undoing a Shear

A shear, by definition, leaves one axis fixed and slides points parallel to it. For example, an x-shear is given by the following relationship:

$$(x, y) \mapsto (x + ky, y)$$

1. Find the orginal matrix for an x-shear with a shear factor of k.
2. Find the inverse matrix for this transformation.

Working Space

Answer on Page 65

4.5 Summary and looking ahead

In this chapter, we have seen how matrices can be used to represent transformations in $\mathbb{R}^2$. We have explored various types of transformations, including scaling, shearing, reflections, and rotations. We have also discussed how to find the inverse of a transformation, which allows us to undo the effects of a transformation.

We want to look at these transformation as effects on directions. Specifically, are there any directions that this transformation does not mix with others?

- Reflections have the property that they flip one direction while keeping the other direction unchanged. For example, a reflection across the x-axis flips the y-direction while keeping the x-direction unchanged. In this, some directions stay the same, while others are flipped. Certain directions are preserved, while others are not.

- Rotations, on the other hand, mix all directions together. There are no directions that remain unchanged under a rotation, except for the trivial case of a 0° rotation (the identity transformation)
- Stretches and scalings also have the property that they stretch or shrink certain directions while keeping others unchanged. For example, a scaling transformation that stretches the x-direction while keeping the y-direction unchanged will preserve the y-direction while stretching the x-direction.

For some transformations, there exist special directions along which vectors are simply scaled. In studying linear transformations, it is natural to ask whether there exist directions that remain unchanged, except for cardinal scaling of its length. Before reading the next chapter, try to find a nonzero vector whose direction is *unchanged* by each transformation. Not every transformation will have such a vector, but some will.

Chapter 5

Eigenvectors and Eigenvalues

Like many specialized disciplines, Linear Algebra uses many unfamiliar terms whose origins you might wonder about. Eigenvectors and eigenvalues are two of them. If you know German, you will recognize that eigen means inherent or a characteristic attribute. Named by the German mathematician David Hilbert, an eigenvector mathematically describes a characteristic feature of an object that remains unchanged after transformation. You can think of an eigenvector as the direction that doesn not change direction. An eigenvector characterizes a linear transformation, whereas its eigenvalue tells how much the vector is scaled. Eigenvalues can be negative or positive. A negative value indicates the direction of the eigenvector is reversed.

Eigenvalues and eigenvectors are a way to break down matrices, which can simplify many calculations and enable us to understand various properties of the matrix. They are widely used in physics and engineering for stability analysis, vibration analysis, and many other applications.

Let's look at a visual example.

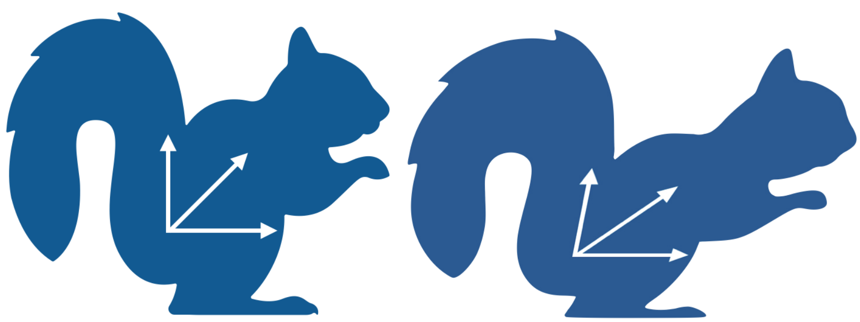


Figure 5.1: Standardized and skewed squirrel image.

You can see that the image on the right is a skewed version of the image on the left. Look closely at the vectors and you will notice that one of the vectors is pointing in the same direction in both images, while the direction of the other two vectors has changed. The eigenvector is the one at the bottom that points to 0 degrees (which you can think of due east) in both images. So, the characteristic attribute of both images is their horizontal direction.

When you overlay the vectors from one image over the other, you will notice that the horizontal vector, while the same direction in both images, is a bit longer in the skewed version. The scale of the stretch is described by an eigenvalue.

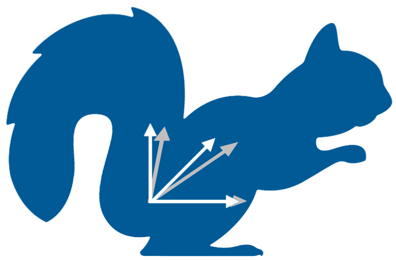


Figure 5.2: Standardized and skewed squirrel image overlayed on eachother.

