Welcome to Week 2
SI 568

Policy Changes

Math !

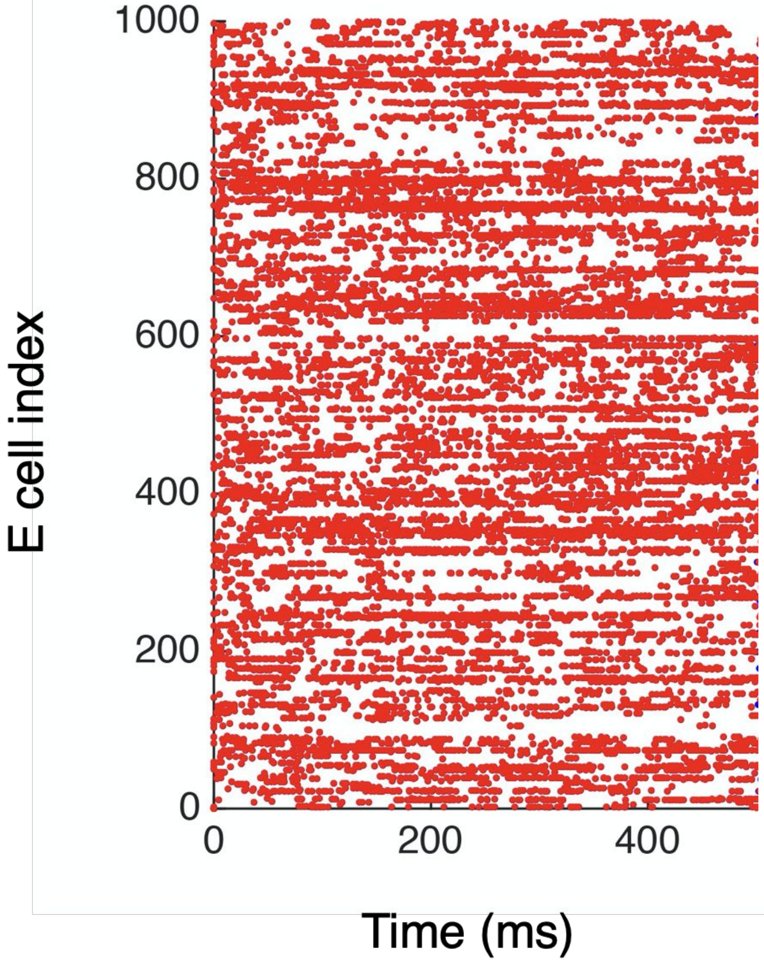
Office Hours

Probability

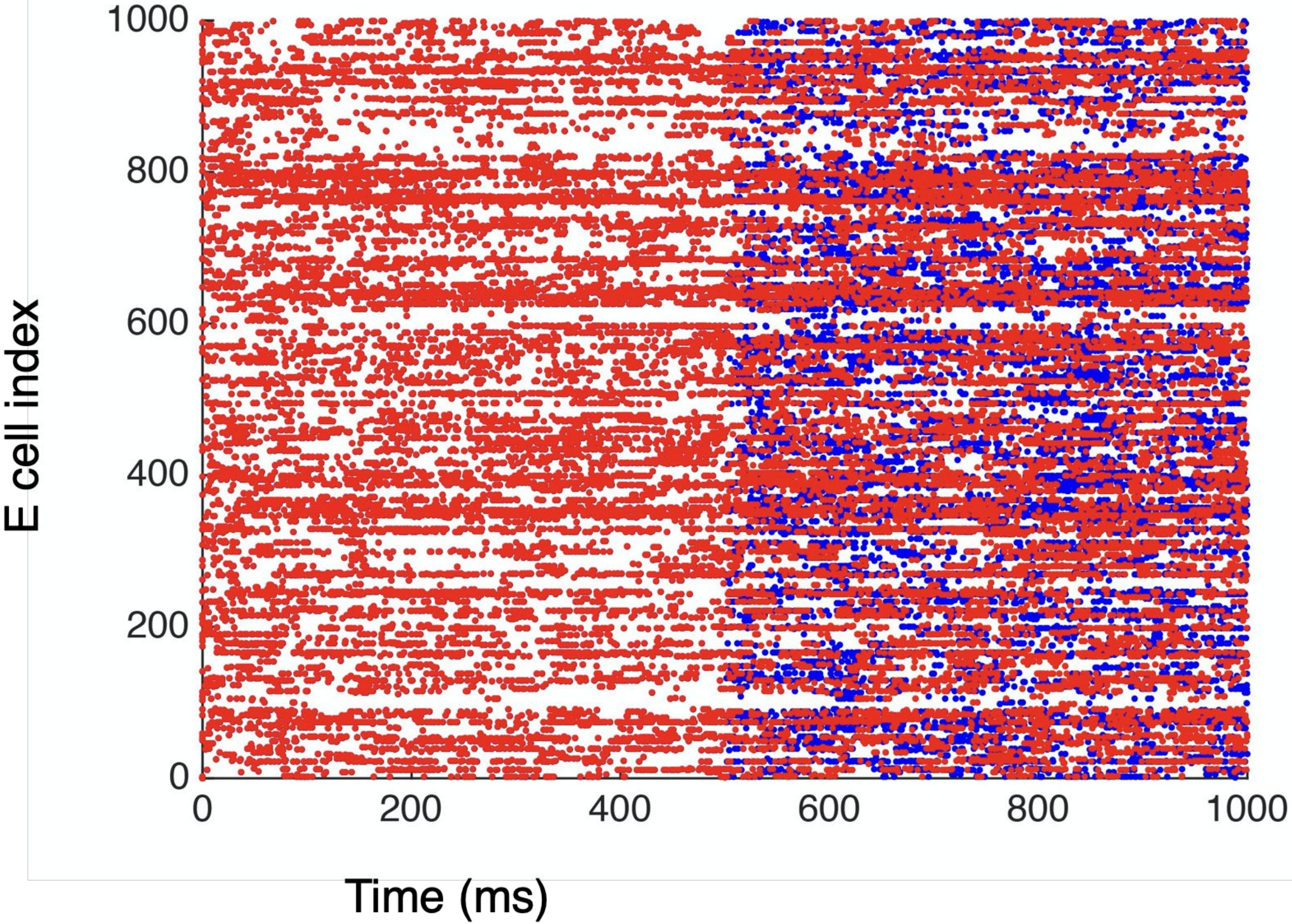
What do we mean by random?

Is it random?

- You flip a coin
- The price of a particular stock on the NYSE
- The weather 10 minutes from now
- The weather 10 days from now
- The weather 10 years from now
- The seat you’re currently sitting in
- You’re walking to class and someone throws a potato at you – is that random?



E cell index
1000
800
600
400
200
0
0
200
400
Time (ms)



E cell index
1000
800
600
400
200
0
0
200
400
600
800
1000
Time (ms)

Slido goes here

Counting!

How many ways can we arrange n objects?

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$$

How many ways can we select k objects from n?

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Editor's Note: Intuitively, there is a very nice explanation for why we should expect $0! = 1$. Because $n!$ enumerates the number of ways to order n objects, we expect $0!$ to enumerate the number of ways to order 0 objects. But there's only one way to rearrange 0 objects, because there are *no* objects to order.

However, we can bring in a little bit more math to extend this mathematical idea of the factorial to a more general class of numbers. To do this, we can define the *gamma function*:

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x}\, dx$$

When we evaluate Γ at a positive integer value, we can show something kind of cool. Using integration by parts (*try it!*), we have

