

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年6月7日現在

機関番号：33920

研究種目：基盤研究(C)

研究期間：2010～2012

課題番号：22500213

研究課題名（和文） ベイズ神経回路網の研究と隠れマルコフ鎖推定への応用

研究課題名（英文） Bayes neural networks and its application to estimation of hidden Markov chains

研究代表者

伊藤 嘉房 (ITO YOSHIFUSA)

愛知医科大学・客員教授

研究者番号：10022774

研究成果の概要（和文）：隠れマルコフ鎖は状態のマルコフ鎖であるが、状態は観測されない。状態に依存して発生する信号の列から、神経回路網を用いて状態の列を推定するのがこの研究の目的である。必要なベイズ判別関数は、神経回路網が学習する。単一の神経回路網に同時に複数のベイズ判別関数を学習させる試みは、モデル実験で成功した。低次元信号の場合、隠れマルコフ鎖推定のシミュレーションは成功した。そのアルゴリズムは高次元信号の場合にも原理的に適用可能である。

研究成果の概要（英文）：A hidden Markov chain is a sequence of states. Though the states cannot be observed, the signals, each of which is generated accompanying the transition of the states, can be observed. We have tried to estimate the sequence of states using a neural network from the observation. For the estimation, Bayesian discriminant functions are necessary. Our neural network can learn several Bayesian discriminant functions simultaneously if necessary. In the case of low-dimensional signals, the network successfully estimated a hidden Markov chain. The algorithm can in principle be applied to higher-dimensional signal cases.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,600,000	480,000	2,080,000
2011年度	600,000	180,000	780,000
2012年度	600,000	180,000	780,000
年度			
年度			
総計	2,800,000	840,000	3,640,000

研究分野：総合領域

科研費の分科・細目：情報学、感性情報学・ソフトコンピューティング

キーワード：隠れマルコフ鎖、前（後）向きアルゴリズム、ベイズ判別関数、神経回路網、学習、ロジスティック活性化関数

1. 研究開始当初の背景

(1) 隠れマルコフ鎖は、複数の状態のマルコフ鎖であるが、状態そのものは観測されない。しかし、各状態において発生する信号は観測され、その確率分布は状態に依存する。こう

して得られる信号の列を解析して、状態の列を推定するのが「隠れマルコフ鎖の推定」で、その結果得られるのは各時点における事象の分布である。この推定には、通常、ベイズ判別関数が用いられ、状態の列がマルコフ鎖

であることが本質的に利用される。

(2) 当研究ではこのベイズ判別関数を神経回路網に学習させる。神経回路網によるベイズ判別関数の学習には多くの先行研究があるが、概して理論的研究にとどまり、用いられる神経回路網（ベイズ神経回路網）の具体的構築については殆ど論文に記載されない状態が続いた。学習の困難によると思われる。Funahashi(1998)はベイズ神経回路網の具体的構築について提案したが、ここにおいてもシミュレーションは行われていない。当代表者は、当初、Funahashiにより提案された神経回路網またはその改良によるベイズ判別関数の学習を試みたが、学習は極めて困難であった。

(3) その理由を解析して、用いる三層神経回路網の隠れ層素子の内部パラメータの学習による最適化の困難が、神経回路網の学習の困難の主な原因と判断した。それゆえ、素子の数の増加と引き換えに内部パラメータの数を削減する方法を提案した (Neural Computation, 2008)。このようにしても最適化が必要なパラメータの数はほとんど変わらず、内部パラメータを外部パラメータに変換した形になった。この方法を採用してから、低次元のベイズ判別関数の学習のシミュレーションは比較的容易に進むようになり、研究結果を学会あるいは Springer の Lecture note 等に発表した。

(4) 当代表者の過去の研究領域に生理学、数学、確率論、確率過程論などが含まれていたため、その結果、1990年頃、数学的方法を駆使した神経回路網の論文に出会って、この分野に興味をもつことになった。それで、神経回路網による関数近似に関する論文を Neural Networks (1991a, b, 1992)などに投稿して、神経回路網の研究に参入した。その後、研究テーマの一つであった確率過程論との関係で、神経回路網の学習の確率過程に興味を持ち、方法論的にも過去の研究経験が有効に働くと考えて、当研究を開始した。

