Pr. MAMOUNI
CRMEF Rabat

Topological Data Analysis
TDA (Overview)

ENSIAS Rabat,

Content

1) History (Overview)
2) Applications in Medecine and Social Science (Overview)
3) Persistent Homology Theory (Overview)
4) Example of research in Biology

1) History :

Born :

Growth :

Explosure :

Research groups :

USA
Germany - Austria
UK
Italy
France
Japan - Korea
Morocco

2) See Slides

3) TDA (overview theoritical)

3.1) ~~Simpli~~ Homology ~~groups~~ theory:

→ **chain-complex**

$$\cdots \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \xrightarrow{d_{n-1}} C_{n-2} \rightarrow \cdots C_{-1} = 0$$

where C_n are $\mathbb{Z}$-modules
(or abelian groups)

st $\boxed{d_{n-1} \circ d_n = 0}$

ie $\boxed{\mathrm{Im}\, d_n \subset \mathrm{Ker}\, d_{n-1}}$

~~We put~~ $C = \oplus C_n$ et $d : C \to C$
$d/C_n = d_n$

→ **cycles** : are elements of Ker d
ie $d\sigma = 0$

→ **bords** : are elements of Im d
ie $\tau = d\sigma$

since $d^2 = 0$ then
bord ⇒ cycle
~~⇐~~

→ **Homology groups**
these are $H_n(C, d) = \mathrm{Ker}\, d_{n-1} / \mathrm{Im}\, d_n$

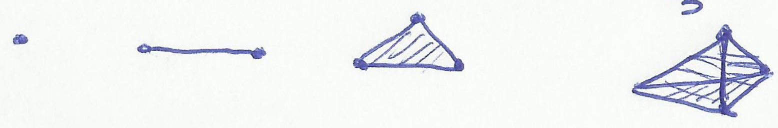
ie $[\sigma] \in H_n(C, d)$
means $d\sigma = 0$
and $[\sigma] = \sigma + \mathrm{Im}\, d_n$

otherwise $[\sigma_1] = [\sigma_2] \Leftrightarrow \sigma_1 = \sigma_2 + \gamma$
bord

3.2) Simplicial Homology

→ **simplices** in $\mathbb{R}^3$
these are $\sigma^n = \mathrm{Aff}(e_0, \ldots, e_n)$ convexe
where $e_0 = (0, 0, 0)$
$e_1 = (1, 0, 0)$
$e_2 = (0, 1, 0)$
$e_3 = (0, 0, 1)$

e_0 e_1 e_2 e_3



→ boundary operator

if $\sigma = [e_0, \dots, e_n]$ (denotation)

we set $\boxed{d\sigma^n = \sum_{i=0}^{n} (-1)^i [e_0, \hat{e_i}, e_n]}$

forget it

we check that $\boxed{d^2 = 0}$

Example:

$d(\cdot) = 0$

$d(\overrightarrow{AB}) = B - A$

$d(\triangle ABC) = AB + BC + CA$

$d^2(\overline{AB}) = dB - dA = 0$

$d^2(\triangle) = d(\overrightarrow{AB}) + d(\overrightarrow{BC}) + d(\overrightarrow{CA})$

