

課題名 (タイトル) :

Statistical mechanics of linear and nonlinear matrix factorization

利用者氏名 : ○坂田 綾香

所属 : 望月理論生物学研究室

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

行列分解問題とは、与えられたデータ行列を、行列の積として表現する問題の総称である。行列分解により、対象とするデータ行列の圧縮表現が得られたり、典型的傾向や性質を抽出したりすることができる。しかし、この問題は非凸であり、また特異性を持つために、厳密に説くことが難しい。そこで、様々な近似アルゴリズムが提案されている。

さらに、データに行列分解問題を当てはめることは、データが線形であることを仮定していることになる。非線形データを含む一般のデータに対して、行列分解を用いて特徴抽出を試みることは、しばしば不適切である。そこで、行列分解をカーネル法と合わせて非線形化する方法が試みられている。これを非線形行列分解と呼ぶ。

本課題では、行列分解・非線形行列分解の問題に対して、性能の良いアルゴリズムを構成することを目標とする。

2. 具体的な利用内容、計算方法

我々は既に行列分解に対して belief propagation と呼ばれるグラフ上の確率伝搬アルゴリズムを構成している。これを非線形行列分解に対して適用することを考えた。非線形行列分解をカーネル関数により定められる特徴空間上での行列分解と位置づけると、カーネル関数の固有値や固有ベクトルから求まるノイズ下での線形行列分解と等価であることが示される。これにより、行列分解に対する belief propagation アルゴリズムを非線形行列分解に拡張することに成功した。

3. 結果

非線形行列分解の中でも、非線形 dictionary learning (DL) と呼ばれる、特徴空間上でのスパース表現を得る問題について、開発したアルゴリズムを適用した。非線形 DL を用いた分類問題やノイズ除去などの例題において、開発した手法はサンプル数が少ない状況下でも、既存手法よりも高い性能を示した。

4. まとめ

非線形行列分解、特に非線形 DL という複雑な問題に対しても、確率伝搬アルゴリズムが良い性能を示すことが示された。スパース性を利用する信号処理の問題において、同様の効率的アルゴリズムが提案できるのではないかと期待される。

5. 今後の計画・展望

行列分解・非線形行列分解問題は、特異モデルと呼ばれるモデル群に属す。これらの問題は、系の対称性に起因する遅い緩和現象を示すことが知られている。この問題点を解決することで、さらに効率の良いアルゴリズムが提案できると考える。

平成 26 年度 RICC 利用報告書

平成 26 年度 RICC 利用研究成果リスト

【学会発表】

- ・坂田綾香、樺島祥介、「belief propagation アルゴリズムによる特徴空間の圧縮」日本物理学会 2014 年秋季大会、2014 年 9 月、中部大学
- ・坂田綾香、樺島祥介、「統計力学的貪欲法による特徴空間のスパース表現」IBIS2014、2014 年 11 月、名古屋大学
- ・坂田綾香、樺島祥介、「統計力学的アルゴリズムによる特徴空間のスパース表現」SITA2014、2014 年 12 月、黒部市(富山)