科研基盤(B) 「生物情報を解明するための統計数学的基礎理論とその応用」
(研究代表者：赤平昌文(筑波大学)) によるシンポジウム

2007 年 9 月 18 日 – 20 日（於：日本女子大学）

条件付独立構造とウィシャート分布について

日本女子大学理学部　　今野 良彦

September 1, 2007

この講演の背景・目的

- 正値対称行列上の確率分布である古典的 Wishart 分布とその性質を等質錐 (homogeneous cones) 上の分布に一般化を行う研究がある.
  - Ishi, H. (J. Math. Soc. Japan, 2001).
  - Andersson, S.A. and Wojnar, G.G. (JTP, 2004).
  - Boutouria, I. (C. R. Acad. Sci. Ser I, 2005).

- 条件付独立構造をもった多変量正規分布モデルの共分散行列の最尤推定量の分布という観点から一般化された Wishart 分布について整理したことを報告する.
  - Dawid, A.P. and Lauritzen, S.L. (AOS, 1993).
  - Roverato, A. (Biometrika, 2000).
  - Letac, G. and Massam, H. (AOS, 2007).

この講演の構成

- 記号と問題設定について
- 条件付独立性と covariance seleciton models について
- 無向グラフによる graphical Gaussian models (UDG モデル) の必要な用語について
- UDG モデルに関わる開凸錐について
- UDG モデルの共分散行列の最尤推定量と一般化された Wishart 分布について

記号

- $\star$ $\mathbb{R}^{r\times p}$: $r\times p$ の行列の空間.

- $\star$ $\mathrm{Sym_p}$, $\mathrm{Sym_p}^+$: $p\times p$ の対称行列と正定値対称行列の空間

- $\star$ $a\in \mathrm{Sym_p}$ に対し, $\mathrm{Det}\,(a)$ を行列式, $\mathrm{Tr}\,(a)$ をトレース, a' を行列 a の転置.

- $\star$ $C_1,\, C_2\subset V := \{1,\, 2,\, \ldots,\, p\}$ と $a=(a_{ij})_{i,\, j=1,\, 2,\, \ldots,\, p}\in \mathrm{Sym_p}$ に対し,

$$a_{C_1,\, C_2} = (a_{ij})_{i\in C_1,\, j\in C_2}.$$

- $\star$ $C\subset V$ と $a=(a_{ij})\in \mathrm{Sym_p}$ に対し, $a_C := a_{C,\, C}$ と書き, $(a_C)^0$ は $p\times p$ の行列で

$$\{(a_C)^0\}_{ij} = \begin{cases} a_{ij} & \text{if } i,\, j\in C \\ 0 & \text{otherwise} \end{cases}.$$

- $\star$ 集合 C の要素の個数を $|C|$ と書く.

問題設定（１）

- 確率（縦）ベクトル $X_1, X_2, \ldots, X_n$ は独立同一に p 変量正規分布 $N_p(0, 2\,\Sigma)$ に従うとする．ただし，$n \geq p$ とする．ただし，$\Sigma \in \mathrm{Sym}_p^+$.
- Σ の最尤推定量は $\widehat{\Sigma}_{mle} = W/(2n)(W := \sum_{i=1}^n X_i X_i')$．すなわち，

$$\widehat{\Sigma}_{mle} = \mathop{\mathsf{argmax}}_{\Sigma \in \mathrm{Sym}_p^+} \frac{1}{(\mathrm{Det}\,\Sigma)^{n/2}} \exp\left\{-\mathrm{Tr}(\Sigma^{-1} W)\right\}$$

- W は母数 $(p, 2\,\Sigma)$ の Wishart 分布に従う ($w_p(n, 2\,\Sigma)$ と記す)：

$$\frac{(\mathrm{Det}\,w)^{n/2} \exp\{-\mathrm{Tr}(w\Sigma^{-1}\}}{\Gamma_p(n/2)(\mathrm{Det}\,\Sigma)^n} \times \frac{\mathbb{1}_{\mathrm{Sym}_p^+}(w)\,dw}{(\mathrm{Det}\,w)^{(p+1)/2}}$$

ただし，$a > (p-1)/2$ に対し，$\Gamma_p(a) = \pi^{(1/4)p(p-1)} \Pi_{i=1}^p \Gamma(a - (i-1)/2)$ で $\Gamma(\,\cdot\,)$ はガンマ関数.

