

円柱データの確率分布と 隠れマルコフモデルについて

東京理科大学・工学部* 塩濱 敬之

Takayuki Shiohama

Department of Information and Computer Technology

Tokyo University of Science

高崎経済大学・経済学部† 宮田 庸一

Yoichi Miyata

Faculty of Economics

Takasaki City University of Economics

南山大学・理工学部‡ 阿部 俊弘

Toshihiro Abe

Department of Systems and Mathematical Sciences

Nanzan University

概要

本稿では、円柱上の非対称な確率分布とそれらを利用した隠れマルコフモデル (HMM) について紹介する。また、それら非対称な確率分布に対して識別可能性を調べ、時系列解析への応用も議論する。特に非対称な円柱分布の HMM における最尤推定量の一致性に関する結果を紹介する。

* 〒125-8585 東京都葛飾区新宿 6-3-1

† 〒370-0801 高崎市上並榎町 1300

‡ 〒466-8673 名古屋市昭和区山里町 18

1 はじめに

円周上に値を取る角度データは、イベント発生時刻を記録したデータや風向や方角等、自然科学や環境科学等様々な分野に表れる。2つの異なる角度データの確率分布を記述するためには、単位円を S で表せば、 $S \times S$ で定義されるトーラス上の確率分布を定義する必要がある。角度変数の同時分布をトーラス上で記述することができれば、回帰分析やマルコフ過程といった統計解析へのアプローチが可能になる。一方、重回帰モデルや高次マルコフモデルへの拡張は実軸上の統計解析と同じように容易に拡張することができない。このように、角度データの時系列解析は、その状態空間を表現する多様体上の確率モデルを用意する必要がある、実軸上の時系列解析とは異なるアプローチが必要になる。

角度変数と実数変数の同時分布を円柱 (シリンダー) 分布という。周期 2π の周期性をもつ角度データと正の実数データの同時分布は、自然科学や環境科学等においてよく現れる。例えば、気象学における風速と風向のデータが円柱データである。風速と風向の時系列モデルの構築は Breckling (1989) や Ailliot et al. (2015) 等で紹介され、主に、風向を表す確率変数 Θ と風速を表す確率変数 R の極座標変換を用いた 2 変量分布 $(X, Y) = (R \cos \Theta, R \sin \Theta) \in \mathbb{R}^2$ を用いて分析されている。また、その時系列モデルには、2 変量自己回帰 (VAR) モデルを用いて分析されることが多かった。

生態学における動物の移動行動の統計モデルには、移動距離と移動角度の観測データを用いた隠れマルコフモデルが利用されることが多い。Morales et al. (2004) は、移動距離の確率分布にワイブル分布、方向転換の角度の確率分布に巻き込みコーシー分布を用いた、移動距離の回転角度の独立性を仮定した隠れマルコフモデル (Hidden Markov Model; HMM) を提案している。動物移動行動に応じた k 個の状態を用意し、各状態推移にマルコフ連鎖を仮定した。このモデルは Multistate Random Walk (MRW) モデルという。その後、Lagona et al. (2015) は円柱分布を利用した HMM を用いて、風速と風向の時系列解析を紹介している。

本稿の構成は以下の通りである。2 章では円柱分布の紹介とその分布特性を紹介し、3 章では、隠れマルコフモデルを含む円柱データの時系列モデルを紹介する。4 章では、円柱分布の隠れマルコフモデルの最尤推定量の漸近特性を紹介する。最後に 5 章で、本稿のまとめと今後の研究の展望について述べる。

2 円柱上の確率モデル

本章では、円柱状に値を取る実変数と角度変数の 2 変量分布を紹介する。紹介するモデルは、Johnson and Wehrly (1978) による指数–von Mises 分布, Abe and Ley (2017) によるワイブル–von Mises 分布, Imoto et al. (2019) による一般化パレート–von Mises 分布である。また、これらの分布に正弦摂動を行うことによって非対称性をもたせた分布も紹介する。

最もよく知られている円柱分布は Johnson–Wehrly 分布 (Johnson and Wehrly (1978)) であろう。本稿では、指数–von Mises 分布と呼ぶことにする。この分布の密度関数は、指数分布と von Mises 分布を合わせた形をしている。密度関数は次のように定義される。

$$f_{\text{ExpVM}}(x, \theta; \sigma, \mu, \kappa) = \frac{\sqrt{1 - \kappa^2}}{2\pi\sigma} \exp \left[-\frac{x}{\sigma} \{1 - \kappa \cos(\theta - \mu)\} \right].$$

ここで、 $\mu \in [-\pi, \pi)$ は円周分布の位置パラメータ、 $\kappa \in [0, 1)$ は円周分布の集中度を表す役割と、角度変数と正の実変数の関連性を記述する役割をもつパラメータである。 σ は指数分布の尺度パラメータである。 $\kappa = 0$ のときに X と Θ の確率分布は独立になる。図 1 には、指数–von Mises 分布の密度関数と等高線図をプロットした。

