$$X^T X + \lambda I$$

why is this always invertible?

square, $\underbrace{n \times n \text{ matrix } A}$

think about linear transformations
from $\mathbb{R}^n \to \mathbb{R}^n$

(non-zero)
eigenvector $\vec{v}$:

$$A\vec{v} = \lambda \vec{v}$$

when multiplied by A, $\vec{v}$'s direction
didn't change; it was just
scaled by a factor of λ

eigenvalue

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

if $\vec{v}$ is an eigenvector of A, so is $c\vec{v}$, for any $c \neq 0$

Let $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$A\vec{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\Rightarrow$ $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ eigenvector of A with $\lambda_1 = 3$

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

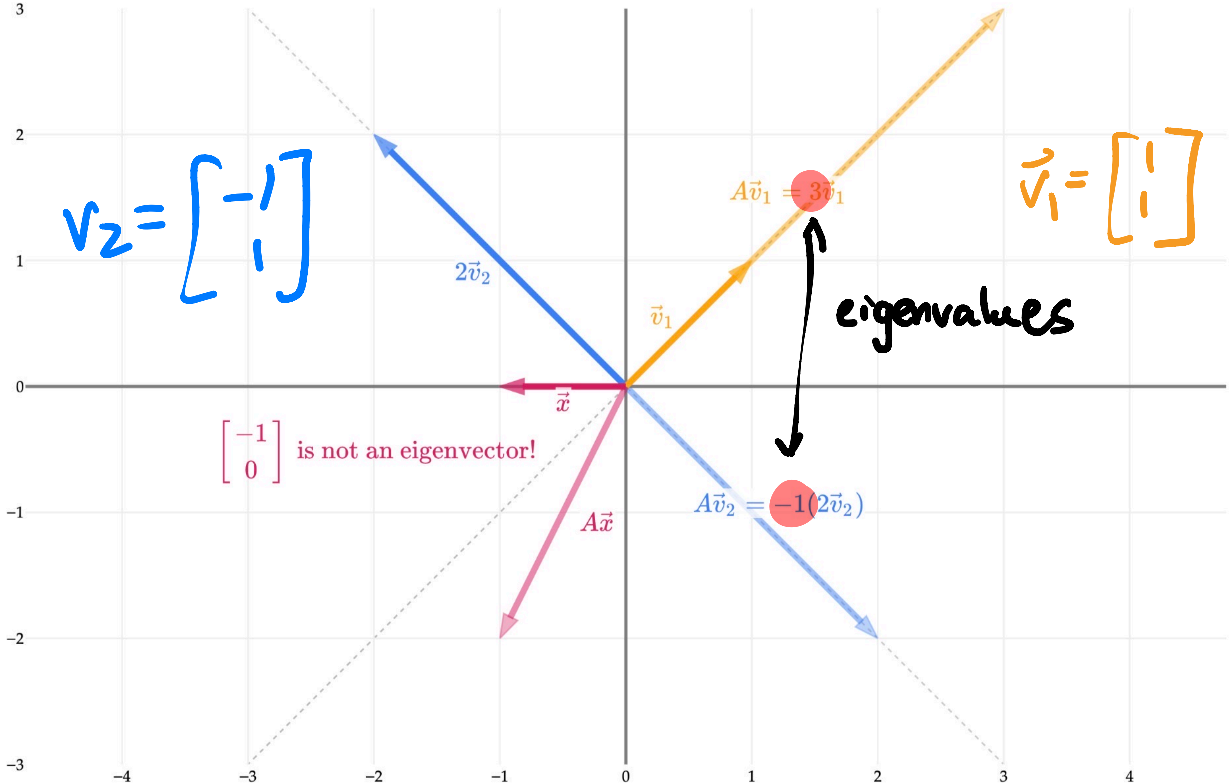
$\Rightarrow$ so $\vec{v} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ is <u>also</u> eigvec of A with $\lambda = 3$

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

$$= (-1) \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

so, $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_2 = -1$

Visualizing the eigenvectors of $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