$$\Gamma(n) = (n-1)!$$

Using this, we can draw an effective equivalence between the gamma function and the factorial. We then expect $\Gamma(1) = 0!$, which we can compute directly as

$$\begin{aligned} 0! = \Gamma(1) &= \int_0^\infty x^0 e^{-x}\, dx = \int_0^\infty e^{-x}\, dx \\ &= \left(-e^{-x}\right)\Big|_0^\infty = 0 - (-1) \\ &= 1 \end{aligned}$$

In particular, this means we can extend the idea of a factorial to non-integer values, to find, for example

$$\left(\frac{1}{2}\right)! = \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$$

- Counting
  - Arranging n objects
  - Arranging k objects out of n
- Probability
  - Experimental
  - Theoretical
  - Complements
  - Containment
  - Adding, Subtracting
- Conditional Probability
- Independence
- Law of total probability
- Berkson’s paradox
- Bayes’ Theorem

Linear Algebra

Code examples are [here]

What is a matrix?

- At its simplest, a matrix is just an array of numbers!

$$M_1 = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} \frac{1}{2} & 3 \\ 2 & 3 \\ 5 & -\frac{1}{2} \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix}$$

```
M1 = np.array([[0,1],[2,-1]])
```

- Notation: the number in the *ith* row and *jth* column is the *(i,j)* entry

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Matrix notation

- A has n rows and m columns
- A is an n × m matrix

- When n = m, we call A a *square* matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

What is a vector?

