METR 5303 – Lecture #12, #13

Statistical Objective Analysis

Reading: Daley, p. 101-106

We shall follow Sec. 4.2 in Daley for the derivation of the Statistical Objective Analysis (SOA) technique, which is obtained via the minimum variance approach. In Sec. 4.1, Daley returns to the one-dimensional approach used in Sec. 2.2 (our Lectures #3 & 4) to show that the minimum variance estimate [eq. (12), Lecture 3] has an expected analysis error variance that is less than the smallest expected observational error variance. Thus each observation, whatever its error, always reduces the expected error variance of the analysis. This indicates that all observations have some usefulness.

Derivation of SOA Algorithm

We start with the basic form of the analysis equation used in the SCM lectures:

$$f_A(\mathbf{r}_i) = f_B(\mathbf{r}_i) + \sum_{k=1}^{K} W_{ik}[\, f_o(\mathbf{r}_k) - f_B(\mathbf{r}_k)\,] \qquad (1)$$

where $f_A(\mathbf{r}_i)$ is the analysis value at grid point i,

$f_B(\mathbf{r}_i)$ is the background field value at grid point i,

$f_o(\mathbf{r}_k)$ is the observation at the station,

$f_B(\mathbf{r}_k)$ is the background value at the station and

W_{ik} is the as yet undetermined *a posteriori* weight.

It is assumed that the points $\mathbf{r}_k$ belong to locations of irregularly-spaced observations, and points $\mathbf{r}_i$ are locations of regularly-spaced grid points.

We now introduce simpler notation: Define

$$\begin{aligned} A_i &= f_A(\mathbf{r}_i) \\ B_i &= f_B(\mathbf{r}_i) \\ O_k &= f_o(\mathbf{r}_k) \\ B_k &= f_B(\mathbf{r}_k) \end{aligned}$$

Note that all of these quantities have error. Thus eq. (1) becomes

$$A_i = B_i + \sum_{k=1}^{K} W_{ik} [\, O_k - B_k \,] \qquad (2)$$

Recall $A_i - B_i$ is the analysis increment, and $O_k - B_k$ is the observation increment.

Now define T_i as the true value of *f* at the grid points, and

T_k as the true value of *f* at the stations. Neither T_i or T_k are known.

[Note that T_i , T_k consist only of truth that can be resolved by the observing network; i.e., the "smooth truth"; see discussion in text]

Now subtract T_i from both sides of eq. (2):

$$A_i - T_i = B_i - T_i + \sum_{k=1}^{K} W_{ik} [O_k - B_k] \qquad (3)$$

Assume background and observational errors are unbiased:

$$< B_i - T_i > = < B_k - T_k > = < O_k - T_k > = 0 \qquad (4)$$

Thus the observation increment $< O_k - B_k >$ will be unbiased, and, via (3), $< A_i - T_i > = 0$ as well (analysis is unbiased).

Note that $< T >$ is the "true climate", and, in general, since the background usually comes from a forecast model, which may have bias, and since observations may have bias, then

$$< B > \neq < T > \quad \text{and} \quad < O > \neq < T >.$$

However, since the bias is usually known, and we have learned how to remove bias (Lecture #4), we will assume that B_i , B_k and O_k are unbiased.

Now we can proceed with the derivation. The goal is to minimize the **expected analysis error variance**

$E_A^2 = < (A_i - T_i)^2 >$, which we need to obtain. Thus, we

square both sides of eq. (3):

$$(A_i - T_i)^2 = [(B_i - T_i) + \Sigma W_{ik} (O_k - B_k)]^2$$

$$= (B_i - T_i)^2 + 2 \sum_{k=1}^{K} W_{ik} (O_k - B_k) (B_i - T_i) + [\sum_{k=1}^{K} W_{ik} (O_k - B_k)]^2$$

We now take expected values of this:

$$< (A_i - T_i)^2 > = < (B_i - T_i)^2 > + 2 \sum_{k=1}^{K} W_{ik} < (O_k - B_k) (B_i - T_i) >$$

$$+ \sum_{k=1}^{K} \sum_{l=1}^{K} W_{ik} W_{il} < (O_k - B_k) (O_l - B_l) > \qquad (5)$$

We define $E_A^2 = < (A_i - T_i)^2 >$ as the <u>expected analysis error variance</u>

and $E_B^2 = < (B_i - T_i)^2 >$ as the <u>expected background error variance</u>

Recall that the <u>covariance</u> between two random variables s, q is

$$\text{Cov}(s, q) = < [s - <s>] [q - <q>] >$$

And that $\text{Cov}(s, s) = \text{Var}\, s = \sigma_s^2$.

