

論文 / 著書情報
Article / Book Information

題目(和文)	スパース行列に対する行列分解の統計力学的解析
Title(English)	Statistical mechanics analysis of low-rank matrix factorization for sparse matrices
著者(和文)	野口千尋
Author(English)	Chihiro Noguchi
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第11671号, 授与年月日:2020年12月31日, 学位の種別:課程博士, 審査員:渡邊 澄夫,金森 敬文,高安 美佐子,三好 直人,山下 真,樺島 祥介
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第11671号, Conferred date:2020/12/31, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第		号	学位申請者氏名		野口千尋	
論文審査 審査員		氏 名		職 名		氏 名	職 名
	主査	渡邊澄夫		教授		山下真	教授
	審査員	金森敬文		教授	審査員	樺島祥介	特定教授
		高安美佐子		教授			
		三好直人		教授			

論文審査の要旨（2000 字程度）

本論文は Statistical mechanics analysis of low-rank matrix factorization for sparse matrices（スパース行列に対する行列分解の統計力学的解析）と題し、本文 5 章および付録 5 章から構成されている。 第 1 章 Introduction（導入）では、研究の背景と本論文の目的を記している。まず、現代社会において日々大量に生み出される高次元データから有用な情報を引き出す方法として、低ランク行列分解による次元圧縮が注目されていることに言及している。これを受けて、本論文の目的は統計力学の概念と技術にもとづき、低ランク行列分解の研究に寄与することである旨を述べている。その上で、本論文で検討する具体的な課題として、低ランク行列分解をもちいたスパース行列に対する行列補完アルゴリズムの開発、および、固有値分解にもとづくコミュニティ検出の性能分析を挙げている。 第 2 章 Preliminaries（準備）では、以降の章で必要となる低ランク行列分解、コミュニティ検出、行列補完に関する基本的な概念と先行研究についてまとめている。 第 3 章 Approximate matrix completion based on cavity method（キャビティ法にもとづく近似的行列補完）では、統計力学で発展した近似的確率推論手法であるキャビティ法にもとづいて低ランク行列分解をもちいた行列補完アルゴリズムを開発している。まず、与えられた行列 Y を適当な 2 つの行列 U, V をもちいて $Y \approx UV^T$ と近似する行列分解の問題に対して、ファクターグラフ表現を導入している。次いで、このファクターグラフにもとづきキャビティ法を適用することで、Y の中で観測された一部の要素から U, V を求めることにより Y の未観測要素を推定するアルゴリズム cavity-based matrix factorization (CBMF) を導出している。更に、CBMF に摂動的な近似を取り入れることで必要なメモリコストを大幅に削減した近似法 approximate cavity-based matrix factorization (ACBMF) を導いている。理論的な分析により、ACBMF の固定点は代表的な行列分解アルゴリズムである alternative least squares (ALS) の固定点と一致することを明らかにしている。加えて、人工データおよび推薦システムに対する標準的なベンチマークデータに適用することで、導出した 2 つのアルゴリズムが ALS および他の代表的なアルゴリズムである stochastic gradient descent と比較して優れた補完性能を示すこと、特に ACBMF は U, V のランク数が比較的大きな場合に計算コストの面でも優位であることを実験的に示している。 第 4 章 Fragility of spectral clustering for networks with an overlapping structure（オーバーラップ構造を持つネットワークに対するスペクトラルクラスタリングの脆弱性）では、行列分解にもとづくコミュニティ検出の方法として隣接行列の固有値分解を利用したスペクトラルクラスタリングを取り上げ、コミュニティ構造の曖昧さが検出性能の劣化を生む原因を統計力学の方法にもとづいて分析している。まず、スペクトラルクラスタリングは、理論解析でしばしばもちいられる stochastic block model (SBM) から生成されたグラフに対しては高い検出性能を示す一方で、実データに対する性能は必ずしも高くないことを指摘している。次いで、実データに対する性能劣化の原因として、複数のコミュニティに属するノードが一定割合存在する“オーバーラッピング構造”に着目し、そうした構造を生み出す SBM を新たに提案している。その上で、提案モデルから生成されたグラフにスペクトラルクラスタリングを適用した際の振る舞いを統計力学のレプリカ法をもちいて解析している。特に、オーバーラッピング構造を特徴づける 2 つのパラメータであるブロックサイズ α、オーバーラッピング強度 σ の影響について詳細に分析し、α の増加に対しては、正しいクラスタリングに対応する孤立固有値が減少しノイズに対応するスペクトルバンドに埋もれてしまうこと、また、σ の増加に対しては、スペクトルバンドが広がることで孤立固有値を飲み込んでしまうこと、が性能劣化の原因であることを明らかにしている。更に、数値実験により、得ら
--

れた結果の妥当性を確認している。

第5章 Conclusion（結論）では、本論文で得られた結果を総括するとともに、今後行うべき研究について展望を述べている。

付録 A-E では、第3章でもちいたベンチマークデータと各章で必要となる計算の詳細を記している。

以上を要するに、本論文は近年ビッグデータの出現とともにその重要性が増している低ランク行列分解に関し、その代表的な用途である行列補完とコミュニティ検出を具体的な対象として、統計力学の手法にもとづいた新規な分析アルゴリズムの開発、および、詳細な性能解析を行なったものであり、理学上、貢献するところが大きい。よって、博士（理学）の学位論文として十分な価値があると認められる。

注意：「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。