

経路積分による個体群動態方程式の解析と系譜木上の状態推定

Path-integral analysis for population dynamics and its application to state inference on lineage tree

杉山 友規 (東京大学生産技術研究所)

Yuki SUGHIYAMA (Institute of industrial science, The University of Tokyo)

yuki.sughiyama@gmail.com

細胞集団の“集団”としての増殖率(集団増殖率)を制御することは様々な分野で現れる重要な問題である。例えば医学分野においては、我々はがん細胞や病原性細菌の集団サイズを抗がん剤や抗生物質などを用いて抑制することを考える。一方、進化生物学の文脈では、細胞(生物)集団は変動する環境の中で集団増殖率を最大化する。そして、最も大きな集団増殖率を持った“種”が生存競争の淘汰の中を生き残る。さて、もう少しマクロな系に視点を移してみよう。有性無性の違いは存在するものの、細胞を個人に置き換えることによりここでの集団増殖率の制御を総人口の制御と読み替えることも出来る。このように至る所で非常に重要になる集団増殖率の制御であるが、実際に実行しようと思うと、細胞(個人)の個体間ダイバーシティが大きく、適切な処方箋を作ることは難しい。少し詳しくこの「難しさ」を表現してみよう。再び、医学分野におけるがん細胞を例としよう。もちろん、我々は“適切な”抗がん剤を用いてこのがん細胞集団のサイズを小さくしたい。しかし、がん細胞集団はヘテロな多数の遺伝型(タイプ)で構成されており、また全てのタイプに効果的な薬剤など存在しない。すなわち、タイプAの抑制には効果的であってもタイプBの抑制には全く寄与しない(下手な場合、活性化する)というような「タイプ特異的」な薬剤しか存在しない。このような状況における“適切な”薬剤の同定はどのように行うべきであろうか?安直に考えると、集団の中で最も多いタイプ(これをタイプAとする)に効果的な薬剤が最も適切な薬剤の様に思えるかもしれない。しかし、これは「真」ではない。なぜならば、がん細胞集団は増殖の過程で頻繁に突然変異を起こし、自身の子孫のタイプを次々に遷移させるからである。この事実は問題を格段に難しくする。たとえ多数派であるタイプAを抑制したとしても、実はタイプAを全体の中で多数派に押し上げていたのは少数派のタイプBであったとすると、全体の増殖抑制の意味ではあまり効果がない。言い換えると、タイプBがタイプAという子孫をたくさん産んでいた場合がこれに当たる。では、タイプAを多数生み出すタイプBを抑制すればよいかと言うと、そうでもない。自身の直下の娘としてタイプAを全く生まないが、1ステップ挟んで孫としてタイプAをタイプB以上に生み出すタイプCのような個体も存在するためである。このような複雑な事情は、細胞増殖だけに存在する問題ではない。人口学においても個人の属性は様々な存在し、それが時間とともに変化していくため、経済政策などを用いて出生数を制御しようとするとき同じような問題に直面することは容易に想像がつくだろう。どちらにしても、もはや自然言語を用いて解決可能な問題の範疇ではなくなっている。そこで、本研究では数理人口学の知識を動員してこの問題にアプローチする。具体的には、細胞系譜木(家系図)データ(以下の図参照)を用いて解析する数理手法を提案する。結果としては、系譜を時間遡及的に辿ったときに得られる統計量が重要な役割を果たすことがわかる。以下で、本講演で紹介する内容について述べよう。

今、各々の細胞(個体)は異なるタイプ(遺伝型や表現型)を持っており、このタイプに依存して分裂までの時間(分裂年齢)や分裂数が確率的に決定される系を考える。この設定は、環境に適したタイプを持つ個体ほど多くの子孫を残せることを意味し、そのタイプが進化的に生き残ることを表す。

