

# 多変量線形回帰モデルにおける一致性を持つ $C_p$ 型 規準が真の変数を選択する確率の収束オーダー\*

広島大学・理学研究科数学専攻<sup>†</sup> 柳原宏和

Hirokazu Yanagihara

Department of Mathematics, Graduate School of Science

Hiroshima University

2016 年 5 月 16 日

## §1. 序

本論文では、正規性を仮定した多変量線形回帰モデル (ここでは正規多変量回帰モデルと呼ぶ) において、変数選択規準を最小にする変数の組を最適なものとする、変数選択規準最小化に基づく変数選択問題を取り扱う。この選択法の場合、どの変数選択規準を用いるかが重要な問題となり、それを決める 1 つの重要な特性として、変数選択規準が一致性を持つかどうかと言うことがある。本論文での一致性とは、厳密に言えば弱一致性のことであり、モデル選択規準によって真の変数が最適な変数として選択される確率が漸近的に 1 になる特性のことを言う。一致性は、多くの場合、標本数  $n$  のみを無限大とする漸近理論である、大標本漸近理論により評価されている。一方で、近年、ハードウェアの発展により、蓄積・解析できるデータの数が爆発的に増大し、目的変数の次元数  $p$  が大きいデータである、高次元データの解析の需要が高まっている。本論文で取り扱う高次元データとは、次元数  $p$  は大きいが標本数  $n$  よりも小さいとする適度な高次元 (moderately high-dimensional) データ (Yao *et al.* [5] 参照) である。このような高次元データでは、大標本漸近理論ではなく、次元数  $p$  も  $p/n$  が 1 未満の定数に収束するという条件の下で  $n$  と共に無限大とする、高次元大標本漸近理論により一致性を評価した方が妥当である。

---

本研究は科学研究費補助金、挑戦的萌芽研究 (課題番号:25540012) の助成を受けたものである。

〒739-8626 広島県東広島市鏡山 1-3-1

近年, Yanagihara [4] で, 以下のような漸近理論に基に, 高次元性を調整した一貫性を持つ一般化  $C_p$  (High-dimensionality-adjusted Consistent Generalized  $C_p$ :  $HCGC_p$ ) 規準が提案された.

$$n \rightarrow \infty, \quad p/n \rightarrow c_0 \in [0, 1). \quad (1)$$

上記の漸近理論は, 次元数  $p$  を無限大にしなくてもよいため, 大標本漸近理論と高次元大標本漸近理論の両方を特別な形として含むものになっている.  $HCGC_p$  規準は, (1) 式に基づく漸近理論で一貫性を保証しているが, 残念ながら, どのようなオーダーで真の変数を選ぶ確率が 1 に近づくかは議論されていない. そこで, 本論文では,  $HCGC_p$  規準が真の変数を最適な変数として選択する確率の 1 への収束オーダーを求めることを目的とする.

以下本論文では, 第 2 章で, 正規多変量線形モデルと  $HCGC_p$  を紹介する. 第 3 章で, 主定理である, 真の変数を選ぶ確率の収束オーダーを求める. 数学的な証明は付録に記載した.

## §2. 高次元性の下でも一貫性を持つ $C_p$ 型規準量

$n$  個の個体に対して,  $p$  個の目的変数を並べた目的変数ベクトル,  $y_1, \dots, y_n$  と  $k$  個の説明変数を並べた説明変数ベクトル  $x_1, \dots, x_n$  が観測されたとし, それらを並べた  $n \times p$ ,  $n \times k$  行列をそれぞれ  $Y = (y_1, \dots, y_n)'$ ,  $X = (x_1, \dots, x_n)'$  とする. ここでは, 推定量の存在を保証するために,  $n > p + k + 1$  とする. 以下,  $k$  個の説明変数のから一部を用いた正規多変量線形回帰モデルを考えるが, どの変数を用いたかを以下の集合の元によって表現する.

$$j \subseteq \omega = \{1, \dots, k\}, \quad k_j = \#(j).$$

このとき,  $X$  から  $j$  の元に対応した列を抜き出し並べた  $n \times k_j$  行列を  $X_j$  とする. 例えば,  $j = \{1, 2, 4\}$  であれば,  $X_j$  は  $X$  の第 1, 2, 4 列を抜き出して並べた行列である. この  $X_j$  を使って, 考える候補のモデルを以下のように記述する.

$$Y \sim N_{n \times p}(X_j \Theta_j, \Sigma_j \otimes I_n). \quad (2)$$

ただし,  $\Theta_j$  は未知回帰係数を並べた  $k_j \times p$  回帰係数行列,  $\Sigma_j$  は  $p \times p$  分散共分散行列で正定値性を仮定する. (2) 式は,  $j$  により定式化されるので, (2) 式を候補のモデル  $j$  と呼ぶことにする. 特に,  $j = \omega$  のときはフルモデルと呼ぶ. このとき,  $X_\omega = X$  である.

真の説明変数は以下の集合の元により表現する.

$$j_* \subseteq \omega, \quad k_* = \#(j_*).$$

ここで,  $X_{j_*} = X_*$  とし, 真のモデル  $j_*$  を以下のように記述する.

$$Y \sim N_{n \times p}(X_* \Theta_*, \Sigma_* \otimes I_n). \quad (3)$$

ただし,  $\Theta_*$  は真の回帰係数を並べた真の  $k_* \times p$  回帰係数行列,  $\Sigma_*$  は真の  $p \times p$  分散共分散行列である. ここで,  $\Sigma_*$  は正定値行列とし,  $\Theta_*$  の各行ベクトルには少なくとも1つ以上零でない値があることを仮定する.

