DSC 140B
Representation Learning

Lecture 05 | Part 1

The Spectral Theorem

Eigenvectors

- Let A be an $n \times n$ matrix. An **eigenvector** of A with **eigenvalue** λ is a nonzero vector $\vec{v}$ such that $A\vec{v} = \lambda\vec{v}$.

Eigenvectors (of Linear Transformations)

- Let $\vec{f}$ be a linear transformation. An **eigenvector** of $\vec{f}$ with **eigenvalue** λ is a nonzero vector $\vec{v}$ such that $f(\vec{v}) = \lambda\vec{v}$.

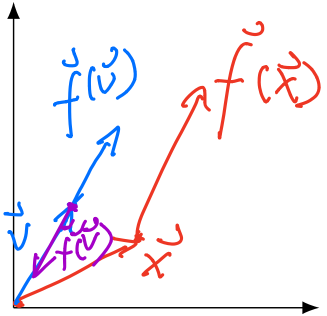
$$f(\vec{v}) = A\vec{v} = \lambda\vec{v}$$

Geometric Interpretation

$\vec{f}(\vec{v}) = \lambda \vec{v}$

< 0

- When $\vec{f}$ is applied to one of its eigenvectors, $\vec{f}$ simply scales it.
    - Possibly by a negative amount.



$$A_{ij} = A_{ji}$$

Symmetric Matrices

- Recall: a matrix A is **symmetric** if $A^T = A$.

$$\begin{bmatrix} 1 & 0 & -100 \\ 0 & 2 & 123 \\ -100 & 123 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix}$$

The Spectral Theorem[1]

- **Theorem**: Let A be an $n \times n$ *symmetric* matrix. Then there exist n eigenvectors of A which are all mutually orthogonal.

[1] for symmetric matrices

What?

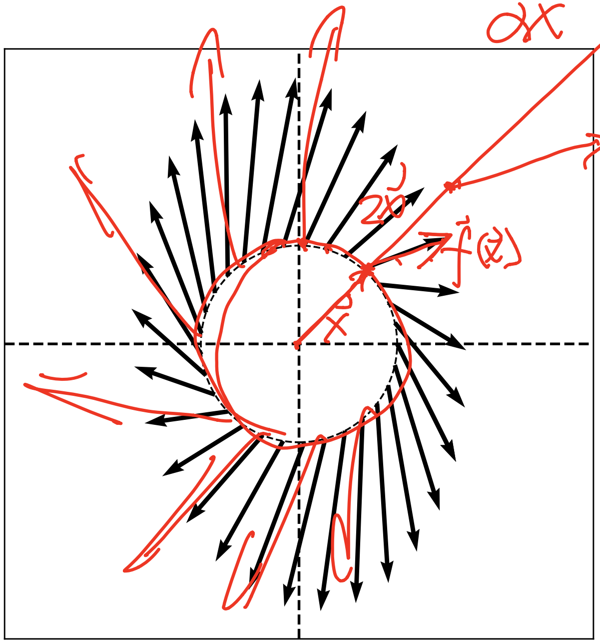
- What does the spectral theorem mean?
- What is an eigenvector, really?
- Why are they useful?

Example Linear Transformation

$\alpha \vec{f}(\vec{x})$

$\alpha \vec{x}$

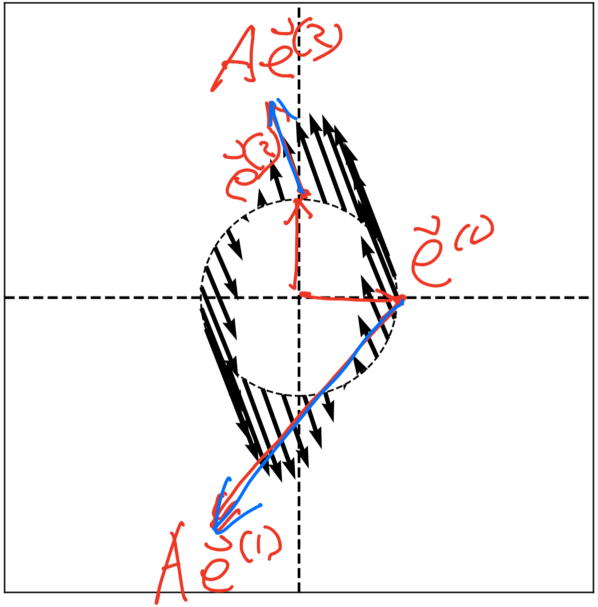
$2\vec{f}(\vec{x})$



$\vec{f}(\vec{e}^{(1)})$ $\vec{f}(\vec{e}^{(2)})$

$$A = \begin{pmatrix} \boxed{\begin{matrix} 5 \\ -10 \end{matrix}} & \boxed{\begin{matrix} 5 \\ 12 \end{matrix}} \end{pmatrix}$$

