

円周上に値をとる時系列データ

円周上のデータ

円周上のデータ ... 個々の観測が角度 $[-\pi, \pi)$ として表すことができるデータ.

例) 風向データ, circular genomeの位置データ, など

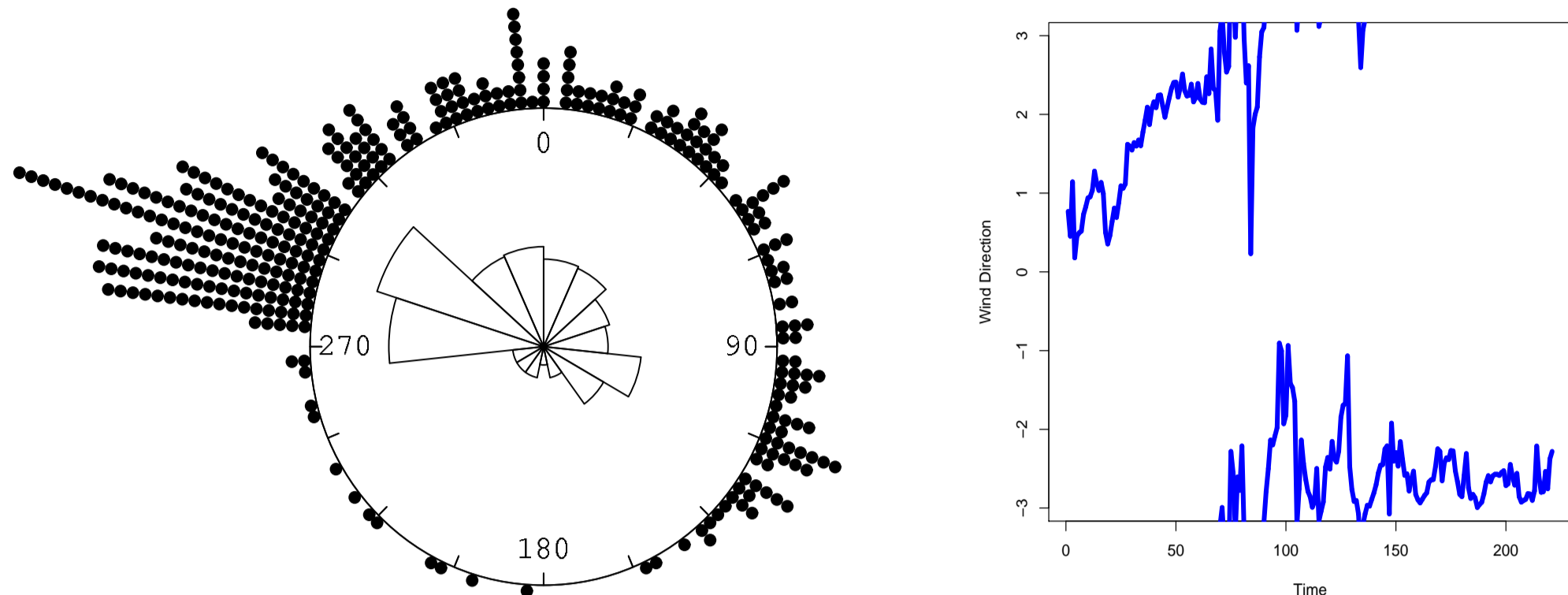


図1. (左) 1995年に岩手県大船渡市綾里の気象観測所にて観測された風向の一日ごとの平均方向のプロット. (右) アメリカ・テキサス州ヒューストンの気象観測所にて1時間ごとに観測された風向の時系列データのプロット.

図1(左)が示すように, 角度 θ は円周上の点 $(\cos \theta, \sin \theta)$ として表されることから, 角度として表されるデータを円周上のデータとよぶ.

円周上の時系列データ

一定時間ごとの風向を記録したデータ(図1(右))のように, 角度(円周上の点)は時系列データとして観測されることがある.

注 図1(右)において, 途中で角度が π から $-\pi$ へ(また $-\pi$ から π へ)ジャンプしているように見えるところがあるが, これは角度の周期性によるものであり, 実際には角度が大きく変化しているというわけではない.

研究の目的

円周上のデータの解析における問題点

● 円周上のデータには周期性があるため, (統計学が通常対象とする) 実数上のデータのための統計的手法がそのまま用いることができない場合がある.

例えば, 実数上のデータのための標本平均の定義は, 円周上のデータにそのまま応用すると不自然な定義となってしまう.

● それは時系列モデルに関しても同様であり, 円周上の時系列データのモデリングにおいて, 実数上の時系列モデルを直接応用することは不適切である.

⇒ 円周上の時系列データのための新たなモデルを考えることが必要!

円周上の時系列モデル

● 既存の時系列モデル ... Breckling (1989), Fisher & Lee (1994), など.

