4DEnVar（4次元アンサンブル変分法）

横　田　　祥*

1．はじめに

表題の4次元アンサンブル変分法（4DEnVar: four-dimensional ensemble-based variational method; Lorenc 2003; Buehner 2005; Zupanski 2005; Liu *et al*. 2008）[†1]は，観測データと数値予報モデルに基づいて大気などの状態を高精度に推定する「4次元データ同化」の手法の一つであり，次世代の実用的なデータ同化手法の候補として国内外で盛んに研究が進められている．本稿では，この手法が現在の4次元データ同化手法の主流である4次元変分法（4D-Var: four-dimensional variational method; Sasaki 1969, 1970; Thompson 1969）や4次元アンサンブルカルマンフィルター（4DEnKF: four-dimensional ensemble Kalman filter; Kalman 1960; Evensen 1994; Hunt *et al*. 2004）とどのように違うのか，また，この手法を用いることにどのようなメリットやデメリットがあるのかについて説明する．

2．4次元データ同化とは

データ同化では，観測値を用いて数値予報モデルで表現された場（第一推定値）を修正（第一推定値と観測値の双方に整合するように最適化）する．この際，3次元データ同化（3Dxxxのように表記する）では，1つの観測値を同化した影響が空間3次元方向に及ぶのに対し，4次元データ同化（4Dxxxのように表記する）では，それに加えて時間方向にも影響が及ぶ．すなわち，4次元データ同化では，観測時刻と異なる時刻の場を修正することができる[†2]．

どのように第一推定値x^bと観測値y^oから解析値x^aを求めるのか，具体的に見てみよう．仮にx^bとy^oがそれぞれ真の値の回りに分散$\sigma_b{}^2$と$\sigma_o{}^2$でガウス分布をしており，互いに独立であるとすると，真の値がxである確率密度関数は，2つのガウス分布関数の積であり，exp［$-J(x)$］に比例する．ここで

$$J(x)=\frac{1}{2\sigma_b{}^2}(x-x^b)^2+\frac{1}{2\sigma_o{}^2}[H(x)-y^o]^2 \qquad (1)$$

であり，この$J(x)$を評価関数と呼ぶ．$H(x)$はxを観測される物理量yに変数変換する演算子である．このとき，xとyの間に時間差がある場合には，$H(x)$に時間推進演算子，つまり数値予報モデルによる予報が含まれる．$H(x)$に時間推進演算子が含まれることが4次元データ同化の大きな特徴である．この$J(x)$が最小値をとる時，確率密度関数exp［$-J(x)$］が最大となり，この時のxが解析値とみなされる．したがって，解析値x^aは，

$$\frac{dJ(x)}{dx}=\frac{1}{\sigma_b{}^2}(x-x^b)+\frac{dH(x)}{dx}\frac{1}{\sigma_o{}^2}[H(x)-y^o]=0 \qquad (2)$$

をxについて解くことによって求められる．

式(1)，(2)は第一推定値と観測値がそれぞれ1つず

* 気象研究所予報研究部．syokota@mri-jma.go.jp

†1 4DEnVarのことをMLEF（maximum likelihood ensemble filter, Zupanski 2005）やEn4DVAR（ensemble-based four-dimensional variational data assimilation, Liu *et al*. 2008）などと呼ぶ場合もある．ただし，En4DVARについては，後述するHybrid 4D-Varのことを指す場合もあるので注意が必要である（Lorenc 2013）．

†2 データ同化と数値予報モデルによる予報を繰り返し行う「解析予報サイクル」のことを4次元データ同化と呼ぶ場合もある（e.g., 三好・本田 2007）が，本稿では解析時刻と異なる時刻の観測値の同化を指すものとする．

