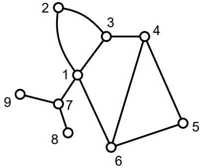
Topological data analysis
Lecture 2

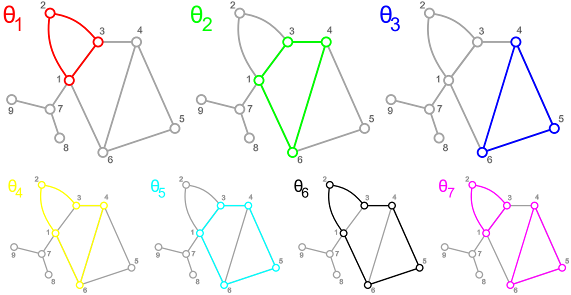
Anton Ayzenberg

ATA Lab, FCS NRU HSE
Noeon Research

Spring 2024
Faculty of Computer Science / Yandex Data School

Counting cycles in a graph





Vector space of cycles

- "The number of cycles" is the number of basic cycles.
- Basic cycles = any collection of cycles, from which all other cycles are expressed uniquely as sums over $\mathbb{Z}_2$ (field with 2 elements).
- In other words, let $Z_1(\Gamma; \mathbb{Z}_2)$ denote the vector space (over $\mathbb{Z}_2$) of all algebraical cycles.
- Algebraical cycle = collection of edges, such that even number of edges meet in each vertex.
- Then the number of cycles = $\dim Z_1(\Gamma; \mathbb{Z}_2)$.

More abstract characterization

- $C_1(\Gamma;\mathbb{Z}_2) = \mathbb{Z}_2\langle\text{edges of }\Gamma\rangle$;
- $C_0(\Gamma;\mathbb{Z}_2) = \mathbb{Z}_2\langle\text{vertices of }\Gamma\rangle$;
- $\partial_1\colon C_1(\Gamma;\mathbb{Z}_2) \to C_0(\Gamma;\mathbb{Z}_2)$.
- where $\partial(\{v_1, v_2\}) = v_1 \oplus v_2$.
- Then $\sigma \in C_1(\Gamma;\mathbb{Z}_2)$ is a cycle iff $\partial_1(\sigma) = 0$.
- Therefore $Z_1(\Gamma;\mathbb{Z}_2) = \operatorname{Ker}\partial_1$.
- We define $H_1(\Gamma;\mathbb{Z}_2) = Z_1(\Gamma;\mathbb{Z}_2)$ and **the first Betti number** $\beta_1(\Gamma) = \dim H_1(\Gamma;\mathbb{Z}_2)$.

First homology, graph case

- We define $H_1(\Gamma;\mathbb{Z}_2) = Z_1(\Gamma;\mathbb{Z}_2)$ and **the first Betti number** $\beta_1(\Gamma) = \dim H_1(\Gamma;\mathbb{Z}_2)$.
- This number counts cycles in a graph.
- In graph theory it is called **circuit rank**.

Remark

The matrix of $\partial_1\colon C_1(\Gamma;\mathbb{Z}_2) \to C_0(\Gamma;\mathbb{Z}_2)$ is the incidence matrix of a graph

$$D_1 = \begin{array}{c|ccc} & \cdots & e & \cdots \\ \hline \vdots & & \vdots & \\ v & \cdots & \varepsilon_{v,e} & \cdots \\ \vdots & & \vdots & \end{array}$$

We have $\beta_1(\Gamma) = \dim \operatorname{Ker} \partial_1 = \#\text{edges} - \operatorname{rk} D_1$. This can be computed by Gauss algorithm.

Graph homology from other invariants

We have $\beta_1(\Gamma) = \dim \operatorname{Ker} \partial_1 = \#\text{edges} - \operatorname{rk} D_1$.

Exercise: $\dim \operatorname{Im} \partial_1 (= \operatorname{rk} D_1) = \#\text{vertices} - \#\text{con.components}$.

Corollary

$$\beta_1(\Gamma) = \#\text{edges} - \#\text{vertices} + \#\text{con.components}.$$

Do you recognize the number $\#\text{vertices} - \#\text{con.components}$?

Graph homology from other invariants

We have $\beta_1(\Gamma) = \dim \operatorname{Ker} \partial_1 = \#\text{edges} - \operatorname{rk} D_1$.

