

2020年度共同利用研究報告書

2021年11月30日

所属・職名 信州大学理学部・准教授

沼田泰英

		整理番号	20200023
1.研究計画題目	位相的データ解析の理論と応用		
2.新規・継続	新規		
3.種別	プロジェクト研究		
4.種目	短期共同研究		
5.研究代表者	氏名	沼田泰英	
	所属 部局名	信州大学理学部	職名 准教授
6.研究実施期間	2021年08月04日(水曜日)～2021年08月06日(金曜日)		
7.キーワード	パーシステントホモロジー; Mapper; Representaion of quiver;		
8.参加者人数	93人		

9.本研究で得られた成果の概要

トポロジ的データ解析(TDA)には様々な手法があるが,その手法を用いた解析という形で応用分野に寄与する一方,応用分野での需要に応じた理論の開発という形で純粋数学へ新しい問題を供給している.また,トポロジ的とはいうものの,その一つ一つの手法はトポロジに限らず多くの数学の分野と繋がっている.

代表的な手法であるパーシステンスホモロジーを例にとると,Quiverの表現という代数的な視点での研究や,ランダム復体などの様な確率論的視点からの研究や,パーシステンスホモロジーのホモロジー類から元の構造を決めるという逆問題に相当する研究など多くの視点での研究がある.

また, もう一つの代表的な手法であるMapperは,与えられた離散的な点データから近似的なReeb graphを構成する手法だと思える.

従って, 得られたグラフの分岐点が重要な意味を持っている可能性があり,複雑ネットワークにおける中心性の議論などとの関連が期待できる.

この様な背景から,本共同利用研究では,TDAをキーワードに様々なバックグラウンドを持つ研究者を集め,Theoreticalな問題からPracticalな問題まで幅広く議論をし,互いの問題意識や技術を共有することで,新たな問題を発見・定式化することを目標として開催した.

当初は参加者による議論を中心とした計画を立てていたが,新型コロナウイルス感染症の感染拡大のため計画変更を余儀なくされ,様々なバックグラウンドをもつ研究者による講演を中心とした研究集会として開催をすることとなった.

オンラインでの開催ということでの困難はあったが,フリーディスカッションの時間には多くの議論や情報交換が行われ,異なる分野の間での問題意識等の共有という当初の目標はある程度達成できた.

九州大学 IMI 共同利用・短期共同研究 位相的データ解析の理論と応用 (Topics on Topological Data Analysis) 成果報告.

沼田泰英 (信州大学理学部数学科)

1 経緯等の概略

本共同利用研究は、当初、2020年夏に直面での実施を予定していた。しかしながら、新型コロナウイルス感染症の感染拡大により、当初の日程での開催は困難となったことから、2020年冬への延期をした。2020年冬の時点でも新型コロナウイルス感染症の対策に関するめど立っておらず、再度の延期を余儀なくされた。その後、2021年夏に直面または対面とオンラインのハイブリッドという形式での開催を計画したが、再度の新型コロナウイルス感染症の感染拡大により、対面を含む形での実施は困難であると判断し、完全オンラインの形式での実施となった。

当初の計画では対面で実施する予定であったので、通常の研究集会のような講演ばかりではなく、参加者によるフリーディスカッションの時間を中心として、スケジュールを立てる予定であった。しかしながら、オンラインでの開催となったため、フリーディスカッションを行うのは、技術的に難しいであろうことが予想されたので、当初の予定よりも講演を重視する形の計画となった。当初の予定とは異なり講演が主体の計画となったことは残念であったが、オンラインでの開催であることの強みを活かし、海外の研究者にも講演

をお願いすることが出来たことは、幸いであった。

また、講演を中心としたスケジュールとはなったものの、当初予定していた計画に近づけるため、参加者からの質問などに自由に講演者が答えられるようスケジュールを作成する際に工夫をした：まず、講演は1日3件程度とし、十分な休憩時間を挟むこととした。実際には、10:00 から、13:00 から、15:00 からとし、講演者には、1時間程度を目安に講演を準備してもらうよう依頼したので、講演と講演の間に1時間程度の時間が確保されており、その時間で質疑応答やディスカッション、休憩などを行うよう計画した。講演者には、講演間に十分な時間が用意されているので、講演中に参加者から質問があった場合は適宜詳しく回答をしてもらうよう依頼した。通常の研究集会では、講演の終了時刻もしくは講演時間が明記されることが多いが、講演の終了時刻等を厳密には決めないこととし、講演者には話しやすい長さでの講演を依頼した。また、講演の終了時刻が明記されていると、参加者の都合などを気遣い、講演をその時刻までに終わらせないといけないという感覚が講演者に生じてしまうことがある。講演途中での質問などで時間を使ってしまった場合、その後の講演内容を省略するなどし時間調整をする講演者も少なくない。講演者には気兼ねなく全ての内容を話してもらえよう、終了時刻などを明記しなかった。講演を録画した場合、参加者が講演途中に質問することに抵抗を覚える可能性があると思われたため、講演は原則として録画しないこととした。講演を録画した場合、講演をリアルタイムでは聴けなかった参加者も、後日講演を聴けるというメリットはあるものの、参加者による積極的な質問を期待し録画は行わなかった。

