

差分型 HMM を用いたデータストリームにおける時系列データ予測

若林 啓[†] 三浦 孝夫[†]

[†] 法政大学 工学研究科 〒184-8584 東京都小金井市梶野町 3-7-2
E-mail: [†]kei.wakabayashi.bq@gs-eng.hosei.ac.jp, ^{††}miurat@k.hosei.ac.jp

あらまし 本研究では時系列データ予測の新しい手法を示す。ここでは各時刻の観測データを事象として解釈し、事象の遷移に基づいてデータの予測を行う。本稿では事象を隠れ状態に対応させた隠れマルコフモデル (HMM) によるモデル化を行い、状態からの出力確率分布を予測値の分布として用いる。また、データストリーム環境においては EM アルゴリズムによる HMM のパラメタ学習を行うことが困難であるため、差分型 Baum-Welch アルゴリズムによるオンライン学習手法を適用し、実験により本手法の有用性を検証する。

キーワード 時系列予測, データストリーム, 隠れマルコフモデル, 差分学習

Forecasting Time-Series on Data Stream using Incremental Hidden Markov Models

Kei WAKABAYASHI[†] and Takao MIURA[†]

[†] Dept. of Elect. & Elect. Engr., HOSEI University 3-7-2, KajinoCho, Koganei, Tokyo, 184-8584 Japan
E-mail: [†]kei.wakabayashi.bq@gs-eng.hosei.ac.jp, ^{††}miurat@k.hosei.ac.jp

Abstract In this paper, we propose a sophisticated technique for time-series prediction. Here we consider an observation as output of the *event*, and estimate future data based on transition of events. In this investigation we formalize event as state of *Hidden Markov Model* (HMM), and observation as output from state. However, HMM approach is not easy to apply *data stream* prediction. This is because Baum-Welch algorithm which is traditional unsupervised HMM learning algorithm scan all historical data many times. Here we apply *incremental Baum-Welch algorithm* which is a on-line training method for HMM to data stream prediction for adapting to new pattern. We show some experimental results to see the validity of the method.

Key words Forecasting, Data Stream, Hidden Markov Model, Incremental Learning

1. 前 書 き

近年、計算機資源の充実化に伴い、より知的な情報処理を計算機で行う研究が盛んになっている。特に、計算機を用いることで膨大な量のデータを扱えるようになったことを背景にして、データベースからの知識発見やデータ予測の技術が注目されている。計算機による時系列データ予測とは、時刻を伴ってリアルタイムに観測されるデータについて、現在までの観測データを用いて未来に観測されるデータを自動的に推定する技術である。

時系列データ予測は、リアルタイムに大量の情報を扱う場合に有用な技術である。例えば、商品の在庫状況の管理を計算機によって行っているシステムにおいて、時系列データ予測が可能となれば、翌日の売り上げを予測することで効果的な商品の仕入れが自動的に行うことができる。また、大量のデータを利用して客観的にデータ予測を行うことで、企業や政府の意思決定を補助する効果も期待される。

時系列データ予測は、これまでに多くの手法が提案されている [4]。指数平滑法 (Exponential Smoothing) は、過去のデータの重み付きの平均値に基づいて予測を行うヒューリスティックなデータ予測手法である。重みは、観測から経過した時刻に対して指数的に減少する。このとき、時刻 t での重み付きの平均値は、ひとつ前の時刻 $t-1$ の重み付き平均値を用いて集約的に求めることができる。

Holt-Winters 法は指数平滑法の一様であり、重み付き平均値の増減の傾向を考慮して予測を行う [3]。ここでは、時刻 t の平滑化した観測値を $\tilde{y}_t$ 、その変化量を F_t として、各時刻で以下の式を用いて $\tilde{y}_t$ および F_t を更新する。

$$\tilde{y}_t = \lambda_1 y_t + (1 - \lambda_1)(\tilde{y}_{t-1} + F_{t-1})$$

$$F_t = \lambda_2(\tilde{y}_t - \tilde{y}_{t-1}) + (1 - \lambda_2)F_{t-1}$$

ただし、 λ_1 および λ_2 は平滑化パラメタと呼ばれる 0.0 から

1.0 までの値であり、 λ が大きいほど過去のデータの重みは急速に減少する。Holt-Winters 法による時刻 $t+h$ の予測は、 $\tilde{y}_t$ を用いて

