Representing Systematic Errors in Retrievals

Anu Dudhia
AOPP, Dept. of Physics
University of Oxford

January 15, 2018

1 Introduction

It is generally understood that 'systematic errors' (i.e., anything apart from random measurement noise) can be incorporated into a retrieval in one of two ways

1. as additional elements in the state vector ('parameter space'), or
2. as additional components in the measurement error covariance matrix ('measurement space'),

and that these two approaches are broadly equivalent.

Von Clarmann *et al.* (2001)[2] is the usually cited reference for this assumption, but the authors of that paper are careful to specify that these two approaches are only equivalent in the strict mathematical sense if systematic error components in parameter space are retrieved in an optimal estimation framework, i.e., with some *a priori* constraint on the uncertainty of the parameters. They show the retrievals are equivalent when the representation of error covariance in measurement space is just the mapping of the *a priori* error covariance from parameter space.

When representing errors in measurement space it is quite common to perform this mapping from parameter space (the alternative being to derive the measurement covariance directly from the observations). For example, uncertainties due to water vapour are calculated first by determining the 1-σ variability in concentration (parameter space), then using a forward model to calculate the difference $\delta\mathbf{y}$ this represents in the measurements (measurement space), and finally adding these, as $(\delta\mathbf{y})(\delta\mathbf{y})^T$, to the measurement error covariance.

However, when representing errors in parameter space, as additional state vector elements, it is frequently done without any *a priori* constraint, i.e., just in a least-squares fit sense. Since this is equivalent to setting the *a priori* covariance to infinity, there can be no explicit equivalent measurement covariance representation, although it turns out there there remains an equivalent *inverse* measurement covariance.

The aims of this note are to

1. review (and re-derive) the von Clarmann *et al.* results;
2. derive a relationship between the *unconstrained* error retrieval in parameter space and the equivalent *inverse* of the measurement covariance matrix;
3. provide a couple of fully-worked examples.

If you are not familiar with the concepts, it may help to start with the examples in the Appendix.

2 Standard Retrievals

Following Rodgers[1], the weighted least-squares-fit (LSF) retrieval of state vector $\mathbf{x}$ given a set of measurements $\mathbf{y}$ is:

$$\mathbf{x} = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{y} \equiv \mathbf{G}\mathbf{y} \tag{1}$$

where $\mathbf{K}$ is the matrix of Jacobians ($K_{ij} = \partial y_i/\partial x_j$), $\mathbf{G}$ is the Gain matrix ($G_{ij} = \partial x_i/\partial y_j$) and $\mathbf{S}_y$ is the measurement error covariance matrix.

The associated *retrieval* error covariance $\mathbf{S}_x$ reduces to just the first part of the expression for $\mathbf{G}$:

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}\right)^{-1}$$

The optimal estimation (OE) retrieval represents a weighted average of measurements and an *a priori* estimate $\mathbf{a}$ with associated uncertainty covariance $\mathbf{S}_a$:

$$\mathbf{x} = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}\left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{y} + \mathbf{S}_a^{-1}\mathbf{a}\right) = \mathbf{G}\mathbf{y} + \mathbf{H}\mathbf{a} \tag{2}$$

The matrix $\mathbf{H}$ ($H_{ij} = \partial x_i/\partial a_j$) acts like $\mathbf{G}$ but containing the dependences of the retrieval on the *a priori*.

The OE retrieval error covariance also reduces to the first part of the OE expression for $\mathbf{G}$:

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T + \mathbf{H}\mathbf{S}_a\mathbf{H}^T = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}$$

As the *a priori* uncertainty increases, $\mathbf{S}_a^{-1} \to 0$ and the OE $\to$ LSF retrieval.

It is also possible to set just some elements of $\mathbf{S}_a^{-1} = 0$ so these elements of $\mathbf{x}$ are retrieved as LSF while others are OE.

