Iteration of the SCM to the Optimal Solution

Reading: Daley, p. 93-96

We have seen that the successive correction method (SCM) using the Barnes weight function has the following features:

a. No background field is needed (although could use one if available)
b. We can pre-determine the amount of detail present in the analysis as a function of the Gaussian filter parameter κ and station spacing Δn.
c. No need for an influence radius except to save computer time
d. Only 2 passes or scans needed (although perhaps not the best procedure)
e. Time-weighting of observations is possible

However, the following problems remain:

a. The method works poorly in both data cluster areas and data sparse regions
b. Treats all data the same.
c. (Here) makes no use of background information

The text and Barnes references have shown that SCM algorithms converge to the observations at the obs. stations. This is desired only if the data are perfect and there is no background field - a situation that never occurs. Thus, with imperfect obs. and a background field (also imperfect), an "optimal analysis" should be a linear combination of both. In the following, we will use the equations and techniques learned so far to derive an equation for this optimal analysis.

Recall the iteration SCM equation (17) from Lecture 8:

$$f_A^{j+1}(\mathbf{r}_i) = f_A^{j}(\mathbf{r}_i) + \underline{W}_i^T [\, \underline{f}_o - \underline{f}_A^{\,j} \,] \tag{1}$$

(this is the scheme that converges to the observations)

where

$$\underline{W}_i^T = \frac{\sum_{k=1}^{K_i} w(r_{ik})}{\sum_{k=1}^{K_i} w(r_{ik}) + \varepsilon_o^2} = \frac{\sum_{k=1}^{K_i} E_b^2 w(r_{ik})}{\sum_{k=1}^{K_i} E_b^2 w(r_{ik}) + E_o^2} \tag{2}$$

since $\varepsilon_o^2 = E_o^2 / E_b^2$, and E_b is not a function of k.

Now multiply (2) by E_o^{-2}/E_o^{-2} and substitute result in Eq. (1):

$$f_A^{j+1}(\mathbf{r}_i) \; - \; f_A^{j}(\mathbf{r}_i) \;=\; \frac{\sum_{k=1}^{K_i} E_b^2 w\,(r_{ik})\, E_o^{-2}\, [\underline{f}_o(\mathbf{r}_k) \; - \underline{f}_A^{j}(\mathbf{r}_k)\,]}{1 \;+\; \sum_{k=1}^{K_i} E_b^2 w\,(r_{ik})\, E_o^{-2}}$$

or

$$f_A^{j+1}(\mathbf{r}_i) \; - \; f_A^{j}(\mathbf{r}_i) \;=\; (1 \;+\; q_i)^{-1} \sum_{k=1}^{K_i} E_b^2\, w\,(r_{ik})\, E_o^{-2}\, [\underline{f}_o(\mathbf{r}_k) \; - \underline{f}_A^{j}(\mathbf{r}_k)\,] \tag{3}$$

where $$q_i \;=\; \sum_{k=1}^{K_i} E_b^2 w\,(r_{ik})\, E_o^{-2}$$

Note that $f_A^{j=0}(\mathbf{r}_i) \;=\; f_B(\mathbf{r}_i)$: the initial background field

$f_A^{j=0}(\mathbf{r}_k) \;=\; f_B(\mathbf{r}_k)$: forward interpolated background value at station

Eq. (3) is still equivalent to eq. (1). Now we are going to modify (3) in **3** significant ways to produce the "optimal" analysis equation. **First**, we add a "fit to background" term:

$$f_A^{j+1}(\mathbf{r}_i) \; - \; f_A^{j}(\mathbf{r}_i) \;=\; \text{RHS of (3)} \;+\; (1 \;+\; q_i)^{-1}\, [f_B(\mathbf{r}_i) \; - \; f_A^{j}(\mathbf{r}_i)\,] \tag{4}$$

Thus the analysis is now a weighted combination of observed and background values. See Fig. 3.10 in Daley for an example of this (when $\varepsilon_o^2 \;=\; 1$).

The **second** step is to specify the weights $w(r_{ik})$; assume they are given by

$$w\,(r_{ik}) \;=\; E_b^{-2} < \varepsilon_B(\mathbf{r}_k)\, \varepsilon_B(\mathbf{r}_i) >$$

where $\varepsilon_B(\mathbf{r}_k)$, $\varepsilon_B(\mathbf{r}_i)$ are background errors at $\mathbf{r}_k$, $\mathbf{r}_i$ resp.