$$\tilde{y}_{t+h|t} = \tilde{y}_t + hF_t$$

として求める。これは、未来のデータが現在の増加量の傾向を維持することを意味する。

しかし、これらの手法は観測値の部分的な線形性を仮定している。イベント会場に出入りした人数の時系列データや、ある地点の風速の時系列データなどは、前の時刻の観測値と直接関係せず、催し物や台風の接近といった「事象」に基づいて観測値が決まる。このような時系列データに対しては、指数平滑法で効果的な予測が可能であるとは考えにくい。

Hassan ら [5] は、隠れ状態をもつ確率過程モデルである隠れマルコフモデルを用いて株価の予測を行う。ここでは過去の観測データから、隠れマルコフモデルの学習により頻出する推移のパターンを抽出する。学習したモデルを用いて、最新の時刻の観測データにどの隠れ状態が対応するかを確率的に求め、次の時刻に観測されるデータの尤度を推定することで予測を行う。隠れ状態の確率を求めることは、観測されるデータに対して、最新の観測データが過去のどの推移パターンと類似するかを推定することに対応する。この手法では、予測値は最新の時刻の観測に直接依存せず、隠れ状態の遷移に基づいて決定する。このことから、隠れた事象を含む時系列データに対しても効果的な予測が期待できる。

しかし、Hassan らは EM アルゴリズムを用いて、オフラインに隠れマルコフモデルの学習を行うため、予測の最中に未知の系列パターンが出現した場合に、新しいパターンをモデルに反映させることができない。このことは、特にデータストリーム環境において大きな問題となる [6] [8]。データストリームではデータは永続的に発生し、また動的にデータの分布が変化するため、学習データに含まれるデータのパターンはデータストリーム全体で発生するパターンのごく一部である。このため、モデルの学習を予測を行いながら（オンラインに）行う必要がある。

隠れマルコフモデルの学習を差分的に行う研究に、Stenger らがある [9]。彼らはラベルなし学習データによる隠れマルコフモデルの学習手法である Baum-Welch アルゴリズムを拡張した、差分型 Baum-Welch アルゴリズム (Incremental Baum-Welch) を提案している。差分型 Baum-Welch アルゴリズムでは、新規のデータに対して推定した隠れ状態の尤度に基づいて、その尤度が最大になるようにパラメタの再推定を行う。この手法は、モデルのパラメタを用いて集約的にパラメタを再推定するため、過去の観測データを記憶することなく、かつ高速に実行できる。Stenger らは、差分型 Baum-Welch アルゴリズムを用いて動画中の物体検出を行っている。

本稿では、差分型 Baum-Welch アルゴリズムを用いた時系列データ予測の手法を提案する。これにより、隠れマルコフモデルによる適応的な時系列データ予測が可能になることを論じる。

2 章では状態遷移に基づくデータ予測について述べる。3 章では本研究で用いる隠れマルコフモデルのパラメタの動的な推

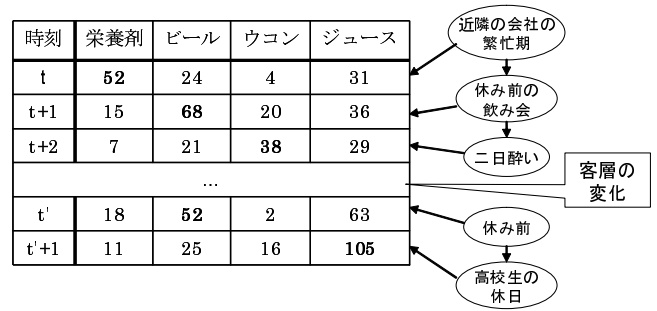


図 1 状態遷移に基づく時系列のモデル化

定法を説明し、本稿で提案するデータ予測のアルゴリズムについて述べる。4 章で実験結果を示し、5 章で結びとする。

2. 状態遷移に基づくデータ予測

図 1 に、本研究で提案する時系列データ予測のモデルを示す。図中の時系列データは、量販店の商品の売り上げ履歴である。ここでは、時刻 t に栄養剤がよく売れ、翌日 (時刻 $t+1$) にはビール、その翌日 (時刻 $t+2$) にはウコンがよく売れている。