3 Systematic Errors

Additional sources of error, whether in the measurement itself (such as radiometric gain uncertainty) or in the forward modelling of the radiances (such as the concentrations of contaminating species) can be represented as a set of error parameters $\mathbf{u}$, representing 1-σ uncertainties about their assumed values (so $\bar{\mathbf{u}} = \mathbf{0}$)

This leads to a covariance matrix $\mathbf{S}_u$, which also allows for any correlations between these errors:

$$\mathbf{S}_u = \langle \mathbf{u}\mathbf{u}^T \rangle \tag{3}$$

$\mathbf{u}$ and $\mathbf{S}_u$ are defined for an arbitrary set of parameters in a variety of units. This covariance in parameter space can be converted to measurement space $\mathbf{S}_s$ (same size as $\mathbf{S}_y$) using the standard covariance transform

$$\mathbf{S}_s = \mathbf{L}\mathbf{S}_u\mathbf{L}^T \tag{4}$$

where $\mathbf{L}$ is the 'error Jacobian' ($L_{ij} = \partial y_i/\partial u_j$). Thus we can express errors as covariances either in parameter space $\mathbf{S}_u$ or in measurement space $\mathbf{S}_s$.

Any reasonable $\mathbf{S}_u$ should be invertible. However, since (usually) the number of measurements is larger than the number of error parameters, $\mathbf{S}_s$ won't be invertible (basically, the mapping has converted the smaller matrix $\mathbf{S}_u$ into the larger matrix $\mathbf{S}_s$, so some degeneracy has been introduced).

4 Retrieval with Errors in Measurement Space

Representing the additional errors in measurement space as covariance $\mathbf{S}_s$ (Eq. 4), the standard LSF retrieval (Eq. 1) is simply modified:

$$\mathbf{x} = \mathbf{G}\mathbf{y} = \left[\mathbf{K}^T\left(\mathbf{S}_y + \mathbf{S}_s\right)^{-1}\mathbf{K}\right]^{-1}\mathbf{K}^T\left(\mathbf{S}_y + \mathbf{S}_s\right)^{-1}\mathbf{y}$$

We assume that $\mathbf{S}_y$ and $\mathbf{S}_y + \mathbf{S}_s$ are both invertible, but not $\mathbf{S}_s$ (see previous section). Standard formula for the inverse of the sum of two matrices (avoiding $\mathbf{B}^{-1}$) is

$$\left(\mathbf{A} + \mathbf{B}\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{A} + \mathbf{B}\right)^{-1}$$

Identifying $\mathbf{A} \equiv \mathbf{S}_y$ and $\mathbf{B} \equiv \mathbf{S}_s$, the solution for $\mathbf{x}$ can then be written as:

$$\mathbf{x} = \left[\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} - \mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{S}_s\left(\mathbf{S}_y + \mathbf{S}_s\right)^{-1}\mathbf{K}\right]^{-1}\mathbf{K}^T\left[\mathbf{S}_y^{-1} - \mathbf{S}_y^{-1}\mathbf{S}_s\left(\mathbf{S}_y + \mathbf{S}_s\right)^{-1}\right]\mathbf{y} \tag{5}$$

As required, as $\mathbf{S}_s \to 0$ this converges to the standard LSF solution (Eq. 1).

5 Retrieval with Errors in Parameter Space

For this case the state vector is separated into two components: the original n target parameters $\mathbf{x}$, and an extra l parameters representing error terms $\mathbf{u}$. The Jacobian matrix is extended $\mathbf{K} \to (\mathbf{K}, \mathbf{L})$ to include the derivatives $\mathbf{L} = \partial\mathbf{y}/\partial\mathbf{u}$ (sec. 3). We also assume that an *a priori* constraint is *only* applied to these extra parameters so, from the OE form (Eq. 2), we assume only the elements of $\mathbf{S}_a^{-1}$ corresponding to $\mathbf{u}$ are non-zero, and set to $\mathbf{S}_u^{-1}$ (Eq. 3). Thus the modified (OE) solution (ignoring the dependence on $\mathbf{a} \equiv \mathbf{u}$, assumed $\mathbf{0}$) is:

$$\begin{aligned}
\begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} = \mathbf{G}\mathbf{y} &= \left[\begin{pmatrix} \mathbf{K}^T \\ \mathbf{L}^T \end{pmatrix} \mathbf{S}_y^{-1} \begin{pmatrix} \mathbf{K}, & \mathbf{L} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{S}_u^{-1} \end{pmatrix} \right]^{-1} \begin{pmatrix} \mathbf{K}^T \\ \mathbf{L}^T \end{pmatrix} \mathbf{S}_y^{-1}\mathbf{y} \\
&= \begin{bmatrix} \mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} & \mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{L} \\ \mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{K} & \mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1} \end{bmatrix}^{-1} \begin{pmatrix} \mathbf{K}^T \\ \mathbf{L}^T \end{pmatrix} \mathbf{S}_y^{-1}\mathbf{y} \equiv \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} \begin{pmatrix} \mathbf{K}^T \\ \mathbf{L}^T \end{pmatrix} \mathbf{S}_y^{-1}\mathbf{y}
\end{aligned}$$

