Inversion of large ill-posed seismic imaging problems

Mauricio D. Sacchi, Department of Physics, University of Alberta*

Summary

In this paper we deal with the inversion of large ill-posed operators which after discretization yields to sparse large system of equations. In particular, we are interested on the numerical inversion of the single-scattering acoustic approximation about a reference medium (Born approximation).

The proposed algorithm uses the Lanczos bidiagonalization to pose the problem into a subspace where the linear operator is bidiagonal. One of the advantages of the bidiagonalization procedure is that the problem is solved in a compressed space and, therefore, the computational cost of the procedure is substantially reduced. Tradeoff curves are efficiently computed by solving several small inverse problems in the subspace obtained by means of the bidiagonalization procedure.

Introduction

An iterative technique based on Lanczos bidiagonalization (O'Leary and Simmons, 1981; Bjorck, 1988) is presented to invert the forward modeling operator derived from the linearized Born approximation to the scattered wavefield due to a perturbation of a reference model (Weglein, 1982; Miller et al., 1987). The forward problem can be represented by

$$\underset{\sim}{\mathbf{L}}\,\mathbf{m} = \mathbf{d}\,. \tag{1}$$

where $\mathbf{d}$ is an $N\times1$ vector that represents the observations ($N = N_{receivers} \times N_{sources} \times N_{samples}$), $\mathbf{m}$ is an $M \times 1$ vector of unknowns (M is the total number of grid points used to describe the perturbed model of the subsurface).

The effect of the operator $\underset{\sim}{\mathbf{L}}$ can be evaluated by analyzing $\mathbf{m}_0 = \underset{\sim}{\mathbf{L}}^T\,\mathbf{d}$ which possesses features very similar to those of $\mathbf{m}$. However, since $\underset{\sim}{\mathbf{L}}$ is a non-orthogonal operator ($\underset{\sim}{\mathbf{L}}^T \neq \underset{\sim}{\mathbf{L}}^{-1}$), the reconstructed data $\underset{\sim}{\mathbf{L}}\,\mathbf{m}_0$ will not properly model the original data.

The least squares solution of equation (1) can be written as

$$\begin{aligned}\hat{\mathbf{m}} &= (\underset{\sim}{\mathbf{L}}^T\,\underset{\sim}{\mathbf{L}})^{-1}\,\underset{\sim}{\mathbf{L}}^T\,\mathbf{d} \\ &= (\underset{\sim}{\mathbf{L}}^T\,\underset{\sim}{\mathbf{L}})^{-1}\,\mathbf{m}_0\,.\end{aligned} \tag{2}$$

The perturbation $\mathbf{m}_0$, which is obtained by applying the transpose operator $\underset{\sim}{\mathbf{L}}^T$ to the data, corresponds to the migrated image of the subsurface (Schuster, 1993). In next section we describe a bidiagonalization procedure to compute an approximation to the operator $(\underset{\sim}{\mathbf{L}}^T\,\underset{\sim}{\mathbf{L}})^{-1}$.

Bidiagonalization

The Lanczos bidiagonalization algorithm factors the $N \times M$ forward operator $\underset{\sim}{\mathbf{L}}$ into

$$\underset{\sim}{\mathbf{U}}^T\underset{\sim}{\mathbf{L}}\underset{\sim}{\mathbf{Q}} = \underset{\sim}{\mathbf{B}}\,, \tag{3}$$

where

$$\underset{\sim}{\mathbf{U}} : N \times N,\;\; \underset{\sim}{\mathbf{B}} : N \times M,\;\; \underset{\sim}{\mathbf{Q}} : M \times M\,.$$

The matrices $\underset{\sim}{\mathbf{U}}$ and $\underset{\sim}{\mathbf{Q}}$ are orthogonal

$$\underset{\sim}{\mathbf{U}}^T\,\underset{\sim}{\mathbf{U}} = \underset{\sim}{\mathbf{I}}\,, \quad \underset{\sim}{\mathbf{Q}}^T\,\underset{\sim}{\mathbf{Q}} = \underset{\sim}{\mathbf{I}},$$

and $\underset{\sim}{\mathbf{B}}$ is bidiagonal:

$$\underset{\sim}{\mathbf{B}} = \begin{pmatrix} \alpha_1 & \beta_1 & & & & \\ & \alpha_2 & \beta_2 & & & \\ & & \alpha_3 & \beta_3 & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \beta_{M-1} \\ & & & & & \alpha_M \\ & & & 0 & & \end{pmatrix}\,. \tag{4}$$