5.1 Definition

Eigenvector and Eigen Equation

Given a square matrix A, a non-zero vector v is an eigenvector of A if multiplying A by v results in a scalar multiple of v. In other words, the *eigenequation* is:

$$Av = \lambda v \tag{5.1}$$

where λ is a scalar known as the eigenvalue corresponding to the eigenvector v.

An eigenvector is a nonzero vector that does not change direction when a linear transformation is applied.

It may stretch, shrink, or flip direction, but it *stays on the same line*.

5.2 Finding Eigenvalues and Eigenvectors

You find the eigenvalues of a matrix A by solving the characteristic equation:

$$\det(A - \lambda I) = 0 \tag{5.2}$$

where $\det(.)$ denotes the determinant, I is the *identity matrix* of the same size as A, and λ is a scalar. In other words, we add λ along the diagonal of a, and solve for its determinant.

$$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

Once your find the eigenvalues, you can find the corresponding eigenvectors by substituting each eigenvalue into the equation $Av = \lambda v$, and solving for v. v is a vector which only gets *scaled*, with no change in direction.

For a non-zero **v** to satisfy this equation,

-
- $\det(A - \lambda I) = 0$ must have a determinant of zero
- $A - \lambda I$ cannot have an inverse
- space is flattened on to a volume of 0.

FIXME visual representation graphically?

5.3 Example

For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the characteristic equation is:

$$(a - \lambda)(d - \lambda) - bc = 0 \tag{5.3}$$

Solving this equation gives the eigenvalues. Substituting each eigenvalue back into the equation $Av = \lambda v$ gives the corresponding eigenvectors.

Let matrix $A =$

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

The characteristic equation is:

$$|A - \lambda I| = 0$$

$$\begin{bmatrix} 5 - \lambda & 4 \\ 1 & 2 - \lambda \end{bmatrix} = 0$$

$$(5 - \lambda)(2 - \lambda) - (4)(1) = 0$$

$$10 - 5\lambda - 2\lambda + \lambda 2 - 4 = 0$$

$$\lambda 2 - 7\lambda + 6 = 0$$

$$(\lambda - 6)(\lambda - 1) = 0$$

$$\lambda = 6, \lambda = 1$$

Now that you have the eigen values you can substitue these values into the equation:

$$|A - \lambda I| = 0$$

For $= 1$:

$$(A - \lambda I)v = O$$

$$\begin{bmatrix} 5-1 & 4 \\ 1 & 2-1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Next, use elementary row transformation by multiplying row 2 by 4, then subtracting row 1.

$$\begin{bmatrix} 4 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Now you can expand as an equation:

$$4x + 4y = 0$$

Assume $y = w$

$$4x = -4w$$

$$x = -w$$

The solution is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -w \\ w \end{bmatrix} = w \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

So the eigenvector is:

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

Now we need to substitute the other eigenvalue, 6, into the equation and follow the same procedure for finding the eigenvector.

$$\begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Next, use elementary row transformation by adding row 1 to row 2.

$$\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Expand as an equation:

$$-x + 4y = 0$$

$$\text{Assume} y = w$$

$$-x + 4w = 0$$

$$x = 4w$$

The solution is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4w \\ w \end{bmatrix} = w \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

So the eigenvector is:

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

In conclusion, the eigenvectors of the given 2 x 2 matrix are:

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{and} \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

5.4 Eigenbasis and Eigenspace

Let A be a square matrix and let λ be an eigenvalue of A.

Eigenspace

The **eigenspace** corresponding to λ is the set of all vectors v such that

$$Av = \lambda v.$$

Equivalently, the eigenspace of λ is the solution set of

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

Note that there will likely be multiple lambdas for a given matrix, and therefore multiple eigenspaces for each matrix.

The eigenspace is a subspace consisting of all eigenvectors associated with λ, together including the zero vector, $\vec{\mathbf{0}}$.

