統計数理(2022)
第 70 巻 第 2 号 195–208

特集「データ同化の方法」
［研究ノート］

共役ベクトルと BFGS 公式を用いた解析誤差共分散行列の導出

丹羽 洋介[1,2]・藤井 陽介[2,3]

（受付 2022 年 1 月 4 日；改訂 6 月 30 日；採択 6 月 30 日）

要　旨

データ同化や逆解析の問題において，4 次元変分法は有効な手法であるが，最適解の誤差が自動的に得られないという欠点がある．本稿は Niwa and Fujii (2020) で示された，4 次元変分法を用いた場合の解析誤差共分散行列の推定手法について解説を行った．本手法では，最適解の探索アルゴリズムとして用いられる BFGS 公式を用いた準ニュートン法を採用しているが，厳密な直線探索やアンサンブル法，直交化を新たに導入することで，BFGS 公式の計算に必要なベクトルの共役性を保ちつつ数を増やして解析誤差共分散行列の推定精度を向上させている．

本稿では，BFGS 公式において前処理を施した場合に観測と同じ数の反復回数（または BFGS 公式で必要なベクトルのペア数）で解析的に解析誤差共分散行列が求められることを解説するとともに，Niwa and Fujii (2020) のアルゴリズムを詳述する．さらに，本手法を CO_2 の逆解析問題に適用した際の結果についても紹介する．

キーワード：解析誤差共分散行列，データ同化，逆解析，4 次元変分法，BFGS 公式，準ニュートン法．

1. はじめに

4 次元変分法（アジョイント法）は問題規模の大きいデータ同化・逆解析の分野において有用な手法であり，広く用いられている手法である（例えば Sasaki, 1969）．しかしながら，アンサンブル・カルマンフィルターなどの手法とは異なり，最適解の誤差（解析誤差）が自動的には得られないという欠点がある．本論文では 4 次元変分法を用いた場合に，どのように解析誤差を得ることができるか，一つの手法を紹介する．

観測データを $\boldsymbol{y}^{\mathrm{o}}$ とし，あるパラメータ $\boldsymbol{x}$ からこの観測データに対応する値を計算するモデル演算を $M(\boldsymbol{x})$ と定義する．ここで，$\boldsymbol{x}, \boldsymbol{y}^{\mathrm{o}}$ はそれぞれ，解析対象とする時空間内のあらゆる時間・場所の制御変数，観測データを含むものとし，要素数は n, m とする．また，モデル演算子 $M(.)$ は時間発展演算子と観測演算子が含まれる．4 次元変分法では，$\boldsymbol{x}$ の初期推定値 $\boldsymbol{x}^{\mathrm{pri}}$ が与えられたとき，条件付確率密度関数 $p(\boldsymbol{x}|\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}})$ が最大となる解 $\boldsymbol{x}^{\mathrm{pos}}$ を求める（maximum a posteriori estimate: MAP 推定）．

[1] 国立環境研究所 地球システム領域：〒 305–8506 茨城県つくば市小野川 16–2
[2] 気象庁気象研究所 全球大気海洋研究部：〒 305–0052 茨城県つくば市長峰 1–1
[3] 統計数理研究所：〒190–8562 東京都立川市緑町 10–3

$$\boldsymbol{x}^{\mathrm{pos}} = \arg\max_{\boldsymbol{x}\in\mathbb{R}^n} p(\boldsymbol{x}|\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}}) \tag{1.1}$$

ここで，ベイズの定理を用いると，

$$p(\boldsymbol{x}|\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}}) = \frac{p(\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}})} \tag{1.2}$$

であり，さらに，パラメータ $\boldsymbol{x}$ が与えられたときに $\boldsymbol{x}^{\mathrm{pri}}$ と $\boldsymbol{y}^{\mathrm{o}}$ が独立に生起するとすると，

$$p(\boldsymbol{x}|\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}}) = \frac{p(\boldsymbol{x}^{\mathrm{pri}}|\boldsymbol{x})p(\boldsymbol{y}^{\mathrm{o}}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{x}^{\mathrm{pri}}, \boldsymbol{y}^{\mathrm{o}})} \tag{1.3}$$

が得られる．式(1.3)の分母は $\boldsymbol{x}$ に依存しないので式(1.1)は

$$\boldsymbol{x}^{\mathrm{pos}} = \arg\max_{\boldsymbol{x}\in\mathbb{R}^n} [p(\boldsymbol{x}^{\mathrm{pri}}|\boldsymbol{x})p(\boldsymbol{y}^{\mathrm{o}}|\boldsymbol{x})p(\boldsymbol{x})] \tag{1.4}$$

と変形することができる．$p(\boldsymbol{x})$ は初期推定値も観測もないときの $\boldsymbol{x}$ の確率密度関数であり，一般に得ることは容易ではないため，ここでは一様分布と仮定すると，

$$\boldsymbol{x}^{\mathrm{pos}} = \arg\max_{\boldsymbol{x}\in\mathbb{R}^n} [p(\boldsymbol{x}^{\mathrm{pri}}|\boldsymbol{x})p(\boldsymbol{y}^{\mathrm{o}}|\boldsymbol{x})] \tag{1.5}$$

となり，MAP 推定は最尤推定(maximum likelihood estimate)に帰着する（露木・川畑, 2008）．4 次元変分法では式(1.5)をもとにして定義した評価関数の勾配をアジョイントモデルとよばれるモデルを用いて計算し，反復しながらその勾配がゼロとなる最適解を探索する．ここでは簡単のため，$p(\boldsymbol{x}^{\mathrm{pri}}|\boldsymbol{x})$ と $p(\boldsymbol{y}^{\mathrm{o}}|\boldsymbol{x})$ がガウス分布であると仮定すると，評価関数は一般的に対数で定義され，

$$J(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{\mathrm{pri}})^{\mathrm{T}}\boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{x}^{\mathrm{pri}}) + \frac{1}{2}(M(\boldsymbol{x}) - \boldsymbol{y}^{\mathrm{o}})^{\mathrm{T}}\boldsymbol{R}^{-1}(M(\boldsymbol{x}) - \boldsymbol{y}^{\mathrm{o}}) \tag{1.6}$$

と表すことができる．4 次元変分法では，この評価関数 J を最小とする解を求める．なお，$\boldsymbol{B}$ および $\boldsymbol{R}$ はそれぞれ初期推定誤差とモデル-観測間の差の共分散行列を表している．さらに，$M(.)$ が十分に線形であるとみなせる場合には，式(1.6)は接線形モデルを表す行列 $\boldsymbol{M}$ を用いて以下のように書き直すことができる．

$$J(\delta\boldsymbol{x}) = \frac{1}{2}\delta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}^{-1}\delta\boldsymbol{x} + \frac{1}{2}(\boldsymbol{M}\delta\boldsymbol{x} - \boldsymbol{d}^{\mathrm{o}})^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{M}\delta\boldsymbol{x} - \boldsymbol{d}^{\mathrm{o}}) \tag{1.7}$$

なお，$\delta\boldsymbol{x} := \boldsymbol{x} - \boldsymbol{x}^{\mathrm{pri}}$ であり $\boldsymbol{d}^{\mathrm{o}} := \boldsymbol{y}^{\mathrm{o}} - M(\boldsymbol{x}^{\mathrm{pri}})$ である．式(1.7)の評価関数が最小となる解は

$$\delta\boldsymbol{x}^{\mathrm{pos}} = \arg\min_{\delta\boldsymbol{x}\in\mathbb{R}^n} J(\delta\boldsymbol{x}) = (\boldsymbol{B}^{-1} + \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{M})^{-1}\boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{d}^{\mathrm{o}} \tag{1.8}$$