以下に, (2) 式のモデル  $j$  における  $HCGC_p$  は以下ようになる (Yanagihara [4] 参照). ただし, 定義 1 内の  $\lim_{n \rightarrow \infty, p/n \rightarrow c_0}$  は, (1) 式の漸近理論での極限を表す.

**定義 1**  $S_j$  をモデル  $j$  における分散共分散行列の不偏推定量,

$$S_j = \frac{1}{n - k_j} Y'(I_n - P_\omega)Y,$$

とする. ただし,  $P_j$  は  $X_j$  の列ベクトルで張る空間への射影行列で,

$$P_j = X_j(X_j'X_j)^{-1}X_j',$$

である. このとき, モデル  $j$  における  $HCGC_p$  は,

$$HCGC_p(j|\alpha) = (n - k_j)\text{tr}(S_j S_\omega^{-1}) + pk_j \alpha, \quad (4)$$

である. ただし,  $\alpha$  は以下を満たす正の数である.

$$\alpha = \frac{n}{n - p} + \beta, \quad \beta > 0 \text{ s.t. } \lim_{n \rightarrow \infty, p/n \rightarrow c_0} \sqrt{p}\beta = \infty, \quad \lim_{n \rightarrow \infty, p/n \rightarrow c_0} \frac{p}{n}\beta = \infty. \quad (5)$$

$\mathcal{J}$  を候補のモデルを集めた集合とし,  $\mathcal{J}$  を真の説明変数を含んでいる過多に記述された (overspecified) モデルを集めた集合  $\mathcal{J}_+$  と真の説明変数を含んでいない過少に記述された (underspecified) モデルを集めた集合  $\mathcal{J}_-$  に分ける. 過少に記述されたモデルは過多に記述されたモデルでないモデルとも言えるので,  $\mathcal{J}_+$  と  $\mathcal{J}_-$  は以下のように定義できる.

$$\mathcal{J}_+ = \{j \in \mathcal{J} | j_* \subseteq j\}, \quad \mathcal{J}_- = \mathcal{J}_+^c \cap \mathcal{J}. \quad (6)$$

ただし,  $A^c$  は集合  $A$  の補集合を示す. さらに, 以下のような非心パラメータ行列を定義する.

$$\Delta_j = \Sigma_*^{-1/2} \Theta_*' X_*'(I_n - P_j) X_* \Theta_* \Sigma_*^{-1/2}. \quad (7)$$

ここで,  $j \in \mathcal{J}_+$  のときは,  $\Delta_j = O_{n,n}$  となり,  $j \in \mathcal{J}_-$  ならば,  $\Delta_j$  は,  $d_j = \text{rank}(\Delta_j)$  とすると,  $d_j \neq 0$  で  $d_j \leq \min\{p, k_{j^c \cap j_*}\}$  となる半正定値行列である. ただし,  $O_{n,n}$  はすべ

ての成分が0である  $n \times n$  行列である。ここで、 $HCGC_p(j|\alpha)$  によって選ばれた変数を  $\hat{j}_\alpha$  とする、つまり、

$$\hat{j}_\alpha = \arg \min_{j \in \mathcal{J}} HCGC_p(j|\alpha), \quad (8)$$

とする。このとき、以下の仮定、

仮定 A1.  $j_* \in \mathcal{J}$ ,

仮定 A2.  $\forall j \in \mathcal{J}_-$  に対して、 $\sup_{(a,b) \in \mathbb{N}^2} \inf_{n \geq a, p \geq b} \frac{1}{n} \text{tr}(\Delta_j) > 0$ ,

が成り立つとき、

$$\lim_{n \rightarrow \infty, p/n \rightarrow c_0} P(\hat{j}_\alpha = j_*) = 1,$$

となる (証明は、Yanagihara [4] 参照)。なお、 $j \in \mathcal{J}_-$  のとき、 $\text{tr}(\Delta_j)$  は  $n > k$  ならば常に正であるので、その結果に仮定 A2 を加えると、

$$\inf_{n > k, p \geq 1} \frac{1}{n} \text{tr}(\Delta_j) > 0, \quad (9)$$

が言える。

### §3. 選択確率の収束オーダー

本章では、(4) 式で定義した  $HCGC_p$  が  $j_*$  を最適な組み合わせとして選択する確率が1に収束するときの収束オーダーを調べる。ここで、候補のモデルとして、すべての組み合わせを考えることにする。つまり、 $\mathcal{J} = \rho(\omega)$  である。さらに、

$$\frac{n-k}{n-k-p-1} = \frac{n}{n-p} + \frac{n+kp}{(n-p)(n-k-p-1)},$$

であることから、(4) 式の  $HCGC_p$  の罰則項  $\alpha$  を、

$$\alpha = \frac{n-k}{n-k-p-1} + \beta, \quad (10)$$

と書き換える。ただし、 $\beta$  は (5) 式と同じ正の数である。

今、

$$P(\hat{j}_\alpha = j_*) = 1 - P(\hat{j}_\alpha \neq j_*) = 1 - P\left(\cup_{j \in \mathcal{J} \setminus \{j_*\}} \{\hat{j}_\alpha = j\}\right),$$

