

# 背景誤差共分散行列に関する補足

伊藤耕介（琉球大学）

## 本解説文について

本資料は、「データ同化実習」でデータ同化の基礎理論について触れたあと、大自由度系のデータ同化を行う際に一度は悩んでしまう背景誤差共分散  $B$  (カルマンフィルタ系列の手法では  $P^f$  と書くことが多い) の扱いについて学ぶためのものである。教科書的には、カルマンフィルタ系列の手法やアジョイント法を実際の問題に適用するためには、 $B^{-1}$  を計算することが必要となるが、大自由度系では直接的に逆行列を得ることが非常に困難であり、そもそも、 $B$  が正則でないために  $B^{-1}$  が存在しない場合も多い<sup>1</sup>。

本資料では、データ同化における背景誤差共分散の必要性について簡単に触れ、現実的な実装において利用されているいくつかの方法について紹介する。簡単のため、観測演算子は線形の観測行列で表されるとし、観測される量はそのままモデル変数に対応しており、内挿やリトリバルを介さないとする。

また、以下の文章では基本的に「データ同化実習」と同じ記号を用いている。

## 1 なぜ背景誤差分散が必要か？

そもそも、背景誤差共分散行列  $B$  は、なぜ必要なのだろう？そのひとつの答えは「解が一つに定まらないから」である。観測値の数は一般に解析値として求められる変数の数よりも圧倒的に少ない。そのため、観測値との差がゼロとなる解析値を探索しようとする、数学的には解が一つに定まらないことになる<sup>2</sup>。また、別の答え方として「観測が存在しない地点で第一推定値からあまりにも大きくずれた解析値が得られた場合には、解析値を不適なものとして扱わなければならない」とも言える。例えば、観測がある地点でぴったりと観測値と解析値が合うからと言って、観測がない地点の地上気温が100度になってはいけな。さらにいうと「背景誤差共分散を用いることにより、観測点から離れた地点でも改善が見込まれる」という言い方もできる。例えば、ある観測点で東西風速の観測が得られて東風を強めるような修正がかかるとき、その周囲でも東風を一定の割合で強めることは自然だし、地衡風平衡・静水圧平衡が期待できる現象・時間スケールを扱うとき、力学的な根拠に基づいて南北風速や温度変化を修正することは情報を有効に生かすことになる。観測値の情報をもとに抽出される解析インクリメント (解析値と第一推定値の差) の構造は4次元データ同化の場合  $M$  と  $B$  の両方に依存するため、 $B$  の適切な設定は欠かすことのできないものである。

実際に背景誤差共分散行列  $B$  をどのように構成するかについては、距離に応じて誤差共分散を小さくするだけの単純なものから、教科書にあるようにコバリアンスマッチングなどを用いるものなどさまざまな方法がある。現業機関の気象学におけるデータ同化では、 $B$  の構成法として、異なる

<sup>1</sup>そういう意味では、私 (伊藤) は  $B^{-1}$  という表現は、単に概念を表しているに過ぎないと考えている。

<sup>2</sup>では、背景誤差共分散行列を加えると、どのような問題を解いたことになるのか、というのは結構ややこしい話になる。興味のある人は、Johnson et al.(2005; QJRM) や私 (伊藤) のホームページにある「逆問題としてのデータ同化」(<http://www.itonwp.sci.u-ryukyu.ac.jp/distribution.html>) を参照いただきたい。

初期時刻から計算を開始して同じ評価時刻で得られた「差」をたくさん集めて統計処理する NMC 法 (Parrish and Derber, 1992) と呼ばれる方法が長らく使われているが、近年ではアンサンブルランの結果を基に時々刻々変化する  $B$  を適用する現業機関も増えてきている (Lorenc, 2003) . 特に、アンサンブルカルマンフィルタの結果をもとに構成された  $B$  を使って 4 次元変分法を実行する同化手法は Hybrid EnKF-4DVAR 法と呼ばれている .