であり，このとき，評価関数の勾配 $\nabla_{\delta\boldsymbol{x}}J$ は

$$\nabla_{\delta\boldsymbol{x}}J = \boldsymbol{B}^{-1}\delta\boldsymbol{x} + \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{M}\delta\boldsymbol{x} - \boldsymbol{d}^{\mathrm{o}}) = \boldsymbol{0} \tag{1.9}$$

を満たす．問題の規模が大きい場合，式(1.8)を解析的に解くことは不可能なため，実際の 4 次元変分法の計算では $\boldsymbol{M}^{\mathrm{T}}$ で表されるアジョイントモデルを用いて評価関数の勾配を計算し，最適解の探索を行う．この時，最適解の誤差共分散行列(解析誤差共分散行列) $\boldsymbol{P}_{\mathrm{a}}$ は必要なく，同時に求められることはないが，実際には $\boldsymbol{P}_{\mathrm{a}}$ は評価関数のヘッセ行列 ($\boldsymbol{A} = \nabla_{\delta\boldsymbol{x}}\nabla_{\delta\boldsymbol{x}}^{\mathrm{T}}J(\delta\boldsymbol{x}) = \boldsymbol{B}^{-1} + \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{M}$) の逆行列と等しく，

$$\boldsymbol{P}_{\mathrm{a}} = \boldsymbol{A}^{-1} = (\boldsymbol{B}^{-1} + \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{M})^{-1} \tag{1.10}$$

と表すことができる．問題規模が大きい場合には式(1.8)と同様に式(1.10)も解析的に解くこと

は不可能なため，$\boldsymbol{P}_{\mathrm{a}}$ を推定する手法が必要となってくる．なお，式(1.7)–(1.10)では，$M(.)$ ではなく線形のモデル $\boldsymbol{M}$ が使われていることに注意が必要である．モデルが線形であるために，評価関数(式(1.7))が 2 次関数であり，極小値が一つしか存在せず，ヘッセ行列が定数行列となっている．したがって，解析誤差共分散行列はモデルが線形である場合にしか定義することができない．本稿では，この解析誤差共分散行列を推定する手法を提案しているが，その適用範囲は，モデルが線形であり，かつ，$p(\boldsymbol{x}^{\mathrm{pri}}|\boldsymbol{x})$ と $p(\boldsymbol{y}|\boldsymbol{x})$ がガウス分布と仮定した線形問題に限っていることに注意されたい．

これまで 4 次元変分法をフレームワークとした $\boldsymbol{P}_{\mathrm{a}}$ の推定には様々な手法が提案されており，ランダム化(randomization)法，モンテ・カルロ法，ランチョス法，Broyden-Fletcher-Goldfarb-Shanno (BFGS)法などがある（Rabier and Courtier, 1992; Fisher and Courtier, 1995; Chevallier et al., 2007）．他にも second-order adjoint 法を用いる手法も存在するが（Ito et al., 2016），制御変数の最適化計算では用いない second-order adjoint モデルを新たに開発する必要があり，モデルが大規模で複雑である場合には開発コストがかかる．ランダム化法やモンテ・カルロ法はその確率論的な性質からサンプル数に精度が大きく依存し，特に共分散($\boldsymbol{P}_{\mathrm{a}}$ の非対角要素)の推定精度に難があることが指摘されている（Fisher and Courtier, 1995）．一方で，ランチョス法および BFGS 法は，それぞれ，4 次元変分法の最適解探索のアルゴリズムとして用いられる共役勾配法や準ニュートン法と密接に関係，またはアルゴリズムの一部として用いられているため，最適解探索のアルゴリズムを応用することで $\boldsymbol{P}_{\mathrm{a}}$ を推定することが可能であるが，その推定精度を上げるためには，さらなる追加の手法が必要となってくる（Bousserez et al., 2015）．

そこで筆者らは，制御変数の最適化に用いる既存の BFGS 法を使用しつつ，実装が簡易なベクトルの共役化やアンサンブルを新たに導入することにより，高精度に $\boldsymbol{P}_{\mathrm{a}}$ を導出する手法を開発した（Niwa and Fujii, 2020）．本論文では，その手法について解説を行う．なお，これからの記述においては，簡単のために $\delta\boldsymbol{x}$ の δ は省略するとともに，慣例的な表記に従って，ベクトル $\boldsymbol{d}$ は制御変数の探索方向，ベクトル $\boldsymbol{y}$ は評価関数の勾配のインクリメントとして表すこととする．

2. BFGS 公式

4 次元変分法で用いられる最適解探索のアルゴリズムの一つとして準ニュートン法がある．この準ニュートン法は，ある k 回目の反復計算における制御変数の更新式

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} + \alpha_k \boldsymbol{d}_{k-1} \tag{2.1}$$

で必要な探索ベクトル $\boldsymbol{d}_{k-1}$ を求める際に，ヘッセ行列の近似逆行列を用いる手法である(なお，α_k はその探索幅で $J(\boldsymbol{x}_{k-1} + \alpha_k \boldsymbol{d}_{k-1})$ が最小となるように決められる)．具体的には，探索ベクトルはヘッセ行列の近似逆行列 $\boldsymbol{H}_k$ と評価関数の勾配ベクトル $\boldsymbol{g}_k$ を使って，

$$\boldsymbol{d}_k = -\boldsymbol{H}_k \boldsymbol{g}_k \tag{2.2}$$

として決定される．この時に必要な $\boldsymbol{H}_k$ の求め方には様々な手法が存在するが，その一つに BFGS 公式があり，以下のような式で書くことができる．

$$\boldsymbol{H}_k = (\boldsymbol{I} - \rho_k \boldsymbol{p}_k \boldsymbol{y}_k^{\mathrm{T}}) \boldsymbol{H}_{k-1} (\boldsymbol{I} - \rho_k \boldsymbol{y}_k \boldsymbol{p}_k^{\mathrm{T}}) + \rho_k \boldsymbol{p}_k \boldsymbol{p}_k^{\mathrm{T}} \tag{2.3}$$

なお，ここで，$\boldsymbol{p}_k := \boldsymbol{x}_k - \boldsymbol{x}_{k-1}$, $\boldsymbol{y}_k := \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$, $\rho_k := 1/(\boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{p}_k)$ であり，$\boldsymbol{I}$ は単位行列を表す．また，最初の反復で必要となる $\boldsymbol{H}_0$ は任意の正定値対称行列を用いることができる．実際の(2.2)と(2.3)の計算では，$\boldsymbol{y}$ および $\boldsymbol{p}$ のベクトルのみを使った再帰的な計算(two-loop recursion scheme)を用いることで，$\boldsymbol{H}_k$ をメモリに保持することなく $\boldsymbol{d}_k$ を計算することができる(例えば Nocedal and Wright, 2006)．

