Lab 11: Diagonalization Solutions

EECS 245, Fall 2025 at the University of Michigan

due by the end of your lab section on Wednesday, November 19th, 2025

Name: ____________________

uniqname: ____________________

Each lab worksheet will contain several activities, some of which will involve writing code and others that will involve writing math on paper. To receive credit for a lab, you must complete all activities and show your lab TA by the end of the lab section.

While you must get checked off by your lab TA **individually**, we encourage you to form groups with 1-2 other students to complete the activities together.

Recap: Eigenvalues and Eigenvectors

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

- An **eigenvector** of A is a non-zero vector $\vec{v}$ such that $A\vec{v} = \lambda\vec{v}$ for some scalar λ. The scalar λ is called the **eigenvalue** corresponding to $\vec{v}$. For A's eigenvectors, multiplying by A is equivalent to multiplying by a scalar.
- The **characteristic polynomial** of A is given by $p(\lambda) = \det(A - \lambda I)$.

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 6-\lambda & 3 \\ 3 & -2-\lambda \end{vmatrix} = (6-\lambda)(-2-\lambda) - 3 \cdot 3 = \lambda^2 - 4\lambda - 21 = (\lambda+3)(\lambda-7)$$

- The eigenvalues of A are the roots of the characteristic polynomial, so $\lambda_1 = -3$ and $\lambda_2 = 7$.
  - The eigenvector $\vec{v}_1$ satisfies $A\vec{v}_1 = -3\vec{v}_1$.

$$\begin{bmatrix} 6 & 3 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = -3 \begin{bmatrix} a \\ b \end{bmatrix} \implies b = -3a$$

    So any vector of the form $\begin{bmatrix} a \\ -3a \end{bmatrix}$ ($a \neq 0$) is an eigenvector of A corresponding to the eigenvalue -3. We could pick $\boxed{\vec{v}_1 = \begin{bmatrix} 2 \\ -6 \end{bmatrix}}$.
  - The eigenvector $\vec{v}_2$ satisfies $A\vec{v}_2 = 7\vec{v}_2$. Another way to find it is to solve for the null space of $A - 7I = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}$. One vector in $\text{nullsp}(A - 7I)$ is $\boxed{\vec{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}}$.

Recap: Diagonalization

- Since A has two linearly independent eigenvectors, it is **diagonalizable**, meaning we can write

$$A = V\Lambda V^{-1} = \underbrace{\begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} -3 & 0 \\ 0 & 7 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 0.05 & -0.15 \\ 0.3 & 0.1 \end{bmatrix}}_{V^{-1}}$$

  where V is an invertible matrix whose columns are the eigenvectors of A and Λ is a diagonal matrix with the eigenvalues of A on the diagonal.

  - The **algebraic multiplicity** of an eigenvalue, $\text{AM}(\lambda_i)$, is the number of times it appears as a root of the characteristic polynomial.

$$p(\lambda) = (\lambda - \lambda_1)^{\text{AM}(\lambda_1)}(\lambda - \lambda_2)^{\text{AM}(\lambda_2)} \cdots (\lambda - \lambda_k)^{\text{AM}(\lambda_k)}$$

  - The **geometric multiplicity** of an eigenvalue, $\text{GM}(\lambda_i)$, is the dimension of the eigenspace corresponding to λ_i.

    $\text{GM}(\lambda_i) = \dim(\text{nullsp}(A - \lambda_i I)) =$ # linearly independent eigenvectors corresponding to λ_i

  - For any λ_i, $1 \leq \text{GM}(\lambda_i) \leq \text{AM}(\lambda_i)$.
  - A is diagonalizable if and only if $\text{AM}(\lambda_i) = \text{GM}(\lambda_i)$ for all eigenvalues λ_i. This ensures that A's eigenvectors form a basis of $\mathbb{R}^n$.

- $A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$ has characteristic polynomial $p(\lambda) = (4 - \lambda)^2$. The eigenvalue $\lambda = 4$ has algebraic multiplicity 2, but the corresponding eigenspace is only 1-dimensional:

$$\text{nullsp}(A - 4I) = \text{nullsp}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

  so, $\lambda = 4$ has geometric multiplicity 1. Since $\text{AM}(4) \neq \text{GM}(4)$, A is **not diagonalizable**.

- $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$ has characteristic polynomial $p(\lambda) = (1 - \lambda)^2(3 - \lambda)$, so $\lambda_1 = 1$ has $\text{AM}(\lambda_1) = 2$ and $\lambda_2 = 3$ has $\text{AM}(\lambda_2) = 1$. The eigenspace for $\lambda_1 = 1$ is 2-dimensional,

$$\text{nullsp}(A - 1I) = \text{nullsp}\left(\begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right)$$

  so $\text{GM}(\lambda_1) = 2$. Since $\text{AM}(\lambda_i) = \text{GM}(\lambda_i)$ for all eigenvalues λ_i, A is **diagonalizable**.

$$A = V\Lambda V^{-1} = \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 0.5 & 0.5 & -0.5 \\ 0.5 & -0.5 & 0.5 \\ -0.5 & 0.5 & 0.5 \end{bmatrix}}_{V^{-1}}$$

- Diagonalizability is not the same as invertibility!