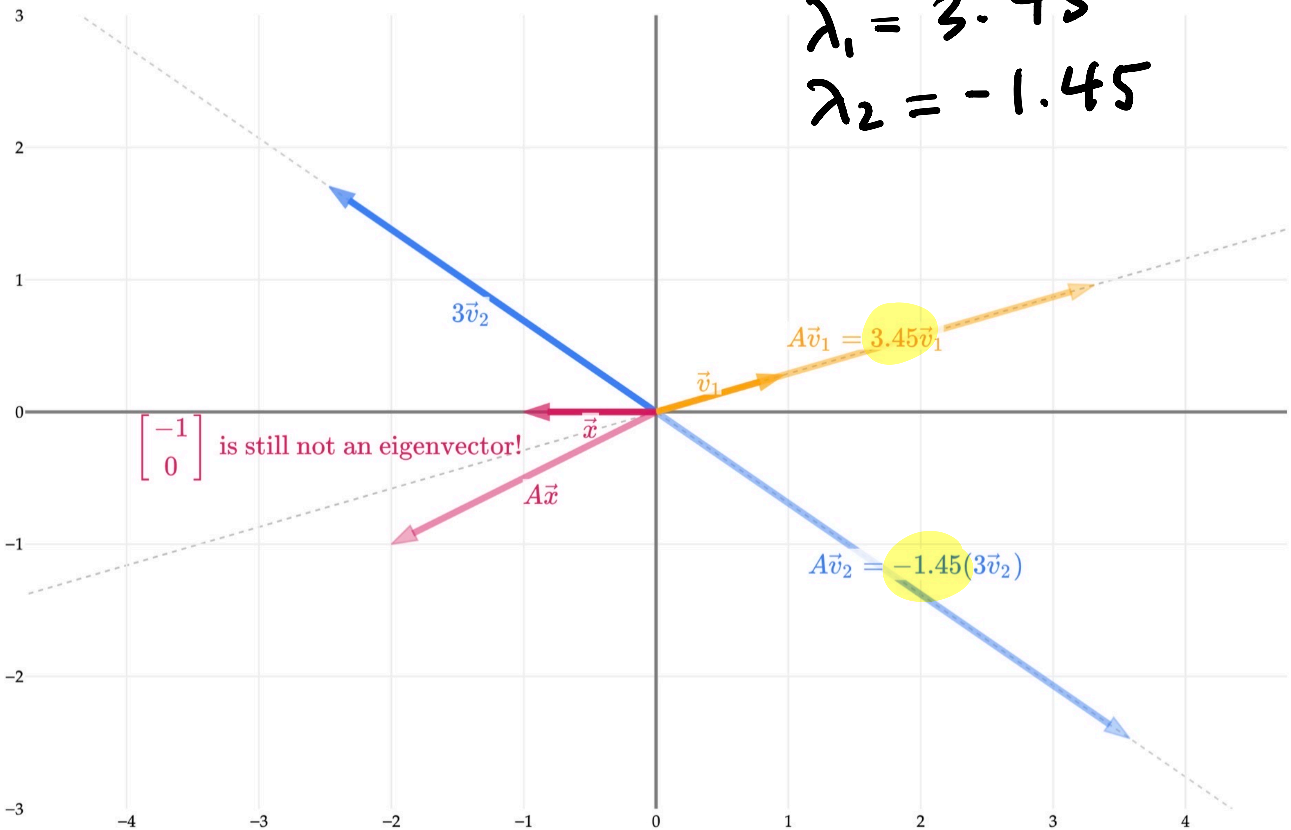


Visualizing the eigenvectors of $B = \begin{bmatrix} 2 & 5 \\ 1 & 0 \end{bmatrix}$

Here,

$\lambda_1 = 3.45$

$\lambda_2 = -1.45$



```python
B = np.array([[2, 5],
              [1, 0]])

np.linalg.eig(B)
```

```
EigResult(eigenvalues=array([ 3.44948974, -1.44948974]), eigenvectors=array([[
0.96045535, -0.82311938],
       [ 0.27843404,  0.56786837]]))
```

cols are the eigvecs

$\vec{v}_1$ $\vec{v}_2$

observe: $\lambda_1 + \lambda_2 = 2$, which is sum of diagonal!

True in general: sum of λ_i's = sum of diagonal

$$A = \begin{bmatrix} 1 & 4 \\ 3 & 12 \end{bmatrix}$$

notice: $\text{rank}(A) = 1$

$\lambda_1 = 13$, $\lambda_2 = \underline{0}$?

$\vec{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$

basis for nullsp(A)!

