QN Inversion of Large-scale MT Data

A. D. Avdeeva[1], **and D. B. Avdeev**[1,2]
[1]Dublin Institute for Advanced Studies, Ireland
[2]IZMIRAN, Russian Academy of Sciences, Russia

Abstract—A limited memory quasi-Newton (QN) method with simple bounds is applied to a 1-D magnetotelluric (MT) problem. The method is used to invert a realistic synthetic MT impedance dataset, calculated for a layered earth model. An adjoint method is employed to calculate the gradients and to speed up the inverse problem solution. In addition, it is shown that regularization stabilizes the QN inversion result. We demonstrate that only a few correction pairs are enough to produce reasonable results. Comparison with inversion based on known L-BFGS-B optimization algorithm shows similar convergence rates. The study presented is a first step towards the solution of large-scale electromagnetic problems with a full 3-D conductivity structure of the Earth.

1. Introduction

Quasi-Newton (QN) methods have become a very popular tool for the numerical solution of electromagnetic (EM) inverse problems (see [8, 7]). The reasoning behind it is that the method requires calculation of gradients only, while at the same time avoiding calculations of second-derivative terms. However, even with the gradients only, the QN methods may require excessively large computational time if the gradients are calculated straightforwardly. An effective way to calculate the gradients is delivered by an adjoint method (see [11, 4]). Also, for large-scale inverse problems the limited memory QN methods have to be applied, since their requirements for storage are not so excessive as for other QN methods. In this paper, as a first step to solving the 3-D EM case, we have applied a limited memory QN method for constrained optimization to solve 1-D magnetotelluric (MT) problems. This optimization method is an extension of previous work [9]. As distinct from this earlier work we implement the Wolfe conditions to terminate the line search procedure, as was recommended in [3]. First we present a simple review of the limited memory QN method for inversion of 1-D MT data. Then, we demonstrate the efficiency of our inversion on a synthetic, but realistic numerical example, along with a comparison with inversion based on the L-BFGS-B optimization method introduced by [3]. The results presented are encouraging and suggest that the method has the potential to handle the more geophysically realistic 3-D inverse problem.

2. 1-D MT Inversion

In the frame of the magnetotelluric method both the electric and magnetic fields are recorded. These fields are then processed to calculate the observed impedance dataset. This dataset is finally inverted to derive a distribution of electrical conductivity in the earth.

Thus, for 1-D MT inversion a layered earth model is considered and conductivities of the layers are sought. This problem is usually solved by minimization of the following objective function

$$\varphi(\mathbf{m}, \lambda) = \varphi_d(\mathbf{m}) + \lambda\varphi_s(\mathbf{m}) \underset{\mathbf{m}}{\longrightarrow} \min, \tag{1}$$

where $\varphi_d(\mathbf{m}) = \frac{1}{2}\sum_{j=1}^{M} \alpha_j \left|Z_j - d_j\right|^2$ is the data misfit. Here $\mathbf{m} = (m_1, ..., m_N)^T$ is the vector consisting of the electrical conductivities of the layers; superscript T means transpose; N is the number of the layers; $Z_j(\mathbf{m})$ and d_j are the complex-valued modeled and observed impedances at the j-th period ($j = 1, ..., M$) respectively; $\alpha_j = \frac{2}{M}\epsilon_j^{-2}\left|d_j\right|^{-2}$ are the positive weights; ϵ_j is the relative error of the impedance $Z_j(\mathbf{m})$ and λ is a Lagrange multiplier. Our choice of λ is a simple variant of an algorithm presented in [6]. As prescribed by the regularization theory (see [12]), the function (1) has a regularized part (a stabilizer) $\varphi_s(\mathbf{m})$. This stabilizer can be chosen by many ways, and moreover, the correct choice of $\varphi_s(\mathbf{m})$ is crucial for a reliable inversion. However, this aspect of the problem is out of the scope of this paper. We consider the following stabilizer (see [5]) $\varphi_s(\mathbf{m}) = \sum_{i=2}^{N}\left(\frac{m_i}{m_i^0} - \frac{m_{i-1}}{m_{i-1}^0}\right)^2$, where $\mathbf{m}^0 = (m_1^0, ..., m_N^0)^T$ is an initial guess model. It is of importance that, as the conductivities $m_i\,(i = 1, \ldots, N)$ must be non-negative and realistic, the optimization problem (1) is subject to bounds

$$l_i \leq m_i \leq u_i, \tag{2}$$

where l_i and u_i are the lower and upper bounds, respectively and $l_i \geq 0\,(i = 1, \ldots, N)$.