## 2 背景誤差共分散の逆行列の扱い方

冒頭で述べたように、背景誤差共分散行列の逆行列  $B^{-1}$  は大自由度系においては求めることが現実的ではない、あるいは、逆行列自体が存在しないことも珍しくない . そのため、これを回避するための方法が必要となる . そのうち、いくつか代表的なものを挙げると

- 単位行列で近似 (簡便ではあるが誤差相関の情報を生かせない)
- $B^{1/2}$  やアンサンブル変換行列を用いた変換
- 特異値分解を利用し一般化逆行列で近似

といったものがある . ここでは変換を用いる方法と一般化逆行列による近似について述べる . 既存のシステムでは、本資料で示す以上の改良がくわえられていたり、ユーザーがこれらの操作を意識しなくてもデータ同化が行えるようになっていることも多いが、いずれにせよ  $B$  の扱いに関して、基礎的な知識を身につけておくことは重要である .

### 2.1 変分法における $B^{1/2}$ を用いた変数変換

背景誤差共分散行列の逆行列  $B^{-1}$  を計算しないで済む方法として、 $B^{1/2}$  を使って、制御変数を変換するという手法がある<sup>3</sup> . ここでは 4 次元変分法の初期値を制御することを例にとって解説する . もし、3 次元変分法に適用したい場合には、以下の式を  $t = 0$  とし、 $M = I$  とすればよい . 具体的な変換方法としては、 $\delta \mathbf{x}_0 \equiv \mathbf{x}_{0,iterk} - \mathbf{x}_{0,iter0}$  ( $\mathbf{x}_{0,iter0}$  は  $\mathbf{x}_b$  に等しい) 及び  $\delta \mathbf{x}_0 = B^{1/2} \delta \mathbf{v}_0$  と変換し、ここで現れた  $\delta \mathbf{v}_0$  を制御変数とする<sup>4</sup> . これにより、評価関数  $J$  及びは評価関数の  $\delta \mathbf{v}_0$  に対する勾配は、

$$J = \frac{1}{2} \delta \mathbf{v}_0^T \delta \mathbf{v}_0 + \frac{1}{2} \sum_t (\mathbf{H} \mathbf{x}_t - \mathbf{y}_t)^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x}_t - \mathbf{y}_t) \quad (1)$$

$$= \frac{1}{2} \delta \mathbf{v}_0^T \delta \mathbf{v}_0 + \frac{1}{2} \sum_t \left( \mathbf{H} \mathbf{M} \mathbf{B}^{1/2} \delta \mathbf{v}_0 - \mathbf{d}_t \right)^T \mathbf{R}^{-1} \left( \mathbf{H} \mathbf{M} \mathbf{B}^{1/2} \delta \mathbf{v}_0 - \mathbf{d}_t \right) \quad (2)$$

$$\frac{\partial J}{\partial \delta \mathbf{v}_0} = \delta \mathbf{v}_0 + \mathbf{B}^{T/2} \sum_t \mathbf{M}^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x}_t - \mathbf{y}_t) \quad (3)$$

$$= \delta \mathbf{v}_0 + \mathbf{B}^{T/2} \sum_t \mathbf{M}^T \mathbf{H}^T \mathbf{R}^{-1} \left( \mathbf{H} \mathbf{M} \mathbf{B}^{1/2} \delta \mathbf{v}_0 - \mathbf{d}_t \right) \quad (4)$$

<sup>3</sup> $B$  が半正定値行列であるときには、 $B = B^{1/2} (B^{1/2})^T$  を満たす平方根行列と呼ばれる行列が存在する .  $B$  が  $\mathbf{U} \mathbf{D} \mathbf{U}^T$  と固有値分解できるなら、 $B^{1/2} = \mathbf{U} \mathbf{D}^{1/2}$  である .

<sup>4</sup>インクリメンタルアプローチを施すときは  $\delta \mathbf{x}_0 \equiv \mathbf{x}_{0,iter(k+1)} - \mathbf{x}_{0,iterk}$  などをとる .

