

令和5年度茨城大学工学部情報工学科卒業研究論文

位相的データ解析を利用した画像の類似度測定

2024年2月2日

指導教員 宮本 賢伍 助教

茨城大学
工学部情報工学科所属

19T4055T 芳賀 光

位相的データ解析を利用した画像の類似度測定

芳賀 光 (19T4055T)

論文要旨

近年、機械学習の需要の高まりから、学習のために画像中の物体や背景の特徴などが類似したデータセットの効率的な分類や選定が切望されている。本論文では位相的データ解析を利用することで、画像中の物体の形状的特徴を考慮した画像間の類似度を測定する手法を提案する。パーシステントホモロジーはデータの幾何的な特徴量を計算するデータ解析手法であり、これによって得られた情報をパーシステント図に保存し、パーシステント図間のボトルネック距離を測ることが可能である。このボトルネック距離を画像の幾何的特徴を捉えた類似度とし、他の類似度指標と比較することで有効性を検証した。実験の結果から部分的な妥当性はみられるものの、実用的な指標とするためには更なる検証が必要である。

目次

第1章 序論	4
第2章 トポロジーの準備	6
2.1 単体複体	6
2.2 環上の加群	7
2.3 単体の向き	10
2.4 ホモロジー群	12
2.5 立方複体	14
第3章 パーシステントホモロジー	16
3.1 連続写像	16
3.2 ホモトピー同値	17
3.3 脈体定理	18
3.4 フィルトレーション	19
3.5 劣位集合フィルトレーション	19
3.6 画像から得られる立方複体	20
3.7 パーシステントホモロジー	20
第4章 有向グラフの表現論	23
4.1 有向グラフ	23
4.2 A_n 型有向グラフ	23
4.3 有向グラフの表現	23
第5章 パーシステントホモロジーによる類似度測定	26
5.1 パーシステントホモロジーの計算	26
5.2 ボトルネック距離	26
5.3 Homcloud	28
5.4 実験手順	28
5.5 実験結果	28
5.6 類似研究との比較	29
5.7 実験のまとめと今後の課題	34

第1章 序論

画像の類似度指標の多くは画像中の画素の信号値の比率や、構造的な関係性を評価することで計算される。広く知られている手法としてはヒストグラム比較法 [SO95] や構造的類似度法 [WBSS04] が存在する。ヒストグラム比較法は、画像の信号値から標本相関係数を求めることで画像比較を行う方法であり、構造的類似度法は小領域ごとの構造と輝度を算出し全体の平均を求めることで類似度とする比較法である。しかしこれらの類似度は、画像中に存在する物体の形状などを考慮しない。そこで本論文では、与えられたデータの図形的性質を抽出する位相的データ解析を利用することで、画像中の物体の形状を考慮した類似度指標を提案する。

位相的データ解析は、トポロジーという分野で培われた発想や道具立てをデータ解析に活かすという、実用的な面と学問的な側面を備えた分野である [大林 17]。データとは多くのサンプルの集合である。数学的な視点からのデータ解析とは、サンプル間の類似性を基に距離を導入し、つながりをネットワークで表すことでデータ集合から構造を抽出すること、そしてその構造を要約する特徴量を定めることである。位相的データ解析は、データ集合から位相的・幾何的な構造を抽出して解析する。すなわち、データの大まかな「形」に着目して解析する方法論であり、主にパーシステントホモロジーという道具により実現される [池 et al. 23, 平岡 13, 平岡 15]。パーシステントホモロジーは、主に材料科学 [平岡 19] や生命科学 [CCR13, ZLKKKW15] などの様々な分野で応用される分野である。ターゲットである図形データに対して空間的なスケールを割り当てることで、連結成分や輪といったトポロジー的な情報を抽出することが可能である。実験ではグレースケール画像データから幾何的な対象を組み合わせた的に取り出した複体を構成する。複体は白黒の濃淡を表す信号値のパラメータを持ち、パラメータに応じた複体を並べることでフィルトレーション $K_\chi: K_1 \subset K_2 \subset \dots \subset K_n$ をなす。これによりフィルトレーションが誘導するホモロジー群と線形写像の系列

$$H_k(\chi): H_k(K_1) \rightarrow H_k(K_2) \rightarrow \dots \rightarrow H_k(K_n)$$

が得られる。

一方で有向グラフの表現論 [ARS95, ASS06, S14] は、与えられた有向グラフの頂点に体 \mathbb{F} 上のベクトル空間、矢印に線形写像を対応させることで、有限次元多元環の加群のなす圏構造を明らかにしようとするものである。有向グラフのなかでは、分岐のない木である A_n 型が最も基本的である。これは n 個の頂点が一列に並び、その間に矢印が一本ずつ配置されたグラフである。位相的データ解析を行う際に現れるフィルトレーションが誘導するホモロジー群と線形写像の系列 $H_k(\chi)$ は、この A_n 型の有向グラフの表現論を用いて代数的に解析できる。有向グラフの頂点は、各ホモロジー群 H_k に、矢印はフィルトレーションが誘導する線形写像に対応させる。

有向グラフの表現論では、与えられた有向グラフの上の表現をより細かい表現に分解することでその構造論を理解する。具体的には、有向グラフの表現を「直既約な表現」に分解する（この分解は一意的である）ことで、その表現のもつ特徴を調べる（Krull–Remak–Schmidt の定理）。そこで、この直既約な表現の同型類が有限個であるときが最も基本的である。Gabriel の定理 [G72] によれば、直既約表現が有限個しかないことと、有向グラフの台グラフ構造が Dynkin 図形の非交和であることが同値であり、このとき直既約表現はグラフのルート系によってパラメトライズされる。

A_n 型の有向グラフの場合のルート系が $\{(i, j) \mid 1 \leq i \leq j \leq n\}$ であり, これらに対応する直既約表現は区間表現と呼ばれる. 従って, 系列 $H_k(\chi)$ は区間表現に一意的に分解される.

系列 $H_k(\chi)$ は区間表現への分解を与えたとき, 分解に現れる区間表現はパラメータの変化による生成, 及び消滅するホモロジー群の図形的性質と対応している [C15]. 区間表現は二次元座標上にプロットされたパーシステント図として表示され, これはパーシステントホモロジーと等価な情報を簡潔にエンコードする. 2つのパーシステント図のプロットされた点同士をマッチングさせることで, マッチした点同士の距離からパーシステント図間の差を定量的に検出するボトルネック距離を定義することができる.

本論文では位相的データ解析のソフトウェアである Homcloud を利用することで, グレースケール画像からパーシステント図を出力する. 各画像間のボトルネック距離を類似度の指標として, ヒストグラム比較法や構造的類似度法から計算された結果と比較することで, ボトルネック距離の類似度指標としての妥当性を考察する. 第 2 章では, パーシステントホモロジーの各構造を形成する要素である単体複体, 数学における代数的構造である加群について説明することで, データの連結成分や輪などのトポロジー的情報を保持したホモロジー群の概念を導入する. 第 3 章では直接計算することが難しいホモロジー群がホモトピー同値であるような単体複体から計算されること, 及びフィルトレーションによって記述されるパーシステント加群が区間加群に分解されることを説明する. 第 4 章ではフィルトレーションによって記述されるホモロジー群とその間の線形写像が有向グラフに対応すること, 有向グラフの表現論から A_n 型有向グラフが区間表現に一意的に分解されること説明する. 第 5 章ではボトルネック距離が安定性定理によってパーシステント図間の差を評価し, ボトルネック距離が画像の類似度の指標として有効であるか, 他の類似度指標も交えて考察した.

第2章 トポロジーの準備

2.1 単体複体

パーシステントホモロジーの各構造を形成する要素である単体複体について説明する. $k \in \mathbb{N}$ とする. \mathbb{R}^d 上の $k+1$ 個の点 v_0, \dots, v_k に対し, k 本のベクトル

$$v_1 - v_0, v_2 - v_0, \dots, v_k - v_0$$

が一次独立のとき, v_0, \dots, v_k は一般の位置にあるという. この定義は原点を v_0 としたとき, 他の点へ向かうベクトルが一次独立である事を表し, この概念は原点の選び方に依存しない. また集合 $C \subset \mathbb{R}^d$ に関して, C の任意の2点を端点とする線分が C に包含である, つまり任意の $x, x' \in C$ と任意の $t \in [0, 1]$ に対して $(1-t)x + tx' \in C$ を満たす時, C を凸集合であるという. 集合 S に関して, S を含む最も小さい凸集合が存在する. 実際 S を含む凸集合からなる集合族を $\mathcal{A} := \{A_i\}_{i \in I}$ で表すと, \mathbb{R}^d は S を含む凸集合のひとつであり, \mathcal{A} は空ではない. 従って, $\bigcap_{i \in I} A_i$ は S を含む凸集合であり包含関係に関して最小である. これを S の凸閉包といい $\text{Conv}(S)$ で表す.

補題 2.1.1. 一般の位置にある $k+1$ 個の点 $v_0, \dots, v_k \in \mathbb{R}^d$ に対して,

$$\text{Conv}(\{v_0, \dots, v_k\}) = \left\{ \sum_{i=0}^k t_i v_k \mid t_i \in \mathbb{R}_{\geq 0}, \sum_{i=0}^k t_i = 1 \right\}$$

が成り立つ.

一般の位置にある $k+1$ 個の点 $v_0, \dots, v_k \in \mathbb{R}^d$ の凸閉包 $\text{Conv}(\{v_0, \dots, v_k\})$ を \mathbb{R}^d 内の k 単体と呼ぶ. また点 v_0, \dots, v_k を単体の頂点と呼び, v_0, \dots, v_k の空でない部分集合 S の凸閉包 $\text{Conv}(S)$ を単体の面という. これより単体複体をここで定めた単体の集まりとして以下のように定める.

定義 2.1.2. \mathbb{R}^d 内の単体の有限集合 K が以下を満たすとき, K を単体複体と呼ぶ.

- (i) $\sigma \in K$ の任意の面は K に属する.
- (ii) $\sigma, \tau \in K$ の共通部分 $\sigma \cap \tau$ が空でなければ, これは σ と τ の両方の面である.

この定義では, \mathbb{R}^d の座標を指定していたが, 単体複体の定義にある組み合わせ構造のみを抜き出すことで, 単体複体を抽象的に扱うことが可能となる. 一般に集合 V に対して 2^V を V の部分集合全体からなる集合として, これを V の冪集合と呼ぶ. 以降では単体複体とは抽象単体複体のことを指し, 区別する場合は略称ではなく正式な名称を用いる.

定義 2.1.3. V を有限集合とする. V の部分集合の集合 $K \subset 2^V$ が以下の3条件を満たすとき, K を抽象単体複体という.

- (i) $\emptyset \notin K$

- (ii) 任意の $v \in V$ に対して $v \in K$
- (iii) $\sigma \in K$ かつ $\emptyset \neq \tau \subset \sigma$ ならば $\tau \in K$

抽象単体複体 K の元を単体と呼び, V の元を頂点という.

単体複体 K に属する単体 σ の次元を σ の要素数で定義し, $\dim(\sigma)$ で表す. また, K の次元は

$$\dim(K) := \max\{\dim(\sigma) \mid \sigma \in K\}$$

で定義する. 単体複体 K の部分集合 L が再び単体複体であるとき, L を K の部分複体という. L が K の部分複体であるとき, 定義から明らかに $\dim(L) \leq \dim(K)$ が成り立つ.

2.2 環上の加群

パーシステントホモロジーを表現するために必要な要素であるホモロジー群の理解のために, 数学における代数的構造と加群について説明する. 一般に集合 G に演算

$$G \times G \ni (a, b) \mapsto a \cdot b \in G$$

が定義され以下の条件を満たすとき, G は群であるという.

- (G1) 任意の $a, b, c \in G$ に対して, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. (結合法則)
- (G2) 任意の元 $a \in G$ に対して $a \cdot e = e \cdot a = a$ となる元 $e \in G$ が存在する. (単位元の存在)
- (G3) 任意の元 $a \in G$ に対して $a \cdot a^{-1} = a^{-1} \cdot a = e$ となる元 $a^{-1} \in G$ が存在する. (逆元の存在)

さらに群 G が $a, b \in G$ に対して次の条件

- (G4) $a \cdot b = b \cdot a$. (交換法則)

を満たすとき, G をアーベル群という. G がアーベル群のときは加法の二項演算を $+$ で表し, 単位元を 0 と表してゼロ元と呼び, $a \in G$ の逆元を $-a$ で表す. 群 G の部分集合 H が G の演算で群になっているとき, H を G の部分群という. ゼロ元からのみなる部分群を 0 で表す. 2つの群 G, G' とその間の写像 $f: G \rightarrow G'$ が任意の $x, y \in G$ に対して

$$f(xy) = f(x)f(y)$$

を満たすとき, f を G から G' への群準同型写像という. アーベル群 R に演算

$$R \times R \ni (a, b) \mapsto a \cdot b \in R$$

が定義され以下の条件を満たすとき, R は環であるという.

- (R1) 任意の $a, b, c \in R$ に対して $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. (結合法則)
- (R2) 任意の $a, b, c \in R$ に対して

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c. \quad (\text{分配法則})$$

- (R3) 任意の $a \in R$ に対して $a \cdot 1 = 1 \cdot a = a$ となる $1 \in R$ が存在する. (乗法の単位元の存在)

さらに環 R が任意の $a, b \in R$ に対して次の条件

$$(R4) \quad a \cdot b = b \cdot a \text{ (交換法則)}$$

を満たすとき, R を可換環という. 以降は環 R に関して $a \cdot b$ を単に ab とかき, R は常に可換環であるとする.