(5) 当研究開始当時、すでに、神経回路網による単純なベイズ判別関数の学習や隠れマルコフ鎖の推定の予備的シミュレーションにある程度成功していたが、学習の成功はパラメータの初期値選択に大きく依存した。予備的シミュレーションの一部はプログラム言語“R”によっていたが、基本的にはエクセルによるシミュレーションに留まり、結果は限定的であった。そのような方法では一般化は困難であると考え、当研究においてその拡張を試みることにした。

2. 研究の目的

(1) ベイズ判別関数を学習し得る神経回路網を構築し、それを用いて隠れマルコフ鎖の状態時系列を推定するのが研究の目的である。ベイズ判別関数は信号を観測した後の各状態の一組の事後確率である。教師信号は信号 $x^{(t)}$ と、それ生成した状態 $\theta^{(t)}$ の対 $(x^{(t)}, \theta^{(t)})$ の列である。当研究では、信号はユークリッド空間の点であり、状態は状態集合 $\Theta = \{\theta_i \mid i=1, \dots, n\}$ の点である。マルコフ鎖が n 個 ($n \geq 2$) の状態を取りえるとき、状態を推定するには一般に $n-1$ 個のベイズ判別関数が必要である。一組の事後確率から、その時点の状態の確率分布が得られる。

状態 θ_i の事後確率 (ベイズ判別関数) を学習する際には、対 $(x^{(t)}, \theta^{(t)})$ において、 $\theta^{(t)} = \theta_i$ のとき、 $\theta^{(t)}$ は 1 に置き換えられ、しからざる場合は 0 に置き換えられる。したがって、神経回路網に提示される教師信号列は $(x^{(t)}, 1)$ と $(x^{(t)}, 0)$ の混合列である。これは教師信号が、独立変数と、関数の値の対 $(x^{(t)}, f(x^{(t)}))$ またはそれに若干の雑音が重畳された通常の学習と状況がかなり異なり、学習の困難をもたらす。このような状況のもとで、目的関数であるベイズ判別関数を学習し得る神経回路網を構築するのが第 1 の目的である。

(2) マルコフ鎖が 3 個以上の状態 ($n \geq 3$) を取りえる場合、複数 ($n-1$) の判別関数が必要であり、神経回路網は複数のベイズ判別関数を学習する必要がある。複数の神経回路網を用意すれば対応できるが、それらは一本のマルコフ鎖にかかわる事後確率である。よって、互いに類似している可能性がある。複数のベイズ判別関数が相互に類似していて、若干の補正により、相互に変換可能な場合、回路網の主要部が判別関数の主要共通部分を学習すれば、それぞれの判別関数は共通部分に若干の補正を加えて得られる可能性がある。

このような学習を、単一の神経回路網に ($n-1$) 個の簡単なモジュールを付加して実現することを考える。補正項はモジュールに蓄えられる。複数の神経回路網により複数のベイズ判別関数を学習する場合、教師信号はそれぞれに振り分けられ、必要な教師信号の不足を招く恐れがある。しかし、単一の神経回路網による学習が実現すると、すべての教師信号がその主要部の学習に用いられる。補正項は単純で、その学習に教師信号の数は少なくても済み、教師信号が振り分けられても数が不足することはない。このように効率がよい学習を試みる。

(3) 事後確率は事前確率に依存する。先行する状態が知られていれば事前確率は推移確率に他ならないが、隠れマルコフ鎖を研究する立場においては先行状態を知ることはで

きない。しかし、先行する状態の確率分布は推定される。したがって、現時点の状態の事前確率は、先行状態の確率分布と推移確率から求められる。こうして求められた推定事前確率と回路網の学習結果を合わせて、現状の事後確率が得られる。これは現在の状態の推定確率分布である。これを繰り返して、推定事後確率分布の列が得られ、したがって、マルコフ鎖における状態の列が推定される。これは前向きアルゴリズムである。このアルゴリズムを実現させるのが研究の目的である。実現すれば、直ちに、この方法を逆方向に適用して後向きアルゴリズムが得られ、さらに両方向の推定を統合したアルゴリズムが得られる。

3. 研究の方法

(1) まずは理論的考察によって、学習の対象となるベイズ判別関数を近似し得る神経回路網の構造を決定し、シミュレーションによって、その学習可能性を確認する。教師信号が(0,1)の二値関数であることに起因する学習の困難があるが、先に述べたように三層神経回路網の隠れ層素子の内部パラメータの削減によりそれを克服する。シミュレーションには、“C プラス”の他に “S プラス” または “R” を用いる。

