Computer Physics Communications 188 (2015) 47–58

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

A computational method for full waveform inversion of crosswell seismic data using automatic differentiation

Danping Cao [a], Wenyuan Liao [b,*]

[a] School of Geosciences, China University of Petroleum, Qingdao, Shandong, 266580, China
[b] Department of Mathematics & Statistics, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

ARTICLE INFO

Article history:
Received 3 December 2013
Received in revised form
28 October 2014
Accepted 3 November 2014
Available online 10 November 2014

Keywords:
Full waveform inversion
Adjoint state method
Automatic differentiation
Numerical optimization
Inverse problem
Crosswell seismic data

ABSTRACT

Full waveform inversion (FWI) is a model-based data-fitting technique that has been widely used to estimate model parameters in Geophysics. In this work, we propose an efficient computational approach to solve the FWI of crosswell seismic data. The FWI problem is mathematically formulated as a partial differential equation (PDE)-constrained optimization problem, which is numerically solved using a gradient-based optimization method. The efficiency and accuracy of FWI are mainly determined by the three main components: forward modeling, gradient calculation and model update which usually involves the gradient-based optimization algorithm. Given the large number of iterations needed by FWI, an accurate gradient is critical for the success of FWI, as it will not only speed up the convergence but also increase the accuracy of the solution. However computing the gradient still remains a challenging task even after the adjoint PDE has been derived. Automatic differentiation (AD) tools have been proved very effective in a variety of application areas including Geoscience. In this work we investigated the feasibility of integrating TAPENADE, a powerful AD tool into FWI, so that the FWI workflow is simplified to allow us to focus on the forward modeling and the model updating. In this paper we choose the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method due to its robustness and fast convergence. Numerical experiments have been conducted to demonstrate the effectiveness, efficiency and robustness of the new computational approach for FWI.

1. Introduction

Full waveform inversion (FWI) is a model-based nonlinear data-fitting procedure taking the seismic waveform data to estimate the model parameters, which usually appear as coefficients (e.g. wave velocity) in the seismic wave equation. The most popular method of FWI is composed of three main steps: forward modeling, which solves the seismic wave equation based on an initial guess of the model parameters; calculating the gradient of the objective function which measures the difference between the synthetic seismogram and the observational field data and updating the model parameters with an optimization method [1–6]. Apparently, the efficiency and accuracy of the FWI are determined by the three time-consuming steps. Among the three steps, the forward modeling has been extensively discussed in the literature, and a variety of accurate and efficient methods have been developed to solve the seismic wave equation. Except for efficiency and accuracy, another difficulty in forward modeling is wave reflection on the domain boundary, which is usually treated by perfectly matched layer (PML) or other type of absorbing boundary condition (ABC). As mentioned early, the gradient calculation is a rather challenging task, despite extensive efforts that have been devoted into this area. It is widely accepted that the adjoint state method is an efficient and cost-effective solution to this problem. Another important component of the proposed computational method is a gradient-based optimization method, which also affects the overall efficiency and accuracy of the inverse problem.

Many seismic FWI techniques utilize the gradient related optimization algorithms to update the model parameters, hence the gradient calculations are unavoidable, especially if the model parameters are in large size. The adjoint state method has been introduced in the theory of inverse problems in the 1970s [7], however it has a pretty long history dating back to Lagrange's work, in which he presented the famous Lagrange identity to define the adjoint operator. Recently, many developments in the adjoint state method have been made to compute the gradient of the objective function in FWI [1,8] and other areas [9,10], to name a few.

* Corresponding author.
E-mail address: wliao@ucalgary.ca (W. Liao).

http://dx.doi.org/10.1016/j.cpc.2014.11.002
0010-4655/

Generally speaking, there are two equivalent ways to implement the discrete adjoint state method: adjoint-then-discretize and discretize-then-adjoint. Extensive comparisons and introductions of continuous and discrete adjoints can be found in [11–17] and references therein. In the first case, one derives the adjoint partial differential equation (PDE) of the seismic wave equation using perturbation theory or Lagrange multiplier method, then numerically solve the adjoint PDE. While in the second approach, one first derives the numerical algorithm to solve the forward problem, then derives the so called discrete adjoint based on the numerical scheme. The two approaches are equivalent in the sense that they produce the same gradient in machine precision, if the same numerical scheme is used to solve the forward and adjoint PDEs. More information regarding the approximation of adjoint gradients on other applications can be found in [18–21]. When the forward model is linear, the adjoint PDE is generally in a simple form and similar to the forward model, hence numerical solution of the adjoint PDE is not very difficult. However in the general case, it is well known that the procedure of numerical solution for the adjoint PDE is error-prone due to hand coding, and time consuming. Moreover, given that PML or other types of ABC is involved in forward modeling, derivation and implementation of adjoint PDE becomes more difficult. Alternatively, this programming work can be substituted by the automatic differentiation technique that calculates the gradient with the adjoint method.