可換環 F のゼロ元を除く任意の元 $a \in F$ に対して $ab = 1$ となるような元 $b \in F$ が存在するとき, F を体という. この b を a の逆元と呼び, a^{-1} で表す. 例えば有理数全体の集合 \mathbb{Q} や実数全体の集合 \mathbb{R} は体である.

可換環 R とアーベル群 M に対して演算

$$R \times M \ni (r, x) \mapsto r \cdot x \in M$$

が定義され, 任意の $r, s \in R$ 及び任意の $x, y \in M$ に対して次の条件を満たすとき, M を R 加群という.

$$(M1) \quad r \cdot (x + y) = r \cdot x + r \cdot y$$

$$(M2) \quad (r + s) \cdot x = r \cdot x + s \cdot x$$

$$(M3) \quad (rs) \cdot x = r \cdot (s \cdot x)$$

$$(M4) \quad 1 \cdot x = x$$

R 加群 M について $r \in R$ と $m \in M$ に対する演算 $r \cdot m$ を単に rm とかく. この演算を M の R からの作用という. また係数である環 R を係数環, R の元をスカラーといい, R による作用をスカラー倍ともいう.

R 上の2つの加群 M, N とその間の群準同型写像 $f: M \rightarrow N$ が任意の $r \in R$ と $x \in M$ に対して

$$f(rx) = rf(x)$$

を満たすとき, f を M から N への R 加群同型写像という. 可換環 R が体のときを線形写像という. R 加群準同型写像 f が全単射であるとき, f は同型写像であるという. このとき M と N は同型であるといい, 記号で $M \simeq N$ と表す. R 加群 M の部分群 N が M による作用で再び R 加群となるとき, N を M の R 部分加群という. R が体のとき, R 加群を R 上のベクトル空間といい, ベクトル空間の部分加群を部分ベクトル空間という.

集合 G 上の二項関係 \equiv を考えよう.

- (a) 任意の $a \in A$ に対して, $a \equiv a$ が成り立つとき, 関係 \equiv は反射律を満たすという.
- (b) 任意の $a, b \in A$ に対して, $a \equiv b$ ならば $b \equiv a$ が成り立つとき, 関係 \equiv は対称律を満たすという.
- (c) 任意の $a, b, c \in A$ に対して, $a \equiv b$ かつ $b \equiv c$ ならば $a \equiv c$ が成り立つとき, 関係 \equiv は推移律を満たすという.

集合 G 上の二項関係 \equiv が上記の3つの条件をすべて満たすとき, それを同値関係という. 集合 G 上の同値関係 \sim があるとき, $x \in A$ と同値関係にある A の元全体

$$\bar{x} = \{y \in A \mid x \equiv y\}$$

を x の同値類といい, x を代表元という. A の同値類を全て集めた集合族を A の同値関係 \equiv に関する商集合といい, A/\equiv で表す.

$$A/\equiv := \{\bar{x} \mid x \in A\}$$

命題 2.2.1. 集合 A 上の関係 \equiv が同値関係であるとする. このとき, 次が成立する.

- (1) 任意の $x \in A$ に対して, $x \in \bar{x}$ である.
- (2) A の元 $x, y \in A$ に対して, $\bar{x} = \bar{y}$ と $x \equiv y$ は同値である.
- (3) A の元 $x, y \in A$ に対して, $\bar{x} \neq \bar{y}$ であれば, $\bar{x} \cap \bar{y} = \emptyset$ である.

証明. (1) 関係 \equiv は A 上の同値関係であるから, 任意の $x \in A$ に対して $x \equiv x$ である. よって $x \in \bar{x}$ となる.

(2) 任意の $x, y \in A$ をとる. $\bar{x} = \bar{y}$ であると仮定する. このとき, (1) より $y \in \bar{y} = \bar{x}$ なので $y \equiv x$ である. 関係 \equiv は A 上の同値関係なので $x \equiv y$ となる. 逆に $x \equiv y$ であると仮定する. 2つの集合 \bar{x} と \bar{y} が等しいことを示そう. 任意に $z \in \bar{x}$ をとれば, $x \equiv z$ であるので対称律を用いて $z \equiv x$ が成り立つ. 仮定から $x \equiv y$ だから推移律を用いて $z \equiv y$ となる. 対称律を用いて $y \equiv z$ だから $z \in \bar{y}$ である. 従って $\bar{x} \subset \bar{y}$ である. x と y の役割を入れ替えれば同様に $\bar{y} \subset \bar{x}$ が成り立つことがわかる. 以上で $\bar{x} = \bar{y}$ である.

(3) A の元 $x, y \in A$ に対して, $\bar{x} \neq \bar{y}$ であると仮定する. このとき, (2) より $x \equiv y$ とはならない. $\bar{x} \cap \bar{y} \neq \emptyset$ と仮定すると, ある元 $z \in \bar{x} \cap \bar{y}$ がとれる. このとき, $x \equiv z$ かつ $y \equiv z$ が成り立つ. 対称律と推移律を用いれば $x \equiv y$ である. これは仮定に反するので矛盾である. 従って $\bar{x} \cap \bar{y} = \emptyset$ である. \square

可換環 R 上の加群 M とその部分加群 $N \subset M$ を考える. このとき $x, y \in M$ に対して, $x - y \in N$ を \equiv を用いて $x \equiv y$ と表したとき, 二項関係 \equiv は M 上の同値関係になる. 命題 2.2.1 により $x, y \in M$ が $x - y \in N$ のとき同じグループに属するとみなすと, M はこのグループ分けにより分割される. このようなグループで全体の成す集合, すなわち M の同値関係 \equiv による商集合を M/N で表すとする. ここで $x \in M$ の属するグループを $[x] \in M/N$ で表すことにし, $[x]$ を $x \in M$ の剰余類という. $x, y \in M, a \in R$ に対して M/N は以下の演算を定めることで, R -加群になる.

$$[x] + [y] := [x + y], \quad a[x] := [ax]$$

この演算によって和とスカラー倍を定めた R -加群 M/N を, M の N による商加群という. R が体であるときは, M/N を商ベクトル空間と呼ぶ.

R 加群 M, N とその間に群準同型写像 $f: M \rightarrow N$ に対して,

$$\begin{aligned} \text{Ker}(f) &= \{x \in M \mid f(x) = 0\}, \\ \text{Im}(f) &= f(M) = \{f(x) \mid x \in M\}, \end{aligned}$$

をそれぞれ f の核, 像という. f の核は M の, f の像は N の部分加群である.

集合 X について $f: X \rightarrow X$ が任意の $x \in X$ に対して

$$f(x) = x$$

を満たすとき, これを恒等写像といい, id_X とかく. また, $N \subset M$ としその間の写像 $\iota: N \rightarrow M$ が, 任意の $n \in N$ に対して

$$\iota(n) = n$$

を満たすとき、これを包含写像という。

集合 I で添字付けられた R 加群の族 $\{M_i\}_{i \in I} = \{M_i \mid i \in I\}$ に対して、組 $(x_i)_{i \in I}$ 全体の集合を $\{M_i\}_{i \in I}$ の直積といい $\prod_{i \in I} M_i$ で表す。任意の $(x_i)_{i \in I}, (y_i)_{i \in I} \in \prod_{i \in I} M_i$ と $r \in R$ に対して、

$$(P1) \quad (x_i)_{i \in I} + (y_i)_{i \in I} = (x_i + y_i)_{i \in I}$$

$$(P2) \quad r \cdot (x_i)_{i \in I} = (rx_i)_{i \in I}$$

という加法と作用を定義することにより $\prod_{i \in I} M_i$ は R 加群となる。また直積の中で

$$\left\{ x = (x_i)_{i \in I} \in \prod_{i \in I} M_i \mid x \text{ の成分 } x_i \text{ は有限個を除き } 0 \right\}$$

を考えるとこれは $\prod_{i \in I} M_i$ の部分加群であり $\{M_i \mid i \in I\}$ の直和といい $\bigoplus_{i \in I} M_i$ で表す。

R -加群 M の部分集合 N を取ったとき、 $\langle N \rangle$ を

$$\langle N \rangle := \{r_0 n_0 + r_1 n_1 + \cdots + r_i n_i \mid i \in \mathbb{Z}_{\geq 0}, n_i \in N, r_i \in R\}$$

と定義する。これは M の部分加群となり $M = \langle N \rangle$ のとき N を M の生成系という。特に N が有限集合であるとき M を有限生成という。また N が R 上の M の生成系かつ、 N に属する有限個の要素が一次独立であるとき、 N を M の R 上の基底という。基底を持つ R 加群を自由加群という。有限生成な R 上の自由加群 M が基底 $N = \{n_1, n_2, \dots, n_k\}$ を持つとき、

$$\overbrace{R \oplus \cdots \oplus R}^k \ni (r_1, r_2, \dots, r_k) \mapsto \sum_{i=1}^k r_i n_i \in M$$

という写像を与えることができる。これは同型写像であり $\overbrace{R \oplus \cdots \oplus R}^k \simeq M$ となる。このとき、 M は階数 n の自由 R 加群と呼ばれる。

2.3 単体の向き

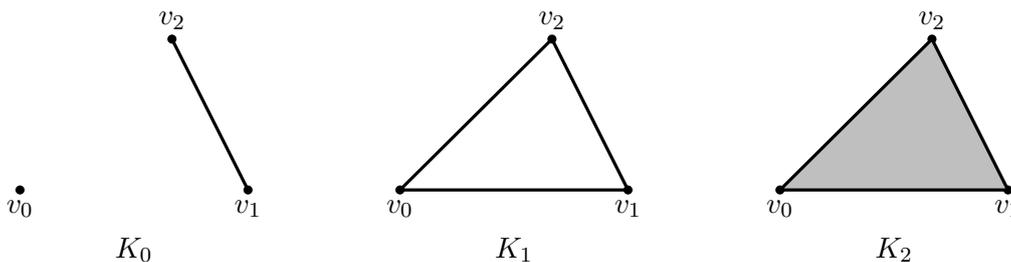
単体複体は幾何学的実現を通すことで、幾何的な対象の組み合わせとして表示され連結成分や輪といった情報を抽出できる。例えば、 $V = \{v_0, v_1, v_2\}$ として、 $K_0, K_1, K_2 \subset 2^V$ を以下で定まる集合としよう。

$$K_0 := \{\{v_0\}, \{v_1\}, \{v_2\}, \{v_1, v_2\}\}$$

$$K_1 := \{\{v_0\}, \{v_1\}, \{v_2\}, \{v_0, v_1\}, \{v_0, v_2\}, \{v_1, v_2\}\}$$

$$K_2 := \{\{v_0\}, \{v_1\}, \{v_2\}, \{v_0, v_1\}, \{v_0, v_2\}, \{v_1, v_2\}, \{v_0, v_1, v_2\}\}$$

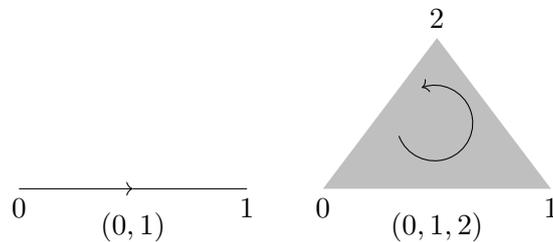
これらは、単体複体の定義を満たしている。これらは、以下のように絵で表すとわかりやすい。



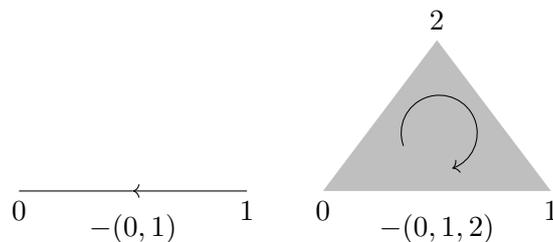
これは V の元を頂点として配置し, $\{v_i, v_j\} (i \neq j)$ を単体複体が含むならば頂点間に辺を引いてグラフをつくる. このとき K_1 や K_2 のように閉路ができる. K_1 と K_2 の違いとして $\{v_0, v_1, v_2\} \notin K_1$, $\{v_0, v_1, v_2\} \in K_2$ であるが, この違いを v_0, v_1, v_2 がつくる閉路の内部を塗ることで表現している. K_1 のような, 内部が塗られていない閉路を輪と呼ぶ. つまり幾何的な性質として K_1 は輪を持ち, K_2 は輪を持たない.

本論文で扱うパーシステントホモロジーでは, 高次元の連結成分や輪を捉えるために直感的な概念を数学的に定式化する. 以下では連結成分や輪より高次元の「穴」を定式化するホモロジーを定義するために, 単体の向きを定義しよう.

単体複体 K の頂点集合 $V(K)$ に 1 つ全順序 $<$ を固定する. $\sigma = \{v_0, v_1, \dots, v_k\} \in K$ を一つ取ると, 固定した全順序によって単体 σ の頂点を昇順に並べることができる. 並び替えを行い, 番号をつけ直すことで改めて $v_0 < \dots < v_k$ であるとしよう. こうして単体 σ を列 (v_0, v_1, \dots, v_k) に対応させることができる. これは以下の図のように単体に向きが付いていると解釈できる.