(2) 応用を考えるとなるべく少数の教師信号で学習が完了するのが望ましい。学習の目標となるベイズ判別関数が類似している場合、単一の神経回路網に、単純なモジュールを付加して複数の判別関数を学習させるアルゴリズムを考え、その有効性をシミュレーションにより確認する。

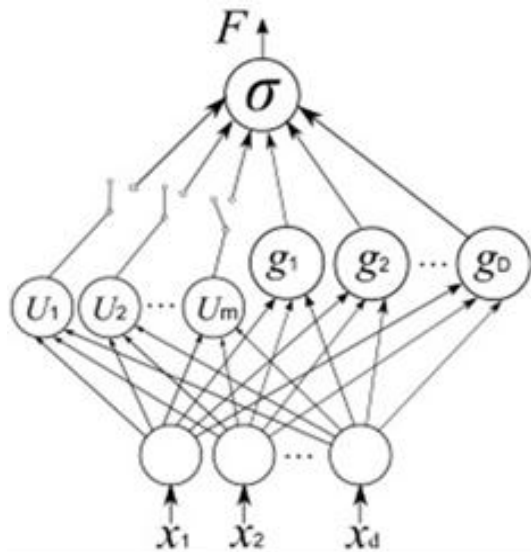


図1 神経回路網の構造

4. 研究成果

(1) 神経回路網に提示される教師信号が(0,1)の2値信号であることによる学習の困難は、隠れ層素子の内部パラメータの削減によりある程度克服可能であることが確認された。この方法を追及する。

しかし、各状態における確率分布がかなり分離されている場合、信号 x にたいする事後確率は0または1に近く、パラメータ空間における勾配はとんどの領域で0に極端に近くなる。しばしば e^{-300} またはそれ以下になる。このような場合、勾配降下法による学習は殆ど進まない。

学習における初期値の選択に失敗すると、信号の確率分布の如何に関わらず、このような現象が起きて、学習が進まない。学習前のパラメータの初期値選択の問題は、重要な問題として今後に残った。

(2) 単一の神経回路網によって、複数の判別関数の学習できることも確認された。シミュレーションで用いられた神経回路網の構造を図1に示す。この神経回路網には上に述べたように複数のモジュール U_1, U_2, \dots, U_m が付加されている。 g_1, g_2, \dots, g_o は非線形素子で本来の隠れ層素子である。出力素子の活性化関数はロジスティック関数 $\sigma(t) = (1+e^{-t})^{-1}$ である。この構造の神経回路網を採用した理由は、ベイズ判別関数の変換が、出力素子の内部ポテンシャルへの補正項の線形加算で済むこと、さらに、以下に述べるように、隠れマルコフ鎖を推定する際に必要な事後確率の変換にも同様の方法が使えることによる。この研究の副産物として、神経回路網が学習したベイズ判別関数をマハラノビス判別関数に変換するアルゴリズムを発見したが、その際にも、この構造が本質的な役割を果たす。

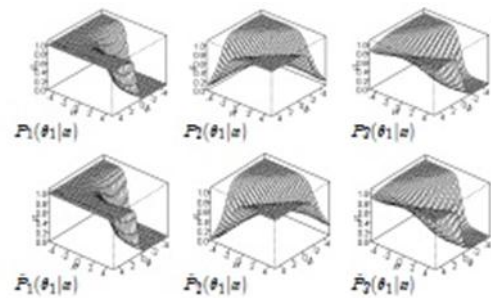


図2 回路網の学習結果

(3) この神経回路網の学習結果の一例を図2に示す。3個の事後確率が学習され、図1において、 $m=3$ に相当する。信号は2次元ユークリッド空間の点で、確率分布が正規分布であるこの場合は $D=3$ である。本来の隠れ層素子は、各事後確率の推定に共通して使われ、

個々のモジュールは、学習結果をそれぞれの事後確率に変換する際に必要な補正項の記憶に用いられた。図2の上段は理論的に求められた事後確率、下段は学習を終えた神経回路網の出力である。

以上のシミュレーションによって得られた判別関数と理論的に得られた判別関数による信号の分類結果を比較したが、両者の結果は98%以上の信号において一致した。確率分布に重なりがあるので、これは信号が正しく分類されたことを意味するものではないが、学習が順調に進んだことを意味する。以上、詳しくは以下の発表論文に記載されている。