Example Linear Transformation

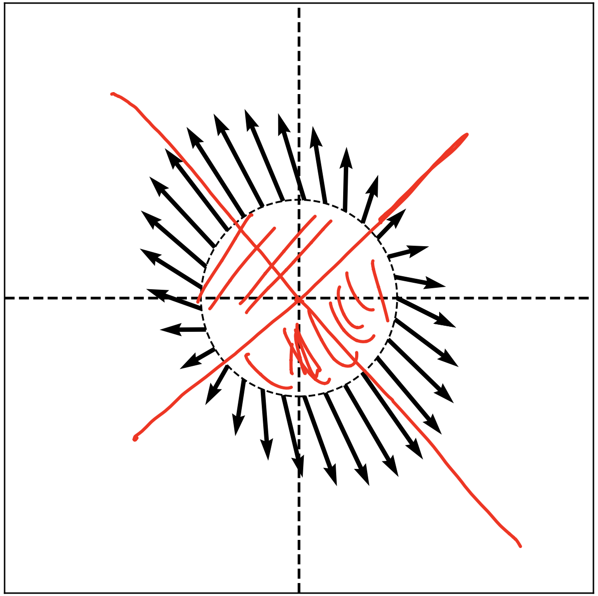


$$A\vec{e}^{(1)} = \begin{pmatrix} -2 \\ -5 \end{pmatrix}$$

$$A\vec{e}^{(2)} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

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

Example Symmetric Linear Transformation

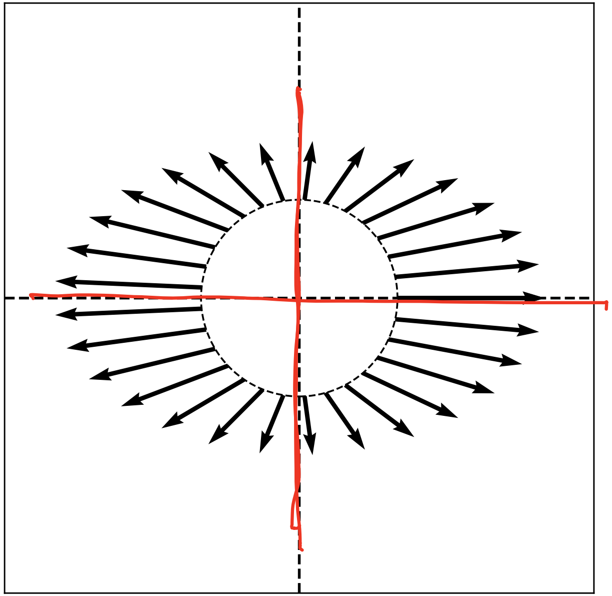


$\mathbb{R}^d$

$d=2$

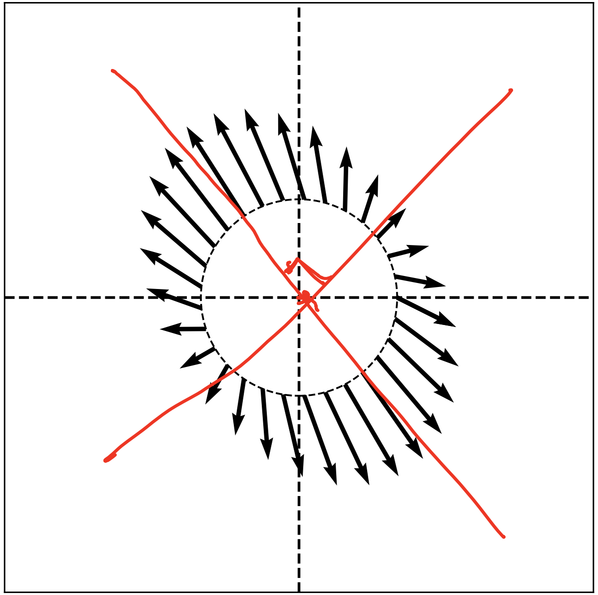
$$A = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$$

