有限混合モデルにおける識別可能性と Fisher 情報行列の正値定符号性

高崎経済大学　宮田庸一

1 はじめに

統計モデルにおける識別可能性は, 漸近理論において必要な仮定となっている. また推定量の漸近正規性を議論すためには, Fisher 情報行列が有限で正値定符号であることは, 重要な条件となっている. しかしながら Fisher 情報行列の有限性, 正値定符号性にはあまり目を向けられてはいない. 例えば最近の Lee and Lee (2009, 誤差項が正規混合モデルになっている GARCH モデル), Svenssona, and Lunab (2010, 打ち切り logistic 回帰の混合) においても, Fisher 情報行列の正値定符号性を仮定している.

Teicher(1963) は有限混合モデルにおける識別可能性に対して, 確認のしやすい十分条件を与えた. 本報告では, Teicher(1963) の方法を拡張することにより, あるパラメーター空間の下で正規混合モデルは Fisher 情報行列が正値定符号であり, 最尤推定量 (MLE) がある種の漸近正規性を持つことを厳密に示す. またこのアプローチは, 片側打ち切り混合モデルの識別可能性, Fisher 情報行列の正値定符号性, 漸近正規性の証明に応用できることも確認した.

2 準備

本報告では有限混合モデルとして, 以下の分布関数 (CDF)$H(x|\boldsymbol{\gamma})$ を考える.

$$H(x|\boldsymbol{\gamma}) = \sum_{j=1}^{g} \pi_j F(x|\boldsymbol{\theta}_j) \tag{1}$$

ここで $\pi_1, ..., \pi_g$ を mixing パラメーターとし, $F(x|\boldsymbol{\theta}_j)$ を各コンポーネントの CDF とする. また各コンポーネントの CDF $F(x|\boldsymbol{\theta}_j)$ は, パラメーター $\boldsymbol{\theta}_j$ に関して識別可能であるとする. このとき, パラメーター空間を

$$\begin{aligned}\Gamma_g = \{(\pi_1, \pi_2, ..., \pi_{g-1}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_g) | 0 < \pi_j < 1, \sum_{j=1}^{g-1} \pi_j < 1 \\ \boldsymbol{\theta}_j \in \Theta (j = 1, 2, ..., g), \boldsymbol{\theta}_i \neq \boldsymbol{\theta}_j \text{ for } \forall i, j (i \neq j)\}\end{aligned} \tag{2}$$

により定義する. ここで Θ は $F(x|\boldsymbol{\theta}_j)$ に対応するパラメーター空間で j に依存しない.

3 識別可能性

有限混合モデルにおいては, 通常の統計理論のように, 異なるパラメーターに対して異なる確率分布が対応するような通常の意味での識別可能性は成り立たない. ここでは, 有限混合モデルにおける識別可能性を紹介する. 詳しくは Tiitterington, Smith, and Makov (1985) を参照のこと.

まず $F(x|\boldsymbol{\theta}_j)$ の有限混合のクラスを $\mathcal{H} = \bigcup_{g=1}^{\infty} \left\{ \sum_{j=1}^{g} \pi_j F(x|\boldsymbol{\theta}_j), \boldsymbol{\gamma} \in \Gamma_g \right\}$ とする.

Definition 1 *For any* $H(x|\boldsymbol{\gamma}),\ H(x|\boldsymbol{\gamma}^*) \in \mathcal{H}$,

$$\sum_{j=1}^{g} \pi_j F(x|\boldsymbol{\theta}_j) = \sum_{j=1}^{g^*} \pi_j^* F(x|\boldsymbol{\theta}_j^*) \tag{3}$$

$\Longrightarrow g = g^*$, ある順列 ρ が存在して $\pi_j = \pi^*_{\rho(j)}$, $\boldsymbol{\theta}_j = \boldsymbol{\theta}^*_{\rho(j)}$, $(j = 1, 2, ..., g)$.
この時, $\mathcal{H}$ は識別可能であると言う.

4 Fisher 情報行列の正値定符号性

Teicher(1963) の方法を拡張することにより, 正規混合モデルにおける情報行列が正値定符号 ($I(\boldsymbol{\gamma}) > 0$ と書く) であることを示した. 最初に Fisher 情報行列を定義する.

$$I(\boldsymbol{\gamma}) = E_{\boldsymbol{\gamma}}\left[\frac{\partial}{\partial \boldsymbol{\gamma}} \log h(X|\boldsymbol{\gamma}) \frac{\partial}{\partial \boldsymbol{\gamma}^T} \log h(X|\boldsymbol{\gamma})\right]$$

Theorem 1 コンポーネント数 g は既知とする. 各コンポーネントが異なる平均 μ_j, 分散 σ_j^2 を持つ有限正規混合モデルにおいて, パラメーター空間 Γ_g の下で, 任意の $\boldsymbol{\gamma} \in \Gamma_g$ に対して, $I(\boldsymbol{\gamma})$ は正値定符号

5 漸近正規性

$Int\Gamma_g$ を Γ_g の内点の集合, $\partial/\partial\gamma_i = \partial_i$, $\partial^2/\partial\gamma_i\partial\gamma_j = \partial_{ij}$, $\partial^3/\partial\gamma_i\gamma_j\gamma_k = \partial_{ijk}$. $L(\boldsymbol{\gamma})$:尤度関数, $\mathcal{M}_n = \{\boldsymbol{\gamma} \in \Gamma_g | L(\boldsymbol{\gamma}) = L(\boldsymbol{\gamma}^n), \boldsymbol{\gamma}^n \text{は} MLE\}$. $\Gamma_g(\boldsymbol{\gamma}_0) = \{(\boldsymbol{\pi}, \boldsymbol{\theta}) \in \Gamma_g | H(x|\boldsymbol{\gamma}) = H(x|\boldsymbol{\gamma}_0)\}$:真のパラメーターと同一視できる集合. $\tilde{\Gamma}_g(\boldsymbol{\gamma}_0) = \{(\pi^0_{\rho_s(1)}, ..., \pi^0_{\rho_s(k-1)}, \boldsymbol{\theta}^0_{\rho_s(1)}, ..., \boldsymbol{\theta}^0_{\rho_s(k)}), s = 1, ..., k\} \cap \Gamma_g$:真の母数をラベルスイッチングすることで得られる有限可算個の集合. ここで $\rho_s : \{1, 2, ..., k\} \to \{1, 2, ..., k\}$, $(s = 1, ..., k!)$ は順列（置換）とする.

Theorem 2 以下の *(C1)-(C6)* を仮定する.
(C1) $\boldsymbol{\gamma}^0 \in Int\Gamma_g$. $\boldsymbol{\gamma}^0$ の近傍 $U_\delta(\boldsymbol{\gamma}^0)$ *and* $N > 0$, *s.t.*, $\forall \boldsymbol{\gamma} \in U_\delta(\boldsymbol{\gamma}^0)$ に対して $\partial_i h(x|\boldsymbol{\gamma})$, $\partial_{ij} h(x|\boldsymbol{\gamma})$, $\partial_{jkl} h(x|\boldsymbol{\gamma})$ が存在 *a.e.*
(C2) $\exists\ M_{jkl}(x)$ *s.t.*, $|\partial_{jkl} h(x|\boldsymbol{\gamma})| \le M_{jkl}(x)$ *for* $\forall \boldsymbol{\gamma} \in U_\delta(\boldsymbol{\gamma}^0)$ *and* $E_{\boldsymbol{\gamma}^0}[M_{jkl}(X)] < \infty$*a.e.*,
(C3) $E_{\boldsymbol{\gamma}^0}\left[\frac{\partial}{\partial \boldsymbol{\gamma}} \log h(X|\boldsymbol{\gamma}^0)\right] = \mathbf{0}$, $E_{\boldsymbol{\gamma}^0}\left[\frac{\partial}{\partial \boldsymbol{\gamma}} \log h(X|\boldsymbol{\gamma}^0) \frac{\partial}{\partial \boldsymbol{\gamma}^T} \log h(X|\boldsymbol{\gamma}^0)\right] = -E_{\boldsymbol{\gamma}^0}\left[\frac{\partial^2}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} \log h(X|\boldsymbol{\gamma})\right] = \mathbf{I}(\boldsymbol{\gamma})$.
(C4) $\mathbf{I}(\boldsymbol{\gamma})$ は有限で, 正値定符号.
(C5) for any MLE $\boldsymbol{\gamma}^n$ *on* Γ_g, $dis\{\boldsymbol{\gamma}^n, \Gamma_g(\boldsymbol{\gamma}^0)\} \to 0$ *w.p.1*
(C6) $\Gamma_g(\boldsymbol{\gamma}^0) = \tilde{\Gamma}_g(\boldsymbol{\gamma}^0)$
$\Longrightarrow \tilde{\boldsymbol{\gamma}}(\boldsymbol{\gamma}^0) = arginf_{\boldsymbol{\gamma} \in \mathcal{M}_n} |\boldsymbol{\gamma} - \boldsymbol{\gamma}^0|$ に対して

$$\sqrt{n}(\tilde{\boldsymbol{\gamma}}(\boldsymbol{\gamma}^0) - \boldsymbol{\gamma}^0) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}(\boldsymbol{\gamma})^{-1}). \tag{4}$$

任意の関数 $g(X) \in L^1(P_{\boldsymbol{\gamma}})$ に対して,

$$\sqrt{n}(E_{\boldsymbol{\gamma}^n}[g(X)] - E_{\boldsymbol{\gamma}^0}[g(X)]) \xrightarrow{L} N\left(0, \frac{\partial}{\partial \boldsymbol{\gamma}^T} E_{\boldsymbol{\gamma}^0}[g(X)] \mathbf{I}(\boldsymbol{\gamma})^{-1} \frac{\partial}{\partial \boldsymbol{\gamma}} E_{\boldsymbol{\gamma}^0}[g(X)]\right)$$

非正規性の下での共通平均の推定量について

井上 淳（早大政経）

1. はじめに

平均 $\mu\mathbf{1}_{n_i}(\mathbf{1}_{n_i}=(1,\ldots,1)'(n_i\times 1))$，分散 $\sigma_i^2\,\mathrm{I}_{n_i}$ (I_{n_i}: n_i 次単位行列，σ_i^2: 未知正数) をもつ n_i 次元観測ベクトル $y_i=(y_{i1},y_{i2},\ldots,y_{in_i})'$ $(i=1,2,\ldots,k)$ に基づいて，未知の実数値 μ を推定する問題を考える．本報告では，次の (1.1) で定められる $\widehat{\mu}_{FG}$（μ の実行可能な一般化最小二乗推定量）を取り扱い，その性質について述べる：

$$\widehat{\mu}_{FG} \;=\; \left(\sum_{i=1}^{k} n_i\,\widehat{\sigma}_i^{\,-2}\,\overline{y}_i\right)\left(\sum_{i=1}^{k} n_i\,\widehat{\sigma}_i^{\,-2}\right)^{-1} \tag{1.1}$$

$\widehat{\mu}_{FG}$ を用いる際に，次式が満たされることが前提となる：

$$\mathrm{V}\left(\widehat{\mu}_{FG}\right) \;\le\; \min_{1\le i\le k}\mathrm{V}(\overline{y}_i)\left(=\;\min_{1\le i\le k}(\sigma_i^2/n_i)\right) \tag{1.2}$$

Shinozaki (1978) は，正規性の条件下で不等式 (1.2) が成り立つための必要十分条件 $2(n_i-1)(n_j-5)\;\ge\;(n_i+1)(n_j-1)$ $(i\ne j)$　を導いている．この条件は $(n_i-3)(n_j-9)\ge 16$ （$i\ne j$) が満たされること，すなわち

$$\left[\,n_i=10 \text{ and } \min_{j\ne i} n_j\ge 19\,\right] \text{ or } \left[\,\min_{1\le i\le k} n_i\ge 11\,\right] \tag{1.3}$$

と同値である．条件 (1.3) が満たされるならば，σ_i^2 $(i=1,2,\ldots,k)$ の未知性にかかわらず ($\overline{y}_i$ $(i=1,2,\ldots,k)$ 単独でなく) $\widehat{\mu}_{FG}$ を用いた方が良い．一方，(1.3) が満たされないならば，標本平均 $\overline{y}_i$ $(i=1,2,\ldots,k)$ の中に $\widehat{\mu}_{FG}$ よりも分散が小さいものが存在することになる．

本報告では，Shinozaki (1978) が導出した条件（正規分布下での条件）を y_i $(i=1,2,\ldots,k)$ が楕円型分布に従う場合に一般化する．結果として，$\widehat{\mu}_{FG}$ の分散が関係式 (1.2) を満たすための条件が，y_i $(i=1,2,\ldots,k)$ がそれぞれ異なるタイプの楕円型分布に従う場合においても提示される．

2. 補題

y_i $(i=1,2,\ldots,k)$ が楕円型分布に従う場合を考え，補題を準備する．以後 y_i は $y_i=\mu\mathbf{1}_{n_i}+\varepsilon_i$, $\varepsilon_i/\sigma_i\;\sim\;\mathrm{Ell}\,(\,\mathbf{0}\,,\mathrm{I}_{n_i},\,g_i)$ $(i=1,2,\ldots,k)$ を満たすものとする．但し，記号 $\mathrm{Ell}\,(\,\mathbf{0}\,,\mathrm{I}_{n_i},\,g_i)$ は平均が $\mathbf{0}$, 分散共分散行列が I_{n_i}, 確率密度関数が g_i の楕円型分布を表すものとし，誤差ベクトル ε_i $(i=1,2,\ldots,k)$ は互いに独立と仮定する．なお，$g_i\ne g_j$ $(i\ne j)$ であれば，ε_i/σ_i と ε_j/σ_j は異なる型の楕円型分布に従うことに注意する．

補題 2.1 ： (1.1) 式で定義される $\widehat{\mu}_{FG}$ は μ の不偏推定量である．

補題 2.2 ： $\widehat{\mu}_{FG}$ の分散は次の形で表される：

$$\mathrm{V}\left(\widehat{\mu}_{FG}\right) \;=\; \mathrm{E}\left[\left\{\sum_{i=1}^{k} n_i\sigma_i^2\widehat{\sigma}_i^{\,-4}\right\}\left\{\sum_{i=1}^{k} n_i\widehat{\sigma}_i^{\,-2}\right\}^{-2}\right]$$

補題 2.3 ： 楕円型分布 $\mathrm{Ell}\,(\,\mathbf{0}\,,\mathrm{I}_n,\,g)$ に従う n 次元確率ベクトル $\mathbf{x}=(\,x_1,x_2,\ldots,x_n\,)'$ を m 次元ベクトル $\mathbf{x}_{(1)}=(x_1,\ldots,x_m)'$ と $(n-m)$ 次元ベクトル $\mathbf{x}_{(1)}=(x_{m+1},\ldots,x_n)'$ に分割し，$\mathbf{x}=(\mathbf{x}_{(1)}',\,\mathbf{x}_{(2)}')'$ と表すものとする．この時，次の等式が任意の実数 $\alpha\,(>-m/2\,)$ に対して成り立つ：

$$\mathrm{E}\left(\|\,\mathbf{x}_{(1)}\,\|^{2\alpha}\right)\times\Gamma\left(\frac{n}{2}+\alpha\right)\Gamma\left(\frac{m}{2}\right) \;=\; \mathrm{E}\left(\|\,\mathbf{x}\,\|^{2\alpha}\right)\times\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}+\alpha\right)$$

