The generalized inverse approach to seismic tomography

Kamal Al-Yahya

INTRODUCTION

This is a supplement to another paper on seismic tomography (Al-Yahya, this SEP report). To keep the length of that paper reasonable, the generalized inverse approach is treated here separately. The reader may need to refer to that paper for more details.

Even though the generalized inverse approach is not practical for use in geophysical tomography, as we shall see, it sheds a light on some aspects of the problem, especially non-uniqueness. The reader is reminded that ray theory is used with its known limitations.

LINEARIZING THE PROBLEM

Let the vector $\mathbf{t}$ represent the direct-arrival travel times measured in an experiment. Each element in this vector is the direct-arrival travel time for a transmitter-receiver pair. Let the region of investigation be divided into nx by nz cells where each cell has a constant slowness (reciprocal of velocity) w_i. These cells can be put in a one-dimensional vector by a suitable ordering (e.g. by the stacking of columns on top of each other). If we have ns transmitters and ng receivers, the relationship between the model vector $\mathbf{w}$ and the data vector $\mathbf{t}$ can be written in matrix form as

$$\mathbf{Lw} = \mathbf{t} \ , \tag{1}$$

where $\mathbf{L}$ is an $ns \times ng$ by $nx \times nz$ matrix. Each row of $\mathbf{L}$ contains elements representing

the lengths of the ray in the corresponding cells. If the ray does not cross that cell the element is understood to be zero.

In equation (1), we want to solve for the slowness $\mathbf{w}$. An essential problem in equation (1) is that there are two unknowns, $\mathbf{L}$ and $\mathbf{w}$. It is common practice to assume that $\mathbf{L}$ is known by linearizing equation (1). To do this, we solve not for $\mathbf{w}$, but for $\Delta\mathbf{w}=\mathbf{w}-\mathbf{w}_0$, where $\mathbf{w}_0$ is a guessed solution, and assume that $\mathbf{L}$ is known by making it a function of $\mathbf{w}_0$. This assumption is reasonable only if the guessed model is reasonably close to the actual one. For an arbitrary model, the difference between the ray path calculated with the actual velocity and the path calculated with a constant velocity is shown in Figure 1. For this model, the difference between the two paths is not big and the approximation is valid. It must be noted however that these two paths can be very far from each other. After solving for $\Delta\mathbf{w}$ we can recompute $\mathbf{L}$ for the new model and again solve for $\Delta\mathbf{w}$ and so on until the solution converges. In a given iteration, our objective is to solve

$$\mathbf{L}\ \Delta\mathbf{w} \ = \ \Delta\mathbf{t}\ , \tag{2}$$

where $\Delta\mathbf{t}=\mathbf{t}-\mathbf{t}_0$ and $\mathbf{t}_0$ is the data corresponding to $\mathbf{w}_0$.

SINGULAR VALUE DECOMPOSITION

Many well known methods of solving the linear system (2) exist. This short paper looks at the singular value decomposition of equation (2). This decomposition is the core of the generalized inverse (or *pseudo*inverse) solution. We note that $\mathbf{L}$ is generally not a square matrix. It can also be singular. It is therefore natural to use the generalized inverse solution to equation (2),

$$\Delta\mathbf{w} = \mathbf{L}_g^{-1}\Delta\mathbf{t} \tag{3}$$

where $\mathbf{L}_g$ is the pseudoinverse of $\mathbf{L}$. For a purely over-specified system, $\mathbf{L}_g$ is the familiar least squares operator $(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\Delta\mathbf{t}$, which minimizes the *square error* $|\Delta\mathbf{t}-\mathbf{L}\mathbf{w}|^2$. However, as we shall see later, in borehole tomography the system is typically both *over*-specified and *under*-determined. In this case, the solution of equation (2) is that which minimizes both the square error $|\Delta\mathbf{t}-\mathbf{L}\mathbf{w}|^2$ and the Euclidean norm $|\Delta\mathbf{w}|^2$.

The direct implementation of equation (3) is not practical, because $\mathbf{L}$ is usually very large. A typical experiment can involve 50 transmitters and 50 receivers. If we divide the region of investigation into 100 by 100 cells, then $\mathbf{L}$ is a 2500 by 10000 matrix! Even though $\mathbf{L}$ is a sparse matrix, it does not have a special structure that alleviates the problem of its size.

