

構造的勾配モデルの許容領域について

東京大学・情報理工学系研究科 清 智也 (Tomonari SEI)
School of Information Science and Technology,
The University of Tokyo

概要

多変量解析のための統計モデルとして提案された構造的勾配モデル (SGM) について, その許容領域を考察する. SGM は, 基準となる確率分布をパラメトリックな勾配写像で引き戻してできる多変量分布族である. 本稿では基準分布を単位超立方体上の一様分布に限定して考える. また実データに対する解析例について述べる.

キーワード グラフィカルモデル, 構造的勾配モデル, フーリエ解析.

1 構造的勾配モデル (SGM) とは

m を固定した正の整数とし, $\mathbb{R}^m$ の勾配作用素を $\nabla = (\partial/\partial x_i)_{i=1}^m$, ヘシアン作用素を $\nabla\nabla^\top = (\partial^2/\partial x_i\partial x_j)_{i,j=1}^m$ と表記する. また行列式を $\det$ と書き, 正定値性を $\succ$, 半正定値性を $\succeq$ で表す. また $\mathbb{Z}_{\geq 0}$ を非負整数全体とする.

定義 1. $\mathbb{Z}_{\geq 0}^m$ の有限部分集合 $\mathcal{U}$ を固定する. 構造的勾配モデル (SGM) $\mathcal{M}$ を次のように定義する.

$$\begin{aligned}\mathcal{M} &= \{p(\cdot|\theta) \mid \theta \in \Theta\}, \\ p(x|\theta) &= \det(\nabla\nabla^\top \psi(x|\theta)), \\ \psi(x|\theta) &= \frac{1}{2}x^\top x + \sum_{u \in \mathcal{U}} \theta_u \prod_{j=1}^m \cos(\pi u_j x_j), \\ x &= (x_j)_{j=1}^m \in [0, 1]^m, \\ \theta &= (\theta_u)_{u \in \mathcal{U}}, \\ \Theta &= \{\theta \mid \nabla\nabla^\top \psi(x|\theta) \succ 0, \quad \forall x \in [0, 1]^m\}.\end{aligned}$$

上で定義される凸集合 Θ を $\mathcal{M}$ の許容領域と呼ぶ. また ψ を SGM のポテンシャル関数と呼ぶ. □

$p(x|\theta)$ は「 $\nabla\psi(X|\theta)$ が一様分布に従うような確率変数 X 」の密度関数である. このように, 勾配写像による密度関数の変数変換に基づいた統計的モデルを一般に勾配モデルと呼ぶ (Sei 2006, 清 2007). 勾配モデルは最適輸送問題 (例えば Villani 2003) と密接な関係がある.

SGM の対数尤度関数 $\log p(x|\theta)$ はパラメータ θ に関して凹関数であるため, 最尤法や事後確率最大化において有利である. しかし, その許容領域 Θ は無限個の制約式を含んでいるため, そのままの形で最適化することは難しい. そこで, 本研究では Θ を内側から近似することを目的とする.

本稿の構成は以下の通りである. まず 2 節において SGM が自然な条件で特徴付けられることを説明する. 3 節, 4 節で主要結果を示す. また応用として, 近似的グラフィカルモデリングを 5 節で述べる. 最後にまとめと今後の課題を 6 節で述べる.

2 SGM の特徴づけ

超立方体 $[0, 1]^m$ から $[0, 1]^m$ へ移す $\nabla\psi$ がフーリエ級数で特徴付けられることを示す. このことから, SGM は自然な条件から導出されていることが分かる.