今後, 単体複体の頂点集合には全順序がひとつ与えられていると仮定する. ここで単体 σ に対して新たに代数的なシンボル $\bar{\sigma}$ を, 全順序によって昇順に並び替えられた頂点の列に対応させる. また, $\bar{\sigma} = (v_0, v_1, \dots, v_k)$ に対して, $-\bar{\sigma} := (v_k, v_{k-1}, \dots, v_0)$ と定める. $-\bar{\sigma}$ は幾何的に向きが逆になっていると解釈される.



単体の代数的シンボルを用いて, 以下のようにチェインの空間を定義する.

定義 2.3.1. \mathbb{F} を体, K を単体複体とする. $k \in \mathbb{N}$ に対して, $C_k(K; \mathbb{F})$ を k 単体の集合を形式的な基底とするような \mathbb{F} 上のベクトル空間と定める. すなわち, k 次元の単体が $\sigma_1, \dots, \sigma_n$ で尽くされるとき, これら代数的シンボルの線形和全体である:

$$C_k(K; \mathbb{F}) := \bigoplus_{\sigma \in K, \dim(\sigma)=k} \mathbb{F}\bar{\sigma} = \{\alpha_1\bar{\sigma}_1 + \dots + \alpha_n\bar{\sigma}_n \mid \alpha_i \in \mathbb{F}, \sigma_i \in K, \dim(\sigma_i) = k\}$$

$C_k(K; \mathbb{F})$ を k 次のチェインの空間といい, $C_k(K; \mathbb{F})$ の元を k 次元チェインという. また $C_{-1}(K; \mathbb{F}) = 0$ と約束する.

単体複体から幾何的な情報である輪を取り出すために、境界を取る操作としてチェインの空間の間に以下のような線形写像を定義する。

定義 2.3.2. K を単体複体とする。 $C_k(K; \mathbb{F})$ の形式的な基底ベクトル $\bar{\sigma} = (v_0, v_1, \dots, v_k)$ に対して、

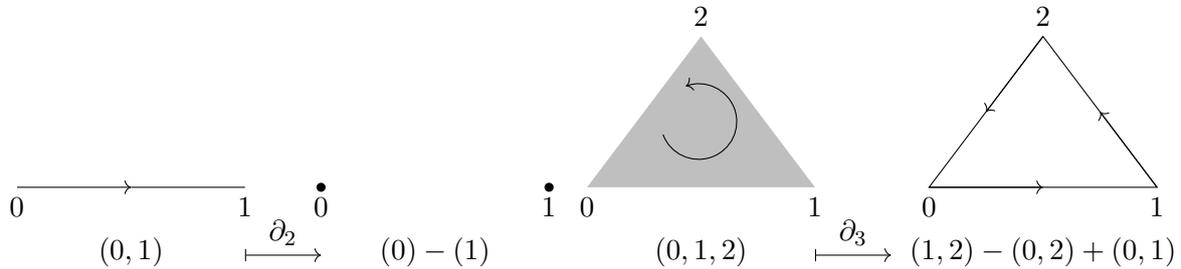
$$\partial_k(\bar{\sigma}) := \sum_{i=0}^k (-1)^i (v_0, \dots, \hat{v}_i, \dots, v_k) \in C_k(K; \mathbb{F})$$

と定める。ここで \hat{v}_i は $\bar{\sigma}$ から v_i を取り除くことを意味する。この対応を \mathbb{F} 上線形に拡張することで線形写像

$$\partial_k: C_k(K; \mathbb{F}) \longrightarrow C_{k-1}(K; \mathbb{F})$$

が定義される。このように定義した ∂_k を k 次の境界準同型という。また $\partial_0 = 0$ と約束する。

境界準同型は、単体複体 K からその境界を抽出する写像と解釈することができる。



2.4 ホモロジー群

前節では「輪」をグラフの閉路と定義したが、高次元の「穴」を定義するにはより一般の次元に対する「輪」を厳密に定義する必要がある。先の例において、1次元チェイン $(1, 2) - (0, 2) + (0, 1)$ は向き付けられた3角形 $(0, 1, 2)$ の外周とみなせるが、これを輪として捉えるための定義を与えたい。この1次元チェイン $(1, 2) - (0, 2) + (0, 1)$ に再度、境界準同型を作用させると

$$\partial_1((1, 2) - (0, 2) + (0, 1)) = (2) - (1) - (2) + (0) + (1) - (0) = 0$$

となる。この観察により、境界準同型を作用させて0になるチェイン全体が輪の候補とみなせるが、前述した単体複体 K_2 は輪がないと考えたい。1次元チェイン $(1, 2) - (0, 2) + (0, 1)$ は2単体 $(0, 1, 2)$ の境界でもあることを考えると、輪とは「境界準同型を作用させて0となるもの」で「何かの境界でないもの」と捉えることができる。この2つはそれぞれ境界準同型 ∂_n の核 $\text{Ker}(\partial_n) = \{c \in C_n(K; \mathbb{F}) \mid \partial_n(c) = 0\}$ と像 $\text{Im}(\partial_{n+1}) = \{\partial_{n+1}(c') \mid c' \in C_{n+1}(K; \mathbb{F})\}$ で定式化できることを述べていこう。

まず、境界準同型については以下の性質を満たす。

定理 2.4.1. K を単体複体として $k \in \mathbb{N}$ とする。このとき、

$$\partial_k \circ \partial_{k+1} = 0$$

が成り立つ。

証明. $k = 0$ の場合は $\partial_0 = 0$ なので等式は成立する. $k \geq 1$ のとき, $k+1$ 単体 $\sigma = \{v_0, v_1, \dots, v_{k+1}\}$ に関する $\bar{\sigma} = (v_0, v_1, \dots, v_{k+1})$ において,

$$\begin{aligned} & \partial_k \circ \partial_{k+1}(\bar{\sigma}) \\ &= \partial_k \sum_{i=0}^{k+1} (-1)^i (v_0, \dots, \hat{v}_i, \dots, v_{k+1}) \\ &= \sum_{i=0}^k \left(\sum_{j<i} (-1)^j (v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_{k+1}) + \sum_{j>i} (-1)^{j-1} (v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_{k+1}) \right) \\ &= \sum_{j<i} (-1)^{i+j} (v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_{k+1}) + \sum_{j>i} (-1)^{i+j-1} (v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_{k+1}) \\ &= 0 \end{aligned}$$

以上より主張が成り立つことが示された. \square

この定理より $\text{Im}(\partial_{n+1}) \subset \text{Ker}(\partial_n)$ となり, これらの商ベクトル空間を考えることで輪を捉えることが可能になる.

定義 2.4.2. K を単体複体とする. $n \in \mathbb{N}$ に対して以下を定義する.

$$\begin{aligned} Z_n(K; \mathbb{F}) &:= \text{Ker}(\partial_n) \subset C_n(K; \mathbb{F}), \\ B_n(K; \mathbb{F}) &:= \text{Im}(\partial_{n+1}) \subset Z_n(K; \mathbb{F}), \\ H_n(K; \mathbb{F}) &:= Z_n(K; \mathbb{F})/B_n(K; \mathbb{F}). \end{aligned}$$

$H_n(K; \mathbb{F})$ を K の n 次 \mathbb{F} 係数ホモロジー群といい, $Z_n(K; \mathbb{F})$ を n 次サイクル群, $B_n(K; \mathbb{F})$ を n 次バウンダリ群という. $Z \in H_n(K; \mathbb{F})$ は商ベクトル空間 $Z_n(K; \mathbb{F})/B_n(K; \mathbb{F})$ における剰余類 $[z]$ を z が代表する n 次ホモロジー類という. 以降は係数体 \mathbb{F} で固定されている場合 \mathbb{F} を省略して $C_n(K), H_n(K)$ と表記する.

次に部分複体と包含写像によって誘導されるホモロジー群の間の写像について考える. 単体複体 K とその部分複体 K' とする. 包含写像 $\iota: K' \hookrightarrow K$ より, この包含写像は各複体のチェインの空間 $C_k(K), C_k(K')$ の間に各 k 毎に線形写像

$$C_k(\iota): C_k(K') \longrightarrow C_k(K)$$

を誘導される. 実際 $C_k(K')$ の元は K' に属する k 単体の形式和と定めたので, これを $C_k(K)$ の元の K に属する k 単体の形式和と対応させればよい. K, K' に関する境界準同型をそれぞれ ∂_k, ∂'_k とすると, 任意の $k \in \mathbb{N}$ に対して,

$$C_{k-1}(\iota) \circ \partial'_k = \partial_k \circ C_k(\iota)$$

が成り立つ. これは次の図式が可換である事を意味する.

$$\begin{array}{ccc} C_k(K') & \xrightarrow{C_k(\iota)} & C_k(K) \\ \partial'_k \downarrow & & \downarrow \partial_k \\ C_{k-1}(K') & \xrightarrow{C_{k-1}(\iota)} & C_{k-1}(K) \end{array}$$

この $C_k(\iota)$ をチェイン写像と呼ぶ.

補題 2.4.3. K を単体複体, K' をその部分複体とする. このとき, 任意の n に対して, チェイン写像 $C_n(\iota) : C_n(K') \rightarrow C_n(K)$ は次の性質をもつ.

$$C_n(\iota)(Z_n(K')) \subset Z_n(K), \quad C_n(\iota)(B_n(K')) \subset B_n(K)$$

証明. 任意の $x \in Z_n(K')$ に対して, 境界準同型との可換性より

$$\partial_n \circ C_n(\iota)(x) = C_{n-1}(\iota) \circ \partial'_n(x) = 0$$

が成り立つ. よって $x \in Z_n(K)$ である. また, 任意の $y \in C_{n+1}(K')$ に対して, 境界準同型との可換性より

$$C_n(\iota) \circ \partial'_{n+1}(y) = \partial_{n+1} \circ C_{n+1}(\iota)(y) \in B_n(K)$$

が成り立つ. よって, $C_n(\iota)(B_n(K')) \subset B_n(K)$ であることがわかった. \square

補題 2.4.3 より, $C_n(\iota)$ は商ベクトル空間の間の線形写像

$$\iota_* : H_n(K') \longrightarrow H_n(K); \quad [x] \longmapsto [C_n(\iota)(x)]$$

を誘導する. この写像をチェイン写像 $C_k(\iota)$ による誘導準同型写像という.

2.5 立方複体

前項までは単体を基礎としてホモロジーを定義したが, 4 角形や立方体, それらの高次元版のホモロジー群を考えることで, 写真や絵イラストなどの画像を扱う際に利用できる. 立方複体に関する詳しい文献は [PR15] があるので, 証明等はそちらを参照されたい.

定義 2.5.1. $j \in \mathbb{Z}$ に対して基本区間 $I \subset \mathbb{R}$ を以下で定義する.

$$I = [j, j+1] \text{ または } \{j\} = [j] \text{ (1 点集合)}$$

また $[j, j+1]$ の基本区間を長さ 1 の区間という. \mathbb{R}^d 内の部分集合 Q が基本立方体であるとは, Q が基本区間の直積として

$$Q = I_1 \times \cdots \times I_d$$

と表されるときをいう. また基本立方体 Q の次元を以下で定める.

$$\dim Q := \#\{i \in \{1, \dots, d\} \mid I_i \text{ は長さ 1 の区間}\}$$

基本立方体の定義によってこれを基礎とした幾何学的対象として以下の立方複体を構成することができる.

定義 2.5.2. \mathbb{R}^d 上の立方複体 K とは \mathbb{R}^d 上の基本立方体からなる有限集合であり, 以下を満たすものである.

- (1) $Q \in K$ かつ $P \subset Q$ ならば $P \in K$
- (2) $P, Q \in K$ ならば $P \cap Q = \emptyset$ または $P \cap Q \in K$

立方複体の場合のチェイン空間を定めよう. \mathbb{R}^d 内の k 次元基本立方体を形式的な基底を持つ体 \mathbb{F} 上のベクトル空間を \mathcal{C}_k^d とおく. これを立方的チェイン空間と呼び, \mathcal{C}_k^d の元を k 次元立方チェインという. ただし, $\mathcal{C}_{-1}^d = 0$ と約束する. また, \mathbb{R}^d 内の立方複体 K に対して, K に属する k 次元の基本立方体全体を形式的基底とするような体 \mathbb{F} 上のベクトル空間を $\mathcal{C}_k^d(K)$ で表す. 定義より明らかに $\mathcal{C}_k^d(K) \subset \mathcal{C}_k^d$ である.

K を立方複体, $Q = I_1 \times \cdots \times I_d \in K$ を次元が k の基本立方体としよう. Q に現れる基本区間のうち, I_{i_1}, \dots, I_{i_s} ($i_1 < \cdots < i_s$) が長さ 1 の区間であって, 各 $j = 1, 2, \dots, s$ に対して, $I_{i_j} = [a_{i_j}, a_{i_j} + 1]$ であると仮定する. また, $\ell = 1, 2, \dots, n$ に対して

$$I'_\ell := \begin{cases} [a_j] & (\ell \in \{i_1, \dots, i_s\}), \\ I_\ell & (\ell \notin \{i_1, \dots, i_s\}), \end{cases} \quad I''_\ell := \begin{cases} [a_j + 1] & (\ell \in \{i_1, \dots, i_s\}), \\ I_\ell & (\ell \notin \{i_1, \dots, i_s\}) \end{cases}$$

と定義する. このとき, 新たな基本立方体 $\lambda_j^0(Q)$ と $\lambda_j^1(Q)$ を

$$\begin{aligned} \lambda_j^0(Q) &:= I_1 \times I_2 \times \cdots \times I_{j-1} \times I'_j \times I_{j+1} \times \cdots \times I_n \\ \lambda_j^1(Q) &:= I_1 \times I_2 \times \cdots \times I_{j-1} \times I''_j \times I_{j+1} \times \cdots \times I_n \end{aligned}$$

と定める. これを用いて, 境界準同型写像 $\partial : \mathcal{C}_k^d(K) \rightarrow \mathcal{C}_{k-1}^d(K)$ を

$$\partial_k(Q) := \sum_{j=1}^s (-1)^{j+1} (\lambda_j^0(Q) - \lambda_j^1(Q))$$

を \mathbb{F} 上線形に拡張することで定義する.