To overcome this difficulty, automatic differentiation (AD) has been introduced for the purpose of gradient calculation. AD is the technique that has been used to generate a computer code based on the forward modeling computer code. Simply speaking, it is a source to source code translator implementing discrete adjoint. Given a computer program evaluating the function $f(x)$, the output of AD is a new computer program evaluating $f'(x)$. AD has been widely used in many areas such as optimization, meteorology, oceanography and Geoscience applications [22–30], however it is still not very popular for seismic inversion problem. Fortunately, AD has attracted increasing attention from Geoscientists and has been introduced into the full waveform inversion domain in the recent years. Tan [31] has shown that AD is an efficient approach to verify the accuracy of the gradient and Hessian operations generated by the adjoint state method. Liao [32] has successfully estimated the acoustic coefficient in a simple 2D acoustic wave equation using the AD tool TAMC. 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.

Another important components in FWI is the optimization procedure. There are many optimization methods that can be used to update the model parameters in the FWI workflow. The method of solving using the Hessian vectors is referred to as the Newton method, which in general is computationally costly. In fact, the full Hessian based pure Newton method is currently not being used in realistic FWI because of the high computational cost [6]. The simple and attractive choice is to substitute the inverse of the Hessian with some simplified approximation. Such modification results in the so-called gradient method or steepest-descent method, which are widely used in solving the gradient-based optimization problem. The mostly used optimization methods include the conjugate-gradient (CG) method, the Quasi-Newton method and Gauss–Newton method in the literature.

In particular, we use TAPENADE [33], 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 Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method. The proposed computational framework is tested on a 2D acoustic wave equation with crosswell seismic datasets. The rest of the paper is organized as follows. In Section 2, the FWI problem is mathematically formulated as a PDE-constrained optimization problem, which is followed by Section 3 in which a brief review of the mathematical foundation of the adjoint state method and a short introduction to the AD tool will be presented. In Section 4 we describe the optimization algorithm for solving the optimization problem, followed by several numerical test cases in Section 5. Finally, the conclusions and possible future extensions are discussed in Section 6.

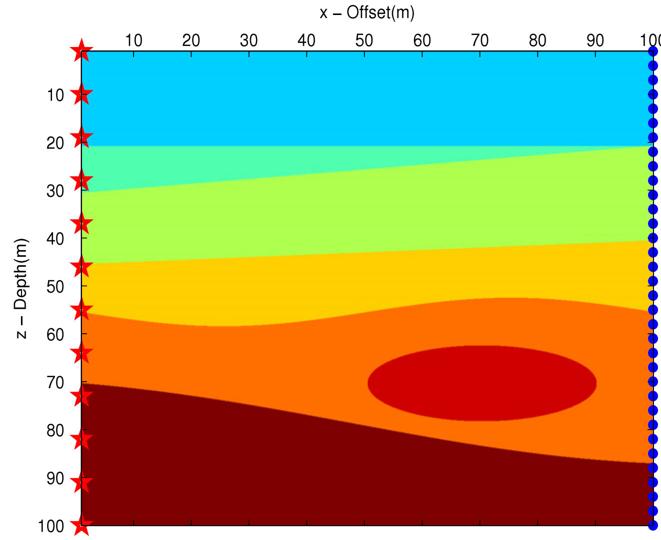

Fig. 1. Velocity model between two vertical wells and acquisition system. The stars represent the sources located at the left side well, the dots represent the receivers placed at the right side well.

2. Mathematical formulation of FWI

In this section we formulate the FWI problem as an optimization problem constrained by the 2-D crosswell seismic model in a heterogeneous medium given by

$$\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} + f(x, z, t), \quad (x, z, t) \in \Omega \times [0, t_f], \quad (1)$$

where $v = v(x, z)$ is the wave velocity, Ω is a 100 m × 100 m square domain shown in Fig. 1, t_f is the end of time domain, $f(x, z, t)$ represents the seismic source generated at the well on the left side, while the receivers are located at the well on the right side. To simplify the description of the numerical procedure, in what follows we ignore the source term, which will be added later.

As shown in Fig. 2, there are three main components in the inverse problem: Forward modeling, gradient calculation and model updating. Starting from an initial guess of the wave velocity, the forward problem is solved to calculate the seismogram, which is then compared with the observational data to calculate the misfit function. Based on the forward problem and the misfit function, the gradient with respect to the model parameters is calculated. Using some gradient-based optimization algorithm, the model is updated, and the procedure moves to the next loop till a good match between the calculated seismogram and observational data is reached. In what follows we describe these components in detail. We first focus on the numerical solution of the seismic wave equation and calculation of the objective function. Many efficient