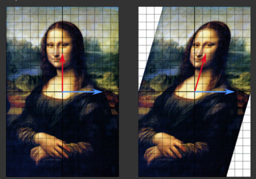
Eigenvalues & Eigenvectors

Lecture 13

Termeh Shafie

linear transformations

A linear transformation is an operation that stretches, squishes, rotates, or otherwise transforms a space



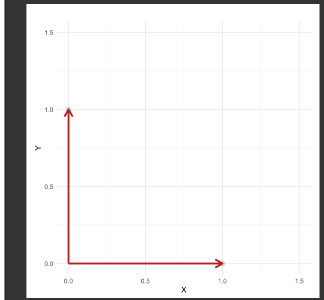
source: https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors

Think of scaling an image, rotating it, or stretching it in one direction.
Eigenvectors and eigenvalues reveal the fundamental "axes" of a transformation.

let's recap some important definitions

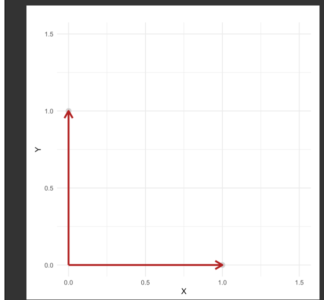
- A set of vectors $\in \mathbb{R}^n$ is called a **basis**, if they are **linearly independent** and every vector $\in \mathbb{R}^n$ can be expressed as a linear combination of these vectors.
- A set of n vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ is linearly independent if no vector in the set can be expressed as a linear combination of the remaining $n-1$ vectors. In other words, the only solution to $c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_n\vec{v_n} = 0$ is $c_1 = c_2 = \ldots = c_n = 0$ (where c_i are scalars)
- In other words, the coefficients $c_1, c_2, \ldots, c_n$ must all be zero for the linear combination to result in the zero vector and properly characterizing linear independence.

example



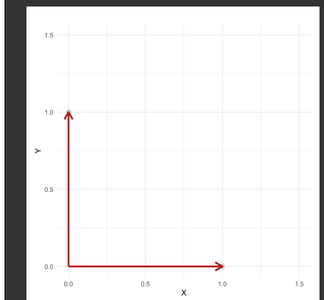
- Consider space $\mathbb{R}^2$
- Consider vectors $\vec{x} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\vec{y} = \begin{bmatrix}0\\1\end{bmatrix}$

example



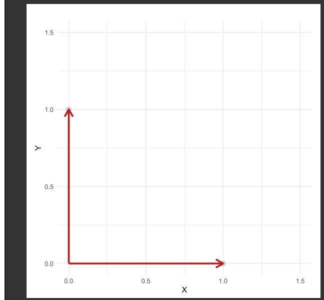
- Consider space $\mathbb{R}^2$
- Consider vectors $\vec{x} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\vec{y} = \begin{bmatrix}0\\1\end{bmatrix}$
- Any vector $\begin{bmatrix}a\\b\end{bmatrix} \in \mathbb{R}^2$ can be expressed as a linear combination of these two vectors:
$$\begin{bmatrix}a\\b\end{bmatrix} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}0\\1\end{bmatrix}$$

example



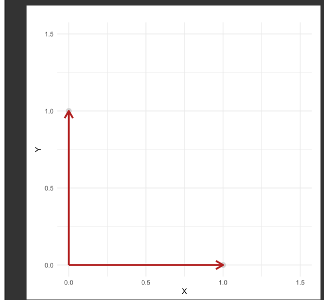
- Consider space $\mathbb{R}^2$
- Consider vectors $\vec{x} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\vec{y} = \begin{bmatrix}0\\1\end{bmatrix}$
- Any vector $\begin{bmatrix}a\\b\end{bmatrix} \in \mathbb{R}^2$ can be expressed as a linear combination of these two vectors:
$$\begin{bmatrix}a\\b\end{bmatrix} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}0\\1\end{bmatrix}$$
- We also note that $\vec{x}$ and $\vec{y}$ are linearly independent since the only solution to
$$c_1\vec{x} + c_2\vec{y} = 0 \text{ is } c_1 = c_2 = 0$$

example



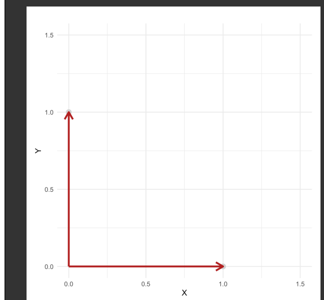
- $\vec{x}$ and $\vec{y}$ are unit vectors in the direction of the coordinate axes
- We are used to representing all vectors in $\mathbb{R}^2$ as linear combinations of these vectors

example



- $\vec{x}$ and $\vec{y}$ are unit vectors in the direction of the coordinate axes
- We are used to representing all vectors in $\mathbb{R}^2$ as linear combinations of these vectors
- We can actually choose any 2 linearly independent vectors in $\mathbb{R}^2$ as basis vectors
- However, an orthogonal basis is the most convenient basis that one can hope for.

example



- $\vec{x}$ and $\vec{y}$ are unit vectors in the direction of the coordinate axes
- We are used to representing all vectors in $\mathbb{R}^2$ as linear combinations of these vectors
- We can actually choose any 2 linearly independent vectors in $\mathbb{R}^2$ as basis vectors
- However, an orthogonal basis is the most convenient basis that one can hope for.

