MATH 311

Topics in Applied Mathematics

Lecture 16: Diagonalization. Euclidean structure in $\mathbb{R}^n$.

Diagonalization

Let L be a linear operator on a finite-dimensional vector space V. Then the following conditions are equivalent:

- the matrix of L with respect to some basis is diagonal;
- there exists a basis for V formed by eigenvectors of L.

The operator L is **diagonalizable** if it satisfies these conditions.

Let A be an $n \times n$ matrix. Then the following conditions are equivalent:

- A is the matrix of a diagonalizable operator;
- A is similar to a diagonal matrix, i.e., it is represented as $A = UBU^{-1}$, where the matrix B is diagonal;
- there exists a basis for $\mathbb{R}^n$ formed by eigenvectors of A.

The matrix A is **diagonalizable** if it satisfies these conditions. Otherwise A is called **defective**.

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

- The matrix A has two eigenvalues: 1 and 3.
- The eigenspace of A associated with the eigenvalue 1 is the line spanned by $\mathbf{v}_1 = (-1, 1)$.
- The eigenspace of A associated with the eigenvalue 3 is the line spanned by $\mathbf{v}_2 = (1, 1)$.
- Eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ form a basis for $\mathbb{R}^2$.

Thus the matrix A is diagonalizable. Namely, $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \qquad U = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Example. $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

- The matrix A has two eigenvalues: 0 and 2.
- The eigenspace corresponding to 0 is spanned by $\mathbf{v}_1 = (-1, 1, 0)$.
- The eigenspace corresponding to 2 is spanned by $\mathbf{v}_2 = (1, 1, 0)$ and $\mathbf{v}_3 = (-1, 0, 1)$.
- Eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ form a basis for $\mathbb{R}^3$.

Thus the matrix A is diagonalizable. Namely, $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad U = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Problem. Diagonalize the matrix $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$.

We need to find a diagonal matrix B and an invertible matrix U such that $A = UBU^{-1}$.

Suppose that $\mathbf{v}_1 = (x_1, y_1)$, $\mathbf{v}_2 = (x_2, y_2)$ is a basis for $\mathbb{R}^2$ formed by eigenvectors of A, i.e., $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some $\lambda_i \in \mathbb{R}$. Then we can take

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad U = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}.$$

Note that U is the transition matrix from the basis $\mathbf{v}_1, \mathbf{v}_2$ to the standard basis.

Problem. Diagonalize the matrix $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$.

Characteristic equation of A: $\begin{vmatrix} 4-\lambda & 3 \\ 0 & 1-\lambda \end{vmatrix} = 0$.

$(4-\lambda)(1-\lambda) = 0 \implies \lambda_1 = 4,\ \lambda_2 = 1.$

Associated eigenvectors: $\mathbf{v}_1 = (1,0),\ \mathbf{v}_2 = (-1,1)$.

Thus $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find A^5.

We know that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Then $A^5 = UBU^{-1}UBU^{-1}UBU^{-1}UBU^{-1}UBU^{-1}$

$$= UB^5U^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1024 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1024 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1024 & 1023 \\ 0 & 1 \end{pmatrix}.$$

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find a matrix C such that $C^2 = A$.

We know that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Suppose that $D^2 = B$ for some matrix D. Let $C = UDU^{-1}$.
Then $C^2 = UDU^{-1}UDU^{-1} = UD^2U^{-1} = UBU^{-1} = A$.

We can take $D = \begin{pmatrix} \sqrt{4} & 0 \\ 0 & \sqrt{1} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.

Then $C = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$.

System of linear ODEs

Problem. Solve a system $\begin{cases} \frac{dx}{dt} = 4x + 3y, \\ \frac{dy}{dt} = y. \end{cases}$

The system can be rewritten in vector form:

$$\frac{d\mathbf{v}}{dt} = A\mathbf{v}, \quad \text{where } A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}, \ \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

We know that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Let $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ be coordinates of the vector $\mathbf{v}$ relative to the basis $\mathbf{v}_1 = (1, 0)$, $\mathbf{v}_2 = (-1, 1)$ of eigenvectors of A. Then $\mathbf{v} = U\mathbf{w} \implies \mathbf{w} = U^{-1}\mathbf{v}$.

