Feasibility of Using Multigrid Methods for Solving Least-Squares Prestack Kirchhoff Migration Equation

Abdolnaser Yousefzadeh
University of Calgary, Calgary, Alberta, Canada
ayousefz@ucalgary.ca
and
John C. Bancroft
University of Calgary, Calgary, Alberta, Canada
bancroft@ucalgary.ca

Summary

The feasibility of different approaches of using the multigrid methods in solving the linear system of Kirchhoff Least-Squares Prestack Time Migration (LSPSTM) equation is investigated.

This study showed that conventional method of multigrid is not viable to solve the Kirchhoff LSPSTM equation for at least two reasons: first, the Hessian matrix is not a diagonally dominant matrix, therefore, standard iterative solvers of the multigrid are not effective, second, Hessian matrix is too large and dense to be loaded in the memory of today's computers.

The performance of Conjugate Gradient (CG) multigrid is discussed. It is shown that because CG does not have a smoothing property, it should not be considered as an effective multigrid iterative solver.

Introduction

Kirchhoff seismic modeling can be defined by

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{m} \ , \qquad (1)$$

where $\boldsymbol{d}$ is the seismic data, $\boldsymbol{m}$ is the earth reflectivity, and $\boldsymbol{G}$ is the Kirchhoff forward modeling operator, a matrix containing diffraction hyperbolas. The inversion process,

$$\boldsymbol{m} = \boldsymbol{G}^{-1}\boldsymbol{d} \ , \qquad (2)$$

recovers the earth reflectivity model from seismic data. Inverting the $\boldsymbol{G}$ matrix is extremely difficult. The transpose of $\boldsymbol{G}$ which is migration may be used as an approximation to the $\boldsymbol{G}^{-1}$:

$$\widehat{\boldsymbol{m}} = \boldsymbol{G}'\boldsymbol{d} \ , \qquad (3)$$

where $\widehat{\boldsymbol{m}}$ is the migrated image and $\boldsymbol{G}'$ is the migration operator. Kirchhoff migration produces some artifacts in the migrated image. These migration artifacts can be attenuated by the Least-Squares Migration (Nemeth et. al., 1999). LSPSTM of the seismic data is obtained by minimizing the following general cost function:

$$J(m) = \|Gm - d\|^2 + \mu^2 \mathcal{R}(m), \qquad (4)$$

where $\mathcal{R}(m)$ is a regularization term, and μ is tradeoff parameter. Euclidian norm is the simplest form of the regularization term, $\mathcal{R}(m) = \|m\|_2^2$, which leads to the Damped Least-Squares solution, m_{DLS}, of the problem obtained by solving following equation:

$$(G'G + \mu I)m_{DLS} = G'd. \qquad (5)$$

Many other regularization functions exist. For example solution to a LSPSTM problem with smoothing the regularization, $\mathbf{m}_{SLS}$, is obtained when $\mathcal{R}(m) = \|D_h m\|_2^2$, where D_h is the first derivative in the offset direction:

$$(G'G + \mu^2 D_h' D_h) m_{SLS} = G'd \quad . \tag{6}$$

With the ability of attenuating of migration artifacts, LSPSTM images have higher resolution than those from migration (Nemeth, 1999). However, two disadvantages are accompanying this method. In addition to the requirement of the accurate velocity information, LSPSTM is more computer time and memory consuming than the migration (Yousefzadeh, 2008).

The performance of least-squares seismic migration usually requires solving a large system of linear equations in the general form of:

$$Au = b \quad , \tag{7}$$

where A is a general $M \times N$ matrix, for example Hessian matrix, $G'G + \mu^2 I$, in Equation 5 or $G'G + \mu^2 D_h' D_h$, in Equation 6, b is a vector with M known elements, for example migration image, and u is the unknown Least-Squares solution. In the LSPSTM equation, the Hessian is a large matrix which is too difficult to be solved using direct methods. Thus, iterative methods replace the direct methods. If a problem is solvable by the multigrid methods, it will be solved faster and with better recovery of the low frequency contents than many other methods such as Successive Over Relaxation (SOR) and CG (Stuben, 2002).

