

〔解 説〕

4011 (数値予報；データ同化；アンサンブル予報；

カルマンフィルタ；アンサンブル・カルマンフィルタ)

アンサンブル・カルマンフィルタ

ーデータ同化とアンサンブル予報の接点ー

三 好 建 正*

1. はじめに

データ同化は、観測データと予報モデルを使い、大気の最も確からしい状態を推定するものである。日々の数値天気予報では、よりよい初期値を与えることが予報精度の向上に重要な役割を果たす。また、気象や気候の研究では、より高精度な大気や海洋の状態を知ることが重要な研究基盤を与える。実際、再解析データは過去の大気の最も確からしい状態を推定したものであるが、これを用いて多くの気象や気候の研究がなされ、重要な発見も多くなされてきている。このように、よりよいデータ同化手法を開発することは、日々の天気予報のみならず、気象や気候の基礎研究にも貢献する、重要な課題と言える。

アンサンブル・カルマンフィルタ (Ensemble Kalman Filter: EnKF) は、最近話題となっている最新のデータ同化手法の1つである。カルマンフィルタは1960年に Kalman が提唱したアルゴリズムで、時間発展が線形モデルで与えられ、誤差の確率分布関数がガウス分布で与えられる場合に、推定誤差を最小とする最適な解を与える (Kalman, 1960)。このアルゴリズムはコンピュータプログラムに適しているため、衛星追跡など幅広い分野で用いられてきた。しかし、自由度が非常に大きい現実的な大気モデルでは、このアルゴリズムを直接実装することはできない。これを解決する先駆的な仕事として、1994年に Evensen がアンサンブル予報を用いてカルマンフィルタを近似的に実現する EnKF を提唱した (Evensen, 1994)。端的に言えば、

アンサンブル予報が提供する予報のばらつきに関する情報を使って、近似的にカルマンフィルタを行うというものである。

本稿は、アンサンブル予報とデータ同化が融合した最新のデータ同化手法、EnKF について、理論的な導入及びその物理的解釈について明快な解説を行おうとするものである (露木 (2004) にも解説がある)。最近の論文に取り組む際の準備となりうることも目標とするが、EnKF に関する研究を網羅的に紹介するレビューではない。レビューは、Hamill (2003) や Evensen (2003) などがある。本稿では、はじめに推定問題の最適解としてのカルマンフィルタを導入し、これから自然に EnKF が導かれることを示す。その後、現実的な大気モデルを考慮した場合の効果的な EnKF の実現法について述べ、最後に EnKF の問題点や利点、その将来性に言及する。

2. カルマンフィルタ

本節では、カルマンフィルタを構成する5つの方程式をなるべく簡潔に導入する。それぞれの式の導出など細かい点は付録 A に記述したので参照頂きたい。また、数学的により厳密な解説等は Jazwinski (1970) を参照頂きたい。簡単な導出や、より詳細な解説は Gelb *et al.* (1974) に記述がある。この他、数値予報モデルのデータ同化におけるカルマンフィルタについては、露木 (1997, 147ページ, 第5.5.8節) にも解説がある。

カルマンフィルタは、線形モデル、ガウス分布の誤差統計、という2つの仮定の下、統計的な推定問題の最適解を与える。そのアルゴリズムは、大きく2つの部分からなる。それは、時間発展及び解析である。時間発展のプロセスは、線形なモデル $\mathbf{M}$ を用いて1時刻前の状態 $\mathbf{x}_{t-1}^a$ を現在の状態 $\mathbf{x}_t^f$ に写す。

気象庁予報部数値予報課 (現在, University of Maryland, College Park 滞在中).
miyoshi@atmos.umd.edu

—2004年6月21日受領—

—2004年11月22日受理—

$$\mathbf{x}_i^f = \mathbf{M}\mathbf{x}_{i-1}^a \quad (1)$$

ここで、下の添え字は時刻を表し、上の添え字は解析 (a)、予報 (f) を表す (以下同じ)。状態と呼んでいるのは、具体的にはモデル変数と同義である。予報及び解析には誤差が含まれると仮定し、その誤差は 2 次モーメント、すなわち、誤差共分散行列 $\mathbf{P}$ として表現される。これも時間発展する。

$$\mathbf{P}_i^f = \mathbf{M}\mathbf{P}_{i-1}^a\mathbf{M}^T + \mathbf{Q} \quad (2)$$

ここで、 $\mathbf{Q}$ はモデルの誤差の 2 次モーメントであり、線形モデルの転置 $\mathbf{M}^T$ はアジョイントモデルと呼ばれるものである。

解析のプロセスについては、予報を観測データ $\mathbf{y}^o$ で修正するプロセス

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_i(\mathbf{y}_i^o - \mathbf{H}\mathbf{x}_i^f) \quad (3)$$

及び、観測データを用いることで状態の推定誤差が小さくなるプロセス

$$\mathbf{P}_i^a = [\mathbf{I} - \mathbf{K}_i\mathbf{H}]\mathbf{P}_i^f \quad (4)$$

からなる。 $\mathbf{I}$ は単位行列である。観測がモデルの全格子点上の全変数について得られることは通常あり得ないため、観測データはモデル変数と必ずしも一致せず、ここでは観測データがモデル変数の線形写像 $\mathbf{H}$ で表されると仮定している。 $\mathbf{H}$ は観測演算子と呼ばれる。 $\mathbf{K}$ は観測の影響を決定する線形変換で、

$$\mathbf{K}_i = \mathbf{P}_i^f\mathbf{H}^T[\mathbf{H}\mathbf{P}_i^f\mathbf{H}^T + \mathbf{R}]^{-1} \quad (5)$$

で与えられる。 $\mathbf{R}$ は観測誤差の 2 次モーメント (観測誤差共分散行列) である。これが、解析の推定誤差を最小とするカルマンフィルタの最適解を与え、カルマンゲイン行列と呼ばれる (カルマンフィルタの最適性の証明については、例えば Jazwinski (1970) や Gelb *et al.* (1974) を参照)。以上に示した第 1 式から第 5 式がカルマンフィルタのアルゴリズムである (付録 A でこれらの式の簡単な導出を行っている)。

大気モデルは非線形であるため、このアルゴリズムは近似的にしか表現されない。その近似を行う単純な方法が接線形化近似である。これを行うには、非線形

モデル M を $\mathbf{x}_0$ のまわりで Taylor 展開する。

$$M(\mathbf{x}_0 + \delta\mathbf{x}) = M(\mathbf{x}_0) + \mathbf{M}_{\mathbf{x}_0}\delta\mathbf{x} + O(\delta\mathbf{x}^2) \quad (6)$$

この 2 次以上の項を無視することで線形化する。すなわち、第 2 式の $\mathbf{M}$ の代わりに非線形モデル M のヤコビアン (接線形モデルという)

$$\mathbf{M}_{\mathbf{x}_0} = \left. \frac{\partial M}{\partial \mathbf{x}} \right|_{\mathbf{x}_0} \quad (7)$$

を用いることで、カルマンフィルタのプロセスを非線形モデルにも適用できる。この近似を行ったものを、拡張カルマンフィルタ (Extended Kalman Filter: EKF) という。