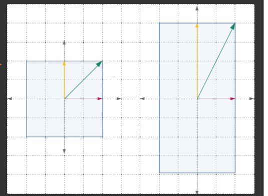
what has all this got to do with eigenvectors?

eigenvectors and eigenvalues: intuitively

example

Applying a vertical scaling of +2 to every vector of a square, will transform the square into a rectangle.

- The horizontal vector remains unchanged (same direction, same length).
- The vertical vector has same direction, but doubled in length.
- The diagonal vector has changed its angle (direction) as well as length.



After vertical scaling of +2, every vector's direction has changed, except the horizontal and vertical ones.
These two vectors are special and are the characteristic of this particular transform. They are called eigenvectors

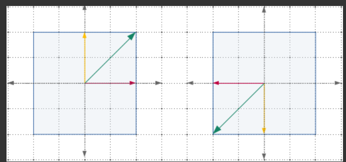
The eigenvalue is how much the eigenvectors are transformed (**stretched or squished**).

- The horizontal vector's length remains same, thus have an eigenvalue of +1.
- The vertical vectors' length doubled, thus have an eigenvalue of +2.

eigenvectors and eigenvalues: intuitively

another example

In 180 degree rotation of square, all vectors are still laying on the same span, but their direction is reversed.
Hence, all vectors are eigenvectors, having an eigenvalue of -1.



Note: In case of 3d rotation transformation of cube, the eigenvector gives the axis of rotation.

eigenvectors and eigenvalues

Let A be a $n \times n$ matrix.

1. An eigenvector of A is a nonzero vector $\vec{v}$ in $\mathbb{R}^n$ such that $A\vec{v} = \lambda\vec{v}$, for some scalar λ
2. An eigenvalue of A is a scalar λ such that the equation $A\vec{v} = \lambda\vec{v}$ has a *non-trivial** solution

If $A\vec{v} = \lambda\vec{v}$ for $\vec{v} \neq 0$, we say that λ is the eigenvalue for $\vec{v}$, and that $\vec{v}$ is an eigenvector for λ.

*means that the solution vector $\vec{v}$ is not the zero vector ($\vec{v} \neq \vec{0}$), and ensures that it represents a meaningful direction in the vector space.

verifying eigenvectors

How to check if a given $\vec{v}$ is the eigenvector of a given matrix A

- multiply $\vec{v}$ by A and see if $A\vec{v}$ is a scalar multiple of $\vec{v}$, i.e. $A\vec{v} = \lambda\vec{v}$
- what happens when a matrix hits a vector?

example

Consider matrix $A = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix}$ and vector $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

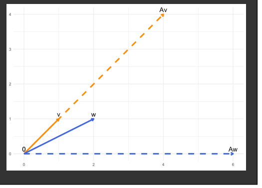
Which are eigenvectors? What are their eigenvalues?

$$A\vec{v} = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\vec{v}$$

$\implies$ $\vec{v}$ is an eigenvector of A

$$A\vec{w} = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

$\implies$ $\vec{w}$ is not an eigenvector of A



verifying eigenvectors

exercise 1

Consider the matrix $A = \begin{bmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{bmatrix}$ and vectors $\vec{v_1} = \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}$ and $\vec{v_2} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$.

Which are eigenvectors? What are their eigenvalues?

the characteristic polynomial

Is there a method for computing all of the eigenvalues of a matrix?

YES!

by finding the roots of the characteristic polynomial (i.e. solving a nonlinear equation in one variable)

> Let A be a $n \times n$ matrix. The characteristic polynomial of A is the function $f(\lambda)$ given by
>
> $$f(\lambda) = \det(A - \lambda I)$$

- The characteristic polynomial is in fact a polynomial
- The point of the characteristic polynomial is that we can use it to compute eigenvalues
- Finding the characteristic polynomial means computing the determinant of the matrix $\det(A - \lambda I)$ whose entries contain the unknown λ

the characteristic polynomial

example

Find the characteristic polynomial of the matrix $A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$.

