Geophysical Journal (1988) **95**, 205–207

RESEARCH NOTE

A note on the calculation of covariance and resolution

Robert L. Nowack and William J. Lutter
Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 49707, USA

Accepted 1988 April 13. Received 1988 April 11; in original form 1988 February 1

SUMMARY
Two formulations of the *a posteriori* covariance using maximum likelihood are considered. These are then related to model resolution. The first formulation is an extension of least squares in which the least squares inverse is replaced by the maximum likelihood inverse. For imperfect resolution, this approach gives a maximum error for diagonal resolutions of 0.5 and decreases for greater or smaller resolutions. The second formulation for the *a posteriori* covariance due to Tarantola (1987) gives a more natural result which reduces to the least squares result for perfect resolution and to the *a priori* covariance in the limit of zero resolution. Many computer algorithms use the first formulation which will systematically underestimate computed model errors as compared to the second formulation in all cases except for perfect resolution.

Key words: covariance, inverse theory, maximum likelihood method, resolution

INTRODUCTION

In this note, two formulations of the *a posteriori* covariance using a maximum likelhood approach are considered. These formulations are then related to the model resolution. The first formulation is an extension of the least squares estimate of covariance in which the least squares inverse is replaced by the maximum likelihood inverse. This extension is given for example by Ellsworth & Koyanagi (1977) and Aki & Richards (1980). Using this formulation, as diagonal resolution goes to zero, *a posteriori* covariance also goes to zero. A maximum error occurs for model parameters with diagonal resolution of 0.5 and the error decreases for model parameters with either greater or smaller diagonal resolution. For this reason Ellsworth & Koyanagi (1977) chose to consider only model parameters with resolutions greater than 0.5 in their teleseismic tomography study. A second formulation of *a posteriori* covariance by Tarantola (1987) gives a more natural result which reduces to the *a priori* covariance in the limit of zero resolution. This behaviour results in a more useful estimate of model errors particularly for lower resolutions. In the case of perfect resolution, both formulations reduce to the least squares estimate. The relevance of this note is that many computer algorithms use the first formulation which will systematically underestimate computed errors relative to the second formulation.

COVARIANCE AND RESOLUTION

The linear forward problem can be written as $Am = d$, where for finite dimensional vector spaces A can be represented as a matrix, m as a model vector and d as a data vector. The maximum likelihood inverse can be written

$$A_{\text{ml}}^{-1} = (A^T C_D^{-1} A + C_M^{-1})^{-1} A^T C_D^{-1}. \tag{1a}$$

Or equivalently (see Aki & Richards, 1980, equation 12.123)

$$A_{\text{ml}}^{-1} = C_M A^T (A C_M A^T + C_D)^{-1}, \tag{1b}$$

where C_D is the observed data covariance, C_M is the *a priori* model covariance and A^T is the conjugate transpose. In the first formulation for the *a posteriori* covariance, C'_M, the least squares estimate of covariance is extended by replacing the least squares inverse by the maximum likelihood inverse, thus

$$C'_M = A_{\text{ml}}^{-1} C_D (A_{\text{ml}}^{-1})^T. \tag{2}$$

Using equation (1a), equation (2) can be written

$$C'_M = (A^T C_D^{-1} A + C_M^{-1})^{-1} A^T C_D^{-1} A (A^T C_D^{-1} A + C_M^{-1})^{-1}. \tag{3}$$

For the case when $C_M = \sigma_M^2 I$ and $C_D = \sigma_D^2 I$, then

$$C'_M = \sigma_D^2 (A^T A + \sigma_D^2/\sigma_M^2 I)^{-1} A^T A (A^T A + \sigma_D^2/\sigma_M^2 I)^{-1}.$$

For a fixed σ_D^2 and no *a priori* model information, then $\sigma_D^2/\sigma_M^2 \to 0$. In addition, if $A^T A$ is invertible (with no nullspace components for A), then this reduces to the standard least squares estimate of covariance

$$C'_M = \sigma_D^2 (A^T A)^{-1}$$

$(A^T A)^{-1}$ simply maps data errors into corresponding model errors.