Example Symmetric Linear Transformation



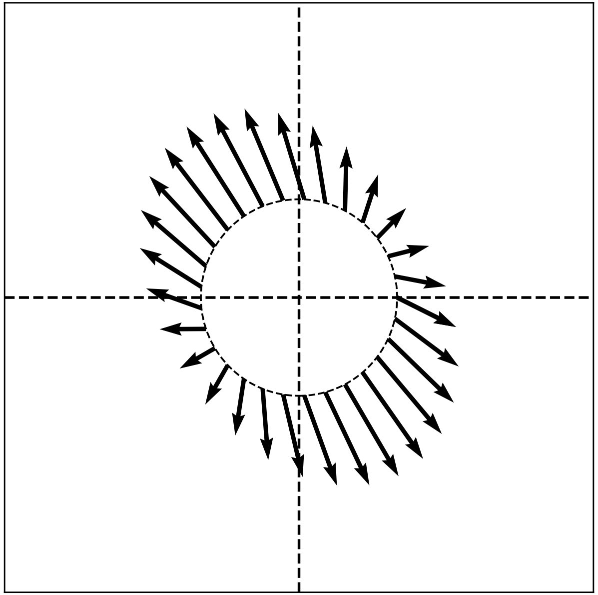
$$A = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}$$

Observation #1



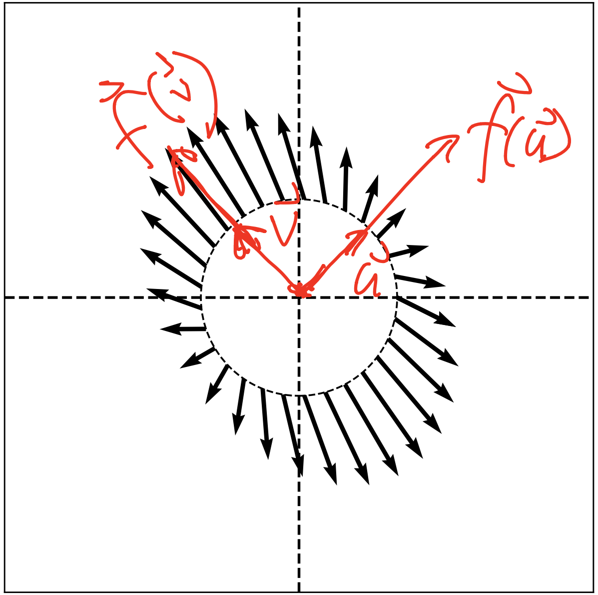
- Symmetric linear transformations have **axes of symmetry**.

Observation #2



- The axes of symmetry are **orthogonal** to one another.

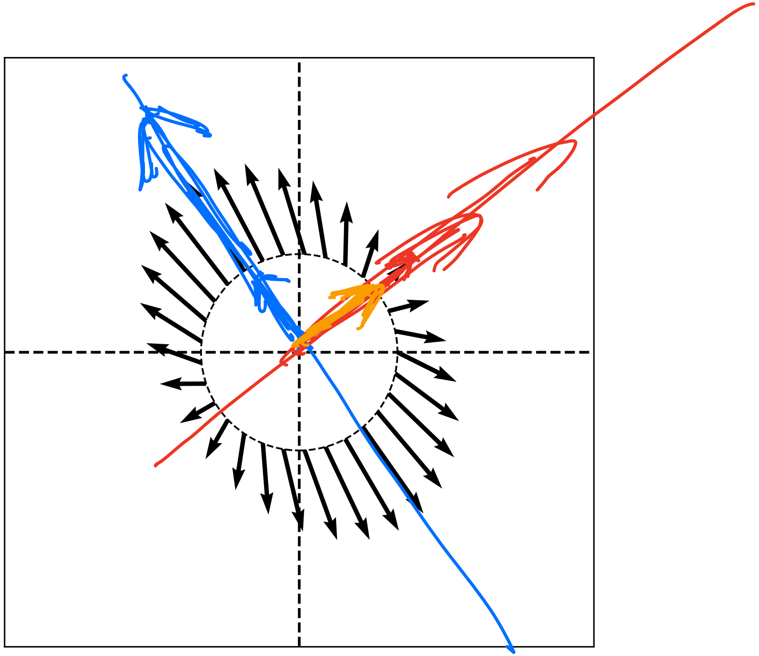
Observation #3



- The action of $\vec{f}$ along an axis of symmetry is simply to scale its input.

Observation #4

$f(\vec{v}) = \lambda \vec{v}$



- The size of this scaling can be different for each axis.