本研究では、時系列データにおいて観測されるデータは、観測できない隠れた状態に依存すると考える。例えば、図 1 の時刻 t ではビールの売り上げが少なく、栄養剤がよく売れていることから、来店する客は仕事で忙しい人が多いと解釈する。時刻 $t+1$ では、栄養剤よりもビールやおつまみがよく売れていることから、翌日が休日などの理由でお酒を飲む人が多い日と解釈する。同様に時刻 $t+2$ ではウコンがよく売れることから、二日酔いの人が多い日と解釈する。

これらの状態の対応づけから、仕事が忙しい人が多い日の翌日はお酒を飲む人が多い日に遷移し、お酒を飲む人が多い日は二日酔いの人が多い日に遷移する。これらのパターンが過去の観測データ系列の中に頻繁に出現していれば、時刻 t の時点で時刻 $t+1$ や時刻 $t+2$ の売り上げの予測が可能である。

ただし、実際にはこれらの状態の解釈は行われる必要はない。隠れ状態の遷移に基づく時系列データのパターンをなんらかのアルゴリズムによって獲得できれば、状態の意味が解釈されなくてもデータ予測は可能である。本研究では、隠れマルコフモデルにより隠れ状態を明示的にモデル化し、状態の意味を事前に与えることなく、観測したデータ系列に基づいてモデルのパラメタを決定する。

図 1 では時間の経過に伴い、近くに高校ができるなどの原因により客層が変化し、これにより状態の意味や遷移パターンが変化する。例えば、高校生の客が増えることで休日のジュースや菓子の売り上げが増加するといった変化が考えられる。このような変化を反映して予測を行うためには、常に新しいパターンの存在を考慮する必要がある。

隠れマルコフモデルにおいて、新しい観測データから未知のパターンをモデルに反映することは容易ではない。これは、モデルが隠れ状態を含むため、観測値から直接最尤のパラメタを推定できないことによる。隠れマルコフモデルのラベルなし学習を行うアルゴリズムには、EM アルゴリズムの一種である

Baum-Welch アルゴリズムが存在する。しかし、Baum-Welch アルゴリズムは尤度が極大になるまで学習データを繰り返し走査してパラメタを推定するため、一般に長大なデータになるほど実行時間は膨大となる。このため、新しい観測値が利用可能になるたびに Baum-Welch アルゴリズムによる再計算を行うのは困難である。

特にデータストリーム環境では、新しい観測を得る度に差分的にパラメタを推定することで動的な新規パターンの反映を実現する必要がある。これは、データ量が膨大であるために過去のデータを保存できないことと、全データの再計算を実行することが計算量の観点から困難であることによる。差分的なパラメタ推定とは、なんらかの集約値を用いて過去のデータを少ない情報量で表現し、これを用いてパラメタを再推定することを指す。差分型 Baum-Welch アルゴリズムは、隠れマルコフモデルの各パラメタを過去のデータの集約値として用いることにより、差分的なパラメタ推定を実現する。

本研究では差分型 Baum-Welch アルゴリズムを適用することにより、新規パターンに適応的な時系列データ予測を実現することを旨とする。

3. 差分型 HMM を用いたデータ予測

3.1 隠れマルコフモデル

隠れマルコフモデル (Hidden Markov Model, HMM) は、確率的に遷移する隠れ状態をもつオートマトンである。隠れ状態は、各時刻で単純マルコフ過程に従って遷移する。状態は観測できないが、各時刻の状態は確率的に観測値を出力する。また、観測値はその時刻の状態にのみ依存する。

隠れマルコフモデルは次の 5 つのパラメタによって定義される [7]。

- (1) $Q = q_1, \dots, q_N$: 状態の有限集合。
- (2) $\Sigma = o_1, \dots, o_M$: 観測値の集合。
- (3) $A = a_{ij}$: 状態遷移確率分布。 a_{ij} は状態 q_i から状態 q_j への遷移確率である。
- (4) $B = b_i(o_t)$: 観測値の出力確率分布。 $b_i(o_t)$ は状態 q_i で観測値 o_t を出力する確率である。
- (5) $\pi = \pi_i$: 初期状態確率分布。 π_i は状態 q_i が初期状態である確率である。