If the factorization is in terms of the first k columns of $\underset{\sim}{\mathbf{U}}$ and $\underset{\sim}{\mathbf{Q}}$, we can write the following equation

$$\underset{\sim}{\mathbf{U}}_k^T\,\underset{\sim}{\mathbf{L}}\,\underset{\sim}{\mathbf{Q}}_k = \underset{\sim}{\mathbf{B}}_k\,, \tag{5}$$

with matrices with the following size

$$\underset{\sim}{\mathbf{U}} : N \times k,\;\; \underset{\sim}{\mathbf{B}} : k \times k,\;\; \underset{\sim}{\mathbf{Q}} : M \times k\,.$$

Now, the matrix $\underset{\sim}{\mathbf{B}}_k$ is bidiagonal and square

$$\mathbf{B}_k = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ & \alpha_2 & \beta_2 & & \\ & & \alpha_3 & \beta_3 & \\ & & & \ddots & \ddots \\ & & & & \ddots & \beta_{k-1} \\ & & & & & \alpha_k \end{pmatrix}\,. \tag{6}$$

The algorithm to retrieve the matrices $\mathbf{U}$, $\mathbf{Q}$ and the elements α_i, β_i is summarized as follows (O'Leary and Simmons, 1981):

Starting with a vector $\mathbf{z}_1$, set the initial elements

$$\mathbf{q}_1 = \frac{\mathbf{z}_1}{||\mathbf{z}_1||},\ \mathbf{y}_1 = \mathbf{L}\,\mathbf{q}_1,\ \alpha_1 = ||\mathbf{y}_1||,\ \mathbf{y}_1 = \frac{\mathbf{y}_1}{\alpha_1}$$

For $i = 1, 2, \ldots, k-1$ {

$$\mathbf{z}_{i+1} = \mathbf{L}^T\,\mathbf{u}_i - \alpha_i\,\mathbf{q}_i\,, \qquad \beta_i = ||\mathbf{z}_{i+1}||$$

$$\mathbf{q}_{i+1} = \frac{\mathbf{z}_{i+1}}{\beta_i}$$

$$\mathbf{y}_{i+1} = \mathbf{L}\mathbf{q}_{i+1} - \beta_i \mathbf{u}_i\,, \qquad \alpha_{i+1} = ||\mathbf{y}_{i+1}||$$

$$\mathbf{u}_{i+1} = \frac{\mathbf{y}_{i+1}}{\alpha_{i+1}}$$

}

The algorithm requires two matrix multiplications per iteration. In each iteration we perform one migration ($\mathbf{L}^T$) and one modeling operation ($\mathbf{L}$). The Lanczos scheme is initiated with $\mathbf{m}_0 = \mathbf{z}_1 = \mathbf{L}^T\mathbf{d}$ (the migrated image).

A problem with Lanczos decomposition is its numerical stability (Kahan and Parlett, 1976). In practice, the Lanczos vectors may loose orthogonality after a few iterations. Different strategies have been adopted to enforce orthogonality. One solution is to perform complete orthogonalization of each vector with all the preceding vectors. However this re-orthogonalization strategy increases the computational cost of the algorithm. A selective re-orthogonalization, where the orthogonalization is carried out with respect to a few preceding vectors (s), is a substantially more efficient scheme (O'Leary and Simmons, 1981)

$$\mathbf{z}_i = \mathbf{z}_i - (\mathbf{z}_i^T \mathbf{q}_j)\mathbf{q}_j \qquad j = i-1, i-2, \ldots, i-s \tag{7}$$

$$\mathbf{y}_i = \mathbf{y}_i - (\mathbf{y}_i^T \mathbf{u}_j)\mathbf{u}_j \qquad j = i-1, i-2, \ldots, i-s \tag{8}$$

In our problem, we usually attempt to retrieve a very small basis (about 10 vectors). In general, we have not found numerical problems for subspaces where $k < 20$.