We have

$$f(\lambda) = \det(A - \lambda I) = \det\left(\begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$$

$$= \det\begin{bmatrix} 5-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} = (5-\lambda)(1-\lambda) - 2\cdot 2$$

$$= \lambda^2 - 6\lambda + 1$$

the characteristic polynomial

Eigenvalues are roots of the characteristic polynomial

Let A be a $n \times n$ matrix and let $f(\lambda) = \det(A - \lambda I)$ be its characteristic polynomial. Then a number λ_0 is an eigenvalue of A if and only if $f(\lambda_0) = 0$.

example cont'd

$$f(\lambda) = \lambda^2 - 6\lambda + 1 = 0$$

$$\implies \lambda = 3 - 2\sqrt{2} \quad \text{and} \quad \lambda = 3 + 2\sqrt{2}\,.$$

To compute the eigenvectors, we solve the homogeneous system of equations $(A - \lambda I)\vec{v} = 0$ for each eigenvalue λ.

the characteristic polynomial: a shortcut

Recall the trace of the square matrix: Let A be an $n \times n$ matrix. The trace of A, denoted $tr(A)$, is the sum of the diagonal elements of A. That is,

$$tr(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

The characteristic polynomial for a 2×2 matrix

Let A be a 2×2 matrix. All coefficients of the characteristic polynomial can be found via

$$f(\lambda) = \lambda^2 + tr(A)\lambda + \det(A)$$

this is generally the fastest way to compute the characteristic polynomial of a 2×2 matrix.

example cont'd

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} \quad \implies \quad f(\lambda) = \lambda^2 - tr(A)\lambda + \det(A) = \lambda^2 - 6\lambda + 1$$

the process so far...

1. **Find Eigenvalues:**
   Determine the eigenvalues λ of the matrix A by solving the characteristic polynomial $\det(A - \lambda I) = 0$ where I is the identity matrix.

2. **Find Eigenvectors:**
   For each eigenvalue, find the corresponding eigenvectors by solving $(A - \lambda I)\vec{v} = 0$ which is equivalent to solving $A\vec{v} = \lambda\vec{v}$

diagonalization

- Diagonal matrices are the easiest kind of matrices to understand: they just scale the coordinate directions by their diagonal entries.
- **Matrix diagonalization** is powerful: it transforms a given square matrix into a diagonal matrix,
  - which is much easier to analyze and compute because their non-diagonal elements are zero
  - e.g. calculations like powers and determinants easy to perform

An $n \times n$ matrix A is **diagonalizable** if it is similar to a diagonal matrix: that is, if there exists an invertible$n \times n$ matrix Q and a diagonal matrix D such that $A = QDQ^{-1}$.

diagonalization

example

$$\begin{bmatrix} -12 & 15 \\ -10 & 13 \end{bmatrix} \text{ is diagonalizable because } \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}^{-1}$$

Note: any diagonal matrix D is diagonalizable because it is similar to itself. For instance,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = I \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} I^{-1}$$

powers of diagonalizable matrices

Multiplying diagonal matrices together just multiplies their diagonal entries:

$$\begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix} \begin{bmatrix} y_1 & 0 & 0 \\ 0 & y_2 & 0 \\ 0 & 0 & y_3 \end{bmatrix} = \begin{bmatrix} x_1y_1 & 0 & 0 \\ 0 & x_2y_2 & 0 \\ 0 & 0 & x_3y_3 \end{bmatrix}$$

so it is easy to take powers of a diagonal matrix:

$$\begin{bmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{bmatrix}^n = \begin{bmatrix} x^n & 0 & 0 \\ 0 & y^n & 0 \\ 0 & 0 & z^n \end{bmatrix}$$

$\Longrightarrow$ if $A = QDQ^{-1}$ where D is the diagonal matrix, then $A^n = QD^nQ^{-1}$

diagonalization theorem

An $n \times n$ matrix A is diagonalizable if if and only if A has n linearly independent eigenvectors.

simply put: a matrix is diagonalizable if it has distinct eigenvalues or, if it has repeated eigenvalues, it still has enough independent eigenvectors to match its dimensionality

- the eigenvalues determine the entries of the diagonal matrix
- the eigenvectors form the columns of a matrix Q
- the transformation reflects how the original matrix A can be simplified, highlighting the intrinsic properties of A

$$A = QDQ^{-1} \quad \text{where} \quad Q = \begin{bmatrix} | & | & | & | \\ \vec{v_1} & \vec{v_2} & \dots & \vec{v_1} \\ | & | & | & | \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

$\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$ are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the corresponding eigenvalues

the process continued...