3. アンサンブル・カルマンフィルタ

本節では、前節に基づき、EnKF を理論的に導入する。物理的解釈については、次節で扱う。

誤差共分散行列の平方根を考える。

$$\mathbf{P}_{i-1}^a = \mathbf{E}_{i-1}^a\mathbf{E}_{i-1}^{aT} \quad (8)$$

共分散行列 $\mathbf{P}$ は非負の実対称行列であるから、実行列の平方根 $\mathbf{E}$ を持つ。なお、この平方根 $\mathbf{E}$ は、 $\mathbf{U}\mathbf{U}^T = \mathbf{I}$ を満たす行列 $\mathbf{U}$ だけの自由度があって、一意に定まるものではない ($\mathbf{E}\mathbf{U}$ も第 8 式を満たし、平方根となる)。誤差共分散行列 $\mathbf{P}$ は誤差の 2 次モーメントで、状態の誤差ベクトル $\delta\mathbf{x}$ の共分散を意味する。式で書くと、

$$\begin{aligned} \mathbf{P} &= \langle \delta\mathbf{x}\delta\mathbf{x}^T \rangle \\ &= \frac{1}{N-1} \sum_{i=1}^N \delta\mathbf{x}^{(i)}\delta\mathbf{x}^{(i)T} \end{aligned} \quad (9)$$

のようになる。ここで、記号 $\langle \bullet \rangle$ は統計期待値を表し、右辺 2 行目は期待値を統計サンプルの平均を使って表したもので、 N はサンプル数、添字 i は統計サンプルを表す。これを用いて、平方根 $\mathbf{E}$ は、具体的には

$$\mathbf{E} = \left[\frac{\delta\mathbf{x}^{(1)}}{\sqrt{N-1}} \quad \cdots \quad \frac{\delta\mathbf{x}^{(N)}}{\sqrt{N-1}} \right] \quad (10)$$

のような行列となる。第 8 式が主張するのは、完全に $\mathbf{P}$ を表現するのに N はモデル変数の次元の数あれば

十分であるということ、 $\delta \mathbf{x}^{(i)}$ は一意には決まらないということである。

第8式を第2式に代入すると、

$$\begin{aligned} \mathbf{P}_i^f &= \mathbf{M}_{\mathbf{x}_{i-1}^a} \mathbf{E}_{i-1}^a \mathbf{M}_{\mathbf{x}_{i-1}^a}^T \\ &= \mathbf{M}_{\mathbf{x}_{i-1}^a} \mathbf{E}_{i-1}^a (\mathbf{M}_{\mathbf{x}_{i-1}^a} \mathbf{E}_{i-1}^a)^T \end{aligned} \quad (11)$$

となる。ここで、実際の大気モデルがそうであるように、非線形モデルを仮定し、そのヤコビアンを使った。また、モデル誤差の共分散行列 $\mathbf{Q}$ は無視した。左辺は予報誤差共分散行列であるが、この平方根を考えると、

$$\mathbf{E}_i^f = \mathbf{M}_{\mathbf{x}_{i-1}^a} \mathbf{E}_{i-1}^a \quad (12)$$

となる。これは、第10式のように、誤差共分散行列 $\mathbf{P}$ の平方根 $\mathbf{E}$ の各列を状態の誤差ベクトルと見なし、それを時間発展させるプロセスを表している。すなわち、第12式は、行列 $\mathbf{E}$ の各列をアンサンブル摂動ベクトルと見なしたアンサンブル予報と考えることができる。ここで、 $\mathbf{E}$ は正方行列であるから、アンサンブルメンバー数（列の数）は、モデル変数の次元の数（行の数）と等しい。ここに至って、第12式の時間発展に、接線形モデルではなく、元々の非線形モデルを用いることができる。非線形モデルを使って誤差ベクトルを時間発展させる場合は、

$$\begin{aligned} \mathbf{M}_{\mathbf{x}_{i-1}^a} \mathbf{E}_{i-1}^a &= \mathbf{M}_{\mathbf{x}_{i-1}^a} (\delta \mathbf{x}_{i-1}^{(1)} \cdots \delta \mathbf{x}_{i-1}^{(N)}) \\ &\cong \left(\frac{M(\mathbf{x}_{i-1}^a + \varepsilon \delta \mathbf{x}_{i-1}^{(1)}) - M(\mathbf{x}_{i-1}^a)}{\varepsilon} \cdots \frac{M(\mathbf{x}_{i-1}^a + \varepsilon \delta \mathbf{x}_{i-1}^{(N)}) - M(\mathbf{x}_{i-1}^a)}{\varepsilon} \right) \end{aligned} \quad (13)$$

のように、接線形モデルの背景場を考慮して計算しなければならないことに注意が必要である。ここで、 ε は接線形化近似が成り立つための、小さな数である。第13式は非線形モデルを用いたアンサンブル摂動ベクトルの時間発展であり、まさにアンサンブル予報を表している。ここで非線形モデルを使うことで、接線形化による近似のデメリットを解消できる場合がある。実際、Evensen (1994) は、非線形の準地衡風モデルを用いて、EnKF が EKF に勝つという結果について言及している。

以上の議論で、アンサンブル予報がいかにカルマンフィルタのアルゴリズムの中に現れてくるかを示した。しかし、第13式ではモデル変数の次元の数だけの

アンサンブルメンバーが必要とされるため、モデル変数の次元が100万を優に超えるような現実的な大気モデルを想定した場合、この計算は実際には不可能である。しかしながら、通常、誤差共分散行列は縮退しており、そのランクはモデル変数の次元よりもずっと小さい。つまり、共分散行列の固有値分解

$$\mathbf{P}^a = \mathbf{S} \mathbf{D} \mathbf{S}^T \quad (14)$$

を行うと、固有値の多くが0に十分近く、有効な固有値の数はモデル変数の次元の数よりもずっと小さい。ここで、 $\mathbf{D}$ は対角行列で各成分は固有値からなり、 $\mathbf{S}$ は各固有値に対応する固有ベクトルからなる直交行列である。この事実を利用すると、数少ないアンサンブルメンバーで誤差共分散行列 $\mathbf{P}$ を十分よく表現できることになる。実際の大気モデルでは、アンサンブルメンバーの数は高々100のオーダーが現実的な計算上の限界であるが、このような少ないメンバーで効率的に $\mathbf{P}$ を表現することについては、第5節で述べる。

また、第5式のカルマンゲイン行列を計算するプロセスでは、観測の数を一辺とした正方行列の逆行列の計算が必要である。EnKFの実装方法によっては、一度に扱う観測の数が膨大となる場合もあり、その場合は、第5式を

$$\begin{aligned} \mathbf{K} &= \mathbf{P} \mathbf{H}^T [\mathbf{H} \mathbf{P} \mathbf{H}^T + \mathbf{R}]^{-1} \\ &= \mathbf{E} (\mathbf{H} \mathbf{E})^T [\mathbf{H} \mathbf{E} (\mathbf{H} \mathbf{E})^T + \mathbf{R}]^{-1} \\ &= \mathbf{E} [\mathbf{I} + (\mathbf{H} \mathbf{E})^T \mathbf{R}^{-1} \mathbf{H} \mathbf{E}]^{-1} [\mathbf{I} + (\mathbf{H} \mathbf{E})^T \mathbf{R}^{-1} \mathbf{H} \mathbf{E}] \\ &\quad (\mathbf{H} \mathbf{E})^T [\mathbf{H} \mathbf{E} (\mathbf{H} \mathbf{E})^T + \mathbf{R}]^{-1} \\ &= \mathbf{E} [\mathbf{I} + (\mathbf{H} \mathbf{E})^T \mathbf{R}^{-1} \mathbf{H} \mathbf{E}]^{-1} (\mathbf{H} \mathbf{E})^T [\mathbf{I} + \mathbf{R}^{-1} \mathbf{H} \mathbf{E} \\ &\quad (\mathbf{H} \mathbf{E})^T] [\mathbf{H} \mathbf{E} (\mathbf{H} \mathbf{E})^T + \mathbf{R}]^{-1} \\ &= \mathbf{E} [\mathbf{I} + (\mathbf{H} \mathbf{E})^T \mathbf{R}^{-1} \mathbf{H} \mathbf{E}]^{-1} (\mathbf{H} \mathbf{E})^T \mathbf{R}^{-1} \end{aligned} \quad (15)$$