Main Idea

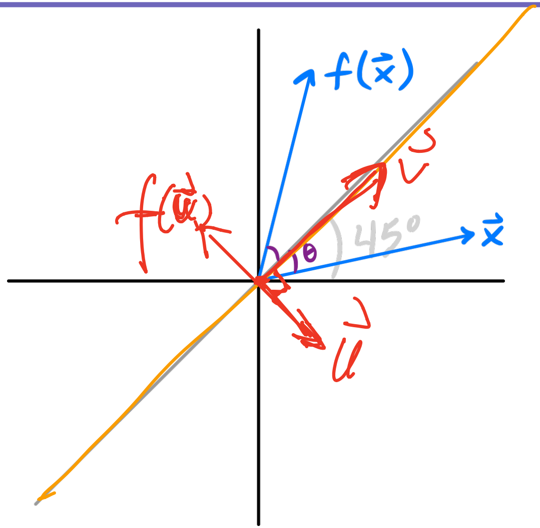
The **eigenvectors** of a symmetric linear transformation (matrix) are its axes of symmetry. The **eigenvalues** describe how much each axis of symmetry is scaled.

Exercise

Consider the linear transformation which mirrors its input over the line of 45°. Give two orthogonal eigenvector of the transformation.

$\boxed{f(\vec{v}) = \lambda \vec{v}}$

$\vec{v}$



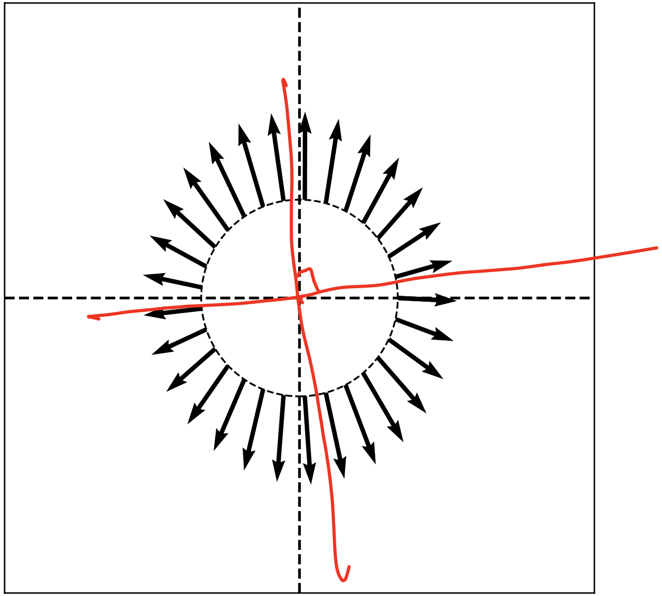
$f(\vec{v}) = \vec{v}$ $\quad \lambda = 1$

$f(\vec{u}) = -\vec{u}$

$\lambda = -1$

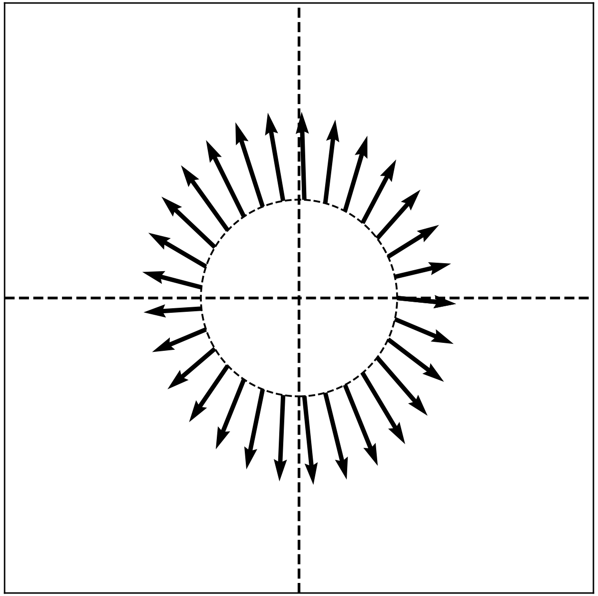
Off-diagonal elements

$A_{ij} = A_{ji}$



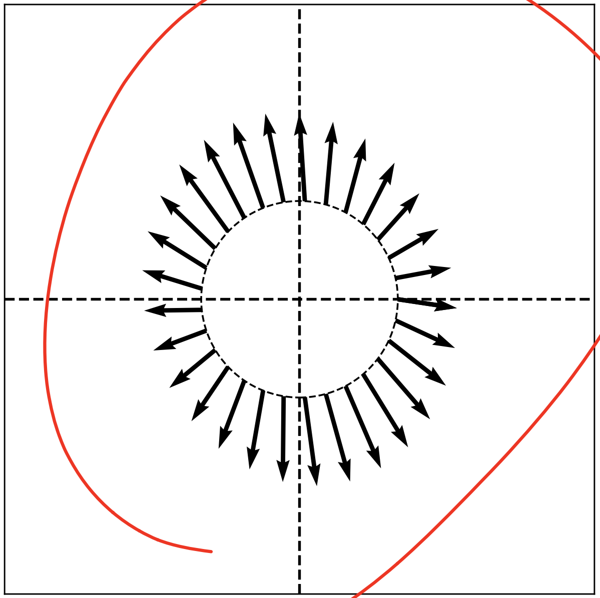
$$A = \begin{pmatrix} 5 & -0.1 \\ -0.1 & 2 \end{pmatrix}$$

