Stanford Exploration Project, Report 77, November 17, 1997, pages 1–185

Resolution for Lanczos and Paige-Saunders inverses in tomography

James G. Berryman[1]

keywords: *resolution, least squares, inversion, modeling*

ABSTRACT

Methods have been developed for computing both the model and data resolution matrices for iterative inverses such as those produced by the Lanczos and Paige-Saunders (LSQR) schemes for constructing linear inverses.

INTRODUCTION

The central role of resolution in inversion has been emphasized by Backus and Gilbert (1968), who base their general inversion methods on techniques designed to optimize the resolution of the resulting geophysical model obtained from processed data. Indeed, any inverse method based on incomplete or imperfect data should be expected to produce a filtered or blurred image of the object or region of interest. Apparent resolution is therefore an important figure of merit for imaging and inversion schemes, since the user of such schemes will eventually want to know what is the smallest object that can be distinguished.

Tomographic reconstruction schemes have well-defined resolution properties (Jackson, 1972). Yet, these properties are seldom computed because of the common misconception (Nolet, 1985) that resolution matrices can only be computed using singular value decomposition, which is generally the most computationally intensive of all matrix inversion methods. Other matrix inversion methods such as conjugate gradients, conjugate directions, Lanczos, and LSQR (van der Sluis and van der Vorst, 1987) are more commonly used in tomography codes, because of their smaller storage and computing requirements. It would therefore clearly improve the state of our knowledge about tomographic resolution in practice if methods for computing resolution matrices were available for all (or at least for the most common) iterative matrix inversion schemes.

I show in this paper how the resolution can be computed for two matrix inversion

[1]**email:** not available

schemes (Lanczos and LSQR). The techniques developed here can be generalized to other iterative inversion schemes as well.

RESOLUTION MATRICES

Resolution matrices for linear inversion problems can be understood most easily by considering a matrix equation of the form

$$\mathbf{M}\mathbf{s} = \mathbf{t} \tag{1}$$

and asking the question: Given $\mathbf{M}$ and $\mathbf{t}$, what $\mathbf{s}$ solves this equation? When the matrix $\mathbf{M}$ is square and invertible, the answer to the question is relatively easy: $\mathbf{s} = \mathbf{M}^{-1}\mathbf{t}$. However, it often happens in geophysical inversion problems that $\mathbf{M}$ is not square, or not invertible even if it is square. In these situations, I need to introduce an approximate inverse $\mathbf{M}^{\dagger}$ called the Moore-Penrose pseudoinverse (Moore, 1920; Penrose, 1955a; Penrose, 1955b). (Although other choices of approximate inverse are known [see Rao (1965)], I will restrict discussion here to the best known approximate inverse.) Then, when I multiply (1) on the left by $\mathbf{M}^{\dagger}$, I find

$$\mathbf{M}^{\dagger}\mathbf{M}\mathbf{s} = \mathbf{M}^{\dagger}\mathbf{t}. \tag{2}$$

If it were true that $\mathbf{M}^{\dagger}\mathbf{M} = \mathbf{I}$ (the identity matrix), then I would have solved the inversion problem exactly. But it is precisely because no such inverse exists in some problems that I need to consider the analysis that follows. Following Backus and Gilbert (1968), I define the matrix coefficient of $\mathbf{s}$ in (2) as the resolution matrix

$$\mathcal{R} \equiv \mathbf{M}^{\dagger}\mathbf{M}. \tag{3}$$

The deviations of $\mathcal{R}$ from the identity matrix $\mathbf{I}$, given by the matrix $\mathbf{I} - \mathcal{R}$, determine the degree of distrust I should place on the components of the solution vector $\mathbf{s}$ that are most poorly resolved.

To be more explicit, consider a crosshole tomography problem. Then, $\mathbf{M}$ is an $m \times n$ ray-path matrix, $\mathbf{t}$ is an m-vector of first arrival traveltimes, and $\mathbf{s}$ is the slowness (inverse velocity) n-vector. I seek the slownesses $\mathbf{s}$ given the measured traveltimes in $\mathbf{t}$ and the estimates of the ray paths between source and receiver locations contained in the matrix $\mathbf{M}$ [see Berryman (1991)]. Then, the resolution matrix defined above is the model resolution, since the slowness vector is the desired model of acoustic wave slowness. I can also define a data resolution matrix. First, multiply (2) on the left by $\mathbf{M}$ so

$$\mathbf{M}\mathbf{M}^{\dagger}\mathbf{M}\mathbf{s} = \mathbf{M}\mathbf{M}^{\dagger}\mathbf{t}, \tag{4}$$

then compare (4) to (1) and note that the matrix product multiplying $\mathbf{t}$ should equal the identity matrix if the approximate inverse $\mathbf{M}^{\dagger}$ is a true inverse. Again, deviations of this matrix from the identity provide information about the degree to which the solution I compute makes use of all the data in $\mathbf{t}$. Thus, I have the data resolution matrix (Wiggins, 1972; Jackson, 1972) is

