CONF- 970847--2

SAN097-1296C

SAND-97-1296C

3D ELECTROMAGNETIC INVERSION USING CONJUGATE GRADIENTS

RECEIVED

JUN 06 1997

G.S.T.I

Gregory A. Newman and David L. Alumbaugh
Sandia National Laboratories
P.O. Box 5800, Albuquerque NM 87185-0750
Tel: (505) 844-8158; Fax (505) 844-7354
Email: ganewma@sandia.gov

ABSTRACT

In large scale 3D EM inverse problems it may not be possible to directly invert a full least-squares system matrix involving model sensitivity elements. Thus iterative methods must be employed. For the inverse problem, we favor either a linear or non-linear (NL) CG scheme, depending on the application. In a NL CG scheme, the gradient of the objective function is required at each relaxation step along with a univariate line search needed to determine the optimum model update. Solution examples based on both approaches will be presented.

INTRODUCTION

Because of computational demands, inversion of frequency-domain electromagnetic data arising from complex geologic media has only been a dream. In fact it is only within the last few years that the ability to forward model complex 3D geology is now emerging. Since modeling complex geology is no longer a barrier, this has also opened up the ability to image it. Important applications of the new imaging technology are now arising in environmental waste characterization and resource exploration.

Within the last two years, we have developed solutions to the 3D inverse problem for frequency-domain applications ([1], [2], [3] and [4]), where 2.5 and 3D imaging codes have been used to reconstruct the conductivity within the subsurface. The inversion codes include an efficient finite-difference forward-modeling algorithm which is necessary to compute predicted data and model sensitivities at high discretization levels, and hence we are able to image complex geological media. In this talk, we will review our efforts in inversion employing iterative conjugate gradients methods.

THE INVERSE PROBLEM

Regularized Least Squares

All least squares solutions begin by minimizing the difference between observed and predicted data, often subject to some sort of constraint needed to stabilize the inversion process. In underdetermined problems, which we plan to consider here, stability is provided with

Tikhonov regularization, where only smoothly varying models are sought. Let's divide the earth into M cells and assign to each cell an unknown conductivity value. Further let $\mathbf{m}$ be a vector that describes these values, which is of length M . We now form an objective functional, ϕ , which combines the data error and model smoothness in the following fashion:

$$\phi = \sum_{j=1}^N \{ (d_j^{\text{obs}} - d_j) / \epsilon_j \}^2 + \lambda \mathbf{m}^T \mathbf{W}^T \mathbf{W} \mathbf{m}. \quad (1)$$

In equation (1) the terms that describe the observed and predicted are d_j^{obs} and d_j , where the summation is over N data points. We also weight the data errors in equation (1) by the standard deviations of the observed data, ϵ_j . This gives noisy data smaller weight when forming ϕ .

The parameters that dictate model smoothness are the regularization matrix $\mathbf{W}$, which consists of a finite difference approximation to the Laplacian (∇^2) operator and the tradeoff parameter, λ , which is used to control the amount of smoothness to be incorporated into the reconstruction. Different strategies for selecting λ can be found in [1] and [5].

In small scale inverse problems it may be feasible to determine the minimum of equation (1) with a brute force search in parameter space. For large scale problems, such as ours, this is not an option. Instead, what is typically done is to set the gradient of the objective function, $\nabla \phi$, with respect to the model parameters to zero, and find by some economical means those model parameters that satisfy $\nabla \phi = 0$. Because the predicted data depend on the model, $\mathbf{m}$, in a non-linear fashion we are forced to solve $\nabla \phi = 0$ using an iterative method.

Derivation of the Jacobian Matrix Elements

Model sensitivity or Jacobian matrix elements play a critical role in the inverse solution and efficient manipulation of these elements is essential for a robust inverse solution. Following the derivation of [1], we derive the model sensitivity elements by first considering a single predicted data point, d_j , consisting of either the electric or magnetic fields. Consider, for example, the x-

dg

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

**Portions of this document may be illegible
in electronic image products. Images are
produced from the best available original
document.**

component of the magnetic field at location j , which can be represented as

$$H_{xj} = h_{g_{j(x)}}^t E. \quad (2)$$

In this expression E is an electric field vector arising from a 3D earth model and has dimension of $NT \times 1$, where NT represents the number of electric field unknowns that are determined from the finite difference forward solution for a given source and discrete frequency [6]. The vector $h_{g_{j(x)}}^t$ is an interpolator vector for the x -component of the magnetic field at the j th measurement point and is of dimension $1 \times NT$ (t here denotes the transpose operator). This vector will interpolate the sampled fields on the forward modeling grid to the measurement point and numerically includes a curl operator that is applied to the electric field. With this definition an element of the Jacobian matrix is written for the x -component of magnetic field as

$$\partial H_{xj} / \partial m_k = h_{g_{j(x)}}^t \partial E / \partial m_k. \quad (3)$$