$\lambda_1 = 13$ ✓ $\begin{bmatrix} 1 & 4 \\ 3 & 12 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 13x \\ 13y \end{bmatrix}$

$x + 4y = 13x \Rightarrow 4y = 12x \Rightarrow y = 3x$

$3x + 12y = 13y \Rightarrow 3x = y$

both same, since any vector on that line works!

e.g. $x = 1$, so $y = 3$,

so $\vec{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

A has an eigenvalue of 0

$\Updownarrow$

A is <u>not</u> invertible

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

$\lambda_1 = 3$ with $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = -1$ with $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

what are the eigenvalues of A^2?

what eigenvectors?

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

$$= \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

if $\vec{v}$ is an eigvec with eigval λ of A, then

$$A^2\vec{v} = AA\vec{v} = A(\lambda\vec{v}) = \lambda(A\vec{v})$$
$$= \lambda(\lambda\vec{v})$$
$$= \lambda^2\vec{v}$$

only one direction!

then, $\vec{v}$ is an eigvec of A^2 with eigval λ^2.

possible for A^2 to have an eigenvector that A doesn't have

think: rotations

$$A\vec{v} = \lambda\vec{v}$$

$$A\vec{v} - \lambda\vec{v} = \vec{0}$$

$$(A - \lambda I)\vec{v} = \vec{0}$$

identity matrix

$\vec{v}$ is in $A - \lambda I$'s null space

$\Rightarrow$ $A - \lambda I$ has a non-trivial null space

$\Rightarrow$ $\det(A - \lambda I) = 0$ $\rightarrow$ use this to find λ's!

characteristic polynomial

$$p(\lambda) = \det(A - \lambda I)$$

degree n polynomial

eigenvalues are solutions to

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

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

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} \right)$$

$$= (1-\lambda)^2 - 4$$

$\lambda = -1,$

$\lambda = 3$

$$= \lambda^2 - 2\lambda + 1 - 4$$

$$= \lambda^2 - 2\lambda - 3 = (\lambda+1)(\lambda-3) = 0$$

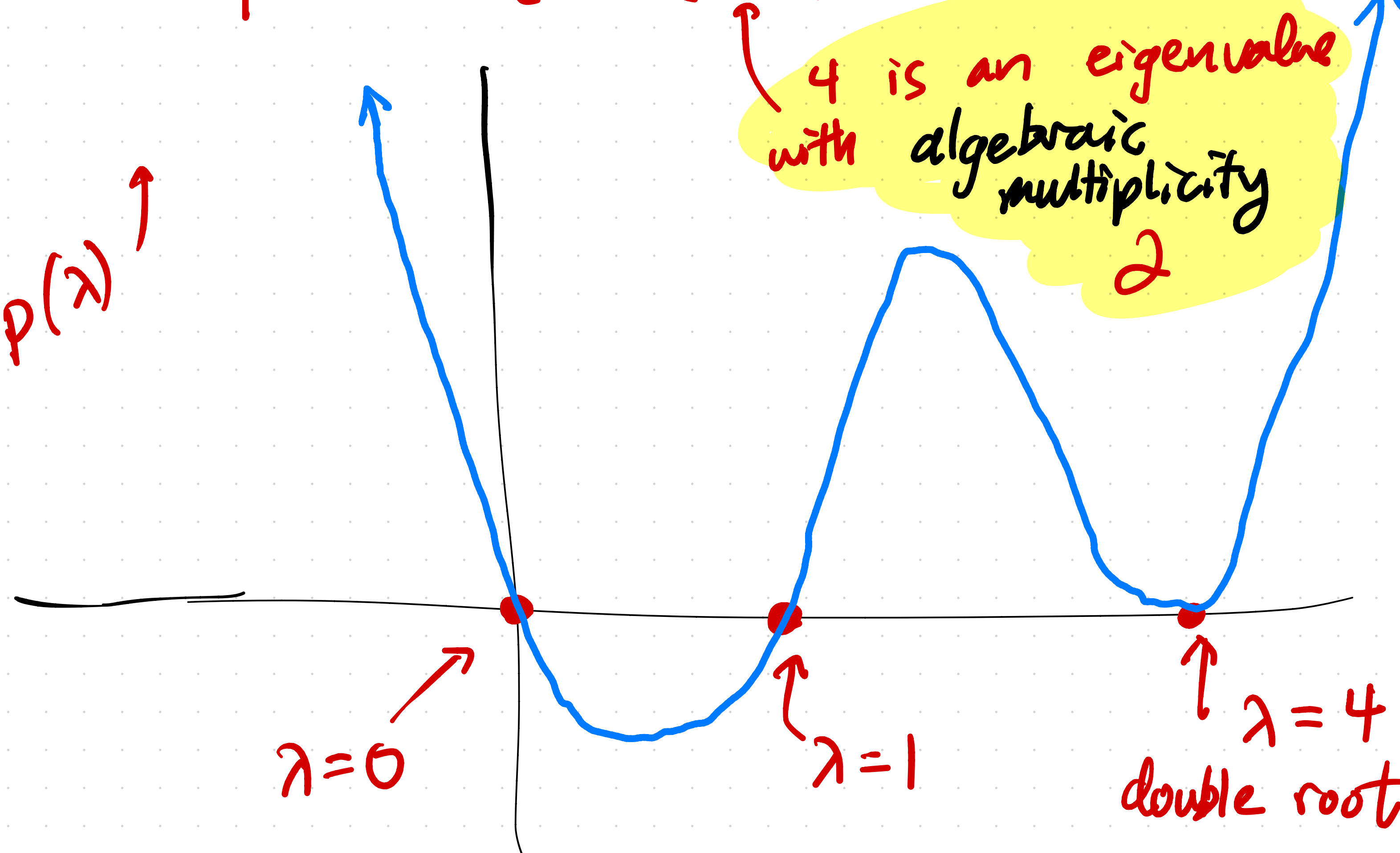
$$A = \begin{bmatrix} 4 & & & 0 \\ & 1 & & \\ & & 0 & \\ 0 & & & 4 \end{bmatrix}$$