補題 1. C^2 級の凸関数 $\psi : [0, 1]^m \rightarrow \mathbb{R}$ に対し, 勾配写像 $\nabla\psi : [0, 1]^m \rightarrow \mathbb{R}^m$ が $[0, 1]^m$ への微分同相写像であると仮定する. このとき, ψ は次のノイマン条件を満たす:

$$(\nabla\psi(x))_j = x_j \quad \text{if } x_j \in \{0, 1\}.$$

Proof. まず, 端点 $x \in \{0, 1\}^m$ に対して $\nabla\psi(x) = x$ であることを示そう. $y = \nabla\psi(x)$ とおき, $y \neq x$ と仮定する. すると, $\nabla\psi$ の全単射性から $\nabla\psi(z) = x$ を満たす $z \in [0, 1]^m$ が存在する. x は超立方体の端点であるから, 各 $j = 1, \dots, m$ に対して $(y_j - x_j)(z_j - x_j) \geq 0$ でなければならず, したがって

$$(\nabla\psi(x) - \nabla\psi(z))^T(x - z) = (y - x)^T(x - z) \leq 0$$

が成り立つ. 一方, ψ は狭義凸であり, また $z \neq x$ であるから

$$(\nabla\psi(x) - \nabla\psi(z))^T(x - z) > 0$$

が成り立つ. これは矛盾である. したがって $y = x$ であることが示された. 次に $[0, 1]^m$ の面 (face) F を考える. $\nabla\psi : [0, 1]^m \rightarrow [0, 1]^m$ は微分同相だから, $\nabla\psi(F)$ は $[0, 1]^m$ のどれかの面でなければならない. ところが, 端点 ($\in \{0, 1\}^m$) はそれ自身に写されるから, $\nabla\psi(F) = F$ でなければならない. よって $x_j \in \{0, 1\}$ である限り $(\nabla\psi(x))_j = x_j$ が成り立つ. $\square$

補題 2. C^2 級の凸関数 $\psi : [0, 1]^m \rightarrow \mathbb{R}$ に対し, 勾配写像 $\nabla\psi : [0, 1]^m \rightarrow \mathbb{R}^m$ が $[0, 1]^m$ への微分同相写像であると仮定する. このとき, ψ は次のフーリエ級数表示を持つ:

$$\psi(x) = \frac{1}{2}x^T x + \sum_{u \in \mathbb{Z}_{\geq 0}^m} \theta_u \prod_{j=1}^m \cos(\pi u_j x_j).$$

Proof. 関数 $\phi(x) = \psi(x) - x^T x/2$ を考える. ϕ は境界条件 $(\nabla\phi(x))_j = 0$ (if $x_j \in \{0, 1\}$) を満たす. さらに ϕ の周期的拡張 $\tilde{\phi}$ (各変数に対して周期 2) を

$$\tilde{\phi}(x) = \phi(|x_1|, \dots, |x_m|) \quad \text{if } x \in [-1, 1]^m$$

により定義すれば $\tilde{\phi}$ はいたるところ連続的微分可能であり, 各変数に対して偶関数である. よって $\tilde{\phi}$ はそのフーリエ級数に一致する:

$$\tilde{\phi}(x) = \sum_{u \in \mathbb{Z}_{\geq 0}^m} \theta_u \prod_{j=1}^m \cos(\pi u_j x_j).$$

$[0, 1]^m$ においては $\tilde{\phi} = \phi$ であるから, 補題が示されたことになる. $\square$

3 許容領域の評価 ($m = 1$ の場合)

$m = 1, \mathcal{U} = \{1, \dots, U\}$ の SGM を考え, 許容領域 Θ を評価しよう. 以下では簡単のため $\psi''(x|\theta)$ を単に $\psi''(x)$ と書く. この場合, Θ は

$$\Theta = \{\theta \mid \psi''(x) > 0 \quad (\forall x \in [0, 1])\}, \quad \psi''(x) = 1 - \sum_{u=1}^U \theta_u \pi^2 u^2 \cos(\pi u x)$$

と書ける. 集合 Θ^{lit} を次式で定義する:

$$\Theta^{\text{lit}} := \left\{ \theta \mid 1 - \sum_{u=1}^U |\theta_u| \pi^2 u^2 > 0 \right\}.$$

補題 3. $\Theta^{\text{lit}} \subset \Theta$ である.