Bidiagonalization and regularization of L

The least squares solution involves the inversion of the operator $\mathbf{L}^T\mathbf{L}$. Since the inverse of this operator is too expensive to compute, we recast our inverse problem in a small subspace obtained by means of the Lanczos bidiagonalization.

The original system of equations (1) can be written as follows,

$$\mathbf{U}_k^T\,\mathbf{L}\,\mathbf{Q}_k\,\mathbf{z} = \mathbf{U}_k^T\,\mathbf{d} \tag{9}$$

where

$$\mathbf{m} = \mathbf{Q}_k\,\mathbf{z}\,. \tag{10}$$

Combining equations (9) and (5) we obtain

$$\mathbf{B}_k\,\mathbf{z} = \mathbf{g}\,, \tag{11}$$

where $\mathbf{g} = \mathbf{U}^T\mathbf{d}$. It can be proved that the singular values of the matrix $\mathbf{B}_k$ are a good a approximation to the large and small singular values of $\mathbf{L}$ (Scales, 1989). It is evident that some kind of regularization is needed. The above equation is solved using penalized least squares

$$\mathbf{z} = (\mathbf{B}_k^T\,\mathbf{B}_k + \mu\mathbf{I})^{-1}\mathbf{B}^T\,\mathbf{g}\,, \tag{12}$$

where the parameter μ denotes the tradeoff parameter or hyper-parameter of the problem. Since $(\mathbf{B}_k^T\,\mathbf{B}_k + \mu\mathbf{I})$ is a tridiagonal matrix, equation (12) is efficiently solved by means of a tridiagonal solver. The solution to equation (12) can also be computed by means of the Singular Value Decomposition (SVD) of the matrix $\mathbf{B}_k$. If the aforementioned strategy is adopted, the solution is given by

$$\mathbf{z} = \sum_{\lambda_i > \Delta} \lambda_i^{-1}\mathbf{v}_i(\mathbf{r}_i^T\,\mathbf{g})\,, \tag{13}$$

where λ_i are the singular values of $\mathbf{B}_k$. The vectors $\mathbf{v}_i$ and $\mathbf{r}_i$ are the eigenvectors of $\mathbf{B}_k^T\,\mathbf{B}_k$ and $\mathbf{B}_k\,\mathbf{B}_k^T$, respectively. The parameter Δ denotes a threshold value.

In our numerical examples we have used penalized least squares, in this case the perturbation to the background model is given by

$$\begin{aligned}
\mathbf{m}_k = \mathbf{Q}_k\,\mathbf{z} &= \mathbf{Q}_k\,(\mathbf{B}_k^T\,\mathbf{B}_k + \mu\mathbf{I})^{-1}\mathbf{U}_k^T\,\mathbf{d} \\
&= \mathbf{Q}_k\,(\mathbf{B}_k^T\mathbf{B}_k + \mu\mathbf{I})^{-1}\mathbf{Q}_k^T\,\mathbf{L}^T\,\mathbf{d} \\
&= \mathbf{Q}_k(\mathbf{B}_k^T\mathbf{B}_k + \mu\mathbf{I})^{-1}\,\mathbf{Q}_k^T\mathbf{m}_0\,.
\end{aligned} \tag{14}$$

By recalling equations (2) and (14) we are now in condition of writing

$$(\mathbf{L}^T\,\mathbf{L})^{-1} \approx \mathbf{Q}_k\,(\mathbf{B}_k^T\,\mathbf{B}_k + \mu\mathbf{I})^{-1}\mathbf{Q}_k^T\,. \tag{15}$$

It is clear, that instead of inverting $\mathbf{L}^T\,\mathbf{L}$, we prefer to invert the tridiagonal matrix $\mathbf{B}_k^T\,\mathbf{B}_k + \mu\mathbf{I}$.