Then $\text{Cov} [(O_k - B_k), (O_l - B_l)] = < [(O_k - B_k) - < (O_k - B_k) >] [(O_l - B_l) - < (O_l - B_l) >] >$.

But since $< (O_k - B_k) > = <(O_l - B_l) > = 0$ (unbiased), $(< (O_k - B_k) >=< (O_k – T_k) – (B_k - T_k) >=0)$

then $< (O_k - B_k) (O_l - B_l) >$ is the covariance between observation increments at station points $\mathbf{r}_k$ and $\mathbf{r}_l$.

Similarly, $< (O_k - B_k) (B_i - T_i) >$ is the covariance between the background error at i and the observation increment at k.

So, to minimize E_A^2, that is, to determine the weights W_{ik} that minimize eq. (5), we differentiate (5) with respect to each of the weights W_{ik} :

$$\frac{\partial E_A^2}{\partial W_{ik}} = 2 < (O_k - B_k) (B_i - T_i) > + 2 \sum_{l=1}^{K} W_{il} < (O_k - B_k) (O_l - B_l) > \Rightarrow 0 \qquad (6)$$

or

$$\sum_{l=1}^{K} W_{il} < (O_k - B_k) (O_l - B_l) > = - < (O_k - B_k) (B_i - T_i) > \qquad (7)$$

Consider the RHS of eq. (7), which can be rewritten (**class exercise**) (hint: add and subtract $T_k (B_i - T_i)$)

$$< (O_k - B_k) (B_i - T_i) > = < (O_k - T_k) (B_i - T_i) > - < (B_k - T_k) (B_i - T_i) > \qquad (8)$$

Terms of the form $< (O_m - T_m) (B_n - T_n) >$ (i.e. – the covariance between observational error and the background error) might be expected to be zero (basically assuming that the background error and the observation error are uncorrelated, which is usually a valid assumption except when background fields are used to obtain observed fields such as in satellite retrievals). Here we will assume they are zero, and so the 2nd term on the RHS of (8) replaces the RHS of (7):

$$\sum_{l=1}^{K} W_{il} < (O_k - B_k) (O_l - B_l) > = < (B_k - T_k) (B_i - T_i) > \qquad (9)$$

Now consider the covariance term on the LHS of (9). We first add two terms assumed to be zero:

$$< (O_k - B_k) (O_l - B_l) > + < (O_k - T_k) (B_l - T_l) > + < (B_k - T_k) (O_l - T_l) >$$

Since the expectation operator is commutative, we can put all terms inside one < > operator, carry out all products, and then regroup the remaining terms to obtain

$$= < (O_k - B_k) (O_l - B_l) + (O_k - T_k) (B_l - T_l) + (B_k - T_k) (O_l - T_l) >$$

$$= < (O_k O_l - O_k B_l - B_k O_l + B_k B_l + O_k B_l - O_k T_l - T_k B_l + T_k T_l + B_k O_l - B_k T_l - T_k O_l + T_k T_l >$$

$$= < (B_k B_l - B_k T_l + T_k T_l - T_k B_l) + (O_k O_l - O_k T_l - T_k O_l + T_k T_l) - O_k B_l - B_k O_l + O_k B_l + B_k O_l >$$

$$= < (B_k - T_k) (B_l - T_l) > + < (O_k - T_k) (O_l - T_l) > \qquad (10)$$

Note that the first term on the RHS of (10) is the background error covariance term and the 2nd term is the observational error covariance term. Using this result in eq. (9), we have

$$\sum_{l=1}^{K} W_{il} [< (B_k - T_k) (B_l - T_l) > + < (O_k - T_k) (O_l - T_l) >] = < (B_k - T_k) (B_i - T_i) > \qquad (11)$$

Therefore, the statistical objective analysis method is the use of eq. (2) with the weights determined by eq. (11).

We now put eq. (11) in matrix form, using the following symbols:

$\mathbf{f}_o$ - a column vector of length K of the observations $f_o(\mathbf{r}_k)$

$\mathbf{f}_B$ - a column vector of length K of background values $f_B(\mathbf{r}_k)$ at stations

$\mathbf{W}_i$ - a column vector of length K of *a posterori* weights (one for each grid point)

$\mathbf{B}_i$ - a column vector of length K of the covariance between background error at stations and at grid points

$\mathbf{B}$ - background error covariance matrix with elements $< \varepsilon_B(\mathbf{r}_k)\ \varepsilon_B(\mathbf{r}_l) >$

$\mathbf{O}$ - observational error covariance matrix with elements $< \varepsilon_o(\mathbf{r}_k)\ \varepsilon_o(\mathbf{r}_l) >$

