

重回帰モデルでの変数選択における 一般化 C_p 規準の一致性の評価*

広島大学・理学研究科数学専攻[†] 柳原宏和

Hirokazu Yanagihara

Department of Mathematics, Graduate School of Science

Hiroshima University

2019 年 6 月 25 日

§1. 序

本論文では、重回帰モデルでの変数選択問題において、Atkinson [1] により提案された一般化 C_p (Generalized C_p ; GC_p) 規準を変数選択規準として用い、その最小化により最適な変数の組を決定する変数選択法を取り扱う。この GC_p 規準は、すべての変数を用いたモデルの下での不偏分散により基準化された最小残差平方和に変数の個数の α 倍を加えることにより定義され、Mallows [4] により提案された C_p 規準や、Fujikoshi and Satoh [2] により提案された Modified C_p (MC_p) 規準を特別な形として含む変数選択規準である。真の変数が変数選択規準により最適な変数として過不足無く選ばれる確率を漸近的に 1 にする性質である一致性、より正確に言えば弱一致性は、変数選択に用いる変数選択規準の重要な特性の一つである。本論文の目的は、一致性を満たすための GC_p 規準の定数 α の十分条件の導出方法をできるだけ詳しく丁寧に紹介することである。

以下本論文では、第 2 章で、重回帰モデルと GC_p 規準を紹介する。第 3 章で、主定理である、一致性を満たすための α の十分条件を求める。数学的な証明は付録に記載した。また、本論文では、行列を大文字の太文字で、列ベクトルを小文字の太文字で表し、また、 $'$ を行列またはベクトルの転置を表す記号とする。

本研究は科学研究費補助金、基盤研究 (C) (課題番号:18K03415) の助成を受けたものである。
〒739-8526 広島県東広島市鏡山 1-3-1

§2. 重回帰モデルと GC_p 規準

n を標本数として、目的変数 y と k 個の説明変数を並べた説明変数ベクトル $\mathbf{x}$ の組が $(y_1, \mathbf{x}'_1), \dots, (y_n, \mathbf{x}'_n)$ と n 個得られているとする. このとき、各個体の目的変数を並べた n 次元ベクトルと各個体の説明変数ベクトルを並べた $n \times k$ 行列をそれぞれ $\mathbf{y} = (y_1, \dots, y_n)'$, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$ とし, $\text{rank}(\mathbf{X}) = k$ を仮定する. 以下, k 個の説明変数のから一部を用いた重回帰モデルを考えるが, どの説明変数を用いたかは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, $\mathbf{X}$ 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 $\mathbf{X}_j$ とする. 例えば, $j = \{1, 2, 4\}$ であれば, $\mathbf{X}_j$ は $\mathbf{X}$ の第 1, 2, 4 列を抜き出して並べた行列である. この $\mathbf{X}_j$ を使って, 考える候補のモデルを以下のように記述する.

$$\mathbf{y} - E[\mathbf{y}] \text{ の成分は互いに独立に同一の分布に従い, } E[\mathbf{y}] = \mathbf{X}_j \boldsymbol{\beta}_j, \text{Cov}[\mathbf{y}] = \sigma_j^2 \mathbf{I}_n. \quad (1)$$

ただし, $\boldsymbol{\beta}_j$ は未知回帰係数を並べた k_j 次元回帰係数ベクトル, σ_j^2 は正值を仮定した未知の分散, さらに $\mathbf{I}_n$ は n 次の単位行列である. (1) 式は, j により定式化されるので, (1) 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, $\mathbf{X}_\omega = \mathbf{X}$ であるに注意する.