$A = QDQ^{-1}$

1. **Find Eigenvalues:**
   Determine the eigenvalues λ of the matrix A by solving the characteristic polynomial $\det(A - \lambda I) = 0$ where I is the identity matrix.
2. **Find Eigenvectors:**
   For each eigenvalue, find the corresponding eigenvectors by solving $(A - \lambda I)\vec{v} = 0$
3. **Form Matrices:**
   Arrange the eigenvectors as columns in a matrix Q and the eigenvalues along the diagonal in a matrix D.
4. **Construct the Diagonalization:**
   If A is diagonalizable you can express it as $A = QDQ^{-1}$

the recipe

$A = QDQ^{-1}$

Let A be an $n \times n$ matrix. To diagonalize A:

1. Find the eigenvalues of A using the characteristic polynomial.
2. For each eigenvalue λ of A, compute the basis B_λ for the $\lambda-$eigenspace.
3. If there are fewer than n total vectors in all of the eigenspace bases B_λ, then the matrix is not diagonalizable.
4. Otherwise, the n vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ in the eigenspace bases are linearly independent, and

$$A = QDQ^{-1} \text{ for } Q = \begin{bmatrix} | & | & | & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_1} \\ | & | & | & | \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

where λ_i is the eigenvalue for $\vec{v_i}$.

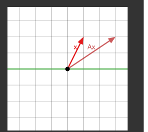
diagonalization

example

Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. The characteristic polynomial of A is $f(\lambda) = (\lambda - 1)^2$ so the eigenvalue of A is 1.

For $\lambda = 1$, solve $A\vec{v} = \lambda\vec{v}$ or $(A - \lambda I)\vec{v} = 0$:

$A - \lambda I = \begin{bmatrix} 1-1 & 1 \\ 0 & 1-1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ which gives equation $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$



From the first row: $0x + 1y = 0 \quad \Longrightarrow \quad y = 0$
and there is no restriction on x so let $x = t$ (a free variable).
The eigenvector is a 1-eigenspace:

$\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ 0 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \end{bmatrix} \Longrightarrow$ Basis: $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$

The 1-eigenspace is exactly the x-axis, so all of the eigenvectors of A lie on the x-axis. It follows that A does not admit two linearly independent eigenvectors, so by the diagonalization theorem, it is not diagonalizable.

diagonalization

exercise 2

Diagonalize the matrix $A = \begin{bmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{bmatrix}$.

eigenvalue decomposition summarized

- Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be the eigenvectors of matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be corresponding eigenvalues
- Consider now a matrix Q whose columns are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$
- We have now

$$AQ = A \begin{bmatrix} | & | & | & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_1} \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | & | \\ A\vec{v_1} & A\vec{v_2} & \ldots & A\vec{v_1} \\ | & | & | & | \end{bmatrix}$$

eigenvalue decomposition summarized

- Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be the eigenvectors of matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be corresponding eigenvalues
- Consider now a matrix Q whose columns are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$
- We have now

$$AQ = A \begin{bmatrix} | & | & | & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_1} \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | & | \\ A\vec{v_1} & A\vec{v_2} & \ldots & A\vec{v_1} \\ | & | & | & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | & | \\ \lambda_1\vec{v_1} & \lambda_2\vec{v_2} & \ldots & \lambda_n\vec{v_1} \\ | & | & | & | \end{bmatrix}$$

eigenvalue decomposition summarized

- Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be the eigenvectors of matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be corresponding eigenvalues
- Consider now a matrix Q whose columns are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$
- We have now

$$AQ = A \begin{bmatrix} | & | & | & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_1} \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | & | \\ A\vec{v_1} & A\vec{v_2} & \ldots & A\vec{v_1} \\ | & | & | & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | & | \\ \lambda_1\vec{v_1} & \lambda_2\vec{v_2} & \ldots & \lambda_n\vec{v_1} \\ | & | & | & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_1} \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \qquad = QD$$

eigenvalue decomposition summarized

- If Q^{-1} exists, then we can write

  $A = QDQ^{-1}$ eigenvalue decomposition

  $Q^{-1}AQ = D$ diagonalization of A

eigenvalue decomposition summarized

- If Q^{-1} exists, then we can write

  $A = QDQ^{-1}$ eigenvalue decomposition

  $Q^{-1}AQ = D$ diagonalization of A

- Under what condition would Q^{-1} exist?
  - If the columns of Q are linearly independent
  - i.e. if A has n linearly independent eigenvectors
  - i.e. if A has n distinct eigenvalues

eigenvalue decomposition summarized

- If Q^{-1} exists, then we can write

  $A = QDQ^{-1}$ eigenvalue decomposition

  $Q^{-1}AQ = D$ diagonalization of A

- Under what condition would Q^{-1} exist?
  - If the columns of Q are linearly independent
  - i.e. if A has n linearly independent eigenvectors
  - i.e. if A has n distinct eigenvalues
- If A is symmetric, we get an even more convenient situation
  - The eigenvalues are orthogonal