Exercise: $\dim \operatorname{Im} \partial_1 (= \operatorname{rk} D_1) = \#\text{vertices} - \#\text{con.components}$.

Corollary

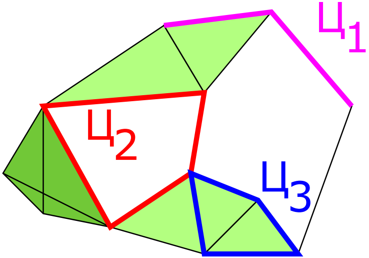
$$\beta_1(\Gamma) = \#\text{edges} - \#\text{vertices} + \#\text{con.components}.$$

Do you recognize the number $\#\text{vertices} - \#\text{con.components}$?

It follows that $\beta_1(\Gamma)$ equals the number of edges remaining after removal of (any) spanning forest.

First homology of a simp. complex

First homology of a simp. complex



	$Ц_1$	$Ц_2$	$Ц_3$
chain?	✓	✓	✓
cycle?	✗	✓	✓
boundary?	✗	✗	✓
homology?	✗	✓	✗

First homology of a simp. complex

- Homology = "cycles, which are not boundaries".
- "Number of holes = number of cycles - number of boundaries".

First homology of a simp. complex

- Homology = "cycles, which are not boundaries".
- "Number of holes = number of cycles - number of boundaries".
- $Z_1(K;\mathbb{Z}_2) = \operatorname{Ker}\partial_1 \colon C_1(K;\mathbb{Z}_2) \to C_0(K;\mathbb{Z}_2)$ the space of 1-cycles.
- $B_1(K;\mathbb{Z}_2) \subset Z_1(K;\mathbb{Z}_2)$: the space generated by boundaries of triangles.

First homology of a simp. complex

- Homology = "cycles, which are not boundaries".
- "Number of holes = number of cycles - number of boundaries".
- $Z_1(K;\mathbb{Z}_2) = \operatorname{Ker} \partial_1 \colon C_1(K;\mathbb{Z}_2) \to C_0(K;\mathbb{Z}_2)$ the space of 1-cycles.
- $B_1(K;\mathbb{Z}_2) \subset Z_1(K;\mathbb{Z}_2)$: the space generated by boundaries of triangles.
- $B_1(K;\mathbb{Z}_2) = \operatorname{Im} \partial_2 \colon C_2(K;\mathbb{Z}_2) \to C_1(K;\mathbb{Z}_2)$ where
- $\partial_2(\{v_1, v_2, v_3\}) = \{v_1, v_2\} \oplus \{v_2, v_3\} \oplus \{v_3, v_1\}$ is the boundary of a triangle.
- $H_1(K;\mathbb{Z}_2) = Z_1(K;\mathbb{Z}_2)/B_1(K;\mathbb{Z}_2)$ (the quotient vector space).

Betti number

- $H_1(K;\mathbb{Z}_2) = Z_1(K;\mathbb{Z}_2)/B_1(K;\mathbb{Z}_2)$.
- $\beta_1(K) = \dim H_1(K;\mathbb{Z}_2)$.
- This number counts "the number of 1-dim. holes in K".

- $H_1(K;\mathbb{Z}_2) = Z_1(K;\mathbb{Z}_2)/B_1(K;\mathbb{Z}_2)$.
- $\beta_1(K) = \dim H_1(K;\mathbb{Z}_2)$.
- This number counts "the number of 1-dim. holes in K".

The definition is consistent with graphs. For graphs we have no triangles, hence $C_2(\Gamma;\mathbb{Z}_2) = 0$ hence $B_1(\Gamma;\mathbb{Z}_2) = 0$ hence $H_1(\Gamma;\mathbb{Z}_2) = Z_1(\Gamma;\mathbb{Z}_2)$.

Betti number

- $H_1(K;\mathbb{Z}_2) = Z_1(K;\mathbb{Z}_2)/B_1(K;\mathbb{Z}_2)$.
- $\beta_1(K) = \dim H_1(K;\mathbb{Z}_2)$.
- This number counts "the number of 1-dim. holes in K".

The definition is consistent with graphs. For graphs we have no triangles, hence $C_2(\Gamma;\mathbb{Z}_2) = 0$ hence $B_1(\Gamma;\mathbb{Z}_2) = 0$ hence $H_1(\Gamma;\mathbb{Z}_2) = Z_1(\Gamma;\mathbb{Z}_2)$.