- One definition: a vector is a matrix that has only a single row or a single column
- Another way to say this: a vector is a list of numbers

- Notation: since vectors only have one dimension, we just use one index:

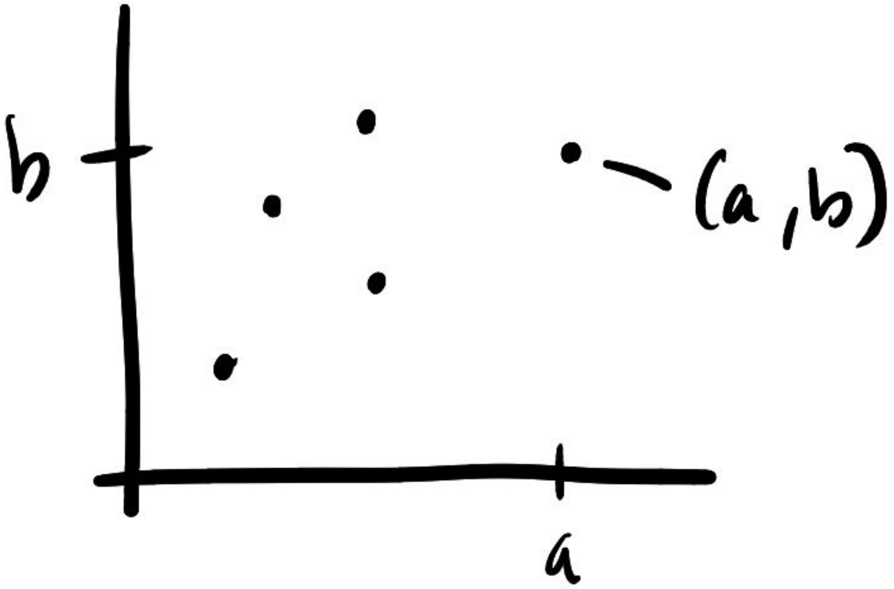
$$u = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

- If a vector is a single row, we call it a *row vector* and if it's a single column we call it a *column vector*. (If not stated, typically assumed to be a column vector)
- Sometimes write a vector with a hat $\vec{v}$ or bold to indicate it's a vector

What is a vector?

- But there are other ways of looking at vectors too!
- We can view a vector as:
  - A point in n-dimensional space
  - An arrow representing geometric displacement
  - Coefficients of a polynomial
  - And many other interpretations!

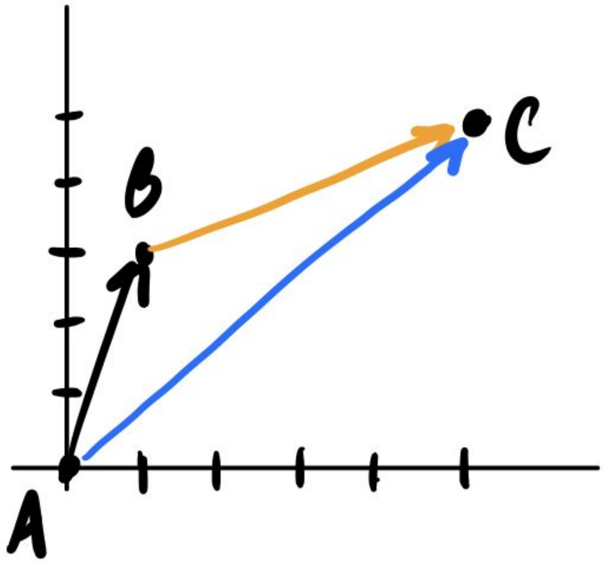
  - For data science, vectors will often be data—a single vector is often a single point of data



What is a vector?

- But there are other ways of looking at vectors too!
- We can view a vector as:
    - A point in n-dimensional space
    - An arrow representing geometric displacement
    - Coefficients of a polynomial
    - And many other interpretations!

    - For data science, vectors will often be data—a single vector is often a single point of data



$A = (0,0)$

$B = (1,3)$

$C = (5,5)$

$$\vec{AB} + \vec{BC} = \vec{AC}$$

$$\begin{array}{ccc} \vec{AB} & \vec{BC} & \vec{AC} \\ \| & \| & \| \\ (1,3) & (4,2) & (5,5) \end{array}$$

What is a vector?

