MET.O.19 BRANCH MEMORANDUM No. 88

INTRODUCTION TO MINIMUM VARIANCE RETRIEVAL METHODS

by

J R EYRE

January 1987

Meteorological Office Unit (Met.O.19c),
Robert Hooke Institute for Co-operative Atmospheric Research,
Clarendon Laboratory, Oxford, OX1 3PU.

INTRODUCTION TO MINIMUM VARIANCE RETRIEVAL METHODS

This note tries to give a gentle introduction to the use of minimum variance techniques for retrieving atmospheric vertical profiles from satellite sounding data. It is intended to supplement the more rigorous treatments of the subject given for example by Rodgers (1976) and Menke (1984).

We start from the premise that, for the satellite measurements we are interested in, it is not possible to obtain an adequate retrieval of the atmospheric variables of interest using the satellite data alone; we need to supplement this with some "first-guess" or "a priori" information on the atmospheric state. Also we have to take proper account of errors, both in the measurements and in the first guess. We are not seeking an exact answer, but we are looking for the best estimate given the measurements and the a priori information. Our problem has a fundamentally statistical nature since we are seeking the "best" estimate in the minimum variance sense, i.e. the solution which minimises the variance of the error in the estimate when averaged over a large number of cases.

Consider first the usual experimental technique for combining two estimates of a variable x to obtain the best value. Given independent measurements:

x_1 (with expected error variance σ_1^2)
and x_2 (with expected error variance σ_2^2),

the "best" (i.e. minimum variance and, for normally distributed errors, maximum probability) estimate of x is given by

$$\hat{x} = \frac{x_1/\sigma_1^2 + x/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2} \tag{1}$$

with expected error variance,

$$\hat{\sigma}^2 = \frac{1}{1/\sigma_1^2 + 1/\sigma_2^2} . \tag{2}$$

Equations 1 and 2 may also be written

$$\hat{x} = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2} \tag{3}$$

and $$\hat{\sigma}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} . \tag{4}$$

The proof that these equations represent the minimum variance solution is given in the Appendix.

Now consider the same calculation from a different point of view: we start with an (a priori) measurement x (with error σ_1^2) and later we obtain a second measurement x_2 (with error σ_2^2). How does x_2 change our previous best estimate (i.e. x_1) ? Equations 3 and 4 can be re-arranged to give

$$\hat{x} = x_1 + \frac{\sigma_1^2 (x_2 - x_1)}{\sigma_1^2 + \sigma_2^2} \tag{5}$$

and

$$\hat{\sigma}^2 = \sigma_1^2 - \frac{\sigma_1^4}{\sigma_1^2 + \sigma_2^2} \, . \tag{6}$$

These equations give $\hat{x}$ and $\hat{\sigma}^2$ in terms of adjustments to x_1 and σ_1^2.

So far we have considered only scalar quantities. Now consider $\underset{\sim}{x_1}$ and $\underset{\sim}{x_2}$ as vectors (of dimension N). Their errors must now be expressed in the form of covariances matrices (dimension NxN), $\underset{\sim}{S_1}$ and $\underset{\sim}{S_2}$ respectively.

It can be shown that the expressions equivalent to equations 5 and 6 are:

$$\underset{\sim}{\hat{x}} = \underset{\sim}{x_1} + \underset{\sim}{S_1} \cdot (\underset{\sim}{S_1} + \underset{\sim}{S_2})^{-1} \cdot (\underset{\sim}{x_2} - \underset{\sim}{x_1}) \tag{7}$$

and

$$\underset{\sim}{\hat{S}} = \underset{\sim}{S_1} - \underset{\sim}{S_1} \cdot (\underset{\sim}{S_1} + \underset{\sim}{S_2})^{-1} \cdot \underset{\sim}{S_1} \tag{8}$$

where $^{-1}$ denotes matrix inverse.

In our case, the first measurement (first guess) $\underset{\sim}{x_1}$ is our knowledge of the atmospheric profile in advance of any satellite measurement. If our first guess is from a forecast, then $\underset{\sim}{S_1}$ represents the error covariance of the forecast. If $\underset{\sim}{x_1}$ is a climatological mean profile, then $\underset{\sim}{S_1}$ represents the covariance of climatology (i.e. the expected variation about the mean profile).

