Topological Data Analysis

2022–2023

Lecture 2

Simplicial Homology

7 November 2022

Let K be a finite ordered abstract simplicial complex with set of vertices $V = \{v_1, \dots, v_N\}$. As usual, we write $(i_0 \cdots i_n)$ instead of $\{v_{i_0}, \dots, v_{i_n}\}$ with the convention that $i_0 < \dots < i_n$.

For $n \geq 0$, we let $C_n(K)$ be the free abelian group on the set of n-faces of K. Elements of $C_n(K)$ are called n-chains in K.

For example, $5(012) + 3(014) - (134)$ is a 2-chain.

We have $C_n(K) = 0$ if $n < 0$ or, more generally, if the set of n-faces of K is empty, since the free abelian group on $\emptyset$ is the trivial group.

New convention:

If $i_0, \dots, i_n$ are not in order, then we define

$$(i_0 \cdots i_n) = \varepsilon(\sigma)\,(i_{\sigma(0)} \cdots i_{\sigma(n)}) \quad \text{where } i_{\sigma(0)} < \dots < i_{\sigma(n)}$$

and $\varepsilon(\sigma) = 1$ if σ is an even permutation while $\varepsilon(\sigma) = -1$ if σ is odd.

For example, $(021) = -(012)$ and $(21043) = (01234)$.

More generally, if R is any commutative ring with 1, we denote by $C_n(K;R)$ the free R-module on the set of n-faces of K.

The default case is $R = \mathbb{Z}$, where we write $C_n(K)$ instead of $C_n(K;\mathbb{Z})$. Note that $\mathbb{Z}$-modules are precisely abelian groups.

If R is a field, then an R-module is a vector space over R. Frequent choices are $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{F}_2$ (the field with 2 elements).

We call R the <u>ring of coefficients</u>. Although we will state the next facts for $R = \mathbb{Z}$ for shortness of notation, everything generalizes to any R, unless explicitly commented.

Boundary operator

The <u>nth</u> <u>boundary</u> is the group homomorphism (or R-module homomorphism if R coefficients are used)

$$\partial_n : C_n(K) \longrightarrow C_{n-1}(K)$$

defined on generators as

$$\partial_n (i_0 \dots i_n) = \sum_{k=0}^{n} (-1)^k (i_0 \dots \hat{i}_k \dots i_n)$$

where $\hat{i}_k$ means that the k^{th} entry is deleted.

The fundamental property of boundary operators is that

$$\boxed{\partial_n \circ \partial_{n+1} = 0} \quad \text{for all } n.$$

The proof is a formalization of the next example:

$$\partial_2(\partial_3(abcd)) = \partial_2\left[(bcd) - (acd) + (abd) - (abc)\right] =$$

$$= \partial_2(bcd) - \partial_2(acd) + \partial_2(abd) - \partial_2(abc) =$$

$$= \cancel{(cd)} - \cancel{(bd)} + \cancel{(bc)} - \cancel{(cd)} + \cancel{(ad)} - \cancel{(ac)} + \cancel{(bd)} -$$

$$- \cancel{(ad)} + \cancel{(ab)} - \cancel{(bc)} + \cancel{(ac)} - \cancel{(ab)} = 0 \quad \checkmark$$

$$\partial_n(\partial_{n+1}(i_0 \dots i_{n+1})) = \partial_n\left(\sum_{k=0}^{n+1} (-1)^k (i_0 \dots \hat{i}_k \dots i_{n+1})\right) = \sum_{k=0}^{n+1} (-1)^k \partial_n (i_0 \dots \hat{i}_k \dots i_{n+1}) =$$

$$= \sum_{k=1}^{n+1} (-1)^k \sum_{\ell=0}^{k-1} (-1)^\ell (i_0 \dots \hat{i}_\ell \dots \hat{i}_k \dots i_{n+1}) + \sum_{k=0}^{n+1} (-1)^k \sum_{\ell=k}^{n} (-1)^\ell (i_0 \dots \hat{i}_k \dots \hat{i}_{\ell+1} \dots i_{n+1})$$

$$= \sum_{k=1}^{n+1} \sum_{\ell=0}^{k-1} (-1)^{k+\ell} (i_0 \dots \hat{i}_\ell \dots \hat{i}_k \dots i_{n+1}) + \sum_{\ell=0}^{n} \sum_{k=0}^{\ell} (-1)^{k+\ell} (i_0 \dots \hat{i}_k \dots \hat{i}_{\ell+1} \dots i_{n+1})$$

$$= \sum_{k=1}^{n+1} \sum_{\ell=0}^{k-1} (-1)^{k+\ell} (i_0 \dots \hat{i}_\ell \dots \hat{i}_k \dots i_{n+1}) + \sum_{r=1}^{n+1} \sum_{k=0}^{r-1} (-1)^{k+r-1} (i_0 \dots \hat{i}_k \dots \hat{i}_r \dots i_{n+1})$$

(Change $\ell+1=r$)

$$= 0 \quad ✓$$

Homology

Note that $\partial_n \circ \partial_{n+1} = 0 \implies \operatorname{Im} \partial_{n+1} \subseteq \operatorname{Ker} \partial_n$.

$$\begin{tikzcd} C_{n+1}(K) \arrow[r, "\partial_{n+1}"] \arrow[rr, bend right, "0"'] & C_n(K) \arrow[r, "\partial_n"] & C_{n-1}(K) \end{tikzcd}$$

We denote $Z_n(K) = \operatorname{Ker} \partial_n$, and call its elements <u>n-cycles</u>;

$B_n(K) = \operatorname{Im} \partial_{n+1}$, and call its elements <u>n-boundaries</u>.