Activity 1: Fundamentals

Let $A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

a) Find the characteristic polynomial of A, and use it to find the eigenvalues of A and their algebraic multiplicities. *Tip: Add the formula for the determinant of a 3×3 matrix to your cheat sheet.*

Solution: The characteristic polynomial of A is

$$\begin{aligned}
p(\lambda) &= \det(A - \lambda I) \\
&= \begin{vmatrix} 3-\lambda & 2 & 0 \\ 2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} \\
&= (3-\lambda) \begin{vmatrix} 3-\lambda & 0 \\ 0 & 5-\lambda \end{vmatrix} - 2 \begin{vmatrix} 2 & 0 \\ 0 & 5-\lambda \end{vmatrix} + 0 \begin{vmatrix} 2 & 3-\lambda \\ 0 & 0 \end{vmatrix} \\
&= (3-\lambda)\left[(3-\lambda)(5-\lambda) - 0 \cdot 0\right] - 2\left[2(5-\lambda) - 0 \cdot 0\right] + 0 \\
&= (3-\lambda)\left[(3-\lambda)(5-\lambda)\right] - 2[2(5-\lambda)] \\
&= (3-\lambda)^2(5-\lambda) - 4(5-\lambda) \\
&= (5-\lambda)\left[(3-\lambda)^2 - 4\right] \\
&= (5-\lambda)\left[9 - 6\lambda + \lambda^2 - 4\right] \\
&= (5-\lambda)\left[\lambda^2 - 6\lambda + 5\right] \\
&= (5-\lambda)(\lambda-1)(\lambda-5) \\
&= (1-\lambda)(5-\lambda)^2
\end{aligned}$$

So, A has eigenvalues $\lambda_1 = 1$ with algebraic multiplicity $\mathrm{AM}(\lambda_1) = 1$ and $\lambda_2 = 5$ with algebraic multiplicity $\mathrm{AM}(\lambda_2) = 2$.

b) Find a basis for the eigenspace corresponding to each eigenvalue.

Solution:

- The eigenspace corresponding to $\lambda_1 = 1$ is

$$\text{nullsp}(A - 1I) = \text{nullsp}\left(\begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right)$$

  If you'd rather see this the systems of equations way, we're looking for a vector $\vec{v}_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ such that

$$\begin{aligned} A \begin{bmatrix} a \\ b \\ c \end{bmatrix} &= (1) \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} &= \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ \begin{bmatrix} 3a + 2b \\ 2a + 3b \\ 5c \end{bmatrix} &= \begin{bmatrix} a \\ b \\ c \end{bmatrix} \end{aligned}$$

  The first two equations both are equivalent, and say that $3a + 2b = a \implies 2a + 2b = 0 \implies b = -a$. The last equation just says $5c = c \implies 4c = 0 \implies c = 0$. The easy solution is to let $a = 1$, which gives $\boxed{\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}$. So, the eigenspace is the line of vectors spanned by $\vec{v}_1$.

- The eigenspace corresponding to $\lambda_2 = 5$ is

$$\text{nullsp}(A - 5I) = \text{nullsp}\left(\begin{bmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \boxed{\text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)}$$

  If you'd rather see this the systems of equations way, we're looking for a vector $\vec{v}_2 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ such that

$$\begin{bmatrix} 3a + 2b \\ 2a + 3b \\ 5c \end{bmatrix} = \begin{bmatrix} 5a \\ 5b \\ 5c \end{bmatrix}$$

  The first two equations both are equivalent and say that $a = b$. The last equation just says that $c = c$. So, the null space of $A - 5I$ is the set of all vectors with equal first and second components, while the third component can be anything. This is a space spanned by two vectors, $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

c) This particular A is diagonalizable. Diagonalize A by finding a matrix V and a diagonal matrix Λ such that $A = V\Lambda V^{-1}$.

Solution: In the previous part, we saw that $\lambda_1 = 1$ has the eigenvector $\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, while $\lambda_2 = 5$ corresponds to the linearly independent eigenvectors $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. So, we can construct V and Λ as follows:

$$V = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

And we have $A = V\Lambda V^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

Activity 2: Rapid Fire

The goal of this activity is to practice spotting eigenvalues and characteristic polynomials quickly. Recall that the sum of the eigenvalues of a matrix is equal to the trace of the matrix (which is the sum of the diagonal entries), and the product of the eigenvalues is equal to the determinant of the matrix.

a) A 2×2 matrix A has $\text{trace}(A) = 5$ and $\det(A) = 6$. What are the eigenvalues of A?

> **Solution:** $\lambda_1 = 2, \lambda_2 = 3$.
> The eigenvalues of A must sum to 5 and multiply to 6. The only possible solution is $\lambda_1 = 2$ and $\lambda_2 = 3$, which indeed satisfy both conditions. There are infinitely many **matrices** that satisfy these conditions, but the eigenvalues in all of them are 2 and 3.

b) Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$. What are the eigenvalues of A? Is A diagonalizable?

Hint: Use the fact that A is an upper triangular matrix. Chapter 5.1 has some hints.