It follows that

$$\frac{d\mathbf{w}}{dt} = \frac{d}{dt}(U^{-1}\mathbf{v}) = U^{-1}\frac{d\mathbf{v}}{dt} = U^{-1}A\mathbf{v} = U^{-1}AU\mathbf{w}.$$

Thus $\quad \dfrac{d\mathbf{w}}{dt} = B\mathbf{w} \quad \Longleftrightarrow \quad \begin{cases} \frac{dw_1}{dt} = 4w_1, \\ \frac{dw_2}{dt} = w_2. \end{cases}$

The general solution: $w_1(t) = c_1 e^{4t}$, $w_2(t) = c_2 e^{t}$, where c_1, c_2 are arbitrary constants. Then

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = U\mathbf{w}(t) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 e^{4t} \\ c_2 e^{t} \end{pmatrix} = \begin{pmatrix} c_1 e^{4t} - c_2 e^{t} \\ c_2 e^{t} \end{pmatrix}.$$

There are **two obstructions** to diagonalization. They are illustrated by the following examples.

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

$\det(A - \lambda I) = (\lambda - 1)^2$. Hence $\lambda = 1$ is the only eigenvalue. The associated eigenspace is the line $t(1, 0)$.

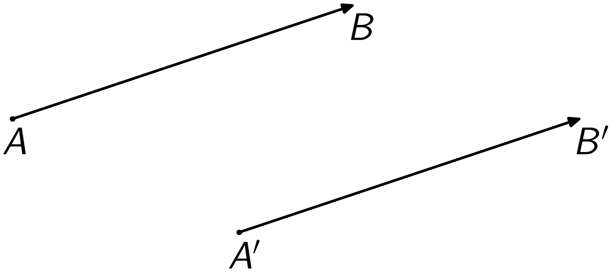
Example 2. $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$\det(A - \lambda I) = \lambda^2 + 1$.

$\Longrightarrow$ no real eigenvalues or eigenvectors

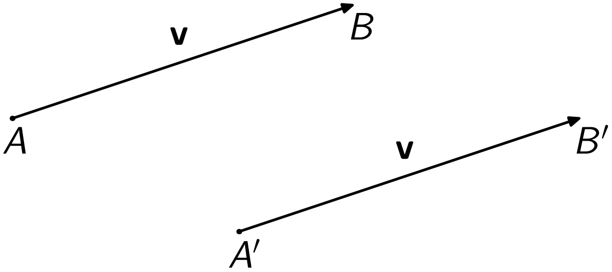
(However there are *complex* eigenvalues/eigenvectors.)

Vectors: geometric approach



- A vector is represented by a directed segment.
- Directed segment is drawn as an arrow.
- Different arrows represent the same vector if they are of the same length and direction.

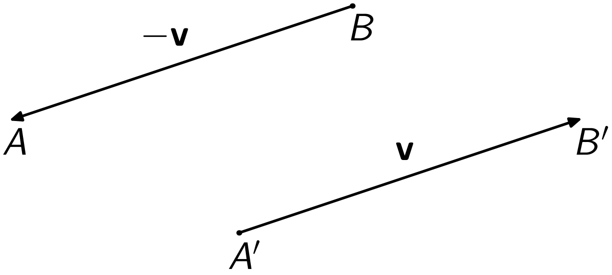
Vectors: geometric approach



$\overrightarrow{AB}$ denotes the vector represented by the arrow with tip at B and tail at A.

$\overrightarrow{AA}$ is called the *zero vector* and denoted $\mathbf{0}$.

Vectors: geometric approach

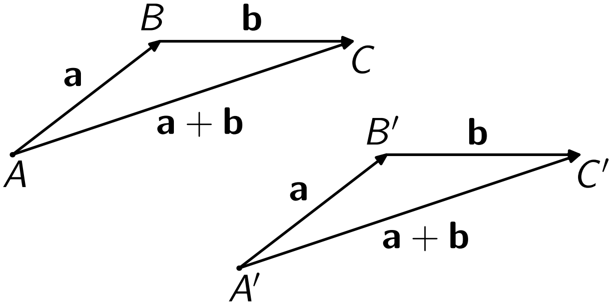


If $\mathbf{v} = \overrightarrow{AB}$ then $\overrightarrow{BA}$ is called the *negative vector* of $\mathbf{v}$ and denoted $-\mathbf{v}$.

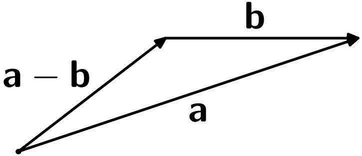
Vector addition

Given vectors **a** and **b**, their *sum* $\mathbf{a} + \mathbf{b}$ is defined by the rule $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.