...しかし, モデルの解釈・解析的な扱いやすさという点で問題.

⇒ これらの欠点を克服した新たな時系列モデルを考える.

円周上に値をとるマルコフ過程

以後便利のため, 円周上の点を複素平面における単位円周上の点 $\cos \theta + i \sin \theta$ として表すことにする. ここで, i は虚数をあらわす($i^2 = -1$).

Fisher & Lee (1994)により提案された円周上の確率過程の導出法と, Kato *et al.* (2008)の回帰モデルを応用することにより, 新たなマルコフモデルを次のように定義する.

モデルの定義 (Kato, to appear)

W_0 : 複素平面の単位円 ∂D 上に値をとる確率変数または定数.

$\{W_n\}_{n=1}^\infty$: 以下で定義される確率変数列とする

$$W_n = \frac{W_{n-1} + \beta}{\beta W_{n-1} + 1} \varepsilon_n, \quad n = 1, 2, \dots, \quad (1)$$

ここに, β は複素平面の単位円盤 D 上に値をとるパラメータとする. また ε_n は, W_0 と独立で, circular Cauchy分布 $C^*(\varphi)$ ($0 \leq \varphi < 1$)に従うランダム標本とする.

このとき, 新たな円周上のマルコフ過程を $\{W_n\}_{n=0}^\infty$ で定義する.

モデルの解釈

● (1)式で用いた変換,

$$\mathcal{M}(w; \beta) = \frac{w + \beta}{\beta w + 1}, \quad w \in \partial D; \quad \beta \in D, \quad (2)$$

はMöbius変換とよばれ, 単位円をそれ自身に写す写像として知られる.

● Möbius変換におけるパラメータ β は, w を $\arg(\beta)$ (β の偏角)の向きに, $|\beta|$ (β の絶対値)の大きさで引き付けるパラメータと解釈することができる(図2).

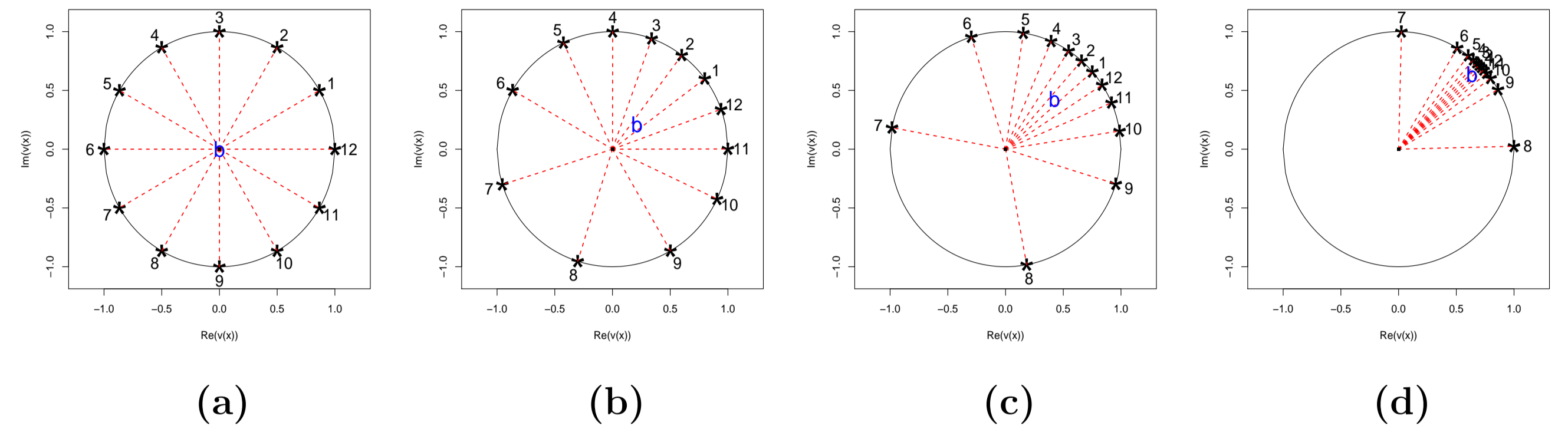


図2. Möbius変換(2) $\mathcal{M}(w; \beta)$ において, $\arg(\beta) = \pi/4$ と固定し, $|\beta|$ の値を(a) 0, (b) 0.3, (c) 0.6, (d) 0.9と変えたときの $\mathcal{M}(w_j; \beta)$ ($w_j = \exp(2\pi j/12)$, $j = 1, \dots, 12$)の値のプロット.

