Stanford Exploration Project, Report 72, December 18, 1997, pages 1–234

Stanford Exploration Project, Report 72, December 18, 1997, pages 1–234

Singular value decomposition for cross-well tomography

Reinaldo J. Michelena[1]

keywords: *modeling, anisotropy, two-dimensional, tomography, traveltime*

ABSTRACT

Singular value decomposition is performed on the matrices that result in tomographic velocity estimation from cross-well traveltimes in isotropic and anisotropic media. The singular vectors in both data and model space are presented along with their corresponding singular values for a simple geometry.

INTRODUCTION

In ray theoretic traveltime tomography, the solution of a linear system of equations is the heart of the problem. By solving this linear system, variations in the traveltimes are transformed into variations in the model parameters. This transformation from variations in the data to variations in the model depends on the properties of the matrix that describes the linear system and singular value decomposition (SVD) is the tool to study such properties.

SVD has been applied in the past to study the structure of the matrices involved in tomographic traveltime inversion problems (White, 1989; Bregman et al., 1989; Pratt and Chapman, 1990). In these studies the results of the SVD have been only *partially* reported because no reference has been made to the properties of the singular vectors in data space.

In this paper, I present the *complete* results of the SVD of the matrices that result after the following three types of parametrization: (a) in regions of constant slowness (McMechan, 1983), (b) in natural pixels (Michelena and Harris, 1991) and (c) in homogeneous anisotropic regions with elliptical velocity dependencies (Michelena and Muir, 1991).

By studying the structure of the corresponding matrices from the SVD results, it is possible to gain some insight on how iterative techniques such as conjugate gradients resolve the data and the model when used to solve tomographic inversions.

[1]**email:** not available

SINGULAR VALUE DECOMPOSITION: SHORT REVIEW

Any MxN-matrix $\underset{\sim}{\mathbf{J}}$ can be decomposed in the following way (Golub and Van Loan, 1989):

$$\underset{\sim}{\mathbf{J}} \;=\; \underset{\sim}{\mathbf{U}}\,\underset{\sim}{\mathbf{L}}\,\underset{\sim}{\mathbf{V}}^T \tag{1}$$

where $\underset{\sim}{\mathbf{U}}$ is an MxM orthogonal matrix of eigenvectors that span the data space, $\underset{\sim}{\mathbf{V}}$ is an NxN orthogonal matrix of eigenvectors that span the model parameters space and $\underset{\sim}{\mathbf{L}}$ is an MxN diagonal matrix whose diagonal elements are the singular values of $\underset{\sim}{\mathbf{J}}$. The columns of $\underset{\sim}{\mathbf{U}}$ are the eigenvectors of $\underset{\sim}{\mathbf{J}}\,\underset{\sim}{\mathbf{J}}^T$ and the columns of $\underset{\sim}{\mathbf{V}}$ are the eigenvectors of $\underset{\sim}{\mathbf{J}}^T\underset{\sim}{\mathbf{J}}$.

When a singular value is zero, the corresponding eigenvector in data (or model) space cannot be mapped into model (or data) space. When a singular value is not zero but is small compared with the largest one (large condition number), the contribution of the corresponding eigenvectors into the solution must be eliminated or attenuated (regularization) because the problem may become unstable.

THE LINEAR SYSTEM

Regardless of how the model is described, the problem of ray theoretic traveltime tomography always reduces to the solution of a system of equations of the form:

$$\underset{\sim}{\mathbf{J}}\,\Delta\boldsymbol{m} \;=\; \Delta\boldsymbol{t}, \tag{2}$$

where $\underset{\sim}{\mathbf{J}}$ is a matrix whose nature depends on how the model is described, $\Delta\boldsymbol{m}$ is a vector that contains the variations in model parameters with respect to a reference model, and $\Delta\boldsymbol{t}$ is the misfit between real and calculated traveltimes.

The model can be described in many different ways. However, in this paper I will only focus on three of them . The first one will be the conventional discretization of the medium in cells of constant slowness S_j (McMechan, 1983):

$$\Delta\boldsymbol{m} \;=\; (\Delta S_1, \Delta S_2, ..., \Delta S_N)^T, \tag{3}$$

where N is the total number of cells.

The second discretization I will consider will be the one based on natural pixels (Michelena and Harris, 1991):

$$\Delta\boldsymbol{m} \;=\; (a_1, a_2, ..., a_M)^T, \tag{4}$$

where the coefficients a_i are used to calculate the slowness perturbation ΔS as a superposition of natural pixels $\phi_i(x,z)$

$$\Delta S \;=\; \sum_{i=1}^{M} a_i \phi_i(x,z).$$

In the previous expressions, M is the total number of traveltimes, which are calculated also along natural pixels.

