

パーシステントホモロジー入門

青木利隆

神戸大学大学院人間発達環境学研究科

20/8/2024

氏名: 青木利隆 (Toshitaka Aoki)

専門: 多元環の表現論・傾理論 (Tilting/Silting theory)

加群圏 $\text{mod } A$. 導来圏 $D^b(\text{mod } A)$.

現在: 2023年2月より神戸大学(エスカラ先生の研究室)に所属.

パーシステントホモロジー解析の表現論的な観点の研究に従事.

主にマルチパラメータ化に向けた理論構築



データ記述科学

データの「かたち」と「うごき」に着目した記述子(=数学言語)の開発

▼ 位相的データ解析 (TDA=Topological Data Analysis)

● パーシステントホモロジー解析 (2000年頃～)

データの「かたち」=> "穴" の持続性(=パーシステンス)に着目
情報を **パーシステンス図** として可視化

応用領域

- ・ 材料科学
- ・ 進化生物学
- ・ 画像解析 等々

数学領域

- ・ 代数/計算トポロジー
- ・ 多元環の表現論
- ・ 確率論
- ・ 機械学習 等々

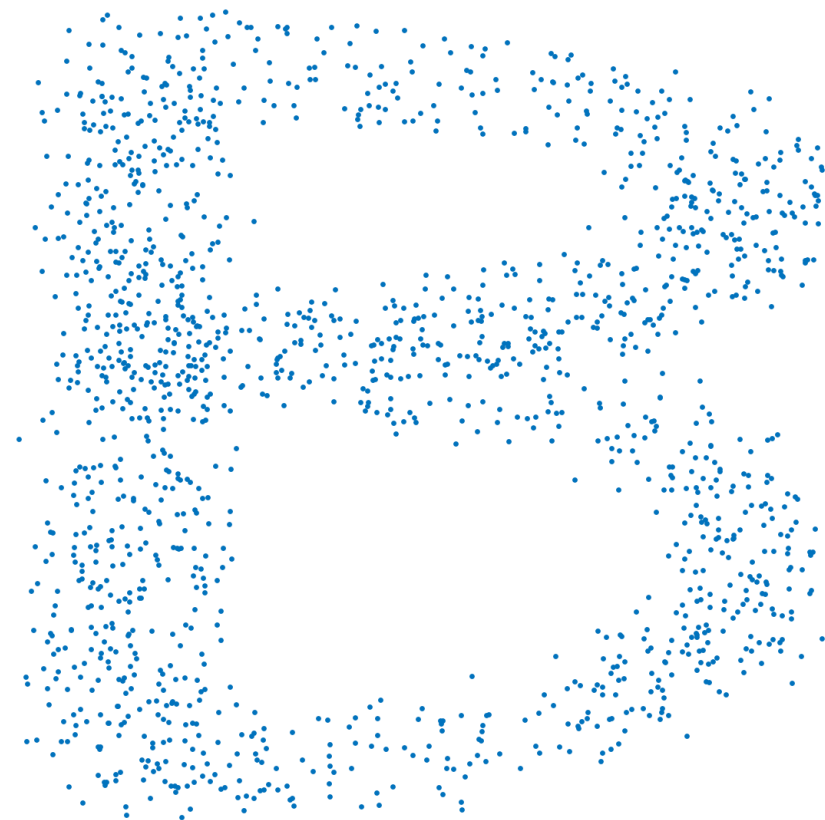
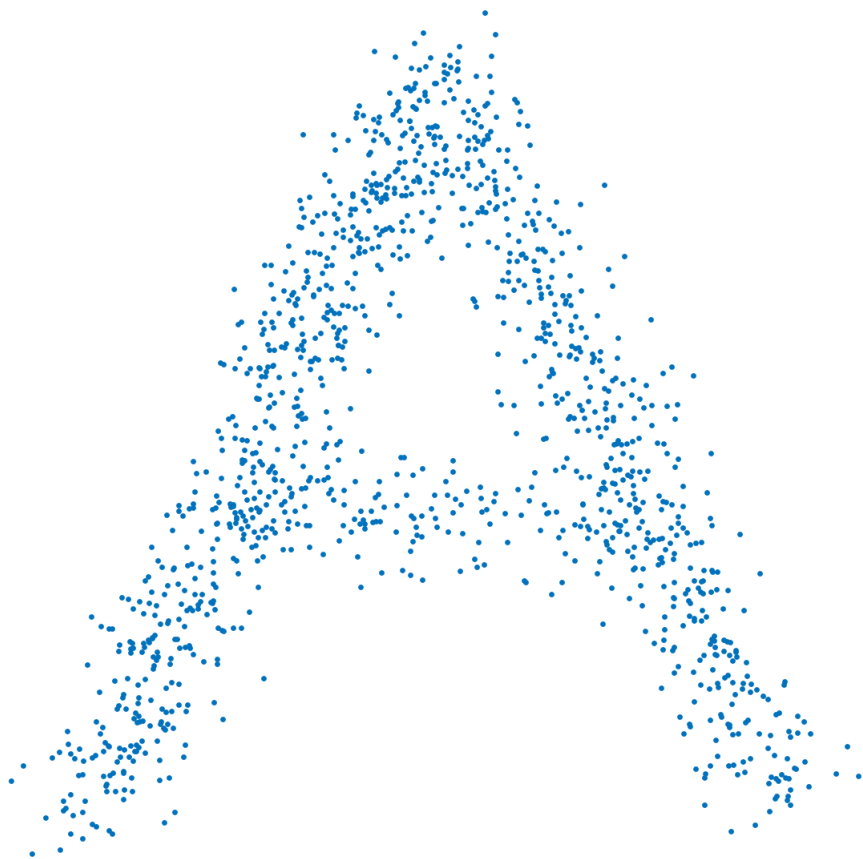
● Mapper解析

▼ 社会における様々な領域(生命科学・医療・産業・AI技術など)での取り組み

Contents

- パーシステントホモロジー解析（1パラメータ）
- 数学的背景
 - ホモロジー群
 - ポセットの表現論・区間表現
 - パーシステンス図の安定性
- マルチパラメータ化に向けた表現論

What's this?





What's this?



Let's try Persistent Homology Analysis!



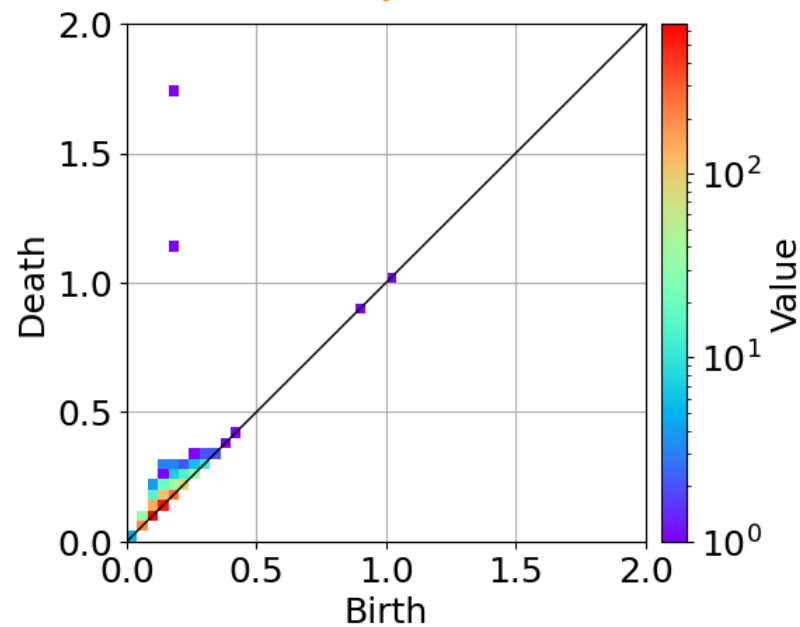
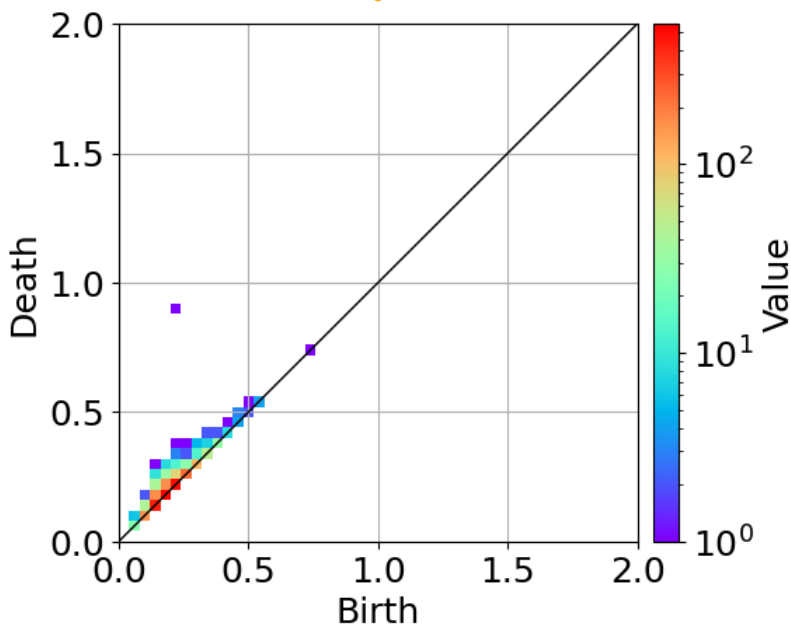
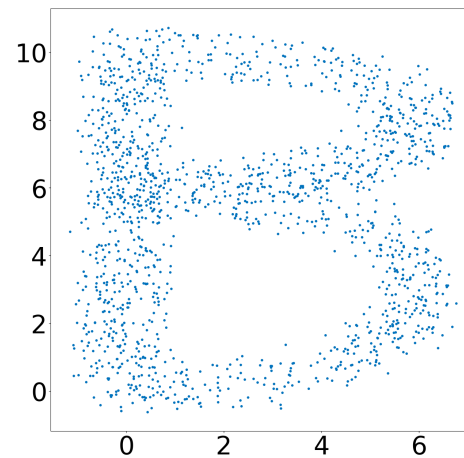
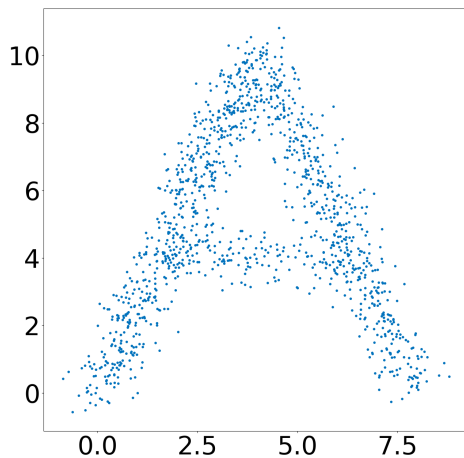
What's this?

n 次元座標
データ

画像データ

パーシステント
ホモロジー解析

パーシステンス図



What's this?