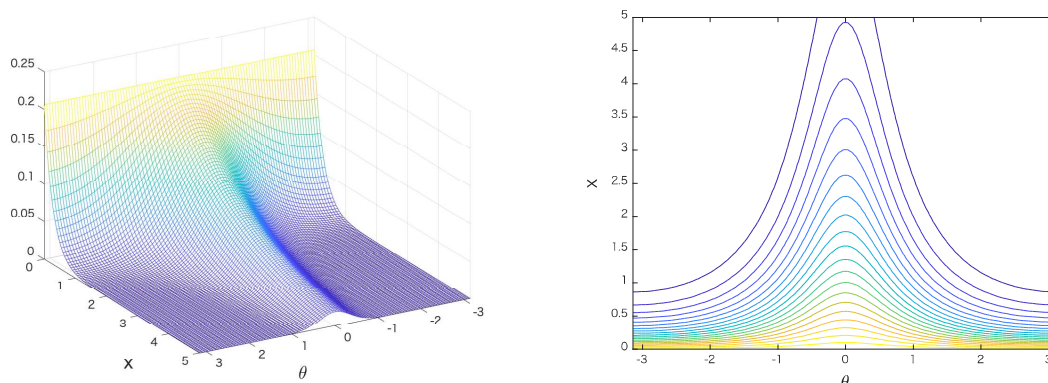


図 1 指数–von Mises 分布の密度関数のプロット (左) とその等高線図 (右)。パラメータは $\kappa = \tanh(1)$, $\sigma = 0.5$, $\mu = 0$ とした。

実変数が大きな値になるほどに、角度データの集中度が大きくなる傾向を指数–von Mises 分布が記述できていることが図 1 から明らかである。このことは、 X による Θ の条件付き密度関数が、位置母数 μ と集中度 $I_1(x\kappa/\sigma)/I_0(x\kappa/\sigma)$ で与えられる von Mises 分布になることから明らかである、ここで $I_p(x)$ は p 次の第 1 種変形 Bessel 関数である。こ

の分布の欠点は, Abe and Ley (2017) が指摘するように, 正の実変数が常に 0 でモードをもつこと, そのために様々な値をモードにもつような実データ解析に不向きである. そこで, Abe and Ley (2017) は正の実変数の確率分布にワイブル分布を用いた次のような円柱分布を提案した.

$$f_{\text{WeiVM}}(x, \theta; \sigma, \delta, \mu, \kappa) = \frac{\sqrt{1 - \kappa^2}}{2\pi\sigma\delta} \left(\frac{x}{\sigma}\right)^{1/\delta-1} \exp\left[-\left(\frac{x}{\sigma}\right)^{1/\delta} \{1 - \kappa \cos(\theta - \mu)\}\right].$$

ここで, $\sigma > 0, \delta > 0, -\pi \leq \mu < \pi, 0 \leq \kappa < 1$ はパラメータである.

図 2 には, ワイブル-von Mises 分布の密度関数と等高線図をプロットした. Θ を与えたときの X の条件付き分布にワイブル分布を用いることで, 様々な値のモードを表現できる.

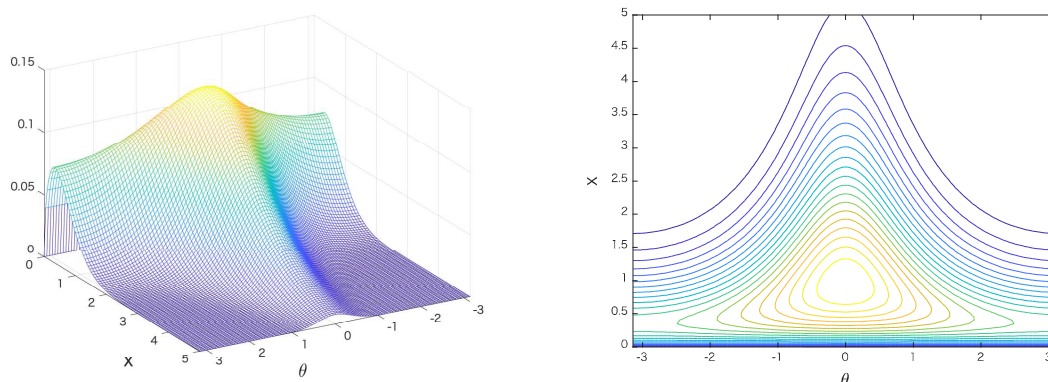


図 2 ワイブル-von Mises 分布の密度関数のプロット (左) とその等高線図 (右). モデルの定式化は (2) に従い, パラメータは $\tilde{\kappa} = 1, \alpha = 1.5, \beta = 0.5, \mu = 0$ とした.

Imoto et al. (2019) は, 実変数の裾の厚さを表現するために, 実数部分の条件付き確率分布を一般化パレート分布に拡張した次の円柱分布を提案している.

$$f_{\text{GPVM}}(x, \theta; \sigma, \delta, \tau, \mu, \kappa) = \frac{\sqrt{1 - \kappa^2}}{2\pi\sigma\delta} \left(\frac{x}{\sigma}\right)^{1/\delta-1} \left[1 + \frac{\tau}{\delta} \left(\frac{x}{\sigma}\right)^{1/\delta} \{1 - \kappa \cos(\theta - \mu)\}\right]^{-(\delta/\tau+1)}, \quad (1)$$

ここで, $\sigma > 0, \delta > 0$ は実数部分の尺度パラメータと形状パラメータであり, τ は実数部分の形状と裾の厚さを表すパラメータである. 図 3 には, 一般化パレート-von Mises 分布の密度関数と等高線図をプロットした. 一般化パレート-von Mises 分布に比べて, 裾の厚さを表現できることがわかる.