Finally, I will consider the discretization of the model in homogeneous orthogonal regions with elliptical velocity dependencies (Michelena and Muir, 1991):

$$\Delta \boldsymbol{m} \;=\; (\Delta S_{x_1}, \Delta S_{x_2}, ..., \Delta S_{x_N}, \Delta S_{z_1}, \Delta S_{z_2}, ..., \Delta S_{z_N})^T. \tag{5}$$

The reasons why I will perform the SVD only for these three particular parametrizations are the following:

- Compare the matrices obtained by discretizing the model in orthogonal homogeneous cells with those obtained by discretizing the model in natural pixels (for the isotropic case).
- Compare the matrices obtained by allowing the model to be either isotropic or anisotropic (when the model is discretized in orthogonal homogeneous cells).

SINGULAR VALUE DECOMPOSITION: APPLICATION

SVD (Dongarra, et al., 1979) was performed on the matrix $\underset{\sim}{\mathbf{J}}$ after parametrizing the model as described by equations (3) , (4) and (5). Pratt and Chapman (1990) suggest a clear way of representing the resultant singular values and singular vectors in model space ($\mathbf{V}$) of the matrices involved in tomographic problems. I will represent my results in a similar fashion including also in the representation the corresponding singular vectors in data space ($\mathbf{U}$). This will help to understand how the data and the model are resolved by using iterative techniques such as conjugate gradients.

The data space (spanned by the columns of $\underset{\sim}{\mathbf{U}}$) is shown in Figure **??**. It is a 2-D space of traveltimes where the horizontal axis contains the receiver depths and the vertical axis contains the source depths. If the origin of both axes is the same, the closer a given point is to the main diagonal (near offset), the closer is to zero the corresponding ray angle.

The geometry used to do the SVD for the different parametrizations is shown in Figure **??** (Pratt and Chapman, 1990): five sources and five receivers in a constant slowness medium. When the model is isotropic, the matrix $\underset{\sim}{\mathbf{J}}$ only depends on the ray geometry and when the model is anisotropic, $\underset{\sim}{\mathbf{J}}$ depends on both the ray geometry and the slowness model (constant in this case).

Figure **??** shows the SVD when the model is discretized using 6 horizontal isotropic layers (equation 3). Notice that the differences among the singular values are small. The largest singular values correspond to singular vectors in data and model space whose components are roughly of the same magnitude. We observe that with this parametrization only some "big structures" (averages) in data space can be explained

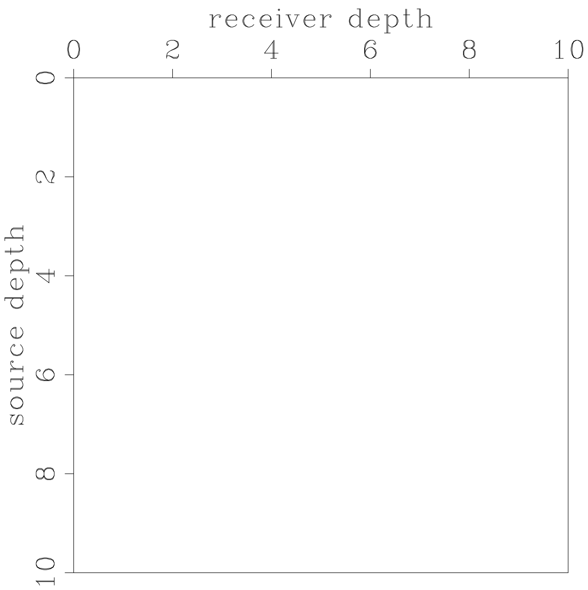

Figure 1: Space used to represent the data: for each source-receiver position the corresponding traveltime is plotted. This space is referred as **U** in subsequent figures. reinaldo-data-space [NR]

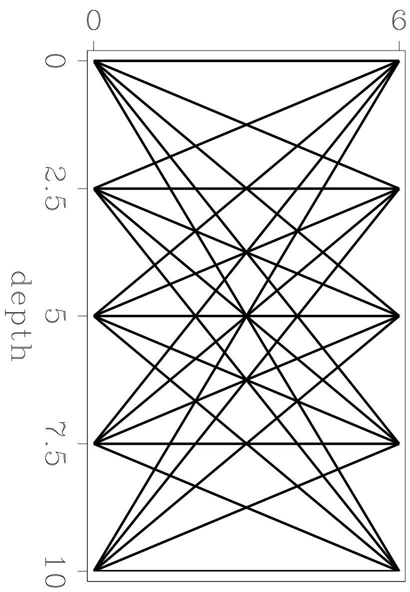

Figure 2: Recording geometry used to do the SVD for the different parametrizations. The slowness is constant. reinaldo-SVD-geometry [NR]

whereas in model space all the parameters can be resolved well. By representing the data space as described by Figure 1, it is possible to identify which traveltimes belong to the null space and therefore, cannot be resolved. Errors in these particular traveltimes (noise) will also have little or no effect in the solution.