$$p(\lambda) = \det(A - \lambda I) = \det\left(\begin{bmatrix} 4-\lambda & & & 0 \\ & 1-\lambda & & \\ & & 0-\lambda & \\ 0 & & & 4-\lambda \end{bmatrix}\right)$$

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

$$= \lambda(\lambda-1)(\lambda-4)^2$$

$$p(\lambda) = \lambda(\lambda-1)(\lambda-4)^2$$

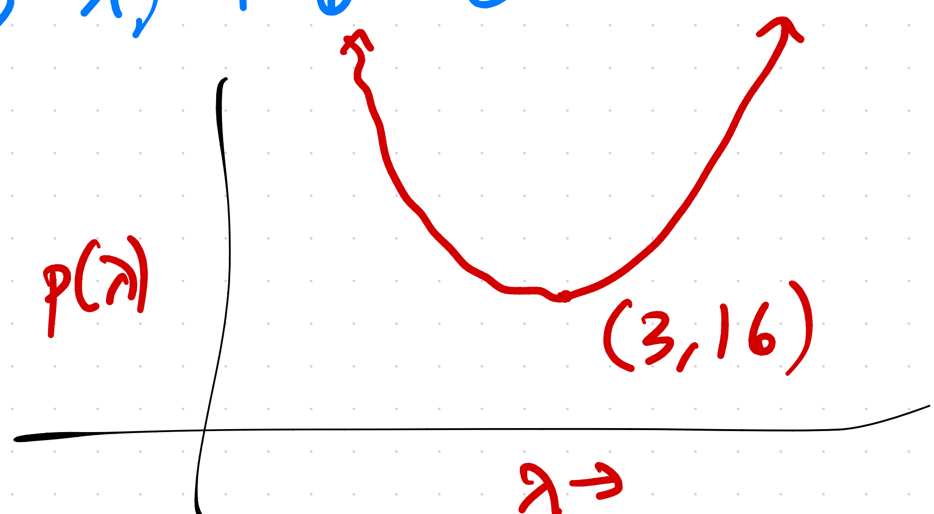


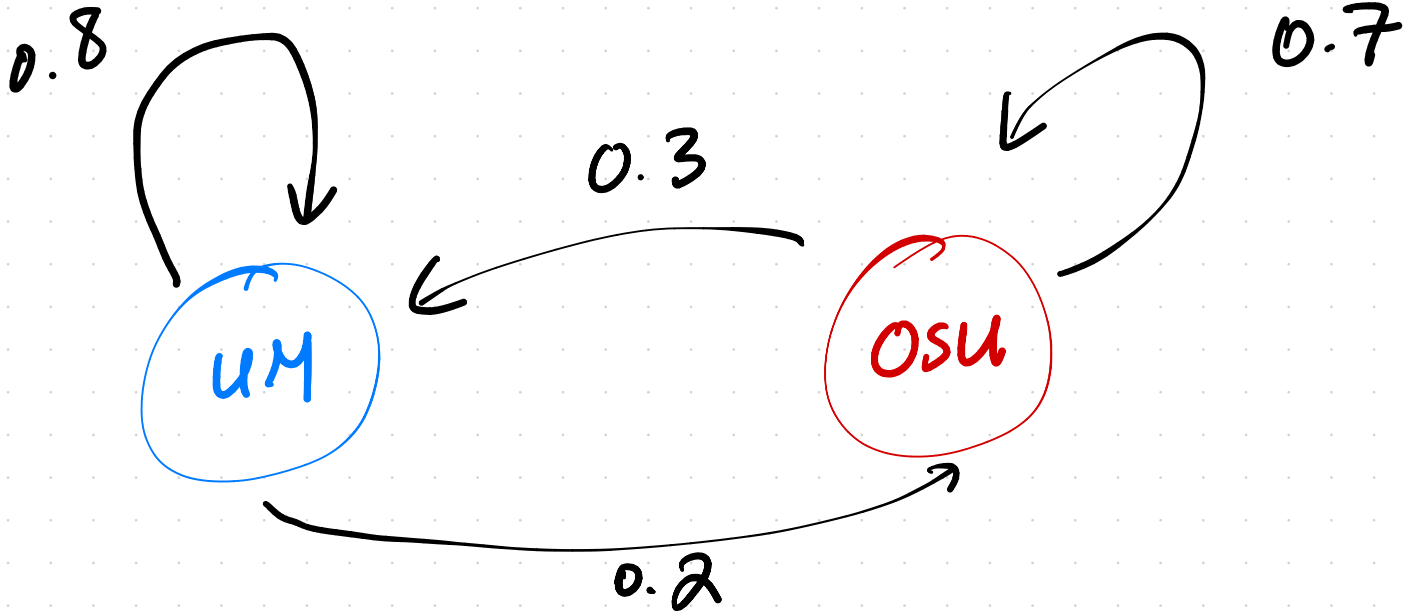
$$A = \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix} = \underbrace{5 \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}}_{\text{rotation, then stretch}}$$

$5e^{i\theta}$ $\quad 5e^{-i\theta}$

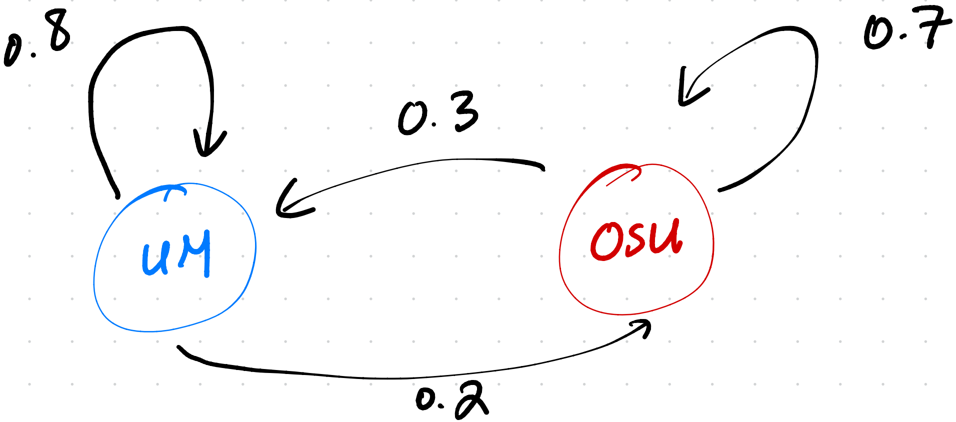
$\theta = \cos^{-1}(3/5)$

$$p(\lambda) = (3-\lambda)^2 + 16 = 0$$





in the long run, what % of games does Michigan win?



$$A = \begin{array}{cc} & \\ \left[\begin{array}{cc} 0.8 & 0.3 \\ 0.2 & 0.7 \end{array}\right] & \begin{array}{l} \to \text{UM} \\ \to \text{OSU} \end{array} \\ \begin{array}{cc} \text{UM}\to & \text{OSU}\to \end{array} & \end{array}$$

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

$$\vec{x}_k = \begin{bmatrix} p(\text{Michigan})_k \\ p(\text{OSU})_k \end{bmatrix}$$

simulate!

$$\vec{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\vec{x}_1 = A\vec{x}_0 = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix}$$

$$\vec{x}_2 = A\vec{x}_1 = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} = A^2\vec{x}_0$$

....

in general,

$$\vec{x}_k = A^k \vec{x}_0$$

multiplying by A steps
one iteration
into the future!

