Erratum on "Inferring unknown unknowns: Regularized bias-aware ensemble Kalman filter"

Andrea Nóvoa[a,c], Alberto Racca[b,c,a], Luca Magri[c,d,a]

[a] Cambridge University Engineering Dept., Trumpington St, Cambridge, CB2 1PZ, UK
[b] Imperial College London, I-X Centre for AI In Science, White City Campus, 84 Wood Lane, London, W12 7SL, UK
[c] Imperial College London, Aeronautics Dept., Exhibition Road, London, SW7 2AZ, UK
[d] The Alan Turing Institute, 96 Euston Rd, London, NW1 2DB, UK

In the paper "Inferring unknown unknowns: regularized bias-aware ensemble Kalman filter. Computer Methods in Applied Mechanics and Engineering, 418, 116502. DOI: 10.1016/j.cma.2023.116502", there are small mathematical typos in Equations (15) and (16). The corrected equation (15) is

$$\boldsymbol{\psi}_j^{\mathrm{a}} = \boldsymbol{\psi}_j^{\mathrm{f}} + \mathbf{K}\left[\left(\mathbb{I} + \mathbf{J}^{\mathrm{f}}\right)^{\mathrm{T}}(\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}}) - \gamma \mathbf{C}_{dd}\mathbf{C}_{bb}^{-1}\mathbf{J}^{\mathrm{f}\,\mathrm{T}}\boldsymbol{b}_j^{\mathrm{f}}\right], \tag{1a}$$

and the corrected equation (16) is

$$\mathbf{K} = \mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\left[\mathbf{C}_{dd} + \left(\mathbb{I} + \mathbf{J}^{\mathrm{f}}\right)^{\mathrm{T}}(\mathbb{I} + \mathbf{J}^{\mathrm{f}})\mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}} + \gamma \mathbf{C}_{dd}\mathbf{C}_{bb}^{-1}\mathbf{J}^{\mathrm{f}\,\mathrm{T}}\mathbf{J}^{\mathrm{f}}\mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\right]^{-1}, \tag{1b}$$

The highlighted terms are the terms differing from the original formulation. The original equations are still exact in the case of a single observation (i.e., $N_q = 1$) in which the Jacobian is a scalar. No conclusion drawn in the paper is affected by this.

1. Derivation of the regularized-bias aware ensemble Kalman Filter

The cost function to minimize is

$$\mathcal{J}[\boldsymbol{\psi}_j] = \left(\boldsymbol{\psi}_j - \boldsymbol{\psi}_j^{\mathrm{f}}\right)^{\mathrm{T}}\mathbf{C}_{\psi\psi}^{\mathrm{f}-1}\left(\boldsymbol{\psi}_j - \boldsymbol{\psi}_j^{\mathrm{f}}\right) + \left(\boldsymbol{y}_j - \boldsymbol{d}_j\right)^{\mathrm{T}}\mathbf{C}_{dd}^{-1}\left(\boldsymbol{y}_j - \boldsymbol{d}_j\right) + \gamma \boldsymbol{b}_j^{\mathrm{T}}\mathbf{C}_{bb}^{-1}\boldsymbol{b}_j \tag{2}$$

where $\boldsymbol{y}_j = \boldsymbol{q}_j + \boldsymbol{b}_j = \mathbf{M}\boldsymbol{\psi}_j + \boldsymbol{b}_j$, and the bias is defined in the observable space, i.e., $\boldsymbol{b}_j = \boldsymbol{b}_j(\boldsymbol{q}_j)$. An analysis state $\boldsymbol{\psi}_j^{\mathrm{a}}$ minimizes the bias-aware cost function if