本稿では、観測値として多次元の連続値ベクトルを考える。ここでは出力確率分布 $b_i(o_t)$ は各次元で独立に正規分布に従う。このため本モデルは、状態毎、および観測値の次元毎に正規分布のパラメタとして平均と分散をもつ。

- (6) μ_{ik} : 観測値平均。 μ_{ik} は状態 q_i において k 次元目の要素の確率分布を与える正規分布の平均である。
 - (7) σ_{ik} : 観測値分散。 σ_{ik} は状態 q_i において k 次元目の要素の確率分布を与える正規分布の分散である。
- 時刻 t の観測ベクトル o_t の次元 k の要素 o_{tk} が出力される確率は、状態 q_i に滞在しているとき、平均 μ_{ik} 、分散 σ_{ik} の正規分布に従う。

$$b_i(o_{tk}) = \frac{1}{\sqrt{2\pi\sigma_{ik}}} e^{-\frac{(o_{tk}-\mu_{ik})^2}{2\sigma_{ik}}}$$

観測ベクトル o_t の出力確率は、それぞれの要素 o_{tk} の出力確率の積として求める。すなわち、 $o_t = (o_{t1}, o_{t2}, \dots, o_{tD})$ が状態 i から出力される確率は、

$$b_i(o_{t1}, o_{t2}, \dots, o_{tD}) = \prod_{k=1}^D b_i(o_{tk})$$

である。

与えられた学習データについて最適な状態遷移確率分布 A 、観測値平均 μ 、観測値分散 σ を推定する際には、ラベルなしデータを用いて準最適なパラメタを求める Baum-Welch アルゴリズムを用いる。本研究では状態の意味はあらかじめ分かっている。Baum-Welch アルゴリズムは EM アルゴリズムの一種であり、与えられた学習データの尤度が大きくなるようにパラメタを繰り返し更新し、準最適なパラメタに収束させる。

3.2 差分型 Baum-Welch アルゴリズム

時間の経過等により新たな観測データを得たとき、モデルのパラメタを再推定する手法について述べる。差分型 Baum-Welch (Incremental Baum-Welch) アルゴリズムは、まず現在のモデルを用いて、新規のデータに対応する時刻 T における状態 i の尤度 $\gamma_T(i)$ を計算する。モデルは現在の時刻 $T-1$ における状態 i の滞在確率 $\gamma_{T-1}(i)$ 、および現在までの全ての時刻の状態 i の滞在確率の和 $\sum_{t=1}^{T-1} \gamma_t(i)$ を保持しているとする。時刻 T で状態 i から状態 j に遷移する確率 $\gamma_T(i, j)$ は、

$$\gamma_T(i, j) = \gamma_{T-1}(i) \frac{a_{ij} b_j(o_T)}{\sum_j a_{ij} b_j(o_T)}$$

で求まる。時刻 T での状態 i の滞在確率は $\gamma_T(i) = \sum_j \gamma_T(i, j)$ である。ここから、各パラメタは次のように推定される [9]。

$$a'_{ij} = \frac{\sum_{t=1}^T \gamma_t(i, j)}{\sum_{t=1}^T \gamma_t(i)} = \frac{\sum_{t=1}^{T-1} \gamma_t(i) a_{ij} + \gamma_T(i, j)}{\sum_{t=1}^T \gamma_t(i)}$$

$$\mu'_i = \frac{\sum_{t=1}^T \gamma_t(i) o_t}{\sum_{t=1}^T \gamma_t(i)} = \frac{\sum_{t=1}^{T-1} \gamma_t(i) \mu_i + \gamma_T(i) o_T}{\sum_{t=1}^T \gamma_t(i)}$$

$$\sigma'_i = \frac{\sum_{t=1}^T \gamma_t(i) (o_t - \mu_i)^2}{\sum_{t=1}^T \gamma_t(i)} = \frac{\sum_{t=1}^{T-1} \gamma_t(i) \sigma_i + \gamma_T(i) (o_T - \mu_i)^2}{\sum_{t=1}^T \gamma_t(i)}$$