The model resolution can be written as $R = A_{\mathrm{ml}}^{-1}A$. Using equation (1a)

$$R = (A^T C_D^{-1} A + C_M^{-1})^{-1} A^T C_D^{-1} A. \tag{4a}$$

For the case when $C_M = \sigma_M^2 I$ and $C_D = \sigma_D^2 I$, equation (4a) reduces to

$$R = (A^T A + \sigma_D^2/\sigma_M^2 I)^{-1} A^T A.$$

For a fixed σ_D^2 and no *a priori* model information, then $\sigma_D^2/\sigma_M^2 \to 0$. In addition, if $A^T A$ is invertible (with no nullspace components for A), then $R = I$, the standard least squares estimate of resolution.

By using equation (1b), the resolution can also be written

$$R = C_M A^T (A C_M A^T + C_D)^{-1} A. \tag{4b}$$

For the case when $C_M = \sigma_M^2 I$ and $C_D = \sigma_D^2 I$, then

$$R = A^T (A A^T + \sigma_D^2/\sigma_{2M} I)^{-1} A.$$

For a fixed σ_M^2 and $\sigma_D^2 \to 0$, then $\sigma_D^2/\sigma_M^2 \to 0$. In addition, if AA^T is invertible (with no perpendicular range space for A), then this reduces to the standard minimum length resolution estimate, $R = A^T(AA^T)A$. For this case $(I - R)$ is the projection onto the null space of A in model space.

The *a posteriori* covariance, C'_M, using equation (2) can be written in terms of the resolution operator R in equation (4a or 4b) as

$$C'_M = R(I - R)C_M. \tag{5}$$

This formula is less useful than equation (2) in the least squares limit where $C_M \to \infty$ and $R \to I$, resulting in possible numerical difficulties of a small number times a large number. The implications of equation (5) are discussed by Ellsworth & Koyanagi (1977) and Aki & Richards (1980, problem 12.9). For example, as the diagonal resolution goes from 0.5 to zero, the resulting errors also tend toward zero. Using this first formulation, a maximum resulting model error results for a resolution of 0.5 and decreases for both smaller and larger resolutions.

In the second formulation for *a posteriori* covariance by Tarantola (1987) the form of the *a posteriori* probability density function is derived (see his equation 1.65). Assuming a linear forward problem and Gaussian statistics for the data and the *a priori* model parameters, the *a posteriori* probability density function can be written

$$PDF'_M = \text{const} \exp\left[-\tfrac{1}{2}[(Am - d_{\mathrm{obs}})^T C_D^{-1}(Am - d_{\mathrm{obs}}) + (m - m_{\mathrm{prior}})^T C_m^{-1}(m - m_{\mathrm{prior}})]\right],$$

where d_{obs} is the observed data vector and m_{prior} is the *a priori* model vector. By defining

$$\langle m \rangle = (A^T C_D^{-1} A + C_M^{-1})^{-1}(A^T C_D^{-1} d_{\mathrm{obs}} + C_M^{-1} m_{\mathrm{prior}})$$

and

$$C'_M = (A^T C_D^{-1} A + C_M^{-1})^{-1} \tag{6}$$

the *a posteriori* probability density function can be rewritten as

$$PDF'_M = \text{const} \exp\{-\tfrac{1}{2}(m - \langle m \rangle)^T C'^{-1}_M (m - \langle m \rangle)\}.$$

The *a posteriori* probability density is again Gaussian with mean $\langle m \rangle$ and *a posteriori* covariance C'_M given by equation (6). For the case when $C_M = \sigma_M^2 I$ and $C_D = \sigma_D^2 I$, then

$$C'_M = \sigma_D^2 (A^T A + \sigma_D^2/\sigma_M^2 I)^{-1}.$$

For a fixed σ_D^2 and no *a priori* model information, then $\sigma_D^2/\sigma_M^2 \to 0$. In addition, if $A^T A$ is invertible (with no nullspace components for A), then this reduces to $C'_M = \sigma_D^2 (A^T A)^{-1}$, the standard least squares estimate.

