Posterior Covariance vs. Analysis Error Covariance in Data Assimilation

F.-X. Le Dimet(1), I. Gejadze(2), V. Shutyaev(3)

(1) Université de Grenoble
(2) University of Strathclyde, Glasgow, UK
(3) Institute of Numerical Mathematics, RAS, Moscow, Russia

ledimet@imag.fr

September 27, 2013
Overview

- Introduction
- Analysis Error Covariance via Hessian
- Posterior Covariance : A Bayesian approach
- Effective Covariance Estimates
- Implementation : some remarks
- Asymptotic Properties
- Numerical Example
- Conclusion
There are two basic approaches for Data Assimilation:

- Variational Methods
- Kalman Filter

Both lead to minimize a cost function:

\[
J(u) = \frac{1}{2} (V_b^{-1}(u - u_b), u - u_b)_X + \frac{1}{2} (V_o^{-1}(C\varphi - y), C\varphi - y)_{Y_o}, \quad (1)
\]

where \( u_b \in X \) is a prior initial-value function (background state), \( y \in Y_o \) is a prescribed function (observational data), \( Y_o \) is an observation space, \( C : Y \to Y_o \) is a linear bounded operator.

We get the same optimal solution \( \bar{u} \)
\( \bar{u} \) is the solution of the Optimality System:

\[
\begin{align*}
\frac{\partial \varphi}{\partial t} &= F(\varphi) + f, \quad t \in (0, T) \\
\varphi|_{t=0} &= u,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial \varphi^*}{\partial t} + (F'(\varphi))^* \varphi^* &= C^* V_o^{-1} (C \varphi - y), \quad t \in (0, T) \\
\varphi^*|_{t=T} &= 0,
\end{align*}
\]

\[
V_b^{-1}(u - u_b) - \varphi^*|_{t=0} = 0
\]
In an analysis there are two inputs:

- The background $u_b$
- the observation $y$

Both have error and the question is what is the impact of these errors on the analysis?

The background can be considered from two different viewpoints:

- Variational viewpoint: the background is a regularization term in the Tykhonov’s sense to make the problem well posed
- Bayesian viewpoint: the background is an a priori information on the analysis
In the linear case we get the same covariance error for the analysis: the inverse of the Hessian of the cost function.

In the non linear case we get two different items:

- Variational approach: Analysis Error Covariance
- Bayesian Approach: Posterior Covariance

Questions:

- How to compute, approximate, these elements?
- What are the differences?
We assume the existence of a "true" solution \( u^t \) and an associated "true" state \( \varphi^t \) verifying:

\[
\begin{align*}
\frac{\partial \varphi^t}{\partial t} &= F(\varphi^t) + f, \quad t \in (0, T) \\
\varphi^t \bigg|_{t=0} &= u^t.
\end{align*}
\]  

(5)

Then the errors are defined by: \( u_b = u^t + \xi_b \), \( y = C\varphi^t + \xi_o \) with covariances \( V_b \) and \( V_o \).
Let

\[ \delta \varphi = \varphi - \varphi^t, \quad \delta u = u - u^t. \]

Then for regular \( F \) there exists \( \tilde{\varphi} = \varphi^t + \tau(\varphi - \varphi^t), \quad \tau \in [0, 1], \) such that

\[
\begin{align*}
\frac{\partial \delta \varphi}{\partial t} - F'(\tilde{\varphi})\delta \varphi &= 0, \quad t \in (0, T), \\
\delta \varphi |_{t=0} &= \delta u,
\end{align*}
\]

(6)

\[
\begin{align*}
\frac{\partial \varphi^*}{\partial t} + (F'(\varphi))^* \varphi^* &= C^* V_o^{-1} (C \delta \varphi - \xi_o), \\
\varphi^* |_{t=T} &= 0,
\end{align*}
\]

(7)

\[
V_b^{-1}(\delta u - \xi_b) - \varphi^* |_{t=0} = 0.
\]

