数理科学実践研究レター　2020–10 September　28，2020

対流性降雨のパーシステントホモロジーを用いる特徴付け

by

キム ミンギュ

UNIVERSITY OF TOKYO
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
KOMABA, TOKYO, JAPAN

対流性降雨のパーシステントホモロジーを用いる特徴付け

キムミンギュ[1]（東京大学大学院数理科学研究科）
Minkyu Kim (Graduate School of Mathematical Sciences, The University of Tokyo)

概 要

対流性降雨とは局所的かつ集中的な降雨のことである．局所的かつ集中的な降雨は時間や地域に対して降雨量のデータの極大値である．連続的変数上のデータは微分を見ることで極値の判定ができることがよく知られているが, 降雨量のデータは離散的変数を定義域にするため工夫が必要である．本研究の目的は対流性降雨を位相的データ分析から特徴付ける方法を提案することである．

1 はじめに

対流性降雨とは局所的かつ集中的な降雨のことである．このように大まかな基準を補正して, 本論文では降雨量のデータ分析より対流性降雨を特徴付ける方法を一つ提案する．局所的かつ集中的な降雨は時間や地域に対して降雨量のデータの極大値になることに注目する．連続的変数上のデータは微分を見ることで極値の判定ができることがよく知られているが, 降雨量のデータは離散的変数（データを収集した地域や時刻）を定義域にするため工夫が要求される．そこで我々はパーシステントホモロジー及びそのノルムを応用する．最近脚光を浴びているトポロジーを用いるデータ分析 [2] にモチベーションがある．

2 有限距離空間上の関数の極値

降雨量のデータは「データを収集した地域や時刻」という有限距離空間から実数 $\mathbb{R}$ への関数を誘導する. 特に対流性降雨をその関数の極大値として扱えると思われる. 本節では有限距離空間上の関数の極値を評価する方法を一つ提案する.2.1 節ではパーシステンスノルムを簡単に復習する.2.2 節では有限距離空間上の関数の極値を判定する方法を提案する．2.3 節では 2.2 節で定義した指標 $I^f_{s_0}$ を正当化するための例を計算する．

2.1 パーシステンスノルム

有限距離空間 (X, d) のパーシステントホモロジーを $M(X, d)$ と書く．$M(X, d)$ は同型を除いて唯一なインターバル分解を持つことが知られている [1]．そのインターバル分解の成分 h からいわゆる生成時刻と消滅時刻の組み (b_h, d_h) が定義される．全ての h に対する $(b_h - d_h)^2$ の和の平方根をパーシステンスノルム $|(X, d)|$, 略して $|X|$ と書く．

2.2 極値の判定法

この節では有限距離空間上の関数の極値判定法を提案する．その準備事項として次のような定義を与える．

定義 1 有限距離空間 (S, d) とその上の関数 $f : S \to \mathbb{R}$ を考える．$s_0 \in S$ とする．実数 a, b に対して, $d(s, s_0) < r$ かつ $a < f(s) < b$ を満たす $s \in S$ の集合を $D^f_{s_0}(a, b\ ; r)$ と表す．$D^f_{s_0}(a, b\ ; r)$ を S の部分距離空間とみなす．有限距離空間 $D^f_{s_0}(a, b\ ; r)$ のパーシステンスノルムを $|D^f_{s_0}(a, b\ ; r)|$ と書く．$|D^f_{s_0}(a, b\ ; r)|$ を以下のような a, b の範囲で積分したものを $E^f_{s_0}(r)$ と定義する．ここで関数 $f(s)$ の $d(s, s_0) < r$ における最小値, 最大値を m_r, M_r とする．

$$E^f_{s_0}(r) = \int_{m_r \le a < b \le M_r} |D^f_{s_0}(a, b\ ; r)| \cdot da \cdot db. \qquad (1)$$

[1] kim@ms.u-tokyo.ac.jp

$E^f_{s_0}(r)$ を f に関して以下のように正規化したものを $F^f_{s_0}(r)$ と書く.

$$F^f_{s_0}(r) = \frac{E^f_{s_0}(r)}{E^c_{s_0}(r)}. \tag{2}$$

ただし、$c : S \to \mathbb{R}$ は定数関数である. この定義は c によらないことに注意する.

r の関数 $F^f_{s_0}(r)$ を $r = 0$ の付近での変化の様子を調べることによって, s_0 における関数 f の極値判定ができると思われる. $r = 0$ の周りで $F^f_{s_0}(r)$ の変化が大きいほど関数 f が s_0 に於いて極値である可能性が高くなるということである. このような動機により以下のような指標 $I^f_{s_0}$ を定義する.

