修士論文 要旨

ガウシアングラフィカルモデルの構造保存群の従順性

数理情報学専攻 48-166220　鈴木大智
指導教員　駒木文保 教授

1　はじめに

ガウシアングラフィカルモデルとは，多変量正規分布の精度行列の非零成分に辺を割り当てたグラフを考えることで変数間の関係を視覚的に表現するモデルである．ガウシアングラフィカルモデルの精度行列の推定は与えられた変数間の関係を見出す上で重要である．ガウシアングラフィカルモデルにたいして，構造を保存して作用する群 G に着目すると，G が従順性を持っているときはミニマックスの意味で良い推定量が構成できることが知られている．さらに，G 上の右不変分布を事前分布としたベイズ解が最良共変推定量と一致することが知られている．本研究ではガウシアングラフィカルモデルに作用する群が，従順群であるための必要十分条件を求め，さらに構造保存群が推移的にモデルに作用している場合の右不変分布と左不変分布の具体形を導出した．

2　既存研究

無向グラフ $\mathcal{G} = ([m], E)$ が与えられたとき，$\mathcal{G}$ に対応する精度行列集合 $\mathcal{S}_{\mathcal{G}}^{+}$ を

$$\mathcal{S}_{\mathcal{G}}^{+} = \bigl\{K \in \mathrm{GL}(\mathbf{R}^m) \mid K \text{ は正定値}, (i,j) \notin E \Rightarrow K_{ij} = 0\bigr\}$$

により定める．ガウシアングラフィカルモデルとは，$\mathcal{S}_{\mathcal{G}}^{+}$ の元を精度行列に持つような多変量正規分布の集合

$$\mathcal{M} = \{p_K(\cdot) \mid K \in \mathcal{S}_{\mathcal{G}}^{+}\}$$

である．X が平均 0, 精度行列 K の正規分布に従うとき，$g \in \mathrm{GL}(\mathbf{R}^m)$ について，gX は平均 0, 精度行列 $(g^{-1})^{\top} K g^{-1}$ の正規分布に従う．よって，ガウシアングラフィカルモデルに対する g の作用を $g \bullet K = (g^{-1})^{\top} K g^{-1})^{-1}$ により定めることができる．$g \bullet \mathcal{S}_{\mathcal{G}}^{+} \subseteq \mathcal{S}_{\mathcal{G}}^{+}$ であるような g の全体 G はガウシアングラフィカルモデルの構造を保存している群であると考えられ，Draisma ら [2] は G の具体形を次のように求めた．まず，$[m]$ 上の擬順序 $\preceq$ を

$$i \preceq j \Leftrightarrow \mathcal{N}(j) \subseteq \mathcal{N}(i) \quad (i,j \in [m])$$

により定める．ただし，$\mathcal{N}(i)$ は i の $\mathcal{G}$ 上の近傍である．そのもとで，

$$G^0 = \{g \in \mathrm{GL}(\mathbf{R}^m) \mid j \not\preceq i \Rightarrow g_{ij} = 0\}$$

と定義する． すると，G は G^0 と $\mathcal{G}$ の自己同型群 $\mathrm{Aut}(\mathcal{G})$ を用いて，$G = \mathrm{Aut}(\mathcal{G})G^0$ と書ける．$\mathcal{G}$ の頂点番号を適当に付け替えることによって，G^0 の元はブロック下三角行列にできることが知られている．

モデルに対して群 G が作用しているとき，作用に対して不変な推定量（共変推定量）を考えることが自然である．群の作用と推定量の関係を考える上で重要な性質の一つが群の従順性である．

定義 1. G を局所コンパクト群，$L^{\infty}(G)$ を G の左不変測度について測度 0 の集合を除いて有界であるような関数 $G \to \mathbf{C}$ がなす線形空間とする．G が従順とは，次を満たす線形関数 $\mu : L^{\infty}(G) \to \mathbf{C}$ が存在することである．

(1) $\mu(\mathbf{1}) = 1$.
(2) $\|\mu\| := \sup_{\|\varphi\|_{\infty} \leq 1} |\mu(\varphi)| = 1$.
(3) ${}_g f(g') := f(gg')$ で定める ${}_g f$ について，$\mu({}_g f) = \mu(f)$.

Kiefer [3] や Wesler [4] は，G が従順性を持つならばミニマックスの意味で良い推定量を共変推定量の中から見つけ出す事ができることを述べている．また，ミニマックスの意味で良い推定量は，構造保存群上の右不変測度に基づく事前分布を用いたベイズ解として構成することができることも知られている (cf.[1]).

3　ガウシアングラフィカルモデルの構造保存群の従順性

以上の背景を踏まえ，ガウシアングラフィカルモデルの構造保存群が従順であるためのグラフの構造に関する次のような必要十分条件を求めた．

定理 1. グラフ $\mathcal{G}$ に基づくガウシアングラフィカルモデルの構造を保存する群を G とする．G が従順となる必要十分条件は $\mathcal{G}$ の任意の辺 $\{i,j\}$ について，頂点 $k \neq i,j$ であって，$\{i,k\}$ と $\{j,k\}$ のうちどちらか一方のみが $\mathcal{G}$ の辺であるようなものが存在することである．