のように変形できる．ここで， $\mathbf{d}_t \equiv \mathbf{y}_t - \mathbf{H}\mathbf{x}_t$  は観測値と第一推定値との差を表すベクトルで，イノベーションと呼ばれる．ここでは，評価関数の一項目を  $J_b$  項とし，二項目を  $J_o$  項と呼ぶことにする． $J_o$  項の  $\mathbf{x}_0$  に対する勾配 (例えば，教科書 p.102) と  $J_o$  項の  $\mathbf{v}_0$  に対する勾配を見比べてみると，

$$\frac{\partial J_o}{\partial \delta \mathbf{v}_0} = \mathbf{B}^{T/2} \frac{\partial J_o}{\partial \delta \mathbf{x}_0} \quad (5)$$

という関係にあることが分かる．

実際の計算手続きとしては以下のようなものが考えられる．

- (1)  $J_o$  項及び  $J_b$  項の  $\delta \mathbf{x}_0$  に対する勾配  $\partial J_o / \partial \delta \mathbf{x}_0$  をアジョイントモデルを用いて計算する．
- (2)  $\partial J_o / \partial \delta \mathbf{v}_0 = \mathbf{B}^{T/2} \partial J_o / \partial \delta \mathbf{x}_0$  を使って，新しい制御変数  $\delta \mathbf{v}_0$  に対する勾配を計算
- (3)  $J_b$  項及び  $J_b$  項の勾配の計算を行う．そのためには， $\delta \mathbf{v}_0$  を計算すればよいが，1 回目のイテレーションにおいては  $\delta \mathbf{v}_0 = 0$  であり ( $\delta \mathbf{v}_0$  の定義を思い出そう)，2 回目のイテレーション以降では，(4) によって計算されたものを用いる．
- (4)  $J$ ， $\delta \mathbf{v}_0$ ， $\partial J / \partial \delta \mathbf{v}_0$  を最適化アルゴリズムに渡して，推定される  $\delta \mathbf{v}_0$  を更新する．
- (5)  $\delta \mathbf{x}_0 = \mathbf{B}^{1/2} \delta \mathbf{v}_0$  によって， $\delta \mathbf{x}_0$  を更新する．

この (1)-(5) の手続きを計算が収束するまで繰り返す．こうすることによって， $\mathbf{B}$  の逆行列  $\mathbf{B}^{-1}$  の計算を回避し， $\mathbf{B}^{1/2}$  を利用した計算を行うことができる．

## 2.2 アンサンブルカルマンフィルタにおける変数変換の例

アンサンブルカルマンフィルタにおいては，複数の変数変換の方法が提案されている．その中でも，ここでは，三好 (2008, 気象研究ノート) に従い，局所アンサンブル変換カルマンフィルタ行列における変数変換について述べる．

この方法では，解析値  $\mathbf{X}_t^a$  と第一推定値  $\mathbf{X}_t^f$  との間に

$$\mathbf{X}_t^a = \mathbf{X}_t^f \mathbf{T} \quad (6)$$

となるような  $N \times N$  の行列  $\mathbf{T}$  を作用させる．この  $\mathbf{T}$  はアンサンブル変換行列とも呼ばれる．ここで，

$$\mathbf{X}_t^a \equiv \frac{1}{\sqrt{N-1}} \left( \mathbf{x}_t^{a,1} - \bar{\mathbf{x}}_t^a, \mathbf{x}_t^{a,2} - \bar{\mathbf{x}}_t^a, \dots, \mathbf{x}_t^{a,N} - \bar{\mathbf{x}}_t^a \right) \quad (7)$$

$$\mathbf{X}_t^f \equiv \frac{1}{\sqrt{N-1}} \left( \mathbf{x}_t^{f,1} - \bar{\mathbf{x}}_t^f, \mathbf{x}_t^{f,2} - \bar{\mathbf{x}}_t^f, \dots, \mathbf{x}_t^{f,N} - \bar{\mathbf{x}}_t^f \right) \quad (8)$$

である．

このとき，カルマンフィルタの更新に関する式，

$$\mathbf{P}_t^a = (\mathbf{I} - \mathbf{K}_t \mathbf{H}) \mathbf{P}_t^f \quad (9)$$

