

AIDOS LAB
AI FOR DATA-ORIENTED SCIENCE

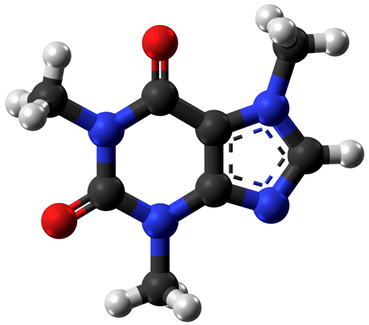
UNI FR

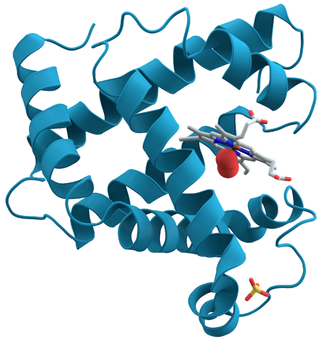
Lecture 2: Geometric and Topological Data Domains

Geometry and Topology in Machine Learning

Inés García Redondo — September 22, 2026

Relational and geometric data are all around us...



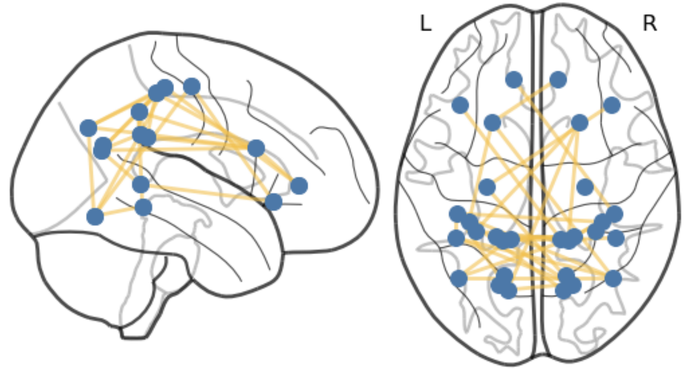


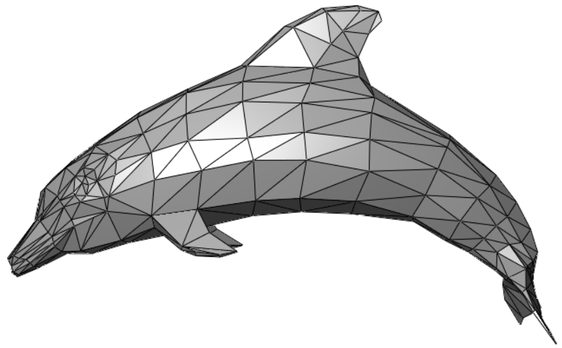
Molecules and proteins

Collaboration networks

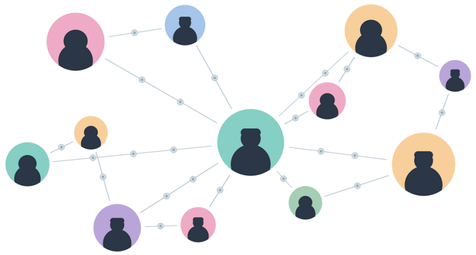


Brain connectome



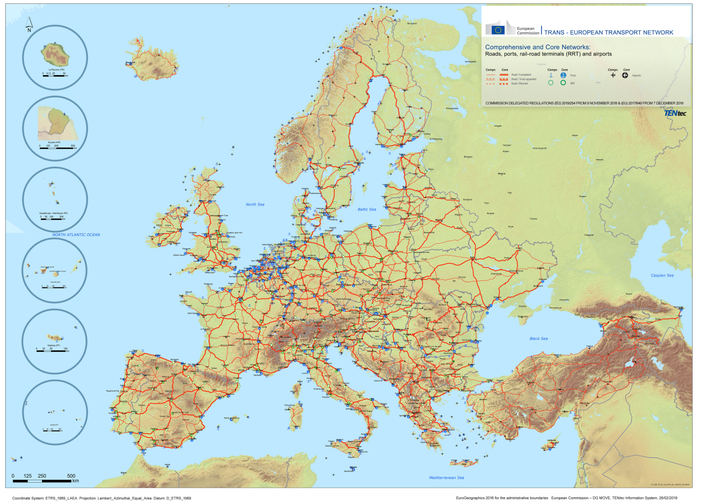


Meshes in computer vision



Social networks

Transportation networks



By European Commission

https://ec.europa.eu/transport/infrastructure/tentec/tentec-portal/site/en/maps.html, CC BY 4.0, https://commons.wikimedia.org/w/index.php?curid=146628735

Session 2 — Topological data and how to analyze them

Outline

- Graphs
  - Matrix representations of graphs
  - Graph Laplacian
  - Graph spectral theory
- Simplicial complexes
  - Simplicial Homology
  - Hodge Laplacian
  - Discrete Hodge theory
- Cell complexes
- Hypergraphs

A small prior on linear algebra

Lecture 2 — Geometric and Topological Data Domains

Linear algebra prior

Vector space

Def. Vector space

A vector space over $\mathbb{R}$ is a triple $(U, +, \cdot)$, where

$$+ : U \times U \to U, \qquad \cdot : \mathbb{R} \times U \to U$$

satisfying, for all $u, v, w \in U$ and $a, b \in \mathbb{R}$:

For the *vector addition*:

- Associativity: $u + (v + w) = (u + v) + w$
- Commutativity: $u + v = v + u$
- Existence of the identity element: $\exists 0 \in U$ s.t. $u + 0 = u$, $\forall u \in U$
- Existence of inverse elements: for all $u \in U$ there exists $-u \in U$ such that $u + (-u) = 0$

For the *scalar multiplication*:

- Associativity: $a \cdot (b \cdot u) = (ab) \cdot u$
- Distributivity of scalar addition: $(a + b) \cdot u = a \cdot u + b \cdot u$
- Distributivity of vector addition: $a \cdot (u + v) = a \cdot u + b \cdot v$
- Identity element for scalar multiplication: $\exists 1 \in \mathbb{R}$ s.t. $1 \cdot u = u$, $\forall u \in U$

Def. Linear map

Mapping $f : U \to W$ between U, W vector spaces such that

- $f(u_1 + u_2) = f(u_1) + f(u_2)$ for all $u_1, u_2 \in U$
- $f(a \cdot u) = a \cdot f(u)$ for all $a \in \mathbb{R}$ and $u \in U$

Linear algebra prior

Bases of vector spaces

Thm + def. Existence of bases

Any finite-dimensional vector space has a *basis*, i.e., there is a subset of elements $\{u_1, \ldots, u_n\} \subset U$ such that $n = \dim(U)$ and for any vector $u \in U$, we can uniquely write

$$u = \sum_{i=1}^{n} a_i \cdot u_i, \quad a_i \in \mathbb{R}$$

Def. Isomorphism of vector spaces

A linear map $f : U \to W$ between vector spaces is an isomorphism iff:

- It is an *injective* map: for all $u_1, u_2 \in U$, $f(u_1) = f(u_2) \implies u_1 = u_2$
- It is a *surjective* map: for all $w \in W$ there exists some $u \in U$ such that

In other words, an isomorphism is a *bijective* linear map

Two vector spaces are isomorphic if they are *essentially the same*, there is a 1-to-1 map between their elements

The definition of basis tells us that the following map is an isomorphism:

$$\begin{array}{ccc} U & \to & \mathbb{R}^n \\ u & \mapsto & (a_1, \ldots, a_n) \end{array}$$

Coordinate vector

In other words, you can forget about this abstraction and just think of $\mathbb{R}^n$, because every finite-dimensional vector space is essentially like it!

Linear algebra prior

The space $\mathbb{R}^n$

- The elements of the space are column vectors: $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$
- Matrices $A \in \mathbb{R}^{n \times m}$ can be seen as *operators* or *linear maps*: $A : \mathbb{R}^m \to \mathbb{R}^n, \quad A\mathbf{x} = \mathbf{y} \in \mathbb{R}^m$
- Square matrices $A \in \mathbb{R}^{n \times n}$ can be seen as *bilinear forms:* $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad \mathbf{y}^\top A\mathbf{x} = \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
- *Kernel of a matrix:* $\ker A = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n$
- *Image of a matrix:* $\operatorname{im} A = \{\mathbf{y} \in \mathbb{R}^m : \exists \mathbf{x} \in \mathbb{R}^n \text{ s.t. } A\mathbf{x} = \mathbf{y}\} \subseteq \mathbb{R}^m$
- *Rank—nullity formula:* $n = \dim \ker A + \dim \operatorname{im} A = \dim \ker A + \operatorname{rank} A$
- The *eigenvectors* and *eigenvalues* of a square matrix $A \in \mathbb{R}^{n \times n}$ are a set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and real numbers $\{\lambda_1, \dots, \lambda_k\}$

$$A\mathbf{u}_i = \lambda_i \mathbf{u}_i, \qquad \forall 1 \leq i \leq k$$

Notice $k \neq n$ in principle

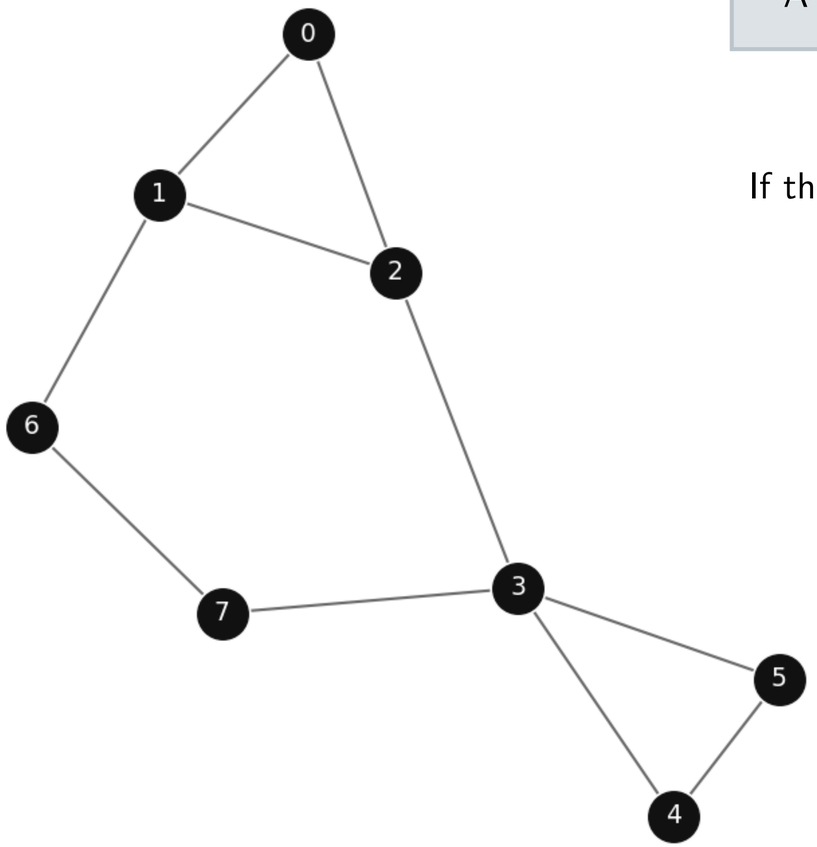
- A matrix $A \in \mathbb{R}^{n \times n}$ is *diagonalizable* if it has n eigenvalues. In that case, you can find a basis of *orthonormal* eigenvectors