3. 結果

前節の結果を用いて，μ の実行可能な一般化最小二乗推定量 $\widehat{\mu}_{FG}$ が不等式 (1.2) を満たすための条件を導いた．以下，次の記号 $\tau_{ia}, \tau_{ib}, \tau_{ic}$ を用いる：

$$\begin{aligned}\tau_{ia} &= \frac{\mathrm{E}\left(\|\varepsilon_i/\sigma_i\|^4\right)}{(n_i+2)n_i}, \quad \tau_{ib} = (n_i-2)(n_i-4)\mathrm{E}\left(\|\varepsilon_i/\sigma_i\|^{-4}\right), \\ \tau_{ic} &= (n_i-2)\mathrm{E}\left(\|\varepsilon_i/\sigma_i\|^{-2}\right) \quad (i=1,2,\ldots,k).\end{aligned}$$

正規性の条件下では $\tau_{ia} = \tau_{ib} = \tau_{ic} = 1 \quad (i=1,2,\ldots,k)$ が成り立つため，Shinozaki (1978) が導出した条件式中には $\tau_{ia}, \tau_{ib}, \tau_{ic}$ の値が現れない．しかし，非正規性の条件下では，これらの値を考慮する必要がある．

定理 3.1 ： 不等式 (1.2) が成立するためには，次の条件 (3.1) が満たされることが必要である：

$$\frac{2\,(n_j-1)}{\tau_{ja}\,(n_j+1)} \geq \frac{\tau_{ib}\,(n_i-1)}{\tau_{ic}\,(n_i-5)} \quad (\,i \neq j\,). \tag{3.1}$$

また，条件 (3.1) と次の条件 (3.2) が同時に満たされることが，不等式 (1.2) の成立のためには十分である：

$$2\frac{n_j-1}{n_j-3}\tau_{jc} - \frac{(n_j-1)^2}{(n_j-3)(n_j-5)}\tau_{jb} \geq \frac{n_i-1}{n_i-3}\tau_{ic} - 1 \quad (\,i \neq j\,). \tag{3.2}$$

4. 適用例

定理 3.1 を Kotz タイプの楕円型分布に適用した例を示す．$k=2$ の場合を考え，$v_i = \|\varepsilon_i/\sigma_i\|^2$ $(\,i=1,\,2\,)$ とおき，ε_i/σ_i が次の確率密度関数をもつものとする：

$$g_i(v_i) = C_{n_i}\,v_i^{N_i-1}\exp(-v_i/2)\,, \tag{4.1}$$

但し，$2N_i+n_i>6$ ，$C_{n_i} = 2^{1-N_i-n_i/2}\Gamma(n_i/2)\Big/\left\{\pi^{n_i/2}\,\Gamma((2N_i+n_i-2)/2)\right\}$.

$\tau_{ia} = \frac{(n_i+2N_i)(n_i+2N_i-2)}{(n_i+2)n_i}$,$\tau_{ib} = \frac{(n_i-2)(n_i-4)}{(n_i+2N_i-4)(n_i+2N_i-6)}$,$\tau_{ic} = \frac{(n_i-2)}{(n_i+2N_i-4)}$ を (3.1) に代入して整理すると，

$$\frac{(n_j+2)\,n_j}{(n_j+2N_j)(n_j+2N_j-2)} \times \frac{2(n_j-1)}{(n_j+1)} \geq \frac{(n_i-4)}{(n_i+2N_i-6)} \times \frac{(n_i-1)}{(n_i-5)} \quad (\,i \neq j\,) \tag{4.2}$$

となり，(3.2) に代入して整理すると，

$$\frac{(n_j-1)(n_j-2)}{(n_j-3)(n_j+2N_j-4)}\left\{2-\frac{(n_j-1)(n_j-4)}{(n_j-5)(n_j+2N_j-6)}\right\} \geq \frac{(n_i-1)(n_i-2)}{(n_i-3)(n_i+2N_i-4)} - 1 \quad (\,i \neq j\,) \tag{4.3}$$

となる．(4.2)((4.3)) は，y_1, y_2 が Kotz タイプの確率密度関数 (4.1) を持つ場合に不等式 (1.2) が成り立つための必要条件 (十分条件) である．

なお，Kotz タイプの確率密度 (4.1) において $N_1=N_2=1$ とすれば，y_1, y_2 に正規性を仮定することになる．この場合，(4.2) は正規性の下で (1.2) が成り立つための必要十分条件 (1.3) と同値である．

参考文献

Fang, K.-T., Kotz, S. and Ng, K.W. (1989). *Symmetric Multivariate and Related Distributions*, Chapman and Hall, New York.

Shinozaki, N. (1978). A note on estimating the common mean of k normal distributions and the Stein problem. *Comm. Statist. Theory Methods* A7(15), 1421–1432.

On parameter estimation based on the contact distance for certain superposed Neyman-Scott cluster fields

統計数理研究所 田中 潮

1　背景

モデルを特徴付けるパラメータ (intensity) が異なる m 種類のクラスターから成る一般化 Neyman-Scott 型クラスター点過程のパラメータ推定につき考察する.

Neyman-Scott クラスター点過程モデルは銀河の点配置を基に, [2] により定義されたクラスター点過程モデルである. モデルの解析法を歴史的にみると, パラメータ推定に関しては最大尤度法が採られていない. それは, Neyman-Scott クラスター点過程の構造が複雑であり, 尤度関数があらわすことができないことに尽きる.

しかし, 近年, Neyman-Scott クラスター点過程の最大尤度解析法によるパラメータ推定が, Palm 型強度関数に基づいた最大 Palm 型尤度法により与えられている ([5]). 直観的に Palm 型強度関数は, クラスター点配置の各点の座標そのものではなく, 相異なる点間の距離内の点の intensity をあらわすと考えられる. 最大尤度解析法によるパラメータ推定のメリットとして, 例えば AIC 等により陽に Neyman-Scott クラスター点過程モデルの比較が可能になることが挙げられる.

尚, 最大 Palm 型尤度法では, 混在するクラスター (intensity が異なる) の種類につき, 2 種類のみが扱われ, 実際, それ以上の混在するクラスターに対しては, 本手法では技術的に困難であることが本研究に於いて注意されている ([3]). 従って, モデル比較をする際, 最大 Palm 型尤度法とは異なる手法が要求される.

ところで, これらの解析法は, 様々なデータ等の解析に向けて [4] により汎用化されている.

2　主結果

さて, 本研究の概要を述べる ([3]).

m 種類のクラスター X_i が混在するクラスター点過程 X につき考察する, ここに,

$$X := \bigcup_{i=1}^{m} X_i.$$

但し, 点同士の重なりはないものとする.

次のパラメータを与える:

$$\sum_{i=1}^{m} a_i = 1, \quad 0 < a_i < 1,$$

ここに, a_i は X の intensity に対する X_i のそれの割合とする.

また, 最大 Palm 型尤度法と異なるパラメータ推定の指標として確率幾何, 空間点過程論等で重要な球状接触距離関数, 最近接距離関数, 及び J-関数の性質を用いる. 最近接距離関数に関する intensity 関数に基づく a_i の最大 NND 尤度法を考える. ここに, 球状接触関数 F, 最近接距離関数 G 及び J-関数は次で定義される ([1]):

$$F(r) := \mathrm{Prob}\{\mathrm{dist}(\psi, \mathbf{X}) \le r\}, \quad \forall r \ge 0,$$

$$G(r) := \mathrm{Prob}[\mathrm{dist}(\psi, \mathbf{X}\backslash\psi) \leq r|\ \{\psi\}], \quad \forall r \geq 0,$$

$$J(r) := \frac{1 - G(r)}{1 - F(r)}, \quad i = 1, \ldots, m, \quad F < 1,$$

ここに, $\mathrm{dist}(\psi, \mathbf{X})$ は任意にとった location ψ からクラスター $\mathbf{X}$ の任意の点までの最短距離をあらわす.

尚, 最近接距離関数に関する intensity 関数は, 等方非一様 Poisson 点過程の intensity 関数とみなす.

主結果として, NND 対数尤度関数が単峰になる為の十分条件を最近接距離関数, 球状接触距離関数 及び J-関数を用いて与える. 応用例として, 特に NND 対数尤度関数がすべての a_i に関して単峰であるときに, 混在する Neyman-Scott クラスター点過程モデルのパラメータが全て一意的に決まることを示す.

参考文献

[1] Baddeley, A., Barany, I., Schneider, R. and Weil, W. (2007). *Stochastic Geometry Lecture Notes in Mathematics 1892*, Springer-Verlag Berlin, Heidelberg.

[2] Neyman, J. and Scott, E. (1958). Statistical approach to problems of cosmology. *J. Roy. Statist. Soc. Ser. B* **20,** 1–43.

[3] Tanaka, U. (2009). On parameter estimation based on the contact distance for certain superposed Neyman-Scott cluster fields. *Research Memorandum.* No. 1104, The Institute of Statistical Mathematics, Tokyo.

[4] Tanaka, U., Ogata, Y. and Katsura, K. (2008). Simulation and Estimation of the Neyman-Scott Type Spatial Cluster Models, *Computer Science Monographs* **34**.

[5] Tanaka, U., Ogata, Y. and Stoyan, D. (2008). Parameter estimation and model selection for Neyman-Scott point processes, *Biom. J* **50,** 43–57.

定常増分過程の WAVELET 係数の局在性と応用

川崎 秀二

1. WAVELET 係数ドメインでの推定問題

確率過程の wavelet 係数は、シフト j およびスケール k の 2 パラメータを持ち、各 j で k の列としての時系列と見做す事ができる．我々は先行研究 [1] において、wavelet 係数ドメインにおけるパラメータ推定の一般形について論じるとともに、シフト k に関する局在性を定式化した．

本研究では、スケール j に関する局在性を定式化する．そこにおいて、クロス・スケール共分散の減衰度を支配する二つの項を与える．それらはともに wavelet の モーメント消滅次数およびスケール・ラグの関数である．この局在性を応用して fractional Brown 運動 (FBM) の Hurst 指数推定に関する中心極限定理 (CLT) の漸近分散の有用な評価を得る．

この推定量は各スケール $j = 1, \cdots, J$ に対する推定量の線形形式であり、推定量の漸近分散は j に関する二次形式となるが、漸近分散への寄与は殆ど対角成分（$j = j'$ の項）からであり、非対角成分（$j \neq j'$ の項）は小さい事が分かる．従って、漸近分散の J に関する挙動は対角成分の挙動のみを用いて十分正確に評価できることとなる．これがスケール局在性の一つの形であると考えられる．特に、最小分散の意味で最適なスケール上限 J を、対角成分の評価から求める事が出来る．

本論文を通して $X^T = \{X_t;\ 0 \leq t \leq T\}$ は平均 0, 分散有限の定常増分かつ H-ss な確率過程とする $(0 < H < 1)$. ψ を、実数値 mother wavelet で次の仮定を満たすものとする．

(ψ1) ψ はコンパクト台を $W = [0, w]$, $w \geq 1$ に持ち、有界関数とする．
(ψ2) ψ はある $\gamma_0 \in \mathbb{N}$ に対し γ_0-次のモーメント消滅性を持つとする：

wavelet 関数族を $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$ により取る．j, k は適当な範囲の整数を取るものとする．このとき、$\psi_{j,k}(t)$ の台は $\mathrm{supp}(\psi_{j,k}) = [2^j k,\ 2^j(k + w)]$ となる．仮定 (ψ1) は k の範囲に制限を課す事になるが、この点については後で触れる．

$N_{T,j} \in \mathbb{N}$ を、各スケール j で、時刻 $t = T > 0$ までに取得可能な wavelet 係数の個数とする：$N_{T,j} = \max\{k;\ 2^j(w + k) \leq T\} = \lfloor 2^{-j}T - w \rfloor$.

サンプルパスの観測長 $T > 0$ に対し、X^T の wavelet 係数を各 $j = J_0 + 1, \cdots, J$ $(J_0 + 1 \leq J)$ で $s_j(k) = \int_{2^j k}^{2^j(w+k)} \psi_{j,k}(t)\, X(t)\, dt,$, $k = 1, \cdots, N_{j,T}$ により定義する．また、$S^T = \{\boldsymbol{s}_j^T;\ j = J_0 + 1, \cdots, J\}$, $\boldsymbol{s}_j^T = \{s_j(k);\ k = 1, \cdots, N_{T,j}\}$ とおく．

X^T が定常増分および平均 0 であれば、各 j で $\boldsymbol{s}_j^T$ は定常列であり平均 0 である事に注意する．また一般に、定常増分過程のクラスは定常過程のクラスを含んでいる．

本論文では、X^T は FBMX_H^T, $0 < H < 1$ とする．良く知られているように、FBM は定常増分および自己相似性を持つ Gauss 過程である．$\boldsymbol{s}_j^T$ の定常性を用いて上述の CLT を得ることになる．従って、$X^T = X_H^T$ の定常増分性は本質的な仮定である．一方、自己相似性および Gauss 性は簡単化のために用いているが、共分散の定式化がある程度できるクラスであれば必ず必要な仮定という訳ではないと考えられる．

さて、X^T のパラメータ $\xi \in \mathbb{R}$ を推定するにおいて、ξ は wavelet 係数ドメインで次のように 書けるものとする：与えられた $f : \mathbb{R}^{\mathcal{J}} \mapsto \mathbb{R}$, $\mathcal{J} = J - J_0$ に対し $\xi = f(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_{J_0+1}, \cdots, \theta_J) \in \mathbb{R}^{\mathcal{J}}$, ただし $\theta_j = \mathbb{E}[\,g(\mathsf{s}_j(1))\,]$, $g : \mathbb{R} \mapsto \mathbb{R}$ はある関数であり、$\mathsf{s}_j(k)$ は正規化された wavelet 係数である：$\mathsf{s}_j(k) = \frac{s_j(k)}{\sigma_j}$, $\sigma_j^2 = \mathrm{Var}[s_j(1)]$. また、$\xi$ の推定量が wavelet 係数ドメインでの推定量 [1] として $\hat{\xi}_T \triangleq f(\hat{\boldsymbol{\theta}}_T)$ と書けることを仮定する．ここで θ_j, $j = J_0 + 1, \cdots, J$ は上述の関数 g による、スケール j での wavelet 係数の汎関数期待値である：$\theta_j = \mathbb{E}[\,g(\mathsf{s}_j(1))\,]$.

$\boldsymbol{\theta}$ の推定量 $\hat{\boldsymbol{\theta}}_T = (\hat{\theta}_{J_0+1,T}, \cdots, \hat{\theta}_{J,T}) \in \mathbb{R}^{\mathcal{J}}$ としては、$\hat{\theta}_{j,T} = N_{j,T}^{-1} \sum_{n=1}^{N_{j,T}} Y_j(n)$, $j = J_0 + 1, \cdots, J$ をとる事にする．ただし、$Y_j(n) = d_j^{-1} \sum_{k=1}^{d_j} g(\mathsf{s}_j(d_j n + k))$, $d_j = 2^{J-j}$.

よく知られているように、$\boldsymbol{s}_j^T$ は各 j でエルゴード定常列である．従って、$\boldsymbol{Y}(n) = (Y_{J_0+1}(n), \cdots, Y_J(n)) \in \mathbb{R}^{\mathcal{J}}$ はエルゴード定常ベクトルである．故に、$\boldsymbol{Y}(n) = (Y_{J_0+1}(n), \cdots, Y_J(n)) \in \mathbb{R}^{\mathcal{J}}$ は $\boldsymbol{\theta}$: $\hat{\boldsymbol{\theta}}_T \to \boldsymbol{\theta}$ の普偏一致推定量である： $\hat{\boldsymbol{\theta}}_T \to \boldsymbol{\theta}$ a.s. as $T \to \infty$ and $\mathbb{E}[\hat{\boldsymbol{\theta}}_T] = \boldsymbol{\theta}$.

