THE STRUCTURE OF PERSISTENCE

(AN INTROSPECTION)

Ulrich Bauer (TUM)

TopAPP Workshop 24-26 April 2019

IST AUSTRIA

PART 2:

2-PARAMETER PERSISTENCE

(CLUSTERS & INDECOMPOSABLES)

joint work with:

Magnus Botnan / Steffen Oppermann / Johan Steen

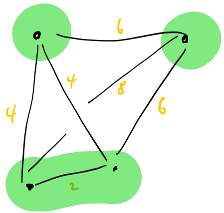
CLUSTERING FUNCTIONS

X : finite set

Clustering function φ :

maps a metric $d : X \times X \to \mathbb{R}$ (distance matrix)

to a partition of X.

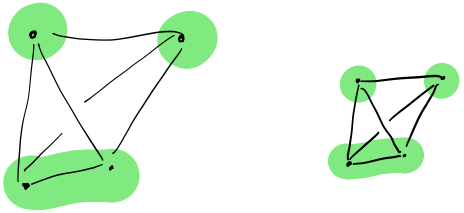


KLEINBERG'S AXIOMS

Desirable properties

- scale invariance :

$$\varphi(d) = \varphi(\lambda \cdot d).$$

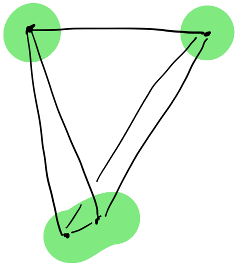


- richness :

every partition is obtained from some d.

- consistency :

decreasing d within clusters /
increasing d across clusters
does not change the result.



KLEINBERG'S IMPOSSIBILITY THEOREM

Thm [Kleinberg 2002] No clustering function satisfies

- scale-invariant,
- richness, and
- consistency.

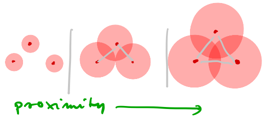
Motivates the use of a scale parameter

⇝ hierarchical clustering

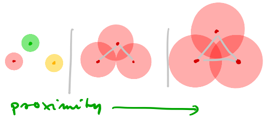
CLUSTERING FROM CONNECTED COMPONENTS

proximity graph

- filter edges by proximity



π_0 (connected components)



single-linkage clustering

H_0 (homology in deg. 0 with coeffs in K)

$$K^3 \to K \to K$$

proximity $\longrightarrow$

persistent homology

$H_0 = F \circ \pi_0$

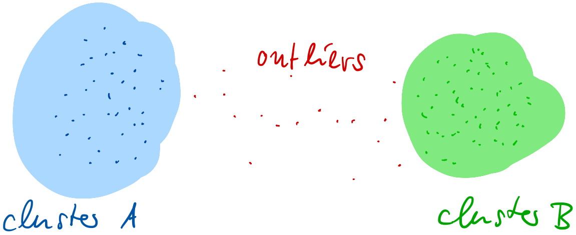
HIERARCHICAL CLUSTERING: EXISTENCE & UNIQUENESS

Thm [Carlsson, Mémoli 2010] Single-linkage clustering is the unique hierarchical clustering method satisfying [... certain axioms similar to Kleinberg's].

But ...

CHAINING EFFECT

Single-linkage clustering is sensitive to outliers

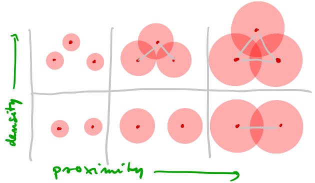


→ not used much in practice!

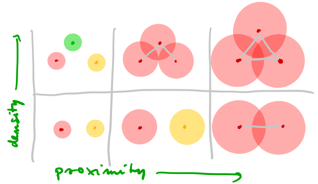
2-PARAMETER CLUSTERING

density-proximity graph

- filter points by density
- filter edges by proximity



π_0 (connected components)



H_0 (homology in deg. 0
with coeffs in K)

$$\begin{tikzcd} K^3 \arrow[r] & K \arrow[r] & K \\ K^2 \arrow[r] \arrow[u] & K^2 \arrow[r] \arrow[u] & K \arrow[u] \end{tikzcd}$$

density

proximity

DECOMPOSING DIAGRAMS OF VECTOR SPACES

$$V \xrightarrow{f} W$$

$$\cong$$

$$\begin{tikzcd}
\ker f \arrow[r] & 0 \\
\oplus & \\
\operatorname{im} f \arrow[r] & \operatorname{im} f \\
\oplus & \\
0 \arrow[r] & \operatorname{coker} f
\end{tikzcd}$$

