

On the Optimum Solution of Bayesian Inversions

*Daisuke Sato¹, Yukitoshi Fukahata¹

1. DPRI, Kyoto University

種々の逆解析で得られる推定値は往々にして食い違う。本発表でも述べるように、モデルパラメタの事前分布と観測方程式を同様に与えた場合でも、最適な推定値の分布(逆問題の最適解)を得るためのモデルパラメタおよび超パラメタの推定法は一意ではない。本研究で我々は特に、地球物理で頻出する観測方程式が連続自由度で記述される逆問題に焦点を当て、この場合に最適解を得るための統計学的指標を明らかにしたい。

地球物理で逆問題を取扱う際にはベイズ推定の枠組みがよく用いられる。ベイズ推定では、観測データとモデルパラメタの関係を記述する観測方程式と、モデルパラメタに関する事前分布とをベイズの定理を用いて結合し、事後分布を構成する。さらに、フルベイズと呼ばれる枠組みでは、モデルパラメタの事前分布を規定する超パラメタについてもその事前分布(超事前分布と呼ぶ)を考慮し、モデルパラメタと超パラメタに関する同時事後分布を構成する。但し、多くのフルベイズの解析では、超事前分布として無情報事前分布(一様分布)が用いられており、その場合、フルベイズと超事前分布を考慮しないベイズとで、同時事後分布の関数形は同じになる。この事後分布からモデルパラメタおよび超パラメタの最適値を得るには情報の縮約が必要となる。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ている。まず1つ目が、超パラメタおよびモデルパラメタの事後確率最大(MAP)を最適値とみなすものである。これは最も直観的な方法と言えるだろう。もう1つが、モデルパラメタを積分することにより、超パラメタに関して周辺化された事後分布を求め、その最大値を最適値とするというものである。赤池ベイズ情報量基準(ABIC)そのものであり、地球物理では幅広く用いられてきた(Yabuki & Matsu'ura, 1996)。なお、ABICで得られる超パラメタの推定値は一致性(真値への漸近収束性)という優れた性質を持つことが知られる(渡辺, 2005)。そしてもう1つは、超パラメタに関して事後分布を積分することにより、モデルパラメタに関して周辺化した事後分布を用いるものである。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。(そのため、この方法により最適値を推定することを、MAP, ABICと対置して以後従来のフルベイズと呼ぶ。)

MAP、ABIC、従来のフルベイズは、その推定値からのばらつきの分布関数(分散など)を含め、多くの問題で似たような結果を与える(e.g., Fukuda & Johnson, 2008)。しかし、モデルパラメタの数を多くした場合にはこれらの解に明瞭な乖離が生じることを発見したので報告する。

まず無情報超事前分布でのフルベイズ推定に関して、MAPおよび従来のフルベイズでの最適解(およびばらつきの分布)の解析的表現を線形逆問題に対して導出した。導出した解析解から、モデルパラメタ自由度とデータ自由度が大きくなるにつれ、モデルパラメタに関して従来のフルベイズの推定値がMAPの推定値に漸近一致することが明らかとなった。さらに、それらの実効的超パラメタの値はABICによる推定値と異なっていることが見出された。

次に、ABICとそれ以外の場合とで得られた2つの異なる推定値のうち、どちらが良い推定値かをsynthetic testで検討した。検討の際は、roughnessが小さいという事前分布(Yabuki & Matsu'ura, 1992)を用いた。検討の結果、MAP推定の分布は、離散化のメッシュサイズを細かくするにつれ、データにオーバーフィットした解にデルタ関数的に収束すること、つまりモデルパラメタの事前分布をほとんど無視する傾向があることがわかった。これは正則化パラメタを0にした場合に劣(混合)決定問題となる設定で一般に生じる問題であることがMAP推定の性質から示唆された。MAP推定でのオーバーフィット傾向はすでに統計学界限では認知されており、ABIC提案の背景となった経緯があるらしい(伊庭, 1992)。一方、ABICを用いた場合には上記の問題は生じず、メッシュサイズを細かくした(離散化誤差の小さい)高解像度な解がフルベイズ推定よりも有意に真の解に

近くなることがいくつかの数値解析例から示された。現時点では、ABICを基準とした最適値推定が最良と示唆される。

誤差を減らしたつもりのメッシュの細かい解にオーバーフィットによる精度劣化が隠れているということは、真値との残差を知り得ない実問題への応用では致命的となりうる。ABICよりも一般的な枠組みと思われるMAPや従来のフルベイズ推定でなぜこのようなことが起きたのだろうか。ラプラス法を用いた鞍点解析を行うと、フルベイズ推定では(一致性を持つ)超パラメタの期待値とMAP/従来フルベイズでの超パラメタの実効値とが(線形非線形にかかわらず)一般に漸近的にすら一致しないという、MAPおよび従来のフルベイズに関する極めてネガティブな性質が導き出された。これがMAPおよび従来のフルベイズ推定でのモデルパラメタのオーバーフィットを生じさせていると考えられる。