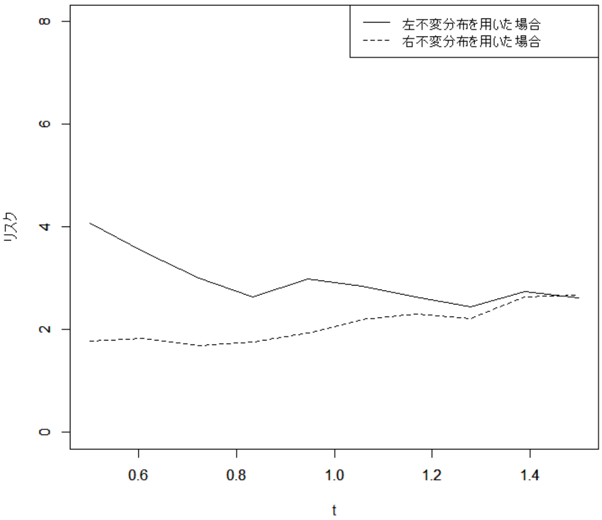

図 1: 右（左）不変分布に基づいた分散共分散行列の事後平均のリスク

G が従順なとき，G^0 の元は下三角行列にすることができる．また，各 $g \in G$ に対して，$g = \sigma_{\mathrm{L}} h_{\mathrm{L}} = h_{\mathrm{R}} \sigma_{\mathrm{R}}$ なる $\sigma_{\mathrm{L}}, \sigma_{\mathrm{R}} \in \mathrm{Aut}(\mathcal{G})$, $h_{\mathrm{L}}, h_{\mathrm{R}} \in G^0$ が一意的に存在する．これらを g の関数 $\sigma_{\mathrm{L}}(g), \sigma_{\mathrm{R}}(g), h_{\mathrm{L}}(g), h_{\mathrm{R}}(g)$ とみなす．すると，次のように G 上の不変分布が得られる．

定理 2. G 上の左不変分布は

$$\pi_{\mathrm{L}}(g) = \frac{1}{\prod_{i=1}^{m} (h_{\mathrm{L}}(g)_{ii})^{i}}$$

と書ける．また，G 上の右不変分布は

$$\pi_{\mathrm{R}}(g) = \frac{1}{\prod_{i=1}^{m} (h_{\mathrm{R}}(g)_{ii})^{m-i+1}}$$

と書ける．

4 数値実験

$\bullet$ —— $\bullet$ —— $\bullet$ のような形をしたグラフ $\mathcal{G}$ について，右（左）不変分布に基づいた事後分布 $\pi_{\mathrm{L}}(g \mid X), \pi_{\mathrm{R}}(g \mid X)$ を用いて分散共分散行列の事後平均を数値的に計算し，右不変分布を用いた場合と左不変分布を用いた場合のリスクを比較する．精度行列の損失関数として，Kullback-Leibler ダイバージェンス

$$\begin{aligned} L(K, K') &= \int p(x \mid K) \log \frac{p(x \mid K)}{p(x \mid K')} dx \\ &= \frac{1}{2} \left\{ \log |K| - \log |K'| + \mathrm{tr}(K' K^{-1}) - m \right\} \end{aligned}$$

を用いると，$L(g \bullet K, g \bullet K') = L(K, K')$ が成り立つ．観測を生み出すパラメータを

$$g_t = \begin{pmatrix} t & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & 0 & 1 \end{pmatrix} \qquad (0.5 \leq t \leq 1.5)$$

として，t の変化に従ってリスクがどのように変化するのかを調べる．まず，$i = 1, \cdots, 200$ について，$X_i \in \mathbf{R}^{3 \times 10}$ を平均 $(0, \cdots, 0)^{\top}$, 分散共分散行列 $g_t g_t^{\top}$ の正規分布にしたがって発生させた 5 個の乱数とする．X_i を用いて，事後分布 $\pi_{\mathrm{L}}(g \mid X_i)$, $\pi_{\mathrm{R}}(g \mid X_i)$ から，1000 個の標本 $g_{\mathrm{L}ij}, g_{\mathrm{R}ij}$ $(j = 1, \cdots, 1000)$ をサンプリングして，分散共分散行列のの事後平均

$$\begin{aligned} \frac{1}{1000} \hat{\Sigma}_{\mathrm{L}t}(X_i) &= \sum_{j=1}^{1000} (g_{\mathrm{L}ij} \bullet I_m)^{-1}, \\ \frac{1}{1000} \hat{\Sigma}_{\mathrm{R}t}(X_i) &= \sum_{j=1}^{1000} (g_{\mathrm{R}ij} \bullet I_m)^{-1} \end{aligned}$$

を求めた．各 X_i について損失 $L(g_t \bullet I_m, \hat{\Sigma}_{\mathrm{L}t}(X_i)^{-1})$, $L(g_t \bullet I_m, \hat{\Sigma}_{\mathrm{R}t}(X_i)^{-1})$ を求め，その i に関する平均を計算した結果を分散共分散行列の事後平均のリスクとする．図 1 は t とリスクとの関係を表したグラフである．右不変分布を用いた場合のリスクは理論的には定数リスクであるが，図 1 をみると，ばらつきが生じている．これは，サンプリングを行う際に MCMC を用いているため，サンプル間に依存関係が生じていることが一因として考えられる．また，事後分布関数は多峰性をもつため，サンプリングを行う際に局所解に陥ってしまった可能性も考えられる．実際に，サンプル数を変化させて実験行っても得られるグラフの形状はほとんど変化しなかった．一方で，全体的に右不変分布を用いた場合のリスクは，左不変分布を用いた場合のリスクよりも小さく，また，定数リスクに近いことを読み取ることができる．

参考文献

[1] J. Berger. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.
[2] J. Draisma, S. Kuhnt, and P. Zwiernik. Groups acting on Gaussian graphical models. *The Annals of Statistics*, Vol. 41, pp. 1944–1969, 2013.
[3] J. Kiefer. Invariance, minimax sequential estimation, and continuous time processes. *The Annals of Mathematical Statistics*, Vol. 28, pp. 573–601, 1957.
[4] O. Wesler. Invariance theory and a modified minimax principle. *The Annals of Mathematical Statistics*, Vol. 30, pp. 1–20, 1959.