ここで注目すべき点は，$\boldsymbol{H}_k$ がヘッセ行列の近似逆行列であるということから，$\boldsymbol{P}_\mathrm{a}$ の近似行列と見なすこともできるということである．したがって，式(2.2)で $\boldsymbol{g}_k$ を $(1,0,\ldots,0)^\mathrm{T}$, $(0,1,\ldots,0)^\mathrm{T},\ldots,(0,0,\ldots,1)^\mathrm{T}$ といったベクトルに置き換えてやれば，$\boldsymbol{P}_\mathrm{a}$ のそれぞれの列を推定することが可能となる．しかし，その推定精度は $\boldsymbol{H}_k$ の計算に必要なベクトル $\boldsymbol{p}$, $\boldsymbol{y}$ や初期推定行列 $\boldsymbol{H}_0$ に依存する．Niwa and Fujii (2020)では，アンサンブル手法を使ってベクトル $(\boldsymbol{p},\boldsymbol{y})$ のペア数を増やしつつ，$(\boldsymbol{p},\boldsymbol{y})$ が本来満たすべき共役性を達成することで $\boldsymbol{P}_\mathrm{a}$ の推定精度が大幅に向上することを実証した．以下の章ではその手法について解説していく．

3. 共役性

問題が線形(評価関数が下に凸な 2 次関数)であるとき，共役性が満たされているとは，以下の式が成り立っていることを意味する．

$$\boldsymbol{p}_i^\mathrm{T}\boldsymbol{A}\boldsymbol{p}_j = 0, \quad \text{for all } i \neq j \tag{3.1}$$

または，線形の場合には $\boldsymbol{A}\boldsymbol{p}_j = \boldsymbol{y}_j$ であることから

$$\boldsymbol{p}_i^\mathrm{T}\boldsymbol{y}_j = 0, \quad \text{for all } i \neq j \tag{3.2}$$

BFGS 公式を用いた反復計算において，この共役性が満たされていれば，制御変数の数 $(=n)$ の回数の反復で解が収束して解析解と等しくなり，また，そこでの $\boldsymbol{H}_k$ もヘッセ行列の逆行列と等しくなる($\boldsymbol{H}_n = \boldsymbol{A}^{-1}$)ことが知られている (Nocedal and Wright, 2006)．逆にいえば，共役性が満たされていなければ，n 以上の回数が解の収束に必要となる．

実際には前処理を行うことで，この収束速度は速めることができる．最適化計算において前処理は一般的に使われるものであるが，最もよく使われるものに Lorenc (1988)の手法がある．この Lorenc (1988)の前処理では，制御変数が

$$\tilde{\boldsymbol{x}} = \boldsymbol{B}^{-1/2}\boldsymbol{x} \tag{3.3}$$

と変換される．この時，変換されたヘッセ行列は

$$\begin{aligned}\tilde{\boldsymbol{A}} &= \boldsymbol{I} + \boldsymbol{B}^{1/2}\boldsymbol{M}^\mathrm{T}\boldsymbol{R}^{-1}\boldsymbol{M}\boldsymbol{B}^{1/2} = \boldsymbol{I} + \boldsymbol{X}\boldsymbol{X}^\mathrm{T}\\ \boldsymbol{X} &:= \boldsymbol{B}^{1/2}\boldsymbol{M}^\mathrm{T}\boldsymbol{R}^{-1/2}\end{aligned} \tag{3.4}$$

と書くことができる．通常のデータ同化または逆解析の問題では，$n > m$ であるため，このヘッセ行列は 1 以外の固有値はたかだか m 個しか存在せず，収束性を速めることができる．なおこの前処理をした後の BFGS 公式において，そのヘッセ近似逆行列の初期値 $\tilde{\boldsymbol{H}}_0$ を最もシンプルな正定値対称行列である $\boldsymbol{I}$ とした場合($\tilde{\boldsymbol{H}}_0 = \boldsymbol{I}$)，前処理前の BFGS 公式において $\boldsymbol{H}_0 = \boldsymbol{B}$ とした場合と同等となる．したがって，通常の BFGS 公式において $\boldsymbol{H}_0 = \boldsymbol{B}$ とすれば，陰に式(3.3)の前処理をしたこととなり，収束を速めることができる．このような前処理をした BFGS 公式においては，観測数 m の数だけの共役性を満たす $(\boldsymbol{y},\boldsymbol{p})$ のベクトルのペアさえあれば，それらによって得られるヘッセ近似逆行列 $\boldsymbol{H}_m$ が解析的に得られる行列，すなわち $\boldsymbol{P}_\mathrm{a}$ と等しくなる．

このことは以下のように証明することができる．まず，共役性が満たされていれば，変換後も

$$\tilde{\boldsymbol{p}}_i^\mathrm{T}\tilde{\boldsymbol{y}}_j = \tilde{\boldsymbol{p}}_i^\mathrm{T}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{p}}_j = 0, \quad \text{for all } i \neq j \tag{3.5}$$

であるため，

$$\begin{aligned}\tilde{\boldsymbol{H}}_k\tilde{\boldsymbol{y}}_j &= [(\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{y}}_k^{\mathrm{T}})\tilde{\boldsymbol{H}}_{k-1}(\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{y}}_k\tilde{\boldsymbol{p}}_k^{\mathrm{T}})+\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{p}}_k^{\mathrm{T}}]\tilde{\boldsymbol{y}}_j \\ &= (\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{y}}_k^{\mathrm{T}})\tilde{\boldsymbol{H}}_{k-1}\tilde{\boldsymbol{y}}_j \\ &= (\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{y}}_k^{\mathrm{T}})(\boldsymbol{I}-\tilde{\rho}_{k-1}\tilde{\boldsymbol{p}}_{k-1}\tilde{\boldsymbol{y}}_{k-1}^{\mathrm{T}})\cdots(\boldsymbol{I}-\tilde{\rho}_{j+1}\tilde{\boldsymbol{p}}_{j+1}\tilde{\boldsymbol{y}}_{j+1}^{\mathrm{T}})\tilde{\boldsymbol{H}}_j\tilde{\boldsymbol{y}}_j, \quad \text{for } j\le k\end{aligned} \tag{3.6}$$

が導かれる．ここで，$\tilde{\boldsymbol{y}}_k := \boldsymbol{B}^{1/2}\boldsymbol{y}_k$, $\tilde{\boldsymbol{p}}_k := \boldsymbol{B}^{-1/2}\boldsymbol{p}_k$, $\tilde{\rho}_k := 1/(\tilde{\boldsymbol{y}}_k^{\mathrm{T}}\tilde{\boldsymbol{p}}_k)$ である．また，

$$\tilde{\boldsymbol{H}}_j\tilde{\boldsymbol{y}}_j = [(\boldsymbol{I}-\tilde{\rho}_j\tilde{\boldsymbol{p}}_j\tilde{\boldsymbol{y}}_j^{\mathrm{T}})\tilde{\boldsymbol{H}}_{j-1}(\boldsymbol{I}-\tilde{\rho}_j\tilde{\boldsymbol{y}}_j\tilde{\boldsymbol{p}}_j^{\mathrm{T}})+\tilde{\rho}_j\tilde{\boldsymbol{p}}_j\tilde{\boldsymbol{p}}_j^{\mathrm{T}}]\tilde{\boldsymbol{y}}_j = \tilde{\boldsymbol{p}}_j, \quad \text{for } j\le k \tag{3.7}$$

により，式(3.6)は

$$\tilde{\boldsymbol{H}}_k\tilde{\boldsymbol{y}}_j = \tilde{\boldsymbol{p}}_j, \quad \text{for } j\le k \tag{3.8}$$