Graphs

Lecture 2 — Geometric and Topological Data Domains

Graphs

Definition



Def. Graph

A graph is a pair $G = (V, E)$ where V is the *set of vertices* and E is the set of edges

If the graph is *undirected*, the edges are *unordered pairs* of vertices $\{i, j\} \in E$ for some $i, j \in V$

Example: undirected graph

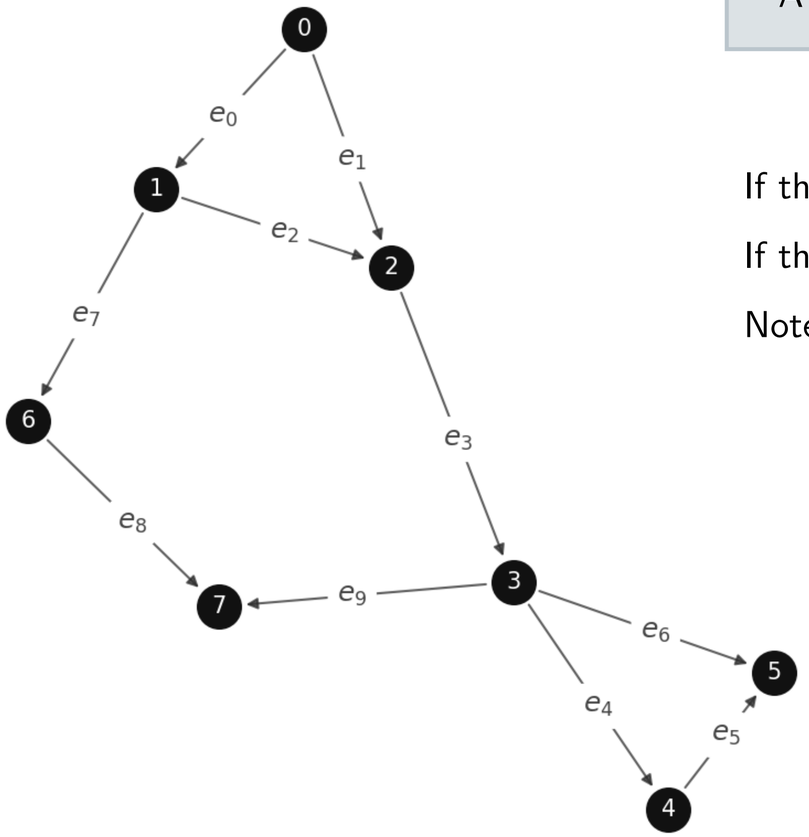
$$V = \{0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7\}$$

$$E = \{\{0,1\}, \{0,2\}, \{1,2\}, \{1,6\}, \{2,3\}, \{3,4\}, \{3,5\}, \{3,7\}, \{6,7\}\}$$

Note that we could equally write $\{0,1\} = \{1,0\}$

Graphs

Definition



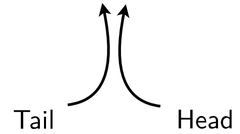
Def. Graph

A graph is a pair $G = (V, E)$ where V is the *set of vertices* and E is the set of edges

If the graph is *undirected*, the edges are *unordered pairs* of vertices $\{i, j\} \in E$ for some $i, j \in V$

If the graph is *directed*, the edges are *ordered pairs* of vertices $(i, j) \in E$ for some $i, j \in V$

Note that, in the second case, $(i, j) \neq (j, i)$



Example: directed graph

$V = \{0, 1, 2, 3, 4, 5, 6, 7\}$

$E = \{(0,1), (0,2), (1,2), (1,6), (2,3), (3,4), (3,5), (3,7), (6,7)\}$

In this case, if we wrote $(0,1)$ it would mean the edge in the opposite direction

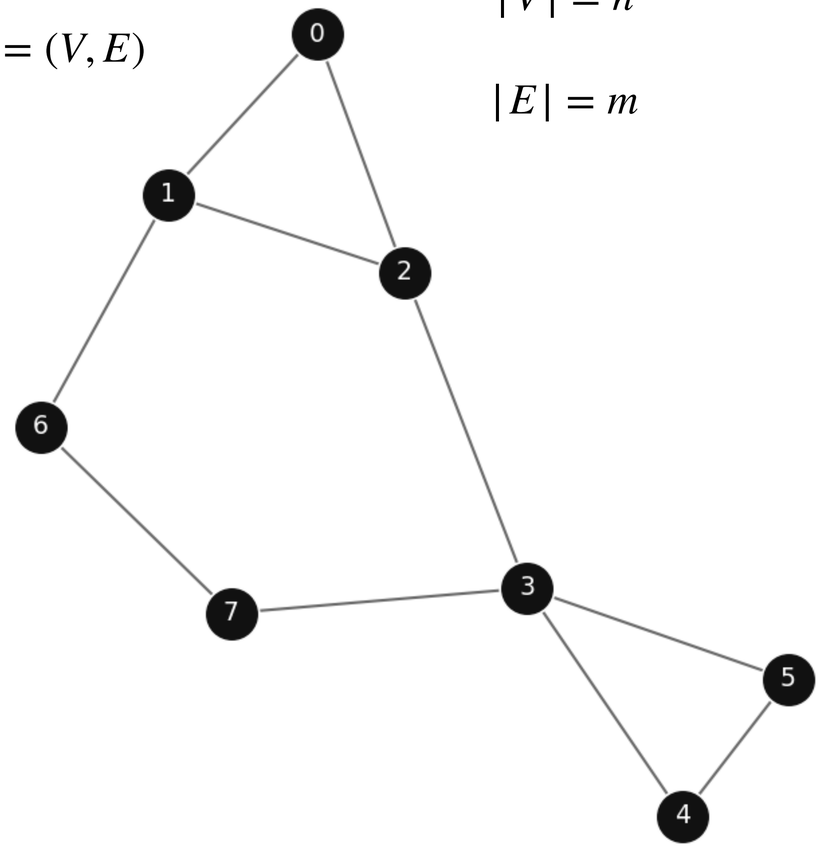
Graphs

Adjacency matrix

$G = (V, E)$

$|V| = n$

$|E| = m$



We can encode the information about the edges present in a graph (also called the adjacencies between vertices) in an **adjacency matrix**

$$A \in \mathbb{R}^{n \times n}, \quad A_{ij} = \begin{cases} 1, & \text{if } (i,j) \in E \\ 0, & \text{otherwise} \end{cases}$$

Note that, for an undirected graph, this matrix is symmetric since $\{i, j\} = \{j, i\}$

Example:

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

$V = \{0, 1, 2, 3, 4, 5, 6, 7\}$

$E = \{\{0,1\}, \{0,2\}, \{1,2\}, \{1,6\}, \{2,3\}, \{3,4\}, \{3,5\}, \{3,7\}, \{6,7\}\}$

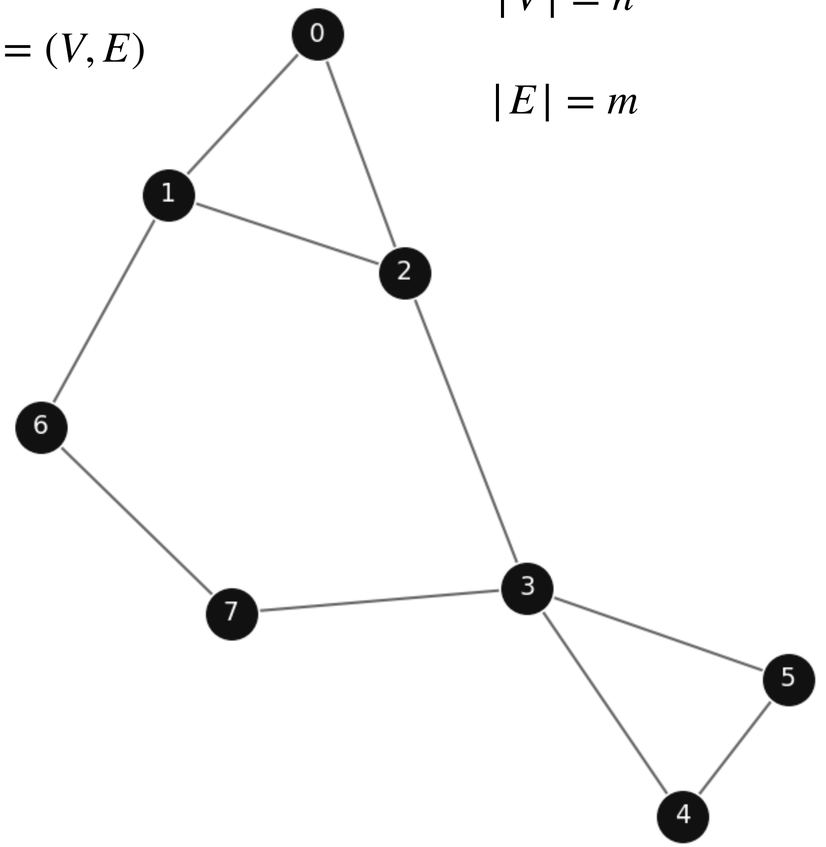
Graphs

Degree matrix

$G = (V, E)$

$|V| = n$

$|E| = m$



Neighbors

$$\mathcal{N}_i = \{ j \in V \mid (i,j) \in E \} = \{ j \in V \mid A_{ij} = 1 \}$$