Proof. 不等式 $\psi''(x) \geq 1 - \sum_{u=1}^U |\theta_u| \pi^2 u^2$ ($\forall x \in [0, 1]$) より明らか. $\square$

整数 $M \geq U + 1$ に対して関数 $\psi''_M(x)$ を

$$\psi''_M(x) := 1 - \sum_{u=1}^U \frac{\theta_u \pi^2 u^2}{1 - u/M} \cos(\pi u x)$$

と定義する. また, 幅 $1/M$ の格子を $L_M = \{v/M\}_{v=0}^M$ とおき, 集合

$$\Theta_M = \{\theta \mid \psi''_M(x) > 0 \quad (\forall x \in L_M)\}$$

を考える. Θ_M は有限個の線形不等式制約からなっていることに注意する. 定理 6 で, Θ_M が Θ の内側からの近似になることを示す. まず, フーリエ級数の正値性に関する Bochner の定理を述べておく (詳細は Feller 1971 参照).

補題 4. M を正の整数とする. $(f_u)_{u \in \mathbb{Z}}$ を周期 $2M$ の実数列とし, $f_{-u} = f_u$ および $f_M = 0$ を満たすものとする. このとき, $\sum_{|u| \leq M-1} f_u \cos(\pi u x)$ が全ての $x \in L_M$ について正になるための必要十分条件は $(f_{u-v})_{u,v \in \mathbb{Z}}$ が正定値であることである.

補題 5. U を正の整数とする. $(f_u)_{u \in \mathbb{Z}}$ を実数列とし, $f_{-u} = f_u$ および $f_u = 0$ (if $|u| \geq U + 1$) を満たすものとする. このとき, $\sum_{|u| \leq U} f_u \cos(\pi u x)$ が全ての $x \in [0, 1]$ について正になるための必要十分条件は $(f_{u-v})_{u,v \in \mathbb{Z}}$ が正定値であることである.

定理 6. 次の等式が成り立つ:

$$\Theta = \bigcup_{M_0 \geq U+1} \bigcap_{M \geq M_0} \Theta_M = \bigcup_{M \geq U+1} \Theta_M.$$

Proof. 以下の 3 つの条件が同値であることを示せばよい.

- (i) 任意の $x \in [0, 1]$ に対して $\psi''(x) > 0$.
- (ii) ある $M_0 \geq U + 1$ が存在して, 任意の $M \geq M_0, x \in L_M$ に対して $\psi''_M(x) > 0$.

(iii) ある $M \geq U + 1$ が存在して, 任意の $x \in L_M$ に対して $\psi''_M(x) > 0$.

まず, (i) から (ii) を示す. (i) を仮定する. 区間 $[0, 1]$ のコンパクト性より,

$$\min_{x \in [0, 1]} \psi''(x) > 0$$

が成り立つ. よって, (ii) を示すには, $\psi''_M(x)$ が $\psi(x)$ に一様収束することを示せば十分である. 実際,

$$\begin{aligned} \max_{x \in [0, 1]} |\psi''_M(x) - \psi(x)| &= \max_{x \in [0, 1]} \left| \sum_{u=1}^U \frac{\theta_u \pi^2 u^3}{M - u} \cos(\pi u x) \right| \\ &\leq \frac{1}{M - U} \sum_{u=1}^U |\theta_u| \pi^2 u^3 \\ &\rightarrow 0 \end{aligned}$$

となり, (ii) を得る. 次に, (ii) から (iii) は明らかである. (iii) から (i) を示す. (iii) を仮定する. 周期 $2M$ の数列 $(f_u)_{u \in \mathbb{Z}}$ として次式を満たすものを考える:

$$f_u = \begin{cases} 1 & \text{if } u = 0 \\ -\frac{\theta_{|u|} \pi^2 u^2}{2(1 - |u|/M)} & \text{if } 1 \leq |u| \leq U \\ 0 & \text{if } U + 1 \leq |u| \leq M \end{cases}$$