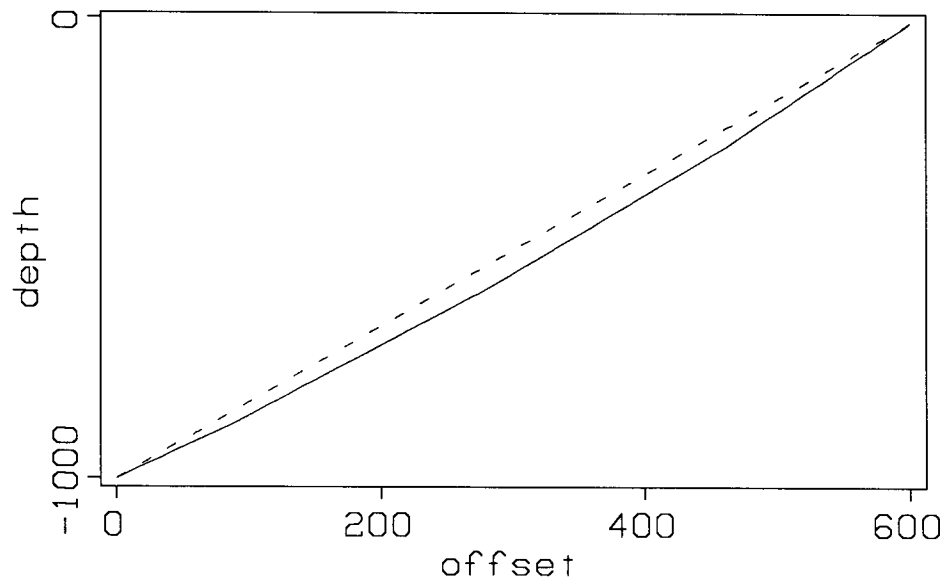

FIG. 1. The ray path between two points. The solid line is the actual one; the dashed line is approximate, made under the assumption of constant velocity.

To understand the nature of the solution given by equation (3) I will perform a singular value decomposition of the matrix $\mathbf{L}$; that is, I write $\mathbf{L}$ as

$$\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$$

where $\mathbf{\Lambda}$ is a diagonal matrix having the eigenvalues λ of these two eigenvalue problems,

$$\mathbf{L}^T\mathbf{L}\mathbf{v}_i = \lambda^2\mathbf{v}_i$$

$$\mathbf{L}\mathbf{L}^T\mathbf{u}_i = \lambda^2\mathbf{u}_i \quad ,$$

in decreasing order along its diagonal, Aki and Richards (1980). $\mathbf{U}$ and $\mathbf{V}$ are matrices whose columns are the corresponding eigenvectors $\mathbf{u}_i$ and $\mathbf{v}_i$ respectively.

TWO EXAMPLES

Let's study two small experiments in which the size of $\mathbf{L}$ is not an issue. The first experiment has two wells 10 meters apart and each 10 meters deep. Let's divide the medium into a 10 by 10 grid of square cells, that is 100 cells, one square meter each, and record with 10 transmitters and 10 receivers, placed at intervals of one meter, starting from one half meter from the surface (Figure 2a). The second experiment has almost the same geometry, with one variation: the transmitters are at the surface (Figure 2b).