> **Solution:** $\lambda_1 = 1, \lambda_2 = 4, \lambda_3 = 6$; A is diagonalizable.
>
> The characteristic polynomial of A is
>
> $$\begin{aligned} p(\lambda) &= \det(A - \lambda I) \\ &= \begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & 4-\lambda & 5 \\ 0 & 0 & 6-\lambda \end{vmatrix} \\ &= (1-\lambda)\begin{vmatrix} 4-\lambda & 5 \\ 0 & 6-\lambda \end{vmatrix} - 2\begin{vmatrix} 0 & 5 \\ 0 & 6-\lambda \end{vmatrix} + 3\begin{vmatrix} 0 & 4-\lambda \\ 0 & 0 \end{vmatrix} \\ &= (1-\lambda)\left[(4-\lambda)(6-\lambda) - 0\cdot 5\right] - 2\left[0\cdot(6-\lambda) - 0\cdot 5\right] + 3\left[0\cdot 0 - 0\cdot(4-\lambda)\right] \\ &= (1-\lambda)[(4-\lambda)(6-\lambda)] - 0 + 0 \\ &= (1-\lambda)(4-\lambda)(6-\lambda) \end{aligned}$$
>
> The eigenvalues of A are the solutions to $p(\lambda) = 0$, which are $\lambda_1 = 1$, $\lambda_2 = 4$, and $\lambda_3 = 6$. Since A has three distinct eigenvalues, it is diagonalizable.
>
> Why does knowing that all of A's eigenvalues are distinct imply that it's diagonalizable? The key is that for any matrix A, **eigenvectors for different eigenvalues must be linearly independent**. Each eigenvector points in a "new" direction relative to eigenvectors for other eigenvalues. An eigenvector can't correspond to multiple eigenvalues.
>
> From the perspective of multiplicities, remember that the minimum geometric multiplicity of an eigenvalue is 1. Since A has three distinct eigenvalues, they all have geometric multiplicities of 1, which match their algebraic multiplicities.

c) A non-invertible 2×2 matrix has an eigenvalue of 5. What is its characteristic polynomial?

Solution: $p(\lambda) = \lambda^2 - 5\lambda$.

Let A be the matrix in question. Since A is not invertible, 0 is one of its eigenvalues. Since it's a 2×2 matrix, it can only have two eigenvalues, so its two eigenvalues are 0 and 5. So, its characteristic polynomial is a polynomial with roots 0 and 5, i.e.

$$p(\lambda) = (\lambda - 0)(\lambda - 5) = \lambda^2 - 5\lambda$$

d) A 5×5 matrix has an eigenvalue of 0 with geometric multiplicity 2. What is its rank?

Solution: $\text{rank}(A) = 3$.

Let A be the matrix in question. Since it has an eigenvalue of 0 with geometric multiplicity 2, it has two linearly independent eigenvectors corresponding to 0. This means that the null space of $A - 0I$, which is the same as the null space of A, has dimension 2.

$$\dim(\text{nullsp}(A)) = 2$$

The rank-nullity theorem tells us that

$$\text{rank}(A) + \dim(\text{nullsp}(A)) = 5$$

Since $\dim(\text{nullsp}(A)) = 2$, we have

$$\text{rank}(A) = 5 - 2 = 3$$

e) A 3×3 matrix A has $\det(A) = 20$ and two unique **positive integer** eigenvalues, one with $\text{AM}(\lambda_1) = 2$ and another with $\text{AM}(\lambda_2) = 1$. What are all possible values of λ_1 and λ_2?

Solution: $\lambda_1 = 2, \lambda_2 = 5$ or $\lambda_1 = 1, \lambda_2 = 20$.

The product of A's eigenvalues — including arithmetic multiplicities — is equal to the determinant of A. Since λ_1 has $\text{AM}(\lambda_1) = 2$, it must appear twice in the product of the eigenvalues. So,

$$\lambda_1^2 \lambda_2 = 20$$

We're told that λ_1 and λ_2 are integers. Note that $20 = 2 \cdot 10 = 2 \cdot 2 \cdot 5$. The only possible solutions are $\lambda_1 = 2$ and $\lambda_2 = 5$ (which means $\lambda_1^2 \lambda_2 = (2)^2 \cdot 5 = 20$) or $\lambda_1 = 1$ and $\lambda_2 = 20$ (which means $\lambda_1^2 \lambda_2 = (1)^2 \cdot 20 = 20$).