であるので、目的を達成するためには、 $P(\cup_{j \in \mathcal{J} \setminus \{j_*\}} \{\hat{j}_\alpha = j\})$  のオーダーを求めれば良い

ことがわかる. ここで,  $\{\hat{j}_\alpha = j\} (\forall j \in \mathcal{J})$  は互いに素であることに注意すれば,

$$\begin{aligned} P\left(\bigcup_{j \in \mathcal{J} \setminus \{j_*\}} \{\hat{j}_\alpha = j\}\right) &= \sum_{j \in \mathcal{J} \setminus \{j_*\}} P(\hat{j}_\alpha = j) \\ &= \sum_{j \in \mathcal{J}_+ \setminus \{j_*\}} P(\hat{j}_\alpha = j) + \sum_{j \in \mathcal{J}_-} P(\hat{j}_\alpha = j), \end{aligned} \quad (11)$$

となる. ここで, 以下にオーダー評価に必要な補助定理を与える (証明は付録 A.1 参照).

**補助定理 1**  $m = n - k - p + 1$  とし,  $r$  は  $r > m/4$  を満たす正の整数とする.  $U$  と  $V$  を互いに独立な確率変数で,  $U \sim \chi_p^2$ ,  $V \sim \chi_m^2$  とする. このとき,  $j \in \mathcal{J}_+$  であれば,

$$P(\hat{j}_\alpha = j) \leq \left(\frac{n-k}{\sqrt{p}\gamma}\right)^{2r} E\left[\left(\frac{U}{V} - \frac{p}{m-2}\right)^{2r}\right], \quad (12)$$

が成り立つ. ただし  $\gamma$  は,

$$\gamma = \sqrt{p}\beta, \quad (13)$$

である. また,  $j \in \mathcal{J}_-$  のとき,  $q_j = k_{j_* \cap j}$  とし,  $U_i$  ( $i = 1, \dots, q_j$ ) を  $V$  と独立な,  $U_i \sim \chi_p^2(\delta_{j,i})$  である確率変数とする. ここで  $\delta_{j,i}$  は,  $\delta_{j,1} + \dots + \delta_{j,q_j} = \text{tr}(\Delta_j)$  であり,  $\delta_{j,i} \geq \lambda_{j,1}/q_j$  を満たし,  $\lambda_{j,1}$  は  $\Delta_j$  の最大固有値とする. このとき,

$$P(\hat{j}_\alpha = j) \leq \sum_{i=1}^{q_j} \left(\frac{n-k}{\delta_{j,i}}\right)^{2r} E\left[\left(\frac{U_i}{V} - \frac{p + \delta_{j,i}}{m-2} - \rho\right)^{2r}\right], \quad (14)$$

が成り立つ. ただし,

$$\rho = \frac{p}{n-k}\beta, \quad (15)$$

である.

さらに, (12), (14) 式内の期待値のオーダーを評価するため, 以下の補助定理を与える (証明は付録 A.2 参照).

**補助定理 2**  $m = n - k - p + 1$  とし,  $\delta$  を

$$\lim_{n \rightarrow \infty, p/n \rightarrow c_0} \delta = \infty, \quad \inf_{n > k, p \geq 1} \frac{\delta}{n} > 0, \quad \frac{p}{\delta} = O(1),$$

を満たす正の数とする. さらに,  $U_1, U_2$  と  $V$  は互いに独立な確率変数で,  $U_1 \sim \chi_p^2$ ,

$U_2 \sim \chi_p^2(\delta)$ ,  $V \sim \chi_m^2$  とする. このとき, 任意の自然数  $r$  に関して, 十分  $n$  が大きければ,

$$E \left[ \left( \frac{U_1}{V} - \frac{p}{m-2} \right)^{2r} \right] = O(p^r n^{-2r}), \quad (16)$$

$$E \left[ \left( \frac{U_2}{V} - \frac{p+\delta}{m-2} - \rho \right)^{2r} \right] = O(\delta^r n^{-2r}), \quad (17)$$

が成り立つ. ただし,  $\rho$  は (15) 式で定義された正の数である.

(5) 式の  $\beta$  の条件から,

$$\lim_{n \rightarrow \infty, p/n \rightarrow c_0} \gamma = \infty, \quad \lim_{n \rightarrow \infty, p/n \rightarrow c_0} \rho = 0.$$

また,

$$\delta = \min_{j \in \mathcal{J}_-} \frac{\lambda_{j,1}}{q_j},$$

と定義すると,  $\lambda_{j,1} \geq \text{tr}(\Delta_j)/q_j$  より,

$$\delta \geq \min_{j \in \mathcal{J}_-} \frac{\text{tr}(\Delta_j)}{q_j^2}. \quad (18)$$

また, (9) 式から,

$$\inf_{n > k, p \geq 1} \min_{j \in \mathcal{J}_-} \frac{\text{tr}(\Delta_j)}{n} = \phi > 0, \quad (19)$$

となる下限が存在し,  $\max_{j \in \mathcal{J}_-} q_j \leq k$  であるので, これら結果と (18), (19) 式から,

$$\frac{p}{\delta} \leq \frac{p}{\min_{j \in \mathcal{J}_-} \text{tr}(\Delta_j)/q_j^2} = \frac{p/n}{\min_{j \in \mathcal{J}_-} \text{tr}(\Delta_j)/(nq_j^2)} \leq \frac{pk^2}{n\phi} \rightarrow \frac{c_0 k^2}{\phi} < \infty,$$