- But there are other ways of looking at vectors too!
- We can view a vector as:
  - A point in n-dimensional space
  - An arrow representing geometric displacement
  - Coefficients of a polynomial
  - And many other interpretations!

  - For data science, vectors will often be data—a single vector is often a single point of data

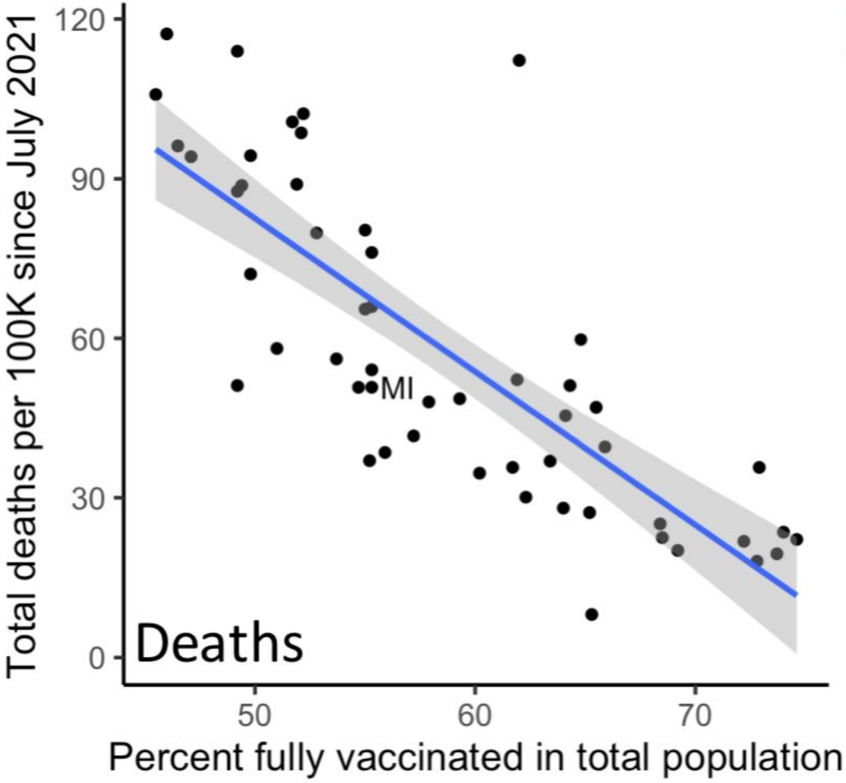
$$v = (1, 2, 5, 3)$$

$$1 + 2x + 5x^2 + 3x^3$$

What is a vector?

- But there are other ways of looking at vectors too!
- We can view a vector as:
  - A point in n-dimensional space
  - An arrow representing geometric displacement
  - Coefficients of a polynomial
  - And many other interpretations!

  - For data science, vectors will often be data—a single vector is often a single point of data



Dots = US states + PR and DC. Each dot/state is a vector (a,b), where a = % vaccinated and b = deaths since July 2021

Basic Matrix & Vector Operations: Addition

Add element-wise!

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$A + B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 4 \\ 3 & -5 \end{pmatrix} \quad \Longrightarrow \quad A + B = \begin{pmatrix} -1 & 7 \\ 5 & 3 \end{pmatrix}$$

Basic Operations: Scalar Multiplication

Scalar multiplication means multiplying a vector or matrix by a constant. For this, we usually have to define where our scalars (constants) are coming from—usually either the real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$

$$c \in \mathbb{R}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \Longrightarrow \quad cA = \begin{pmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{pmatrix}$$

$$c = 3, \quad A = \begin{pmatrix} 1 & 3 \\ 2 & 8 \end{pmatrix} \quad \Longrightarrow \quad 3A = \begin{pmatrix} 3 & 9 \\ 6 & 24 \end{pmatrix}$$

```
c = 3
A = np.array([[1, 3], [2, 8]])

c*A
```

```
array([[ 3,  9],
       [ 6, 24]])
```

Basic Operations: Matrix Multiplication

Now it starts to get a little more tricky! Let's start with two vectors:

$$uv = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

A 1xn matrix (vector) times an nx1 matrix (vector) gave us a number—i.e. a 1x1 matrix!