Eigenbasis

An *eigenbasis* for a matrix A is a basis of the vector space consisting entirely of eigenvectors of A.

$$\{v_1, v_2, \ldots, v_n\}$$

is an eigenbasis for A if:

$$Av_i = \lambda_i v_i \quad \text{for each } i = 1, 2, \ldots, n$$

Exercise 9 A 2×2 matrix

Working Space

Find the eigenbasis of the matrix

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

Answer on Page 66

5.5 Eigenvalues and Eigenvectors in Python

Create a file called `vectors_eigen.py` and enter this code:

```
# import numpy to perform operations on vector
import numpy as np
from numpy.linalg import eig

a = np.array([[2, 2, 4],
              [1, 3, 5],
              [2, 3, 4]])
eigenvalue,eigenvector = eig(a) #return a tuple of eigenvalues [l1, l2,..., ln],
↪  eigenvectors [ [...], [...], ...]

# The values are not in any particular order
print('Eigenvalues:', eigenvalue)

# The eig function returns the normalize vectors
print('Eigenvectors:', eigenvector)
```

5.6 Connection to transformations

In the previous chapter, we studied matrix transformations such as scaling, reflection, shear, and rotation. Each of these transformations changes vectors in space by stretching, shrinking, flipping, or rotating them. Most vectors change both their length and/or their direction under such transformations.

Eigenvectors arise naturally when a transformation leaves at least one direction unchanged. Different types of transformations exhibit different eigenvector behavior:

- **Scaling transformations** preserve all directions. Every nonzero vector is an eigenvector.
- **Reflections** preserve the direction of vectors along the axis of reflection and reverse vectors perpendicular to it. Vectors lying on the axis of reflection are unchanged by the transformation and therefore are eigenvectors with eigenvalue 1. Vectors perpendicular to the axis of reflection are flipped to point in the opposite direction, making them eigenvectors with eigenvalue -1.
- **Shear transformations** preserve exactly one direction, even though most vectors change direction.
- **Rotations in two dimensions** generally preserve no directions, and therefore have no real eigenvectors (except in the trivial cases of $\theta = \frac{\pi}{2}$ or 0).

These observations explain why some matrices have many eigenvectors, some have only a few, and some have none at all. Eigenvectors exist precisely when a transformation has invariant directions.

Understanding eigenvectors as preserved directions allows us to connect algebraic computations to geometric behavior. When a matrix has enough independent eigenvectors, it becomes possible to describe the transformation entirely in terms of simple scaling along these special directions.

5.7 Summary

In summary, lets take a *transformation*:

Something that stretches, squishes, flips, or rotates space. This transformation could be represented by a matrix.

Most vectors in space will get changed in both direction and length when you apply the transformation.

But some special vectors don't change direction at all — they *only* get stretched or shrunk (and maybe flipped).

These special "unchanging-direction" vectors are called eigenvectors. The amount they get stretched (or shrunk) is their eigenvalue.

5.8 Where to Learn More

Watch this video from Khan Academy, *Introduction to Eigenvectors*: https://rb.gy/mse7i

Chapter 6

Projections

The word ”projection” has two main meanings in everyday life. One is a projection as a forecast or estimate of something in the future based on the current situation; another is the result of shining a light to cast a shadow or show a movie. Both these definitions apply to mathematical projection.

Projections are used in many fields, such as science, math, engineering, and finance. Here are a few examples:

- Investors evaluate risk and return of a portfolio by projecting an asset's return onto a reference portfolio.
- Astronomers analyze the motion of stellar objects by projecting the object's true motion onto the plane of the sky.
- Robotics engineers use projections to prevent robots from running into obstacles by projecting the robot's position onto the optimal path.

Mathematically, a projection describes the relationship of one vector to another in terms of direction and orthogonality. Given two vectors, $\mathbf{u}$ and $\mathbf{v}$, the projection of $\mathbf{u}$ onto $\mathbf{v}$ separates $\mathbf{u}$ into two components. The first component signifies how much $\mathbf{u}$ lies in the direction of $\mathbf{v}$. The second signifies the component of $\mathbf{u}$ that is orthogonal (perpendicular) to $\mathbf{v}$.

The figure 6.1 depicts a projection. The perpendicular line dropped from the end of $\mathbf{u}$ is the *orthogonal* component. The portion of $\mathbf{u}$ that lies in the direction of $\mathbf{v}$ is the blue segment.