- この分布を $w_p(n, 2\Sigma)$ と書く.

問題設定（２）

- $X_1, X_2, \ldots, X_n$ は独立同一に p 変量正規分布 $N_p(0, 2\Sigma)$ に従うとする．

- Covariance Selection Models　$K = (k_{ij}) = \Sigma^{-1}$ の適当な非対角成分は 0 であると仮定する．たとえば，$p = 3$ のとき，下のような

$$K = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & 0 \\ k_{31} & 0 & k_{33} \end{pmatrix}$$

- $p \geq 3$ の整数とし，制限されて Σ の最尤推定量（存在すれば）の分布を考える：

$$\widehat{\Sigma}_{mle} = \underset{\text{restricted } \Sigma}{\operatorname{argmax}} \frac{1}{(\mathrm{Det}\,\Sigma)^{n/2}} \exp\left\{-\mathrm{Tr}(\Sigma^{-1}W)\right\}$$

注意：Σ の制約について

- $K = (K_{ij}) = \Sigma^{-1}$, $\Sigma = (\Sigma_{ij})$ と $V = \{1,\,2,\,\ldots,\,p\}$ であった．

  $i \neq j$ に対し，

$$K_{ij} = 0 \Longleftrightarrow \Sigma_{ij} = \Sigma_{i,\,V\setminus\{i,\,j\}}\{\Sigma_{V\setminus\{i,\,j\},\,V\setminus\{i,\,j\}}\}^{-1}\Sigma_{V\setminus\{i,\,j\},\,j}$$

条件付き独立性(CI)

- f：確率変数が係わる確率密度関数の generic symbol とする.

- 確率ベクトル X, Y, Z に対して,

$$X \perp\!\!\!\perp Y \,|\, Z \iff f(x,\, y,\, z)f(z) = f(x,\, z)f(y,\, z)$$

CI の性質

確率ベクトル X, Y, Z, W に対して,

(C1)	$X \perp\!\!\!\perp Y \,\vert\, Z$	then	$Y \perp\!\!\!\perp X \,\vert\, Z$
(C2)	$X \perp\!\!\!\perp Y \,\vert\, Z$ and $U = g(Y)$	then	$X \perp\!\!\!\perp U \,\vert\, Z$
(C3)	$X \perp\!\!\!\perp Y \,\vert\, Z$ and $U = g(Y)$	then	$X \perp\!\!\!\perp Y \vert\, (Z,\, U)$
(C4)	$X \perp\!\!\!\perp Y \,\vert\, Z$ and $X \perp\!\!\!\perp W \,\vert\, (Y,\, Z)$	then	$Y \perp\!\!\!\perp (Y,\, W) \,\vert\, Z$

Covariance selection models（1）

- $X = (X_1,\, X_2,\, \ldots,\, X_p)' \sim N_p(0,\, 2\Sigma)$ と $K = \Sigma^{-1}\,(\Sigma \in \mathrm{Sym}_p^+)$ であった．
  $i, \neq j\,(i,\, j \in V = \{1,\, 2,\, \ldots,\, p\})$ に対して，

$$
\begin{aligned}
K_{ij} = 0 \quad &\Longleftrightarrow \quad X_i \perp\!\!\!\perp X_j \,|\, X_{V\setminus\{i,j\}} \\
&\Longleftrightarrow \quad \Sigma_{i,\,V\setminus\{i,j\}}\{\Sigma_{V\setminus\{i,j\},\,V\setminus\{i,j\}}\}^{-1}\Sigma_{V\setminus\{i,j\},\,j}
\end{aligned}
$$