That is, choose points A, B, C so that $\overrightarrow{AB} = \mathbf{a}$ and $\overrightarrow{BC} = \mathbf{b}$. Then $\mathbf{a} + \mathbf{b} = \overrightarrow{AC}$.

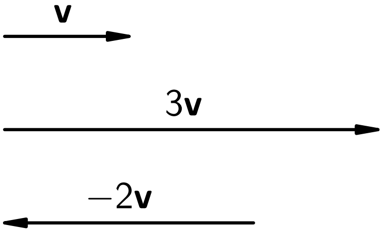


The *difference* of the two vectors is defined as
$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$.



Scalar multiplication

Let **v** be a vector and $r \in \mathbb{R}$. By definition, $r\mathbf{v}$ is a vector whose magnitude is $|r|$ times the magnitude of **v**. The direction of $r\mathbf{v}$ coincides with that of **v** if $r > 0$. If $r < 0$ then the directions of $r\mathbf{v}$ and **v** are opposite.



Beyond linearity: length of a vector

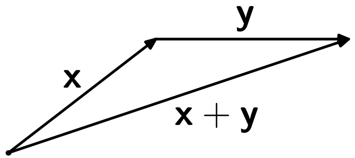
The **length** (or the **magnitude**) of a vector $\overrightarrow{AB}$ is the length of the representing segment AB. The length of a vector $\mathbf{v}$ is denoted $|\mathbf{v}|$ or $\|\mathbf{v}\|$.

Properties of vector length:

$|\mathbf{x}| \geq 0$, $|\mathbf{x}| = 0$ only if $\mathbf{x} = \mathbf{0}$ (positivity)

$|r\mathbf{x}| = |r|\,|\mathbf{x}|$ (homogeneity)

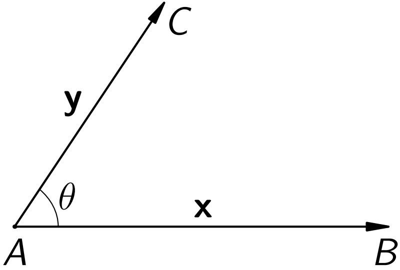
$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$ (triangle inequality)

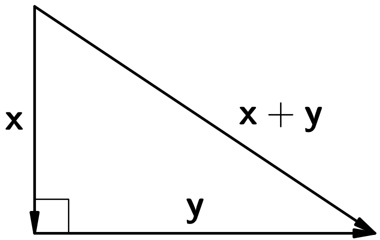


Beyond linearity: angle between vectors

Given nonzero vectors $\mathbf{x}$ and $\mathbf{y}$, let A, B, and C be points such that $\overrightarrow{AB} = \mathbf{x}$ and $\overrightarrow{AC} = \mathbf{y}$. Then $\angle BAC$ is called the **angle** between $\mathbf{x}$ and $\mathbf{y}$.

The vectors $\mathbf{x}$ and $\mathbf{y}$ are called **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if the angle between them equals 90^{o}.





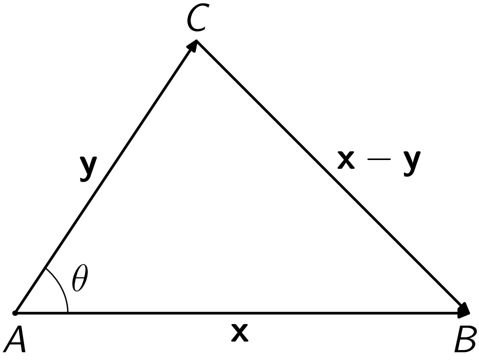
Pythagorean Theorem:

$$\mathbf{x} \perp \mathbf{y} \implies |\mathbf{x}+\mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2$$

3-*dimensional Pythagorean Theorem:*

If vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are pairwise orthogonal then

$$|\mathbf{x}+\mathbf{y}+\mathbf{z}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 + |\mathbf{z}|^2$$



Law of cosines:

$$|\mathbf{x}-\mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2|\mathbf{x}|\,|\mathbf{y}|\,\cos\theta$$

Beyond linearity: dot product

The **dot product** of vectors $\mathbf{x}$ and $\mathbf{y}$ is

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}|\,|\mathbf{y}|\,\cos\theta,$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

The dot product is also called the **scalar product**.
Alternative notation: $(\mathbf{x}, \mathbf{y})$ or $\langle \mathbf{x}, \mathbf{y} \rangle$.

The vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

Relations between lengths and dot products:

- $|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$
- $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}|\,|\mathbf{y}|$
- $|\mathbf{x} - \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2\,\mathbf{x}\cdot\mathbf{y}$

Vectors: algebraic approach

An n-dimensional coordinate vector is an element of $\mathbb{R}^n$, i.e., an ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of real numbers.

Let $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ be vectors, and $r \in \mathbb{R}$ be a scalar. Then, by definition,

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n),$$

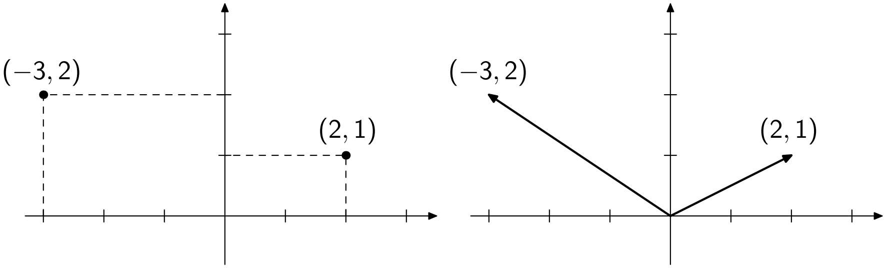
$$r\mathbf{a} = (ra_1, ra_2, \ldots, ra_n),$$

$$\mathbf{0} = (0, 0, \ldots, 0),$$

$$-\mathbf{b} = (-b_1, -b_2, \ldots, -b_n),$$

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}) = (a_1 - b_1, a_2 - b_2, \ldots, a_n - b_n).$$

Cartesian coordinates: geometric meets algebraic



Once we specify an *origin* O, each point A is associated a *position vector* $\overrightarrow{OA}$. Conversely, every vector has a unique representative with tail at O.

Cartesian coordinates allow us to identify a line, a plane, and space with $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$, respectively.

Length and distance

Definition. The **length** of a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ is

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

The **distance** between vectors/points $\mathbf{x}$ and $\mathbf{y}$ is $\|\mathbf{y} - \mathbf{x}\|$.

Properties of length:

$\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0$ only if $\mathbf{x} = \mathbf{0}$ (positivity)

$\|r\mathbf{x}\| = |r|\,\|\mathbf{x}\|$ (homogeneity)

$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality)

Scalar product

Definition. The **scalar product** of vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ is

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{k=1}^{n} x_k y_k.$$

Properties of scalar product:

$\mathbf{x} \cdot \mathbf{x} \geq 0$, $\mathbf{x} \cdot \mathbf{x} = 0$ only if $\mathbf{x} = \mathbf{0}$ (positivity)

$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ (symmetry)

$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$ (distributive law)

$(r\mathbf{x}) \cdot \mathbf{y} = r(\mathbf{x} \cdot \mathbf{y})$ (homogeneity)

Relations between lengths and scalar products:

$\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}}$

$|\mathbf{x}\cdot\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$ (Cauchy-Schwarz inequality)

$\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\,\mathbf{x}\cdot\mathbf{y}$

By the Cauchy-Schwarz inequality, for any nonzero vectors $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ we have

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \quad \text{for some} \quad 0 \leq \theta \leq \pi.$$

θ is called the **angle** between the vectors **x** and **y**. The vectors **x** and **y** are said to be **orthogonal** (denoted $\mathbf{x} \perp \mathbf{y}$) if $\mathbf{x}\cdot\mathbf{y} = 0$ (i.e., if $\theta = 90^{\text{o}}$).

Problem. Find the angle θ between vectors $\mathbf{x} = (2, -1)$ and $\mathbf{y} = (3, 1)$.

$\mathbf{x} \cdot \mathbf{y} = 5$, $\|\mathbf{x}\| = \sqrt{5}$, $\|\mathbf{y}\| = \sqrt{10}$.

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|} = \frac{5}{\sqrt{5}\,\sqrt{10}} = \frac{1}{\sqrt{2}} \implies \theta = 45^{\mathrm{o}}$$

Problem. Find the angle ϕ between vectors $\mathbf{v} = (-2, 1, 3)$ and $\mathbf{w} = (4, 5, 1)$.

$$\mathbf{v} \cdot \mathbf{w} = 0 \implies \mathbf{v} \perp \mathbf{w} \implies \phi = 90^{\mathrm{o}}$$