のように変形すると、アンサンブルメンバーの数を1辺とした正方行列の逆行列の計算ですむようになる。なお、通常異なる観測の間に相関はなく、観測の総数から見て非常に限られた数の観測誤差相関を考慮すればよい。このため、観測誤差共分散行列 $\mathbf{R}$ はほぼ対角行列に近いブロック対角行列として扱われることが通常であり、逆行列の計算は難しくない。

誤差共分散行列の平方根を考える更なる利点として、観測演算子 $\mathbf{H}$ が非線形であってもそのまま対応できることがある。通常、非線形の観測演算子は、第5

式でその転置行列（アジョイント演算子）が必要とされるため、線形化する必要がある。しかし、第15式中には、観測演算子は $\mathbf{H}\mathbf{E}$ の形でしか現れない。すなわち、アンサンブルメンバーを観測空間に写す必要がないため、観測演算子は便宜的に元々の非線形写像のまま計算できることになる。この取り扱いは、第13式で接線形モデルの計算を便宜的に非線形モデルで置き換えたのと同様である。第13式と同様、アンサンブル摂動ベクトルをそのまま非線形写像で写すのではなく、基本場のまわりで写す必要がある点に注意する必要がある。すなわち、非線形の H に対し、基本場 $\mathbf{x}_i^f$ のまわりで線形化したヤコビアン $\mathbf{H}_{x_i^f}$ を用いるべきところを、

$$\begin{aligned} \mathbf{H}_{x_i^f} \mathbf{E}_i^f &= \mathbf{H}_{x_i^f} (\delta \mathbf{x}_i^{(1)} \cdots \delta \mathbf{x}_i^{(N)}) \\ &\cong \left(\frac{H(\mathbf{x}_i^f + \varepsilon \delta \mathbf{x}_i^{(1)}) - H(\mathbf{x}_i^f)}{\varepsilon} \cdots \frac{H(\mathbf{x}_i^f + \varepsilon \delta \mathbf{x}_i^{(N)}) - H(\mathbf{x}_i^f)}{\varepsilon} \right) \end{aligned} \quad (16)$$

とすることで、非線形の H をそのまま用いることができる。

実際にデータ同化を行う際は、 $\mathbf{E}$ と $\mathbf{H}\mathbf{E}$ という形で計算機のメモリ上に情報を蓄え、第15式2行目または5行目により $\mathbf{K}$ を計算し、この $\mathbf{K}$ を使って第3式のデータ同化の式を計算する。これにより、モデル変数の次元の2乗という膨大な大きさを持つ $\mathbf{P}$ を直接メモリ上に蓄えることなく、データ同化の計算ができるのである。

最後に問題となるのは、第4式の解析誤差共分散行列を計算するプロセスである。このプロセスは、次のアンサンブルメンバーを作り出すプロセス（アンサンブル・アップデート）に相当する。このプロセスの実装方法はテクニカルであって、様々な方法で実現される。これらの方法は大きく2つに分けられ、1つは観測摂動法（Perturbed Observation method：PO法）と呼ばれるもので、もう1つは第4式を直接的に解く平方根フィルタ（Square Root Filter：SRF）と呼ばれるものである。PO法では、それぞれのアンサンブルメンバーに対し、独立した解析予報サイクルを適用する。同じ観測データを用いて、独立した解析予報サイクルを行うと、

$$\mathbf{E}^a = [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{E}^f \quad (17)$$

となる（導出は付録Bを参照）。これは

$$\mathbf{P}^a = [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{P}^f [\mathbf{I} - \mathbf{K}\mathbf{H}]^T \quad (18)$$

と同値であり、第4式より $[\mathbf{I} - \mathbf{K}\mathbf{H}]$ が一つ多く、解析誤差が必要以上に小さく見積もられてしまう。これは、観測誤差が考慮されていないためで、実際は、

$$\begin{aligned} \mathbf{P}^a &= [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{P}^f [\mathbf{I} - \mathbf{K}\mathbf{H}]^T + \mathbf{K} \mathbf{R} \mathbf{K}^T \\ &= [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{P}^f \end{aligned} \quad (19)$$

のように、第4式と一致すべきものである（第A-10、A-16式を参照）。これを満たすために、観測誤差共分散 $\mathbf{R}$ を満たすように、観測に摂動を与える必要がある。この方法は、それぞれのアンサンブルメンバーが独立した解析予報サイクルで計算されるという点で簡便であり、Evensen (1994) や Houtekamer and Mitchell (1998; 2001) など、初期の EnKF の研究ではよく用いられてきた。観測に摂動を与える必要性については、Burgers *et al.* (1998) が理論的な導出を行っている。しかしながら、Whitaker and Hamill (2002) が指摘するように、観測データに摂動を与えるという新たなサンプリングエラーを導入しており、観測に摂動を与えない SRF の方法に比べて性能が劣る可能性が示されており、最近の動向は以下に述べる SRF が主なものとなっている。

SRF は第4式を直接解く方法であり、アンサンブル・アップデートを線形写像として与える。例えば、行列 $\mathbf{T}$ を用いた線形変換

$$\mathbf{E}^a = \mathbf{E}^f \mathbf{T} \quad (20)$$

でアンサンブル・アップデートを与えると、これが第4式を満たすには、第20式を第4式に代入して、

$$\mathbf{E}^f \mathbf{T} (\mathbf{E}^f \mathbf{T})^T = (\mathbf{I} - \mathbf{K}\mathbf{H}) \mathbf{E}^f \mathbf{E}^{fT} \quad (21)$$

が満たされる必要がある。これを $\mathbf{T}$ について解くと、アンサンブル・アップデートを与える公式が得られる。この解は第8式の解が一意でなかったのと同様、 $\mathbf{U}\mathbf{U}^T = \mathbf{I}$ を満たす行列 $\mathbf{U}$ だけの自由度があって、一意ではない ($\mathbf{T}\mathbf{U}$ も第21式を満たす)。これを実際に実現する方法はいくつか知られている。SRF は、Andrews (1968) により初めて提案され、この手法を Whitaker

and Hamill (2002) が気象の問題に適用した。他にも、Anderson (2001) が提案したアンサンブル調節カルマンフィルタ (Ensemble Adjustment Kalman Filter: EAKF) や、Bishop *et al.* (2001) が提案したアンサンブル変換カルマンフィルタ (Ensemble Transform Kalman Filter: ETKF)、また、Ott *et al.* (2002; 2004) による局所 EnKF (Local Ensemble Kalman Filter: LEKF) で用いられるアンサンブル・アップデートは、SRF の一種である。これら様々な手法の紹介及び解説は煩雑であって、本稿の目的からはずれるので、ここでは省略する。様々なアンサンブル SRF (EnSRF) については Tippett *et al.* (2003) や Miyoshi (2004) にまとめられているので、これらを参照して頂きたい。