条件 (iii) と補題 4 より, $(f_{u-v})_{u, v \in \mathbb{Z}}$ は正定値であり, したがって $(f_{u-v} 1_{\{u, v \in \{0, \dots, M-1\}\}})_{u, v \in \mathbb{Z}}$ は半正定値である. さらに, これをずらして平均をとった行列

$$\frac{1}{M} \sum_{t \in \mathbb{Z}} f_{(u-t) - (v-t)} 1_{\{u-t, v-t \in \{0, \dots, M-1\}\}} = \begin{cases} 1 & \text{if } u - v = 0, \\ -\frac{\theta_{|u-v|} \pi^2 (u-v)^2}{2} & \text{if } 1 \leq |u - v| \leq U \\ 0 & \text{if } U + 1 \leq |u - v| \end{cases}$$

は正定値である. よって, 補題 5 から条件 (i) が成り立つ. $\square$

以下, これまでの結果を具体例で考えてみる. 式の簡単のため $F_u := -\theta_u \pi^2 u^2$ とおく.

例 1 (標本化の影響). $U = 3$, $F_1 = 0$, $F_2 = 0.5$, $F_3 = 1$ とおくと, 関数 ψ'' は

$$\psi''(x) = 1 + 0.5 \cos(2\pi x) + \cos(3\pi x)$$

となる. この ψ'' は閉区間 $[0, 1]$ において正値関数でない (実際, 例えば $\psi''(1/3) = -1/4$ である). 一方, 格子点 $L_4 = \{v/4\}_{v=0}^4$ においては関数値がそれぞれ

$$\psi''(0) = \frac{5}{2}, \quad \psi''\left(\frac{1}{4}\right) = 1 - \frac{1}{\sqrt{2}}, \quad \psi''\left(\frac{1}{2}\right) = \frac{1}{2}, \quad \psi''\left(\frac{3}{4}\right) = 1 + \frac{1}{\sqrt{2}}, \quad \psi''(1) = \frac{1}{2}$$

となり, 正値である. つまり単なる標本化では, 許容性の十分条件が得られないことが分かる. $\square$

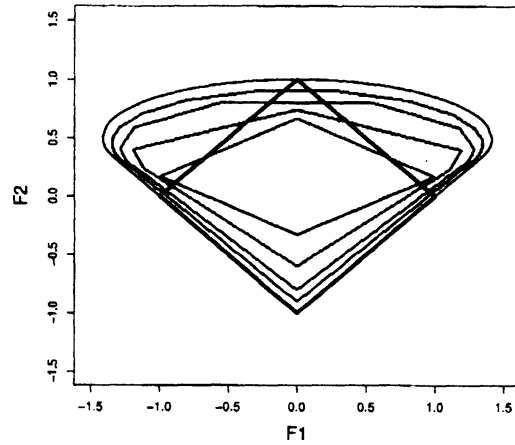


図 1: 許容領域を座標 (F_u) で表した図. 細線は, 内側から順に $M = 3, 5, 10, 20, \infty$ (厳密解) に対応する Θ_M の境界を表す. 太線は, Θ^{lit} の境界を表す.

例 2. $m = 1, U = 2, \mathcal{U} = \{1, 2\}$ の場合, 許容領域は次式で与えられる:

$$|F_1| < 1 + F_2 \leq \frac{|F_1|}{4} \quad \text{or} \quad \left(F_2 - \frac{1}{2}\right)^2 + \frac{F_1^2}{8} < \frac{1}{4}.$$

許容領域 Θ, Θ_M ($M = 3, 5, 10, 20$), Θ^{lit} を図 1 に示す.

例 3. $m = 1, \mathcal{U} = \{1, 3\}$ の場合の許容領域を図 2 (a) に, また $\mathcal{U} = \{1, 4\}$ の場合の許容領域を図 2 (b) に示す.