A limited memory quasi-Newton method. We notice that the problem posed in (1)– (2) is a typically constrained optimization problem with simple bounds. To solve this problem Newton-type iterative methods are commonly applied. However, most of these methods are not applicable to large-scale optimization problems because the storage and computational requirements become excessive. To overcome this, a limited memory quasi-Newton method has been developing (see [10] for a good introduction). Let us now describe our implementation of such a technique.

At each iteration step k the search direction, vector $\mathbf{p}^{(k)}$, is calculated as $\mathbf{p}^{(k)} = -\mathbf{G}^{(k)}\mathbf{g}^{(k)}$, where the symmetric matrix $\mathbf{G}^{(k)}$ is an approximation to the inverse Hessian matrix and $\mathbf{g}^{(k)}$ is the gradient $\mathbf{g} = (\frac{\partial\varphi}{\partial m_1}, \ldots, \frac{\partial\varphi}{\partial m_N})^T$ calculated at $\mathbf{m} = \mathbf{m}^{(k)}$. An explicit expression for the matrix $\mathbf{G}^{(k)}$ is given in [10, p. 225, formula (9.5)]. It is important that the matrix $\mathbf{G}^{(k)}$ is stored implicitly using n_{cp} correction pairs $\{\mathbf{s}^{(n)}, \mathbf{y}^{(n)}\}$ $(n = k - n_{cp}, \ldots, k - 1)$ previously computed as $\mathbf{s}^{(n)} = \mathbf{m}^{(n+1)} - \mathbf{m}^{(n)}$, $\mathbf{y}^{(n)} = \mathbf{g}^{(n+1)} - \mathbf{g}^{(n)}$. The main idea behind this approach is to use information from only the most recent iterations and the information from earlier iterations is discarded in the interests of saving storage. In [10, p. 225] it is advocated that n_{cp} between 3 and 20 may produce satisfactory results.

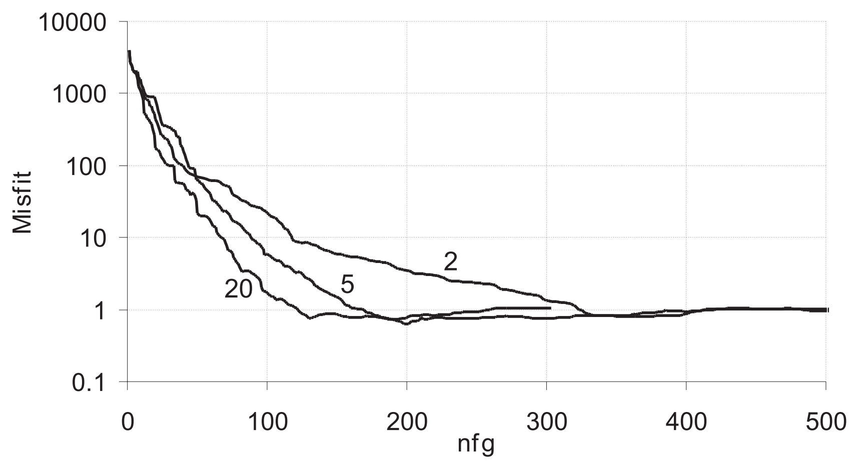

Figure 1: Comparison of LMQNB inversions for 2, 5 and 20 correction pairs.

The next iterate $\mathbf{m}^{(k+1)}$ is then found as $\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha^{(k)}\mathbf{p}^{(k)}$, where the step length $\alpha^{(k)}$ is computed by an inexact line search procedure. This procedure finds a step length that delivers an adequate decrease in the objective function φ along the search direction $\mathbf{p}^{(k)}$. Let us demonstrate how in our implementation we provide the positive definiteness of the matrix $\mathbf{G}^{(k)}$, required to guarantee the descent direction $\mathbf{p}^{(k)}$. When the vectors $\mathbf{s}^{(k-1)}$ and $\mathbf{y}^{(k-1)}$ satisfy the curvature condition $\mathbf{s}^{(k-1)T}\mathbf{y}^{(k-1)} > 0$, it can be shown that the matrix $\mathbf{G}^{(k)}$ is positive definite. This condition is guaranteed to hold if we use the Wolfe conditions (see [10]) to terminate the line search. But the Wolfe conditions may not be reached inside the feasible region defined within the bounds set by equation (2). In this case, we modify $\mathbf{s}^{(k-1)}$ as prescribed in [9, p. 1513] to guarantee that matrix $\mathbf{G}^{(k)}$ is positive definite.