真のモデルは以下の様に定義する.

$$E[\mathbf{y}] = \boldsymbol{\mu}_* = (\mu_1^*, \dots, \mu_n^*)', \text{Cov}[\mathbf{y}] = \sigma_*^2 \mathbf{I}_n, \\ \boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)' = (\mathbf{y} - \boldsymbol{\mu}_*)/\sigma_*, \varepsilon_1, \dots, \varepsilon_n \sim i.i.d. \varepsilon, E[\varepsilon] = 0, \text{Var}[\varepsilon] = 1. \quad (2)$$

ただし, $\boldsymbol{\mu}_*$ は真の平均を並べた n 次元ベクトル, σ_*^2 は真の分散である. ここでの σ_*^2 と ε の確率分布は, n に依存しない, つまり, n によって変化しないとす. さらに, $\boldsymbol{\mu}_*$ の各成分は数列 $\{\mu_m^*\}_{m=1,2,\dots}$ に対応するとする. 付け加えて, 第 ℓ 番目 ($\ell = 1, \dots, k$) の説明変数の値は数列 $\{x_{m\ell}\}_{m=1,2,\dots}$ に対応するとする.

$\mathcal{J}$ を考えうる候補のモデルをすべて集めた集合とする. つまり, $\wp(A)$ を集合 A の冪集合とすると, $\mathcal{J} = \wp(\omega)$ である. ここで, $\mathcal{J}$ を過多に記述された (overspecified) モデルを集めた集合 $\mathcal{J}_+$ と過少に記述された (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, $\mathbf{P}_j$ を $\mathbf{X}_j$ で張る空間への直交射影行列, つまり,

$$\mathbf{P}_j = \mathbf{X}_j(\mathbf{X}_j' \mathbf{X}_j)^{-1} \mathbf{X}_j',$$

とおくと, $\mathcal{J}_+$ と $\mathcal{J}_-$ は以下のように定義される.

$$\mathcal{J}_+ = \{j \in \mathcal{J} \mid (\mathbf{I}_n - \mathbf{P}_j)\boldsymbol{\mu}_* = \mathbf{0}_n\}, \quad \mathcal{J}_- = \mathcal{J}_+^c \cap \mathcal{J}. \quad (3)$$

ただし, $\mathbf{0}_n$ はすべての成分が 0 である n 次元ベクトル, A^c は集合 A の補集合を表す. $\mathcal{J}_+ \neq \emptyset$ のとき, 以下のように定義される集合を真の変数の組と定義する.

$$j_* = \arg \min_{j \in \mathcal{J}_+} k_j. \quad (4)$$

よって, $\mathcal{J}_+ \neq \emptyset$ のとき j_* を真のモデルと呼ぶ.

(1) 式のモデル j における GC_p 規準は, 以下の様に定義される.

$$GC_p(j|\alpha) = \frac{\mathbf{y}'(\mathbf{I}_n - \mathbf{P}_j)\mathbf{y}}{s_\omega^2} + \alpha k_j. \quad (5)$$

ただし, α はモデルの複雑さに対する罰則を調整する正数であり, s_ω^2 はフルモデルの下での分散の不偏推定量, つまり,

$$s_\omega^2 = \frac{1}{n-k} \mathbf{y}'(\mathbf{I}_n - \mathbf{P}_\omega)\mathbf{y}.$$

GC_p 規準は既存の変数選択規準を特別な形として含む変数選択規準である. それら既存の変数選択規準との関係は以下の通りである.

$$\alpha = \begin{cases} 2 & (C_p) \\ 2 \left(\frac{n-k}{n-k-2} \right) & (MC_p) \end{cases}.$$

このとき, GC_p 規準によって選ばれたモデルを $\hat{j}_\alpha$ と書く. つまり,

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

であるとする. これから, $\hat{j}_\alpha$ と j_* を使うと, 本論文で取り扱う一致性とは, $\mathcal{J}_+ \neq \emptyset$ のとき,

$$\lim_{n \rightarrow \infty} P \left(\hat{j}_\alpha = j_* \right) = 1, \quad (6)$$

を満たすことであり, (6) 式を満たすための α の十分条件を求めることが目的となる.