最後に、アンサンブル・アップデートの注意点に言及する。通常、カルマンフィルタでは、たとえモデルが完全であっても、主にモデルの非線形性が考慮されていないなどの理由で、誤差共分散を小さく見積もってしまう。言い換えれば、モデルの非線形性がランダム誤差を導入し、カルマンフィルタの見積もる誤差共分散よりも大きな誤差共分散を与える。カルマンフィルタが誤差共分散を小さく見積もると、観測の持つ情報が必要以上に軽く見積もられることになり、モデルの結果が現実から離れていってしまう。これを防ぐため、誤差共分散を人為的に大きくしてやる必要がある。具体的には、第 4 式で解析誤差共分散行列を計算した後、

$$\hat{\mathbf{P}}^a = \mathbf{P}^a(1 + \delta) \quad (22)$$

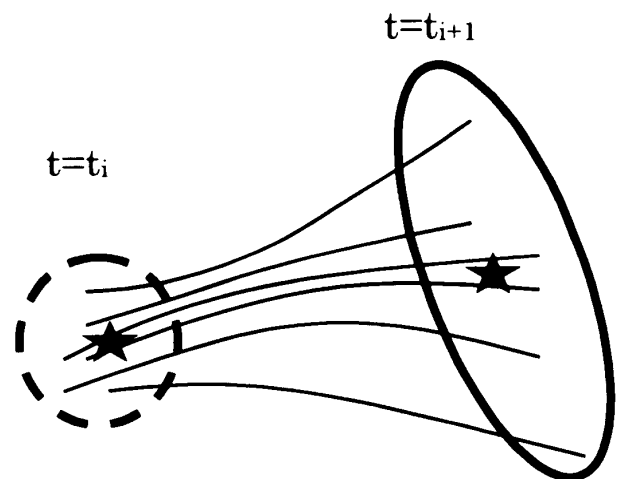
を行う。ここで、 δ は小さな正数である。このプロセスは共分散膨張 (covariance inflation) と呼ばれ、非線形モデルでカルマンフィルタを行う際の常套手段である。EnKF では、アンサンブルの摂動を全体的に少し大きくすることに相当する。別の方法として、第 2 式にあるモデル誤差の共分散行列 $\mathbf{Q}$ を考慮する場合もある。この際は、 $\mathbf{Q}$ としてランダムノイズを与えることが多い。これは、モデルの非線形性に起因する誤差がランダムだと仮定することに相当する。しかしながら、カルマンフィルタが誤差を小さく見積もる原因すべてがモデル誤差だけで説明されるわけではなく、 $\mathbf{Q}$ の導入がより明確な理論的背景を持っているとは言えない。ポストプロセスとして第 22 式の代わりに小さなランダムノイズを与えるという手法もある。これは第

2 式における $\mathbf{Q}$ の導入に近いが、第 4 式の後にランダムノイズを加えているため、これとは異なる。Breeding 法 (BGM 法とも言われる: Toth and Kalnay, 1993; 1997) を用いたアンサンブル予報において、Breeding サイクルごとにランダムノイズを加えることが有益であることが知られており (Wang, 2004, 私信; Miyoshi and Kalnay, 2004, 私信)、これに相当するものと言える。これら 3 つの方法のうち、第 22 式が用いられることが多いようである。

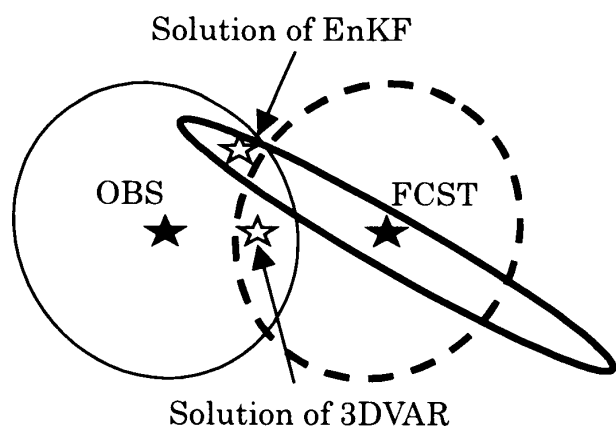
4. 物理的解釈

アンサンブル予報のそもそもの目的を考えると、決定論的ではなく、確率論的な予報を与えることである。つまり、初期値には誤差が含まれることを想定し、その誤差がどのように時間発展するかを予報する。一方、カルマンフィルタが必要とするのは、誤差の時間発展と解析のプロセスであり、アンサンブル予報がこれを補う。具体的には、以下の通りである。

アンサンブル予報により、流れに依存した誤差の情報が得られ、これが誤差の時間発展に相当する情報を与える。これを模式的に示したのが第 1 図である。時刻 t_i から t_{i+1} に時間発展する際、初期時刻 t_i の解析誤



第 1 図 アンサンブル予報が日々の流れに応じた予報誤差の分布を与えることを示すモデル位相空間上の模式図。星は初期時刻 t_i 及び終端時刻 t_{i+1} における最も確からしい状態を示し、円はそれらの誤差の広がり (確率分布関数の等値線) を示す。初期時刻での誤差の広がり (太点線) に対し、終端時刻では誤差の広がり大きく扁平になっている (太実線)。細実線は各アンサンブルメンバーの時間発展の軌道を示す。



第2図 3次元変分法(3DVAR)とEnKFの違いを示すモデル位相空間上の模式図。黒い星は、左は観測データ(OBS)、右は予報値(FCST)を示しており、円はそれぞれの誤差の広がり(確率分布関数の等値線)を示す。観測データの誤差は細実線で示されている。予報値の誤差については、3DVARでは統計的平均(太点線)が用いられるが、EnKFでは日々変化する(太実線)。

差(太点線)を考慮したアンサンブルメンバーが、予報時刻 t_{i+1} において予報誤差の分布(太実線)を与える。データ同化において、日々変化する予報誤差の情報を考慮する意味を模式的に示したのが第2図である。左の黒い星は観測データを表し、右のそれは予報値を表している。それぞれの誤差の広がり(確率分布関数の等値線)は、円で示されている。観測誤差の広がりには観測に固有のもので事前に与えられており、細実線で示されている。3次元変分法(3DVAR)など誤差の時間発展を考慮しない方法では、誤差の広がりには統計的な平均で与えられる(太点線)。一方、EnKFでは、アンサンブル予報の結果から誤差はその日特有の分布をすることが分かっており、これを太実線で示している。データ同化の解は、それぞれ白い星で示したとおりである。双方とも観測データからは同じくらい離れているが、異なる場所に位置している。EnKFの解は、その日に特有の誤差の広がりを考慮している点で、3DVARより優れたものと言える。

5. 効果的なアンサンブル・カルマンフィルタ

実際にEnKFを用いる際重要となるのは、数少ないアンサンブルメンバーで十分精度の高い誤差共分散行列を再現することである。より少ないメンバーであればあるほど、モデルの時間積分の計算回数を減らすことができ、より効率的に計算できる。誤差共分散行列

のランクがモデル変数の次元よりもずっと小さいと言っても、実際の大気モデルのモデル変数は100万次元を優に超えることが多く、ランクが100のオーダーまで落ちていることは期待しにくい。数少ないアンサンブルメンバーでは、誤差共分散行列を見積もる際のサンプリングエラーが大きく、特に離れた点同士の相関が不自然に大きく見積もられることが知られている(Houtekamer and Mitchell, 1998; Hamill and Snyder, 2000)。このことが問題となり、EnKFはなかなか現実的なモデルでは期待されるような働きをしないことがある。

そこで、いかにサンプリングエラーの影響を抑え、数少ないアンサンブルメンバーの持つ情報を最大限に活用するか、という工夫が必要となる。その解決法は大きく2通りあり、一つはハイブリッド法、もう一つは局所化である。ハイブリッド法とは、日々変化する誤差の情報を3DVARに部分的に取り込むことで、3DVARを改善させるものである(Hamill and Snyder, 2000; Corazza *et al.*, 2002, 2003)。3DVARは、評価関数

$$J(\mathbf{x}) = \frac{1}{2} [(\mathbf{x} - \mathbf{x}^f)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^f) + (\mathbf{y}^o - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H}\mathbf{x})] \quad (23)$$