オンラインで実施する際には、Zoom のミーティングを使用した。参加者同士の議論・情報交換が積極的に行われるよう以下を試みた：参加者同士の議論が可能なように、zoom の機能を使いブレイクアウトルームを事前に複数用意し、自由に入退室ができるようにし、参加者にはアナウンスをした。しかしながら、このブレイクアウトルームを使用した参加者はいなかった。また、情報共有や講演中のチャットによる議論のため、Slack というサービスを利用した。Slack のワークスペースを世話人の方で用意し、参加登録時にその URL をアナウンスしていた。Zoom にもチャットの機能はあるが、自分が参加するよりも前に流れている情報は見ることができない、一旦ログアウトすると情報は失われてしまうなど

といった特徴がある。これらはセキュリティの観点からは有益ではあると思われるが、研究集会における情報共有という点では不便な特徴である。そのため、Slack のワークスペースを用意し、スライドファイルの共有や参考となる URL と言った情報の共有、講演中の議論などは、Slack のチャンネルの中で行うようアナウンスをし、zoom のチャットに投稿された情報は、適宜世話人が slack の方へ転記した。Slack は、講演中の質問や議論も含め、それなりに活用されていたように見受けられる。

今回のスケジュールは、一つ一つの講演の開始時間は 2 時間以上離れている。そのため、講演がいくつか連続して行われる通常の研究集会とは異なり、むしろ、単独のセミナー・講演が集まったような形式となっている。この様な形式であることや、各講演での話題も幅広いこともあり、参加者は各講演ごとに大きく異なっているように見受けられた。

2 講演等の概略

パーシステンスホモロジー等といった位相的データ解析の手法を中心的なトピックとした講演が行われた。その内容は、理論研究や応用研究、データ解析のためのツールの実装といった幅広いものであり、様々な視点からの話題提供があった。

また、各講演においては、講演直後からはフリーディスカッションの時間としたが、講演で扱われた話題を元に、多くの議論が行われていた。特定の問題をとりあげて議論をするわけではなく、フリーディスカッションという形での議論であったので、ここで行われた議論から直接論文になるような成果が得られたわけではないが、参加者の興味に応じて情報交換や問題提起が積極的に行われていた。

講演は、以下のものが行われた:

- 講演者: 濱田直希 (KLab 株式会社開発推進部機械学習グループ)

タイトル: ゲーム開発における機械学習と TDA

多目的最適化の問題に関わるトピック、パレート集合と呼ばれる多目的最適化の問題にとってキーとなる概念の構造、について解説がされた。また、ゲーム開発の現場での実際の問題を例に、多目的最適化の問題について説明された。

- 講演者: 池祐一 (東京大学大学院情報理工学系研究科)

タイトル: TDA のニューラルネットワーク解析への応用

ニューラルネットワークによる分類問題での過誤を解析するというテーマについて解説がされた。特に、ニューラルネットワークの一部の層を取り出し、発火しているネットワークの 0 次パーシステントホモロジーつまりクラスタリングに着目し調べる手法について解説された。
- 講演者: 梅田裕平 (富士通株式会社人工知能研究所)

タイトル: 企業での TDA 実践適用の取り組みと新しい取り組み

脳波の解析に用いられるトポロジカルデータ解析の手法などを例に挙げ、データの幾何構造等の解析などに関して、企業で取り組まれている実践的な応用を踏まえつつ解説がされた。
- 講演者: 吉脇理雄 (大阪市立大学数学研究所/理化学研究所革新知能統合研究センター)

タイトル: パーシステンス加群の導来圏と代数的安定性定理

通常のパーシステントホモロジーは半径パラメータという 1 つのパラメータで統制されている加群の列である。例えば半径パラメータの他に時刻のパラメータなどという形で複数のパラメータをもつ形の加群の列を考えるとといった拡張を考える場合には、本質的に問題が難しくなることが理論的に知られている。既存のパーシステントホモロジーの理論を導来圏という概念を用い再構築し、安定性定理を捉え直し証明するという理論的な研究トピックについて解説がされた。
- 講演者: 田中康平 (信州大学経法学部)