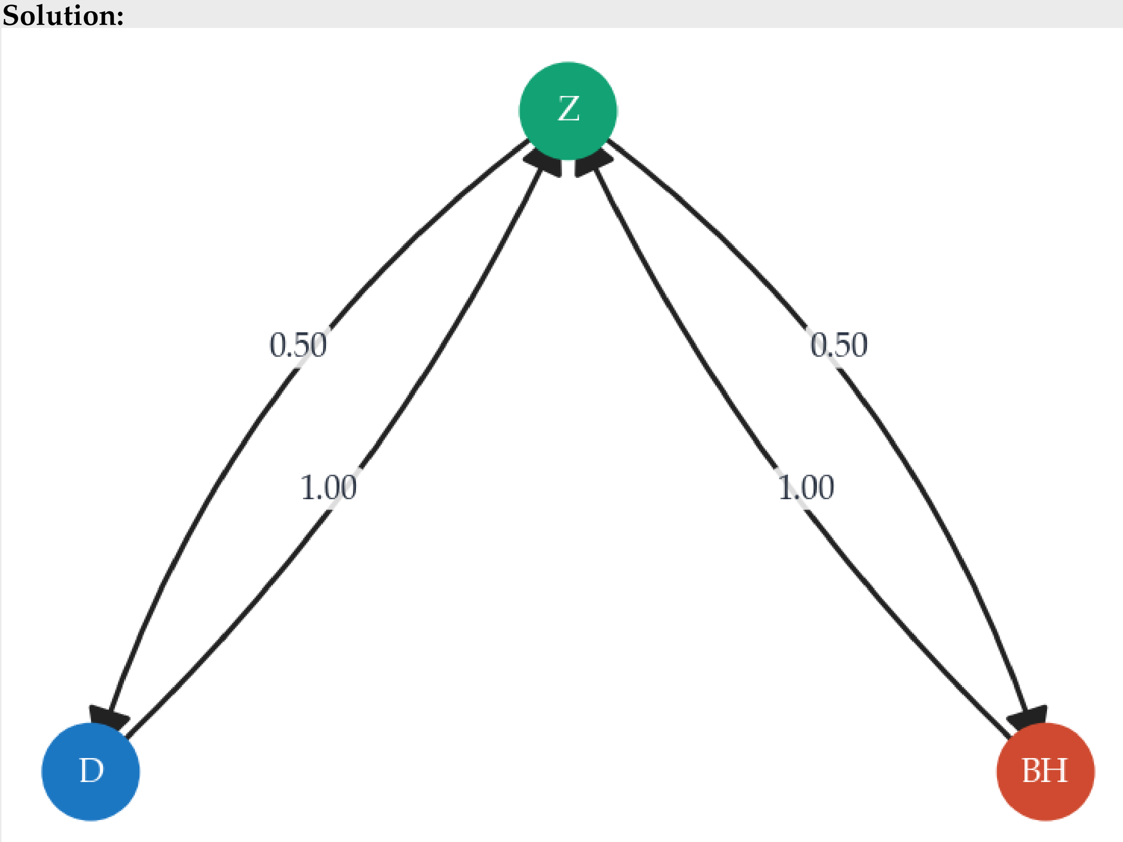
Activity 3: Adjacency Matrices

Suppose that a Wolverine moves between three classic Ann Arbor spots: the Diag, Zingerman's, and the Big House.

- From the Diag, it **always** walks to Zingerman's.
- From the Big House, it **always** walks to Zingerman's.
- From Zingerman's, it is **equally likely** to walk to the Diag or the Big House.

a) Draw a state diagram (like the one in Chapter 5.1, Part 2) for this Markov chain. Make sure to clearly label the edges with their transition probabilities.

Solution:



b) Find the adjacency matrix A.

Solution: $A = \begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 1/2 \\ 1 & 1 & 0 \end{bmatrix}.$

Let the Diag be state 1, so the first column corresponds to movement **from** the Diag. The second column corresponds to movement from the Big House, and the third column corresponds to movement from Zingerman's.

c) In the long run, what fraction of time does the Wolverine spend at each of the three spots? What does this have to do with the eigenvalues and eigenvectors of A?

Solution: **Important:** This Markov chain **does not converge!** This problem is more complicated than initially intended, but provides for a good learning opportunity.

In theory, to answer the question, we need to find the eigenvector of A corresponding to the eigenvalue 1. This will give us the vector $\vec{x}$ such that $A\vec{x} = \vec{x}$, i.e. where advancing one step in time doesn't change the vector.

Let this steady state vector be $\vec{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Then, we're looking for

$$\begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 1/2 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

which is equivalent to

$$\begin{aligned} 0a + 0b + 1/2c &= a \\ 0a + 0b + 1/2c &= b \\ 1a + 1b + 0c &= c \end{aligned}$$

The first equation gives us $a = c/2$ and the second equation gives us $b = c/2$ as well, so $a = b$. The third equation contributes no new information, but verifies what we've already found, since it requires $a + b = c$, which is true if $a = b = c/2$.

So, any vector of the form $\begin{bmatrix} a \\ a \\ 2a \end{bmatrix}$, i.e. on the line spanned by $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, is an eigenvector of A corresponding to the eigenvalue 1. The specific vector we're looking for is the one whose components sum to 1, so that we can interpret those components as probabilities. We need to find a such that $a + a + 2a = 1$, which gives us $a = 1/4$. So, one would think that the steady state vector is $\begin{bmatrix} 1/4 \\ 1/4 \\ 1/2 \end{bmatrix}$.

And, in most cases, this would be the correct answer. But there's something special about this Markov chain that makes it not converge.

It has to do with what A's other eigenvalues and eigenvectors are. Let's try and find them. can find the other eigenvalues and eigenvectors of $A = \begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 1/2 \\ 1 & 1 & 0 \end{bmatrix}$. We could find A's characteristic polynomial, but since we already know one of the eigenvalues, the trace and determinant "checks" might be easier.

- We know that all three eigenvalues need to add to the trace, $0+0+0=0$. Since one eigenvalue is 1, the other two – say, λ_2 and λ_3 – must add to -1.
- We know that the product of the eigenvalues is equal to the determinant. Since the matrix isn't invertible (columns 1 and 2 are the same), the determinant is 0. So, the other two eigenvalues must multiply to 0.
- The other two eigenvalues must add to -1 and multiply to 0, so they must be $\lambda_2 = 0$ and $\lambda_3 = -1$.