--- big idea

(see notes for code):

```
x_1 = [0.8 0.2]
x_2 = [0.7 0.3]
x_3 = [0.65 0.35]
x_4 = [0.625 0.375]
x_5 = [0.6125 0.3875]
x_6 = [0.60625 0.39375]
x_7 = [0.603125 0.396875]
x_8 = [0.6015625 0.3984375]
x_9 = [0.60078125 0.39921875]
x_10 = [0.60039063 0.39960938]
x_11 = [0.60019531 0.39980469]
x_12 = [0.60009766 0.39990234]
x_13 = [0.60004883 0.39995117]
x_14 = [0.60002441 0.39997559]
```

$\vec{x}_k \to$ an eigen vec of A

(corresponding to $\lambda = 1$)

$$A\vec{x} = 1\vec{x}$$

$$\vec{x} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$$

$$\begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\left.\begin{aligned} 0.8x_1 + 0.3x_2 &= x_1 \\ 0.2x_1 + 0.7x_2 &= x_2 \end{aligned}\right\} \Rightarrow \vec{x} = c\begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

add cond that

$$x_1 + x_2 = 1$$

Big idea: If A is an adjacency matrix,

$$A^k \vec{x}_0 \to \text{an eigenvector for the eigenvalue } 1$$

conceptually, the long-run distribution vector $\vec{x}$ satisfies

$$A\vec{x} = \underset{1}{\uparrow}\vec{x}$$

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

Why do adjacency matrices always have an eigenvalue of 1?

- observe: cols add up to 1
- observe: $A^T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
- key: A and A^T have the same eigenvalues (same $p(\lambda)$ for both)

<u>Important</u> : It is true that all adjacency matrices have $\lambda = 1$ as an eigenvalue.

$\Rightarrow$ what's also true is $\lambda = 1$ is the largest eigenvalue for any adjacency matrix.

$\Rightarrow$ theorem linked in notes

$A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$ $\lambda_1 = 5$ $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

not an adjacency matrix! $\lambda_2 = 2$ $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Q: what does $A^k \vec{x}_0$ approach as $k \to \infty$?

A: The components of $A^k \vec{x}_0$ approach ∞, but the direction of $A^k \vec{x}_0$ approaches the direction of the dominant eigenvalue's eigenvector