Alternative ways to deal with such boundary constrained QN optimization can be found in [3].

Speeding up the solution. As presented above, at each iteration step k the inverse problem solution requires calculating the gradient $\mathbf{g}^{(k)}$. However for large-scale problems (when N is large) a straightforward calculation may be prohibitive in terms of computational time. One can significantly speed up the calculation by using an adjoint method (see [11, 4]). We have applied such a method to the 1-D MT case. The derivatives are numerically calculated as $\frac{\partial\varphi_d}{\partial\sigma_i} = -Re\left(\sum_{j=1}^{M}\alpha_j\Gamma_{ij}\left(\overline{Z_j} - \overline{d_j}\right)Z_j^2\right)$, where $i = 1, \ldots, N$, the sign Re means the real part of its complex argument and all the coefficients Γ_{ij} are found by solving a single adjoint forward problem. Thus, our calculation of the gradient requires the solution of a single forward and adjoint problem. This approach may be extended with some effort to the 3-D case. It is also noteworthy that for the 1-D MT case the gradient can also be calculated using the chain-rule (see [5]).

We have therefore implemented the limited memory QN method with simple bounds (hereinafter, referred as LMQNB), which is described above. It should be noted here that our implementation differs from that of [9], in that the LMQNB uses the Wolfe conditions to terminate the line search.

3. Model Examples

Let us study on a synthetic 1-D MT example the convergence rate of the LMQNB inversion for a various number n_{cp} of correction pairs. A 7-layered earth model (see Fig. 3) was compiled from the models (see [1]) derived from a seafloor MT and a global GDS long-period dataset collected in the North Pacific Ocean. To complicate the inversion process we subdivided the three upper layers in this model (to depth of 394 km) into 197 equally thick sublayers ($N = 197$). For this 201-layered model we inverted the impedance $d_j = Z_j(\mathbf{m})$ calculated at $M = 30$ periods from 10 s to 10800 s.

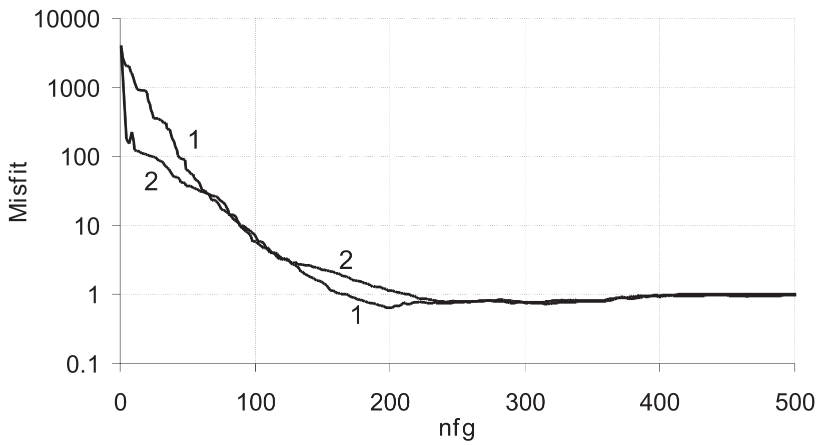

Figure 2: Comparison of inversions based on the LMQNB (1) and L-BFGS-B (2) optimisations with $n_{cp} = 5$ and $\lambda = 320$.

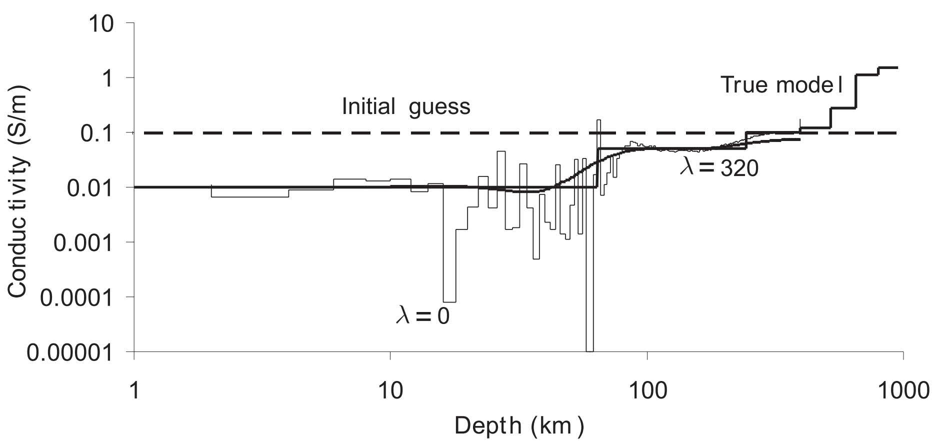

Figure 3: The conductivity models obtained from inversion based on the LMQNB optimization with $n_{cp} = 5$ for $\lambda = 0$ and $\lambda = 320$. The true model is shown by the solid line and the initial guess by the dashed line.

