

多様なグラフ上の感染症ダイナミクスとパーシステントホモロジー

Epidemic Dynamics and Persistent Homology on Various Graphs

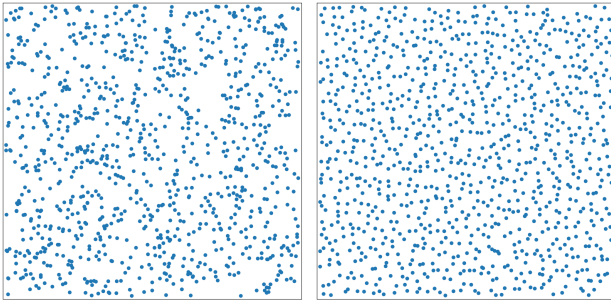
物理学専攻 石井 理矩斗

ISHII Rikuto

1 先行研究

先行研究 [1], [2] では PPP と GPP 上で連続パーコレーションを行い, 得られたグラフ上で SIR 感染シミュレーションを行っている. そして各感染率の違いによる累計感染者数 N_I に注目した議論を行っている.

1.1 PPP と GPP 上の連続パーコレーション



(a) PPP

(b) GPP

図 1: 点の数 $N = 1000$

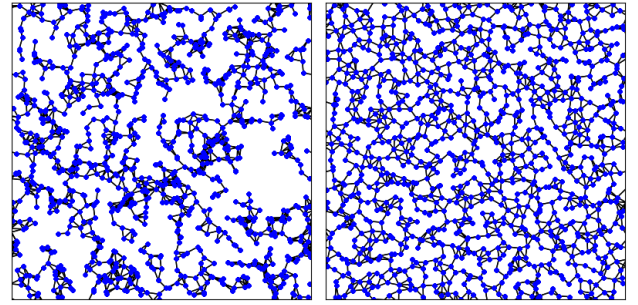
PPP は点同士の相関がなく空間的に一様に分布し, 図 1 (a) のように疎密がみられる. この疎密は揺らぎ, 偶然によるものである. 一方, GPP は各点同士に反発作用が働き, 密度や揺らぎが抑制される超一様性 [5] とよばれる偏りのない点過程なので, 図 1 (b) のように疎密がみられない. これらはそれぞれ 2 次元平面 $[0, 1]^2$ 上, 密度 N の点過程 Ξ_N^{PPP} , Ξ_N^{GPP} である.

次に得られた点過程に対して連続パーコレーションを考える. ある 2 点 $x, y \in \Xi$ が隣り合う点であるというのは, ある正の数 r を用いて $|x - y| < 2r$ となるときと考える. つまり x を中心とした半径 r の円盤 $B_r(x)$ を考え, $B_r(x) \cap B_r(y) \neq \emptyset$ となるときである. 連続パーコレーション・モデルでは密度 ρ , 円盤の面積 $a = \pi r^2$ を用いた

$$\kappa := \rho a = \rho \pi r^2 \quad (1)$$

を定義し, これを充填率という. 本研究では $\kappa = 1.3$

に固定して研究を行った. 図 1.4 が $N = 1000$, $\kappa = 1.3$ で連続パーコレーションを行い得られた最大クラスター上のグラフ \mathcal{G}_{1000} の一例である. 周期的境界条件を課しているため, 右端の点が左端に, 上端の点が下端にそれぞれ繋がっている.



(a) PPP

(b) GPP

図 2: $\kappa = 1.3$ のときに得られる $\mathcal{G}_{1000}^{\text{PPP}}$ と $\mathcal{G}_{1000}^{\text{GPP}}$ の例

1.2 SIR モデル感染シミュレーション

SIR モデルとは人口集団における感染症流行を記述する数理モデルであり, 未感染者 (Susceptible), 感染者 (Infected), 回復者 (Recovered) の頭文字をとっている. これは未感染者 S から感染者 I へ, 感染者 I から回復者 R へと一方向にのみ状態が変化し, 総人口の変化のないモデルである. このモデルを用いて \mathcal{G}_N 上で感染シミュレーションを行う.