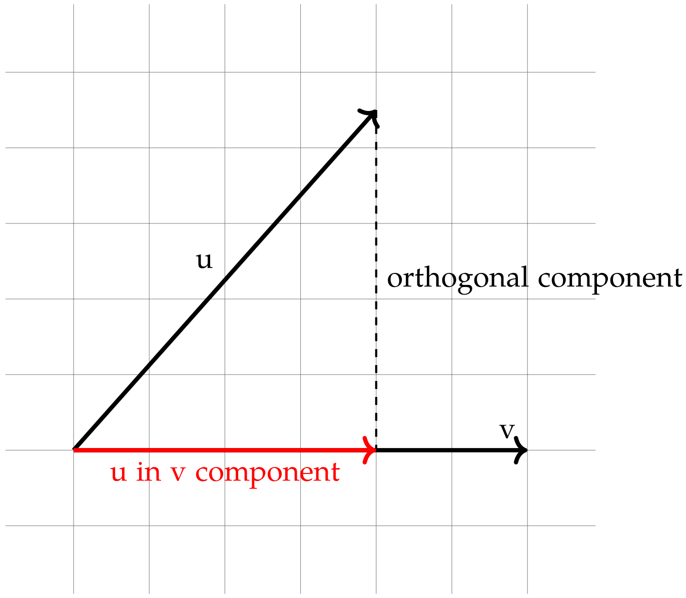


Figure 6.1: The projection of u onto v

You can also think of a projection as the shadow cast by one vector onto the other by an overhead light. See Figure 6.2

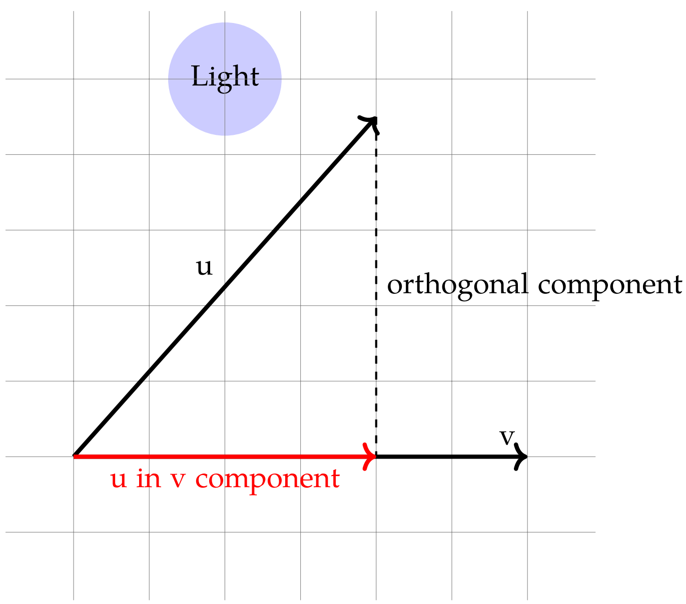


Figure 6.2: Projection of u onto v with a light included to simulate a shadow.

The projected vector can be in any direction and its length can extend beyond the vector onto which it is projecting. See Figure 6.3.

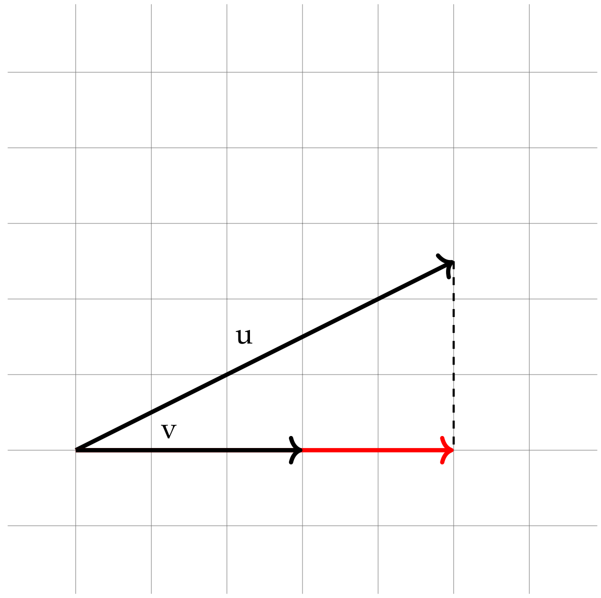


Figure 6.3: Projection vector extended beyond v.