が最小となるような $\mathbf{x}$ を探すアルゴリズムで、誤差がガウス分布だと仮定した場合の最尤推定解を与える。ここで、 $\mathbf{B}$ は背景誤差共分散行列で、様々なケースで当てはまるような統計的平均から決まる一定値である。 $\mathbf{B}$ が、カルマンフィルタで見積もられた日々変化する $\mathbf{P}^f$ で与えられるとき、3DVARの解はカルマンフィルタの解と一致する。しかし、数少ないアンサンブルメンバーから見積もられた $\mathbf{P}^f$ にはサンプリングエラーが大きいことはもちろん、たとえ完全な $\mathbf{P}^f$ が得られたとしても、自由度が大きい大気モデルでは、3DVARでそのまま用いることはできない。そこで3DVARで用いられる背景誤差共分散行列を、

$$\mathbf{B}' = (1 - \alpha) \mathbf{P}^f + \alpha \mathbf{B} \quad (24)$$

で置き換える。ここで、 α は0と1の間の実数で、0に近いほど日々変化する誤差の情報を多く取り込むことになるような、重み係数である。ここでの要点は、

解析場の基本的な構造を、3DVAR で使われる流れに依存しない背景誤差共分散行列で与えることで、アンサンブルメンバーがもたらすサンプリングエラーの影響を最小限に抑えようというものである。実際、ETKF を Lorenz の40変数モデル (Lorenz, 1996; Lorenz and Emanuel, 1998) に適用した簡単な実験では、ETKF そのものでは40メンバーが必要とされる状況でも、ハイブリッド法を使えば5メンバー程度でも3DVAR の結果を上回り、むしろ40メンバーを使ったETKF に近いことが確認されている (Yang, 2004, 私信)。もちろん40メンバーを用いた完全な EnKF と比べると成績は劣るものの、非常に少ないメンバー数で、3DVAR を大きく改善し、流れに依存したデータ同化を部分的にはあるが実現することの利点は大きい。

局所化による方法は、サンプリングエラーの大きい離れた点同士の相関を始めから無視することで、その影響を最小化しようとするものである。局所化を行わずにサンプリングエラーを減らすには、より多くのアンサンブルメンバーが必要とされるが、メンバー数を増やしても、離れた点同士の相関が小さくなるだけで、大きな構造にはあまり影響を与えないことが多い。このため、離れた点同士の相関は0で置き換えるなど、始めから物理的に有意な方法で局所化する。こうすることで、始めから比較的次元の元々ランクが小さい行列を対象とすることになる。局所化すればするほど必要とされるメンバーは少なくなるが、必要以上の局所化は、大きな構造に影響を与え、解析値の力学的バランスを崩してしまうため、取り組む問題に適切な局所化を行うことが重要となる。

Patil *et al.* (2001) は、米国環境予測センター (National Centers for Environmental Prediction: NCEP) の Breeding 法の摂動ベクトルの張る空間が、局所的には小さな次元であることを確かめている。Ott *et al.* (2002; 2004) はこの事実を利用して局所化を行った LEKF を考案した。NCEP の全球モデルを用いた理想実験で、LEKF がよく働くことが確認されている (Szunyogh *et al.*, 2004)。

6. 問題点とその対応、実用上の利点、将来性

これまで、EnKF が、流れに依存した誤差の情報を活用することで、優れたデータ同化スキームを与えることを述べた。しかしながら、以上の議論は、モデルが完全で、誤差がガウス分布に従うことを仮定していることを忘れてはならない。実際、モデルに小さなバ

イアスがある場合、3DVAR や最適内挿法のような流れに依存しないデータ同化を行った場合は比較的安定にデータ同化サイクルが実現されるのに対し、EnKF では不安定ですぐに誤差が急激に増大してしまう。これは、与えられた条件下で厳しく最適性を求め、得られる情報を最大限に活用して最高の結果を得ようとするために起こるものである。すなわち、少しでも条件が満たされないと、その最適性を失い、悪い結果を与えてしまう。これを簡単に解決するのが、EnKF の最適性を緩和することである。具体的には、第22式で導入された共分散膨張のパラメータ δ を大きくし、予報誤差を故意に大きく見積もり、観測の影響が大きくなるようにする。また、ハイブリッド法も同じような働きをする。いずれも、カルマンフィルタの最適性を失わせる処理であるが、このことにより、カルマンフィルタの前提条件が多少満たされない場合にもその働きを拡張することができる。一方、カルマンフィルタの最適性をなるべく失わせることなく、モデルバイアスの問題に対処する方法も模索されている。Dee and da Silva (1998) は、カルマンフィルタを用いてモデルバイアスを動的に推定する手法を提案した。この他、複数の異なるモデルを使うことで、ある一つのモデルに特有のバイアスを打ち消し合わせようとするマルチモデル・アンサンブルも、アンサンブル予報の研究で最近話題となっている。これまでのところ、これらの新しい技術を EnKF に適用しようと試みた例はないようだが、最適なカルマンフィルタを最適なまま用いるための、将来の重要な課題の一つと言えよう。

もう一つの問題点として、EnKF は3DVAR と同様、一度に1時刻の観測しか扱えないということがある。例えば、6時間毎にデータ同化を行うことを考えた場合、6時間毎の観測データは使えるが、それ以外の毎時の観測データ等は使えないことになる。このような非定時の観測データは観測システムの進歩により近年増加してきており、これらを有効に活用することは実用上重要である。この問題を解決する方法として、Hunt *et al.* (2004) が4次元 EnKF (4D-EnKF) を提唱している。この方法は、アンサンブル予報を用いて観測の時刻を移動するというものであり、EnKF においても非定時の観測データを効率的に使うことができるようになる。

EnKF の実用上の利点は、第3節でも述べたように、時間発展モデルと観測演算子のそれぞれの接線形コード及びアジョイントコードが必要とされないことであ

る。EnKF と同様に流れに依存したデータ同化を行う手法で、現在主流となりつつある 4 次元変分法 (4DVAR) では、モデルの接線形コード及びアジョイントコードが必要とされ、これらの開発に多大なコストがかかる。その上、非線形の前報モデルは絶え間ない開発により随時更新され、これに追従して接線形コード及びアジョイントコードも更新する必要がある、その開発コストは継続的なものとなる。しかし、EnKF では非線形の前報モデルしか使われないため、頻繁に更新される前報モデルに対応した継続的な開発は必要とされない。

また、EnKF はアンサンブル予報の結果を使うため、アンサンブル予報を定常的にやっている環境があれば、データ同化のために新たにモデルの計算を行う必要がなく、効率的である。4DVAR では、モデルの時間発展を含んだ繰り返し計算が必要とされるため、1 回のデータ同化に多くの予報計算が必要とされる。このため、解析のために粗い格子を用意するなど、計算コストを下げるための工夫が必要となる。これに対し、EnKF では、データ同化をアンサンブル予報と一体化することにより、既に得られている情報を効率的に活用して、余分な予報計算を生じない。また、データ同化のために粗い格子を用意する必要もなく、アンサンブル予報で使われる格子系のままデータ同化を行うことができる。

以上に述べたように、EnKF は、そのいくつかの問題点については既に解決法が提案されており、その上、開発コストと計算コストの点で、実用上の重要な利点がある。このため、EnKF は先端研究者の間で広く将来性が認められ、研究も盛んに行われているものと言えよう。