Thus $B_n(K) \subseteq Z_n(K)$ for all n, and we define

$$\boxed{H_n(K) = Z_n(K) \big/ B_n(K)}$$

nth homology of K.

If coefficients in a ring R are used, then $Z_n(K;R)$ is an R-submodule of $C_n(K;R)$ and $B_n(K;R)$ is an R-submodule of $Z_n(K;R)$. Hence $H_n(K;R)$ acquires an R-module structure.

- If R is a field, then $H_n(K;R)$ is an R-vector space of finite dimension for all n, and it is zero beyond the dimension of K.
- If $R = \mathbb{Z}$, then $Z_n(K)$ and $B_n(K)$ are free abelian groups, since $C_n(K)$ is free and every subgroup of a free abelian group is free. However, $H_n(K)$ can have torsion. Moreover, $Z_n(K)$ is finitely generated and therefore $H_n(K)$ is finitely generated as well. Hence

$$H_n(K) \cong \mathbb{Z}^r \oplus \mathbb{Z}/p_1^{\alpha_1} \oplus \dots \oplus \mathbb{Z}/p_m^{\alpha_m}$$

for some primes $p_1, \dots, p_m$ and $\alpha_i \geq 1$. Moreover,

$$r = \operatorname{rank} H_n(K) = \dim_{\mathbb{Q}} H_n(K;\mathbb{Q}).$$

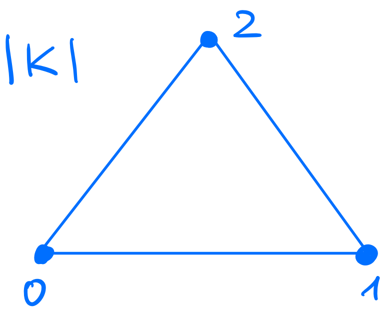
- For an arbitrary ring R, the R-module $Z_n(K;R)$ need neither be free nor finitely generated. However, if R is a <u>principal ideal domain</u> (PID) then $Z_n(K;R)$ is finitely generated and free, and hence $H_n(K;R)$ is also finitely generated. The structure theorem for finitely generated modules over a PID implies that

$$H_n(K;R) \cong R/(d_1) \oplus \dots \oplus R/(d_m)$$

for some ideals (d_i) of R with $(d_{i+1}) \subseteq (d_i)$ for all i. The free part corresponds to those indices with $d_i = 0$. If R is a <u>field</u>, then $d_i = 0$ for all i, since all R-modules are free.

In general, the ideals (d_i) are uniquely determined by $H_n(K;R)$, but the isomorphism is not unique.

Example: Let K have maximal faces $(01), (02), (12)$.



$$C_0(K) = \mathbb{Z}(0) \oplus \mathbb{Z}(1) \oplus \mathbb{Z}(2)$$

$$C_1(K) = \mathbb{Z}(01) \oplus \mathbb{Z}(02) \oplus \mathbb{Z}(12)$$

$$0 \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} 0$$

Matrix of ∂_1:

$$\begin{array}{c|ccc|c} & (01) & (02) & (12) & (12)-(02)+(01) \\ \hline (0) & -1 & -1 & 0 & 0 \\ (1) & 1 & 0 & -1 & 0 \\ (2) & 0 & 1 & 1 & 0 \end{array}$$

$$\partial_1(01) = (1) - (0)$$
$$\partial_1(02) = (2) - (0)$$
$$\partial_1(12) = (2) - (1)$$

By column-reduction we find that $\operatorname{rank} \partial_1 = 2$ and $\operatorname{Ker} \partial_1$ is generated by $(12)-(02)+(01)$. Hence

$$\left.\begin{array}{l} \operatorname{Ker} \partial_0 = C_0(K) = \langle (0), (1), (2) \rangle \\ \operatorname{Im} \partial_1 = \langle (1)-(0), (2)-(0) \rangle \end{array}\right\} \quad H_0(K) = \frac{\operatorname{Ker}\partial_0}{\operatorname{Im}\partial_1} \cong \mathbb{Z}$$

since $[0]=[1]=[2]$

$$(2)-(1) = (2)-(0)-((1)-(0))$$

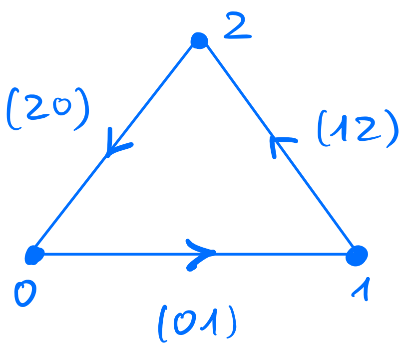
generated by the class $[0]$ of (0).

$$\left.\begin{array}{l} \operatorname{Ker} \partial_1 = \langle (12) - (02) + (01) \rangle \\ \operatorname{Im} \partial_2 = 0 \end{array}\right\} \quad H_1(K) = \frac{\operatorname{Ker} \partial_1}{\operatorname{Im} \partial_2} \cong \mathbb{Z}$$

generated by the class of the 1-cycle

$$z = (12) - (02) + (01).$$

A 1-cycle can be viewed geometrically as a closed edge path:



Note that $(20) = -(02)$