Degree

$$d_i = |\mathcal{N}_i| = \sum_j A_{ij}$$

Degree matrix

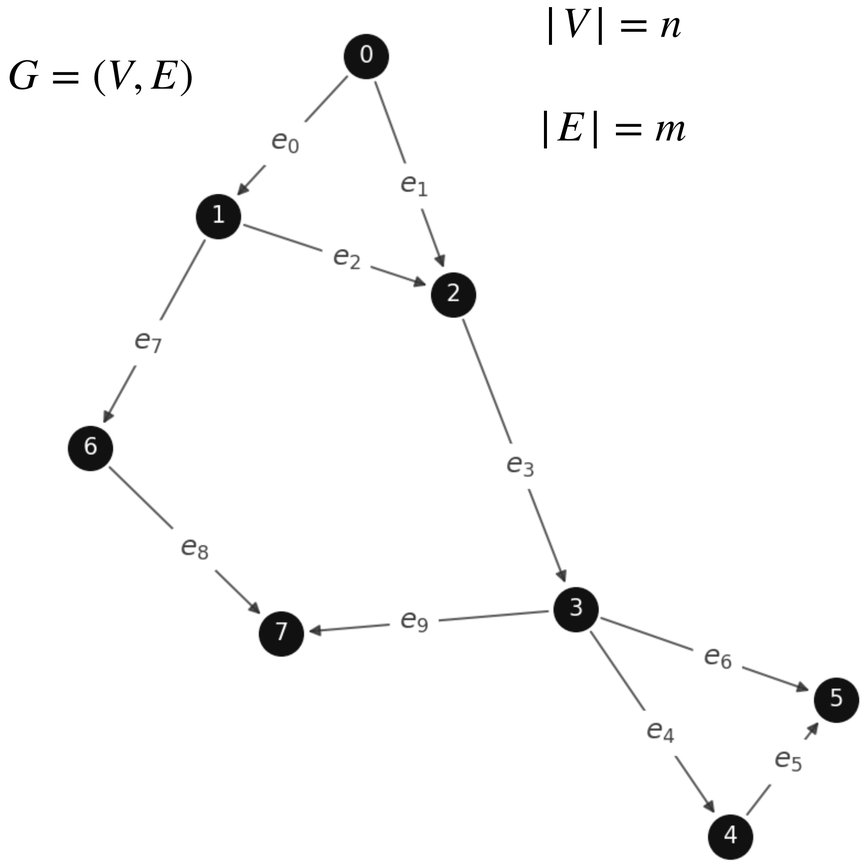
Diagonal matrix $D \in \mathbb{R}^{n \times n}$ storing the degrees of the vertices

Example:

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

Graphs

Incidence matrix



For an undirected graph, we choose an *orientation* of the edges and we can define the **(signed) incidence matrix** as

$$B \in \mathbb{R}^{n \times m}, \qquad B_{ve} = \begin{cases} -1 & \text{if vertex } v \text{ is the tail of edge } e, \\ 1 & \text{if vertex } v \text{ is the head of edge } e, \\ 0 & \text{otherwise.} \end{cases}$$

Example:

$$B = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

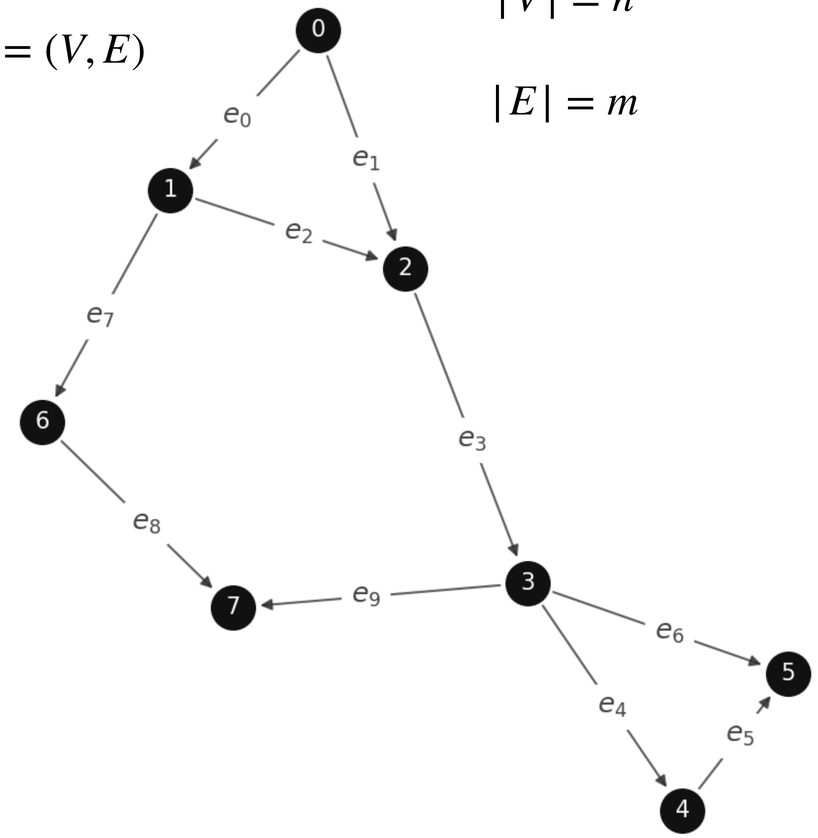
Graphs

Graph Laplacian

$G = (V, E)$

$|V| = n$

$|E| = m$



With the incidence matrix, we define one of the most important matrices in graph theory: the **graph Laplacian**

$$L = BB^\top$$

Example:

$$L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & 0 & 0 & -1 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & -1 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 2 \end{bmatrix}$$

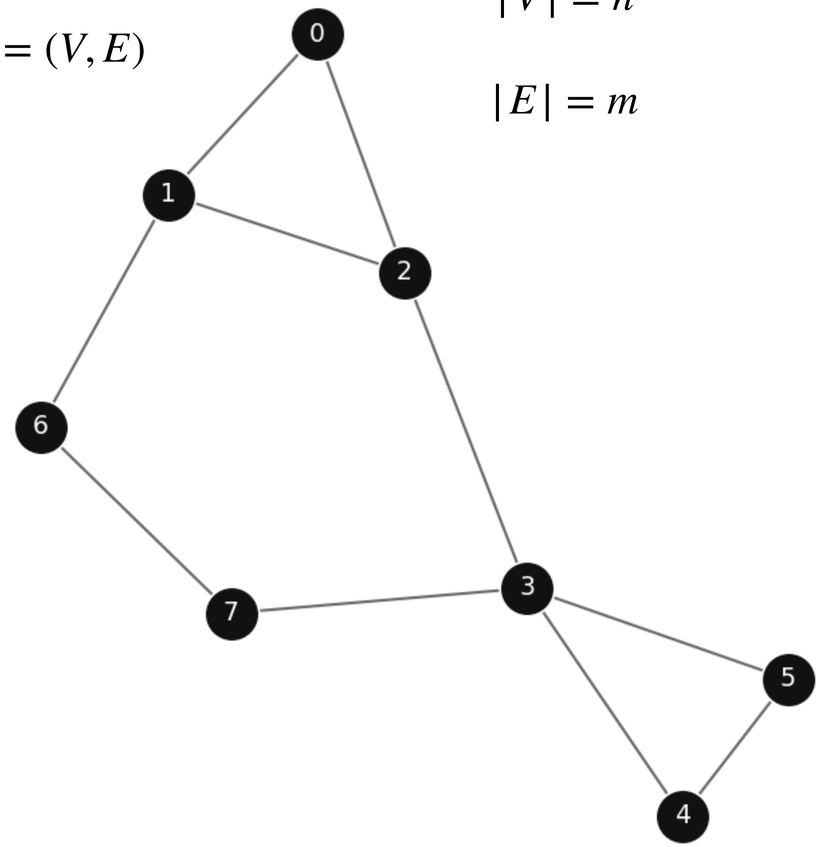
Graphs

Graph Laplacian

$G = (V, E)$

$|V| = n$

$|E| = m$



Lemma: the graph Laplacian has the two following (equivalent) definitions

$$L = BB^\top \qquad L = D - A$$

Proof:

$$(BB^\top)_{uv} = \sum_{e \in E} B_{ue} B_{ve}$$

If $u = v$

$$\sum_{e \in E} B_{ue}^2 = d(u)$$

If $u \neq v$

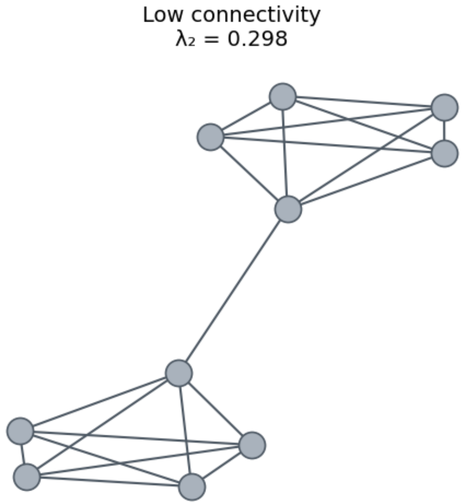
$$B_{ue} B_{ve} = \begin{cases} -1 & (u, v) \in E \\ 0 & \text{otherwise} \end{cases}$$

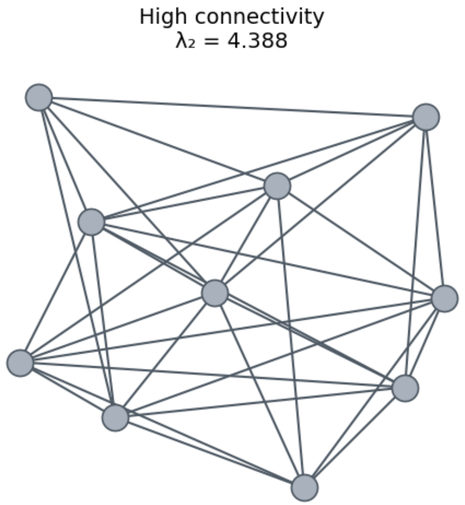
$$\sum_{e \in E} B_{ue} B_{ve} = -1 \iff (u, v) \in E$$

