

Proper reduction and asymptotic sampling difficulty of posterior distributions in fully Bayesian inversions

*Daisuke Sato¹, Yukitoshi Fukahata¹

1. DPRI, Kyoto University

前回の地震学会で、モデルパラメータ数が大きくなるにつれて従来のフルベイズ推定と赤池ベイズ情報量規準 (ABIC) によるベイズ推定とでモデルパラメータ推定値が乖離し、これまでフルベイズの近似とわれてきたABICの方がより適切な推定値となるという問題を報告した (佐藤・深畑, 2019). 本発表では、その原因となっていた同時事後分布の性質を精査する。

ベイズ推定では、観測方程式とモデルパラメータに関する事前分布とをベイズの定理を介して結合し、事後分布を構成する。フルベイズ推定では、さらにモデルパラメータの事前分布を規定する超パラメータについてもその事前分布 (超事前分布) を考慮し、モデルパラメータと超パラメータに関する同時事後分布を構成する。

ここで、フルベイズ推定の本来の解は同時事後分布であるのだが、多次元の確率分布そのままでは有用な情報を引き出せない (松浦, 1990). そこで、同時事後分布からモデルパラメータの有用な統計量を得るために情報の縮約が必要となる。適切な縮約の方法は自明ではなく、少なくとも3つの流儀がこれまで用いられてきた。1つは、同時事後分布から超パラメータとモデルパラメータの組を直接的に同時推定する。この立場では、事後確率最大 (MAP) 解が推定値として通常採用される。もう1つは、同時事後分布からモデルパラメータのみを推定する。これは、超パラメータに関して事後分布を積分し、モデルパラメータに関して周辺化した事後分布を用いることに相当する。Fukuda & Johnson (2008) などフルベイズ推定を行っている研究では、しばしばこの指標が用いられている。そして、最後の1つはABICで、モデルパラメータを積分することにより超パラメータに関して周辺化された事後分布を求め、その分布を基にモデルパラメータを推定する。これは、同時事後分布を超パラメータの周辺事後分布とモデルパラメータの条件付き事後分布とに因数分解し、これら2つの確率値に基づく二段推定を行うことに相当する。この方法は、地球物理では広く用いられてきた (Yabuki & Matsu'ura, 1992).

これらの縮約に関して、モデルパラメータ数の増加につれて、モデルパラメータの周辺事後分布がMAP解にデルタ関数的に収束する一方で、超パラメータの周辺事後分布はABIC推定値にデルタ関数的に収束することを我々は解析的に導いた。これは、ABIC型の多段推定を行うかどうかでフルベイズ推定のモデルパラメータ推定値がABIC型とMAP型に二分されるということを示す。

それらの二分された推定値のうちどちらが良いかをsynthetic testで検討した。検討の際は、roughnessが小さいという事前分布を用いた。その結果、従来の報告通り、同時事後分布が、MAP解として通常採用される局所最大に加え、underfit解 [モデルパラメータ数が大きい場合にはさらにoverfit解 (伊庭, 1996)] に大域最大を持つ多峰分布であることを確認した。また、モデルパラメータ数が大きい場合に局所最大解がunderfit側にバイアスされ、さらには消失していくことも新たに観察された。対照的に、超パラメータの周辺分布 (ABIC) は、一貫して最適値付近にピークを持つ単峰分布だった。

このようにしてABIC型の多段推定がそれ以外の縮約された分布に比べてより適切だとわかった。しかし、いずれの縮約も同時事後分布から生成される分布関数を用いている以上、これらの分布から得られる平均値は同一でなければならない。この点で、三通りに縮約された分布関数が異なる推定値にデルタ関数的に収束するとした我々の結果は、一見奇妙である。

そこで、事後平均 (EAP) 推定値の解析的表現を導いた。その結果、モデルパラメータのABIC推定値とEAP推

定値とが同様の値をとることが見出されると共に、EAP推定値の同時事後分布での出現確率がモデルパラメータ数の指数関数として急減少していくことが明らかになった。従来、モンテカルロサンプリングは多次元標本空間がパラメータ数の指数で増加する問題（次元の呪い）を確率値の高い事象のみをサンプルすることで克服するとしたが、同時事後分布では確率の高い推定値（MAP解近傍）は適切でない以上、この処方是有効に機能しない。従来のモンテカルロサンプリングによるEAP推定値の数値近似解はMAP解の側にバイアスされていた可能性が高い。

一方、超パラメータの周辺事後分布については、ほぼ無限大の個数で且つほぼ確率ゼロの事象が平均を支配するという同時事後分布の困難を、字義通り無限個の事象を積算することで解決していた、と考えられる。MAPのABICからのずれに相当する、同時事後分布と超パラメータの周辺事後分布の対数差の主要項が、シャノンエントロピーと呼ばれる量（確率変数に対して拡張された事象の個数の対数（エントロピー）を表す量）であることがわかり、この予想が裏付けられた。