

データ同化

Data Assimilation

村 上 章 (むらかみ あきら)

京都大学 教授 農学研究科地域環境科学専攻

1. データ同化とは

「同化」を辞書でひけば、「外から得た知識などを理解して自分のものとする」とある¹⁾。データ同化とはこれに由来して、「数値シミュレーションに実測データを埋め込み、馴染ませること」とされる²⁾。概念的には事象の観測値を数値シミュレーションに取り込み、現象に対する数値予測の精度を向上させることを言い、主として地球科学の分野で用いられてきた用語である。とりわけ、数値天気予報の手段として知られる。

もう少し具体的に見ると、有限要素法などによる数値シミュレーションと観測値に乖離が見られる時、観測値に基づいて数値シミュレーションモデルに含まれるパラメータや初期/境界条件を同定することと、状態変数について一時刻先予測を補正することを同時に行う。地盤工学の問題では変位や間隙水圧が状態変数となり、天気予報などのような短期予測に不向きであるため、しばしば前者（パラメータ同定）のみが行われる。いわゆる逆解析がこれにあたる。

2. データ同化の目的と手段

2.1 データ同化の目的³⁾

- 1) 予測のための最適な初期条件/境界条件を求める
- 2) シミュレーションモデル内に含まれるパラメータの最適値を求める
- 3) 観測されていない時間・空間点における観測値を補間する（観測値に基づき、シミュレーションの予測値補正を実時間で行う）

地盤工学では、2)を地盤挙動観測値に適用し、同定値を用いた再計算を以後の施工に反映することが行われる。

2.2 データ同化の手段^{2)~6)}

(1) 三次元データ同化

数値天気予報が実用化され始めた1950年代には、関数あてはめ法（状態変数を基底関数の線形結合で表し、その展開係数を観測値から最小二乗法により求める）、逐次修正法（第一推定値を与えて、解析点と周囲の観測点との距離に依存する重み及び観測値から推定値を修正する）、最適内挿法（第一推定値と観測値との差である解析誤差につき、その分散が最小となるように重みを与える：Optimal Interpolation、以後「OI」と呼ぶ）が提案された。そのほか、OIの目的関数最小化を、共役勾

配法や準ニュートン法などにより行う三次元変分法（3D-VAR；「3D」は空間三次元を表す）がある。

以上を簡単なモデル（シミュレーションモデルが有限要素法による場合）で説明しよう。時刻 t における節点での状態変数（例えば、変位や間隙水圧）をベクトル表記し、 $\mathbf{x}_t$ と表す。一方、同時刻における観測値もベクトル $\mathbf{y}_t$ で表記する。

初期値境界値問題を例にとると、状態変数の時刻 $t-1$ から t への時間遷移は、有限要素モデルを用いて一般的に次のように表せる。これは状態式と呼ばれる。

$$\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}) + \mathbf{v}_t \quad \dots\dots\dots (1)$$

ここに、 $\mathbf{v}_t$ はシステムノイズである。観測は状態変数の一部についてなされるので、その関係を次の観測式として記述する。

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{w}_t \quad \dots\dots\dots (2)$$

ここに、 $\mathbf{H}_t$ は観測行列と呼ばれ、観測式がどこで何を観測するかを表す場合、0 または 1 から成る行列となる。 $\mathbf{w}_t$ は観測ノイズである。

a) 関数あてはめ法

関数あてはめ法で用いられる最小二乗法は、観測値と推定値との差を最小とすることにより、

$$\|\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_t\|^2 \rightarrow \min. \quad \dots\dots\dots (3)$$

次式の解 $\hat{\mathbf{x}}_t$ を得る。

$$\hat{\mathbf{x}}_t = (\mathbf{H}_t^T \mathbf{H}_t)^{-1} \mathbf{H}_t^T \mathbf{y}_t \quad \dots\dots\dots (4)$$

b) 逐次修正法

観測が空間的に偏っている場合など条件が良くない時は、第一推定値（だいたい、解はこの辺りであろう） $\mathbf{x}_{t,b}$ も併用して、次の基準による最小値を採用する。

$$\|\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_t\|^2 + \alpha \|\mathbf{x}_t - \mathbf{x}_{t,b}\|^2 \rightarrow \min. \quad \dots\dots\dots (5)$$

