

独立成分分析を用いた潜在変数をもつ グラフィカルモデルの構造推定

数理情報学専攻 48196204 石濱 和樹

指導教員 駒木 文保 教授

1 はじめに

グラフィカルモデリングは、多変量データの関係構造を表現する統計モデルをグラフによって表現する多変量解析手法である。観測データから構造や変数間の依存関係を推定する様々な手法が研究されており、ノイズ分布に非ガウス性を仮定した線形モデルである LiNGAM[4] もその推定対象の一つである。

本研究では、潜在変数を含むグラフィカルモデルの一つである線形潜在モデルを対象として、LiNGAM モデルに対する推定手法と独立成分分析の性質を援用することで、既存手法の改善を提案した。

2 独立成分分析

独立成分分析 (Independent Component Analysis) は信号成分が互いに独立であるという情報を基に、信号の混合から元の成分を復元する手法である。ICA モデルは観測 $\mathbf{x}$ が独立成分 $\mathbf{s}$ の $\mathbf{A}$ による線形混合として得られる以下のように定義される。

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (1)$$

簡単のため混合行列 $\mathbf{A}$ は正方とする。右辺の分解は一意ではないため、ICA の解は行の置換とスケールの自由度を持つ。ICA モデルの混合行列の推定は、独立性の規準を導入し、各成分の独立性が最大となるような観測 $\mathbf{x}$ の射影 $\mathbf{w}^\top \mathbf{x}$ を逐次的に求める手順となる。主要な推定アルゴリズムとしては、不動点法を用いた fastICA[2] が提唱されている。

応用的な ICA モデルとして、観測に分散共分散行列 Σ が既知のガウスノイズ $\mathbf{n}$ が乗った以下の Noisy-ICA モデルを考えることができる。

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n} \quad (2)$$

NoisyICA モデルにおける混合行列の推定は、前処理として中心化と準白色化を行ったデータに対して、fastICA と同様の導出で得られる偏差除去アルゴリズム (Algorithm 1) を用いて行うことが可能である。

Algorithm 1 偏差除去 ICA アルゴリズム [1]

- 1: 観測 $\mathbf{x}$ に中心化と準白色化を行い $\mathbf{z}$ を得る
- 2: $p = 1$ とする
- 3: $\mathbf{w}_p$ をノルムが 1 となるように乱数で初期化
- 4: $\mathbf{w} \leftarrow \mathbf{E}[\mathbf{z}g(\mathbf{w}^\top \mathbf{z})] - (\mathbf{I} + \tilde{\Sigma})\mathbf{w}\mathbf{E}[g'(\mathbf{w}^\top \mathbf{z})]$
- 5: $\mathbf{w}_p \leftarrow \mathbf{w}_p - \sum_{j=1}^{p-1}(\mathbf{w}_p^\top \mathbf{w}_j)\mathbf{w}_j$
- 6: $\mathbf{w}_p \leftarrow \mathbf{w}_p / \|\mathbf{w}_p\|$
- 7: 収束していなければ 4 に戻る
- 8: $p < m$ ならば $p = p + 1$ として 3 に戻る

3 LiNGAM

LiNGAM(Linear Non-Gaussian Acyclic Model)[4] は加法性と線形性、非巡回性を仮定し、ノイズ分布として非ガウス分布を仮定した構造方程式モデルである。LiNGAM の構造方程式は変数を因果的順序に並べ替えることにより、厳密な下三角正方行列 $\mathbf{B}$ を用いて以下のように表現される。

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{e} \quad (3)$$

$\mathbf{e}$ は互いに独立な非ガウス分布に従う。この式は変形することにより、ICA モデルとして表現することが可能である。

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{e} \iff \mathbf{x} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{e}$$

$\mathbf{B} - \mathbf{I}$ は対角成分が 1 である下三角行列となるため、この ICA モデルは一意に解くことが可能となる。したがって、LiNGAM は識別可能なモデルとなる。

4 提案手法

本研究で推定の対象とする線形潜在モデルは各潜在変数から 3 つ以上の観測変数が得られる構造である。

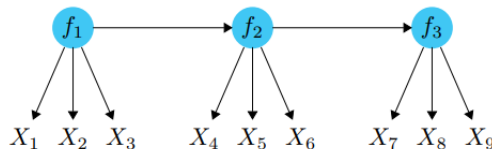


図 1. 線形潜在モデル。

潜在変数間の構造方程式ではノイズが非ガウス分布に従い、観測ノイズをガウス分布とした場合、このモデルの構造方程式は以下の表現が可能となる。

$$\mathbf{f} = \mathbf{B}\mathbf{f} + \mathbf{d} \quad (4)$$

$$\mathbf{x} = \mathbf{G}\mathbf{f} + \mathbf{e} \quad (5)$$

$\mathbf{d}, \mathbf{e}$ は非ガウス分布とガウス分布にそれぞれ従うノイズで、(4) 式は潜在変数の関係を表現する LiNGAM モデルであり、(5) 式は観測モデルを表す。 $\mathbf{A} = \mathbf{G}(\mathbf{I} - \mathbf{B})^{-1}$ とすると、この 2 式から

$$\mathbf{x} = \mathbf{A}\mathbf{d} + \mathbf{e} \quad (6)$$

という表現が得られ、これは NoisyICA モデルとなる。既存手法 [3] では、PCA により次元削減を行ったうえで近似的に fastICA アルゴリズムを利用して $\mathbf{A}$ の推定を行う。