(8)
Let us introduce the operator $R(\varphi) : X \rightarrow Y$ as follows:

$$R(\varphi)v = \psi, \quad v \in X,$$

where $\psi$ is the solution of the tangent linear problem

$$\frac{\partial \psi}{\partial t} - F'(\varphi)\psi = 0, \quad \psi|_{t=0} = v. \quad (10)$$

Then, the system for errors can be represented as a single operator equation for $\delta u$:

$$H(\varphi, \tilde{\varphi})\delta u = V_b^{-1}\xi_b + R^*(\varphi)C^*V_o^{-1}\xi_o, \quad (11)$$

where

$$H(\varphi, \tilde{\varphi}) = V_b^{-1} + R^*(\varphi)C^*V_o^{-1}CR(\tilde{\varphi}). \quad (12)$$
The operator $H(\varphi, \tilde{\varphi}) : X \rightarrow X$ can be defined by

$$\frac{\partial \psi}{\partial t} - F'(\tilde{\varphi})\psi = 0, \quad \psi|_{t=0} = v,$$

$$-\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^*\psi^* = -C^*V_o^{-1}C\psi, \quad \psi^*|_{t=T} = 0,$$

$$H(\varphi, \tilde{\varphi})v = V_b^{-1}v - \psi^*|_{t=0}.\quad \quad (13), \quad (14), \quad (15)$$

The operator $H(\varphi, \tilde{\varphi})$ is neither symmetric, nor positive definite.

$\varphi = \tilde{\varphi} = \theta$, it becomes the Hessian $H(\theta)$ of the cost function $J_1$ in the following auxiliary DA problem: find $\delta u$ and $\delta \varphi$ such that

$$J_1(\delta u) = \inf_v J_1(v),$$

$$J_1(\delta u) = \frac{1}{2}(V_b^{-1}(\delta u - \xi_b), \delta u - \xi_b)_X + \frac{1}{2}(V_o^{-1}(C\delta \varphi - \xi_o), C\delta \varphi - \xi_o)_Y,$$

and $\delta \varphi$ satisfies the problem

$$\frac{\partial \delta \varphi}{\partial t} - F'(\theta)\delta \varphi = 0, \quad \delta \varphi|_{t=0} = \delta u.$$

$$\quad \quad (16), \quad (17)$$
The optimal solution (analysis) error $\delta u$ is assumed to be unbiased, i.e. $E[\delta u] = 0$, and

$$V_{\delta u} \cdot = E[(\cdot, \delta u)X \delta u] = E[(\cdot, u - u^t)X (u - u^t)].$$

(18)

The best value of $\varphi$ and $\tilde{\varphi}$ independent of $\xi_o, \xi_b$ is apparently $\varphi^t$ and using

$$R(\tilde{\varphi}) \approx R(\varphi^t), \ R^*(\varphi) \approx R^*(\varphi^t),$$

(19)

the error equation reduces to

$$H(\varphi^t)\delta u = V^{-1}_b \xi_b + R^*(\varphi^t)C^*V^{-1}_o \xi_o, \ H(\cdot) = V^{-1}_b + R^*(\cdot)C^*V^{-1}_o CR(\cdot).$$

(20)

We express $\delta u$ from equation (??)

$$\delta u = H^{-1}(\varphi^t)(V^{-1}_b \xi_b + R^*(\varphi^t)C^*V^{-1}_o \xi_o)$$

and obtain for the **analysis error covariance**

$$V_{\delta u} = H^{-1}(\varphi^t)(V^{-1}_b + R^*(\varphi^t)C^*V^{-1}_o CR(\varphi^t))H^{-1}(\varphi^t) = H^{-1}(\varphi^t).$$

(21)
In practice the 'true' field $\varphi^t$ is not known, thus we have to use an approximation $\bar{\varphi}$ associated to a certain optimal solution $\bar{u}$ defined by the real data $(\bar{u}_b, \bar{y})$, i.e. we use

$$V_{\delta u} = H^{-1}(\bar{\varphi}).$$  \hfill (22)

