Asymptotic Theory of Empirical and Variational Bayes Learning
（経験および変分ベイズ学習の漸近論）

論文要旨

東京工業大学
総合理工学研究科
知能システム科学専攻

中島　伸一

2006

パラメータ空間のすべての点において Fisher 情報行列が正定値を持つモデルは正則モデルと呼ばれ，その汎化特性を一般的に説明する理論が確立されている．しかし，近年，機械学習の分野において利用される多くのモデル，例えば神経回路網，ベイジアンネットワーク，混合分布，隠れマルコフモデルなどは，階層的なパラメータ構造を持つため，Fisher 情報行列が縮退する特異点を持っている．このようなモデルは特異モデルと呼ばれ，正則モデルにおいて成立する一般的な理論が成立しないことが知られている．特異モデルの解析は一般に困難であったが，近年の研究により，いくつかの性質が明らかにされている．特異モデルの最も重要な性質のひとつとして，以下が挙げられる．特異モデルにおいては，学習方法として最尤法を用いた場合には過学習しやすく，ベイズ法を用いた場合には過学習しにくい．

従って，特異モデルの学習においてはベイズ法を用いることが望ましいのであるが，ベイズ事後分布を正確に実現することは難しい．特異モデルのベイズ事後分布を精度良く近似する方法としては，マルコフ鎖モンテカルロ（Markov chain Monte Carlo）法がよく利用される．この方法は任意のモデルに対して用いることができるが，多くの計算量を必要とすることが欠点であると言える．これに対し，近年，より少ない計算コストでベイズ事後分布を近似する方法として，変分ベイズ（variational Bayes）法が提案され，注目を集めている．多くのアプリケーションにおいて，最尤法を実現する目的で提案された expectation-maximization（EM）アルゴリズムと比較して，変分ベイズ法が良い汎化性能を示すことが実験的に知られているが，その性質は理論的には十分解明されていない．事後分布の近似精度の指標となり，また，モデル選択等でも利用される変分自由エネルギーの漸近形が，いくつかのモデルにおいて解明されているが，ベイズ法の場合と異なり，この結果が直ちに汎化誤差の漸近形を与えるということにはなっていない．これまでのところ，すべての特異モデルにおいて，変分ベイズ法の汎化性能は理論的に解明されていなかった．

本論文の目的のひとつは，最も簡単な特異モデルである 3 層線形神経回路網（linear neural networks）において，変分ベイズ法の振る舞いを解析し，理論的にその汎化誤差を求めることにある．また，経験ベイズ（empirical Bayes）法を拡張することによって導出した別の近似法である，部分空間ベイズ（subspace Bayes）法についてもその振る舞いを解析する．部分空間ベイズ法は，変分ベイズ法ほど適用範囲が広くはないのであるが，その解析結果は，変分ベイズ法の解析結果を解釈する上で重要な知見をもたらす．

本論文ではまず始めに，部分空間，および変分ベイズ法を線形神経回路網に適用し，その解析解を導出する．次に，汎化性能を特徴付ける量，すなわち，汎化誤差，学習誤差，自由エネルギーを理論的に解明する．線形神経回路網においては，最尤法およびベイズ法の汎化性能が既に解明されており，それらと比較することによって，ベイズ法の近似法が，どのような場合に，どのような理由で良い汎化性能を示すのかについて議論することができる．本論で明らかになった主要な知見を以下に挙げる．

1. 3 層線形神経回路網において，部分空間ベイズ法および変分ベイズ法は Jame-Stein 型縮小推定と漸近的に等価であり，互いに似た結果をもたらす．従って，真の分布を実現するために必要な成分は最尤推定量と漸近等価になるが，冗長な成分は漸近的にも等価ではない．
2. 多くの場合，ベイズ法に匹敵する汎化性能を示す．
3. ベイズ法に似た特徴だけはでなく，最尤法に似た特徴も併せ持つ．特に，特異点上で 0 でない事前分布を用いた場合，特異モデルの汎化誤差が正則モデルの汎化誤差を漸近的に超えることがないという，ベイズ法の最も良い性質のうちのひとつはこれらの近似法では成立しない．
4. ベイズ法においては自由エネルギーと汎化誤差の漸近的振る舞いには単純な関係があるが，変分ベイズ法においては，それらの振る舞いは大きく異なる．特に，自由エネルギーは定義によりベイズ法の場合より小さくなることはないが，汎化誤差はベイズ法よりも小さくなる場合がある．

本論文ではさらに，真の分布と特異点とのカルバック擬距離が学習データ数の逆数と同程度の大きさであるような場合（*delicate* situations）についても解析する．特異モデルにおいて，このような場合についての解析は困難であることが多いため，漸近解析においては考慮されない場合が多いが，実際に有限サンプルを用いて統計的推測を行う上では非常に重要である．本論では，このような場合の線形神経回路網における汎化性能を，部分空間ベイズ法，変分ベイズ法において解明した．なお，この解析は本論文の結果が，ベイズ法の許容性と矛盾しないことを示すためにも必要である．