To compute β_1, one needs $\operatorname{rk} D_1$ (the incidence matrix between edges and vertices) and $\operatorname{rk} D_2$, where D_2 is "the incidence matrix" between triangles and edges.

Exercise: $\beta_1(K) = \#\text{edges} - \operatorname{rk} D_1 - \operatorname{rk} D_2$.

General definitions

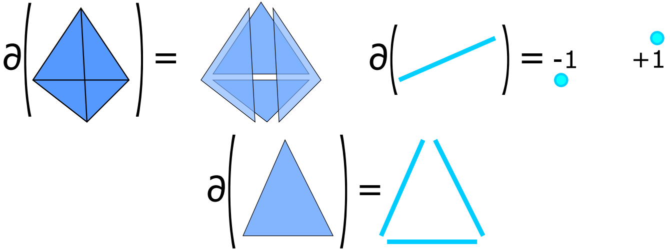
Let K be a simplicial complex, and $\Bbbk$ be a field.

- $C_j(K;\Bbbk)$: the $\Bbbk$-vector space spanned freely by j-dimensional simplices of K.
- $C_j(K;\Bbbk)$ is called the space of j-dim simplicial chains of K.
- $\partial_j\colon C_j(K;\Bbbk) \to C_{j-1}(K;\Bbbk)$: j-th boundary operator, also called simplicial differential.
- $\partial_j(\{v_0, v_1, \dots, v_j\}) =$
  $\{v_1, \dots, v_j\} - \{v_0, v_2, \dots, v_j\} + \cdots + (-1)^j\{v_0, v_1, \dots, v_{j-1}\}$.

Exercise: Prove that $\partial_j \circ \partial_{j+1} = 0$. "There is no boundary of a boundary".

Boundaries

- $\partial_j(\{v_0, v_1, \ldots, v_j\}) =$
  $\{v_1, \ldots, v_j\} - \{v_0, v_2, \ldots, v_j\} + \cdots + (-1)^j\{v_0, v_1, \ldots, v_{j-1}\}.$
- "There is no boundary of a boundary".



General definitions

- $C_j(K;\Bbbk)$: the vector space of j-dim chains.
- $\partial_j\colon C_j(K;\Bbbk) \to C_{j-1}(K;\Bbbk)$: j-th boundary operator.
- $Z_j(K;\Bbbk) = \operatorname{Ker}\partial_j$: the vector space of j-dim cycles.
- $B_j(K;\Bbbk) = \operatorname{Im}\partial_{j+1}$: the vector space of j-dim boundaries.

General definitions

- $C_j(K; \Bbbk)$: the vector space of j-dim chains.
- $\partial_j : C_j(K; \Bbbk) \to C_{j-1}(K; \Bbbk)$: j-th boundary operator.
- $Z_j(K; \Bbbk) = \operatorname{Ker} \partial_j$: the vector space of j-dim cycles.
- $B_j(K; \Bbbk) = \operatorname{Im} \partial_{j+1}$: the vector space of j-dim boundaries.
- Exercise implies $B_j(K; \Bbbk) \subseteq Z_j(K; \Bbbk)$.

Definition

The quotient space $H_j(K; \Bbbk) = Z_j(K; \Bbbk)/B_j(K; \Bbbk)$ is called the ***j*-th simplicial homology module** of K.

Homology

$$\text{Homology: } H_j(K;\Bbbk) = Z_j(K;\Bbbk)/B_j(K;\Bbbk)$$

j-th Betti number

$$\beta_j(K) = \dim H_j(K;\Bbbk) = \dim Z_j(K;\Bbbk) - \dim B_j(K;\Bbbk).$$

- $\beta_j(K)$ counts the number of j-dim holes in K.

Homology

$$\text{Homology: } H_j(K;\Bbbk) = Z_j(K;\Bbbk)/B_j(K;\Bbbk)$$

j-th Betti number

$$\beta_j(K) = \dim H_j(K;\Bbbk) = \dim Z_j(K;\Bbbk) - \dim B_j(K;\Bbbk).$$

- $\beta_j(K)$ counts the number of j-dim holes in K.
- **Exercise 1:** $\beta_0(K)$ equals the number of connected components of K.