となり，シンプルな関係性が得られる（Dennis, Jr. and Moré, 1977; Fisher and Courtier, 1995）．さらに $\boldsymbol{X}$ の固有ベクトル $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_m\}$ のいずれとも直交する任意のベクトルを $\bar{\boldsymbol{u}}$ とすると，

$$\begin{aligned}\tilde{\boldsymbol{H}}_k\bar{\boldsymbol{u}} &= [(\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{y}}_k^{\mathrm{T}})\tilde{\boldsymbol{H}}_{k-1}(\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{y}}_k\tilde{\boldsymbol{p}}_k^{\mathrm{T}})+\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{p}}_k^{\mathrm{T}}]\bar{\boldsymbol{u}} \\ &= (\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{y}}_k^{\mathrm{T}})\tilde{\boldsymbol{H}}_{k-1}\bar{\boldsymbol{u}} \\ &= (\boldsymbol{I}-\tilde{\rho}_k\tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{y}}_k^{\mathrm{T}})(\boldsymbol{I}-\tilde{\rho}_{k-1}\tilde{\boldsymbol{p}}_{k-1}\tilde{\boldsymbol{y}}_{k-1}^{\mathrm{T}})\cdots(\boldsymbol{I}-\tilde{\rho}_1\tilde{\boldsymbol{p}}_1\tilde{\boldsymbol{y}}_1^{\mathrm{T}})\tilde{\boldsymbol{H}}_0\bar{\boldsymbol{u}} \\ &= \bar{\boldsymbol{u}}\end{aligned} \tag{3.9}$$

という関係が得られる．なお，ここで

$$\tilde{\boldsymbol{y}}_k \in \mathrm{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_m\}, \quad \tilde{\boldsymbol{p}}_k \in \mathrm{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_m\}, \quad \text{for any } k\le m \tag{3.10}$$

を満たしており，

$$\mathrm{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_m\} = \mathrm{span}\{\tilde{\boldsymbol{p}}_1,\tilde{\boldsymbol{p}}_2,\ldots,\tilde{\boldsymbol{p}}_m\} \tag{3.11}$$

であることから（これらの詳しい証明は Niwa and Fujii, 2020 を参照のこと），$1\le k\le m$ について $\tilde{\boldsymbol{p}}_k^{\mathrm{T}}\bar{\boldsymbol{u}}=0$ かつ $\tilde{\boldsymbol{y}}_k^{\mathrm{T}}\bar{\boldsymbol{u}}=0$ であること，さらに $\tilde{\boldsymbol{H}}_0=\boldsymbol{I}$ であることを用いた．

ここで，制御変数の空間の任意のベクトル $\boldsymbol{b}$ は式(3.11)により

$$\begin{aligned}\boldsymbol{b} &= \sum_{i=1}^{m} a_i\tilde{\boldsymbol{p}}_i + \bar{\boldsymbol{u}} = \tilde{\boldsymbol{P}}\boldsymbol{a}+\bar{\boldsymbol{u}} \\ \tilde{\boldsymbol{P}} &:= (\tilde{\boldsymbol{p}}_1,\tilde{\boldsymbol{p}}_2,\ldots,\tilde{\boldsymbol{p}}_m), \boldsymbol{a} := (a_1,a_2,\ldots,a_m)^{\mathrm{T}}\end{aligned} \tag{3.12}$$

と表すことができ，その $\boldsymbol{b}$ は以下の式を満たす．

$$\tilde{\boldsymbol{H}}_m\tilde{\boldsymbol{A}}\boldsymbol{b} = \tilde{\boldsymbol{H}}_m(\tilde{\boldsymbol{Y}}\boldsymbol{a}+\bar{\boldsymbol{u}}) = \tilde{\boldsymbol{P}}\boldsymbol{a}+\bar{\boldsymbol{u}} = \boldsymbol{b} \tag{3.13}$$

ここで $\tilde{\boldsymbol{Y}} := (\tilde{\boldsymbol{y}}_1,\tilde{\boldsymbol{y}}_2,\ldots,\tilde{\boldsymbol{y}}_m)$ であり，式(3.4)，(3.8)，(3.9)，また，$\tilde{\boldsymbol{y}}=\tilde{\boldsymbol{A}}\tilde{\boldsymbol{p}}$, $\boldsymbol{X}^{\mathrm{T}}\bar{\boldsymbol{u}}=\boldsymbol{0}$ であることを用いた．さらに式(3.13)は

$$(\tilde{\boldsymbol{H}}_m\tilde{\boldsymbol{A}}-\boldsymbol{I})\boldsymbol{b}=\boldsymbol{0} \tag{3.14}$$

と書き換えることができ，この式が任意のベクトル $\boldsymbol{b}$ に対して成り立つことから，$\tilde{\boldsymbol{H}}_m=\tilde{\boldsymbol{A}}^{-1}$ であり，前処理前の行列においても

$$\boldsymbol{H}_m = \boldsymbol{A}^{-1} = \boldsymbol{P}_{\mathrm{a}} \tag{3.15}$$

が成り立つ．以上のことから，共役性が保たれていれば，m 個の $(\boldsymbol{y}, \boldsymbol{p})$ のペアによって $\boldsymbol{P}_{\mathrm{a}}$ が解析的に導出できることが示されたが，m 個だけあれば良いとはいえ，観測データの数は通常，可能な反復回数よりも大きい．したがって，$(\boldsymbol{y}, \boldsymbol{p})$ のベクトルのペアを如何に増やすか，また，それらの $(\boldsymbol{y}, \boldsymbol{p})$ のベクトルについて，どうやって共役性を保たせることができるか，これらの問題に対応した手法を次節で紹介する．

4. 共役性のある $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルの導出

上記では解析誤差共分散行列 $\boldsymbol{P}_{\mathrm{a}}$ の推定のために BFGS 公式を用いることを提案しているが，そこで必要な $(\boldsymbol{y}, \boldsymbol{p})$ の作成には必ずしも BFGS 公式を用いた準ニュートン法を用いる必然性はない．しかし，Niwa and Fujii (2020) では BFGS 公式を用いた準ニュートン法をベースとしたアルゴリズム POpULar (Preconditioned Optimizing Utility for Large-dimensional analyse) (Fujii and Kamachi, 2003; Fujii, 2005) を用いて $(\boldsymbol{y}, \boldsymbol{p})$ のペアを作成しており，ここで，新たに厳密な直線探索 (exact line search) を導入して $(\boldsymbol{y}, \boldsymbol{p})$ の共役性が保たれるようにしている．さらに，$(\boldsymbol{y}, \boldsymbol{p})$ のペア数を増やすためにアンサンブルを導入し，アンサンブルメンバー間の $(\boldsymbol{y}, \boldsymbol{p})$ についても共役性が保たれるように直交化を行っている．

4.1 厳密な直線探索

問題が線形であり，評価関数が下に凸な二次関数である場合，式 (2.1) において $J(\boldsymbol{x}_{k-1} + \alpha_k \boldsymbol{d}_{k-1})$ が最小となる α_k は解析的に

$$\alpha_k = -\frac{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{g}_{k-1}}{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{d}_{k-1}} \tag{4.1}$$

と表すことができ，この α_k を用いた直線探索を厳密な直線探索 (exact line search) と呼ぶ．問題が線形な場合，この厳密な直線探索が BFGS 公式を用いた準ニュートン法で行われれば，得られる $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルが共役性を持つことが知られている (Dennis, Jr. and Moré, 1977)．しかしながら，式 (4.1) は $\boldsymbol{A}$ が基本的に未知であり，また，既知であったとしても，その行列のサイズの大きさ $(n \times n)$ から容易に計算することはできない．しかし，新たな変数

$$\boldsymbol{f}_k = \boldsymbol{e}_{k-1} + \boldsymbol{M}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{M} \boldsymbol{d}_{k-1} \tag{4.2}$$