We did this essentially element-wise, but this captures the main idea for larger matrices

Matrix Multiplication

Let’s try a bigger example:

$$M_2 = \begin{pmatrix} \frac{1}{2} & 3 \\ 2 & 3 \\ 5 & -\frac{1}{2} \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix}$$

$$M_2 M_3 = \begin{pmatrix} \frac{1}{2} & 3 \\ 2 & 3 \\ 5 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

3x2 2x4 3x4

Matrix Multiplication

1st row

1st col

1,1 spot

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

$$\frac{1}{2}\cdot 1 + 3\cdot 1 = \frac{7}{2}$$

$$M_2M_3 = \begin{pmatrix}\frac{1}{2} & 3\\ 2 & 3\\ 5 & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}1 & 0 & 2 & 0\\ 1 & 3 & 1 & 0\end{pmatrix} = \begin{pmatrix}\frac{7}{2} & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\end{pmatrix}$$

Matrix Multiplication

$$M_2M_3 = \begin{pmatrix} \frac{1}{2} & 3 \\ 2 & 3 \\ 5 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} & \cdot & \cdot & \cdot \\ 5 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$2 \cdot 1 + 3 \cdot 1 = 5,$$

Matrix Multiplication

$$M_2M_3 = \begin{pmatrix} \frac{1}{2} & 3 \\ 2 & 3 \\ 5 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} & 9 & 4 & 0 \\ 5 & 9 & 7 & 0 \\ \frac{9}{2} & -\frac{3}{2} & \frac{19}{2} & 0 \end{pmatrix}$$

@ = matrix multiplication

```
A = np.array([[0.5, 3],[2, 3],[5,-0.5]])
B = np.array([[1,0,2,0],[1,3,1,0]])

A@B
```

```
array([[ 3.5,  9. ,  4. ,  0. ],
       [ 5. ,  9. ,  7. ,  0. ],
       [ 4.5, -1.5,  9.5,  0. ]])
```

Matrix Multiplication

Matrix times a vector

$$\underset{3\times 3}{\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ 5 & 0 & 0 \end{pmatrix}} \underset{3\times 1}{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}} = \underset{3\times 1}{\begin{pmatrix} 11 \\ 2 \\ 5 \end{pmatrix}}$$

Takes a vector to another vector!

Matrix Multiplication

Note that the dimensions have to work out for this to work!

Can’t multiply a 2x3 times a 4x4 for example (or 2x3 times a 2x4!)

Matrix multiplication is not commutative (you can’t switch the order!)

2x3 times 3x2 = 2x2

But

3x2 times 2x3 = 3x3 —not the same!

What is a matrix?

Now that we understand a little more about how matrices work, we can understand them a little better

A **matrix** is a transformation or a function that acts on a vector

In other words, a **matrix** is a rule for taking one vector and getting a new vector

Linear algebra is really all about looking at and taking advantage of how vectors (data) get transformed by matrices

How does a matrix act on a vector?

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

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

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

The columns of the matrix are just where it sends each of these special vectors!

(these vectors are called the unit basis vectors, more on why in a bit)

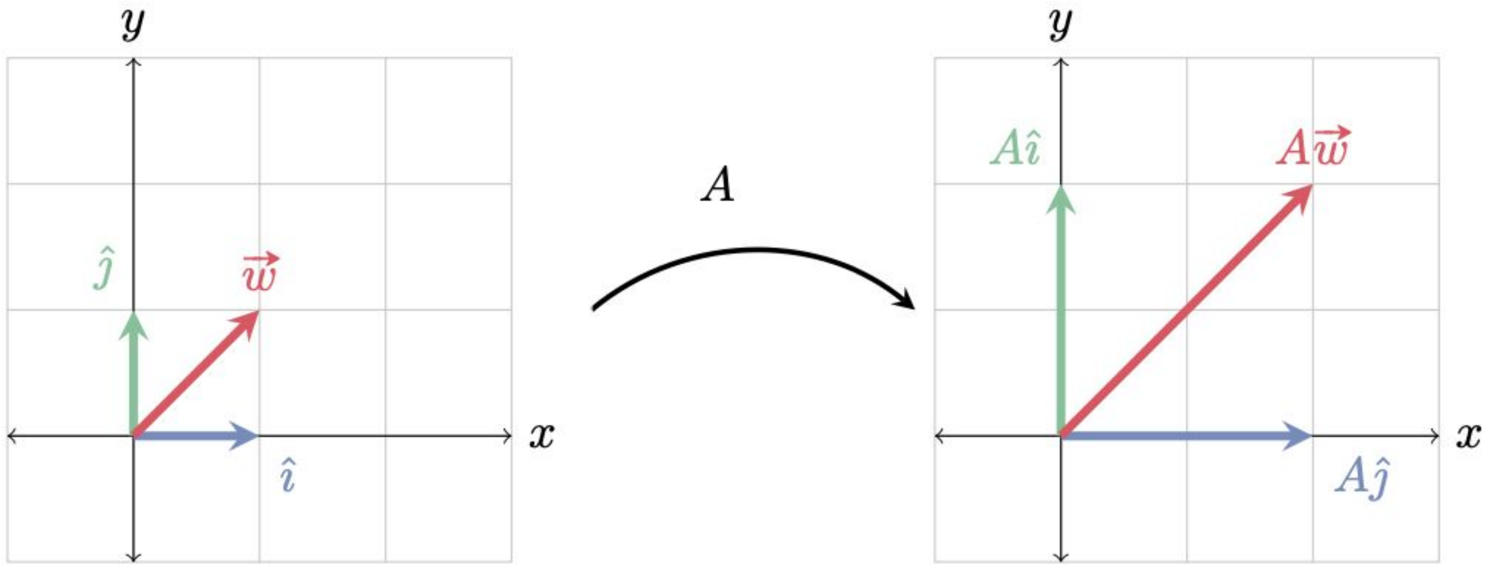
This shows how the matrix acts on each dimension of the space

