周期時系列データの多重整列法

森 源太 †　　岩田 一貴 ‡　　林 朗 ‡　　末松 伸朗 ‡

† 広島市立大学情報科学学部　　‡ 広島市立大学大学院情報科学研究科

1　概要

ガウス過程事前分布を利用した時系列整列法 [1] をもとに，周期時系列の多重整列法を提案する．本研究では，特に既知の周期を持つ時系列データについて多重整列することを目指す．

複数の周期時系列データについて，標準時系列に時間伸縮とノイズ，振幅の変異を加えることで観測値が得られると仮定する生成モデルを考え，周期的な標準時系列と時間変換についてベイズ解析を行うことで多重整列を実現する．この手法では，標準時系列と時間変換関数に対してガウス過程事前分布をおき，マルコフ連鎖モンテカルロ法によりそれらの事後平均を得ることで推定を行う．

提案手法の有効性は，人工データを用いた実験により検証される．

2　提案手法

ガウス過程 [2] とは，関数の確率分布であり，平均関数と共分散関数により，完全に指定される．本研究では，標準関数 $f(t)$ と各時系列 i の時間変換関数 $s_i(t)$ に対してガウス過程事前分布を仮定する．従って，そのモデルは，

$$\begin{aligned} f(t) &\sim \mathcal{GP}(m_f(t), k_f(t,t')) \\ s_i(t) &\sim \mathcal{GP}(m_s(t), k_s(t,t')) \\ y_{i,j} &\sim \mathcal{N}(f(s_i(t_j)), \sigma_\epsilon^2) \end{aligned} \tag{1}$$

と書ける．ここで，$\{y_{i,j}|j = 1,\ldots,T\}$ は i 番目の時系列データであり，$t_1,\ldots,t_T$ は時系列の j 番目のデータが観測された時刻である．

本研究では，周期的なデータを対象とするので，$f(t)$ も $s_i(t)$ も周期性をもたなければならない．そのため，共分散関数は 2 乗指数型共分散関数に基づいて得られる

$$k(t,t') = \sigma^2 \exp\left(-\frac{2\sin^2(\frac{t-t'}{2})}{l^2}\right) \tag{2}$$

を用いる [2]．ここで σ^2 はある 1 点の値 $f(t)$ の周辺分布の分散，l はスケールパラメータである．

本研究では，モデル (1) のための MCMC アルゴリズムを開発し，利用する．

3　実験

人工データを整列する実験を行った．人工データは提案手法の生成モデルに従って生成した 1 セット 5 本の図 1 左のような周期時系列を 100 セット用意した．提案手法と Ramsay の手法 [3] との比較を行う．本実験では，MCMC を 10,000 ステップ行い，後半の 5,000 サンプルより事後平均を求めた．図 1 右に推定結果を用いて整列した例を示す．評価は，推定した時系列データ間の時間の対応を示すワーピングパスと真のワーピングパスとのずれの面積を用いた．100 セットの結果を図 2 の箱ひげ図にまとめる．

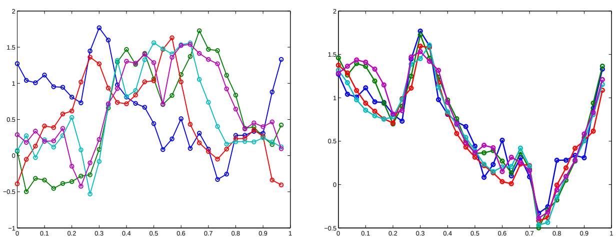

図 1　整列例

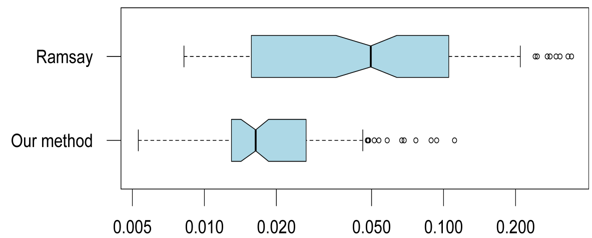

図 2　評価の箱ひげ図

4　今後の課題

ガウス過程事前分布のパラメータは既知のものとして実験を行ったが，これらのパラメータについても推定できるように手法を改良したい．また，様々な実データについての実験も行う必要がある．

参考文献

[1] 秋本 真治，末松 伸朗，林 朗，岩田 一貴，電子情報通信学会論文誌. D, 情報・システム J96-D(3), 587-595, 2013-03-01

[2] C.E. Rasmussen and C.K.I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[3] J.O. Ramsay and X. Li, "Curve registration," Journal of the Royal Statistical Society. Series B (Statistical Methodology), vol.60, no.2, pp.351-363, 1998.

Copyright © 2014 IEICE