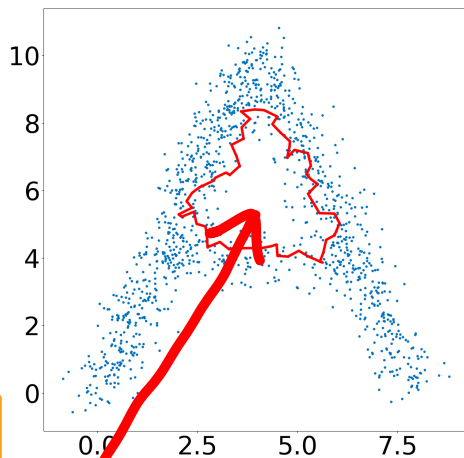
n 次元座標
データ

画像データ

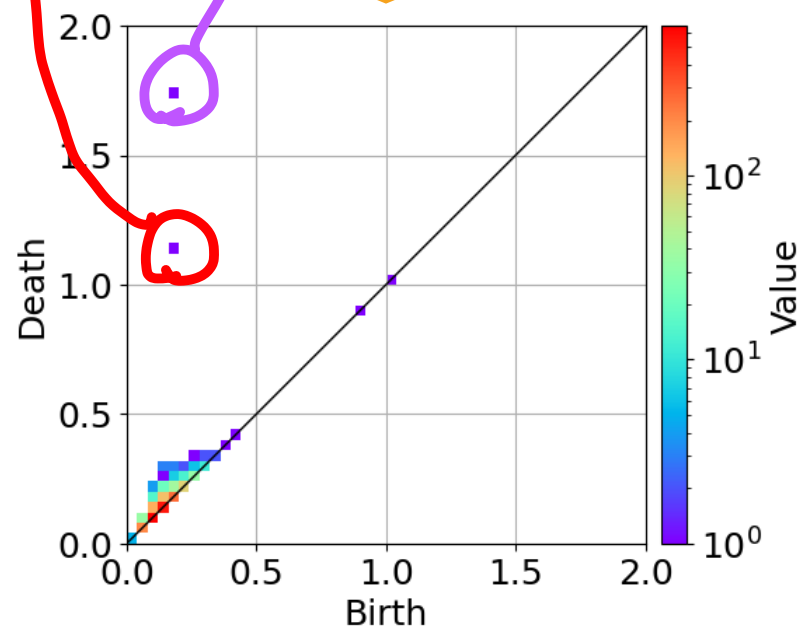
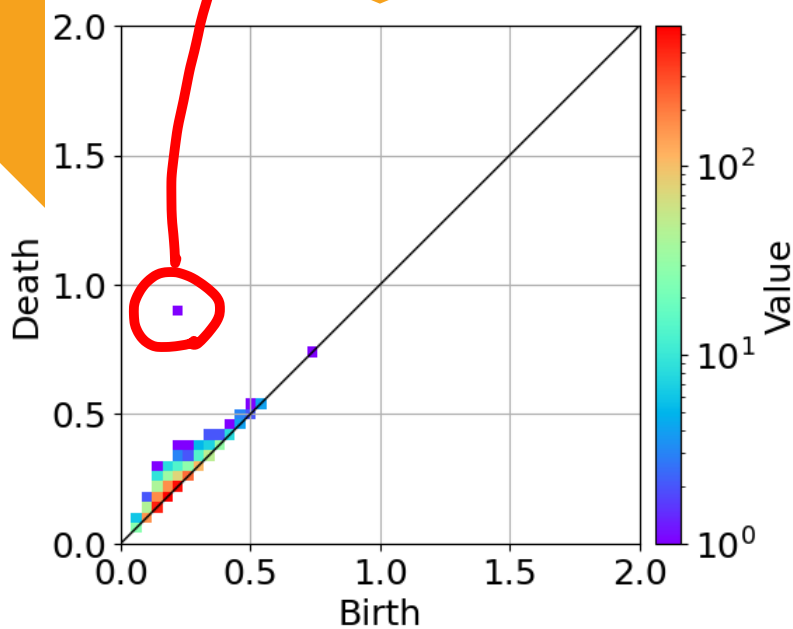
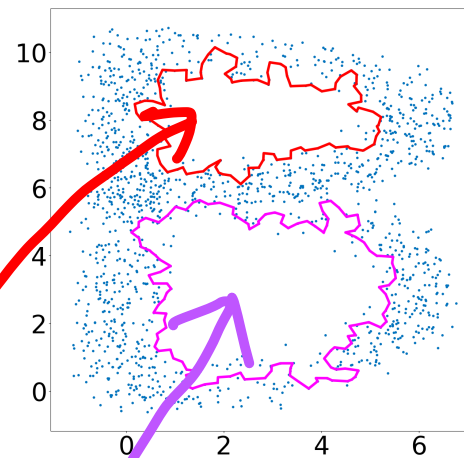
逆解析

パーシステント
ホモロジー解析

パーシステンス図



Let's try PH!





What are you doing?

What are you doing?

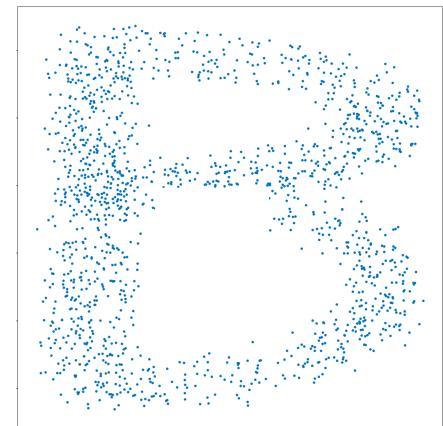
Let's try PH!

点群 $P \subseteq \mathbb{R}^N$ に対し, 各点の周りで半径 r の球体を考える.

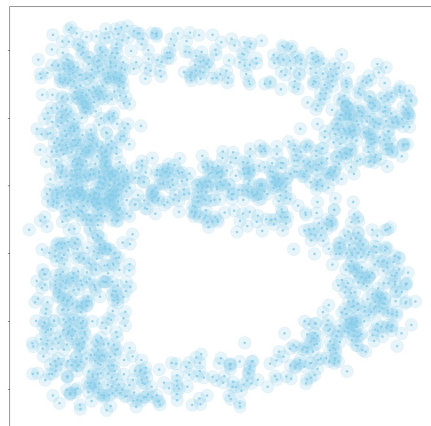
$$\bigcup_{x \in P} B(x, r)$$

~> 半径 r をどんどん大きくしていく.

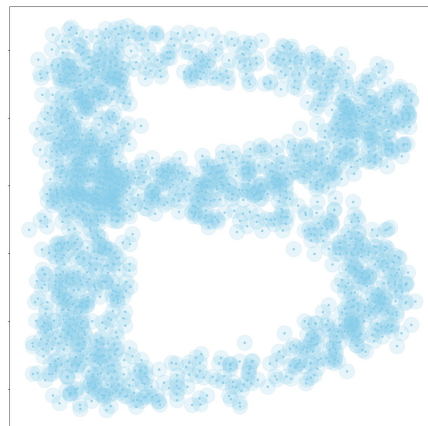
~> “穴”の発生・消滅時刻のペア(寿命)がデータの形を表わしている(と解釈).



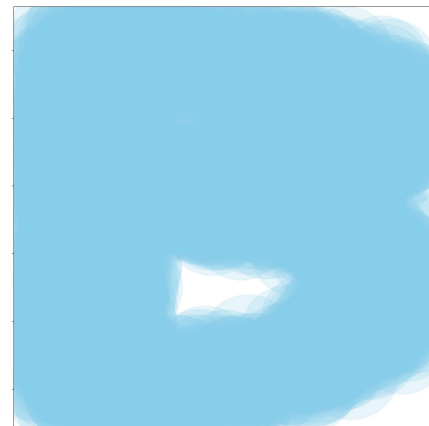
半径 $r = 0$



$r = 0.160$



$r = 0.166$



$r = 1.122$



$r = 1.725$



What are you doing?

Let's try PH!

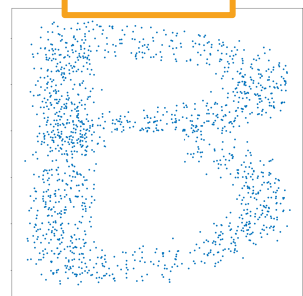
点群 $P \subseteq \mathbb{R}^N$ に対し, 各点の周りで半径 r の球体を考える.

$$\bigcup_{x \in P} B(x, r)$$

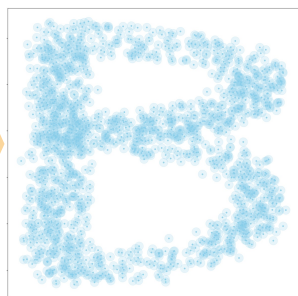
⇒ 半径 r をどんどん大きくしていく.

⇒ “穴”の発生・消滅時刻のペア(寿命)がデータの形を表わしている(と解釈).

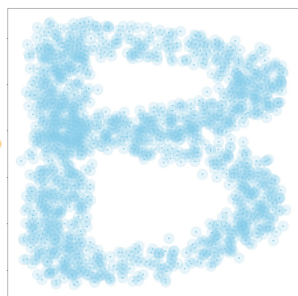
入力



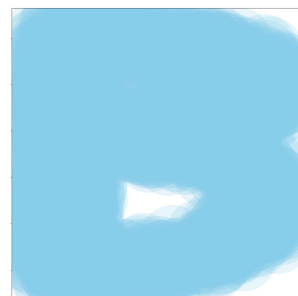
半径 $r = 0$



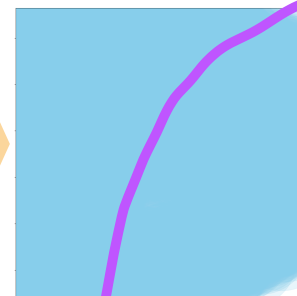
$r = 0.160$



$r = 0.166$

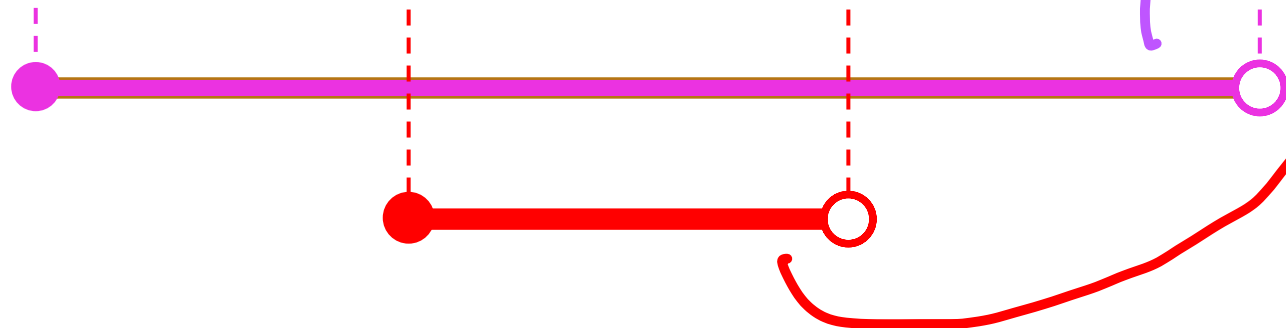
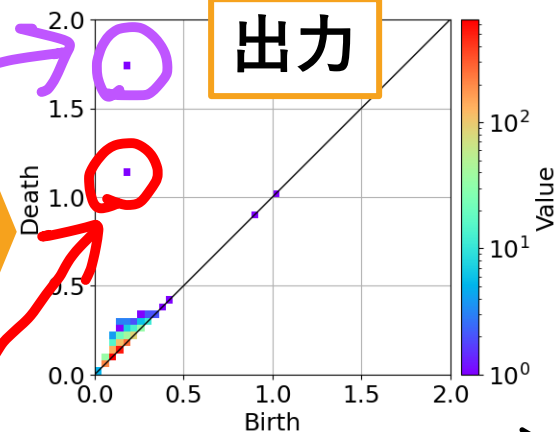


$r = 1.122$



$r = 1.725$

出力



大きい穴(下部)の寿命

区間 $[0.160, 1.725)$

小さい穴(上部)の寿命

区間 $[0.166, 1.122)$

What are you doing?