From the forward problem, [6], the electric fields are determined from the linear system,

$$KE = S, \quad (4)$$

where K is the sparse finite-difference stiffness matrix with 13 non-zero entries per row and depends linearly on the electrical parameters we desire to estimate. Because the forward problem can specify boundary conditions and sources that depend on the model parameters, the source vector, S , can also depend on the model parameters. Thus differentiating equation (4) with respect to m_k yields,

$$\partial E / \partial m_k = K^{-1} (\partial S / \partial m_k - \partial K / \partial m_k E). \quad (5)$$

The derivatives $\partial S / \partial m_k$ and $\partial K / \partial m_k$ in equation (5) are rapid to compute analytically; the interested reader is referred to [1] for the specific details. Finally, an element of the Jacobian matrix for the x -component of magnetic field can be written as

$$\partial H_{xj} / \partial m_k = h_{g_{j(x)}}^t K^{-1} (\partial S / \partial m_k - \partial K / \partial m_k E). \quad (6)$$

Similar expressions can be derived for the other electric and magnetic field components.

Inverse Solution via Conjugate Gradients

Because of the size of the 3D inverse problem, direct methods cannot be used to repeatedly invert a full least-squares system matrix involving model sensitivity elements. However, gradient methods are feasible. The steepest-descent method is the most widely known of the

gradient methods. Unfortunately the method usually converges very slowly near the minimum of the objective functional. A better approach can be the conjugate-gradient (CG) method. Linear and non-linear CG schemes are known and we have investigated both types for the 3D EM inverse problem.

Linear Conjugate Gradients

Minimization of equation (1) with respect to m , given a background model $m^{(0)}$ needed to linearize the problem for iteration (i), yields the model update,

$$m = [(DA^{(0)})^t (DA^{(0)}) + \lambda(W)^t(W)]^{-1} (DA^{(0)})^t (D\delta d^{(0)}) \quad (7)$$

with

$$\delta d^{(0)} = (d^{obs} - d^{(0)} + A^{(0)} m^{(0)}), \quad (8)$$

where the weighting matrix, D , is diagonal and consists of the reciprocal of the data standard deviations. The Jacobian matrix is denoted by $A^{(0)}$ in the above expressions and its individual elements are given from the derivation above. The fact that the Jacobian matrix and the predicted data, $d^{(0)}$, depend on the iteration count implies that they must be re-computed after each model step.

Linear conjugate gradients can be applied directly to linear systems that are symmetric positive definite, including the least squares normal equations given by equation (7). In a linear conjugate gradient method, all one needs is one matrix-vector multiplication per relaxation step. However, because the matrix given by this operation is $[(DA^{(0)})^t (DA^{(0)}) + \lambda(W)^t(W)]$, there are several other matrix-vector multiplications to be considered. First, the matrix product of $(DA^{(0)})^t$ with $DA^{(0)}$ requires two matrix-vector multiplications. In addition, the regularization-matrix product with its transpose requires two more matrix-vector multiplications. Since the latter matrix-vector multiplications are easy to implement and compute, no further elaboration will be given to them here.

For the Jacobian matrix-vector multiplications, $DA^{(0)}$ and $(DA^{(0)})^t$, we have

$$y = DA^{(0)} u \quad (9)$$

and

$$z = (DA^{(0)})^t y, \quad (10)$$

where u is an arbitrary real vector, known as a CG search direction vector. Because the data weighting and Jacobian matrices are real (we treat real and imaginary components of the data separately), the vector y is real with dimension $2N$, where N is the number of complex data points used in the inversion. The vector z is real since the model parameters are assumed to be real valued. [1] give compact and

computationally efficient forms for the two matrix-vector multiplications for dipolar source fields, which are also used to treat the matrix-vector multiplications given in equations (7) and (8), i.e. $A^{(0)}m^{(0)}$ and $(DA^{(0)})^t (D\delta d^{(0)})$, which are needed to initialize the CG solver at each iteration of the inversion.