Examples

For each of the examples below, we're going to follow the same basic script. We start with the standard basis unit vectors $\hat{\imath} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\hat{\jmath} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, as well as one additional vector $\vec{w} = \hat{\imath} + \hat{\jmath}$. We will plot these 3 vectors, as well as where they end up after applying the matrix A, and try to understand the action of A geometrically.

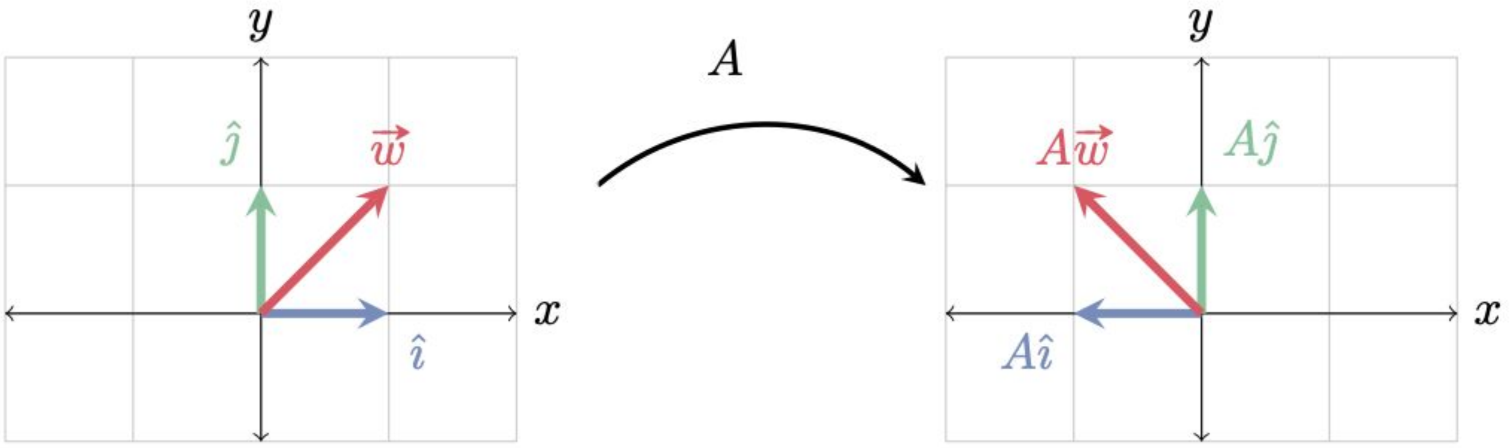
$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

This matrix stretches a vector to have twice the length, but point in the same direction.