When lateral variations are allowed in the previous parameterization results a matrix $\underset{\sim}{\mathbf{J}}$ whose SVD is shown in Figure **??**. The largest singular value corresponds roughly to horizontal layers (in model space) and non-horizontal rays (in data space). As the singular values decrease, the eigenvectors in model space tend to contain more horizontal variations and the eigenvectors in data space tend to expand near and far offsets alternatively (diagonal and non-diagonal structures). The null space of the problem corresponds to "pure" horizontal variations in model space and high frequency variations data space (rapid changes among nearby traveltimes). These vectors in the null space of the data are probably not a problem in real applications because if the real velocities do not change rapidly, high frequency variations are unlikely to be found in the measured traveltimes. The presence of noise introduces high frequency variations in the data. Unfortunately, these type of variations are not confined to the null space of the data but are present in singular vectors corresponding to larger singular values. Therefore, in some applications it might be necessary to damp the effect of singular values larger than those contained in the null space in order to attenuate the effect of certain components of the noise.

Figure **??** shows the SVD of the matrix that result after discretizing the model and computing the data using natural pixels (equation 4). We notice immediately that the matrix for the natural pixels is much better conditioned than the one obtained with square pixels (Figure **??**). The singular vectors in data space that correspond to the largest singular value are almost identical for both discretizations. However, the ways those data are distributed in the model are different for the different discretizations. For all the eigenvectors in Figure **??** there is a clear correspondency between the structures represented in data and model space, unlike the eigenvectors in Figure **??**. With the natural pixels all the parameters and data can be resolved because the discretization does not introduce another null space in the problem (different to the null space of the measurements), whereas the discretization of the model in square pixels does.

The SVD for a 1-D anisotropic parametrization (equation 5) is shown in Figure **??**. The upper half of each eigenvector in model space corresponds to S_x and the lower half corresponds to S_z. We observe that the largest singular values correspond to S_x in model space and in data space the behavior is similar to the isotropic 1-D case. The null space of the model is spanned by vectors that only contain information about S_z.

Figure **??** shows the SVD when the model is 2-D anisotropic. The result is nearly a combination of the previous ones in model space, namely, vertical variations in S_x correspond to the largest singular values and horizontal variations in S_z to the smallest ones. The null space of the model (Figure **??**) is formed almost entirely by "unit"