変数選択問題で重要な役割を果たすものに非心パラメータがある. モデル j における非心パラメータは以下のように定義される.

$$\delta_j = \frac{1}{\sigma_*^2} \boldsymbol{\mu}'_*(\mathbf{I}_n - \mathbf{P}_j)\boldsymbol{\mu}_*. \quad (7)$$

非心パラメータという名前は, (2) 式での ε の確率分布が正規分布であるとき, $\mathbf{y}'(\mathbf{I}_n - \mathbf{P}_j)\mathbf{y}/\sigma_*^2$ が自由度 $n - k_j$, 非心パラメータ δ_j を持つ非心 χ^2 分布に従うことに由来する. $\mathcal{J}_+$ の定義より, $j \in \mathcal{J}_+$ のときは $\delta_j = 0$ となる. 反対に, $j \in \mathcal{J}_-$ ときは, $\delta_j > 0$ となる. また, 非心パラメータは以下の性質を持つ (証明は付録 A.1 を参照).

定理 1 すべての $j \in \mathcal{J}_-$ に対して, δ_j は n に関して単調増加である.

§3. GC_p 規準の一致性

本章では, (5) 式の α に関して, 実際に (6) 式を満たすための十分条件を求める. そのため, 以下の仮定を用意する.

仮定 A1 $\mathcal{J}_+ \neq \emptyset$. ただし, $\mathcal{J}_+$ は (3) 式で与えられた過多に記述されたモデルを集めた集合である.

仮定 A2 $\mathbf{X} = (\mathbf{a}_1, \dots, \mathbf{a}_k)$ とすると, $\forall \ell \in \omega$ に対して,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \mathbf{a}'_\ell (\mathbf{I}_n - \mathbf{P}_{\omega \setminus \{\ell\}}) \mathbf{a}_\ell > 0.$$

仮定 A1 は真のモデルが候補のモデルに含まれていることを保証するもので, 一致性を評価するときには必ず必要となる仮定である. また, 仮定 A2 が満たされているとき, (7) 式の非心パラメータは以下のような漸近性質を持つことになる (証明は付録 A.2 を参照).

$$\forall j \in \mathcal{J}_-, \liminf_{n \rightarrow \infty} \frac{\delta_j}{n} > 0, \quad (8)$$

定理 1 より δ_j は n に関して単調増加であることがわかっていたが, 仮定 A2 を追加することにより, δ_j は n の増加とともに ∞ に発散し, 逆数が 0 に収束するオーダーを仮定していることになる. つまり, (8) 式より,

$$\forall j \in \mathcal{J}_-, \limsup_{n \rightarrow \infty} \frac{n}{\delta_j} < \infty,$$

であるので,

$$\frac{n}{\delta_j} = O(1), \quad (9)$$

が言える.

次に, 一致性に関する主定理を証明するために必要な補助定理を用意する (証明はそれぞれ, 付録 A.3, A.4, A.5 を参照).

補助定理 1 h_j を $\liminf_{n \rightarrow \infty} h_j > 0$ を満たすモデルに依存する正数とする. このとき以下が成り立つ.

$$\begin{aligned} \forall j \in \mathcal{J} \setminus \{\ell\}, \quad \frac{1}{h_j} \{GC_p(j|\alpha) - GC_p(\ell|\alpha)\} &\geq T_j \xrightarrow{p} \tau_j > 0 \\ &\Rightarrow \lim_{n \rightarrow \infty} P(\hat{j}_\alpha = \ell) = 1. \end{aligned}$$

補助定理 2 仮定 A1 が成り立っているとする. このとき, j_* を (4) 式で与えられた真のモデルとすると,

$$GC_p(j|\alpha) - GC_p(j_*|\alpha) = \frac{1}{v} \left(\delta_j - 2\sqrt{\delta_j} t_j + w_{j_*} - w_j \right) + \alpha(k_j - k_{j_*}),$$

と変形できる. ただし,

$$\gamma_j = \begin{cases} \mathbf{0}_n & (\delta_j = 0) \\ \frac{1}{\sigma_* \sqrt{\delta_j}} (\mathbf{I}_n - \mathbf{P}_j) \boldsymbol{\mu}_* & (\delta_j \neq 0) \end{cases},$$

