



Analysis error via Hessian in variational data assimilation

Francois-Xavier Le Dimet* — Victor Shutyaev** — Igor Gejadze*

*LMC-IMAG, Université Joseph Fourier,
BP 51, 38051 Grenoble Cedex 9, France
ledimet@imag.fr; igor.gejadze@imag.fr

**Institute of Numerical Mathematics,
Russian Academy of Sciences,
119991 Gubkina 8, Moscow, Russia
shutyaev@inm.ras.ru



RÉSUMÉ. Le problème de l'assimilation variationnelle de données pour un modèle non linéaire d'évolution est formulé comme un problème de contrôle optimal par rapport à la condition initiale. En utilisant le Hessien de la fonction coût et l'adjoint au second ordre, on dérive une équation gouvernant la propagation des statistiques d'erreur des entrées du problème vers la condition initiale. La dépendance de l'opérateur de covariance de l'erreur d'analyse est exprimé en fonction de celui de la covariance des erreurs des entrées du modèle (erreur d'ébauche et erreur d'observation). Des algorithmes sont proposés pour la construction de la covariance de l'analyse à partir de la covariance des entrées.

ABSTRACT. The problem of variational data assimilation for a nonlinear evolution model is formulated as an optimal control problem to find the initial condition function. Based on the Hessian of the cost functional and the second-order adjoint techniques, the equation for the error of the optimal solution (analysis) is derived through the statistical errors of the input data. The covariance operator of the analysis error is expressed through the covariance operators of the input errors (background and observation errors). Numerical algorithms are developed to construct the covariance operator of the analysis error using the covariance operators of the input errors.

MOTS-CLÉS : l'assimilation de données, contrôle optimal, l'erreur d'analyse, le Hessien, l'opérateur de covariance

KEYWORDS : data assimilation, optimal control, analysis error, Hessian, covariance operator



1. Introduction

The methods of data assimilation were designed to combine models and observational data as sources of information. From the mathematical point of view, these problems may be formulated as optimal control problems (e.g. [5]). A major advantage of this technique is the derivation of an optimality system which contains all the available information. In practice the optimality system includes background errors and observation errors. The error of the optimal solution (analysis) may be derived through the errors of the input data using the Hessian of the cost functional and the second-order adjoint techniques. For deterministic case it was done in [6]. Here we present the developments of the ideas of [6] for the case of statistical errors.

2. Statement of the problem

Consider the mathematical model of a physical process that is described by the evolution problem

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases} \quad (2.1)$$

where $\varphi = \varphi(t)$ is the unknown function belonging for any t to a Hilbert space X , $u \in X$, F is a nonlinear operator mapping X into X . Let $Y = L_2(0, T; X)$, $\|\cdot\|_Y = (\cdot, \cdot)_Y^{1/2}$, $f \in Y$. Suppose that for given $u \in X$, $f \in Y$ there exists a unique solution $\varphi \in Y$ to (2.1).

Let us introduce the functional

$$S(u) = \frac{1}{2}(V_1(u - u_0), u - u_0)_X + \frac{1}{2}(V_2(C\varphi - \varphi_{obs}), C\varphi - \varphi_{obs})_{Y_{obs}}, \quad (2.2)$$

where $\alpha = const \geq 0$, $u_0 \in X$ is a prior initial-value function (background state), $\varphi_{obs} \in Y_{obs}$ is a prescribed function (observational data), Y_{obs} is a Hilbert space (observation space), $C : Y \rightarrow Y_{obs}$ is a linear bounded operator, $V_1 : X \rightarrow X$ and $V_2 : Y_{obs} \rightarrow Y_{obs}$ are symmetric positive definite operators.

Consider the following data assimilation problem with the aim to identify the initial condition : find $u \in X$ and $\varphi \in Y$ such that they satisfy (2.1), and on the set of solutions to (2.1), the functional $S(u)$ takes the minimum value, i.e.