ここに、 α は正則化（または、適切化）パラメータであり、次式の解を得る。

$$\hat{\mathbf{x}}_t = \mathbf{x}_{t,b} + (\alpha \mathbf{I} + \mathbf{H}_t^T \mathbf{H}_t)^{-1} \mathbf{H}_t^T (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_{t,b}) \quad \dots\dots\dots (6)$$

上で $\mathbf{x}_{t,b} = 0$ であれば、ティホノフの正則化法に相当する。

c) OI

式(5)をベイズの定理により導出し

$$\begin{aligned} & \frac{1}{2} (\mathbf{x}_t - \mathbf{x}_{t,b})^T \mathbf{M}_t^{-1} (\mathbf{x}_t - \mathbf{x}_{t,b}) \\ & + \frac{1}{2} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_t)^T \mathbf{R}_t^{-1} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_t) \rightarrow \min. \quad \dots\dots\dots (7) \end{aligned}$$

を得る。ここに、 $\mathbf{M}_t$ は推定誤差共分散行列、 $\mathbf{R}_t$ は観測誤差共分散行列であり、この解として次式を得る。

$$\hat{\mathbf{x}}_t = \mathbf{x}_{t,b} + \mathbf{M}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{M}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_{t,b}) \dots \dots (8)$$

あるいは降下法により解を求める手段がOIである。また(5)式と同様、(7)式第1項に正則化パラメータ α を乗じ事前情報と観測情報間の相対的な重み付けを行うことを「拡張ベイズ法」と呼んでいる⁷⁾。

(2) 四次元データ同化

三次元データ同化手法は、添字に時刻 t が入っているものの、その時刻内で推定値の更新が行われる。それに対してある期間の観測値を合理的に取り込むために、異なる時刻にわたって推定値の更新を進めることを四次元データ同化(空間三次元+時間一次元)と呼び、四次元変分法(4D-VAR, アジョイント法)とカルマンフィルタが主たる手段となる。四次元データ同化は、あるデータ同化期間の観測を一括して処理する「非逐次型データ同化」(4D-VAR)と、観測を得るたびに状態変数を修正し、適切なものに収束させる「逐次型データ同化」(カルマンフィルタ)とに分けられる。

a) 4D-VAR, アジョイント法

アジョイント法における評価関数は、同化期間を $t=0$ から T とすると次のようになる。

$$J = \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_{0,b})^T \mathbf{M}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_{0,b}) + \sum_{t=0}^T \frac{1}{2} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_t)^T \mathbf{R}_t^{-1} (\mathbf{y}_t - \mathbf{H}_t \mathbf{x}_t) \dots \dots \dots (9)$$

上式を状態式(1)と連立させ降下法により最小化する際、勾配 $\mathbf{g} = \partial J / \partial \mathbf{x}_0$ を求める必要がある。評価関数を最小とする $\mathbf{x}_0$ には初期値のほか、観測値と最も整合のとれたパラメータや外力を含め、これを同定することもできる。

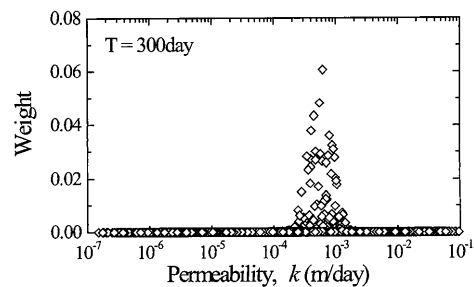
b) カルマンフィルタ

カルマンフィルタは一時刻ごとに、ベイズ法による状態変数の観測的な更新と、状態式(1)に基づく時間的な更新を組み合わせたものである。詳しくは土と基礎の講座⁸⁾を参照されたい。カルマンフィルタは元来、線形の状態式・観測式に対するものであるが、これらは非線形で記述されることが多い。その場合は、テラー展開による線形化を行い線形フィルタにあてはめる拡張カルマンフィルタが多く用いられ、広い範囲にわたる適用例が見られる。カルマンフィルタの適用は、気象学・地盤工学分野のいずれについても1980年代半ばに嚆矢が見られる。