を導入することにより，式 (4.1) は

$$\alpha_k = -\frac{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{g}_{k-1}}{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{f}_k} \tag{4.3}$$

となり，ベクトルのみで計算することが可能となる．なお，ここで，$\boldsymbol{e} := \boldsymbol{B}^{-1} \boldsymbol{d}$ である．式 (4.2) 式に $\boldsymbol{M}$ と $\boldsymbol{M}^{\mathrm{T}}$ があるため，$\boldsymbol{f}$ の計算には接線形モデルとアジョイントモデルによる計算が必要となっているが，一方で評価関数の勾配 $\boldsymbol{g}$ は

$$\boldsymbol{g}_k = \boldsymbol{g}_{k-1} + \alpha_k \boldsymbol{f}_k \tag{4.4}$$

という式で更新することができるため，通常の最適化計算において必要であった $\boldsymbol{g}$ の導出における $\boldsymbol{M}$ および $\boldsymbol{M}^{\mathrm{T}}$ の計算は必要ではなくなり，全体的な計算量は厳密な直線探索を導入しても変わらないようになっている．

実はこの $\boldsymbol{f}$ を用いた厳密な直線探索は Derber and Rosati (1989) で提案されている．ただし，その最適化手法は準ニュートン法ではなく共役勾配法が用いられている．この共役勾配法では，先ほど述べた Lorenc (1988) の前処理を処しているものの，漸化式を用いることによって，

前処理に必要な $\boldsymbol{B}^{1/2}$ や $\boldsymbol{B}^{-1}$ の計算は必要ないというのがもう一つの特徴である．POpULar は，非線形の場合にも対応できるように，この Derber and Rosati (1989)の手法を BFGS 公式を用いた準ニュートン法に拡張したもので，直線探索も厳密なものではないものに置き換えられている．しかし，Niwa and Fujii (2020)はこの POpULar をベースとしつつも，直線探索を上記の厳密なものに戻して共役性のある $(\boldsymbol{y},\boldsymbol{p})$ ベクトルのペアを作成するようにした(Linear POpULar)．そのアルゴリズムは以下のようになる．

Algorithm 1 Linear POpULar

$\boldsymbol{x}_0 \leftarrow \boldsymbol{0}$;
$\boldsymbol{g}_0 \leftarrow -\boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{d}^0$;
$\boldsymbol{d}_0 \leftarrow -\boldsymbol{B}\boldsymbol{g}_0$;
$\boldsymbol{e}_0 \leftarrow -\boldsymbol{g}_0$;
$k \leftarrow 1$;
repeat
  $\boldsymbol{f}_k = \boldsymbol{e}_{k-1} + \boldsymbol{M}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{M}\boldsymbol{d}_{k-1}$;
  $\boldsymbol{x}_k \leftarrow \boldsymbol{x}_{k-1} + \alpha_k \boldsymbol{d}_{k-1}$,
  ここで，α_k は厳密な直線探索(式 4.3)で得られる;
  $\boldsymbol{g}_k = \boldsymbol{g}_{k-1} + \alpha_k \boldsymbol{f}_k$;
  $\boldsymbol{d}_k \leftarrow -\boldsymbol{H}_k \boldsymbol{g}_k$,
  $\boldsymbol{e}_k \leftarrow \boldsymbol{B}^{-1}\boldsymbol{d}_k$,
  ここで，two-loop recursion scheme を用いた BFGS 公式の計算により $\boldsymbol{d}_k$ のみならず $\boldsymbol{e}_k$ の計算も行う．その際，$\boldsymbol{B}^{-1}$ の計算を行わずとも $\boldsymbol{e}_k$ を計算することができる;
  $k \leftarrow k+1$;
until convergence

なお，BFGS 公式において，最初のヘッセ行列の逆行列の推定($\boldsymbol{H}_0$)に $\boldsymbol{B}$ を用いることで Lorenc (1988)の前処理を導入しているが，上記のアルゴリズムで見て分かる通り，$\boldsymbol{B}^{1/2}$ や $\boldsymbol{B}^{-1}$ の計算は必要ない．

問題が線形で厳密な直線探索が使われている場合，共役勾配法と BFGS 公式を用いた準ニュートン法は同一であることが知られているが(例えば Nazareth, 1979)，共役勾配法は丸め誤差の影響を受けやすいため (Golub and van Loan, 1996)，安定して共役性の保たれた $(\boldsymbol{y},\boldsymbol{p})$ のペアを得るためには，準ニュートン法を用いた方が適している．また，ここで注意しておきたいのが，オリジナルの POpULar の BFGS 公式で使われており，また，一般的にもよく使われているメモリ節約は Niwa and Fujii (2020)では用いていないという点である．メモリ節約とは式(2.3)の $\boldsymbol{H}$ の更新式において，全ての $(\boldsymbol{y},\boldsymbol{p})$ ベクトルを用いるのではなく，直近のあらかじめ定めた数のペアのみを用いることであるが，ここでの目的は，共役性の保たれた $(\boldsymbol{y},\boldsymbol{p})$ ベクトルのペアをできるだけ多く作るということであるため，わざわざベクトルの数を減らすメモリ節約は行わない．なお，近年の計算機性能の向上により，全ての $(\boldsymbol{y},\boldsymbol{p})$ のデータを保持することは容易であり(必ずしもメモリに載せる必要はなく，ディスクに格納できれば良い)，全ての $(\boldsymbol{y},\boldsymbol{p})$ のデータを保持することに大きな障害は生じない．

最終的な反復回数を K，モデル計算にかかるコストを C とすると，Algorithm 1 全体の計算量は $O(nK^2)+O(KC)$ と表すことができる．第一項は反復回数の二乗に比例しているが，大気・海洋などの問題でモデル計算のコストが圧倒的に大きい場合には，第二項が卓越し，反復回数に比例して計算量が増加する．また，理論的に必要とされる $(\boldsymbol{y},\boldsymbol{p})$ の数(= 反復回数)はたかだか m であるが，$n \gg m$ であったとしても，相当の反復回数が必要となる場合がある．特

にモデル計算が大規模な場合には，アルゴリズムが逐次的であるために，実行可能な反復回数に限界が生じる(計算時間(wall-clock time)がかかりすぎる)．そこで次節では，実行可能な反復回数で Algorithm 1 の計算を打ち切る一方，同様の計算を複数，同時に行うことで，効率よく $(\boldsymbol{y}, \boldsymbol{p})$ の数を増やすことを考える．

4.2 アンサンブルと直交化

BFGS 公式の式(2.3)を見て分かる通り，$\boldsymbol{H}$ は $(\boldsymbol{y}, \boldsymbol{p})$ の順序には関係ない．したがって，$\boldsymbol{Ap} = \boldsymbol{y}$ である $(\boldsymbol{y}, \boldsymbol{p})$ のベクトルの数を何らかの方法で増やせばよく，そのため，Niwa and Fujii (2020)では，準ニュートン法による最適化計算を同時に複数行うアンサンブル手法を用いた．先行研究である Bousserez et al. (2015)でも同様に，BFGS 公式とアンサンブルを組み合わせた手法を提案しているが，Niwa and Fujii (2020)では，さらに，得られた $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルに直交化を施すプロセスを入れている．これは，それぞれのアンサンブルメンバーは独立に最適化計算を行うことから，メンバーをまたいでは $(\boldsymbol{y}, \boldsymbol{p})$ の共役性が保たれないためで，直交化は，これらのベクトルを変換して共役性を満たすようにする．この直交化は以下のようにして行う．