In addition to the forward solutions necessary to determine E for each source and frequency, the matrix-vector multiplications in equations (9) and (10) require solving a series of forward problems corresponding to the total number of unique data measurements locations, where

$$v_j^t = g_j^t K^{-1}, \quad (11)$$

or since $K^t = K [1]$,

$$K v_j = g_j. \quad (12)$$

A unique measurement location comprises the interpolator vector g_j , which is based on the measurement of a specific field component or data type made at the site, independent of the source. Thus the total number of forward solutions needed for each model update is given by $N_s + N_r$, where N_s and N_r are the total number of sources and unique receiver positions used in the inversion at a given frequency. Note that multiple frequency data will require additional forward solutions for both the source and unique receiver positions.

To launch the inversion using the linear CG approach, we assume an initial background model and compute the predicted data for all the different sources. At the first iteration the matrix-vector multiplications needed to solve the linear least-squares system of equations are computed and equation (7) is solved using the linear CG method. We proceed to the next iteration if the data error (sum of squared errors) is above the estimated noise of the data set. If this is true the model is linearized again about the new model m , new predicted data and electric fields are computed from the updated background model, and the new model update is determined once the tradeoff parameter is specified.

A 3D inversion scheme using linear conjugate gradients has been successfully applied to crosswell 20 kHz electromagnetic data collected at the Richmond Field Station near Berkeley, California [2]. By comparing images of the data collected before injection of 50 000 gallons of salt water, a 3D image of the plume has been developed, which shows the location of zones of maximum permeability surrounding the injection well through which the salt water has migrated (Figure 1). A resolution analysis has determined that the location of the plume is fairly accurate. As this example clearly

demonstrates, diffusive EM fields possesses the ability to map permeability changes in the subsurface.

Nonlinear Conjugate Gradients

We have found that the linear CG scheme discussed above is optimal when many sources are employed as in a crosswell survey. For natural source fields (magnetotelluric applications) there is, however, a better approach based on non-linear conjugate gradients.

To use non-linear conjugate gradients for solving $\nabla\phi = 0$ demands that we make two subsidiary calculations at each iteration of the procedure. First we calculate the gradient of the objective functional in equation (1)

$$\nabla\phi = \nabla\phi_d + \lambda \nabla\phi_m, \quad (13)$$

where ϕ_d and ϕ_m are functionals that relate to the data misfit and the model constraint. Second we minimize ϕ along a specified ray, that is, find the value of α that minimizes the expressions $\phi(m + \alpha u)$ for specified model parameters m and conjugate search direction u .

Consider first computing a specific element of $\nabla\phi_d$ and let us define the difference between observed and predicted data as

$$\Delta d_j = \{(d_j^{\text{obs}} - d_j)/\varepsilon_j\}, \quad (14)$$

thus

$$\partial\phi_d / \partial m_k = -2 \sum_{j=1}^N \Delta d_j \partial d_j / \partial m_k. \quad (15)$$

Therefore, in general, we can then write from equation (6)

$$\partial\phi_d / \partial m_k = -2 \sum_{j=1}^N \Delta d_j g_j^t \cdot K^{-1} (\partial S_1 / \partial m_k - \partial K / \partial m_k E) \quad (16)$$

where the interpolator vector g_j^t is based on the data type. Evaluation of $\nabla\phi_m$ leads to

$$\nabla\phi_m = 2W^t Wm. \quad (17)$$

It is now possible to show that the number of forward solves needed to evaluate the gradient in equation (16) is only two for each source and frequency. One solve is needed to obtain the electric fields E of a specified source and another solve is needed to compute the fields arising from the following source

$$g^t = -2 \sum_{j=1}^N \Delta d_j g_j^t. \quad (18)$$

Fields for this source distribution are obtained following equations (11) and (12).

Extension of equation (16) to include multiple sources is easy, where the gradients for a full set of data are just the sum of gradients for each independent source position and frequency. Thus the total number of forward modeling applications necessary to evaluate (16) is twice the number of sources at each discrete frequency.

Because a line search is needed to minimize $\phi(\mathbf{m} + \alpha \mathbf{u})$, this will require additional forward solutions. [7] provides a number of strategies to carry out the line search, where the number of forward modeling applications varies. A reliable approach to conduct the line search is to use both functional and derivative information at two points, fit a cubic through these points and step immediately to the minimum. The number of forward modeling applications required for this type of line search is four per frequency. However, half of the required information is already known due to the previous model update. Thus only two additional applications of the forward modeling code at each source and frequency are required.