- グラフに対応させる：$G = (V,\, E)$ で
  - V を頂点集合
  - 辺集合 $E \subset V \times V \setminus \{\{i,\, i\};\, i \in V\}$ を

$$K_{ij} \neq 0 \Longleftrightarrow \{i,\, j\} \in E$$

Covariance selection models（2）

- $X \sim N_p(0,\, 2\Sigma)$ と $\Sigma = K^{-1}$ で

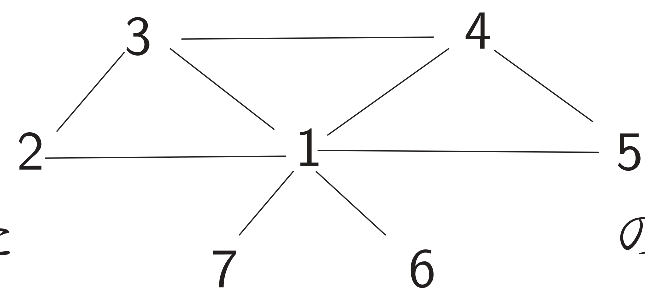
$$K = \begin{pmatrix} K_{11} & K_{12} & 0 & 0 \\ K_{21} & K_{22} & K_{23} & 0 \\ 0 & K_{32} & K_{33} & K_{34} \\ 0 & 0 & K_{43} & K_{44} \end{pmatrix} \qquad 1 — 2 — 3 — 4$$

  $G = (V,\, E)$ において，$V = \{1,\, 2,\, 3,\, 4\}$ と $E = \{\{1,\, 2\},\, \{2,\, 3\},\, \{3,\, 4\}\}$ となる．

- このように，K によって与えられるグラフを $G(K) = (V,\, E(K))$ と書くことにする．

グラフィカルモデルの用語について（１）

- $G = (V, E)$ はループと多重辺を持たない有限の頂点のグラフ
- 連結グラフ：どの2頂点も道（辺による経路）でむすばれているもの
- $A \subset E$ とし，$E_A = E \cap (A \times A)$ を頂点集合と辺集合とするグラフ $G_A = (A, E_A)$ を A により誘導された G の部分グラフという.



【例】$A = \{1, 2, 4\}$ により誘導された の部分グラフは 2—1—4 となる.

グラフィカルモデルの用語について（２）

- グラフ $G = (V, E)$ において，異なる頂点 $v_0, v_1, \ldots, v_k,$ に対して，$\{v_{i-1}, v_i\}, \{v_k, v_0\} \in E\,(i = 1, 2 \ldots, k)$ かつ $\{v_0, v_k\}$ を除いて添え字が連続する頂点間以外は辺をもたないとき，列 $v_0, v_1, \ldots, v_k, v_0$ を長さ $(k+1)$ の cycle という.

- グラフ $G = (V, E)$ が長さ 4 以上の cycle を誘導部分グラフとしてもたないとき，G は C_4−free(chordal グラフ，decomposable グラフ, traiangulated グラフ) という.

- グラフ $G = (V, E)$ は誘導部分グラフ ○—○—○—○ を持たないとき，G は A_4−free（等質グラフ）という.

グラフィカルモデルの用語について（３）

- グラフ $G = (V, E)$ において，$C \subset V$ のすべての頂点間に辺が存在するとき，C は完全（complete）であるという.
- $C \subset V$ は完全で，C を含む任意の V の部分集合が完全ではないとき，C を maximal clique という．以後，maximal cliques の個数を k と書く.
- $\{C_1, C_2, \dots, C_k\}$ をグラフ $G = (V, E)$ の maximal cliques の列とし，$j = 2, \dots, k$ に対し，

$$H_j = \cup_{i=1}^{j} C_i, \qquad R_j = C_j \setminus H_{j-1}, \qquad S_j = H_{j-1} \cap C_j$$

とおく．このとき，$\{C_1, C_2, \dots, C_k\}$ が perfect 列とは,

  - すべての $i > 1$ に対し，ある $j < i$ が存在し，$S_i \subset C_j$
  - $S_i \, (i = 2, \dots, k)$ は完全.

- S_i を minimal seperator という.