Let's try PH!

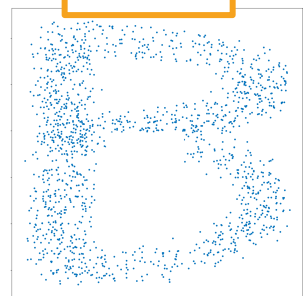
点群 $P \subseteq \mathbb{R}^N$ に対し, 各点の周りで半径 r の球体を考える.

$$\bigcup_{x \in P} B(x, r)$$

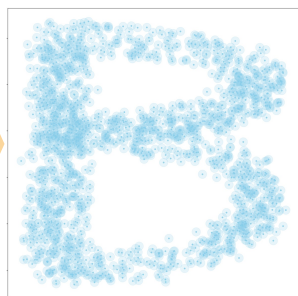
⇒ 半径 r をどんどん大きくしていく.

⇒ “穴”の発生・消滅時刻のペア(寿命)がデータの形を表わしている(と解釈).

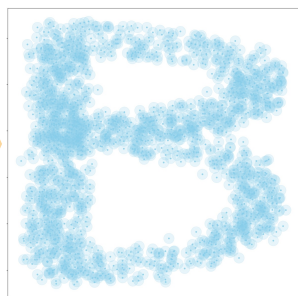
入力



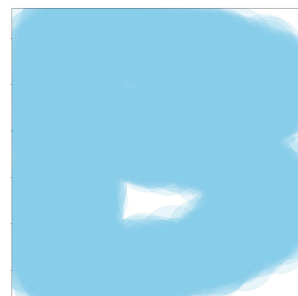
半径 $r = 0$



$r = 0.160$



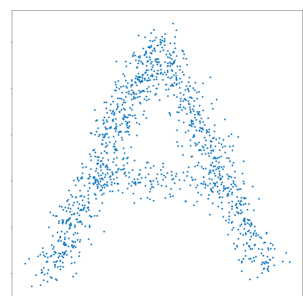
$r = 0.166$



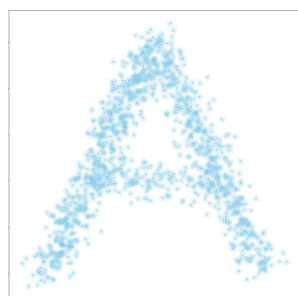
$r = 1.122$



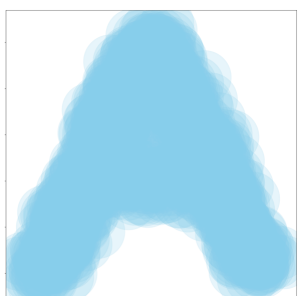
$r = 1.725$



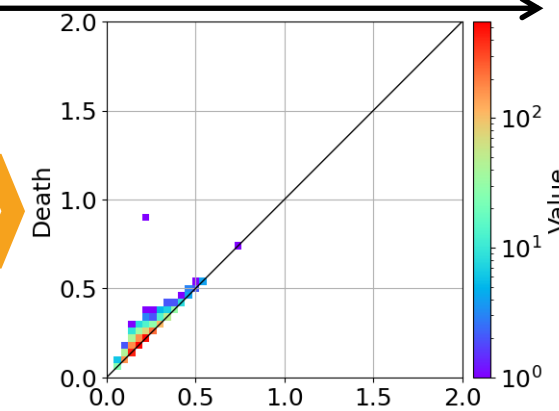
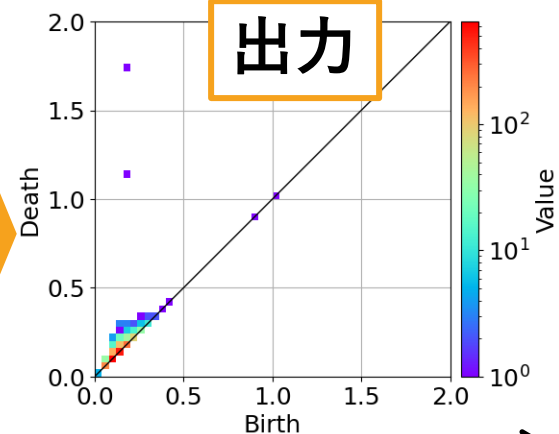
$r = 0.205$



$r = 0.881$



出力



材料科学における適用例

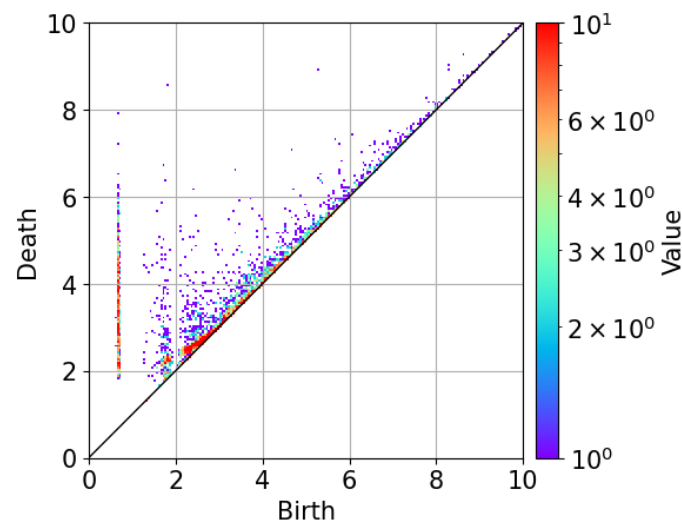
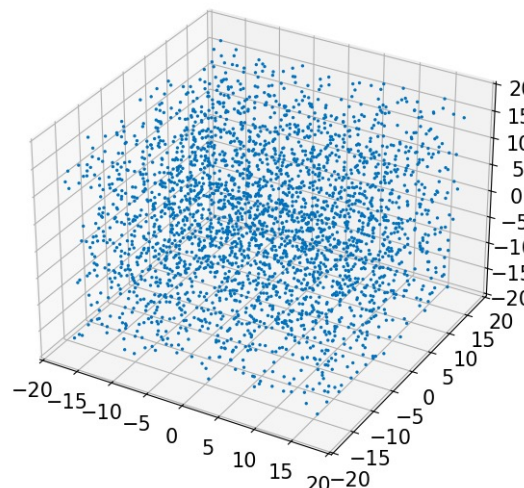
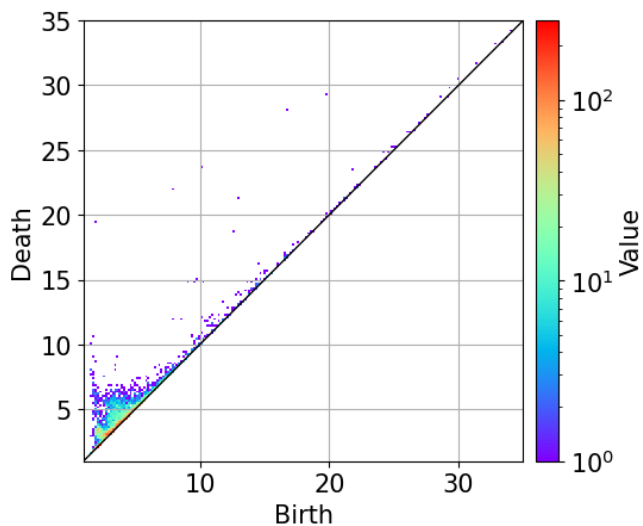
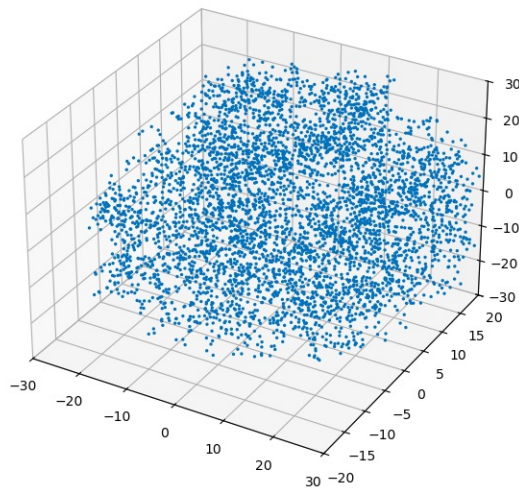
ヘモグロビン (DEOXY HUMAN HEMOGLOBIN)

出典: Protein Data
Bank Japan

PDBエントリーID
1A3N

標準アミノ酸を
構成している
原子のx, y, z座標
[Å]を図示

Å: オングストローム
(1Å = 0.1 ナノメートル)



アモルファスシリカ (AMORPHOUS SILICA)

出典: S. L. Roux, V. Petkov,
ISAACS – Interactive
structure analysis
of amorphous and
crystalline systems.
J. Appl. Cryst. 43,
181–185 (2010).

<https://isaacs.sourceforge.io/ex.html>

材料科学における適用例

平岡先生(京都大学)らによる ガラスの原子配置の解析(2016)

PH解析の材料科学における最初期の成功例

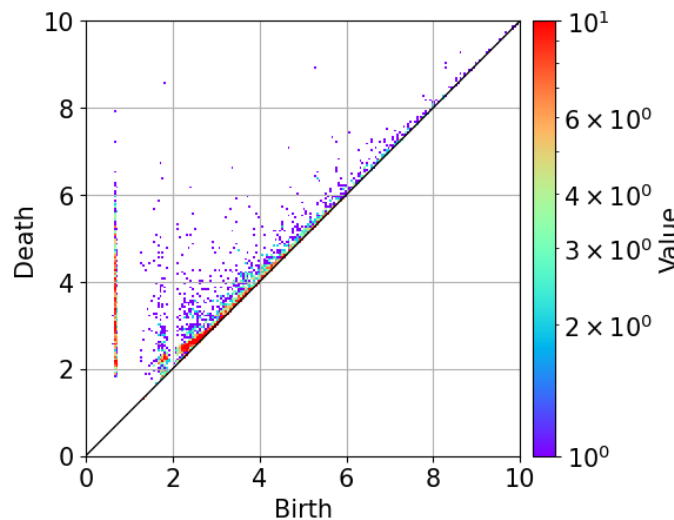
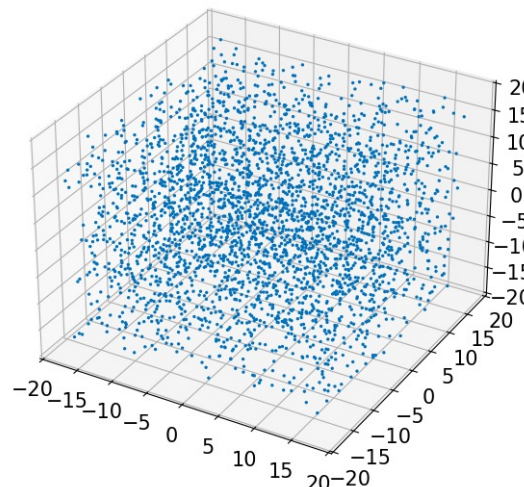
パーシステントホモロジー解析により

- 中距離秩序構造の記述および
- 液体とガラス状態の内部構造の違いの幾何学的特徴付けに成功.