Off-diagonal elements



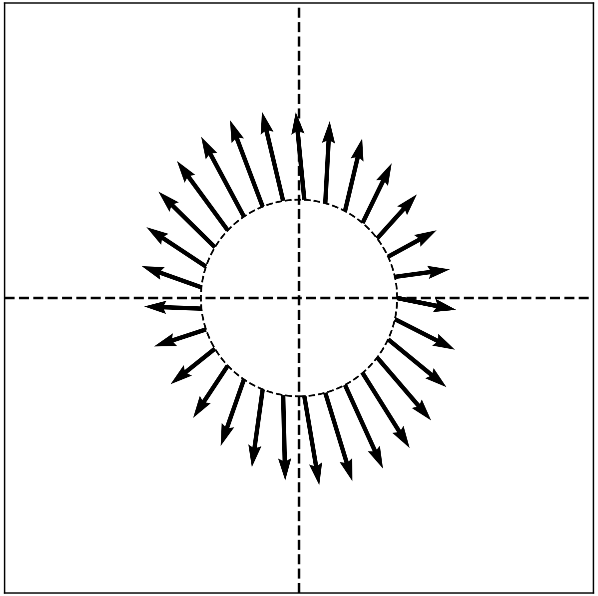
$$A = \begin{pmatrix} 5 & -0.2 \\ -0.2 & 2 \end{pmatrix}$$

Off-diagonal elements



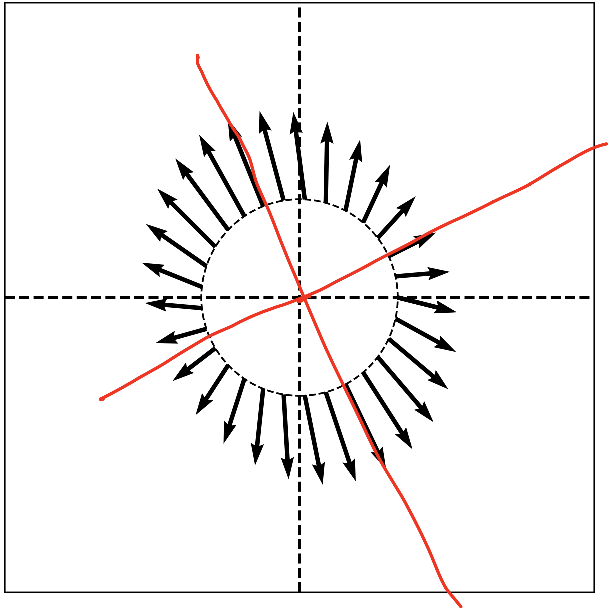
$$A = \begin{pmatrix} 5 & -0.3 \\ -0.3 & 2 \end{pmatrix}$$

Off-diagonal elements



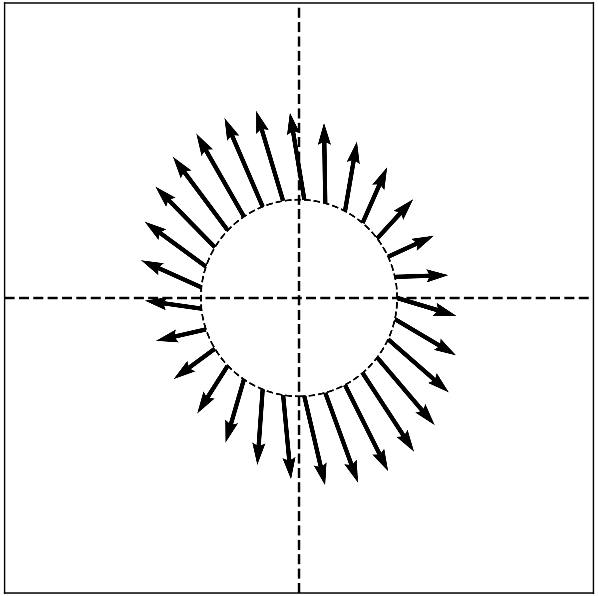
$$A = \begin{pmatrix} 5 & -0.4 \\ -0.4 & 2 \end{pmatrix}$$