That is, we are assuming that the weight between station k and grid point i is given by the <u>background error covariance normalized by the expected background error variance.</u>

We will now put this back into full matrix form.

Define $\underline{\underline{Q}}$ as a diagonal obs. error covariance matrix with elements $< \varepsilon_o(\mathbf{r}_k)\, \varepsilon_o(\mathbf{r}_i) >$

Define $\underline{B}_i$ as a background error column vector of length K (no. of stations) with elements $< \varepsilon_B(\mathbf{r}_k)\, \varepsilon_B(\mathbf{r}_i) >$.

Now rewrite eq. (4) as

$$f_A^{j+1}(\mathbf{r}_i) - f_A^{j}(\mathbf{r}_i) = (1 + q_i)^{-1} \{ \underline{B}_i^T\, \underline{\underline{Q}}^{-1} [\, \underline{f}_o - \underline{f}_A^{j}] + [f_B - f_A^{j}] \} \quad (5)$$

Now rewrite eq. (5) for the case where the analysis is performed at the observation stations:

$$\underline{f}_A^{j+1} - \underline{f}_A^{j} = (\underline{\underline{I}} + \underline{\underline{Q}})^{-1} \{ \underline{\underline{B}}\, \underline{\underline{Q}}^{-1} [\, \underline{f}_o - \underline{f}_A^{j}] + [\underline{f}_B - \underline{f}_A^{j}] \} \qquad (6)$$

where $\underline{\underline{B}}$ is a background error covariance matrix with elements

$$b_{kl} = < \varepsilon_B(\mathbf{r}_k)\, \varepsilon_B(\mathbf{r}_l) > \quad \text{where } \mathbf{r}_k \text{ and } \mathbf{r}_l \text{ are observation locations.}$$

and $\underline{\underline{Q}}$ is a diagonal matrix with elements $q_k = \sum_{l=1}^{K} b_{kl}\, E_o^{-2}$

In the limit, if eq. (6) converges, $\underline{f}_A^{j} = \underline{f}_A^{j+1} = \underline{f}_A^{\infty}$ and

$$\underline{f}_A^{\infty} = \underline{f}_B + \underline{\underline{B}} [\, \underline{\underline{B}} + \underline{\underline{Q}}]^{-1} [\, \underline{f}_o - \underline{f}_B] \qquad (7)$$

Deriving (7) is an optional class exercise, since the convergence proof follows the discussion in Sec. 3.5 in Daley. Recall that eq. (7) is an analysis equation valid at observation locations, and is the same as eq. (6) in Lecture 5.

Now go back to eq. (5) (an analysis valid at grid points), and we also have in the limit if the analysis converges

$$f_A^{j}(\mathbf{r}_i) = f_A^{j+1}(\mathbf{r}_i) = f_A^{\infty}(\mathbf{r}_i) \quad \text{and eq. (5) becomes}$$

$$f_A^{\infty}(\mathbf{r}_i) = f_B(\mathbf{r}_i) + \underline{B}_i^T\, \underline{\underline{Q}}^{-1} [\, \underline{f}_o - \underline{f}_A^{\infty}] \qquad (8)$$