An eigenvector for $\lambda_2 = 0$ is $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, since $A\vec{v}_2 = \vec{0}$ (I found this by noticing that column 1 – column 2 $= \vec{0}$). An eigenvector for $\lambda_3 = -1$ is $\vec{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$, which you can find by solving the system of equations $A\vec{v}_3 = -1\vec{v}_3$, or by inspection.

So, to recap, we have

$$\lambda_1 = 1, \vec{v}_1 = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/2 \end{bmatrix}, \qquad \lambda_2 = 0, \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \qquad \lambda_3 = -1, \vec{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

You might remember from Chapter 5.1, Part 2 that the vector $A^k\vec{x}$, for large k, converges to the eigenvector corresponding to the dominant eigenvalue, which is the eigenvalue with the largest magnitude (absolute value). You might also remember that adjacency matrices are guaranteed to have a largest eigenvalue of 1. **What complicates this situation is that it also has an eigenvalue of -1, which has the same magnitude as 1!** So, the vector $A^k\vec{x}$ doesn't actually converge to $\vec{v}_1$; rather, it "bounces" or "oscillates" between vectors that depend on the initial state vector $\vec{x}$. See our simulation on the next page.

Conceptually, this happens because this Markov chain is **periodic**, which means that the Wolverine must always travel in a specific cycle of states — if it's at Zingerman's at time t, we know for sure it will be at Zingerman's at time $t+2$, $t+4$, etc. This can happen whenever an adjacency matrix has an eigenvalue of -1.

Check out our simulation below. We encourage you to test this out yourself too!

```python
A = np.array([
    [0, 0, 1/2],
    [0, 0, 1/2],
    [1, 1,   0 ]
])

x = np.array([[1/3], [1/3], [1/3]])
for k in range(50, 55):
    x_k = np.linalg.matrix_power(A, k) @ x
    print(f'x_{k} = ', x_k.flatten())
```

✓ 0.0s

```
x_50 =  [0.33333333 0.33333333 0.33333333]
x_51 =  [0.16666667 0.16666667 0.66666667]
x_52 =  [0.33333333 0.33333333 0.33333333]
x_53 =  [0.16666667 0.16666667 0.66666667]
x_54 =  [0.33333333 0.33333333 0.33333333]
```

```python
# If we start with the eigenvector for lambda=1,
# it has already converged.
x = np.array([[1/4], [1/4], [1/2]])
for k in range(50, 55):
    x_k = np.linalg.matrix_power(A, k) @ x
    print(f'x_{k} = ', x_k.flatten())
```

✓ 0.0s

```
x_50 =  [0.25 0.25 0.5 ]
x_51 =  [0.25 0.25 0.5 ]
x_52 =  [0.25 0.25 0.5 ]
x_53 =  [0.25 0.25 0.5 ]
x_54 =  [0.25 0.25 0.5 ]
```

```python
# But if we start somewhere else, it oscillates.
x = np.array([[1/2], [1/4], [1/4]])
for k in range(50, 55):
    x_k = np.linalg.matrix_power(A, k) @ x
    print(f'x_{k} = ', x_k.flatten())
```

✓ 0.0s

```
x_50 =  [0.375 0.375 0.25 ]
x_51 =  [0.125 0.125 0.75 ]
x_52 =  [0.375 0.375 0.25 ]
x_53 =  [0.125 0.125 0.75 ]
x_54 =  [0.375 0.375 0.25 ]
```

Activity 4: Quadratic Forms Return

Open Desmos in 3D mode at **desmos.com/3d** and write $z = x^2 + 2bxy + 16y^2$. This should show you a 3D surface along with a slider for b. Drag the slider to see how the shape of the surface changes for different b's. You should notice that depending on the value of b, the surface may or may not have a global minimum. Let's explore!

a) z is a quadratic form, $f(\vec{x}) = \vec{x}^T A \vec{x}$, where $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and A is a symmetric matrix. Find A.

> **Solution:** $A = \begin{bmatrix} 1 & b \\ b & 16 \end{bmatrix}$.
>
> To see where this comes from, let's expand $f(\vec{x}) = \vec{x}^T A \vec{x}$:
>
> $$\begin{aligned} f(\vec{x}) &= \vec{x}^T A \vec{x} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & b \\ b & 16 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} x + by \\ bx + 16y \end{bmatrix} \\ &= x(x + by) + y(bx + 16y) \\ &= x^2 + bxy + bxy + 16y^2 \\ &= x^2 + 2bxy + 16y^2 \end{aligned}$$
>
> So, $A = \begin{bmatrix} 1 & b \\ b & 16 \end{bmatrix}$.

b) For a vector-to-scalar function $f : \mathbb{R}^n \to \mathbb{R}$, the **Hessian** of f, denoted $\nabla^2 f$, is the $n \times n$ matrix of second partial derivatives of f. Find $\nabla^2 f$ for $f(\vec{x}) = \vec{x}^T A \vec{x}$.

Solution: $\nabla^2 f = \begin{bmatrix} 2b & 2b \\ 2b & 32 \end{bmatrix}$.