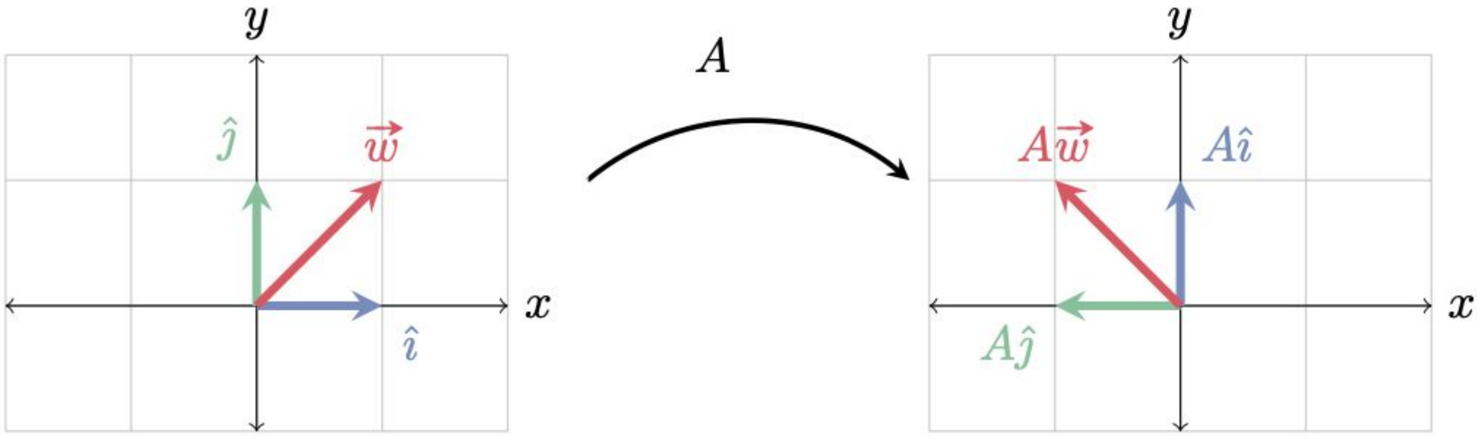
$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

This matrix reflects a vector about the y-axis.



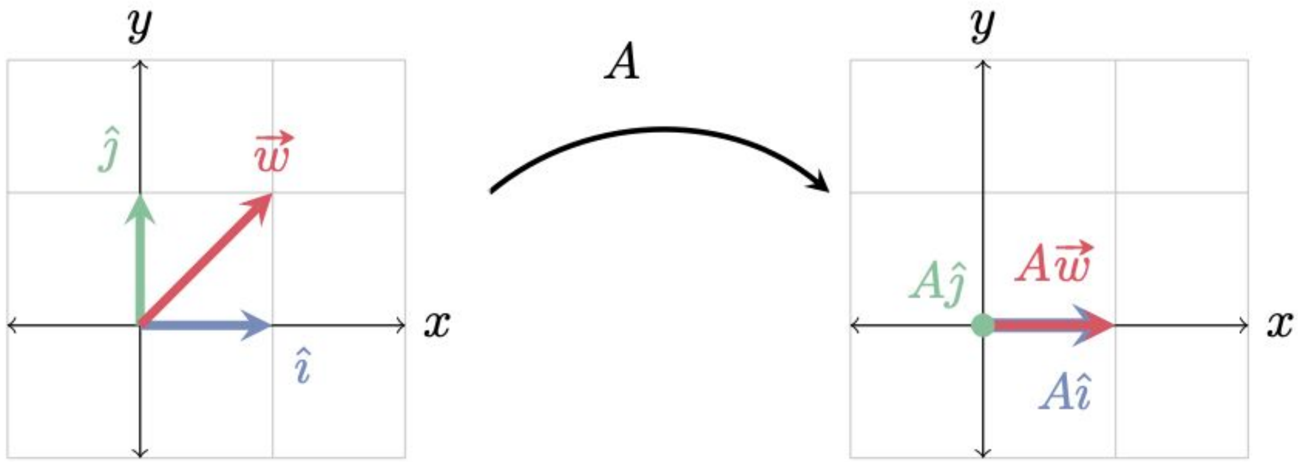
$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

This matrix rotates a vector counter-clockwise by $\pi/2$.