各点 $x \in \mathcal{G}_N$ において, 確率変数 $\eta(x) \in \{S, I, R\}$ とおく. そして連続時間のマルコフ過程 $\eta_t := \{\eta_t(x) : x \in \mathcal{G}\}$ を考える. 点の状態 $\eta := \{\eta(x) : x \in \mathcal{G}\}$ から始まるマルコフ過程の確率 P^η とする. このとき感染率 λ , 正の関数 $c(x, \eta)$, 回復率 μ , 未感染者 S と繋がっている感染者 I の数を n とすると

$$\begin{aligned} P^\eta(\eta_t(x) = I, \eta(x) = S) &= \lambda n t + o(t) \\ P^\eta(\eta_t(x) = R, \eta(x) = I) &= \mu t + o(t) \quad \text{as } t \rightarrow 0 \end{aligned} \quad (2)$$

と与えられる. これを Gillespie アルゴリズム [15] とよばれる連続時間におけるマルコフ過程のシミュレーション

ン方法を用いて、 G_N^{PPP} と G_N^{GPP} 上でシミュレーションする。初期状態はある感染率 λ に対してランダムに1点を状態 I とし、状態 I の点がなくなるまで行う。同じ G_N 上で 100 回繰り返し、平均累計感染者数を得る。これを異なる G_N 上で 10 回行い、得られた 10 個の平均累計感染者数の平均を G_N 上の累計感染者数の代表値 N_I とし、横軸に感染率 λ 、縦軸に代表値 N_I としてプロットする。以降では N_I を単に累計感染者数と称する。

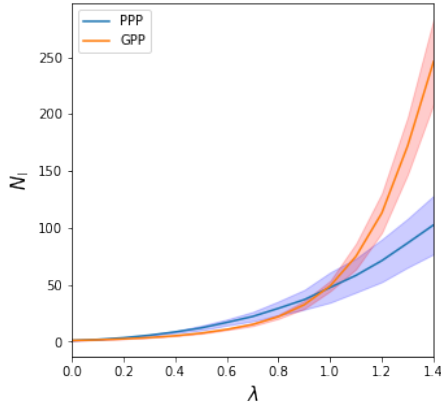


図 3: λ と N_I の関係

図 3 において、線が累計感染者数、薄く色が塗られた影の部分は 10 個の平均累計感染者数の標準偏差である。感染者推移が交錯しているところの感染率を λ^* とおく。すると $0 < \lambda < \lambda^*$ の範囲では PPP の累計感染者数の方が多くなっているのに対して、 $\lambda^* < \lambda$ の範囲では GPP の累計感染者数の方が多くなっている。これを逆転現象という。

2 パーシステントホモロジー

パーシステントホモロジーとはデータの背後に隠された特徴的な構造を捉えることができる強力なツールであり、近年様々な分野で急速に応用が進んでいる。

2.1 パーシステント図

パーシステントホモロジーは単体複体 K^t , $t = 0, 1, \dots$ のフィルトレーション

$$\mathbb{K} : K^0 \subset K^1 \subset \dots \subset K^t \subset \dots \quad (3)$$

に対して定義される。ここでフィルトレーション内の

単体複体 K^t を支配する添字 t を時刻とよぶことにする。フィルトレーション \mathbb{K} は、ある非負整数 Θ が存在し $K^j = K^\Theta$, $j \geq \Theta$ が成り立つとき、有限型であるという。またこの性質を満たす Θ の最小値を、フィルトレーションの飽和時刻とよぶことにする。このときパーシステントホモロジー $H_k(\mathbb{K})$ は、各 K^t でのホモロジー $H_k(K^t)$ とその間に誘導される線形写像の列

$$H_k(\mathbb{K}) : H_k(K^0) \rightarrow H_k(K^1) \rightarrow \dots \rightarrow H_k(K^\Theta) \quad (4)$$

として定められる。また、パーシステントホモロジー $H_k(\mathbb{K})$ に対して $b_i, d_i \in \{1, \dots, n\}$ ($b_i \leq d_i$) および非負整数 l_k が一意に定まり

$$H_k(\mathbb{K}) \simeq \bigoplus_{i=1}^{l_k} I[b_i, d_i] \quad (5)$$

の形に分解できることが知られている。ここで区間 $I[b, d]$ は時刻 b で k 次元の穴が発生し、時刻 d まで存続した後消滅することを表している。これはパーシステントホモロジーの一意分解定理として知られるものである。これにより、パーシステントホモロジーは発生と消滅の組 (b_i, d_i) を指定することで同型の差を除いて決まることが分かる。そこでこの組を平面上の多重集合