となるので,  $p/\delta = O(1)$  であることがわかる. さらに, 仮定 A2 より,  $\lim_{n \rightarrow \infty, p/n \rightarrow c_0} \text{tr}(\Delta) = \infty$  であるので,  $\lim_{n \rightarrow \infty, p/n \rightarrow c_0} \delta = \infty$  が言える. よって補題 A1・A2 より, 任意の自然数  $r$  に対して,

$$P(\hat{j}_\alpha = j) = \begin{cases} O(\gamma^{-r}) & (j \in \mathcal{J}_+ \setminus \{j_*\}) \\ O(\delta^{-r}) & (j \in \mathcal{J}_-) \end{cases}. \quad (20)$$

最後に,  $O(\gamma^{-r})$  と  $O(\delta^{-r})$  のどちらの収束速度が速いかを判定する.  $\beta$  の定義式 (5) から  $p\beta/n \rightarrow 0$  であるので,  $\sqrt{p}\beta/n = (p\beta/n)/\sqrt{p} \rightarrow 0$ . この結果と (19) 式より,

$$\frac{\gamma}{\delta} \leq \frac{\sqrt{p}\beta/n}{\min_{j \in \mathcal{J}_-} \text{tr}(\Delta_j)/(nq_j^2)} \leq \frac{\sqrt{p}\beta}{n} \cdot \frac{k^2}{\phi} \rightarrow 0.$$

よって,  $O(\delta^{-r})$  の方が  $O(\gamma^{-r})$  よりも 0 に収束する速度が速いことがわかる. 以上の結果と, (11), (20) 式により, 以下の定理を得ることができる.

定理 1 仮定 A1・A2 の下で,  $HCGC_p$  が真の変数を選択する確率は以下ようになる.

$$P(\hat{j}_\alpha = j_*) = 1 + O(p^{-r/2}\beta^{-r}).$$

ただし,  $r$  は任意の自然数である.

## A. 付録

### A.1. 補助定理 1 の証明

まず,  $j \in \mathcal{J}_+ \setminus \{j_*\}$  のときを考える. このとき,  $j_- = j \setminus \{a\}$ ,  $a \in j \cap j_*$  とおくと,  $\mathcal{J} = \rho(\omega)$  であるので  $j_- \in \mathcal{J}$  である. よって,

$$\begin{aligned} P(\hat{j}_\alpha = j) &= P(\cap_{\ell \in \mathcal{J} \setminus \{j\}} \{HCGC_p(\ell|\alpha) - HCGC_p(j|\alpha) > 0\}) \\ &\leq P(HCGC_p(j_-|\alpha) - HCGC_p(j|\alpha) > 0), \end{aligned} \quad (\text{A.1})$$

となる. ここで,  $k_{j_-} = k_j - 1$  であり,

$$(n - k_j) \text{tr}(\mathbf{S}_j \mathbf{S}_\omega^{-1}) = (n - k)p + (n - k) \text{tr} \{ \mathbf{Y}'(\mathbf{P}_\omega - \mathbf{P}_j) \mathbf{Y} \{ \mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_\omega) \mathbf{Y} \}^{-1} \}, \quad (\text{A.2})$$

であることに注意すれば,

$$\begin{aligned} D_1(j) &= HCGC_p(j_-|\alpha) - HCGC_p(j|\alpha) \\ &= (n - k) \text{tr} \{ \mathbf{Y}'(\mathbf{P}_j - \mathbf{P}_{j_-}) \mathbf{Y} \{ \mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_\omega) \mathbf{Y} \}^{-1} \} - p\alpha, \end{aligned}$$

となる. ここで,

$$\mathbf{W}_1 = \Sigma_*^{-1/2} \mathbf{Y}'(\mathbf{P}_j - \mathbf{P}_{j_-}) \mathbf{Y} \Sigma_*^{-1/2}, \quad \mathbf{W}_2 = \Sigma_*^{-1/2} \mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_\omega) \mathbf{Y} \Sigma_*^{-1/2},$$

とおくと,

$$D_1(j) = (n - k) \text{tr}(\mathbf{W}_1 \mathbf{W}_2^{-1}) - p\alpha, \quad (\text{A.3})$$

である. 今,  $(\mathbf{I}_n - \mathbf{P}_\omega)(\mathbf{P}_j - \mathbf{P}_{j_-}) = \mathbf{O}_{n,n}$  なので, ウィッシュヤート分布の性質より,

$$\mathbf{W}_1 \perp \mathbf{W}_2, \quad \mathbf{W}_1 \sim W_p(1, \mathbf{I}_p), \quad \mathbf{W}_2 \sim W_p(n - k, \mathbf{I}_p),$$

が成り立つ. ここでさらに, ウィッシュヤート分布の性質より,  $\mathbf{0}_p$  をすべての成分が 0 である  $p$  次元ベクトルとし,  $\mathbf{z} \sim N_p(\mathbf{0}_p, \mathbf{I}_p)$  とすると,  $\mathbf{W}_1 = \mathbf{z} \mathbf{z}'$  と書き換えることができ, それらを用いると,

$$\mathbf{z}' \mathbf{W}_2^{-1} \mathbf{z} = \frac{\mathbf{z}' \mathbf{z}}{\{(\mathbf{z}' \mathbf{z})^{-1/2} \mathbf{z}' \mathbf{W}_2^{-1} \mathbf{z} (\mathbf{z}' \mathbf{z})^{-1/2}\}^{-1}},$$

