Integration geoConvention 2013 geoscience engineering partnership

Full waveform inversion of crosswell seismic data using automatic differentiation

Danping Cao*, China University of Petroleum, Qingdao, China and University of Calgary, Calgary, Canada
caodp@upc.edu.cn
and
Wenyuan Liao, University of Calgary, Calgary, Canada, wliao@ucalgary.ca

Summary

Full waveform inversion (FWI) is an effective and efficient data-fitting technique that has been widely used to produce accurate estimation of model parameters in Geophysics. The efficiency and accuracy of FWI are determined by the three main components: numerical solution for forward problem, gradient calculation and model update which usually involves the optimization method. The success of the adjoint-based field data inversion relies on the choice of the methods for solving these time consuming components. In this paper, we introduce the automatic differentiation (AD) tool TAPENADE to compute the gradient of the objective function, hence the FWI workflow is simplified so we can focus mainly on the forward modeling and the model updating. Based on the result from forward modeling, and the observational data, the objective function is calculated as the misfit function, which is then minimized by utilizing an optimization algorithm. A variety of methods can be used to minimize the misfit function, whereas in this work we choose the limited memory BFGS (L-BFGS) method due to the low requirement on memory usage. Numerical test for a 2D model shows that the combination of the AD tool and L-BFGS method is effective and efficient to solve the full waveform inversion problem.

Introduction

Full waveform inversion (FWI) is a nonlinear data-fitting procedure based on the seismic waveform data to estimate the model parameters, which usually appear as coefficients in a wave equation. The most popular method of FWI is composed of forward modeling which solves the wave equation with the initial model, calculating the gradient of objective function which measures the difference between the synthetic data and field data, updating the model parameters with a valid optimization method (Tarantola, 1984; Lailly, 1983; Virieux, 2009; Fichtner, 2011). Consequently, the three time-consuming parts influence the efficiency and accuracy of the FWI. The numerical method for the forward problem has been extensively discussed in the literature. So, we focus on how to calculate the gradient and the choice of an optimization method in this paper.

Most of the seismic FWI techniques utilize the gradient related optimization algorithms to update the model parameters so the gradient of the objective function is needed at each iteration step. The efficiency of the FWI method greatly depends on the accuracy and efficiency of the gradient computation. The adjoint state method has been introduced in the theory of inverse problems in the 1970s (Chavent, 1974), which is well-known to be an effective and efficient technique for computing the gradient of the objective function in FWI (Plessix, 2006; Fichtner, 2011). However, this approach is still computational expensive, especially for the problem with large size parameter, more over, the procedure of numerical solution of the adjoint equation is error-prone due to hand coding. Alternatively, this programming work can be substituted by the automatic differentiation technique that calculates the gradient with the adjoint method.

Automatic differentiation (AD) is the technique whereby the output variables of a complicated computer code can be differentiated with respect to the input variables automatically and efficiently. AD has been widely used in areas such as optimization, meteorology and oceanography but appear to have little impact in the Geosciences (Sambridge, 2007). Fortunately, AD has attracted increasing attention from Geoscientists and been introduced into the full waveform inversion domain in the recent years. Tan (2010) has shown that the automatic differentiation is an efficient approach to verify the accuracy of the gradient and Hessian operations generated by the adjoint state method. Liao (2011) has successfully estimated the acoustic coefficient in a 2D acoustic wave equation using the AD method. We believe that with the increasing popularity of adjoint state method and the increase of computing power, it should be a good practice to make full use of AD tool to solve the gradient of the objective function in the FWI workflow. In fact, there are a variety of AD tools that can be used to generate the adjoint code for the purpose of gradient calculation, although they vary in all aspects such as efficiency, easiness for use, etc.

To minimize the objective function, there are many optimization methods that can be used to update the model parameters in the FWI workflow. The method solving the Hessian vectors is referred to as the Newton method, which in general is computationally costly. In fact, the pure Newton method based on the full Hessian is currently not being used in realistic FWI because of the high computational cost (Virieux, 2009). The simple and attractive choice is to substitute the inverse of the Hessian with a carefully chosen step length. Such modification results in the so called gradient method or steep-descent method, which are widely used in solving the gradient-based optimization problem. The mostly used optimization methods include the conjugate-gradient method, the quasi-Newton method and Gauss-Newton method in the literature.

