修士論文 要旨

マルコフ同値類を考慮したガウシアン有向グラフィカルモデルの推定

数理情報学専攻 48196207　荻野 正親
指導教員　清智也 准教授

1　はじめに

有向グラフィカルモデルは, 変数間に成り立つ条件付き独立性を DAG(directed acyclic graph, 有向非巡回グラフ) によって表現するものである. DAG の構造から導かれる変数集合間の全ての条件付き独立性は大域的マルコフ性と呼ばれる. ここで, 異なる DAG が等しい大域的マルコフ性を規定する場合があり, 大域的マルコフ性により DAG に同値類を与えたものはマルコフ同値類と呼ばれる. 特殊な状況を除いて, データからはマルコフ同値類以上の推定がなされるべきではないことが述べられてきた.

本研究では, この観点から DAG Wishart 分布 [1] に対するハイパーパラメータの考察を行う. またマルコフ同値類を考慮した MCMC(マルコフ連鎖モンテカルロ法) の手法について提案する.

2　マルコフ同値類の特徴

定理 1 ([2]). 2 つの DAG がマルコフ同値である必要十分条件はスケルトンと v 構造の集合が等しいことである.

定理 2 ([3]). 2 つの DAG がマルコフ同値である必要十分条件はスケルトンが等しく coverd edge の反転のみで移りあえることである.

3　尤度等価性

DAG に対してある値を返す関数に対して, マルコフ同値な DAG に同じ値を与える性質を尤度等価性という.

4　DAG Wishart 分布と, DAG の事後確率が尤度等価性を持つハイパーパラメータ設定

DAG $\mathscr{D}$ に対して,
$\mathbb{D}_+^p := \{D \in \mathbb{R}^{p\times p} \mid D$ は全対角成分が正の対角行列 $\}$
$L_{\mathscr{D}} := \{L \in \mathbb{R}^{p\times p} \mid L$ は全対角成分が 1 かつ
$i \notin \mathrm{pa}(j)$ となる $i \neq j$ について, $L_{ij} = 0\}$

を定める. $(D, L) \in (\mathbb{D}_+^p \times L_{\mathscr{D}})$ に対する以下の分布

$$\pi_{U,\alpha}^{\Theta_{\mathscr{D}}}(D, L) := \frac{\exp\left\{-\frac{1}{2}\mathrm{tr}((LD^{-1}L^\top)U)\right\}\prod_{i=1}^p D_{ii}^{-\frac{1}{2}\alpha_i}}{z_{\mathscr{D}}(U, \alpha)}$$

を DAG Wishart 分布と呼ぶ. D, L はガウス分布の精度行列の修正コレスキー分解に対応する.

DAG Wishart 分布は正定値行列 U とベクトル α をハイパーパラメータに持つ. その望ましい設定について以下のことを示した.

定理 3. DAG の事前分布を尤度等価なものとする. 各 $\mathscr{D}$ において (D, L) の事前分布を, $U_{\mathscr{D}} = U, \{\alpha_{\mathscr{D}}\}_i = 2 * \mathrm{pa}_i + d$ とした DAG Wishart 分布 $\pi_{U_{\mathscr{D}},\alpha_{\mathscr{D}}}^{\Theta_{\mathscr{D}}}$ とする. ここで, $U, d(> 2)$ は $\mathscr{D}$ によらない値とする. このとき観測 Y の後の DAG の事後分布 $p(\mathscr{D}|Y)$ は尤度等価である.

この尤度等価な事後確率は次に述べるマルコフ同値類を考慮した MCMC を行う際に使用できる.

5　有向グラフィカルモデルの MCMC

有向グラフィカルモデルにおいて, ある特徴要素 (e.g. ある変数間の直接の因果効果の有無) を推定する際, MCMC によって得た DAG のサンプルに従って特徴要素の有無を重みづけて推定を行う手法が一つの有力な手法である. ただし, DAG に一様事前分布を仮定するようなサンプリング手法は, 多くの DAG が属するマルコフ同値類を過大評価することにつながる. ここでは, マルコフ同値類に対して一様事前分布を仮定できるような MCMC, 特にメトロポリスヘイスティング法の手法について論じる.