つの場合の式であるが，一般のデータ同化では，空間3次元に広がりを持つ多変数の第一推定値に複数の観測値を同化して状態を推定する．このため，第一推定値と観測値をそれぞれベクトル $\mathbf{x}^b$，$\mathbf{y}^o$で表し，式(1)，(2)をそれぞれ多変数に拡張した

$$J(\mathbf{x})=\frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)+\frac{1}{2}[H(\mathbf{x})-\mathbf{y}^o]^T\mathbf{R}^{-1}[H(\mathbf{x})-\mathbf{y}^o] \tag{3}$$

$$\nabla J(\mathbf{x})=\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)+\mathbf{H}(\mathbf{x})^T\mathbf{R}^{-1}[H(\mathbf{x})-\mathbf{y}^o]=0 \tag{4}$$

という式を用いて解析値 $\mathbf{x}^a$を求めることになる．ここで，$\mathbf{B}$ と $\mathbf{R}$ はそれぞれ第一推定値 $\mathbf{x}^b$，観測値 $\mathbf{y}^o$の成分間の誤差共分散を表す対称行列である．また，$\mathbf{H}(\mathbf{x})$は $H(\mathbf{x})$のヤコビ行列［$\mathbf{H}(\mathbf{x})=\partial H(\mathbf{x})/\partial\mathbf{x}$］であり，これを転置した $\mathbf{H}(\mathbf{x})^T$を $\mathbf{H}(\mathbf{x})$のアジョイント演算子と呼ぶ．

3．カルマンフィルターと変分法

式(4)は一般に，解析的に解くことは出来ない．しかし，$H(\mathbf{x})$を

$$H(\mathbf{x})\approx H(\mathbf{x}^b)+\mathbf{H}(\mathbf{x}^b)(\mathbf{x}-\mathbf{x}^b) \tag{5}$$

のように線形近似すると，解析値を

$$\mathbf{x}^a=\mathbf{x}^b+\mathbf{B}\mathbf{H}(\mathbf{x}^b)^T\left[\mathbf{H}(\mathbf{x}^b)\mathbf{B}\mathbf{H}(\mathbf{x}^b)^T+\mathbf{R}\right]^{-1}\left[\mathbf{y}^o-H(\mathbf{x}^b)\right] \tag{6}$$

のように求めることが出来る[†3]．この式(6)を用いて陽に解析値を求める手法を拡張カルマンフィルター（EKF: extended Kalman filter）と呼ぶ．「拡張」という単語を用いているのは，元々 $H(\mathbf{x})$が線形の場合について提案されたカルマンフィルターを，非線形の場合にも適用できるように「拡張」されたものだからである．一方，式(6)を用いず，式(4)の $\nabla J(\mathbf{x})$が0に近づくように $\mathbf{x}$ を調整することによって陰に解析値を求める手法を変分法（Var）と呼ぶ．

式(5)の近似を用いない Var は，EKF より厳密な式を解いているように思えるが，$J(\mathbf{x})$が複雑な形をしている場合には，$J(\mathbf{x})$を最小にする $\mathbf{x}$ を求めるのが困難である上，そのような $\mathbf{x}$ が最適とも限らない（Lorenc and Payne 2007）．このため，必ずしも EKF より Var の方が優れた手法というわけではない．なお，式(5)の近似を用いつつ陰に解析値を求める Var をインクリメント法と呼び，同じ $\mathbf{B}$ を用いた場合，EKF と同じ解析値が得られる．

4．アンサンブルカルマンフィルターとアンサンブル変分法

上記の EKF や Var で解析値を求める際，$\mathbf{H}(\mathbf{x})$の計算が必要になるが，4次元データ同化の場合は，$H(\mathbf{x})$が数値予報モデルによる予報を含み，一般に非線形・不連続かつ複雑であるため，$\mathbf{H}(\mathbf{x})$を計算するのは簡単ではない．ところが，アンサンブル予報を用いて $\mathbf{B}^{1/2}$と $\mathbf{H}(\mathbf{x})\mathbf{B}^{1/2}$を

