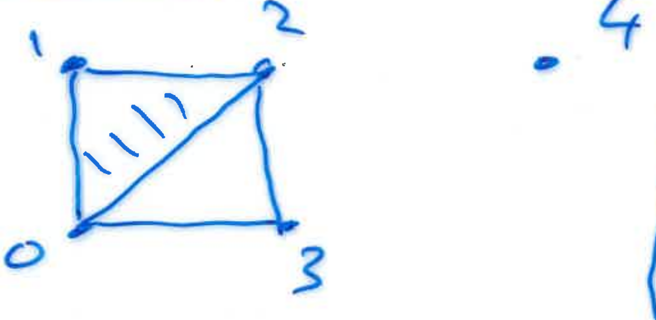
Recall

K a simplicial complex

Example



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

$K = \{\{0\},\{1\},\{2\},\{3\},\{4\},$
$\{0,1\},\{0,2\},$
$\{0,3\},\{1,2\},$
$\{2,3\},$
$\{0,1,2\}\}$

$C_n K$ a vector space (over say $\mathbb{R}$) with one basis element e_σ for each n-simplex $\sigma = \{v_0, v_1, \dots, v_n\} \in K$.

$d_n : C_n K \to C_{n-1} K$ is a linear hom$^{\text{sm}}$ given by

$$d_n(e_\sigma) = \sum_{i=0}^{n} (-1)^i e_{\sigma \setminus \{v_i\}}$$

Defn The degree n Betti number of K is

$$\beta_n = \mathrm{Dim}(\mathrm{Ker}\, d_n) - \mathrm{Dim}(\mathrm{Image}\, d_{n+1})$$

It is convenient to view β_n as the dimension of a certain vector space.

Lemma

$$d_n(d_{n+1} e_\sigma) = 0$$

for all $n+1$-simplices σ

Proof exercise

Thus

$$\mathrm{Im}\, d_{n+1} \subseteq \mathrm{Ker}\, d_n$$

Defn The degree n homology of K is the vector space

$$H_n(K) = \frac{\ker d_n}{\operatorname{im} d_{n+1}}.$$

From this definition we see

$$\boxed{\beta_n = \dim(H_n(K))}$$

Observation: If L is a subcomplex of the simplicial complex K (i.e. L is a simplicial complex with $L \subset K$) then

$C_n L$ is a sub vector space of $C_n K$.

In fact we have a diagram

$$\begin{tikzcd}
\vdots \arrow[d] & \vdots \arrow[d] \\
C_3 L \arrow[r, hook] \arrow[d, "d_3"] & C_3 K \arrow[d, "d_3"] \\
C_2 L \arrow[r, hook] \arrow[d, "d_2"] & C_2 K \arrow[d, "d_2"] \\
C_1 L \arrow[r, hook] \arrow[d, "d_1"] & C_1 K \arrow[d, "d_1"] \\
C_0 L \arrow[r, hook] & C_0 K
\end{tikzcd}$$

This diagram induces a hom$^{\text{sm}}$ of vector spaces

$$H_n(L) \longrightarrow H_n(K)$$

which is <u>not</u> in general injective.

Data Analysis

Suppose given a distance matrix $D = (d_{ij})$ of distances between n items. Suppose we have chosen a sequence of thresholds

$$\varepsilon_1 < \varepsilon_2 < \varepsilon_3 < \cdots < \varepsilon_t$$

we then get a sequence of clique simplicial complexes

$$K_{\varepsilon_1} \subseteq K_{\varepsilon_2} \subseteq K_{\varepsilon} \subseteq \cdots \subseteq K_t$$

with K_{ε_i} a subcomplex of $K_{\varepsilon_{i+1}}$.

This yields a sequence of linear homomorphisms

$$H_n(K_{\varepsilon_1}) \to H_n(K_{\varepsilon_2}) \to \cdots \to H_n(K_{\varepsilon_t}).$$

In data analysis this sequence of homomorphisms is represented using bar codes.

In this module we are interested only in:

$n = 0$ (connected components)

$n = 1$ (1-dimensional holes)