差分 Baum-Welch アルゴリズムは、モデルのパラメタを再推定式の集約値として用いることにより、過去の観測データを保持することなく、高速に再推定を行うことができる。

3.3 差分型 HMM によるデータ予測

本稿で提案する、差分型 HMM を用いたデータ予測の手法を示す。

差分 Baum-Welch アルゴリズムは EM アルゴリズムではないため、乱数で初期化したモデルのパラメタの収束は遅い。ここでは、最初から差分 Baum-Welch アルゴリズムによる学習を適用せず、まず既知である観測系列を用いて Baum-Welch アルゴリズムにより初期モデルを生成する。これは、パラメタの収束が極端に遅くなるのを防ぐためである。

まず、状態数 N の HMM のパラメタとして、状態遷移確率分布、観測値平均を乱数で決定し、観測値分散を全て 1.0 とする。こ

のモデルについて、既知である観測系列を用いて Baum-Welch アルゴリズムを実行する。パラメタが収束したら、時刻 T までの各時刻 t で状態 i に滞在する確率 $\gamma_t(i)$ を計算し、その和 $\sum_{t=1}^T \gamma_t(i)$ および最終時刻 T での滞在確率 $\gamma_T(i)$ を保存する。

得られた初期モデルを用いて、次の時刻 $T+1$ の観測値の予測を行う。 $T+1$ の観測値が未知であるとき、時刻 $T+1$ で状態 i に滞在する尤度 $\gamma_{T+1}(i)$ は、

$$\gamma_{T+1}(i) = \sum_{j=1}^N \gamma_T(j) a_{ji}$$

で与えられる。状態 i に滞在するとき、時刻 $T+1$ の最も尤もらしい観測値は、 $\mu_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{iD})$ である。以上のことから、時刻 $T+1$ の観測値の期待値は、

$$\sum_{j=1}^N \gamma_{T+1}(i) \mu_i$$

であり、この値を時刻 $T+1$ の予測値とする。

時刻が経過し、時刻 $T+1$ の観測値 o_{T+1} を観測したとき、差分型 Baum-Welch アルゴリズムに従ってパラメタの再推定を行う。

新しく推定したパラメタを用いて、時刻 $T+1$ の状態の滞在確率を求める。すなわち、

$$\gamma_{T+1}(i) = \sum_{j=1}^N \gamma_T(j) a'_{ji} b'_i(o_{T+1})$$

として滞在確率を求め、次の時刻の予測に用いる。

4. 実験

4.1 実験方法

本稿では「気象データひまわり」の1996年から2000年の5年分、1828日分のデータベースから、「日最大瞬間風速 (m/s)」「日照時間 (hour)」「日降水量 (mm)」の3項目の予測を行う。ここでは、東京、大阪、福岡の3観測地点についてのデータを実験に用いる。表1に、東京の観測データ系列を示す。

1828日のデータのうち、学習データとして系列の始めから10%から50%のデータを用いて残りの観測データの予測を行い、結果を比較する。また、予測アルゴリズムとして、差分 Baum-Welch を使用しない HMM による予測法 (HMM)、差分 Baum-Welch を使用した HMM による予測法 (Inc HMM)、およびベースラインとして Holt-Winters の Exponential Smoothing 法 (H-W) の3つで予測を行い、精度を比較する。予測精度の尺度は平均二乗誤差 (Mean Square Error; MSE) を用いる。時刻 t の予測値を p_t 、実際の観測値を x_t とすると、MSE は次のように定義される。

$$MSE = \frac{1}{T} \sum_{t=1}^T (p_t - x_t)^2$$

また、全実験データの90%について、差分 HMM と通常の HMM を用いて予測を行い、実行時間を比較する。実行時間は10回の試行の平均により評価する。

本稿では HMM の状態数は15として実験を行う。

4.2 実験結果

表2に、各アルゴリズムによる福岡の最大瞬間風速、日照時間、降水量の予測の MSE を示す。ここに示したのは、学習データと

時刻	最大瞬間風速	日照時間	降水量
1996 年 6 月 29 日	16.1	4.8	0.0
1996 年 6 月 30 日	11.0	1.4	14.0
1996 年 7 月 1 日	8.4	2.6	0.5
1996 年 7 月 2 日	7.6	0.8	0.0
1996 年 7 月 3 日	7.9	0.9	6.5
1996 年 7 月 4 日	10.5	0.1	0.0
1996 年 7 月 5 日	16.8	0.7	3.5