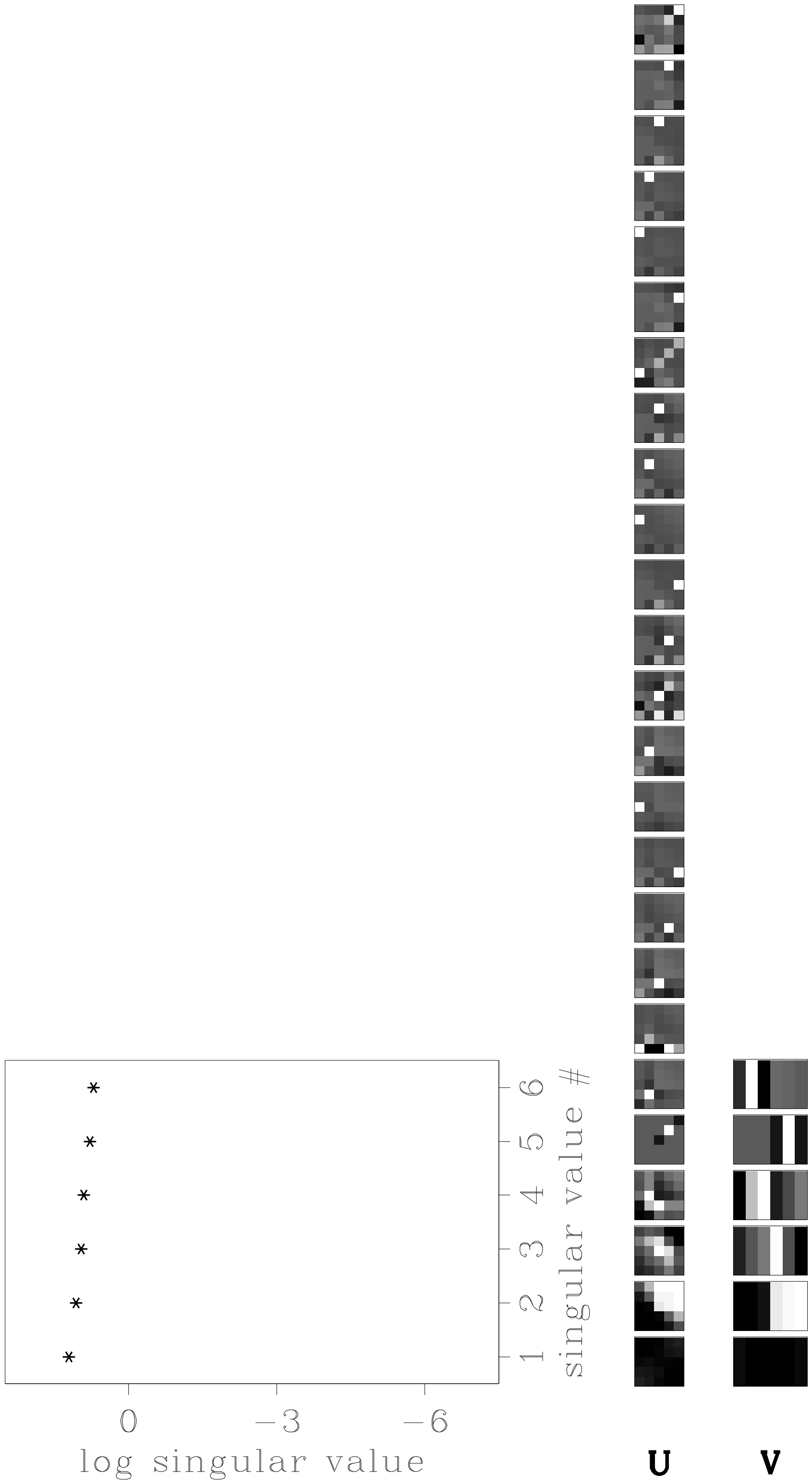

Figure 3: Singular value decomposition when the model is described as a superposition of 6 horizontal isotropic layers (1-D isotropic). **U** represents the singular vectors in data space and **V** represents the singular vectors in model space. The origin in data space in the upper left corner. The axes in model space are depth (perpendicular to the stratification) and horizontal distance. The gray scale goes from black (negative) to white (positive). reinaldo-SVD-iso-1d [NR]