$$\mathcal{R}_{data} \equiv \mathbf{M}\mathbf{M}^{\dagger} \tag{5}$$

and the model resolution matrix (Backus and Gilbert, 1968; Jackson, 1972) defined previously is

$$\mathcal{R}_{model} \equiv \mathbf{M}^{\dagger}\mathbf{M}. \tag{6}$$

I can clarify the significance of these two matrices by considering the singular value decomposition (SVD) of the matrix $\mathbf{M}$, given by

$$\mathbf{M} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \mathbf{v}_i^T, \tag{7}$$

where the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are the eigenvectors of $\mathbf{M}$ determined by

$$\mathbf{M}\mathbf{v}_i = \lambda_i \mathbf{u}_i \quad \text{and} \quad \mathbf{u}_i^T \mathbf{M} = \lambda_i \mathbf{v}_i^T \tag{8}$$

and the λ_is are the eigenvalues. The eigenvectors also satisfy orthonormality conditions $\mathbf{u}_i^T\mathbf{u}_j = \delta_{ij}$ and $\mathbf{v}_i^T\mathbf{v}_j = \delta_{ij}$. The rank r of $\mathbf{M}$ (the number of nonzero eigenvalues) has a value $r \le m, n$. The Moore-Penrose pseudoinverse is then known to be given by

$$\mathbf{M}^{\dagger} = \sum_{i=1}^{r} \lambda_i^{-1} \mathbf{v}_i \mathbf{u}_i^T, \tag{9}$$

so the resolution matrices are

$$\mathcal{R}_{model} = \sum_{i=1}^{r} \mathbf{v}_i \mathbf{v}_i^T \tag{10}$$

and

$$\mathcal{R}_{data} = \sum_{i=1}^{r} \mathbf{u}_i \mathbf{u}_i^T. \tag{11}$$

When expressed in this form, it is clear that the resolution matrices simply determine the completeness of the resolved model or data spaces respectively.

Now it is important to recognize that, although the resolution matrices are *defined* by the equations (6) and (5), these matrices may nevertheless be *computed* in other ways. For example, consider the SVD of the normal matrix (from least-squares analysis)

$$\mathbf{M}^T\mathbf{M} = \sum_{i=1}^{r} \lambda_i \mathbf{v}_i \mathbf{u}_i^T \sum_{j=1}^{r} \lambda_j \mathbf{u}_j \mathbf{v}_j^T = \sum_{i=1}^{r} \lambda_i^2 \mathbf{v}_i \mathbf{v}_i^T. \tag{12}$$

Then, I find easily that

$$\mathcal{R}_{model} = \left(\mathbf{M}^T\mathbf{M}\right)^\dagger \mathbf{M}^T\mathbf{M} = \sum_{i=1}^{r} \mathbf{v}_i\mathbf{v}_i^T. \tag{13}$$

It is also straightforward to show that

$$\mathcal{R}_{model} = \mathbf{M}^T \left(\mathbf{M}\mathbf{M}^T\right)^\dagger \mathbf{M}. \tag{14}$$

Similarly, I find that the data resolution is given by

$$\mathcal{R}_{data} = \mathbf{M}\mathbf{M}^\dagger = \mathbf{M}\mathbf{M}^T \left(\mathbf{M}\mathbf{M}^T\right)^\dagger = \mathbf{M}\left(\mathbf{M}^T\mathbf{M}\right)^\dagger \mathbf{M}^T. \tag{15}$$

In the following discussion, I will show how these equivalent formulas for the resolution matrices aid in their computation during the process of finding an approximate inverse of $\mathbf{M}$ using iterative methods.

METHOD OF LANCZOS

One common method of solving linear inversion problems of the form $\mathbf{Ms} = \mathbf{t}$ is Lanczos's method (Lanczos, 1950; Golub and Van Loan, 1983; van der Sluis and van der Vorst, 1987). This method introduces a sequence of orthonormal vectors $\mathbf{z}^{(k)}$ through a process of tridiagonalization. Applying this method to the normal equations $\mathbf{M}^T\mathbf{Ms} = \mathbf{M}^T\mathbf{t}$, Lanczos's method is a projection procedure equivalent to the following:

$$\mathbf{z}^{(1)} \left(\mathbf{z}^{(1)}\right)^T \mathbf{M}^T\mathbf{t} = \mathbf{M}^T\mathbf{t}, \tag{16}$$

$$\left[\mathbf{z}^{(1)} \left(\mathbf{z}^{(1)}\right)^T + \mathbf{z}^{(2)} \left(\mathbf{z}^{(2)}\right)^T\right] \mathbf{M}^T\mathbf{M}\mathbf{z}^{(1)} = \mathbf{M}^T\mathbf{M}\mathbf{z}^{(1)}, \tag{17}$$

and, for $k \geq 2$,

$$\left[\mathbf{z}^{(k-1)} \left(\mathbf{z}^{(k-1)}\right)^T + \mathbf{z}^{(k)} \left(\mathbf{z}^{(k)}\right)^T + \mathbf{z}^{(k+1)} \left(\mathbf{z}^{(k+1)}\right)^T\right] \mathbf{M}^T\mathbf{M}\mathbf{z}^{(k)} = \mathbf{M}^T\mathbf{M}\mathbf{z}^{(k)}. \tag{18}$$