と書くことができる. 上式の右辺の分母分子は独立でそれぞれ,  $\chi_p^2, \chi_m^2$  に従うことから,

$$\text{tr}(\mathbf{W}_1 \mathbf{W}_2^{-1}) = \frac{U}{V},$$

と書き換えることができる. ただし,  $U$  と  $V$  は補助定理 1 で定義した確率変数である. この式と (A.3) 式を用いれば,

$$P(D_1(j) > 0) = P((n-k)U/V - p\alpha > 0) = P(U/V > p\alpha/(n-k)).$$

ここで, (13) 式で与えられた  $\gamma$  を用いると,

$$\frac{p}{n-k}\alpha - \frac{p}{m-2} = \frac{p\beta}{n-k} = \frac{\sqrt{p}\gamma}{n-k},$$

となるので,  $U/V - p/(m-2) \leq |U/V - p/(m-2)|$  に注意すれば,

$$\frac{U}{V} - \frac{p}{m-2} > \frac{\sqrt{p}\gamma}{n-k} \Rightarrow \left| \frac{U}{V} - \frac{p}{m-2} \right| > \frac{\sqrt{p}\gamma}{n-k}.$$

よって, 上記の関係式とマルコフの不等式を用いれば,

$$\begin{aligned} P(U/V > p\alpha/(n-k)) &= P(U/V - p/(m-2) > \sqrt{p}\gamma/(n-k)) \\ &\leq P(|U/V - p/(m-2)| > \sqrt{p}\gamma/(n-k)) \\ &= P(|U/V - p/(m-2)|^{2r} > \{\sqrt{p}\gamma/(n-k)\}^{2r}) \\ &\leq \left(\frac{n-k}{\sqrt{p}\gamma}\right)^{2r} E \left[ \left(\frac{U}{V} - \frac{p}{m-2}\right)^{2r} \right]. \end{aligned}$$

よって, (12) 式が示せた.

次に,  $j \in \mathcal{J}_-$  のときを考える. 今,  $j_+ = j \cup j_*$  とおくと,  $\mathcal{J} = \wp(\omega)$  から  $j_+ \in \mathcal{J}$  が言え, (A.1) 式と同様な手法により,

$$P(\hat{j}_\alpha = j) \leq P(HCGC_p(j_+|\alpha) - HCGC_p(j|\alpha) > 0),$$

が言える.  $k_{j_+} = k_j + k_{j \cap j_*}$  であることと (A.2) 式を用いれば,

$$\begin{aligned} D_2(j) &= HCGC_p(j_+|\alpha) - HCGC_p(j|\alpha) \\ &= -(n-k)\text{tr} \{ \mathbf{Y}'(\mathbf{P}_{j_+} - \mathbf{P}_j) \mathbf{Y} \{ \mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_\omega) \mathbf{Y} \}^{-1} \} + pq_j \alpha, \end{aligned}$$

となる. ただし,  $q_j = k_{j \cap j_*}$  である. ここで,

$$\mathbf{W}_1 = \Sigma_*^{-1/2} \mathbf{Y}'(\mathbf{P}_{j_+} - \mathbf{P}_j) \mathbf{Y} \Sigma_*^{-1/2}, \quad \mathbf{W}_2 = \Sigma_*^{-1/2} \mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_\omega) \mathbf{Y} \Sigma_*^{-1/2},$$

とおくと,

$$D_2(j) = -(n-k)\text{tr}(\mathbf{W}_1\mathbf{W}_2^{-1}) + pq_j\alpha, \quad (\text{A.4})$$

となる. 今,  $(\mathbf{I}_n - \mathbf{P}_\omega)(\mathbf{P}_{j_+} - \mathbf{P}_j) = \mathbf{O}_{n,n}$  なので, ウィッシュヤート分布の性質より,

$$\mathbf{W}_1 \perp \mathbf{W}_2, \quad \mathbf{W}_1 \sim W_p(q_j, \mathbf{I}_p; \Delta_j), \quad \mathbf{W}_2 \sim W_p(n-k, \mathbf{I}_p), \quad (\text{A.5})$$

が成り立つ. ただし,  $\Delta_j$  は (7) 式で与えられた非心パラメータ行列である. (A.5) 式に付録 A.3 の補助定理 A.1 を適用すれば, (A.4) 式より,

$$P(D_2(j) > 0) = P(\text{tr}(\mathbf{W}_1\mathbf{W}_2^{-1}) < q_j p \alpha / (n-k)) \leq \sum_{i=1}^{q_j} P(U_i/V < p\alpha / (n-k)),$$

となる. ただし,  $U_i$  は補助定理 1 で定義した確率変数である. ここで, (15) 式で与えられた  $\rho$  を用いると,

$$\frac{p}{n-k}\alpha - \frac{p + \delta_{j,i}}{m-2} = \frac{p\beta}{n-k} - \frac{\delta_{j,i}}{m-2} = \rho - \frac{\delta_{j,i}}{m-2},$$

であるので,

$$-\frac{U_i}{V} + \frac{p + \delta_{j,i}}{m-2} + \rho \leq \left| \frac{U_i}{V} - \frac{p + \delta_{j,i}}{m-2} - \rho \right|,$$