Standard formula for inverse of block matrix is

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & -\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}$$

(Note the repetition of $(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$ in all four blocks of the inverse.) However we are only interested in the top 'row' of $\mathbf{G}$ which gives $\mathbf{x}$

$$\begin{aligned}
\mathbf{x} &= \left[\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right]^{-1} \begin{pmatrix} \mathbf{I}_n, & -\mathbf{B}\mathbf{D}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{K}^T \\ \mathbf{L}^T \end{pmatrix} \mathbf{S}_y^{-1}\mathbf{y} \\
&= \left[\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} - \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{L}\right)\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1}\right)^{-1}\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{K}\right)\right]^{-1} \\
&\quad \times \left(\mathbf{K}^T\mathbf{S}_y^{-1} - \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{L}\right)\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1}\right)^{-1}\mathbf{L}^T\mathbf{S}_y^{-1}\right)\mathbf{y} \qquad (6)
\end{aligned}$$

As $\mathbf{S}_u \to 0$, i.e., errors are reduced, $\mathbf{S}_u^{-1} \to \infty$ and this also converges on the standard LSF solution (Eq. 1).

6 Equivalence

Both LSF (Eq. 5) and OE (Eq. 6) solutions for $\mathbf{x}$ incorporating systematic errors are in the form

$$\mathbf{x} = \left[\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} - \mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{M}\mathbf{K}\right]^{-1}\left(\mathbf{K}^T\mathbf{S}_y^{-1} - \mathbf{M}\right)\mathbf{y}$$

and converge to the standard LSF solution if $\mathbf{M} \to 0$, so that $\mathbf{M}$ contains the influence of the systematic errors. The two retrievals are equivalent if the two forms for $\mathbf{M}$ are the same:

$$\mathbf{S}_s\left(\mathbf{S}_y + \mathbf{S}_s\right)^{-1} \stackrel{?}{=} \mathbf{L}\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1}\right)^{-1}\mathbf{L}^T\mathbf{S}_y^{-1} \qquad (7)$$

At this point, von Clarmann *et al.* introduce the square root of matrix $\mathbf{S}_u$, but the Rodgers (sec 4.1 in the book) approach seems more elegant. Starting from an expression (which can be confirmed by simply multiplying out):

$$\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1}\right)\mathbf{S}_u\mathbf{L}^T = \mathbf{L}^T\mathbf{S}_y^{-1}\left(\mathbf{S}_y + \mathbf{L}\mathbf{S}_u\mathbf{L}^T\right)$$

pre-multiply both sides by $\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1}\right)^{-1}$ and post-multiply both sides by $\left(\mathbf{S}_y + \mathbf{L}\mathbf{S}_u\mathbf{L}^T\right)^{-1}$, leading to

$$\mathbf{S}_u\mathbf{L}^T\left(\mathbf{S}_y + \mathbf{L}\mathbf{S}_u\mathbf{L}^T\right)^{-1} = \left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L} + \mathbf{S}_u^{-1}\right)^{-1}\mathbf{L}^T\mathbf{S}_y^{-1}$$

Pre-multiplying both sides by $\mathbf{L}$ and noting (Eq. 4) $\mathbf{S}_s = \mathbf{L}\mathbf{S}_u\mathbf{L}^T$ leads to Eq. (7), hence demonstrating the equivalance.

7 Unconstrained Errors in Parameter Space

It has now been shown that a retrieval using errors represented in parameter space with an *a priori* covariance $\mathbf{S}_u$, is equivalent to a retrieval for which $\mathbf{S}_u$ mapped into measurement space, $\mathbf{S}_s$, is incorporated into the measurement error covariance.