In (Rabier and Courtier, 1992), the error equation is derived in the form

$$(V_b^{-1} + R^*(\varphi)C^*V_o^{-1}CR(\varphi))\delta u = V_b^{-1}\xi_b + R^*(\varphi)C^*V_o^{-1}\xi_o.$$  \hfill (23)

The error due to transitions $R(\bar{\varphi}) \rightarrow R(\varphi^t)$ and $R^*(\varphi) \rightarrow R^*(\varphi^t)$; we call it the 'linearization' error. The use of $\bar{\varphi}$ instead of $\varphi^t$ in the Hessian computations leads to another error, which shall be called the 'origin' error.
Given $u_b \sim \mathcal{N}(\bar{u}_b, V_b)$, $y \sim \mathcal{N}(\bar{y}, V_o)$, the following expression for the posterior distribution of $u$ is derived from the Bayes theorem:

$$p(u|\bar{y}) = C \cdot \exp \left( -\frac{1}{2}(V_b^{-1}(u - \bar{u}_b), u - \bar{u}_b) \right) \cdot \exp \left( -\frac{1}{2}(V_o^{-1}(C\varphi - \bar{y}), C\varphi - \bar{y}) \right).$$

(24)

The solution to the variational DA problem with the data $y = \bar{y}$ and $u_b = \bar{u}$ is equal to the mode of $p(u, \bar{y})$ (see e.g. Lorenc, 1986; Tarantola, 1987). Accordingly, the Bayesian posterior covariance is defined by:

$$\mathcal{V}_{\delta u} = E[(\cdot, u - E[u])X (u - E[u])]$$

(25)

with $u \sim p(u|\bar{y})$. 
In order to compute $\mathcal{V}_{\delta u}$ by the Monte Carlo method, one must generate a sample of pseudo-random realizations $u_i$ from $p(u|\bar{y})$. We will consider $u_i$ to be the solutions to the DA problem with the perturbed data $u_b = \bar{u}_b + \xi_b$, and $y = \bar{y} + \xi_o$, where $\xi_b \sim \mathcal{N}(0, V_b)$, $\xi_o \sim \mathcal{N}(0, V_o)$. Further we assume that $E[u] = \bar{u}$, where $\bar{u}$ is the solution to the unperturbed problem in which case $\mathcal{V}_{\delta u}$ can be approximated as follows

$$\mathcal{V}_{\delta u} = E[(\cdot, u - \bar{u}) \times (u - \bar{u})] = E[(\cdot, \delta u) \times \delta u].$$  

(26)
Unperturbed O.S. with \( u_b = \bar{u}_b, \ y = \bar{y} \):

\[
\frac{\partial \bar{\varphi}}{\partial t} = F(\bar{\varphi}) + f, \ \varphi |_{t=0} = \bar{u}, \tag{27}
\]

\[
\frac{\partial \bar{\varphi}^*}{\partial t} + (F'(\bar{\varphi}))^* \bar{\varphi}^* = C^* V_o^{-1}(C \bar{\varphi} - \bar{y}), \ \bar{\varphi}^* |_{t=T} = 0, \tag{28}
\]

\[
V_b^{-1}(\bar{u} - \bar{u}_b) - \bar{\varphi}^* |_{t=0} = 0 \tag{29}
\]

With perturbations: \( u_b = \bar{u}_b + \xi_b, \ y = \bar{y} + \xi_o \), where \( \xi_b \in X, \xi_o \in Y_o \).

