

論文の内容の要旨

論文題目 Bayesian theory for kernel machines and network models

(カーネルマシン及びネットワークモデルのベイズ理論)

氏名 小林 景

本論文では、カーネルマシンやネットワークモデル、回帰分析等のベイズ理論とその応用を中心として、著者が博士課程で行った研究内容を紹介している。本論文で扱った主題は大きく以下の4つに分けられる。

1. 楽観性, 悲観性を導入したネットワーク決定問題の効率的な解法

ネットワークモデルにおける意思決定問題は、確率モデルと効用関数および観測値が与えられたときに、効用関数の条件付期待値を最大化するような決定を求める理論である。特に、確率モデルや効用関数の独立関係を表すグラフの形状を生かして、最適決定を効率よく計算することができる。

しかし、正規分布に従う変数のみでモデルが構成されるときに計算方法は知られているが、離散変数を含む一般のモデルでの最適な決定を、グラフの特徴を生かして効率的に計算する方法は知られていなかった。

本研究では、まず正規変数と離散変数とともに含むようなモデルにおいて、**Junction Tree Algorithm** と呼ばれる手法を応用した最適決定アルゴリズムを提案し、グラフの特徴を利用した計算量の減少を実現した。その際に **Lauritzen** 等により提案された **strong root** の議論を用いた。さらに線形指数二次ガウス (**LEQG**)モデルと呼ばれる最適制御理論をグラフィカルモデルに拡張することによって、決定者が楽観性, 悲観性をもつ場合の最適決定の理論を構築した。

次に、ロバスト性をもつ制御として知られる H^∞ 最適制御理論を拡張した、一般化 H^∞ 最適決定の理論を提案した。さらに一般化 H^∞ 最適決定が、楽観性をもつモデルの最適決定列のある極限と一致することを証明した。また、コレスキー分解を用いることにより、一般化 H^∞ 最適決定を直接求める方法も示した。

本研究は離散変数と連続変数の混在モデルの最適決定アルゴリズムを提案し、楽観性を決定問題に導入したことによって、ネットワークモデルの最適決定の理論を大きく拡張することに成功した。

2. 情報量規準によるカーネルマシンのパラメータ選択

判別分析およびパターン認識に用いられるサポートベクトルマシンとカーネルロジスティック回帰の二つのカーネルマシンに関して、情報量規準を用いた正則化パラメータの選

択手法を提案した。

サポートベクトルマシンは、その汎化能力の高さと二次計画に問題が帰着できる点から現在最も注目されている判別学習機械のひとつである。一方カーネルロジスティック回帰は、損失関数の連続性から自然な統計モデルを構成できるうえに、近年分解アルゴリズムによる計算手法が提案されたことにより応用上も注目されている学習機械である。しかし、その性能は正則化パラメータと呼ばれるパラメータの値に大きく依存する。よって、そのパラメータの最適化が大きな課題となっている。

本研究では、以上の二つのカーネルマシンのベイズ統計モデルを構成する。さらに、それを正則化統計モデルと解釈し、正則化情報量規準 (R I C) を用いてカーネルマシンのパラメータを最適化することを提案した。ただし、カーネルマシンでは、カーネル関数とよばれる正定値関数を用いて特徴空間の内積のみを定義するため、R I Cの直接の計算は不可能である。そこで本研究では、R I Cの計算を一旦ある固有値問題に帰着することで、その導出を可能にした。しかし、R I Cの計算にはサンプル数の3乗オーダーの計算量が必要であるという問題点があった。そこで、本研究では、さらにグラム行列の近似計算に用いられる *Nystrom* 法をR I Cの計算に用いることを提案し、サンプル数の線形オーダーの計算量での近似計算を可能にした。

実験の結果、R I Cによるパラメータの最適化は、*Wahba* の *GACV* や *xi-alpha* アルゴリズムなどの高速なパラメータ選択規準や、ベイズ統計学を用いた *evidence* による規準に比べて、有意にテスト判別率が高いことが示された。さらに、広く用いられる *10-fold* クロスヴァリデーションと同程度の判別性能を、数百程度のサイズの訓練データセットに関して、10倍~1000倍の速さで実現可能であることが示された。