Since k is usually small ($k \approx 10$), the vectors $\mathbf{u}_i$, $\mathbf{q}_i$ and the elements of the bidiagonal matrix $\mathbf{B}_k$ can be saved and used to accommodate our inversion to a desire misfit value. This is achieved by modifying μ or by using a smaller subspace, i.e., $k' < k$. In summary, tradeoff curves can be efficiently computed by solving several small problems of the type given by equations (12) or (13).

For certain problems it might be better to solve

$$min||\underset{\sim}{\mathbf{W}}(\underset{\sim}{\mathbf{L}}\,\mathbf{m}-\mathbf{d})||_2^2\,, \qquad (16)$$

where $\underset{\sim}{\mathbf{W}}$ is a diagonal matrix that accounts for the quality of the data or irregular sampling. In this case, the Lanczos algorithm uses the following inner products $\underset{\sim}{\mathbf{W}}\,\underset{\sim}{\mathbf{L}}\,\mathbf{x}$ (forward or modeling operator) and $\underset{\sim}{\mathbf{L}}^T\,\underset{\sim}{\mathbf{W}}^T\,\mathbf{y}$ (transpose operator). In the synthetic simulations the matrix $\underset{\sim}{\mathbf{W}}$ is designed to emphasize the fact that the sources are irregularly distributed. A similar idea is used in Feichtinger et al.(1995) in the context of non-uniform sampling theory. The matrix $\underset{\sim}{\mathbf{W}}$ may also serve to remove the seismic source from the data. In other words, the inversion can contemplate a wavelet source deconvolution term. This issue is under investigation.

Example

Now, I will consider the problem of recovering the perturbation to the background velocity based on observed seismograms from a multi-source multi-receiver acoustic experiment. The total number of grid points in the model (Figure 1) is $M = NX \times NZ$ ($NX = 600$, $NZ = 400$). The number of time samples is $NT = 1000$. Twelve sources are located along the surface of the model, randomly spaced from $x = 0$ to $x = 2000\,$m. The array of receivers consists of 64 equally spaced geophones distributed along the surface of the model. The size of the operator is $N \times M$ where $N = 12 \times 64 \times 1000$. The migrated image is portrayed in Figure 2. This image is severely contaminated by sampling and aperture artifacts.

The forward modeling operator was inverted using equation (14) in a subspace composed of 15 Lanczos vectors ($\mathbf{u}_i$, $\mathbf{q}_i$, $i = 1 \dots 15$). The inverted image shows an important attenuation of sampling artifacts.

The tradeoff curve of the problem is displayed in Figure 4. The vertical axis is the misfit error (a measure of how well the inverted image reproduces the observations). The horizontal axis is the norm of the inverted perturbation $||\mathbf{m}_k||$.

Conclusions

I have presented a bidiagonalization algorithm to invert large ill-posed problems which arise in the context of acoustic imaging using the Born approximation. To alleviate the computational burden of the inversion, the inverse problem is posed in a subspace obtained by means of the Lanczos bidiagonalization.

Finally, it should be stressed that without much problem, this algorithm can be applied to more complicated problems as long as the proper inner products exists. An example is an operator that not only models the data, but also performs other tasks, i.e, filtering and/or source processing.

Acknowledgements

This research was supported by funding from Pan Canadian Petroleum and Geo-X Systems Ltd.

References

Bjorck, A., 1988, A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations, BIT, **28**, 659-670.

Feichtinger, H. G., Gröchenig, K., and Strohmer, T., 1995, Efficient numerical methods in non-uniform sampling theory: Numer. Math., **69**, 423-440.

Kahan, W., and Parlett, B. N, 1976, How far should you go with the Lanczos process?, in sparse matrix computations, Academic Press, New York.

Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943-964.

O'Leary D. P., and Simmons, J. A., 1981, A bidiagonalization-regularization procedure for large scale discretization of ill-posed problems, SIAM J. Sci. Stat. Comput., **2**, 474-488.

Scales, J. A., 1989, On the use of conjugate gradient to calculate the eigenvalues and singular values of large, sparse matrices: Geophys. Journal, **97**, 1790-1793.

Schuster, G., T., 1993, Least squares cross-well migration: 63rd Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 110-113.