提案手法においては、(6) 式が PCA を行った上でも NoisyICA モデルに従う性質を利用し、偏差除去アルゴリズム [1] を援用する。この際、各観測ノイズの分散 σ_i の推定が必要となるため、確認的因子分析による $\mathbf{G}$ の推定結果を用いて

$$\hat{\sigma}_i = V[X_i] - \|\hat{\mathbf{g}}_i\|^2 \quad (7)$$

として計算する。ただし、 $\hat{\mathbf{g}}_i$ は $\hat{\mathbf{G}}$ の i 行目である。

以上の計算に基づき、本研究では以下のアルゴリズムを提案した。なお、実際に構造推定を行う場合はこの前の手順として因果探索手法である BPC アルゴリズム [5] を用いるが、本研究の改善部分に関与しないため省略する。

Algorithm 2 線形潜在構造方程式モデルの推定

- 1: 確認的因子分析により、 $\mathbf{G}$ の推定量 $\hat{\mathbf{G}}$ を得る
 - 2: PCA の射影行列 $\mathbf{E}$ を計算
 - 3: (7) 式に基づき次元削減後の観測ノイズの分散共分散行列を計算
 - 4: 潜在因子の数の次元に削減を行い、Algorithm 1 を利用し $\mathbf{A}$ の推定量 $\hat{\mathbf{A}}$ を得る
 - 5: $\tilde{\mathbf{W}} = (\hat{\mathbf{G}}^\top \hat{\mathbf{A}})^{-1} (\hat{\mathbf{G}}^\top \hat{\mathbf{G}})$
 - 6: $\tilde{\mathbf{W}}$ を適切に行置換、スケール変換を行い $\mathbf{I} - \mathbf{B}$ の推定量 $\hat{\mathbf{W}}$ を得る。
 - 7: $\hat{\mathbf{B}} = \mathbf{I} - \hat{\mathbf{W}}$
-

5 数値実験

提案手法の性質を確認するため、既存手法 [3] との比較実験を行った。一様乱数から生成したパス係数の行列 $\mathbf{B}$ と $\mathbf{G}$ を用いて図 1 の構造に従う人工データを生成し、潜在変数間のパス係数の推定を行った。なお、 $\mathbf{G}$ の推定結果は両手法で同じものを用いている。標本数の変化に対する係数の RMSE (500 回の試行の平均) をプロットした結果を示す。

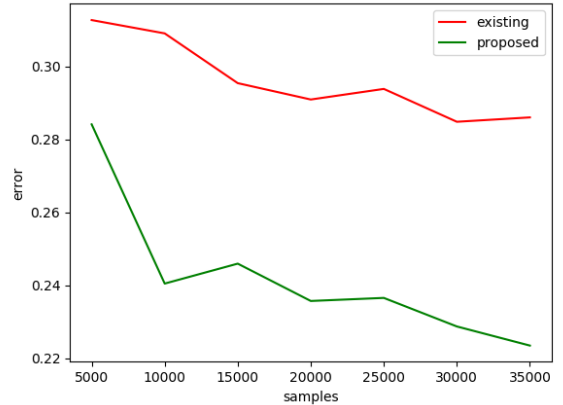


図 2. 標本数を変化させた場合の推定精度の変化。赤線は既存手法 [3] で緑線は提案手法。

図 2 の結果から、既存手法と比較して提案手法の方が推定誤差が小さく、標本数の増加とともに推定精度が向上することが確認された。

参考文献

- [1] A. Hyvärinen. Fast ICA for noisy data using gaussian moments. *IEEE International Symposium on Circuits and Systems*, Vol 5, pp. 57–61, 1999.
- [2] A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, Vol. 10, No. 3, pp. 626–634, 1999.
- [3] S. Shimizu, P. O. Hoyer and A. Hyvärinen. Estimation of linear non-Gaussian acyclic models for latent factors. *Neurocomputing*, Vol 72, No. 7-9, pp. 2024–2027, 2009.
- [4] S. Shimizu, P. O. Hoyer, A. Hyvärinen and A. Kerminen. A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, Vol. 7, pp. 2003–2030, 2006.
- [5] R. Silva, R. Scheine, C. Glymour and P. Spirtes. Learning the structure of linear latent variable models. *Journal of Machine Learning Research*, Vol 7, Feb, pp. 191–246, 2006.