Remark: it does not depend on the choice of orientation

Graphs

Graph Laplacian properties





- L is **symmetric** and **positive semi-definite**

  Proof: $L^\top = (BB^\top)^\top = (B^\top)^\top B^\top = BB^\top = L$ (symmetry)

  $\mathbf{x}^\top L\mathbf{x} = \mathbf{x}^\top BB^\top\mathbf{x} = (B^\top\mathbf{x})^\top B^\top\mathbf{x} = \|B^\top\mathbf{x}\|^2 \geq 0,\ \forall\mathbf{x} \in \mathbb{R}^n$ (positive semi-definiteness)

  $\implies L$ is diagonalizable and all its eigenvalues are non-negative

- All rows and columns sum to 0

Proof: $\sum_{u\in V} L_{uv} = \sum_{u\in V} (D-A)_{uv} = d(u) - |N(u)| = 0$

- 0 is always an eigenvalue:

Proof: $L\mathbf{1} = D\mathbf{1} - A\mathbf{1} = 0$

We will prove these in the course

- $\dim\ker L = \#\{\text{connected components } G\}$

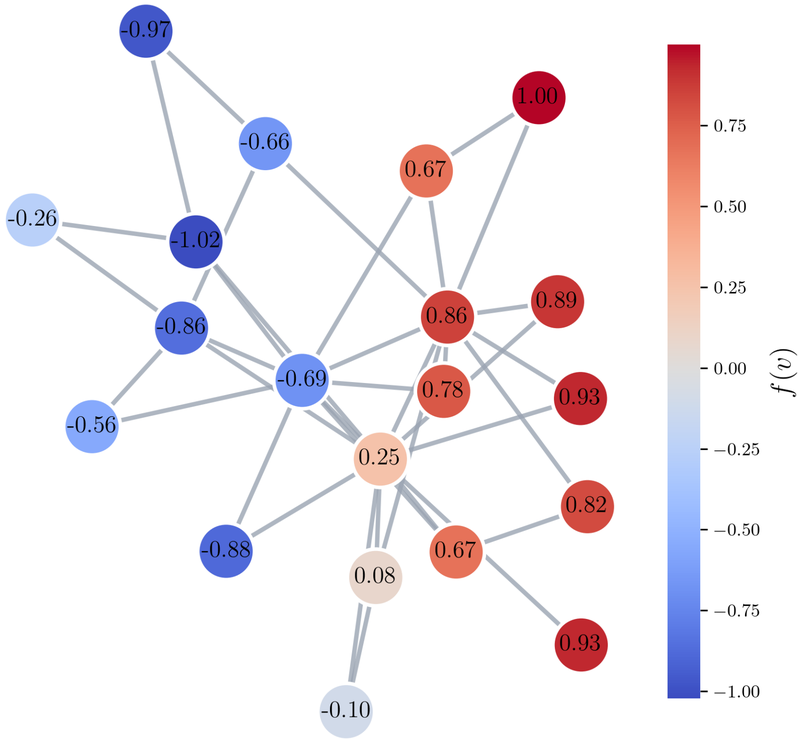
- **Fiedler value** (λ_2): smallest non-zero eigenvalue, tells us how "well connected" the graph is

  1. Small $\lambda_2 \implies$ graph almost disconnected
  2. Large $\lambda_2 \implies$ graph well knit together

Graphs

Graph Laplacian for signal processing

L is usually understood as acting over *signals* on the graph, i.e., functions $f : V \to \mathbb{R}$



The space of signals forms a *vector space*, whose elements we can represent by coordinate vectors as:

$$\mathbf{f} = (f_1, \ldots, f_n) \in \mathbb{R}^n \qquad f_i = f(i) \text{ for } i \in V$$

The Laplacian can be seen as:

1. A *matrix operator* acting over signals:

$$L : \mathbb{R}^n \to \mathbb{R}^n \qquad \mathbf{f} \in \mathbb{R}^n \mapsto L\mathbf{f} \in \mathbb{R}^n$$

$$(L\mathbf{f})_i = d_i f_i - \sum_{j \sim i} f_j$$

How much the signal differs on average with respect to the neighbors

2. A *quadratic form* acting over signals:

$$\mathbf{f}^\top L\mathbf{f} = \sum_{(i,j) \in E} (f_i - f_j)^2$$

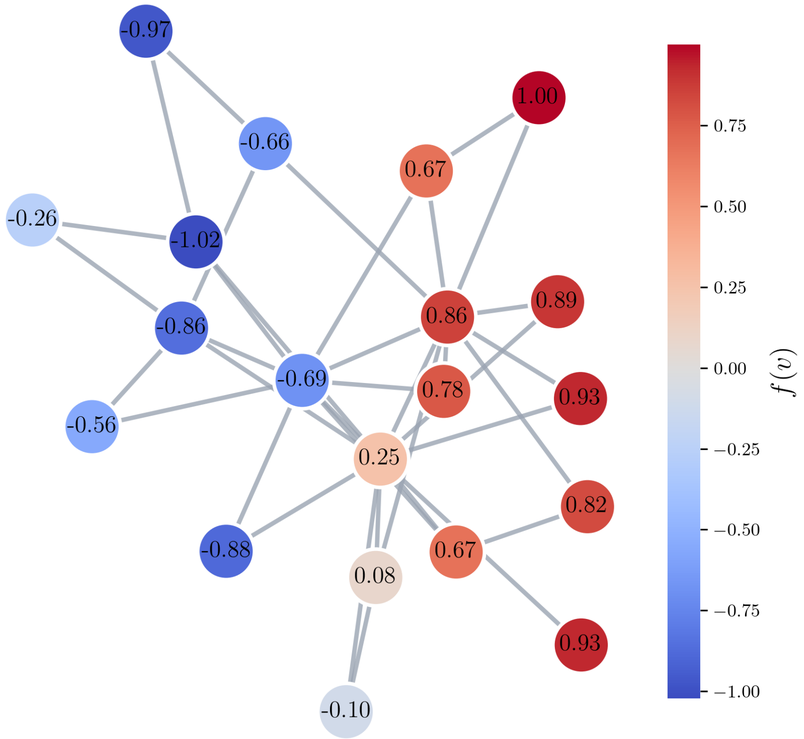
How smooth the signal is

Refer to the exercises

Graphs

Graph Fourier Decomposition

In the same way that we can decompose periodic functions into their Fourier series, we can use the Laplacian to decompose signals (functions over the vertices) using the eigenvector decomposition of the graph Laplacian



Graph eigenvectors and eigenvalues:

$$Lu_i = \lambda_i u_i,\ u_i \in \mathbb{R}^n \text{ and } \lambda_i \in \mathbb{R} \text{ for } 1 \leq i \leq n$$

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

The u_i can also be seen as signals, called *modes* of *frequency* λ_i, and form a basis of the space of signals. If they are orthonormal:

$$u_i^\top L u_i = \lambda_i \|u_i\|^2 = \lambda_i$$

- Low frequency modes correspond to smoother signals (recall, $u_1 = \mathbf{1}$)
- High frequency modes correspond to signals that fluctuate a lot

Graph Fourier decomposition:

$$\mathbf{f} = \sum_{i=1}^{n} a_i u_i$$

You can *low-pass filter*: keep only some of the low frequency components

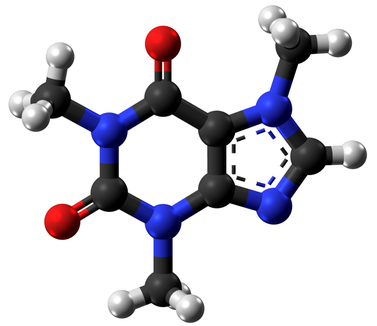
Going beyond pairwise interactions

What happens if we **add interactions between three or more entities at the same time**, instead of only pairwise connections as we have in graphs?



Papers, often times, have more than 2 authors.
How can we encode that?

Can we add information about the presence of *aromatic rings* in a molecule to our data representation?