An *unconstrained* retrieval is represented by the limit of $\mathbf{S}_u \to \infty$. This presents no problems in parameter space, which only (Eq. 6) involves $\mathbf{S}_u^{-1}$, so remains finite and well-defined.

The measurement space solution (Eq. 5), however, involves $\mathbf{S}_s$ which becomes infinite. But since Eq. (7) still holds, so as $\mathbf{S}_u^{-1} \to 0$, then

$$\mathbf{S}_s(\mathbf{S}_y + \mathbf{S}_s)^{-1} \to \mathbf{L}\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L}\right)^{-1}\mathbf{L}^T\mathbf{S}_y^{-1}$$

From sec. 4 we can identify the inverse of the measurement covariance

$$(\mathbf{S}_y + \mathbf{S}_s)^{-1} \equiv \mathbf{S}_y^{-1}\left(\mathbf{I} - \mathbf{S}_s(\mathbf{S}_y + \mathbf{S}_s)^{-1}\right)$$

Therefore, in the limit of the unconstrained retrieval in parameter space, the equivalent inverse of the measurement error covariance converges to

$$(\mathbf{S}_y + \mathbf{S}_s)^{-1} \to \mathbf{S}_y^{-1} - \mathbf{S}_y^{-1}\mathbf{L}\left(\mathbf{L}^T\mathbf{S}_y^{-1}\mathbf{L}\right)^{-1}\mathbf{L}^T\mathbf{S}_y^{-1} \tag{8}$$

So, although we cannot define an equivalent measurement space covariance corresponding to the unconstrained parameter error retrieval, its *inverse* covariance can be defined.

A A Simple Example

A.1 Standard BTD Retrieval

Consider the basic brightness temperature difference (BTD) retrieval using two measurements (y_1, y_2) to derive a target molecule concentration x (with known Jacobian $k = dy/dx$) and a background radiance b:

$$\begin{aligned} y_1 &= kx + b \\ y_2 &= \quad 0 + b \end{aligned} \quad\Rightarrow\quad \begin{aligned} x &= (y_1 - y_2)/k \\ b &= y_2 \end{aligned}$$

Putting this into matrix form

$$\mathbf{y} = \mathbf{K}\mathbf{x} = \begin{pmatrix} k & 1 \\ 0 & 1 \end{pmatrix}\mathbf{x} \quad ; \quad \mathbf{x} = \mathbf{G}\mathbf{y} = \mathbf{K}^{-1}\mathbf{y} = \frac{1}{k}\begin{pmatrix} 1 & -1 \\ 0 & k \end{pmatrix}\mathbf{y}$$

If we regard the background b as an error term, this is equivalent to retrieving the error in parameter space but in the LSF sense, i.e., without any *a priori* constraint on b (effectively $\mathbf{S}_u^{-1} \to 0$). Note also that since b is already in measurement units, the equivalent components of $\mathbf{L}$ (second column of $\mathbf{K}$) are 1.

Although $\mathbf{S}_y$ hasn't been specified, $\mathbf{G}$ does conform to weighted least-squares-fit solution (Eq. 1) since, in this case, $\mathbf{K}$ (and therefore also $\mathbf{K}^T$) is invertible:

$$\mathbf{G} = \left[\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}\right]^{-1}\mathbf{K}^T\mathbf{S}_y^{-1} = \left[\left(\mathbf{K}^T\right)^{-1}\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}\right]^{-1}\mathbf{S}_y^{-1} = \left[\mathbf{S}_y^{-1}\mathbf{K}\right]^{-1}\mathbf{S}_y^{-1} = \mathbf{K}^{-1}$$

i.e., with two measurements and two unknowns, only the exact solution is possible and measurement errors do not affect the solution. However, we can still evaluate the error in the solution. If we assume measurement error covariance of the form $\mathbf{S}_y = S_y\mathbf{I}$, the retrieval error is given by:

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T = \frac{S_y}{k^2}\begin{pmatrix} 2 & -k \\ -k & k^2 \end{pmatrix}$$

A.2 Retrieval with Errors in Parameter Space

To add an *a priori* constraint to the background term in the BTD retrieval, define:

$$\mathbf{a} = 0 \quad ; \quad \mathbf{S}_a^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & S_b^{-1} \end{pmatrix} \quad ; \quad \alpha = \frac{S_y}{S_b}$$