タイトル: 小圏のチェック複体と脈体定理

サンプリングされた離散データからパーシステントホモロジーを計算しサンプリング元の幾何的情報を調べるというデータ解析手法の正当性の基礎となるものの一つに、脈体定理というものがある。それを一般化した、小圏のチェック複体における脈対定理について解説がされた。小圏は Poset といった基本的な組合せ論的対象を含む広いクラスであり、応用が期待される。

- 講演者: 大林一平 (岡山大学サイバーフィジカル情報応用研究コア)

タイトル: パーシステントホモロジーによるデータ解析ソフトウェア HomCloud

HomCloud は, パーシステントホモロジーを計算しデータ解析を行うための代表的なソフトウェアの一つである. HomCloud について, その開発者による解説が行われた.
- Speaker: Vidit NANDA (Mathematical Institute, University of Oxford)

Title: Principal Components along Quiver Representations

Quiver の表現を一つ固定する. この表現から定まる切断のなす空間が, どの様に記述できるかに関する理論的な研究成果について解説された. この講演については, オンライン開催のメリットを活かし, 講演者には現地 (米国) から現地時刻の昼 (日本時刻夜間) にオンラインでの講演であった.
- 講演者: Emerson ESCOLAR (神戸大学大学院人間発達環境学研究科/理化学研究所革新知能統合研究センター)

タイトル: パーシステンス加群の区間分解可能性、区間近似、及びその計算

通常のパーシステントホモロジーは半径パラメータという 1 つのパラメータで統制されている加群の列であり, 区間表現と呼ばれる簡単な構造をもって記述できるが, パラメータが複数ある場合には, 本質的に難しくなることが理論的にわかっている. そのためパラメータが複数ある場合は単純な一般化は難しいが, 区間表現を用いて近似を行う手法について, その理論と実装について解説された.
- 講演者: 中村壮伸 (産業技術総合研究所 CD-FMat)

タイトル: パーシステントホモロジーを用いた環状高分子の貫通の定量化

絡みあう環状高分子はその絡み合い方で性質が変わってくるため, どのような絡み合い方をしているかを分類する必要がある. ある環状高分子を別の環状高分子が貫通しているという状態の判別に関し, 従来の曲率を用いた方法ではなく, パーシステントホモロジーを用いた新しい手法が提案された. また, この問題を例として, 複数の種類の点のあるポイントクラウドに対する, パーシステントホモロジーによるデータ解析手法の開発に関して問題が提起された.

位相的データ解析の理論と応用

Topics on Topological Data Analysis

日時： 2021年8月4日（水）10:00～6日（金）16:00
場所： Zoom ミーティングによるオンライン開催



8月4日（水）

10:00-

濱田直希（KLab 株式会社開発推進部機械学習グループ）
ゲーム開発における機械学習と TDA

13:00-

池祐一（東京大学大学院情報理工学系研究科）
TDA のニューラルネットワーク解析への応用

15:00-

梅田裕平（富士通株式会社人工知能研究所）
企業での TDA 実践適用の取り組みと新しい取り組み

8月5日（木）

10:00-

吉脇理雄（大阪市立大学数学研究所/理化学研究所革新知能統合研究センター）
パーシステンス加群の導来圏と代数的安定性定理

13:00-

田中康平（信州大学経法学部）
小圏のチェック複体と脈体定理

15:00-

大林一平（岡山大学サイバーフィジカル情報応用研究コア）
パーシステントホモロジーによるデータ解析ソフトウェア HomCloud

22:00- (13:00 UTC -)

Vidit NANDA (Mathematical Institute, University of Oxford)
Principal Components along Quiver Representations

8月6日（金）

10:00-

Emerson ESCOLAR（神戸大学大学院人間発達環境学研究科/理化学研究所革新知能統合研究センター）
パーシステンス加群の区間分解可能性、区間近似、及びその計算

15:00-

中村壮伸（産業技術総合研究所 CD-FMat）

パーシステントホモロジーを用いた環状高分子の貫通の定量化

各講演の後、休憩等を挟まずそのまま自由討論の時間とすることを予定しています。

また、講演者には、1時間前後の講演をお願いしていますが、講演中の質問などの状況により大幅にずれる可能性があります。

8/5(木)22:00からの講演は、英語での講演となります。

講演中の情報共有や急なスケジュール変更などの連絡のために、slackのワークスペースを
用います。URLなどは、参加登録時にお知らせいたします。

※研究実施期間：2020年8月4日(水)～8月6日(金)