ここに紹介した 3 つの円柱分布の関係は, ワイブル-von Mises 分布において $\delta = 1$ とす

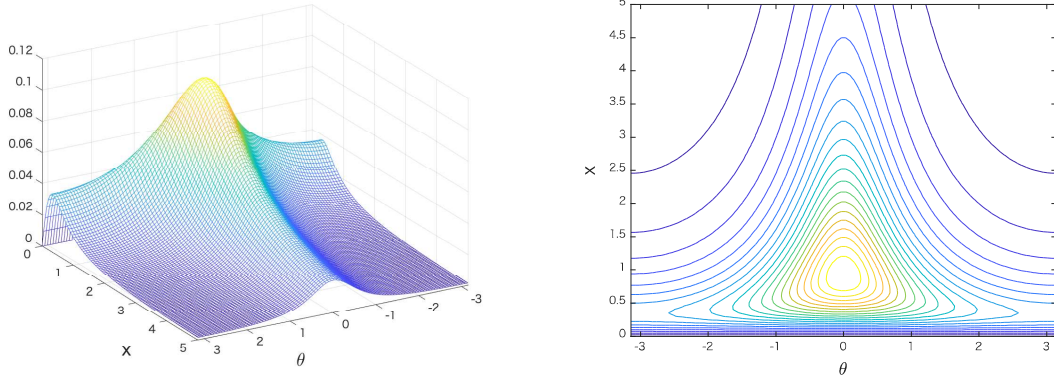


図3 一般化パレート-von Mises 分布の密度関数のプロット (左) とその等高線図 (右). パラメータは $\kappa = \tanh(1)$, $\sigma = 1$, $\tau = 1.5$, $\delta = 0.5$, $\mu = 0$ とした.

れば指数-von Mises 分布に帰着するし、一般化パレート分布-von Mises 分布においては、 $\tau \rightarrow 0$ とすればワイブル-von Mises 分布に帰着する。

Abe and Ley (2017) や Imoto et al. (2019) では、上述の円柱分布の角度変数の確率分布に正弦摂動によって非対称性を取り入れた円柱分布の正弦摂動による確率分布を提案している。今、円柱分布のパラメータを $\boldsymbol{\eta} \in \mathbf{H}$ とし、 θ に関して対称な任意の円柱分布の密度関数を $f_{\text{Cyl}}(x, \theta; \boldsymbol{\eta})$ と表すと、正弦摂動による円柱分布は、非対称性を表すパラメータ $\lambda \in [-1, 1]$ を用いて次のように定義される。

$$f_{\text{SSCyl}}(x, \theta; \lambda, \boldsymbol{\eta}) = \{1 + \lambda \sin(\theta - \mu)\} f_{\text{Cyl}}(x, \theta; \boldsymbol{\eta}).$$

正弦摂動による円柱分布は、ワイブル-von Mises 分布の場合も、一般化パレート分布-von Mises 分布の場合においても、 Θ と X の周辺密度関数と条件付き密度関数が陽に与えられることである。 Θ の周辺分布は次の巻き込みコーシー分布の正弦摂動の確率分布 (Sine-skewed Wrapped Cauchy (SSWC) 分布) に従う。

$$f_{\text{SSWC}}(\theta; \mu, \kappa, \lambda) = (1 + \lambda \sin(\theta - \mu)) \frac{\sqrt{1 - \kappa^2}}{2\pi \{1 - \kappa \cos(\theta - \mu)\}}.$$

$\lambda = 0$ の場合、対称な巻き込みコーシー分布になり、その集中度は $\rho_1 = \kappa / (1 + \sqrt{1 - \kappa^2})$ で与えられる。

指数-von Mises 分布とワイブル-von Mises 分布はそれぞれ次のような表現も可能である。

$$f_{\text{ExpVM}}(x, \theta; \beta, \mu, \tilde{\kappa}) = \frac{\beta}{2\pi \cosh(\tilde{\kappa})} \exp\{-\beta x (1 - \tanh(\tilde{\kappa}) \cos(\theta - \mu))\},$$

ここで, $\beta > 0$ は実変数の指数分布の形状パラメータであり, $\tilde{\kappa} > 0$ は, 円周分布の集中度パラメータである.

$$f_{\text{WeiVM}}(x, \theta; \alpha, \beta, \mu, \tilde{\kappa}) = \frac{\alpha\beta^\alpha}{2\pi \cosh(\tilde{\kappa})} x^{\alpha-1} \exp\{-(\beta x)^\alpha(1 - \tanh(\tilde{\kappa}) \cos(\theta - \mu))\}, \quad (2)$$

$\alpha > 0, \beta > 0$ は実変数のワイブル分布の形状パラメータと尺度パラメータである.