となることに注意すれば,

$$-\frac{U_i}{V} + \frac{p + \delta_{j,i}}{m-2} + \rho > \frac{\delta_{j,i}}{n-k} \Rightarrow \left| \frac{U_i}{V} - \frac{p + \delta_{j,i}}{m-2} - \rho \right| > \frac{\delta_{j,i}}{n-k}.$$

よって, 上記の関係式とマルコフの不等式を用いれば,

$$\begin{aligned} & P(U_i/V < p\alpha / (n-k)) \\ &= P(-U_i/V + (p + \delta_{j,i}) / (m-2) + \rho > \delta_{j,i} / (n-k)) \\ &\leq P(|U_i/V + (p - \delta_{j,i}) / (m-2) - \rho| > \delta_{j,i} / (n-k)) \\ &= P(|U_i/V + (p - \delta_{j,i}) / (m-2) - \rho|^{2r} > \{\delta_{j,i} / (n-k)\}^{2r}) \\ &\leq \left( \frac{n-k}{\delta_{j,i}} \right)^{2r} E \left[ \left( \frac{U_i}{V} - \frac{p + \delta_{j,i}}{m-2} - \rho \right)^{2r} \right]. \end{aligned}$$

よって, (14) 式が示せた.

## A.2. 補助定理 2 の証明

$B = 1/V$  とし,  $\xi_1 = E[U_1]$ ,  $\nu = E[B]$  とおく. このとき,  $p/(m-2) = \xi_1\nu$  であるので,

$$\frac{U_1}{V} - \frac{p}{m-2} = U_1B - \xi_1\nu = (U_1 - \xi_1)(B - \nu) + \xi_1(B - \nu) + \nu(U_1 - \xi_1), \quad (\text{A.6})$$

が成り立つ. よって,

$$\begin{aligned} & E \left[ \left( \frac{U_1}{V} - \frac{p}{m-2} \right)^{2r} \right] \\ &= \sum_{\substack{a+b+c=2r \\ 0 \leq a, b, c \leq 2r}} \frac{(2r)!}{a!b!c!} E \left[ \xi_1^b \nu^c (U_1 - \xi_1)^{a+c} (B - \nu)^{a+b} \right] \\ &= \sum_{\substack{a+b+c=2r \\ 0 \leq a, b, c \leq 2r}} \frac{(2r)!}{a!b!c!} \xi_1^b E \left[ (U_1 - \xi_1)^{a+c} \right] \nu^c E \left[ (B - \nu)^{a+b} \right]. \end{aligned} \quad (\text{A.7})$$

ここで, 付録 A.4 の補助定理 A.2 の (A.13) 式より,

$$E[(U_1 - \xi_1)^a] = \begin{cases} 1 & (a = 0) \\ 0 & (a = 1) \\ O(p^{[a/2]}) & (a \geq 2) \end{cases}, \quad (\text{A.8})$$

を得る. ただし,  $[ \ ]$  はガウス記号である. また,

$$\begin{aligned} E[(B - \nu)^a] &= E \left[ \left( \frac{1}{V} - \frac{1}{m-2} \right)^a \right] \\ &= \sum_{i=0}^a \frac{a!}{i!(a-i)!} \left( -\frac{1}{m-2} \right)^{a-i} E \left[ \left( \frac{1}{V} \right)^i \right] \\ &= \sum_{i=1}^a \frac{a!}{i!(a-i)!} \left( -\frac{1}{m-2} \right)^{a-i} \prod_{d=1}^i \frac{1}{m-2d} + \left( -\frac{1}{m-2} \right)^a \\ &= \left( -\frac{1}{m-2} \right)^a \prod_{d=1}^a \frac{1}{m-2d} \left[ 1 + \right. \\ &\quad \left. \sum_{i=0}^{a-1} \frac{a!}{i!(a-i)!} \{-(m-2)\}^i \prod_{d=1}^{a-i} \{m-2a+2(d-1)\} \right]. \end{aligned}$$

ここで,  $T \sim \chi_{m-2a}^2$  とすると,

$$E[T^{a-i}] = \begin{cases} 1 & (i = a) \\ \prod_{d=1}^{a-i} \{m-2a+2(d-1)\} & (i \leq a-1) \end{cases}$$

となるので,  $h = \{-(m-2)\}^{-a} \prod_{d=1}^a (m-2d)^{-1}$  とおけば,

$$\begin{aligned} E[(B-\nu)^a] &= h \sum_{i=0}^a \frac{a!}{i!(a-i)!} \{-(m-2)\}^i E[T^{a-i}] \\ &= h E[\{T-(m-2)\}^a] \\ &= h \sum_{i=0}^a \frac{a!}{i!(a-i)!} (-2a)^i E[\{T-(m-2a)\}^{a-i}]. \end{aligned}$$

ここで,  $m-2d = O(n)$  より,  $h = O(n^{-2a})$ . また, 補助定理 A.2 の (A.13) 式から,

$$E[\{T-(m-2a)\}^{a-i}] = \begin{cases} 1 & (i=a) \\ 0 & (i=a-1) \\ O(n^{[(a-i)/2]}) & (i \leq a-2) \end{cases},$$

を得る.  $i = 0, 1, \dots, a$  であるので, 上記の式の最大オーダーは  $O(n^{[a/2]})$  となり, この結果から,

$$E[(B-\nu)^a] = \begin{cases} 1 & (a=0) \\ 0 & (a=1) \\ O(n^{-2a+[a/2]}) & (a \geq 2) \end{cases}, \quad (\text{A.9})$$

を得る.  $\xi_1 = O(p)$ ,  $\nu = O(n^{-1})$  より, (A.8), (A.9) 式を用いると, (A.7) 式内の最大オーダーは,  $a = b = 0$ ,  $c = 2r$  のときなので, (16) 式を得ることができる.