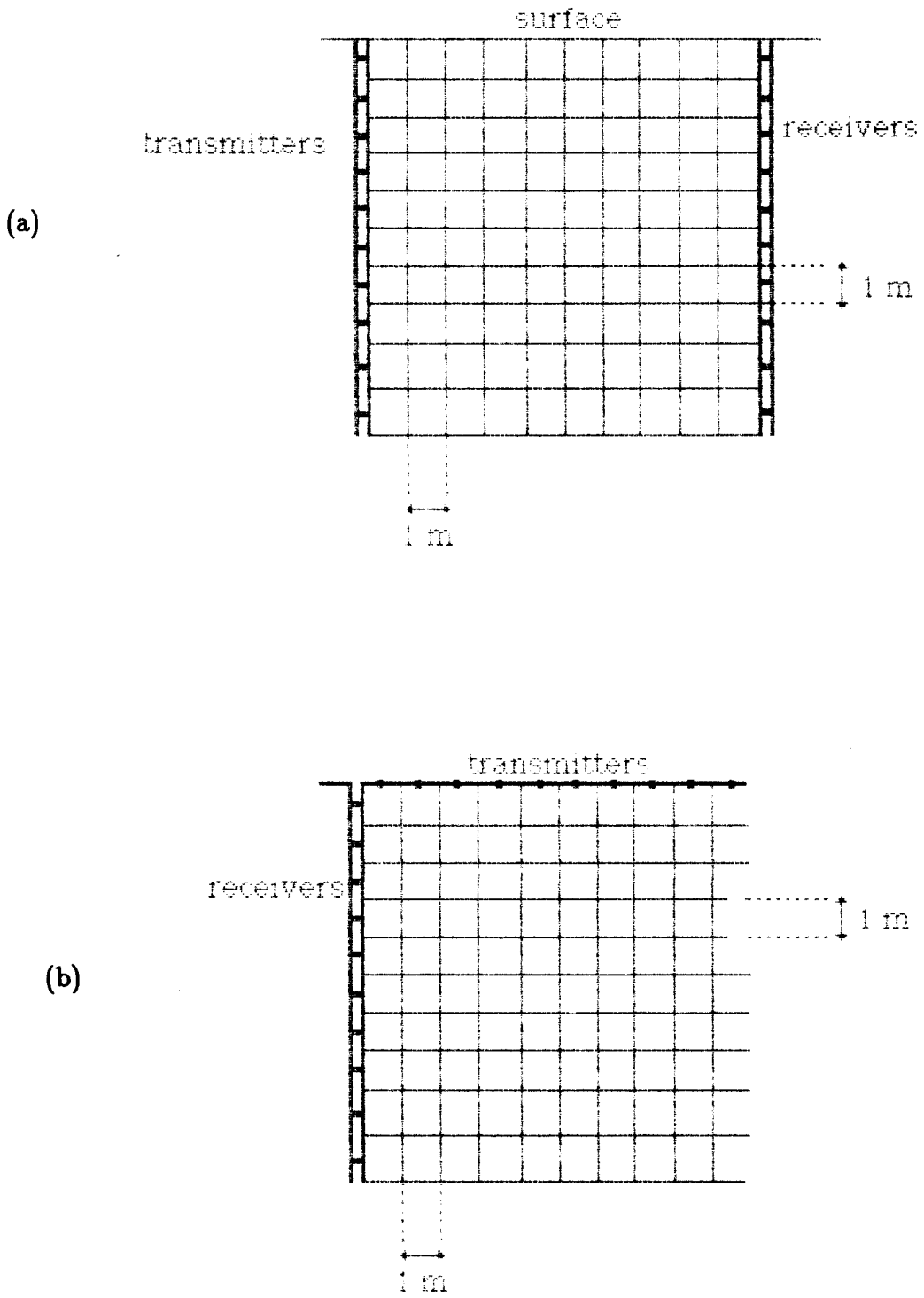

FIG. 2. a) Cross-hole geometry. b) Surface-to-hole geometry.

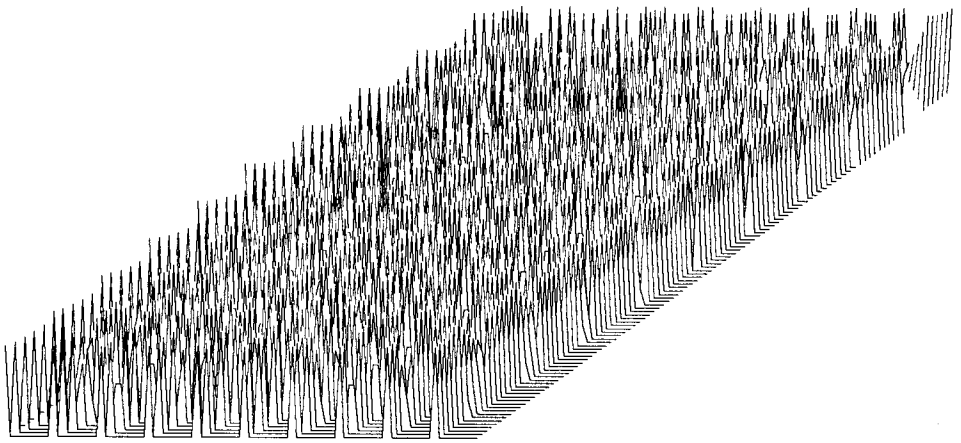

FIG. 3. The matrix **L** of equation (1) for the cross-hole example.

We assume that the ray paths are straight and construct the matrix **L** accordingly for each experiment. Figure 3 shows **L** for the cross-hole example where it is seen to be sparse. The singular values of **L** for the two setups are shown in Figure 4.