\( \delta u = u - \bar{u}, \ \delta \varphi = \varphi - \bar{\varphi}, \ \delta \varphi^* = \varphi^* - \bar{\varphi}^* \).

\[
\frac{\partial \delta \varphi}{\partial t} = F(\varphi) - F(\bar{\varphi}), \ \delta \varphi |_{t=0} = \delta u, \tag{30}
\]

\[
\frac{\partial \delta \varphi^*}{\partial t} + (F'(\varphi))^* \delta \varphi^* = [((F'(\bar{\varphi}))^* - F'(\varphi))^*] \bar{\varphi}^* + C^* V_o^{-1}(C \delta \varphi - \xi_o), \tag{31}
\]

\[
V_b^{-1}(\delta u - \xi_b) - \delta \varphi^* |_{t=0} = 0. \tag{32}
\]
Introducing $\tilde{\varphi}_1 = \bar{\varphi} + \tau_1 \delta \varphi$, $\tilde{\varphi}_2 = \bar{\varphi} + \tau_2 \delta \varphi$, $\tau_1, \tau_2 \in [0, 1]$, we derive the exact system for errors:

$$\frac{\partial \delta \varphi}{\partial t} = F'(\tilde{\varphi}_1) \delta \varphi, \quad \delta \varphi \bigg|_{t=0} = \delta u,$$  \hspace{1cm} (33)

$$\frac{\partial \delta \varphi^*}{\partial t} + (F'(\varphi))^* \delta \varphi^* = [(F'(\tilde{\varphi}_2))^* \bar{\varphi}^*]' \delta \varphi + C^* V_o^{-1}(C \delta \varphi - \xi_o),$$  \hspace{1cm} (34)

$$V^{-1}_b(\delta u - \xi_b) - \delta \varphi^* \bigg|_{t=0} = 0,$$  \hspace{1cm} (35)

equivalent to a single operator equation for $\delta u$:

$$\mathcal{H}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2) \delta u = V^{-1}_b \xi_b + R^*(\varphi) C^* V_o^{-1} \xi_o,$$  \hspace{1cm} (36)

where

$$\mathcal{H}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2) = V^{-1}_b + R^*(\varphi)(C^* V_o^{-1} C - [(F'(\tilde{\varphi}_2))^* \bar{\varphi}^*]') R(\tilde{\varphi}_1).$$  \hspace{1cm} (37)
Posterior Covariance: Operator $\mathcal{H}$

$\mathcal{H}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2) : X \to X$ is defined by solving:

$$
\frac{\partial \psi}{\partial t} = F'(\tilde{\varphi}_1)\psi, \quad \psi\big|_{t=0} = \nu, \tag{38}
$$

$$
-\frac{\partial \psi^*}{\partial t} - (F'(\varphi))^*\psi^* = [(F'(\tilde{\varphi}_2))^*\bar{\varphi}^*]'\psi - C^* V_o^{-1} C \psi, \tag{39}
$$

$$
\mathcal{H}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2)\nu = V_b^{-1} \nu - \psi^*\big|_{t=0}. \tag{40}
$$

If we had $\varphi = \tilde{\varphi}_1 = \tilde{\varphi}_2$, $\mathcal{H}(\varphi)$ becomes the Hessian of the cost function in the original DA problem, it is symmetric and positive definite if $u$ is a minimum of $J(u)$. The equation is often referred as the 'second order' adjoint (Le Dimet et al., 2002).

As above, we assume that $E(\delta u) \approx 0$, and we consider the following approximations

$$
R(\tilde{\varphi}_1) \approx R(\bar{\varphi}), \quad R^*(\varphi) \approx R^*(\bar{\varphi}), \quad [(F'(\tilde{\varphi}_2))^*\bar{\varphi}^*]' \approx [(F'(\bar{\varphi}))^*\bar{\varphi}^*]'. \tag{41}
$$
Posterior Covariance via Hessian

The exact error equation (42) is approximated as follows

$$\mathcal{H}(\bar{\phi})\delta u = V_b^{-1}\xi_b + R(\bar{\phi})^*C^*V_o^{-1}\xi_o,$$  \hspace{1cm} (42)

where

$$\mathcal{H}(\cdot) = V_b^{-1} + R^*(\cdot)(C^*V_o^{-1}C - [(F^*(\cdot))\bar{\phi}^*]'R(\cdot)).$$  \hspace{1cm} (43)