Off-diagonal elements



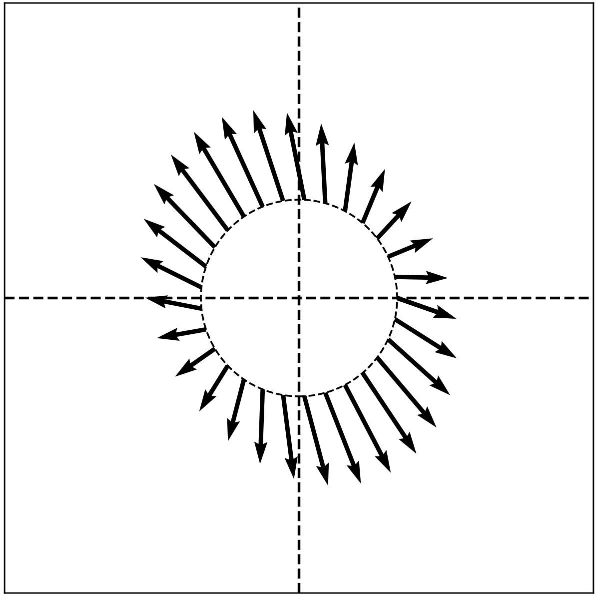
$$A = \begin{pmatrix} 5 & -0.5 \\ -0.5 & 2 \end{pmatrix}$$

Off-diagonal elements



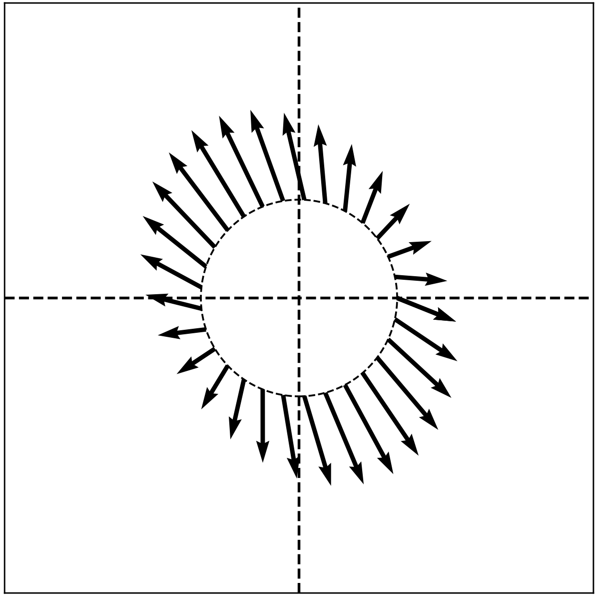
$$A = \begin{pmatrix} 5 & -0.6 \\ -0.6 & 2 \end{pmatrix}$$

Off-diagonal elements



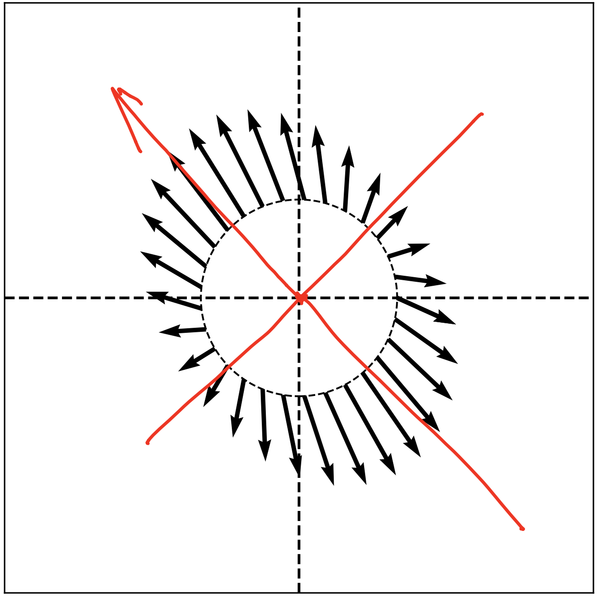
$$A = \begin{pmatrix} 5 & -0.7 \\ -0.7 & 2 \end{pmatrix}$$