We see that there are some zero singular values indicating that the rank of **L** is less than 100 (which is the size of **L**); so **L** is not a full-rank matrix which implies that there is a *null space*. That is, the system we are trying to solve is under-determined: there are not enough independent equations to make the system uniquely determined. In the cross-hole experiment there are 17 zero eigenvalues which correspond to 17 undetermined components. In the surface-to-hole experiment, there are about 50 eigenvalues which is expected because one-half of the medium has not been covered.

To make the system less undetermined, we have to decrease the number of unknowns by decreasing the number of cells which means reducing the resolution of the solution. Experiment has shown that unless too few components in the horizontal direction are solved for, we would expect to have undetermined components.

Looking at the eigenvectors **u** and **v** is interesting and informative. The **u**'s are the *model* space eigenvectors, and the **v**'s are the *data* space eigenvectors. These two eigenvectors are coupled only through non-zero singular values. Because the number of

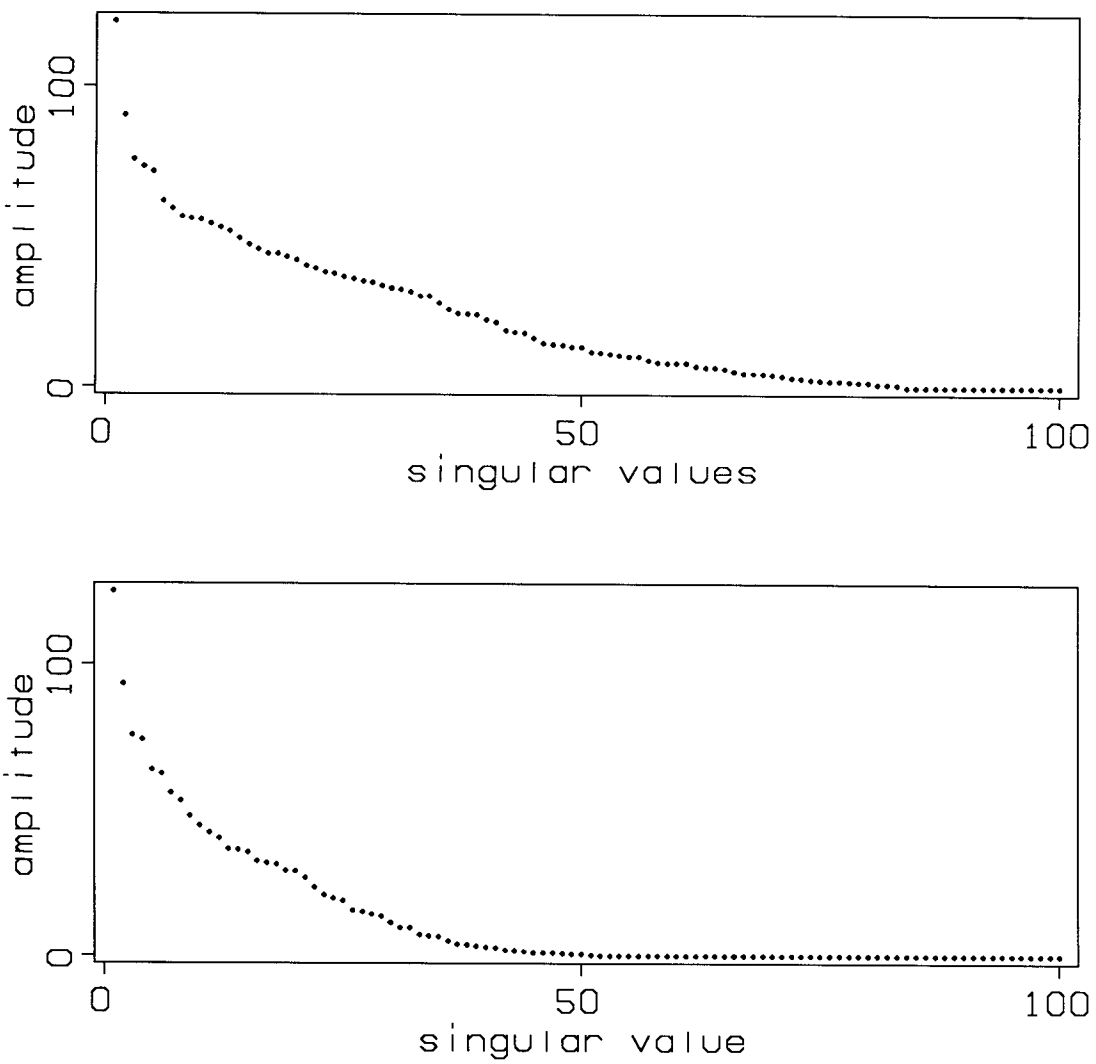

FIG. 4. Singular values of the matrix **L** using cross-hole geometry (top) and surface-to-hole geometry (bottom).