4 許容領域の評価 ($m > 1$ の場合)

一般の SGM ($m \geq 1, \mathcal{U} \subset \mathbb{Z}_{\geq 0}^m$ の場合) を考え, 許容領域 Θ を評価しよう. 以下では簡単のため $\nabla \nabla^\top \psi(x|\theta)$ を単に $\nabla \nabla^\top \psi(x)$ と書く. 許容領域 Θ は

$$\Theta = \left\{ \theta \mid \nabla \nabla^\top \psi(x) \succ 0 \quad (\forall x \in [0, 1]^m) \right\}, \quad \psi(x) = \frac{x^\top x}{2} + \sum_{u \in \mathcal{U}} \theta_u \prod_{j=1}^m \cos(\pi u_j x_j)$$

であった. 集合 Θ^{lit} を次式で定義する:

$$\Theta^{\text{lit}} := \left\{ \theta \mid 1 - \sum_{u \in \mathcal{U}} |\theta_u| \pi^2 u_j^2 > 0 \quad (\forall j = 1, \dots, m) \right\}.$$

$m = 1$ の場合は前節で定義した Θ^{lit} に一致することに注意する.

定理 7. $\Theta^{\text{lit}} \subset \Theta$ である.

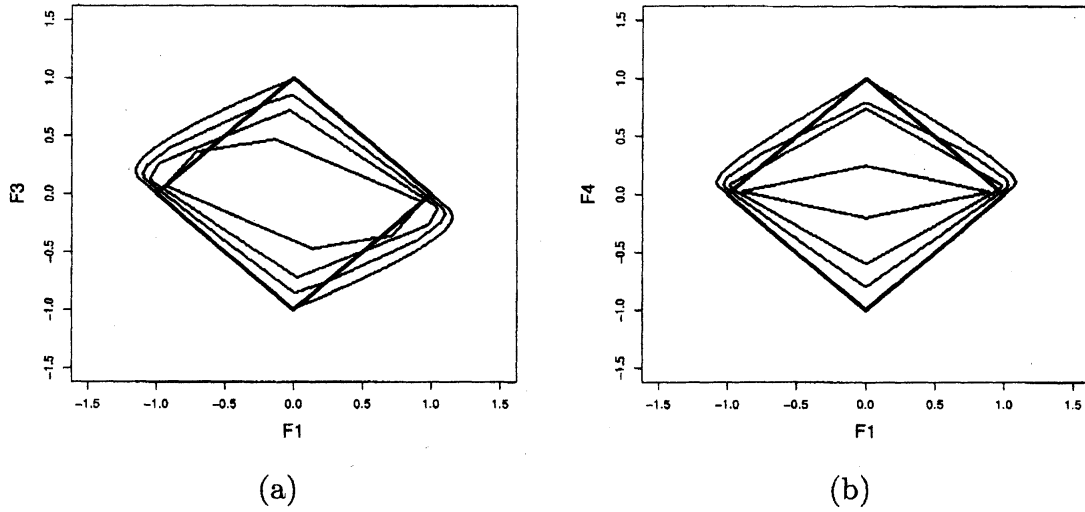


図 2: 許容領域を座標 (F_u) で表した図. 細線は, 内側から順に $M = 5, 10, 20, 500$ に対応する Θ_M の境界を表す. 太線は, Θ^{lit} の境界を表す. (a) $\mathcal{U} = \{1, 3\}$ の場合, (b) $\mathcal{U} = \{1, 4\}$ の場合.

