Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Optimization of computations for adjoint field and Jacobian needed in 3D CSEM inversion

Rahul Dehiya, Arun Singh, Pravin K. Gupta, M. Israil*

Department of Earth Sciences, Indian Institute of Technology Roorkee, Roorkee 247667, India

ARTICLE INFO

Article history:
Received 14 July 2016
Received in revised form 17 November 2016
Accepted 19 November 2016
Available online 23 November 2016

Keywords:
Inverse modeling
Controlled-Source Electromagnetic
Numerical optimization

ABSTRACT

We present the features and results of a newly developed code, based on Gauss-Newton optimization technique, for solving three-dimensional Controlled-Source Electromagnetic inverse problem. In this code a special emphasis has been put on representing the operations by block matrices for conjugate gradient iteration. We show how in the computation of Jacobian, the matrix formed by differentiation of system matrix can be made independent of frequency to optimize the operations at conjugate gradient step. The coarse level parallel computing, using OpenMP framework, is used primarily due to its simplicity in implementation and accessibility of shared memory multi-core computing machine to almost anyone. We demonstrate how the coarseness of modeling grid in comparison to source (comp'utational receivers) spacing can be exploited for efficient computing, without compromising the quality of the inverted model, by reducing the number of adjoint calls. It is also demonstrated that the adjoint field can even be computed on a grid coarser than the modeling grid without affecting the inversion outcome. These observations were reconfirmed using an experiment design where the deviation of source from straight tow line is considered. Finally, a real field data inversion experiment is presented to demonstrate robustness of the code.

1. Introduction

Controlled-Source ElectroMagnetic (CSEM) is considered to be a credible tool for subsurface electrical resistivity imaging. It continues to evolve steadily, and has a potential to be placed after 3D seismic in hierarchy of importance for hydrocarbon exploration. Strack (2014) argued that CSEM method is almost at the end of its conceptual phase, and last decade has seen development of numerous technologies out of which only the operationally mature ones have survived. During this period, many surveys acquiring 3D CSEM data have been carried out. Interpretation of these data sets requires the development of 3D inversion algorithms. Considerable efforts have been invested in the development of accurate and computationally efficient algorithms (Newman and Alumbaugh, 1997; Sasaki, 2001; Zhang, 2003; Commer and Newman, 2008).

In the linearized scheme of inversion algorithms, the model parameters are updated via an iterative procedure. These updates require solving many simulations, sometimes in the order of hundreds to thousands of forward problems per iteration for the 3D data set. These forward calls typically are the most computationally intensive part of inverse modeling. Iterative solvers are generally preferred for these computations due to the highly sparse matrix of forward kernel and the less memory requirement. Recently, some studies have demonstrated certain advantages of direct solver for both forward (e.g., Streich, 2009) and inverse modeling (e.g., Grayver et al., 2013). Another study using a direct solver in combination with Schur compliments method demonstrated the advantages in scenarios where inversion domain is small as compared to the modeling one (Jaysaval et al., 2014).

In a nonlinear optimization problem, Newton method and its variants such as Gauss-Newton (GN), quasi-Newton etc. have been implemented in many geophysical problems. The Newton method converges quadratically, but it has not been implemented widely because the computation of the second order derivative of the predicted response (Hessian), needed in this method, is considered prohibitively expensive (Haber et al., 2000). GN provides a good compromise between computational cost and performance. GN only requires the computation of the first order derivatives of the residuals (building the Jacobian matrix), and the convergence speed is far better than the one achieved by gradient-based methods. In most studies, explicit formation of the Jacobian is avoided because of its large storage requirement and it is bypassed through Krylov subspace based iterative solver like conjugate gradient (CG) (e.g. Mackie and Madden, 1993;Newman and Alumbaugh, 1997). Some studies have illustrated the advantage of explicitly forming the

* Corresponding author.
E-mail address: mohammad.israil@gmail.com (M. Israil).

http://dx.doi.org/10.1016/j.jappgeo.2016.11.018
0009-2541/© 2016 Elsevier B.V. All rights reserved.

Jacobian for case studies like land CSEM (Grayver et al., 2013). In both cases, whether the Jacobian is formed or bypassed (using CG), the adjoint technique (McGillivray and Oldenburg, 1990;Abubakar et al., 2008) can be used for efficient computation. Another popular scheme is the non-linear conjugate gradient (NLCG) method, where the gradient is used to find a new conjugate direction for model update and step length is estimated using line search methods. This algorithm requires three forward calls per source per iteration. The efficiency achieved due to less number of forward calls per inversion iteration is offset by the larger number of inversion iterations it takes to converge to a desired level of accuracy. Rodi and Mackie (2001) showed, for 2D MT case, that NLCG without pre-conditioner is not as efficient as GN-CG (GN with CG). Pre-conditioned NLCG has also been implemented to improve the convergence (Newman and Alumbaugh, 2000; Rodi and Mackie, 2001; Newman and Boggs, 2004) but it increases the cost per iteration. The computation cost of pre-conditioned NLCG and GN-CG is found to be comparable by Rodi and Mackie (2001). Superiority of NLCG over GN-CG is a debatable issue. For detailed review of these methods the readers are referred to Avdeev (2005) and Siripunvaraporn (2012).

In this contribution, we first briefly discuss the forward modeling algorithm, and then we revisit the GN-CG inversion scheme. We discuss the Jacobian (or its transpose) matrix-vector multiplication in detail where the Jacobian is represented using block matrices. The Jacobian component formed due to system matrix differentiation is made frequency independent. This reformulation adds efficiency at the step of model updating via CG. We also briefly discuss the parallel implementation of our code on shared memory system under the framework of OpenMP. Through numerical experiments, we discuss the reasoning of computing the adjoint field at selected grid nodes rather than at receiver positions, in cases where grid spacing is larger than the receiver spacing. This numerical test also studies the impact of such an implementation on adjoint field computations when there exists source drifting from the straight tow line.