また,  $L = U_2/V$ ,  $\mu = E[T]$  とおくと,

$$E[(L-\mu-\rho)^{2r}] = \sum_{i=0}^{2r} \frac{(2r)!}{i!(2r-i)!} E[(L-\mu)^{2r}] (-\rho)^{2r-i}. \quad (\text{A.10})$$

ただし,  $\rho$  は (15) 式で与えられた正の数であり, (5) 式から  $\rho = o(1)$  が言える.  $\xi_2 = E[U_2]$  おくと,  $\mu = \xi_2 \nu$  なので, (A.6) と同様な変形を行い, (A.7) 式と同様な展開により,  $E[(L-\mu)^a]$  のオーダーを求める. 仮定より,  $p/\delta = O(1)$  なので, 補助定理 A.2 の (A.14) 式より,

$$E[(U_2 - \xi_2)^a] = \begin{cases} 1 & (a=0) \\ 0 & (a=1) \\ O(\delta^{[a/2]}) & (a \geq 2) \end{cases}, \quad (\text{A.11})$$

よって,  $\xi_2 = O(\delta)$  より, (A.9), (A.11) 式より,

$$E[(L-\mu)^a] = \begin{cases} 1 & (a=0) \\ 0 & (a=1) \\ O(\delta^{[a/2]} n^{-a}) & (a \geq 2) \end{cases}, \quad (\text{A.12})$$

を得ることができる.  $\rho = o(1)$  であるので, (A.10), (A.12) 式より, (17) 式を得ることができる.

### A.3. 補助定理 A.1 とその証明

補助定理 A.1  $\Delta$  を  $p \times p$  半正定値行列とし,  $d = \text{rank}(\Delta) \leq \min\{p, q\}$ ,  $d \neq 0$  とする. 今,  $W_1$  と  $W_2$  を互いに独立な  $p \times p$  確率変数行列で,  $W_1 \sim W_p(q, I_p; \Delta)$ ,  $W_2 \sim W_p(n-k, I_p)$  とし, さらに,  $U_i$  と  $V$  を互いに独立な確率変数で,  $U_i \sim \chi_p^2(\delta_i)$ ,  $V \sim \chi_{n-k-p+1}^2$  ( $i = 1, \dots, q$ ) とする. ただし,  $\delta_1 + \dots + \delta_q = \text{tr}(\Delta)$  で,  $\lambda_1$  を  $\Delta$  の最大固有値とすると,  $\delta_i \geq \lambda_1/q$  を満たすものとする. このとき,

$$P(\text{tr}(W_1 W_2^{-1}) < \tau) \leq \sum_{i=1}^q P(U_i/V < \tau/q),$$

が成り立つ.

証明)  $Q$  を  $\Delta$  を対角化する  $p$  次直交行列とする. つまり,  $\lambda_i$  を  $\Delta$  の  $i$  番目に大きい固有値とし,  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$  とすると,  $Q' \Delta Q = \Lambda$  である. ただし, 仮定より,  $\lambda_1 > 0$  で,  $\lambda_{d+1} = \dots = \lambda_p = 0$  である. ここで,  $B_1 = Q' W_1 Q$ ,  $B_2 = Q' W_2 Q$  とおくと, ウィッシュヤート分布の特性より,

$$B_1 \perp B_2, \quad B_1 \sim W_p(q, I_p; \Lambda), \quad B_2 \sim W_p(n-k, I_p),$$

となる. 今,  $\Lambda_0 = \text{diag}(\lambda_1, \dots, \lambda_d)$ ,  $\mathcal{E} \sim N_{q \times p}(\mathbf{O}_{q,p}, I_p \otimes I_q)$ ,

$$\Gamma = \begin{pmatrix} \Lambda_0^{1/2} & \mathbf{O}_{d,p-d} \\ \mathbf{O}_{p-d,d} & \mathbf{O}_{p-d,p-d} \end{pmatrix},$$

とし,  $G = \mathcal{E} + \Gamma$  とおく. このとき,  $B_1$  の分布と  $G'G$  の分布は同じものになるので, 以下,  $G'G$  の分布を調べる.  $H = (h_1, \dots, h_q)$  を  $q$  次直交行列とし  $h_1 = \mathbf{1}_q/\sqrt{q}$  とする. ただし,  $\mathbf{1}_q$  はすべての成分が 1 である  $q$  次元ベクトルである. このとき,  $(\eta_1, \dots, \eta_q)' = H\Gamma$  とおけば,

$$(\eta_1, \dots, \eta_q)' = H \begin{pmatrix} \Lambda_0^{1/2} & \mathbf{O}_{d,p-d} \\ \mathbf{O}_{p-d,d} & \mathbf{O}_{p-d,p-d} \end{pmatrix} = (\sqrt{\lambda_1} h_1, \dots, \sqrt{\lambda_d} h_d, \mathbf{O}_{q,p-d}),$$