Listed below is an outline of the nonlinear CG method based on the algorithm [8] that will be used in this analysis:

- (1) choose $\mathbf{m}_{(1)}$ and select $\mathbf{u}_{(1)} = -\nabla\phi(\mathbf{m}_{(1)})$
- (2) find $\alpha(i)$ that minimizes $\phi(\mathbf{m}_{(i)} + \alpha(i) \mathbf{u}_{(i)})$
- (3) set $\mathbf{m}_{(i+1)} = \mathbf{m}_{(i)} + \alpha(i) \mathbf{u}_{(i)}$ and $\mathbf{r}_{(i+1)} = -\nabla\phi(\mathbf{m}_{(i+1)})$
- (4) $\beta_{(i+1)} = \max\{(\mathbf{r}_{(i+1)}^T \mathbf{r}_{(i+1)} - \mathbf{r}_{(i+1)}^T \mathbf{r}_{(i)}) / \mathbf{r}_{(i)}^T \mathbf{r}_{(i)}, 0\}$
- (5) $\mathbf{u}_{(i+1)} = \mathbf{r}_{(i+1)} + \beta_{(i+1)} \mathbf{u}_{(i)}$
- (6) Stop when $|\mathbf{r}_{(i+1)}|$ is sufficiently small, otherwise go to step (2).

The convergence of this scheme can also be accelerated by including preconditioning. The interested reader is referred to [9] for the details.

ACKNOWLEDGMENTS

This work was performed at Sandia National Laboratories, which is operated for the United States Department of Energy (DOE). Funding for the work was provided by DOE's office of Basic Energy Sciences, Division of Engineering and Geoscience, under contract DE-AC04-94AL85000.

REFERENCES

- [1] Newman G. A., and Alumbaugh D. L., 1997, Three-dimensional massively parallel electromagnetic inversion - I. Theory: *Geophysical Journal International*, **128**, 345-354.

Figure 1. Difference image between pre- and post-injection of 50,000 gallons of 1 Ω .m saltwater near 30 m depth. Two views of the same reconstruction are included such that horizontal slices, showing how the conductivity has changed, can be shown at nine different depths, with the back and left-side panels showing how the change varies continuously with depth. The four observation wells are shown surrounding the injector well shown at the center. Vertical magnetic field receivers were deployed in the observation well from the surface to 60 m depth at 5 m intervals. An 18.5 kHz vertical magnetic dipole was used as the transmitter and was placed in the injector well. The transmitter was moved in 2.5 m increments from 5 to 60 m depth in the crosswell experiment. Refer to [2] for additional details. Notice that the salt-water plume is clearly observed in the difference image near 30 m depth.

[2] Alumbaugh D. L., and Newman G. A., 1997, Three-dimensional electromagnetic inversion - II. Analysis of a crosswell electromagnetic experiment: *Geophysical Journal International*, **128**, 355-363.

[3] Newman G. A., and Alumbaugh D. L., 1997, 3D Electromagnetic Modeling and Inversion on Masively Parallel Computers: In Press, In *3D Electromagnetics, Investigations in Geophysics Series*, Society of Exploration Geophysicists, Tulsa OK.

[4] Alumbaugh, D. L., and Newman, G. A., 1997b, 3-D electromagnetic inversion for environmental site characterization, In R. S. Bell, editor, *Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems*, SAGEEP '97, March 23 to 26, 1997, Reno, NV., Environment Engineering Geophysical Society, Wheat Ridge, CO, 355-364.

[5] Wang T., Ortaglio M., Tripp A., and Hohmann G., 1994, Inversion of diffusive transient electromagnetic data by a conjugate-gradient method: *Radio Science*, **29**, 1143-1156.

[6] Alumbaugh D. L., Newman G. A., Prevost L., and Shadid, J. N., 1996, Three-dimensional wideband electromagnetic modeling on massively parallel computers: *Radio Science*, **31**, 1-23.

[7] Acton F. S., 1970, *Numerical Methods that Work*, Harper and Row, New York.

[8] Polyak E., and Ribiere, 1969, Note sur la convergence des méthodes conjuguées: *Rev. Fr. Inr. Rech. Oper.*, **16**, 35-43.

[9] Shewchuk, J. R., An introduction to the conjugate gradient method without agonizing pain: School of Computer Sciences, Carnegie Mellon University, Pittsburg, PA. Electronic copy of the article is available by anonymous FTP to WARP.CS.CMU.EDU (IP address 128.2.209.103) under the file name quake-papers/painless-conjugate-gradient-ps.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.