It is now clear that (16) defines $\mathbf{z}^{(1)}$ as the unit vector in the direction of $\mathbf{M}^T\mathbf{t}$, and can have no terms from the right null space of $\mathbf{M}$. Equation (17) defines $\mathbf{z}^{(2)}$ as the unit vector found by removing the component of $\mathbf{M}^T\mathbf{M}\mathbf{z}^{(1)}$ in the direction of $\mathbf{z}^{(1)}$ and then normalizing. Equation (18) determines $\mathbf{z}^{(k+1)}$ as the unit vector along the component of $\mathbf{M}^T\mathbf{M}\mathbf{z}^{(k)}$ that is orthogonal to both $\mathbf{z}^{(k)}$ and $\mathbf{z}^{(k-1)}$. By construction,

$$\left(\mathbf{z}^{(i)}\right)^T \mathbf{z}^{(j)} = \delta_{ij}, \tag{19}$$

so the vectors are orthonormal.

Defining the constants

$$N_1 = \left|\mathbf{M}^T\mathbf{t}\right| = \left(\mathbf{z}^{(1)}\right)^T \mathbf{M}^T\mathbf{t}, \tag{20}$$

$$D_k = \left(\mathbf{z}^{(k)}\right)^T \mathbf{M}^T\mathbf{M}\mathbf{z}^{(k)} \quad \text{for} \quad k = 1, 2, \ldots, \tag{21}$$

and

$$N_{k+1} = \left(\mathbf{z}^{(k+1)}\right)^T \mathbf{M}^T\mathbf{M}\mathbf{z}^{(k)} \quad \text{for} \quad k = 1, 2, \ldots, \tag{22}$$

it is then clear that the equations (16)–(18) determine a tridiagonal system of the form

$$\begin{pmatrix} 0 & \ldots & 0 & \mathbf{z}^{(k+1)} \end{pmatrix} N_{k+1} + \mathbf{Z}_k \begin{pmatrix} D_1 & N_2 & & & \\ N_2 & D_2 & N_3 & & \\ & N_3 & D_3 & N_4 & \\ & & \ddots & \ddots & \ddots \\ & & & N_k & D_k \end{pmatrix} = \mathbf{M}^T\mathbf{M}\mathbf{Z}_k \quad \text{for} \quad 2 \le k \le r, \tag{23}$$

where the orthogonal matrix composed of the resulting orthonormal vectors is

$$\mathbf{Z}_k = \begin{pmatrix} \mathbf{z}^{(1)} & \mathbf{z}^{(2)} & \mathbf{z}^{(3)} & \ldots & \mathbf{z}^{(k)} \end{pmatrix}. \tag{24}$$

The process stops when $k = r$ (the rank of the matrix) because then $N_{r+1} = 0$, or is numerically negligible. It follows that this tridiagonalization process results in the identity

$$\mathbf{M}^T\mathbf{M} = \mathbf{Z}_r\mathbf{T}_r\mathbf{Z}_r^T, \tag{25}$$

where the tridiagonal matrix of coefficients is defined by

$$\mathbf{T}_k = \begin{pmatrix} D_1 & N_2 & & & \\ N_2 & D_2 & N_3 & & \\ & N_3 & D_3 & N_4 & \\ & & \ddots & \ddots & \ddots \\ & & & N_k & D_k \end{pmatrix} \quad \text{for} \quad 2 \le k \le r. \tag{26}$$

Since $\mathbf{T}_r$ is invertible (by the definition of the rank r of $\mathbf{M}$),

$$\left(\mathbf{M}^T\mathbf{M}\right)^\dagger = \mathbf{Z}_r \left(\mathbf{T}_r\right)^{-1} \mathbf{Z}_r^T. \tag{27}$$

The solution to the least-squares inversion problem may therefore be written as

$$\mathbf{s} = \mathbf{Z}_r \left(\mathbf{T}_r\right)^{-1} \mathbf{Z}_r^T\mathbf{M}^T\mathbf{t} = N_1\mathbf{Z}_r \left(\mathbf{T}_r\right)^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{28}$$

where I used (16), (20), and the fact that

$$\mathbf{Z}_r^T \mathbf{z}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{29}$$

It follows that only the first column of the inverse of $\mathbf{T}_r$ is needed to solve this inversion problem.