Homology

$$\text{Homology: } H_j(K;\Bbbk) = Z_j(K;\Bbbk)/B_j(K;\Bbbk)$$

j-th Betti number

$$\beta_j(K) = \dim H_j(K;\Bbbk) = \dim Z_j(K;\Bbbk) - \dim B_j(K;\Bbbk).$$

- $\beta_j(K)$ counts the number of j-dim holes in K.
- **Exercise 1:** $\beta_0(K)$ equals the number of connected components of K.
- **Exercise 2:** $\beta_j(K) = \#j\text{-dim simplices} - \operatorname{rk} D_j - \operatorname{rk} D_{j+1}$, where D_j is the matrix of ∂_j. This is the incidence matrix between j-dim simplices and $(j-1)$-dim simplices.

Therefore we need $C_{j-1}(K;\Bbbk)$, $C_j(K;\Bbbk)$, $C_{j+1}(K;\Bbbk)$, ∂_j, and ∂_{j+1} to compute $\beta_j(K)$.

Old but instructive meme

Homology is an invariant of homotopy equivalence

- K : a simplicial complex;
- $|K|$: its geometrical realization in some $\mathbb{R}^d$ (a picture).

Fact:

1. Homeomorphism. If $|K| \cong |L|$ then $H_j(K;\Bbbk) \cong H_j(L;\Bbbk)$.
2. Homotopy equivalence. If $|K| \simeq |L|$ then $H_j(K;\Bbbk) \cong H_j(L;\Bbbk)$.

We will not prove it here, but we will be using this fact.

Examples and calculations

Example 0: if pt is the one-point space, then $\beta_0(\mathrm{pt}) = 1$ and $\beta_j(\mathrm{pt}) = 0$ for $j > 0$.

Homology are invariants

If $X \simeq Y$ then $\beta_j(X) = \beta_j(Y)$. Homology do not depend on triangulations of X and Y.

Examples and calculations

Example 0: if pt is the one-point space, then $\beta_0(\text{pt}) = 1$ and $\beta_j(\text{pt}) = 0$ for $j > 0$.

Homology are invariants

If $X \simeq Y$ then $\beta_j(X) = \beta_j(Y)$. Homology do not depend on triangulations of X and Y.

Example 1: $D^n = \{x \in \mathbb{R}^n \mid \|x\| \leqslant 1\}$ is called an n-disk. Exercise: $D^n \simeq \text{pt}$. So far, disk does not have interesting homology.

Examples and calculations

Example 0: if pt is the one-point space, then $\beta_0(\mathrm{pt}) = 1$ and $\beta_j(\mathrm{pt}) = 0$ for $j > 0$.

Homology are invariants

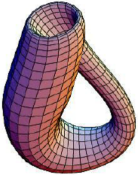
If $X \simeq Y$ then $\beta_j(X) = \beta_j(Y)$. Homology do not depend on triangulations of X and Y.

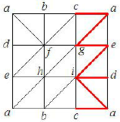
Example 1: $D^n = \{x \in \mathbb{R}^n \mid \|x\| \leqslant 1\}$ is called an n-disk. Exercise: $D^n \simeq \mathrm{pt}$. So far, disk does not have interesting homology.

Example 2: $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$. We have $S^{n-1} = \partial D^n$. Exercise: prove that $\beta_j(S^{n-1}) = 1$ if $j = 0$ or $n-1$, and all other Betti numbers are zero.

Hint: Represent S^{n-1} as the boundary $\partial\Delta^n$ of a simplex. This is a simplicial complex on $n+1$ vertices whose simplices are $2^{[n+1]} \backslash [n+1]$. Understand the relation between chain spaces of $\partial\Delta^n$ and that of Δ^n.

Examples and calculations





Exercise: Triangulate Klein bottle somehow, and compute its homology with coefficients in $\mathbb{Z}_2$, $\mathbb{Z}_3$, and $\mathbb{Q}$.

Exercise: Find $H_j(K;\mathbb{Z}_2)$, where K is the 2-skeleton of 4-dim simplex (simp.comp. on $[5]=\{1,2,3,4,5\}$, which consists of all subsets of cardinality $\leqslant 3$).

Exercise: Consider the simp.comp. U_3, whose vertex set is the set of all nonzero vectors of $\mathbb{Z}_2^3$ (7 vertices in total), and simplices are exactly those subsets, which correspond to linearly independent collections of vectors. Compute $\beta_j(U_3;\mathbb{Z}_2)$.