2.1 識別可能性

次に円柱分布の正弦摂動の確率分布の識別可能性について考える. 今, パラメータベクトルを $\gamma = (\lambda, \boldsymbol{\eta}^T)^T \in \boldsymbol{\Gamma}$ とする. ここで $\boldsymbol{\Gamma}$ はパラメータ空間である. パラメータ $\gamma \in \boldsymbol{\Gamma}$ をもつ確率密度関数 $f_{\text{SSCyI}}(x, \theta|\gamma)$ の族を考える. このとき, $f_{\text{SSCyI}}(x, \theta|\gamma)$ は $\boldsymbol{\Gamma}$ において識別可能であるとは, 任意の $\gamma_1, \gamma_2 \in \boldsymbol{\Gamma}$ に対して, $\gamma_1 \neq \gamma_2$ のときに $f_{\text{SSCyI}}(x, \theta|\gamma_1) \neq f_{\text{SSCyI}}(x, \theta|\gamma_2)$ が成り立つことをいう. SSWC 分布のパラメータベクトルを $\gamma_{\text{SSWC}} = (\mu, \kappa, \lambda)^T$ とし, そのパラメータ空間を

$$\boldsymbol{\Gamma}_{\text{SSWC}} = \{(\mu, \kappa, \lambda) \mid -\pi \leq \mu < \pi, \kappa > 0, -1 \leq \lambda \leq 1\}$$

と表し, ワイブル-von Mises 分布の正弦摂動の確率分布のパラメータ空間を次のように表す.

$$\boldsymbol{\Gamma}_{\text{WeiSSvM}} = \gamma_{\text{SSWC}} \times \{(\alpha, \beta) \mid \alpha > 0, \beta > 0\}.$$

また, 一般化パレート分布と von Mises 分布の正弦摂動の確率分布のパラメータ空間を次のように表す.

$$\boldsymbol{\Gamma}_{\text{GPSSvM}} = \gamma_{\text{SSWC}} \times \{(\delta, \tau, \sigma) \mid \delta > 0, \tau > 0, \sigma > 0\}.$$

このとき, 以下の結果が成り立つ.

定理 1 (Miyata et al.(2020))

- (i) ワイブル-von Mises 分布の正弦摂動の密度関数は, パラメータ空間 $\boldsymbol{\Gamma}_{\text{WeiSSvM}}$ において識別可能である.
- (ii) 一般化パレート分布と von Mises 分布の正弦摂動の密度関数は, パラメータ空間 $\boldsymbol{\Gamma}_{\text{GPSSvM}}$ において識別可能である.

これらの定理は Miyata et al. (2020) において示されている. Miyata et al. (2019) では, 正弦摂動の円周分布の混合モデルに対する識別可能性を議論しているが, 円柱分布の混合モデルの識別可能性に関しては明らかにされていない.

3 円柱データの時系列解析

時刻 t において円柱上に値を取る確率変数 $\mathbf{X}_t = (X_t, \Theta_t)^T$, $X_t \in [0, +\infty)$, $\Theta_t \in [-\pi, \pi)$ の時系列 $\mathbf{X}_{1:T} = (\mathbf{X}_1, \dots, \mathbf{X}_T)$ を考える.

まずは、隠れマルコフモデルを定式化する. 今, K 個の状態を推移するマルコフ連鎖に従って, X_t が生成されるとする. マルコフ連鎖は離散時間に値を取る多項過程とする. $\mathbf{Z}_t = (z_{t,1}, \dots, z_{t,K})^T$ は, 時刻 t において, \mathbf{X}_t が k 番目の状態から得られたときに, $z_{t,k} = 1$, それ以外の場合は $z_{t,k} = 0$ をとる潜在過程とする. 観測される時系列は, 潜在状態の確率変数 \mathbf{Z}_t を用いて次のように表す

$$\mathbf{X}_t | (\mathbf{Z}_t = \mathbf{e}_k) \sim k\text{th state with density } f_{\text{SSCLy}}(\mathbf{X}_t | \boldsymbol{\eta}_k),$$

ここで, $\boldsymbol{\eta}_k$ は, 状態 k における円柱分布の正弦摂動の確率密度関数のパラメータベクトルであり, \mathbf{e}_k は第 k 要素が 1, それ以外の要素が 0 であるベクトルである. また, マルコフ連鎖を次のように表現する

$$p(\mathbf{Z}_t = \mathbf{e}_k | \mathbf{Z}_{t-1} = \mathbf{e}_j) = a_{jk},$$

ここで $a_{jk} \in [0, 1)$ は状態遷移行列 \mathbf{A} の (j, k) 要素であり, $\sum_{k=1}^K a_{jk} = 1$ を満たす.

初期状態の確率分布を表すベクトルを $\mathbf{q} = (q_1, \dots, q_K)^T$, $q_k = p(z_{1,k} = 1)$, $k = 1, \dots, K$, $\sum_k q_k = 1$ とすると, 状態遷移の同時分布は次のように表すことができる

$$p(\mathbf{Z}_{1:T}; \mathbf{q}, \mathbf{A}) = p(\mathbf{Z}_1 | \mathbf{q}) \prod_{t=2}^T p(\mathbf{Z}_t | \mathbf{Z}_{t-1}; \mathbf{A}) = \prod_{k=1}^K q_k^{z_{1,k}} \prod_{t=2}^T \prod_{j=1}^K \prod_{k=1}^K a_{jk}^{z_{t-1,j} z_{t,k}}. \quad (3)$$

