

## 論文の内容の要旨

### 論文題目

Reconstruction of cell signaling dynamics by statistical time series analysis  
(一細胞時系列解析によるシグナル伝達機構の抽出と再構成)

氏名： 近藤 洋平

本論文では、力学系モデルの時系列データからの再構成、および得られたモデルの解析によって細胞内シグナル伝達系のダイナミクスを理解することを目指す。力学系は、気象から化学反応まで、様々なダイナミクスの解析に使われてきたが、近年では生物学分野における重要性が認知されつつある。この一因として、生細胞イメージングによって情報伝達を担う分子システムの一細胞レベルでのダイナミクスがかつてない時空間分解能で明らかにされてきていることがある。しかし、観測過程と細胞の挙動の両方に強い雑音ないし確率性が存在すること、多次元の系に対して部分的な観測しか行えないこと、および細胞状態を精密に制御しえないこと、これらの困難は未だ避けがたいものとしてある。この結果、構成されるモデルはしばしば信頼できないものとなっている。

このデータの不完全性という問題に対処するため、統計学的枠組みに基づくモデル推定が行われ始めている。統計的手法を用いることによってダイナミクスのノイズに内在する情報の活用が可能になるため、一細胞レベルでの時系列データへの応用は特に有益であると期待されている。しかしそれでもなお、細胞シグナル伝達系の生化学的に精確なモデルの構築は非常に困難を伴う。そこで代わりに、分岐などの、システム生物学において重要視されてきたモデル特

徴を再現する低次元のモデルを統計的学習によって構成したい。この統計的学習と、力学系理論による。以上を踏まえ、機械学習手法のデザインと適用、さらに得られたモデルの解析までを行った。

まず第1章にて歴史的背景および先行研究に関する要約を与えたのち、第2章では、力学系の基礎的性質の一つである分岐に着目した。分岐は、力学系の挙動が定性的に変化するとき、その変化の様式を特徴づける概念である。これに基づくことで、系の全てが観測できない場合でも、分子的詳細に依存しない形で細胞挙動を理解することができる。それゆえ真の系と完全に同一でなくとも、同じ分岐タイプを持つモデルが得られるならば観測対象の理解に大きく寄与することが期待できる。そこで **Expectation-Maximization (EM)** アルゴリズムと粒子フィルタによって、分岐が発生する点をまたいだ複数のパラメータ値において得られた時系列データによって力学系を最尤推定する手法を実装した。ここでいう複数のパラメータ値とは、異なる環境での実験に対応していると考えられる。検証のため、最もよく理解されている分子システムの一つである細胞周期制御系のモデルのうち異なる分岐を持つ2つを採用し、人工データを生成した。また、細胞のような複雑な系を記述する際に不可避免的に伴うモデルの低次元化を考慮し、データから学習するモデルはデータ生成モデルより低い状態次元を持つと仮定する。学習の結果得られた低次元モデルは、元の系の分岐タイプを正しく受け継いでいることがわかった。この結果は頑強であり、少数のパラメータ値からのデータであっても推定は定性的には変化しない。また、分岐点の片側のみからデータが得られている場合でも、力学系の変化が適切に外挿されることで分岐タイプが推定されることが確認できた。

第3章では、スパース推定によって、データに基づき力学系モデルを単純化する方法を検証する。これによって、多数の可能な項を含む複雑なモデルから必要な部分モデルを抽出できるため、不確実性が多い細胞モデリングに対し有用であると期待できる。そこで **least absolute shrinkage and selection operator (lasso)** と呼ばれる、パラメータベクトルの  $L_1$  ノルムを用いて対数尤度に罰則をかける方法を採用する。これは第2章の手法に **Bayesian lasso** と呼ばれる事前分布を導入することで統計的枠組み内で実装することができる。この手法の人工データによる検証を次の2つの例で行い、いずれもモデル同定が成功することを確認した。ひとつはカオス力学系である。**Rössler** モデルを用い、その倍周期分岐カスケード周辺のいくつかの点でデータを生成した。カオスが存在する場合には時系列が複雑になるが、これはむしろ相空間の広い範囲から情報が得られることを意味するため、推定には助けとなりうる。もうひとつの検証例は適応力学系である。この場合には時間変化する入力のもとでの遷移的ダイナミクスから情報を抽出する必要がある、学習の観点からは困難な問題になりうる。しかし生物の情報伝達系における適応ダイナミクスの存在は普遍的であり機能的にも重要であるため、このタイプの力学系の学習が可能であることは応用上必須である。