In this study, feasibility of using multigrid properties for solving LSPSTM in order to reduce the computational cost or enhance the resolution of the resulted image is investigated. It has been shown the reason that multigrid with its conventional solvers is not viable to solve the mentioned problem and CG multigrid is not an effective method.

Solving LSPSTM using standard multigrid

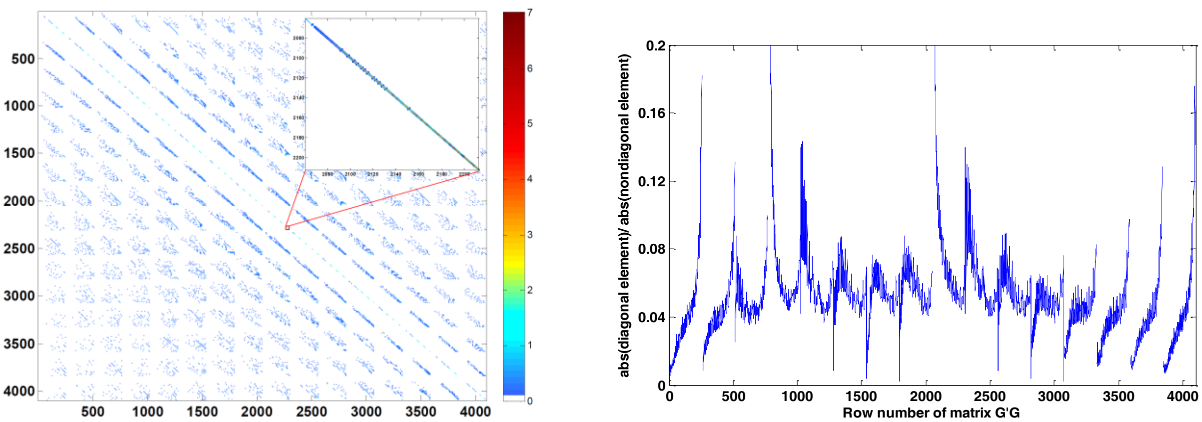
It can be shown that the Jacobi and Gauss-Seidel methods converge to the solution if matrix A in Equation 7 be diagonally dominant and the convergence rate is slower for lower frequencies of the solution (Strang, 1986). Removing high frequency contents from residuals in the Jacobi (or the Gauss-Seidel) first few iterations produces a smooth error vector including mostly low frequency contents (Briggs et. al., 2000).

In the method of multigrid, an iterative solver produces low frequency contents in residual after a few iterations on Equation 7. By restriction, the kernel of the main problem and its residual are transferred to a coarser grid, where the low frequency components act as the high frequency components. Solving the original equation with a solution to the problem in the coarse grid as the starting point returns an answer which contains more low frequency contents than the solving equation with a vector of zeros as the initial guess (Strang, 1986). This algorithm called a v-cycle multigrid. The algorithm can be extended to the finer grids (V-cycle), or to perform more iteration on the coarser grids (W-cycle) (Strang, 1986; Briggs et. al., 2000).

Jacobi and Gauss-Seidel are conventional multigrid solvers. In both, it is necessary to extract the diagonal elements of the matrix A and invert it. Therefore, in order to apply multigrid to the LSPSTM, it is necessary to have its Hessian matrix in the explicit form. The size of matrix G in LSPSTM equals to the size of the data multiplied by the size of the migration image. Therefore, G can be large enough to be impossible to be loaded in the memory of today's computers.