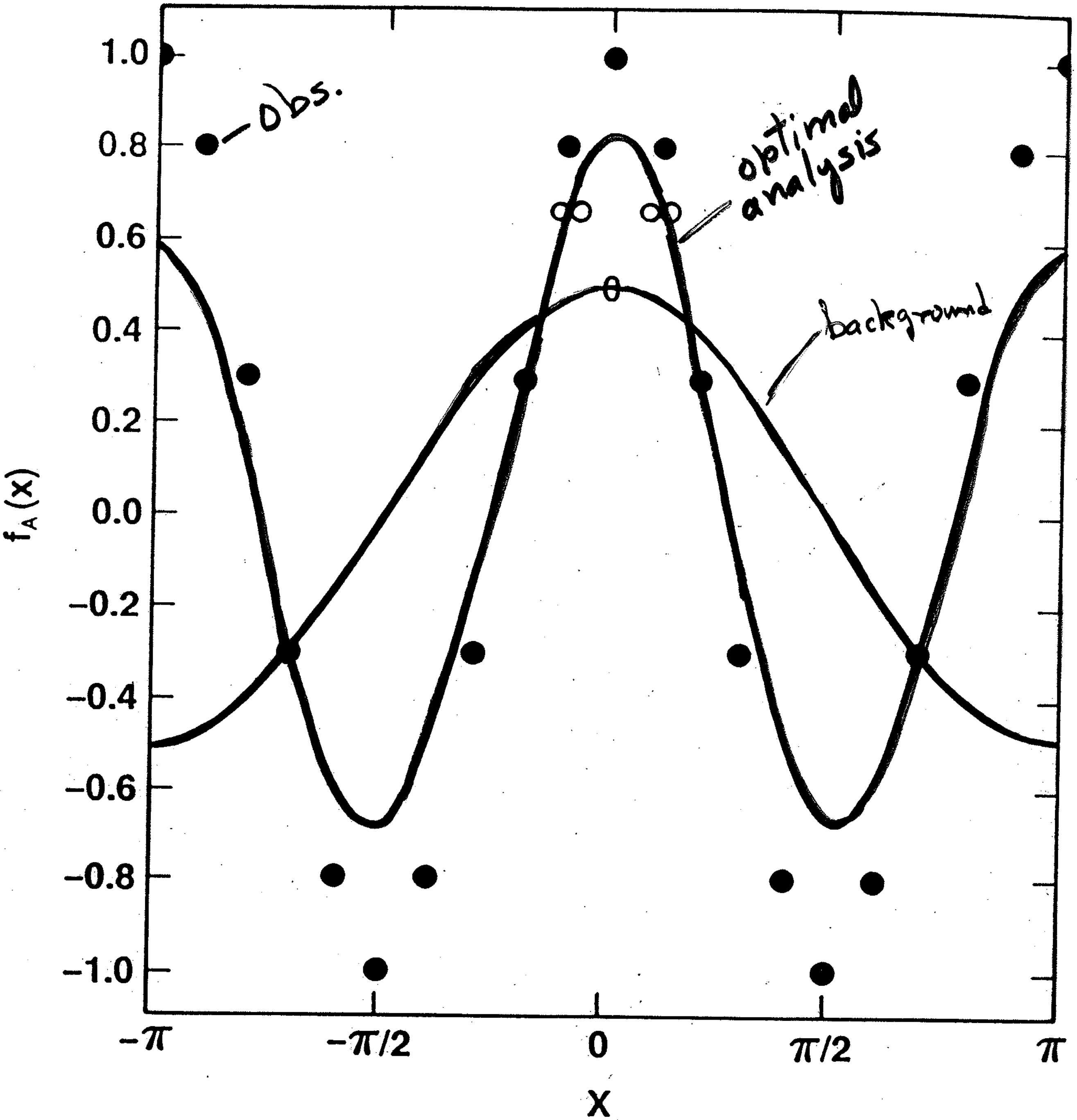
Now the **third** modification is to introduce a revised forward interpolation. That is, we obtain the values of f_A^j at the obs. stations by applying the analysis algorithm directly at the obs. locations via eq. (7). Thus, substituting (7) in for the $\underline{f}_A$ term in (8), we have, after "some algebra":

$$f_A^{\infty}(\mathbf{r}_i) \quad = \quad f_B(\mathbf{r}_i) + \underline{\mathbf{B}}_i^{\mathbf{T}} \, [\, \underline{\underline{\mathbf{B}}} + \underline{\underline{\mathbf{Q}}} \,]^{-1} \, [\, \underline{f}_o - \underline{f}_B \,] \qquad (9)$$

Eq. (9) is our desired "optimal analysis" obtained by iterating a specific form of the SCM algorithm.

Bratseth (1986) has also shown how a SCM algorithm can converge to an optimal analysis. See Appendix F in Daley.

Chap. 4 in Daley will show how to obtain eq. (9) directly using minimum variance estimation methods. Kalnay's book also derives this equation, and we will see how it is related to the "3D-Var." analysis method.



ıre 3.10 Illustration of the application of an SCM algorithm that converges t optimal analysis. The format is the same as for Figure 3.2(a), with the backgr denoted 0 and the asymptotic analysis ∞.