$$\mathbf{B}^{1/2}\approx\frac{1}{\sqrt{m-1}}\left(\mathbf{x}_1{}^b-\overline{\mathbf{x}^b},\cdots,\mathbf{x}_m{}^b-\overline{\mathbf{x}^b}\right) \tag{7}$$

$$\mathbf{H}(\mathbf{x})\mathbf{B}^{1/2}\approx\frac{1}{\sqrt{m-1}}\left[H(\mathbf{x}_1{}^b)-\overline{H(\mathbf{x}^b)},\cdots,H(\mathbf{x}_m{}^b)-\overline{H(\mathbf{x}^b)}\right] \tag{8}$$

のように近似すれば，これらを式(4)や式(6)に代入することにより，$\mathbf{H}(\mathbf{x})$を直接計算することなく解析値を求めることが出来る．ここで，m はアンサンブルメンバー数，$\mathbf{x}_i{}^b$はメンバー i（$1\leq i\leq m$）の第一推定値であり，文字の上のバーはアンサンブル平均を表す．このように，アンサンブル予報を用いて $\mathbf{H}(\mathbf{x})$を間接的に計算し，解析値を求める EKF と Var を，それぞれアンサンブルカルマンフィルター（EnKF），アンサンブル変分法（EnVar）と呼ぶ．

$\mathbf{H}(\mathbf{x})$を直接計算する必要が無いことは，EnKF や EnVar の大きなメリットである．しかし，アンサンブルメンバー数が少ないと，$\mathbf{B}^{1/2}$と $\mathbf{H}(\mathbf{x})\mathbf{B}^{1/2}$を精度よく近似することが出来ない．このため，実用の際には，距離の離れた変数間の誤差共分散を小さくすることによってサンプリング誤差を抑える「局所化」が行われることが多い[†4]．ただし，「局所化」は物理法則

[†3] 式(6)は，ガウス分布を仮定しない不偏最小分散推定によって導出することも出来る（Kalman 1960）．

[†4] EnKF では一般に解析変数と観測変数の距離に応じて局所化を行う．EnVar でもこの局所化を行うことは出来るが（Yokota *et al*. 2016），一般的な EnVar では解析変数同士の距離に応じて局所化を行う．この局所化の違いが EnKF と EnVar の本質的な違いであるという考え方もある．

に基づいた操作ではないため，問題点も多い（e.g., Cohn *et al*. 1998）.

EKF は **B** を前解析からの線形時間発展で，Var は **B** を気候値で与えて解析値を求めるのが一般的であるのに対し，EnKF と EnVar は **B** を式(7)で（アンサンブル予報を用いて）与えるのが一般的である．アンサンブル予報を用いて **B** を与えると，気候値で与えた場合に比べて，時間とともに変化する場に応じた（流れに依存した）解析を行うことが出来るが，この **B** は，アンサンブル予報による期待値と気候値の重み付き平均で与えること（ハイブリッドデータ同化と呼ぶ）も可能であり，より精度の高い解析が可能とされている（e.g., Hamill and Snyder 2000）.

5．まとめ

以上のデータ同化手法の分類を第1表にまとめた．表題の4DEnVar は，下記の特徴を持つデータ同化手法と言える．

(i) $H(x)$に数値予報モデルによる予報を含む (4D)．
(ii) アンサンブル予報を用いて $\mathbf{H}(\mathbf{x})$を近似する (En)．
(iii) $\nabla J(\mathbf{x})=0$ を陰に解く（Var）.

