

角度データのための混合効果モデル

加藤 昇吾 数理・推論研究系 准教授

はじめに

背景

様々な学術分野において、角度 $[-\pi, \pi)$ として表される観測が存在する。
(例. 風向, 動物の移動方向, 事件が起こる時刻, など.)

角度の観測を含むデータに対して、線形データのための回帰モデルをそのまま当てはめると、**誤った解析結果**を得てしまうことある。

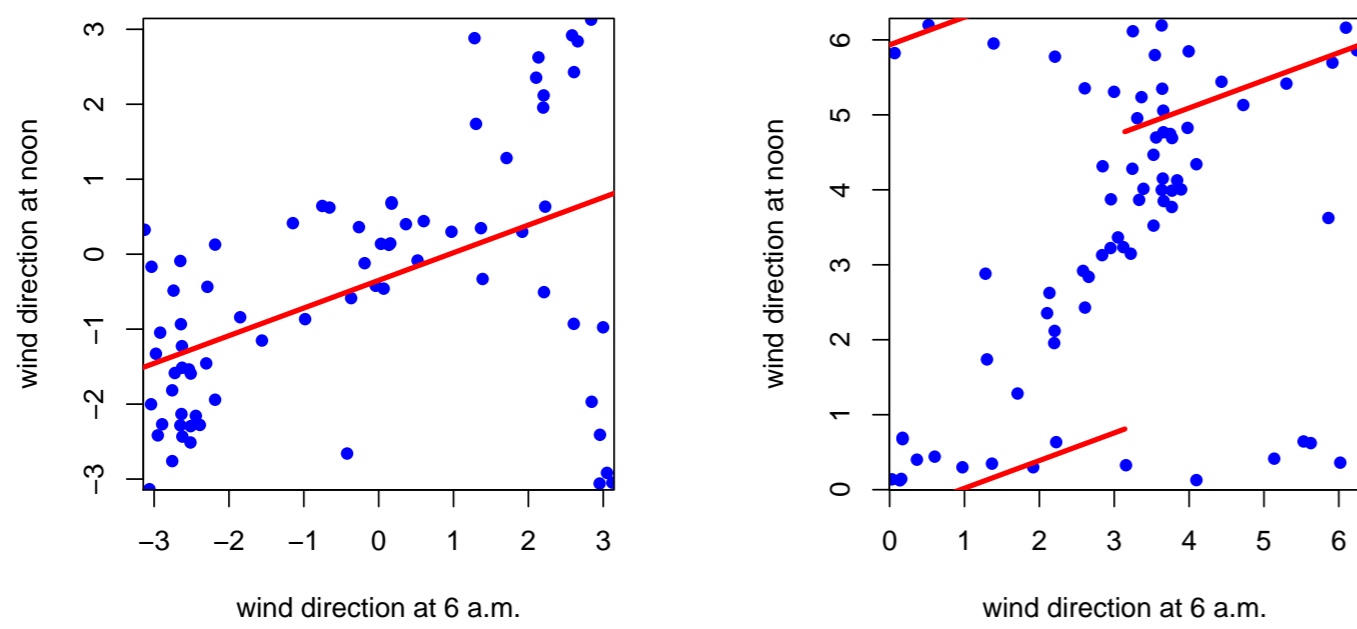


図1. 2変量の角度データ $\{(\theta_{1j}, \theta_{2j})\}_{j=1}^{n_i}$ に線形回帰モデルを当てはめたときのプロット. 角度のデータは、 $[-\pi, \pi)^2$ (左), $[0, 2\pi)^2$ (右) で表されている。

研究の目的

本研究では、**角度の観測を含むデータのための混合効果モデルを提案し、その統計的性質を議論する。**

具体的には、**被説明変数が角度、説明変数が実数値と角度からなる多変量データのための混合効果モデル**を議論する。

なお、本研究は、Louis-Paul Rivest 教授 (Université Laval, Canada) との共同研究である。

角度データのための混合効果モデル

モデルの定義

Y_{ij} : 被説明変数 (角度 $[-\pi, \pi)$ として表される),
 x_{ij} : 角度 $[-\pi, \pi)$ の説明変数,
 z_{ij} : 実数 \mathbb{R} または正の実数 \mathbb{R}^+ に値をとる説明変数,
 $(i = 1, \dots, m, j = 1, \dots, n_i)$.

このとき、提案する混合効果モデルは以下で定義される。

$$Y_{ij} = \mu(x_{ij}, z_{ij}; \beta) + a_i + e_{ij}. \quad (1)$$

ただし、 $\mu(x_{ij}, z_{ij}; \beta)$: 予測方向, β : 固定効果のベクトル,
 $a_i \sim i.i.d. \text{VM}(0, \kappa_a)$: 変量効果,
 $e_{ij} \sim i.i.d. \text{VM}(0, \kappa_e)$: 誤差, $a_i \perp e_{ij}$.