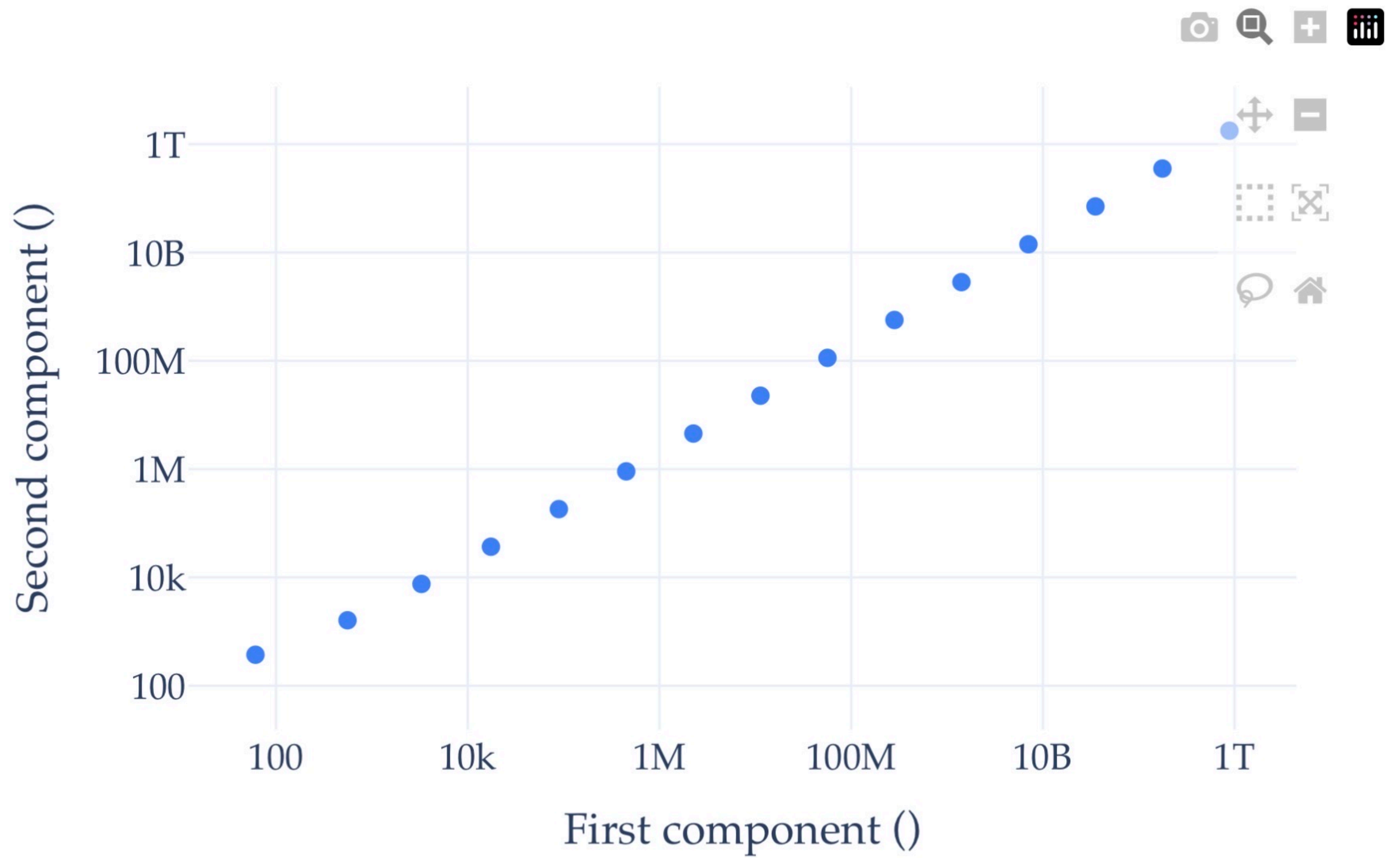
```
simulate_steps(
    A = np.array([[3, 1],
                  [2, 4]]),
    x0 = np.array([[-13], [100]])
)
```

```
x_1 = [ 61 374]
x_2 = [ 557 1618]
x_3 = [3289 7586]
x_4 = [17453 36922]
x_5 = [ 89281 182594]
x_6 = [450437 908938]
x_7 = [2260249 4536626]
x_8 = [11317373 22667002]
x_9 = [ 56619121 113302754]
x_10 = [283160117 566449258]
x_11 = [1415929609 2832117266]
x_12 = [ 7079906093 14160328282]
x_13 = [35400046561 70801125314]
x_14 = [177001264997 354004594378]
x_15 = [ 885008389369 1770020907506]
```

as $k \to \infty$,

$A^k \vec{x}_0 \to$ a scaled version of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$A^{15} \begin{bmatrix} -13 \\ 100 \end{bmatrix}$



1T
10B
100M
1M
10k
100
Second component ()
100
10k
1M
100M
10B
1T
First component ()

Why did $A^k \vec{x}_0$ converge in direction to the eigenvector with the largest eigenvalue?

Suppose $\vec{x} \in \mathbb{R}^2$

$A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$ $\lambda_1 = 5$ $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

not an adjacency matrix! $\lambda_2 = 2$ $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\vec{v}_1, \vec{v}_2$ are linearly independent and span all of $\mathbb{R}^2$,

so it must be possible to write

$$\vec{x} = c_1 \vec{v}_1 + c_2 \vec{v}_2$$

$$\vec{x} = c_1 \vec{v}_1 + c_2 \vec{v}_2$$

Q: what happens when I multiply $\vec{x}$ by A?

$$A\vec{x} = A(c_1\vec{v}_1 + c_2\vec{v}_2)$$

$$= c_1 A\vec{v}_1 + c_2 A\vec{v}_2$$

$$= c_1 \lambda_1 \vec{v}_1 + c_2 \lambda_2 \vec{v}_2$$

$$A^2\vec{x} = c_1 \lambda_1^2 \vec{v}_1 + c_2 \lambda_2^2 \vec{v}_2$$

$$A^k\vec{x} = c_1 \lambda_1^k \vec{v}_1 + c_2 \lambda_2^k \vec{v}_2$$

remember, $\vec{v}_1$, $\vec{v}_2$ are eigenvectors of A!

