

## 「特集 データ同化の方法」について

上野 玄太<sup>1,2,3</sup> (オーガナイザー)

データ同化は、観測データを用いてシミュレーションを改善する方法である。「データ同化」という言葉は、シミュレーションありきの前提で定義され意味を持つ。データ同化をしないのが通常のシミュレーションであり、データ同化の操作を追加することでシミュレーションの精度を改善することが可能になる。

この改善の方法にも歴史があり、古くはシミュレーションによる変数の計算結果を観測データと置き換えた上で、予報の計算を進めるといった素朴なものであった。観測データで置き換えるのだからその時点の変数の値としては完璧で、それゆえその後の予報の精度も高い、とはならないところが悩ましいところである。その理由は、シミュレーションが実際の現象を表現するには力不足であるから、また、観測データには誤差が含まれるからである。そのため、一度に置き換えずに複数の時間ステップに分割し徐々に置き換える方法、そして現在主流であるベイズ的なアプローチへと変遷を遂げてきた。

データ同化の実行に際しては、計算コストやシミュレーションの複雑さなどに応じて、いくつかのアルゴリズムがある。実装が平易なアンサンブルカルマンフィルタ、超大規模なシミュレーションも扱える4次元変分法、非線形性の強い現象にも柔軟に対応する粒子フィルタなどが代表的である。アンサンブルカルマンフィルタ・粒子フィルタは各時点の状態の周辺分布を推定するためのアルゴリズム、4次元変分法は全時点の状態の同時分布のMAP解を得るためのアルゴリズムである。

国内でデータ同化の教科書が続けて出版された(露木・川畑, 2008; 淡路 他, 2009; 樋口, 2011) 時期から10年が経ち、データ同化の利用者や応用分野の拡大が見られる。本号では、現在のデータ同化の方法についての記事を特集した。

川畑論文、三好論文は、粒子フィルタによるデータ同化を扱っている。シミュレーションは一般的に非線形微分方程式に立脚しているため、非線形状態空間モデルの状態推定の問題としてとらえ、粒子フィルタを適用することは自然である。アンサンブルカルマンフィルタのようにみなしガウス分布のような便宜的な操作を施さず、シミュレーションによる状態変数の非線形発展をそのまま生かす粒子フィルタの特徴は魅力である。しかし、粒子フィルタにより適切に状態推定を行うためには粒子数 $N$ を大きくとる必要があることから、大きな計算コストが要求されることが問題である。川畑論文は、観測ノイズの共分散行列を動的に推定した上で粒子フィルタを実施する方法を提案するとともに( $N = 1000$ )、得られる確率分布の解析の方法についても述べている。三好論文は、計算コストを要しないシミュレーションを採用することで、粒子フィルタの適用に加え、より大きい粒子数を必要とする粒子スムーザを適用し( $N = 10000$ )、推定されたシミュレーション内のパラメータの時間変化の物理的考察につなげ

<sup>1</sup> 統計数理研究所：〒190-8562 東京都立川市緑町 10-3

<sup>2</sup> 総合研究大学院大学 複合科学研究科統計科学専攻：〒190-8562 東京都立川市緑町 10-3

<sup>3</sup> データサイエンス共同利用基盤施設 データ同化研究支援センター：〒190-0014 東京都立川市緑町 10-3

ている。

アンサンブルカルマンフィルタは変数の分布をガウス分布とみなして推定を行うことで、粒子フィルタより少ない粒子数  $N$  (アンサンブルメンバー数と同義) でも精度が向上しやすい。  $N = 200$  でのアンサンブル計算が可能であれば、アンサンブルカルマンフィルタがそのまま有効に機能すると予想する。  $N$  が十分ではない (100 以下) 場合は、アンサンブル共分散行列にサンプリング誤差が含まれることを考慮して、見せかけの相関を小さくする共分散局所化の操作を追加して対処する。相関関係を認めるのは地理的に近い場所で定義された変数間に限定し、離れた2地点の変数間の相関 (共分散) は積極的に減少させることで、データ同化の効果が適切に得られるようにする操作である。樋口 (2011) による記法を用いれば、一期先予測のアンサンブル共分散行列  $\hat{V}_{t|t-1}$  の代わりに、地理的な距離に応じて減衰する共分散関数  $\rho$  を作用させた  $\rho \circ \hat{V}_{t|t-1}$  を用いてフィルタの操作を行うことに対応する ( $\circ$  は要素積)。関山論文では、  $N = 20$  でアンサンブルカルマンフィルタを実施するに際して、この共分散局所化の考え方を進めて、変数局所化を検討している。変数局所化とは、物理法則上関連が低い変数間の相関をゼロとする操作である。

