

変分ベイズにおける最適解探索効率の検証

岡 威久馬

近年、確率モデルの学習アルゴリズムを用いたデータ分析が注目されている。ビッグデータの背後に隠れたパターンを発見し、解釈可能な知識を抽出することは重要な課題である。この課題は、対象とするデータに応じて、類似したデータのグループを発見するクラスタリングや、系列データからパターンを見つける系列データマイニングなどと呼ばれている。このような課題に対して、確率モデルによる統計的データ分析に基づく手法が有効と考えられており、クラスタリングをおこなう際に用いられる混合ガウスモデル(GMM)や系列データを扱う際に用いられる隠れマルコフモデル(HMM)のようなモデルが提案されている。

しかし、これらのモデルを用いたデータ分析では、局所最適解に陥るという問題を抱えている。この問題を解決するための手法として摂動を加えるというアプローチが存在する。例えば、反復的数値最適化手法であるEMアルゴリズムや変分ベイズにおいては、期待値のモンテカルロ近似や、確率的勾配降下法を用いた手法などが提案されている。これらの手法は局所最適解から抜け出すためのアプローチとして有効であるということは示されているが、それぞれの手法がどれだけ効率的な最適解探索をおこなっているかは示されていない。本研究では、変分ベイズにおいて、摂動を加えない通常の変分ベイズ、期待値のモンテカルロ法による近似、確率的勾配法の3つの手法を比較した。その結果、期待値のモンテカルロ法を用いた近似が最も良い性能を示した。

今後の課題としては、GMMやHMM以外のモデルにおける挙動を調査することや、本研究で用いた手法以外の手法を用いて比較実験をおこない、一般的な結論を得ることである。

(指導教員 若林啓)