In this article, these main computational parts will be discussed in detail while the workflow of FWI using AD tool will be presented. In particular, we use TAPENADE (Hascoët, 2004), one of the most powerful AD tools to compute the gradient of the objective function in the full waveform inversion, in which the objective function is defined as the difference in 2-norm between the observational data and synthetic seismogram. The optimization problem is then solved by the limited memory BFGS (L-BFGS) method. The proposed computational framework is tested on a 2D acoustic wave equation with crosswell seismic datasets.

Method

FWI is an unconstrained optimization problem in which we minimize the least squares difference between the observed data and the synthetic seismogram, which is the result of the forward model using an initial model **m**,

$$E(\mathbf{m}) = \frac{1}{2}\left(\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})\right)^T \left(\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})\right) \tag{1}$$

Where $\mathbf{d}_{obs}$ is the recorded seismic data, $\mathbf{d}_{cal}(\mathbf{m})$ is the predicted seismic data according to the model $\mathbf{m}$ at the designed shot-receivers position. The model $\mathbf{m}$ represents the physical parameters of the subsurface discretized over the computational domain.

Using the second-order Taylor-Lagrange expansion of the misfit function (1) at the vicinity of initial model $\mathbf{m}_0$, one can obtain the derivative with respect to the model parameter, then the perturbation model vector at the initial model $\mathbf{m}_0$ can be expressed as (Virieux, 2009),

$$\Delta\mathbf{m} = -\left[\frac{\partial^2 E(\mathbf{m}_0)}{\partial \mathbf{m}^2}\right]^{-1} \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} \tag{2}$$

The perturbation model is searched in the opposite direction of the steepest ascent of the misfit function at the initial model $\mathbf{m}_0$. The second derivative of the misfit function is the Hessian that defines the curvatures of the objective function. Then, the updated model $\mathbf{m}$ can be written as the sum of the initial model $\mathbf{m}_0$ plus the perturbation model,

$$\mathbf{m}=\mathbf{m}_0+\Delta\mathbf{m} \tag{3}$$

With the objective function and the model update equation, the workflow of full waveform inversion can be illustrated as the following (Figure 1)