To calculate the projection of **v** onto **u**, use this formula:

$$\text{Projection of } \mathbf{u} \text{ onto } \mathbf{v}: \quad \text{proj}_{\mathbf{v}}(\mathbf{u}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v}$$

where $\mathbf{u} \cdot \mathbf{v}$ is the dot product of $\vec{u}$ and $\vec{v}$. Note that either form of the equation works, and the **v** being multiplied by the dot product quotient cannot be cancelled because it is a vector, not a scalar.

Note that the denominator is the magnitude squared of vector **v**.

$$\left(\sqrt{a_1^2 + a_2^2 + ... + a_n^2}\right)^2$$

You learned previously that this is the same as the dot product of a vector with itself.

$$v \cdot v$$

In the examples that follow, we will simplify to the dot product notation.

Let's look at a specific example:

$$u = (1, 4, 6)$$

$$v = (-2, 6, 2)$$

$$\mathbf{proj_v}(\mathbf{u}) = \frac{u \cdot v}{\| v \|^2} v$$
$$\mathbf{proj_v}(\mathbf{u}) = \frac{(1,4,6) \cdot (-2,6,2)}{(-2,6,2) \cdot (-2,6,2)}(-2,6,2)$$
$$\mathbf{proj_v}(\mathbf{u}) = (\frac{34}{44}(-2,6,2)$$
$$\mathbf{proj_v}(\mathbf{u}) = (-1.545, 4.64, 1.545)$$

As you work your way through this course, you will have a chance to apply the calculations you learn in this chapter to a variety of problems. Specifically, the next chapter shows how to transform a set of linearly independent vectors into a set of orthogonal ones. Projections are essential to that transformation.

Exercise 10 Projections

Find the projection of **a** on **b** where:

$$a = (1,3)$$
$$b = (-4,6)$$

Working Space

Answer on Page 67

6.1 Projections in Python

Create a file called `projections.py` and enter this code:

```
import numpy as np

# define two vectors
a = np.array([1, 4, 6])
b = np.array([-2, 6, 2])

# use np.dot() to calculate the dot product
projection_a_on_b = (np.dot(a, b)/np.dot(b, b))*b

print("The projection of vector a on vector b is:", projection_a_on_b)
```

6.2 Where to Learn More

Watch this Introduction to Projections from Khan Academy from your digital resources: https://rb.gy/yf0i3

Appendix A

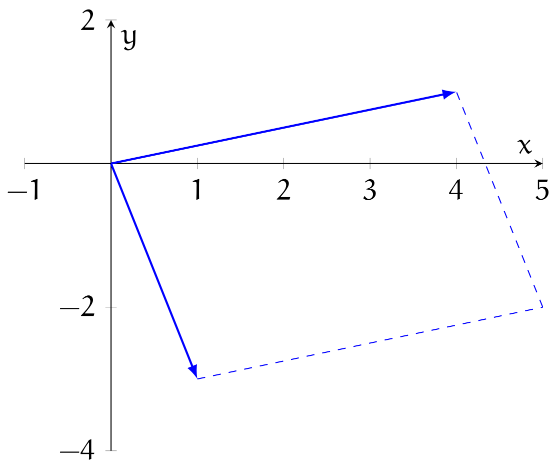
Answers to Exercises

Answer to Exercise 1 (on page 6)

1. Since the second vector is a scalar multiple of the first, the span of $S = \{[1, 2, 4], [-2, -4, -8]\}$ is a *line*.
2. Since the second vector is not a scalar multiple of the first, the span of $S = \{[2, 0, 0,], [0, 1, 3]\}$ is a *plane*.
3. None of the three vectors are scalar multiples or linear combinations of the other two. Therefore, the span of $S = \{[3, 0, 0], [0, 3, 3], [3, 3, 2]\}$ is $\mathbb{R}^3$.