3. 回帰分析の縮小ベイズ予測理論

本研究では、正規ノイズを仮定した線形回帰分析に関して、ある縮小型事前分布のクラスを用いたベイズ予測分布が、*Kullback-Leibler* 損失に関して、一様事前分布を用いたベイズ予測分布を優越することを証明した。また、その縮小型事前分布のベイズ予測のミニマクス性も示した。さらに、これらの縮小型事前分布のうちのいくつかは説明変数によらないことから、任意の説明変数に関して、一様事前分布によるベイズ予測を優越するような事前分布が構成できる。

ここで、一様事前分布を用いたベイズ予測は、最尤推定値のプラグインを優越することが知られているので、その一様事前分布を用いたベイズ予測をさらに優越する予測を提案することは意味がある。

証明の概要を説明する。まず線形回帰問題を、平均が未知、共分散行列は既知であるが訓練標本と未来の標本で変化する場合の、多変量正規分布の予測問題に帰着させた。共分散行列が一定の場合の多変量正規分布の予測問題に関する、縮小型事前分布を用いたベイズ予測の有効性は *Komaki* や *George* らの研究によって既に知られているが、本研究では彼

らの結果を共分散行列が、訓練標本と未来の標本で変化する場合に拡張した。具体的には、訓練標本の共分散行列および未来の標本の共分散行列によって表されるある条件をみたすような計量のクラスを定義し、そのクラスに属する任意の計量に関する Stein 事前分布を用いたベイズ予測が、一様事前分布を用いたベイズ予測を優越することを証明した。ただし、その証明方法は Green の定理を用いるもので、共分散行列が不変の場合の Komaki や George らの証明とは本質的に異なる。また、同様のクラスに属する計量に関して、密度関数が優調和関数となる事前分布に関するベイズ予測がミニマックスになることも証明した。

以上の結果を用いると、回帰分析に関する縮小ベイズ予測の有効性も示すことができる。縮小事前分布による Kullback-Leibler リスクの改良は、データが高次元であるほど大きくなるので、一般に高次元な特徴空間をもつカーネルマシンへの応用が期待される。特に Stein 事前分布による縮小は、現在広く用いられているリッジ回帰の縮小に比べて、真の平均パラメータに関するロバスト性があることが本研究の結果からいえる。

4. トーナメントによる全順位決定問題の組み合わせ論的解析

スポーツの大会などで通常行われるトーナメントでは、実力が一番の選手と二番の選手が一回戦で対戦してしまうと、二番の選手は全体の半数以下の順位になってしまう。本発表では、このような問題を解決するために、並列マッチング法とよばれる新しいトーナメント構成方式を提案した。並列マッチング方式では、実力が上の選手が必ず勝つと仮定すると (sure winner model) , 最終順位が実力通りに並ぶ。平行して複数回の試合が行われることも可能であるという仮定の下で、並列マッチング方式に必要なラウンド数の期待値が、チーム数の平方根のオーダーであることを証明した。これは、総当り戦による決定方式に必要なラウンド数がチーム数の線形オーダーであるのに対して、並列マッチング方式がより効率的な順位決定方式であることを意味する。

証明の際には、ラウンド数の期待値を求める問題が、確率論の破産問題および投票問題に帰着されることを利用した。

一方、試合の結果が確率 $1/2$ のベルヌイ試行に従う場合には (totally random model) , 並列マッチング方式に必要なラウンド数の期待値が、チーム数の対数の平方根のオーダーであることを証明した。また、より現実的なモデルとして、sure winner model と totally random model モデルの中間にあたる Bradley-Terry モデルの並列マッチング法のラウンド数の期待値を数値的に評価した。

この問題の解析には、グラフ理論の Hasse 図や統計学の極値理論、Kolmogorov-Smirnov 統計量の知識を用いるという点が特徴的である。また、本研究は並列計算アルゴリズムの理論と共通点を持ち、並列計算の新しいアルゴリズムの開発および理論の構築につながることを期待される。