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \\ S(u) = \inf_v S(v). \end{cases} \quad (2.3)$$

The necessary optimality condition reduces the problem (2.3) to the following system [7], [1] :

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \end{cases} \quad (2.4)$$

$$\begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* = -C^* V_2 (C\varphi - \varphi_{obs}), & t \in (0, T) \\ \varphi^*|_{t=T} = 0, \end{cases} \quad (2.5)$$

$$V_1(u - u_0) - \varphi^*|_{t=0} = 0 \quad (2.6)$$

with the unknowns φ, φ^*, u , where $(F'(\varphi))^*$ is the adjoint to the Frechet derivative of F , and C^* is the adjoint to C defined by $(C\varphi, \psi)_{Y_{obs}} = (\varphi, C^*\psi)_Y$, $\varphi \in Y, \psi \in Y_{obs}$. We assume that the system (2.4)–(2.6) has a unique solution.

Suppose that $u_0 = \bar{u} + \xi_1$, $\varphi_{obs} = C\bar{\varphi} + \xi_2$, where $\xi_1 \in X$, $\xi_2 \in Y_{obs}$, and $\bar{\varphi}$ is the ("true") solution to the problem (1.1) with $u = \bar{u}$:

$$\begin{cases} \frac{\partial \bar{\varphi}}{\partial t} = F(\bar{\varphi}) + f, & t \in (0, T) \\ \bar{\varphi}|_{t=0} = \bar{u}. \end{cases} \quad (2.7)$$

The functions ξ_1, ξ_2 may be treated as the errors of the input data u_0, φ_{obs} (background and observation errors, respectively). For V_1 and V_2 in (2.2), one usually has $V_1 = V_{\xi_1}^{-1}$, $V_2 = V_{\xi_2}^{-1}$, where V_{ξ_i} is the covariance operator of the corresponding error ξ_i , $i = 1, 2$.

Having supposed that the solution of the problem (2.4)–(2.6) exists, we will study the influence of the errors ξ_1, ξ_2 on the optimal solution u and develop the theory presented in [6] for the case of statistical errors. We derive the covariance operator of the optimal solution error through the covariance operators of the input errors. Numerical algorithms are developed to construct the covariance operator of the optimal solution error using the covariance operators of the input errors.

3. Error analysis via Hessian

The system (2.4)–(2.6) with the three unknowns φ, φ^*, u may be treated as an operator equation of the form

$$\mathcal{F}(U, U_d) = 0, \quad (3.1)$$

where $U = (\varphi, \varphi^*, u)$, $U_d = (u_0, \varphi_{obs}, f)$, and the action of \mathcal{F} is defined by

$$\mathcal{F}(U, U_d) = \begin{cases} \frac{\partial \varphi}{\partial t} - F(\varphi) - f, \\ \varphi|_{t=0} - u, \\ -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* + C^* V_2 (C\varphi - \varphi_{obs}), \\ \varphi^*|_{t=T}, \\ V_1 (u - u_0) - \varphi^*|_{t=0}. \end{cases}$$

Thus, $\mathcal{F}(U, U_d)$ is linear in u, u_0, φ_{obs}, f . The following equality holds for the "exact solution" ("true state") :

$$\mathcal{F}(\bar{U}, \bar{U}_d) = 0, \quad (3.2)$$

with $\bar{U} = (\bar{\varphi}, \bar{\varphi}^*, \bar{u})$, $\bar{U}_d = (\bar{u}, C\bar{\varphi}, f)$, $\bar{\varphi}^* = 0$. The system (3.2) is the necessary optimality condition of the following minimization problem : find u and φ such that

$$\begin{cases} \frac{\partial \varphi}{\partial t} = F(\varphi) + f, & t \in (0, T) \\ \varphi|_{t=0} = u, \\ \bar{S}(u) = \inf_v \bar{S}(v), \end{cases}$$

where

$$\bar{S}(u) = \frac{1}{2} (V_1 (u - \bar{u}), u - \bar{u})_X + \frac{1}{2} (V_2 (C\varphi - C\bar{\varphi}), C\varphi - C\bar{\varphi})_{Y_{obs}}.$$