$$PD_k(\mathbb{K}) = \{(b_i, d_i) \mid i = 1, \dots, l_k\} \quad (6)$$

として表示したものを、 \mathbb{K} の k 次パーシステント図と定める。具体例として、PPP と GPP の 1 次パーシステント図を以下に示す。

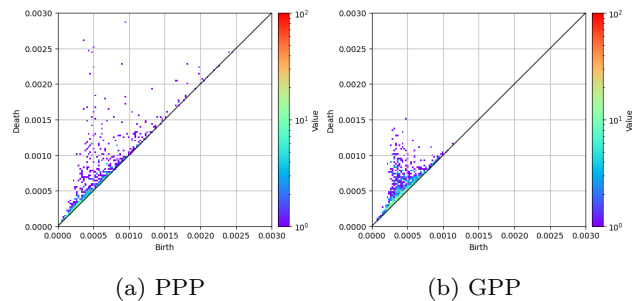


図 4: 1 次パーシステント図

パーシステント図の作成については Python のデータ解析ソフトウェアである HomCloud [7] を用いている。図 4 (a) では点が点在しているので様々なサイズの穴が

存在しており、PPP の特徴である疎密について表していることが分かる。一方、図 4 (b) では点が密集しているので同じくらいのサイズの穴が存在しており、GPP の特徴であるソーシャルディスタンスを保つ構造について表していることが分かる。ただしここで考えている穴とは、1 次パーシステント図のため 1 次元の穴、つまりリング構造のことである。これらは元のデータ同士を比較するときと比べ、より容易に特徴を捉え比較することができる。

3 多様なグラフ上の感染者推移

3.1 2次元格子上の感染者推移

ハニカム格子、正方格子、三角格子といった 2 次元格子上で感染シミュレーションを行い、PPP、GPP も加えて比較する。

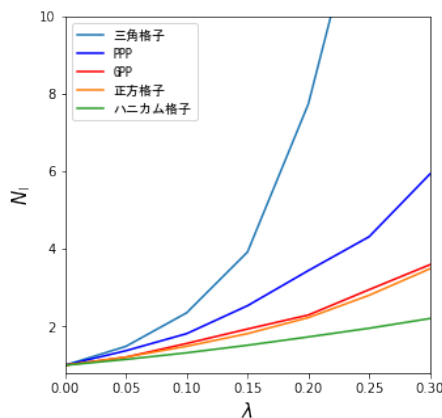


図 5: λ と N_I の関係

図 5 の凡例は $\lambda = 0.3$ 時の累計感染者数 N_I の降順になっている。PPP の平均次数が 5.2、GPP の平均次数が 4.2 なので、この順番は平均次数の降順と一致している。つまり感染率の十分小さいところではグラフの平均次数に依存した成長をすることが分かった。

3.2 ランダムグラフ上の感染者推移

HPP, PTL, Halton 列, mix 型, ガウス分布型のランダムグラフ上で感染シミュレーションを行い、PPP, GPP も加えて比較する。図 6 のようにランダムグラフの種類によって様々な感染者推移がみられるが、PPP のように緩やかな成長がみられるグラフと、GPP のよ

うに急激な成長がみられるグラフの 2 種類に分類することができる。前者を PPP 型、後者を GPP 型とよぶと、表 1 のように分類できる。

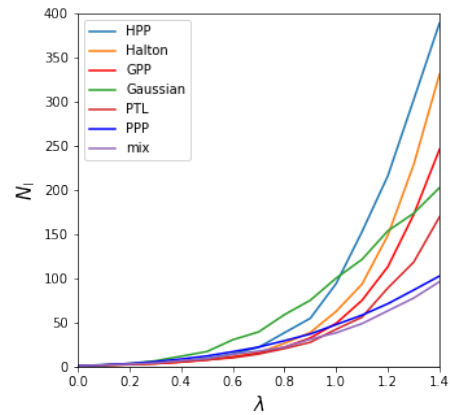


図 6: λ と N_I の関係

表 1: ランダムグラフの分類

PPP 型	GPP 型
mix 型	HPP
ガウス分布型	PTL
	Halton 列

ここで、図 6 をみると、PPP 型と GPP 型で逆転現象が発生しているため、逆転現象は PPP と GPP 特有の現象ではなく、無秩序な集団とみなせる系とソーシャルディスタンスを保つ集団とみなせる系との間で顕著に発生する現象であることが分かった。