Künneth formula

Consider the generating function of Betti numbers

$$\mathrm{P}(X;t) = \sum_j \beta_j(X)t^j,$$

called Poincaré polynomials (or Hilbert–Poincaré polynomials) of a space X.

Künneth formula

Consider the generating function of Betti numbers

$$\mathrm{P}(X;t) = \sum_j \beta_j(X)t^j,$$

called Poincaré polynomials (or Hilbert–Poincaré polynomials) of a space X.

Exercise: $\mathrm{P}(X \sqcup Y;t) = \mathrm{P}(X;t) + \mathrm{P}(Y;t)$.

Künneth formula

Consider the generating function of Betti numbers

$$\mathrm{P}(X;t) = \sum_j \beta_j(X)t^j,$$

called Poincaré polynomials (or Hilbert–Poincaré polynomials) of a space X.

Exercise: $\mathrm{P}(X \sqcup Y;t) = \mathrm{P}(X;t) + \mathrm{P}(Y;t)$.

Example: $\mathrm{P}(S^n;t) = 1 + t^n$.

Künneth formula

Consider the generating function of Betti numbers

$$\mathrm{P}(X;t) = \sum_j \beta_j(X)t^j,$$

called Poincaré polynomials (or Hilbert–Poincaré polynomials) of a space X.

Exercise: $\mathrm{P}(X \sqcup Y;t) = \mathrm{P}(X;t) + \mathrm{P}(Y;t)$.

Example: $\mathrm{P}(S^n;t) = 1 + t^n$.

Künneth theorem

$$\mathrm{P}(X \times Y;t) = \mathrm{P}(X;t) \cdot \mathrm{P}(Y;t)$$

Künneth formula

Consider the generating function of Betti numbers

$$\mathrm{P}(X;t) = \sum_j \beta_j(X)t^j,$$

called Poincaré polynomials (or Hilbert–Poincaré polynomials) of a space X.

Exercise: $\mathrm{P}(X \sqcup Y;t) = \mathrm{P}(X;t) + \mathrm{P}(Y;t)$.

Example: $\mathrm{P}(S^n;t) = 1 + t^n$.

Künneth theorem

$$\mathrm{P}(X \times Y;t) = \mathrm{P}(X;t) \cdot \mathrm{P}(Y;t)$$

Example: $T^n = S^1 \times \cdots \times S^1$ is called the *n*-dimensional torus. We have $\mathrm{P}(T^n;t) = (1+t)^n$, so $\beta_j(T^n) = \binom{n}{j}$.

Functoriality

- Let $L \subset K$ be a simplicial subcomplex:
- which means every simplex of L is also a simplex of K.

Functoriality

- Let $L \subset K$ be a simplicial subcomplex:
- which means every simplex of L is also a simplex of K.
- Then every j-cycle in L is also a j-cycle in K that is $Z_j(L;\Bbbk) \subset Z_j(K;\Bbbk)$.
- And every j-boundary in L is also a j-boundary in K that is $B_j(L;\Bbbk) \subset B_j(K;\Bbbk)$.

Functoriality

- Let $L \subset K$ be a simplicial subcomplex:
- which means every simplex of L is also a simplex of K.
- Then every j-cycle in L is also a j-cycle in K that is $Z_j(L;\Bbbk) \subset Z_j(K;\Bbbk)$.
- And every j-boundary in L is also a j-boundary in K that is $B_j(L;\Bbbk) \subset B_j(K;\Bbbk)$.
- We have a linear map

$$\begin{tikzcd} H_j(L;\Bbbk) \arrow[r, "f_{L\subset K}"] \arrow[d, equal] & H_j(K;\Bbbk) \arrow[d, equal] \\ Z_j(L;\Bbbk)/B_j(L;\Bbbk) \arrow[r] & Z_j(K;\Bbbk)/B_j(K;\Bbbk) \end{tikzcd}$$

called **the induced map of inclusion** $L \hookrightarrow K$.

Sources

- H. Edelsbrunner, J. L. Harer, Computational Topology: An Introduction, 2010.
- R. Ghrist, Elementary applied topology, available online.
- A. Hatcher, Algebraic topology, 2002.
- A. J. Zomorodian, Topology for computing, 2005.