に， $\mathbf{K}_t = \mathbf{X}_t^f (\mathbf{X}_t^f)^T \mathbf{H}_t^T (\mathbf{R}_t + \mathbf{H} \mathbf{X}_t^f (\mathbf{X}_t^f)^T \mathbf{H}^T)^{-1}$ ， $\mathbf{P}_t^f = \mathbf{X}_t^f (\mathbf{X}_t^f)^T$ ， $\mathbf{P}_t^a = \mathbf{X}_t^a (\mathbf{X}_t^a)^T$  を代入し，Sherman-Morrison-Woodbury の式<sup>5</sup>を用いて，アンサンブル変換行列についてこれを整理すると，

$$\mathbf{T} \mathbf{T}^T = \left[ \mathbf{I} + (\mathbf{H} \mathbf{X}_t^f)^T \mathbf{R}_t^{-1} \mathbf{H} \mathbf{X}_t^f \right]^{-1} \quad (10)$$

<sup>5</sup> $(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{C}^{-1} + \mathbf{V} \mathbf{A}^{-1} \mathbf{U})^{-1} \mathbf{V} \mathbf{A}^{-1}$  という恒等式．教科書 (A.12) 式に等価な式が載っている．

となる．この時点では，まだ，逆行列の計算が残っているが，すでに解くべき逆行列のサイズは  $N \times N$  となっている．さらに，

$$\mathbf{I} + (\mathbf{H}\mathbf{X}_t^f)^T \mathbf{R}_t^{-1} \mathbf{H}\mathbf{X}_t^f = \mathbf{U}\mathbf{D}\mathbf{U}^T \quad (11)$$

となる固有値分解を施し， $\mathbf{U}\mathbf{U}^T = \mathbf{I}$ ，すなわち， $\mathbf{U}^{-1} = \mathbf{U}^T$  となることを使えば，

$$\mathbf{T} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{U}^T \quad (12)$$

として  $\mathbf{T}$  を得ることができる．結局，逆行列を解かずに，ただかアンサンブルメンバー数程度のサイズの行と列を持つ行列の固有値分解を解くという問題に帰着される．

## 2.3 特異値分解を用いた一般化逆行列の利用

$\mathbf{B}$  を特異値分解し，Moore-Penrose 型の一般化逆行列 (疑似逆行列ともいう；以下，単に一般化逆行列<sup>6</sup>と言う)  $\mathbf{B}^{-I}$  を逆行列  $\mathbf{B}^{-1}$  の代わりに使うのも現実的な実装方法の一つである．

ここでは， $\mathbf{B}$  が

$$\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{V}^T = \sum_{l=1}^m \mathbf{u}_l \lambda_l \mathbf{v}_l^T \quad (13)$$

のように特異値分解されるとする． $\mathbf{u}_l, \mathbf{v}_l$  は特異値  $\lambda_l$  に対応した左特異ベクトルと右特異ベクトルである．以下，簡単のため， $\lambda_l$  は非負の特異値が大きいものから順に並べられているとする．また， $\mathbf{D}$  は対角成分以外はゼロ行列で，対角成分には非負の特異値  $\lambda_l$  が大きいものから順に収められている． $\mathbf{U}$  と  $\mathbf{V}$  は，対応した特異ベクトルからなる行列である．

このとき，一般化逆行列  $\mathbf{B}^{-I}$  は

$$\mathbf{B}^{-I} = \mathbf{V}\mathbf{D}^{-I}\mathbf{U}^T = \sum_{l=1}^m \mathbf{v}_l \lambda_l^{-1} \mathbf{u}_l^T \quad (14)$$

と表される．また， $\mathbf{D}^{-I}$  は  $\mathbf{D}$  のゼロ成分以外を逆数で置き換え，それ以外をゼロとする行列である． $\mathbf{B}$  に逆行列が存在しないような場合であっても，一般化逆行列は必ず存在し，計算することができる．