2. 主結果

2つのスケール j, j', $J_0+1 \le j \le j' \le J$ をとり、$j' = j+m$, $m = 0, 1, \cdots, J-j$ と書くことにする．m はスケールのラグに相当する．X^T の wavelet 係数の共分散は

$$\text{(1)}\quad \begin{aligned} \operatorname{Cov}[\, s_j(k_1),\ s_{j'}(k_2)\,] &= 2^{(2H+1)j+(m/2)}\Big(-\frac{1}{2}\Big)\iint_{W^2}\psi(s)\psi(t)\left|s-2^m t+k_1-2^m k_2\right|^{2H} dsdt \\ &= 2^{(2H+1)j}\cdot\operatorname{Cov}[\, s_0(k_1),\ s_m(k_2)\,] \triangleq 2^{(2H+1)j}\, r(m,n) \end{aligned}$$

ただし $n = k_1 - 2^m k_2 \in \mathbb{Z}$, と書ける．特に $j = j'$ とすれば $\operatorname{Cov}[\, s_j(k_1),\ s_j(k_2)\,] = 2^{(2H+1)j}\cdot\operatorname{Cov}[\, s_0(k_1),\ s_0(k_2)\,]$ である．さらに $k_1 = k_2$ とすれば $\sigma_j^2 \triangleq \operatorname{Var}[s_j(0)] = 2^{(2H+1)j}\cdot\operatorname{Var}[s_0(0)] = 2^{(2H+1)j}\sigma_0^2$ となる．

上述の関数 $f : \mathbb{R}^{\mathcal{J}} \mapsto \mathbb{R}$ として、線形形式 $f(\boldsymbol{\theta}) = \sum_{j=J_0+1}^{J} a_j\, \varphi_j(\theta_j)$, $\{a_j\} \subset \mathbb{R}$, $\varphi_j : \mathbb{R} \mapsto \mathbb{R}$. をとる事にする．先ず次の中心極限定理に注意する．

Proposition 1 ([1] からの修正版). $(\gamma_0 - H)p > 1$ とする．このとき、$\mathcal{J}$-次元 CLT $\sqrt{T}\,[\, \hat{\boldsymbol{\theta}}'_T - \boldsymbol{\theta}'\,] \Longrightarrow \mathcal{N}(0, \Sigma)$ および 1 次元-$CLT\sqrt{T}\,\left[\, f(\hat{\boldsymbol{\theta}}'_T) - f(\boldsymbol{\theta}')\,\right] \Longrightarrow \mathcal{N}(0, v_{H,J}^2)$ がそれぞれ成り立つ．漸近共分散行列 $\Sigma = (\Sigma_{j,j'})$ および 漸近分散 $v_{H,J}^2$ はそれぞれ

$$\text{(2)}\qquad \Sigma_{j,j+m} = 2^{-m/2} R_H(m) \quad \textit{with } R_H(m) = \rho^l(m) + 2\sum_{k\in\mathbb{N}} \rho^l(m,k),$$

$v_{H,J}^2 = \xi_J R_H + 2\sum_{m=1}^{\mathcal{J}-1} \xi_J(m)\, R_H(m)$ により与えられる．ただし、$\xi_J(m) \triangleq \sum_{j=J_0+1}^{J-m} 2^j a_j a_{j+m}\, [\theta_j \dot{\varphi}_j(\theta_j)]\, [\theta_{j+m}\dot{\varphi}_j(\theta_{j+m})]$. ここで、$\xi_J \triangleq \xi_J(0) = \sum_{j=J_0+1}^{J} 2^j a_j^2\, \left[\,\theta_j\, \dot{\varphi}_j(\theta_j)\,\right]^2$ および $R_H \triangleq R_H(0)$ とおいた．

本論文における j-局在性の意義は、wavelet 係数共分散の、各 $m = 1, 2, \cdots$ に対する減衰度の評価にある．これまでは $n = k_1 - 2^m k_2 \to \infty$ の漸近的な評価のみが知られていたが、[1] での最尤推定量と比較して漸近分散がさほど遜色がない事を示す応用において、また、以下のように、Hurst 指数推定に用いるスケールの個数を最小分散の意味で最適化する応用においては、実際に漸近的評価のみでは足りず、「各点」評価が必要である．

Theorem 2. ψ を MRA から生成された $wavelet$ とする．$0 < H < 1$, $m = 1, 2$, $\gamma \in \mathbb{N}$ に対し $\gamma \in \mathbb{N}$,

$$\text{(3)}\qquad 2^{-m/2}\,\frac{r(m,n)}{\sigma_0^2} \le \Psi_{\gamma,H}(m)\, K_\gamma(m,n), \qquad n \in \mathbb{N}_0,$$

が成り立つ．ただし、$K_\gamma(m,n)$ と $\Psi_{\gamma,H}$ は次のように与えられる：

(i) $m = 1$ については $K_\gamma^2(1,n) \triangleq \sum_{k\in\mathbb{N}} \left\{1 + \left(n + 2^m k\right)^2\right\}^{-2\gamma}\Big|_{m=1}$, であり、$\Psi_{\gamma,H}(1) \triangleq A_\gamma(1) + 2^{2H}\left\{\frac{5}{(1+2\gamma)^2}\right\}^{\gamma} B_\gamma(1)$; ここで $A_\gamma^2(1) = \pi^{-1}\int_{(\pi,\,2\pi]}\left|\frac{\hat{\psi}(2\lambda)}{\hat{\psi}(\lambda)}\right|^2 d\lambda$ and $B_\gamma^2(1) = \pi^{-1}\int_{(0,\,\pi]}\left|\frac{\hat{\psi}(\lambda)}{\hat{\psi}(2\lambda)}\right|^2 d\lambda$ である．

Theorem 3. $l \ge p$, $m = 1, 2$ に対し、次が成り立つ：

$$\text{(4)}\qquad R_H(m) \le \frac{\left[\,\Psi_{\gamma,H}(m)\, K_\gamma(m)\,\right]^l}{C_*^{(l)}}\cdot R_H\,,$$

ただし $C_*^{(l)} = (z_l - 1)/z_l$ であり、$z_l = \sup\left\{z \;\middle|\; \sum_{k\in\mathbb{N}}\left[\,\rho_k\,\right]^{2l} \le \frac{1}{z^2}\right\}$ である．ここに $K_\gamma^l(m) \triangleq K_\gamma^l(m,0) + 2\sum_{n\in\mathbb{N}} K_\gamma^l(m,n)$. である．

これより、漸近分散 $v_{H,J}^2$ はその「対角」成分 R_H のみを用いて十分精確に評価されることになる．実際、「非対角」成分 $R_H(m)$, $m = 1, 2, \cdots$ は小さい事が分かる．

References

[1] S. Albeverio and S. Kawasaki, *On a Localization Property of Wavelet Coefficients of Processes with Stationary-Increments and Applications. I: Localization with Respect to Shift*, in *Adv. Appl. Prob.*, **37(4)**, 938-962, 2005.

[2] J.-M. Bardet, G. Lang, E. Moulines and P. Soulier (2000). Wavelet Estimator of Long-Range Dependent Processes, *Stat. Infer. Stochast. Process.*, **3**, 85-99.

球面対称分布の下での適応的リッジ推定量について

森　一将　　　　倉田　博史
東京大学大学院総合文化研究科　東京大学大学院 情報学環・学際情報学府

1. はじめに

$\mathbf{y}=\mathbf{A\beta}+\mathbf{e}$， $\mathrm{E}(\mathbf{e})=\mathbf{0},\mathrm{V}(\mathbf{e})=\sigma^2\,\mathbf{I}$ を満たす回帰モデルを考える。ここで $\mathbf{y}$ は $n\times 1$ の観測ベクトル、$\mathbf{A}$ は $n\times p$ かつランク p の計画行列である。また、$\boldsymbol{\beta}$ は $p\times 1$ で真値が未知の回帰係数、$\mathbf{e}$ は $n\times 1$ の誤差ベクトルである。回帰係数 $\boldsymbol{\beta}$ の推定には最小二乗推定(LSE) $\hat{\boldsymbol{\beta}}=(\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\mathbf{y}$ が用いられるが、この推定量の分散は $\mathrm{V}(\hat{\boldsymbol{\beta}})=\sigma^2(\mathbf{A}^t\mathbf{A})^{-1}$ であるため、$\mathbf{A}$ の各ベクトルの相関が高くなり、$\mathbf{A}^t\mathbf{A}$ が特異に近くなると $\hat{\boldsymbol{\beta}}$ の分散は大きくなり、推定値は安定しない。また、 $p\geq 3$ のとき LSE は許容的でないことが知られている。

この問題に対して、Hoerl and Kennard (1970)は、$\hat{\boldsymbol{\beta}}(k)=(\mathbf{A}^t\mathbf{A}+k\mathbf{I})^{-1}\mathbf{A}^t\mathbf{y}$ なる形の推定量とその一般化

$$\hat{\boldsymbol{\beta}}(\mathbf{K})=\Phi^t(\Lambda+\mathbf{K})^{-1}\Phi\mathbf{A}^t\mathbf{y} \tag{1}$$

を提案した。これを一般化リッジ推定量という。ここで $\mathbf{K}=\mathrm{diag}(k_1,\cdots k_p)$ は適切に選択された定数、$\boldsymbol{\Lambda}=diag(\lambda_1\cdots\lambda_{\mathrm{p}})$, $\lambda_1\geq\lambda_2\cdots\geq\lambda_{\mathrm{p}}$ と $\boldsymbol{\Phi}=(\boldsymbol{\varphi}_{1,}\cdots\boldsymbol{\varphi}_p)$ はそれぞれ $\mathbf{A}^t\mathbf{A}$ の固有値と固有ベクトルである。(1)の推定量は $k_i^*=\sigma^2/\beta_i^2$ のときに最小二乗誤差を最小化することができるが、未知母数が含まれるため、通常は何らかの別の推定量や統計量で k_i が代替される。

代替されるものの一例として Vinod & Ullah (1980)が提案した適応的推定量の改良(Wang & Chow ,1990)を考える。

$$\hat{k}_{i(l_1,l_2)}=l_1\hat{\sigma}^2\Big/(\hat{\boldsymbol{\beta}}^t\hat{\boldsymbol{\beta}}-l_2\hat{\sigma}^2/\lambda_i)\,. \tag{2}$$

ここで l_1 と l_2 は定数であり、 $\hat{\sigma}^2=\|\mathbf{y}-\mathbf{A}\hat{\boldsymbol{\beta}}\|/m$ 、 $m=n-p$ である。彼らは MSE、誤差の正規性の下で(2)の推定量が LSE を改善することを示している。その一方で、誤差をより一般的な球面対称分布に仮定した場合の同様の結果はまだ知られてない。他の一般化リッジ推定量についても、球面対称分布下での改善についてはわずかに Maruyama&Strawderman(2005)が知られている限りである。本研究では、(2)の推定量を $\sigma^{-n}f(\mathbf{e}^t\mathbf{e}/\sigma^2)$ を密度関数に持つ球面対称分布の下で考え、LSE を改善する条件を導

出した。

2. 主な結果

得られた定理は以下のとおり。

【定理】

リスクを損失関数 $\mathrm{L}_j(\mathbf{b},\boldsymbol{\beta},\sigma^2)=(\mathbf{b}-\boldsymbol{\beta})^t(\mathbf{A}^t\mathbf{A})^j(\mathbf{b}-\boldsymbol{\beta})/\sigma^2$ を用いて $\mathrm{R}_j(\mathbf{b},\boldsymbol{\beta},\sigma^2)=\mathrm{E}[\mathrm{L}_j(\mathbf{b},\boldsymbol{\beta},\sigma^2)]$ と定義する。また回帰モデルの誤差項は球面対称分布に従うものとする。このとき(2)による適応的リッジ推定量を考える。

すると、l_1,l_2 が以下の条件を満たすとき

(i) $0<l_1<\dfrac{2(n-p)}{n-p+2}\left(\dfrac{\lambda_p^2}{\lambda_1^j}\sum_{i=1}^{p}\lambda_i^{\,j-2}-2\right)$

(ii) $l_1>l_2$

$\boldsymbol{\beta}\in\mathbf{R}^p$ かつ $\sigma^2>0$ に対して(2)による推定量は LSE を改善する。

3. シミュレーション

p=6、n=20、$\mathbf{A}^t\mathbf{A}$ の固有値 $\Lambda=\mathrm{diag}(1,0.01,0.0019,0.0017,0.0015,0.001)$、誤差項 $\mathbf{N}_6(0,\mathbf{I})$ および $\mathrm{t}_6(10)$ で回帰係数を推定した。推定方法は LSE と(2)による適応的リッジ推定量の 2 種類で、相対的効率 $\mathrm{R}_j(\hat{\boldsymbol{\beta}}(\mathbf{K}),\boldsymbol{\beta},\sigma^2)/\mathrm{R}_j(\hat{\boldsymbol{\beta}},\boldsymbol{\beta},\sigma^2)$ で評価を行った。結果の一部は以下のとおり。

(表 1)多変量正規誤差の場合の評価

percentile of l_1	beta 2.0, l_2 = 0	−1	−2
	risk	risk	risk
25%	0.965	0.979	0.984
50%	0.934	0.958	0.968
75%	0.906	0.938	0.952
100%	0.880	0.920	0.937

(表 2)多変量 t 分布誤差の場合の評価

percentile of l_1	beta 2.0, l_2 = 0	−1	−2
	risk	risk	risk
25%	0.965	0.979	0.984
50%	0.934	0.958	0.968
75%	0.906	0.938	0.952
100%	0.880	0.920	0.937

(2)による適応的リッジ推定量は多変量 t 分布誤差の場合にも LSE を改善していること、定数 l_1,l_2 は定理が保証する範囲で大きくとることにより改善の度合いが強くなることが示された。

参考文献

Hoerl A.E. and Kannard, R.W. (1970a), *Technometrics*, Vol 12,55-67.
Hoerl A.E. and Kannard, R.W. (1970b), *Technometrics*, Vol 12, 69-82.
Marymama,Y. and Strawderman, W.E. (2005), Annals of Statistics, Vol 33,1753-1770.
Vinod, H.D., and Ullah A.,(1980), *Recent Advances in Regression Models*
Wang S.G. and Chow, S.C. (1990), *Statistics and Probability Letters,* Vol 10,17-21.