観測変数 \mathbf{X}_t と潜在変数 \mathbf{Z}_t は条件付き独立であるとする, 状態過程を与えたときの観測値の条件付き分布は次のように表すことができる.

$$f(\mathbf{X}_{1:T} | \mathbf{Z}_{1:T}; \boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_K) = \prod_{t=1}^T \prod_{k=1}^K f_{\text{SSCLy}}(x_t, \theta_t; \boldsymbol{\eta}_k)^{z_{t,k}}. \quad (4)$$

正弦摂動の円柱分布のすべての状態におけるパラメータベクトルを $\boldsymbol{\eta}_{1:K} = (\boldsymbol{\eta}_1^T, \dots, \boldsymbol{\eta}_K^T)^T$ と表すと, 観測値と潜在変数の同時分布 (完全尤度) はつぎのように表すこ

とができる.

$$\begin{aligned}
p(\mathbf{X}_{1:T}, \mathbf{z}_{1:T} | \mathbf{q}, \mathbf{A}, \boldsymbol{\eta}_{1:K}) &= p(\mathbf{z}_1 | \mathbf{q}) \left[\prod_{t=2}^T p(\mathbf{z}_t | \mathbf{z}_{t-1}; \mathbf{A}) \right] \prod_{t=1}^T f(\mathbf{X}_t | \mathbf{z}_t; \boldsymbol{\eta}_{1:K}) \\
&= \prod_{k=1}^K q_k^{z_{1,k}} \left[\prod_{t=2}^T \prod_{j=1}^K \prod_{k=1}^K a_{jk}^{z_{t-1,j} z_{t,k}} \right] \prod_{t=1}^T \prod_{k=1}^K f_{\text{SSCly}}(x_t, \theta_t; \boldsymbol{\eta}_k)^{z_t} \quad (5)
\end{aligned}$$

また, 観測変数の周辺分布は, 次のようになる.

$$p(\mathbf{X}_{1:T} | \mathbf{q}, \mathbf{A}, \boldsymbol{\eta}_{1:K}) = \sum_{\mathbf{z}_{1:T}} p(\mathbf{X}_{1:T}, \mathbf{z}_{1:T} | \mathbf{q}, \mathbf{A}, \boldsymbol{\eta}_{1:K}). \quad (6)$$

一般に, 潜在過程と観測過程の完全尤度の最大化は困難であるため, 最尤法による未知母数の推定は次の EM アルゴリズムを用いる. (5) 式で表される完全尤度から, 対数完全尤度は次のように表すことができる.

$$\begin{aligned}
&\ln L(\boldsymbol{\eta}_{1:K}; \mathbf{X}_{1:T}, \mathbf{Z}_{1:T}, \mathbf{q}) \\
&= \ln p(\mathbf{Z}_1 | \mathbf{q}) + \sum_{t=2}^T \ln p(\mathbf{Z}_t | \mathbf{Z}_{t-1}; \mathbf{A}) + \sum_{t=1}^T \ln f_{\text{SSCly}}(\mathbf{X}_t | \mathbf{Z}_t; \boldsymbol{\eta}_{1:K}). \quad (7)
\end{aligned}$$

潜在過程 $\mathbf{Z}_{1:T}$ は観測できないため, (7) 式の最大化によって $\boldsymbol{\eta}_{1:K}$ と \mathbf{A} の最尤推定量を得ることができない. EM アルゴリズムは, 条件付き完全対数尤度を最大化しながら, 最尤推定量を得る繰り返しアルゴリズムである. $j-1$ における条件を与えたときに, 条件付き完全対数尤度は次式で与えられる.

$$\begin{aligned}
&Q(\mathbf{q}, \mathbf{A}, \boldsymbol{\eta}_{1:K} | \mathbf{q}^{(j-1)}, \mathbf{A}^{(j-1)}, \boldsymbol{\eta}_{1:K}^{(j-1)}) \\
&= \sum_{k=1}^K \mathbb{E}(z_{1,k} | \mathbf{X}_{1:N}, \mathbf{q}^{(j-1)}, \mathbf{A}^{(j-1)}, \boldsymbol{\eta}_{1:K}^{(j-1)}) \ln q_k \\
&\quad + \sum_{t=2}^T \sum_{j=1}^K \sum_{k=1}^K \mathbb{E}(z_{t-1,j} z_{t,k} | \mathbf{X}_{1:T}, \mathbf{q}^{(j-1)}, \mathbf{A}^{(j-1)}, \boldsymbol{\eta}_{1:K}^{(j-1)}) \ln a_{jk} \\
&\quad + \sum_{t=1}^T \sum_{k=1}^K \mathbb{E}(z_{t,k} | \mathbf{X}_{1:T}, \mathbf{q}^{(j-1)}, \mathbf{A}^{(j-1)}, \boldsymbol{\eta}_{1:K}^{(j-1)}) \ln f_{\text{SSCly}}(x_t, \theta_t; \boldsymbol{\eta}_k) \\
&= \sum_{k=1}^K \gamma(z_{1,k}) \ln q_k + \sum_{t=2}^T \sum_{j=1}^K \sum_{k=1}^K \xi(z_{t-1,j}, z_{t,k}) \ln a_{jk} + \sum_{t=1}^T \sum_{k=1}^K \gamma(z_{t,k}) \ln f(x_t, \theta_t; \boldsymbol{\eta}_k) \quad (8)
\end{aligned}$$