このように立方複体に対するチェイン空間と境界準同型写像を構成することで, 単体複体と同様にホモロジー群を構成することができる.

以降は単体複体, 立方複体のどちらにも適用される性質を論じる際にはどちらの概念もまとめたセル複体という用語を使い, その構成要素である単体と基本立方体をセルという.

第3章 パーシステントホモロジー

3.1 連続写像

前章で定義したホモロジー群を使用することで点群や画像からセル複体を作ることが可能になる。さらにそれらの構成からセル複体の増大の様子を記述したフィルトレーションの概念が得られる。その説明のために重要な概念であるホモトピー同値、及びその基本概念として連続写像について定義する。

集合 X に対して関数 $d: X \times X \rightarrow \mathbb{R}_{\geq 0}$ が以下の条件を満たすとき、 d を X 上の距離という。

(d1) 任意の $x, y \in X$ に対して $d(x, y) = 0 \Leftrightarrow x = y$ (非退化性)

(d2) 任意の $x, y \in X$ に対して $d(x, y) = d(y, x)$ (対称性)

(d3) 任意の $x, y, z \in X$ に対して $d(x, z) \leq d(x, y) + d(y, z)$ (3角不等式)

またこのとき組 (X, d) を距離空間という。ベクトル空間 \mathbb{R}^d に対して $d_E: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$ を $v = (v_1, \dots, v_d)$, $w = (w_1, \dots, w_d)$ に対して

$$d_E(v, w) = \|v - w\| := \sqrt{\sum_{i=1}^d (v_i - w_i)^2}$$

で定義すると、 (\mathbb{R}^d, d_E) は距離空間となり、距離 d_E をユークリッド距離という。距離空間 (\mathbb{R}^d, d_E) は d 次元ユークリッド空間と呼ばれる。より一般に、ベクトル空間 \mathbb{R}^d と $p \in \mathbb{R} \cup \{\infty\}$ に対して d_p を以下のように定める。

$$d_p(v, w) = \|v - w\|_p := \left(\sum_{i=1}^d |v_i - w_i|^p \right)^{\frac{1}{p}}$$
$$d_\infty = \|v - w\|_\infty := \max \{|v_1 - w_1|, |v_2 - w_2|, \dots, |v_d - w_d|\}$$

これにより $p \geq 1$ または $p = \infty$ に対して $d_p: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$ は \mathbb{R}^d 上の距離を定め、 d_p を l^p 距離という。 $p = 2$ のときはユークリッド距離である。

X をユークリッド空間 \mathbb{R}^d への部分集合とする。このとき、組 $(X, d_E|_X)$ は再び距離空間となることが容易にわかる。ここで、 $d_E|_X$ は d_E の定義域を X に制限したものである。この $d_E|_X$ を \mathbb{R}^d からの誘導距離という。本論文を通して、 \mathbb{R}^d はユークリッド空間であるとして、その部分集合には共同距離を用いて距離空間の構造が入っていると仮定する。また、 $X, Y \subset \mathbb{R}^d$ に対して、写像 $d_{X \times Y}: (X \times Y) \times (X \times Y) \rightarrow \mathbb{R}_{\geq 0}$ を、

$$d_{X \times Y}((x_1, y_1), (x_2, y_2)) := d_E|_X(x_1, x_2) + d_E|_Y(y_1, y_2)$$

で定めれば、組 $(X \times Y, d_{X \times Y})$ は距離空間となる。これを X と Y の直積距離空間という。

定義 3.1.1. $(X, d_X), (Y, d_Y)$ を距離空間, f を X から Y への写像 $f: X \rightarrow Y$ とし, $a \in X$ とする. 任意の $\varepsilon > 0$ に対してある $\delta > 0$ が存在し, 以下を満たすとき f は a で連続であるという.

$$d_X(x, a) < \delta \text{ ならば } d_Y(f(x), f(a)) < \varepsilon$$

f が任意の $a \in X$ で連続なとき, f は連続写像という.

3.2 ホモトピー同値

連続写像を用いることで幾何学的対象の粗い分類であるホモトピー同値という概念を考えることができる. ホモトピー同値は大雑把な表現として「連続的に伸ばしたり縮めることで写りあえる」を意味する同値関係である.

定義 3.2.1. $X, Y \subset \mathbb{R}^d$ とし, f, g を X から Y への連続写像とする. $x \in X$ とし, 閉区間 $I = [0, 1]$ と X との直積距離空間から Y への連続写像 $F: X \times [0, 1] \rightarrow Y$ が以下を満たすとき, f と g をホモトピックであるといい $f \simeq g$ で表す.

$$F(x, 0) = f(x), \quad F(x, 1) = g(x)$$

この定義よりホモトピックな写像は f, g は F を通じて変化量 $t \in [0, 1]$ を使うことで連続的に写りあえることになり, F を f から g へのホモトピーという.

補題 3.2.2. $X, Y \subset \mathbb{R}^d$ とし, $\text{Conti}(X, Y)$ を X から Y への連続写像全体のなす集合とする. このとき, $\text{Conti}(X, Y)$ 上の二項関係 \simeq は同値関係である.

証明. (反射律) 任意に $f \in \text{Conti}(X, Y)$ をとる. このとき, F を

$$F(t, x) = f(x) \quad (t \in [0, 1], x \in X)$$

と定義すれば, f が連続写像であるから F も連続写像である. よって, $f \simeq f$ が成り立つ.

(対称律) $f, g \in \text{Conti}(X, Y)$ が $f \simeq g$ であるとする. 仮定より f から g へのホモトピー $F: [0, 1] \times X \rightarrow Y$ が存在する. そこで, 写像 $G: [0, 1] \times X \rightarrow Y$ を

$$G(t, x) = F(1 - t, x)$$

と定義すれば, F が連続写像なので G も連続写像である. このとき, 任意の $x \in X$ に対して

$$G(0, x) = F(1, x) = g(x), \quad G(1, x) = F(0, x) = f(x)$$

だから, G は g から f へのホモトピーである. よって, $g \simeq f$ が成り立つ.

(推移律) $f, g, h \in \text{Conti}(X, Y)$ が $f \simeq g$ かつ $g \simeq h$ であるとする. 仮定より f から g へのホモトピー $F: [0, 1] \times X \rightarrow Y$ と g から h へのホモトピー $G: [0, 1] \times X \rightarrow Y$ が存在する. そこで, 写像 $H: [0, 1] \times X \rightarrow Y$ を

$$H(t, x) = \begin{cases} F(2t, x) & (0 \leq t \leq \frac{1}{2}) \\ G(2t - 1, x) & (\frac{1}{2} \leq t \leq 1) \end{cases}$$

と定義すれば, F, G が連続写像であり, $t = \frac{1}{2}$ での値は等しいから H も連続写像である. \square

定義 3.2.3. $X, Y \subset \mathbb{R}^d$ に対して, X と Y をホモトピー同値であるとは, ある連続写像 $f: X \rightarrow Y$, $g: Y \rightarrow X$ で

$$g \circ f \simeq \text{id}_X, f \circ g \simeq \text{id}_Y$$

を満たすものが存在するときをいう. X と Y がホモトピー同値であるとき, 記号で $X \simeq Y$ で表す. ホモトピー同値という関係は同値関係である.

3.3 脈体定理

ホモロジー群に関して $X, Y \subset \mathbb{R}^d$ がホモトピー同値 $X \simeq Y$ であるとき, そのホモロジー群は同型であることが知られている (例えば, [柘田 02] などを参照されたい). このように, $X \subset \mathbb{R}^d$ のある数学的な量や対象 e_X があつたとき, $X \simeq Y$ ならば e_X と e_Y が本質的に同一視できるとき, このような量をホモトピー不変量という. ホモロジー群はホモトピー不変量のひとつの例である. 直接ホモロジー群を計算することが困難な幾何学的対象に対して, ホモトピー同値であるような単体複体を構成できれば計算機を使ってホモロジー群を計算できる. そこで球の集まりとホモトピー同値になる単体複体を構成することで, 点群データなどへの応用が可能となる. 脈体及び脈体定理はそのための重要な概念である.

定義 3.3.1. 空でない部分集合 $X \subset \mathbb{R}^d$ が有限個の部分集合の集まり $\Phi = \{B_i \subset \mathbb{R}^d \mid i = 1, \dots, m\}$ を用いて以下のように表せるとき, X が Φ で被覆されているという.

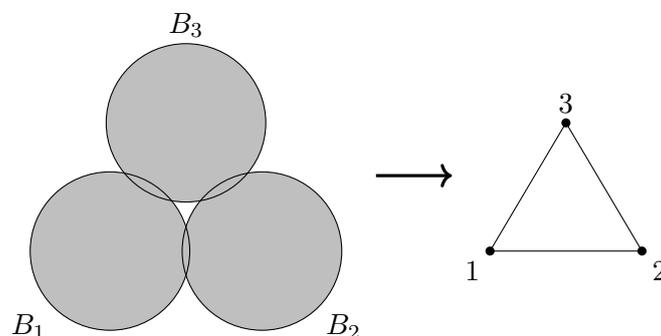
$$X = \bigcup_{i=1}^m B_i$$

頂点集合 $V = \{1, \dots, m\}$ とする. これに対して, $N(\Phi)$ を

$$N(\Phi) = \left\{ \{i_0, \dots, i_k\} \mid \bigcap_{j=0}^k B_{i_j} \neq \emptyset \right\}$$

で定義すれば, $N(\Phi)$ は抽象単体複体となり, これを Φ の脈体という.

以下の図は凸集合の集まり $X = B_1 \cup B_2 \cup B_3$ が構成する脈体の例であり, これらはホモトピー同値となる.



この例は脈体定理によって一般化される.

定理 3.3.2 (脈体定理). 集合 $X \subset \mathbb{R}^d$ が凸閉集合の有限個の集まり $\Phi = \{B_i \mid i = 1, \dots, m\}$ で被覆されているとする. このとき X と脈体 $N(\Phi)$ はホモトピー同値となる.

3.4 フィルトレーション

半径 r のような実数パラメータを設定する事で、それに応じて異なるセル複体を得られる。しかし実用的な利用を考えたとき、実数で添え字付けられた複体の族に対して特徴量を記述することがパーシステントホモロジーの目的である。パラメータ変化による複体の幾何的な情報の遷移や持続性を扱うためにフィルトレーションという増大列を定義する。

定義 3.4.1. K をセル複体とする。 $\tau, \sigma \in K$ において、各セルに実数 \mathbb{R} を対応させる関数 $\phi: K \rightarrow \mathbb{R}$ が $\tau \subset \sigma$ ならば $\phi(\tau) \leq \phi(\sigma)$ を満たすとき、関数 ϕ を K のフィルトレーションという。

セル複体 K のフィルトレーション ϕ が与えられたとき、各 $r \in \mathbb{R}$ で添え字付けられた K の部分複体 K_r を以下のように定める。

$$K_r := \{\sigma \in K \mid \phi(\sigma) \leq r\}$$

このとき部分複体 K_r の族 $(K_r)_{r \in \mathbb{R}}$ は以下の条件を満たす。

(K1) $r \leq s$ ならば $K_r \subset K_s$

(K2) $\bigcup_{r \in \mathbb{R}} K_r = K$

この部分複体の族も K のフィルトレーションという。これは有限セル複体のフィルトレーションがとりうる値が有限個なため、 ϕ の像である有限集合を R とし R での部分複体 K_r ($r \in R$) を表示するだけでフィルトレーションが決定されるからである。つまり ϕ のとりうる値を小さな値から並べたものを $\{r_0, \dots, r_T\}$ とすると以下を満たす。

(r1) $s \leq r_0$ ならば $K_s = \emptyset$

(r2) $r_i \leq s < r_{i+1}$ ならば $K_{r_i} = K_s \subset K_{r_{i+1}}$

(r3) $r_T \leq s$ ならば $K_{r_T} = K_s$

これより添え字付けられた部分複体の増大列

$$K_{r_0} \subset \dots \subset K_{r_i} \subset \dots \subset K_{r_T} = K$$

に全ての情報が含まれている。

3.5 劣位集合フィルトレーション

K をセル複体とする。0次元上の集合 V に定義された実数値関数 $f: V \rightarrow \mathbb{R}$ が与えられたとき、 K のフィルトレーションが以下により定まる。

$$\phi(\sigma) := \max_{v \in \sigma} f(v)$$

このフィルトレーション ϕ を f に関する劣位集合フィルトレーションという。

一方それ自身以外には包含されないような極大なセルの集合を F とする。 F 上に定義された実数値関数 $g: F \rightarrow \mathbb{R}$ が与えられたとき、 K のフィルトレーションが以下で定まる。

$$\phi(\sigma) := \min_{\tau \in F, \sigma \subset \tau} g(\tau)$$

このフィルトレーション ϕ を f に関する優位集合フィルトレーションという。

3.6 画像から得られる立方複体

画像からセル複体とするフィルトレーションを構成する方法は主に2種類存在する. ここでは画像は画素ごとに1つの画素値を持つモノクロ諧調画像を対象とする. $W_1, W_2 \in \mathbb{N}_{>0}$ とし拡張実数を $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ で表すと2次元画像は以下の関数で表される.

$$I: \{0, 1, \dots, W_1 - 1\} \times \{0, 1, \dots, W_2 - 1\} \longrightarrow \bar{\mathbb{R}}$$

これは座標 $(x, y) \in \{0, 1, \dots, W_1 - 1\} \times \{0, 1, \dots, W_2 - 1\}$ に対してその画素値 $f(x, y) \in \bar{\mathbb{R}}$ を対応させる関数とみなせる. この画像に対して T 構成と V 構成を定める.