であるので,  $\delta_i = \|\eta_i\|^2$  とすると,  $\delta_i \geq \lambda_1/q$  を満たす. さらに,  $Z = (z_1, \dots, z_q)' = H\mathcal{E}$  とおくと,  $z_1, \dots, z_q \sim N_p(\mathbf{0}_p, I_p)$  であるので,

$$G'G = G'H'HG = (H\mathcal{E} + H\Gamma)'(H\mathcal{E} + H\Gamma) = \sum_{i=1}^q (z_i + \eta_i)(z_i + \eta_i)'$$

以上により,

$$\begin{aligned} \text{tr}(\mathbf{W}_1 \mathbf{W}_2^{-1}) &= \sum_{i=1}^q (\mathbf{z}_i + \boldsymbol{\eta}_i)' \mathbf{W}_2^{-1} (\mathbf{z}_i + \boldsymbol{\eta}_i) \\ &= \sum_{i=1}^q \frac{\|\mathbf{z}_i + \boldsymbol{\eta}_i\|^2}{\{\|\mathbf{z}_i + \boldsymbol{\eta}_i\|^{-1} (\mathbf{z}_i + \boldsymbol{\eta}_i)' \mathbf{W}_2^{-1} (\mathbf{z}_i + \boldsymbol{\eta}_i) \|\mathbf{z}_i + \boldsymbol{\eta}_i\|^{-1}\}^{-1}} \\ &= \sum_{i=1}^q T_i. \end{aligned}$$

ここで, 上記の式の二段目の右辺の分母分子は互いに独立で, それぞれ  $\chi_p^2(\delta_i)$ ,  $\chi_m^2$  に従うことから,  $T_i = U_i/V$  と書き換えることができる. ここで,

$$\bigcap_{i=1}^q \{T_i \geq \tau/q\} \Rightarrow \sum_{i=1}^q T_i \geq \tau,$$

なので,

$$\sum_{i=1}^q T_i < \tau \Rightarrow \bigcap_{i=1}^q \{T_i < \tau/q\},$$

となり, その結果,

$$P(\sum_{i=1}^q T_i < \tau) \leq \sum_{i=1}^q P(T_i < \tau/q),$$

と言える. よって題意が示せた. ■

#### A.4. 補助定理 A.2 とその証明

補助定理 A.2  $X_1 \sim \chi_f^2$  とすると,

$$E[(X_1 - f)^a] = \begin{cases} 1 & (a = 0) \\ 0 & (a = 1) \\ O(f^{[a/2]}) & (a \geq 2) \end{cases}, \quad (\text{A.13})$$

となる. また,  $X_2 \sim \chi_f^2(\delta)$ ,  $f/\delta = O(1)$  とすると,

$$E[\{X_2 - (f + \delta)\}^a] = \begin{cases} 1 & (a = 0) \\ 0 & (a = 1) \\ O(f^{[\delta/2]}) & (a \geq 2) \end{cases}, \quad (\text{A.14})$$

となる.

証明)  $a = 0, 1$  のときは明らか. Stuart & Ord [2] の 3.14 節より,  $a_1 + \dots + a_d = a$ ,  $2 \leq a_1, \dots, a_d \leq 2r$  とすると, 平均周りの  $a$  次モーメントは,  $a_1, \dots, a_d$  次キュムラントの積の線形結合で表現できる. Lancaster [1] より,  $X_1 - f$  の  $a$  次キュムラントは,

$$\kappa_{a,1} = 2^{a-1}(a-1)!f,$$

であり, Tiku [3] より,  $X_2 - (f + \delta)$  の  $a$  次キュムラントは,

$$\kappa_{a,2} = 2^{a-1}(a-1)!(f + a\delta),$$

である.  $f/\delta = O(1)$  であることに注意すれば,  $\kappa_{a,2} = O(\delta)$  である. よって, それぞれの平均周りの  $a$  次モーメントの最大次数の項は,  $a$  が偶数であれば,  $\kappa_{2,i}^{a/2}$  であり,  $a$  が奇数であれば,  $\kappa_{2,i}^{(a-1)/2-1} \kappa_{3,i}$  となる ( $i = 1, 2$ ). よって, 題意を得る. ■

## 引用文献

- [1] Lancaster, H. O. (1982). Chi-square distribution. In *Encyclopedia of Statistical Sciences, Vol. 1* (eds. S. Kotz & N. L. Johnson), 439–442, John Wiley & Sons, New York.
- [2] Stuart, A. & Ord, J. K. (1994). *Kendall's Advanced Theory of Statistics. Vol. 1. Distribution Theory* (6th ed.). Edward Arnold, London; distributed in the United States of America by Oxford University press, New York.
- [3] Tiku, M. (1985). Noncentral chi-square distribution. In *Encyclopedia of Statistical Sciences, Vol. 6* (eds. S. Kotz & N. L. Johnson), 276–280, John Wiley & Sons, New York.
- [4] Yanagihara, H. (2016). A high-dimensionality-adjusted consistent  $C_p$ -type statistic for selecting variables in a normality-assumed linear regression with multiple responses (submitted for publication).
- [5] Yao, J., Zheng, S. Bai, Z. (2015). *Large Sample Covariance Matrices and High-dimensional Data Analysis*. Cambridge University Press, New York.