

樋口 知之

情報・システム研究機構 統計数理研究所・副所長/教授

先端的データ同化手法と適応型シミュレーションの研究

1. 研究実施の概要

シミュレーションなどの数値モデルによる対象状態の時間発展更新と、装置からの部分的な観測に基づく状態補正の二つを適切に組み合わせる作業はデータ同化と呼ばれる。本研究では先端的なデータ同化手法を開発し、さらにこの技法をもとに、シミュレーションモデルを複数走らせ、データ適応的に数値モデルを切り替え、あるいはそれらを統合するようなメタシミュレーションモデルを創出するプラットフォームを創造することが最終的な目標である。

シミュレーションモデルが含む変数の次元は、数百次元から 100 万次元程度の超大規模なものであり、データ同化の研究はいろいろな意味で計算限界への挑戦である。データ同化は逐次型と非逐次型の方法があり、逐次型においてはアンサンブルカルマンフィルタなどが主に用いられているが、粒子フィルタの適用もわずかながらなされてきている。我々データ同化グループでは、アンサンブルカルマンフィルタ、粒子フィルタ、混合カルマンフィルタを中心に、逐次データ同化とよばれる同化手法の研究と応用を行っている。近年、統計的推測の視点と経済効率的なアプリケーションソフト作成の観点から、逐次型の優位性はますます明らかになりつつある。

実際にデータ同化手法の研究を行うには、具体的なシミュレーションモデルとデータセットの二つ、つまり具体的テーマの選定が必要である。そのテーマは、手法の新展開の観点から未解決の数理的及び数値的課題を内包し、併せて応用分野における知識発見や予測性能向上の観点からも興味深いものでなければならない。この観点から、昨年度に引き続き、大気・海洋、津波、海洋潮汐、宇宙空間(リングカレント)、ゲノムデータ同化の5つの領域における具体的テーマに取り組んできたが、今年度から“ものづくり”データ同化プロジェクトを新たに開始し、また有限要素法へのデータ同化適用プロジェクトも本格的に実施した。このようにさまざまな研究領域においてデータ同化は、特に予測性能向上という視点からシミュレーション技術にブレイクスルーを与えるものとして大いに期待が高まっている。

2. 研究実施内容(文中にある参照番号は 4.(1)に対応する)

プロジェクトも終盤に近づき、メタシミュレーションの開発に重点を置くとともに、これまで成功し

てきた新規応用分野の開発調査を引き続き行った^{3), 6)}。

通常の粒子フィルタは、フィルタリングを繰り返すうちに状態の確率分布をうまく表現できなくなってしまうため、高次元の問題に対してはあまり有効とは言えない。そのため、予測分布アンサンブルから抽出した複数の粒子を組み合わせ、高次元の非線型問題に対してもうまく状態の推定ができる方法を探った⁵⁾。その検討過程で、限界を探る上でも計算機のリソースを最大限に活用し、“試しに”一億個の粒子数を用いて再度パラメータを推定してみた。その結果、既発表の十万個数の粒子フィルタによる推定結果と比較して、格段に推定性能が向上していることが確認できた。特に本年度作成した、超多数粒子の特性を多面的に表示する可視化ソフトウェアによって、一億粒子フィルタの優越性を明瞭に視認することができた。一億個の粒子フィルタを並列計算機に実装するにはそれなりの工夫が必要であるが¹²⁾、この性能向上は予想を大きく上回るものであった。これまで退化の問題に対して数多くの緩和策が提案されてきたが、このスケールのパラメータ推定問題においてこれほどの性能を示す手法は存在しない。現在、粒子フィルタは多岐にわたる応用領域で用いられているが、大抵の場合、利用する粒子数は高々数万オーダー程度であったが、「一億個に増やす」というある意味安易なアイデアによってアルゴリズムのポテンシャルが飛躍的に高まったと言える。