また、 $\text{VM}(\mu, \kappa)$ は、以下の確率密度関数を持つ von Mises 分布である：

$$f(\theta) = \frac{1}{2\pi \mathcal{I}_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad -\pi \leq \theta, \mu < \pi, \kappa \geq 0, \\ \mathcal{I}_j: \text{第1種変形ベッセル関数.}$$

モデルの性質

以後、簡単のため、混合効果モデル(1)において、 $\mu_{ij} = \mu(x_{ij}, z_{ij}; \beta)$ と書くことにする。

性質

混合効果モデル(1)において、 $(Y_{i1}, \dots, Y_{in_i})$ の同時確率密度関数は

$$f_{n_i}(y_{i1}, \dots, y_{in_i}) = \frac{\mathcal{I}_0(\sigma_i)}{(2\pi)^{n_i} \mathcal{I}_0(\kappa_e)^{n_i} \mathcal{I}_0(\kappa_a)}, \quad y_{i1}, \dots, y_{in_i} \in [-\pi, \pi),$$

で与えられる。ただし、

$$\sigma_i = \left\| \begin{pmatrix} \kappa_a \\ 0 \end{pmatrix} + \kappa_e \sum_{j=1}^{n_i} \begin{pmatrix} \cos(y_{ij} - \mu_{ij}) \\ \sin(y_{ij} - \mu_{ij}) \end{pmatrix} \right\|.$$

この性質を用いることにより、以下を導くことが可能である：

- trigonometric moment は陽に表すことができる ($p_1, p_2 \in \mathbb{Z}; \mu_{ij} = 0$):

$$\begin{aligned} E[\cos(p_1 Y_{i1} + p_2 Y_{i2})] &= A_{p_1}(\kappa_e) A_{p_2}(\kappa_e) A_{p_1+p_2}(\kappa_a), \\ E[\sin(p_1 Y_{i1} + p_2 Y_{i2})] &= 0, \quad (A_p(\kappa) = \mathcal{I}_p(\kappa) / \mathcal{I}_0(\kappa)). \end{aligned} \quad (2)$$

- 角度の相関係数 (Rivest, 1982, etc.) は、以下で与えられる：

$$\rho_{SS} \equiv \frac{E\{\sin(Y_{i1}) \sin(Y_{i2})\}}{E\{\sin^2(Y_{i1})\}} = \frac{A_1^2(\kappa_e) \{1 - A_2(\kappa_a)\}}{1 - A_2(\kappa_a) A_2(\kappa_e)}.$$

- $Y_{ij} = y_{ij}$ ($j = 1, \dots, n_i$) を所与としたときの a_i の条件付分布は von Mises 分布となる。その分布を用いて a_i の予測を行うことが可能。

モデルの推定

混合効果モデル(1)の μ_{ij} が、

$$\mu_{ij} = \arctan^* \left(\frac{\gamma_2}{\gamma_1} \right), \quad \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \sum_{k=0}^p \beta_k z_{ijk} \begin{pmatrix} \cos x_{ijk} \\ \sin x_{ijk} \end{pmatrix}, \quad (3)$$

の形で与えられるとき、以下のステップで最尤推定することができる：

1. 固定効果ベクトル β の予備的な推定値を、Rivest *et al.* (2016) のアルゴリズムにより推定する。
2. (2)式に基づく (κ_e, κ_a) のモーメント推定値を計算 (陽に表現できる)。
3. 上記の $(\beta, \kappa_e, \kappa_a)$ の推定値を初期値として用いて、最尤推定値を数値的に求める。

データへの応用

モデル(1)を、複数の条件を考慮した上でのハマトビムシが跳躍する方向のデータへと応用する。 μ_{ij} には(3)式を仮定する。

y_{ij} : i 匹目のハマトビムシが j 回目に跳躍する方向 ($1 \leq i \leq 59, 1 \leq j \leq 5$)

$$\{(x_{ij}, z_{ij})\} = \{(Azi, 1), (Azi + \pi/2, 1), (0, 1), (\pi/2, 1), (DirW, 1), \\ (DirW + \pi/2, 1), (DirW, SpeeW), (DirW + \pi/2, SpeeW), \\ (0, Eye), (\pi/2, Eye), (0, T), (\pi/2, T)\},$$

ただし、 Azi : 太陽方位, $DirW$: 風向, $SpeeW$: 風速,
 Eye : ハマトビムシの目の非対称性尺度の値,
 $T (= j - 1)$: j 回目の跳躍における反時計回りの影響。

backward elimination (Rivest *et al.*, 2016) と Wald 検定により変数選択し、最尤推定によりパラメータ推定した結果、以下のモデルが推定された：

$$\begin{aligned} \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} &= \begin{pmatrix} \cos(Azi) \\ \sin(Azi) \end{pmatrix} - 0.619 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 0.004 \cdot SpeeW \begin{pmatrix} \cos(DirW) \\ \sin(DirW) \end{pmatrix} \\ &\quad + 2.742 \cdot Eye \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 0.170 \cdot T \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \hat{\kappa}_e = 4.184, \quad \hat{\kappa}_a = 4.338. \end{aligned}$$