# Strong Consistency of the Maximum Likelihood Estimator for Finite Mixtures of Location-Scale Distributions When Penalty is imposed on the Ratios of the Scale Parameters

*Naoya Hieda*

*2017-06-01*

基本的にはタイトルの英文の全訳。

## Abstract

有限混合な location-scale(位置-尺度) 分布において、パラメータに制約やペナルティがないと、尤度が閉じていないため、最尤推定値は存在しない。この問題を回避するために、ペナルティ付きの尤度を考える。ここでのペナルティとは、尺度パラメータとサンプルサイズの比の最小値の関数である。ペナルティ付き最尤推定値が強一致性を持つことがわかる。我々は、尺度パラメータのペナルティつき最尤推定値が、一致性を持つことを調べる。

## 1.Introduction

この論文では、単変量の location-scale(位置-尺度) 分布から得られる標本の、ペナルティ付き最尤推定値の強一致性 Ciuperca(2003) を証明する。この特別な場合として、Hathway(1985) の未解決問題を解く。
1969 年より、有限混合な location-scale(位置-尺度) 分布の尤度関数が閉じていないために、MLE(最尤推定値) は、存在しない。ということで、我々は簡単なケースとして、二つの正規分布による混合分布 $(\alpha_1\phi(x;\mu_1,\sigma_1)+\alpha_2\phi(x;\mu_2,\sigma_2)$ を考える。ここで、得られるサンプルは真の分布からの i.i.d.$X_1, X_2 \ldots, X_n$ を考える。
推定値が $\mu_1 = X_1$ で、$\sigma_1 \to 0$ となったとすると、尤度が $\infty$ に発散してしまう。尤度関数が閉じていないとはこういうことである。
この問題に対する単純な打開策は、正定値以下 (below by a positive constant) で $\sigma$ の推定値に最小値を設定することである。Render(1981) の 6 章より、パラメータの範囲が真のパラメータの範囲内ならば、最尤推定値が強一致性を持った上で、パラメータ空間が一部に制限される。
もう一つの方法が、ペナルティつき最尤推定である。しかし、もしペナルティが適切でない場合、尤度関数は閉じない。Ciuperca(2003) らは、ペナルティが分散をうまく調整できた場合を考え、ペナルティつき最尤推定が一致性を持つことを証明した。Ciuperca らが得たこの結果は、正規分布の混合分布のパラメータ推定において、ペナルティの仮定のチェックと実装を容易にしたため、非常に有益である。この論文では、彼らの一致性に関する結果を正規分布でない場合でペナルティがサンプルサイズに依存する場合に拡張していく。

正規分布の混合分布では、Hatheway(1985) が尤度の発散を避ける次の制約を考えた。Hatheway(1985)

$$\min_{m,m'} \frac{\sigma_m}{\sigma_{m'}} \geq b$$

これは、構成要素の分散の比率を最小にすることを制限している。彼は、真の分布が上記の式を満たすなら、最尤推定値の強一致性が保持されることを示した。直感的に、サンプルサイズが少ないときほど、尤度が発散

することを避けるために、より強い制約が必要である。なぜなら、分散が非常に小さい構成要素は、一つの観測値から大きな影響を受けるからである。ゆえに、一致性をもつ制約は、サンプルサイズが増えると弱めることができる。この導入によって、次の二つの疑問があがる

- 一致性を保守しながら、n を $\infty$ に増やすことで、b を 0 まで減らすことは可能か?
- 上記が可能だとして、b はどのくらいの勢いで 0 に近づくのか。

これらは、Hathaway(1985) が唱え、MacLachala & Peel(2000) が扱った未解決問題である。
この議題は、篩法 (sieve method) に密接に関連する。Grenander(1981) と Geman & Hwang(1982) を見てほしい。正規混合分布での、篩法に基づいている最尤推定値の収束率は、Genovese & Wasserma(2000) や Ghosal & van der Vaart(2001) で調べられている。Tanaka & Takemura(2006) では、location-scale(位置-尺度) の混合分布が、尺度パラメータが、($c_n = e^{-n^d}(0 < d < 1)$) で制限されているとき、最尤推定値が強一致性を持つことを証明した。しかしながら、構成要素の分散の比率の制限がある、上記の最初の疑問は解けていない。
この論文では、より一般的で統一された骨組み (framework) でこの問題を解く。混合正規分布では、尺度パラメータとサンプルサイズの比率を最小化するペナルティ付き尤度を考える。ペナルティの効果は、尺パラメータの比率の最小値が 0 に近づくほど、より強くなる。ペナルティがサンプルサイズに依存することに注意する。サンプルサイズの n が $\infty$ に近づくほど、ペナルティの効果が弱められる。定理 1 では、ペナルティ付きの最尤推定が一致性を保持すること示す。推論 1 として、Hathaway と MacLachala & Peel(2000) の問題の解法を定理 1 の特別な場合として解く。尺度パラメータのペナルティ付き最尤推定値の一致性を分析する。定理 2 でえられた結果は、Ciuperca の推論 1 の一般化である。
この論文を通して、真の分布は混合 location-scale(位置-尺度) 分布を仮定し、構成要素は分かっているものとする。
2 章で、記法と規則的な条件を提示する。主な結果を 3 章で述べる。4 章は証明にさく。5 章でこの論文をまとめる。

## 2.Preliminaries

### 2.1 Notation

M 混合な location-scale(位置-尺度) 分布は次のように定式化できる

$$f(x;\theta) = \sum_{m=1}^{M} \alpha_m f_m(x;\mu_m,\sigma_m)$$

$\alpha_m$ は重み係数であり、$\alpha_m \geq 0$かつ $\sum \alpha_m = 1$ の条件を持つ。また、密度 f が次を満たすと仮定する。

$$f_m(x,\mu_m,\sigma_m) = \frac{1}{\sigma_m} f_m(\frac{x-\mu_m}{\sigma_m};0,1) \quad 1 \leq m \leq M$$

ここで $\mu_m$ と $\sigma_m$ がそれぞれ、位置と尺度のパラメータである。今後、$(\alpha_1,\mu_1,\sigma_1,\ldots,\alpha_M,\mu_M,\sigma_M)$ を、$\theta$ と略し、$(\mu_m,\sigma_m)$ を $\theta_m$ と略す。また真のパラメータを $\theta_0$ とする。

$\Omega_m = \{(\mu_m\sigma_m)|\mu_m \in R, \sigma_m \in (0,\infty)\}$ と、第 m 構成要素のパラメータ空間を表す。すると、パラメータ空間全体 $\Theta$ が次のように再現できる。

$$\Theta = \left\{(\alpha_1,\ldots,\alpha_m)| \sum_{m=1}^{M} \alpha_m = 1, \alpha_m \geq 0\right\} \times \prod_{m=1}^{M} \Omega_m$$

$f(x;\theta_0)$ からのサンプル $X = (X_1,\ldots,X_n)$