強い識別条件での混合分布のギブスサンプラーの非正則なふるまい

東京大学 大学院数理科学研究科　　鎌谷研吾

序

本講演では，混合分布モデルについて，ギブスサンプラーの幾つかの非正則なふるまいを研究する．本講演での目的は，ギブスサンプラーの非正則なふるまいの例を紹介し，その対処法について検討する事である．[Kamatani(2009)] は正則なギブスサンプラーの漸近的な収束等の振る舞いを調べた．とくに正則なパラメトリック族については，ギブスサンプラーの振る舞いは AR 過程の分布に収束する事を示した．しかし，非正則なパラメトリック族では必ずしも AR 過程の分布には収束せず，以下の三つの場合がある．a) ギブスサンプラーの振る舞いはタイトネスを持ち，極限はエルゴード性を持つが，AR 過程ではない，b) タイトネスを持ち，極限はエルゴード性を持たない，c) タイトネスを持たない．強い識別条件のもとの混合分布モデルであっても，真値がパラメータ空間の境界上にある場合，ギブスサンプラーは正則なパラメトリック族と異なるふるまいをする事がある．識別条件の強さによって a) および b) の二つの場合が現れる．本講演ではこれらの収束，および b) の場合の対処法についての検討を行う．

1　ギブスサンプラーの空間と距離

本項目では簡単にギブスサンプラーの動く空間を考える．まず (S, d) は完備可分の距離空間で，$d(a, b) \leq 1$ となるとする．可算個の S の直積を S^∞ と書き，その上での距離 d_∞ を

$$d_\infty(s, t) = \sum_{i=1}^{\infty} 2^{-i} d(s(i), t(i)) \tag{1}$$

とさだめる．ただし $s = (s(1), s(2), \ldots)$ および $t = (t(1), t(2), \ldots)$ は S^∞ の元である．完備可分な距離空間 (S^∞, d_∞) 上の確率分布全体の集合を $\mathcal{P}(S^\infty)$ と書き，その上での Prohorov 距離を ρ と書く事にする．本論文で考える確率測度は，完備可分な距離空間 $(\mathcal{P}(S^\infty), \rho)$ の上の確率測度であり，その全体を $\mathcal{P}^*(S^\infty)$ と書き，またその上の Prohorov 距離を ρ^* とかく事にする．すると再び $(\mathcal{P}^*(S^\infty), \rho^*)$ 完備可分な距離空間である．ギブスサンプラーの振る舞いは $(\mathcal{P}^*(S^\infty), \rho^*)$ の元として表現できる．

Example 1.1 ギブスサンプラーのふるまいは $\mathcal{P}^*(\Theta^\infty)$ の元で表される．ただし，Θ はパラメータ空間である．ここで X および Y は状態空間で，X は観測の空間，Y はギブスサンプラーのための人工的な空間とする．$\Theta \times X \times Y$ 上の σ-有限な測度 p は分解

$$p(d\theta, dx, dy) = p_{X,Y}(dx, dy) p_{\Theta|X,Y}(d\theta|x, y) = p_{X,\Theta}(dx, d\theta) p_{Y|X,\Theta}(dy|x, \theta) \tag{2}$$

をもつとする．ただし，$p_{\Theta|X,Y}$ および $p_{Y|X,\Theta}$ は確率トランジションカーネルである．与えられた観測 $x \in X$ のもと，状態空間が $\Theta \times Y$ であるようなマルコフチェインが，ルール $y(i) \sim p_{Y|X,\Theta}(dy|x, \theta(i))$ および $\theta(i+1) \sim p_{\Theta|X,Y}(d\theta|x, y(i))$ かつ出発点 $\theta(1) \sim \nu_{\Theta|X}(d\theta|x)$ のもと生成されるとする．すると Θ へのマージナル $\theta(1), \theta(2), \ldots$ もやはりマルコフチェインで，その分布を，x に依存する事を明示して，$\omega(x)$ と書く．これは $\mathcal{P}(\Theta^\infty)$ の元である．さらに，$\omega(x)$ の法則は $\mathcal{P}^*(\Theta^\infty)$ の元である．

2 非正則なギブスサンプラーのふるまい

非正則なギブスサンプラーの一例を見る．まず $\Theta = [0,1]$, $X = \mathbf{R}^{\mathbf{d}}$, $Y = \{0,1\}$ とし，$\mathcal{X}$ は X のボレル σ-代数で，$\mathcal{Y} = 2^{\{0,1\}}$．また p_Θ は $\mathcal{P}(\Theta)$ の元であり，$F_0, F_1 \in \mathcal{P}(X)$ を固定すると $p_{X|Y,\Theta}(\cdot|y,\theta) = F_y$ $(y=0,1)$ および $p_{Y|\Theta}(\{y\}|\theta) = (1-\theta)^{1-y}\theta^y$ であるものとする．独立同分布を考え，$p^n \in \mathcal{P}(\Theta \times X_n \times Y_n)$ を次のように定めるが，ここで X_n および Y_n は X および Y の n 回コピーである．

$$p^n(d\theta, dx_n, dy_n) = p^n_{X_n|Y_n,\Theta}(dx_n|y_n,\theta) p^n_{Y_n|\Theta}(dy_n|\theta) p_\Theta(d\theta) \tag{3}$$

ただし $x_n = (x^1, \ldots, x^n)$，$y_n = (y^1, \ldots, y^n)$ であり，$p^n_{X_n|Y_n,\Theta}(dx_n|y_n,\theta) = \prod_{i=1}^n p_{X|Y,\Theta}(dx^i|y^i,\theta)$, $p^n_{Y_n|\Theta}(dy^n|\theta) = \prod_{i=1}^n p_{Y|\Theta}(dy^i|\theta)$ である．

Proposition 2.1 真値が $\theta = 0$ のとき，$p^n \in \mathcal{P}(\Theta \times X_n \times Y_n)$ から生成されたギブスサンプラーは，以下の条件のもと，適当なスケーリングのもと，*Poisson-Gamma* ギブスサンプラーに収束する．

1. きわめて強い識別条件を持つ：$\alpha = F_1((\mathrm{suppF_0})^{\mathrm{c}}) > 0$.
2. 事前分布 p_Θ は正則である．
3. ある $\epsilon > 0$ があって，

$$0 < \int_X (f_1(x)/f_0(x))^{1+\epsilon} F_0(dx) < \infty \tag{4}$$

となる．ただし f_i は F_i の，$F_1 + F_2$ に対する微分．

ここで

$$I := \int (f_1(x)/f_0(x) - 1)^2 F_0(dx), \tag{5}$$

とする．

Proposition 2.2 真値が $\theta = 0$ のとき，$p^n \in \mathcal{P}(\Theta \times X_n \times Y_n)$ から生成されたギブスサンプラーは，以下の条件のもと，一点にとどまり，動かないギブスサンプラーに収束する．

1. 強い識別条件を持つ: $\alpha = F_1((\mathrm{suppF_0})^{\mathrm{c}}) = 0$ かつ $I \in (0,\infty)$.
2. 事前分布 p_Θ は正則である．

参考文献

[Kamatani(2009)] Kengo Kamatani. On Some Asymptotic Properties of the Gibbs Sampler (submitted). 2009.

セミパラメトリックモデルの経験確率過程による解析

久保木 久孝（電通大）　鈴木 武（早稲田大）

1. 序

セミパラメトリックモデル $\mathscr{P} = \{P_{\theta,\eta} : \theta \in \Theta,\ \eta \in H\}$ （θ は $\mathbb{R}^k$ の部分集合、H は無限次元空間）が与えられたとき，母数 θ, あるいは η の推定問題を考える．母数の推測においてスコア，有効スコア，有効推定方程式，接空間が重要な働きをする．さらに推定量の漸近的性質を議論する際に，経験確率過程の理論が非常に有効であることが近年認識されてきている．このような観点に立って，本報告では特に線形モデルに限定して，母数の推定量の漸近的性質，有効推定量の構成に関する研究の現状を紹介する．

2. 線形回帰モデル

a) 残差と共変量が独立でない場合：回帰モデル $Y = \beta' Z + e,\ \beta \in \mathbb{R}^k,\ Z \in \mathbb{R}^k$ を考える．この場合 β の有効スコアは $\tilde{\ell}_{\beta,\eta}(Y,Z) = (I - \Pi_{\beta,\eta})\dot{\ell}_{\beta,\eta} = \dfrac{Z(Y-\beta' Z)}{P_{\beta,\eta}(e^2|Z)}$ となる．上式の右辺の分母は一般に Z に関係し，したがって $\tilde{\ell}_{\beta,\eta}(Y,Z)$ は $Z(Y-\beta' Z)$ とは比例しなくなる．このことからも，サンプル $\{(Y_i, Z_i), i = 1,\ldots n\}$ にもとづく β の最小２乗推定量 $\hat{\beta} = (\sum_{i=1}^n Z_i Z_i')^{-1}(\sum_{i=1}^n Z_i' Y_i)$ は β の有効推定量に必ずしもならないことが予想される．

b) 残差と共変量が独立の場合：この場合は β の有効スコアは $\tilde{\ell}_{\beta,\eta} = -\left(\dfrac{\dot{\eta}}{\eta}\right)(e)Z - h(e) = -\left(\dfrac{\dot{\eta}}{\eta}\right)(e)(Z-\mu) + \dfrac{e\mu}{\sigma^2}$ となる．さらに β の有効情報量は $\tilde{I}_{\beta,\eta} = P_{\beta,\eta}[\tilde{\ell}_{\beta,\eta}\tilde{\ell}'_{\beta,\eta}] = P_{\beta,\eta}\left[\left(\dfrac{\dot{\eta}}{\eta}\right)^2(e)\right]\cdot P\{(Z-\mu)(Z-\mu)'\} + \dfrac{\mu\mu'}{\sigma^2}$ である．β の最小２乗推定量 $\hat{\beta}$ の漸近分散は $\sigma^2\{P_{\beta,\eta}(ZZ')\}^{-1}$ であるから $[\sqrt{n}(\hat{\beta}-\beta)$ の漸近分散$] \geq \tilde{I}_{\beta,\eta}^{-1} \geq I_{\beta,\eta}^{-1}$ なる不等式が成立する．これから，$\hat{\beta}$ は必ずしも β の有効推定量とならなく，その改良の存在が示唆される．残差 e が正規分布 $\mathrm{N}(0,\sigma^2)$ にしたがうときは $\hat{\beta}$ の漸近分散は $\tilde{I}_{\beta,\eta}^{-1}$ に等しくなり，$\hat{\beta}$ は β の有効推定量となることがわかる．$\hat{F}$ を，残差 $\{Y_1 - \hat{\beta}' Z_1, \ldots, Y_n - \hat{\beta}' Z_n\}$ にもとづく経験推定量 $\hat{F}(t) = \dfrac{1}{n}\sum_{i=1}^n 1\{Y_i - \hat{\beta}' Z_i \leq t\}$ とする．このとき確率過程として $\sqrt{n}(\hat{F} - F)$ はタイトな平均 0 のガウス過程に弱収束することが示される．このことは，関数族 $\{1\{Y - b' Z \leq v\} : b \in \mathbb{R}^k, v \in \mathbb{R}\}$ が VC-クラス（したがって Donsker クラス）となっていることを用いて示される．

3. 回帰分析に関する有効推定方程式

回帰モデル $Y = g_\theta(Z) + e;\ E(e|Z) = 0,\ \theta \in \mathbb{R}^k$ を考える．共変量 Z と誤差 e は必ずしも独立ではないとする．このとき θ の有効スコアは $\tilde{\ell}_{\theta,\eta}(y,z) = \dfrac{\dot{g}_\theta(z)(y - g_\theta(z))}{V(z)}$ で与えられる．ここで，$V(z) = \mathrm{Var}(e|Z = z) = E(e^2|Z = z)$ である．有効スコアー関数は θ の有効推定量を求めるための推定方程式を導くのに使われる．しかし一般に，$V(z)$ は η に依存する未知の関数である．これを $\hat{V}(z)$ で推定するとき，θ の推定方程式 $\hat{S}_n(\theta) = \mathbb{P}_n\left[\dfrac{\dot{g}_\theta(Z)(Y - g_\theta(Z))}{\hat{V}(Z)}\right] = \dfrac{1}{n}\sum_{i=1}^n \dfrac{\dot{g}_\theta(Z_i)(Y_i - g_\theta(Z_i))}{\hat{V}(Z_i)}$ が得られる．もし，$\hat{V}$ が $\tilde{V} \not\equiv V$ に確率収束するなら，これは漸近的に推定方程式 $\tilde{S}_n(\theta) = \mathbb{P}_n\left[\dfrac{\dot{g}_\theta(Z)(Y - g_\theta(Z))}{\tilde{V}(Z)}\right]$ と同等になり，$\hat{S}_n(\theta) = 0$ の解と $\tilde{S}_n(\theta) = 0$ の解は，同じ漸近分散をもつオーダー $\sqrt{n}$ の一致推定量となることが期待できる．推定方程式 $\hat{S}_n(\theta) = 0$ と $\tilde{S}_n(\theta) = 0$ にもとづく 2 つの推定量の漸近同等性は，以下の正則条件の下で示すことができる．

(1) 推定量 $\hat{\theta}_n$ は $\hat{S}_n(\hat{\theta}_n) = o_P(n^{-1/2})$ および $\hat{\theta}_n \xrightarrow{\mathrm{P}} \theta$ を満たす．

(2) 任意の $\varepsilon > 0$ に対して，次を満たす P-Donsker クラス $\mathscr{G}$ が存在する：$P_*(\hat{V} \in \mathscr{G}) > 1 - \varepsilon$.

(3) ある $\tau > 0$ に対して，次の $\mathscr{F}_1$ は P-Donsker クラス，$\mathscr{F}_2$ は P-Glivenko-Cantelli クラスである：

$$\mathscr{F}_1 = \left\{ \frac{\dot{g}_{\theta_1}(z)(y - g_{\theta_2}(z))}{W(z)} : \|\theta_1 - \theta\| \le \tau,\ \|\theta_2 - \theta\| \le \tau,\ W \in \mathscr{G} \right\}$$

$$\mathscr{F}_2 = \left\{ \frac{\dot{g}_{\theta_1}(z)\dot{g}_{\theta_2}(z)'}{W(z)} : \|\theta_1 - \theta\| \le \tau,\ \|\theta_2 - \theta\| \le \tau,\ W \in \mathscr{G} \right\}$$

(4) $\hat{\theta}_n$ と任意の $\tilde{\theta}_n \xrightarrow{\mathrm{P}} \theta$ に対して．次が成り立つ：

$$E\left[\frac{\dot{g}_{\hat{\theta}_n}(Z)}{\hat{V}(Z)} - \frac{\dot{g}_\theta(Z)}{\tilde{V}(Z)}\right]^2 \xrightarrow{\mathrm{P}} 0, \qquad \left| E\left[\frac{\dot{g}_{\hat{\theta}_n}(Z)\dot{g}_{\tilde{\theta}_n}(Z)'}{\hat{V}(Z)}\right] - E\left[\frac{\dot{g}_\theta(Z)\dot{g}_\theta(Z)'}{\tilde{V}(Z)}\right]\right| \xrightarrow{\mathrm{P}} 0$$

これらの条件の下，$w(z) = \dfrac{V(z)}{\tilde{V}(z)}$ を考え，$\tilde{\tilde{I}}_{\theta,\eta} = E[w(Z)U(Z)] = E\left[\dfrac{\dot{g}_\theta(Z)\dot{g}_\theta(Z)'}{\tilde{V}(Z)}\right]$ とおくと，$\hat{\theta}_n = \theta + \mathbb{P}_n\left\{\tilde{\tilde{I}}_{\theta,\eta}^{-1}\tilde{\ell}_{\theta,\eta}(Y,Z)w(Z)\right\} + o_P(n^{-1/2})$ が成り立つ．したがって，$\sqrt{n}(\hat{\theta}_n - \theta) \rightsquigarrow \mathrm{N}\left(0, \tilde{\tilde{I}}_{\theta,\eta}^{-1}\right)$，$\Sigma_{\theta,\eta}^w = \tilde{\tilde{I}}_{\theta,\eta}^{-1} \ge \tilde{I}_{\theta,\eta}^{-1}$ となる．とくに，条件 (1)-(4) が $\tilde{V} = V$ に対して満たされるなら，$\hat{\theta}_n$ は有効影響関数 $\tilde{I}_{\theta,\eta}\tilde{\ell}_{\theta,\eta}(y,z)$ をもつ漸近的線形推定量 $\hat{\theta}_n = \theta + \mathbb{P}_n\{\tilde{I}_{\theta,\eta}^{-1}\tilde{\ell}_{\theta,\eta}(Y,Z)\} + o_P(n^{-1/2})$ であり，したがって θ の漸近的有効推定量である．

4．単回帰係数の漸近有効推定

単回帰モデル $E(Y|Z=z) = g_\theta(z) = \alpha + \beta z$, $\theta = (\alpha,\beta)' \in \mathbb{R}^2$ を考える．ただし，共変量 Z の範囲は $[a,b]$ であるとする．サンプル $(Y_1,Z_1),\ldots,(Y_n,Z_n)$ にもとづき，F と H を $\hat{F}_n(u) = \frac{1}{n}\sum_{i=1}^n 1\{Z_i \le u\}$，$\hat{H}_n(u) = \frac{1}{n}\sum_{i=1}^n \hat{e}_i^2 1\{Z_i \le u\}$ で推定する．ここで，$\hat{e}_i = Y_i - \hat{\alpha} - \hat{\beta}Z_i$ であり，$\hat{\theta} = (\hat{\alpha},\hat{\beta})'$ は θ の通常の最小 2 乗推定量である．このとき，$V(z)$ の推定量 $\hat{V}(z)$ を次で定義する．ここで L および h は適当にとられたカーネルおよびバンド幅である：

$$\hat{V}(z) = \begin{cases} \dfrac{\displaystyle\int_{\mathbb{R}} \frac{1}{h} L\Big(\frac{z-u}{h}\Big)\, d\hat{H}_n(u)}{\displaystyle\int_{\mathbb{R}} \frac{1}{h} L\Big(\frac{z-u}{h}\Big)\, d\hat{F}_n(u)} = \dfrac{\displaystyle\sum_{i=1}^n \hat{e}_i^2 L\Big(\frac{z-Z_i}{h}\Big)}{\displaystyle\sum_{i=1}^n L\Big(\frac{z-Z_i}{h}\Big)} & z \in [a+h, b-h] \\ \hat{V}(a+h) & z \in [a, a+h) \\ \hat{V}(b-h) & z \in (b-h, b] \end{cases}$$

このとき，$\mathscr{F}_1$，$\mathscr{F}_2$ の P−Donsker 性，および条件 (1)-(4) を確かめることにより，$\hat{V}$ を用いて構成される推定方程式 $\hat{S}_n(\theta) = 0$ の解 $\check{\theta}$ は θ の一致推定量であり，漸近的有効推定量となることが示される．

参考文献

Bickel, P.J., Klaassen, C.A., Ritov, Y., and Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models.* Springer-Verlag, New York.