定義 3.6.1. T 構成は画素を正方形のマスと捉え, $W_1 \times W_2$ 個のマスからなる正方格子をセル複体とみなす. フィルトレーションは2セル(正方形)では I の値, 1セル(辺)と0セル(頂点)ではそれを含む正方形のフィルトレーション値の最小値を対応させる優位集合フィルトレーションである. ピクセルの連結性は上下左右と斜めも加えた8近傍を考える.

定義 3.6.2. V 構成は画素を格子点と捉え, $W_1 \times W_2$ 個の頂点と $(W_1 - 1) \times (W_2 - 1)$ 個のマスからなる正方格子をセル複体とみなす. フィルトレーションは0セル(頂点)では I の値, 1セル(辺)と2セル(正方形)ではそれが含む頂点のフィルトレーション値の最大値を対応させる劣位集合フィルトレーションである. ピクセルの連結性は上下左右の4近傍を考える.

この2つは双対の関係にあり, 他方からもう一方のパーシステントホモロジーを計算することもできる. 実用上は差がないとされており, 使用されているソフトウェアによって使用できる構成は異なるため注意が必要である.

3.7 パーシステントホモロジー

セル複体のフィルトレーション $(K_r)_{r \in \mathbb{R}}$ は空間の増大の様子を記述したが, パーシステントホモロジーは各 K_r のホモロジーを r の相互関係も含めて計算する. まずパーシステントホモロジーと等価な情報を簡潔にエンコードするパーシステント図について説明し, その後パーシステントホモロジーをホモロジーの列として導入する. 以下の定理はパーシステントホモロジーの構造定理というパーシステントホモロジーにおいて最も重要な定理の1つである.

定理 3.7.1. $(K_r)_{r \in \mathbb{R}}$ をセル複体 K のフィルトレーションとする. n 次サイクル $z \in Z_n(K_r)$ の $H_r(K_r)$ におけるホモロジー類を $[z]_r$ と表す. このとき n 次元チェイン $\{z_i \in C_n(K)\}_{i=1}^m$ と拡張実数 $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ の多重集合 $\{(b_i, d_i)\}_{i=1}^m$ であって以下を満たすものが存在する.

- (PH1) $r < b_i$ ならば $z_i \notin Z_n(K_r)$
- (PH2) $b_i \leq r$ ならば $z_i \in Z_n(K_r)$
- (PH3) $b_i \leq r < d_i$ ならば $[z]_r \neq 0$, すなわち $z_i \notin B_n(K_r)$
- (PH4) $d_i \leq r$ ならば $[z]_r = 0$, すなわち $z_i \in B_n(K_r)$
- (PH5) $\{[z_i]_r \mid i \text{ は } b_i \leq r < d_i \text{ 全体を満たすもの全体}\}$ は $H_n(K_r)$ の基底

条件 (1) は $r < b_i$ ならば $z_i \notin (K_r)$ であるか K_r のサイクルに属さないこと, (2) は $r = b_i$ で z_i が K_r のサイクルになることを意味する. (3) は z_i が $b_i \leq r < d_i$ の間は 0 でないホモロジーの元であることを, (4) はそれが $r = d_i$ で消滅することを表す. (5) は $\{(b_i, d_i)\}_{i=1}^m$ と $\{z_i\}_{i=1}^m$ がフィルトレーション $(K_r)_{r \in \mathbb{R}}$ のホモロジーの情報を相互関係を含めて完全に保持することを保証する.

このときの拡張実数 $\bar{\mathbb{R}}$ のペア $\{(b_i, d_i)\}_{i=1}^m$ を生成消滅対といい, 生成消滅対を集めた多重集合 $D_n((K_r)_{r \in \mathbb{R}}) = \{(b_i, d_i)\}_{i=1}^m \subset \bar{\mathbb{R}}^2$ をフィルトレーション $(K_r)_{r \in \mathbb{R}}$ の n 次パーシステント図という. また各生成消滅対または z_i に対して $d_i - b_i$ を寿命という.

フィルトレーション $(K_r)_{r \in \mathbb{R}}$ は r によって添え字づけられた包含な関係の族であった. フィルトレーション $K_r \subset K_s$ 間の包含な写像を $\iota_{s,r}$ とすると, これはホモロジー群間の準同型写像 ι_{s,r_*} を誘導する. この誘導準同型写像を用いて以下を定義する.

定義 3.7.2. $(K_r)_{r \in \mathbb{R}}$ をセル複体 K のフィルトレーションとする. このとき $n \in \mathbb{N}$ に対して n 次のホモロジー群と包含写像が誘導する線形写像の組 $(H_n(K_r))_{r \in \mathbb{R}}, (\iota_{s,r_*})_{r \leq s}$ を $(K_r)_{r \in \mathbb{R}}$ の n 次パーシステントホモロジーという.

この定義より任意の $r \leq s \leq t$ で添え字付けられたホモロジー群について, 線形写像の族とは線形写像が誘導する準同型写像が以下の図式を可換とするような無限の列と解釈できる.

$$\begin{array}{ccccc} H_n(K_r) & \xrightarrow{\iota_{s,r_*}} & H_n(K_s) & \xrightarrow{\iota_{t,s_*}} & H_n(K_t) \\ & & & \searrow & \\ & & & \text{---} \iota_{t,s_*} \text{---} & \end{array}$$

フィルトレーション $(K_r)_{r \in \mathbb{R}}$ を K_r の変化が起こる \mathbb{R} の有限部分 $R = \{r_0, \dots, r_T\}$ で添え字付けられた部分複体の増大族 $(K_r)_{r \in R}$ としても表示できた. $r_{-1} = -\infty, r_{T+1} = \infty$ とすると $r_i \leq s < r_{i+1}$ ならば $H_n(K_s) = H_n(K_{r_i})$ であり, $r_i \leq s \leq t < r_{i+1}$ ならば $\iota_{t,s_*} = \text{id}_{H_n(K_s)}$ である. したがって $r \leq s$ である $r, s \in R$ についてのみ $H_n(K_r)$ と ι_{t,s_*} を表示すれば \mathbb{R} 全体で添え字付けられた $(K_r)_{r \in \mathbb{R}}$ のパーシステントホモロジーの情報を記述でき, 以下の系列を得る.

$$H_n(K_{r_0}) \rightarrow \cdots \rightarrow H_n(K_{r_i}) \rightarrow \cdots \rightarrow H_n(K_{r_T})$$

これはパーシステントホモロジーの抽象化であるパーシステント加群として定義される.

定義 3.7.3. R を \mathbb{R} の部分集合とする. R で添え字づけられたベクトル空間の族 $(V_r)_{r \in R}$ と $r \leq s$ である任意の $r, s \in R$ に対して定義された線形写像 $V_s, r: V_r \rightarrow V_s$ の族 $(V_s, r)_{r \leq s}$ の組 $\mathbf{V} = ((V_r)_{r \in R}, (V_s, r)_{r \leq s})$ が以下の条件を満たすとき, \mathbf{V} をパーシステント加群という. R 上のパーシステント加群 \mathbf{V} について, 任意の $r \in R$ に対して V_r が有限次元であるとき \mathbf{V} は各点で有限次元であるという.

$$(V1) \quad V_{r,r} = \text{id}_{V_r}$$

$$(V2) \quad r \leq s \leq t \text{ に対して } V_{t,s} = V_{t,s} \circ V_{s,r}$$

R 上の 2 つのパーシステント加群 $\mathbf{V} = ((V_r)_{r \in R}, (V_s, r)_{r \leq s}), \mathbf{W} = ((W_r)_{r \in R}, (W_s, r)_{r \leq s})$ に対してパーシステント加群の間の線形写像を $f_r: V_r \rightarrow W_r$ とし, その族を $F = (f_r)_{r \in R}$ で表す. $r \leq s$ である任意の $r \in R$ に対して $f_s \circ V_{s,r} = W_{s,r} \circ f_r$ が成り立つ, つまり以下の図式が可換になるとき F をパーシステント加群の射という.

$$\begin{array}{ccc}
 V_r & \xrightarrow{V_{s,r}} & V_s \\
 f_r \downarrow & & \downarrow f_s \\
 W_r & \xrightarrow{W_{s,r}} & W_s
 \end{array}$$

射 $\text{id}_V: V \rightarrow V$ を $\text{id}_V = (\text{id}_{V_r})_{r \in R}$ により定めて V の恒等射という. R 上の2つのパーシステント加群の射 $F = (f_r)_{r \in R}: V \rightarrow W$, $G = (g_r)_{r \in R}: U \rightarrow V$ に対してその合成は $F \circ G = (f_r \circ g_r)_{r \in R}$ で定義される.

R 上の2つのパーシステント加群 V, W に対してパーシステント加群の射 $F = (f_r)_{r \in R}: V \rightarrow W$, $G = (g_r)_{r \in R}: W \rightarrow V$ が存在して $G \circ F = \text{id}_V$ かつ $F \circ G = \text{id}_W$ を満たすとき V と W は同型であり $V \simeq W$ と表す.

3.7節ではパーシステントホモロジーのホモロジーの相互関係の情報を保持する生成消滅対, それを集めた多重集合をパーシステント図といった. パーシステント加群においても一定の条件を満たしたパーシステント加群が単純なブロックに分解されることでパーシステント図が定義される. 分解については有向グラフの章で説明するため, ここではパーシステント加群の直和と区間加群のみ定義する.

定義 3.7.4. $R \subset \mathbb{R}$ とする. A で添え字づけられた R 上のパーシステント加群 $(V_\alpha)_{\alpha \in A}$, $V_\alpha = ((V_\alpha)_r)_{r \in R}$, $((V_\alpha)_{s,r})_{r \leq s}$ に対して以下を定める.

$$V_r := \bigoplus_{\alpha \in A} (V_\alpha)_r, \quad V_{s,r} := \bigoplus_{\alpha \in A} (V_\alpha)_{s,r}$$

これにより R 上のパーシステント加群 V が得られる. これを $(V_\alpha)_{\alpha \in A}$ の直和といい, $\bigoplus_{\alpha \in A} V_\alpha$ で表す.

定義 3.7.5. 空集合, \mathbb{R} 全体, または $b \leq d$ を満たす $b, d \in \mathbb{R}$ によって $[b, d]$, $[b, d)$, $(b, d]$, (b, d) または $[b, \infty)$, (b, ∞) , $(-\infty, d]$, $(-\infty, d)$ の形で書ける \mathbb{R} の部分集合を \mathbb{R} の区間といい, \mathbb{R} の区間と R の共通集合のことを R の区間という. I を R の区間とし R 上のパーシステント加群 $\mathbb{F}_I = (((\mathbb{F}_I)_r)_{r \in R}, ((\mathbb{F}_I)_{s,r})_{r \leq s})$ を以下で定める.

$$(\mathbb{F}_I)_r := \begin{cases} \mathbb{F} & (r \in I) \\ 0 & (\text{それ以外}), \end{cases} \quad (\mathbb{F}_I)_{s,r} := \begin{cases} \text{id}_{\mathbb{F}} & (r, s \in I) \\ 0 & (\text{それ以外}) \end{cases}$$

\mathbb{F}_I を I の区間加群という. パーシステント加群の同型の定義より $\mathbb{F}_I \simeq \mathbb{F}_J$ が成り立つことと $I = J$ であることは同値である.

第4章 有向グラフの表現論

4.1 有向グラフ

有効グラフについて説明する. 有向グラフは以下のように定義される.

定義 4.1.1. 2つの集合 Q_0, Q_1 とその間の写像 $s, t: Q_1 \rightarrow Q_0$ からなる4つ組 $Q = (Q_0, Q_1, s, t)$ を有向グラフという. 集合 Q_0 の元を Q の頂点, 集合 Q_1 の元を Q の矢印という. また各矢 $\alpha \in Q_1$ に対して $S(\alpha)$ を矢 α の始点, $t(\alpha)$ を矢 α の終点という. また矢印 α を以下で表す

$$\alpha: s(\alpha) \longrightarrow t(\alpha)$$

有向グラフ Q の頂点 Q_0 と矢印 Q_1 が有限集合であるとき, Q は有限であるという. 本論文では常に Q は有限とする.