Then, for the OE solution:

$$\begin{aligned} \begin{pmatrix} x \\ b \end{pmatrix} &= \mathbf{G}\mathbf{y} = \left[\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right]^{-1}\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{y} = \begin{bmatrix} k^2 & k \\ k & 2+\alpha \end{bmatrix}^{-1}\begin{pmatrix} k & 0 \\ 1 & 1 \end{pmatrix}\mathbf{y} \\ &= \frac{1}{k^2(1+\alpha)}\begin{pmatrix} 2+\alpha & -k \\ -k & k^2 \end{pmatrix}\begin{pmatrix} k & 0 \\ 1 & 1 \end{pmatrix}\mathbf{y} = \frac{1}{k(1+\alpha)}\begin{pmatrix} (1+\alpha) & -1 \\ 0 & k \end{pmatrix}\mathbf{y} \end{aligned}$$

For the OE retrieval error covariance:

$$\mathbf{S}_x = \left[\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right]^{-1} = \frac{S_y}{k^2(1+\alpha)}\begin{pmatrix} (2+\alpha) & -k \\ -k & k^2 \end{pmatrix}$$

As the *a priori* constraint is relaxed, $S_b \to \infty$, $\alpha = S_y/S_b \to 0$, and both $\mathbf{G}$ and $\mathbf{S}_x$ tend to the standard BTD solutions in sec. A.1.

As the *a priori* constraint is strengthened (i.e., background radiance b is externally specified) $\alpha \to \infty$ and the x-component (top left) of $\mathbf{S}_x \to S_y/k^2$, i.e. the error expected from just using y_1.

A.3 Retrieval with Errors in Measurement Space

Set up the retrieval for a single parameter x with the background represented as an additional measurement covariance $\mathbf{S}_s$. In this case, keeping just the first column, $\mathbf{K} = (k, 0)$:

$$x = \mathbf{G}\mathbf{y} = \left[\begin{pmatrix} k & 0 \end{pmatrix}(\mathbf{S}_y + \mathbf{S}_s)^{-1}\begin{pmatrix} k \\ 0 \end{pmatrix}\right]^{-1}\begin{pmatrix} k & 0 \end{pmatrix}(\mathbf{S}_y + \mathbf{S}_s)^{-1}\mathbf{y}$$

As before $\mathbf{S}_y = S_y\mathbf{I}$ and, for the background errors, $\mathbf{L} = (1,1)$ so

$$\mathbf{S}_s = \mathbf{L}\mathbf{S}_b\mathbf{L}^T = \begin{pmatrix} 1 \\ 1 \end{pmatrix} S_b \begin{pmatrix} 1 & 1 \end{pmatrix} = S_b\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$(\mathbf{S}_y + \mathbf{S}_s)^{-1} = \begin{pmatrix} S_b + S_y & S_b \\ S_b & S_b + S_y \end{pmatrix}^{-1} = \frac{1}{S_b}\begin{pmatrix} 1+\alpha & 1 \\ 1 & 1+\alpha \end{pmatrix}^{-1} = \frac{1}{S_y(2+\alpha)}\begin{pmatrix} 1+\alpha & -1 \\ -1 & 1+\alpha \end{pmatrix}$$

Noting that the $1/S_y(2+\alpha)$ terms cancel:

$$x = \left[k^2(1+\alpha)\right]^{-1} k \begin{pmatrix} (1+\alpha) & -1 \end{pmatrix}\mathbf{y} = \frac{1}{k(1+\alpha)}\begin{pmatrix} (1+\alpha) & -1 \end{pmatrix}\mathbf{y}$$

For the retrieval covariance

$$\mathbf{S}_x = \left[\mathbf{K}^T(\mathbf{S}_y + \mathbf{S}_s)^{-1}\mathbf{K}\right]^{-1} = \frac{S_y}{k^2(1+\alpha)}(2+\alpha)$$

Which match the x-components of $\mathbf{G}$ and $\mathbf{S}_x$ derived in sec. A.2.