人口学とのアナロジーを述べると、タイプを個人が所有する資産として定義し、資産を多く持つ個人ほど多くの子供を産むと仮定すると上記の系と同じ数理で解析できる。このような系に対して、集団増殖率（人口増加率）を評価することは言うまでもなく最も基礎的な問題の一つであろう。特に、環境が変化したときの集団増殖率の応答を評価することは応用上とても興味深い。人口学的に述べれば、社会状況の変化により資産価値が変動したときの人口増加率の変化を計算することに対応する。これらの問題に対して、本講演では細胞系譜（家系図）データを用いて解析する。具体的には、細胞系譜を時間前向きに辿った母子関係を表す経路と時間後ろ向きに辿ったその違いに着目することが解析のエッセンスとなる（以下の図参照）。結果を数学的に述べると、時間前向きのパス上に定義される大偏差関数を増殖バイアスにより Legendre 変換することで集団増殖率が評価される。また、この構造を用いることにより、環境変動に対する応答は、時間後ろ向きのパス上で計算される統計量から計算できることが明らかになる。さらに、これらの結果はただの理論ではなく、近年では実験的に検証可能なものであることも紹介する

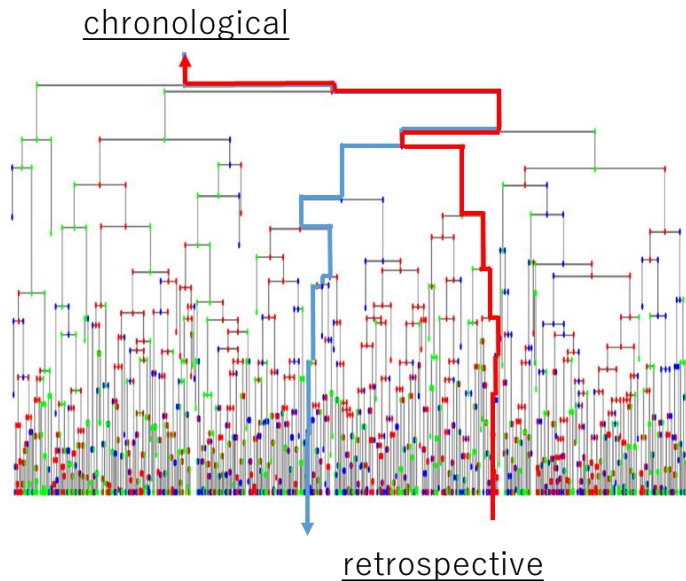
本講演の前半では以上の話を紹介するが、後半ではさらに踏み込んで、そもそもタイプとは如何に定義されるべきかということについて議論する。上述の細胞の話ではタイプとして遺伝型を仮定して述べていたが、実際の遺伝型は非常に多数存在し、そのすべてが増殖に寄与しているわけ

では無い。すなわち、実際の実験データの解析から応用につなげようと思うと、非常に広大な遺伝型空間を増殖に寄与する部分だけに粗視化する必要がある。そして、この粗視化したタイプが実際の上記の図に色として表れているタイプである。人口学の文脈ではこの問題はさらに本質的になる。そもそも、個人の増殖に寄与する属性とは何なのだろうかと言う社会学的な問いとも密接に関わってくる。本講演では、この属性（タイプ）というものを恣意性なく、系譜木から数理的に推定することを試みる。具体的には、色付けの無い系譜木に推定を用いて色付けする作業を行う。結果としては、隠れ Markov モデル+EM アルゴリズムを用いることで、効果的なタイプ推定が可能であることを見出した。

本講演は数理をベースとした理論的なものではあるが、数理を専門としない人にもファミリーなものになるよう心掛ける。また、理論自体は細胞集団を対象とした“生命科学的”なものではあるが、人口問題とのアナロジーを意識して構成しようと考えている。

<参考文献>

- [1] S. Nakashima, Y. Sughiyama and T. J. Kobayashi, Bioinformatics (2018). (To be published)
DOI: <https://doi.org/10.1093/bioinformatics/btaa040>
- [2] Y. Sughiyama, S. Nakashima and T. J. Kobayashi, Phys. Rev. E Vol. 99, 012413 (2019).
DOI: <https://doi.org/10.1103/PhysRevE.99.012413>



図：系譜木の上に置かれた色は細胞のタイプを表す。また、青線は先祖から子孫へと辿った時に得られる時間前向きのパスであり、赤線は子孫から先祖へと辿った時に得られる時間後ろ向きのパスである。