

研究終了報告書

「増殖系と隠れ Markov モデルの対応に基づく学習の考察」

研究期間：2019 年 10 月～2022 年 3 月

研究者：中島 蒼

1. 研究のねらい

生命はランダムに変化する環境に適応し生き残るため、個体として経験から学習すると同時に、集団としても適者生存繁栄の自然選択を通じた進化を活用している。この「個の学習」と「集団の進化」は数学や情報学でも重要な研究対象になってきた。実際、個の学習は勾配法などの反復型の最適化アルゴリズムなどとして現れ、学習理論や強化学習などの分野を動機づけてきた。また集団の進化も進化計算や粒子フィルタなどの増殖を利用したアルゴリズムの設計原理になってきた。生物学と数学・情報学の双方で個の学習と集団の進化はどちらか一方に注目して研究されるのが主だったものの、両者が協調する系についての研究が始まっている。生物学においては、個の学習（個体の情報処理）が進化や適応（小進化）に与える影響が考察されてきた。特に個体が自らや先祖の経験から学習することで進化が加速されるのかが考察されてきた。また情報学においても、進化計算に勾配法などの個の学習（反復法による最適化）を組み合わせた memetic algorithm が提案されている。しかしながら、双方の分野で未解決な問題が残されている。生物学においては個体の経験学習が進化・適応の観点から価値があるのかは未解明である。また情報学においても memetic algorithm の性能解析は難しい課題である。

この研究課題では個の学習と集団の進化が協調する系についてのそれぞれの分野での課題に関して統一的な視点からの解決を試みる。その際、増殖を通じ進化する系に共通して現れる Feynman-Kac 公式を通じた数理的な対応関係を活用する。生物学においては（課題 1-1）進化を加速させるような個体の経験学習は可能なのかを考察し、加えて（課題 1-2）そのような学習に必要な情報の特徴づけを試みる。また（課題 1-3）個体の経験学習が進化を加速させる効果を定量化し、集団として増殖している状況下で個の学習を行う意義の定量的な予測や考察を可能にすることを目指す。この考察においては、隠れ Markov モデルや粒子フィルタの知見を Feynman-Kac 公式を通じ移行し利用する。数学・情報学においては（課題 2）memetic algorithm の性能解析をする新たな手法を理論生物学の知見を移行することで提案することを試みる。

。

2. 研究成果

（1）概要

まずは生物学における成果について説明する。（課題 1-1）第一に先祖の経験を利用した先祖学習という(ancestral learning)学習法により進化が加速されることを示した。具体的には先祖学習が隠れ Markov モデルのオンライン勾配法と数学的に等価であり、先祖学習により最適な形質発現戦略を獲得できることを示した。また数値実験で先祖学習がある場合と無い場合



に対して形質戦略の獲得を比較し、前者の方が早いことを示した。(課題 1-2) 第二に先祖学習で用いる先祖の経験の情報は学習に十分であることを示した。すなわち集団適応度の勾配の不変推定量を構築できることを確認した。これは個体間の意思疎通なしで学習ができることを意味し、生物学的な含意もあるものである。(課題 1-3) 最後に先祖学習による進化の加速を Fisher の基本定理を拡張することで定量化した。Fisher の基本定理は、元々は進化による平均適応度の増加を定量化する関係式である。拡張された Fisher の基本定理は先祖学習による集団適応度の増分を与えている。これは先祖学習がどのような状況で有利かを予測可能にするもので、実験生物学への貢献も期待できるものである。

(課題 2) 情報学においては個体群動態という理論進化学の手法を用いて memetic algorithm の性能を理論的に解析する枠組みを提案した。第一に memetic algorithm についても Fisher の基本定理を拡張することで、memetic algorithm と個の学習の並列実行との性能の差を比較した。拡張された Fisher の基本定理からは、memetic algorithm は個の学習の並列実行より性能が良いことが示され、増殖ダイナミクスを最適化アルゴリズムに組み込む効果を評価できている。第二に拡張された Fisher の基本定理と比べ具体例での計算が簡単な方法として、遡及過程という手法を個体群動態から取り入れた。遡及過程は増殖の影響を加味しバイアスされた個の学習である。確率的勾配法の遡及過程を線形近似の下で計算することで、確率的勾配法を用いた memetic algorithm の性能の評価を行った。結果として勾配の推定の S/N 比が改善するという形で増殖の効果が新たな視点から評価できている。

