

Title	Study on Learning Invariant Patterns with Second Order Statistics
Author(s)	Hara, Satoshi
Citation	大阪大学, 2013, 博士論文
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/27563">https://hdl.handle.net/11094/27563</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	原 聡
博士の専攻分野の名称	博士 (工学)
学位記番号	第 26225 号
学位授与年月日	平成25年3月25日
学位授与の要件	学位規則第4条第1項該当 工学研究科電気電子情報工学専攻
学位論文名	Study on Learning Invariant Patterns with Second Order Statistics (2次統計量に基づく不変パターンの学習に関する研究)
論文審査委員	(主査) 教授 鷲尾 隆 (副査) 教授 滝根 哲哉 教授 北山 研一 教授 馬場口 登 教授 三瓶 政一 教授 井上 恭 教授 河崎 善一郎 准教授 古崎 晃司

## 論文内容の要旨

本論文では異なる時点や複数の環境下で得られたデータセット群を対象とし、これら複数のデータセットに共通する不変パターンを学習する手法を提案した。我々はこのような学習手法の構築のために、特に多変量データを表現する代表的なパラメータである2次統計量に着目した。本論文は以下の6章で構成される。

第1章は序論であり、研究目的及び背景となる既存研究、そして2次統計量に基づく基本的なモデルとして主成分分析モデル (PCAモデル) とグラフィカル・ガウシアン・モデル (GGM) の2つを紹介する。

第2章では定常部分空間分析モデルと呼ばれる、PCAモデルに代表される線形混合モデルの一種を対象とした。定常部分空間分析の目的は、線形混合モデルに基づいて複数の共分散行列に内在する不変パターンを抽出することである。これに対し、我々は一般化固有値問題を解いて問題の解を得るアルゴリズムを提案した。提案法は計算効率性、解の最適性保証の2点において従来法よりも優れた手法である。これら2つの有効性について理論的解析及び数値実験に基づく比較を行った。

第3-5章では変数間の依存性を表現するGGMを対象とした。第3章では以降の章の基礎となる凸最適化手法DAL-ADMMを提案した。この手法の有用性はその柔軟性にある。従来のGGM学習法の多くは11正則化を用いた最も単純な問題を想定してきた。これに対し、DAL-ADMMはグループ正則化を初めとするより多様な正則化項を扱うことが可能である。この柔軟性は4,5章でより複雑な問題を扱う上で欠かすことができない重要な特性である。

第4章では2つ目の学習手法の核となる部分、複数のGGMの不変パターンを学習する問題を扱う。我々はこの問題を、11正則化とグループ正則化という2つの技術を用いることで凸最適化問題として定式化した。さらに、この問題がDAL-ADMMにより扱えることを示した。また、提案法の有効性を確認するために数値実験及び実データを用いた異常箇所同定実験を行った。

第5章ではGGMを用いた異常箇所同定問題を扱う。本章では前章での問題の定式化を異常箇所同定へより特化したものへと置き換える。新たな定式化の基礎となるのは、異常箇所同定においては異常前後での精度行列 (共分散行列の逆行列) 間の行・列単位での変化が重要な役割を果たすということである。我々はこの考えに基づいて新たな正則化項を導入し、対象を凸最適化問題として定式化した。さらに、問題に適切な変形を加えることでDAL-ADMMでの扱いが可能となることを導いた。最後に、実データを用いた異常箇所同定実験において提案法の有

効性を確認した。

第6章は結論であり、本研究で得られた成果を総括する。

## 論文審査の結果の要旨

通信ネットワーク技術やセンシングネットワーク技術の発展に伴い、複雑な構造を有する膨大なデータが、オンライン・オフラインを問わず計算機に蓄積される時代となっている。

このような技術的、社会的状況を背景として、本論文では異なる時点や複数の環境下で得られたデータセット群を対象とし、特に多変量データを表現する代表的なパラメータである2次統計量に着目し、これら複数のデータセットに共通する不変パターンを学習する手法を提案している。主たる研究成果を要約すると以下ようになる。

(1) 定常部分空間分析モデルと呼ばれる、PCAモデルに代表される線形混合モデルの一種である線形混合モデルを対象とし、複数の共分散行列に内在する不変パターンを抽出する定常部分空間分析手法を提案している。特に、一般化固有値問題を解いて問題の解を得るアルゴリズムを提案し、提案法は計算効率性、解の最適性保証の2点において従来法よりも優れた手法であることを、理論的解析及び数値実験に基づき明らかにしている。

(2) 従来のGraphical Gaussian Model (GGM) 学習法の多くは11正則化を用いた最も単純な問題を想定してきたのに対し、グループ正則化を初めとするより多様な正則化項を扱うことが可能な凸最適化手法DAL-ADMMを提案している。

(3) 変数間の依存関係を表す複数のGGMに亘る不変な依存関係を学習する問題を扱っている。この問題を、11正則化とグループ正則化という2つの技術を用いることで凸最適化問題として定式化し、更に上述のDAL-ADMMにより扱えることを示し、数値実験及び実データを用いた異常箇所同定実験を通じてその有効性を確認している。

(4) 異常箇所同定問題を新たな正則化項の導入により凸最適化問題に定式化できることを示し、上述のDAL-ADMMでの扱いが可能となることを導き、実データを用いた異常箇所同定実験において本定式化手法の有効性を確認している。

以上のように、本論文は大量の変数から構成されるデータにおいて、変数間の複雑な依存関係を刻々と変化する非定常部分と定常な部分に分離して体系的に解析を行う手段を提供している。データマイニングや統計解析の分野においてこのような方法論は斬新かつ有用であると考えられ、当該分野研究の今後の発展に寄与するところ大きい。よって、本論文は博士論文として価値あるものと認める。