Now, we express $\delta u$:

$$\delta u = \mathcal{H}^{-1}(\bar{\phi})(V_b^{-1}\xi_b + R(\bar{\phi})^*C^*V_o^{-1}\xi_o),$$

and obtain an approximate expression for the **posterior error covariance**

$$\mathcal{V}_1 = \mathcal{H}^{-1}(\bar{\phi})(V_b^{-1} + R^*(\bar{\phi})V_o^{-1}R(\bar{\phi}))\mathcal{H}^{-1}(\bar{\phi}) = \mathcal{H}^{-1}(\bar{\phi})H(\bar{\phi})\mathcal{H}^{-1}(\bar{\phi}),$$ \hspace{1cm} (44)

where $H(\bar{\phi})$ is the Hessian of the cost function $J_1$ computed at $\theta = \bar{\phi}$.

Other approximations of the posterior covariance:

$$\mathcal{V}_2 = \mathcal{H}^{-1}(\bar{\phi}), \quad \mathcal{V}_3 = H^{-1}(\bar{\phi}).$$ \hspace{1cm} (45)
Posterior Covariance: "Effective" Estimates

To suppress the linearization errors, the 'effective' inverse Hessian may be used for estimating the analysis error covariance (see Gejadze et al., 2011):

\[ V_{\delta u} = E \left[ H^{-1}(\varphi) \right] . \]  

(46)

The same is true for the posterior error covariance \( V_{\delta u} \):

\[ V_{\delta u} = E \left[ \mathcal{H}^{-1}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2) H(\varphi) \mathcal{H}^{-1}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2) \right] . \]

First, we substitute a possibly asymmetric and indefinite operator \( \mathcal{H}(\varphi, \tilde{\varphi}_1, \tilde{\varphi}_2) \) by \( \mathcal{H}(\varphi) \):

\[ V_{\delta u} \approx V_{1e} = E \left[ \mathcal{H}^{-1}(\varphi) H(\varphi) \mathcal{H}^{-1}(\varphi) \right] . \]  

(47)

Next, by assuming \( H(\varphi)\mathcal{H}^{-1}(\varphi) \approx I \) we get

\[ V_{\delta u} \approx V_{2e} = E \left[ \mathcal{H}^{-1}(\varphi) \right] . \]  

(48)

Finally, by assuming \( \mathcal{H}^{-1}(\varphi) \approx H^{-1}(\varphi) \) we obtain yet another approximation

\[ V_{\delta u} \approx V_{3e} = E \left[ H^{-1}(\varphi) \right] . \]  

(49)
Numerical Example

\( \varphi(x, t) \) is governed by the 1D Burgers equation with a nonlinear viscous term:

\[
\frac{\partial \varphi}{\partial t} + \frac{1}{2} \frac{\partial (\varphi^2)}{\partial x} = \frac{\partial}{\partial x} \left( \nu(\varphi) \frac{\partial \varphi}{\partial x} \right), \tag{50}
\]

\( \varphi = \varphi(x, t), \ t \in (0, T), \ x \in (0, 1), \)

with the Neumann boundary conditions

\[
\frac{\partial \varphi}{\partial x} \bigg|_{x=0} = \frac{\partial \varphi}{\partial x} \bigg|_{x=1} = 0 \tag{51}
\]

and the viscosity coefficient

\[
\nu(\varphi) = \nu_0 + \nu_1 \left( \frac{\partial \varphi}{\partial x} \right)^2, \ \nu_0, \nu_1 = \text{const} > 0. \tag{52}
\]

Two initial conditions \( u^t = \varphi^t(x, 0) \) are considered (case A and case B).
Numerical Example: Cas A
Numerical Example: Cas B
Numerical Example: Squared Riemann distance

\[ \mu(A, B) = \left( \sum_{i=1}^{M} \ln^2 \gamma_i \right)^{1/2}. \]