論文情報: Y. Hiraoka, T. Nakamura, A. Hirata, E. G. Escobar, K. Matsue, and Y. Nishiura, "Hierarchical structures of amorphous solids characterized by persistent homology", Proceedings of the National Academy of Sciences.

参考資料: ガラスの「形」を数学的に解明～トポロジーで読み解く無秩序の中の秩序～

<https://www.jst.go.jp/pr/announce/20160614/index.html>

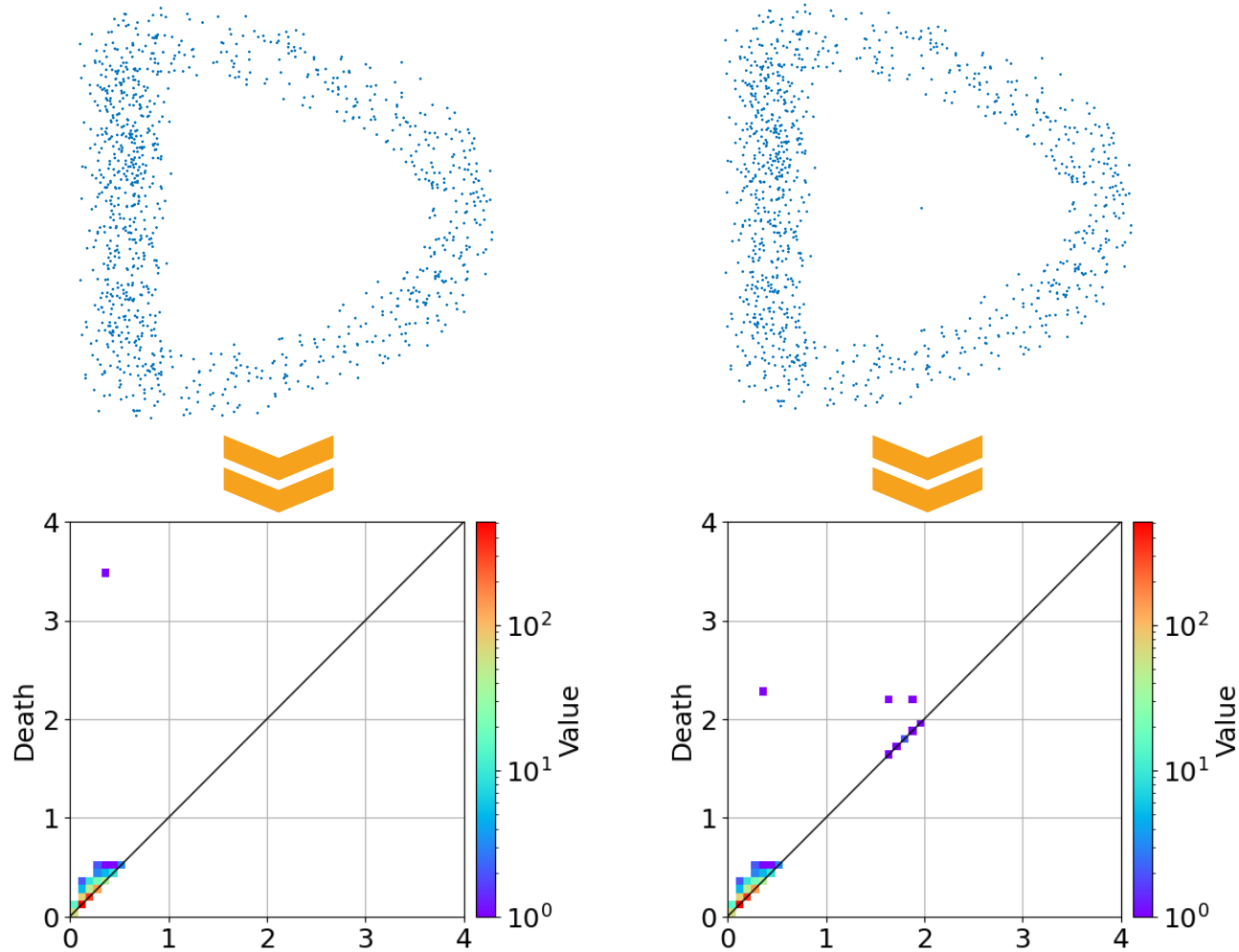


アモルファスシリカ (AMORPHOUS SILICA)

出典: S. L. Roux, V. Petkov, ISAACS – Interactive structure analysis of amorphous and crystalline systems. J. Appl. Cryst. 43, 181–185 (2010).

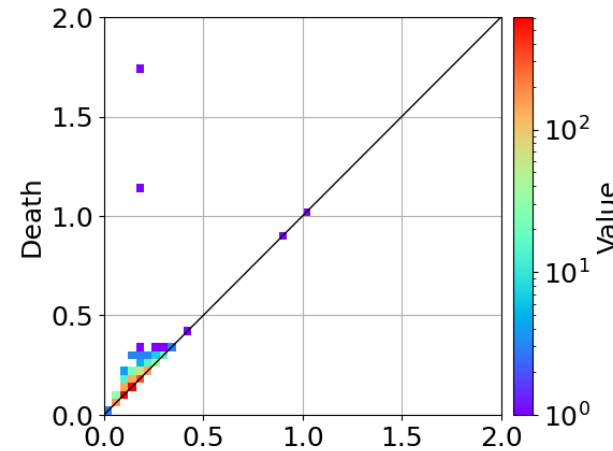
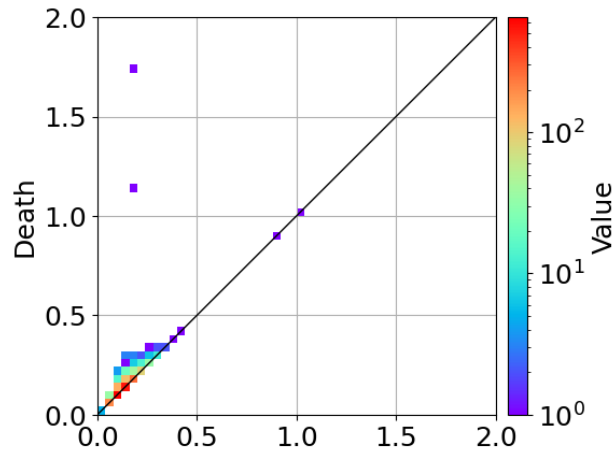
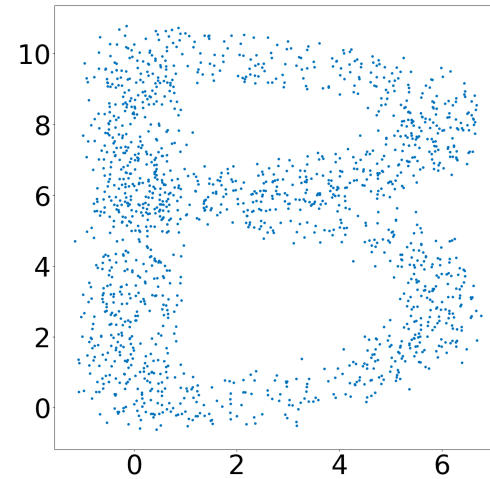
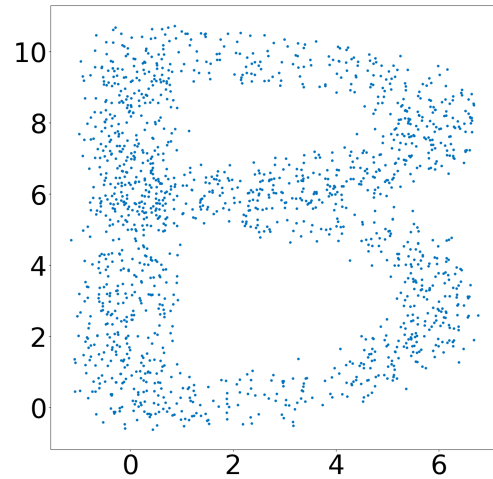
<https://isaacs.sourceforge.io/ex.html>

Hmm... Their PD look quite different.



PH is not robust and unstable to outliers.

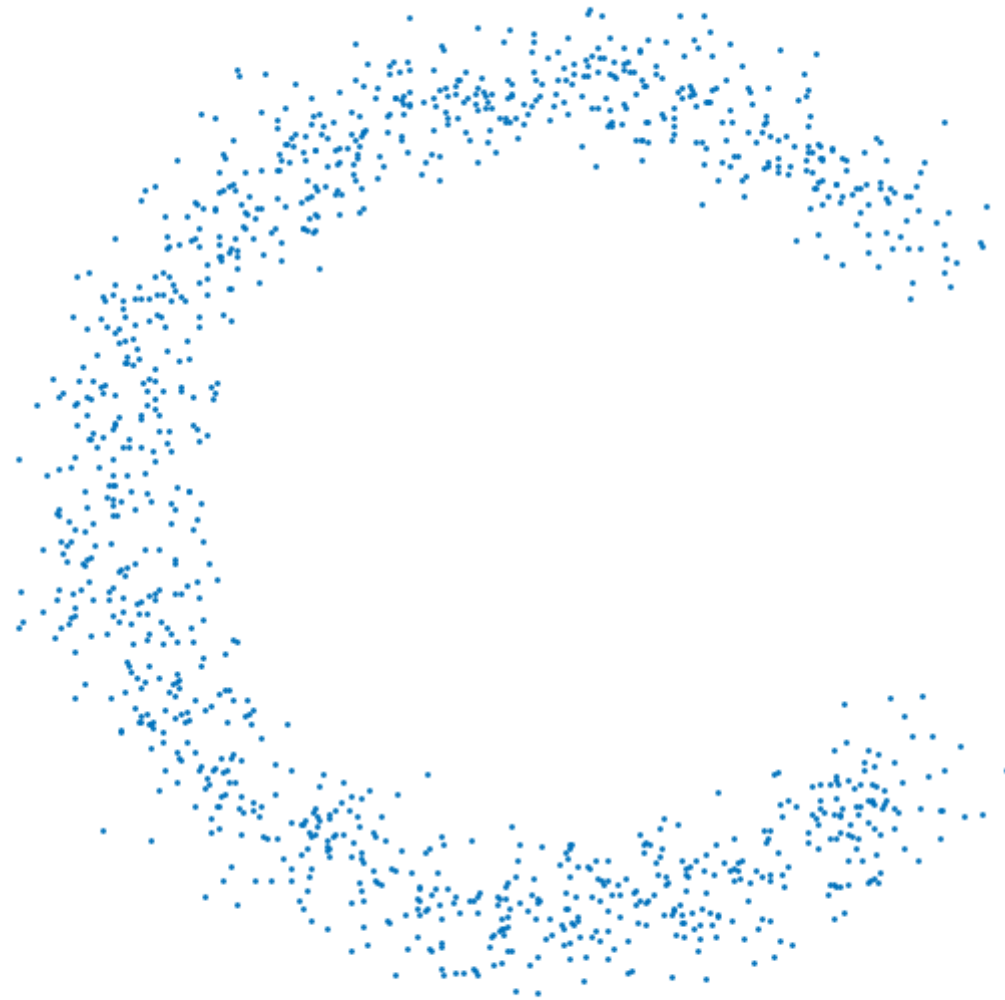
Hmm... Are they similar to each other?



