Estimation for Multivariate Stable Distributions with Generalized Empirical Likelihood Method

小方浩明 (首都大学東京)

1　はじめに

統計的解析において主要な役割を担っているのは明らかに正規分布である．しかしながら，経済データなどの多くの実データは，正規分布では表せないような非対称で裾の重いような分布に従うことが経験的に知られている．そのような特徴を表すことのできる分布の一つとして，安定分布があげられる．安定分布は一部を除き密度関数が陽に書けない，平均や分散を持たない場合がある，などの難点があり，パラメータの推定が困難である．それでも 1 次元の場合は多数の研究結果があるが，多次元の場合はまだまだ少ない．本報告では，多次元安定分布のパラメータ推定を，一般経験尤度法を用いて行うことを考える．

2　多次元安定分布

確率ベクトル $\boldsymbol{X} = (X_1, \dots, X_d)^\top$ が d 次元安定分布に従うとは，その特性関数が

$$\phi(\boldsymbol{\nu}) = E\big[\exp\{i\langle \boldsymbol{X}, \boldsymbol{\nu}\rangle\}\big] = \exp\left\{-\int_{S_d} \psi(\langle \boldsymbol{s}, \boldsymbol{\nu}\rangle)\Gamma(d\boldsymbol{s}) + i\langle \boldsymbol{\nu}, \boldsymbol{\mu}\rangle\right\}$$

のように表現できることである．ここに $S_d = \{\boldsymbol{s} : \| \boldsymbol{s} \| = 1\}$ は $\mathbb{R}^d$ における単位球面，記号 $\langle\cdot,\cdot\rangle$ は内積，Γ は S_d 上の有限スペクトル測度，$\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^\top \in \mathbb{R}^d$ は位置ベクトルを表す．また関数 $\psi(\cdot)$ は

$$\psi(u) = \begin{cases} |u|^\alpha\left(1 - i\,\mathrm{sign}(u)\tan\dfrac{\pi\alpha}{2}\right) & (\alpha \neq 1) \\ |u|\left(1 + i\dfrac{2}{\pi}\mathrm{sign}(u)\ln|u|\right) & (\alpha = 1), \end{cases}$$

で定義され，$\alpha \in (0, 2]$ は裾の重さを表すパラメータである．以後，この d 次元安定分布を $S^d(\alpha, \Gamma, \boldsymbol{\mu})$ で表す．本報告ではスペクトル測度を以下のような離散型スペクトル測度

$$\Gamma^* = \sum_{\ell=1}^{L} \gamma_\ell \delta_{\boldsymbol{s}_\ell}, \qquad \boldsymbol{s}_\ell \in S_d$$

で近似し，パラメータ $\boldsymbol{\theta} = (\alpha, \gamma_1, \dots, \gamma_L, \mu_1, \dots, \mu_d)^\top \in (0, 2] \times \mathbb{R}_+^L \times \mathbb{R}^d$ を同時に推定することを考える．

推定は，データから得られる経験特性関数と，理論的な特性関数をマッチングさせることによって行う．$\{\boldsymbol{X}_j\}_{j=1}^n$ を $S^d(\alpha, \Gamma^*, \boldsymbol{\mu})$ から得られる独立な観測系列とし，推定関数を

$$h(\boldsymbol{\nu}, \boldsymbol{X}_j, \boldsymbol{\theta}) = \exp\{i\langle \boldsymbol{\nu}, \boldsymbol{X}_j\rangle\} - \phi_{\boldsymbol{\theta}}(\boldsymbol{\nu})$$

で定義する．真のパラメータを $\boldsymbol{\theta}_0$ で表した場合明らかに $E[h(\boldsymbol{\nu}, \boldsymbol{X}_j, \boldsymbol{\theta}_0)] = 0$ が任意の周波数 $\boldsymbol{\nu}$ で成り立つ．適当な周波数 $\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_K \in \mathbb{R}^d$ を選び，さらに実部と虚部を分けて，推定関数を以下のように定義しなおす.