Both $\mathbf{B}$ and $\mathbf{O}$ are symmetric K x K matrices. Thus eq. (2) becomes

$$f_A(\mathbf{r}_i) = f_B(\mathbf{r}_i) + \mathbf{W}_i^T [\, \mathbf{f}_o - \mathbf{f}_B \,] \qquad (12)$$

and eq. (11) is

$$[\, \mathbf{B} + \mathbf{O} \,]\, \mathbf{W}_i = \mathbf{B}_i \rightarrow \mathbf{W}_i = [\, \mathbf{B} + \mathbf{O} \,]^{-1}\, \mathbf{B}_i \rightarrow \mathbf{W}_i^T = \mathbf{B}_i^T [\, \mathbf{B} + \mathbf{O} \,]^{-1} \qquad (13)$$

Thus $\mathbf{W}_i$ are the <u>optimal weights</u> and the use of (13) into (10) gives us the optimal analysis equation:

$$f_A(\mathbf{r}_i) = f_B(\mathbf{r}_i) + \mathbf{B}_i^T [\, \mathbf{B} + \mathbf{O} \,]^{-1} [\, \mathbf{f}_o - \mathbf{f}_B \,] \qquad (14)$$

which is the same as the optimal analysis equation we obtained at the end of Lecture #11.

Now that we have a statistical objective analysis algorithm, eq. (14), in which $f_A(\mathbf{r}_i)$, provides a linear unbiased estimate of $f_T(\mathbf{r}_i)$, we should also formulate an expression for the expected analysis error variance, E_A^2 . This requires manipulation of our previous equation for E_A^2, eq. (5) given earlier. The result, without derivation is

$$E_A^2 = E_B^2 - \mathbf{W}_i^T \mathbf{B}_i = E_B^2 - \mathbf{B}_i^T [\, \mathbf{B} + \mathbf{O} \,]^{-1} \mathbf{B}_i \qquad (15)$$

All terms on the RHS of (15) are known from quantities used in the analysis procedure. Thus the expected analysis error variance is the background error variance minus a positive term proportional to the weights used on the observations - which means that the analysis error is smaller than the background error in regions where there are observations.

In summary, eq. (12) is the minimum variance estimate of $f_T(\mathbf{r}_i)$, and if the optimal weights given by eq. (13) are used for W_i , then the system (12) and (13) is called optimal interpolation *provided* $\mathbf{B}$, $\mathbf{O}$, $\mathbf{B}_i$ *are correct.* Since these are not perfectly known, Daley prefers the term statistical interpolation or SOA. The estimation of $\mathbf{B}$, $\mathbf{O}$, $\mathbf{B}_i$ is still a subject of research, especially for mesoscale and storm-scale data assimilation.

Normalized form of the SOA Equations

We now obtain a non-dimensional (normalized) version of the SOA analysis equations. We will need to introduce some new notation.

Define $\boldsymbol{\sigma}_B = < (B_k - T_k)^2 >^{1/2}$ to be a diagonal matrix of background error standard deviations at observation stations. We also define a new set of normalized weights $\boldsymbol{\omega}_i$ via

$$\mathbf{W}_i^T = E_B \, \boldsymbol{\omega}_i^T \boldsymbol{\sigma}_B^{-1} \;, \quad \text{where } E_B = < \varepsilon^2_B(\mathbf{r}_i) >^{1/2} = < (B_i - T_i)^2 >^{1/2}$$

is the background error standard deviation at the grid points.

If we use this expression in eq. (12), we have

$$f_A(\mathbf{r}_i) = f_B(\mathbf{r}_i) + E_B \, \boldsymbol{\omega}_i^T \boldsymbol{\sigma}_B^{-1} [\, \mathbf{f}_O - \mathbf{f}_B \,] \qquad (16)$$

Dividing (16) by $E_B = < (B_i - T_i)^2 >^{1/2}$ and reverting back to summation notion, we can rewrite (16) as

$$\frac{A_i - B_i}{< (B_i - T_i)^2 >^{1/2}} = \sum_{k=1}^{K} \frac{\omega_{ik} \left[O_k - B_k \right]}{< (B_k - T_k)^2 >^{1/2}} \qquad (17)$$

This is a normalized (non-dimensional) version of eq. (12). We now wish to rewrite the weight equation (13) using our normalized weights ω_i :

$$[\mathbf{B} + \mathbf{O}] E_B \, \omega_i^T \, \sigma_B^{-1} = \mathbf{B}_i$$

Now left multiply by σ_B^{-1} and divide by E_B to get

$$\sigma_B^{-1} [\mathbf{B} + \mathbf{O}] \, \omega_i^T \, \sigma_B^{-1} = E_B^{-1} \, \sigma_B^{-1} \, \mathbf{B}_i$$