However, the experiences with the explicit form of Hessian matrices, for different data acquisition geometries show that they are not diagonally dominant. For instance, for a modeling operator with two sources with 180 m interval spacing and five receivers per source with 72 m interval spacing and with a model 632 m long distance and 0.512 seconds depth, matrix $G'G$ has 10% nonzero elements as shown in Figure 1a. With eight bites per word precision, $G'G$ needs more than 120 megabyte memory for this small synthetic example. Figure 1b shows the ratio of absolute values of diagonal elements to the sum of absolute values of nondiagonal elements for each row of the $G'G$ matrix. Examples show that adding a reasonably large amount to the diagonal elements of the Hessian matrix or applying restriction to that does not change it to a diagonally dominant matrix.

Splitting $G'G$ matrix to be correspondent to columns of LSPATM image in order to solve the problem by inverting one column at each time does not produce diagonally dominant matrices.



(a) (b)

Figure 1. (a) Non-zero elements of matrix $G'G$. (b) The ratio of absolute values of diagonal elements to the sum of absolute values of nondiagonal elements for each row of $G'G$. Values higher than 1 are diagonally dominant rows.

Multigrid LSPSTM with CG as the solver

CG is a powerful method for solving a system of linear equation. Without requirement of large memory size, after a few iterations, CG Least-Squares (CGLS), an extension of CG for the normal equations, retrieves a high resolution image of the earth subsurface reflectivity. However, CG does not have a smoothing property. In fact, CG methods are roughers and not smoothers (Shewchuk, 1994). This property is shown on a LSPSTM problem. Figures 2a and 2b show the convergence rate of CGLS for a synthetic model with different dominant frequencies in the data. These figures show no better convergence when data include higher frequency content in both damped and smoothed LSPSTM. Convergence of CG method for LSPSTM does not depend on the frequency contents of data.

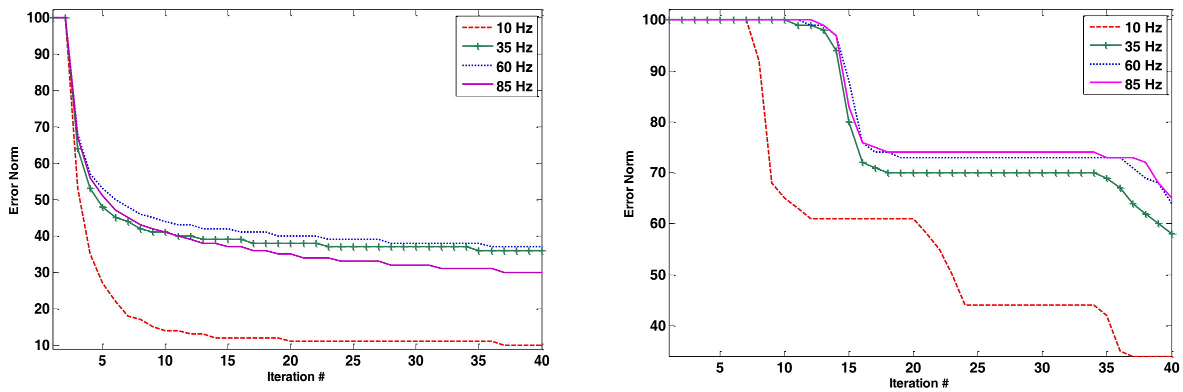
When an iterative method dose not converge faster for data with higher frequency contents than the data containing lower frequency contents, it is not able to leave low frequency contents in the residuals and act as the smoother. Therefore, CG methods should not be an effective solver for the multigrid.

There are three main approaches to applying multigrid to a LSPSTM equation. In order to transfer the problem to a higher or lower grid sizes, restriction and interpolation of a LSPSTM problem can be applied in each, horizontal or distance, vertical or time, or both directions of the model. As shown in the previous section, applying multigrid in the vertical (time) direction should not improve the performance of LSPSTM since the convergence is not faster for the lower frequency components of the data. The analyses of using multigrid CGLS with restriction and interpolation in the distance direction is shown by comparison between multigrid CGLS and CGLS (Figure 3). Restriction to a coarser grid is performed by deleting half of the traces (leavening one trace and removing next one) from the migration image. Figure 3a is the true model, Figure 3b shows the image of Kirchhoff LSPSTM with five iterations on the CGLS, and same results is obtained from performing a full multigrid CGLS with five iterations on each grid (Figure 3c).