<table>
<thead>
<tr>
<th>Case</th>
<th>( \mu^2(V_3, \hat{V}) )</th>
<th>( \mu^2(V_2, \hat{V}) )</th>
<th>( \mu^2(V_1, \hat{V}) )</th>
<th>( \mu^2(V_3^e, \hat{V}) )</th>
<th>( \mu^2(V_2^e, \hat{V}) )</th>
<th>( \mu^2(V_1^e, \hat{V}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>3.817</td>
<td>3.058</td>
<td>4.738</td>
<td>2.250</td>
<td>1.418</td>
<td>1.151</td>
</tr>
<tr>
<td>A2</td>
<td>17.89</td>
<td>18.06</td>
<td>21.50</td>
<td>2.535</td>
<td>1.778</td>
<td>1.602</td>
</tr>
<tr>
<td>A5</td>
<td>5.832</td>
<td>5.070</td>
<td>5.778</td>
<td>3.710</td>
<td>2.886</td>
<td>2.564</td>
</tr>
<tr>
<td>A8</td>
<td>1.133</td>
<td>0.585</td>
<td>1.419</td>
<td>1.108</td>
<td>0.466</td>
<td>0.246</td>
</tr>
<tr>
<td>A9</td>
<td>20.18</td>
<td>20.65</td>
<td>24.52</td>
<td>2.191</td>
<td>1.986</td>
<td>1.976</td>
</tr>
<tr>
<td>A10</td>
<td>10.01</td>
<td>8.521</td>
<td>8.411</td>
<td>3.200</td>
<td>2.437</td>
<td>2.428</td>
</tr>
<tr>
<td>B1</td>
<td>7.271</td>
<td>6.452</td>
<td>6.785</td>
<td>2.852</td>
<td>1.835</td>
<td>1.476</td>
</tr>
<tr>
<td>B3</td>
<td>9.937</td>
<td>10.70</td>
<td>17.70</td>
<td>4.125</td>
<td>3.636</td>
<td>3.385</td>
</tr>
<tr>
<td>B4</td>
<td>6.223</td>
<td>5.353</td>
<td>11.50</td>
<td>2.600</td>
<td>1.773</td>
<td>1.580</td>
</tr>
<tr>
<td>B5</td>
<td>10.73</td>
<td>9.515</td>
<td>9.875</td>
<td>4.752</td>
<td>3.178</td>
<td>2.530</td>
</tr>
<tr>
<td>B6</td>
<td>6.184</td>
<td>4.153</td>
<td>8.621</td>
<td>4.874</td>
<td>2.479</td>
<td>1.858</td>
</tr>
</tbody>
</table>
The reference mean deviation $\hat{\sigma}(x)$ (related to $\hat{V}$). Cases A2, A8.

The mean deviation vector $\sigma$ is defined as follows: $\sigma(i) = \sqrt{1/2}(i, i)$. 
The reference mean deviation $\hat{\sigma}(x)$ (related to $\hat{V}$). Cases B6, B9.

![Graph showing the reference mean deviation $\hat{\sigma}(x)$ for cases B6 and B9. The graph displays two curves, one for each case, with $x$ on the x-axis and $\sigma$ on the y-axis. The black solid line represents case B6, and the red dashed line represents case B9.](image-url)
Absolute errors in the correlation matrix: $\epsilon_3$, $\epsilon^e_3$, $\epsilon^e_1$ (case A2)
Absolute errors in the correlation matrix: $\epsilon_3$, $\epsilon_3^e$, $\epsilon_1^e$ (case A8)
Absolute errors in the correlation matrix: $\epsilon_3$, $\epsilon_3^e$, $\epsilon_1^e$ (case B6)
Variational and Bayesian approaches of DA lead to the minimization of the same function.

For Error Estimation we obtain two different concepts:

- Analysis Error Covariance for the Variational Approach
- Posterior Covariance for the Bayesian Approach

In the linear case the covariances coincide.

In the nonlinear case strong discrepancies can occur.

Algorithms for the estimation and the approximation of these covariances are proposed.