とし, $\boldsymbol{\varepsilon}$ を (2) 式で与えられる n 次元確率変数ベクトルとすると,

$$v = \frac{1}{n-k} \boldsymbol{\varepsilon}' (\mathbf{I}_n - \mathbf{P}_\omega) \boldsymbol{\varepsilon}, \quad t_j = \gamma_j' \boldsymbol{\varepsilon}, \quad w_j = \boldsymbol{\varepsilon}' \mathbf{P}_j \boldsymbol{\varepsilon}. \quad (10)$$

補助定理 3 仮定 A1 が成り立っているとし, ρ を $\lim_{n \rightarrow \infty} \rho = \infty$ を満たす正数とする. . このとき, (2) 式の真のモデルの下で, (10) 式の v, t_j, w_j は以下を満たす.

$$v \xrightarrow{p} 1, \quad \frac{1}{\rho} t_j \xrightarrow{p} 0, \quad \frac{1}{\rho} w_j \xrightarrow{p} 0,$$

補助定理 1 は, Yanagihara (2015) で与えられたものと同じ, 一般的な状況に適用可能な補助定理となっている. しかしながら, 主定理は, $\{GC_p(j|\alpha) - GC_p(j_*|\alpha)\}/h_j$ が正の値に確率収束することが言えるため, T_j のような下限を使用することなく証明できる.

以上の補助定理を用いて, 以下のような主定理を得る.

定理 2 仮定 A1, A2 が成り立っているとする. このとき,

$$\lim_{n \rightarrow \infty} \alpha = \infty, \quad \lim_{n \rightarrow \infty} \frac{\alpha}{n} = 0, \quad (11)$$

を満たすとき, GC_p 規準は一致性を持つ.

証明) 定数 h_j を以下のように定義する.

$$h_j = \begin{cases} \alpha & (j \in \mathcal{J}_+ \setminus \{j_*\}) \\ \delta_j & (j \in \mathcal{J}_-) \end{cases}.$$

このとき, (8) 式と (11) 式より, $\lim_{n \rightarrow \infty} h_j = \infty$ である. 今,

$$C_j = \begin{cases} \frac{1}{v} \left(\frac{w_{j_*} - w_j}{\alpha} \right) + k_j - k_{j_*} & (j \in \mathcal{J}_+ \setminus \{j_*\}) \\ \frac{1}{v} \left(1 - 2 \frac{t_j}{\sqrt{\delta_j}} + \frac{w_{j_*} - w_j}{\delta_j} \right) + \frac{\alpha}{\delta_j} (k_j - k_{j_*}) & (j \in \mathcal{J}_-) \end{cases}, \quad (12)$$

とすれば, h_j と補助定理 2, さらに $j \in \mathcal{J}_+$ のとき $\delta_j = 0$ となることを用いれば,

$$\frac{1}{h_j} \{GC_p(j|\alpha) - GC_p(j_*|\alpha)\} = C_j, \quad (13)$$

となる. ここで補助定理 3 より,

$$v \xrightarrow{p} 1, \quad \frac{w_{j_*} - w_j}{h_j} \xrightarrow{p} 0, \quad \frac{t_j}{\sqrt{\delta_j}} \xrightarrow{p} 0, \quad (14)$$

を満たす. また, (9) 式と (11) 式より,

$$\frac{\alpha}{\delta_j} = \frac{\alpha}{n} \cdot \frac{n}{\delta_j} \rightarrow 0, \quad (15)$$

と言える. よって, (14) 式と (15) 式を (12) 式に代入すると,

$$C_j \xrightarrow{p} \tau_j = \begin{cases} k_j - k_{j_*} & (j \in \mathcal{J}_+ \setminus \{j_*\}) \\ 1 & (j \in \mathcal{J}_-) \end{cases},$$

を得る. $j \in \mathcal{J}_+ \setminus \{j_*\}$ のとき $k_j - k_{j_*} > 0$ であるので, $\tau_j > 0$. この結果を (13) 式に代入すると,

$$\frac{1}{h_j} \{GC_p(j|\alpha) - GC_p(j_*|\alpha)\} \xrightarrow{p} \tau_j > 0.$$