From (3.1)–(3.2), we get

$$\mathcal{F}(U, U_d) - \mathcal{F}(\bar{U}, \bar{U}_d) = 0. \quad (3.3)$$

Let $\delta U = U - \bar{U}$, $\delta U_d = U_d - \bar{U}_d$. Then (3.3) gives

$$\mathcal{F}(\bar{U} + \delta U, \bar{U}_d + \delta U_d) - \mathcal{F}(\bar{U}, \bar{U}_d) = 0. \quad (3.4)$$

From (3.4), for regular \mathcal{F} , there exists $\tilde{U} = (\tilde{\varphi}, \tilde{\varphi}^*, u)$ such that $\tilde{\varphi} = \bar{\varphi} + \tau(\varphi - \bar{\varphi})$, $\tilde{\varphi}^* = \bar{\varphi}^* + \tau(\varphi^* - \bar{\varphi}^*)$, $\tau \in \mathbf{R}$, and

$$\mathcal{F}'_{\tilde{U}}(\tilde{U}, U_d)\delta U + \mathcal{F}'_{U_d}(\tilde{U}, U_d)\delta U_d = 0, \quad (3.5)$$

where $\mathcal{F}'_{\tilde{U}}$, \mathcal{F}'_{U_d} are the Gateaux derivatives with respect to U and U_d .

Let $\delta\varphi = \varphi - \bar{\varphi}$, $\delta u = u - \bar{u}$; then $\delta U = (\delta\varphi, \varphi^*, \delta u)$, $\delta U_d = (\xi_1, \xi_2, 0)$. By calculating the derivatives $\mathcal{F}'_{\tilde{U}}$, \mathcal{F}'_{U_d} , it is easily seen that equation (3.5) is equivalent to the system :

$$\begin{cases} \frac{\partial \delta\varphi}{\partial t} - F'(\tilde{\varphi})\delta\varphi = 0, & t \in (0, T), \\ \delta\varphi|_{t=0} = \delta u, \end{cases} \quad (3.6)$$

$$\begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\tilde{\varphi}))^*\varphi^* = (F''(\tilde{\varphi})\delta\varphi)^*\tilde{\varphi}^* - C^*V_2(C\delta\varphi - \xi_2), \\ \varphi^*|_{t=T} = 0, \end{cases} \quad (3.7)$$

$$V_1(\delta u - \xi_1) - \varphi^*|_{t=0} = 0. \quad (3.8)$$

The problem (3.6)–(3.8) is a linear data assimilation problem ; for fixed $\tilde{\varphi}$, $\tilde{\varphi}^*$ it is the necessary optimality condition to the following minimization problem : find u and φ such that

$$\begin{cases} \frac{\partial \varphi}{\partial t} - F'(\tilde{\varphi})\varphi = 0, & t \in (0, T) \\ \varphi|_{t=0} = u \\ S_1(u) = \inf_v S_1(v), \end{cases} \quad (3.9)$$

where

$$S_1(u) = \frac{1}{2}(V_1(u - \xi_1), u - \xi_1)_X + \frac{1}{2}(V_2(C\varphi - \xi_2), C\varphi - \xi_2)_{Y_{obs}} - \frac{1}{2}(F''(\tilde{\varphi})\varphi\varphi, \tilde{\varphi}^*)_Y. \quad (3.10)$$

Consider the Hessian H of the functional (3.10) ; it is defined by the successive solutions of the following problems :

$$\begin{cases} \frac{\partial \psi}{\partial t} - F'(\tilde{\varphi})\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases} \quad (3.11)$$

$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\tilde{\varphi}))^*\psi^* = (F''(\tilde{\varphi})\psi)^*\tilde{\varphi}^* - C^*V_2C\psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases} \quad (3.12)$$

$$Hv = V_1v - \psi^*|_{t=0}. \quad (3.13)$$

Note that for $\tilde{\varphi} = \varphi$, $\tilde{\varphi}^* = \varphi^*$, where φ, φ^* are the solutions of (2.4)–(2.6), the operator H coincides with the Hessian of the original functional $S(u)$. (In this case the equation (3.5) is satisfied with an accuracy of the second order in δU for regular \mathcal{F} .)