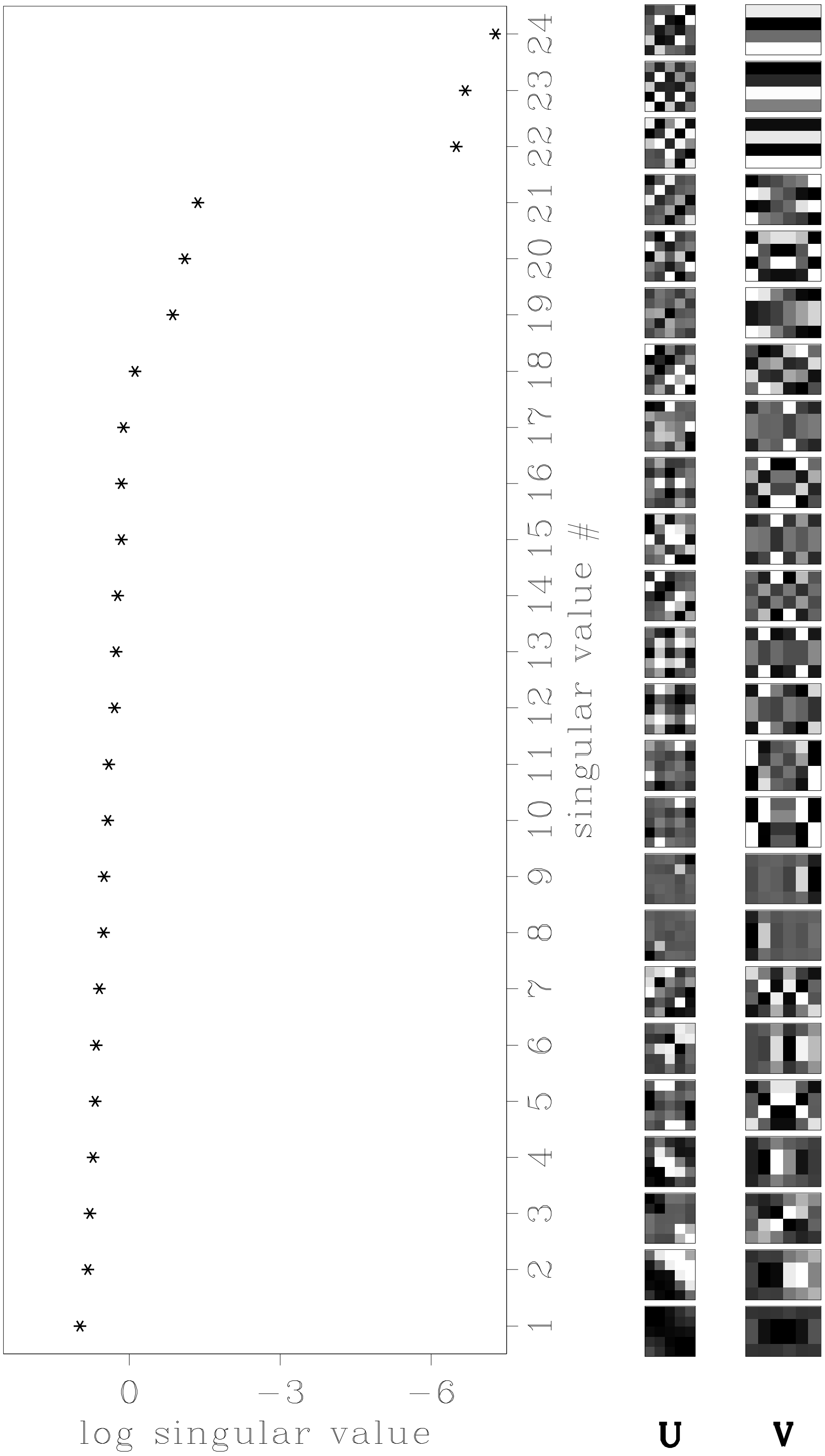

Figure 4: Singular value decomposition when the model is described as a superposition of 6x4 homogeneous isotropic squared regions (2-D isotropic). **U** represents the singular vectors in data space and **V** represents the singular vectors in model space. The origin in data space is in the upper left corner. The gray scale goes from black (negative) to white (positive). reinaldo-SVD-iso-2d [NR]

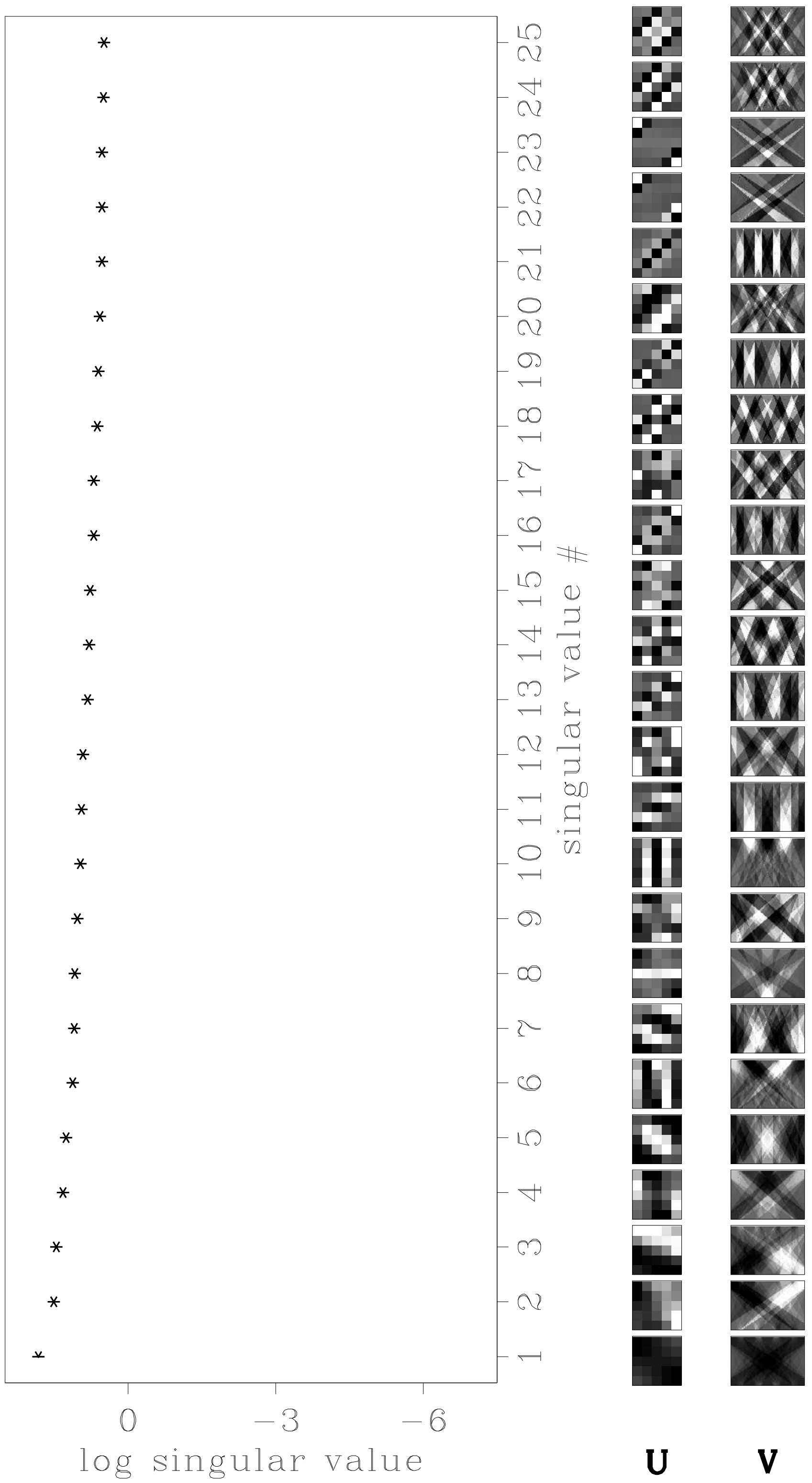

Figure 5: Singular value decomposition when the model is 2-D isotropic and is described as a superposition of 25 natural pixels. **U** represents the singular vectors in data space and **V** represents the sum of the components of each singular vector in model space along the corresponding natural pixels. The origin in data space is in the upper left corner. The gray scale goes from black (negative) to white (positive). reinaldo-SVD-fatrays [NR]