Since Lanczos's method directly produces a sequence of orthonormal vectors in the model space, it is straightforward to see that the model resolution matrix for this method is given by

$$\mathcal{R}_{model} = \mathbf{M}^T\mathbf{M}\left(\mathbf{M}^T\mathbf{M}\right)^{\dagger} = \mathbf{Z}_r\mathbf{Z}_r^T = \sum_{k=1}^{r} \mathbf{z}^{(k)}\left(\mathbf{z}^{(k)}\right)^T, \tag{30}$$

which is also clearly symmetric. It is however more difficult to compute the data resolution matrix. Using the fact that

$$\mathcal{R}_{data} = \mathbf{M}\mathbf{M}^{\dagger} = \mathbf{M}\left(\mathbf{M}^T\mathbf{M}\right)^{\dagger}\mathbf{M}^T \tag{31}$$

together with (27), I find that

$$\mathcal{R}_{data} = \mathbf{M}\mathbf{Z}_r\left(\mathbf{T}_r\right)^{-1}\mathbf{Z}_r^T\mathbf{M}^T. \tag{32}$$

It is clear that both of the resolution matrices are symmetric when the full Lanczos inverse has been computed. If the process is terminated early so that $k < r$ by setting the constant N_{k+1} equal to zero, then a Lanczos approximate inverse is given by

$$\mathbf{X}_k = \mathbf{Z}_k\left(\mathbf{T}_k\right)^{-1}\mathbf{Z}_k^T\mathbf{M}^T, \tag{33}$$

so that $\mathbf{X}_k\mathbf{M}\mathbf{X}_k = \mathbf{X}_k$ for all k, but $\mathbf{M}\mathbf{X}_k\mathbf{M} \neq \mathbf{M}$ if $k < r$. The effective resolution matrices are given by $\mathcal{E}_{model} = \mathbf{X}_k\mathbf{M}$ and $\mathcal{E}_{data} = \mathbf{M}\mathbf{X}_k$, or are found by replacing $\mathbf{Z}_r$ and $\mathbf{T}_r$ in (30) and (32) by $\mathbf{Z}_k$ and $\mathbf{T}_k$, respectively. Clearly, the effective resolutions are also symmetric.

An alternative method of computing the resolution matrices involves noting that the vector sequence $\mathbf{z}^{(k)}$ may also be used as the starting set of vectors for the method of conjugate directions. Then, I can produce a set of conjugate vectors from this orthogonal set and easily compute both the model resolution matrix and the data resolution matrix if desired. This alternative requires additional work, however, to produce the sequence of conjugate vectors and it also generally produces a model resolution that is not symmetric, and therefore difficult to interpret. Thus, the Lanczos procedure appears to be superior to conjugate directions/gradients from this point of view. The main disadvantage of the Lanczos method is that the amount of storage increases with the iteration number k, since all the vectors $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(k)}$ must be stored until the final calculation of $\mathbf{s}^{(k)} = \mathbf{X}_k\mathbf{t}$. In contrast, conjugate gradients requires a fixed amount of storage, but an easily interpretable model resolution must be sacrificed to gain this advantage.

LSQR ALGORITHM OF PAIGE AND SAUNDERS

The method of Lanczos may be applied to any square, symmetric matrix inversion problem. To solve the least-squares inversion problem, I applied the method to the normal matrix $\mathbf{M}^T\mathbf{M}$ in the preceding section. But, I could have "completed the square" with $\mathbf{M}$ and considered instead the problem

$$\begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{O} \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{s} \end{pmatrix} = \begin{pmatrix} \mathbf{t} \\ 0 \end{pmatrix}, \tag{34}$$

where $\mathbf{r} = \mathbf{t} - \mathbf{Ms}$ is the residual vector. The resulting modification of Lanczos's method is the LSQR algorithm of Paige and Saunders (1982).

To derive this algorithm, first apply Lanczos's method to (34) and divide the $(m+n)$-vector $\mathbf{z}^{(k)}$ into the two pieces: $\mathbf{g}^{(k)}$ in the m-dimensional data space and $\mathbf{h}^{(k)}$ in the n-dimensional model space. Thus,

$$\mathbf{z}^{(k)} = \begin{pmatrix} \mathbf{g}^{(k)} \\ \mathbf{h}^{(k)} \end{pmatrix}. \tag{35}$$

Then, the tridiagonalization process takes the form

$$\mathbf{z}^{(1)} \left(\mathbf{z}^{(1)}\right)^T \begin{pmatrix} \mathbf{t} \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{t} \\ 0 \end{pmatrix}, \tag{36}$$

$$\left[\mathbf{z}^{(1)} \left(\mathbf{z}^{(1)}\right)^T + \mathbf{z}^{(2)} \left(\mathbf{z}^{(2)}\right)^T\right] \begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{O} \end{pmatrix} \mathbf{z}^{(1)} = \begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{O} \end{pmatrix} \mathbf{z}^{(1)}, \tag{37}$$

and, for $k > 2$,

$$\begin{aligned} \left[\mathbf{z}^{(k-1)} \left(\mathbf{z}^{(k-1)}\right)^T + \mathbf{z}^{(k)} \left(\mathbf{z}^{(k)}\right)^T + \mathbf{z}^{(k+1)} \left(\mathbf{z}^{(k+1)}\right)^T\right] \begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{O} \end{pmatrix} \mathbf{z}^{(k)} \\ = \begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{O} \end{pmatrix} \mathbf{z}^{(k)}. \end{aligned} \tag{38}$$