ここで

$$\begin{aligned}\gamma(z_{t,k}) &= \mathbb{E}(z_{t,k}) = \sum_{\mathbf{z}_{1:T}} \gamma(\mathbf{z}_t) z_{t,k}, \\ \xi(z_{t-1,j}, z_{t,k}) &= \mathbb{E}(z_{t-1,j} z_{t,k}) = \sum_{\mathbf{z}_{1:T}} \xi(\mathbf{z}_{t-1}, \mathbf{z}_t) z_{t-1,j} z_{t,k}.\end{aligned}$$

(8) 式を $\boldsymbol{\eta}_{1:K}$ に関して最大化し, $Q_k(\boldsymbol{\eta}, \boldsymbol{\eta}^{(j-1)}) = \sum_{k=1}^K \gamma(z_{t,k}) \ln f_{\text{SSCly}}(x_t, \theta_t; \boldsymbol{\eta}_k)$ と置けば, j 回目の繰り返しにおける E ステップによる状態変数の更新は次のようになる.

$$q_k^{(j)} = \frac{\gamma(z_{1,k})}{\sum_{k=1}^K \gamma(z_{1,k})}, \quad a_{jk}^{(j)} = \frac{\sum_{t=2}^T \xi(z_{t-1,j}, z_{t,k})}{\sum_{k=1}^K \sum_{t=2}^T \xi(z_{t-1,j}, z_{t,k})}$$

また, M ステップにおける推定値はつぎのようになる.

$$\boldsymbol{\eta}_k^{(j)} := \arg \max_{\boldsymbol{\eta}_k} Q_j(\boldsymbol{\eta}_k, \boldsymbol{\eta}_k^{(j-1)}).$$

この E ステップと M ステップを, 収束条件が満たされるまで繰り返す.

4 最尤推定量の漸近的性質

円柱分布の正弦摂動の確率分布に対する HMM の最尤推定量の一致性から調べる. $\boldsymbol{\eta}_k = (\mu_k, \lambda_k, \boldsymbol{\gamma}_k^T)^T$ は状態 k におけるパラメータベクトルであった. これより推定すべきパラメータは $\boldsymbol{\eta} = (a_{11}, \dots, a_{KK}, \boldsymbol{\eta}_1^T, \dots, \boldsymbol{\eta}_K^T)^T$ である. HMM モデルのパラメータ空間を次のように表す.

$$\mathbf{H} = \left\{ \boldsymbol{\eta} \left| \sum_{k=1}^K a_{jk} = 1, \quad a_{jk} > 0, \quad -\pi \leq \mu_k < \pi, \quad -1 \leq \lambda_k \leq 1, \quad \boldsymbol{\gamma}_k \in \boldsymbol{\Gamma}, \quad (k = 1, \dots, K) \right. \right\}.$$

$\boldsymbol{\eta}^0 = (a_{11}^0, \dots, a_{KK}^0, \underbrace{\mu_1^0, \lambda_1^0, \boldsymbol{\gamma}_1^{0T}}_{\boldsymbol{\eta}_1^0}, \dots, \underbrace{\mu_K^0, \lambda_K^0, \boldsymbol{\gamma}_K^{0T}}_{\boldsymbol{\eta}_K^0})^T$ と表す. また, 最尤推定量 $\hat{\boldsymbol{\eta}}_{ML}$ は次のように定義された.

$$\hat{\boldsymbol{\eta}}_{ML} = \arg \max_{\boldsymbol{\eta}} L_{z_0, T}(\boldsymbol{\eta})$$

$$L_{z_0, T}(\boldsymbol{\eta}) = \sum_{\mathbf{z}_{1:T}} q(z_0) f_{\text{SSCly}}(\theta_0, x_0; \boldsymbol{\eta}_0) \prod_{t=1}^T a_{z_{t-1}, z_t} \prod_{k=1}^K f_{\text{SSCly}}(\theta_t, x_t; \boldsymbol{\eta}_k)^{z_t}.$$

今, $\psi(\mu) = (\cos \mu, \sin \mu)^T$ と定義し, 関数 $\Pi(y) = (1 + y)/2$ と連続写像 $h_{\eta_k} : [-1, 1] \rightarrow \mathbb{R}^+$, $g_{\eta_k} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ を定義すると, 正規化定数 C_{η_k} を用いて, 状態 k の正弦摂動による円

柱分布の密度関数は次のように表すことができる。

$$\begin{aligned} f_{SSCly}^\psi(x, \theta; \boldsymbol{\eta}_k) &= (1 + \lambda_k \sin(\theta - \mu_k)) f_{Cly}(x, \theta; \boldsymbol{\eta}_k) \\ &= 2C_{\eta_k} g_{\eta_k}(x) h_{\eta_k}(\psi(\theta)^T \psi(\mu_k)) \Pi \left\{ \lambda_k \psi(\mu_k)^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \psi(\theta) \right\}, \end{aligned}$$