Kosorok, M.R. (2008). *Introduction to Empirical Processess and Semiparametric Inference.* Springer, New York.

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications in Statistics.* Springer-Verlag, New York.

遺伝統計学と最近のゲノム研究事情

株式会社スタージェン　上辻茂男
kamitsuji@stagen.co.jp

生物の設計図である染色体は、人間の場合１人あたり 30 億個のゲノム配列情報を含んでいる。そしてゲノム配列の違いが個々の多様性を生みだす原因の１つであると考え、たとえば、病気発症の有無、薬剤の効果や副作用発症の有無などに注目し、それらの原因となるゲノム配列箇所を探索する研究の総称をゲノム研究とよぶ。

ゲノム研究の歴史はまさに統計学の歴史である。ゴールトン、カール・ピアソン、フィッシャーらは遺伝学者であり、遺伝的現象を説明するためにさまざまな新しい統計手法を生み出した。ゴールトン、ピアソンらは生物計測学派、フィッシャーはメンデル学派とよばれ、両学派の間で激しい対立が続いたことは有名な話であろう。それは「観測データから原因と結果の間のメカニズムを推測する」生物計測学派に対して、メンデル学派は「結果はメンデルの法則と呼ばれる遺伝継承法則に基づいて得られる、つまり原因と結果の間のメカニズムは遺伝継承法則そのものであり、推測するものではない」と考えていたからだ。しかしメンデル学派が考える原因、つまりゲノム配列が観測できるようになったのは 1970 年になってからである。

ゲノムの配列の観測は、1952 年のワトソン・クリックの染色体 2 重螺旋構造の発見、1970 年代のゲノム配列観測法（サンガー法）の開発によって可能となった。そして 1980 年代には世界中でゲノム配列が観測され、様々な疾患の原因遺伝子（遺伝子とは特定のゲノム配列領域の名称。人間の染色体上には約 2 万数千個ほどの遺伝子があると言われている。）が特定された。このとき、再現可能な結果を導いた研究はすべて遺伝継承法則に基づいた解析結果のみであった（このような病気はメンデル型遺伝病と呼ばれる）。そしてゲノム研究において、遺伝継承法則に基づいて多様性を説明する統計学を遺伝統計学（Statistical Genetics）と呼んだ。

ゲノム配列を解析の対象とする以上、遺伝継承法則を無視することはできない。しかし生活習慣病など、多様性の対象が遺伝的要因だけでなく環境要因も強く影響するものも多く、遺伝継承法則だけでは説明できない対象も増えてきた。現在は、遺伝継承法則に加えて、統計モデルを導入した解析方法や、情報科学的アプローチも多く見られるようになった。バイオインフォマティクスと呼ばれることもある。

またゲノム配列情報は、計測器や計算機の技術革新によって大量に観測されるようになった。ゲノムワイド関連研究（Genome-Wide Association Study, GWAS）は、集団においてゲノム配列が異なると分かっている場所を、1 人あたり 100 万個観測し、原因を探索する研究で、現在のゲノム研究の主流である。海外などでは 1 万人規模の研究が盛んに行われる。さらに最新の技術では 1 人の 30 億個のゲノム配列をすべて観測する次世代シーケンサと呼ばれる測定器も登場した。しかし情報が増えると有用なものだけでなく擬陽性も増える。遺伝統計学に基づく解析では、原因と結果の間に遺伝継承法則が成立するか、ゲノム配列１つ１つに対して検定を行うため、タイプ I の誤りの多重性問題が深刻である。

本報告では、ゲノム研究の上で無視することができない遺伝継承法則とその特徴を紹介し、現在のゲノム研究の統計的問題を紹介する。

遺伝統計学と最近のゲノム研究事情

株式会社スタージェン　上辻茂男
kamitsuji@stagen.co.jp

● ゲノム研究の対象

ゲノム研究において、原因を「アレル（allele)」、結果を「形質（trait)」と呼ぶことが多い。アレルとは、世代から世代へ安定して伝達される物体を示し、父と母から１つずつ伝達される。アレルの存在する場所を座位（locus, loci）とよび、父と母が伝達されたアレルの組み合わせを遺伝子型（genotype）と呼ぶ。ただし、座位の定義によって、対象となるアレルが変わるが、ゲノム配列そのものを対象としたことに変わりはない。一方、形質とはいわば目的変量である。特に、形質の観測値を表現型（phenotype）と呼ぶ。

● 遺伝継承法則（laws of inheritance）

メンデルの３つの法則と１つの例外からなる法則で、アレルの伝達から表現型（phenotype）の取り方までを制御する確率法則である。アレルを研究対象とする以上、遺伝継承法則を無視することはできない。

● 遺伝継承法則に基づいた解析

観測されたデータが遺伝継承法則に適合するかどうか検定する。またゲノム情報と形質との間の関連性の有無の検定と関連性の強さの推定が主な解析目的である。家系情報を用いた解析を連鎖解析（linkage analysis）と、個体間のつながりがない集団における関連解析に分かれる。

● 候補遺伝子研究とゲノムワイド関連研究

いくつかの注目する遺伝子領域の座位について関連性を議論する研究を候補遺伝子研究（candidate gene approach）とよび、全染色体の中から網羅的に大量の座位について関連性を調べる研究をゲノムワイド関連研究（Genome-Wide Association Study, GWAS）とよぶ。GWAS では、100 万座位のゲノム情報との関連性について検討することが多く、タイプＩの誤りの多重性が問題となる。

● 次世代シーケンサ

生物の全塩基配列を観測するための機械をさす。人であれば 1 人につき 30 億塩基（30 億座位と考えてよい）の観測が得られる。

以上

B-スプライン非線形回帰推定量の漸近挙動

島根大学大学院総合理工学研究科　吉田 拓真
島根大学総合理工学部　内藤 貫太

1　はじめに

複雑なデータに対して平滑化を行うとき, データに非線形回帰モデルを当てはめ解析する方法が考えられる．中でも，切断べき関数を用いたスプライン平滑化は非線形回帰の代表的な手法となっている．本講演では，B-スプライン非線形回帰推定量の漸近分布について議論する．

スプラインの漸近理論については，Silverman(1984), Hall and Opsomer(2005) など少数の結果しか得られていない．しかし，近年 Li and Ruppert(2008) によって罰則付きスプラインの漸近分布が示された．これは，p 次 B-スプライン非線形回帰において，罰則関数として m 次差分行列を用いた罰則付きスプラインに関する結果であり，$(p,m)=(0,1)$, $(0,2)$, $(1,1)$ としたもとでの罰則付きスプライン $\hat{f}(x)$ は kernel 回帰における Nadaraya-Watson 推定量と漸近同等であるというものである．つまり，$\hat{f}(x)\sim(1/n)\sum_{i=1}^{n}H_{h_n}(x-x_i)y_i$ となる．ここで，$H_{h_n}(x)=(1/h_n)H(x/h_n)$ はある kernel 関数であり，h_n は equivalent bandwidth である．しかしながら，この結果は上記の (p,m) のときでしか議論されておらず，p が高次の場合，その漸近分布についてはまだ導出されていない．

一方で，B-スプラインによる平滑化を考えたときに，一般には $p=3$, 特に $(p,m)=(3,2)$ がよく用いられる．そこで，罰則付きスプラインの漸近分布も $p=0,1$ などの場合でなく，より高次のものが求められるであろう．

2　問題設定

得られたデータ $\{(y_i,x_i):i=1,\cdots,n\}$ に対して，

$$y_i=f(x_i)+\varepsilon_i=\sum_{j=1}^{K(n)+p}B_j^{[p]}(x_i)b_j+\varepsilon_i \tag{1}$$

を考える．ここで，$x_i=i/n$ とし，ノット $\kappa_j=j/K(n)(j=1,\cdots,K(n)+p)$ に対し，$B_j^{[p]}(x)$ を p 次の B-スプライン基底関数とする．また，$b_j(j=1,\cdots,K(n)+p)$ は未知パラメータであり，$K(n)$ は，$K(n)\to\infty(n\to\infty)$ であり，$M=n/K(n)\in\mathbb{Z}$ と仮定する．さらに，$E[\varepsilon_i]=0, V[\varepsilon_i]=\sigma^2(x_i)$ とし，$\sigma^2(x)$ は連続関数，ε_i は互いに独立な確率変数であるとする．

(1) は，$\boldsymbol{y}=(y_1\ \cdots\ y_n)'$, $X^{[p]}=(B_j^{[p]}(x_i))_{ij}$, $\boldsymbol{b}=(b_1\ \cdots\ b_{K(n)+p})'$, $\boldsymbol{\varepsilon}=(\varepsilon_1\ \cdots\ \varepsilon_n)'$ を用いると

$$\boldsymbol{y}=X^{[p]}\boldsymbol{b}+\boldsymbol{\varepsilon}$$

と書ける．このとき，$\boldsymbol{b}$ の推定は m 次差分行列を罰則関数として用いたときの罰則付き最小二乗法によって得るものとする．すなわち，$\boldsymbol{b}$ の推定量 $\hat{\boldsymbol{b}}$ は (p,m) を 1 組み与えたもとで，

$$((X^{[p]})'X^{[p]}+\lambda_n^*(D^m)'D^m)\hat{\boldsymbol{b}}=(X^{[p]})'\boldsymbol{y}$$

を満たすものとして得る．ここで，λ_n^* は平滑化パラメータであり，$(D^m)'D^m$ は m 次差分行列である．この $\hat{\boldsymbol{b}}$ を用いて，各点 $x\in(0,1)$ に対して，回帰関数 $f(\cdot)$ の推定量を

$$\hat{f}(x)=\sum_{j=1}^{K(n)+p}B_j^{[p]}(x)\hat{b}_j$$

と構成する．これを罰則付きスプラインという (Eilers and Marx(1996)).

本稿では，与えられた (p,m) に対して，$\hat{f}(x) \sim (1/n)\sum_{i=1}^{n} H_{h_n}(x - x_i)y_i$ となることに論点を置く．これは，$q = \max\{p,m\}$ とするとき，行列 $\Lambda = ((X^{[p]})'X^{[p]} + \lambda_n^*(D^m)'D^m) = (\Lambda_1 \ \cdots \ \Lambda_{K(n)+p})$ の部分行列 $(\Lambda_{q+1} \ \cdots \ \Lambda_{K(n)+p-q})$ が Toeplitz 構造を持っているという性質を利用して導かれる．つまり，$\Lambda_j \ (j = q+1, \cdots, K(n)+p-q)$ は

$$(0 \ \cdots \ 0 \ \omega_q \ \cdots \ \omega_1 \ \omega_0 \ \omega_1 \ \cdots \ \omega_q \ 0 \ \cdots \ 0)$$

という構造を持っている．ただし，各 Λ_j に対して j 行目が ω_0 である．

Li and Ruppert(2008) では，$(p,m) = (0,1),(1,1)$ のとき，$H(x) = (1/2)\exp[-|x|], h_n = hn^{-1/5}$，$h > 0$ であること，$(p,m) = (0,2)$ のとき，$H(x) = L^{-1}\exp[-|x|]\{\cos(x)+\sin(|x|)\}, h_n = hn^{-1/9}$，$h > 0$ であることが示されている．ここで，L は normalized constant である．そこで，我々の目的は Li and Ruppert(2008) が示した $(p,m) = (0,1),(0,2)$ の場合の理論を応用し，主に $(p,m) = (2,2),(3,1)$ としたときの罰則付きスプラインの漸近分布を導出することである．

3 罰則付きスプラインの漸近分布

$(p,m) = (3,1)$ のときの罰則付きスプラインは特定の仮定，条件のもとで

$$n^{2/5}\{\hat{f}(x) - f(x)\} \to N(\beta_1(x), \nu_1(x))$$

を満たす．ただし，$\beta_1(x) = h^2 f^{(2)}(x)$, $\nu_1(x) = (4h)^{-1}\sigma^2(x)$ である (Wand and Jones(1995)).

また，$(p,m) = (2,2)$ のときは

$$n^{4/9}\{\hat{f}(x) - f(x)\} \to N(\beta_2(x), \nu_2(x))$$

を満たす．ただし，$H(x)$ は第 2 節において，$(p,m) = (0,2)$ のときのものとして，

$$\beta_2(x) = 24^{-1}h^4 f^{(4)}(x)\int_{-\infty}^{\infty} x^4 H(x)dx, \ \nu_2(x) = h^{-1}\sigma^2(x)\int_{-\infty}^{\infty} H(x)^2 dx$$

である．

参考文献

[1] Eilers,P.H.C. and Marx,B.D.(1996). Flexible smoothing with B-splines and penalties(with Discussion). *Statist.Sci.***11**,89-121.

[2] Hall,P and Opsomer,J.D.(2005).Theory for penalized spline regression. *Biometrika* **92**,105-118.

[3] Li,Y and Ruppert,D (2008).On the asymptotics of penalized splines : *Biometrika* **95**,2, 415-436.

[4] Silverman,B,W.(1984). Spline smoothing:the equivalent variable kernel method.*Ann.Statist* **12**, 898-916.

[5] Wand,M.P. and Jones,M.C(1995). *Kernel Smoothing.* Chapman & Hall/CRC.