Solving equations (36)–(38) after substituting (35) shows that

$$\mathbf{g}^{(1)} \left(\mathbf{g}^{(1)}\right)^T \mathbf{t} = \mathbf{t} \quad \text{and} \quad \mathbf{h}^{(1)} = 0, \tag{39}$$

$$\mathbf{g}^{(2)} = 0 \quad \text{and} \quad \mathbf{h}^{(2)} \left(\mathbf{h}^{(2)}\right)^T \mathbf{M}^T\mathbf{g}^{(1)} = \mathbf{M}^T\mathbf{g}^{(1)}, \tag{40}$$

and, for $k \geq 1$,

$$\left[\mathbf{g}^{(2k-1)} \left(\mathbf{g}^{(2k-1)}\right)^T + \mathbf{g}^{(2k+1)} \left(\mathbf{g}^{(2k+1)}\right)^T\right] \mathbf{Mh}^{(2k)} = \mathbf{Mh}^{(2k)} \quad \text{and} \quad \mathbf{h}^{(2k+1)} = 0, \tag{41}$$

and

$$\mathbf{g}^{(2k+2)} = 0 \quad \text{and} \quad \left[\mathbf{h}^{(2k)} \left(\mathbf{h}^{(2k)}\right)^T + \mathbf{h}^{(2k+2)} \left(\mathbf{h}^{(2k+2)}\right)^T\right] \mathbf{M}^T\mathbf{g}^{(2k+1)} = \mathbf{M}^T\mathbf{g}.^{(2k+1)} \tag{42}$$

Thus, the tridiagonal system of Lanczos has been reduced to two coupled bidiagonal systems in LSQR.

Defining the constants

$$b_1 = |\mathbf{t}| = \left(\mathbf{g}^{(1)}\right)^T \mathbf{t}, \tag{43}$$

and, for $k \geq 1$,

$$b_{2k} = \left(\mathbf{h}^{(2k)}\right)^T \mathbf{M}^T \mathbf{g}^{(2k-1)} = \left(\mathbf{g}^{(2k-1)}\right)^T \mathbf{M}\mathbf{h}^{(2k)}, \tag{44}$$

and

$$b_{2k+1} = \left(\mathbf{h}^{(2k)}\right)^T \mathbf{M}^T \mathbf{g}^{(2k+1)} = \left(\mathbf{g}^{(2k+1)}\right)^T \mathbf{M}\mathbf{h}^{(2k)}. \tag{45}$$

Then, the orthogonal matrix of the vectors $\mathbf{g}^{(2k+1)}$ is defined by

$$\mathbf{G}_k = \left(\mathbf{g}^{(1)} \quad \mathbf{g}^{(3)} \quad \dots \quad \mathbf{g}^{(2k-1)} \right), \tag{46}$$

while the corresponding matrix of the vectors $\mathbf{h}^{(2k)}$ is defined by

$$\mathbf{H}_k = \left(\mathbf{h}^{(2)} \quad \mathbf{h}^{(4)} \quad \dots \quad \mathbf{h}^{(2k)} \right). \tag{47}$$

Now the two bidiagonal systems may be written, for $2 \leq k \leq r$, as

$$\mathbf{H}_k \begin{pmatrix} b_2 & b_3 & & & \\ & b_4 & b_5 & & \\ & & b_6 & b_7 & \\ & & & \ddots & \ddots \\ & & & & b_{2k} \end{pmatrix} = \mathbf{M}^T \mathbf{G}_k, \tag{48}$$

and

$$\left(0 \quad \dots \quad 0 \quad \mathbf{g}^{(k+1)} \right) b_{2k+1} + \mathbf{G}_k \begin{pmatrix} b_2 & & & & \\ b_3 & b_4 & & & \\ & b_5 & b_6 & & \\ & & \ddots & \ddots & \\ & & & b_{2k-1} & b_{2k} \end{pmatrix} = \mathbf{M}\mathbf{H}_k. \tag{49}$$

Then, introducing the $k \times k$ lower bidiagonal matrix

$$\mathbf{B}_k = \begin{pmatrix} b_2 & & & & \\ b_3 & b_4 & & & \\ & b_5 & b_6 & & \\ & & \ddots & \ddots & \\ & & & b_{2k-1} & b_{2k} \end{pmatrix}, \tag{50}$$

it is clear that, when $b_{2r+1} = 0$ (or is numerically negligible), multiplying (49) on the right by $\mathbf{H}_r^T$ yields the formal relation[2]

$$\mathbf{M} \simeq \mathbf{G}_r \mathbf{B}_r \mathbf{H}_r^T, \tag{51}$$

and also that

$$\mathbf{X}_r = \mathbf{H}_r \left(\mathbf{B}_r\right)^{-1} \mathbf{G}_r^T. \tag{52}$$