transmitters, receivers, rows, and columns of the model are all 10, we can display the columns of both eigenvectors as 10×10 matrices and we have a set of 100 pairs of these matrices. For non-zero singular values, we can visually perform the operation that **L** does on **w** in the model space and see the result in the data space. This operation is simple summation along the ray path. For zero singular values there is no coupling between the two spaces. It is most interesting to see this operation as a movie on a graphic terminal, unfortunately something I cannot show the reader. Figure 5 is an attempt to show the operation using a variable-intensity plot. The figure shows the first two eigenvectors corresponding to the two largest singular values for the cross-hole example.

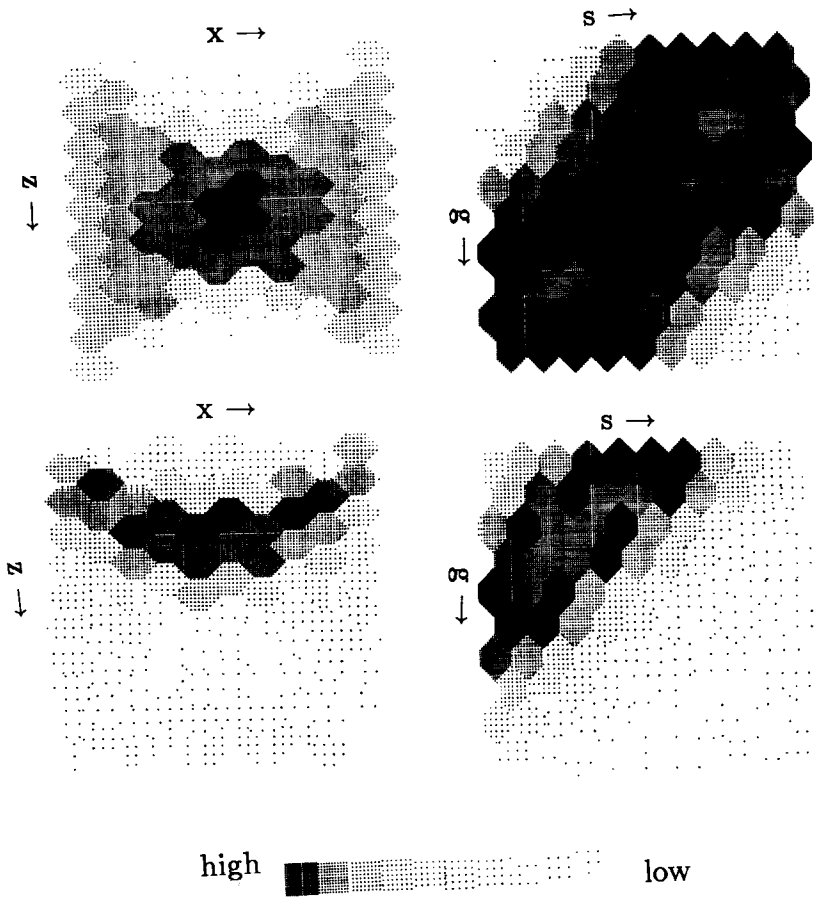

FIG. 5. Top: eigenvectors corresponding to the largest singular value for the cross-hole example; left is the model space eigenvectors and right is the data space eigenvectors. Integration from a transmitter to a receiver in the model space gives the corresponding point in the data space. Bottom: same as top for the second singular value.

CONCLUSIONS

We saw that borehole tomography is generally an under-determined inverse problem. Even when simple geometry is used, the problem is too large to solve using direct matrix inversion.

The problem was an intersting personal experience.

ACKNOWLEDGMENTS

John Toldi encouraged me to work on the this problem. His valuable suggestions and critical review of the paper were very helpful.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology, theory and methods, W. H. Freeman and Company, p. 684-689.

Al-Yahya, K., 1985, An iterative solution to seismic tomography. This SEP report.