Answer to Exercise 2 (on page 13)

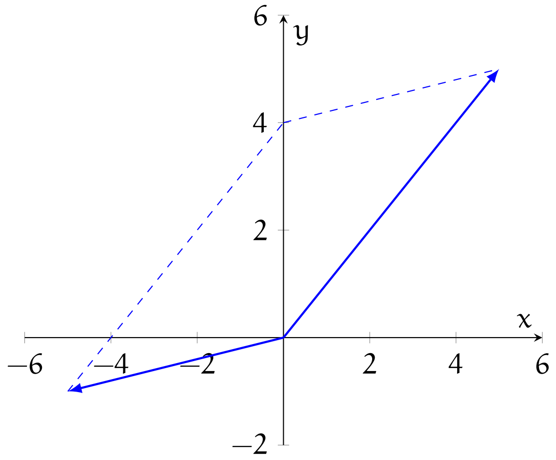
1. Our two vectors from the columns of the matrix are $[1, -3]$ and $[4, 1]$. Plotting:



The area of this parallelogram is the same as the determinant of the matrix:

$$\det\left(\begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix}\right) = 1 \cdot 1 - (4 \cdot -3) = 1 + 12 = 13$$

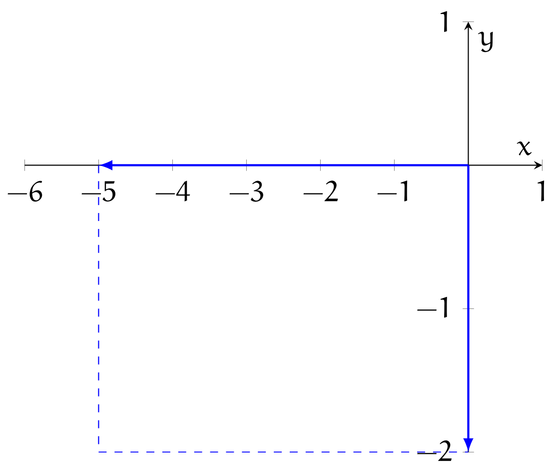
2. Our two vectors from the columns of the matrix are $[5, 5]$ and $[-5, -1]$. Plotting:



The area of this parallelogram is the same as the determinant of the matrix:

$$\det\left(\begin{bmatrix} 5 & -5 \\ 5 & -1 \end{bmatrix}\right) = 5 \cdot -1 - (-5 \cdot 5) = -5 + 25 = 20$$

3. Our two vectors from the columns of the matrix are $[0, -2]$ and $[-5, 0]$. Plotting:



This is a rectangle, and we can see the area is $5 \cdot 2 = 10$. However, the determinant is:

$$\det\left(\begin{bmatrix} 0 & -5 \\ -2 & 0 \end{bmatrix}\right) = 0 \cdot 0 - (-5 \cdot -2) = 0 - 10 = -10$$

We will discuss this unusual response in a future chapter.

Answer to Exercise 3 (on page 17)

First, find the determinant of A:

$$\det(A) = (1)(-1) - (2)(0) = -1$$

So our inverse is given by:

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

Answer to Exercise 4 (on page 18)

First, check that A has an inverse:

$$\det A = 1 \begin{vmatrix} 1 & 1 \\ 4 & -3 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 \\ 2 & -3 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 \\ 2 & 4 \end{vmatrix} = 1(-3-4) - 2(0-2) - 2(0-2) = -7+4+4 = 1$$

Since $\det(A) \neq 0$, A must have an inverse. We find it's inverse by the augmented matrix method:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 4 & -3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 \leftarrow R_2 - R_3} \left[\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & -2 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 \leftarrow R_1 + 2R_3} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -3 & 0 & 2 \\ 0 & 1 & 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & -2 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 \leftarrow R_1 - 2R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -7 & -2 & 4 \\ 0 & 1 & 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & -2 & 0 & 1 \end{array}\right].$$

This gives us our inverse on the right side, $A^{-1} = \begin{bmatrix} -7 & -2 & 4 \\ 2 & 1 & -1 \\ -2 & 0 & 1 \end{bmatrix}$

To check our answer, we can use $AA^{-1} = I_3$:

$$AA^{-1} = \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & 1 \\ 2 & 4 & -3 \end{bmatrix} \begin{bmatrix} -7 & -2 & 4 \\ 2 & 1 & -1 \\ -2 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1(-7)+2(2)+(-2)(-2) & 1(-2)+2(1)+(-2)(0) & 1(4)+2(-1)+(-2)(1) \\ 0(-7)+1(2)+1(-2) & 0(-2)+1(1)+1(0) & 0(4)+1(-1)+1(1) \\ 2(-7)+4(2)+(-3)(-2) & 2(-2)+4(1)+(-3)(0) & 2(4)+4(-1)+(-3)(1) \end{bmatrix}$$

$$= \begin{bmatrix} -7+4+4 & -2+2+0 & 4-2-2 \\ 0+2-2 & 0+1+0 & 0-1+1 \\ -14+8+6 & -4+4+0 & 8-4-3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

Answer to Exercise 5 (on page 19)

No. Each row is a scalar multiple of the first row:

- Row 2 = 2 × Row 1
- Row 3 = 4 × Row 1
- Row 4 = 8 × Row 1

So the rows are *linearly dependent*, meaning the matrix has rank 1, and its determinant is 0. A matrix with determinant 0 is singular, so no inverse exists.