The coupled equations (48) and (49) are equivalent to

$$\mathbf{H}_r \mathbf{B}_r^T \mathbf{B}_r = \mathbf{M}^T \mathbf{M} \mathbf{H}_r, \tag{53}$$

which should be compared to (23). If the tridiagonal matrix $\mathbf{T}_r$ satisfies

$$\mathbf{T}_r = \mathbf{B}_r^T \mathbf{B}_r \tag{54}$$

(so that $D_k = b_{2k}^2 + b_{2k+1}^2$ and $N_k = b_{2k-1} b_{2k}$), then Lanczos's orthogonal matrix for the least-squares problem satisfies $\mathbf{Z}_r = \mathbf{H}_r$.

The solution of the inversion problem may now be written as

$$\mathbf{s} = \mathbf{X}_r \mathbf{t} = \mathbf{H}_r \left(\mathbf{B}_r\right)^{-1} \mathbf{G}_r^T \mathbf{t} = b_1 \mathbf{H}_r \left(\mathbf{B}_r\right)^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{55}$$

where I used (39), (43), and the fact that

$$\mathbf{G}_r^T \mathbf{g}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{56}$$

It follows that only the first column of the inverse of $\mathbf{B}_r$ is needed to solve the inversion problem.

Once I have the approximate generalized inverse via (52), I can write the model resolution matrix as

$$\mathcal{R}_{model} = \mathbf{M}^\dagger \mathbf{M} = \mathbf{X}_r \mathbf{M} = \mathbf{H}_r \mathbf{H}_r^T. \tag{57}$$

Although it is tempting to think that the data resolution is given similarly by

$$\mathcal{R}_{data} = \mathbf{M}\mathbf{M}^\dagger = \mathbf{M}\mathbf{X}_r \simeq \mathbf{G}_r \mathbf{G}_r^T, \tag{58}$$

this is not true in general! It is easy to see that

$$\mathbf{G}_r \mathbf{G}_r^T \mathbf{t} = \mathbf{t}, \tag{59}$$

[2]I explain later why this relation is formal and not exact.

by construction. Yet the absurd result seemingly implied by (59) that the data resolution is always perfect for this method cannot hold generally for $\mathcal{R}_{data}$, since it implies that the data can be perfectly matched *in all problems* (all possible choices of $\mathbf{M}$ and $\mathbf{t}$) which is clearly false. The difficulty is that the initial vector $\mathbf{g}^{(1)}$ in LSQR is proportional to the data vector itself, which may contain contributions from the data null space if $\mathbf{t}$ and $\mathbf{M}$ are inconsistent (*i.e.*, there may exist a vector $\mathbf{n}$ such that $\mathbf{n}^T\mathbf{M} = 0$ but $\mathbf{n}^T\mathbf{t} \neq 0$). Since the LSQR process never modifies this initial vector and never multiplies it by $\mathbf{M}$, these terms in the data null space never get stripped off during the LSQR computation. This difficulty is equivalent to the observation that $\mathbf{X}_r\mathbf{M}\mathbf{X}_r \neq \mathbf{X}_r$ for the LSQR algorithm if $\mathbf{t}$ and $\mathbf{M}$ are inconsistent.

Modified LSQR

I can resolve this difficulty at the end of the calculation by computing the data resolution using the formula

$$\mathcal{R}_{data} = \mathbf{M}\mathbf{X}_r = \mathbf{M}\mathbf{H}_r\left(\mathbf{B}_r\right)^{-1}\mathbf{G}_r^T, \tag{60}$$

instead of (58). Formula (60) implies that I actually multiply the matrices $\mathbf{M}$ and $\mathbf{X}_r$ to find the true resolution, instead of taking the inappropriate shortcut indicated in (58). It is not obvious whether (60) is symmetric or not, although I expect that it is. Another approach is to use the Moore-Penrose uniqueness condition for the generalized inverse to strip off the parts of $\mathbf{X}_r$ that lie in the null space of $\mathbf{M}$, *i.e.*,

$$\mathbf{M}^\dagger \equiv \mathbf{X}_r\mathbf{M}\mathbf{X}_r, \tag{61}$$

where again the matrix multiplies appearing in the formula must actually be executed in order to achieve the required stripping effect. I call (61) the *Modified LSQR* inverse. Once $\mathbf{M}^\dagger$ has been constructed this way, (55) is replaced by

$$\mathbf{s} = \mathbf{M}^\dagger\mathbf{t} = \mathbf{X}_r\mathbf{M}\mathbf{X}_r\mathbf{t} \tag{62}$$

and I can use the standard formula $\mathcal{R}_{data} = \mathbf{M}\mathbf{M}^\dagger$ to compute the fully symmetric data resolution matrix, if this is required.