Q: As $k \to \infty$, what happens to $A^k \vec{x}$?

$$A^k \vec{x} = c_1 \lambda_1^k \vec{v}_1 + c_2 \lambda_2^k \vec{v}_2$$

Just to illustrate, use $\lambda_1 = 5$, $\lambda_2 = 2$

$$\frac{A^k \vec{x}}{5^k} = \frac{c_1 5^k \vec{v}_1 + c_2 2^k \vec{v}_2}{5^k}$$

$$\frac{A^k \vec{x}}{5^k} = c_1 \vec{v}_1 + c_2 \left(\frac{2}{5}\right)^k \vec{v}_2$$

As $k \to \infty$, $\left(\frac{2}{5}\right)^k \to 0$

so direction of $A^k \vec{x} \to$ direction of $\vec{v}_1$!

In our analysis, we assumed that

$$\vec{x} = c_1 \vec{v}_1 + c_2 \vec{v}_2$$

$\vec{v}_1, \vec{v}_2$ were linearly independent and spanned $\mathbb{R}^2$

→ what if a matrix doesn't have n linearly independent eigenvectors?

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow$$ Find its eigenvalues and eigenvectors

Last time: we showed that if A is an $n \times n$ matrix and $\vec{x} \in \mathbb{R}^n$, then

as $k \to \infty$,

absolute value

$A^k \vec{x} \to$ a scaled version of the eigenvector with the largest $|$eigvalue$|$.

Our analysis required us to assume that A's eigenvectors spanned all of $\mathbb{R}^n$, i.e. A has n linearly independent eigenvectors

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

Find eigenvalues and eigenvectors

$$p(\lambda) = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

$\lambda = 1$ is a double root

$$\lambda_1 + \lambda_2 = 2$$

$$\lambda_1 \lambda_2 = 1$$

$$\Rightarrow \lambda_1 = 1, \; \lambda_2 = 1$$

$\Rightarrow$ repeated eigenvalue!

eigenvector must satisfy

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{v}} = 1 \begin{bmatrix} a \\ b \end{bmatrix}$$

$$a + b = a$$

$$b = b$$

$\Rightarrow$ $b = 0$, a free

$$\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$\mathbb{R}^2$

$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ only has one line of eigenvectors for $\lambda = 1$

$\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\Rightarrow$ A's eigenvectors don't span all of $\mathbb{R}^2$!

$\Rightarrow$ A is <u>not</u> diagonalizable!

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

also has $\lambda = 1$
repeated eigenvalue

$$p(\lambda) = \begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix}$$

$$= (1-\lambda)^2$$

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

$$a = a$$

$$b = b$$

both variables are free!

the set of all eigenvectors of A with $\lambda = 1$ is a 2-dimensional subspace of $\mathbb{R}^2$ (ignore the $\vec{0}$)

Eigenvalue decomposition of a matrix

Suppose A is an $n \times n$ matrix with n linearly independent eigenvectors

eigenvectors

$$A\vec{v}_1 = \lambda_1 \vec{v}_1$$

$$A\vec{v}_2 = \lambda_2 \vec{v}_2$$

$$\vdots$$

$$A\vec{v}_n = \lambda_n \vec{v}_n$$

some λ_i's may be the same!

$$A \underbrace{\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}}_{V}$$

$$
A \underbrace{\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}}_{V} = \begin{bmatrix} | & | & & | \\ A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_n \\ | & | & & | \end{bmatrix}
$$

$$
= \begin{bmatrix} | & | & & | \\ \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \\ | & | & & | \end{bmatrix}
$$

$$
= \underbrace{\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}}_{\Lambda}
$$

$$AV = V\Lambda$$

If A has n linearly independent eigenvectors,

then $V = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$

is invertible!

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

eigenvalue decomposition of A

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

Matrix powers are easy!

$$A^2 = V\Lambda \underbrace{V^{-1} V}_{I} \Lambda V^{-1}$$

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

$$= V\Lambda^2 V^{-1}$$

$$\Lambda^2 = \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{bmatrix}$$

$$A^{100} = V \Lambda^{100} V^{-1}$$

$$= V \begin{bmatrix} \lambda_1^{100} & & & \\ & \lambda_2^{100} & & \\ & & \ddots & \\ & & & \lambda_n^{100} \end{bmatrix} V^{-1}$$

way more efficient than multiplying
A by itself 100x !