Yes, they are similar by stability!

Oh... I can imagine.

Exercise.



計算ソフトウェア

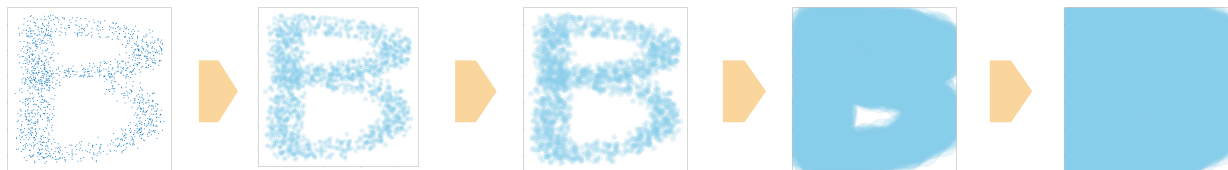
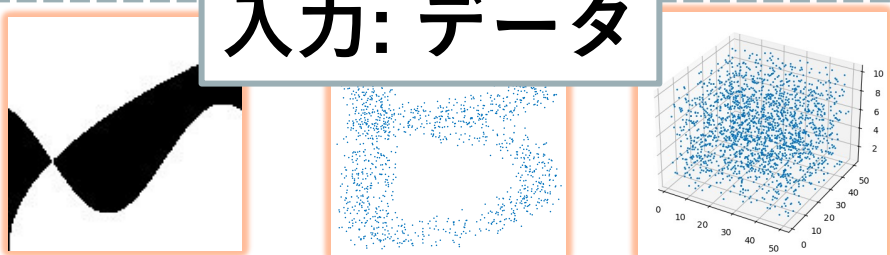
- GUDHI
- HomCloud (逆解析の機能、国産) 今回の資料作成に使用
- Ripser
- RIVET (2パラメータの不変量計算)

など他にも多数

パーシステントホモロジー解析の流れ

- ・ 原子等
- ・ 物体の配置
- ・ 特徴量
- ・ CT画像
- ・ …

入力: データ

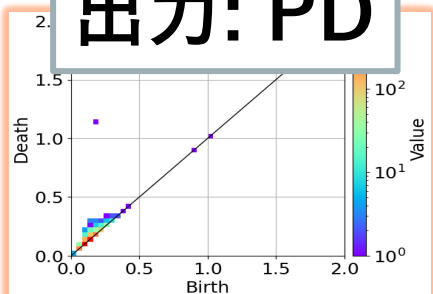


$$F: X_0 \subseteq X_1 \subseteq X_2 \subseteq X_3 \subseteq X_4$$

$$H_q(P) \rightarrow H_q(X_1) \rightarrow H_q(X_2) \rightarrow H_q(X_3) \rightarrow H_q(X_4)$$

$$H_q(F) \cong \bigoplus_{i=1}^m V_{[a_i, b_i]}$$

出力: PD



Data

Filtration

Persistence
Modules

Persistence
Diagrams

形を捉える

単体的複体の
ホモロジー

マルチスケール
で捉える

フィルトレーション
付き単体的複体

マルチスケールで
形を捉える

パーシステンス
加群

MATHEMATICAL FOUNDATIONS

単体的複体

ホモロジー群

フィルトレーション

PH Analysis

Definition (単体的複体)

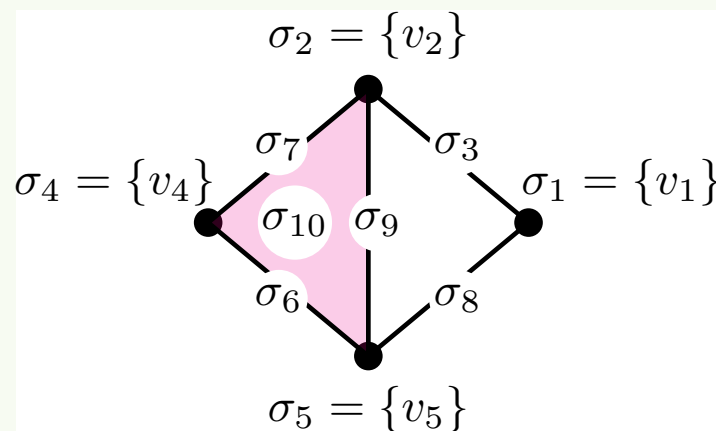
有限集合 V の部分集合の集まり $K \subseteq 2^V$ が

- $v \in V \implies \{v\} \in K$
- $\tau \subseteq \sigma \subseteq V, \sigma \in K \implies \tau \in K$

を満たすとき, K を頂点に V を持つ (抽象的) 単体的複体と呼ぶ.

Example

$V = \{v_1, v_2, v_4, v_5\}$,
 $K := \{\sigma_1, \dots, \sigma_{10}\}$, where
 $\sigma_i := \{v_i\}$ ($v_i \in V$),
 $\sigma_3 := \{v_1, v_2\}$,
 $\sigma_{10} := \{v_2, v_4, v_5\}$, etc.



0-単体 頂点
1-単体 辺
2-単体 面
⋮

Data

Filtration

Persistence
Modules

Persistence
Diagrams

単体的複体

ホモロジー群

フィルトレーション

K : 単体的複体. k : 体 (e.g., 複素数体 \mathbb{C} , 有限体 $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$)

鎖複体 $C(K; k) = (C_q(K; k), \partial_q)_{q \in \mathbb{Z}}$, where

- $C_q(K; k) := \text{span}_k \{ \sigma \mid \sigma \in K, \#\sigma = q + 1 \}$.
- $\partial_q: C_q(K; k) \rightarrow C_{q-1}(K; k)$ s.t. $\partial_{q-1} \circ \partial_q = 0$,
 $\sigma = \{v_0, \dots, v_q\} \mapsto \sum_{r=0}^q (-1)^r \{v_0, \dots, v_{r-1}, v_{r+1}, \dots, v_q\}$.

Definition (q 次ホモロジー群)

$$H_q(K; k) := \text{Ker } \partial_q / \text{Im } \partial_{q+1}.$$

ホモロジーは関手的. $H_q: \text{Simp} \rightarrow \text{vect } k$.

$f: K \rightarrow L$: 単体写像に対し, $H_q(f): H_q(K) \rightarrow H_q(L)$: 線型写像.

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

単体的複体

ホモロジー群

フィルトレーション

K : 単体的複体. k : 体 (e.g., 複素数体 \mathbb{C} , 有限体 $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$)

鎖複体 $C(K; k) = (C_q(K; k), \partial_q)_{q \in \mathbb{Z}}$, where

- $C_q(K; k) := \text{span}_k \{ \sigma \mid \sigma \in K, \#\sigma = q + 1 \}$.
- $\partial_q: C_q(K; k) \rightarrow C_{q-1}(K; k)$ s.t. $\partial_{q-1} \circ \partial_q = 0$,
 $\sigma = \{v_0, \dots, v_q\} \mapsto \sum_{r=0}^q (-1)^r \{v_0, \dots, v_{r-1}, v_{r+1}, \dots, v_q\}$.

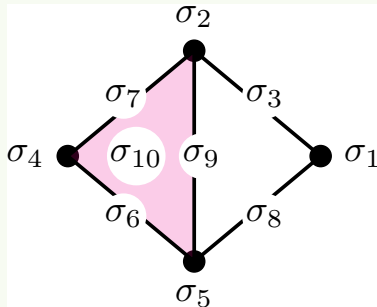
Definition (q 次ホモロジー群)

$$H_q(K; k) := \text{Ker } \partial_q / \text{Im } \partial_{q+1}.$$

H_0 連結成分

H_1 穴

H_2 空洞



$$\dim_k H_q(K, k) = \begin{cases} 1 & q = 0, \\ 1 & q = 1, \\ 0 & q \geq 2. \end{cases}$$

PH Analysis

Data

Filtration

Persistence Modules

Persistence Diagrams

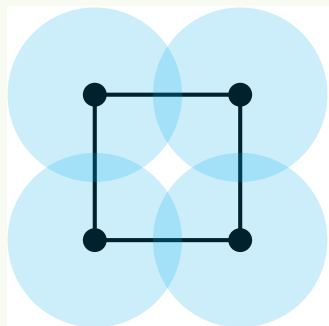
単体的複体

ホモロジー群

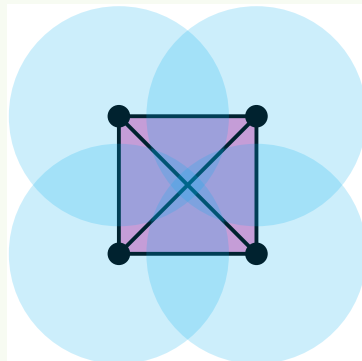
フィルトレーション

チェック複体

$$\bigcup_{x \in P} B(x; r) \simeq |\check{C}(P, r)| \implies H_q\left(\bigcup_{x \in P} B(x; r)\right) \simeq H_q(\check{C}(P, r))$$



$r = 0.6$



$r = 0.8$

$\dim_k H_0$

1

1

$\dim_k H_1$

1

0

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

単体的複体

ホモロジー群

フィルトレーション

$F: \mathbb{R}$ -filtration $\stackrel{\text{def}}{\iff} F: \mathbb{R} \rightarrow \text{Top/Simp}$ s.t. $F(r) \subseteq F(s)$ ($\forall r \leq s$)

Definition (sub-level set filtration)

X : 位相空間 or 単体的複体. 与えられた関数 $f: X \rightarrow \mathbb{R}$ に対し,
 X の部分空間の族 $\mathcal{S}^\uparrow(f) := (\mathcal{S}^\uparrow(f)_r)_{r \in \mathbb{R}}$:

$$\mathcal{S}^\uparrow(f)_r := f^{-1}((-\infty, r]) \subseteq X$$

を f による X の sub-level set filtration と呼ぶ. 実際, 以下が成立.

$$r \leq s \implies \mathcal{S}^\uparrow(f)_r \subseteq \mathcal{S}^\uparrow(f)_s.$$

ただし, X が単体的複体の場合は以下の条件を課すものとする.

$$\tau \subseteq \sigma \implies f(\tau) \leq f(\sigma).$$

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

単体的複体

ホモロジー群

フィルトレーション

Example (点を膨らませる構成はsub-level set filtrationである.)