Below we introduce two auxiliary operators R_1, R_2 . Let $R_1 = V_1$. Let us introduce the operator $R_2 : Y_{obs} \rightarrow X$ acting on the functions $g \in Y_{obs}$ according to the formula

$$R_2 g = \theta^*|_{t=0}, \quad (3.14)$$

where θ^* is the solution to the adjoint problem

$$\begin{cases} -\frac{\partial \theta^*}{\partial t} - (F'(\tilde{\varphi}))^* \theta^* &= V_2 C^* g, \quad t \in (0, T) \\ \theta^*|_{t=T} &= 0. \end{cases} \quad (3.15)$$

>From (3.11)–(3.15) we conclude that the system (3.6)–(3.8) is equivalent to the single equation for δu :

$$H \delta u = R_1 \xi_1 + R_2 \xi_2. \quad (3.16)$$

The Hessian H acts in X as a self-adjoint operator with domain of definition $D(H)=X$. We will suppose that H is positive definite, and hence invertible.

As follows from (3.16), the influence of the errors ξ_1, ξ_2 on the value of the error δu of the optimal solution is determined by the operators $H^{-1}R_1, H^{-1}R_2$, respectively. The values of the norms of these operators may be considered as sensitivity coefficients : the less is the norm of the operator $H^{-1}R_i$, the less impact on δu is given by the corresponding error ξ_i . This criteria was used for deterministic error analysis in [4], [6]. Here, assuming the statistical structure of the errors ξ_1, ξ_2 , we will derive the covariance operator of the optimal solution error through the covariance operators of the input errors and develop numerical algorithms to construct the covariance operator of the optimal solution error using the covariance operators of the input errors.

4. Covariance operators

The analysis-error covariances through the Hessian in variational data assimilation were considered by many authors (e.g. [9, 10, 8, 2, 11, 3]) usually for a linearized model (so-called tangent linear hypothesis). We will use the equation (3.16) to derive the formulas for the covariance operator of the optimal solution errors involving the Hessian of the functional $S(u)$ of the original nonlinear data assimilation problem (2.3).

Consider the error equation (3.16). Under the hypotheses that H is invertible, we may rewrite it as

$$\delta u = T_1 \xi_1 + T_2 \xi_2, \quad (4.1)$$

where $T_i = H^{-1}R_i$, $T_1 : X \rightarrow X$, $T_2 : Y_{obs} \rightarrow X$.

Below we suppose that in (3.16) $\tilde{\varphi} = \varphi$, $\tilde{\varphi}^* = \varphi^*$, where φ, φ^* are the solutions of the original optimality system (2.4)–(2.6). As we have mentioned above, in this case the operator H coincides with the Hessian of the original functional $S(u)$ and it can be defined through φ, φ^*, u by the successive solutions of the following problems (for a given $v \in X$) :

$$\begin{cases} \frac{\partial \varphi}{\partial t} &= F(\varphi) + f, \quad t \in (0, T) \\ \varphi|_{t=0} &= u, \end{cases} \quad (4.2)$$

$$\begin{cases} -\frac{\partial \varphi^*}{\partial t} - (F'(\varphi))^* \varphi^* &= -C^* V_2 (C\varphi - \varphi_{obs}), \quad t \in (0, T) \\ \varphi^*|_{t=T} &= 0, \end{cases} \quad (4.3)$$

$$\begin{cases} \frac{\partial \psi}{\partial t} - F'(\varphi)\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases} \quad (4.4)$$