(4) 隠れマルコフ鎖推定のシミュレーションは、信号が低次元の場合に成功している。神経回路網が学習するのはマルコフ鎖の平衡分布を事前確率とする一組の事後確率である。この事後確率を変換して、現事象の事後確率を得る。その際、先行事象の推定分布と推定遷移確率により事前確率を求めて使用する。その結果、現事象の推定確率分布が得られる。ここでも、出力素子の活性化関数がロジスティック関数であることが重要で、変換は出力素子の内部ポテンシャルに常数を加算すれば済む。

これを連続的に繰り返して、各段階における事象の事後確率が求まり、隠れマルコフ鎖が推定される。このアルゴリズムはいわゆる前向きアルゴリズムを神経回路網により実現したものである。推定結果は、事象の平衡分布を単純に事前確率として用いた推定に比べてかなり良好で、多くの場合、理論的に計算した前向きアルゴリズムによる推定と殆ど変わらない結果であった。しかし、信号の確率分布や不適切な回路網のパラメータの初期値選択によって、学習が順調に進まない場合もあり、問題が残っている。以上の結果は、近く、国際学会にて発表の予定である。ここで用いたアルゴリズムは原理的に高次元信号の場合に適用可能である。

(5) 当研究において構築した神経回路網によりマハラノビス判別関数の学習が可能であることを発見した。信号の確率分布が正規分布の場合、学習されたベイズ判別関数は容易にマハラノビス判別関数に変換される。ここでも、出力素子の活性化関数がロジスティック関数であることが重要である。さらに、信号の確率分布が正規分布でない場合にも、中心極限定理と大数の法則を利用したアルゴリズムによりマハラノビス判別関数の学習は可能である。当研究の副産物としての成果なので、ここでは詳細を省くが、以下の発表論文の一部に詳しく記載されている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計6件)

- ① Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Simultaneous learning of several Bayesian discriminant functions by a neural network with additional nodes, Australian Journal of Intelligent Information Processing Systems, 査読有, Vol.11, 2010, 1-7.
- ② Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Simultaneous learning of several Bayesian and Mahalanobis discriminant functions by a neural network with additional nodes, Proceedings of IJCNN2011, 査読有, 2011, 733-740.
- ③ Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, A new algorithm for learning Mahalanobis discriminant functions by a neural network, Lecture Note in Computer Science, Vol.7063, 査読有, 2011, 590-605.
- ④ Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Learning of Mahalanobis discriminant functions in extremely non-normal distribution cases, Proceedings of 22th JNNS, 査読有, 2012, 2 pages.
- ⑤ Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Learning of Mahalanobis discriminant functions by a neural network with memory nodes. Lecture Note in Computer Science, 査読有, Vol.7667, 2012, 590-605.
- ⑥ 泉寛幸, 伊藤嘉房, Cidambi Srinivasan, 神経回路網によるマハラノビス判別関数の学習——信号が正規分布にしたがわない場合——, 電気通信情報学会, 査読無, Vol.112, 2012, 79-84.

[学会発表] (計6件)

- ① Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Simultaneous learning of several Bayesian discriminant functions by a neural network with additional nodes, ICONIP(国際神経情報処理学会), 2010年11月22日, Sydney (Australia).
- ② Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Simultaneous learning of several Bayesian and Mahalanobis discriminant functions by a neural network with additional nodes, IJCNN(国際総合神経回路網学会) 2011年8月1

- 日、San Jose (America).
- ③ Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, A new algorithm for learning Mahalanobis discriminant functions by a neural network, ICONIP (国際神経情報処理学会), 2011年11月15日、上海(中華人民共和国).
 - ④ Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Learning of Mahalanobis discriminant functions in extremely non-normal distribution cases, Proceedings of 22th JNNS, 2012年9月12日、名古屋工業大学
 - ⑤ Yoshifusa Ito, Hiroyuki Izumi, Cidambi Srinivasan, Learning of Mahalanobis discriminant functions by a neural network with memory nodes. ICONIP (国際神経情報処理学会), 2012年11月22日、Doha (Qatar).
 - ⑥ 泉寛幸, 伊藤嘉房, Cidambi Srinivasan, 神経回路網によるマハラノビス判別関数の学習——信号が正規分布にしたがわない場合—— 電気通信情報学会 NC 研究会、2012年12月12日、豊橋技術科学大学

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計0件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

[その他]

ホームページ等

6. 研究組織

(1) 研究代表者

伊藤 嘉房 (ITO YOSHIFUSA)
愛知医科大学・客員教授
研究者番号：22500213

(2) 研究分担者
()

研究者番号：

(3) 連携研究者
()

研究者番号：