7. まとめ

本稿では、EnKF がカルマンフィルタから実に自然な形で導入されるものであることを述べた。そのエッセンスは、誤差共分散行列の平方根を誤差ベクトルと見なすところにあった。しかしながら、この平方根は一意ではなく、また、アンサンブル・アップデートのプロセスとしても多くの方法が提案されている。また、効果的な実装方法も様々模索されている。無数の実装方法がある中で、どの方法が最良の結果を与えるかは、Tippett *et al.* (2003) も述べているように、自明ではない。次世代のデータ同化手法として将来を約束されているように例えられる EnKF だが、同じく流れに依

存する情報を取り込むデータ同化法で、最近主流となってきた 4DVAR と比べて、優れているのかどうかも自明ではない。このように、様々な手法を比較することが、今後の課題の 1 つとなる。

本稿の議論では、単純化のため、アンサンブル予報をデータ同化に活用するという観点からのみ EnKF を記述した。しかし、EnKF は、第 4 式のアンサンブル・アップデートのプロセスで、観測の情報も考慮した最良のアンサンブル摂動ベクトルを作り出す。Breeding 法では、観測データの分布も考慮したマスクをかけて摂動ベクトルを生成することがアンサンブル予報の改善をもたらすが、データ同化は 3DVAR など古いままで、流れに依存した情報が使われなくても、EnKF により与えられるアンサンブル・アップデートを行うだけで、アンサンブル予報が改善することも確かめられている (Wang and Bishop, 2003; Wang *et al.*, 2004)。また、従来のプラス・マイナスのペアというアンサンブルはメンバー数の半分の数の方向しか持たないのに対し、EnKF はすべてが異なる方向を持った摂動ベクトルを作り出す。これにより、特にアンサンブルメンバーの数が限られている場合、EnKF のアンサンブル・アップデートがアンサンブル予報を改善することが知られている (Wang *et al.*, 2004)。このように、EnKF がもたらすものは、アンサンブル予報を利用したデータ同化の改善だけでなく、データ同化を利用したアンサンブル予報の改善も含めた、双方向の改善なのである。

EnKF は、1994 年の Evensen に始まり、現在大きな広がりを見せているホットなトピックである。確率論的な予報を与えるアンサンブル予報が必要とされる限り、それをデータ同化に活用しようとする EnKF の方向性は効率的なものと言え、今後は実用化を視野に入れ、ますます研究が盛んとなる分野と言えよう。本稿が、読者の EnKF に対する理解を深め、最新の論文に取り組む際の一助となれば、著者の喜びである。

謝 辞

本稿の執筆に当たり、気象庁予報部数値予報課の露木 義課長、小泉 耕予報官、気象研究所の斉藤和雄予報第二研究室長の各氏から、貴重なコメントをいただいた。また、本稿は、著者が人事院長期在外研究員として米国メリーランド大学カレッジパーク校滞在中に執筆したものである。メリーランド大学の Eugenia Kalnay 教授, Joaquim Ballabrera 博士, Shu-Chih

Yang, John Harlim の各氏との有意義な議論が本稿の随所で役立っている。本稿の投稿に際しては、気象庁予報部数値予報課の石田純一、荒波恒平の両氏にご協力頂いた。最終稿を仕上げるに当たり、担当編集委員の大淵 清氏と匿名の査読者には、大変貴重な助言をいただいた。本稿の執筆を可能にした、これらすべての援助等に心から感謝の意を表したい。

付 録

A. カルマンフィルタの導出

本付録では、第2節で導入したカルマンフィルタを構成する5つの方程式（第1～5式）の簡単な導出を行う。より数学的に厳密な導出や解説は、Jazwinski (1970)を参照して頂きたい。簡単な導出や、より詳細な解説は Gelb *et al.* (1974) に記述がある。本節中で導入された記号は、本付録中では導入なしに用いる。

第1式は予報モデルの定義とも言える。予報モデル $\mathbf{M}$ は、1時刻前の解析値（初期値）から現在の予報値への写像を与える時間発展演算子である。

$$\mathbf{x}_i^f = \mathbf{M}\mathbf{x}_{i-1}^a \quad (\text{A-1})$$

解析値や予報値の誤差を、

$$\delta\mathbf{x}^{a,f} = \mathbf{x}^{a,f} - \mathbf{x}^t \quad (\text{A-2})$$

で定義する。ここで、添字 t は真の状態 (truth) を表し、添字 a, f は解析 a または予報 f のいずれにも同じ式が成り立つことを表す。これらの誤差の共分散行列 $\mathbf{P}^{a,f}$ は、

$$\mathbf{P}^{a,f} = \langle \delta\mathbf{x}^{a,f} \delta\mathbf{x}^{a,fT} \rangle \quad (\text{A-3})$$

のように書かれる。ここで、記号 $\langle \bullet \rangle$ は統計期待値を表す。真の時間発展 $\mathbf{M}^t$ を

$$\mathbf{x}_i^t = \mathbf{M}^t \mathbf{x}_{i-1}^t \quad (\text{A-4})$$

で定義する。モデル $\mathbf{M}$ が完全なときは、 $\mathbf{M}$ は $\mathbf{M}^t$ と一致するが、一般にはそうではなく、

$$\mathbf{x}_i^f = \mathbf{M}\mathbf{x}_{i-1}^f - \mathbf{w}$$

のように、誤差 $\mathbf{w}$ を含む。ここでは簡単のため、モデ

ルの誤差 $\mathbf{w}$ はバイアスがないとする ($\langle \mathbf{w} \rangle = 0$)。これらを用いて、予報誤差の共分散行列 $\mathbf{P}^f$ は、

$$\begin{aligned} \mathbf{P}_i^f &= \langle \delta\mathbf{x}_i^f \delta\mathbf{x}_i^{fT} \rangle \\ &= \langle (\mathbf{x}_i^f - \mathbf{x}_i^t) (\mathbf{x}_i^f - \mathbf{x}_i^t)^T \rangle \\ &= \langle (\mathbf{M}\mathbf{x}_{i-1}^a - \mathbf{M}^t \mathbf{x}_{i-1}^t) (\mathbf{M}\mathbf{x}_{i-1}^a - \mathbf{M}^t \mathbf{x}_{i-1}^t)^T \rangle \\ &= \langle (\mathbf{M}(\mathbf{x}_{i-1}^a - \mathbf{x}_{i-1}^t) + \mathbf{w}) (\mathbf{M}(\mathbf{x}_{i-1}^a - \mathbf{x}_{i-1}^t) + \mathbf{w})^T \rangle \\ &= \langle (\mathbf{M}\delta\mathbf{x}_{i-1}^a + \mathbf{w}) (\mathbf{M}\delta\mathbf{x}_{i-1}^a + \mathbf{w})^T \rangle \\ &= \langle (\mathbf{M}\delta\mathbf{x}_{i-1}^a) (\mathbf{M}\delta\mathbf{x}_{i-1}^a)^T + (\mathbf{M}\delta\mathbf{x}_{i-1}^a) \mathbf{w}^T + \mathbf{w} (\mathbf{M}\delta\mathbf{x}_{i-1}^a)^T + \mathbf{w}\mathbf{w}^T \rangle \\ &= \langle \mathbf{M}\delta\mathbf{x}_{i-1}^a \delta\mathbf{x}_{i-1}^{aT} \mathbf{M}^T \rangle + \langle \mathbf{w}\mathbf{w}^T \rangle \\ &= \mathbf{M}\mathbf{P}_{i-1}^a \mathbf{M}^T + \mathbf{Q} \end{aligned} \quad (\text{A-5})$$

のように変形される。ここで、解析誤差 $\delta\mathbf{x}^a$ とモデル誤差 $\mathbf{w}$ の間に相関はないと仮定し ($\langle \delta\mathbf{x}^a \mathbf{w}^T \rangle$ などのクロスタームは0)、 $\mathbf{Q}$ はモデル誤差の共分散行列とした ($\mathbf{Q} = \langle \mathbf{w}\mathbf{w}^T \rangle$)。これが、本節中第2式で与えた誤差共分散行列の時間発展の式の導出である。