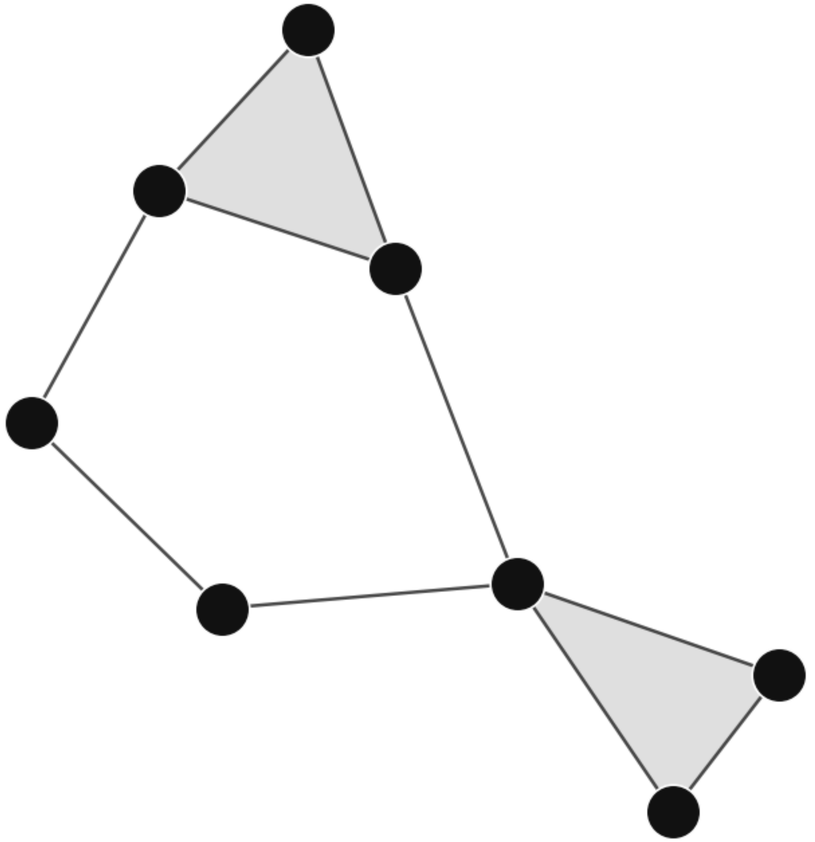


Simplicial complexes

Lecture 2 — Geometric and Topological Data Domains

Simplicial complexes

Intuition



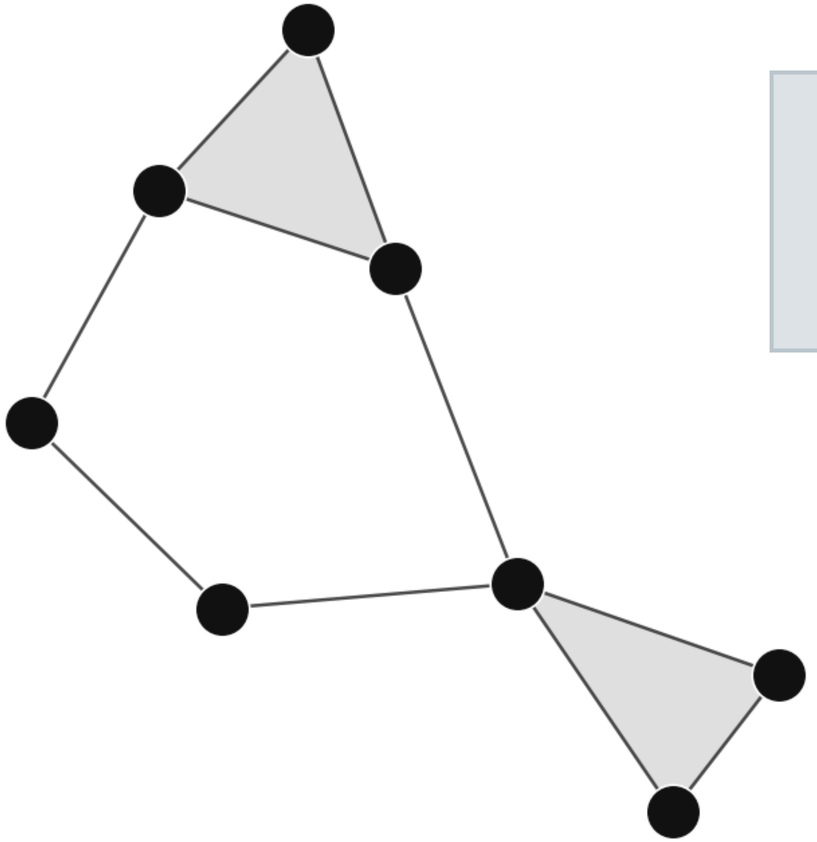
The next natural step is to add interactions between three entities, that is, **triangles**

We can request that, whenever we have a triangle, we have all the pairwise interactions, as if we were filling in the graph

The definition of simplicial complexes is based on this general principle

Simplicial complexes

Definition



Def. (Abstract) Simplicial Complex

Given a discrete set $S = \{s_1, \ldots, s_n\}$, a simplicial complex X is a set of subsets $\sigma \subset S$ *closed under inclusion* (if $\sigma \in X$, then for all $\tau \subseteq \sigma$ we must have $\tau \in X$)

We can *stratify* a simplicial complex by dimension:

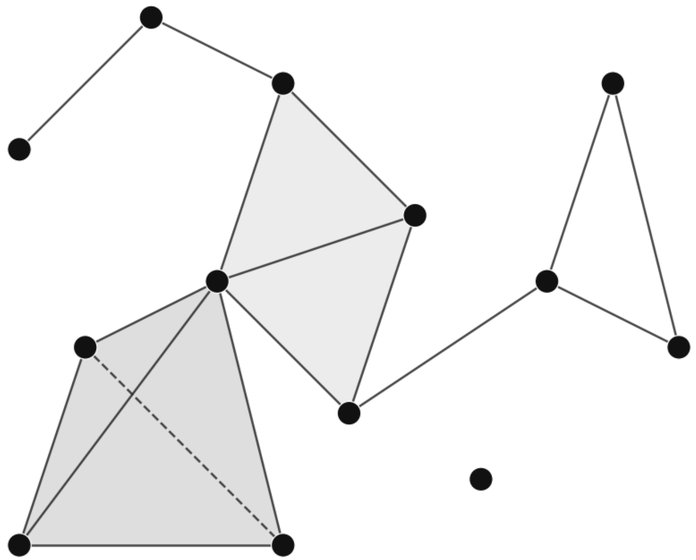
Vertices: 0-dimensional simplices

Edges: 1-dimensional simplices

Triangles: 2-dimensional simplices

Simplicial complexes

General definitions



In general:

- k-**dimensional simplex**: $\sigma = \{s_{i_0}, \ldots, s_{i_k}\}$
- X_k: set of k-dimensional simplices
- X has dimension $d =$ the dimension of its highest dimensional simplex
- k-**skeleton** $(X_{\leq k})$: all simplices with dimension $\leq k$

Example: structure on the left

We have a set of vertices (X_0), a set of edges (X_1), a set of triangles (X_2) and a set with one tetrahedron (X_3)

Note that when a triangle in present, all the edges that define it are also present

The presence of the tetrahedron also means that all triangles and all edges in it are also in the simplicial complex

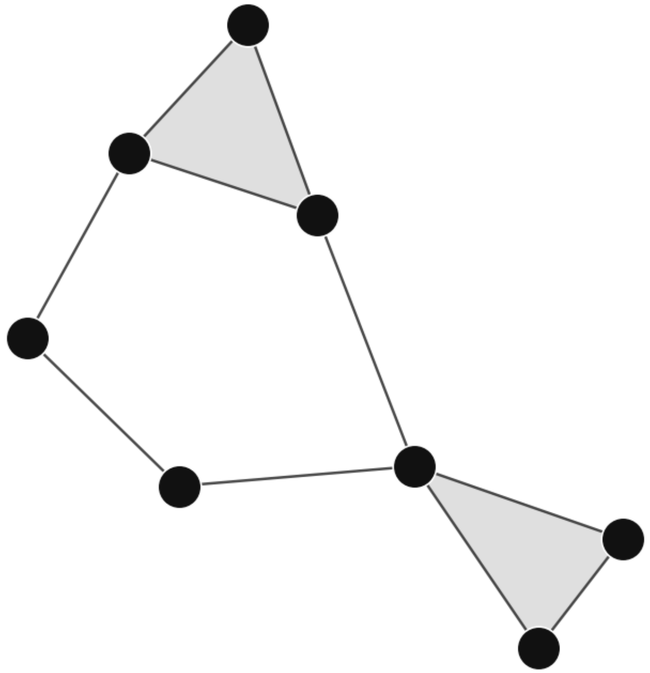
Simplicial homology

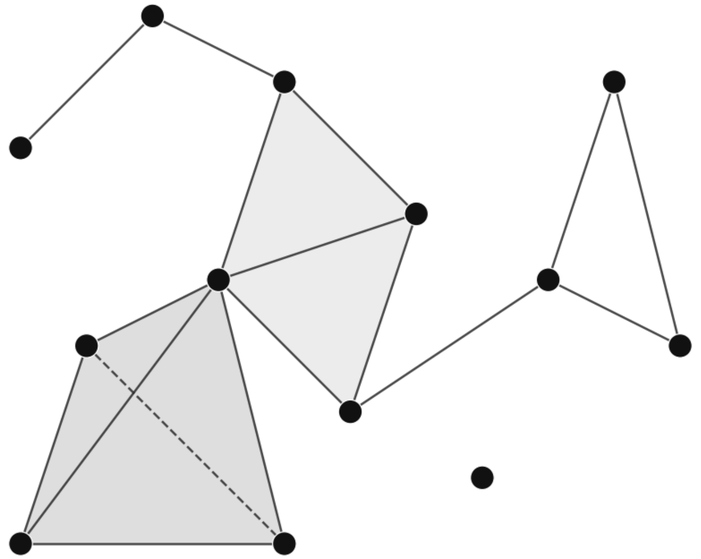
Lecture 2 — Geometric and Topological Data Domains

Simplicial homology

Motivation

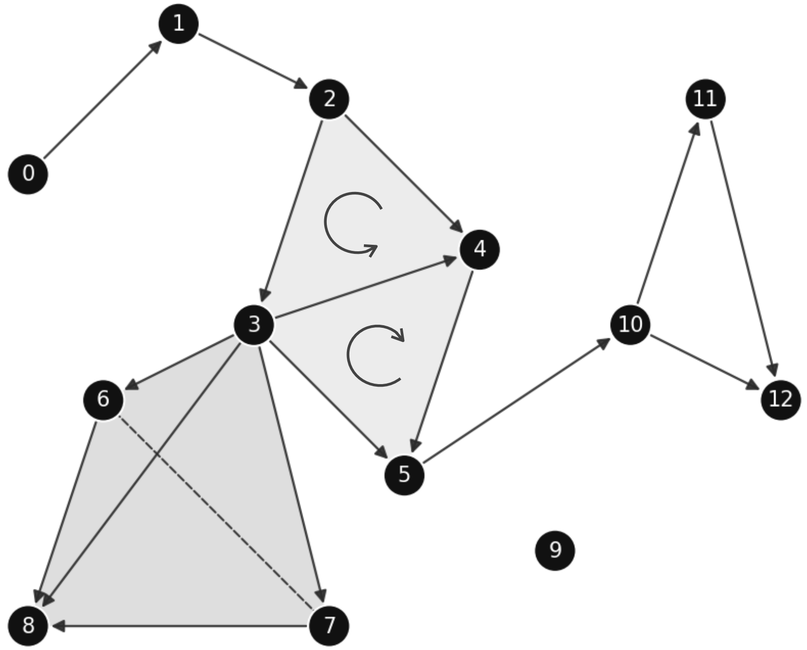
We want to be able to say that the following figures have *holes* in different dimensions