(2) 詳細

研究テーマ 1 個の経験学習の適応的意義の考察

数理生物学において、個体の情報処理と集団の進化の関係は古くから注目されつつも、十分には検討されてきていなかった。最近では個体の情報処理により増殖が早くなるかという適応的価値に興味を持たれ、この研究の流れにおいて、個体が自身の経験から形質を変更する経験学習を行うことで進化や適応が加速される可能性が考えられてきた。Xue と Leibler は先祖の形質を学習(Xue の学習則)により集団適応度が最大になる形質発現戦略が学習できることを示している。しかしながら、学習が進化を加速しているかは明らかになっていない。また進化を加速させるうえで学習に必要な情報が何なのかも特徴づけられていない。加えて、進化がどの程度加速されるかの定量的な予測も分かっていない。

本研究課題ではこれら三つの課題について、隠れ Markov モデルとの Feynman-Kac 公式を通じた対応関係を利用して解析を行った。第一に、先祖学習という先祖をまねる学習法(Xue の学習則の拡張)により進化が加速されることを理論・数値実験の双方から証明した。理論的には、先祖学習は隠れ Markov モデルにおけるオンライン勾配法や EM アルゴリズムと対応することを確認し、最適な形質発現戦略が学習できることを証明した。数値実験では、最適戦略の獲得の早さを先祖学習がある場合とない場合で比較し、前者の方が早く獲得されることを確認した(図1)。

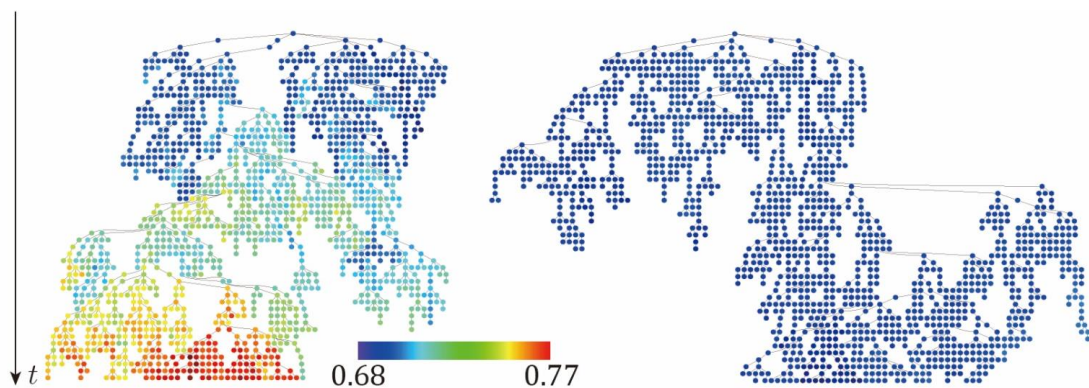


図1. 学習による進化の加速. 左は先祖学習がある場合の個体の系譜木であり、右は先祖学習がない場合である. 色で各個体の形質発現戦略の集団適応度を表している. 先祖学習がある場合の方が早く集団適応度が上昇するのが確認できる.

第二に進化を加速させる学習に必要な情報の特徴づけを行った. 先祖学習では先祖の形質の発現の経験分布(以下先祖情報と呼ぶ)を利用し学習が行われている. 本研究課題では先祖情報から集団適応度の十分推定量を構築できることを証明した. これは、適応度勾配情報を使わないランダムな変異と比較して、祖先学習が適応度勾配の1次の情報を活用している意味で進化を加速しており、また先祖情報が学習とその加速に十分な情報であるという特徴づけになっている. 特に個体間での意思疎通なしに学習できるというのは、生物学的な示唆もある特徴づけになっていると考えている.

第三に先祖学習によりどの程度進化が加速されるかを Fisher の基本定理を拡張することで定量化した. Fisher の基本定理は、元々は理論進化学において進化による平均適応度の上昇を定量化する関係式である. 本研究課題ではこの定理を先祖学習に拡張し、先祖学習による集団適応度の増分を定量化する関係式を証明した. 拡張された Fisher の基本定理は集団が増殖している状況下での個体の情報処理(先祖学習)の役割を集団適応度の増分として定量化したという点で、「集団の進化に加えて個の学習を行う意義は何か?」という研究計画時に提示した問いに答えたものになっている. 生物学においても、どのような状況で先祖学習が有利かの定量的な考察を可能にし、実験データの解釈や新たな実験系の設計に使えるという点で意義がある拡張であると考えている.