$= (B - A) + (C - B) + (A - C)$

$= 0$

3.3) Persistent homology

3.3.1) Rips – Vietoris – Cech – Delaunay Complexes

let $\mathcal{C}$ = cloud of points $(A_k)_{k \in \mathbb{N}}$

→ $C_0 = \mathcal{C}$

→ We consider balls $B_k(\varepsilon) = B(A_k, \varepsilon)$

where ε_p increase

when two balls $B_k(\varepsilon) \cap B_p(\varepsilon) \neq \emptyset$

we set $C_1 = C_0 \cup \overline{A_k A_p}$

We get $C_0 \subset C_1$

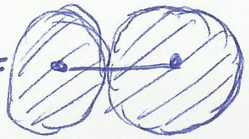
and $C_1 \xrightarrow{d_1} C_0$

that is the first construction of what we call filtration

3.3.2) Constructive examples

Example 1 $\mathcal{E} = \{A, B\}$ $d(AB) = 1$

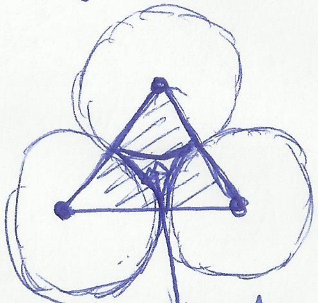
$C_0 = \cdot \quad \cdot$

$C_1 =$  when $\varepsilon = \frac{1}{2}$

Example 2 $\mathcal{E} = \{A, B, C\}$

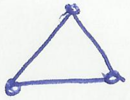
$d(A,B) = 1$
$d(A,C) = 1$
$d(B,C) = 1$

$C_0 =$ 

$C_1 =$ 

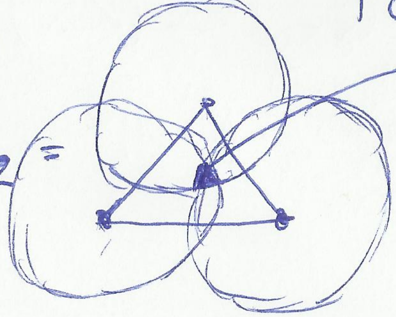
empty

$\varepsilon = \frac{1}{2}$

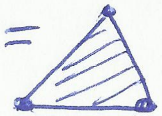
$C_1 =$ 

View video

full

$C_2 =$ 

$\varepsilon = \frac{\sqrt{2}}{2}$

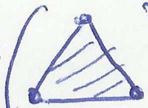
$C_2 =$ 

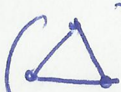
3.3.3) Cycle and bord

$d(\cdot) = 0$ means that $\cdot$ cycle is born

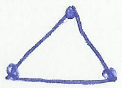
$d(\text{—}) = \{\cdot \ \cdot\}$ was two cycle become a bord its homolog class is null

we say that $\cdot \ \cdot$ is killed when — is born

$d($$) =$ 

$d($$) = d^2($$) = 0$

cycle born

the same cycle is killed when  is born

3.3.4) Persistent Betti Number

β_n^{ε} = Nbr of homotopical class of n-simplics that still on live at time ε

Example 1 : $\mathcal{E} = \{A, B\}$

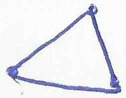
$0 < \varepsilon < \frac{1}{2}$  two 0-cycles born

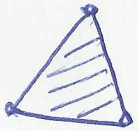
$\varepsilon = 1$  the two cycle killed

Bar Codes

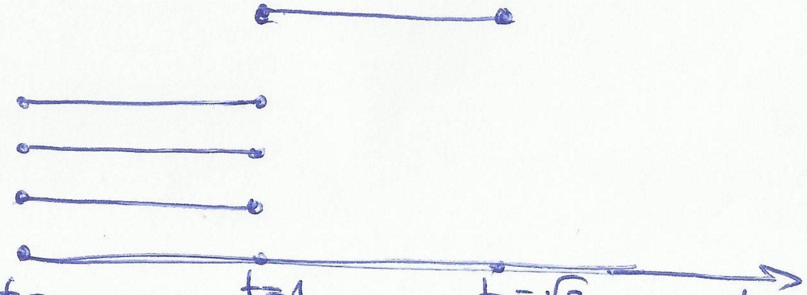
Example 2 $\mathcal{E} =$ 

$C_0 =$  3 0-cycles $0 \leq \varepsilon < \frac{1}{2}$

$C_1 =$  $\varepsilon = \frac{1}{2}$ the 0-cycles are killed and the  1-cycle is born

$C_2 =$  $\varepsilon = \frac{\sqrt{2}}{2}$ the 1-cycle  is killed by 

Barcodes



3.3.5) Boundary matrix

Example 1 $\mathcal{C} = \{ \overset{A}{\cdot} \quad \overset{B}{\cdot} \}$

the complexe will be

$\{ \underset{\sigma_1}{A}, \underset{\sigma_2}{B}, \underset{\sigma_3}{[AB]} \}$ $\qquad \partial\sigma_3 = \sigma_2 - \sigma_1$