$$\boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}) = (\mathfrak{Re}[h(\boldsymbol{\nu}_1, \boldsymbol{X}_j, \boldsymbol{\theta})], \dots, \mathfrak{Re}[h(\boldsymbol{\nu}_K, \boldsymbol{X}_j, \boldsymbol{\theta})], \mathfrak{Im}[h(\boldsymbol{\nu}_1, \boldsymbol{X}_j, \boldsymbol{\theta})], \dots, \mathfrak{Im}[h(\boldsymbol{\nu}_K, \boldsymbol{X}_j, \boldsymbol{\theta})])^\top.$$

再び，$E[\boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}_0)] = \mathbf{0}$ が成り立つ．この推定関数を用い，次節に述べる一般経験尤度法によってパラメータ推定を考える．

3　一般経験尤度法

経験尤度法は Owen (1988) によって提唱された，非母数的な推測方法の一つである．$\boldsymbol{\theta}$ における経験尤度比は

$$\mathcal{R}(\boldsymbol{\theta}) = \max_{(p_1,\ldots,p_n)} \left\{ \prod_{j=1}^{n} np_j \;\middle|\; \sum_{j=1}^{n} p_j \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}) = \mathbf{0}, \sum_{j=1}^{n} p_j = 1, p_j \geq 0 \right\}, \tag{1}$$

で定義される．(1) における最大化はラグランジュの方法によってなされ，対数経験尤度比は

$$l(\boldsymbol{\theta}) = -\log \mathcal{R}(\boldsymbol{\theta}) = \sup_{\boldsymbol{\lambda}} \sum_{j=1}^{n} \log(1 + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}))$$

のように表される．ここに $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\theta}) \in \mathbb{R}^{2K}$ はラグランジュ乗数である．最大経験尤度推定量 $\hat{\boldsymbol{\theta}}_{\mathrm{EL}}$ は

$$\hat{\boldsymbol{\theta}}_{\mathrm{EL}} = \arg\min_{\boldsymbol{\theta}} \sup_{\boldsymbol{\lambda}} \sum_{j=1}^{n} \log(1 + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta})). \tag{2}$$

で定義される．Smith (1997, 2001) はこの推定量の一般化を考察した．$\rho(v)$ を，0 を含む開区間 $\mathcal{V} \subset \mathbb{R}$ 上で定義された凹関数とし，一般経験尤度推定量

$$\hat{\boldsymbol{\theta}}_{\mathrm{GEL}} = \arg\min_{\boldsymbol{\theta}} \sup_{\boldsymbol{\lambda} \in \hat{\Lambda}_n(\boldsymbol{\theta})} \sum_{j=1}^{n} \rho(\boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta})) \tag{3}$$

を与えた．ここに $\hat{\Lambda}_n(\boldsymbol{\theta}) = \{\boldsymbol{\lambda} : \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}) \in \mathcal{V}, j = 1, \ldots, n\}$ である．(2) は $\rho(v) = \log(1+v)$, $\mathcal{V} = (-1, \infty)$ とおいたときの特別な場合である．$\boldsymbol{\Omega} = E[\boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}) \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta})^\top]$，$\boldsymbol{G} = E[\partial \boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta}) / \partial \boldsymbol{\theta}]$ とおき，以下の仮定 (A1)–(A5) をおいた後，(3) の一致性並びに漸近正規性に関する定理を述べる．

(A1) $\boldsymbol{\theta}_0 \in \mathrm{int}(\Theta)$．ここに $\Theta = [\epsilon, 1-\epsilon] \cup [1+\epsilon, 2-\epsilon] \times [\epsilon, M]^L \times [-M, M]^d$ であり，$\epsilon > 0$ は十分小さく $M > 0$ は十分大きい数である．

(A2) $\boldsymbol{\theta}_0 \in \Theta$ は $E[\boldsymbol{g}(\boldsymbol{X}_j, \boldsymbol{\theta})] = \mathbf{0}$ の唯一の解である．

(A3) $\boldsymbol{\Omega}$ は正則である．

(A4) $\rho(v)$ は 0 の近傍で連続 2 回微分可能である．

(A5) $\mathrm{rank}(\boldsymbol{G}) = 1 + L + d$.

定理
(A1)–(A5) の下，$\hat{\boldsymbol{\theta}}_{\mathrm{GEL}} \xrightarrow{p} \boldsymbol{\theta}_0$ であり，

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{\mathrm{GEL}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}\big(\mathbf{0}, (\boldsymbol{G}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{G})^{-1}\big)$$

である．