Proof. $\theta \in \Theta^{\text{lit}}$ とする. 任意の $x \in [0, 1]^m$ に対して $\nabla \nabla^\top \psi(x) \succ 0$ を示せばよい. まず $\cos$ の積を和に直すと

$$\prod_{j=1}^m \cos(\pi u_j x_j) = 2^{-m} \sum_{b \in \{-1, 1\}^m} \cos(\pi b^\top D(u)x)$$

となる. ただし $D(u)$ は u を対角成分とする対角行列である. よって

$$\begin{aligned} \nabla \nabla^\top \psi(x) &= \nabla \nabla^\top \left(\frac{1}{2} x^\top x + \sum_{u \in \mathcal{U}} \theta_u 2^{-m} \sum_{b \in \{-1, 1\}^m} \cos(\pi b^\top D(u)x) \right) \\ &= I_m - \sum_{u \in \mathcal{U}} \theta_u 2^{-m} \pi^2 \sum_{b \in \{-1, 1\}^m} \cos(\pi b^\top D(u)x) D(u) b b^\top D(u) \\ &\succeq I_m - \sum_{u \in \mathcal{U}} |\theta_u| 2^{-m} \pi^2 \sum_{b \in \{-1, 1\}^m} D(u) b b^\top D(u) \\ &= I_m - \sum_{u \in \mathcal{U}} |\theta_u| \pi^2 D(u)^2 \\ &\succ 0 \end{aligned}$$

を得る. ただし I_m は m 次単位行列である. □

次に, $\mathcal{U}$ を特殊な場合に限定すると, Θ^{lit} と Θ が一致することを示す.

定理 8. $\mathcal{U} = \{u(1), \dots, u(m)\}$ とし, $u(1), \dots, u(m)$ は modulo 2 の意味で一次独立と仮定する. このとき $\Theta^{\text{lit}} = \Theta$ である.

Proof. 前の定理より, $\Theta \subset \Theta^{\text{lit}}$ を示せばよい. $\theta = (\theta_u)_{u \in \mathcal{U}} \in \Theta$ とする. 格子点 $x \in \{0, 1\}^m$ において $\nabla \nabla^\top \psi(x)$ を評価しよう. 格子点では $\partial_j^2 (\prod_{j=1}^m \cos(\pi u_j x_j)) = -\pi^2 u_j^2 (-1)^{u^\top x}$ となることに注意する. また, $\mathcal{U}$ の一次独立性より,

$$u^\top x \equiv 1_{\{\theta_u < 0\}} \pmod{2} \quad \forall u \in \mathcal{U}$$

を満たす $x \in \{0, 1\}^m$ が存在する. この x に対して

$$\begin{aligned} 0 &< (\nabla \nabla^\top \psi(x))_{jj} \\ &= 1 - \sum_{u \in \mathcal{U}} \theta_u \pi^2 u_j^2 (-1)^{u^\top x} \\ &= 1 - \sum_{u \in \mathcal{U}} |\theta_u| \pi^2 u_j^2 \end{aligned}$$

が成り立つ. よって $\theta \in \Theta^{\text{lit}}$ である. □

5 応用例：近似的グラフィカルモデリング

$\mathcal{U} = \{0, 1\}^m$ として SGM を考えると, 近似的にグラフィカルモデルを構成することができる. 詳細は Sei (2008) を参照せよ. またグラフィカルモデリングに関しては宮川 (1997) を参照せよ.

例えば, $m = 3$ としてポテンシャル関数

$$\psi(x|\theta) = \frac{1}{2} x^\top x + \theta \cos(\pi x_1) \cos(\pi x_2) \cos(\pi x_3)$$

を考える (簡単のため $\theta_{(1,1,1)}$ を θ と略記した). 対応する密度関数 $p(x|\theta)$ は

$$p(x|\theta) = \det \begin{pmatrix} 1 - \theta \pi^2 c_1 c_2 c_3 & \theta \pi^2 s_1 s_2 c_3 & \theta \pi^2 s_1 c_2 s_3 \\ \theta \pi^2 s_1 s_2 c_3 & 1 - \theta \pi^2 c_1 c_2 c_3 & \theta \pi^2 c_1 s_2 s_3 \\ \theta \pi^2 s_1 c_2 s_3 & \theta \pi^2 c_1 s_2 s_3 & 1 - \theta \pi^2 c_1 c_2 c_3 \end{pmatrix}$$