この手法の4D-Var との違いは(ii)にあり，メンバー数が少ない場合はアンサンブル予報による近似が困難であるが，$\mathbf{H}(\mathbf{x})$を直接計算せずに流れに依存した解析が行える点は大きな利点である．また，4DEnKF との違いは(iii)にあり，$J(\mathbf{x})$の最小値を見つける困難さはあるものの，$H(\mathbf{x})$の線形近似を行わずに解析値が求められる点は利点と言える．

多くの現業数値予報センターや研究機関で既に実用化されている4D-Var や4DEnKF に比べて，4DEnVar を実用化した例はまだ少ない．今後4DEnVar が世界の主流になるかもしれないし，ならないかもしれない（我々の研究に懸かっている！？）．いずれにしても，全ての同化手法にはメリットとデメリットがあり，「どの手法が良いか」は目的によって大きく異なる．データ同化を行う際や，データ同化によって作成された解析値を使う際には，各手法の違いを正しく認識しておくことが極めて重要である．

謝　辞

本稿の作成にあたり，気象研究所予報研究部・台風研究部の皆様に多くの有益なコメントを頂きました．深く感謝いたします．

参　考　文　献

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. Quart. J. Roy. Meteor. Soc., **131**, 1013-1043.

Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz and D. Lamich, 1998: Assessing the effects of data selection with the DAO physical-space statistical analysis system. Mon. Wea. Rev., **126**, 2913-2926.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., **99**, 10 143-10 162.

Hamill, T. M. and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. Mon. Wea. Rev., **128**, 2905-2919.

Hunt, B. R. *et al*., 2004: Four-dimensional ensemble Kalman filtering. Tellus, **56A**, 273-277.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. Trans. ASME J. Basic Eng., **82**, 35-45.

Liu, C., Q. Xiao and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme. Part I: Technical formulation and preliminary test. Mon. Wea. Rev., **136**, 3363-3373.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP: A comparison with 4D-Var. Quart. J. Roy. Meteor. Soc., **129**, 3183-3203.

Lorenc, A. C., 2013: Recommended nomenclature for EnVar data assimilation methods. CAS/JSC Working

第1表　4次元データ同化手法の分類．

<table>
<tr><th>同化手法</th><th>Bの与え方</th><th>$\mathbf{H}(\mathbf{x})$の計算方法</th><th>$\nabla J(\mathbf{x})=0$ の解法</th></tr>
<tr><td>4DEKF</td><td>前解析からの線形時間発展</td><td>直接計算</td><td rowspan="3">$H(\mathbf{x})$を線形近似して陽に解く</td></tr>
<tr><td>4DEnKF</td><td>アンサンブルで計算</td><td rowspan="2">アンサンブルで計算</td></tr>
<tr><td>Hybrid 4DEnKF</td><td>アンサンブル＋気候値</td></tr>
<tr><td>4D-Var</td><td>気候値</td><td rowspan="2">直接計算</td><td rowspan="4">陰に解く</td></tr>
<tr><td>Hybrid 4D-Var</td><td>アンサンブル＋気候値</td></tr>
<tr><td>4DEnVar</td><td>アンサンブルで計算</td><td rowspan="2">アンサンブルで計算</td></tr>
<tr><td>Hybrid 4DEnVar</td><td>アンサンブル＋気候値</td></tr>
</table>

Group on Numerical Experimentation: Research Activities in Atmospheric and Oceanic Modelling, (43), 1.7-1.8.

Lorenc, A. C. and T. Payne, 2007: 4D-Var and the butterfly effect: Statistical four-dimensional data assimilation for a wide range of scales. Quart. J. Roy. Meteor. Soc., 133, 607-614.

三好建正，本田有機，2007: 気象学におけるデータ同化．天気，54，287-290.

Sasaki, Y., 1969: Proposed inclusion of time variation terms, observational and theoretical, in numerical variational objective analysis. J. Meteor. Soc. Japan, 47, 115-124.

Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. Mon. Wea. Rev., 98, 875-883.

Thompson, P. D., 1969: Reduction of analysis error through constraints of dynamical consistency. J. Appl. Meteor., 8, 738-742.

Yokota, S., M. Kunii, K. Aonashi and S. Origuchi, 2016: Comparison between four-dimensional LETKF and ensemble-based variational data assimilation with observation localization. SOLA, 12, 80-85.

Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. Mon. Wea. Rev., 133, 1710-1726.