Answer to Exercise 6 (on page 24)

Setting up the matrix, we get:

$$\begin{bmatrix} \vec{\mathbf{i}} & \vec{\mathbf{j}} & \vec{\mathbf{k}} \\ 4 & 5 & 6 \\ 3 & 7 & -8 \end{bmatrix}$$

$$(5(-8)-(7)(6))\,\vec{\mathbf{i}} - ((4)(-8)-(6)(3))\vec{\mathbf{j}} + ((4)(7)-(5)(3))\vec{\mathbf{k}}$$

$$(-40 - 46)\vec{\mathbf{i}} - (-32 - 18)\vec{\mathbf{j}} + (28 - 15)\vec{\mathbf{k}}$$
$$-86\vec{\mathbf{i}} + 50\vec{\mathbf{j}} + 13\vec{\mathbf{k}} \implies \langle -82, 50, 13 \rangle$$

Answer to Exercise 7 (on page 40)

The matrix that represents this transformation is:

$$A = \begin{bmatrix} 1.5 & 0 \\ -2 & -1 \end{bmatrix}$$

We find this by first noticing that the house and door components are flipped over the x-axis, which means that the y-values of the points are flipped. This gives us a negative one value in the e_2 basis of the matrix.

Next we can see that the house is stretched horizontally in the x-direction. The orange dots end at 4.5, which is 1.5 times the original length of 3, so we know that the x-values are scaled by a factor of 1.5. This gives us the value of 1.5 in the first row of the matrix.

Finally, the roof is slanted, which means that there is a shearing component to this transformation. Notice that the roof (door corner) is moved up, so the e_1 also has a negative value y-component. This gives us the value of -2 in the second row, first column of the matrix.

Answer to Exercise 8 (on page 45)

1. The original matrix for an x-shear with a shear factor of k is:

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

2. The inverse matrix for this transformation is:

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

   (the determinant is 1, and the inverse is found by swapping the diagonal elements and negating the off-diagonal elements, which in this case results in negating k).

Answer to Exercise 9 (on page 52)

First, write the $A - \lambda I$ equation

$$\begin{bmatrix} 1-\lambda & 2 \\ 2 & -2-\lambda \end{bmatrix}$$

Our goal is to get the determinant of $A-\lambda I$ to be equal to zero. For a 2×2, the determinant is $ad - bc$:

$$\begin{aligned} (1-\lambda)(-2-\lambda) - 4 &= 0 \\ (-2+2\lambda-\lambda+\lambda^2) - 4 &= 0 \\ \lambda^2+\lambda-6 &= 0 \\ \lambda_1 = 2, &\quad \lambda_2 = -3 \end{aligned}$$

Then we solve for the two eigenvectors, $\mathbf{e}_n$, using $A - \lambda I$

For $\lambda_1 = 2$:

$$\lambda_1 = 2 : A - 2I = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$$

We solve the following using row-column multiplication (rows of numbers and columns of x_i)

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

We notice, $x_1 = 2x_2$. This means every eigenvector is a multiple of $\mathbf{e}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (using $x_2 = 1$)
.

For $\lambda_2 = -3$:

$$\lambda_2 = -3 : A + 3I = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} 4x_1 + 2x_2 = 0 \\ 2x_1 + x_2 = 0 \end{cases}$$

This results in $x_2 = -2x_1$, meaning $\mathbf{e}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

Answer to Exercise 10 (on page 58)

Compute dot product of **a** and **b**:

$$1 * -4 + 3 * 6 = -4 + 18 = 14$$

Compute the dot product of **b** and **b**

$$16 + 36 = 52$$

$$14/52 * (-4, 6) = (-1.076, 1.61)$$

Index