4.2 A_n 型有向グラフ

正の自然数 $n \in \mathbb{N}_{>0}$ に対して長さ n の向き τ を, 記号「 f 」または記号「 b 」を元とする長さ $n-1$ の列 $\tau = (\tau_1, \dots, \tau_{n-1})$ をとる. この τ に対して有向グラフ $A_n(\tau)$ を以下で定める.

$$A_n(\tau): 1 \overset{\alpha_1}{\longleftarrow} 2 \overset{\alpha_2}{\longleftarrow} \dots \overset{\alpha_{n-1}}{\longleftarrow} n$$

ただし i 番目の矢 α_i の方向は次のように τ_i で決める. $\tau_i = f$ ならば $\alpha_i: i \rightarrow i+1$, $\tau_i = b$ ならば $\alpha_i: i \leftarrow i+1$ とする. $A_n(\tau)$ を A_n 型有向グラフという. また $\tau = (f, \dots, f)$ の場合 $A_n(\tau)$ を \vec{A}_n と表す.

4.3 有向グラフの表現

定義 4.3.1. \mathbb{F} を体, Q を有向グラフとする. このとき頂点集合 Q_0 で添え字づけられた \mathbb{F} ベクトル空間の族を $(V(i))_{i \in Q_0}$, 矢の集合 Q_1 で添え字づけられた線形写像 $V(\alpha): V(s(\alpha)) \rightarrow V(t(\alpha))$ の族を $(V(\alpha))_{\alpha \in Q_1}$ とすると, 2つの組 $V = ((V(i))_{i \in Q_0}, (V(\alpha))_{\alpha \in Q_1})$ を有向グラフ Q の \mathbb{F} 表現という.

表現 V の全てのベクトル空間 $V(i)$ が有限次元のとき, V は有限次元表現という. 各 $i \in Q_0$ に対して $V(i) = 0$ を満たす有向グラフ Q の表現を零表現といい, 単に 0 で表す.

例えば $R = \{1, 2, \dots, n\}$ について R 上のパーシステント加群は $\vec{A}_n = A_n(f \cdots f)$ の表現に他ならず, 各点で有限次元な R 上のパーシステント加群は \vec{A}_n の有限次元表現である. A_n 型有向グラフ $A_n(\tau)$ の表現をジグザグパーシステント加群という.

定義 4.3.2. Q を有向グラフとし, $V = (V(i), V(\alpha)), V' = (V'(i), V'(\alpha))$ を Q の2つの表現とする. 表現 V から V' への射 $f: V \rightarrow V'$ とは, 任意の $i \xrightarrow{\alpha} j \in Q_1$ に対して,

$$f_j \circ V(\alpha) = V'(\alpha) \circ f_i$$

となるような線形写像の組 $(f_i)_{i \in Q_0}$ である。以下、簡単のため、表現の射 $(f_i)_{i \in Q_0}$ を (f_i) と略記する。もし、各 f_i が全単射であれば、 V と V' は同型とよばれる。このとき、 $V \simeq V'$ と表す。表現 V に同型なすべての表現を集めたクラスを V の同型類と呼ぶ。

例えば、次のような例を考えよう。 Q を有向グラフ $1 \rightarrow 2$ とする。 Q の表現 V, V' を次で定義する。

$$V: \mathbb{F} \xrightarrow{1} \mathbb{F}, \quad V': k \xrightarrow{0} 0$$

という表現に対して、その間の射 $f = (f_1, f_2): V \rightarrow V'$ を考えよう。このとき、 f_1 は $a \in k$ に対するスカラー倍写像で f_2 は零写像となる。

$$\begin{array}{ccc} V & & \mathbb{F} \xrightarrow{1} \mathbb{F} \\ \downarrow f & & \downarrow f_1 \quad \downarrow f_2 \\ V' & & \mathbb{F} \longrightarrow 0 \end{array}$$

Q を有向グラフとする。 Q の表現 $V = (V(i), V(\alpha)), M' = (V'(i), V'(\alpha))$ を Q の表現とすると、

$$V \oplus V' := \left(V(i) \oplus V'(i), \begin{bmatrix} V(\alpha) & 0 \\ 0 & V'(\alpha) \end{bmatrix} \right)_{i \in Q_0, \alpha \in Q_1}$$

を V と V' の直和という。

定義 4.3.3. 有向グラフ Q 表現 V が零表現ではなく、 V が2つの非零表現の直和として表せない場合、すなわち、 Q の表現 V, V'' が $V \simeq V' \oplus V''$ であるとき、 $V' = 0$ または $V'' = 0$ が成り立つとき、 V は直既約であると呼ばれる。

直既約表現は「それ以上分解できない」ような表現のことである。直既約表現が同型を除いて有限個であるとき、有向グラフ Q は有限表現型という。

以下の主張によって、有向グラフの表現は直既約なものに分解されるので、直既約な表現を考察すれば十分であることがわかる。

定理 4.3.4 (Krull–Remak–Schmidt の定理). 有向グラフ Q の任意の有限次元表現 $V = ((V(i))_{i \in Q_0}, (V(\alpha))_{\alpha \in Q_1})$ は同型を除いて一意的な直既約表現の直和として表される。つまり以下の条件を満たす。

- (1) V はいくつかの直既約表現 $(V_1(i), V_1(\alpha)), \dots, (V_n(i), V_n(\alpha))$ の直和で表される。

$$(V(i), V(\alpha)) = (V_1(i), V_1(\alpha)) \oplus (V_2(i), V_2(\alpha)) \oplus \cdots \oplus (V_n(i), V_n(\alpha))$$

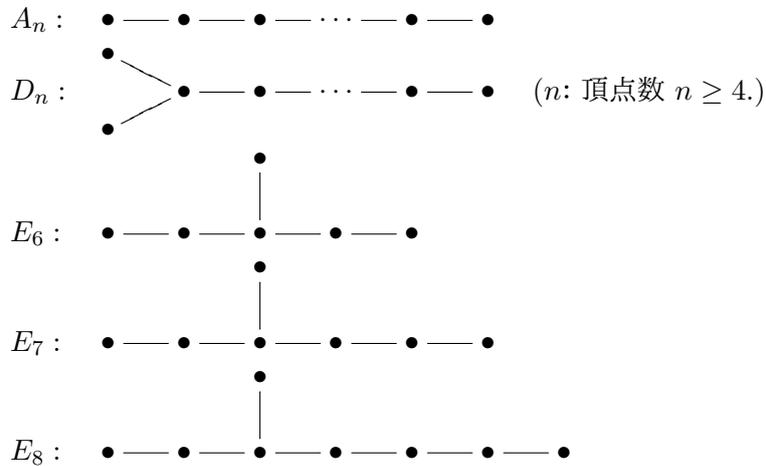
- (2) 2通りの直既約表現の直和

$$\begin{aligned} (V(i), V(\alpha)) &= (V_1(i), V_1(\alpha)) \oplus (V_2(i), V_2(\alpha)) \oplus \cdots \oplus (V_n(i), V_n(\alpha)) \\ &= (V'_1(i), V'_1(\alpha)) \oplus (V'_2(i), V'_2(\alpha)) \oplus \cdots \oplus (V'_m(i), V'_m(\alpha)) \end{aligned}$$

で表される時 $n = m$ であり、適当に順序をいれかえることで自己への同型な射 $\text{id}_{V(i)}$ を用いて以下を満たす。

$$\text{id}_{V(i)}(V_j(i), V_j(\alpha)) = (V'_j(i), V'_j(\alpha)) \quad (1 \leq j \leq n)$$

Krull–Remak–Schmidt の定理によって、有向グラフ Q の非零表現は直既約表現の直和で表される。従って、直既約表現の全体が有限個しかない場合が基本的である。そのような有向グラフ Q は Gabriel によって 1980 年代に与えられており、以下に表す有限 **Dynkin** 図形と呼ばれるもので特徴付けがされる。



特に A 型の有向グラフに関するジグザグパーシステント加群 ($A_n(\tau)$ の表現) について、次の定理は重要である。

定理 4.3.5 (Gabriel の定理). τ を任意の向きとすると、有向グラフ $A_n(\tau)$ の直既約表現は同型を除いて以下のように定義される区間表現 $I[b, d]$ ($b \leq d \in \{1, 2, \dots, n\}$) となる。

$$\begin{aligned}
 i \in Q_0 \text{ に対して } I[b, d](i) &:= \begin{cases} \mathbb{F} & (i \in \{b, \dots, d\}), \\ 0 & (\text{それ以外}), \end{cases} \\
 \alpha \in Q_1 \text{ に対して } I[b, d](\alpha) &:= \begin{cases} \text{id}_{\mathbb{F}} & (s(\alpha), t(\alpha)) \in \{b, \dots, d\}, \\ 0 & (\text{それ以外}). \end{cases}
 \end{aligned}$$

これより $A_n(\tau)$ は有限表現型であり、任意のジグザグパーシステント加群 V は区間表現の直和 $V \simeq \bigoplus_{i=1}^N I[b_i, d_i]$ で一意的に表される。したがってジグザグパーシステント加群に対してパーシステント図を定義することができ、同様に各点で有限次元な R 上のパーシステント加群も区間加群に一意的に分解されパーシステント図を定義することができる。

第5章 パーシステントホモロジーによる類似度測定

5.1 パーシステントホモロジーの計算

画像データから複体とフィルトレーションを構成する手法は3.6章で説明したが、実際の画像データからパーシステントホモロジーを計算するためには画像データが1つの画素値を持った画像である必要がある。本論文ではカラー画像をグレースケールに変換したものを対象とする。使用する構成は V 構成を用い、フィルトレーションの値を白黒の濃淡を示す画素値と対応させる。そしてグレースケール画像から得られたパーシステント図を比較し距離を求めることで、画像同士の類似度とする。パーシステントホモロジーは形の特徴量を計算することが可能であり、この性質は画素値によって構成される被写体の形状や濃淡を考慮した、新しい類似度の指標となるのではないかと考察する。

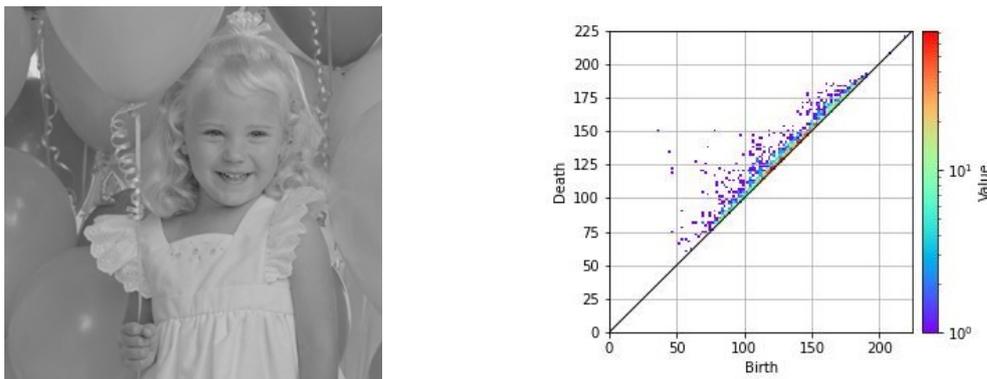


図 5.1: グレースケール画像と生成されたパーシステント図

5.2 ボトルネック距離

パーシステント図間の差を定量的に検出するためにパーシステント図の空間に距離を定義する。2つのパーシステント図の点同士を互いにマッチングさせ、2点間の距離の最大値を考える。距離の最大値をマッチングのコストとし、コストが最小になるようにマッチングを選んだときのコストをパーシステント図間の距離と定義したい。この定義の課題として以下が挙げられる。

- (1) 2つのパーシステント図の点の個数が等しいとは限らない
- (2) 重要度が異なると想定される対角線に近い点のと遠い点とがマッチングされる可能性がある。

そこで点同士を部分的にマッチングさせ、残りの点是对角線上の存在する点とマッチングすると考える。以下パーシステント図の点を $q = (b, d)$ とし、 \mathbb{R}^2 上の l^∞ 距離を $\bar{\mathbb{R}}^2$ に拡張して、 $q = (b, d), q' = (b', d')$ に対して次のように定める。

$$d_\infty(q, q') = \|q - q'\| := \max\{|b - b'|, |d - d'|\}$$

ここで $a, a' \in \bar{\mathbb{R}}$ に対して以下を約束する。

$$|a - a'| := \begin{cases} |a - a'| & (a, a' \in \mathbb{R}) \\ 0 & (a = \infty, a' = \infty \text{ または } a = -\infty, a' = -\infty) \\ \infty & (\text{それ以外}) \end{cases}$$

$$c \in \mathbb{R} \cup \{\infty\} \text{ に対して } \max\{\infty, c\} = \infty$$

定義 5.2.1. D, D' を 2 つのパーシステント図とする。 D, D' の間の部分マッチングを以下を満たすような $D \times D'$ の部分集合 M で定める。

- (1) 任意の $q \in D$ に対して $(q, q') \in M$ を満たす $q' \in D'$ は高々 1 つ。
- (2) 任意の $q' \in D'$ に対して $(q, q') \in M$ を満たす $q \in D$ は高々 1 つ。

このとき部分マッチング M は $M: D \leftrightarrow D'$ と表す。 $(q, q') \in M$ となる対をマッチングされた対といい、マッチングされた対の成分にならない D と D' の非交和 $D \sqcup D'$ の点をマッチングされない点という。

定義 5.2.2. $q = (b, d) \in D, q' = (b', d') \in D'$ とする。 D と D' の間の部分マッチング $M: D \leftrightarrow D'$ に対して、そのコストを $c(M)$ とし以下で定める。

$$c(M) := \max \left\{ \sup_{(q, q') \in M} \|q - q'\|_\infty, \sup_{q \in D \sqcup D'} \frac{|d - b|}{2} \right\}$$