● circular Cauchy分布 $C^*(\phi)$ は, 確率密度関数

$$f(y) = \frac{1}{2\pi} \frac{1 - |\phi|^2}{|y - \phi|^2}, \quad y \in \partial D; \quad \phi \in D,$$

で定義される円周上の確率分布である. この分布は単峰形で, $y = \pm \phi/|\phi|$ に関して対称となる. $\arg(\phi)$ は分布の平均方向を調節し, $|\phi|$ は分布の集中度に影響する.

モデルの性質

定理1

$$W_n | (W_0 = w_0) \sim C^*\{\phi_n(w_0)\},$$

ここに,

$$\phi_n(w_0) = \left(\frac{\varphi + \beta \varphi}{\beta + 1} \right)^n \circ w_0, \quad \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \circ z = \frac{a_{00}z + a_{01}}{a_{10}z + a_{11}}, \quad n \geq 0.$$

定理2

$$W_n | (W_0 = w_0) \xrightarrow{d} C^*(\phi_\infty) \text{ as } n \rightarrow \infty,$$

ただし,

$$\phi_\infty = \begin{cases} \frac{\varphi - 1 + \sqrt{\{(1 - \varphi)^2 + 4\varphi|\beta|^2\}}}{2|\beta|} \frac{\beta}{|\beta|}, & \beta \in D \setminus \{0\}, \\ 0, & \beta = 0. \end{cases}$$

また上記の結果以外にも, $C^*(\phi_\infty)$ がuniqueな不変分布となることや, エルゴード性が成り立つこと, などのモデルの性質を示すことができる(Kato, to appear).

風向の時系列データへの応用

● ドイツ・ヴァルトホーフにおいて観測された風向の時系列データ ($n = 96$) に対して, 今回提案した(定常)モデルとFisher & Lee (1994)のモデルを最尤法による当てはめを行う.

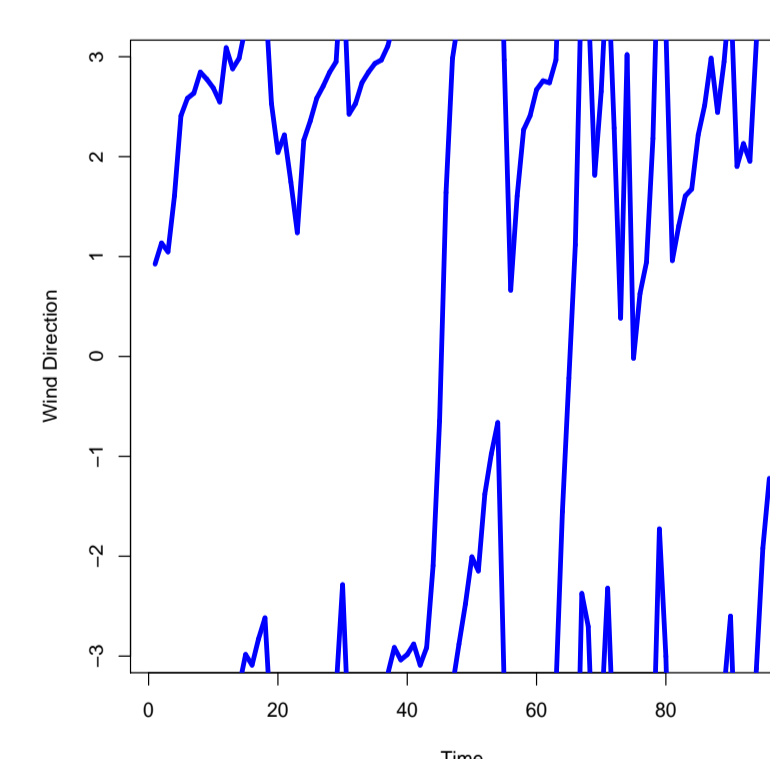


表1. 図3の風向データに対して当てはめを行った2つのモデルの最大対数尤度とAIC.

モデル	最大対数尤度	AIC
提案したモデル	-117.8	241.6
Fisher & Lee (1994)	-130.1	266.2

図3. 2007年5月23日午後10時から27日午後9時にかけて, ドイツ・ヴァルトホーフの気象観測所にて1時間ごとに観測された風向の時系列データのプロットと, 最尤法により推定された定常モデル.

● 最大対数尤度, AICを比較すると, 提案したモデルの方がFisher & Lee (1994)のモデルよりも良い当てはめを与えていることがわかる.

References

- [1] BRECKLING, J. (1989). *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*. Springer, London.
- [2] FISHER, N.I. & LEE, A.J. (1994). Time series analysis of directional data. *Journal of the Royal Statistical Society, Ser. B*, **56**, 327-339.
- [3] KATO, S. A Markov process for circular data. *Journal of the Royal Statistical Society, Ser. B*, to appear.
- [4] KATO, S., SHIMIZU, K. & SHIEH, G.S. (2008). A circular-circular regression model. *Statistica Sinica*, **18**, 633-645.