これらの研究成果はプレプリント (arxiv:2106.03367) として公開されており、現在 Physical Review Research において査読中である. また国内・国際学会で発表を行っており、特に数理生物学会においてはシンポジウムを企画し関連研究者と議論を行った(5-(3)).

研究テーマ 2 個体群動態の手法を用いた memetic algorithm の解析

進化計算の性能の理論的な解析は数多くの研究者によって取り組まれてきた重要な課題である. しかしながら、進化計算には増殖と組み換えという2つの個体間の相互作用があることもあり、解析はいまだ完全なものではない. また、勾配法と進化計算を組み合わせた memetic algorithm の解析も同様に難しい課題である. 本研究課題では、Feynman-Kac 公式を通じた対応関係を活用し、個体群動態の手法を進化計算の解析に応用することで新たな理論的な手法

を提案することを試みた。具体的には Fisher の基本定理と遡及過程 (retrospective process) という二つの枠組みを用いた解析手法を提案した。これらの成果は現在論文を執筆中である。

3. 今後の展開

本研究課題の数理生物学における成果は、個体の情報処理と集団の進化の関係性を一般的に捉えたものであり、また定量的な予測を与えているため実験的な検証や新たな実験系の設計など理論生物学を越えた実験生物学への波及も期待できる。情報学への成果は進化計算の理論解析をする新たな手法を与えたものである。進化計算は最近でも強化学習と組み合わせるなど統計・機械学習コミュニティで使われており、理論解析・アルゴリズム設計の双方で波及効果があると期待している。

4. 自己評価

本研究課題では、申請時点に掲げた課題を概ね達成できたと考えている。生物学においては掲げた課題を達成できている。先祖学習が進化を加速させることを理論・数値実験の双方から示し、また勾配法との関連を示すことで生物の学習戦略の解明はできたと考えている。また Fisher の基本定理という形で自然選択に加えて学習をする利得を評価できたことも、個の学習と集団の進化の協調を捉えたいという当初の動機に応えられたものと考えている。加えて Feynman-Kac 公式を通じた情報学への応用に関しても、当初想定していた粒子フィルタではないものの進化計算の理論解析という形で貢献できたと考えている。

研究の進め方も概ね計画通りであった。情報学への応用に関しては当初想定していた隠れ Markov モデルや粒子フィルタではなく進化計算を取り扱ったものの、個体群動態での手法や考察を情報学へ応用するという点において計画と一致した解析ができたと考えている。国内外の研究者との議論や共同研究については、コロナ禍もあり想定通りにはいかなかったものの、オンライン学会を通じてある程度の機会は確保できたと考えている。

研究成果の波及効果に関しては、「3. 今後の展開」で述べた通り、生物学・数学・情報科学のすべてに渡って波及すると考えている。特に生物学においては、数理生物学会においてシンポジウムを主宰し(5-(3))関連研究者との交流や議論を深めており、今後の進展が期待できると考えている。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数:0件

ga

(2) 特許出願



該当なし

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

- 中島蒼, 杉山友規, 小林徹也「細胞の内部状態を探り出す系譜木解析」
実験医学増刊 Vol.38 No.20「機械学習を生命科学に使う！」
- 日本数理生物学会 2020 年 9/22
中島蒼・小林徹也「増殖系におけるこの学習がもつ役割の考察」
学会内シンポジウム「増殖系と系譜解析の連携と発展」を主宰
- 2020 SMB Annual Meeting, 2020 年 8/17
So Nakashima, Yuki Sughiyama & Tetsuya J. Kobayashi “Lineage EM Algorithm for Inferring Latent States from Cellular Lineage Trees”
- the 20th IUPAB Congress, 45th Annual SBBf Meeting, and 50th Annual SBBq Meeting
2021 年 10/4~10/8
So Nakashima & Tetsuya J. Kobayashi “Acceleration of Evolutionary Processes by Learning and Extended Fisher’s Fundamental Theorem”