Interchanging ω_i^T and σ_B^{-1} (because σ_B^{-1} is a diagonal matrix) and bringing σ_B^{-1} inside bracket, we have

$$[\sigma_B^{-1} \mathbf{B} \, \sigma_B^{-1} + \sigma_B^{-1} \mathbf{O} \, \sigma_B^{-1}] \, \omega_i = E_B^{-1} \, \sigma_B^{-1} \, \mathbf{B}_i \qquad (18)$$

Note that $\sigma_B^{-1} \mathbf{B} \, \sigma_B^{-1}$ is the definition of the <u>background error correlation matrix</u> at stations, denoted ρ_B ($-1 \leq \rho_B \leq 1$), and written in summation form as

$$\sigma_B^{-1} \mathbf{B} \, \sigma_B^{-1} = \frac{< (B_k - T_k) (B_l - T_l) >}{[< (B_k - T_k)^2 > < (B_l - T_l)^2 >]^{1/2}} = \rho_B$$

Similarly,

$$E_B^{-1}\,\boldsymbol{\sigma}_B^{-1}\,\mathbf{B}_i = \frac{< B_k - T_k)\,(B_i - T_i) >}{[< (B_k - T_k)^2 > < (B_l - T_l)^2 >]^{1/2}} = \boldsymbol{\rho}_{iB}$$

is the background error correlation vector (between stations and grid point $\mathbf{r}_i$).

Therefore, we can now write eq. (18) as

$$[\,\boldsymbol{\rho}_B + \boldsymbol{\sigma}_B^{-1}\,\mathbf{O}\,\boldsymbol{\sigma}_B^{-1}\,]\,\boldsymbol{\omega}_i = \boldsymbol{\rho}_{iB} \qquad (19)$$

We will use a version of eq. (19) later for simple analysis examples.

The normalized equation for the expected analysis error variance is

$$\varepsilon^2_A = 1 - \boldsymbol{\omega}_i^T\,\boldsymbol{\rho}_{iB} \quad \text{where} \quad \varepsilon^2_A = E_A^2 / E_B^2$$

So now our normalized SOA algorithm is eq. (16) or (17) written as

$$\frac{f_A(\mathbf{r}_i) - f_B(\mathbf{r}_i)}{E_B} = \frac{\boldsymbol{\omega}_i^T\,[\,\mathbf{f}_o - \mathbf{f}_B\,]}{\boldsymbol{\sigma}_B}$$ with weights given by eq. (19).

Conversion between Daley and Kalnay Notation

In Sec. 5.6 of Daley's book, he summarizes some of the limitations of the SOA algorithm as presented in his Chap. 4 and 5. Many of these limitations are valid today, but some can be addressed by provided the proper formulation of $\underline{f}_B$, the values of the background field at the stations, which, up to now, have been assumed to be obtained

simply by interpolation from surrounding grid points. However, suppose the observations are not the same variable as the ones on the grid (e.g., radiances instead of temperature), then we need to introduce a linear operator that acts on the model variables to both convert those variables to the observed variables and interpolates to the observation locations. In Daley, the symbol Ω is used to denote this operator, whereas Kalnay uses the conventionally accepted symbol H. Thus the analysis equation (12) becomes

$$f_A(\mathbf{r}_i) = f_B(\mathbf{r}_i) + \mathbf{W}_i^T [\, \mathbf{f}_o - \mathbf{\Omega}\, \mathbf{f}_{\underline{B}} \,] \qquad (1)$$

which is the same as eq. 5.4.1 in Kalnay:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{W} [\, \mathbf{y}_o - H(\mathbf{x}_b) \,] \qquad (2)$$

where $\mathbf{x}_a$ is the maximum likelihood analysis on the grid, $\mathbf{x}_b$ is the background field on the grid, and $\mathbf{y}_o$ contains the observations at the stations. The $H(\mathbf{x}_b)$ term operates on the background fields (H could simply be interpolation formulas or the equations of radiative transfer) to produce background estimates at the stations.

Daley then repeats the SOA derivation for a new expression for $\mathbf{W}_i^T$, which expresses the weights associated with the analysis at the i^{th} grid point. The analysis weights for _all_ analysis grid points produces a (K x I) rectangular matrix **W** that replaces eq. (13) and is written

$$\mathbf{W} = [\mathbf{O} + \mathbf{\Omega}\, \mathbf{B}\, \mathbf{\Omega}^T]^{-1}\, \mathbf{\Omega}\, \mathbf{B} \qquad (3)$$

which is equivalent to eq. 5.4.19a in Kalnay:

$$\mathbf{W} = \mathbf{B}\, \mathbf{H}^T [\mathbf{R} + \mathbf{H}\, \mathbf{B}\, \mathbf{H}^T]^{-1} \qquad (4)$$

where **R** is the observation error covariance matrix.