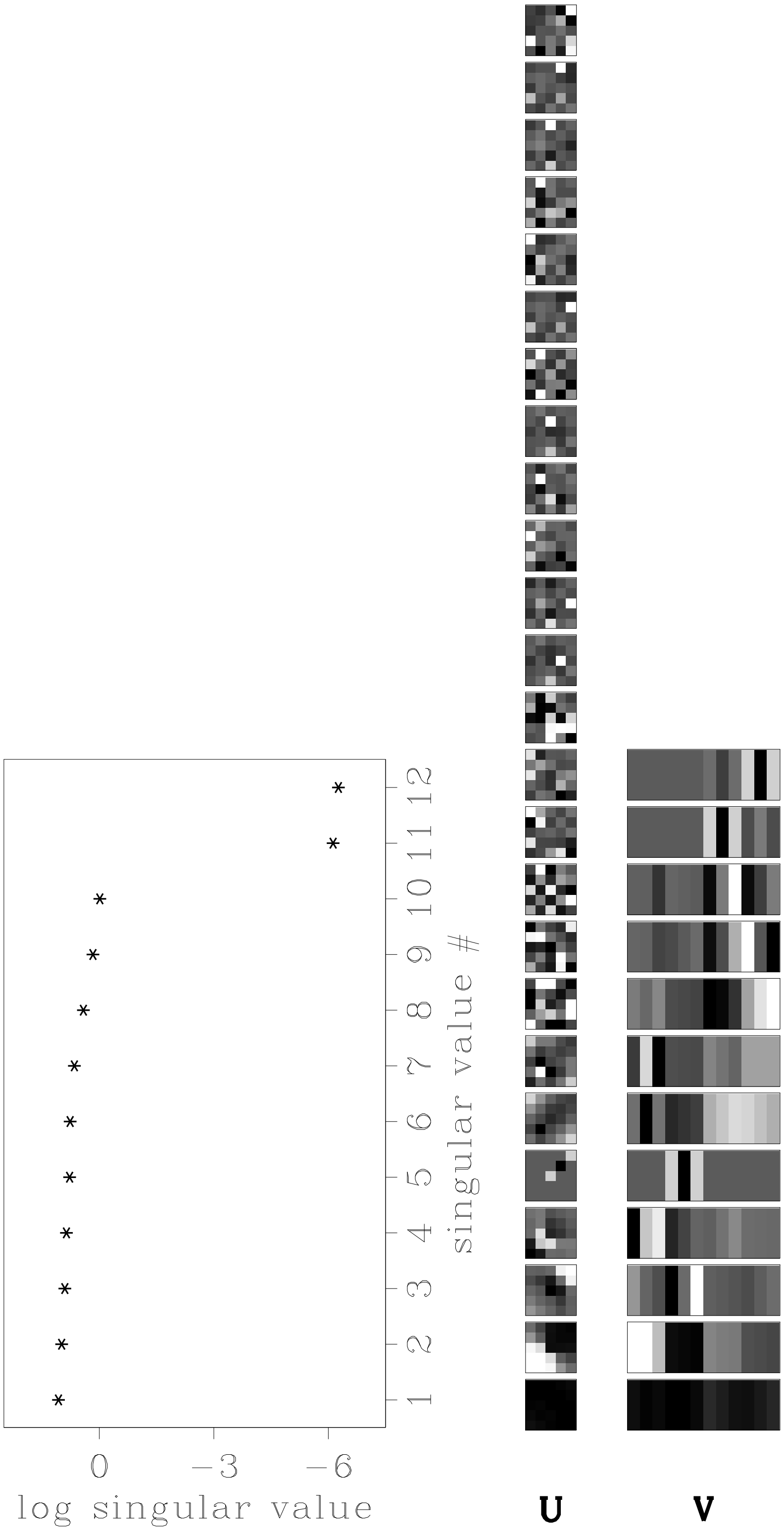

Figure 6: Singular value decomposition when the model is horizontally layered and anisotropic (6x2 model parameters). **U** represents the singular vectors in data space and **V** represents the singular vectors in model space. The upper half of each eigenvector in model space corresponds to S_x and the lower half corresponds to S_z. The origin in data space is in the upper left corner. The gray scale goes from black (negative) to white (positive). reinaldo-SVD-ani-1d [NR]