と書ける. ただし $c_i := \cos(\pi x_i)$, $s_i := \sin(\pi x_i)$ とおいた. $|\theta| \ll 1$ として展開すると

$$p(x|\theta) = 1 - 3\theta \pi^2 c_1 c_2 c_3 + O(\theta^2)$$

となり, 特にパラメータ θ と 3 次モーメントは近似的に比例関係を持つ:

$$\int_{[0,1]^3} p(x|\theta) (x_1 - 1/2)(x_2 - 1/2)(x_3 - 1/2) dx = \frac{24\theta}{\pi^4} + O(\theta^2).$$

同様に $\mathcal{U} = \{0, 1\}^m$ のとき, θ_u は変数集合 $\{x_j \mid u_j = 1\}$ のキュムラントと近似的比例関係を持つことが示される.

Mardia et al. (1973) に載っている 5 科目の成績データに関して, 最尤法と AIC に基づきモデル選択した結果を図 3 に示す. SGM のポテンシャル関数は次式で与えられる:

$$\psi(x|\theta) = \frac{1}{2} x^\top x + \sum_{u \in \{0,1\}^5} \theta_u \prod_{j=1}^5 \cos(\pi u_j x_j).$$

最尤法を実行する際, 4 節で定義した Θ^{lit} を利用した. なお, ここでは SGM の基準分布を一様分布でなく正規分布とした (詳細は省略).

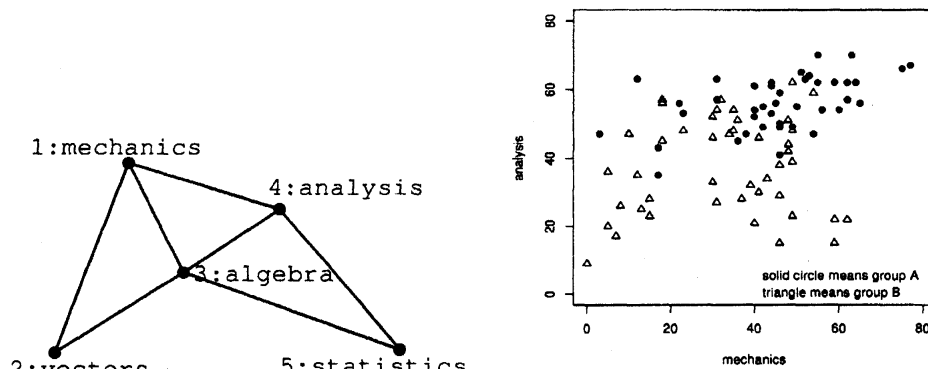


図 3: 成績データの解析例. (a) 選択された無向独立グラフ, (b) algebra の成績で層別した mechanics と analysis の散布図.

6 まとめと今後の課題

本稿では, SGM の許容領域 Θ を内側から評価した. 特に $m = 1$ の場合, 有限個の線形不等式制約 Θ_M の列によって Θ を内側から近似できることが分かった. また $m \geq 1$ の場合に集合 Θ^{lit} を定義し, $\Theta^{\text{lit}} \subset \Theta$ を示した. 今後の課題は, $m > 1$ の場合に Θ を内側から近似する集合列を与えることである.

参考文献

- [1] Feller, (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, John Wiley & Sons, Canada.
- [2] Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*, Academic Press, London.
- [3] 宮川 雅巳 (1997). *グラフィカルモデリング*, 朝倉書店.
- [4] Sei, T. (2006). Parametric modeling based on the gradient maps of convex functions, Technical Report METR2006-51, Department of Mathematical Engineering and Information Physics, The University of Tokyo.
- [5] 清 智也 (2007). Gradient modeling and information geometry, RIMS 共同研究 “Statistical Decision for Multiple Comparison and Its Related Topics”, 数理解析研究所講究録 1560, 109–120.
- [6] Sei, T. (2008). Gradient modeling for multivariate analysis, submitted.
- [7] Villani, C. (2003). *Topics in Optimal Transportation*, AMS, Providence.