2. Forward modeling algorithm

CSEM forward problem is solved using the vector Helmholtz equation. We have implemented the scattered source (primary/secondary decomposition) approach (Alumbaugh et al., 1996). In the frequency domain, equation for secondary electric field $\mathbf{E}^s$ can be written as,

$$\triangledown \times \triangledown \times \mathbf{E}^s - i\omega\mu_0\sigma^*\mathbf{E}^s = i\omega\mu_0(\sigma^* - \sigma^{p*})\mathbf{E}^p, \tag{1}$$

where $\mathrm{i} = \sqrt{-1}$, superscripts p and s denote the primary and secondary quantities, ω represents the angular frequency, magnetic permeability of free space in denoted by μ_0, the complex conductivity $\sigma^* = \sigma + i\omega\varepsilon$ consists of conductivity σ and permittivity ε, while σ^{p^*} represents the background complex conductivity. The background medium is considered as 1D layered model for which the primary field is computed using semi analytical methods given by Løseth and Ursin (2007).

We implemented the finite difference method on a staggered grid (Yee, 1966) to solve the forward modeling problem under the boundary condition that the secondary electrical field parallel to boundary faces vanishes. This leads to a linear system as,

$$\mathbf{A}\mathbf{e}^s = \mathbf{b}, \tag{2}$$

where $\mathbf{A}$ is a $N_e \times N_e$ finite difference system matrix with maximum 13 non-zero elements per row, $\mathbf{e}^s$ is a vector containing secondary field for all internal nodes and $\mathbf{b}$ contains the scattered source information. Both $\mathbf{A}$ and $\mathbf{b}$ depend on conductivity of the medium. The system matrix is transformed to a symmetric form by pre-multiplying it with a diagonal matrix whose elements depend on cell volume (Fomenko and Mogi, 2002). At boundary faces,

$$\mathbf{e}^s \times \hat{n} = 0, \tag{3}$$

where $\hat{n}$ is the unit vector normal to the boundary face.

The system matrix Eq. (2) is solved using pre-conditioned Bi-conjugate gradient stabilized (BiCGStab) iterative solver. Incomplete Cholesky factorization of sub block of the system matrix (Mackie et al., 1994) with zero level of filling (ICLU(0)) is used as a pre-conditioner. At the static limit $\omega \rightarrow 0$, the terms containing the conductivity in left hand side of Eq. (1) tend to zero. This adds non-uniqueness to the problem and also slows down the convergence considerably. To overcome this issue, a static correction, proposed by Smith (1996), is applied periodically. Once the secondary field is computed, the total electric field is given as

$$\mathbf{e} = \mathbf{e}^p + \mathbf{e}^s. \tag{4}$$

Using the Maxwell's equation, the computation of the magnetic field from electric field is straight forward. It requires a simple transformation matrix which is a discrete approximation of scaled curl operator where the scaling factor is reciprocal of $i\omega\mu_0$.

3. Inverse modeling algorithm

3.1. Formulation of the problem

The inverse problem is posed as the optimization of the objective functional $\phi(\mathbf{m})$ defined as,

$$\phi(\mathbf{m}) = \phi_d(\mathbf{m}) + \lambda\phi_m(\mathbf{m}, \mathbf{m}^{ref}), \tag{5}$$

where ϕ_d is defined as the scalar product of weighted misfit between observed data and predicted data. The ϕ_m depends on model parameters and is generally defined as model smoothness. ϕ_d can be expressed as,

$$\phi_d(\mathbf{m}) = \frac{1}{2}\left[\mathbf{d}^{obs} - \mathbf{f}(\mathbf{m})\right]^H \mathbf{W}_d^T\mathbf{W}_d\left[\mathbf{d}^{obs} - \mathbf{f}(\mathbf{m})\right], \tag{6}$$

where superscript H represents the Hermitian transpose. $\mathbf{d}^{obs} = (d_1^{obs}, d_2^{obs}, \ldots, d_{N_d}^{obs})^T$ is a complex vector containing N_d observed data points, $\mathbf{f(m)}$ denotes predicted data, where $\mathbf{f}$ is forward operator which maps real model parameters to complex data values and $\mathbf{W}_d$ is a $N_d \times N_d$ diagonal matrix representing data weighting. Generally, the diagonal elements of $\mathbf{W}_d$ are reciprocal of the data standard deviation (or amplitudes). Though the prime objective is to minimize $\phi_d(\mathbf{m})$ one has to redefine the problem as minimization of $\phi(\mathbf{m})$ because the minimization of first term ($\phi_d(\mathbf{m})$) is an unstable process and the second term $\phi_m(\mathbf{m}, \mathbf{m}^{ref})$, known as regularization functional, stabilizes the optimization (Tikhonov and Arsenin, 1977; Constable et al., 1987). The $\mathbf{m}^{ref}$ contains a *priori* information. λ is the scalar trade off parameter which controls the influence of regularization term over misfit.

The regularization functional is defined as,

$$\phi_m\left(\mathbf{m}, \mathbf{m}^{ref}\right) = \left(\mathbf{m} - \mathbf{m}^{ref}\right)\mathbf{W}_m^T\mathbf{W}_m\left(\mathbf{m} - \mathbf{m}^{ref}\right), \tag{7}$$

where $\mathbf{W}_m$ defines the model smoothness which is taken as finite difference approximation of Laplacian ($\triangledown^2$).