In addition, we added 0.5% random noise to the impedance data. The relative error ϵ_j (see α_j on page 1) of the impedance was taken as 0.01. A 10 ohm-m uniform half-space was used as an initial guess $\mathbf{m} = \mathbf{m}^0$. The convergence rate curves are shown in Fig. 1 as a function of the total number (n_{fg}) of function φ and gradient $\mathbf{g}$ evaluations for $\lambda = 320$. It is surprising that so small a number of pairs ($n_{cp} = 5$) can be chosen sufficiently to get a relatively reasonable result.

Let us again consider the 201-layered model and the dataset, described in the previous example. In our next example (see Fig. 2) we present a comparison of two solutions of a 1-D MT inverse problem. The first solution is based on the optimization method offered in this paper and the second one exploits the L-BFGS-B optimization code from [3]. The comparison is presented for $n_{cp} = 5$ correction pairs and for $\lambda = 0$ and $\lambda = 320$. Both solutions converge in a similar way and also produce similar models. The inversion results are presented in Fig. 3.

4. Conclusion

This paper described a limited memory QN method applied to solve a 1-D MT inverse problem. The method is also valid for large-scale problems, and may be equally applied for 1-D, 2-D, or 3-D MT cases. In the numerical examples presented we have demonstrated that a few correction pairs are enough to obtain reasonable inversion results. To speed-up the inversion the adjoint method has been applied to calculate the gradients. The non-trivial problem of such a calculation of gradients in the 3-D MT case is presented in a companion paper [2]. Another finding of our numerical experiments is that the LMQNB solution converges similarly to the solution based on the L-BFGS-B method introduced in [3].

Acknowledgment

This work was carried out in the Dublin Institute for Advanced Studies and was supported by the Cosmogrid project, which is funded by the Programme for Research in Third Level Institutions under the National Development Plan and with assistance from the European Regional Development Fund. Thanks to Brian O'Reilly for corrections to English. We acknowledge helpful comments of an anonymous reviewer. AA also acknowledges the constructive support of her supervisors Alan Jones and Colin Brown.

REFERENCES

1. Avdeev, D. B., H. Utada, A. V. Kuvshinov, and T. Koyama, "Three-dimensional electromagnetic modeling of the Hawaiian swell," *AGU Fall Meeting, 13-17 December, San Francisco, CA*, F630, 2004.
2. Avdeev, D. B. and A. D. Avdeeva, "A rigorous 3-D MT inversion," this volume.
3. Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Scientific Computing*, Vol. 16, 1190–1208, 1995.
4. Chen, J., D. W. Oldenburg, and E. Haber, "Reciprocity in electromagnetics: application to modeling marine magnetometric resistivity data," *Phys. Earth Planet. Inter.*, Vol. 150, 45–61, 2005.
5. Constable, S. C., R. L. Parker, and C. G. Constable, "Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data," *Geophysics*, Vol. 52, 289–300, 1987.
6. Haber, E. and D. W. Oldenburg, "Joint inversion: a structural approach," *Inverse Problems*, Vol. 13, 63–77, 1997.
7. Haber, E., "Quasi-Newton methods for large-scale electromagnetic inverse problems," *Inverse Problems*, Vol. 21, 305–323, 2005.
8. Newman, G. A. and P. T. Boggs, "Solution accelerators for large-scale three-dimensional electromagnetic inverse problem," *Inverse Problems*, Vol. 20, s151–s170, 2004.
9. Ni, Q. and Y. Yuan, "A subspace limited memory quasi-Newton algorithm for large-scale nonlinear bound constrained optimization," *Mathematics of Computation*, 66/220, 1509–1520, 1997.
10. Nocedal, J. and S. J. Wright, *Numerical Optimization*, Springer, 1999.
11. Rodi, W. and R. L. Mackie, "Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion," *Geophysics*, Vol. 66, 174–187, 2001.
12. Tikhonov, A. N. and V. Y. Arsenin, *Solutions of Ill-posed Problems*, Wiley, New York, 1977.