商品先物市場のモデル化によるリスク評価（１）
–問題提起ならびに統計的データ解析–

株式会社 QUICK 青木 義充

はじめに

商品先物市場における取引参加者，商品取引員（売買取引の仲介者）が考慮すべきリスクを精緻に計測するためには，実際の商品先物市場の仕組みを十分に考慮したモデルを構築することが必要である．本報告では，以下の２点に焦点を当てて報告を行なった．

はじめに，商品先物取引市場に上場されている商品先物の価格変動に焦点をあて，価格変動のモデル化とその推定方法について説明を与えた．取引参加者が注視すべき価格変動リスク（市場リスク）の指標である Value at Risk(VaR) の算出法も合わせて導出した．

次に，商品先物取引の特徴である「証拠金取引制度」に焦点を絞り，取引参加者，商品取引員がともに考慮すべき取引追証拠金（追証）が発生する確率，商品取引員が考慮すべきデフォルトリスクの算出法をそれぞれ導出した．

なお，商品先物市場の制度については，日本商品先物取引協会が発行する [1]，[2] に基づいており，実際の商品としては，東京工業品取引所に上場されている，金，銀，白金，パラジウム，アルミニウム，ガソリン，灯油，原油，ゴムのうち，2008 年中に限月を迎えたものを対象とした．

価格のモデル化と価格変動リスク

実際の商品先物市場では，対象商品の受給事情の変化や急激な先行き見通しの修正などの不確定的な事象によって，相場が急激に変動することを防ぐため，1 日における値動きについて，前日価格に対し一定金額を加減した制限値段が定められている．このため，対象商品の価値が高騰しているとしても，同日の価格には制限値段を越えた部分の価値が反映されない．すなわち，価格のヒストリカルデータからでは，対象商品の価値に関する情報を観測できない場合がある．我々は，この制限値段幅が与える影響を重視し，価格変動のモデルに導入した．観測可能である日々の価格とは別に，対象商品の潜在的な価値を「真の価値」として定義した．真の価値は一部観測不能という負の側面を持つが，その変動に値幅制限という制約がないために，価格よりも自由にモデルを設定できるという利点を持つ．実際の価格データを解析したところ，価格の前日差が自己回帰性を示すという結果が得られた．この性質は株式には見られなかったため，「価格前日差が自己回帰性を有する」という性質を商品先物の価格特有のものと考え，モデル化に組み込んだ．

なお，価格変動モデルを導出する理論的枠組みについては，「商品先物市場のモデル化によるリスク評価（２）–データ解析のための統計理論–」にて，別途報告を行なっている．

証拠金取引のモデル化とデフォルトリスク

商品取引員によって課せられる取引本証拠金は，あらかじめ取引所によって決められた取引本証拠金基準額を基準として決定される．取引本証拠金基準額は，商品の価格変動を取引所が勘案して決定されるもので，毎月見直し・公表をしている．商品先物価格の下落によって，取引参加者の資産における損失が取引本証拠金基準額の半額を上回った場合には追証が発生する．取引参加者は，当該商品の取引を続行する場合には指定された追証を預託せねばならない．一方，取引参加者が追証を納めない場合には，仲介者である商品取引員は該当取引参加者の建玉を次営業日以降に清算せねばならない．このとき，価格変動の状況によっては，元々積み立てられていた取引本証拠金を上回る損失が起こる場合がある．本報告では，この超過損失が起こるリスクをデフォルトリスクとして定義した．具体的には，取引参加者が追証を預託する／しないに二項確率モデルを導入し，実際の価格データを元に超過損失金額をシミュレーションを用いて算出した．

参考文献

[1] 日本商品先物取引協会 [2007],「商品先物取引委託のガイド (第 15 版)」

[2] 日本商品先物取引協会 [2007],「商品先物取引委託のガイド別冊 (第 27 版)」

商品先物市場のモデル化によるリスク評価（2）

–データ解析のための統計理論–

山形大・地域教育文化　加藤　剛　 E-mail: tkato@e.yamagata-u.ac.jp

1　目的

先物取引とは，価格が変動する商品について，現物の受け渡しは数ヶ月先に実行することとして，売買の約定を結ぶ取引のことである．対義語が現物取引で，これは，売買契約の成立と同時または数日後に現品の受渡しを行う取引である（詳細は [1] を参照）．

先物取引を行いたい場合，通常は取引業者（商品取引所の会員）に証拠金と呼ばれる担保を預け，取引を委託する．売買を行う過程で計算上の損失が一定額以上になった場合，委託者が取引を継続する場合は，追証拠金（新たな担保）を取引業者に支払わなければならない．取引を精算（決済）する場合は，追証拠金を負担する必要はない．損失が生じた状態で委託者が決済を行った場合，取引業者はその損失を負う．したがって，帳入値段（最終約定値）が将来どの程度変動して，どの程度の損失を被る可能性があるかを予測することは，取引業者にとって切実な問題である．本研究は，帳入値段のデータに対して時系列モデル化を行い，マルコフ連鎖モンテカルロ法 (MCMC) を利用してモデルの未知パラメータを推定し，帳入値段の将来の確率分布を算出することを目的とする．

2　モデル化

t 時点の帳入値段（最終約定値）を P_t とする．これは，すべての時刻において観測可能な変量である．他方，モデル化を容易にするために，帳入値段とは別に，対象商品の潜在的な価値を「真の価値」とし，これを X_t で表すことにする．いま，値幅制限を L とするとき，P_t と X_t の間には，次の関係が成立する．

$$P_{t+1} = \begin{cases} P_t + L & : \quad X_{t+1} \geq P_t + L \\ X_{t+1} & : \quad P_t - L < X_{t+1} < P_t + L \\ P_t - L & : \quad X_{t+1} \leq P_t - L \end{cases} \tag{1}$$

実際の $\{P_t\}$ のデータをもとに単位根検定とモデルのあてはめを行うと，前日差 $\Delta P_t = P_t - P_{t-1}$ をとることによって非定常性をもつ可能性が消え，ΔP_t が自己回帰性をもつことが発見できた．そこで，真の価値 ΔX_t についても前日差を考えることにして，AIC 基準をもとに，$\{\Delta X_t\}$ に対して AR(1)

$$\Delta X_t = \phi\,\Delta X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \quad t = 2, 3, \ldots, \tag{2}$$

を仮定する．ここで，$|\phi| < 1$ と $\sigma > 0$ が未知パラメータとなる．

3　パラメータ推定

ΔX_t に対するモデルの仮定から，σ^2 と ϕ，および，初期値 $\Delta X_1 = \Delta x_1$ が与えられたと仮定するとき，$\Delta \boldsymbol{X} = (\Delta X_2, \Delta X_3, \ldots, \Delta X_n)$ の同時確率密度関数は，

$$f(\Delta \boldsymbol{x} | \sigma^2, \phi, \Delta x_1) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^{n} (\Delta x_t - \phi\,\Delta x_{t-1})^2 \right\} \tag{3}$$

となる．

他方，σ^2 の事前分布として，$\alpha_0 > 0$ を形の母数，$\beta_0 > 0$ を尺度母数とする逆ガンマ分布 $\sigma^2 \sim \mathrm{IG}(\alpha_0, \beta_0)$ を選ぶとする．このとき，ϕ と Δx_1, $\Delta\boldsymbol{X} = \Delta\boldsymbol{x}$ が与えられたときの σ^2 の条件付事後分布は，

$$\sigma^2|(\phi, \Delta x_1, \Delta\boldsymbol{x}) \ \sim\ \mathrm{IG}(\alpha_1, \beta_1), \quad \alpha_1 = \alpha_0 + \frac{n}{2}, \quad \beta_1 = \beta_0 + \frac{1}{2}\sum_{t=2}^{n}(\Delta x_t - \phi\,\Delta x_{t-1})^2 \tag{4}$$

である．さらに，ϕ の事前分布として $(-1,1)$ 上の一様分布 $U(-1,1)$ を仮定すると，σ^2 と $\Delta x_1, \Delta\boldsymbol{X} = \Delta\boldsymbol{x}$ が与えられたときの ϕ の条件付事後分布は，

$$\begin{gathered}\phi|(\sigma^2, \Delta x_1, \Delta\boldsymbol{x}) \ \sim\ TN_{(-1,1)}(\mu_\phi,\, \sigma_\phi^2), \\ \mu_\phi = \sum_{t=2}^{n}\Delta x_t \Delta x_{t-1} \Bigg/ \sum_{t=2}^{n}\Delta x_{t-1}^2\,, \quad \sigma_\phi^2 = \sigma^2 \Bigg/ \sum_{t=2}^{n}\Delta x_{t-1}^2\end{gathered} \tag{5}$$

で与えられる．

(1) より，

- $|P_t - P_{t-1}| < L$ かつ $|P_{t-1} - P_{t-2}| < L \ \Rightarrow\ \Delta X_t = \Delta P_t \ \Rightarrow\ \Delta X_t$ は観測可能
- $|P_t - P_{t-1}| = L$ または $|P_{t-1} - P_{t-2}| = L \ \Rightarrow\ \Delta X_t \neq \Delta P_t \ \Rightarrow\ \Delta X_t$ は観測不可能

が成り立つので，ΔX_t に対するモデル (2) は，トービット・モデル（打ち切りのある線形モデル）でもあることがわかる．そこで，(3)，(4)，(5) で求めた確率分布と，トービット・モデルにおけるデータ補間の MCMC アルゴリズム，AR(1) におけるパラメータ推定の MCMC アルゴリズムをうまく組み合わせると，観測不可能な ΔX_t を補間しつつ，未知パラメータの ϕ と σ^2 の推定量 $\widehat{\phi}$ と $\widehat{\sigma^2}$ を得ることができる（[2] 参照）．これより，ΔX_t が ΔP_t と一致して確実な ΔX_t の値がわかっている最新時刻を

$$t_0 = \arg\max\{\Delta X_t = \Delta P_t : t = 3, 4, \ldots, n\}$$

とすると，観測可能なデータ $\mathcal{P}_n = \{P_t\}_{t=1}^n$ が与えられたとき，t_0 時点から $\Delta T (\geq 1)$ 期間の価格変動の確率分布は，

$$(X_{t_0+\Delta T} - X_{t_0})|\mathcal{P}_n \ \sim\ N\left(\Delta P_{t_0}\sum_{j=1}^{\Delta T}\widehat{\phi}^j,\, \widehat{\sigma^2}\sum_{j=1}^{\Delta T} j\widehat{\phi}^{2(j-1)}\right)$$

として求めることができる．また，$t_0 + \Delta T$ 期先における VaR $100 \times \alpha\%$ 点は，Φ^{-1} を標準正規分布関数の逆関数とすると，

$$\mathrm{VaR}_{100\times\alpha} = \Delta P_{t_0}\sum_{j=1}^{\Delta T}\widehat{\phi}^j + \left(\widehat{\sigma^2}\sum_{j=1}^{\Delta T} j\widehat{\phi}^{2(j-1)}\right)^{1/2}\Phi^{-1}(\alpha)$$

で与えることができる．

参考文献

[1] 青木義充，横内大介，加藤剛．商品先物市場のモデル．2009. (投稿中)
[2] 和合肇編著.『ベイズ計量経済分析』．東京経済新報社．2005.

L_q 罰則による one-step 推定量の平滑化パラメータ選択

勘場 大 [1]　　内藤 貫太 [2]

[1] 島根大学大学院総合理工学研究科

[2] 島根大学総合理工学部　　E-mail: naito@riko.shimane-u.ac.jp

1 はじめに

Zou and Li (2008) は SCAD 罰則や L_q 罰則という罰則関数を用いて変数選択する方法を考えた．彼らが提案した one-step 推定量 (One Step Estimator;OSE) はオラクルプロパティーと呼ばれる性質を持つ．オラクルプロパティーとは，その推定量が漸近正規性とスパース性を有することを意味する．

しかしながら，OSE がオラクルプロパティーを持つためには然るべき条件を満足する平滑化パラメータが必要である．Zou and Li (2008) はその平滑化パラメータを 5-fold CV で選択した．しかしながら 5-fold CV で決定した平滑化パラメータでは，その条件を満足するとは限らず，そもそも 5-fold CV の実装それ自体のコンピュータコストも大きくなる．そこで本講演においては，OSE がオラクルプロパティーを有するような平滑化パラメータを簡単に決定する方法を提案する．

2 問題設定

データ $\{(y_i, \boldsymbol{x}_i) | i = 1, \cdots, n\}$ が観測されたとする．ただし，y_i は応答変数で，$\boldsymbol{x}_i$ は p 次元説明変数ベクトルである．$\boldsymbol{x}_i$ を与えたときの y_i が指数型分布族の密度を持つと仮定する．そのときの対数尤度を $\ell(\boldsymbol{\beta}, \phi)$ で表す．

Zou and Li(2008) は，罰則関数 p_{λ_n} として SCAD 罰則関数や L_q 罰則関数を考え，罰則付き尤度を $\boldsymbol{\beta}$ について最大化する問題を考えている．ただし，λ_n は平滑化パラメータである．特に，平滑化パラメータ λ_n と罰則関数 $p(\cdot)$ が分離できる場合は，

$$\ell(\boldsymbol{\beta}, \phi) - n\lambda_n \sum_{j=1}^{p} p(|\beta_j|) \tag{1}$$

を $\boldsymbol{\beta}$ について最大化する．

実際は (1) を最大化することは困難となるので，対数尤度と罰則関数を近似したあと -1 倍して最小化する．それを最小にする $\boldsymbol{\beta}$ を OSE といい，

$$\widehat{\boldsymbol{\beta}}(\mathrm{ose}) = \mathrm{argmin} \left\{ \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)})^T [-\nabla^2 \ell(\boldsymbol{\beta}^{(0)})](\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)}) + n \sum_{j=1}^{p} p'_{\lambda_n}(|\beta_j^{(0)}|)|\beta_j| \right\}$$

で表わされる．

平滑化パラメータ λ_n と罰則関数 $p(\cdot)$ が分離できるとき，この推定量は

条件 1. $p'(\cdot)$ は $(0, \infty)$ で連続である．

条件 2. $t \to 0_+$ のとき，$p'(t) = O(t^{-s})$ を満たす $s > 0$ が存在する．

条件 3. $n^{(1+s)/2}\lambda_n \to \infty$ かつ $\sqrt{n}\lambda_n \to 0$ である．

という条件の下で，オラクルプロパティーと呼ばれる性質を持つ (Zou and Li(2008)).

3　平滑化パラメータの決定

$\beta = (\beta_1, \cdots, \beta_p)$ が事前分布として指数型分布族の密度

$$g(\beta;\theta) = d(\theta)\exp\left(\frac{1}{\theta}\sum_{j=1}^{p} T(\beta_j)\right)$$

を持つとする．ただし，θ はパラメータ，$d(\theta)$ は $g(\beta;,\theta)$ を積分して 1 となる定数，$T(\cdot)$ は可測関数である．このとき，$(y_1, \cdots, y_n)$ の同時密度の対数は，

$$\ell(\beta,\phi) + \frac{1}{\theta}\sum_{j=1}^{p} T(\beta_j) + \log d(\theta) \tag{2}$$

となる．この式を β について最大化することを考える．

このとき，$\theta = 1/n\lambda_n$，$T(\beta_j) = -p(|\beta_j|)$ とした (2) の最大化と (1) の最大化は同値であることに注意する．これよりパラメータ θ を用いると，平滑化パラメータは，

$$\lambda_n = \frac{1}{n\theta} \tag{3}$$

で与えられることがわかる．

3.1　罰則関数と平滑化パラメータの条件

(3) に含まれるパラメータ θ を定数とすると，$0 < s < 1$ に対しては我々の提案した平滑化パラメータでは条件 3 を満たさない．条件 3 を満たすために，我々の提案した平滑化パラメータに含まれる θ を n に依存させて $\theta = 1/n^{\alpha}$ とする．ただし，$\alpha \in \mathbb{R}$ である．このとき，

$$\frac{1}{2} > \alpha > \frac{1-s}{2}, \quad (0 < s < 1)$$

を満たす α を取ってくれば，$\theta = n^{-\alpha}$ とした (3) は条件 3 を満たすことが確認できる．

特に，L_q 罰則 $p(|\beta|) = |\beta|^q (0 < q < 1)$ について考える．このとき L_q 罰則は条件 1 と条件 2 を満たすことがわかる．ただし，条件 2 に含まれる s は $s = 1 - q$ である．ここで α が満たすべき条件は，$1/2 > \alpha > q/2$ である．したがって，例えば $q < 1/2$ に対しては $\alpha = q$ と取ってくればよい．

4　まとめと展望

OSE を構成する際に必要な平滑化パラメータとして，我々の提案した平滑化パラメータは 5-fold CV で決定した平滑化パラメータよりもより良い振る舞いをしていることがわかった．

ここでのモデルは説明変数とパラメータの線形性を仮定している．ここでの議論はスプラインを用いたセミパラメトリック回帰モデルなどでも用いることができる．一般化加法モデルなども含むような広いモデルにおいて同様な理論の整備が期待される．

参考文献

[1] Zou,H. and Li,R.(2008).One-step sparse estimates in nonconcave penalized likelihood models. *The Annals of Statistics* **36**: 1509-1533.