以上により, 補助定理 1 より題意を得る. ■

A. 付録

A.1. 定理 1 の証明

$\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_n)'$ を $n \times h$ ($n > h$) 行列で, $\text{rank}(\mathbf{H}) = h$ とし, $\mathbf{P}_\mathbf{H}$ を $\mathbf{H}$ で張られる空間への直交射影行列, つまり, $\mathbf{P}_\mathbf{H} = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'$ とおく. このとき, 題意を示すには,

$$d_n = \frac{1}{\sigma_*^2} \boldsymbol{\mu}'_*(\mathbf{I}_n - \mathbf{P}_\mathbf{H})\boldsymbol{\mu}_*,$$

が n に関して単調増加列であることを示せばよい. $\boldsymbol{\theta} \in \mathbb{R}^h$ とし,

$$R_n(\boldsymbol{\theta}) = (\boldsymbol{\mu}_* - \mathbf{H}\boldsymbol{\theta})'(\boldsymbol{\mu}_* - \mathbf{H}\boldsymbol{\theta}) = \sum_{i=1}^n (\mu_i^* - \mathbf{h}_i' \boldsymbol{\gamma})^2,$$

とおくと, 簡単な線形代数から,

$$\sigma_*^2 d_n = \min_{\boldsymbol{\theta} \in \mathbb{R}^h} R_n(\boldsymbol{\theta}),$$

であると言える.

$$\hat{\boldsymbol{\theta}}_n = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^h} R_n(\boldsymbol{\theta}),$$

とすると,

$$R_{n+1}(\hat{\boldsymbol{\theta}}_{n+1}) = R_n(\hat{\boldsymbol{\theta}}_{n+1}) + (\mu_{n+1}^* - \mathbf{h}_{n+1}' \hat{\boldsymbol{\theta}}_{n+1})^2 \geq R_n(\hat{\boldsymbol{\theta}}_{n+1}) \geq R_n(\hat{\boldsymbol{\theta}}_n),$$

となり, $R_n(\hat{\boldsymbol{\theta}}_n) \leq R_{n+1}(\hat{\boldsymbol{\theta}}_{n+1})$ を得る. よって, $d_n \leq d_{n+1}$ となり題意を得る. ■

A.2. (8) 式の証明

仮定 A1 が成り立っているとする. $h \in j_*$ とし, β_h^* を h 番目の説明変数に対する真の回帰係数とすると, $\boldsymbol{\mu}_*$ は,

$$\boldsymbol{\mu}_* = \sum_{h \in j_*} \beta_h^* \mathbf{a}_h,$$

と書くことができる. ただし, $\mathbf{a}_h$ は仮定 A2 で定義された $\mathbf{X}$ の第 h 列ベクトルである. ここで,

$$(\mathbf{I}_n - \mathbf{P}_j) \mathbf{a}_h \quad \begin{cases} = \mathbf{0}_n & (h \in j_* \cap j \text{ のとき}) \\ \neq \mathbf{0}_n & (h \in j_* \cap j^c \text{ のとき}) \end{cases},$$

であり, $\forall \ell \in j_* \cap j^c$ に対して $j \subset \omega \setminus \{\ell\}$ であることに注意すれば, (7) 式の非心パラメータは,

$$\begin{aligned} \delta_j &= \frac{1}{\sigma_*^2} \left(\sum_{h \in j_* \cap j^c} \beta_h^* \mathbf{a}_h \right) (\mathbf{I}_n - \mathbf{P}_j) \left(\sum_{h \in j_* \cap j^c} \beta_h^* \mathbf{a}_h \right) \\ &\geq \frac{1}{\sigma_*^2} \left(\sum_{h \in j_* \cap j^c} \beta_h^* \mathbf{a}_h \right) (\mathbf{I}_n - \mathbf{P}_{\omega \setminus \{\ell\}}) \left(\sum_{h \in j_* \cap j^c} \beta_h^* \mathbf{a}_h \right) \\ &= \frac{\beta_\ell^{*2}}{\sigma_*^2} \mathbf{a}_\ell' (\mathbf{I}_n - \mathbf{P}_{\omega \setminus \{\ell\}}) \mathbf{a}_\ell, \end{aligned}$$