有限集合 $P \subset \mathbb{R}^N$ に対し, $X := \mathbb{R}^N$ 上の関数 f_P を次で定める.

$$\begin{array}{ccc} f_P: & \mathbb{R}^N & \longrightarrow & \mathbb{R} \\ & \cup & & \cup \\ & \mathbf{y} & \longmapsto & \inf_{\mathbf{x} \in P} \|\mathbf{x} - \mathbf{y}\|. \end{array}$$

このとき, sub-level set filtration $\mathcal{S}^\uparrow(f_P)$ について以下が成立.

$$\mathcal{S}^\uparrow(f_P)_r = \bigcup_{\mathbf{x} \in P} B(\mathbf{x}, r) \quad (\simeq |\check{C}(P, r)|).$$

他にも Vietoris-Rips filtration や, super-level bi-filtration などとも sub-level set filtration として構成できる.

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

単体的複体

ホモロジー群

フィルトレーション

PH Analysis

Observation

$f: X \rightarrow \mathbb{R}$ に対し (e.g. $f = f_P$), sub-level set filtration

$$S^\uparrow(f) = (S^\uparrow(f)_r)_{r \in \mathbb{R}} \quad \text{where} \quad S^\uparrow(f)_r = f^{-1}((-\infty, r])$$

を得た. 特に, 任意の増大列 $r_0 < r_1 < r_2 < \dots$ に対し

$$S^\uparrow(f)_{r_0} \xrightarrow{\iota_0} S^\uparrow(f)_{r_1} \xrightarrow{\iota_1} S^\uparrow(f)_{r_2} \xrightarrow{\iota_2} \dots$$

H_q の関手性より, 次のベクトル空間と線型写像の列を得る.

$$H_q(S^\uparrow(f)_{r_0}) \xrightarrow{H_q(\iota_0)} H_q(S^\uparrow(f)_{r_1}) \xrightarrow{H_q(\iota_1)} H_q(S^\uparrow(f)_{r_2}) \xrightarrow{H_q(\iota_2)} \dots$$

Data

Filtration

Persistence
Modules

Persistence
Diagrams

単体的複体

ホモロジー群

フィルトレーション

PH Analysis

Observation

$f: X \rightarrow \mathbb{R}$ に対し (e.g. $f = f_P$), sub-level set filtration

$$S^\uparrow(f) = (S^\uparrow(f)_r)_{r \in \mathbb{R}} \quad \text{where} \quad S^\uparrow(f)_r = f^{-1}((-\infty, r])$$

を得た. 特に, 任意の増大列 $r_0 < r_1 < r_2 < \dots$ に対し

$$S^\uparrow(f)_{r_0} \xrightarrow{\iota_0} S^\uparrow(f)_{r_1} \xrightarrow{\iota_1} S^\uparrow(f)_{r_2} \xrightarrow{\iota_2} \dots$$

H_q の関手性より, 次のベクトル空間と線型写像の列を得る.

$$H_q(S^\uparrow(f)_{r_0}) \xrightarrow{H_q(\iota_0)} H_q(S^\uparrow(f)_{r_1}) \xrightarrow{H_q(\iota_1)} H_q(S^\uparrow(f)_{r_2}) \xrightarrow{H_q(\iota_2)} \dots$$

$$\mathbb{R} \xrightarrow{S^\uparrow(f)} \text{Top/Simp} \xrightarrow{H_q(-;k)} \text{vect } k \quad : \text{ functor.}$$

\mathbb{R} -パーシステンス加群(表現).

Data

Filtration

Persistence
Modules

Persistence
Diagrams

ポセット P の表現の圏 $\text{rep}_k(P) := \text{Fun}(P, \text{vect } k)$

Objects: $M = (M(x), M(x \leq y))_{x,y \in P}$ s.t.

- 各元 $x \in P$ に対し, $M(x)$ は有限次元 k ベクトル空間.
- 各 $x \leq y$ に対し, $M(x \leq y): M(x) \rightarrow M(y)$ は k 線型写像.
- $M(x \leq x) = \text{id}_{M(x)}$ ($\forall x \in P$).
- $M(y \leq z) \circ M(x \leq y) = M(x \leq z)$ ($\forall x \leq y \leq z$)

$$\begin{array}{ccc}
 M(x) & \xrightarrow{M(x \leq y)} & M(y) \\
 \downarrow f_x & \circlearrowleft & \downarrow f_y \\
 N(x) & \xrightarrow{N(x \leq y)} & N(y)
 \end{array}$$

Morphisms: $f = (f_x)_{x \in P}: M \rightarrow N$ s.t.

- $N(x \leq y) \circ f_x = f_y \circ M(x \leq y)$ ($\forall x \leq y$).

Example

$$P = \begin{array}{ccccc}
 & & \leq & & \\
 & & \nearrow & & \searrow \\
 x & & & y & & z \\
 & & \searrow & & \nearrow \\
 & & \leq & & \\
 & & & w & & \\
 & & \leq & & \\
 & & & z & & \leq \\
 & & & & & v,
 \end{array}
 \quad
 M := \begin{array}{ccccc}
 & & \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \rightarrow & k^2 & \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} & \rightarrow & k^3 & \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} & \rightarrow & k \\
 & & & \circlearrowleft & & & & & & & \\
 k & & \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \rightarrow & k^3 & \xrightarrow{\text{id}} & k^3 & & & &
 \end{array}$$

PH Analysis

Data

Filtration

Persistence Modules

Persistence Diagrams

ポセット P の表現の圏 $\text{rep}_k(P) := \text{Fun}(P, \text{vect } k)$

Objects: $M = (M(x), M(x \leq y))_{x,y \in P}$ s.t.

- 各元 $x \in P$ に対し, $M(x)$ は有限次元 k ベクトル空間.
- 各 $x \leq y$ に対し, $M(x \leq y): M(x) \rightarrow M(y)$ は k 線型写像.
- $M(x \leq x) = \text{id}_{M(x)}$ ($\forall x \in P$).
- $M(y \leq z) \circ M(x \leq y) = M(x \leq z)$ ($\forall x \leq y \leq z$)

Morphisms: $f = (f_x)_{x \in P}: M \rightarrow N$ s.t.

- $N(x \leq y) \circ f_x = f_y \circ M(x \leq y)$ ($\forall x \leq y$).

$$\begin{array}{ccc} M(x) & \xrightarrow{M(x \leq y)} & M(y) \\ \downarrow f_x & \circlearrowleft & \downarrow f_y \\ N(x) & \xrightarrow{N(x \leq y)} & N(y) \end{array}$$

- $\text{rep}_k(P)$ はアーベル圏.
- $\text{rep}_k(P)$ 上で直既約分解の一意性が成立(Kull-Schmidtの定理).
- P が有限集合のとき, $\text{rep}_k(P)$ は隣接代数 $k[P]$ の加群圏と圏同値.

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

ポセット P の表現の圏 $\text{rep}_k(P) := \text{Fun}(P, \text{vect } k)$

Definition (区間)

P の部分集合 I が区間であるとは、次を満たすときをいう。

- (凸性) $\forall x, y \in I, \forall z \in P, x < z < y \implies z \in I$.
- (連結性) $\forall x, y \in I, \exists x = z_0, z_1, \dots, z_{m-1}, z_m = y \stackrel{e \in I}{\text{s.t.}}$
 $z_j < z_{j+1}$ or $z_j > z_{j+1}$ ($0 \leq j \leq m - 1$).

Definition (区間表現)

P の区間 I に対し、区間表現 $V_I \in \text{rep}_k(P)$ を以下で定義する。

$$V_I(x) = \begin{cases} k & \text{if } x \in I, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad V_I(x \leq y) = \begin{cases} \text{id}_k & \text{if } x, y \in I, \\ 0 & \text{otherwise.} \end{cases}$$

- V_I は直既約表現である。実際、 $\text{End}(V_I) \cong k$ が成立。
- V_I の部分表現や商表現は区間表現の直和である。

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

$P = (\mathbb{R}, \leq)$ の表現の圏 $\text{rep}_k(\mathbb{R})$

Theorem [Crawley-Boevey '15] (cf. Gabriel's thm for \mathbb{A}_n)

$\text{rep}_k(\mathbb{R})$ の任意の直既約表現は区間表現である. したがって, 任意の表現 $M \in \text{rep}_k(\mathbb{R})$ は以下の形の直和に同型となる.

$$M \cong \bigoplus_{I \in \mathcal{B}(M)} V_I.$$

(\mathbb{R}, \leq) 上の区間は以下の形のいずれか ($I = \langle a, b \rangle$ で表わす).

$[a, b], (a, b], [a, b), (a, b), [a, \infty), (a, \infty), (-\infty, b], (-\infty, b).$

Definition (パーシステンス図)

$\bar{\mathbb{R}} := \mathbb{R} \sqcup \{\pm\infty\}$ とおく. $\bar{\mathbb{R}}^2$ 上の多重集合

$$D(M) := \{(a, b) \in \bar{\mathbb{R}}^2 \mid I = \langle a, b \rangle \in \mathcal{B}(M)\}$$

を表現 M のパーシステンス図と呼ぶ.