また q に対して q からパーシステント図の $b = d$ である対角線へと直交するように射影した時の点を $\pi(q)$ とするとき、 $\frac{|d-b|}{2}$ は $\|q - \pi(q)\|_\infty$ と等しい。

これらの定義から、 D から D' との間の距離 $d_B(D, D')$ を以下で定める。

$$d_B(D, D') := \inf_{M: D \leftrightarrow D'} c(M)$$

ここで M は D と D' の間の部分マッチング全体を渡る。この $d_B(D, D')$ をボトルネック距離という。ボトルネック距離は直感的にはマッチングの中で最も離れた対の距離をはかっている、つまりボトルネック距離は 1 点の差のみで決まるため、その意味での頑健性はない。ただしパーシステントホモロジーはデータの「大まかな形」、すなわちトポロジーを抽出して出力するため、入力データのある種の変化に対しては出力があまり変わらずにノイズに頑健であることが期待される。これは数学的な定理として述べるのが安定性定理であり、パーシステント図間のボトルネック距離が入力データの何らかの距離で上から評価される、という形で主張される。安定性定理の証明に関する詳しい文献は [CEH07] があるので、証明等はそちらを参照されたい。

定理 5.2.3 (劣位集合フィルトレーションに関する安定性定理). $X \subset \mathbb{R}^d$ としその間の連続写像を $f, g: X \rightarrow \mathbb{R}$ とする。その一様ノルムを以下で定める。

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

また 2 つの関数 f, g 間の間の距離は $\|f - g\|_\infty$ で表される. 劣位集合フィルトレーション $(X(f)_r)_{r \in \mathbb{R}}, (X(g)_r)_{r \in \mathbb{R}}$ を持つような有限次元なホモロジー群が存在するとする. このとき劣位集合フィルトレーションのパーシステント図 $D(f), D(g)$ の間のボトルネック距離と関数の一様ノルムの間以下に不等式が成り立つ.

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

この定理により入力データ f と g のズレが小さいならば, パーシステント図のズレもボトルネック距離の意味で小さいということを意味し, パーシステントホモロジーのある種のノイズ耐性を示している. 本論文ではこの性質を利用し, ボトルネック距離を類似度の尺度として利用できないか考察するものとする.

5.3 Homcloud

パーシステントホモロジーの計算及び実験はソフトウェア「*Homcloud*」を用いた. *Homcloud* は材料科学への応用を目標とし大林らによって開発されたソフトウェアで, 入力データとして利用可能なものとして 2, 3 次元ポイントクラウド, n 次元画像, 距離行列, 抽象有限複体などがある. *Homcloud* は可視化, 機械学習, 逆解析といった応用的な機能に焦点を当てて開発が進められており, パーシステント図の計算やベクトル化, ボトルネック距離の計算などの基本的な機能も揃えられている.

5.4 実験手順

実際にどのように類似度を測定し, 比較するのか説明する. まず画像データとして Microsoft の提供する猫と犬のデータセット「Kaggle Cats and Dogs Dataset」を使用した. これは Web サービスの保護のために, 人間には判定しやすくコンピューターには判定しにくい課題として使用されるデータとして配布された. 犬か猫かのどちらであるか判定するためのデータセットとしても使用されるが, 今回はある程度のグレースケールの濃淡を持ち単純でない形状を持つ物体が含まれる画像として生物が含まれる画像である本データを選択した. データの構造として各画像は x 座標, y 座標に整数で表される濃度値 $I = \{0, 1, \dots, 255\}$ を格納した画素で構成され, 0 に近いほど黒く 255 に近いほど白い画素として表される. まずデータセットより各 4 枚の猫と犬の画像を選択し画像サイズを全てのデータで統一した後, 立方複体の V 構成によるフィルトレーションを行うことで 5.1 に示したような生成消滅対をプロットしたパーシステント図を出力する. 次に得られた 8 枚のパーシステント図同士のボトルネック距離を計算する.

5.5 実験結果

以下に実験に使用した画像を示す. 各画像はカラー画像をグレースケール化したものであり, 実験の際には画像のサイズを 300×200 に変換したものを使用した.

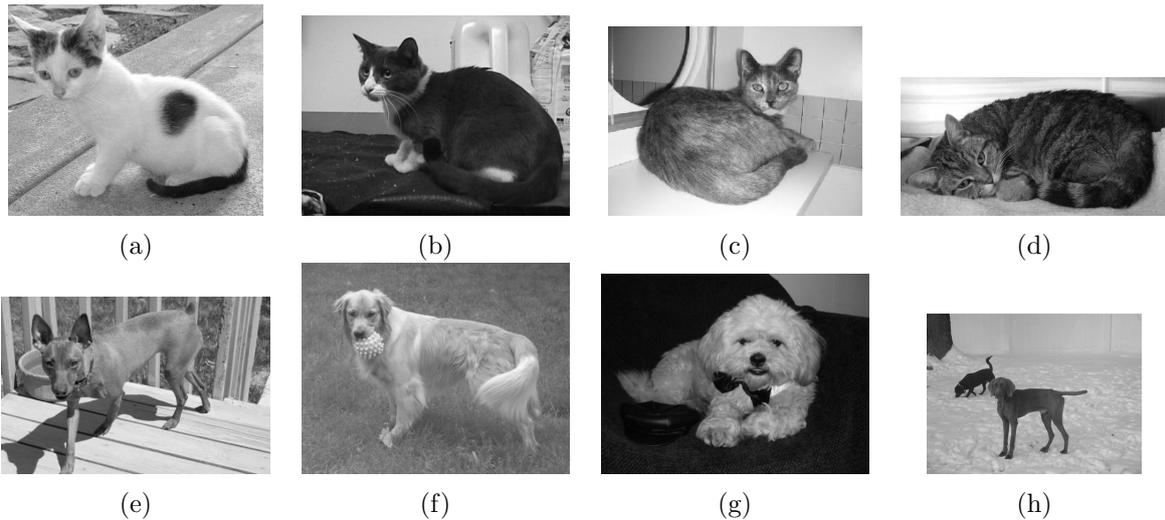


図 5.2: 使用したグレースケール画像

表 5.1 に各グレースケール画像とのボトルネック距離を表にまとめた. 表 5.1 を上段の行から調べていくと (a) は (b), (e) との距離が短く (b) に関しては幾何的な側面で見れば猫の形状などから類似していると考えられ, (e) とは被写体は猫と犬で異なるものの構図などから類似していると考えられる. また (b) は (e) と最も距離が近いが, これは被写体の形状が類似しているとは考えられない. (c) は (d), (e), (f), (g), (h) との距離が近く, 例えば (d) や (g) に関して直感的には類似しているが (e), (f), (h) に関しては類似していないと考えられる. (f) は (h) との距離とも近く, これは構図や形状は近いが, 画像全体の評価としてはあまり類似していないと考えられる.

表 5.1: 各画像のボトルネック距離

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
(a)	0	33	49	62	35	47	58	63
(b)	33	0	50	64	30	56	59	64
(c)	49	50	0	36	33	31	32	34
(d)	62	64	36	0	50	42	49	43
(e)	35	30	33	50	0	43	40	44
(f)	47	56	31	42	43	0	46	33
(g)	58	59	32	49	40	46	0	51
(h)	63	64	34	43	44	33	51	0

5.6 類似研究との比較

画像の類似度の主要な測定手法としてヒストグラム比較法と構造的類似度法が存在する. ヒストグラム比較法は2つの画像の色や明るさを表す画素値から作られたヒストグラムに対して, 標本相関係数を求めることで類似度を測定する手法である. すなわち N 個のビンを持つ2つのヒストグラムを H_1, H_2 で表し, 各ビンの度数をそれぞれ以下で表すとする.

$$\{H_k(I) \mid k \in \{1, 2\}, I \in \{1, 2, \dots, N\}\}$$

このときのヒストグラム比較法による類似度 $d(H_1, H_2)$ は以下で示される.

$$\bar{H}_k = \frac{1}{N} \sum_j H_k(j)$$

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{\sum_I (H_1(I) - \bar{H}_1)^2 (H_2(I) - \bar{H}_2)^2}}$$

このときの類似度 $d(H_1, H_2)$ は $[1, -1]$ の範囲で表され 2 つのヒストグラムが近いほど 1, 遠いほど -1 に近づく.

構造的類似度法は人間の画質劣化の知覚に画像構造の類似度が寄与するものとした指標で, 画像をいくつかの小領域に分割して領域ごとの構造や輝度など類似度を測り, 最終的に全体の類似度を測定する手法である. 2 つの画像 X_1, X_2 の画素値を x_1, x_2 とすると小領域ごとの類似度 $SSIM(x, y)$ は以下で表される.

$$SSIM(x_1, x_2) = \frac{(2\mu_{x_1}\mu_{x_2} + C_1)(2\sigma_{x_1x_2} + C_2)}{(\mu_{x_1}^2 + \mu_{x_2}^2 + C_1)(\sigma_{x_1}^2 + \sigma_{x_2}^2 + C_2)}$$

また各変数は以下で表される.

- μ_{x_k} : 画像 X_k の小領域内の画素値の平均画素値
- σ_{x_k} : 画像 X_k の小領域内の画素値の標準偏差
- $\sigma_{x_1x_2}$: 画像 X_1, X_2 の小領域内の画素値の共分散
- c_1, c_2 : 分母の値が小さ過ぎる場合に出力を安定させるための定数

これにより小領域ごとの $SSIM(x_1, x_2)$ を算出し, その平均値を最終的な類似度とする. このときの類似度は $[1, 0]$ の範囲で表され 2 つの画像が類似しているほど 1 に近づく.

ヒストグラム比較法, 構造的類似度法による類似度をボトルネック距離との結果と交えて各表の上段から考察し, 被写体の形状から考えた類似度と照らし合わせることで各評価の妥当性を考える.

表 5.2: (a) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	1	0.290555	0.162498	-0.015281	0.333433	0.0611564	-0.0708951	0.50153
構造的類似度法	1	0.155352	0.166189	0.12418	0.145583	0.228602	0.117895	0.222747
ボトルネック距離	0	33	49	62	35	47	58	63

(a) はヒストグラム比較法では (h), 構造的類似度法では (f), (h), ボトルネック距離では (b), (e) の類似度が高く評価され, ボトルネック距離の結果はヒストグラム比較法, 構造的類似度法ともに同一ではない. また被写体の形状から (a) と (b), (e) の類似度は高いと考えられる. よってボトルネック距離の評価が妥当である.

表 5.3: (b) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	0.290555	1	0.440479	0.120375	-0.0102455	-0.137818	-0.10101	0.290833
構造的類似度法	0.155352	1	0.211848	0.226042	0.128791	0.227928	0.19908	0.259932
ボトルネック距離	33	0	50	64	30	56	59	64

(b) はヒストグラム比較法では (c), 構造的類似度法では (h), ボトルネック距離では (a), (e) の類似度が高く評価され, ボトルネック距離の結果はヒストグラム比較法, 構造的類似度法ともに同一ではない. また被写体の形状から (b) と (a), (c) の類似度は高く, (b) と (e), (h) は低いと考えられる. よってヒストグラム比較法の評価が妥当であり, ボトルネック距離の評価も部分的に妥当である.

表 5.4: (c) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	0.162498	0.440479	1	0.27434	-0.0214228	-0.00837501	-0.0864179	-0.0092573
構造的類似度法	0.166189	0.211848	1	0.205483	0.175432	0.273543	0.123294	0.341057
ボトルネック距離	49	50	0	36	33	31	32	34

(c) はヒストグラム比較法では (b), 構造的類似度法では (f),(h), ボトルネック距離では (d), (e), (f), (g), (h) の類似度が高く評価され, 部分的にボトルネック距離と構造的類似度法の結果が同一である. また被写体の形状から (c) と (b), (d), (g) の類似度は高く, (c) と (h) は低いと考えられる. よってヒストグラム比較法の評価が妥当であり, ボトルネック距離の評価も部分的に妥当である.

表 5.5: (d) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	-0.015281	0.120375	0.27434	1	-0.173423	-0.147884	-0.115108	0.0487279
構造的類似度法	0.12418	0.226042	0.205483	1	0.11382	0.183923	0.127353	0.210716
ボトルネック距離	62	64	36	0	50	42	49	43

(d) はヒストグラム比較法では (c), 構造的類似度法では (b), (c), (h), ボトルネック距離では (c) の類似度が高く評価され, ボトルネック距離とヒストグラム比較法の結果が同一であり, 構造的類似度法とは部分的に同一である. また被写体の形状から (d) と (b), (c) の類似度は高く, (d) と (h) は低いと考えられる. よって各類似度手法の評価は部分的に妥当である.

表 5.6: (e) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	0.333433	-0.0102455	-0.0214228	-0.173423	1	0.472212	0.0172264	0.0823765
構造的類似度法	0.145583	0.128791	0.175432	0.11382	1	0.184234	0.0843167	0.195397
ボトルネック距離	35	30	33	50	0	43	40	44

(e) はヒストグラム比較法では (a), (f), 構造的類似度法では (h), ボトルネック距離では (a), (b), (c) の類似度が高く評価され, 部分的にボトルネック距離とヒストグラム比較法の結果が同一である. また被写体の形状から (e) と (a), (f) の類似度は高く, (e) と (b), (c), (h) は類似度が低いと考えられる. よってヒストグラム比較法の評価が妥当であり, ボトルネック距離も部分的に妥当である.