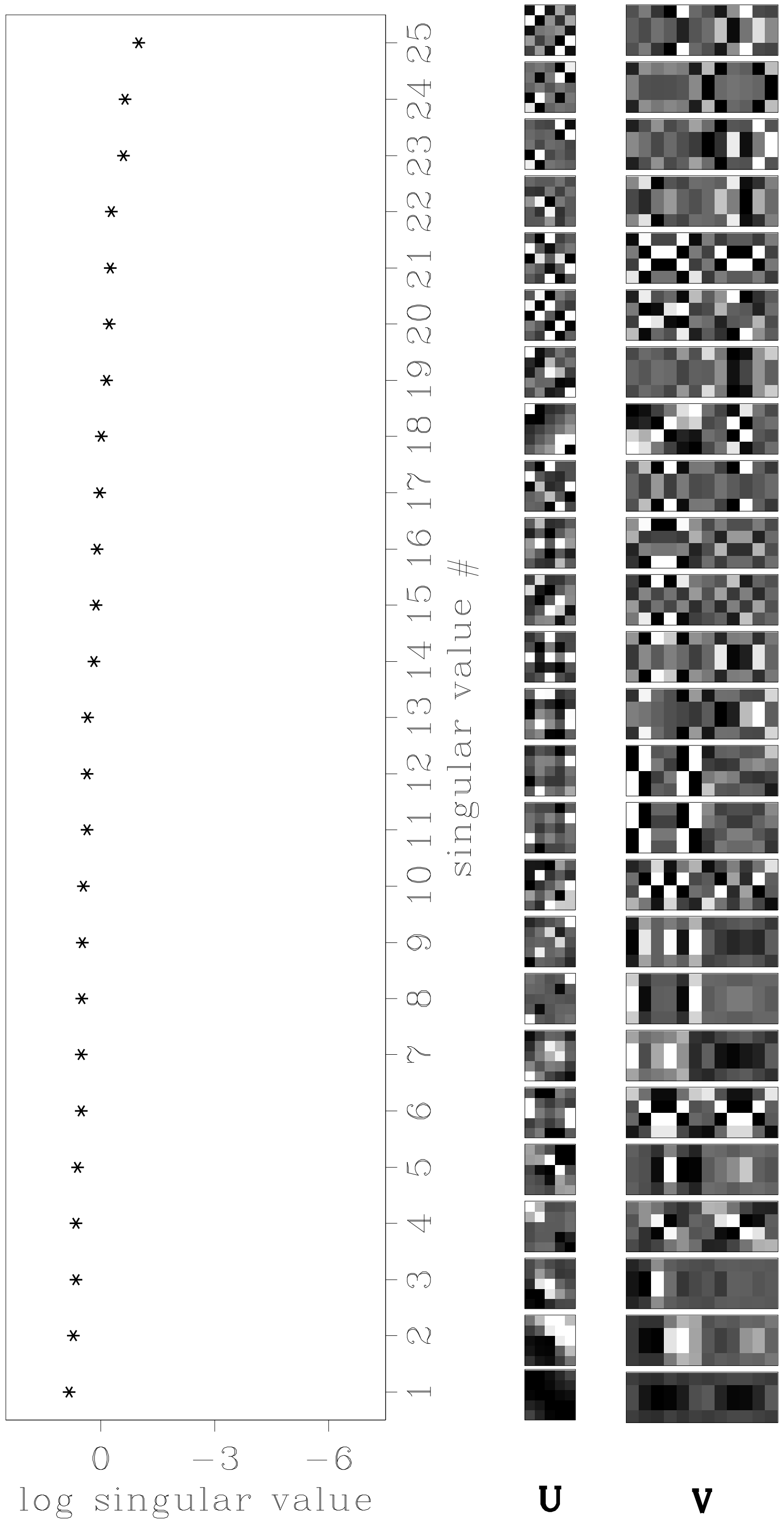

Figure 7: Singular value decomposition when the model is 2-D anisotropic. (24x2 model parameters). **U** represents the singular vectors in data space and **V** represents the singular vectors in model space. The upper half of each eigenvector in model space corresponds to S_x and the lower half corresponds to S_z. The origin in data space is in the upper left corner. The gray scale goes from black (negative) to white (positive). reinaldo-SVD-ani-2d [NR]