第4章では社会性アメーバ *Dictyostelium discoideum* (以下、粘菌) の細胞間シグナル伝達系の学習を行う。特に、一細胞レベルでの入出力に着目するが、これは次の観点から一般性を持つ。細胞の集団が協動的なダイナミクスを示すことは頻繁に観測される現象であるが、要素である細胞間のシグナル伝達が常にその基礎にある。ここでもし、その一細胞レベルでの入出力ダイナミクスを再構成できたなら、内部の分子システムの生化学的詳細が不明であっても多細胞の振る舞いが予測、理解できると期待される。これは原理的には発生ダイナミクスの全てを含む広大な応用を持ちうるものであり、ここでは粘菌の細胞間シグナル伝達を担う環状アデノシンリン酸 (cAMP) に関する入出力ダイナミクスの学習によってその一歩目とする。この系は一細胞レベルでは適応を示すが、無性発生と呼ばれる協同的過程の初期において、細胞内 cAMP 濃度の特徴的な時空間パターンを形成することが知られている。そこで観測される化学波は細胞直径である数 10 マイクロメートルに対して数ミリメートルという長波長であり、紛れもなく多細胞現象である。

共同研究者である神野圭太氏は一細胞の入力と出力をそれぞれ細胞外と内の cAMP と定義し、その定量的関係を測定した。実験系では入力は還流系によって時間的に制御される一方、出力は Fluorescence resonance energy transfer (FRET) に基づくプローブによる生細胞イメージングによって計測される。測定によって、様々なステップ入力に対する適応時系列が得られた。さらに定量的解析は粘菌の cAMP 応答が、Fold-Change Detection (FCD) と呼ばれる性質を持つことを示唆した。これはステップ入力における入力レベル変化前後の比によって応答が決まるという性質であり、一細胞レベルでは初めて直接的に観測されたものである。

FCD は、力学系においては特定の対称性に対応していることが知られている。しかし、そのような対称性を持つモデルは既に複数提案されているため、FCD 性のみを根拠として一つのモデルを選択することはできない。この問題に対処するため、定性的、定量的な2つのアプローチをとった。

まず、FCD を示すモデルに共通し、かつ定量的性質によらない帰結を調べた。そのために、FCD 性を示す細胞内の cAMP シグナル伝達に加え、細胞外の cAMP を考えることで多細胞ダイナミクスのモデルを構成し、解析した。すると、一細胞レベルでの FCD 性は、多細胞レベルでの振動ダイナミクスにおいて、細胞密度の違いに対する不変性を生むことが分かった。興味深いことに、粘菌の無性発生は細胞密度の変化に対して非常に頑強であることが知られている。この現象に対して FCD に基づく多細胞ダイナミクスの不変性は、最終的なものではないにせよ、有望な仮説を与えると考えられる。

次に、cAMP 入出力に関する時系列データを用いた学習によって、観測時系列をよく再現する力学系モデルを得た。しかしそのモデルは複雑であり解析困難であった。そこでスパース推定によって、得られたモデルからその基本的な構造を抜き出すことを試みた。抽出された構造

は、粘菌の cAMP シグナル伝達系に関する先行研究と整合的であるが、データの不足による不確定性が多く残る結果となった。

第 5 章では、一細胞のモデルに基づく多細胞ダイナミクスの予測・理解について、粘菌を例にとり、より詳細に議論した。FCD 性を持つ極度に単純化したモデルによって、粘菌の協同ダイナミクスの多くが定性的に再現されることを示し、さらに、cAMP 時空間パターン形成以降の無性発生を頑強に遂行するためのメカニズムとして走化性の FCD モデルを提案した。また一方では、第 4 章における学習によって再構成された一細胞モデルに基づく協同的ダイナミクスの予測を行った。学習した力学系を空間的に拡張した多細胞モデルによって、実験的に観測されるものに類似した cAMP 波が再現された。

以上で提案された、統計的手法によるモデル学習と力学系理論によるモデル解析を組み合わせた方法論は、第 6 章にて議論するように、細胞さらには生命のダイナミクスを将来より深く理解する際に有用なものになると期待される。