PH Analysis

Data

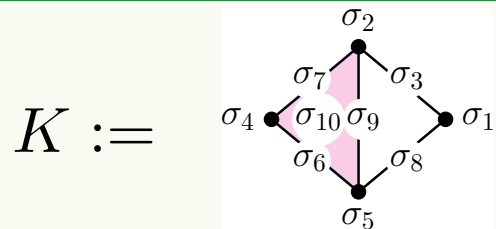
Filtration

Persistence
Modules

Persistence
Diagrams

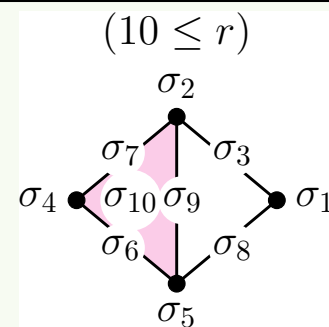
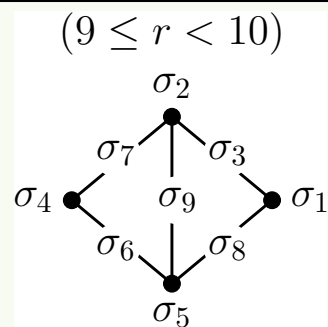
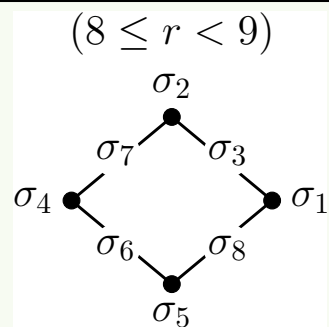
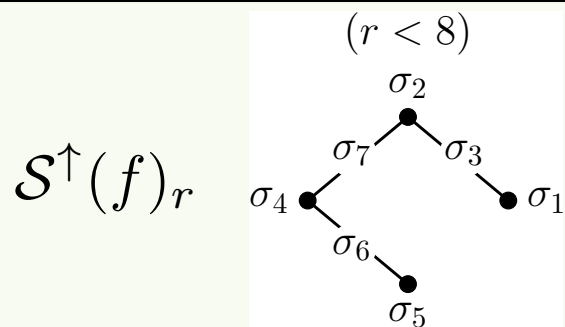
$P = (\mathbb{R}, \leq)$ の表現の圏 $\text{rep}_k(\mathbb{R})$

Example (over $k := \mathbb{F}_2$.)



$$f: K \rightarrow \mathbb{R}, f(\sigma_t) = t \quad (1 \leq t \leq 10).$$

$$\mathcal{S}^\uparrow(f)_r = \{\sigma_t \in K \mid t \leq r\}.$$



$H_1(\mathcal{S}^\uparrow(f)_r)$
の基底

\emptyset

$[z := \sigma_3 + \sigma_8 + \sigma_6 + \sigma_7]$

$[z]$

$[z]$

$[\sigma_6 + \sigma_7 + \sigma_9]$

$$H_1(\mathcal{S}^\uparrow(f)) \quad \dots \quad 0 \xrightarrow{0} k \xrightarrow{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} k^2 \xrightarrow{\begin{bmatrix} 1 & 0 \end{bmatrix}} k \xrightarrow{\text{id}_k} \dots$$

$$\cong V_{[8, \infty)} \oplus V_{[9, 10)}.$$

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

$P = (\mathbb{R}, \leq)$ の表現の圏 $\text{rep}_k(\mathbb{R})$

パーシステンス図を得るためには
表現の直既約分解の計算が必要…
というわけではない!

実用上, 効率のよい計算アルゴリズムが既実装されている.

e.g. 単体的複体の列から直接パーシステンス図が計算できる:

行列演算による標準アルゴリズムでは $O(n^3)$. (n は単体の個数)

cf. $O(n^\omega)$ for $\omega := \text{matrix multiplication time} < 2.371552$.

PH Analysis

Data

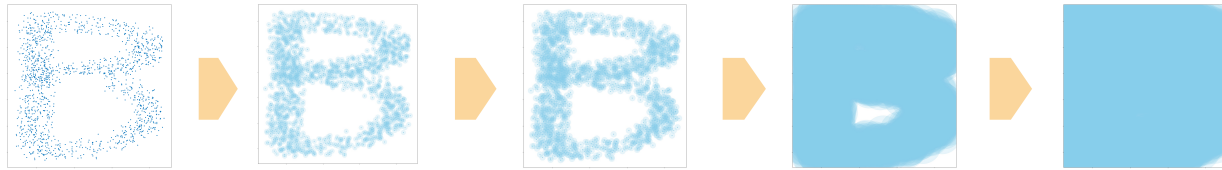
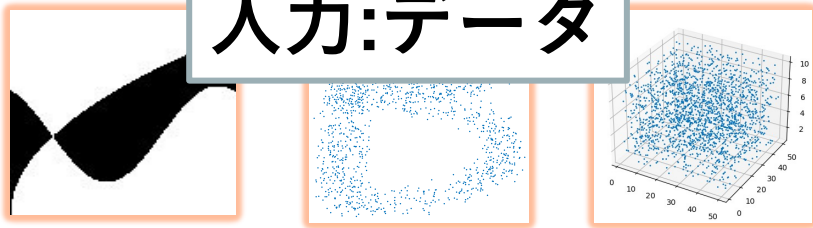
Filtration

Persistence
Modules

Persistence
Diagrams

パーシステントホモロジー解析の流れ(再訪)

入力:データ

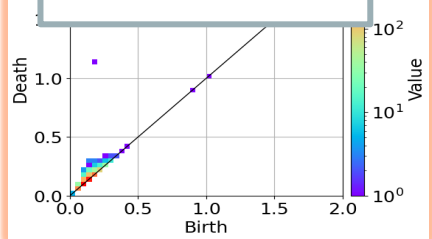


$$F: X_0 \subseteq X_1 \subseteq X_2 \subseteq X_3 \subseteq X_4$$

$$H_q(P) \rightarrow H_q(X_1) \rightarrow H_q(X_2) \rightarrow H_q(X_3) \rightarrow H_q(X_4)$$

$$H_q(F) \cong \bigoplus_{i=1}^m V_{[a_i, b_i]}$$

出力:PD



Data

Filtration

Persistence
Modules

Persistence
Diagrams

$$P \subseteq \mathbb{R}^N.$$

X : Top/Simp.

$$f_P: \mathbb{R}^N \rightarrow \mathbb{R}.$$

$$f: X \rightarrow \mathbb{R}.$$

$$\mathcal{S}^\uparrow(f) = (\mathcal{S}^\uparrow(f)_r)_{r \in \mathbb{R}} : \text{Sub-level set filt.}$$

$$V_f := H_q(\mathcal{S}^\uparrow(f)).$$

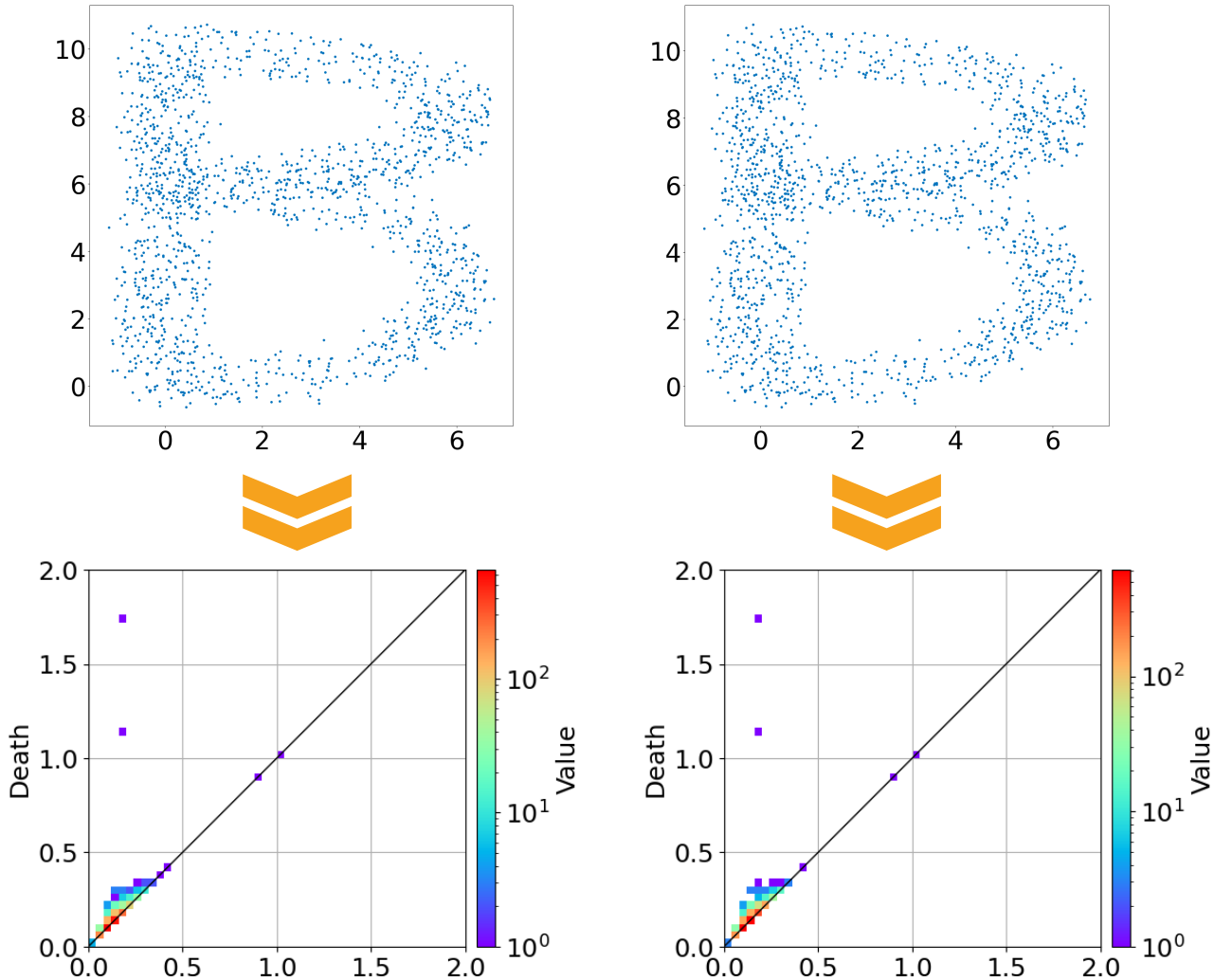
$$V_f \cong \bigoplus_{I \in \mathcal{B}(V_f)} V_I \text{ in } \text{rep}_k(\mathbb{R}).$$

V_I : 区間表現.

$D(V_f)$: パーシステンス図.

パーシステンス図の安定性 Stability

Hmm... Are they similar to each other?



Yes, they are similar by stability!

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

安定性(stability)

$\overline{\mathbb{R}} := \mathbb{R} \sqcup \{\pm\infty\}$: 拡張実数

Definition (bottleneck distance)

$\overline{\mathbb{R}}^2$ 上の多重集合 \mathcal{A}, \mathcal{B} の間の部分マッチングとは, 次の形の全単射 M :

$$\mathcal{A} \supset A \xrightarrow[M]{} B \subset \mathcal{B}$$

のことである. さらに M のコスト $c(M)$ を以下で定める.

$$c(M) := \max \left\{ \sup_{p \in A} \|\mathbf{p} - M(\mathbf{p})\|_{\infty}, \sup_{q \notin A \sqcup B} \frac{|q_2 - q_1|}{2} \right\}.$$

ボトルネック距離とは, 次で定義される拡張擬距離:

$$d_B(\mathcal{A}, \mathcal{B}) := \inf_M c(M).$$

“拡張” … 値に ∞ を取り得る.

“擬距離” … $d_B(\mathcal{A}, \mathcal{B}) = 0 \not\Rightarrow \mathcal{A} = \mathcal{B}$.