To find the second partial derivatives of f, we first find the first partial derivatives:

$$f(\vec{x}) = \vec{x}^T A \vec{x} = x^2 + 2bxy + 16y^2$$
$$\frac{\partial f}{\partial x} = 2x + 2by$$
$$\frac{\partial f}{\partial y} = 2bx + 32y$$

Since there are two first partial derivatives, there are $2 \times 2 = 4$ second partial derivatives. We can find them all by taking the partial derivatives of the first partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(2x + 2by) = 2$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(2x + 2by) = 2b$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(2bx + 32y) = 2b$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(2bx + 32y) = 32$$

Note that $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, which is a general property of second partial derivatives — it doesn't matter which order we take the derivatives in. So, the Hessian is

$$\nabla^2 f = \begin{bmatrix} 2 & 2b \\ 2b & 32 \end{bmatrix}$$

Notice that this is just $2A$!

c) A symmetric matrix A is **positive semidefinite** (PSD) if $\vec{v}^T A \vec{v} \geq 0$ for all $\vec{v} \in \mathbb{R}^n$. In English, this says that A is positive semidefinite if the quadratic form $f(\vec{v}) = \vec{v}^T A \vec{v}$ is always non-negative for all $\vec{v} \in \mathbb{R}^n$. Two relevant facts:

- A differentiable vector-to-scalar function f is **convex** if its Hessian is PSD.
- A symmetric matrix A is PSD if and only if all of its eigenvalues are non-negative.

Using the facts above, find the range of values b for which f is convex, and verify your answer by dragging the slider on Desmos.

Solution: As long as $-4 \leq b \leq 4$, the Hessian is PSD, and f is convex.

Recall from the previous part that the Hessian is

$$\nabla^2 f = \begin{bmatrix} 2 & 2b \\ 2b & 32 \end{bmatrix} = 2A$$

f is convex if and only if $2A$ is PSD, which means that all both of $2A$'s eigenvalues are non-negative. Let λ_1 and λ_2 be $2A$'s eigenvalues. Then, we need $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. How do we ensure this?

Let's recall the facts about the trace and determinant introduced at the start of Activity 2:

- The sum of $2A$'s eigenvalues, $\lambda_1 + \lambda_2$, is equal to the trace of $2A$, trace$(2A)$, which is the sum of the diagonal entries of $2A$. Here, this is $2 + 32 = 34$.
- The product of $2A$'s eigenvalues, $\lambda_1 \lambda_2$, is equal to the determinant of $2A$, $\det(2A)$, which is $2 \cdot 32 - (2b)^2 = 64 - 4b^2$.

The fact that $\lambda_1 + \lambda_2 = 34$ means that no matter what b is, at least one of the eigenvalues is non-negative, since you can't add two negative numbers and get a positive sum.

So, let's focus our attention on the product of the eigenvalues, $\lambda_1 \lambda_2 = 64 - 4b^2$. We need this to be ≥ 0, because if it were negative, then one of the eigenvalues would be negative (since multiplying a positive number by a negative number gives a negative number).

Using this reasoning, we need $64 - 4b^2 \geq 0$, which simplifies to $4b^2 \leq 64$, which simplifies to $b^2 \leq 16$, which simplifies to $-4 \leq b \leq 4$. If $|b| > 4$, then one of the eigenvalues would be negative, $2A$ would not be PSD, and f would not be convex.

Notice that in this case, the "test" for convexity ended up simplifying to checking whether the determinant of A is non-negative. That is true, in general, for quadratic forms $f(\vec{x}) = \vec{x}^T A \vec{x}$ where A is a 2×2symmetric matrix. But if A is 3×3 or larger, the determinant test alone isn't sufficient; it's possible to have a positive determinant and trace but a negative eigenvalue in a 3×3 matrix. What **is** true in general is that if A is a symmetric $n \times n$ matrix, then $f(\vec{x}) = \vec{x}^T A \vec{x}$ is convex if and only if all of A's eigenvalues are non-negative.

We don't expect you to finish all of the following activities in lab, but you should complete them afterwards for practice; they're in scope for the Final Exam.

Activity 5: Symmetric Matrices

Suppose A is a symmetric $n \times n$ matrix, meaning $A = A^T$. Symmetric matrices have a special property that we'll use more in Thursday's lecture: **the eigenvectors corresponding to different eigenvalues are orthogonal**. Below, we provide a proof of this fact.

1. Let $\vec{v}_i$ be an eigenvector of A with eigenvalue λ_i, so $A\vec{v}_i = \lambda_i \vec{v}_i$.
2. Let $\vec{v}_j$ be an eigenvector of A with eigenvalue λ_j, where $\lambda_i \neq \lambda_j$, so $A\vec{v}_j = \lambda_j \vec{v}_j$.
3. The dot product of $\vec{v}_i$ and $A\vec{v}_j$ is $\vec{v}_i \cdot (A\vec{v}_j) = \vec{v}_i \cdot (\lambda_j \vec{v}_j) = \lambda_j(\vec{v}_i \cdot \vec{v}_j)$.
4. But also, $\vec{v}_i \cdot (A\vec{v}_j) = \vec{v}_i^T A \vec{v}_j = \vec{v}_i^T A^T \vec{v}_j = (A\vec{v}_i)^T \vec{v}_j = \lambda_i \vec{v}_i^T \vec{v}_j = \lambda_i(\vec{v}_i \cdot \vec{v}_j)$.
5. The final expressions in both cases are equal, so $\lambda_j(\vec{v}_i \cdot \vec{v}_j) = \lambda_i(\vec{v}_i \cdot \vec{v}_j)$.
6. Equivalently, $(\lambda_j - \lambda_i)(\vec{v}_i \cdot \vec{v}_j) = 0$. But since $\lambda_i \neq \lambda_j$, we must have $\vec{v}_i \cdot \vec{v}_j = 0$.

a) In which line did we use the fact that A is symmetric? Select the line below, and then in that line, circle the specific part of the equation that uses the fact that $A = A^T$.