Simplicial homology

Chain spaces and boundary operators



Choose an **orientation** for the simplices

For example:

- Choose an ordering of the vertices $S = \{1,\ldots,n\}$
- Orient higher order simplices accordingly: $(i_0, \ldots, i_k)$, $i_0 < \ldots < i_k$

Chain space: let $k \geq 0$, the k-chain space is the *vector space* spanned by the *oriented simplices*

$$C_k(X) = \operatorname{span}\{(i_0, \ldots, i_k) : i_0 < \cdots < i_k\} \cong \mathbb{R}^{|X_k|}$$

$$c = \sum_{\sigma \in X_k} a_\sigma \cdot \sigma, \quad a_\sigma \in \mathbb{R}$$

Boundary operator: *linear map* connecting chain spaces of consecutive dimension $\partial_k : C_k(X) \to C_{k-1}(X)$ defined on the basis elements as

$$\partial_k(i_0, \ldots, i_k) = \sum_{j=0}^{k} (-1)^j \ (i_0, \ldots, \widehat{i_j}, \ldots i_k)$$

$$\partial_k \, \partial_{k+1} = 0 \implies \ker \partial_k \subseteq \operatorname{im} \partial_{k+1}$$

$C_0(X, \mathbb{R}) = \mathbb{R}^{13}$

$C_1(X, \mathbb{R}) = \mathbb{R}^{17}$

$C_2(X, \mathbb{R}) = \mathbb{R}^{6}$

$C_3(X, \mathbb{R}) = \mathbb{R}$

$$\partial_1(0,1) = (1) - (0)$$
$$\partial_2\,(2,3,4) = (3,4) - (2,4) + (2,3)$$
$$\partial_3(3,6,7,8) = (6,7,8) - (3,7,8) + (3,6,8) - (3,6,7)$$

Mathematical interlude

Quotient vector spaces

Def. Quotient vector space

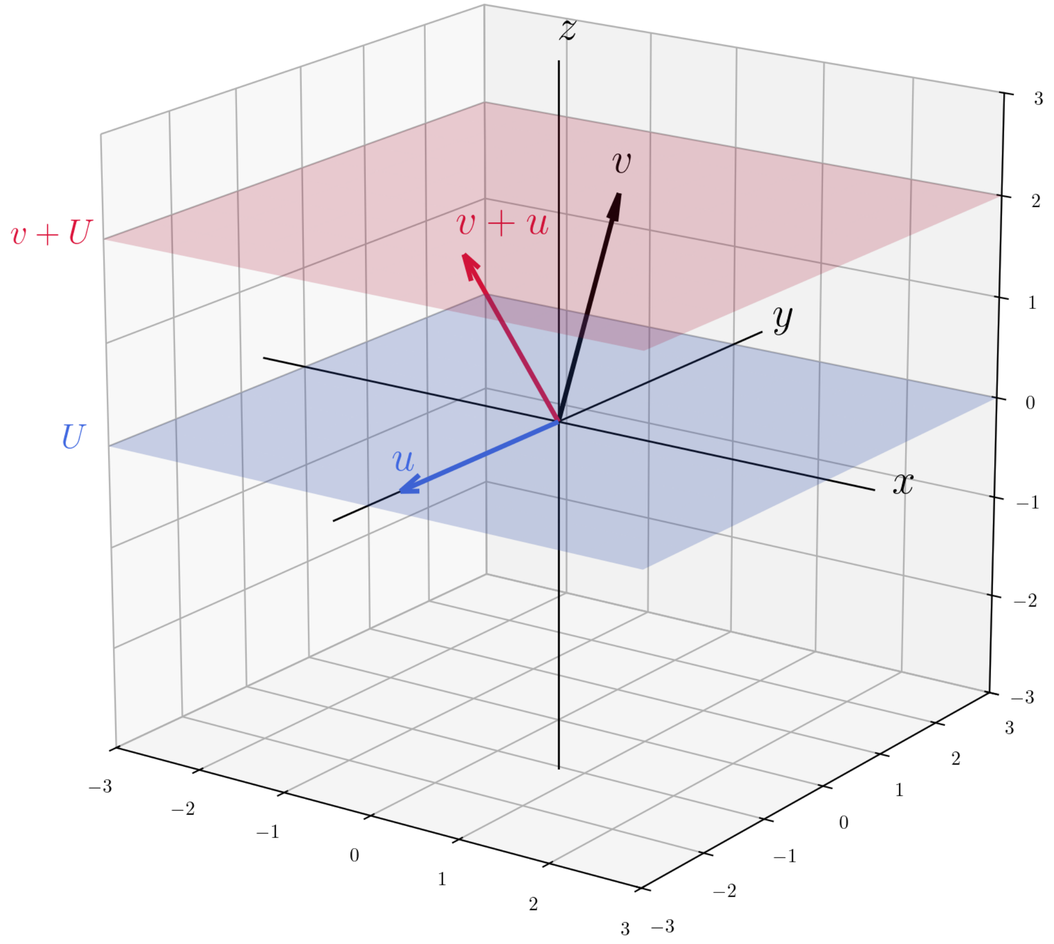
Given two subspaces $U, V \subset \mathbb{R}^n$ with $U \subseteq V$, the quotient space $\dfrac{V}{U}$ is defined as the vector space given by the *equivalence classes*:

$$v + U \coloneqq \{v + u : u \in U\}$$

This is indeed a vector space as it is closed under addition and multiplication by scalars

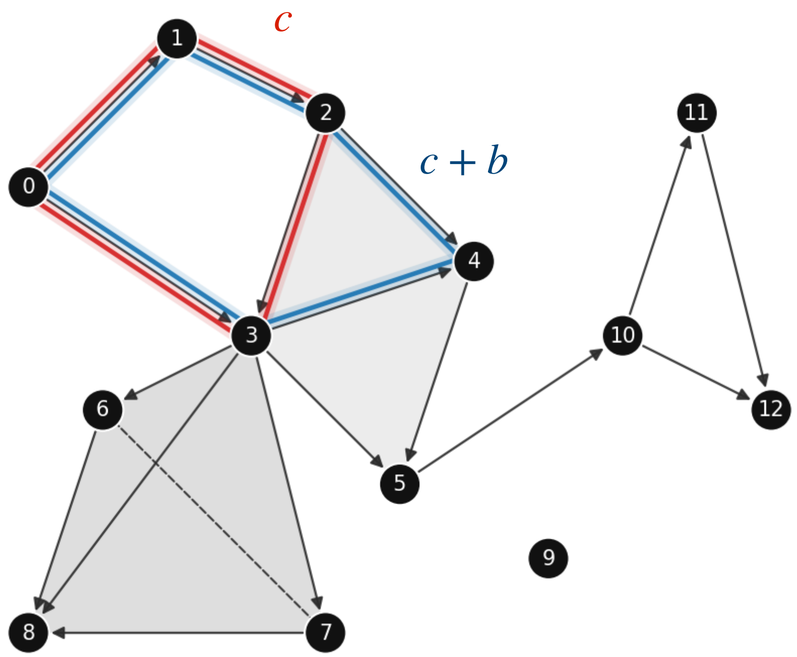
If you want to sum the elements $v_1 + U$ and $v_2 + U$ you take representatives of the classes, sum them, and see what class you end up in

$$(v_1 + u_1) + (v_2 + u_2) = (v_1 + v_2) + (u_1 + u_2) \in (v_1 + v_2) + U$$



Simplicial homology

Homology groups



If I can move through filled in triangles, I stay inside of the homology class

Def. Homology group

Cycles

$$H_k(X) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}$$

Boundaries

Example: What are the elements of the homology group?

First, we **find a cycle:**

$c = (0,1) + (1,2) + (2,3) - (0,3) \in \ker \partial_1$

$\partial_1 (c) = (1) - (0) + (2) - (1) + (3) - (2) - (3) + (0) = 0$

An element of $H_1(X)$ is given by $c + \operatorname{im} \partial_2 = \{c + \partial_2 \tau : \tau \in C_2(X)\}$

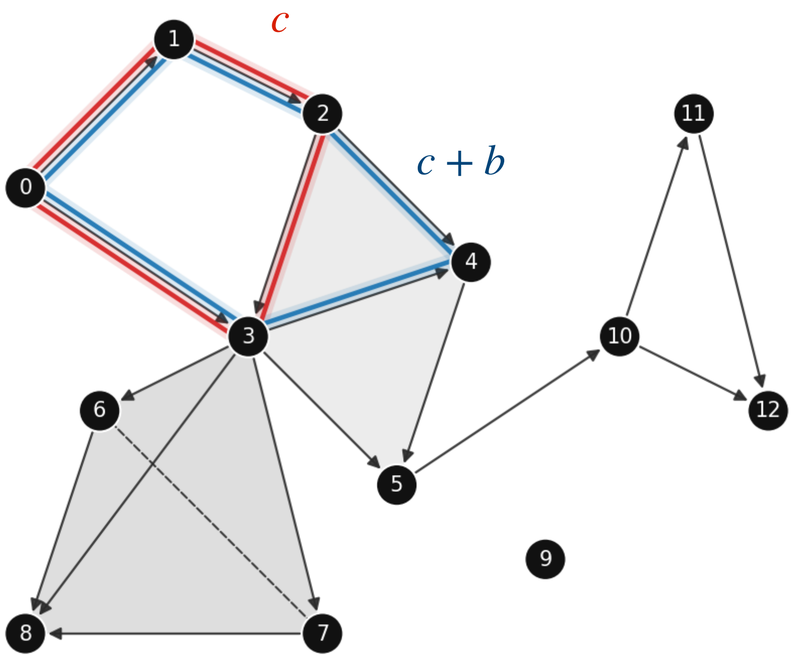
For instance, an **element of the class** is:

$c + b = (0,1) + (1,2) + (2,4) - (3,4) - (0,3)$

$b = -(2,3) - (3,4) + (2,4) = \partial_2[-(2,3,4)] \in \operatorname{im} \partial_2$

Simplicial homology

Betti numbers



Def. Betti number

$$\beta_k(X) = \dim H_k(X)$$

"Number of k-dimensional holes in X"

- $k = 0$: number of connected components
- $k = 1$: number of loops
- $k = 2$: number of cavities or bubles