TWO MAPS

$$
\begin{tikzcd}
U \arrow[r, "g"] & V \arrow[r, "f"] & W \\[-1em]
\cong & & \\[-1em]
K^a \arrow[r] & 0 \arrow[r] & 0 \\[-1em]
\oplus & & \\[-1em]
0 \arrow[r] & K^b \arrow[r] & 0 \\[-1em]
\oplus & & \\[-1em]
0 \arrow[r] & 0 \arrow[r] & K^c \\[-1em]
\oplus & & \\[-1em]
K^d \arrow[r] & K^d \arrow[r] & 0 \\[-1em]
\oplus & & \\[-1em]
0 \arrow[r] & K^e \arrow[r] & K^e \\[-1em]
\oplus & & \\[-1em]
K^f \arrow[r] & K^f \arrow[r] & K^f
\end{tikzcd}
$$

SEQUENCES OF MAPS

$$V: \quad V_0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow \cdots \longrightarrow V_n$$

decomposes into summands of the form

$$\cdots \longrightarrow 0 \longrightarrow \underbrace{K \longrightarrow \cdots \longrightarrow K}_{i \qquad\qquad j} \longrightarrow 0 \longrightarrow \cdots$$

$V \cong$ "collection of intervals"

$\rightsquigarrow$ persistence barcode.

DROZD'S TRICHOTOMY

Given a finite indexing poset P. (𝕜 : algebr. closed)

What are the indecomposable diagrams with shape P?

3 cases (representation types):

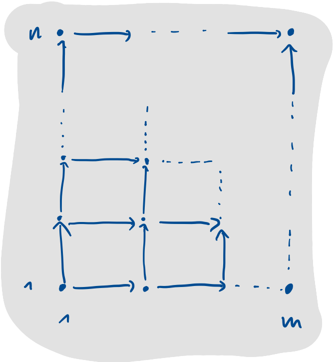
ⓐ A finite list. finite type

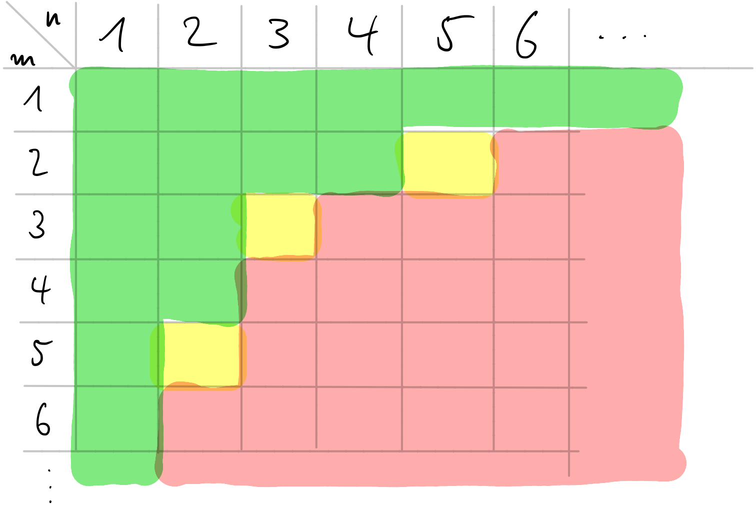
ⓑ A finite list (of 1-param. families). tame

ⓒ It's complicated. wild

(as complicated as modules over any finite-dim. algebra; including undecidable problems)

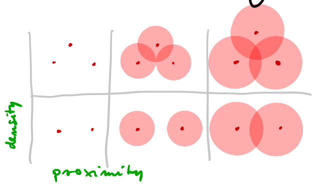
REPRESENTATION TYPES OF COMMUTATIVE GRIDS



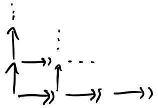


GRID DIAGRAMS FROM CLUSTERING

Consider again 2-parameter clustering (proximity / density)



This yields diagrams of the form



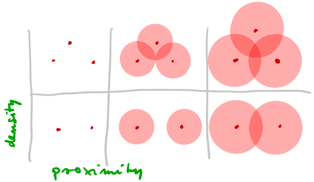
Horizontal maps are surjective!