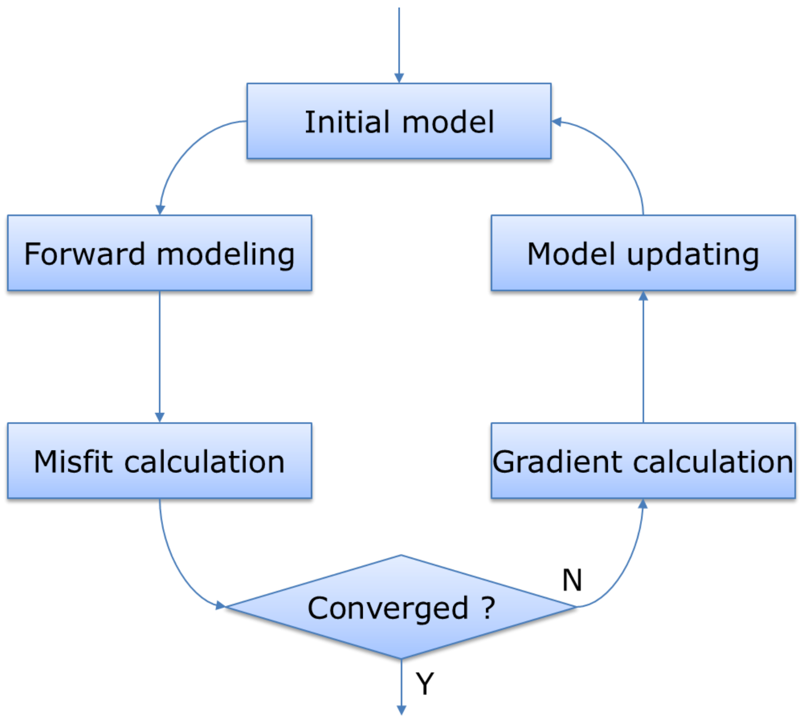


Figure 1: Workflow of full waveform inversion

As shown in Figure 1, the FWI workflow consists of four parts: forward modeling, misfit calculation, gradient calculation and model updating. With the given initial model and field data, we first calculate the misfit between synthetic data and field data. If convergence is reached, the iteration stops, otherwise, the gradient will be calculated and used to minimize the misfit function. The result from the minimization will then be used to update the model, then a new iterative step starts until convergence is reached. Apparently, most of the computational resources are used on the three parts: forward modeling, gradient calculation and optimization. How to simplify the workflow and the selection of a valid method to solve the optimization problem are critical to the overall efficiency and accuracy of the seismic inversion problem. In what follows, we will give a detailed description for each part of the computational framework.

The forward modeling is the basis of the whole inversion problem. In this paper we produce the synthetic data by solving the forward acoustic wave equation, based on an initial velocity model. In general, one can use any numerical method to solve the forward problem, however higher-order methods are preferred due to its' accuracy and low numerical dispersion. Here we use a finite difference method that is second-order in time and fourth-order in space. To reduce the effect from the reflection occurred at the boundary, the absorbing boundary condition is implemented in the forward modeling. In the future, finite difference method with fourth-order accuracy in time will be considered.

The gradient calculation is fairly important, as the optimization algorithm relies on the derivative of the objective function to update the model parameters. The gradient can be computed efficiently using the adjoint state method. As mentioned above, the calculation of the derivatives may be problematic and troublesome especially if the procedure solving the adjoint equation is hand-coded. Alternatively, the automatic differentiation (AD) is a good choice because of its quick and simple implementation. The user just needs to provide the AD tool with a source program that solves the forward problem, the AD tool will generate a new program which will compute the gradient of the objective function. All of these procedures are done automatically. The computational efficiency for calculating the gradient by AD tool is determined by the forward numerical modeling program. This feature allows the user to focus on improving the forward numerical modeling without worrying about the implementation and efficiency of the gradient calculation.

Automatic differentiation (AD) is a set of techniques based on the mechanical application of the chain rule to obtain derivatives of a function given as a computer program. There are about 40 AD tools such as TAF, ADIFOR, TAMC, OpenAD, TAPENADE, ADOL-C, ADiMat etc. In this paper, we use TAPENADE to generate the adjoint code and calculate the gradient of the objective function. This AD tool will be introduced into the workflow of FWI to simplify the inversion procedure and improve the efficiency of the whole inversion problem. To ensure that the gradient calculated by TAPENADE (shown in figure 2(a)) is accurate, we compare it with the gradient calculated by central difference quotient (shown in Figure 2(b)). Both gradient results are calculated according to the initial model (shown in Figure 4(a)) and the synthetic data of the true model (shown in Figure 3(a)) that the source located at the middle of the left side. This comparison clearly indicates that the gradient by TAPENADE is accurate, as both results show the same pattern.

