Joint IMD-WMO Group Fellowship Training

on

Numerical Weather Prediction

04 Oct - 10 Nov 2021

Lecture Note
on
Introduction to cost function for 3-DVAR as well as for 4-DVAR data assimilation and its minimization, giving rise to analyzed field

by

Suryakanti Dutta
National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences
A-50, Sector-62
Noida-201309
Uttar Pradesh, India

1. ***Background:*** We have been introduced to a brief introduction to Data Assimilation in the lecture note on Gridpoint Statistical Interpolation (GSI) scheme & Concept of Observation Operator. Variational approaches fall under the class of *Constant Statistical Methods* of data assimilation. *3-Dimensional Variational* (3D-Var) and *4-Dimensional Variational* (4D-Var) data assimilation are the two commonly used variational approaches in NWP. For a data assimilation cycle, we require to have a background information file or first guess (x_b) (in general, which is a short-range forecast from the previous assimilation cycle) and, observation or observed variables (y_o). The first guess of observations is obtained by interpolating the NWP model first guess forecast values to the observation location and converting the model variables to observation variables (y_o) (if the observed variables are different from the model forecast variables like satellite radiances, etc). The first guess of observations are mathematically denoted as $H(x_b)$, where *'H'* is called the *Observation Operator*. The observation operator (H) performs the required transformation (model to observation variables) and interpolation to observation location or observation space as is termed in data assimilation. Once the model first guess is mapped into observation space, *'observation increments'* or *'innovations'* $\left(y_o - H(x_b)\right)$ is computed, which is the difference between the model first guess and the observations. The analysis (x_a) which is the best estimate of the current atmospheric state and serves as the initial condition for running the NWP model forecast is obtained by adding the observation increments or innovations to the NWP model first guess. Estimated statistical error covariances of the observations assimilated and the model forecast used as the first guess are applied in the form of Weights *'W'* to compute the analysis, given by:

$$x_a = x_b + W\,[y_o - H(x_b)\,] \tag{1}$$

Different data assimilation schemes are based on eq-1. Differences between the schemes arise by the difference in the approach they take in combining the observations and the background values to produce the analysis. The matrix of weights '*W*' is a function of the distance between the observations used for assimilation and the NWP model grid point. Optimal interpolation and variational methods are data assimilation schemes that come under the constant statistical approach. In Optimal Interpolation (Gandin 1963; Kalnay 2003), '*W*' is determined by minimizing the analysis errors at each grid point. In variational methods, a *cost function* is defined which is proportional to the square of the distance between the analysis and both the observations used for assimilation and the background values (Sasaki 1970; Kalnay 2003). The analysis is then computed by minimizing the cost function. The cost function in variational data assimilation is formulated as:

$$J_{var}(x) = J_b + J_o + J_c \tag{2}$$

where,

$J_{var} \Rightarrow$ The cost function

$J_b \Rightarrow$ Fit to background

$J_o \Rightarrow$ Fit to observation

$J_c \Rightarrow$ Constraint terms

2. *3D-Var:* In 3D-Variational Data Assimilation, the cost function (eq-2) is represented as:

$$J_{3DVar}(x) = \frac{1}{2}(x - x_b)^T B^{-1} (x - x_b) + \frac{1}{2}\big(y_o - H(x)\big)^T R^{-1} \big(y_o - H(x)\big) + J_c \tag{3}$$

$J_{3DVar}(x)$ ⇒ the cost function (J_{var})

$\frac{1}{2}(x - x_b)^T\, B^{-1}\, (x - x_b)$ ⇒ fit to background (J_b)