$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^* \psi^* = (F''(\varphi)\psi)^* \varphi^* - C^* V_2 C \psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases} \quad (4.5)$$

$$Hv = V_1 v - \psi^*|_{t=0}. \quad (4.6)$$

We suppose that the errors ξ_1, ξ_2 are normally distributed, unbiased, and mutually uncorrelated. By V_{ξ_i} we denote the covariance operator of the corresponding error ξ_i , $i = 1, 2$, i.e. $V_{\xi_1} \cdot = E[(\cdot, \xi_1)_X \xi_1]$, $V_{\xi_2} \cdot = E[(\cdot, \xi_2)_{Y_{obs}} \xi_2]$, where E is the expectation. By $V_{\delta u}$ we denote the covariance operator of the optimal solution (analysis) error : $V_{\delta u} \cdot = E[(\cdot, \delta u)_X \delta u]$. From (4.1) we get

$$V_{\delta u} = T_1 V_{\xi_1} T_1^* + T_2 V_{\xi_2} T_2^*. \quad (4.7)$$

To find the covariance operator $V_{\delta u}$, we need to construct the operators $T_i V_{\xi_i} T_i^*$, $i = 1, 2$.

Consider the operator $T_1 V_{\xi_1} T_1^*$. Since $T_1 = H^{-1} R_1 = H^{-1} V_1 = T_1^*$, we have $T_1 V_{\xi_1} T_1^* = H^{-1} V_1 V_{\xi_1} V_1 H^{-1}$. Moreover, if $V_1 = V_{\xi_1}^{-1}$, then

$$T_1 V_{\xi_1} T_1^* = H^{-1} V_1 H^{-1} = H^{-1} V_{\xi_1}^{-1} H^{-1}. \quad (4.8)$$

Thus, the algorithm for finding $w = T_1 V_{\xi_1} T_1^* v$, $v \in X$, consists in the the following :

- 1) solve the equation $Hp = v$;
- 2) compute $V_1 p$;
- 3) solve the equation $Hw = V_1 p$.

Consider the operator $T_2 V_{\xi_2} T_2^*$. Since $T_2 = H^{-1} R_2$, then

$$T_2 V_{\xi_2} T_2^* = H^{-1} R_2 V_{\xi_2} R_2^* H^{-1}.$$

To determine R_2^* , consider the inner product $(R_2 g, p)_X$, $g \in Y_{obs}$, $p \in X$. From (3.14)–(3.15),

$$(R_2 g, p)_X = (\theta^*|_{t=0}, p)_X = (C^* V_2 g, \phi)_Y = (g, R_2^* p)_{Y_{obs}},$$

where $R_2^* p = V_2 C \phi$, and ϕ is the solution to the problem

$$\begin{cases} \frac{\partial \phi}{\partial t} - F'(\varphi)\phi = 0, & t \in (0, T), \\ \phi|_{t=0} = p. \end{cases} \quad (4.9)$$

Thus, the operator $T_2 V_{\xi_2} T_2^*$ is defined by successive solutions of the following problems (for a given $v \in X$) :

$$Hp = v, \quad (4.10)$$

$$\begin{cases} \frac{\partial \phi}{\partial t} - F'(\varphi)\phi = 0, & t \in (0, T), \\ \phi|_{t=0} = p, \end{cases} \quad (4.11)$$

$$\begin{cases} -\frac{\partial \theta^*}{\partial t} - (F'(\varphi))^* \theta^* = C^* V_2 V_{\xi_2} V_2 C \phi, & t \in (0, T) \\ \theta^*|_{t=T} = 0, \end{cases} \quad (4.12)$$

$$Hw = \theta^*|_{t=0}, \quad (4.13)$$

then

$$T_2 V_{\xi_2} T_2^* v = w. \quad (4.14)$$

If $V_2 = V_{\xi_2}^{-1}$, then $C^* V_2 V_{\xi_2} V_2 C = C^* V_2 C$ and from (4.11)–(4.12) we obtain that

$$\theta^*|_{t=0} = H_0 p - V_1 p,$$

where H_0 is the Hessian of the tangent linear approximation, that is, the Hessian of the functional $S(u)$ when the nonlinear model in (2.3) is replaced by its tangent linear, and it is defined by successive solutions of the following problems (for a given $v \in X$):

$$\begin{cases} \frac{\partial \psi}{\partial t} - F'(\varphi)\psi = 0, & t \in (0, T), \\ \psi|_{t=0} = v, \end{cases} \quad (4.15)$$

$$\begin{cases} -\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^* \psi^* = -C^* V_2 C \psi, & t \in (0, T) \\ \psi^*|_{t=T} = 0, \end{cases} \quad (4.16)$$

$$H_0 v = V_1 v - \psi^*|_{t=0}. \quad (4.17)$$

Then we get

$$R_2 V_{\xi_2} R_2^* = H_0 - V_1$$

and

$$T_2 V_{\xi_2} T_2^* = H^{-1}(H_0 - V_1)H^{-1}. \quad (4.18)$$

Thus, the algorithm for finding $w = T_2 V_{\xi_2} T_2^* v$, $v \in X$, consists in the the following :