系統的モデル設計法やメタシミュレーションの開発の基礎となるのは、モデルの評価法の研究である。逐次データ同化のアルゴリズムはすべて統計モデルで表現できるので、超高次元統計モデルの尤度計算の高精度化がモデル評価法研究の正攻法である。観測モデル中の共分散行列が正則でない場合、その行列式や逆行列を必要とする尤度の計算が原理的に不可能であり、データ同化モデルの良し悪しの比較ができなかった。そこで、昨年度は特異ではあるものもとの共分散行列の構造をある程度維持しながら、正則性を満たすような行列の設定法を構築した。本年度はそれに続く形で、実際にデータ同化の問題が対象とする数千次元以上の行列に適用するための有効なアルゴリズムを開発した。これにより、大規模なデータ同化システムに対しても尤度の定義と数値的に安定な計算が可能となり、情報量規準に基づくデータ同化モデル比較への道が整備された。

複数のシミュレーションモデルを同時に走らせ、データ適応的にシミュレーションモデルを切り替え・混合する、言わばメタシミュレーション法の研究開発を行った。特に、シミュレーションモデルの固有の特性として境界条件に注目し、シミュレーションモデル毎に境界条件が異なる設定で多数のシミュレーションモデルを競合させた。同種の実験はすでに津波データ同化でも行ってきたが、上述した大規模尤度計算に一定の道筋が見えてきたので、尤度にもとづくメタシミュレーションを実行した。具体的には、これまで行ってきた海洋潮汐データ同化プロジェクトで、シミュレーションの境界条件として与える海底地形データセットの不確定性を見積もりにメタシミュレーション法を適用した。海底地形データセットには人工衛星の重力観測に基づき全球にわたって推定された水深データを用いることが多い。一方で、船舶の航行時の音波による直接観測に基づく水深データは、海洋を部分的にしかカバーできないが、人工衛星起源のデータよりも高い精度を有する。船舶起源のデータがない場合にも現実的な海洋潮汐シミュレーションモデルを得るために、人工衛星起源の水深データを領域分割し、海域毎に多様に変化させた固有の境界条件を持つシミュレーションモデルを多数予め用意する。最適化アルゴリズムの一つである **Evolution Algorithm** を用いて、最適な境界条件を自己組織的に構成することに成功した⁸⁾。

これまで行ってきた、大気・海洋、津波⁷⁾、海洋潮汐、宇宙空間(リングカレント)²⁾、ゲノムデータ同化^{1),4),10),11)}の各領域において、データ同化実験を引き続き推進した。また、新規分野の開拓として二つのテーマの研究を開始した。一つは、複数衛星の直接観測による、磁気圏の多点観測システムへのデータ同化の応用である^{2),3)}。本年度は、多点観測データをデータ同化により統合する手法の研究に着手し、プログラムの開発は概ね終了した。現在はそのテストとして、10機程度の人工衛星が磁気圏で観測を行っていた場合を想定し、仮想的な人工衛星で粒子観測を得た場合にプラズマ粒子分布の時間・空間変化をどの程度うまく推定できるのかについて、双子実験により検討を始めた。電場ポテンシャルパターンが変化しない簡単な場合については、電場分布の大域的な構造が良く推定できることをこれまでに確認している。

もう一つの新規分野は、データ同化の“ものづくり”への応用である。製造業では製品性能の向上やその均一化のため作業工程の効率化が図られている一方、製品の設計に関しては熟練工の経験と勘に頼っているのが現状である。そこで、それらの問題の一助とするために、今まで熟練工の経験と勘で設計されていた機器内部のパラメータ推定のためデータ同化の手法を応用した⁹⁾。具体的には、高圧ガス用の圧力調整器を例として取り上げ、直接計測できないパラメータ値をデータ同化により推定可能であることを示した。得られた結果は、圧力調整器の設計へフィードバックすることで設計現場にとって有益な知見となることが期待できる。また、本手法は他の振動機器へも広く応用することが可能である。

