角度時系列を含む多変量観測データに対する隠れマルコフモデルの応用

甫喜本 司 (東京大学), 清水邦夫 (慶應義塾大学)

1. はじめに

確率分布の時間的な変化を予測するためのモデルは, 様々な分野におけるデータ分析の中で必要となることが多い。近年, このようなモデルを考える上で隠れマルコフモデル (HMM) がしばしば応用される。一般的な時系列データに対する HMM の研究については, 推移確率の変化の構造に関するモデル化 (Raftery(1985a)), 分布のパラメータに集中した時間の関数に基づくモデル化 (マルコフ回帰モデル, Zeger and Qaqish(1988)) をはじめとして提案と検討が行われてきた。本研究では, 角度の属性をもつ時系列データを取り扱うための HMM の手法に関心がある。この中で, 角度時系列を含む多変量観測データに基づいて予測を行う HMM の方法に関する検討と, その有効性について実際の観測データに基づく一つの検証を行う。

2. 角度時系列を含む多変量データに対するモデルの記述

$t(t=1,\ldots)$ を時間, $s(s=1,\ldots,m)$ を状態に関するパラメータとする。以下では, 非定常な時系列データ $\{Y_t^{(L)}\},\{Y_t^{(S)}\}$ と角度時系列データ $\{Y_t^{(D)}\}$ とを同時に観測した状況を考える。本研究では一定時間幅の観測データに HMM をあてはめて推定された状態推移確率 $\gamma_{ij}^{(\cdot)}(t|t-1)(i,j=1,\ldots,m)$ と状態依存確率分布のパラメータ $\theta_{t,s}^{(\cdot)}$ の変化に基づいて確率的なモデル化を試みる。具体的には, 状態推移確率の変化が状態依存確率分布のパラメータの変化へ与える影響に着目し, 状態毎に次の多変量自己回帰モデルを考える。

$$\begin{aligned}\mathbf{\Theta}_{t,s} &= B_{1,s}\mathbf{\Theta}_{t-1,s}+\cdots+B_{p,s}\mathbf{\Theta}_{t-p,s}+\boldsymbol{\xi}_{t,s},\ \ s=1,\ldots,m, \qquad (1)\\ &\mathbf{\Theta}_{t,s}\equiv(\nabla\theta_{t,s}^{(L)},\nabla\theta_{t,s}^{(D)},\nabla\theta_{t,s}^{(S)},p_{t|t-1}^{(L)},p_{t|t-1}^{(D)},p_{t|t-1}^{(S)})',\\ &p_{t|t-1}^{(\cdot)}\equiv\log\left(\frac{\prod_{i\neq j}^{m}\gamma_{ij}^{(\cdot)}(t|t-1)}{1-\prod_{i\neq j}^{m}\gamma_{ij}^{(\cdot)}(t|t-1)}\right).\end{aligned}$$

ここで ∇ は 1 時点前との階差をとるオペレータを意味し, $B_{\cdot,s}$ は状態毎に定義される未知の係数行列, $\xi_{t,s}$ は攪乱項で時間と変数間の双方に関して無相関な構造をもつものとする。

$\{Y_t^{(D)}\}$ の確率分布は円周上の確率分布に従うと仮定する。これを状態依存確率とした HMM を $\{Y_t^{(D)}\}$ にあてはめた際の状態依存確率分布のパラメータと状態推移確率を $\Theta_{t,s}$ に定義する。例として, $\{Y_t^{(D)}\}$ が von Mises 分布に従う場合にはその確率密度関数が

$$f(y;\kappa,\mu)=\frac{1}{2\pi I_0(\kappa)}\exp(\kappa(y-\mu)),\quad 0<\kappa<\infty,\quad 0\le\mu<2\pi$$

で与えられる ($I_0(\cdot)$ は 0 次の第 1 種変形ベッセル関数) ので, $\Theta_{t,s}$ として

$$\begin{aligned}\mathbf{\Theta}_{t,s}&\equiv(\nabla\mu_L(t,s),\nabla\mu(t,s),\nabla\mu_S(t,s),p_{t|t-1}^{(L)},p_{t|t-1}^{(D)},p_{t|t-1}^{(S)})',\\ \mathbf{\Theta}_{t,s}&\equiv(\nabla\log\sigma_L^2(t,s),\nabla\log\kappa(t,s),\nabla\log\sigma_S^2(t,s),p_{t|t-1}^{(L)},p_{t|t-1}^{(D)},p_{t|t-1}^{(S)})'\end{aligned}$$

をそれぞれ定義して (1) を同定する。データからモデルを同定するためには, 一定時間間隔の時系列データへ HMM をあてはめて得られる推定値に基づいて $\{\hat{\Theta}_{t,s}\}$ を構成し, これに

$$\hat{\mathbf{\Theta}}_{t,s}=C_{1,s}\hat{\mathbf{\Theta}}_{t-1,s}+\cdots+C_{p,s}\hat{\mathbf{\Theta}}_{t-p,s}+\boldsymbol{\varepsilon}_{t,s},\ \ s=1,\ldots,m,$$