表 5.7: (f) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	0.0611564	-0.137818	-0.00837501	-0.147884	0.472212	1	0.0534812	-0.115584
構造的類似度法	0.228602	0.227928	0.273543	0.183923	0.184234	1	0.16537	0.311082
ボトルネック距離	47	56	31	42	43	0	46	33

(f) はヒストグラム比較法では (e), 構造的類似度法では (c), (h), ボトルネック距離では (c), (h) の類似度が高く評価され, ボトルネック距離と構造的類似度法の結果が同一である. また被写体の形状から (f) と (c), (e) の類似度は高く, (f) と (h) は低いと考えられる. よって各類似度手法の評価は部分的に妥当である.

表 5.8: (g) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	-0.0708951	-0.10101	-0.0864179	-0.115108	0.0172264	0.0534812	1	-0.0897113
構造的類似度法	0.117895	0.19908	0.123294	0.127353	0.0843167	0.16537	1	0.16923
ボトルネック距離	58	59	32	49	40	46	0	51

(g) はヒストグラム比較法での評価が平均的に低く, 構造的類似度法では (b), ボトルネック距離では (c) の類似度が高く評価され, ボトルネック距離の結果はヒストグラム比較法, 構造的類似度法ともに同一ではない. また被写体の形状から (g) と (b), (c) は類似度は高いと考えられる. よって構造的類似度法, ボトルネック距離の評価は妥当である.

表 5.9: (h) の各類似度評価

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
ヒストグラム比較法	0.50153	0.290833	-0.0092573	0.0487279	0.0823765	-0.115584	-0.0897113	1
構造的類似度法	0.222747	0.259932	0.341057	0.210716	0.195397	0.311082	0.16923	1
ボトルネック距離	63	64	34	43	44	33	51	0

(h) はヒストグラム比較法では (a), 構造的類似度法では (c), (f), ボトルネック距離では (c), (f) の類似度が高く評価され, ボトルネック距離と構造的類似度法の結果が同一である. また被写体の形状から (h) と (a), (c), (f) の類似度は低いと考えられる. よって各評価類似度手法の評価は妥当ではない.

最後にヒストグラム比較法, 構造的類似度法による評価を表にまとめたものと, 各グレイスケール画像から生成されたパーシステント図を以下に示す.

表 5.10: 各画像のヒストグラム比較

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
(a)	1	0.290555	0.162498	-0.015281	0.333433	0.0611564	-0.0708951	0.50153
(b)	0.290555	1	0.440479	0.120375	-0.0102455	-0.137818	-0.10101	0.290833
(c)	0.162498	0.440479	1	0.27434	-0.0214228	-0.00837501	-0.0864179	-0.0092573
(d)	-0.015281	0.120375	0.27434	1	-0.173423	-0.147884	-0.115108	0.0487279
(e)	0.333433	-0.0102455	-0.0214228	-0.173423	1	0.472212	0.0172264	0.0823765
(f)	0.0611564	-0.137818	-0.00837501	-0.147884	0.472212	1	0.0534812	-0.115584
(g)	-0.0708951	-0.10101	-0.0864179	-0.115108	0.0172264	0.0534812	1	-0.0897113
(h)	0.50153	0.290833	-0.0092573	0.0487279	0.0823765	-0.115584	-0.0897113	1

表 5.11: 各画像の構造的類似度

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
(a)	1	0.155352	0.166189	0.12418	0.145583	0.228602	0.117895	0.222747
(b)	0.155352	1	0.211848	0.226042	0.128791	0.227928	0.19908	0.259932
(c)	0.166189	0.211848	1	0.205483	0.175432	0.273543	0.123294	0.341057
(d)	0.12418	0.226042	0.205483	1	0.11382	0.183923	0.127353	0.210716
(e)	0.145583	0.128791	0.175432	0.11382	1	0.184234	0.0843167	0.195397
(f)	0.228602	0.227928	0.273543	0.183923	0.184234	1	0.16537	0.311082
(g)	0.117895	0.19908	0.123294	0.127353	0.0843167	0.16537	1	0.16923
(h)	0.222747	0.259932	0.341057	0.210716	0.195397	0.311082	0.16923	1

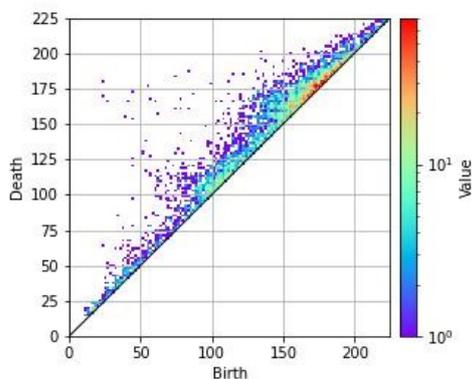


図 5.3: (a) のパーシステント図

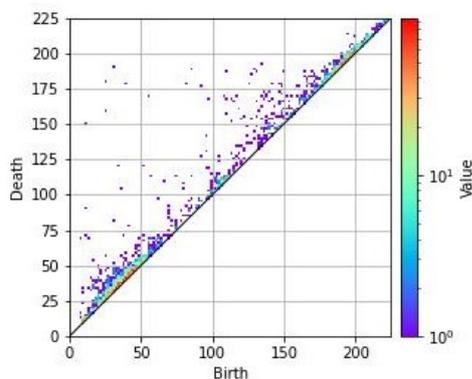


図 5.4: (b) のパーシステント図

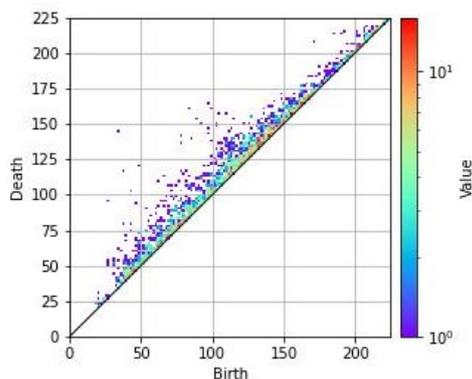


図 5.5: (c) のパーシステント図

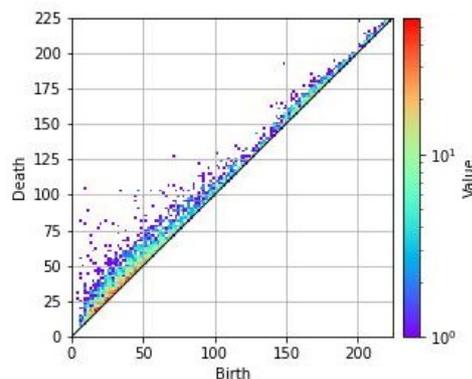


図 5.6: (d) のパーシステント図

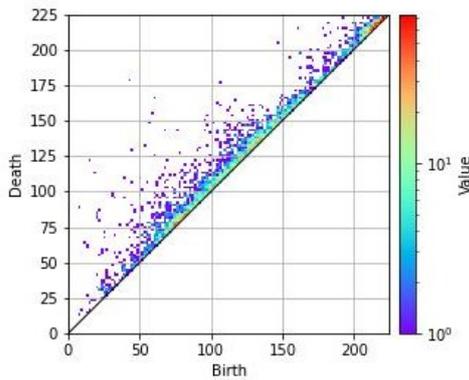


図 5.7: (e) のパーシステント図

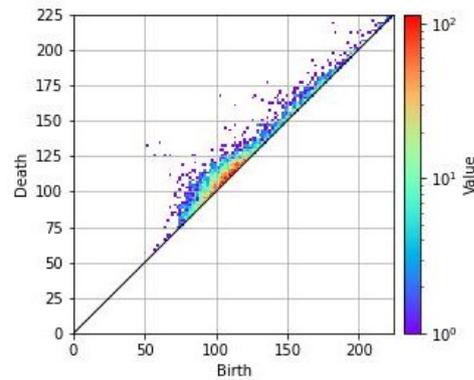


図 5.8: (f) のパーシステント図

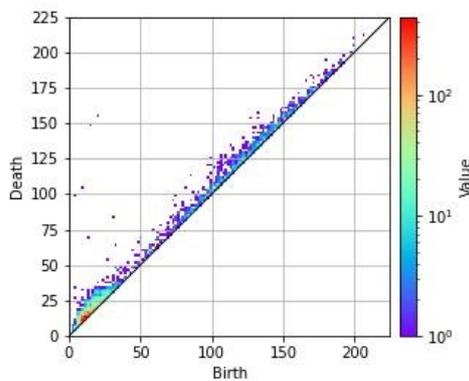


図 5.9: (g) のパーシステント図

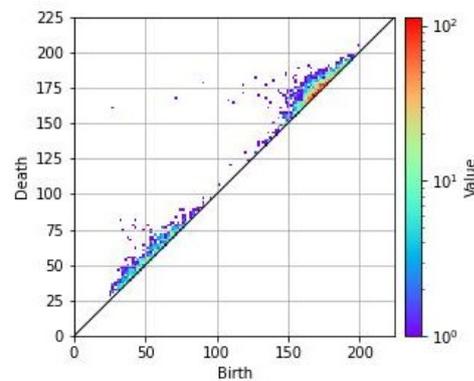


図 5.10: (h) のパーシステント図

5.7 実験のまとめと今後の課題

今回の実験ではグレースケール画像からパーシステント図を生成し、パーシステント図間のボトルネック距離を画像間の類似度とみなせるか検証した。従来の類似度指標は画像中の被写体の形状的特徴などを評価しないため、データの幾何的な特徴量を計算するパーシステントホモロジーが有効な評価指標になるのではないかと考察した。実験ではボトルネック距離は類似度の指標として部分的な妥当性を示しつつも、その有効性を示したわけではない。例えば図 5.2(h) の画像は被写体の形状から類似している画像はなく、距離が遠いながらも (f) との距離が最も近いと予想したが、(c), (f) との距離が近いという結果だった。今後の課題としてはリターゲティングなどを用いて画像中の物体にマスクやフィルターを施し、画像の図形的性質を強調することで有意なパーシステント図とボトルネック距離を抽出できるかを今後の課題として進めていきたい。

参考文献

- [池 et al. 23] 池祐一, E.G. エスカラ, 大林一平, 鍛冶静雄, 位相的データ解析かた構造発見へ—パーシステントホモロジーを中心に—, サイエンス社, 2023.
- [岩永 02] 岩永恭雄, 佐藤真久, 環と加群のホモロジー代数的理論, 日本評論社, 2002.
- [大林 17] 大林一平, 位相的データ解析の現在, 数理解析研究所講究録 2057, 34–50, 2017.
- [大林 20] 大林一平, パーシステントホモロジーに基づくデータ解析パッケージ Homcloud の紹介, 数理解析研究所講究録 2166, 103–107, 2020.
- [平岡 13] 平岡裕章, タンパク質構造とトポロジー—パーシステントホモロジー群入門—, 共立出版, (2013).
- [平岡 15] データに潜む幾何構造～パーシステントホモロジー～, 数理科学 2015 年 6 月号, サイエンス社, 2015.
- [平岡 19] 平岡裕章, 大林一平: パーシステントホモロジーの基礎と材料工学への適用例, 日本金属学会会報「まてりあ」, Volume 58 (1), 17–22, 2019.
- [栞田 02] 栞田幹也, 代数的トポロジー, 朝倉書店, (2022).
- [ARS95] M. Auslander, I. Reiten and S. Smalø, *Representation Theory of Artin Algebras*, Cambridge studies in advanced mathematics **36**, Cambridge University Press, 1995.
- [ASS06] I. Assem, D. Simson and A. Skowroński, *Elements of the Representation Theory of Associative Algebras*, London Mathematical Society Student Texts **65**, Cambridge University Press, 2006.
- [CCR13] J. M. Chan, G. Carlsson, and R Rabadan, *Topology of viral evolution*, Proceedings of the National Academy of Sciences, 110 (**46**), 18566–18571, 2013.
- [CEH07] D. Cohen-Steiner, Herbert Edelsbrunner, and John Harer, *Stability of persistence diagrams*, Discrete & Computational Geometry, **37** (1), 103–120, 2007.
- [C15] W. Crawley-Boevey, Decomposition of pointwise finite-dimensional persistence modules, J. Algebra Appl., **14**, 2015.
- [G72] P. Gabriel, *Unzerlegbare Darstellungen. I*, Manuscripta Mathematica **6**, 71–103, 1972.
- [PR15] P. Pilarczyk and P. Real, *Computation of cubical homology, cohomology, and (co)homological operations via chain contraction*, Advances in Computational Mathematics, **41**, 253–275, 2015.

-
- [S14] R. Schiffler, *Quiver Representations*, CMS Books in Mathematics, Springer, 2014.
- [SO95] M. Stricker and M. Orengo, *Similarity of Color Images*, Storage and Retrieval for Image and Video Databases III, 2420, pp.381–392, 1995.
- [WBSS04] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, *Image Quality Assessment: From Error Visibility to Structural Similarity*, IEEE TRANSACTIONS ON IMAGE PROCESSING, 13 (4), pp.600–612, 2004.
- [ZLKKKW15] C. Zixuan, M. Lin, W. Kedi, O. Kristopher, X. Kelin, and G. Wei. *topological approach for protein classification* Computational and Mathematical Biophysics, 3, 110–162, 2015.