Conclusion

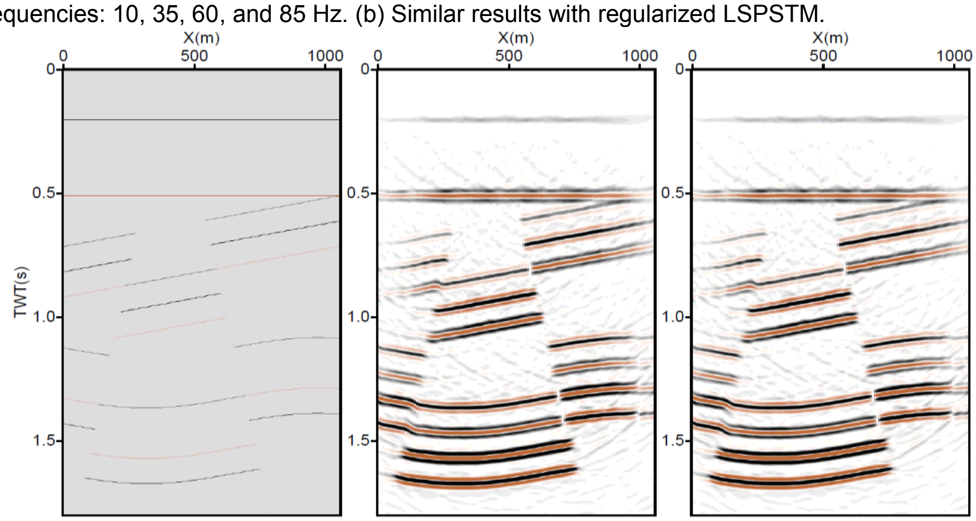
Numerical examples showed that time domain LSPSTM problem is not solvable by the Jacobi or Gauss-Seidel iterations. Consequently, the conventional multigrid is not viable to the mentioned problem. Requirement of large memory size is another problem associated with this method.

CG is an effective solver. However, CG is not a smoother. Therefore, using CG as the multigrid solver does not increase the speed of convergence or provide a better recovery of the low frequency contents. Using multigrid with CG as the iterative solver may slightly reduces the number of iterations for the same rate of convergence in comparison to the CGLS by introducing an initial value. However, it may not reduce the total computational cost.



(a) (b)

Figure 2. (a) Convergence of CGLS to solve damped LSPSTM for a synthetic model with wavelets having different dominant frequencies: 10, 35, 60, and 85 Hz. (b) Similar results with regularized LSPSTM.



(a) (b) (c)

Figure 3. Comparison between true model (a), CGLS (b), and Multigrid CGLS (c) of a synthetic seismic dataset.

Acknowledgements

Authors wish to acknowledge the sponsors of CREWES project for their continuing support. They also appreciate the computer technicians in the CREWES project.

References

Briggs, W. L., Henson, V. E., and McCormick, S. F., 2000, A multigrid tutorial, second edition, SIAM.

Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflection data: Geophysics, 64, no. 1, 208-221.

Shewchuk, J. R., 1994, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain: School of Computer Science, Carnegie Mellon University.

Strang, G., 1986, Introduction to applied mathematics: Wellesley-Cambridge Press.

Stuben, K. 2002, Multigrid Tutorial, Institute for Pure & Applied Mathematics.
http://www.ipam.ucla.edu/publications/es2002/es2002_1449.pdf (20/08/2009).
Yousefzadeh, A. 2008, Seismic Data Reconstruction with Regularized Least-Squares Prestack Kirchhoff Time Migration, MSc. thesis, University of Alberta.