定義 2 定義 *1* の設定を考える. 形式的な定義であるが, $I^f_{s_0}$ を以下のように定義する.

$$I^f_{s_0} = \left. \frac{dF^f_{s_0}(r)}{dr} \right|_{r=0}. \tag{3}$$

ただし, $F^f_{s_0}(r)$ は r に対して離散的なので微分ができないことに注意する.

以上の準備の下で以下の主張を述べる. この主張の証明はまだできていないが, 根拠となる例を次の章で与える.

主張 3 $I^f_{s_0}$ がゼロでなければ（あるいは, 十分大きければ）$s_0 \in S$ で関数 f は極大値か極小値をもつ.

注意 4 主張 *3* により鞍点は排除できる. しかし極大値と極小値のどちらか判定することはできないことに注意する.

2.3 例

この節では主張 3 の根拠として典型的な例を示す. 図 1 で特定の距離空間 (S, d) に対して $F^f_{s_0}(r)$ の様子を r に対してプロットした. 距離空間 (S, d) はユークリッド距離付きの 2 次元整数格子 $\mathbb{Z} \times \mathbb{Z}$ の部分距離空間である. S は原点を中心とする十分大きい半径を持つ円板にする. 関数 $f : S \to \mathbb{R}$ と $s_0 \in S$ として正則点, 極小値, 鞍点を与える代表的なものを考えている. 図 1 では関数 f が $s_0 \in S$ で極小値を持つ場合 $F^f_{s_0}(r)$ が急激に増加していることがわかる. 極大値の場合も同様であることに注意する. 主張 3 の根拠はこのような例である.

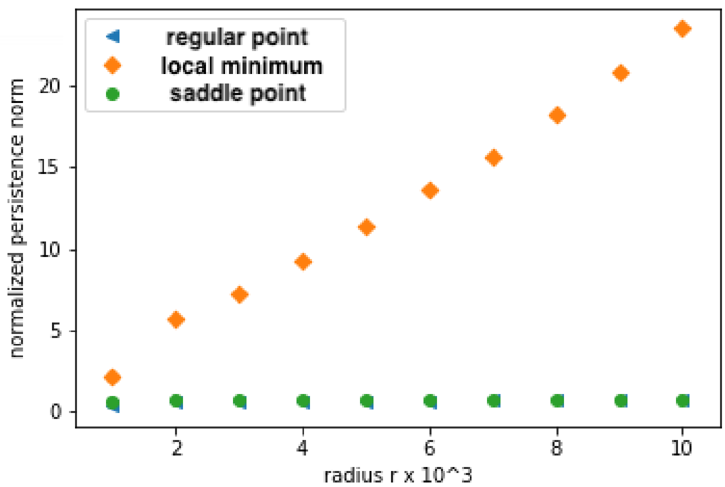

図 1: 整数格子 $S = \mathbb{Z} \times \mathbb{Z}$ と $s_0 = (0, 0)$ に対する $F^f_{s_0}(r)$ のグラフ. 正則点 (regular point) は $f(x, y) = x + y$ の場合である. 極小値 (local minimum) は $f(x, y) = x^2 + y^2$ の場合である. 鞍点 (saddle point) は $f(x, y) = x^2 - y^2$ の場合である.

3　終わりに

本論文では有限距離空間上の関数に対する極値の判定法を例に基づいて考察した．また，典型例である低次関数の場合を計算し，主張 3 の妥当性を数値的に検証した. 現時点における問題点はパーシステントノルムの計算が膨大なことである. そのため，まだ定義 2 の指標を実際の降雨量のデータに対して計算できていない. 定義 2 の指標を実際のデータに適用し極大点を検出することが今後の課題である.

本論文では有限距離空間を考慮しているが局所的に有限な距離空間に適用できる．局所的に有限な距離空間上の関数に対する数学的な応用も考えていきたいと思う.

参考文献

[1] Crawley-Boevey, W. (2015). Decomposition of pointwise finite-dimensional persistence modules. Journal of Algebra and its Applications, 14(05), 1550066.

[2] Zomorodian, A. (2012). Topological data analysis. Advances in applied and computational topology, 70, 1-39.