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

安定性(stability)

$\bar{\mathbb{R}} := \mathbb{R} \sqcup \{\pm\infty\}$: 拡張実数

PH Analysis

Definition (bottleneck distance)

$\bar{\mathbb{R}}^2$ 上の多重集合 \mathcal{A}, \mathcal{B} の間の部分マッチングとは, 次の形の全単射 M :

$$\mathcal{A} \supset A \xrightarrow[M]{\sim} B \subset \mathcal{B}$$

のことである. さらに M のコスト $c(M)$ を以下で定める:

$$c(M) := \max \left\{ \sup_{p \in A} \|p - M(p)\|_{\infty}, \sup_{q \notin A \sqcup B} \frac{|q_2 - q_1|}{2} \right\}.$$

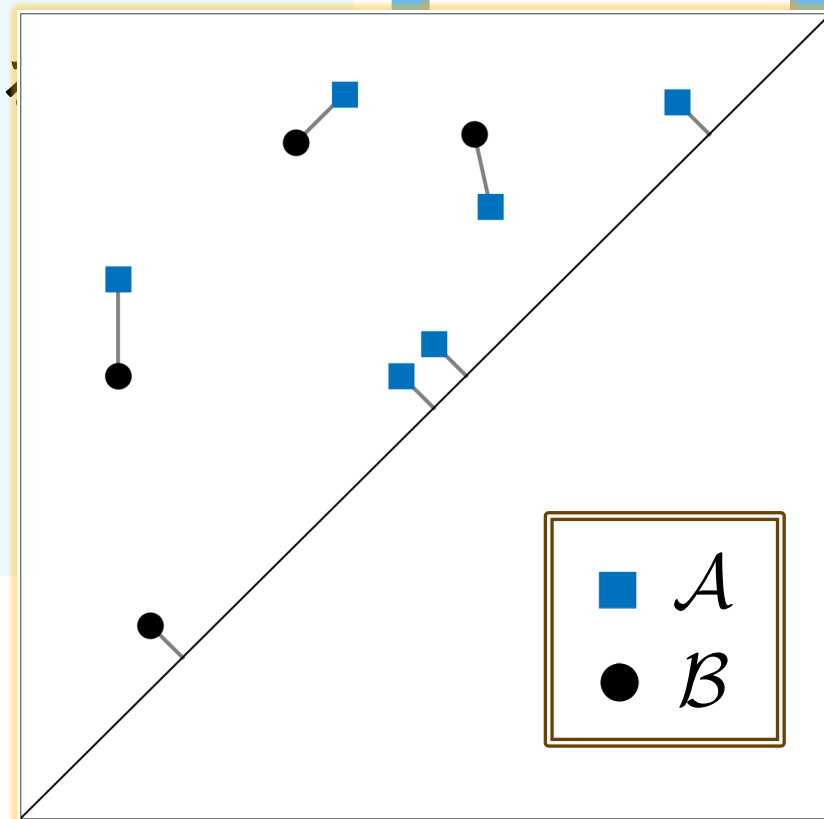
ボトルネック距離とは, 次で定義される拡張擬距離:

$$d_B(\mathcal{A}, \mathcal{B}) := \inf_M c(M).$$

“拡張” ... 値に ∞ を取り得る.

“擬距離” ... $d_B(\mathcal{A}, \mathcal{B}) = 0 \not\Rightarrow \mathcal{A} = \mathcal{B}$.

Data



安定性(stability)

X : 位相空間. $f, g: X \rightarrow \mathbb{R}$: 関数.

$V_f := H_q(\mathcal{S}^\uparrow(f)), V_g := H_q(\mathcal{S}^\uparrow(g)) \in \text{rep}_k(\mathbb{R})$.

$D(V_f), D(V_g) \subseteq \overline{\mathbb{R}^2}$: パーシステンス図 (多重集合) .

Theorem (安定性定理)

$$d_B(D(V_f), D(V_g)) \leq \|f - g\|_\infty.$$

$$\|f - g\|_\infty := \sup_{x \in X} |f(x) - g(x)|.$$

Corollary

有限部分集合 $P, P' \subseteq \mathbb{R}^N$ に対し, 以下が成立.

$$d_B(D(V_{f_P}), D(V_{f_{P'}})) \leq \|f_P - f_{P'}\|_\infty.$$

PH Analysis

Data

Filtration

Persistence
Modules

Persistence
Diagrams

まとめ

- パーシステントホモロジー解析はデータの「かたち」を捉える解析手法。

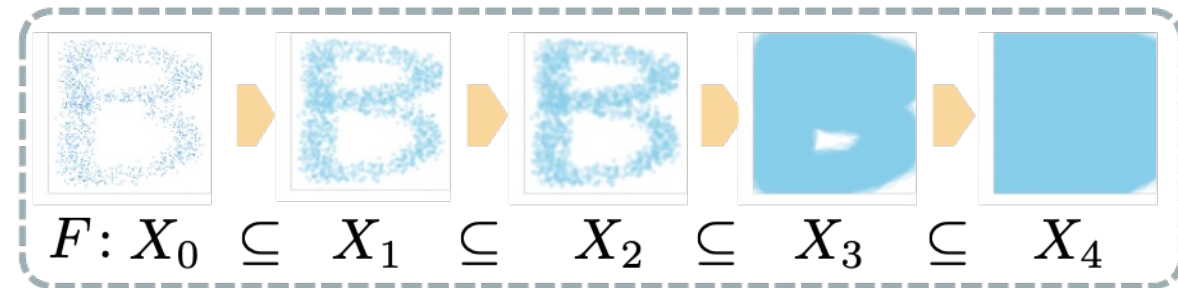
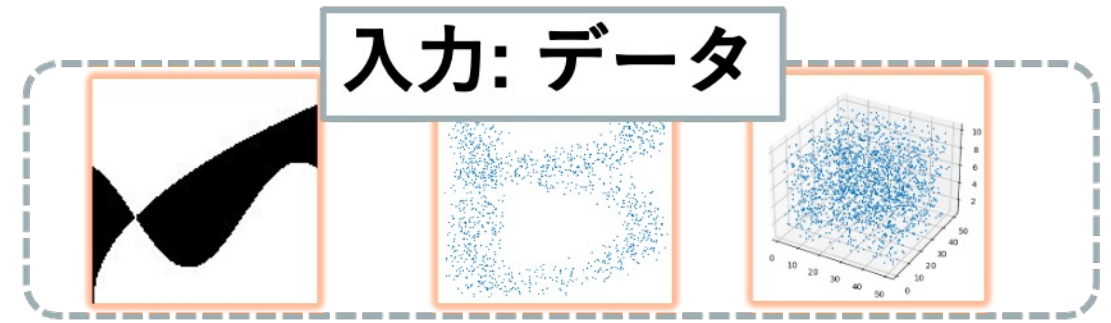
情報をパーシステンス図として出力。

「かたち」 = "穴"の持続性(パーシステンス)

- 逆解析により"穴"の特定・解釈が可能。
- 表現論が自然に現れる。
- 摂動に対する安定性定理が成り立つ。
ただし、外れ値に対しては頑強ではない。
- ◆ 特に材料科学への応用。
- ◆ 最近では機械学習との組合せ。

～ Special Thanks ～

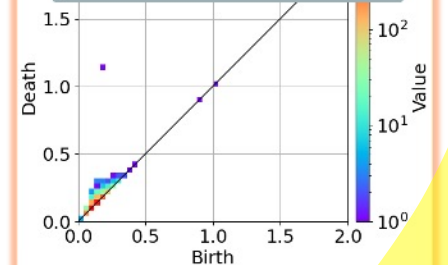
資料提供 多田駿介さん(神戸大D3)
with HomCloud and python.



$$H_q(X_0) \rightarrow H_q(X_1) \rightarrow H_q(X_2) \rightarrow H_q(X_3) \rightarrow H_q(X_4)$$

$$H_q(F) \cong \bigoplus_{i=1}^m V_{[a_i, b_i]}$$

出力: PD



パーシステントホモロジー解析の マルチパラメータ化(高次元化)に向けて

e.g. 時系列PH解析 $\left[\begin{array}{l} \text{半径パラメータ } r \\ \text{時間パラメータ } t \end{array} \right.$

$P = \mathbb{R}^N$ 上の表現は
一般に区間では表わせない。
表現の分類すら困難(wild表現型)。



“パーシステンス図”
が定義できない！
“穴”の解釈が困難！

現在は表現論的な観点から性質や不変量を研究。

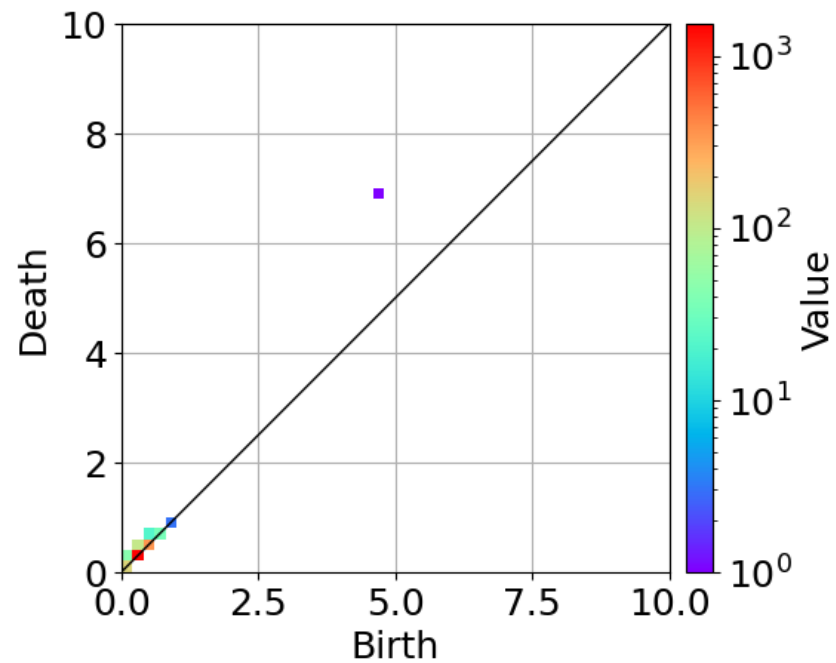
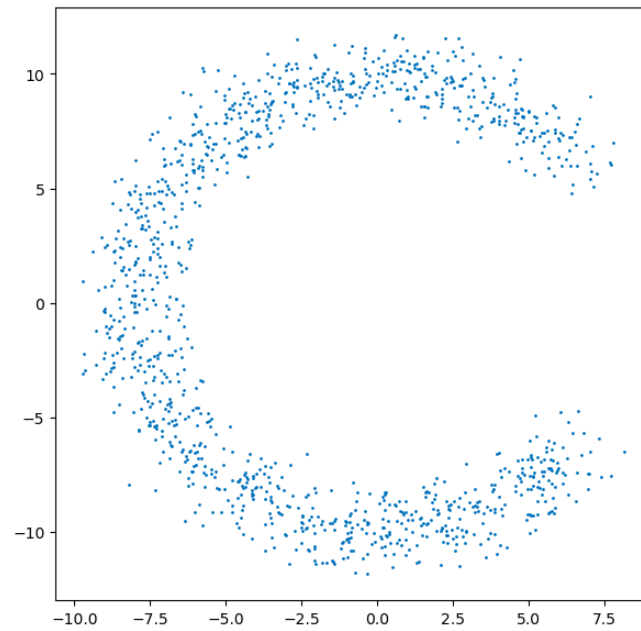
$\text{rep}_k(P)$ 上の表現論へ

パーシステントホモロジー解析の マルチパラメータ化(高次元化)に向けて

- **安定性定理の代数的理解**
インターリービング距離
- **区間表現の理論**
区間近似・区間分解次元

Answer is...

Exercise.



Answer is...

Exercise.

逆解析