と書き換えることができる. 以上により, $\beta_\ell^* \neq 0$ であることに注意すれば, 仮定 A2 が成り立てば,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \delta_j = \frac{\beta_\ell^{*2}}{\sigma_*^2} \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbf{a}'_\ell (\mathbf{I}_n - \mathbf{P}_{\omega \setminus \{\ell\}}) \mathbf{a}_\ell > 0.$$

よって題意を得る. ■

A.3. 補助定理 1 の証明

基本的な確率の定理から, 以下が成り立つ.

$$\begin{aligned} P(\hat{j}_\alpha = j_*) &= 1 - P(\hat{j}_\alpha \neq j_*) \\ &= 1 - P(\cup_{j \in \mathcal{J} \setminus \{j_*\}} \{GC_p(j|\alpha) < GC_p(j_*|\alpha)\}) \\ &\geq 1 - \sum_{j \in \mathcal{J} \setminus \{j_*\}} P(GC_p(j|\alpha) < GC_p(j_*|\alpha)). \end{aligned} \quad (\text{A.1})$$

ここで, $\forall j \in \mathcal{J} \setminus \{j_*\}$ で $T_j \xrightarrow{P} \tau_j > 0$ が成り立つならば, $\forall \epsilon > 0$ に対して,

$$P(|T_j - \tau_j| > \epsilon) \rightarrow 0, \quad (\text{A.2})$$

が成り立つ. また, $\{GC_p(j|\alpha) - GC_p(j_*|\alpha)\}/h_j \geq T_j$ であるので,

$$GC_p(j|\alpha) - GC_p(j_*|\alpha) < 0 \implies T_j < 0,$$

を満たす. 以上により,

$$\begin{aligned} P(|T_j - \tau_j| > \tau_j) &= P(\{T_j > 2\tau_j\} \cup \{T_j < 0\}) \geq P(T_j < 0) \\ &\geq P(\{GC_p(j|\alpha) - GC_p(j_*|\alpha)\}/h_j < 0) \\ &= P(GC_p(j|\alpha) < GC_p(j_*|\alpha)). \end{aligned}$$

上記の左辺は (A.2) 式より 0 に収束するので, $P(GC_p(j|\alpha) < GC_p(j_*|\alpha)) \rightarrow 0$. この式を (A.1) 式に代入すれば, 題意を得る. ■

A.4. 補助定理 2 の証明

簡単な計算から,

$$\begin{aligned} \frac{1}{\sigma_*^2} \mathbf{y}' (\mathbf{I}_n - \mathbf{P}_j) \mathbf{y} &= \left(\frac{1}{\sigma_*} \boldsymbol{\mu}_* + \boldsymbol{\varepsilon} \right)' (\mathbf{I}_n - \mathbf{P}_j) \left(\frac{1}{\sigma_*} \boldsymbol{\mu}_* + \boldsymbol{\varepsilon} \right) \\ &= \delta_j + 2\sqrt{\delta_j} \boldsymbol{\gamma}'_j \boldsymbol{\varepsilon} + \|\boldsymbol{\varepsilon}\|^2 - w_j, \end{aligned} \quad (\text{A.3})$$

と変形できる. 仮定 A1 が成り立っているとき, $\forall j \in \mathcal{J}_+$ に対して $\delta_j = 0$ であるので, (A.3) 式より,

$$\frac{1}{\sigma_*^2} \mathbf{y}'(\mathbf{I}_n - \mathbf{P}_{j_*}) \mathbf{y} = \|\varepsilon\|^2 - w_{j_*}, \quad \frac{s_\omega^2}{\sigma_*^2} = \frac{1}{n-k} (\|\varepsilon\|^2 - w_\omega) = v,$$