まず得られた全ての $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルをそれぞれ並べて以下のような行列を作る．

$$\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \ldots, \boldsymbol{p}_N) \tag{4.5}$$

$$\boldsymbol{Y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \ldots, \boldsymbol{y}_N) \tag{4.6}$$

ここで N は全てのアンサンブルメンバーから得られた $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルの数である．次に，$\boldsymbol{P}$ と $\boldsymbol{Y}$ をかけて行列 $\boldsymbol{Q}$ を作成する．

$$\boldsymbol{Q} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{Y} \tag{4.7}$$

$\boldsymbol{p}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{p}^{\mathrm{T}}\boldsymbol{Ap}$ であり $\boldsymbol{A}$ がそもそも対称行列であることから，$\boldsymbol{Q}$ も対称行列であり，次のように固有値分解することができる．

$$\boldsymbol{Q} = \boldsymbol{ULU}^{\mathrm{T}} \tag{4.8}$$

ここで $\boldsymbol{U}$ は固有ベクトルからなる直交行列，$\boldsymbol{L}$ は固有値を対角成分にもつ対角行列である．なお，$N \leq m$ であり，$\boldsymbol{Q}$ は $N \times N$ の次元である．m が十分に小さい場合には(実際に大気・海洋などのデータ同化・逆問題などでは，$n \gg m$ であることが多い)，式(4.8)の固有値分解は容易に実行可能である．この固有値分解において，理論的には $\boldsymbol{Q}$ は正定値行列であるために，得られる固有値も全て正となるべきだが，実際の計算では計算誤差により非常に小さい，または負の固有値が計算されてしまう．そこで Niwa and Fujii (2020)では，計算を安定化させるため，最大の固有値と比して 1.0×10^{-6} より小さい固有値を除外している．その除外された後の固有ベクトル，固有値からなる行列をそれぞれ，$\boldsymbol{U}'$, $\boldsymbol{L}'$ とし，それらを使って $\boldsymbol{P}$ と $\boldsymbol{Y}$ を以下のように変換する．

$$\boldsymbol{P}' = (\boldsymbol{p}'_1, \boldsymbol{p}'_2, \boldsymbol{p}'_3, \ldots, \boldsymbol{p}'_{N'}) = \boldsymbol{PU}'\boldsymbol{L}'^{-1/2} \tag{4.9}$$

$$\boldsymbol{Y}' = (\boldsymbol{y}'_1, \boldsymbol{y}'_2, \boldsymbol{y}'_3, \ldots, \boldsymbol{y}'_{N'}) = \boldsymbol{YU}'\boldsymbol{L}'^{-1/2} \tag{4.10}$$

ここで，N' は除外後の有効な固有値の数であり，$\boldsymbol{U}'$ は $N \times N'$, $\boldsymbol{L}'$ は $N' \times N'$ の次元となっている．なお，変換後でも

$$\boldsymbol{P}'^{\mathrm{T}}\boldsymbol{Y}' = \boldsymbol{I} \tag{4.11}$$

が成り立ち，共役性

$$\boldsymbol{p}_i'^{\mathrm{T}}\boldsymbol{y}_j' = \boldsymbol{p}_i'^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_j' = 0, \quad \text{for all } i \neq j \tag{4.12}$$

が満たされていることに注目されたい．この変換後の $(\boldsymbol{y}', \boldsymbol{p}')$ を BFGS 公式の計算に用いるが，単一の反復計算で得られる数よりもはるかに多くの数で，かつ，共役性が保たれているため，$\boldsymbol{P}_{\mathrm{a}}$ を高精度に推定することが可能となる．

なお，アンサンブルで得られた複数の解析値のばらつきを解析誤差とするのではなく，単純に $(\boldsymbol{y}, \boldsymbol{p})$ のベクトルのペア数を増やせれば良いので，アンサンブルメンバーの作成はシンプルにできる．式(1.10)で表せられるように，$\boldsymbol{P}_{\mathrm{a}}$ は誤差共分散行列 $\boldsymbol{B}$ と $\boldsymbol{R}$，またモデルの計算 $\boldsymbol{M}$ にしかよらないため，これらの行列さえ共通であれば，特段，摂動の与え方に制約はない．そこで，Niwa and Fujii (2020)では，単純にランダムに生成した $\boldsymbol{d}^{\mathrm{o}}(= \boldsymbol{y}^{\mathrm{o}} - M(\boldsymbol{x}^{\mathrm{pri}}))$ を入力として用いている．

Niwa and Fujii (2020)の手法では，4.1 節の厳密な直線探索を使った準ニュートン法の計算をアンサンブルで行い，そこで得られた $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルを直交化して BFGS 公式を計算している．最終的に直交化するため，理論的には各アンサンブルメンバーの計算で厳密な直線探索を用いる必然性はないが($\boldsymbol{A}\boldsymbol{p} = \boldsymbol{y}$ さえ満たされていれば良い)，あらかじめ，厳密な直線探索で各メンバー内の $(\boldsymbol{y}, \boldsymbol{p})$ の共役性が保たれていた方が安定的に $\boldsymbol{P}_{\mathrm{a}}$ を推定できることを実際の実験で経験している．また，直線探索を厳密にすることで発生する追加の計算コストはないため，安全のためにも直線探索を厳密に行っておく方が良い．なお，以上のアルゴリズムでは全て問題が線形であることを前提としているので，モデル計算も線形であることが必須条件である．線形の現象を扱っているモデルであっても，精度向上のために，しばしば非線形の処理がなされる場合があるが，そういった非線形の部分はあらかじめ外しておく必要がある．

5. $\mathbf{CO_2}$ 逆解析問題への適用

上記で紹介したアルゴリズムの有用性を確認するため，Niwa and Fujii (2020)では，大気二酸化炭素(CO_2)濃度の観測データから地表面フラックスを推定するという CO_2 逆解析問題において，上記アルゴリズムを使って解析誤差共分散行列 $\boldsymbol{P}_{\mathrm{a}}$ の導出を試みた．用いた数値モデルや実験設定の詳細は Niwa and Fujii (2020)を参照されたい．

図 1 は BFGS 公式で推定した $\boldsymbol{P}_{\mathrm{a}}$ が反復回数を重ねるごとに解析解に近づいていく様子を示している．図 1(a)，(b)では一つの BFGS 公式による計算結果を示し，図 1(c)，(d)は 50 メンバーのアンサンブルを使った場合を示す．また，ここで解析解との差をノルムと Ueno and Nakamura (2016)で定義された divergence の 2 通りで示すことで，メトリックによる違いも確認している．図 1 でわかることは，反復を重ねるとともにほぼ単調に推定誤差が減少すること，また，アンサンブルを用いることで，誤差の減少速度を大幅に短縮していることがわかる(上段と下段で横軸のスケールが異なることに注意)．このことから，BFGS 公式で用いる $(\boldsymbol{y}, \boldsymbol{p})$ の数を増やすことで，$\boldsymbol{P}_{\mathrm{a}}$ の推定精度を高められていることがわかる．さらに，ここでは，共役性のある $(\boldsymbol{y}, \boldsymbol{p})$ を用いた場合に加え(実線)，共役性のない場合(点線)も示している．共役性のない場合では，通常，よく用いられる厳密ではない直線探索を使っており，また，直交化も行なっていないが，精度を高めるために BFGS 公式の計算を 10 回繰り返している(以後同じ．詳細は Niwa and Fujii, 2020 を参照のこと)．この両者の違いから，共役性を保つことで推定精度が高められていることがわかる．