This iterative procedure may be terminated early (*i.e.*, for $k < r$) by setting $b_{2k+1} = 0$. Then, an approximate LSQR inverse is given by

$$\mathbf{X}_k = \mathbf{H}_k\left(\mathbf{B}_k\right)^{-1}\mathbf{G}_k^T. \tag{63}$$

It is easy to check that $\mathbf{X}_k\mathbf{M}\mathbf{X}_k \neq \mathbf{X}_k$ and that $\mathbf{M}\mathbf{X}_k\mathbf{M} \neq \mathbf{M}$ if $k < r$. The effective model resolution matrix is given easily for the LSQR method by

$$\mathcal{E}_{model} = \mathbf{H}_k\mathbf{H}_k^T = \sum_{i=1}^{k}\mathbf{h}^{(2i)}\left(\mathbf{h}^{(2i)}\right)^T. \tag{64}$$

The same problems that were discussed above for $\mathcal{R}_{data}$ apply to $\mathcal{E}_{data}$ as well, so the formula for effective data resolution in this situation is

$$\mathcal{E}_{data} = \mathbf{M}\mathbf{X}_k = \mathbf{M}\mathbf{H}_k \left(\mathbf{B}_k\right)^{-1} \mathbf{G}_k^T. \tag{65}$$

Effective resolution $\mathcal{E}_{model}$ is clearly symmetric for any choice of k. I expect $\mathcal{E}_{data}$ is also symmetric, but it is again not immediately obvious from the formula. Similarly, the Modified LSQR for partial iteration then uses $\mathbf{X}_k\mathbf{M}\mathbf{X}_k$ as the inverse, rather than $\mathbf{X}_k$ itself.

Not surprisingly, when $\mathbf{t}$ and $\mathbf{M}$ are consistent, the LSQR algorithm has the same qualitative characteristics as the Lanczos method from which it derives. The main advantage of this method is that it produces the generalized inverse of $\mathbf{M}$ directly, rather than producing $\left(\mathbf{M}^T\mathbf{M}\right)^\dagger$. Thus, the dynamic range problems due to poor conditioning of $\mathbf{M}$ are aggravated when the Lanczos method is applied directly to the normal equations since the eigenvalues are squared in $\mathbf{M}^T\mathbf{M}$; in contrast, no such aggravation occurs in the LSQR algorithm. Both algorithms are qualitatively similar to truncated SVD, but they are easier to implement and more economical than many SVD routines. Of course, both methods can be turned into SVD routines simply by including an additional step to find the eigenvalues and eigenvectors of the tridiagonal matrix $\mathbf{T}_r = \mathbf{B}_r^T\mathbf{B}_r$. However, such a step is not necessary either to solve the inversion problem or to compute the resolution matrices.

One major difference between LSQR and the method of Lanczos is that storage is greater for LSQR by a factor of $(m+n)/n$, since both $\mathbf{G}_k$ and $\mathbf{H}_k$ must be stored, especially when the resolution matrices must be computed. This difference can be significant if the data dimension m is much greater than the model dimension n as it is for overdetermined problems.

LSQR FOR DAMPED LEAST-SQUARES

If the ray-path matrix $\mathbf{M}$ is very poorly conditioned, it may be advisable to modify the normal equations with a damping term. Then, (34) changes to

$$\begin{pmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{M}^T & -\mu\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ \mathbf{s} \end{pmatrix} = \begin{pmatrix} \mathbf{t} \\ 0 \end{pmatrix}, \tag{66}$$

where $\mathbf{r} = \mathbf{t} - \mathbf{M}\mathbf{s}$ is still the residual vector, but now the solution vector satisfies

$$\mathbf{s} = \left(\mathbf{M}^T\mathbf{M} + \mu\mathbf{I}\right)^{-1} \mathbf{M}^T\mathbf{t}. \tag{67}$$

The inverse matrix in (67) exists for any value of $\mu > 0$, since the eigenvalues of $\mathbf{M}^T\mathbf{M}$ are nonnegative and the addition of the diagonal term μ shifts all the eigenvalues up by μ.

Applying the Lanczos method to (66), I quickly discover that the terms proportional to μ always cancel out of the equations so the hard part of the analysis is the same as for the undamped problem. Thus, I find

$$\mathbf{M} \simeq \mathbf{G}_r \mathbf{B}_r \mathbf{H}_r^T \quad \text{and} \quad \mathbf{M}^\dagger \simeq \mathbf{H}_r \mathbf{B}_r^{-1} \mathbf{G}_r^T, \tag{68}$$

as before. The orthogonal matrices $\mathbf{G}_k$ and $\mathbf{H}_k$ are all identical to those for the undamped case.