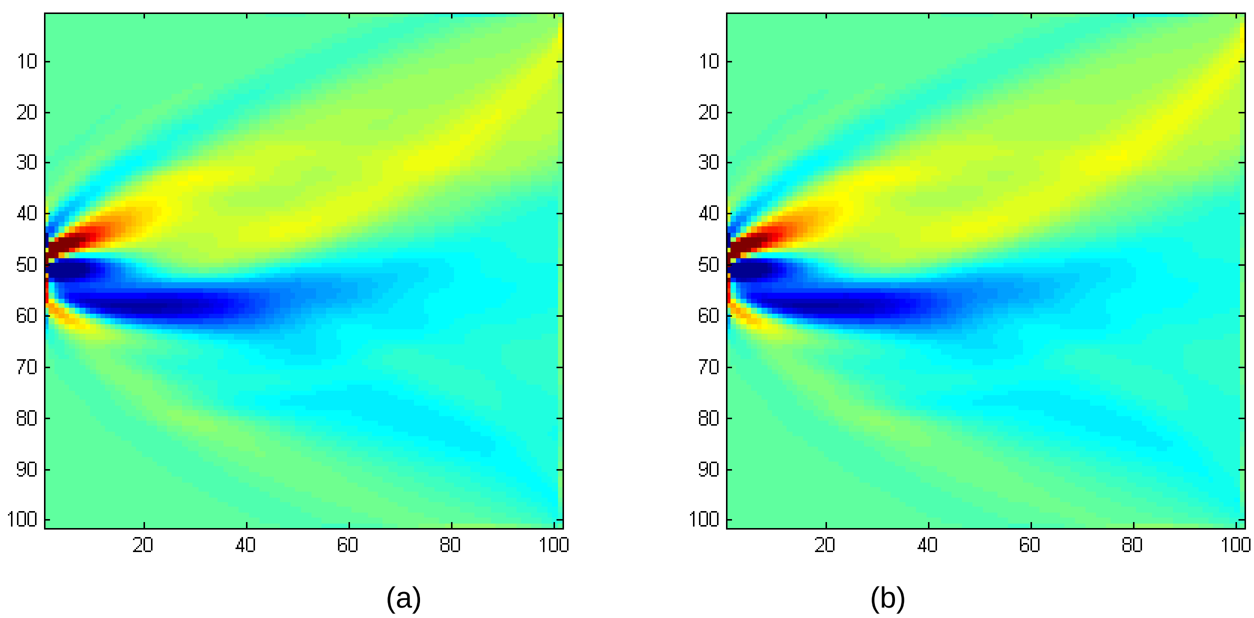


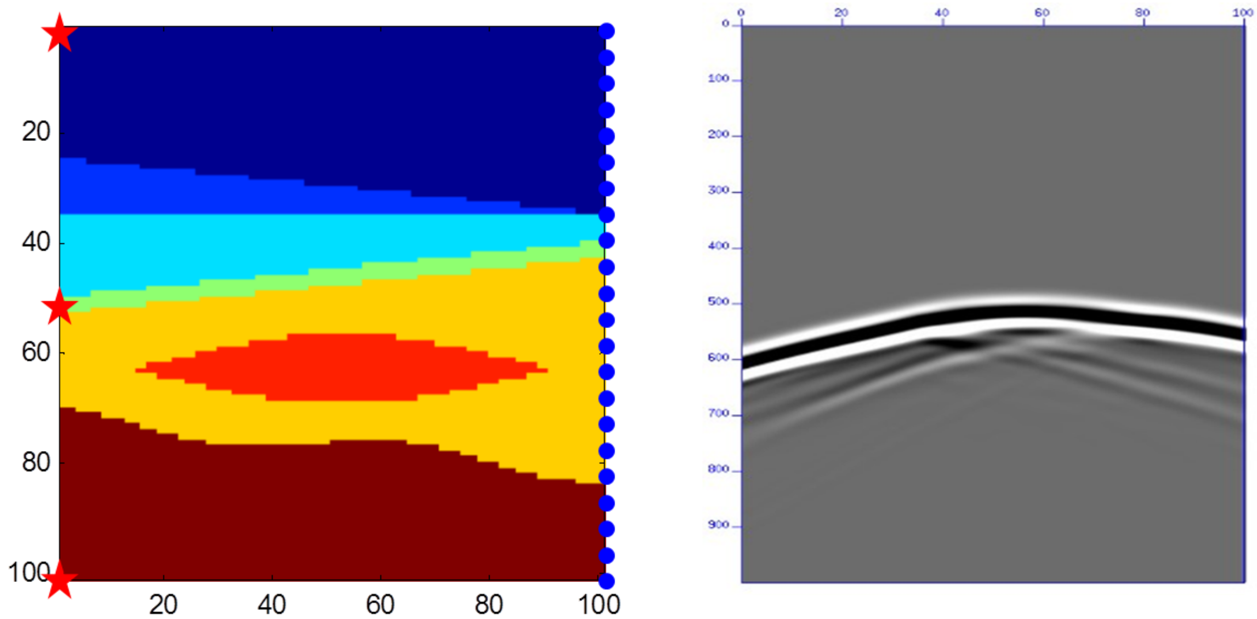
Figure 2: Gradient by TAPENADE (a) and Gradient by central difference quotient (b)