A.4 Inverse Measurement Covariance

From the previous section we already have an explicit expression for the inverse measurement covariance for the errors included in measurement space which we can evaluate as the *a priori* constraint is relaxed ($\alpha \to 0$):

$$(\mathbf{S}_y + \mathbf{S}_s)^{-1} = \frac{1}{S_y(2+\alpha)}\begin{pmatrix} 1+\alpha & -1 \\ -1 & 1+\alpha \end{pmatrix} \to \frac{1}{2S_y}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

which, as expected, is singular. We can also evaluate this using the derived expression, Eq. (8). Recalling that $\mathbf{L} = \partial\mathbf{y}/\partial b$ is the column vector (1,1):

$$\begin{aligned}(\mathbf{S}_y + \mathbf{S}_s)^{-1} &= \frac{1}{S_y}\mathbf{I} - \frac{1}{S_y}\mathbf{L}\left[\mathbf{L}^T\mathbf{L}\right]^{-1}\mathbf{L}^T = \frac{1}{S_y}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{S_y}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\left[\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]^{-1}\begin{pmatrix} 1 & 1 \end{pmatrix} \\ &= \frac{1}{S_y}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{2S_y}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2S_y}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\end{aligned}$$

B A Slightly More Complicated Example

B.1 Standard BTD Retrieval

For a 'genuine' LSF solution (as opposed to the 'exact solution' in the previous example), add another background measurement to the BTD case, so now there are 3 measurements and 2 unknowns. The straightforward solution (assuming equal noise on each measurement) is just to average the two background measurements and then solve as before:

$$\begin{array}{rl} y_1 = & kx+b \\ y_2 = & 0+b \\ y_3 = & 0+b \end{array} \quad \Rightarrow \quad \begin{array}{rl} x = & (y_1 - \frac{1}{2}(y_2+y_3))/k \\ b = & \frac{1}{2}(y_2+y_3) \end{array}$$

However, we can also apply the standard LSF solution for $\mathbf{y} = \mathbf{K}\mathbf{x}$ where:

$$\mathbf{K} = \begin{pmatrix} k & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \quad ; \quad \mathbf{K}^T = \begin{pmatrix} k & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Assuming, as before, $\mathbf{S}_y \equiv S_y\mathbf{I}$, the standard LSF solution is given by:

$$\mathbf{G} = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{S}_y = \begin{pmatrix} k^2 & k \\ k & 3 \end{pmatrix}^{-1} \begin{pmatrix} k & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \frac{1}{2k}\begin{pmatrix} 2 & -1 & -1 \\ 0 & k & k \end{pmatrix}$$

which gives the same solution for x, b as averaging the two background measurements. $\mathbf{S}_y$ again cancels out but, unlike the two-measurement case, this time it is because we assumed the simplified form $S_y\mathbf{I}$ rather than because the solution fundamentally does not depend on it.

The retrieval error covariance is given by

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}\right)^{-1} = S_y\begin{pmatrix} k^2 & k \\ k & 3 \end{pmatrix}^{-1} = \frac{S_y}{2k^2}\begin{pmatrix} 3 & -k \\ -k & k^2 \end{pmatrix}$$

Comparing this with the 2-measurement case (sec. A.1) we see that not only has the variance in the retrieved background (bottom right term) been reduced from S_y to $S_y/2$ as a result of having two measurements instead of one, but this has also resulted in a reduction in the variance in retrieved x (top left term) from $2S_y/k^2$ to $3S_y/2k^2$.

B.2 Retrieval with Errors in Parameter Space

For the (constrained) 2-parameter OE retrieval, $\mathbf{S}_b^{-1}$ is included in the denominator as in the simple example

$$\begin{aligned} \mathbf{G} &= \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_b^{-1}\right)^{-1}\mathbf{K}^T\mathbf{S}_y = \left[\frac{1}{S_y}\begin{pmatrix} k^2 & k \\ k & 3 \end{pmatrix} + \frac{1}{S_b}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]^{-1}\mathbf{K}^T\mathbf{S}_y^{-1} = \begin{bmatrix} k^2 & k \\ k & 3+\alpha \end{bmatrix}^{-1}\mathbf{K}^T\mathbf{I} \\ &= \frac{1}{(2+\alpha)k^2}\begin{pmatrix} 3+\alpha & -k \\ -k & k^2 \end{pmatrix}\begin{pmatrix} k & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \frac{1}{k(2+\alpha)}\begin{pmatrix} 2+\alpha & -1 & -1 \\ 0 & k & k \end{pmatrix} \end{aligned}$$

The error covariance for the OE solution is given by

$$\mathbf{S}_x = \left(\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_b^{-1}\right)^{-1} = \frac{S_y}{k^2(2+\alpha)}\begin{pmatrix} 3+\alpha & -k \\ -k & k^2 \end{pmatrix}$$

In the limit $S_b \to \infty$, $\alpha \to 0$, both $\mathbf{G}$ and $\mathbf{S}_x$ converge to the unconstrained retrieval values in sec. B.1 as expected.