4 まとめ

感染者推移は緩やかな成長をする PPP 型と急激な成長をする GPP 型の 2 種類に分類でき、この間で顕著な逆転現象が発生している。この違いは無秩序な集団なのか、ソーシャルディスタンスを保つ集団なのかで判断できる。また図 6 をみると、同じ GPP 型であっても感染者推移には大きな差がみられる。つまり、感染者推移の傾向を調べる場合はパーシステントホモロジーなどを用いて元のデータを比較すれば分かるが、より詳細な感染者推移を調べたい場合はグラフそのものの特徴に注目しなければならないということになる。

本研究では2種類に分類できたことで、ソーシャルディスタンスを保つという戦略がその詳細な保ち方によらず一般に同様な傾向を示すことが分かった。つまり、全体として密度や揺らぎが抑制される GPP のような集団、他人と必ず一定の距離を離すといった HPP のような集団、完全に秩序を保つ正方格子のような集団などでも感染抑制の効果が期待できるということになる。しかしこれらの集団は極端に感染率の大きい感染症である場合、むしろ感染が次々と広がりやすくなるため感染抑制の効果を期待できなくなる。従って、感染率が十分小さいときは感染者の周辺に注意すべきであり、そうでないときは PPP 型や GPP 型のように全体の構造、配置に注意した行動が求められる。

参考文献

- [1] Machiko Katori, Makoto Katori, Continuum percolation and stochastic epidemic models on Poisson and Ginibre point processes. *Phys. A. Stat. Mech. its Appl.* 581 (2021),126191
- [2] Machiko Katori, Makoto Katori, Spreading and suppression of infection clusters on the Ginibre continuum percolation clusters. (2021), arXiv:2105.044142
- [3] 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.
- [4] T.Tomé, R. M. Ziff, On the critical behavior of the Susceptible-Infected-Recovered (SIR) model on a square lattice. *Phys. Rev. E* 82, 051921 (2010).
- [5] 松井貴都, 数分散の厳密表式に基づく超一様行列式点過程の研究, 2020 年度中央大学大学院理工学研究科物理学専攻修士論文.
- [6] 赤池誠, ランダムな点過程上の連続パーコレーションと感染症ダイナミクス, 2022 年度中央大学大学院理工学研究科物理学専攻修士論文.
- [7] Obayashi, I., Nakamura, T. & Hiraoka, Y. Persistent Homology Analysis for Materials Research and Persistent Homology Software: HomCloud. *Journal of the Physical Society of Japan* 91, doi:10.7566/jpsj.91.091013 (2022).
- [8] Ghosh, S., Miyoshi, N. & Shirai, T. Disordered Complex Networks: Energy Optimal Lattices and Persistent Homology. *Ieee Transactions on Information Theory* 68, 5513-5534, doi:10.1109/tit.2022.3163604 (2022).
- [9] Kusano, G., Fukumizu, K. & Hiraoka, Y. Persistence weighted Gaussian kernel for topological data analysis. in *33rd International Conference on Machine Learning*. (2016).
- [10] Hofer, R. A metric result for special sequences related to the Halton sequences. *Journal of Complexity* 44, 23-29, doi:10.1016/j.jco.2017.09.001 (2018).
- [11] WOLFRAM, “HardcorePointProcess”, wolfram 言語 & システムドキュメントセンター, <https://reference.wolfram.com/language/ref/HardcorePointProcess>, (2024-01-10).
- [12] WPI-AIMR, ISM, JST, “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〜”, JST プレスリリース一覧, <https://www.jst.go.jp/pr/announce/20160614/>, (2024-01-10).
- [13] 平岡 裕章, 西浦 廉政, ランダムの中に見る秩序-パーシステントホモロジーとその応用-, *日本物理学会誌*, 2017, 72 巻, 9 号, p. 632-640, 公開日 2018/07/25, Online ISSN 2423-8872, Print ISSN 0029-0181, https://www.jstage.jst.go.jp/article/butsuri/72/9/72_632/_article/-char/ja/.
- [14] Huber, M. Generating From the Strauss Process Using Stitching, in *14th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (MCQMC)*. 241-251 (2022).
- [15] D.T.Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* 22 (4) (1976) 403-434, [https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3).