For the model updating method, in principle, there are many different optimization methods that can be used to solve the minimum of the objective function $E(\mathbf{m})$. In this paper the limited memory BFGS (L-BFGS) is implemented to minimize the objective function as it is very efficient in solving the optimization problem of large size parameters. It has been shown that the L-BFGS is very robust and converges faster than many other popular methods such as the standard conjugate gradient method

(Liu, 1989). More over, the memory requirement of L-BFGS is considerably lower than that of other Newton-type methods. The combination of the AD tool (TAPENADE) and L-BFGS method is tested on the 2D model with the geometry of crosswell seismic inversion. For the sake of simplicity, we assume constant density and isotropic medium in the numerical test.

Example

We created a 2D velocity model corresponding to the crosswell survey geometry shown in Figure 3(a) to test our method. The typical sand body such as wedge body and lens-shaped body has been designed in this 2D model. The medium is discretized using uniform vertical and horizontal grid spacing of 1m, which results in a 101×101 grid. The synthetic waveform records were computed using the popular time domain finite difference code. The Ricker wavelet with a central frequency 180Hz is used as the source. To reduce the artificial reflection, the PML absorbing boundary condition is implemented in the forward numerical modeling. Red stars and blue circles in the diagram represent the source and receiver locations, respectively. The synthetic record correspond to the source located at the middle of the model left side is shown at the Figure 3(b).



(a) (b)

Figure 3: Geological model (a) and one synthetic profile (b)

Using the synthetic data of the 2D model, the AD tool mentioned above is used to calculate the gradient and the L-BFGS is used to minimize the objective function. Starting from the initial velocity model shown in Figure 4(a), we recovered the velocity models using different number of shots. The results shown in Figure 4(b) – (d) clearly show that with the increase of the number of shots, the results are getting more accurate with the same number of L-BFGS iterations been taken for each case.