表 1 東京の実験データ

手法	最大風速	日照時間	降水量	平均
HMM	13.29	18.79	196.82	76.30
差分型 HMM	13.14	15.89	190.99	73.34
Holt-Winters	21.33	23.35	316.43	120.37

表 2 各アルゴリズムによる福岡の観測の予測誤差 (MSE)

事前学習データ	HMM	差分型 HMM	Holt-Winters
10%	70.79	69.10	118.37
20%	57.79	56.81	96.80
30%	60.39	59.52	98.79
40%	60.75	61.09	101.15
50%	65.18	64.50	106.12

表 3 東京の平均 MSE

事前学習データ	HMM	差分型 HMM	Holt-Winters
10%	49.51	47.91	79.54
20%	49.05	47.43	78.44
30%	51.51	51.62	83.00
40%	51.65	49.99	79.35
50%	52.79	45.46	73.39

表 4 大阪の平均 MSE

して全実験データの20%にあたる365日分の観測データを用いたときの結果である。表中の平均は、最大瞬間風速の MSE、日照時間の MSE、降水量の MSE を合計して属性数3で割った値である。

この条件においては、各属性とも差分型 Baum-Welch を使用した HMM による予測で最も MSE が小さくなる。Holt-Winters 法と比較すると、差分型 HMM の MSE は平均で約39%の改善がみられる。差分型 Baum-Welch を使用しない HMM との比較では、平均で約4%、最大では日照時間の予測において約15%の MSE の改善となる。

表3,4,5に、それぞれ東京、大阪、福岡の MSE の、各属性の平均値を示す。事前学習データは、オフライン学習に用いた系列の長さが全体の実験データの内で占める割合である。これらの結果から、本実験データについて HMM を用いた手法は Holt-Winters 法よりも小さい誤差で予測が可能であるといえる。また、通常の HMM 手法と差分型 HMM 手法では予測誤差に劇的な違いはないが、特に学習データが少ないときの予測精度で差分型 HMM 手法による精度の改善がみられる。

表6に、予測の実行時間を示す。これは全実験データの90%にあたる1646件の観測データの予測に要した時間である。1観測あたりの実行時間は、実行時間を観測データ数1646で割っ

事前学習データ	HMM	差分型 HMM	Holt-Winters
10%	71.30	68.65	115.24
20%	76.30	73.34	120.37
30%	75.85	73.71	123.15
40%	61.01	60.39	102.09
50%	54.88	55.70	92.55

表 5 福岡の平均 MSE

手法	実行時間 (ms)	1 観測あたりの実行時間 (ms)
HMM	623.3	0.38
差分型 HMM	892.1	0.54

表 6 予測の実行時間

た値である。パラメタの再推定を実行する分、差分型 HMM 手法は実行時間がかかるが、その差は 1 観測あたり 0.16(ms) とほとんど差がない。差分 Baum-Welch アルゴリズムは、動的なパラメタ推定を極めて高速に行うことができると言える。

4.3 考察

実験結果の考察と、提案アルゴリズムの評価を行う。

表 3,4,5 から、Holt-Winters 法の予測精度は HMM を用いた手法に比べて非常に悪い。これは、観測値の時間方向の差分値に基づいて予測を行うことが原因である。このため、本実験データのように一日ごとに観測値の差分が大きく変化するデータに対しては効果的な予測ができない。また、各次元の観測値を独立に予測を行うため、予測に利用できる情報が少ないことも予測精度を悪化させる要因である。