図 2 は $\boldsymbol{P}_{\mathrm{a}}$ の非対角成分から計算した，東アジアのある場所のフラックスに対する誤差相関の分布を示す．ここでは，50 メンバーを用いて 15 回の反復を行った結果について，共役性が

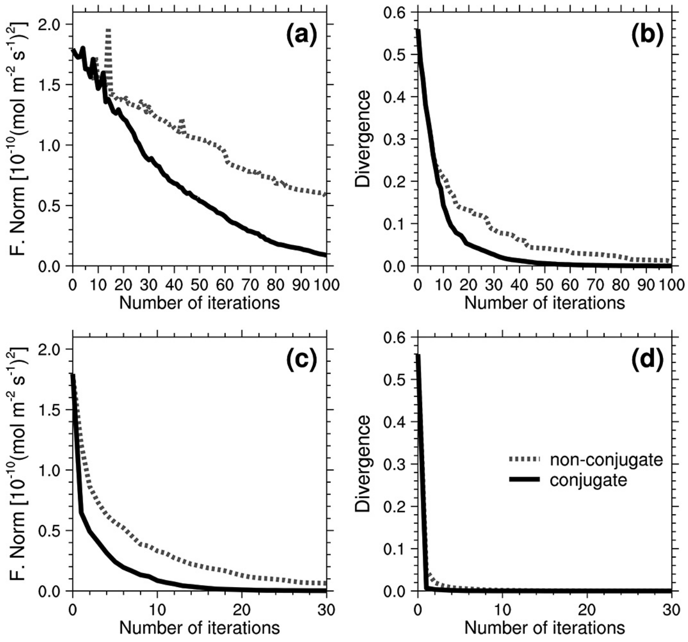

図 1．CO_2 逆解析問題において，BFGS 公式で導出した解析誤差共分散行列 $\boldsymbol{P}_{\mathrm{a}}$ とその解析解との差のノルム（a，c），また，Ueno and Nakamura（2016）で定義された divergence（b，d）の反復回数との関係．上段は一つの BFGS 公式の計算によるもの，下段は 50 メンバーを用いた場合を示す．また，実線は $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルに共役性がある場合，点線はない場合を示す．Niwa and Fujii（2020）の Fig. 6 および Fig. 9 より改変．

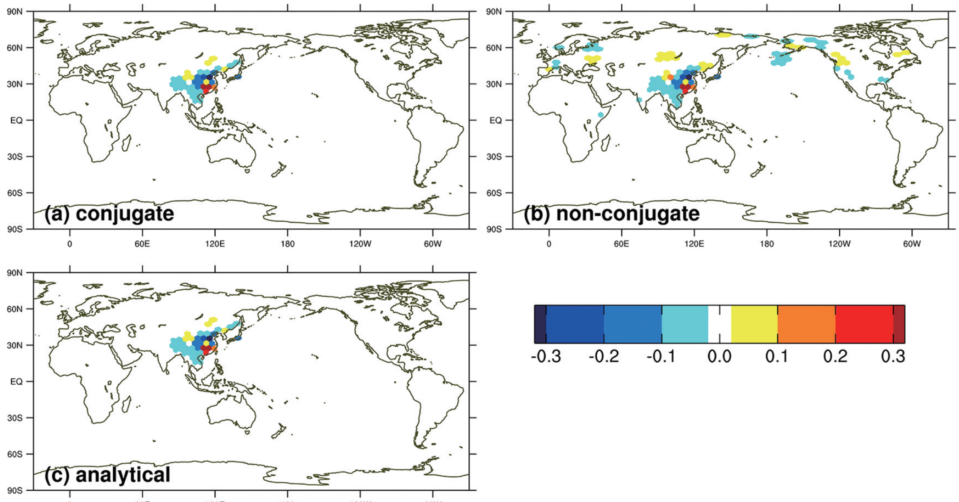

図 2．$\boldsymbol{P}_{\mathrm{a}}$ から求めた東アジアのある場所のフラックスに対する誤差相関の分布．共役性のある $(\boldsymbol{y}, \boldsymbol{p})$（a）とない $(\boldsymbol{y}, \boldsymbol{p})$（b）から推定したものと解析的に求めたもの（c）．Niwa and Fujii（2020）の Fig. 8 より改変．

保たれた場合と保たれていない場合の 2 通りを示している．一般的には対角成分よりも非対角成分の方が導出が困難であるが，共役性が保たれている場合には，解析解の分布をより正確に再現していることが見て取れ，本手法の精度の高さがうかがえる．なお，この時，直交化を経て実際の BFGS 公式に用いた有効な $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルのペア数は 125 個であり，一方で，用いた観測の数は 260 個である．3 節で，前処理をした BFGS 公式では，共役性のある $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルを観測の数だけ用意すれば，解析的に得たものと同等の $\boldsymbol{P}_{\mathrm{a}}$ が得られることを示したが，この実験においては，およそ半分の数でほぼ十分な精度で推定できていることがわかる．このことは，多くの観測が互いに相関していることに起因していると考えられる．

なお，実際の解析では真の解析誤差共分散行列を知り得ないが，図 1 で示すように，2 つのメトリックは単調に減少しているため，その減少幅が十分に小さくなったところでの推定値が解析解とほぼ同等のものであると期待できる．なお，図 2 のケースでは，解析解と同等の分布を得るには 50 アンサンブルのケースで 15 回の反復回数が必要であったことから，より収束が遅く見えるノルムで判断した方が良いことがわかる．

6. まとめと今後の課題

本稿では，Niwa and Fujii (2020) による解析誤差共分散行列 $\boldsymbol{P}_{\mathrm{a}}$ の新たな推定手法について解説を行った．この手法は，データ同化または逆解析において 4 次元変分法を用いた際に，その最適解の誤差情報を正確に得ることを目的としている．Niwa and Fujii (2020) の手法では，まず，収束速度を速めるために，ヘッセ行列の近似逆行列の初期値に $\boldsymbol{B}$ を用いる．さらに，BFGS 公式の計算で必要な $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルの共役性を保つために厳密な直線探査を用いる．また，計算時間を短縮するためにアンサンブルで $(\boldsymbol{y}, \boldsymbol{p})$ の数を増やす場合には，得られた全ての $(\boldsymbol{y}, \boldsymbol{p})$ が共役性を保つように直交化を行う．Niwa and Fujii (2020) では，以上の手法を $\mathrm{CO_2}$ 逆解析の問題に適用して，実際に高い精度で $\boldsymbol{P}_{\mathrm{a}}$ を推定できることを実証した．