が成り立つ. よって,

$$\begin{aligned} & GC_p(j|\alpha) - GC_p(j_*|\alpha) \\ &= \frac{\sigma_*^2}{s_\omega^2} \left\{ \frac{\mathbf{y}'(\mathbf{I}_n - \mathbf{P}_j) \mathbf{y}}{\sigma_*^2} - \frac{\mathbf{y}'(\mathbf{I}_n - \mathbf{P}_{j_*}) \mathbf{y}}{\sigma_*^2} \right\} + \alpha(k_j - k_{j_*}) \\ &= \frac{1}{v} \left\{ \delta_j + 2\sqrt{\delta_j} \gamma_j' \varepsilon + \|\varepsilon\|^2 - w_j - (\|\varepsilon\|^2 - w_{j_*}) \right\} + \alpha(k_j - k_{j_*}) \\ &= \frac{1}{v} \left(\delta_j + 2\sqrt{\delta_j} \gamma_j' \varepsilon + w_{j_*} - w_j \right) + \alpha(k_j - k_{j_*}). \end{aligned}$$

よって, 題意を得る. ■

A.5. 補助定理 3 の証明

まず, t_j について考える. $j \in \mathcal{J}_+$ のときは明らかであるので, $j \in \mathcal{J}_-$ のときを考える. このとき $\|\gamma_j\| = 1$ に注意すると,

$$E[t_j] = 0, \quad \text{Var}[t_j] = \gamma_j' E[\varepsilon \varepsilon'] \gamma_j = \gamma_j' \mathbf{I}_n \gamma_j = \|\gamma_j\|^2 = 1,$$

を得るので, $t_j = O_p(1)$ が成り立つ. よって, $t_j/\rho = O_p(\rho^{-1})$ となるので, $t_j/\rho \xrightarrow{P} 0$ を得る. 次に w_j について考える.

$$E[w_j] = E[\text{tr}(\mathbf{P}_j \varepsilon \varepsilon')] = \text{tr}(\mathbf{P}_j \mathbf{I}_{k_j}) = \text{tr}(\mathbf{P}_j) = k_j,$$

であり, $w_j \geq 0$ なのでマルコフの不等式より, $\forall \epsilon > 0$ に対して,

$$P(w_j/\rho > \epsilon) \leq \frac{E[w_j]}{\epsilon \rho} = \frac{k_j}{\epsilon \rho} \rightarrow 0.$$

よって, $w_j/\rho \xrightarrow{P} 0$. 最後に, v について考える. このとき,

$$r = \frac{1}{n} \sum_{i=1}^k \varepsilon_i^2,$$

とおくと,

$$v = \left(\frac{n}{n-k} \right) r - \frac{w_\omega}{n-k}. \quad (\text{A.4})$$

今, $E[\varepsilon^2] = 1 < \infty$ であるので, 大数の法則より, $r \xrightarrow{P} E[\varepsilon^2] = 1$ である (1 次モーメントの存在しか課さない大数の法則は, 例えば, 藤越ら [3] を参照). さらに, $w_j/\rho \xrightarrow{P} 0$ なので, $w_\omega/(n-k) \xrightarrow{P} 0$. 以上の結果を (A.4) 式に代入すれば, $v \xrightarrow{P} 1$. よって, 題意を得る.

■

引用文献

- [1] Atkinson, A. C. (1980). A note on the generalized information criterion for choice of a model. *Biometrika*, **67**, 413–418.
- [2] Fujikoshi, Y. & Satoh, K. (1997). Modified AIC and C_p in multivariate linear regression. *Biometrika*, **84**, 707–716.
- [3] 藤越康祝, 若木宏文, 柳原宏和 (2011). 確率・統計の数学的基礎. 広島大学出版会, 広島.
- [4] Mallows C. L. (1973). Some comments on C_p . *Technometrics*, **15**, 661–675.
- [5] Yanagihara, H. (2015). Conditions for consistency of a log-likelihood-based information criterion in normal multivariate linear regression models under the violation of normality assumption. *J. Japan Statist. Soc.*, **45**, 21–56.