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}\mathcal{J}}{\mathrm{d}\boldsymbol{\psi}_j}\bigg|_{\boldsymbol{\psi}_j^{\mathrm{a}}} &= \mathbf{C}_{\psi\psi}^{\mathrm{f}-1}\left(\boldsymbol{\psi}_j^{\mathrm{a}} - \boldsymbol{\psi}_j^{\mathrm{f}}\right) + \frac{\mathrm{d}\boldsymbol{y}_j}{\mathrm{d}\boldsymbol{\psi}_j}\bigg|_{\boldsymbol{\psi}_j^{\mathrm{a}}}^{\mathrm{T}}\mathbf{C}_{dd}^{-1}\left(\boldsymbol{y}_j^{\mathrm{a}} - \boldsymbol{d}_j\right) + \gamma\frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{\psi}_j}\bigg|_{\boldsymbol{\psi}_j^{\mathrm{a}}}^{\mathrm{T}}\mathbf{C}_{bb}^{-1}\boldsymbol{b}_j^{\mathrm{a}} \\
&= \mathbf{C}_{\psi\psi}^{\mathrm{f}-1}\left(\boldsymbol{\psi}_j^{\mathrm{a}} - \boldsymbol{\psi}_j^{\mathrm{f}}\right) + \left(\mathbf{M}^{\mathrm{T}} + \frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{\psi}_j}\bigg|_{\boldsymbol{\psi}_j^{\mathrm{a}}}^{\mathrm{T}}\right)\mathbf{C}_{dd}^{-1}\left(\boldsymbol{y}_j^{\mathrm{a}} - \boldsymbol{d}_j\right) + \gamma\frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{\psi}_j}\bigg|_{\boldsymbol{\psi}_j^{\mathrm{a}}}^{\mathrm{T}}\mathbf{C}_{bb}^{-1}\boldsymbol{b}_j^{\mathrm{a}} \\
&= \mathbf{C}_{\psi\psi}^{\mathrm{f}-1}\left(\boldsymbol{\psi}_j^{\mathrm{a}} - \boldsymbol{\psi}_j^{\mathrm{f}}\right) + \mathbf{M}^{\mathrm{T}}\left\{\left(\mathbb{I} + \frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{q}_j}\bigg|_{\boldsymbol{q}_j^{\mathrm{a}}}^{\mathrm{T}}\right)\mathbf{C}_{dd}^{-1}\left(\boldsymbol{y}_j^{\mathrm{a}} - \boldsymbol{d}_j\right) + \gamma\frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{q}_j}\bigg|_{\boldsymbol{q}_j^{\mathrm{a}}}^{\mathrm{T}}\mathbf{C}_{bb}^{-1}\boldsymbol{b}_j^{\mathrm{a}}\right\} = \boldsymbol{0}
\end{aligned} \tag{3}$$

Assuming that the analysis observable state $\boldsymbol{q}_j^{\mathrm{a}}$ is sufficiently close to the forecast observable state $\boldsymbol{q}_j^{\mathrm{f}}$, we linearise the analysis bias term as

$$\boldsymbol{b}_j^{\mathrm{a}} \approx \boldsymbol{b}_j^{\mathrm{f}} + \mathbf{J}^{\mathrm{f}}\left(\boldsymbol{q}_j^{\mathrm{a}} - \boldsymbol{q}_j^{\mathrm{f}}\right), \quad \text{such that} \quad \frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{q}_j}\bigg|_{\boldsymbol{q}_j^{\mathrm{a}}} \approx \frac{\mathrm{d}\boldsymbol{b}_j}{\mathrm{d}\boldsymbol{q}_j}\bigg|_{\boldsymbol{q}_j^{\mathrm{f}}} \triangleq \mathbf{J}^{\mathrm{f}}. \tag{4}$$

Introducing (4) into (3) and grouping the terms in $\boldsymbol{\psi}_j^{\mathrm{a}}$,

$$\left\{\mathbf{C}_{\psi\psi}^{\mathrm{f}^{-1}} + \mathbf{M}^{\mathrm{T}}\mathcal{W}\mathbf{M}\right\}\boldsymbol{\psi}_j^{\mathrm{a}} = \mathbf{C}_{\psi\psi}^{\mathrm{f}^{-1}}\boldsymbol{\psi}_j^{\mathrm{f}} + \mathbf{M}^{\mathrm{T}}\mathcal{Q}, \tag{5}$$