また，Niwa and Fujii (2020) では，前処理を施した場合には，観測と同じ数の共役性のある $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルを用意すれば解析値と同等の $\boldsymbol{P}_{\mathrm{a}}$ が得られることを示したが，実際の実験ではその観測数よりもはるかに少ない数の $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルで，解析的に得られるものとほぼ同等の $\boldsymbol{P}_{\mathrm{a}}$ が得られていた．しかし，それでも必要とされる $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルの数は大きく，問題サイズが大きくなればなるほど，その数も大きくなるため，アルゴリズムのさらなる効率化が望まれる（本稿で示している実験では，比較のための解析解が容易に得られるように，実際の問題よりもサイズを落としている）．特に，直交化後の有効な $(\boldsymbol{y}, \boldsymbol{p})$ ベクトルの数 N' の N に対する割合は，N の数が大きければ大きいほど小さくなってしまうことが，一つの問題としてある．図 3 は，反復回数に対する N'/N の推移を示しているが，回数を重ねるごとに有効なベクトルの取得率が低下していることがわかる．また，アンサンブル数が大きい方が，取得率が低いことも図 3 で示されており，この取得率の面にだけ注目すれば，単一の反復計算によるものが効率が良い．したがって，本手法を実際の問題に適用する際には，利用可能なトータルの計算資源と計算時間との兼ね合いでアンサンブル数と個々の反復数を決める必要がある．この N'/N の減少について，一つの原因として考えられるのは，$\boldsymbol{y}$ や $\boldsymbol{p}$ の有効桁が反復計算を重ねるにつれて減少することである（アンサンブル数を増やすと N'/N がより減少するのは，直交化で $\boldsymbol{y}$ と $\boldsymbol{p}$ をかけている（式(4.7)）ためかもしれない）が，現在のところ確実な原因の解明には至っていない．今後は，その原因を突き止め，N'/N を大きくする手法を開発することが，さらなる効率化の 1 つのアプローチであると考えられる．

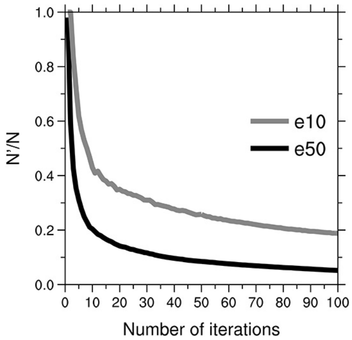

図 3. 反復回数を増やしていった場合の N'/N の推移. N はアンサンブルで得られた $(\boldsymbol{y}, \boldsymbol{p})$ の合計の数. N' は直交化後の有効な $(\boldsymbol{y}, \boldsymbol{p})$ の数. アンサンブルメンバー数が 10(灰色)と 50(黒)の場合を示す. Niwa and Fujii (2020)の Fig. 10 より改変.

謝 辞

査読者から有益なコメントを頂きました. 感謝申し上げます. 本研究は統計数理研究所共同研究プログラム(28-共研-2004, 29-共研-2007, 30-共研-2007, 2019-ISMCRP-2030, 2020-ISMCRP-2049, 2021-ISMCRP-2023)の助成を受けたものです.

参 考 文 献

Bousserez, N., Henze, D. K., Perkins, A., Bowman, K. W., Lee, M., Liu, J., Deng, F. and Jones, D. B. A. (2015). Improved analysis-error covariance matrix for high-dimensional variational inversions: Application to source estimation using a 3D atmospheric transport model, *Quarterly Journal of the Royal Meteorological Society*, **141**(690), 1906–1921, http://dx.doi.org/10.1002/qj.2495.

Chevallier, F., Bréon, F.-M. and Rayner, P. J. (2007). Contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework, *Journal of Geophysical Research*, **112**(D09307), http://dx.doi.org/10.1029/2006JD007375.

Dennis, Jr., J. E. and Moré, J. J. (1977). Quasi-Newton methods, motivation and theory, *SIAM Review*, **19**(1), 46–89, http://dx.doi.org/10.1137/1019005.

Derber, J. and Rosati, A. (1989). A global oceanic data assimilation system, *Journal of Physical Oceanography*, **19**, 1333–1347.

Fisher, M. and Courtier, P. (1995). Estimating the covariance matrices of analysis and forecast error in variational data assimilation, Technical Memorandum, No.220, ECMWF, Reading, UK.

Fujii, Y. (2005). Preconditioned Optimizing Utility for Large-dimensional analyses (POpULar), *Journal of Oceanography*, **61**(1), 167–181, http://dx.doi.org/10.1007/s10872-005-0029-z.

Fujii, Y. and Kamachi, M. (2003). A nonlinear preconditioned quasi-Newton method without inversion of a first-guess covariance matrix in variational analyses, *Tellus A*, **55**(5), 450–454, http://dx.doi.org/10.1034/j.1600-0870.2003.00030.x.

Golub, G. H. and van Loan, C. F. (1996). *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, Maryland.

Ito, S., Nagao, H., Yamanaka, A., Tsukada, Y., Koyama, T., Kano, M. and Inoue, J. (2016). Data assimilation for massive autonomous systems based on a second-order adjoint method, *Physical Review E*, **94**, p.043307, http://dx.doi.org/10.1103/PhysRevE.94.043307.

Lorenc, A. C. (1988). Optimal nonlinear objective analysis, *Quarterly Journal of the Royal Meteorological Society*, **114**(479), 205–240, http://dx.doi.org/10.1002/qj.49711447911.

Nazareth, L. (1979). A Relationship between the BFGS and Conjugate Gradient Algorithms and Its Implications for New Algorithms, *SIAM Journal on Numerical Analysis*, **16**(5), 794–800, http://dx.doi.org/10.1137/0716059.

Niwa, Y. and Fujii, Y. (2020). A conjugate BFGS method for accurate estimation of a posterior error covariance matrix in a linear inverse problem, *Quarterly Journal of the Royal Meteorological Society*, **146**(732), 3118–3143, http://dx.doi.org/10.1002/qj.3838.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*, 2nd ed, Springer, New York.

Rabier, F. and Courtier, P. (1992). Four-dimensional assimilation in the presence of baroclinic instability, *Quarterly Journal of the Royal Meteorological Society*, **118**(506), 649–672, http://dx.doi.org/10.1002/qj.49711850604.

Sasaki, Y. (1969). Proposed inclusion of time variation terms, observational and theoretical, in numerical variational objective analysis, *Journal of the Meteorological Society of Japan*, **47**(2), 115–124, http://dx.doi.org/10.2151/jmsj1965.47.2_115.

露木 義, 川畑拓矢 編 (2008).『気象学におけるデータ同化』, 気象研究ノート 217 号, 日本気象学会, 東京.

Ueno, G. and Nakamura, N. (2016). Bayesian estimation of the observation-error covariance matrix in ensemble-based filters, *Quarterly Journal of the Royal Meteorological Society*, **142**(698), 2055–2080, http://dx.doi.org/10.1002/qj.2803.

Estimation of a Posterior Error Covariance Matrix Using Conjugate Vectors and the BFGS Formula

Yosuke Niwa[1,2] and Yosuke Fujii[2,3]

[1]Earth System Division, National Institute for Environmental Studies
[2]Department of Atmosphere, Ocean, and Earth System Modeling Research, Meteorological Research Institute, Japan Meteorological Agency
[3]The Institute of Statistical Mathematics

A four-dimensional variational method is commonly used for data assimilation/inverse problems. However, it cannot automatically provide a posterior error. Niwa and Fujii (2020) developed a technique to estimate a posterior error covariance matrix within the framework of the four-dimensional variational method. Their technique adopts a quasi-Newton method with the BFGS formula, which is a conventional optimizing method. To enhance the estimation accuracy of a posterior error covariance matrix, their technique also employs an exact line search, an ensemble method, and orthogonalization to increase the number of conjugate vectors used in the BFGS formula.

This report explains the fact that with preconditioning in the BFGS formula, an analytical posterior error covariance matrix can be obtained from the same number of iterations (or vector pairs used in the BFGS formula) as observations followed by a detailed description of Niwa and Fujii's technique (2020). Finally, the results obtained by applying this technique to an inverse problem of atmospheric CO_2 are demonstrated.

Key words: Posterior error covariance matrix, data assimilation, inverse analysis, four-dimensional variational method, BFGS formula, quasi-Newton method.