3. 研究実施体制

(1)「樋口」グループ

①研究分担グループ長 樋口 知之 (情報・システム研究機構 副所長/教授)

②研究項目

- ・ PF 等の逐次データ同化法の高次元化
- ・ 系統的モデル設計法の検討
- ・ メタシミュレーションの開発
- ・ 新規応用分野の開発調査

4. 研究成果の発表等

(1) 論文発表 (原著論文)

1. O. Hirose, R. Yoshida, S. Imoto, R. Yamaguchi, T. Higuchi, D. S. Charnock-Jones, C. Print, and S. Miyano, Statistical inference of transcriptional module-based gene networks from time course gene expression profiles by using state space models, *Bioinformatics*, Vol.24, No.7, 932-942, doi:10.1093/bioinformatics/btm639, 2008.
2. S. Nakano, G. Ueno, Y. Ebihara, M.-C. Fok, S. Ohtani, P.C. Brandt, D.G. Mitchell, K.Keika, and T. Higuchi, A method for estimating the ring current structure and the electric potential distribution using ENA data assimilation, *Journal of Geophysical Research*, Vol.113, A05208, doi:10.1029/2006JA011853, 2008.

3. Y. Futaana, S. Nakano, M. Wieser, and S. Barabash, ENA Occultation: New remote sensing technique to study the lunar exosphere, *Journal of Geophysical Research*, V. 113, A11204, doi:10.1029/2008JA013356, 2008.
4. R. Yoshida, M. Nagasaki, R. Yamaguchi, S. Imoto, S. Miyano, and T. Higuchi, Bayesian learning of biological pathways on genomic data assimilation, *Bioinformatics* Vol.24, No.22, 2592–2601, 2008.
5. 中野慎也, 上野玄太, 中村和幸, 樋口知之, Merging particle filter とその特性, *統計数理* Vol.56, No.2, 225–234, 2008.
6. 佐藤忠彦, 樋口知之, 動的個人モデルによる消費者来店行動の解析(討論付), *日本統計学会誌*, Vol.38, No.1, 1–38, 2008.
7. K. Nakamura, N. Hirose, B.H. Choi and T. Higuchi, Particle filtering in data assimilation and its application to boundary condition of tsunami simulation model, *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications* (in press), Springer, 2009.
8. D. Inazu, T. Sato, S. Miura, Y. Ohta, K. Nakamura, H. Fujimoto, C. F. Larsen, and T. Higuchi, Accurate ocean tide modeling in southeast Alaska and large tidal dissipation around Glacier Bay, *Journal of Oceanography* (in press), 2009
9. T. Ishigaki, and T. Higuchi, Parameter identification of a pressure regulator with a nonlinear structure using a particle filter based on the nonlinear state space model, *The Proceedings of 11th International Conference on Information Fusion*, 886–891, 2008
10. O. Hirose, R. Yoshida, R. Yamaguchi, S. Imoto, T. Higuchi, S. Miyano, Analyzing time course gene expression data with biological and technical replicates to estimate gene networks by state space models, /*Proc. 2nd Asia International Conference on Modeling & Simulation*/, 940–946.(AMS2008: Refereed conference), 2008
11. R. Yamaguchi, S. Imoto, M. Yamauchi, M. Nagasaki, R. Yoshida, T. Shimamura, Y. Hatanaka, K. Ueno, T. Higuchi, N. Gotoh, S. Miyano, Predicting differences in gene regulatory systems by state space models, *Genome Informatics*, 21:101–113, 2008
12. K. Nakamura, R. Yoshida, M. Nagasaki, S. Miyano, and T. Higuchi, Parameter Estimation of In Silico Biological Pathways with Particle Filtering Towards a Petascale Computing, *The Proceedings of The fourteenth Pacific Symposium on Biocomputing*, 227–238, 2009

(2) 特許出願

平成 20 年度 国内特許出願件数 : 0 件 (CREST 研究期間累積件数 : 0 件)