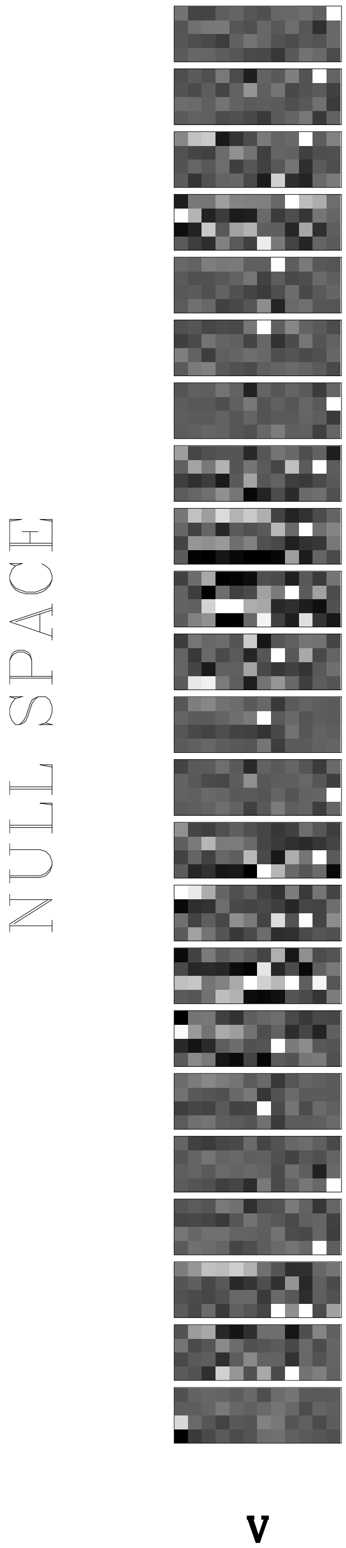

Figure 8: Vectors that span the null space of the model for the SVD shown in Figure 7. Most vectors are contained in the S_z space and therefore, S_z cannot be estimated at the same resolution of S_x from cross-well traveltimes alone. reinaldo-model-null [NR]

vectors that describe S_z. For this reason, in the synthetic and real data applications shown in Michelena and Muir (1991), the resolution of the vertical component of the slowness is poor when compared with the resolution of the horizontal component. As expected, S_z is poorly defined by the cross-well recording geometry. Note that by increasing the number of model parameters the condition number of the matrix diminishes when compared with the one of the matrix for the isotropic case (Figure 4).

CONCLUDING REMARKS

In 2-D isotropic models, the discretization of the model and prediction of the data using natural pixels produces a matrix that is well conditioned. Data and model parameters can be well resolved. Unfortunately, I don't know a formulation similar to the natural pixels one valid also for anisotropic models and therefore, for estimating slowness anisotropy in 2-D, I still have to deal with the inconveniences inherent to the square pixels discretization.

From the previous results we have confirmed that it is a more difficult problem to find horizontal variations in S_z (smaller singular values) than it is to find vertical variations in S_x (larger singular values).

If the components of the model in the null space are eliminated, the problem in one dimension is better conditioned than the problem in two dimensions for both the isotropic and anisotropic cases. This result is not surprising but it has received little attention. For cross-well geometries, most tomographic traveltime inversions try to estimate 2-D variations even in cases when the medium is known to be horizontally layered. The extra degrees of freedom allowed in the inversion have to be penalized appropriately in the objective function. When the velocity logs are available and the medium is horizontally layered and isotropic, 1-D inversions are not interesting and the intrinsic advantages of the parametrization (fewer unknowns and better conditioning) are not useful. However, if the medium is anisotropic and horizontally layered, a 1-D instead of a 2-D parametrization can be the difference between recovering or not (accurately) both slowness components.

The results in data space tell us also something new: when using iterative techniques to solve the problem, the data are not resolved "evenly" as iterations proceed but, instead, some iterations resolve part of the data not resolved by other iterations. This has to be taken into account when interpreting traveltime residuals plotted in this fashion because their spatial distribution may depend on when the iterations are stopped.

REFERENCES

Bregman, N.D., Bailey, R.C., and Chapman, C.H., 1989, Ghosts in tomography: the effects of poor angular coverage in 2-D seismic traveltime inversion: Can. J. Expl. Geophy., **25**, 7–27.

Dongarra, J.J., Bunch, J.R., Moler, C.B., and Stewart, G.W., 1979, Linpack: Users' Guide: Soc. Ind. Appl. Math., Chapter 11.

Golub, G.H., and Van Loan, C.F., 1989, Matrix Computations. The Johns Hopkins University Press.

McMechan, G.A., 1983, Seismic tomography in boreholes: Geophys. J. Roy. Astr. Soc., **74**, 601–612.

Michelena, R.J., and Harris, J.M., 1991, Tomographic traveltime inversion using natural pixels: Geophysics, **56**, 635–644.

Michelena, R.J., and Muir, F., 1991, Anisotropic tomography: SEP–**70**, 201–226.

Pratt, R.G., and Chapman, C.H., 1990, Traveltime tomography in anisotropic media: Application: Submitted to Geophys. J. Intl.

White, D.J., 1989, Two-dimensional seismic refraction tomography: Geophysical Journal, **97**, 223–245.