Weglein, A., B., 1982, Multidimensional seismic analysis: migration and inversion: Geoexploration, **20**, 47-60.

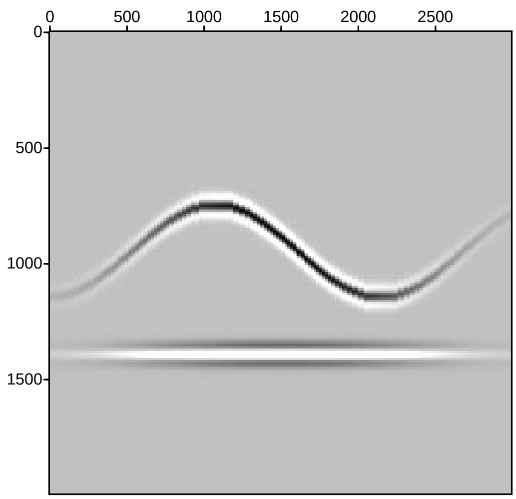

Fig. 1: Syn thetic image. The horizontal and v ertical coordinates are in meters.

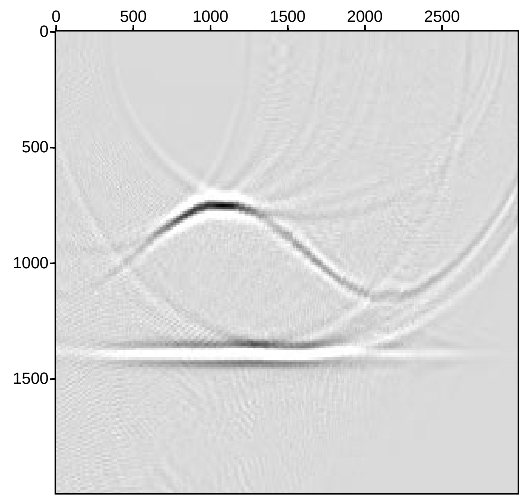

Fig. 3: In verted image. The in version is performed in a subspace composed of $k = 15$ Lanczos vectors.

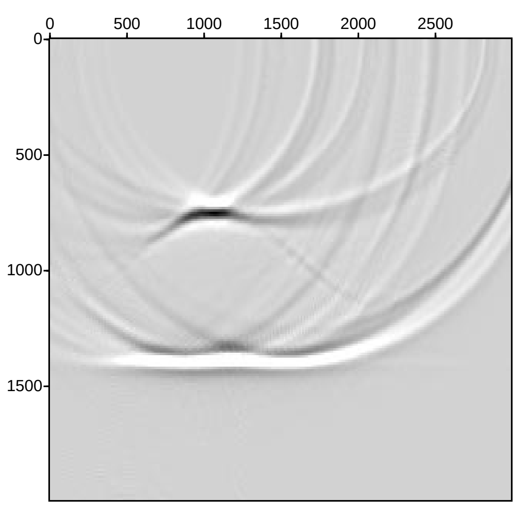

Fig. 2: Migrated image using the transpose operator.

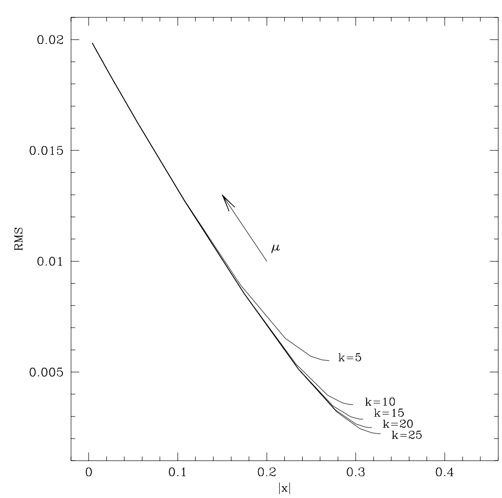

Fig. 4: T radeoff curv e. The parameter k is the size of the Lanczos subspace and μ a hyper-parameter. The vertical axis is the misfit function, the horizontal axis denotes the norm of the in verted model perturbation.