Example:

$$\beta_0(X) = 2 \qquad \beta_1(X) = 2$$

$$\beta_2(X) = 1 \qquad \beta_k(X) = 0,\ k \geq 3$$

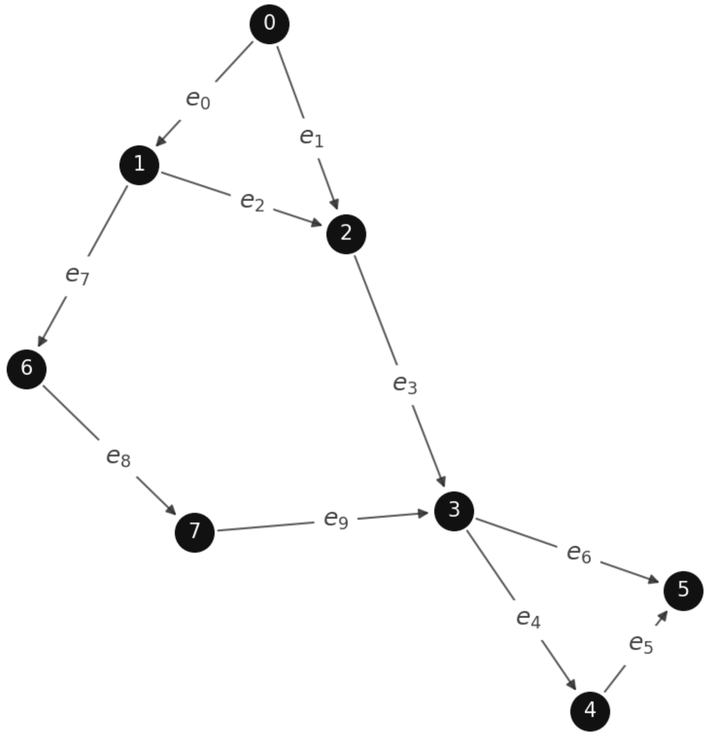
Because the tetrahedron is filled!

There are no simplices with dimension greater than 3 so

$$C_k(X) = 0,\ k \geq 3 \implies H_k(X) = 0,\ k \geq 3$$

Simplicial homology

Incidence matrices revisited



$$C_1(G) \xrightarrow{\partial_1} C_0(G)$$

$$C_1(G) \cong \mathbb{R}^{|E|} \qquad C_0(G) \cong \mathbb{R}^{|V|}$$

$$B \in \mathbb{R}^{n \times m}, \qquad B_{ve} = \begin{cases} 1 & \text{if vertex } v \text{ is the tail of edge } e, \\ -1 & \text{if vertex } v \text{ is the head of edge } e, \\ 0 & \text{otherwise.} \end{cases}$$

$$B = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \partial_1$$

$$\partial_1(u, v) = \{v\} - \{u\}$$

Boundary maps extend the incidence matrix to higher dimensions!

Hodge Theory

Lecture 2 — Geometric and Topological Data Domains

Hodge Theory

Inner products

So far we have only used the *vector space* structure of the chain spaces

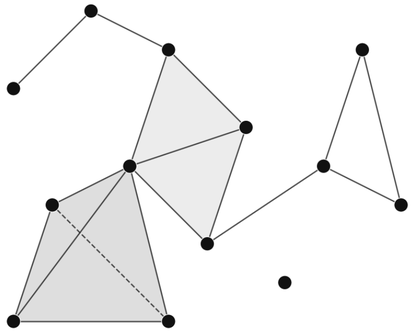
We can also endow these spaces with *inner products* $\langle \bullet, \bullet \rangle_{C_k} : C_k(X) \times C_k(X) \to \mathbb{R}$

Example: standard inner product (also called ℓ^2) $\langle \bullet, \bullet \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}; \quad \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y} \quad \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$

Having and inner product allows us to have norms: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\mathbf{x}^\top \mathbf{x}}$

And normalize vectors: $\tilde{u} = \dfrac{u}{\|u\|}$

Example: weighted simplicial complexes $(X, w),$ with $w : X \to \mathbb{R}_{>0}$



$$C_k(X) = \operatorname{span}\{(i_0, \dots, i_k) : i_0 < \cdots < i_k\} \cong \mathbb{R}^{|X_k|}$$

$$\langle \sigma, \tau \rangle_{C_k} = \begin{cases} w(\sigma) & \text{if } \sigma = \tau \\ 0 & \text{otherwise} \end{cases}$$

The k-simplices form an orthogonal basis of $C_k(X)$ with norms prescribed by their weights:

$$\|\sigma\| = \sqrt{\langle \sigma, \sigma \rangle} = \sqrt{w(\sigma)}$$

Hodge Theory

Combinatorial Hodge Laplacians

Having inner products allows the definition of *adjoint maps*

Def. Adjoint map

A map $f : \mathbb{R}^n \to \mathbb{R}^m$ has an adjoint map $f^* : \mathbb{R}^m \to \mathbb{R}^n$ which satisfies $\langle \mathbf{x}, f(\mathbf{y}) \rangle_{\mathbb{R}^m} = \langle f^*(\mathbf{x}), \mathbf{y} \rangle_{\mathbb{R}^n}, \ \forall \mathbf{x} \in \mathbb{R}^m, \ \mathbf{y} \in \mathbb{R}^n$

Example: **adjoint boundary maps** $\partial_k^* : C_{k-1}(X) \to C_k(X)$ $\langle \sigma, \partial_k \tau \rangle_{C_{k-1}} = \langle \partial_k^* \sigma, \tau \rangle_{C_k}, \quad \forall \sigma \in C_{k-1}(X), \ \tau \in C_k(X)$

For an unweighted complex (all weights equal 1) $\langle \bullet, \bullet \rangle_{C_k}$ is the standard inner product and $\partial_k^* = \partial_k^\top$

With these boundary maps and their adjoints, we have a sequence of connecting maps: $C_{k+1}(X) \underset{\partial_{k+1^*}}{\overset{\partial_{k+1}}{\rightleftarrows}} C_k(X) \underset{\partial_k^*}{\overset{\partial_k}{\rightleftarrows}} C_{k-1}(X)$

Def. Hodge Laplacian

A map $L_k : C_k(X) \to C_k(X)$ with matrix given by: $L_k = \partial_{k+1} \, \partial_{k+1}^* + \partial_k^* \, \partial_k = L_k^{\text{up}} + L_k^{\text{down}}$

For unweighted simplicial complexes, simply: $L_k = \partial_{k+1} \, \partial_{k+1}^\top + \partial_k^\top \, \partial_k = L_k^{\text{up}} + L_k^{\text{down}}$

Hodge Theory

Hodge decomposition

Hodge Laplacian operator: $L_k = \partial_{k+1}\ \partial_{k+1}^* + \partial_k^*\ \partial_k = L_k^{\text{up}} + L_k^{\text{down}}$

Theorem. Hodge decomposition

The chain space has the following orthogonal decomposition

$$C_k(X) = \operatorname{im} \partial_{k+1} \oplus \ker L_k \oplus \operatorname{im} \partial_k^*$$

Refer to the exercises

Theorem.

$$\ker L_k = \ker \partial_k \cap \ker \partial_{k+1}^* \cong H_k(X)$$

The elements of $\ker L_k$ are called *harmonic representatives*

This proves that: $\dim \ker L_0 = \dim H_0(X) = \beta_0(X) =$ # *connected components*

Example: graph case $G = (V, E)$

$$0 \underset{0}{\overset{0}{\rightleftarrows}} C_1(G) \underset{\partial_1^*}{\overset{\partial_1}{\rightleftarrows}} C_0(X)$$

$$L_0 = L_0^{\text{up}} = \partial_1\ \partial_1^*$$
$$L_1 = L_1^{\text{down}} = \partial_1^*\ \partial_1$$

ℓ^2 inner product

$$\partial_1 = B_1, \qquad \partial_1^* = B_1^\top$$
$$L_0 = B_1\, B_1^\top \qquad L_1 = B_1^\top\, B_1$$

Graph Laplacian

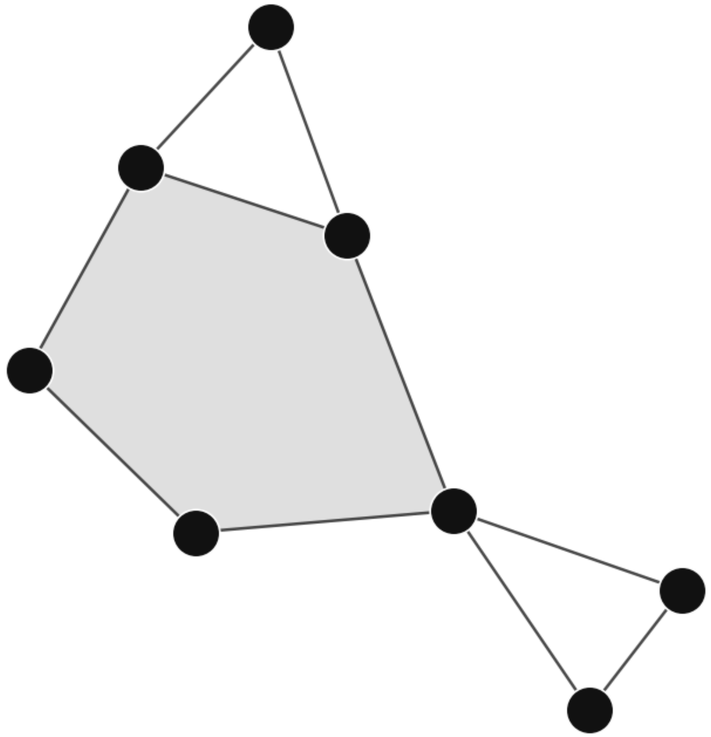
Lim, L.-H. (2020). Hodge Laplacians on Graphs. SIAM Review, 62(3), 685–715. https://doi.org/10.1137/18M1223101

Cell complexes

Lecture 2 — Geometric and Topological Data Domains

Cell complexes

Definition



What if we now wanted to add the pentagon as a quintic interaction, but without requiring all the quartic and triadic interactions as it happens in a simplicial complex?