○ 1 ○ 2 ○ 3 ● 4 ○ 5 ○ 6

> **Solution:** Line 4 is the only line that uses the fact that $A = A^T$, but it's the most important line in the proof. Specifically, here's where we used that fact:
>
> $$\vec{v}_i^T A \vec{v}_j = \vec{v}_i^T A^T \vec{v}_j$$
>
> By using $A = A^T$ here, we were able to write $\vec{v}_i \cdot (A\vec{v}_j)$ as $\vec{v}_i^T A^T \vec{v}_j = (A\vec{v}_i) \cdot \vec{v}_j$, which was key to our final expression.

b) The **spectral theorem** states that any symmetric $n \times n$ matrix A can be diagonalized by an orthogonal matrix Q through

$$A = Q\Lambda Q^T$$

Explain how this relates to the "regular" diagonalization $A = V\Lambda V^{-1}$.

Solution: A standard (not necessarily symmetric) matrix A that is diagonalizable can be written as

$$A = V\Lambda V^{-1}$$

If A is symmetric and can be written $A = Q\Lambda Q^T$, then the fact that we have Q^T instead of Q^{-1} stems from the fact that orthogonal matrices satisfy $Q^T = Q^{-1}$. This comes from the fact that $Q^T Q = QQ^T = I$. Transposes are easier to compute than inverses, and easier to interpret, too: if Q is a rotation, then Q^T is the inverse rotation, i.e. a rotation by the opposite amount.

Activity 6: Positive Semidefinite Matrices

As mentioned in Activity 4, a symmetric matrix A is **positive semidefinite** if $\vec{v}^T A\vec{v} \geq 0$ for all $\vec{v} \in \mathbb{R}^n$.

a) Prove that if A is positive semidefinite, then all of its eigenvalues are non-negative. *Hint: You can assume that all of A's eigenvalues are real numbers; this is true for all symmetric matrices. Start by considering $\vec{v}_i^T A\vec{v}_i$ for an eigenvector $\vec{v}_i$ of A with eigenvalue λ_i.*

> **Solution:** As the hint states, let's suppose that A is positive semidefinite, and let λ_i be an eigenvalue of A with eigenvector $\vec{v}_i$. Then, we have
>
> $$A\vec{v}_i = \lambda_i \vec{v}_i$$
>
> Similar to one of the proofs in Chapter 2.8, where we proved that $\text{rank}(X) = \text{rank}(X^T X)$, let's multiply both sides by $\vec{v}_i^T$ on the left and see where this gets us.
>
> $$\vec{v}_i^T A\vec{v}_i = \lambda_i \vec{v}_i^T \vec{v}_i$$
>
> On the right, we have $\vec{v}_i^T \vec{v}_i$, which is just $\vec{v}_i \cdot \vec{v}_i$, which is just $\|\vec{v}_i\|^2$.
>
> $$\vec{v}_i^T A\vec{v}_i = \lambda_i \|\vec{v}_i\|^2 \implies \lambda_i = \frac{\vec{v}_i^T A\vec{v}_i}{\|\vec{v}_i\|^2}$$
>
> Since A is positive semidefinite, we have $\vec{v}_i^T A\vec{v}_i \geq 0$, so the numerator on the right side is non-negative. But $\|\vec{v}_i\|^2$ is always non-negative (since it's the square of the magnitude of a vector), so $\lambda_i \geq 0$.

b) Prove that if all of A's eigenvalues are non-negative, then A is positive semidefinite. *Hint: A is symmetric, which means it has a special version of the eigenvector decomposition, hinted in the previous activity. Consider $\vec{x}^T A\vec{x}$ for some arbitrary (not necessarily an eigenvector) vector $\vec{x}$, and then replace A with its special eigenvector decomposition.*

Solution: Let's let the hint guide us. Since A is symmetric (which is a prerequisite for being positive semidefinite), **its eigenvectors for different eigenvalues are orthogonal, and so it can be written as $A = Q\Lambda Q^T$, where Q is an orthogonal matrix and Λ is a diagonal matrix with the eigenvalues of A on the diagonal.** This is the eigenvector decomposition of A.

On top of that, suppose that all of A's eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_n$, are non-negative.

As the hint suggests, let $\vec{x}$ be some arbitrary vector (not necessarily an eigenvector) in $\mathbb{R}^n$. Then, eventually we need to show that $\vec{x}^T A\vec{x} \geq 0$, regardless of what $\vec{x}$ is. Let's start by expanding $\vec{x}^T A\vec{x}$ using the fact that $A = Q\Lambda Q^T$.

$$\vec{x}^T A\vec{x} = \vec{x}^T (Q\Lambda Q^T)\vec{x} = (\vec{x}^T Q)\Lambda(Q^T\vec{x})$$