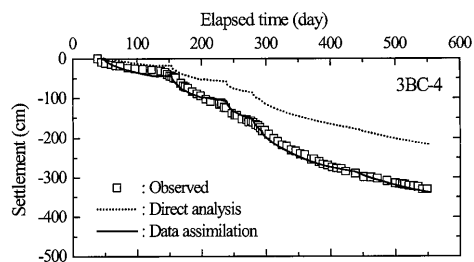
3. 非線形カルマンフィルタによるデータ同化

非線形状態式・観測式をそのまま用いる非線形カルマンフィルタ(アンサンブルカルマンフィルタ及び粒子フィルタ)が、2000年前後に相次いで提案された^{2)~6)}。これらはサンプルあるいは粒子と呼ばれる多数の状態変数の候補を用意し、その集合により状態変数の確率密度分布を近似するとともに、時間変化を記述する。コンピュータの発達により非常に多くのサンプル/粒子計算が可能となり、非線形・非ガウス性の問題にも適用できることから、データ同化は著しく発展した。

図一1に粒子フィルタを用いたデータ同化の例(神戸



(a) 同定パラメータの尤度分布



(b) データ同化後の予測挙動

図一1 データ同化の実例⁹⁾

空港人工島の地盤挙動)⁹⁾を示す。観測値を用いた刻々のデータ同化により、粒子群の尤度は変化する。図一1(a)は300日経過後の透水係数各粒子の重み(尤度)分布であり、観測が進むにつれ、当初の一樣分布からピークを有する形状へ変化している。この分布を用いた重み付き平均が刻々のパラメータ同定値を与える。図一1(b)は同定パラメータを用いた予測計算値(実線)が、当初設計時の予測(破線)に比べ、観測値に馴染んでいる様子を示す。詳細は、文献9)、10)を参照されたい。

参 考 文 献

- 1) 大辞林, 第三版, 2006.
- 2) 細川敬祐編: データ同化の考え方とその方法(講演: 中野慎也), MTIハンドブック, pp. 9~17, 2009.
- 3) 樋口知之編: データ同化入門一次世代のシミュレーション技術, 朝倉書店, 2011.
- 4) 露木 茂・川畑拓矢編: 気象学におけるデータ同化, 気象研究ノート, 第217号, 2008.
- 5) 淡路敏之・蒲池政文・池田元美・石川洋一編: データ同化, 京都大学学術出版会, 2009.
- 6) 中村和幸・上野玄太・樋口知之: データ同化: その概念と計算アルゴリズム, 統計数理, Vol. 53, No. 2, pp. 211~229, 2005.
- 7) 本城勇介: 講座「地盤工学における逆解析」5. 逆解析の手法, 土と基礎, Vol. 43, No. 5, pp. 67~72, 1995.
- 8) 佐藤忠信: 講座「地盤工学における逆解析」3. 逆解析の手法, 土と基礎, Vol. 43, No. 5, pp. 67~72, 1995.
- 9) Murakami, A., Shuku, T., Nishimura, S., Fujisawa, K., Nakamura, K.: Data assimilation using the particle filter for identifying the elasto-plastic material properties of geomaterials. Int. J. Numer. Anal. Methods Geomech., Vol. 37, pp. 1642~1669, 2013.
- 10) 珠玖隆行・村上 章・藤澤和謙・西村伸一: 地盤解析におけるデータ同化, 地盤工学会誌, Vol. 59, No. 3, pp. 8~11, 2011.

(原稿受理 2015.4.8)