Def. (2-dimensional) Abstract cell complex

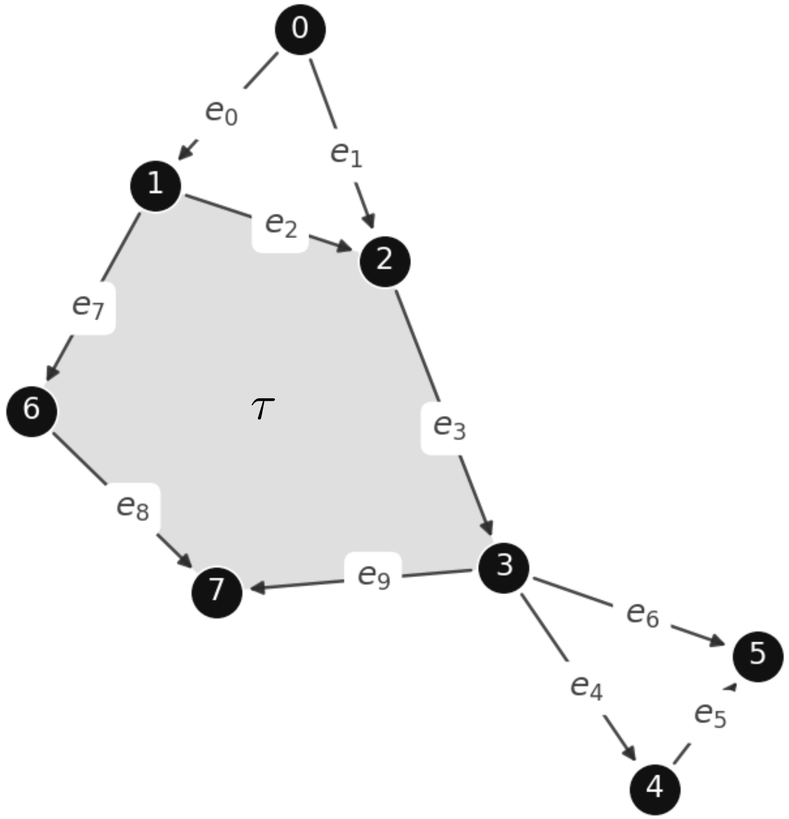
An abstract regular cell complex C of dimension 2 is given by:

- A set $C_0 = \{c_0^1, \dots c_0^n\}$ of 0-cells
- A set $C_1 = \{c_1^1, \dots, c_1^m\}$ of 1-cells
- A set $C_2 = \{c_2^1, \dots, c_2^l\}$ of 2-cells
- Boundary (incidence) matrices
  - $B_1 \in \{0, \pm 1\}^{n \times m}$ such that each column has exactly one $+1$ and one -1 entry
  - $B_2 \in \{0, \pm 1\}^{m \times l}$ such that each column forms a loop in C_1 with the sign corresponding to the orientation of the 1-cell in the loop

Hoppe, J., Grande, V. P., & Schaub, M. T. (2026). Don't be Afraid of Cell Complexes! An Introduction from an Applied Perspective. https://doi.org/10.48550/arXiv.2506.09726

Cell complexes

Example



Oriented edges!

Example: $C_0 = V, \quad C_1 = E, \quad C_2 = \tau$

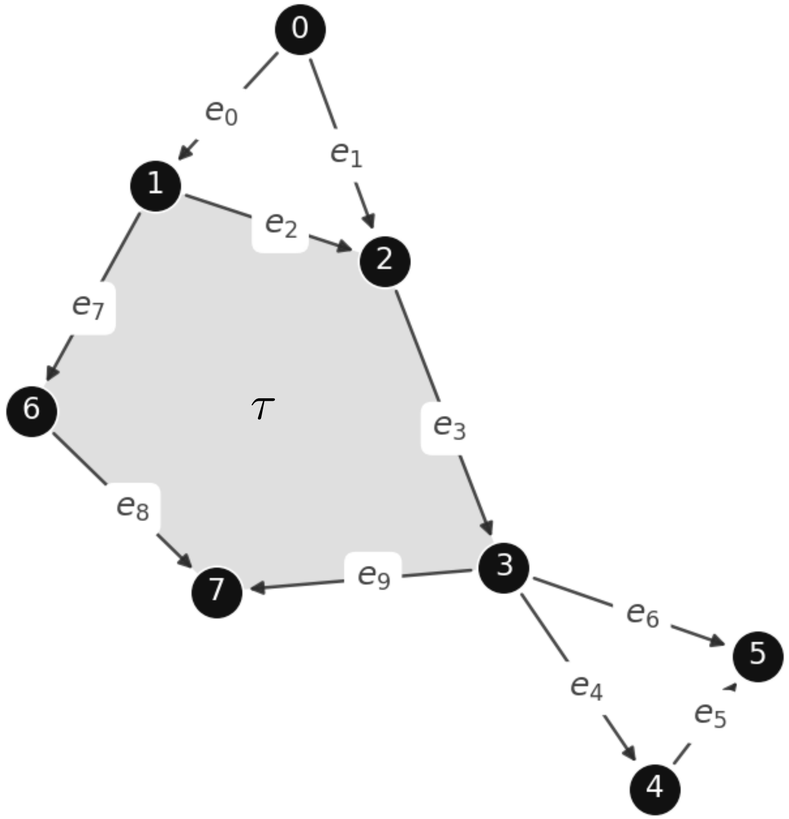
$$B_1 = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad B_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

We still find $C_2 \xrightarrow{B_2} C_1 \xrightarrow{B_1} C_0 \qquad B_1 B_2 = 0$

Hoppe, J., Grande, V. P., & Schaub, M. T. (2026). Don't be Afraid of Cell Complexes! An Introduction from an Applied Perspective. https://doi.org/10.48550/arXiv.2506.09726

Cellular Homology

Definition



We still find $C_2 \xrightarrow{B_2} C_1 \xrightarrow{B_1} C_0$ $B_1 B_2 = 0$

Def. Homology groups

$$H_0(C) = \ker B_1 \qquad H_1(C) = \frac{\ker B_1}{\operatorname{im} B_2}$$

Connected components

Holes/loops

$$\beta_i = \dim H_i(C)$$

Hoppe, J., Grande, V. P., & Schaub, M. T. (2026). Don't be Afraid of Cell Complexes! An Introduction from an Applied Perspective. https://doi.org/10.48550/arXiv.2506.09726

Hypergraphs

Lecture 2 — Geometric and Topological Data Domains

Hypergraphs

Intuition

What if we don't want to impose **any closure under inclusion**?

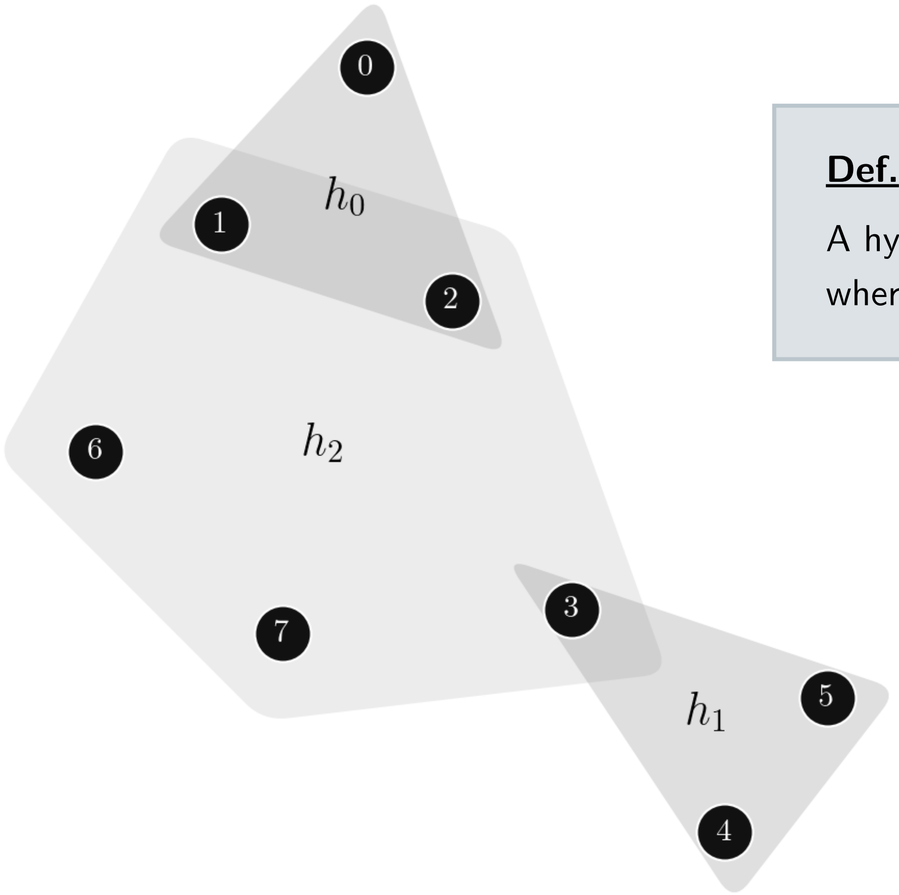


For example, if we want to model the network of papers where nodes are authors and shared papers are the relations, we can have groups of 3 authors that have papers together, where the pairs of authors have never published between them in isolation

This breaks the key defining property of simplicial and cell complexes, where we always ask for closure under inclusion

Hypergraph

Definition



Def. Hypergraph

A hypergdaph $H = (V, E)$ is given by a set of vertices V and a set of hyperedges H where a hyperedge is simply a subset of vertices

Example: $V = \{0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7\}$

$$H = \{\{0, 1, 2\},\ \{1, 2, 3, 6, 7\},\ \{3, 4, 5\}\}$$

To sum up

Today we have learnt how to...

- How to describe graphs through matrices
- How to add higher-order interactions via simplicial complexes
- How to measure topological features using homology
- The definition of the Hodge Laplacian and Hodge decomposition
- How to add higher-order interactions via cellular complexes
- How to add higher-order interactions via hypergraphs

Thank you for your attention!



AIDOS LAB
AI FOR DATA-ORIENTED SCIENCE

UNI
FR