次に、解析の式を導入する。解析は、予報 $\mathbf{x}^f$ を観測 $\mathbf{y}^o$ で修正するプロセスで、それぞれの信頼度 (誤差の大きさ) に応じた重み付き平均をとる。簡単のため1次元の場合を考えると、変数 x と変数 y の重み付き平均は、 y の重みを k ($0 \leq k \leq 1$) としたとき、 $(1-k)x + ky$ で与えられる。これと同様に、同じ空間に属するベクトル変数 $\mathbf{x}$ と $\mathbf{y}$ の重み付き平均は、 $\mathbf{y}$ の重みを $\mathbf{K}$ としたとき、 $(\mathbf{I} - \mathbf{K})\mathbf{x} + \mathbf{K}\mathbf{y}$ で与えられる。今、予報 $\mathbf{x}^f$ と観測 $\mathbf{y}^o$ は異なる空間に属し、それらは写像 $\mathbf{H}$ (観測演算子) で結びつけられている。この場合の重み付き平均は、観測 $\mathbf{y}^o$ の重みを $\mathbf{K}$ としたとき、 $(\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{x}^f + \mathbf{K}\mathbf{y}^o$ で与えられる。つまり、解析 $\mathbf{x}^a$ は、予報 $\mathbf{x}^f$ と観測 $\mathbf{y}^o$ の重み付き平均として、本節中の第3式

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f) \quad (\text{A-6})$$

で与えられる。推定誤差を最小にする観測の重み $\mathbf{K}$ は、解析誤差の分散を最小にすることで得られる。第A-2式と同様に、観測の誤差 $\delta\mathbf{y}$ を

$$\delta\mathbf{y} = \mathbf{y}^o - \mathbf{H}\mathbf{x}^t \quad (\text{A-7})$$

で定義する。これを用いると、 $\mathbf{y}^o - \mathbf{H}\mathbf{x}^f$ は、

$$\begin{aligned} \mathbf{y}^o - \mathbf{H}\mathbf{x}^f &= \mathbf{y}^o - \mathbf{H}\mathbf{x}^t + \mathbf{H}\mathbf{x}^t - \mathbf{H}\mathbf{x}^f \\ &= \delta\mathbf{y} - \mathbf{H}\delta\mathbf{x}^f \end{aligned} \quad (\text{A-8})$$

のように変形できる. 第 A-6 式の両辺から $\mathbf{x}^t$ を引いて, 第 A-8 式を用いて変形すると,

$$\delta\mathbf{x}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\delta\mathbf{x}^f + \mathbf{K}\delta\mathbf{y} \quad (\text{A-9})$$

が得られる. この 2 次モーメントを計算すると,

$$\begin{aligned} \mathbf{P}^a &= \langle \delta\mathbf{x}^a \delta\mathbf{x}^{aT} \rangle \\ &= \langle ((\mathbf{I} - \mathbf{K}\mathbf{H})\delta\mathbf{x}^f + \mathbf{K}\delta\mathbf{y}) \\ &\quad ((\mathbf{I} - \mathbf{K}\mathbf{H})\delta\mathbf{x}^f + \mathbf{K}\delta\mathbf{y})^T \rangle \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T \end{aligned} \quad (\text{A-10})$$

となる. ここで, 予報誤差と観測誤差には相関がないものと仮定し, $\langle \delta\mathbf{x}^f(\mathbf{K}\delta\mathbf{y})^T \rangle$ などのクロスタームは 0 とした. 解析誤差の分散は, 2 次モーメントのトレース和で与えられる. $\text{trace}(\mathbf{P}^a)$ を重み $\mathbf{K}$ の関数と見て, これが最小となるような $\mathbf{K}$ を探すには,

$$\frac{\partial}{\partial \mathbf{K}}(\text{trace}(\mathbf{P}^a)) = 0 \quad (\text{A-11})$$

を解けばよい. この際, 数学の公式

$$\frac{\partial}{\partial \mathbf{A}}(\text{trace}(\mathbf{A}\mathbf{B}\mathbf{A}^T)) = \mathbf{A}(\mathbf{B} + \mathbf{B}^T) \quad (\text{A-12})$$

$$\frac{\partial}{\partial \mathbf{A}}(\text{trace}(\mathbf{A}\mathbf{B})) = \mathbf{B}^T \quad (\text{A-13})$$

が有用である (Gelb *et al.*, 1974, 第 2.1-73, 2.1-72 式を参照). 第 A-10 式を第 A-11 式に代入し, 第 A-12, A-13 式を用いて計算すると,

$$-2(\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f\mathbf{H}^T + 2\mathbf{K}\mathbf{R} = 0 \quad (\text{A-14})$$

となる. ここで, 共分散行列 $\mathbf{P}$ 及び $\mathbf{R}$ が対称行列であることを用いた. これを $\mathbf{K}$ について解けば, 本節中の第 5 式

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1} \quad (\text{A-15})$$

が得られる. これを第 A-10 式に代入し, 簡単のため $\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R} = \mathbf{S}$ と表記すると,

$$\begin{aligned} \mathbf{P}^a &= (\mathbf{I} - \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H})\mathbf{P}^f(\mathbf{I} - \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H})^T \\ &\quad + \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{R}(\mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1})^T \\ &= \mathbf{P}^f - 2\mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f + \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f \\ &\quad + \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{R}\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f \\ &= \mathbf{P}^f - 2\mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f \\ &\quad + \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f \\ &= \mathbf{P}^f - 2\mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f + \mathbf{P}^f\mathbf{H}^T\mathbf{S}^{-1}\mathbf{S}\mathbf{S}^{-1}\mathbf{H}\mathbf{P}^f \\ &= \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{P}^f \\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f \end{aligned} \quad (\text{A-16})$$

のように本節中の第 4 式が得られる.

B. 観測摂動法 (PO 法)

本付録では, PO 法における本節中第 17 式の導出を行う. $\delta\mathbf{x}^a$, $\delta\mathbf{x}^f$, $\delta\mathbf{y}^o$ を, それぞれ, 変数 $\mathbf{x}^a$, $\mathbf{x}^f$, $\mathbf{y}^o$ の中心値 (アンサンブル平均) $\bar{\mathbf{x}}^a$, $\bar{\mathbf{x}}^f$, $\bar{\mathbf{y}}^o$ からの摂動とする. つまり,

$$\begin{aligned} \mathbf{x}^a &= \bar{\mathbf{x}}^a + \delta\mathbf{x}^a \\ \mathbf{x}^f &= \bar{\mathbf{x}}^f + \delta\mathbf{x}^f \\ \mathbf{y}^o &= \bar{\mathbf{y}}^o + \delta\mathbf{y}^o \end{aligned} \quad (\text{B-1})$$

が成り立つ. PO 法では, それぞれのアンサンブルメンバーについて独立に解析を行うので, $\mathbf{x}^a$, $\mathbf{x}^f$, $\mathbf{y}^o$ について本節中の第 3 式で与えられるデータ同化の式

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f) \quad (\text{B-2})$$

が成り立つ. すなわち,

$$\bar{\mathbf{x}}^a + \delta\mathbf{x}^a = \bar{\mathbf{x}}^f + \delta\mathbf{x}^f + \mathbf{K}[\bar{\mathbf{y}}^o + \delta\mathbf{y}^o - \mathbf{H}(\bar{\mathbf{x}}^f + \delta\mathbf{x}^f)] \quad (\text{B-3})$$