Off-diagonal elements



$$A = \begin{pmatrix} 5 & -0.8 \\ -0.8 & 2 \end{pmatrix}$$

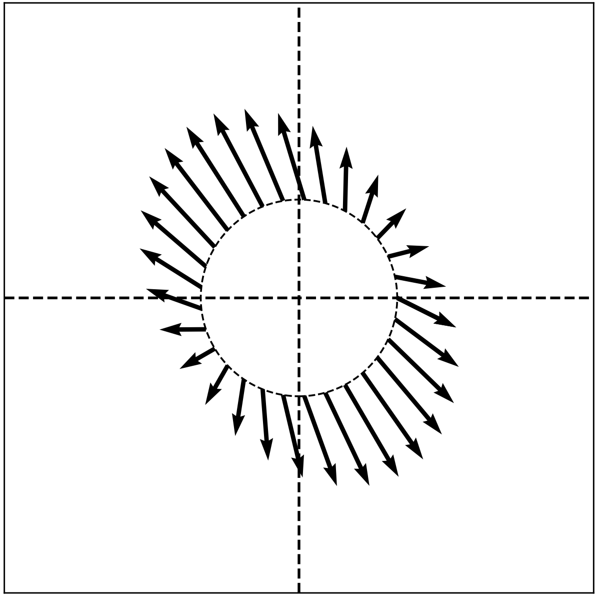
Off-diagonal elements



$$A = \begin{pmatrix} 5 & -0.9 \\ -0.9 & 2 \end{pmatrix}$$

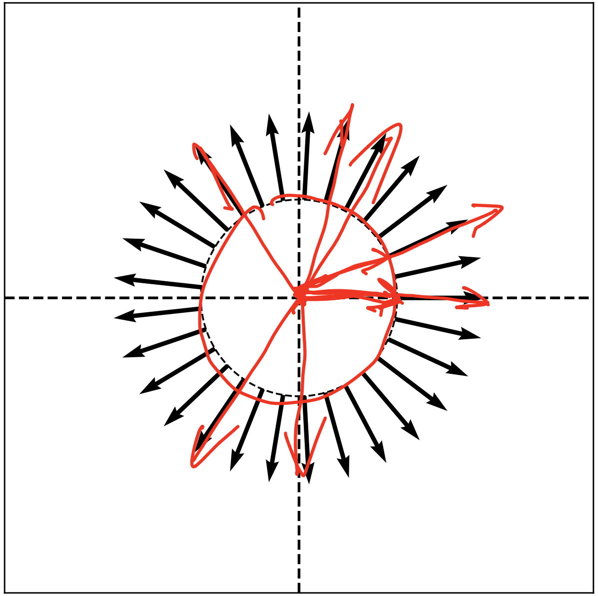
The Spectral Theorem[2]

- **Theorem**: Let A be an $n \times n$ *symmetric* matrix. Then there exist n eigenvectors of A which are all mutually orthogonal.



[2] for symmetric matrices

What about total symmetry?



$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 \\ 3x_2 \end{pmatrix}$$

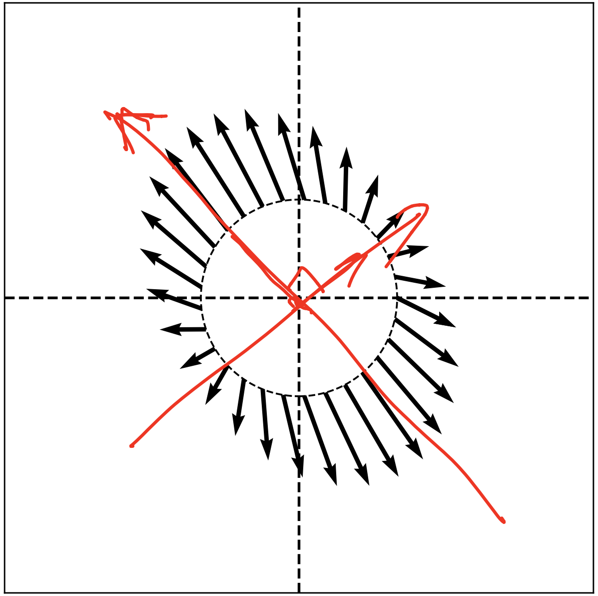
- Every vector is an eigenvector. $= 3\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \qquad \lambda = 3$$

$$A_{ij} = A_{ji}$$

Computing Eigenvectors

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



```
>>> A = np.array([[2, -1], [-1, 3]])
>>> np.linalg.eigh(A)
(array([1.38196601, 3.61803399]),
 array([[-0.85065081, -0.52573111],
        [-0.52573111,  0.85065081]]))
```

λ_1 λ_2

V_1 V_2

DSC 140B
Representation Learning

Lecture 05 | Part 2

Why are eigenvectors useful?

OK, but why are eigenvectors[3] useful?

- Eigenvectors are nice “building blocks” (basis vectors).
- Eigenvectors are **maximizers** (or minimizers).
- Eigenvectors are **equilibria**.