グラフィカルモデルの用語について（４）

Lauritzen (1996)

$X \sim N_p(0,\, 2K^{-1})$ と連結なグラフ $G(K) = (V,\, E(K))$ のとき，つぎは同値：

- ⋆ $G(K)$ は C_4−free
- ⋆ maximal cliques のある列は perfect.
- ⋆ 完全な maximal clique 列 $\{C_1,\, C_2,\, \dots,\, C_k\}$ と $j = 2,\, \dots,\, k$ に対し，
  $H_j = \cup_{i=1}^{j} C_i, \qquad R_j = C_j \setminus H_{j-1}, \qquad S_j = H_{j-1} \cap C_j$ としたとき，

$$X_{C_j \setminus H_{j-1}} \perp\!\!\!\perp X_{H_{j-1} \setminus C_j} \,|\, X_{S_j}$$

Letac and Massam (AOS, 2007)

つぎは同値：

- ⋆ 連結な $G(K)$ は C_4−free かつ A_4−free
- ⋆ 任意の maximal cliques の列は perfect.

UDG モデルに関わる開凸錐（１）

- $V=\{1,\,2,\,\ldots,\,p\}$，$N_{|V|}(0,\,2\,\Sigma)$，$K=\Sigma^{-1}$ と $G(K)=(V,\,E(K))$ で

$$i\neq j \text{ かつ } \{i,\,j\}\notin E(K)\Leftrightarrow K_{ij}=0,\qquad\qquad (i,\,j\in V).$$

- 以下では，$G(K)$ は連結で C_4−free を仮定する．
- $\mathscr{C}=\{C_1,\,C_2,\,\ldots,\,C_k\}$：maximal cliques の完全列とする．
- $j=2,\,3,\,\ldots,\,k$ に対し，

$$H_j=\cup_{i=1}^{j}C_i,\ R_j=C_j\setminus H_{j-1},\ S_j=C_j\cap H_{j-1}$$

としたとき，すべての j に対し，ある $i<j$ が存在し，$S_j\subset C_i$ となる．

- $\mathscr{S}$ を minimal seperators $S_2,\,S_3,\,\ldots,\,S_k$ の集合とする．
- $S\in\mathscr{S}$ に対し，$\nu(S)=\#\{S_i=S\,(i=2,\,3,\,\ldots,\,k)\}$ とする．

UDG モデルに関わる開凸錐（２）

- 対称行列空間の部分集合

  ⋆ $Z_G := \{x = (x_{ij}) \in \mathrm{Sym}_{|V|};\, i \neq j$ かつ $\{i, j\} \notin E(K) \Rightarrow x_{ij} = 0\} \subset \mathrm{Sym}_{|V|}$.

  ⋆ I_G を $\{i, j\} \in E(K)$ および $i = j$ に対し，$x_{ij} = x_{ij}$ のみが定義されている $G-$imcomplete symmetric matrices の集合とする.

- 開凸錐

  ⋆ $P_G := Z_G \cap \mathrm{Sym}^+_{|V|} \subset \mathrm{Sym}^+_{|V|}$.

  ⋆ $Q_G := \{x \in I_G;\, x_{C,C} > 0$ for $\forall C \in \mathscr{C}\} \subset I_G$.

  ⋆ $R_G := \tau(\mathrm{Sym}^+_{|V|}) \subset I_G$.
  ただし，“$>$” は正定値の意味，τ は $\mathrm{Sym}_{|V|}$ から I_G への自然な射影.

- 定義から $R_G \subset Q_G$ である．逆は？

UDG モデルに関わる開凸錐（３）

Gröner *et al.* (1984)

グラフ $G(K)$ は連結で C_4−free (chordal) のとき，$\forall x \in Q_G$ に対し，一意的に $\tilde{x} \in \mathrm{Sym}^+_{|V|}$ ($G-$ p.d. matrix completion) が存在し，

(i) $\{i, j\} \in E(K)$ と $i = j$ に対し，$x_{ij} = \tilde{x}_{ij}$　　(ii) $\tilde{x}^{-1} \in P_G$

- すなわち，$G(K)$ が連結で C_4−free (chordal) のとき，$R_G = Q_G$.

