Optimal Interpolation and Nudging

Brian Hunt

4 June 2013

Thought Experiment

- You see two clocks, one says 9:30 and one says 9:40. What is your best estimate of the current time?
- For now, assume the clocks look identical and no other information is available.
- One possible answer: 9:35.
- Another possisible answer: about 9:30 with probability $1/2$, and about 9:40 with probability $1/2$.

Thought Experiment

- If we perceive the clock that reads 9:40 to be more accurate (or less likely to have failed) than the other, we could consider 9:40 to be more probable than 9:30.
- If we make a nonprobabilistic estimate (focus of this lecture), it should be closer to 9:40 than 9:30.
- How much closer depends on quantifying the uncertainties in the clock readings and what we mean by “best estimate”.
- One notion of “best” is the minimum variance unbiased estimator (MVUE).

Unbiased Estimators

- Suppose we have two independent observations y_1 and y_2 of an unknown scalar quantity x, with different accuracies.
- More specifically, assume each y_j is independently sampled from a distribution with (unknown) mean x and (known) standard deviation σ_j.
- An estimator for x is a scalar function $f(y_1, y_2)$.
- It is unbiased if the mean (over different samples) of $f(y_1, y_2)$ is x.

MVUE

- For all λ, the statistic $\lambda y_1 + (1 - \lambda) y_2$ is an unbiased estimator of x.
- The variance of this estimator is $\lambda^2 \sigma_1^2 + (1 - \lambda)^2 \sigma_2^2$, which is minimized when
$$\lambda = \sigma_1^{-2} / (\sigma_1^{-2} + \sigma_2^{-2}).$$
- The MVUE for x is
$$(\sigma_1^{-2} y_1 + \sigma_2^{-2} y_2) / (\sigma_1^{-2} + \sigma_2^{-2}).$$
- This is also the maximum likelihood estimate for x if (e.g.) the distributions are Gaussian and the “prior” is uniform.

Observation Bias

- On the previous slides, I assumed that the observations y_j are unbiased: the mean of the error $y_j - x$ is 0.
- Real observations are likely to be biased.
- Conundrum:
- If we know what the bias (the mean of $y_j - x$) is, we can subtract it from y_j to get an unbiased observation.
- If we don't know what the bias is, then what is the relationship between y_j and x?

Observation (Forward) Operator

- A possible answer to the previous question is to assume that y_j is sampled from a distribution with mean $H_j(x)$; the function H_j is called an observation operator or forward operator.
- If the same procedure is used to "observe" x at many different times, one can try to adjust the function H_j to improve the accuracy of this assumption.
- One may be tempted instead to assume that some function of y_j has mean x, but this approach is less flexible.

Role of Data Assimilation

- Suppose x is a time-varying vector, and the available observations at a given time form a vector y.
- Data assimilation is particularly useful in cases when y does not contain enough information to uniquely determine x, perhaps b/c it is lower dimensional.
- On the other hand, if we are trying to predict future observations, it helps to have a model whose state x has enough information to determine y.

Scalar Optimal Interpolation

- Let's return to the case when x is a scalar, but plan ahead for the vector case.
- Assume each H_j is linear: y_j is sampled from a distribution with mean $H_j x$.
- The MVUE for x is

$$\frac{(\sigma_1/H_1)^{-2}y_1/H_1 + (\sigma_2/H_2)^{-2}y_2/H_2)}{(\sigma_1/H_1)^{-2} + (\sigma_2/H_2)^{-2}}$$

$$= \frac{H_1\sigma_1^{-2}y_1 + H_2\sigma_2^{-2}y_2}{H_1^2\sigma_1^{-2} + H_2^2\sigma_2^{-2}}.$$

Scalar Optimal Interpolation

- Now assume $H_1 = 1$ and y_1 is a background estimate x^b of x, which may be based in part on some previous observations, while y_2 is a new observation y^o.
- Replace σ_1^2 with P^b, σ_2^2 with R^o, and H_2 with H; the MVUE is then

$$\frac{(P^b)^{-1}x^b + H(R^o)^{-1}y^o}{(P^b)^{-1} + H^2(R^o)^{-1}}$$

$$= x^b + \frac{H(R^o)^{-1}(y^o - Hx^b)}{(P^b)^{-1} + H^2(R^o)^{-1}}$$

Vector Optimal Interpolation (OI)

- Next, assume x and y are vectors and H, P^b, and R^o are matrices.
- The MVUE can then be written

$$x^a = x^b + G(y^o - Hx^b)$$

where

$$\begin{aligned} G &= [(P^b)^{-1} + H^T(R^o)^{-1}H]^{-1}H^T(R^o)^{-1} \\ &= P^bH^T[HP^bH^T + R^0]^{-1} \end{aligned}$$

- Here x^a is the analysis ("after") estimate that takes into account the background ("before") x^b and the new observations y^o.

Data assimilation cycle

- Consider now a sequence of observations and estimates indexed by n, representing times $t_n = n\Delta t$.
- Let M_n be a model such that the true state satisfies $x^t_{n+1} \approx M_n(x^t_n)$.
- A data assimilation (DA) cycle is

$$x^a_n = x^b_n + G_n(y^o_n - Hx^b_n)$$
$$x^b_{n+1} = M_n(x^a_n)$$

- An OI cycle, like 3DVar, typically uses a background covariance P^b, and hence a gain G, that is independent of n.

Nudging and Direct Insertion

- If G is constant in time but formulated in a more ad hoc manner than the OI gain, I'll call the resulting DA cycle discrete-time nudging; nudging is usually formulated in continuous time (limit as $\Delta t \to 0$).
- If $H = I$ and $G = I$, then $x_n^a = y_n^o$; this is direct insertion.
- Direct insertion also refers to the case when H maps x to a subset of its coordinates; then $x_n^a = y_n^o$ for observed coordinates and $x_n^a = x_n^b$ for unobserved coordinates.

Canonical OI Citations

- Eliassen, A. 1954. Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Report 5. Videnskaps-Akademiet, Institut for Vaer och Klimaforskning (Norwegian Academy of Sciences, Institute of Weather and Climate Research), Oslo, Norway.
- Gandin, L. 1963. Objective analysis of meteorological fields. Leningrad: Gidromet; English translation Jerusalem: Israel Program for Scientic Translations, 1965.

OI vs. 3DVar

- The OI analysis computes the minimum of the 3DVar cost function in the case that the observation operator h is linear (H).
- 3DVar often refers to a minimization implementation that allows nonlinear h and may include e.g. preconditioning.
- OI is often presumed to include localization, where the estimate of x at a given geographical location is done only with nearby observations.
- OI with localization was used at NCEP in 1980s [e.g., Derber, Parrish, Lord 1991].

Some Further OI Perspectives

http://www.ecmwf.int/newsevents/training/meteorological_presentations/pdf/DA/AssAlg_2.pdf

http://www.ecmwf.int/newsevents/training/rcourse_notes/DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts8.html