Equations 7 and 8 cannot be used directly because our second measurement is not a measurement of the atmospheric profile but is a measurement of radiation (radiance or brightness temperature) usually at a number of wavelengths or "channels". We therefore have two problems:

- to "map" the radiances from "measurement space" into "atmospheric profile space", and

- to combine the mapped measurement with the a priori profile.

Again let us look first at the scalar version of the problem:

our first guess is x_1, with error σ_1^2,
our measurement is y with error σ_y^2.

Let a perfect measurement y be related to the true value of x through the equation

$$y = k x + c \tag{9}$$

where k and c are assumed constant and known (and so dy/dx = k).

From our imperfect measurement, y_2, we can therefore estimate x_2:

$$x_2 = \frac{y_2 - c}{k} \tag{10}$$

with $$\sigma_2^2 = \sigma_y^2 / k^2 . \tag{11}$$

Substituting equations 10 and 11 into 5 and 6 we obtain

$$\begin{aligned}\hat{x} &= x_1 + \sigma_1^2 . \left(\sigma_1^2 + \frac{\sigma_y^2}{k^2}\right)^{-1} . \left(\frac{y_2 - c}{k} - x_1\right) \\ &= x_1 + \sigma_1^2 k . (k \sigma_1^2 k + \sigma_y^2)^{-1} . (y_2 - \{k x_1 + c\})\end{aligned} \tag{12}$$

and $$\hat{\sigma}^2 = \sigma_1^2 - \sigma_1^2 k . (k \sigma_1^2 k + \sigma_y^2)^{-1} . k \sigma_1^2 . \tag{13}$$

Furthermore, we can write equation 12 as

$$\hat{x} = x_1 + \sigma_1^2 k . (k \sigma_1^2 k + \sigma_y^2)^{-1} . (y_2 - y\{x_1\}) \tag{14}$$

where $y\{x_1\} = k.x_1 + c$, i.e. the value of y corresponding the first guess value x_1.

When equations 13 and 14 are generalised to combine a first-guess vector $\underset{\sim}{x_1}$ (dimension N) of error covariance $\underset{\sim}{S_1}$ with a measurement vector $\underset{\sim}{y_2}$ (dimension M) of error covariance $\underset{\sim}{S_y}$, we obtain

$$\boxed{\begin{aligned}\hat{\underset{\sim}{x}} &= \underset{\sim}{x_1} + \underset{\sim}{S_1} . \underset{\sim}{K}^T . (\underset{\sim}{K} . \underset{\sim}{S_1} . \underset{\sim}{K}^T + \underset{\sim}{S_y})^{-1} . (\underset{\sim}{y_2} - \underset{\sim}{y}\{\underset{\sim}{x_1}\}) \quad (15) \\ \hat{\underset{\sim}{S}} &= \underset{\sim}{S_1} - \underset{\sim}{S_1} . \underset{\sim}{K}^T . (\underset{\sim}{K} . \underset{\sim}{S_1} . K^T + \underset{\sim}{S_y})^{-1} . \underset{\sim}{K} . \underset{\sim}{S_1} \quad (16)\end{aligned}}$$

where $\underset{\sim}{y}\{x_1\}$ is the measurement vector corresponding to the first guess $\underset{\sim}{\text{vector}}$,
and $\underset{\sim}{K}$ contains the partial derivations of the elements of y with respect to the elements of x,
and ^{T} represents matrix transpose.

Equations 15 and 16 represent the minimum variance ("best") estimate of $\underset{\sim}{x}$ and the estimate of its error covariance. They are the forms of the minimum variance solution most commonly used in the retrieval problem for atmospheric profile remote sensing. They look complicated but they are only the multi-dimensional, two-space equivalents to equations 1 and 2, i.e. the usual equations for obtaining the best estimate given two independent measurements and their respective errors.