ここで $\psi(\theta) = (\cos \theta, \sin \theta)^T$ また $\boldsymbol{\mu}_k = (\cos \mu_k, \sin \mu_k)^T$ である。対応する HMM の完備なパラメータ空間は次のようになる。

$$\mathbf{H}^\psi = \left\{ \boldsymbol{\eta} \left| \sum_{k=1}^K a_{jk} = 1, \quad a_{jk} > 0, \quad \boldsymbol{\mu}_k \in \mathcal{S}^1, \quad -1 \leq \lambda_k \leq 1, \quad \gamma_k \in \Gamma, \quad (k = 1, \dots, K) \right. \right\}.$$

次のような \mathbf{H} から \mathbf{H}^ψ への全単射の写像を考える。

$$\Psi \{(\mu_1, \gamma_1^T, \dots, \mu_K, \gamma_K^T)^T\} = (\psi(\mu_1)^T, \gamma_1^T, \dots, \psi(\mu_K)^T, \gamma_K^T)^T.$$

また, $\boldsymbol{\eta}_\psi^0 := \Psi(\boldsymbol{\eta}^0)$ は \mathbf{H}^ψ における真のパラメータベクトルとする。最尤推定量の一致性のためには次の仮定が必要になる。

- 仮定 1** [A1] (コンパクト性) Γ は \mathbb{R}^g におけるコンパクト部分集合とする。任意の $\theta \in [-\pi, \pi)$ とすべての $k \in \{1, \dots, K\}$ に対して, $\sup_{\eta_k \in \mathcal{S}^1 \times \Gamma} f_{SSCly}^\psi(x, \theta | \gamma_k) < \infty$ が成り立つ, またパラメータベクトルの真値 $\boldsymbol{\eta}^0$ は \mathbf{H}^ψ の元とする。
- [A2] (有界性) 任意の $\boldsymbol{\eta}_k^* \in \mathcal{S}^1 \times [-1, 1] \times \Gamma$ に対して $\boldsymbol{\eta}_j^*$ の近傍 $U(\boldsymbol{\eta}_j^*)$ が存在して, $E_{\boldsymbol{\eta}_k^0} \left[\log \sup_{\boldsymbol{\eta}_k \in U(\boldsymbol{\eta}_k^*)} \left\{ \max(f_{SSCly}^\psi(X, \Theta | \boldsymbol{\eta}_k), 1) \right\} \right] < \infty, (k = 1, \dots, K)$ が成り立つ。
- [A3] (クロスエントロピーは下に有界) すべての $\theta \in [-\pi, \pi)$ と $x \in \mathbb{R}^+$ に対して $E_{\boldsymbol{\eta}_k^0} \left\{ \log f_{SSCly}^\psi(X, \Theta | \boldsymbol{\eta}_k) \right\} > -\infty$.
- [A4] (連続性) すべての $\theta \in [-\pi, \pi)$ と $x \in \mathbb{R}^+$ に対して $f_{SSCly}^\psi(\theta, x | \boldsymbol{\eta}_k)$ は $\boldsymbol{\eta}_k$ に対して連続。
- [A5] $\mathcal{N}(\boldsymbol{\eta}_k) := \{\theta | f^\psi(x, \theta | \boldsymbol{\eta}_k) = 0\}$ とする。任意の $\boldsymbol{\eta}_k \in \mathcal{S}^1 \times [-1, 1] \times \Gamma$ に対して, $P_0\{\mathcal{N}(\boldsymbol{\eta}_k)\} = 0$ である, ここで P_0 真の確率密度関数 $f_{SSCly}^\psi(x, \theta | \boldsymbol{\eta}_k^0)$ に対応する確率測度である。

次の仮定は, マルコフ連鎖が定常でエルゴードであるための条件である。

仮定 2 マルコフ連鎖の遷移確率行列 $\mathbf{A} = [a_{jk}]_{j,k=1, \dots, n}$ は非再帰的で非周期的である。

また, 円柱分布の正弦摂動の確率分布の K 個の混合モデルは識別可能であることを仮定する。

仮定 3 次の本質的な識別可能性 (generic identifiability) が成立する. すなわち

$$\sum_{k=1}^K \alpha_k f_{\text{SSCl}_y}(x, \theta; \boldsymbol{\eta}_k) = \sum_{k=1}^K \alpha'_k f_{\text{SSCl}_y}(x, \theta; \boldsymbol{\eta}'_k)$$

であるならば, $k = 1, \dots, K$ に対して $\alpha_k = \alpha'_k$ また $\boldsymbol{\eta}_k = \boldsymbol{\eta}'_k$ が成り立つ.

Holzmann et al. (2006) は, HMM において混合分布モデルの識別可能性がどのように適用されるかに関して説明している. 円周分布の有限混合分布の識別可能性に関しては, Fraser et al. (1981) や Holzmann et al. (2004) において議論されているが, 円柱分布の混合分布に関する識別可能性はまだ明らかにされていない. 識別可能性が証明されると, Leroux (1992) や Bickel et al. (1998) で示されるように, 最尤推定量 (MLE) の一致性と漸近正規性を示すことができる.

HMM は潜在状態の並べ替えに関して識別可能でないため, 次のような真のモデルによって定式化されるパラメータ集合を考える.

$$\mathbf{H}^\psi(\boldsymbol{\eta}^0) = \left\{ \boldsymbol{\eta} \in \mathbf{H}^\psi \left| \sum_{k=1}^K \pi_{\boldsymbol{\eta}, k} f_{\text{SSCl}_y}(\theta, x; \boldsymbol{\eta}_k) = \sum_{k=1}^K \pi_{\boldsymbol{\eta}^0, k} f_{\text{SSCl}_y}(\theta, x; \boldsymbol{\eta}_k^0) \right. \right\}$$

このとき次の最尤推定量の一致性が成り立つ.

