

(続紙 1)

京都大学	博士 (情報 学)	氏名	長谷川 嵩矩
論文題目	Reconstructing Biological Systems Incorporating Multi-Source Biological Data via Data Assimilation Techniques データ同化手法を用いた多種生体内データの統合による生体内システム再構築の研究		
(論文内容の要旨) <p>本論文はデータ同化という統計的手法に基づき、遺伝子発現時系列データ、および、既知のネットワーク情報などから、遺伝子ネットワークなどの生体内システムを再構築するために手法について述べられており、8章から構成されている。</p> <p>第1章では、研究の背景と動機、本論文で提案している各手法の位置づけ、データ同化の概念的説明、論文の構成と各章の概要について述べている。</p> <p>第2章では、提案手法の基盤となる数理モデル、および、基本アルゴリズムについて説明している。具体的には線形状態空間モデル、非線形状態空間モデル、カルマンフィルター、粒子フィルター、EM (Expectation Maximization) アルゴリズム、情報量基準などについて説明している。</p> <p>第3章では、部分的に制御関係が判明している遺伝子を対象に、遺伝子発現時系列データに加えて薬物動態や文献由来制御構造に関する情報も取り込んだ上で、遺伝子ネットワーク構造を推定する手法を提案している。具体的には線形状態空間モデルに基づき、カルマンフィルターとL1制約つきEMアルゴリズムを組み合わせた推定手法を提案している。シミュレーションデータおよびベンチマークデータを用いて既存手法との比較を行い有効性を示すとともに、ラット骨筋細胞に対してコルチコステロイド刺激を与えた場合の薬物動態に関するネットワーク解析に適用している。</p> <p>第4章では、非線形な要素を取り入れた状態空間モデルのもとで、与えられた不完全な初期ネットワークから開始して遺伝子発現時系列データを用いてネットワーク構造を改善していく手法を提案している。具体的にはアンセンテッド・カルマンフィルター、EMアルゴリズム、さらに、ベイズ情報量基準を組み合わせた手法を提案している。そして、シミュレーションデータ、および、実際の遺伝子発現時系列データを用いて既存手法との比較を行い有効性を示している。</p> <p>第5章では、第4章の非線形状態空間モデルと類似のモデルを用いつつも、隠れ変数の条件付き分布の一次と二次のモーメント、さらに三次と四次のモーメントを維持することが可能なhigher moment ensemble particle filterという手法を提案している。シミュレーションデータ、および、前述のコルチコステロイド刺激経路を用いて既存手法との比較を行い有効性を示している。</p> <p>第6章では、微分方程式で記述された状態空間モデルを用いて、既存のネットワークの妥当性を評価し、さらに、より良い予測ができるように網羅的に候補モデルを作成、評価し、ネットワーク構造やパラメータを改善する手法を提案している。モデルの評価にあたっては、粒子フィルターを用いたモンテカルロ法により隠れ変数やパラメータの確率分布を計算し、それにベイズ情報量基準を組み合わせている。そしてラットの肝臓に関する大規模遺伝子データに適用し、対象とする刺激に対して比較的小規模な経路を持つ遺伝子に関してその推定が可能であることを示している。</p> <p>第7章では、第6章で提案した手法では多数の候補モデルを扱うため多大な計算機コストがかかるという問題点を解決するため、観測データの予測に関する制御構造の類似関係を利用することで、全ての候補モデルを評価すること無しに、効率的かつ選択的に候補モデルを評価可能な手法を提案している。そして提案手法をコルチコステロイド刺激経路とそれに関する遺伝子発現時系列データに適用し、生物学的に妥当と考えられる結果が得られることを示している。</p>			

第8章は結論であり、本研究をまとめるとともに、今後の研究の方向性や課題について述べている。

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し
審査結果の要旨は日本語500～2,000字程度で作成すること。

(論文審査の結果の要旨)

本論文は、遺伝子発現時系列データ、薬物投与データ、文献からのネットワークデータを統合することにより遺伝子ネットワークを中心とする生体内ネットワークを精密に推定することを目的として、シミュレーションと統計的推定を組み合わせたデータ同化という概念に基づいて新たに開発した一連の計算手法について述べたものであり、得られた成果は以下のとおりである。

(1) 線形状態空間モデルに基づき、カルマンフィルターとL1制約つきEMアルゴリズムを組み合わせた推定手法を提案した。本手法は既存のベクトル回帰モデルや状態空間モデルに基づく推定手法と異なり、遺伝子制御システムの基本的な要素を網羅した上で疎構造の条件下において推定を実行でき、かつ、薬物動態や文献由来制御構造を取り込むことが可能であるという特徴を持つ。ベンチマークデータなどを用いて既存手法に対する優位性を示すとともに、実データ解析に適用し、新たな制御関係を推定した。

(2) 非線形な要素を取り入れた状態空間モデルに基づき、ネットワーク構造を改善していく手法を提案した。本手法はアンセンテッド・カルマンフィルター、EMアルゴリズム、ベイズ情報量基準を組み合わせた独創的なものであり、さらに隠れ変数の条件付き分布の一次と二次のモーメント、さらに三次と四次のモーメントを維持することが可能なhigher moment ensemble particle filterへの拡張も行った。そして、シミュレーションデータ、および、実際の遺伝子発現時系列データを用いて既存手法との比較を行い有効性を示した。

(3) 微分方程式で記述された状態空間モデルに基づき、既存のネットワークの妥当性を評価し、かつ、より良い予測ができるように候補モデルを効率的に作成、評価する手法を提案した。提案手法をコルチコステロイド刺激経路とそれに関する遺伝子発現時系列データに適用し、生物学的に妥当と考えられる結果が得られることを示した。

以上、本論文ではバイオインフォマティクス、および、統計科学における重要な研究課題であるネットワーク推定問題に対し、新規で有用な手法をいくつか提案するとともに、シミュレーションデータおよび実際の遺伝子発現時系列データを用いた計算機実験により、それらの有効性を示している。提案手法のいずれもが新規性、有用性、拡張性の高いものであり、当該分野の発展のために十分な寄与をしている。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。また、平成26年11月20日論文内容とそれに関連した口頭試問を行った結果合格と認めた。

注) 論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。更に、試問の結果の要旨(例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」)を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。
要旨公開可能日： 年 月 日以降