Review of variational methods

NCEO Intensive Course on Data Assimilation

University of Reading, 2022

1 Review of Theory

Variational data assimilation methods produce MAP (maximum-a-posteriori) estimators. They use optimisation techniques to minimise cost-functions, which are often quadratic forms. The cost-function can be interpreted as an estimator of the negative logarithm of the posterior pdf. The analysis state $\mathbf{x}^a \in \mathcal{R}^{N_x}$ is the minimizer of the cost-function, i.e. $\mathbf{x}^a = \operatorname{argmin} \mathcal{J}(\mathbf{x})$. This cost-function is different for 3D-Var and 4D-Var.

1.1 3D-Var

3D-Var assimilates observations at a given time point, i.e. there is no time component in the minimisation. The 3D-Var cost-function is:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x} - \mathbf{x}^b\right)^{\mathbf{T}} \mathbf{B}^{-1} \left(\mathbf{x} - \mathbf{x}^b\right) + \frac{1}{2}\left(\mathbf{y} - h(\mathbf{x})\right)^{\mathbf{T}} \mathbf{R}^{-1} \left(\mathbf{y} - h(\mathbf{x})\right) \tag{1}$$

where $\mathbf{x}^b \in \mathcal{R}^{N_x}$ is the background estimate of the model state variables, $\mathbf{B} \in \mathcal{R}^{N_x \times N_x}$ is the background error covariance matrix, $\mathbf{y} \in \mathcal{R}^{N_y}$ is a vector of observations, $\mathbf{R} \in \mathcal{R}^{N_y \times N_y}$ is the observation error covariance matrix, and $h : \mathbf{R}^{N_x} \rightarrow \mathbf{R}^{N_y}$ is the observation operator. Here we apply the 3D-Var algorithm sequentially, which means that the model is evolved one step at a time and the observations are assimilated in order. Each time a new set of observations becomes available they are combined with the current model forecast (the background) and the cost function (1) is minimised, producing an updated analysis state. The model is then propagated forward to the time of the next observations, using the analysis as the initial state, and the assimilation process is repeated.

1.2 (Strong constraint) 4D-Var

4D-Var assimilates observations distributed over a time window, also called the assimilation window. Hence, there is a time component in the minimisation. In the strong-constraint setting (which is the one we explore), the model is considered perfect (without errors) and the problem is reduced to one of finding the *optimal* initial state $\mathbf{x}_0^a \in \mathcal{R}^{N_x}$. For observations every δ model time-steps, and an assimilation window of length K observation times, the cost-function is:

$$\begin{aligned}\mathcal{J}(\mathbf{x}_0) = &\frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right)^{\mathbf{T}} \mathbf{B}^{-1} \left(\mathbf{x}_0 - \mathbf{x}_0^b\right) \\ &+ \frac{1}{2}\sum_{k=1}^{K} \left(\mathbf{y}_k - h_k\left(m_{0\rightarrow\delta.k}(\mathbf{x}_0)\right)\right)^{\mathbf{T}} \mathbf{R}_k^{-1} \left(\mathbf{y}_k - h_k\left(m_{0\rightarrow\delta.k}(\mathbf{x}_0)\right)\right)\end{aligned} \tag{2}$$

where $\mathbf{x}_0^b \in \mathcal{R}^{N_x}$ is the background estimate for the initial state at the start of the assimilation window, and $\mathbf{B} \in \mathcal{R}^{N_x \times N_x}$ is defined as in 3D-Var. The observation term in the cost-function becomes a sum, with one term for each observation time within the assimilation window. The k^{th} observation vector is $\mathbf{y}_k \in \mathcal{R}^{N_y}$ (we keep the size constant), $h_k : \mathcal{R}^{N_x} \to \mathcal{R}^{N_y}$ is the k^{th} observation operator, $\mathbf{R}_k \in \mathcal{R}^{N_y \times N_y}$ is the observation error covariance for the k^{th} observation vector, and the model operator $m_{0 \to \delta.k} : \mathcal{R}^{N_x} \to \mathcal{R}^{N_y}$ evolves the initial state from the start of the assimilation window to the time of the k^{th} observation. In these experiments we will assume that $h_k = h$ and $\mathbf{R}_k = \mathbf{R}$ fixed.