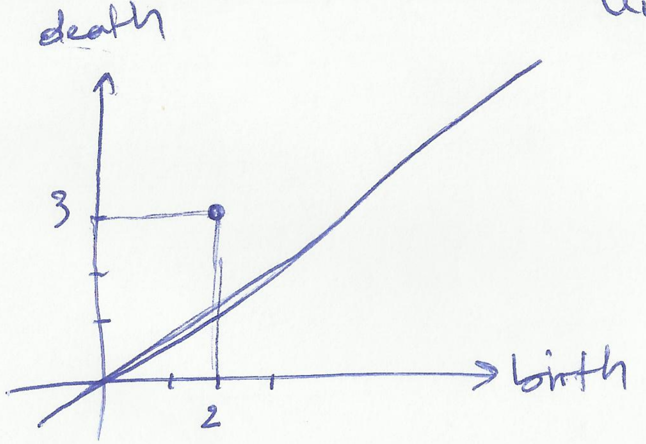
$\partial\sigma_1 = \partial\sigma_2 = 0$

$$\begin{array}{c} \\ \text{birth} \end{array} \quad \partial = \begin{array}{c|ccc} & \partial\sigma_1 & \partial\sigma_2 & \partial\sigma_3 \\ \hline \sigma_1 & 0 & 0 & 1 \\ \sigma_2 & 0 & 0 & 1 \\ \sigma_3 & 0 & 0 & 0 \end{array} \leftarrow \text{low coeff}$$

death

the matrix is reduced

since there is no lows in the same level (the same low line).



Persistent diagram

Example 2 $\mathcal{C} = \{ \underset{A}{\cdot} \quad \overset{\cdot C}{\underset{B}{\cdot}} \}$

the complex will be

$\{ \underset{1}{A}, \underset{2}{B}, \underset{3}{C}, \underset{4}{AB}, \underset{5}{AC}, \underset{6}{BC}, \underset{7}{ABC} \}$

$\partial 1 = 0, \ \partial 2 = 0, \ \partial 3 = 0, \ \partial 4 = 2 - 1$

$\partial 5 = 3 - 2, \ \partial 6 = 3 - 1, \ \partial 7 = 4 + 5 + 6$

the boundary matrix with $\mathbb{Z}/2\mathbb{Z}$ coeff is

$\partial =$

	1	2	3	4	5	6	7
1	0	0	0	1		1	
2				1	1		
3					1	1	
4							1
5							1
6							1
7							

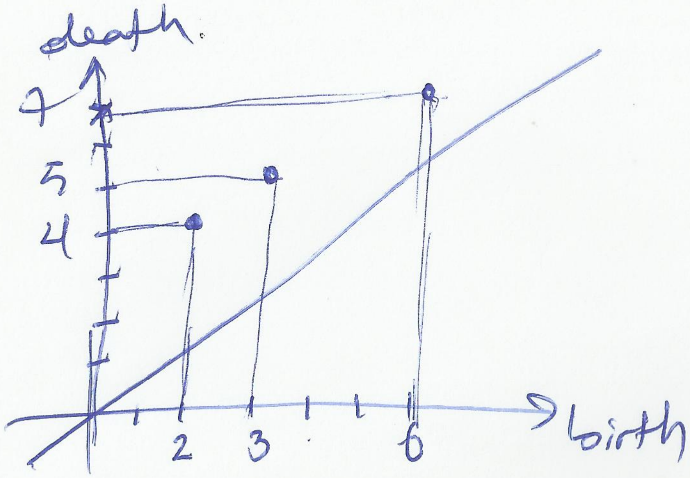
low

two double lows

low

$C_6 \leftarrow C_6 + C_5 + C_4$ the reduced boundary matrix is

	1	2	3	4	5	6	7
1				1			
2				1	1		
3					1		
4							1
5							1
6							1
7							



4) - Research in Biology

view Model 1
view Model 2
view Edel - Haver end book

EnD