4次元変分法では、MAP解を得るために同時分布の最大化と同値な目的関数の最小化問題を解く。これは非線形最適化を行うことに相当し、準ニュートン法や共役勾配法が用いられることが多い。このような勾配法のアルゴリズムを用いて、より少ない反復回数で解への収束を達成するには、変数変換をして目的関数のヘッセ行列の条件数を1に近づける、いわゆる前処理が有効である。勾配の計算にはシミュレーションの時間積分が必要であるため、シミュレーションが超大規模であるという理由で4次元変分法を選択している場合は特に、反復回数の減少は重要である。前処理には、背景誤差共分散行列 ( $B$  と書かれることが多い (露木・川畑, 2008; 淡路 他, 2009) が、樋口 (2011) の記法では4次元変分法の場合は  $V_{1|0}$ 、3次元変分法の場合は  $V_{t|t-1}$  である) の平方根行列  $U$  を用いることが多い。石橋論文は、背景誤差共分散行列を予測アンサンブル共分散行列と局所化を用いて  $\rho \circ \hat{V}_{t|t-1}$  と設定する場合に、  $\rho \circ \hat{V}_{t|t-1} = UU'$  を満たす  $U$  の一般的な定式化のレビューである。

さて、4次元変分法はMAP解を得ることに注力したおかげで超大規模なシミュレーションに対してもデータ同化を可能にしたが、MAP解  $x_t^*$  以外の確率分布の情報は得られない。しかし、MAP解以外の情報、具体的には、MAP解周りの分散・共分散の情報は特に望まれる。丹羽論文では、4次元変分法でMAP解を推定する際に用いる準ニュートン法 (BFGS 公式) のアルゴリズムで用いる変数の利用により、副次的にMAP解周りの状態変数の共分散行列を推定する方法を紹介している。この行列は、線形ガウス状態空間モデルの場合はフィルタ共分散行列  $V_{t|t}$  に対応するものである (このとき、MAP解  $x_t^*$  はフィルタ平均ベクトル  $x_{t|t}$  と一致する)。藤井論文では、アンサンブルメンバー  $x_{t|t}^{(i)}$  を、同じく準ニュートン法 (BFGS 公式) を用いて生成する方法を提案している。アンサンブルメンバーはデータ同化以降時点のアンサンブル予報に利用できるが、予報の時間積分には計算コストがかかるため、予報結果が散らばるメンバーを優先的に生成し、予報幅を得たいという狙いがある。

一方、アンサンブルカルマンフィルタは、観測モデルがガウス分布でない分布で表現されている場合には原理的に適用が難しい。中野論文では、観測モデルがポアソン分布で表現される場合を想定し、アンサンブルを用いた4次元変分法のアルゴリズム (同時分布のMAP解を求めるアルゴリズム) を提案している。通常の4次元変分法は、シミュレーションコードに即した勾配計算プログラム (アジョイントコード) を書く必要があるが、アンサンブルを用いる提案法では、シミュレーションコードはブラックボックスとみなして勾配を求め、目的関数の最小化が可能である。

冒頭で述べたように、データ同化はシミュレーションを立脚点として、推定精度・予報精度

を上げるために、フィルタや最適化の手法を取り入れて発展してきた。杉浦論文は、まだデータ同化では取り入れられていない手法として、シグネチャ法の入門を紹介している。シグネチャ法の利用により、系列データおよび経路の視点での観測モデルの設計、目的関数の設定の可能性が示唆されている。

本特集の編集に当たっては、原稿の執筆を引き受けていただいた著者の方々にまずお礼申し上げます。共同研究、共同研究集会で顔を合わせる研究者に声を掛けさせていただいた。同時に、査読者の方々に何よりお礼申し上げます。データ同化は手法と応用分野が交錯する分野であるが、今回は手法に関する記述に重点を置いた特集号とするために、原稿の著者とはやや異なる専門分野の研究者に査読を引き受けていただいた。本特集号で、上野が共著者となっている原稿については、統計数理研究所の中野慎也准教授に編集を協力いただいた。合わせてお礼申し上げます。

### 参 考 文 献

- 淡路敏之, 蒲地政文, 池田元美, 石川洋一(編著) (2009). 『データ同化—観測・実験とモデルを融合するイノベーション—』, 京都大学学術出版会, 京都.
- 樋口知之(編著) (2011). 『データ同化入門』, 予測と発見の科学, 6, 朝倉書店, 東京.
- 露木義, 川畑拓矢(編著) (2008). 『気象学におけるデータ同化』, 気象研究ノート第 217 号, 日本気象学会, 東京.