- 1) solve the equation $Hp = v$;
- 2) compute $(H_0 - V_1)p$;
- 3) solve the equation $Hw = (H_0 - V_1)p$.

>From (4.8), (4.18) it follows the result for $V_{\delta u}$:

$$V_{\delta u} = T_1 V_{\xi_1} T_1^* + T_2 V_{\xi_2} T_2^* = H^{-1} H_0 H^{-1}. \quad (4.19)$$

The last formula gives the analysis-error covariance operator through the Hessian H of the original nonlinear data assimilation problem and the Hessian H_0 of the tangent linear approximation. If the tangent linear hypothesis is valid, then omitting $F''(\varphi)$ in (4.5), we have $H = H_0$, and the right-hand side of (4.19) gives

$$H^{-1} H_0 H^{-1} = H_0^{-1} H_0 H_0^{-1} = H_0^{-1},$$

i.e. the covariance operator is the inverse Hessian in accordance with the well-known results (e.g. [9, 8]).

Acknowledgements

This work was carried out within the IDOPT project (CNRS, INRIA, UJF, INPG) and the project 04-05-64481 of the Russian Foundation for the Basic Research and within the programme ECO-NET (EGIDE).

5. Bibliographie

- [1] V. AGOSHKOV, G. MARCHUK « On solvability and numerical solution of data assimilation problems », *Russ. J. Numer. Analys. Math. Modelling*, vol. 8, n° 1 : 1–16, 1993.
- [2] J.R. GUNSON, P. MALANOTTE-RIZZOLI, « Assimilation studies of open-ocean flows. 2. Error measures with strongly nonlinear dynamics », *J. Geophys. Res.*, vol. 101, n° C12 : 28473–28488, 1996.
- [3] F.-X. LE DIMET, I.M. NAVON, D.N. DAESCU, « Second-order information in data assimilation », *Monthly Weather Review*, vol. 130, n° 3 : 629–648, 1996.
- [4] F.-X. LE DIMET, P. NGNEPIEBA, V. SHUTYAEV, « On error analysis in data assimilation problems », *Russ. J. Numer. Anal. Math. Modelling*, vol. 17, n° 1 : 71–97, 2002.
- [5] F.-X. LE DIMET, O. TALAGRAND, « Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects », *Tellus*, vol. 38A : 97–110, 1986.
- [6] F.-X. LE DIMET, V. SHUTYAEV, « On deterministic error analysis in variational data assimilation », *Nonlinear Processes in Geophysics*, vol. 12 : 481–490, 2005.
- [7] J.-L. LIONS, « Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles », *Paris : Dunod*, 1968.
- [8] F. RABIER, P. COURTIER, « Four-dimensional assimilation in the presence of baroclinic instability », *Quart. J. Roy. Meteorol. Soc.*, vol. 118 : 649–672, 1992.
- [9] W.C. THACKER, « The role of the Hessian matrix in fitting models to measurements », *J. Geophys. Res.*, vol. 94, n° C5 : 6177–6196, 1989.
- [10] J.N. THEPAUT, P. COURTIER, « Four-dimensional variational assimilation using the adjoint of a multilevel primitive equation model », *Quart. J. Roy. Meteorol. Soc.*, vol. 117 : 1225–1254, 1991.
- [11] F. VEERSÉ, « Variable-storage quasi-Newton operators as inverse forecast/analysis error covariance matrices in variational data assimilation », *INRIA Technical Report*, n° 3685, 1999.