Then, I must write the regularized normal matrix in terms of these orthogonal matrices and find its inverse. The resulting matrix is given by

$$\mathbf{M}^T\mathbf{M} + \mu\mathbf{I} = \mathbf{H}_r\mathbf{B}_r^T\mathbf{B}_r\mathbf{H}^T + \mu\mathbf{I} = \mathbf{H}_r\left(\mathbf{B}_r^T\mathbf{B}_r + \mu\mathbf{I}\right)\mathbf{H}_r^T + \mu\left(\mathbf{I} - \mathbf{H}_r\mathbf{H}_r^T\right). \tag{69}$$

The inverse matrix is easily shown to be

$$\left(\mathbf{M}^T\mathbf{M} + \mu\mathbf{I}\right)^{-1} = \mathbf{H}_r\left(\mathbf{B}_r^T\mathbf{B}_r + \mu\mathbf{I}\right)^{-1}\mathbf{H}_r^T + \frac{1}{\mu}\left(\mathbf{I} - \mathbf{H}_r\mathbf{H}_r^T\right). \tag{70}$$

Substituting this result into (67) gives

$$\mathbf{s} = \mathbf{H}_r\left(\mathbf{B}_r^T\mathbf{B}_r + \mu\mathbf{I}\right)^{-1}\mathbf{B}_r^T\mathbf{G}_r^T\mathbf{t} = b_1\mathbf{H}_r\left(\mathbf{B}_r^T\mathbf{B}_r + \mu\mathbf{I}\right)^{-1}\mathbf{B}_r^T\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \tag{71}$$

which should be compared to (55). If the data vector $\mathbf{t}$ and the ray-path matrix $\mathbf{M}$ are inconsistent, then (71) must be modified in analogy to the modification made in transforming from (55) to (62). Following the same analytical path in the Lanczos method for damped least-squares, the result is

$$\mathbf{s} = \mathbf{Z}_r\left(\mathbf{T}_r + \mu\mathbf{I}\right)^{-1}\mathbf{Z}_r^T\mathbf{M}^T\mathbf{t} = D_1\mathbf{Z}_r\left(\mathbf{T}_r + \mu\mathbf{I}\right)^{-1}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \tag{72}$$

which should be compared to (28).

The resulting resolution matrices are symmetric. The same results hold for partial versions of these formulas, where the process is terminated early at some $k < r$ and rs are replaced by ks everywhere in the formulas (71) and (72).

CONCLUSIONS

The methods of computing resolution matrices that have been illustrated here may be easily generalized to a variety of other iterative and approximate inversion methods. In many situations, all these algorithms need to be modified to compute only the diagonal components of these resolution matrices since the full matrices can be of enormous size. These other methods and related practical issues will be discussed in greater detail elsewhere.

ACKNOWLEDGMENTS

I thank F. Muir and W. W. Symes for helpful discussions.

REFERENCES

Backus, G., and Gilbert, F., 1968, The resolving power of gross earth data: *Geophys. J. R. Astron. Soc.*, **16**, 169–205.

Berryman, J. G., 1991, *Lecture Notes on Nonlinear Inversion and Tomography: I. Borehole Seismic Tomography*, LLNL UCRL-LR-105358, October, 1991.

Golub, G. H., and Van Loan, C. F., 1983, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, Chapters 9 and 10, pp. 322–379.

Jackson, D. D., 1972, Interpretation of inaccurate, insufficient and inconsistent data: *Geophys. J. R. Astron. Soc.*, **28**, 97–109.

Lanczos, C., 1950, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators: *J. Res. Nat. Bur. Stand.*, **45**, 255–282.

Moore, E. H., 1920, *Bull. Amer. Math. Soc.*, **26**, 394–395.

Nolet, G., 1985, Solving or resolving inadequate and noisy tomographic systems: *J. Comp. Phys.*, **61**, 463–482.

Paige, C. C., and Saunders, M. A., 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: *ACM Trans. Math. Software*, **8**, 43–71.

Penrose, R., 1955a, A generalized inverse for matrices: *Proc. Cambridge Philos. Soc.*, **51**, 406–413.

Penrose, R., 1955b, On best approximation solutions of linear matrix equations: *Proc. Cambridge Philos. Soc.*, **52**, 17–19.

Rao, C. R., 1965, *Linear Statistical Inference and Its Applications*, Wiley, New York, pp. 24–26.

van der Sluis, A., and van der Vorst, H. A., 1987, Numerical solution of large, sparse linear algebraic systems arising from tomographic problems: in *Seismic Tomography – With Applications in Global Seismology and Exploration Geophysics*, G. Nolet (ed.), Reidel, Dordrecht, Holland, Chapter 3, pp. 49–83.

Wiggins, R. A., 1972, The general linear inverse problem: Implications of surface waves and free oscillations for Earth structure: *Rev. Geophys. Space Phys.*, **10**, 251–285.