Estimating function approach for CHARN models

金井裕臣 / 早稲田大学
小方浩明 / 首都大学東京
谷口正信 / 早稲田大学

1　はじめに

Godambe (1960, 1985) や Hansen (1982) らによって，最小 2 乗推定量と最尤推定量の架け橋となる推定関数による推定法が提唱された．ここでは，多くの非線形時系列モデルを特別な場合として含む CHARN モデルに対して推定関数法を適用することを考える．推定関数法は必ずしも有効推定量を導くとは限らないため，我々は漸近有効となる最適な推定関数を提案する．さらに CHARN モデルにおける経験尤度法も考察する．

2　CHARN モデル

以下のモデル（CHARN モデル）

$$\boldsymbol{X}_t = \boldsymbol{F}_{\boldsymbol{\theta}}(\boldsymbol{X}_{t-1}, \ldots, \boldsymbol{X}_{t-p}) + \boldsymbol{H}_{\boldsymbol{\theta}}(\boldsymbol{X}_{t-1}, \ldots, \boldsymbol{X}_{t-q})\boldsymbol{U}_t \tag{1}$$

によって定義される多変量確率過程 $\{\boldsymbol{X}_t = (x_{1,t}, \ldots, x_{m,t})^T, t \in \boldsymbol{N}\}$ を考える．ここに $\boldsymbol{F}_{\boldsymbol{\theta}} : \boldsymbol{R}^{mp} \to \boldsymbol{R}^m$ はベクトル値可測関数，$\boldsymbol{H}_{\boldsymbol{\theta}} : \boldsymbol{R}^{mq} \to \boldsymbol{R}^m \quad \boldsymbol{R}^m$ は正定値可側関数，$\{\boldsymbol{U}_t = (U_{1,t}, \ldots, U_{m,t})^T\}$ は i.i.d. で平均 $\boldsymbol{0}$，可積分な確率ベクトルの列である．また，ベクトルパラメーター $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_r)^T$ は開集合 $\boldsymbol{\Theta} \subset \boldsymbol{R}^r$ の内点に存在するとし，その真値を $\boldsymbol{\theta}^0$ で表す.

3　推定関数法

一般的なフレームワークでの推定関数法について述べる．$\boldsymbol{X}^{(n)} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^T$ を，確率過程の形をなす確率ベクトルの集まり，$\mathcal{F}$ を $\boldsymbol{R}^{mn}$ 上の確率分布 F の族とし，$\boldsymbol{\theta} = \boldsymbol{\theta}(F)$, $F \in \mathcal{F}$ を実母数ベクトルとする．また，$\boldsymbol{X}^{(n)}$ の確率分布は $\mathcal{F}$ に属するものとする．このとき，以下の等式

$$E_F[\boldsymbol{G}_{\boldsymbol{\theta}^0}(\boldsymbol{X}^{(n)})] = \boldsymbol{0}, \quad F \in \mathcal{F}.$$

を満たすような $\boldsymbol{X}^{(n)}$ と $\boldsymbol{\theta}$ の実ベクトル関数 $\boldsymbol{G}$ を不偏推定関数と呼ぶ．$\boldsymbol{\theta}^0$ の推定量として，推定方程式 $\boldsymbol{G}_{\boldsymbol{\theta}} = \boldsymbol{0}$ の根を考え，それを $\hat{\boldsymbol{\theta}}_{EF}$ で表す．またすべての不偏推定関数 $\boldsymbol{G}$ の中で，すべての $F \in \mathcal{F}$ に対して以下の量

$$\left(E_F\left[\frac{\partial}{\partial \boldsymbol{\theta}^T}\boldsymbol{G}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(n)})\right]\right)^{-1} E_F[\boldsymbol{G}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(n)})\boldsymbol{G}_{\boldsymbol{\theta}}^T(\boldsymbol{X}^{(n)})] \left(E_F\left[\frac{\partial}{\partial \boldsymbol{\theta}^T}\boldsymbol{G}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(n)})\right]^T\right)^{-1}$$

を最小にするようなものを $\boldsymbol{G}^*$ で表し，最適推定関数と呼ぶ.

4　CHARN モデルにおける推定関数法

以下では，$\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ は CHARN モデル (1) に従っているとする．確率ベクトル $\{\boldsymbol{X}_s\,;\, t-l \le s \le t-1\}$ によって生成される σ 集合体を $\mathcal{F}^l_{t-1}$ で表し，$l = t-1$ の場合は簡単のため $\mathcal{F}_{t-1}$ と記述する．$\mathcal{G}$ を以下のような形

$$\boldsymbol{G}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(n)}) = \sum_{t=l+1}^{n} \boldsymbol{A}^l_{\boldsymbol{\theta}}(t-1)\boldsymbol{h}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t)})$$

で書けるベクトル値推定関数の族とする．　ここに $\boldsymbol{h}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t)}) = \boldsymbol{X}_t - E[\boldsymbol{X}_t|\mathcal{F}_{t-1}]$ であり，$\boldsymbol{A}^l_{\boldsymbol{\theta}}(t-1)$ は $\boldsymbol{\theta}$ に依存する任意の $\mathcal{F}^l_{t-1}$ 可測 $r \quad m$ 行列である．

次の定理は $\hat{\boldsymbol{\theta}}_{EF}$ の漸近正規性について述べたものである．

Theorem 1 適当な正則条件のもと，以下の *2* 条件を満たす推定量の列 $\hat{\boldsymbol{\theta}}_{EF}$ が存在する．

(i) $\hat{\boldsymbol{\theta}}_{EF} \overset{a.s.}{\to} \boldsymbol{\theta}^0$

(ii) 任意の $\quad > 0$ に対し $P(E) > 1 -$ 　なる集合 E と自然数 n_0 が存在し，E 上では $n > n_0$ なる n に対して $\boldsymbol{G}_{\hat{\boldsymbol{\theta}}_{EF}} = \boldsymbol{0}$ が成り立つ．

さらに，

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{EF} - \boldsymbol{\theta}^0) \xrightarrow{d} N(\boldsymbol{0}, U^{-1}R(U^{-1})^T).$$

が成り立つ．*(*行列 U, R の形は省略*)*.

次の定理は，最適推定関数 $\boldsymbol{G}_{\boldsymbol{\theta}} = \boldsymbol{G}^*_{\boldsymbol{\theta}} \in \mathcal{G}$ について述べたものである．

Theorem 2 適当な正則条件のもと，$\mathcal{G}$ に属する推定関数 $\boldsymbol{G}_{\boldsymbol{\theta}}$ の中で，$\boldsymbol{G}_{\boldsymbol{\theta}} = \boldsymbol{G}^*_{\boldsymbol{\theta}}$ のときに $\sqrt{n}(\hat{\boldsymbol{\theta}}_{EF} - \boldsymbol{\theta}^0)$ の漸近分散が最小となる．ここに

$$\boldsymbol{G}^*_{\boldsymbol{\theta}}(\boldsymbol{X}^{(n)}) = \sum_{t=l+1}^{n} \boldsymbol{A}^{l\,*}_{\boldsymbol{\theta}}(t-1)\boldsymbol{h}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t)}).$$

であり，

$$\boldsymbol{A}^{l\,*}_{\boldsymbol{\theta}}(t-1) = \boldsymbol{C}\left(\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{F}^T_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t-1)})\right)(\boldsymbol{H}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t-1)})\boldsymbol{H}^T_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t-1)}))^{-1} \quad (\boldsymbol{C} : r \quad r \text{ 定数行列})$$

である．

最後に，経験尤度法によるアプローチについて述べる．推定関数を

$$\boldsymbol{m}_{\theta}(\mathbf{X}^{(t)}) = \boldsymbol{A}^l_{\boldsymbol{\theta}}(t-1)\boldsymbol{h}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(t)})$$

とおき，$\boldsymbol{\theta}$ における経験尤度比関数を

$$\mathcal{R}_n(\boldsymbol{\theta}) = \max_{\boldsymbol{p}}\left\{\prod_{t=l+1}^{n}(n-l)p_t \;\middle|\; \sum_{t=l+1}^{n} p_t\boldsymbol{m}_{\theta}(\mathbf{X}^{(t)}) = \boldsymbol{0},\ \sum_{t=l+1}^{n} p_t = 1,\ p_t \ge 0\right\}$$

で定義すると以下の定理を得る．

Theorem 3 適当な正則条件のもと，$-2\log\mathcal{R}_n(\boldsymbol{\theta}^0) \xrightarrow{d} \chi^2_r \quad (n \to \infty)$ が成り立つ．

On the unit root process with locally stationary innovation

新潟大学　理学部　蛭川　潤一
新潟大学　理学部　定方　真子

ランダムウォークは 1 次の自己回帰（AR(1)）モデルの係数が 1 の場合に相当し、非定常過程のクラスである単位根過程に含まれる。ランダムウォークの部分和過程は標準ブラウン運動に分布収束することがよく知られている。この事実は汎関数中心極限定理 (FCLT) と呼ばれる。Beveridge-Nelson (B-N) 分解と呼ばれる手法を使うと、ランダムウォークのイノベーション過程を i.i.d. 過程から、線形過程に一般化することが出来る。本講演では、局所定常過程 (LSP) をイノベーション過程に持つ、単位根過程や単位根近接モデルについて FCLT を導いた。局所定常過程も非定常過程であるので、これらのモデルは、二つの異なるタイプの非定常性を表現することができる。

ランダムウォークはその非定常性により、最小二乗推定量 (LSE) が漸近正規性を持たないことが知られている。そこで、LSP イノベーションを持つ単位根過程の LSE の極限分布も導いた。また LSE を特別な場合に含むような一般化した推定量のクラスも提案する。更に、LSP イノベーションを持つ単位根過程の limiting Gaussian functional (LGF) 性についても述べた。

LSP イノベーションを持つ単位根過程：

$$\begin{aligned} x_{j,T} &= x_{j-1,T} + u_{j,T} \\ &= x_{0,T} + \sum_{i=1}^{j} u_{i,T} \end{aligned}$$

ここに

$$u_{j,T} = \sum_{l=0}^{\infty} \alpha_l \left(\frac{j}{T}\right) \varepsilon_{j-l} = \int_{-\pi}^{\pi} A\left(\frac{j}{T}, \lambda\right) e^{ij\lambda} d\xi(\lambda)$$

$x_{0,T} = \sigma\sqrt{T}X(0)$, $X(0) \sim N(\gamma_X, \delta_X^2)$, $X(0) \perp \{\varepsilon_j\}$

部分和過程：

$$X_T(t) = \frac{1}{\sigma\sqrt{T}} x_{j,T} + T\left(t - \frac{j}{T}\right) \frac{1}{\sigma\sqrt{T}} u_{j,T} \qquad \left(\frac{j-1}{T} \le t \le \frac{j}{T}\right)$$

定理 1 (LSP イノベーションを持つ単位根過程についての FCLT).

$$X_T \ \Rightarrow \ h^{(1)}(X(0), W) \equiv X$$

ここに

$$h_t^{(1)}(x, y) = x + \alpha(t, 1)\, y(t) - \int_0^t \alpha'(\nu, 1)\, y(\nu) d\nu$$

$$\Longrightarrow \ dX(t) = A(t, 0)\, dW(t)$$

AR(1) モデル $x_{j,T} = \rho x_{j-1,T} + u_{j,T}$ の最小二乗推定量：

$$\hat{\rho} = \frac{\sum_{j=2}^{T} x_{j-1,T} x_{j,T}}{\sum_{j=2}^{T} \left(x_{j-1,T}\right)^2}$$

$S_{1,T} = T\left(\hat{\rho} - 1\right)$ の極限分布：

$$\begin{aligned}
&\mathcal{L}\left(S_{1,T}\right) = \mathcal{L}\left(T\left(\hat{\rho}-1\right)\right) \\
&\to \mathcal{L}\left(\frac{\frac{1}{2}\left\{X\left(1\right)^2 - X\left(0\right)^2 - \int_0^1 \sum_{l=0}^{\infty} \alpha_l\left(\nu\right)^2 d\nu\right\}}{\int_0^1 X\left(\nu\right)^2 d\nu}\right) \\
&= \mathcal{L}\left(\frac{\int_0^1 X\left(\nu\right) dX\left(\nu\right) + \frac{1}{2}\int_0^1 \left[\left\{\sum_{l=0}^{\infty} \alpha_l\left(\nu\right)\right\}^2 - \sum_{l=0}^{\infty} \alpha_l\left(\nu\right)^2\right] d\nu}{\int_0^1 X\left(\nu\right)^2 d\nu}\right)
\end{aligned}$$

疑似正規対数尤度：

$$T\left\{L_T\left(\boldsymbol{x}_T;\theta,\rho_T\right) - L_T\left(\boldsymbol{x}_T;\theta_0,1\right)\right\}$$

$$\begin{aligned}
&= \frac{1}{2}\log\frac{\left|\Sigma_{\theta_0}\right|}{\left|\Sigma_\theta\right|} + \frac{1}{2}\boldsymbol{u}_T'\left\{\Sigma_{\theta_0}^{-1} - \Sigma_\theta^{-1}\right\}\boldsymbol{u}_T \\
&\quad - \beta\frac{1}{T}\boldsymbol{u}_T'\Sigma_\theta^{-1}\tilde{\boldsymbol{x}}_T - \frac{\beta^2}{2}\frac{1}{T^2}\tilde{\boldsymbol{x}}_T'\Sigma_\theta^{-1}\tilde{\boldsymbol{x}}_T
\end{aligned}$$

$$\begin{aligned}
&\equiv T\left\{L_T\left(\boldsymbol{u}_T;\theta\right) - L_T\left(\boldsymbol{u}_T;\theta_0\right)\right\} \quad \text{(LAN)} \\
&\quad - \beta G_T^{(1)} - \frac{\beta^2}{2}G_T^{(2)}
\end{aligned}$$

LGF 性：

$$\begin{aligned}
&-\beta G_T^{(1)} \Rightarrow g^{(1)}\left(W,X\right) \\
&= -\frac{\beta\sigma_{\theta_0}^2}{\sigma_\theta^2}\int_0^1 B\left(\nu,0\right) X\left(\nu\right) dW\left(\nu\right) \\
&= -\frac{\beta\sigma_{\theta_0}^2}{\sigma_\theta^2}\int_0^1 A_{\theta_0}\left(u,\lambda\right)\left|A_\theta\left(u,\lambda\right)\right|^{-2} X\left(\nu\right) dW\left(\nu\right) \\
&= -\frac{\beta\sigma_{\theta_0}^2}{\sigma_\theta^2}\int_0^1 \frac{X\left(\nu\right)}{\left|A_\theta\left(\nu,0\right)\right|^2} dX\left(\nu\right) \\
&-\frac{\beta^2}{2}G_T^{(2)} \Rightarrow g^{(2)}\left(X\right) \\
&= -\frac{\beta^2\sigma_{\theta_0}^2}{2\sigma_\theta^2}\int_0^1 \frac{X\left(\nu\right)^2}{\left|A_\theta\left(\nu,0\right)\right|} d\nu
\end{aligned}$$

$\theta = \theta_0 \Longrightarrow$ 測度変換（ギルサノフの定理）：

$$\begin{aligned}
\frac{d\mu_Y}{d\mu_X}\left(X\right) &:= \exp\left\{g^{(1)}\left(W,X\right) + g^{(2)}\left(X\right)\right\} \\
&= \exp\left\{-\beta\int_0^1 \frac{X\left(\nu\right)}{\left|A_{\theta_0}\left(\nu,0\right)\right|^2} dX\left(\nu\right) - \frac{\beta^2}{2}\int_0^1 \frac{X\left(\nu\right)^2}{\left|A_{\theta_0}\left(\nu,0\right)\right|} d\nu\right\}
\end{aligned}$$

Higher Order Asymptotic Bond Price Valuation for Interest Rates with Non-Gaussian Dependent Innovations

東京理科大学 本田 哲弘
早稲田大学 玉置 健一郎
東京理科大学 塩濱 敬之

1 はじめに

債券価格や金利デリバティブの理論価格の導出には, 適切な金利のモデル化が重要であり, これまで数多くの金利変動を表すモデルが提案されてきた. 中でも Vasicek(1977) モデルは遷移密度が正規分布に従うためパラメータの推定が容易であることに加え, 債券価格が解析解を持つという利点がある. しかし Vasicek モデルでは, イノベーションの独立性や正規性の仮定などにより, 実際の短期金利変動を十分に表現できないことが多くの実証分析で報告されている.