Does this simplify the picture?

EPIMORPHISMS

- $\mathrm{Rep}(m,n)$: all commutative dgms over $m \times n$ grid
- $\mathrm{Rep}^{\twoheadrightarrow}(m,n)$: epis in horizontal direction
- $\mathrm{Rep}^{\uparrow\twoheadrightarrow}(m,n)$: epis in both directions.

Lemma $\mathrm{Rep}^{\twoheadrightarrow}(m,2)$ is finite type.



$$H_0 \rightsquigarrow \begin{tikzcd} K^3 \arrow[r, two heads, "(1\,1\,1)"] & K \arrow[r, two heads] & K \\ K^2 \arrow[u, "\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}"] \arrow[r, two heads] & K^2 \arrow[u, "(1\,1)"'] \arrow[r, two heads, "(1\,1)"'] & K \arrow[u] \end{tikzcd}$$

$$\cong \begin{tikzcd} K \arrow[r, two heads] & K \arrow[r, two heads] & K \\ K \arrow[u] \arrow[r, two heads] & K \arrow[u] \arrow[r, two heads] & K \arrow[u] \end{tikzcd} \oplus \begin{tikzcd} K \arrow[r, two heads] & 0 \arrow[r, two heads] & 0 \\ K \arrow[u] \arrow[r, two heads] & K \arrow[u] \arrow[r, two heads] & 0 \arrow[u] \end{tikzcd} \oplus \begin{tikzcd} K \arrow[r, two heads] & 0 \arrow[r, two heads] & 0 \\ 0 \arrow[u] \arrow[r, two heads] & 0 \arrow[u] \arrow[r, two heads] & 0 \arrow[u] \end{tikzcd}$$

EPIC GRIDS & WILD THINGS

Thm [B, Botnan, Oppermann, Steen]

$$\begin{array}{ccc} \mathrm{Rep}^{\uparrow \twoheadrightarrow}(m,n) & \sim & \mathrm{Rep}^{\uparrow}(m,n-1) \\ \wr & & \wr \\ \mathrm{Rep}^{\twoheadrightarrow}(m-1,n) & \sim & \mathrm{Rep}(m-1,n-1) \end{array}$$

same representation type

Corollary $\mathrm{Rep}^{\twoheadrightarrow}(m,n)$ is

- finite type for $n \leq 2$ and $(n=3,\ m \leq 4)$
- tame for $(n=3,\ m=5)$ and $(n=4,\ m=3)$
- wild otherwise.

TORSION PAIRS

Example: Abelian groups (fin. generated)

$$G = \underbrace{\mathbb{Z}}_{F \text{ (free)}} \oplus \underbrace{\mathbb{Z}_{q_1} \oplus \cdots \oplus \mathbb{Z}_{q_n}}_{T \text{ (torsion)}}$$

- canonically: $T \hookrightarrow G \twoheadrightarrow F$ (short exact sequence)
- The only homomorphism $\phi : T \to F$ is $x \mapsto 0$.

Category: Ab. Subcategories: $\mathcal{T}$ (torsion groups) $\mathcal{F}$ (free Abelian groups)

Then $(\mathcal{T}, \mathcal{F})$ is a torsion pair:

- $\mathrm{Hom}(\mathcal{T}, \mathcal{F}) = 0$
- For any $G : Ab$, there is a short ex. seq $T \hookrightarrow G \twoheadrightarrow F$ $(T : \mathcal{T},\ F : \mathcal{F})$

COTORSION PAIRS

$\mathcal{A}$: Abelian category

$\mathcal{C}, \mathcal{D}$: subcategories (closed under direct summands)

$(\mathcal{C}, \mathcal{D})$ is cotorsion pair if

- $\mathrm{Ext}^1(\mathcal{C}, \mathcal{D}) = 0$ (any short exact $D \hookrightarrow E \twoheadrightarrow C$ split),
- For any $A \in \mathcal{A}$, there are short exact sequences

$D \hookrightarrow C \twoheadrightarrow A$ and $A \hookrightarrow \tilde{D} \twoheadrightarrow \tilde{C}$ $(D, \tilde{D} \in \mathcal{D};\ C, \tilde{C} \in \mathcal{C})$.