where we have defined

$$\begin{aligned}\mathcal{W} &= \left(\mathbb{I} + \mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\right)\mathbf{C}_{dd}^{-1}\left(\mathbb{I} + \mathbf{J}^{\mathrm{f}}\right) + \gamma\mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\mathbf{C}_{bb}^{-1}\mathbf{J}^{\mathrm{f}}, \quad \text{and}\\ \mathcal{Q} &= \left(\mathbb{I} + \mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\right)\mathbf{C}_{dd}^{-1}\left(\boldsymbol{d}_j + \mathbf{J}^{\mathrm{f}}\mathbf{M}\boldsymbol{\psi}_j^{\mathrm{f}} - \boldsymbol{b}_j^{\mathrm{f}}\right) + \gamma\mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\mathbf{C}_{bb}^{-1}\left(\mathbf{J}^{\mathrm{f}}\mathbf{M}\boldsymbol{\psi}_j^{\mathrm{f}} - \boldsymbol{b}_j^{\mathrm{f}}\right)\end{aligned}$$

and making use of the Woodbury formula[1]

$$\begin{aligned}\boldsymbol{\psi}_j^{\mathrm{a}} &= \left[\mathbf{C}_{\psi\psi}^{\mathrm{f}} - \mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\left(\mathcal{W}^{-1} + \mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\right)^{-1}\mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\right]\left[\mathbf{C}_{\psi\psi}^{\mathrm{f}^{-1}}\boldsymbol{\psi}_j^{\mathrm{f}} + \mathbf{M}^{\mathrm{T}}\mathcal{Q}\right]\\ &= \boldsymbol{\psi}_j^{\mathrm{f}} + \mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\left(\mathbb{I} + \mathcal{W}\mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\right)^{-1}\left[\mathcal{Q} - \mathcal{W}\mathbf{M}\boldsymbol{\psi}_j^{\mathrm{f}}\right] \qquad (6)\\ &= \boldsymbol{\psi}_j^{\mathrm{f}} + \mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\left(\mathbf{C}_{dd} + \mathbf{C}_{dd}\mathcal{W}\mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\right)^{-1}\left[(\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}}) + \mathbf{C}_{dd}\mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\left(\mathbf{C}_{dd}^{-1}(\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}}) - \gamma\mathbf{C}_{bb}^{-1}\boldsymbol{b}_j^{\mathrm{f}}\right)\right]\\ &= \boldsymbol{\psi}_j^{\mathrm{f}} + \mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}}\left(\mathbf{C}_{dd} + \mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}} + \mathcal{V}\right)^{-1}\left[\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}} + \mathcal{U}\right], \qquad (7)\end{aligned}$$

where the model bias is encapsulated in the terms

$$\mathcal{U} = \mathbf{C}_{dd}\mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\left(\mathbf{C}_{dd}^{-1}(\boldsymbol{d}_j - \boldsymbol{y}_j^{\mathrm{f}}) - \gamma\mathbf{C}_{bb}^{-1}\boldsymbol{b}_j^{\mathrm{f}}\right) \tag{8}$$

$$\mathcal{V} = \left[\mathbf{J}^{\mathrm{f}} + \mathbf{C}_{dd}\mathbf{J}^{\mathrm{f}^{\mathrm{T}}}\mathbf{C}_{dd}^{-1}\left(\mathbb{I} + \mathbf{J}^{\mathrm{f}} + \gamma\mathbf{C}_{dd}\mathbf{C}_{bb}^{-1}\mathbf{J}^{\mathrm{f}}\right)\right]\mathbf{M}\mathbf{C}_{\psi\psi}^{\mathrm{f}}\mathbf{M}^{\mathrm{T}} \tag{9}$$

If the observation error matrix is a identity diagonal matrix (i.e., the observations are assumed uncorrelated), such that $\mathbf{C}_{dd} = \sigma_d\mathbb{I}$, then the r-EnKF equations simplify to (1).

Acknowledgement. We thank our colleague Defne E. Ozan for scrutinizing the equations.

[1] $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}$