Assuming a resolution operator of the same form as equation (4a or 4b), then the *a posteriori* covariance operator in equation (6) using the second formulation of Tarantola (1987, equation 7.165) can be written

$$C'_M = (I - R)C_M. \tag{7}$$

Note however, that Tarantola (1987, p. 494) then gives the incorrect interpretation that when $R \approx I$ then $C'_M \approx 0$. In fact, this is only true if, for fixed σ_M^2, σ_D^2 tends to zero (perfect data). If, instead, σ_D^2 is fixed, $R \approx I$ only when $\sigma_M^2 \to \infty$, and then, the standard least squares estimate of C'_M results, which is in general nonzero. Using equation (7), as the diagonal resolution decreases from 0.5 to zero, or as $R \to 0$, the *a posteriori* covariance tends toward the *a priori* covariance.

A final expression for the maximum likelihood *a posteriori* covariance is given by Menke (1984, equation 5.23) who in turn refers to Tarantola & Valette (1982). This is written as the following linear combination

$$C'_M = A_{\mathrm{ml}}^{-1} C_D (A_{\mathrm{ml}}^{-1})^T + (I - R)C_M(I - R)^T. \tag{8}$$

The first term on the right hand side is just that used in the first formulation above, where in the least squares limit this maps data errors into model space. In the minimum norm limit (with no perpendicular range space for A), the second term in equation (8) projects the *a priori* model errors onto the nullspace of A. In the general maximum likelihood case, the situation is more complicated with C_D and C_M included in both terms of equation (8). Although Menke (1984) does not elaborate, equation (8) can be reduced using equation (5) for the first term to the simpler expression, $C'_M = (I - R)C_M$, and correspondingly equation (6). Thus equation (8) is equivalent to the second formulation and is primarily useful in comparing the two formulations.

CONCLUSION

Two formulations for *a posteriori* covariance using maximum likelihood are considered. These are then related to model resolution. The first formulation is an extension of the least squares estimate of covariance where the least squares inverse is replaced by the maximum likelihood inverse. The *a posteriori* covariance from the first formulation is given by equation (3) or in terms of resolution by equation (5). For imperfect resolution, this formulation gives a maximum error for resolutions of 0.5 and decreases for both greater or smaller resolutions. The second formulation for *a posteriori* covariance by Tarantola (1987) is given by equation (6) or in terms of the resolution operator by equation (7). This formulation gives more natural results for lower resolutions. In particular as resolution goes to zero, the *a posteriori* covariance tends to the *a priori* covariance. Thus the inverse problem using the second formulation does nothing to reduce the uncertainty for model parameters with zero resolution. In the least squares limit, both formulations agree. Since many computer codes use algorithms similar to the first formulation, model errors may be underestimated as compared to the second formulation for all cases except perfect resolution.

ACKNOWLEDGMENT

This work was supported by NSF Grant No. EAR-8518147.

REFERENCES

Aki, K. & Richards, P., 1980. *Quantitative Seismology: Theory and Methods,* W. H. Freeman, San Francisco.

Ellsworth, W. L. & Koyanagi, R. Y., 1977. Three-dimensional crust and mantle structure of Kilauea volcano, Hawaii, *J. geophys. Res.,* **82,** 5379–5394.

Menke, W., 1984. *Geophysical Data Analysis: Discrete Inverse Theory,* Academic Press, Orlando, FA.

Tarantola, A., 1987. *Inverse Problem Theory,* Elsevier, Amsterdam.

Tarantola, A. & Valette, B. 1982. Generalized non-linear problems solved using the least-squares criterion, *Rev. Geophys. Space Phys.* **20,** 219–232.