差分型 HMM と通常の HMM の比較では、学習量の全データに対する割合が小さいほど差分型 HMM 手法の方が通常の HMM の予測精度を上回る傾向がみられる。これは、少ない学習データ中には出現しなかった系列パターンを差分型 Baum-Welch によって動的にモデルに反映させているためである。データストリーム環境においては、一般に学習データとして利用可能な系列は予測を行うデータに比べ非常に少ないため、この結果は提案手法がデータストリーム環境において効果的であることを示している。

大阪のデータの予測においては、学習データの割合が大きいにも関わらず差分型 HMM 手法が通常の HMM の予測精度を大きく上回っている。これは主に、学習量を 50 %にしたことで通常 HMM 手法の降水量の予測精度が悪化したことによる。学習データを増加させたにも関わらず予測精度が悪化したのは、学習に用いたデータの特性に起因する。実験データの解析から、大阪では 98 年の降水量が他の年に比べて少なく、また雨が続かない特徴的な系列を有している。学習量を 40 %から 50 %に増加させたことで 98 年の系列を学習データに取り込んだため、翌年以降の予測精度を悪化させたと考えられる。

一方、差分型 Baum-Welch を行う HMM 手法では、通常 HMM 手法と同じモデルから予測を開始していながら、翌年以降の傾向をモデルに反映させることで予測精度の悪化を防いでいる。このことから、提案手法が学習データの局所的な特性に依存することなく、未来のデータについて適応的に予測を行うことが可能であるといえる。

また、実行時間の観点から、差分型 Baum-Welch アルゴリズムは通常の HMM 手法と比較してわずかな計算で実行可能である。本手法は実行速度の面でも、データストリーム上での予測に十分適用可能であるといえる。

5. 結 論

本研究では、時系列データの予測を行う新しい手法として、差分型 HMM を用いた手法を提案した。本手法により、学習データが少ない環境においても適応的にオンライン学習を行い、高速に予測を行えることを実験によって確かめた。

本研究では、モデルの状態数はあらかじめ与えられているものとした。しかし、差分パラメタ推定に加えて状態数を増減することにより、過去に例を見ないパターンの学習をより素早く行えるものとなる。

本研究では観測値は数値データとしたが、本手法は確率過程モデルを用いているため非数値データへの適用も可能である。例えばニュースストリームに適用することにより、事象の予測を行うといった応用も考えられる。

文 献

- [1] Cavalin P.R., Sabourin R., Suen C.Y. and Britto Jr., A.S.: Evaluation of Incremental Learning Algorithms for An HMM-Based Handwritten Isolated Digits Recognizer, The 11th International Conference on Frontiers in Handwriting Recognition (ICFHR 2008), Montreal, August 19-21, 2008.
- [2] S. Duan and S. Babu: Processing forecasting queries. In Proc. of the 2007 Intl. Conf. on Very Large Data Bases, 2007.
- [3] Sarah Gelper, Roland Fried, and Christophe Croux: Robust Forecasting with Exponential and Holt-Winters Smoothing, (June 2007)
- [4] Jan G. de Gooijer, Rob J. Hyndman: "25 Years of IIF Time Series Forecasting: A Selective Review," Tinbergen Institute Discussion Papers 05-068/4, 2005.
- [5] Md. Rafiul Hassan , Baikunth Nath: StockMarket Forecasting Using Hidden Markov Model: A New Approach, Proceedings of the 5th International Conference on Intelligent Systems Design and Applications, 2005.
- [6] Jian, N. and Gruenwald, L.: Research Issues in Data Stream Association Rule Mining, SIGMOD Record 35-1, 2006, pp.14-19
- [7] 北 研二: “確率的言語モデル”, 東京大学出版会, 1999.
- [8] S.Muthukrishnan: Data streams: algorithms and applications, Proceedings of the fourteenth annual ACM-SIAM symposium on discrete algorithms (2003).
- [9] B. Stenger and V. Ramesh and N. Paragios: Topology free Hidden Markov Models: Application to background modeling, In IEEE International Conference on Computer Vision, 2001.
- [10] Kei Wakabayashi, Takao Miura: Topics Identification Based on Event Sequence Using Co-occurrence Words, 13th Intn'l Conf. on Applications of Natural Language to Information Systems (NLDB),2008