B.3 Retrieval with Errors in Measurement Space

For the single parameter, correlated error LSF, the covariance $\mathbf{S}_y$ is constructed as

$$(\mathbf{S}_y + \mathbf{S}_s) = S_y\mathbf{I} + S_b\mathbf{J} = S_b\begin{pmatrix} 1+\alpha & 1 & 1 \\ 1 & 1+\alpha & 1 \\ 1 & 1 & 1+\alpha \end{pmatrix} \quad \Rightarrow \quad (\mathbf{S}_y + \mathbf{S}_s)^{-1} = \frac{1}{(3+\alpha)S_y}\begin{pmatrix} 2+\alpha & -1 & -1 \\ -1 & 2+\alpha & -1 \\ -1 & -1 & 2+\alpha \end{pmatrix}$$

The Jacobian matrix $\mathbf{K}$ reduces to the column vector $(k, 0, 0)$ so (cancelling out the determinants of $(\mathbf{S}_y + \mathbf{S}_s)^{-1}$):

$$\begin{aligned} \mathbf{G} &= \left(\mathbf{K}^T(\mathbf{S}_y + \mathbf{S}_s)^{-1}\mathbf{K}\right)^{-1}\mathbf{K}^T(\mathbf{S}_y + \mathbf{S}_s)^{-1} \\ &= \left[\begin{pmatrix} k & 0 & 0 \end{pmatrix}\begin{pmatrix} 2+\alpha & -1 & -1 \\ -1 & 2+\alpha & -1 \\ -1 & -1 & 2+\alpha \end{pmatrix}\begin{pmatrix} k \\ 0 \\ 0 \end{pmatrix}\right]^{-1}\begin{pmatrix} k & 0 & 0 \end{pmatrix}\begin{pmatrix} 2+\alpha & -1 & -1 \\ -1 & 2+\alpha & -1 \\ -1 & -1 & 2+\alpha \end{pmatrix} \\ &= \left[k\begin{pmatrix} 2+\alpha & -1 & -1 \end{pmatrix}\begin{pmatrix} k \\ 0 \\ 0 \end{pmatrix}\right]^{-1} k\begin{pmatrix} 2+\alpha & -1 & -1 \end{pmatrix} = \frac{1}{k(2+\alpha)}\begin{pmatrix} 2+\alpha & -1 & -1 \end{pmatrix} \end{aligned}$$

For the retrieval covariance,

$$S_x = \left[\mathbf{K}^T(\mathbf{S}_y + \mathbf{S}_s)^{-1}\mathbf{K}\right]^{-1} = \frac{S_y}{k^2(2+\alpha)}(3+\alpha)$$

which again match the x-components of $\mathbf{G}$ and $\mathbf{S}_x$ derived in the previous section.

B.4 Inverse Measurement Covariance

As before we already have an explicit expression for $(\mathbf{S}_y + \mathbf{S}_s)^{-1}$ in sec. B.3, but we can also calculate it for the limit of an unconstrained retrieval ($\alpha \to 0$) from Eq. (8). So following the same pattern as in sec. A.4 but with $\mathbf{L}$ now a *three*-element column vector containing 1's:

$$(\mathbf{S}_y + \mathbf{S}_s)^{-1} = \frac{1}{S_y}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3S_y}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \frac{1}{3S_y}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

References

[1] Clive D. Rodgers. *Inverse Methods for Atmospheric Sounding, Theory and Practice*. World Scientific, 2000.

[2] Thomas von Clarmann, Udo Grabowski, and Michael Kiefer. On the role of non-random errors in inverse problems in radiative transfer and other applications. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 71(1):39–46, 2001.