COTORSION TORSION TRIPLES

Theorem $(\mathcal{T},\mathcal{F})$ torsion pair, $(\mathcal{C},\mathcal{T})$ cotorsion pair.
Then $\mathcal{F}$ is equivalent to $\dfrac{\mathcal{C}}{\mathcal{C}\cap\mathcal{T}}$.

(We call $(\mathcal{C},\mathcal{T},\mathcal{F})$ a cotorsion torsion triple.)

"tilting subcategory"; already determines $(\mathcal{C},\mathcal{T},\mathcal{F})$

APPLICATION TO GRID REPRESENTATIONS

Corollary

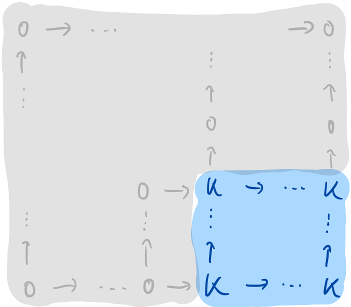
$$\frac{\mathrm{Rep}^{\twoheadrightarrow}(m,n)}{\mathrm{Rep}^{\hat{\imath}\twoheadrightarrow}(m,n)} \simeq \mathrm{Rep}(m,n-1)$$

$\mathcal{C}$

$\mathcal{C}\cap\mathcal{J}$

$\mathcal{F}$

indecomposables are of the form



$\Rightarrow$ finite type

THE EQUIVALENCE, MADE EXPLICIT

$$
\begin{array}{ccc}
\boxed{V_{n1} \twoheadrightarrow \cdots \twoheadrightarrow V_{nm}} & = & I_n \\
 & & \uparrow f_{n-1} \\
\vdots \qquad \vdots \qquad \vdots & & I_{n-1} \\
\uparrow \qquad \uparrow \qquad \uparrow & & \vdots \\
 & & \uparrow f_2 \\
\boxed{V_{21} \twoheadrightarrow \cdots \twoheadrightarrow V_{2m}} & = & I_2 \\
\uparrow \qquad \uparrow \qquad \uparrow & & \uparrow f_1 \\
\boxed{V_{11} \twoheadrightarrow \cdots \twoheadrightarrow V_{1m}} & = & I_1
\end{array}
\quad \rightsquigarrow \quad
\begin{array}{c}
0 \\
\uparrow \\
\ker(f_{n-1}) \\
\vdots \\
\uparrow \\
\ker(f_{n-1} \circ \cdots \circ f_2) \\
\uparrow \\
\ker(f_{n-1} \circ \cdots \circ f_2 \circ f_1)
\end{array}
$$

REALIZATIONS

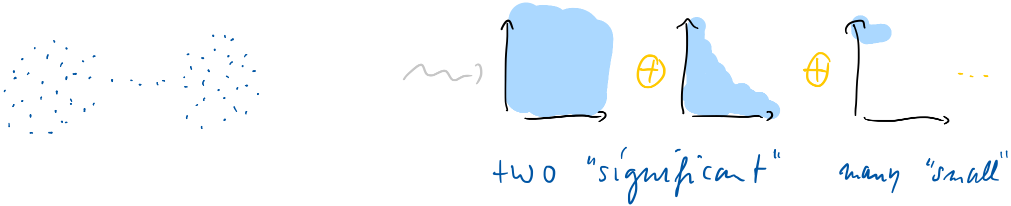
Thm [Carlsson, Zomorodian] Any Rep(m,n) (m×n diagram of $\mathbb{Z}_2$- vector spaces) can be realized as pth-homology (H_p of an m×n diagram of top. spaces), for any $p > 0$.

What about H_0?

Thm [BBOS]
Not every $\mathrm{Rep}^{\uparrow}_{\rightarrow}(m,n)$ can be realized as H_0 (counterexample), not even as a summand.

INDECOMPOSABLES VS CLUSTERS

- Indecomposables of H_0 from density/proximity do not correspond directly to clusters
- Rather: "linear combinations of components"
- Topological features, help to survey parameter space
- Typical example:



FURTHER READING

- Barcodes (1 parameter) arXiv 1610.10085
- Clusters (2 parameters) arXiv 1904.07322