Hybrid OCCAM-Conjugate Gradients Inversion Algorithms with Applications to Marine CSEM data

Gary D. Egbert[1], Naser Meqbel[2] and Paulo Werdt[2]
[1]CEOAS, Oregon State University, Corvallis OR, USA
[2]Observatório Nacional, Rio de Janiero, Brazil

SUMMARY

Egbert (2012) developed a hybrid multi-transmitter (mTX) OCCAM-Conjugate Gradient (CG) algorithm and illustrated basic ideas on a simple 2D MT inverse problem. Here we present results of application of these ideas to a more computationally challenging problem, 3D marine controlled-source EM (mCSEM). There are two main aspects of our approach. First, we save results of calculations required for an iterative solution of the data-space Gauss-Newton (GN) normal equations and use these to construct a low dimensional approximation of the full Jacobian. Once computed, this allows rapid computation of trial inverse solutions for a range of regularization parameters, so Occam-type schemes become practical even for very large problems. The hybrid is based on the Golub-Kahan bidiagonalization of the Jacobian matrix. Second, every transmitter (Tx; i.e., different frequency, or location) requires solution of a separate adjoint problem, associated with the gradient of the data misfit for that separate Tx dataset. By saving results of these distinct calculations (instead of summing them, to compute the gradient of the total data misfit) we can form a more complete approximation to the Jacobian. Furthermore, with suitable modifications to the iterative CG algorithm (Egbert, 2012), more rapid (fewer adjoint and forward solves) and stable solution of the G-N equations can be achieved. We demonstrate the effectiveness of our methods using synthetic datasets based on a realistic 3D resistivity model constructed for the Campos Basin on the Brazilian margin. We also demonstrate how the mTX algorithm can be useful for joint inversion of multiple EM data types, specifically for combining MT and mCSEM data. The hybrid scheme allows for efficient exploration of relative weights for the different measured data types, using the approximate Jacobian. The approximated Jacobian can also be used for linearized uncertainty and resolution analysis of the solutions obtained.

Keywords: EM Geophysics, Inversion Methods, mCSEM

INTRODUCTION

In this work we discuss some new approaches to minimization of the quadratic penalty functional

$$\Phi(\mathbf{m},\mathbf{d}) = \left(\mathbf{d}-\boldsymbol{f}(\mathbf{m})\right)^{\mathrm{T}}(\mathbf{d}-\boldsymbol{f}(\mathbf{m})) + \lambda\mathbf{m}^{T}\mathbf{m} \qquad (1)$$

where $\mathbf{m}$ is the model parameter $\mathbf{d}$ the data vector, and $f(\mathbf{m})$ the forward mapping. Note that as our focus here is on minimization algorithms, we ignore data and model "covariances" in the data misfit and model regularization terms in (1). In practice this simple generic form may be obtained by suitable transformation of model and data space vectors. Gradient-based linearized search schemes such as NLCG or LBFGS have proven to be an effective and efficient way to minimize the quadratic penalty functional. However, there are potential advantages to a Gauss-Newton (GN) approach such as OCCAM (Constable et al., 1987), as we discuss. In the simplest approach a GN scheme requires computation of the full Jacobian (sensitivity) matrix $\mathbf{J}$, followed by computation of a cross-product matrix such as $\mathbf{J}^T\mathbf{J}$, and then solution of a large system of linear normal equations. In a data-space variant of the OCCAM scheme (Siripunvaraporn et al., 2005) the normal equations take the form

$$(\mathbf{J}\,\mathbf{J}^{T} + \lambda\mathbf{I})\mathbf{b} = \hat{\mathbf{d}} = \mathbf{d} - \boldsymbol{f}(\mathbf{m}_n) + \mathbf{J}\mathbf{m}_n \qquad (2)$$

with the updated model parameter at iteration $n+1$ computed as $\mathbf{m}_{n+1} = \mathbf{J}^{T}\mathbf{b}$. In the OCCAM scheme the parameter λ is varied, initially to minimize data misfit, then to find the smallest model consistent with the desired misfit tolerance. OCCAM thus automatically optimizes the regularization (tradeoff) parameter.

While OCCAM has been widely used in 1D and 2D inverse problems, application in 3D is more challenging, due to the need to compute the full Jacobian, to form the large cross-product matrices, and to solve a large system of normal equations. These equations can also be solved using conjugate gradients (CG) without first computing $\mathbf{J}$. Each step requires multiplication of arbitrary model vectors (e.g.,m_n) by $\mathbf{J}$, and data vectors (e.g., $\mathbf{b}$) by $\mathbf{J}^T$. These operations involve one forward/adjoint solution, respectively, of the 3D EM PDE. While this CG approach makes a GN approach more practical, it is not immediately obvious how to implement OCCAM – seemingly, the CG iterations must be run repeatedly for different values of the regularization parameter λ.