定理 2 仮定 1–3 の下で, 最尤推定量 $\hat{\boldsymbol{\eta}}_{ML}$ の強一致性が成り立つ.

$$P_0 \left[\lim_{n \rightarrow \infty} \text{dis}(\hat{\boldsymbol{\eta}}_{ML}, \mathbf{H}^\psi(\boldsymbol{\eta}^0)) = 0 \right] = 1,$$

ただし, 2つの集合に対する距離は次のように定める.

$$\text{dis}(\Omega_1, \Omega_2) = \inf_{x \in \Omega_1} \inf_{y \in \Omega_2} \|x - y\|.$$

ここで, 仮定 2 は, Miyata et al. (2019) における円周分布の正弦振動の混合分布における MLE の一致性の条件を円柱分布に拡張した仮定に相当する. 一致性の証明は, Cappé et al. (2005) の定理 130 や, Douc et al. (2014) の Proposition 13.13 および 13.14 と同様にして示すことができる.

5 まとめ

本論文では, 既存の円柱分布の紹介をし, 非対称化について述べた. また, その円柱分布のクラスにおける識別可能性の現状について議論し, 円柱データに関する時系列解析の手

法について紹介した。さらに、最尤推定量の一致性の定理と証明の概略を述べた。非対称な円柱分布の HMM において最尤推定量を求めるための EM アルゴリズムには、Miyata et al. (2019) による非対称パラメータの取り扱いの工夫の余地や、加速法 (Louis, 1982) などの計算時間を改善する取り組みが必要である。また、ベイズ流のアプローチについて考える必要がある。

今後の課題として、円柱分布に対する混合モデルの識別可能性が挙げられる。また、Abe et al. (2020) で議論された尺度変換によって非対称性をもたせた確率分布に関して、円柱分布への拡張の余地や HMM モデルのパラメータ推定問題について分析する必要がある。

謝辞

本研究集会にて講演の機会を与えていただきました筑波大学の小池健一先生に厚くお礼申し上げます。また本研究は、科学研究費 (課題番号:18K01706, 19K11863, 19K11869) の助成を受けている。

参考文献

- Abe, T. and C. Ley (2017). A tractable, parsimonious and flexible model for cylindrical data, with applications. *Econometrics and Statistics* 4, 91–104.
- Abe, T., Y. Miyata, and T. Shiohama (2020). Bayesian estimation for mode and anti-mode preserving circular distributions. *Submitted*.
- Ailliot, P., J. Bessac, V. Monbet, and F. Pene (2015). Non-homogeneous hidden Markov-switching models for wind time series. *Journal of Statistical Planning and Inference* 160, 75–88.
- Bickel, P. J., Y. Ritov, and T. Ryden (1998). Asymptotic normality of the maximum-likelihood estimator for general hidden markov models. *The Annals of Statistics* 26(4), 1614–1635.
- Breckling, J. (1989). *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, Volume 61. Springer Science & Business Media.
- Cappé, O., E. Moulines, and T. Rydén (2005). *Inference in Hidden Markov Models*. Springer Science & Business Media.
- Douc, R., E. Moulines, and D. Stoffer (2014). *Nonlinear Time Series: Theory, Methods, and Applications with R Examples*. Chapman and Hall/CRC.

- Fraser, M., Y. S. Hsu, and J. Walker (1981). Identifiability of finite mixtures of von Mises distributions. *The Annals of Statistics* 9, 1130–1131.
- Holzmann, H., A. Munk, and B. Stratmann (2004). Identifiability of finite mixtures with applications to circular distributions. *Sankhyā: The Indian Journal of Statistics* 66, 440–449.
- Holzmann, H., A. Munk, M. Suster, and W. Zucchini (2006). Hidden Markov models for circular and linear-circular time series. *Environmental and Ecological Statistics* 13(3), 325–347.
- Imoto, T., K. Shimizu, and T. Abe (2019). A cylindrical distribution with heavy-tailed linear part. *Japanese Journal of Statistics and Data Science* 2(1), 129–154.
- Johnson, R. A. and T. E. Wehrly (1978). Some angular-linear distributions and related regression models. *Journal of the American Statistical Association* 73(363), 602–606.
- Lagona, F., M. Picone, and A. Maruotti (2015). A hidden Markov model for the analysis of cylindrical time series. *Environmetrics* 26(8), 534–544.
- Leroux, B. G. (1992). Maximum-likelihood estimation for hidden Markov models. *Stochastic Processes and Their Applications* 40(1), 127–143.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 44(2), 226–233.
- Miyata, Y., T. Shiohama, and T. Abe (2019). Estimation of finite mixture models of skew-symmetric circular distributions. *Metrika*, <https://doi.org/10.1007/s00184-019-00756-z>.
- Miyata, Y., T. Shiohama, and T. Abe (2020). Identifiability of asymmetric circular and cylindrical distributions. *arXiv:1908.09114*.
- Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* 85(9), 2436–2445.