アンサンブル平均についても第 B-2 式が成り立つので,

$$\delta\mathbf{x}^a = \delta\mathbf{x}^f + \mathbf{K}[\delta\mathbf{y}^o - \mathbf{H}\delta\mathbf{x}^f] \quad (\text{B-4})$$

も成り立つことになる. 観測に摂動がないとき ($\delta\mathbf{y}^o = 0$), 第 B-4 式は,

$$\delta \mathbf{x}^a = [\mathbf{I} - \mathbf{K}\mathbf{H}] \delta \mathbf{x}^f \quad (\text{B-5})$$

となる。すべてのアンサンブルメンバーについて第 B-5 式が成り立つので、各列がアンサンブルメンバーから構成される行列 $\mathbf{E}$ について、本節中第 17 式の

$$\mathbf{E}^a = [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{E}^f \quad (\text{B-6})$$

が成り立つ。

略語一覧

3DVAR : 3 次元変分法

4D-EnKF : 4 次元アンサンブル・カルマンフィルタ

4DVAR : 4 次元変分法

Breeding 法 (BGM 法) : Breeding of Growing Modes
(成長モードを育成するアンサンブル誤差ベクトル生成法)

EAKF : Ensemble Adjustment Kalman Filter (アンサンブル調節カルマンフィルタ)

EKF : Extended Kalman Filter (拡張カルマンフィルタ)

EnKF : Ensemble Kalman Filter (アンサンブル・カルマンフィルタ)

EnSRF : Ensemble Square Root Filter (アンサンブル平方根フィルタ)

ETKF : Ensemble Transform Kalman Filter (アンサンブル変換カルマンフィルタ)

LEKF : Local Ensemble Kalman Filter (局所アンサンブル・カルマンフィルタ)

NCEP : National Centers for Environmental Prediction (米国環境予測センター)

PO 法 : Perturbed Observation method (摂動観測法)

SRF : Square Root Filter (平方根フィルタ)

参 考 文 献

Anderson, J. L., 2001 : An ensemble adjustment Kalman filter for data assimilation, *Mon. Wea. Rev.*, **129**, 2884-2903.

Andrews, A., 1968 : A square root formulation of the Kalman covariance equations, *AIAA J.*, **6**, 1165-1168.

Bishop, C. H., B. J. Etherton and S. J. Majumdar, 2001 : Adaptive sampling with the ensemble transform Kalman Filter. Part I : Theoretical aspects, *Mon. Wea. Rev.*, **129**, 420-436.

Burgers, G., P. J. van Leeuwen and G. Evensen, 1998 : Analysis scheme in the ensemble Kalman filter,

Mon. Wea. Rev., **126**, 1719-1724.

Corazza, M., E. Kalnay, D. J. Patil, E. Ott, J. Yorke, I. Szunyogh and M. Cai, 2002 : Use of the breeding technique in the estimation of the background error covariance matrix for a quasigeostrophic model, *AMS Symposium on Observations, Data Assimilation and Probabilistic Prediction*, Orlando, Florida, 154-157.

Corazza, M., E. Kalnay, D. J. Patil, S.-C. Yang, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt and J. A. Yorke, 2003 : Use of the breeding technique to estimate the structure of the analysis "errors of the day", *Nonlinear Proc. Geophys.*, **10**, 1-11.

Dee, D. P. and A. M. da Silva, 1998 : Data assimilation in the presence of forecast bias, *Quart. J. Roy. Meteor. Soc.*, **124**, 269-295.

Evensen, G., 1994 : Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, **99C5**, 10143-10162.

Evensen, G., 2003 : The ensemble Kalman filter : theoretical formulation and practical implementation, *Ocean Dynamics*, **53**, 343-367.

Gelb, A., J. F. Kasper, R. A. Nash, C. F. Price and A. A. Sutherland, 1974 : Applied optimal estimation, The M. I. T. Press, 374pp.

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

Hamill, T. M., 2003 : Ensemble-based atmospheric data assimilation : A tutorial, NOAA-CIRES Climate Diagnostics Center, 46pp.

Houtekamer, P. L. and H. L. Mitchell, 1998 : Data assimilation using an ensemble Kalman filter technique, *Mon. Wea. Rev.*, **126**, 796-811.

Houtekamer, P. L. and H. L. Mitchell, 2001 : A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Wea. Rev.*, **129**, 123-137.

Hunt, B. R., E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Sauer, I. Szunyogh, J. A. Yorke and A. V. Zimin, 2004 : Four-dimensional ensemble Kalman filter, *Tellus*, **56A**, 273-277.

Jazwinski, A. H., 1970 : Stochastic processes and filtering theory, Academic Press, 376pp.

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

Lorenz, E., 1996 : Predictability - A problem partly solved, *Proc. Seminar on Predictability*, Reading, UK, ECMWF, 1-18.

- Lorenz, E. and K. Emanuel, 1998 : Optimal sites for supplementary weather observations, *J. Atmos. Sci.*, **55**, 399-414.
- Miyoshi, T., 2004 : A review of the deterministic ensemble Kalman filtering methods and related techniques, Masters Scholarly Paper. Department of Meteorology, University of Maryland, College Park, 20pp.
- Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil and J. A. Yorke, 2002 : Exploiting local low dimensionality of the atmospheric dynamics for efficient Kalman filtering, *ArXiv : arc-ive/paper 0203058*, <http://arxiv.org/abs/physics/0203058>.
- Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil and J. A. Yorke, 2004 : A local ensemble Kalman filter for atmospheric data assimilation, *Tellus*, **56A**, 415-428.
- Patil, D. J., B. R. Hunt, E. Kalnay, J. A. Yorke and E. Ott, 2001 : Local low dimensionality of atmospheric dynamics, *Phys. Rev. Lett.*, **86**, 5878-5881.
- Szunyogh, I., E. J. Kostelich, G. Gyarmati, D. J. Patil, B. R. Hunt, E. Kalnay, E. Ott and J. A. Yorke, 2004 : Assessing a local ensemble Kalman filter : Perfect model experiments with the NCEP global model, *Tellus*, submitted.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003 : Ensemble square root filters, *Mon. Wea. Rev.*, **131**, 1485-1490.
- Toth, Z. and E. Kalnay, 1993 : Ensemble forecasting at NMC : The generation of perturbations, *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.
- Toth, Z. and E. Kalnay, 1997 : Ensemble forecasting at NCEP and the breeding method, *Mon. Wea. Rev.*, **125**, 3297-3319.
- 露木 義, 1997 : 変分法によるデータ同化, 数値予報課報告・別冊第43号, 気象庁予報部, 102-165.
- 露木 義, 2004 : データ同化とアンサンブル予報の統合に向けて, 月刊海洋, in press.
- Wang, X. and C. H. Bishop, 2003 : A Comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, *J. Atmos. Sci.*, **60**, 1140-1158.
- Wang, X., C. H. Bishop and S. J. Julier, 2004 : Which is better, and ensemble of positive-negative pairs or centered spherical simplex ensemble?, *Mon. Wea. Rev.*, **132**, 1590-1605.
- Whitaker, J. S. and T. M. Hamill, 2002 : Ensemble data assimilation without perturbed observations, *Mon. Wea. Rev.*, **130**, 1913-1924.

Ensemble Kalman Filtering

—A Meeting Point between Data Assimilation and Ensemble Forecasting—

Takemasa MIYOSHI

Department of Meteorology, University of Maryland, College Park, Maryland 20742, USA
Numerical Prediction Division, Japan Meteorological Agency, 1-3-4 Otemachi, Chiyoda-ku, Tokyo
100-8122, Japan

(Received 21 June 2004 : Accepted 22 November 2004)