5.1　既存研究

マルコフ同値類は, ある無向辺と有向辺からなるグラフ (completed PDAG [4]) に対応づけられることが知られている. [5] による手法は, completed PDAG を対象とした MCMC を行うことでマルコフ同値類に対する一様事前分布を実現する. 近傍への遷移は completed PDAG に対する局所構造の変化によって行う. ただし, メトロポリスヘイスティング法の運用のため遷移が後退可能なものとなるように, 一部の遷移が制限されてい

る．このため, 近傍空間が小さいものになるという懸念がある．

5.2 提案手法

以上の背景を踏まえ, 本研究ではマルコフ同値類を対象とする別の近傍遷移による MCMC を提案する．提案手法における遷移は以下の流れによる．

1. 現在のマルコフ同値類 C に属する DAG $\mathscr{D}$ を一様ランダムに選ぶ．
2. DAG $\mathscr{D}$ に対し, 局所的な辺構造を変化させる作用を与え, グラフ $\mathscr{D}'$ を得る．
3. $\mathscr{D}'$ が DAG である場合には, $\mathscr{D}'$ の属するマルコフ同値類を C' とし, 採択率 $\alpha_{CC'}$ で C' への遷移を行う．採択されなかった場合, または $\mathscr{D}'$ が DAG でない場合, 元のマルコフ同値類 C を遷移先とする．

手順 1 において, マルコフ同値類内の DAG の一様サンプリングは, マルコフ同値類内の DAG の全探索によって行う．これは定理 2 に基づき covered edge の探索により行える．

手順 2 では, 2 頂点 i, j を一様ランダムに選び, i, j 間に辺があれば辺の削除または反転の一方, i, j 間に辺がなければ両方向の辺の追加のうちの一方, を一様ランダムに選び作用させる．

このとき, 詳細釣り合い条件が成り立つには, 手順 3 での採択率 $\alpha_{CC'}$ を

$$\alpha_{CC'} = \min\left\{1, \frac{p(Y|C')|C|}{p(Y|C)|C'|}\right\}$$

で定めればよい．ただし, $|C|$ はマルコフ同値類 C に属する DAG の数である．遷移が拒否される場合の $|C'|$ の計算を除き, 同値類のサイズの計算は手順 2 を通して行えるため, 余分な計算を必要としない．

近傍について, 以下のことが示される．

定理 4. 提案手法の近傍は PDAG MCMC [5] の近傍を包含する．

6 4 頂点モデルでの MCMC の模擬実験

図 1 の DAG に対するマルコフ性を持つあるガウス分布からデータの観測を行い, DAG Wishart 事前分布を用いて全てのマルコフ同値類の事後確率を求める．この事後分布に収束する MCMC のサンプルを提案手法, PDAG MCMC [5] で得る．MCMC によるマルコフ同値類の分布と真の分布の全変動距離 dif の比較を行う．

観測数 $n = 0, 100$ における結果を示す．

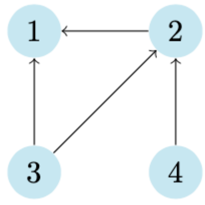

図 1. 真の DAG

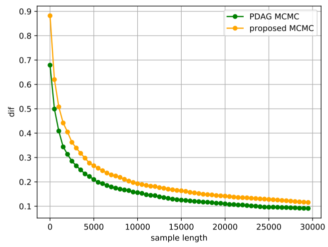

図 2. PDAG MCMC との比較 (n=0)

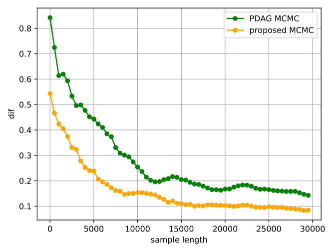

図 3. PDAG MCMC との比較 (n=100)

観測数が大きくなると, 提案手法がより良い収束性を示した．確率上位のマルコフ同値類間の近傍関係を比較することにより, これは近傍空間の拡張により多峰性の問題が改善されたためであることが示唆された．

参考文献

[1] Emanuel Ben-David, Tianxi Li, Helene Massam, and Bala Rajaratnam. High dimensional Bayesian inference for Gaussian directed acyclic graph models. *arXiv preprint arXiv:1109.4371*, 2011.

[2] Thomas Verma and Judea Pearl. An algorithm for deciding if a set of observed independencies has a causal explanation. In *Uncertainty in Artificial Intelligence*, pp. 323–330. Elsevier, 1992.

[3] David Maxwell Chickering. A transformational characterization of equivalent Bayesian network structures. *arXiv preprint arXiv:1302.4938*, 2013.

[4] David Maxwell Chickering. Learning equivalence classes of Bayesian-network structures. *Journal of Machine Learning Research*, Vol. 2, No. Feb, pp. 445–498, 2002.

[5] Yangbo He, Jinzhu Jia, and Bin Yu. Reversible MCMC on Markov equivalence classes of sparse directed acyclic graphs. *The Annals of Statistics*, Vol. 41, No. 4, pp. 1742–1779, 2013.