ただし，式 (14) において非常に小さい特異値の逆数に関わる計算をしてしまうと，計算安定性の観点からあまりよろしくない (本来ノイズであるべきモードが増幅される)．そこで，実際には， $\mathbf{B}$  のうち，特異値の大きなものから  $m'$  個を選び， $m' + 1$  番目以降の小さな特異値を 0 として再構成した  $\hat{\mathbf{B}}$  を代わりに用いることも一般的である<sup>7</sup>．すなわち，

$$\hat{\mathbf{B}} = \mathbf{U}\hat{\mathbf{D}}\mathbf{V}^T = \sum_{l=1}^{m'} \mathbf{u}_l \lambda_l \mathbf{v}_l^T \quad (15)$$

という  $\mathbf{B}$  の近似行列  $\hat{\mathbf{B}}$  に対し，

$$\hat{\mathbf{B}}^{-I} = \mathbf{V}\hat{\mathbf{D}}^{-I}\mathbf{U}^T = \sum_{l=1}^{m'} \mathbf{v}_l \lambda_l^{-1} \mathbf{u}_l^T \quad (16)$$

<sup>6</sup> $\mathbf{A}$  の Moore-Penrose 型一般化逆行列  $\mathbf{A}^{-I}$  とは， $\mathbf{A}$  の随伴行列を  $\mathbf{A}^*$  と表すことにすると， $\mathbf{A}\mathbf{A}^{-I}\mathbf{A} = \mathbf{A}$ ， $\mathbf{A}^{-I}\mathbf{A}\mathbf{A}^{-I} = \mathbf{A}^{-I}$ ， $(\mathbf{A}\mathbf{A}^{-I})^* = \mathbf{A}\mathbf{A}^{-I}$ ， $(\mathbf{A}^{-I}\mathbf{A})^* = \mathbf{A}^{-I}\mathbf{A}$  を満たすものであり，ただ一つに定まる． $\mathbf{A}$  が正則のときには，逆行列  $\mathbf{A}^{-1}$  がこの条件を満たしていることから，逆行列の一般化となっていることが分かる．また，一般化逆行列と呼ばれるものには何種類かあるが，Moore-Penrose 型一般化逆行列が最もよく使われる (?) ．

<sup>7</sup>近似行列のランクを定めたとき， $\hat{\mathbf{B}}$  は，もとの行列  $\mathbf{B}$  との行列の要素の差の二乗和が最小となる行列であることが証明できる (証明は略) ．

という一般化逆行列を考え、これを逆行列の代わりとして用いる。ただし、 $\hat{D}^{-I}$  は  $D^{-I}$  の対角成分のうち  $m' + 1$  番目より小さい特異値をすべてゼロとした行列である。このような低ランク行列による近似には、計算安定性を確保するだけでなく計算時間を短縮できるという点でも実用的な利点がある。

### 3 Lorenz96 モデル

4次元変分法において  $B^{1/2}$  を用いて変数変換する方法について、サンプルプログラムに示している。基礎方程式は以下の式で示される Lorenz(1996) の式である。これを時間方向にオイラー法で差分している。

$$\frac{dx_i}{dt} = -x_{i-2}x_{i-1} + x_{i-1}x_{i+1} - x_i + F \quad (17)$$

自由度は 40,  $F=8$  とする。初期条件としては、あらかじめ積分して計算結果を得ておき、積分時間が異なる 2 つの結果を真の値・第一推定値として与えている。これを基にして、いくつかの問題に挑戦してみよう。

- 3次元変分法との比較
- 背景誤差共分散行列の平方根の設定を変更し、実際に結果が変わることを確認
- 局所アンサンブル変換カルマンフィルタや特異値分解による背景誤差共分散行列の設定の実装
- ハイブリッド法の実装

### 参考文献

- [1] Johnson et al., 2005: A singular vector perspective of 4D-Var: Filtering and interpolation. *Quart. J. Roy. Meteor. Soc.*, **131**(605), 1–19.
- [2] Lorenc, A. C., 2003: Modelling of error covariances by 4D-Var data assimilation. *Quart. J. Roy. Meteor. Soc.*, **129**, 3167–3182.
- [3] Lorenz, 1996: Predictability: A problem partly solved. Seminar on Predictability, **1**, ECMWF.
- [4] 三好, 2008: カルマンフィルタ. 気象学におけるデータ同化, 気象研究ノート, **217**, 69–96.
- [5] Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.