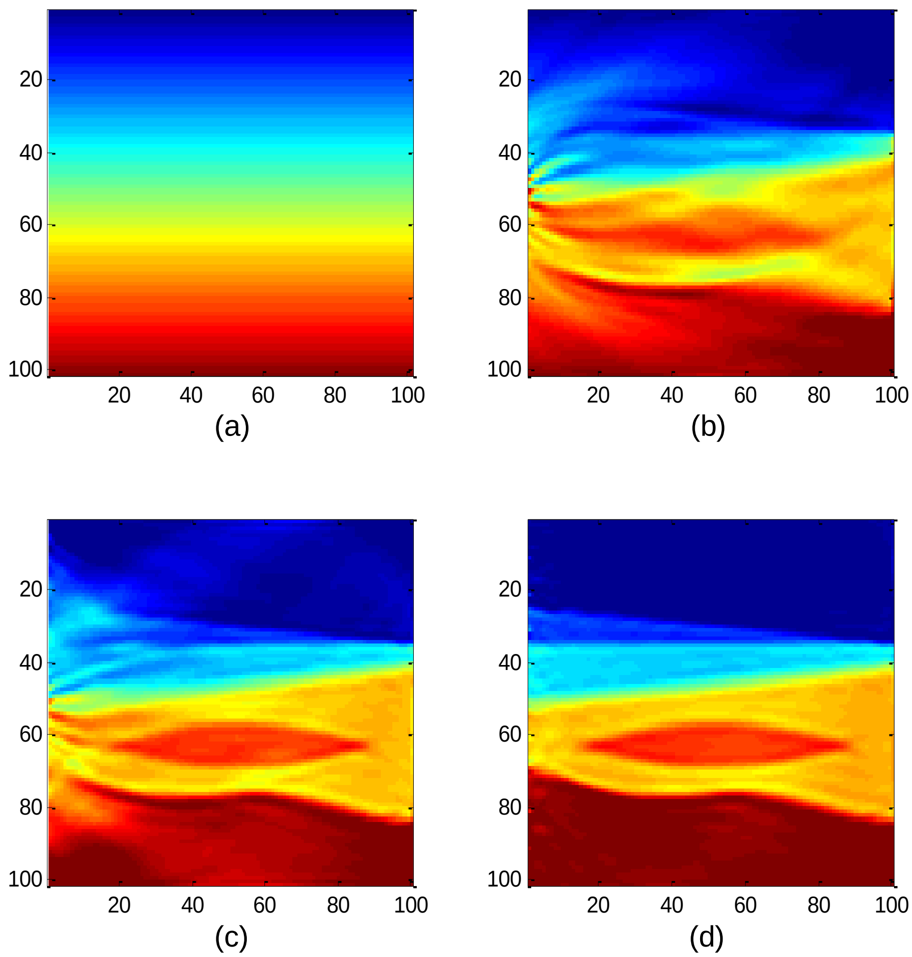


Figure 4: The contrast of initial model (a), the inversion result with 1 shot data (b), the inversion result with 3 shots data (c) and the inversion result with 11 shots data (d).

Clearly the inverted velocity models are reasonably close to the true model without any extra processing to the initial model and the observed data. It is obvious that with the increase in the numbers of shots, the result will be more accurate because of more observational data. The shape and velocity information have been inverted clearly with 11 shots data that cover the whole model. In fact, more observed data will lead to more accurate inversion result and this is true for more complicated model and field data.

Conclusions

Full waveform inversion (FWI) is computationally intensive procedure as each of the three components: the forward modeling, gradient calculation and model updating with optimization method are time consuming tasks. A popular finite difference scheme that is 2^{nd}-order in time and 4^{th}-order in space has been employed for the forward modeling. The automatic differentiation (AD) tool (TAPENADE) has been used to calculate the gradient of the objective function with respect the model parameters automatically and efficiently. The result is then combined with the limited memory BFGS (L-BFGS) method to update the model parameter. The significant reduction of effort in calculating the gradient allows us to focus on the forward modeling part such as the development of more efficient and accurate numerical schemes to solve the wave equation will improve the overall efficiency of FWI. The presented 2D numerical example shows that the proposed computational framework is a promising approach to solve the full waveform inversion of crosswell seismic data.

Acknowledgements

The work of the first author is supported by the Chinese Scholarship Council, National Nature Science Foundation of China (No. 41004050) and 973 Program of China (No. 2013CB228604). The work of the second author is supported by NSERC through his Individual Discovery Grant (RT734491). Both authors would like to thank the generous supports from CREWES, POTSI and the Department of Mathematics and Statistics, University of Calgary.

References

Chavent, G., 1974, Identification of parameter distributed systems, in R.Goodson and M. Polis, eds., Identification of function parameters in partial differential equations: American Society of Mechanical Engineers, 31–48.

Fichtner, A., 2011, Full seismic waveform modeling and inversion, Springer Verlag.

Hascoët, L. and Pascual, V., 2004, TAPENADE 2.1 user's guide, http://tapenade.inria.fr:8080/tapenade/index.jsp

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Tulsa, Oklahoma: Philadelphia, SIAM: 206-220.

Liao, W., 2011, An accurate and efficient algorithm for parameter estimation of 2D acoustic wave equation: International Journal of Applied Physics and Mathematics, 1, 96-100.

Liu, D.C., and Nocedal, J., 1989, On the Limited Memory Method for Large Scale Optimization: Mathematical Programming B, 45, 503-528.

Nocedal, J., 1980, Updating Quasi-Newton Matrices with Limited Storage: Mathematics of Computation, 35, 773-782.

Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, 167, 495–503.

Sambridge, M., Rickwood, P., Rawlinson, N., etc., 2007, Automatic differentiation in geophysical inverse problems: Geophysical Journal International, 167, 1-8.

Tan, L., Brytik, V., Baumstein, A., etc., 2010, Verification of Gradient and Hessian computation for full wavefield inversion using automatic differentiation: SEG Annual Meeting Abstract, 2762-2766.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, 49, 1259-1266.

Virieux, J. and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, WCC1-WCC26