[3] of symmetric matrices

Eigendecomposition

- Any vector $\vec{x}$ can be written in terms of the eigenvectors of a symmetric matrix.

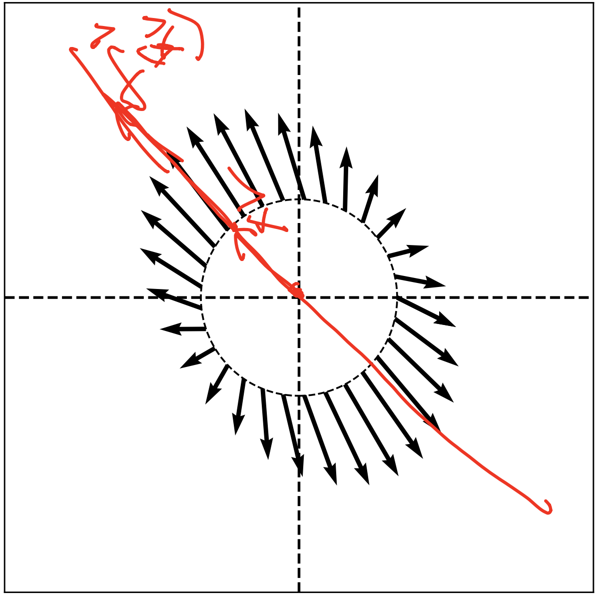
- This is called its **eigendecomposition**.

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

$$\vec{x} = x_1 \vec{e}^{(1)} + x_2 \vec{e}^{(2)}$$

$$= z_1 \vec{v}_1 + z_2 \vec{v}_2$$

Observation #1



- $\vec{f}(\vec{x})$ is longest along the "main" axis of symmetry.
  - In the direction of the eigenvector with largest eigenvalue.

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

Main Idea

To maximize $\|\vec{f}(\vec{x})\|$ over unit vectors, pick $\vec{x}$ to be an eigenvector of $\vec{f}$ with the largest eigenvalue (in abs. value).

Main Idea

To minimize $\|\vec{f}(\vec{x})\|$ over unit vectors, pick $\vec{x}$ to be an eigenvector of $\vec{f}$ with the smallest eigenvalue (in abs. value).

Proof $\|\vec{f}(\vec{x})\|$

Show that the maximizer of $\|A\vec{x}\|$ s.t., $\|\vec{x}\| = 1$ is the top eigenvector of A.

$$\max_{\vec{x},} \|A\vec{x}\|$$

$$\text{s.t. } \|\vec{x}\| = 1$$

$\vec{x}^* =$ top eigenvector of A

Corollary

To maximize $\vec{x} \cdot A\vec{x}$ over unit vectors, pick $\vec{x}$ to be top eigenvector of A.

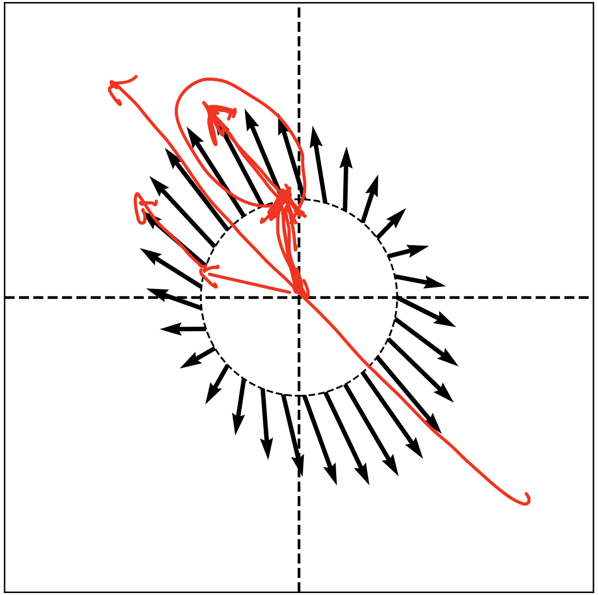
Example

$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

- Maximize $4x_1^2 + 2x_2^2 + 3x_1x_2$ subject to $x_1^2 + x_2^2 = 1$

$\|x\| = 1$

Observation #2



- $\vec{f}(\vec{x})$ rotates $\vec{x}$ towards the "top" eigenvector $\vec{v}$.
- $\vec{v}$ is an equilibrium.

The Power Method

- Method for computing the top eigenvector/value of A.
- Initialize $\vec{x}^{(0)}$ randomly
- Repeat until convergence:
  - Set $\vec{x}^{(i+1)} = A\vec{x}^{(i)} / \|A\vec{x}^{(i)}\|$