$\frac{1}{2}\big(y_o - H(x)\big)^T R^{-1}\, \big(y_o - H(x)\big)$ ⇒ fit to observation (J_o)

where,

x ⇒ analysis/control variable

x_b ⇒ background vector

B ⇒ background error covariance matrix

H ⇒ observation operator

R ⇒ observation error covariance matrix (instrument error + representative error)

y_o ⇒ observation vector

The background error covariance matrix 'B' acts as a weighting to the correction made to the background fields. It also helps to spread information spatially and among the control variables. NMC method is one of the popular methods used for the computation of the 'B' matrix. It is a static method that uses lagged pairs of forecast values (i.e. 24 and 48-hour forecasts valid at the same time), sampled over a significantly long period. This method assumes linear error growth rate and same model bias. The ensemble method is a dynamical flow-dependent method used for updating the background error covariance which uses ensemble differences of forecasts. It works on the assumption that ensembles represent actual background errors. The observation error covariance matrix 'R' includes the error incorporated due to the observation measuring equipment and the errors of representativeness. The representative errors represent the errors of the observing system seen from an NWP model point of view (Huang et al. 2002). The error covariance matrices 'B' & 'R' are typical of the order of $N^2_{NWP\_variables}$ and N^2_{obs}.

The cost function in eq-3 attains a minimum value for $x = x_a$, called the analysis, such that

$$\frac{\partial J}{\partial x} = \nabla_x J(x_a) = 0 \tag{4}$$

Assuming that the analysis is close to the truth, we can write:

$x = [x_b + (x - x_b)]$, where $(x - x_b)$ is assumed to be small.

Then we can linearize the observation operator:

$$\left(y_o - H(x)\right) = y_o - H\left(x_b + (x - x_b)\right) = \left(y_o - H(x_b)\right) - H(x - x_b) \tag{5}$$

Substituting eq-5 into eq-3, we get:

$$\begin{aligned} 2J_{3DVar}(x) = {} & (x - x_b)^T B^{-1} (x - x_b) \\ & + \left[\left(y_o - H(x_b)\right) - H(x - x_b)\right]^T R^{-1} \left[\left(y_o - H(x_b)\right) - H(x - x_b)\right] + J_c \end{aligned} \tag{6}$$

Expanding the products and ignoring the constraint term, we get:

$$\begin{aligned} 2J_{3DVar}(x) = {} & (x - x_b)^T B^{-1} (x - x_b) + (x - x_b)^T H^T R^{-1} H(x - x_b) \\ & - \left(y_o - H(x_b)\right)^T R^{-1} H(x - x_b) - (x - x_b)^T H^T R^{-1} \left(y_o - H(x_b)\right) \\ & + \left(y_o - H(x_b)\right)^T R^{-1} \left(y_o - H(x_b)\right) + J_c \end{aligned} \tag{7}$$

The cost function equation is now a quadratic function of the analysis increments $(x - x_b)$. To minimize the cost function, we need to compute the gradient of J_{3DVar}.

Combining the first two terms of eq-7, we get:

$$\begin{aligned} 2J_{3DVar}(x) = {} & (x - x_b)^T (B^{-1} + H^T R^{-1} H)(x - x_b) - \left(y_o - H(x_b)\right)^T R^{-1} H(x - x_b) \\ & - (x - x_b)^T H^T R^{-1} \left(y_o - H(x_b)\right) + (Term\ independent\ of\ x) \end{aligned} \tag{8}$$

From eq-8, we get the gradient of the cost function J_{3DVar} with respect to x.

$$\nabla J_{3DVar}(x) = (B^{-1} + H^T R^{-1} H)(x - x_b) - H^T R^{-1}\left(y_o - H(x_b)\right) \quad (9)$$

Now,

$\nabla J_{3DVar}(x) = 0$ for $x = x_a$

So, for $= x_a$, eq-9 can be written as:

$$(B^{-1} + H^T R^{-1} H)(x - x_b) - H^T R^{-1}\left(y_o - H(x_b)\right) = 0 \quad (10)$$

$$\Rightarrow (B^{-1} + H^T R^{-1} H)(x_a - x_b) = H^T R^{-1}\left(y_o - H(x_b)\right)$$

$$\Rightarrow x_a = x_b + (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}\left(y_o - H(x_b)\right)$$

So, the solution of the 3-Dimensional Variational Analysis problem is:

$$\boxed{x_a = x_b + (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}\left(y_o - H(x_b)\right)} \quad (11)$$

Comparing eq-11 with eq-1 gives the weight matrix for 3D-Var as:

$$W = (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} \quad (12)$$

The solution of x_a involves in practice, minimization algorithms for $J_{3DVar}(x)$ using iterative methods for minimization such as conjugate gradient method or quasi-Newton method.