Laruritzen (1996)

グラフ $G(K)$ は連結で C_4−free (chordal) のとき，$\forall x \in Q_G$ に対し，

$$\tilde{x} = \left\{ \sum_{C \in \mathscr{C}} [(x_C)^{-1}]^0 - \sum_{S \in \mathscr{S}} \nu(S)[(x_S)^{-1}]^0 \right\}^{-1}$$

UDG モデルに関わる開凸錐（４）

- $(x, y) \in I_G \times Z_G$ に対し，$\langle x, y \rangle = \displaystyle\sum_{\{i,j\} \in E(K),\, i=j} x_{ij} y_{ij}$
- P_G の双対開凸錐：$P_G^* := \{x \in I_G; \langle x, y \rangle > 0 \text{ for } \forall y \in \bar{P}_G \setminus \{0\}\}$
- Q_G の双対開凸錐：$Q_G^* := \{y \in Z_G; \langle x, y \rangle > 0 \text{ for } \forall x \in \bar{Q}_G \setminus \{0\}\}$

ただし， $\bar{P}_G$ と $\bar{Q}_G$ は P_G と Q_G の閉包

Letac and Massam (2007)

$\star$ グラフ $G(K)$ は連結で C_4–free (chordal) のとき，P_G と Q_G は互いに双対．すなわち，$P_G^* = Q_G = R_G \subset I_G$ かつ $Q_G^* = P_G \subset Z_G$.
$\star$ C_4–free かつ A_4-free のとき，P_G は等質（自己同型群が推移的に P_G に作用）

UDG モデルの共分散行列の最尤推定量と H-Wishart 分布（１）

- $G(K)$ は C_4-free とし，$|V| \geq 3$ の整数とする．$X_1, X_2, \ldots, X_n$ は独立同一に $|V|$ 変量正規分布 $N_{|V|}(0, 2\,\Sigma)$ $(\Sigma^{-1} \in P_G)$ の Σ の最尤推定量は

$$\widehat{\Sigma}_{mle} = \tilde{\hat{\sigma}}_{mle}, \qquad \hat{\sigma}_{mle} = \operatorname*{argmax}_{x \in Q_G = R_G} \frac{1}{(\mathrm{Det}\,\tilde{x})^{n/2}} \exp\left\{-\langle \tau(W), \{\tilde{x}\}^{-1} \rangle\right\}$$

ただし，$n \geq \max_{C \in \mathscr{C}} |C|$ であり，$\tau : \mathrm{Sym}_{|V|} \to I_Z$ は射影で，$\tilde{x} \in \mathrm{Sym}^+_{|V|}$ $(\tilde{x}^{-1} \in P_G)$ は $x \in Q_G$ の $G-$ p.d. matrix completion.

- $\tilde{\hat{\sigma}}_{mle} = \left\{ \sum_{C \in \mathscr{C}} [(x_C)^{-1}]^0 - \sum_{S \in \mathscr{S}} \nu(S) [(x_S)^{-1}]^0 \right\}^{-1}$ は $\tau(W)$ のみ依存.

- $\widehat{\Sigma}_{mle}$ の分布とは，$\hat{\sigma} \in Q_G$ の分布(Q_G 上の分布)！

UDG モデルの共分散行列の最尤推定量と H-Wishart 分布（２）

- CI の性質と maximal cliques 列が完全（入れ子構造）であることから

Dawid and Lauritzen (AOS, 1993)