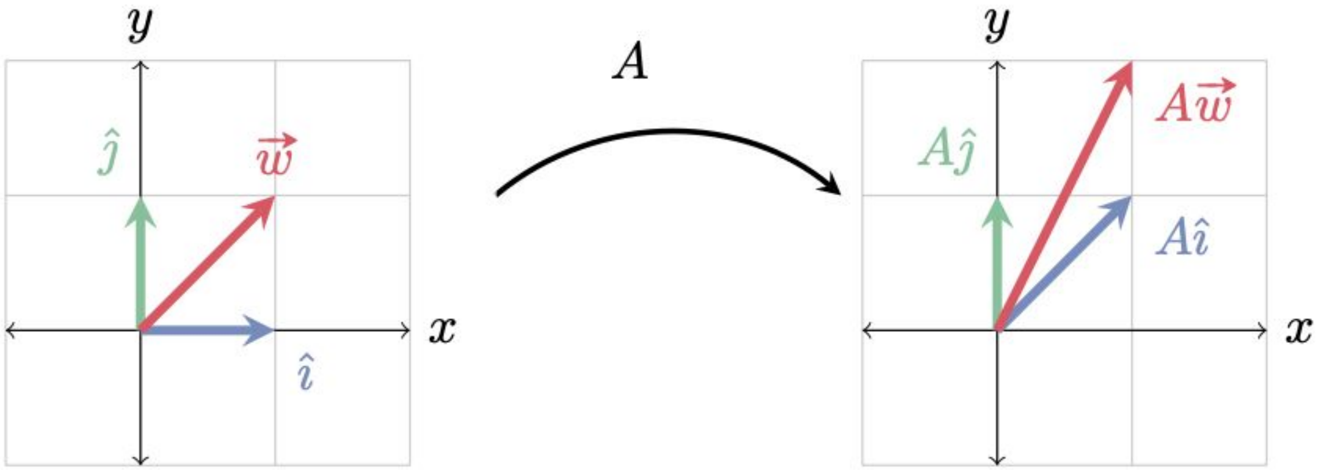
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

This matrix projects a vector onto the x-axis (i.e. zeros out the y-component).



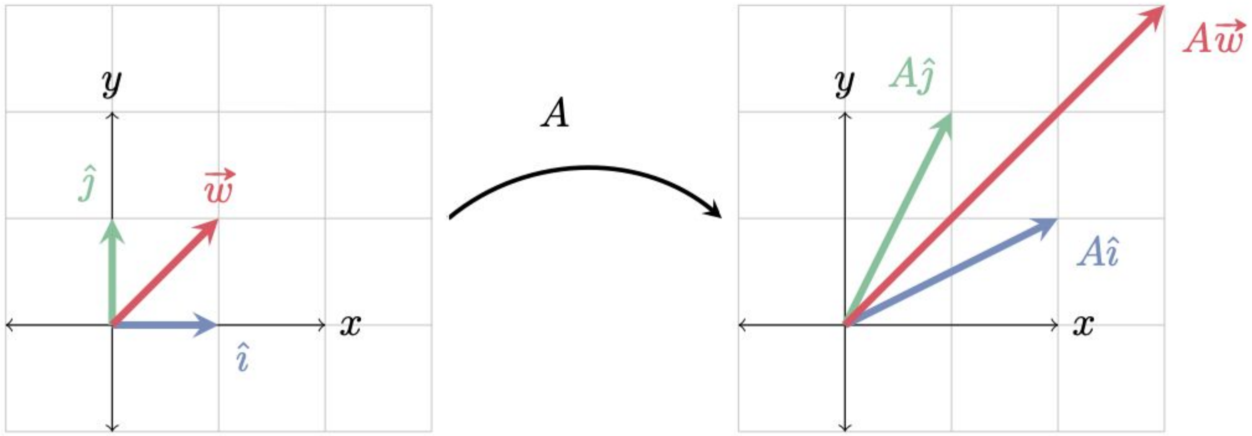
$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

This matrix adds the x-component onto the y-component.



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

This matrix is a bit harder to parse geometrically. I trust you all to understand how we generate the picture below, but we'll leave this example floating around until we've discussed eigenvalues and eigenvectors below. Clearly something nice is happening with $\vec{w}$, but what is happening with $\hat{\imath}$ and $\hat{\jmath}$ is not so obvious.



Revisit matrix operations—what do they mean for the matrix as a transformation?

Addition

Scalar multiplication

Matrix multiplication

Why did we pick this weird rule for matrix multiplication? This is why!

What is a vector space?

A vector space is a set of vectors that can be added and can be multiplied by constants called scalars.

$$\{(a,b) \mid a,b \in \mathbb{R}\}$$

$$\text{and} \quad c = \mathbb{R}$$

(Vector spaces are what we've been working with this whole time!)

How do matrices fit in to this?

What is a vector space?

Matrices move vectors around the vector space, or even take vectors from one vector space to another!

Span and generating from vectors