We can also write equation 15 as

$$\hat{\underset{\sim}{x}} - \underset{\sim}{x_1} = \underset{\sim}{W} . (\underset{\sim}{y_2} - \underset{\sim}{y}\{\underset{\sim}{x_1}\}) . \tag{17}$$

This is a more general form of the retrieval equation which shows how the "inverse matrix" $\underset{\sim}{W}$ maps differences between real measurements $\underset{\sim}{y_2}$

and those which would have been calculated from the first guess into the differences between the retrieval $\hat{\underset{\sim}{x}}$ and the first guess $\underset{\sim}{x_1}$. For the minimum variance solution $\underset{\sim}{W}$ takes the particular form,

$$\underset{\sim}{W} = \underset{\sim}{S_1} \cdot \underset{\sim}{K}^T \cdot (\underset{\sim}{K} \cdot \underset{\sim}{S_1} \cdot \underset{\sim}{K}^T + \underset{\sim}{S_y})^{-1} . \tag{18}$$

Also, equation 16 may be written as

$$\begin{aligned} \hat{\underset{\sim}{S}} &= \underset{\sim}{S_1} - \underset{\sim}{W} \cdot \underset{\sim}{K} \cdot \underset{\sim}{S_1} \\ &= (\underset{\sim}{I} - \underset{\sim}{W} \cdot \underset{\sim}{K}) \cdot \underset{\sim}{S_1} \end{aligned} \tag{19}$$

where $\underset{\sim}{I}$ is a unit matrix.

REFERENCES

Rodgers C D, 1976.
Rev Geophys Space Phys, 14, 609-624.
Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation.

Menke W, 1984.
Academic Press.
Geophysical data analysis: discrete inverse theory.

APPENDIX. Proof of minimum variance solution.

An individual measurement x_1 has error ϵ_{x_1}, such that over many cases the error variance is $\sigma^2_{x_1}$:

$$\sigma^2_{x_1} = \frac{1}{I} \sum_{i=1}^{I} (\epsilon_{x_1})^2_i . \tag{A.1}$$

Similarly, measurement x_2 has error ϵ_{x_2}, with error variance $\sigma^2_{x_2}$, and we assume ϵ_{x_1} and ϵ_{x_2} uncorrelated. Our best estimate of x is some linear combination of x_1 and x_2:

$$\hat{x} = w_1 x_1 + w_2 x_2 . \tag{A.2}$$

Our problem is to find the weights w_1 and w_2 which minimise the error in $\hat{x}$, which is given by

$$\epsilon_{\hat{x}} = w_1 \cdot \epsilon_{x_1} + w_2 \epsilon_{x_2} . \tag{A.3}$$

Averaging over many cases, we obtain

$$\hat{\sigma}^2 = \frac{1}{I} \sum_{i=1}^{I} \epsilon^2_{\hat{x}_i} = \frac{1}{I} \sum_{i=1}^{I} (w_1 \epsilon_{x_{1i}} + w_2 \epsilon_{x_{2i}})^2 . \tag{A.4}$$

We minimise A.4 with respect to w_1 (or w_2) under the constraint

$$w_1 + w_2 = 1 . \tag{A.5}$$

[This is obtained by considering A.2: if $x_1 = x_2$, then $\hat{x} = x_1 = x_2$, and so $w_1 + w_2 = 1$.]

Therefore

$$\hat{\sigma}^2 = \frac{1}{I} \sum_{i=1}^{I} \{ w_1 \epsilon_{x_{1i}} + (1-w_1) \epsilon_{x_{2i}} \}^2 = \text{minimum}, \qquad \text{A.6}$$

or

$$\frac{d\hat{\sigma}^2}{dw_1} = 2 w_1 \frac{1}{I} \sum_{i=1}^{I} (\epsilon^2_{x_{1i}} + \epsilon^2_{x_{2i}}) - 2 \frac{1}{I} \sum_{i=1}^{I} \epsilon^2_{x_{2i}}$$

$$= 2 w_1 (\sigma_1^2 + \sigma_2^2) - 2 \sigma_2^2 = 0. \qquad \text{A.7}$$

Therefore

$$w_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \qquad \text{A.8}$$

Similarly

$$w_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}. \qquad \text{A.9}$$

Therefore

$$\hat{x} = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2}. \qquad \text{A.10 (cf. Equ. 3)}$$