$\Sigma^{-1} \in P_G$ のとき，$\hat{\sigma}_{mle} \in Q_G$ の分布は

$$\frac{\Pi_{C \in \mathscr{C}} w_{|C|}(n,\, 2\Sigma_{C,\,C})}{\Pi_{S \in \mathscr{S}} \{w_{|S|}(n,\, 2\Sigma_{S,\,S})\}^{\nu(S)}}$$

$$\propto \frac{\Pi_{C \in \mathscr{C}} \dfrac{(\mathrm{Det}\, x_{C,\,C})^{(n-|C|-1)/2}}{(\mathrm{Det}\, \Sigma_{C,\,C})^{n/2}} \exp\{-\mathrm{Tr}\,(x_{C,\,C}(\Sigma_{C,\,C})^{-1})\}}{\Pi_{S \in \mathscr{S}} \left(\dfrac{(\mathrm{Det}\, x_{S,\,S})^{(n-|S|-1)/2}}{(\mathrm{Det}\, \Sigma_{S,\,S})^{n/2}} \exp\{-\mathrm{Tr}\,(x_{S,\,S}(\Sigma_{S,\,S})^{-1})\} \right)^{\nu(S)}} \mathbb{1}_{Q_G}(x)\, dx$$

UDG モデルの共分散行列の最尤推定量と H-Wishart 分布（３）

Roverato(Biometrika, 2000) and Letac and Massam (2007)

$\Sigma^{-1} \in P_G$ のとき，$\hat{\sigma}_{mle} \in Q_G$ の分布は

$$\frac{1}{\Pi_{i=1}^{k} \pi^{(|C_i|-|S_i|)|S_i|/2} \Gamma_{|C_i|-|S_i|}\left(\frac{n-|S_i|}{2}\right)} \times \frac{(\mathrm{Det}\,\tilde{x})^{n/2} \exp\{-\langle x,\, \Sigma^{-1}\rangle\}}{(\mathrm{Det}\,\Sigma)^{n/2}} \mu_G(dx)$$

ただし，$|S_1| = 0$ で

$$\mu_G(dx) = \mathbb{1}_{Q_G}(x)\, dx \Bigg/ \frac{\Pi_{i=1}^{k} (\mathrm{Det}\, x_{C,\,C})^{(|C_i|+1)/2}}{\Pi_{i=2}^{k} (\mathrm{Det}\, x_{S,\,S})^{(|S_i|+1)/2}}$$

参考文献

Andersson, S.A., Madigan, D., Perlman, M.D., and Triggs, C.M., A graphical characterization of lattice conditional independece models, *Ann. Math. Artifical Intelligence* **21** (1997) 27–50.

Andersson, S.A. and Perlman, M.D., Lattice models for conditional independence in a multivariate normal models, *Ann. Statist.* **21** (1993), 525–572.

Andersson, S.A. and Wojnar, G.G., Wishart distributions on homogeneous cones, *J. Theoret. Probab.* **17** (2004) 781–818.

Boutouria, I., Characterization of the Wishart distributions on homogeneous cones, *C. R. Acad. Sci. Paris, Ser. 1* **341** (2005) 43–48.

Dawid, A. and Lauritzen, S.L., Hyper-Markov laws in the statistical analysis of decomposable graphical models, *Ann. Statist.* **21** (1993) 1272–1317.

Gröne, R., Johnson, C.R., De Sá, E. and Wolfowicz, H., Positive definite completions of partial Hermitian matrices, *Linear Algebra Appl.* **58** 109–124.

Ishi, H., Positive Rietz distributions on homogeneous cones, *J. Math. Soc. Japan* **52** (2000) 161–181.

Konno, Y., Inadmissibility of the maximum likelihood estimator of normal covariance matrices with the lattice conditional independence, *J. Multivariate Anal.* **79** (2001) 33–51.

Lauritzen, S.L., Graphical Models, Oxford Univ. Press (1996).

Letac, G.G. and Massam, H., Extremal rays and duals fro cones fo positive definite matrices with prescribed zeros, *Linear Algebra Appl.* **418** (2006) 737–750.

Letac, G.G. and Massam, H., Wishart distributions for decomposable graphs, *Ann. Statist.* **35** (2007) 1278–1323.

Maassam, H. and Neher, E., Estimation and testing for lattice conditional independence models on Euclidean Jordan algebras, *Ann. Statist.* **26** 1051–1082.

Neher, E., Transformations groups of the Andersson-Perlman cone, *J. Lie Theory* **9** (1999) 203–313.

Roverat, A., Cholesky decomposition of hyper inverse Wishart matrix, *Biometrika* **87** (2000) 99–112.