3. ***4D-Var:*** In 4D-Variational Data Assimilation, the cost function (eq-2) is represented as:

$$J_{4DVar}(x_0) = \frac{1}{2}(x_0 - x_b)^T B^{-1} (x_0 - x_b) + \frac{1}{2}\sum_{i=0}^{N} \left(y_i - H\big(M_i(x_0)\big)\right)^T R_i^{-1} \left(y_i - H\big(M_i(x_0)\big)\right) + J_c \quad (13)$$

where,

$subscript\ '0'$ ⇒ at time t_0

$subscript\ 'i'$ ⇒ at time t_i

N ⇒ Number of Observational Vectors y_i distributed over time.

In 3D-Var, the observations for assimilation are considered within a single time window centred on the analysis time. For each observational site, the assimilation system will select the observation closest in time to the analysis time. Other observations from the site, if present is rejected. Whereas in 4D-Var, the observations are binned into time slots 'i'. $'M_i'$ is the NWP model operator which converts the state vector x_0 into its forecast values at time 'i'. This enables the use of observations from the same site obtained at different times. Also, it minimizes the offset in time between the time of the observations and the valid time for the forecast fields against which the observations are compared (Huang et al. 2002).

The atmospheric flow is governed by several dynamical and physical laws. The NWP model can be symbolically written as: $x_{i+1} = M_{i+1,i}(x_i)$ (14)

where $M_{i+1,i}$ is the non-linear NWP model from time $t_i\ to\ t_{i+1}$. A perturbation of the atmospheric state is evolved by the tangent linear model: $\delta x_{i+1} = \mathbf{M_{i+1,i}}\,(x_i)\,\delta x_i$ (15)

Substituting eq-14 in eq-13, the gradient of $J_{4DVar}(x_0)$ with respect to x_0 becomes:

$$\nabla_{x_0} J_{4DVar}(x_0) = B^{-1}(x_0 - x_b) + \sum_{i=0}^{N} \mathbf{M}_{i,0}^T \mathbf{H}_i^T R_i^{-1} \left(y_i - H(\boldsymbol{x_i})\right) = 0 \quad (16)$$

where,

$\mathbf{H_i}$ ⇒ the tangent linear operator of the observation operator H_i ($\mathbf{H_i^T}$ is the corresponding adjoint operator and is simply the complex-conjugate transpose of the tangent linear by definition)

$\mathbf{M}_{i,0}^{\mathrm{T}} = \mathbf{M}_{1,0}^T \mathbf{M}_{2,1}^T \ldots \mathbf{M}_{i,i-1}^T$ ⇒ the adjoint model and is a backward integration from time $t_i \; to \; t_0$

In 4D-variational data assimilation too, the gradient of the cost function (eq-16) is solved iteratively. In the first step of the iteration, moving forward in time the first guess trajectory and the observation departures or innovations $\left(y_i - H(\boldsymbol{x_i})\right)$ are computed by integrating the non-linear model in eq-14. Using the first guess trajectory, the adjoint model is integrated backwards in time and, the observation forcing $\mathbf{H}_i^T R_i^{-1} \left(y_i - H(\boldsymbol{x_i})\right)$ is added to the adjoint variable computed in the previous step. The integrated final value of the adjoint variable plus the background term gradient $B^{-1}(x_0 - x_b)$ is the gradient ∇J_{4DVar} of the cost function with respect to the control variable x_0 for the present iteration step. The first guess is updated with the computed gradient of the cost function. These steps are iterated starting with the updated first guess value till the convergence criteria are fulfilled.

References

Gandin, L. S., 1963: Objective analysis of meteorological fields. Gidrometerologicheskoe Izdatelstvo, Leningrad. English translation by Israeli Program for Scientific Translations, Jerusalem, 1965.

Huang, Xiang-Yu & Henrik Vedel, 2002: An Introduction to Data Assimilation.

Kalnay, E., 2003: Atmospheric modeling, data assimilation and predictability; Cambridge University Press

Rabier F. and Z. Liu, 2003: Variational Data Assimilation: theory and overview.

Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. Mon. Wea. Rev. 98, 875-883.