をあてはめる。予測量は

$$\hat{\mathbf{\Theta}}_{t+l,s}=\hat{C}_{1,s}\hat{\mathbf{\Theta}}_{t+l-1,s}+\cdots+\hat{C}_{p,s}\hat{\mathbf{\Theta}}_{t+l-p,s},\qquad l=1,\ldots,L,\ \ s=1,\ldots,m.$$

ただし, $\hat{C}_{\cdot,s}$ は $\{\hat{\Theta}_{t,s}\}$ に基づいて最小 2 乗法より推定する。

3. 角度時系列データに基づくモデルの効果

本研究では, 上記のモデルの構造が分布の予測にとって効果的であるかという点について実際のデータを用いて検証した。分析データとして海象に関する観測値である海面水位, 風向, 風速の同時計測データを使用し, 3 変量のデータに基づいて海面水位の分布の変化を予測することを目標とする。それぞれの観測データを 60 のクラスに分類し, これらへ上記の方法を適用した場合に海面水位分布のパラメータの予測精度がどのようになるかを評価した。海面水位と風速については正規分布を状態依存確率分布とする HMM を用いる (状態数は 3 変量共に 2 とする)。

予測実験は, 複数の予測量に基づいて外挿によるパラメータの予測を行った後, 精度の比較検討を行った。対象とした予測量は, P1) 海面水位変動のみから $\Theta_{t,s}$ を定義した場合, P2) 3 変量全ての状態推移確率 (6 種類) を $\Theta_{t,s}$ で定義した場合, P3) モデル (1) ($\{Y_t^{(D)}\}$ の状態依存確率分布が正規分布に従う場合), P4) モデル (1) ($\{Y_t^{(D)}\}$ の状態依存確率分布が Wrapped Cauchy 分布に従う場合), P5) モデル (1) ($\{Y_t^{(D)}\}$ の状態依存確率分布が von Mises 分布に従う場合), の 5 つである。モデルの構造上の検討を行うため, 全てのモデルの次数を 1 に固定して実験を行った。

表は上記の予測実験を予測開始時点を変えながら繰り返し 70 回行った後, 予測対象である海面変動の HMM に対する各状態のパラメータ毎に予測値の精度を評価した結果の一例である。予測値と実際に推定された値との間で得られた平均 2 乗予測誤差 (MAE), 及び両者の相関係数 (CORR) の例が示されている。1 変数のみの HMM に基づく予測量 P1 の結果は他の予測量よりも予測性能が悪くなる傾向があり, モデル (1) が 3 変量の情報を反映して海面水位分布の予測精度の改善に貢献していることが示唆されている。状態推移確率を全て $\Theta_{t,s}$ に組み込んだ P2 は P3 から P5 の各予測量よりも予測精度が低下する傾向にあるが, 未知係数を多く含んでいるために予測精度にデメリットを与える可能性が高いことが示唆される。P3 から P5 の 3 種類の予測量を比較すると, P3 に比べて P4 や P5 が予測精度として改善される傾向が観察される。この結果は風向データに円周上の確率分布を定義することで, 角度データのもつ情報をより効果的に予測へ反映できる効果が期待できることを示唆している。

表 状態依存確率分布のパラメータに対する予測性能の例 (1 期先予測と 5 期先予測, 実験回数:70 回)

	MAE(1)					CORR(1)				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
μ-I	2.223	9.446	0.740	0.647	0.822	0.520	0.568	0.650	0.656	0.602
σ^2-I	10.8073	0.3942	0.0062	0.0057	0.0083	-0.097	0.698	0.991	0.993	0.991
μ-II	1.199	2.976	0.682	0.566	0.541	0.535	0.312	0.676	0.694	0.693
σ^2-II	0.00590	0.05980	0.00014	0.00017	0.00015	0.748	0.285	0.991	0.990	0.992

	MAE(5)					CORR(5)				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
μ-I	3.354	1.928	2.982	2.245	1.784	0.025	0.158	0.259	0.244	0.181
σ^2-I	0.928	0.047	0.102	0.047	0.071	0.514	0.936	0.844	0.938	0.927
μ-II	1.957	2.210	2.720	1.924	2.334	0.222	0.167	0.278	0.298	0.265
σ^2-II	0.01366	0.00058	0.00051	0.00054	0.00053	0.576	0.963	0.963	0.967	0.970

参考文献

Holzmann, H., Munk, A. Suster, M. and Zucchini, W. (2006), Hidden Markov models for circular and linear-circular time series, Environ. Ecol. Stat., 325-347.
MacDonald, I. L. and Zucchini, W. (1997), Hidden Markov and Other Models for Discrete-valued Time Series, Chapman & HALL/CRC.
MacDonald, I. L. and Zucchini, W. (2009), Hidden Markov Models for Time Series, CRC Press.
Raftery, A. E. (1985a), A model for high-order Markov chains, J. R. Statistical Soc., B47, 528-539.
Zeger, S. L. and Qaqish, B. (1988), Markov regression models for time series: a quasi-likelihood approach, Biometrica, 44, 1019-1031.