Data and Uncertainty - Data assimilation

Dr Amos S. Lawless (a.s.lawless@reading.ac.uk)

1. Suppose that we have an observation T_o of temperature in a room, with error variance σ_o^2, and we also have a background estimate of the temperature T_b, with error variance σ_b^2. Assume that the background and observation errors are unbiased and uncorrelated.

   Let the analysed temperature T_a be given by

   $$T_a = \alpha T_b + (1-\alpha)T_o \tag{1}$$

   with

   $$\alpha = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2}.$$

   (a) Show that the analysed temperature T_a is unbiased.

   [5 marks]

   (b) Show that the variance of the analysis error σ_a^2 is given by

   $$\sigma_a^2 = \frac{\sigma_b^2 \sigma_o^2}{\sigma_b^2 + \sigma_o^2}. \tag{2}$$

   [6 marks]

   (c) A big uncertainty in assimilation schemes is the specification of the background error. Suppose that in the coefficients of (1) the background error variance σ_b^2 is incorrectly specified by the value $\tilde{\sigma}_b^2$. Find an expression for the true variance of the analysis error σ_a^2 when this value is used. Show that this will be greater than the perceived analysis error variance, obtained by substituting $\tilde{\sigma}_b^2$ into (2), for values of $\tilde{\sigma}_b^2$ less than the true background error variance σ_b^2.

   [9 marks]

2. Let the model state vector $\mathbf{x}$ consist of two variables r, θ defined at a single spatial point, so that $\mathbf{x} = (r, \theta)^T$. Suppose that we have a background field $\mathbf{x}_b = (r_b, \theta_b)^T$ with background error covariance matrix given by

   $$\mathbf{B} = \begin{pmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{pmatrix}.$$

   Suppose that we have two observations: y_1, which is an observation of $r \sin\theta$, and y_2, which is an observation of $r\cos\theta$, each with error variance σ_o^2, that we wish to assimilate using a 3D-Var algorithm.

   [*Recall that 3D-Var is the same as 4D-Var, but without the time dimension*].

   (a) Write down a cost function for this problem, defining any symbols that you use.

   [6 marks]

   (b) Calculate the gradient of the cost function at the point $\mathbf{x} = (1, \pi/2)^T$.

   [14 marks]

3. Suppose that we have a model of temperature T at a single grid point, with

$$T^{n+\Delta t} = \alpha \Delta t T^n,$$

where n is the time level, Δt is the model time step and α is a scalar constant. Suppose further that we have a background temperature T_b at time t_b, with error variance σ_b^2, and a single observation T_o at time $t_b + k\Delta t$ where k is a positive integer, with error variance σ_o^2. We run a 4D-Var assimilation to convergence over the time window $[t_b, t_b + k\Delta t]$.

(a) Write down the cost function for this problem. By minimising this function find the value of the analysed temperature T_a at the start of the time window.

(b) Suppose now that we have two observations T_1, T_2 valid at times $t_b + \Delta t$, $t_b + 2\Delta t$ respectively, both with observation error variance σ_o^2. What is the gradient of the observation part of the cost function on the first iteration, assuming that the first guess is equal to the background?

4. Suppose we have a model state $\mathbf{x} = (x, y)^T$ whose evolution is described by the equations

$$\begin{aligned} x_{k+1} &= \alpha x_k \\ y_{k+1} &= x_k + y_k, \end{aligned}$$

where the subscript k indicates the time level. Suppose further that we have a background field of $\mathbf{x}$ at time t_0 given by $\mathbf{x}_f = (x_b, y_b)^T$, with error covariance matrix

$$\mathbf{P}_f(t_0) = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$$

and assume we have observations $\tilde{x}_k, \tilde{y}_k$ of x and y at every time level t_k and that all observations have constant error variance σ_o^2. We analyse the observations sequentially using a Kalman filter.

(a) Calculate the analysis at the initial time t_0.

[11 marks]

(b) Find the analysis error covariance matrix associated with the analysis at time t_0.

[4 marks]

(c) Using your answer from part (b), and assuming that the model is perfect, show that the forecast error covariance matrix at time t_1 is given by

$$\mathbf{P}_f(t_1) = \begin{pmatrix} \alpha^2 \frac{\sigma_x^2 \sigma_o^2}{\sigma_x^2 + \sigma_o^2} & \alpha \frac{\sigma_x^2 \sigma_o^2}{\sigma_x^2 + \sigma_o^2} \\ \alpha \frac{\sigma_x^2 \sigma_o^2}{\sigma_x^2 + \sigma_o^2} & \frac{\sigma_x^2 \sigma_o^2}{\sigma_x^2 + \sigma_o^2} + \frac{\sigma_y^2 \sigma_o^2}{\sigma_y^2 + \sigma_o^2} \end{pmatrix}.$$

[5 marks]

5. Suppose we have a model state $\mathbf{x} = (u, v)^T$ whose evolution is described by the equations

$$\begin{aligned} u_{k+1} &= 2u_k + 2v_k \\ v_{k+1} &= 2u_k + v_k, \end{aligned}$$

where the subscript indicates the time level. Suppose further that we have a background field of $\mathbf{x}$ at time t_0 given by $\mathbf{x}_0^b = (u_b, v_b)^T$, with error covariance matrix

$$\mathbf{B} = \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}.$$

Assume that we have observations y_0, y_1 of the quantity u^2 at times t_0, t_1 repectively, each with error variance σ_o^2, that we wish to assimilate using the method of 4D-Var.

(a) Write down the 4D-Var cost function for this problem, defining any symbols that you use.

(b) Using the method of Lagrange multipliers, find the gradient of the cost function with respect to the initial condition $\mathbf{x}_0$.