そこで本研究では, 離散型 Vasicek モデルのイノベーションの独立性, 正規性の仮定を緩め, それらが債券価格やスポットレートに与える影響を明らかにする. 具体的には, Tamaki and Taniguchi(2007) のアプローチを用いる. 彼らはオプション価格評価において, 対数収益率が定常非正規過程に従うとしてオプション価格の高次評価をし, 非正規性, 従属性が債券価格に与える影響を明らかにした. 同様にして, 離散型 Vasicek モデルのイノベーションが定常非正規過程に従うとして, 債券価格とスポットレートの高次評価をする.

2 モデル

短期金利 $r_0, r_1, r_2, \cdots, r_n$ は離散間隔 Δ で観測されたとする. 本研究で考察する短期金利モデルを次のように定義する.

$$r_j - r_{j-1} = \kappa(\mu - r_{j-1})\Delta + \Delta^{1/2} X_j.$$

$\{X_j\}$ が独立, 同一に $N(0,\sigma^2)$ に従えば, このモデルは (オイラー近似による) 離散型 Vasicek モデルに一致する. 実証分析の立場から $\{X_j\}$ の非正規性, 従属性が報告されていることを踏まえて, $\{X_j\}$ に次の仮定をおく.

仮定 1 $\{X_j\}$ は 4 次定常過程で, 4 次までのキュムラントをもつとする.
仮定 2 $k = 2,3,4, j = 1,\cdots,k-1$ に対して, $\{X_j\}$ の k 次キュムラントは以下を満たすとする.

$$\sum_{u_1,\cdots,u_{k-1}=-\infty}^{\infty} (1 + |u_j|^{2-k/2})|c_{X,k}(u_1,\cdots,u_{k-1})| < \infty$$

3 債券価格評価

現時点を 0, 満期時点を T として, 債券価格は次のように表すことができる.

$$P(0,T) = \exp\left[-\frac{\kappa\mu\Delta^2}{2}\left(\sum_{i=1}^{n} a_i\right) - \frac{\Delta}{2}\{1 + a_n(1-\kappa\Delta)\} r_0\right] \times E^*\left[\exp\left(-\frac{\sqrt{T}}{\sqrt{n}}\sum_{i=1}^{n} b_i X_{n-i+1}\right)\right].$$

ここで, $n\Delta = T$, $a_i = \frac{2}{\kappa\Delta}\left[1-(1-\kappa\Delta)^{i-1}\left(1-\frac{\kappa\Delta}{2}\right)\right]$, $b_i = \frac{1}{\kappa}\left\{1-\left(1-\frac{\kappa T}{n}\right)^{i-1}\left(1-\frac{\kappa T}{2n}\right)\right\}$. Y_n を, $Y_n \equiv \frac{1}{\sqrt{n}}\sum_{i=1}^{n} b_i X_{n-i+1}$ とし, Y_n の 2,3,4 次キュムラントを,

$$\begin{aligned}
\mathrm{cum}\left(Y_n, Y_n\right) &\equiv c_{21}^{(n)} + n^{-1}c_{22}^{(n)} + o\left(n^{-1}\right), \\
\mathrm{cum}\left(Y_n, Y_n, Y_n\right) &\equiv n^{-1/2}c_{3}^{(n)} + o\left(n^{-1}\right), \\
\mathrm{cum}\left(Y_n, Y_n, Y_n, Y_n\right) &\equiv n^{-1}c_{4}^{(n)} + o\left(n^{-1}\right).
\end{aligned}$$

と表現する. $Y \equiv (c_{21}^{(n)})^{-1/2}Y_n$ の密度関数を Edgeworth 展開すると, 次の定理が示される.

定理　満期時点 T の債券価格は次のように評価できる.

$$\begin{aligned}
P(0,T) &= \exp\left[-\frac{\kappa\mu\Delta^2}{2}\left(\sum_{i=1}^{n} a_i\right) - \frac{\Delta}{2}\left\{1 + a_n(1-\kappa\Delta)\right\}r_0 + \frac{1}{2}Tc_{21}^{(n)}\right] \\
&\quad \times \left[1 - T^{3/2}\left(c_{21}^{(n)}\right)^{3/2}\frac{c_3^{(n)}}{6\sqrt{n}} + Tc_{21}^{(n)}\frac{c_{22}^{(n)}}{2n} + T^2\left(c_{21}^{(n)}\right)^2\frac{c_4^{(n)}}{24n} + T^3\left(c_{21}^{(n)}\right)^3\frac{\left(c_3^{(n)}\right)^2}{72n}\right] + o\left(n^{-1}\right),
\end{aligned}$$

また, スポットレート $R(0,T)$ は, $R(0,T) = -\frac{1}{T}\log P(0,T)$ より導かれる.

定理より, $\{X_j\}$ が i.i.d. 確率変数の場合, 次の系が示される.

系　$\{X_j\}$ が i.i.d. 確率変数で, 平均 0, 分散 σ^2, 歪度 γ_1, 尖度 γ_2 を持つ場合, 債券価格は次のように表現できる.

$$\begin{aligned}
P(0,T) &= \exp\left[-\frac{\kappa\mu\Delta^2}{2}\left(\sum_{i=1}^{n} a_i\right) - \frac{\Delta}{2}\left\{1 + a_n(1-\kappa\Delta)\right\}r_0 + \frac{1}{2}TG_{1,n}\sigma^2\right] \\
&\quad \times \left[1 - T^{3/2}\left(G_{1,n}\sigma^2\right)^{3/2}\frac{G_{2,n}\sigma^3\gamma_1}{6\sqrt{n}} + T^2\left(G_{1,n}\sigma^2\right)^2\frac{G_{3,n}\sigma^4(\gamma_2-3)}{24n} + T^3\left(G_{1,n}\sigma^2\right)^3\frac{\left(G_{2,n}\sigma^3\gamma_1\right)^2}{72n}\right] + o(n
\end{aligned}$$

ここで, $G_{1,n} = \frac{1}{n}\sum_{i=1}^{n} b_i^2$, $G_{2,n} = \frac{1}{n}\sum_{i=1}^{n} b_i^3$, $G_{3,n} = \frac{1}{n}\sum_{i=1}^{n} b_i^4$.

これより, イノベーションの歪度, 尖度の意味での非正規性が債券価格評価に与える影響を調べることが可能である. なお, $\{X_j\}$ が i.i.d. 正規確率変数で, 平均 0, 分散 σ^2 を持ち, 満期時点 T 固定で, $n \to \infty, \Delta \to 0$ とすると, 債券価格は Vasicek(1977) と一致する. また, 従属性による影響について, $\{X_j\}$ が GARCH(1,1) と Gaussian AR(1) に従う特殊な 2 つの場合について, 同様の債券価格評価の近似公式を導いた.

4　さいごに

本研究では, 離散型 Vasicek モデルのイノベーションが定常非正規過程に従うとして, 債券価格を高次まで評価した近似公式を導出した. また, イノベーションが skewed t 分布, GARCH(1,1), Gaussian AR(1) に従う場合のイールドカーブを図示し, 非正規性, 従属性が債券価格の評価に大きな影響を与えることを明らかにした.

参考文献

[1] Tamaki, K. and Taniguchi, M. (2007). Higher order asymptotic option valuation for non-Gaussian dependent returns, *Journal of Statistical Planning and Inference*, **137**, 1043-1058.

[2] Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics*, **5**, 177-188.

Second order properties of the distributions of test statistics derived under Whittle measure

早稲田大学政治経済学部　玉置 健一郎

時系列解析では，推定や検定に Whittle likelihood がよく用いられる．Whittle likelihood は計算が容易であり，正規過程の場合，対数尤度の近似になっており，最尤推定量と同等な性質をもつ推定量を構成できる．それ故，経済学等の様々な分野でも用いられている．また，近年，正規過程に対して Whittle measure の Contiguity が示された (Choudhuri et al., 2004)．これより，Whittle measure のもとで得られた漸近的な結果を真の分布のもとでの結果に変換することが出来る．

ここでは，Whittle measure のもとでの検定統計量の性質から，真の measure のもとでの性質を導くための高次漸近理論を考える．さらに，Whittle likelihood に基づく推定量の性質についても述べる．

まず

$$\lambda_j = 2\pi j/n,\ r_j = \exp(i\lambda_j),\ v_j = n^{-1/2}(r_j, r_j^2, \ldots, r_j^n)' \quad \text{for } j = 1, \ldots, n-1,$$

とし、$c_0 = n^{-1/2}(1, 1, \ldots, 1)'$ (n-vector),

$$c_j = (v_j + v_{n-j})/\sqrt{2},\ s_j = -i(v_j - v_{n-j})/\sqrt{2} \quad \text{for } j = 1, \ldots, [n/2],$$

とする．さらに，次の $n \times n$-直交行列 P_n を定義する．

$$P_n = \begin{cases} \left(c_0, c_1, s_1, \ldots, c_{n/2-1}, s_{n/2-1}, 2^{-1/2}c_{n/2}\right)', & \text{if } n \text{ is even}, \\ \left(c_0, c_1, s_1, \ldots, c_{[n/2]}, s_{[n/2]}\right)', & \text{if } n \text{ is odd}. \end{cases}$$

$X \sim N(0, \Gamma_\theta)$, $((\Gamma_\theta)_{jk} = \gamma_\theta(j-k))$, $f_\theta(\lambda) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_\theta(j)\exp(-ij\lambda)$ とする．このとき，$Z = P_n X \sim P_{n,\theta} = N(0, P_n\Gamma_\theta P_n')$ となる．また，Z の Whittle measure $Q_{n,\theta}$ は，$N\{0, D_n(f_\theta)\}$ である．ここに，

$$D_n(f) = \begin{cases} 2\pi\,\mathrm{diag}\{f(0), f(\lambda_1), f(\lambda_1), \ldots, f(\lambda_{n/2})\}, & \text{if } n \text{ is even}, \\ 2\pi\,\mathrm{diag}\{f(0), f(\lambda_1), f(\lambda_1), \ldots, f(\lambda_{[n/2]}), f(\lambda_{[n/2]})\}, & \text{if } n \text{ is odd}. \end{cases}$$

Whittle measure と真の measure の尤度関数をそれぞれ $\tilde{L}(\theta)$, $L(\theta)$ とし，$\tilde{L}(\theta)$ に基づく以下の検定統計量のクラスを考える．

$$\mathcal{S} = \Big\{t \,|\, t = I^{-1/2}\tilde{Z}_1 + n^{-1/2}(a_1\tilde{Z}_1^2 + a_2\tilde{Z}_1\tilde{Z}_2) + O(n^{-1}; Q_{n,\theta}),$$
$$\text{where } a_1 \text{ and } a_2 \text{ are nonrandom constants}\Big\},$$

ここに、

$$\tilde{Z}_1(\theta) = \frac{1}{n^{1/2}}\left\{\frac{\partial}{\partial\theta}\log\tilde{L}(\theta)\right\},$$
$$\tilde{Z}_2(\theta) = \frac{1}{n^{1/2}}\left[\frac{\partial^2}{\partial\theta^2}\log\tilde{L}(\theta) - E_{Q_{n,\theta}}\left\{\frac{\partial^2}{\partial\theta^2}\log\tilde{L}(\theta)\right\}\right].$$

$\tilde{u}_1 \in \mathcal{S}$, $\tilde{u}_2 = \log \tilde{L}(\theta + n^{-1/2}\varepsilon)/\tilde{L}(\theta)$ とし，また，$\tilde{U} = (\tilde{u}_1, \tilde{u}_2)$ とする．さらに，$Q_{n,\theta}$ のもとで次が成り立つとする．

$$
\begin{aligned}
\mathrm{cum}(\tilde{u}_1; Q_{n,\theta}) &= n^{-1/2}\tilde{\kappa}_1^{(1)} + O(n^{-1}),\\
\mathrm{cum}(\tilde{u}_2; Q_{n,\theta}) &= \kappa_2^{(0)} + n^{-1/2}\kappa_2^{(1)} + O(n^{-1}),\\
\mathrm{cum}(\tilde{u}_j, \tilde{u}_k; Q_{n,\theta}) &= \kappa_{jk}^{(1)} + n^{-1/2}\kappa_{jk}^{(2)} + O(n^{-1}),\\
\mathrm{cum}(\tilde{u}_j, \tilde{u}_k, \tilde{u}_l; Q_{n,\theta}) &= n^{-1/2}\kappa_{jkl}^{(1)} + O(n^{-1}), \quad j,k,l = 1,2.
\end{aligned}
$$

定理 1. $u_1 \in \mathcal{S}$, $u_2 = \log L(\theta + n^{-1/2}\varepsilon)/L(\theta)$ とし，また，$U = (u_1, u_2)$ とする．このとき $P_{n,\theta}$ のもとで次が成り立つ．

$$
\begin{aligned}
\mathrm{cum}\{u_1; P_{n,\theta}\} &= n^{-1/2}\kappa_1^{(1)} + o(n^{-1/2}),\\
\mathrm{cum}\{u_2; P_{n,\theta}\} &= \kappa_2^{(0)} + n^{-1/2}\kappa_2^{(1)} + O(n^{-1}),\\
\mathrm{cum}\{u_j, u_k; P_{n,\theta}\} &= \kappa_{jk}^{(1)} + n^{-1/2}\kappa_{jk}^{(2)} + O(n^{-1}),\\
\mathrm{cum}\{u_j, u_k, u_l; P_{n,\theta}\} &= n^{-1/2}\kappa_{jkl}^{(1)} + O(n^{-1}), \quad j,k,l = 1,2,
\end{aligned}
$$

ここに，

$$
\begin{aligned}
\kappa_1^{(1)} &= \tilde{\kappa}_1^{(1)} - \frac{1}{4\pi}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta} f_\theta(\lambda)\right\} b_\theta(\lambda) f_\theta(\lambda)^{-2} d\lambda,\\
b_\theta(\lambda) &= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} |j| \gamma_\theta(j) \exp(-ij\lambda).
\end{aligned}
$$

これより，検定統計量の Whittle measure の下での検出力から，真の measure の下での検出力が導かれる．また，$L(\theta)$ に基づく検定統計量との比較も行う．

References

CHOUDHURI, N., GHOSAL, S. & ROY, A. (2004). Contiguity of the Whittle measure for a Gaussian time series. *Biometrika* **91**, 211–218.