Suppose that $\vec{y} = Q^T\vec{x}$. Then, $\vec{y}^T = (Q^T\vec{x})^T = \vec{x}^T Q$. This seems like an arbitrary maneuver, but check out what it buys us:

$$\begin{aligned}
\vec{x}^T A\vec{x} &= \vec{y}^T\Lambda\vec{y} \\
&= \begin{bmatrix} y_1 & y_2 & \ldots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\
&= \begin{bmatrix} y_1 & y_2 & \ldots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 y_1 \\ \lambda_2 y_2 \\ \vdots \\ \lambda_n y_n \end{bmatrix} \\
&= \sum_{i=1}^{n} \lambda_i y_i^2
\end{aligned}$$

Remember that none of the λ_i's are negative. So, each term $\lambda_i y_i^2$ is non-negative too, meaning the entire sum $\sum_{i=1}^{n} \lambda_i y_i^2$ is non-negative. What we've shown is that if $\vec{x}$ is **any** vector in $\mathbb{R}^n$, then $\vec{x}^T A\vec{x}$ ends up being a sum of this form, which is always non-negative, so $\vec{x}^T A\vec{x} \geq 0$, and thus A is positive semidefinite.

Activity 7: Diagonalization in Action

Let A be a 3×3 with:

- eigenvalue $\lambda_1 = 3$ with eigenvector $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.
- eigenvalue $\lambda_2 = -2$ with the 2-dimensional eigenspace: span $\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right)$.

a) Find A. Feel free to use `numpy` to find an inverse for you and verify your answer.

Solution: $A = \begin{bmatrix} 0.5 & -2.5 & 5 \\ 0 & -2 & 0 \\ 1.25 & -1.25 & 0.5 \end{bmatrix}$.

The eigenvector decomposition of A is $A = V \Lambda V^{-1}$, where V is the matrix whose columns are the eigenvectors of A and Λ is the diagonal matrix with the eigenvalues of A on the diagonal, arranged in the same order as the eigenvectors in V. We know that $\lambda_1 = 3$ with eigenvector $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ and that $\lambda_2 = \lambda_3 = -2$ with eigenvectors $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$. Placing this information into V and Λ gives us

$$V = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

So,

$$A = V \Lambda V^{-1} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \left(\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \right)^{-1}$$

A bit of help from `numpy` shows that $A = \begin{bmatrix} 0.5 & -2.5 & 5 \\ 0 & -2 & 0 \\ 1.25 & -1.25 & 0.5 \end{bmatrix}$.

```
>>> V = np.array([
    [2, 1, 0],
    [0, 1, 2],
    [1, 0, 1]
])
>>> V @ np.diag([3, -2, -2]) @ np.linalg.inv(V)
array([[ 0.5 , -2.5 ,  5.  ],
       [ 0.  , -2.  ,  0.  ],
       [ 1.25, -1.25,  0.5 ]])
```

b) Let V be the matrix whose columns are the eigenvectors of A and Λ be the diagonal matrix with the eigenvalues of A on the diagonal. In terms of V and Λ, what is A^8?

Solution: $A^8 = V\Lambda^8 V^{-1}$.
This comes from the fact that $A = V\Lambda V^{-1}$, so

$$A^8 = V\Lambda(V^{-1}V)\Lambda(V^{-1}V)\Lambda V^{-1}...V\Lambda^8 V^{-1} = V\Lambda V^{-1}$$

The sequential multiplication of V^{-1} with V cancels out, and we stack together 8 copies of Λ.
Since Λ is a diagonal matrix, we can raise each diagonal element to the 8th power.

$$\Lambda^8 = \begin{bmatrix} 3^8 & 0 & 0 \\ 0 & (-2)^8 & 0 \\ 0 & 0 & (-2)^8 \end{bmatrix}$$

So, $A^8 = V \begin{bmatrix} 3^8 & 0 & 0 \\ 0 & (-2)^8 & 0 \\ 0 & 0 & (-2)^8 \end{bmatrix} V^{-1}$.

c) Suppose $\vec{x} \in \mathbb{R}^3$ is some vector. As $k \to \infty$, what does $A^k\vec{x}$ approach? Explain your answer in English.

Solution: As $k \to \infty$, $A^k\vec{x}$ approaches the eigenvector corresponding to $\lambda_1 = 3$ (the dominant eigenvalue), so some scalar multiple of $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.

Remember that $A^k = V\Lambda^k V^{-1}$. As k increases, the diagonal entries of Λ^k — $3^k, (-2)^k, (-2)^k$ — all grow exponentially larger, but 3^k grows faster than $(-2)^k$. So, the contribution of the second and third eigenvectors to $A^k\vec{x}$ diminishes, and $A^k\vec{x}$ approaches a scalar multiple of $\vec{v}_1$.

To be clear, the numbers in $A^k\vec{x}$ will approach infinity; it's the **direction** of $A^k\vec{x}$ that approaches the direction of $\vec{v}_1$. Simulate this in `numpy` to see for yourself!

```
>>> x = np.array([[1], [1], [1]])
>>> A = np.array([
    [0.5, -2.5, 5],
    [0, -2, 0],
    [1.25, -1.25, 0.5]
])
>>> x_k = np.linalg.matrix_power(A, 50) @ x
>>> x_k # Massive numbers!
array([[7.17897988e+23],
       [1.12589991e+15],
       [3.58948994e+23]])
>>> x_k / np.linalg.norm(x_k) # Unit vector makes numbers smaller.
array([[8.94427191e-01],
       [1.40275570e-09],
       [4.47213596e-01]])
# Notice that the unit vector is roughly [2, 0, 1] multiplied by a scalar.
```