Here we make two points: First, we can save results of calculations required for an iterative CG solution of the data-space Gauss-Newton (GN) normal equations and use these to construct a low dimensional approximation of the full Jacobian. Once computed, this allows rapid computation of trial inverse solutions for a range of regularization parameters, so Occam-type schemes become practical even for very large problems. Second, every transmitter (Tx; i.e., different frequency, or location) requires solution of a separate adjoint problem, associated with the gradient of the data misfit for that separate Tx dataset. By saving results of these distinct calculations (instead of summing them, to compute the gradient of the total data misfit) we can form a more complete approximation to the Jacobian, and solve the G-N equations with fewer adjoint and forward computations.

METHODS

The Basic Hybrid Algorithm (BDORTH): After K steps Lanczos bi-diagonalization of the Jacobian matrix produces the decomposition

$$\mathbf{J}^T\mathbf{U}_K = \mathbf{V}_K\mathbf{B}_K \qquad (3)$$

Here $\mathbf{U}_K = [\boldsymbol{u}_1, \dots, \mathbf{u}_K]$ and $\mathbf{V}_K = [\,\mathbf{v}_1, \dots\ , \mathbf{v}_K]$ are orthogonal matrices whose columns are, respectively, data and model space vectors, and $\mathbf{B}_K$ is bidiagonal. Each step in the algorithm requires multiplication by $\mathbf{J}^T$ and $\mathbf{J}$. Saving these matrices (i.e., K data and model space vectors) we can project the Jacobian and data vectors, into a K dimensional subspace and solve the equivalent of Eq. (2)

$$(\mathbf{U}_K^T\mathbf{J}\,\mathbf{J}^T\mathbf{U}_K + \lambda\mathbf{I})\mathbf{b}_\lambda = (\mathbf{B}_K\mathbf{B}_K^T \quad + \lambda\mathbf{I}) = \mathbf{U}_K^T\,\hat{\mathbf{d}} \qquad (4)$$

The approximate model space solution is then computed as $\mathbf{m}_\lambda = \mathbf{V}_K\mathbf{b}_\lambda$. This can be computed for any λ making an approximate OCCAM scheme possible. We refer to this algorithm here as "BDORTH". This is an example of what has often been referred to as a hybrid algorithm for the iterative solution of linear equations.

Multi-transmitter Extension (BDMTX): In this variant we save computations (both data and model space vectors $\mathbf{U}_K$ and $\mathbf{V}_K$) separately for each transmitter (Tx; different source dipole source location or different frequency) separately. Because each Tx requires independent forward and adjoint calculations (the costliest step in the inversion) the additional computational burden is minimal (although memory requirements are generally increased). Egbert (2012) describes a modified Lanczos scheme that works well. We summarize briefly here:

We have $j = 1, \dots, J$ transmitters. At each step $k = 1, \dots K$ in the iterative algorithm we compute and save J data space vectors $\mathbf{u}_{jk}$, and model space vectors $\mathbf{v}_{Jk}$ one each for each Tx, and for each iterative step. We collect these as

$$\mathbf{U}_k = [\mathbf{u}_{11} \dots \mathbf{u}_{1k}\,|\dots\ |\ \mathbf{u}_{Jk} \dots \mathbf{u}_{Jk}] = [\mathbf{W}_1 \dots \mathbf{W}_J] \qquad (5)$$

$$\mathbf{V}_k = [\mathbf{J}_1^T\mathbf{u}_{11} \dots \mathbf{J}_1^T\mathbf{u}_{1k}\,|\dots\ |\ \mathbf{J}_J^T\mathbf{u}_{Jk} \dots \mathbf{J}_J^T\mathbf{u}_{Jk}] \qquad (6)$$

The first step ($k = 1$) is initialized as $\boldsymbol{u}_{j1} = \boldsymbol{d}_j/||\boldsymbol{d}_j||$ where $\mathbf{d}_j$ is the data vector for Tx j. To compute data space vectors for t step k we solve the projected normal space equations

$$(\mathbf{V}_k^T\mathbf{V}_k + \lambda\mathbf{I})\tilde{\mathbf{b}} = [\mathbf{d}_1^T\mathbf{U}_{1k} \dots \mathbf{d}_{N_{Tx}}^T\mathbf{U}_{N_{Tx}k}]^T$$

The coefficient matrix of the symmetric system is $Jk \times Jk$, so even with hundreds of Tx the computational cost is negligible. Then compute the "trial solution" $\mathbf{m}_k = \mathbf{V}_k\tilde{\mathbf{b}}$, multiply by Jacobian for eaxh Tx, $\mathbf{c}_{jk} = \mathbf{J}_j\boldsymbol{m}_k$ and finally compute $\mathbf{e}_{jk} = \mathbf{c}_{jk} - \mathbf{W}_j\ \mathbf{W}_j^T\boldsymbol{c}_{jk}$ and $\boldsymbol{u}_{jk+1} = \boldsymbol{e}_{jk}/||\boldsymbol{e}_{jk}||$. This is iterated until a sufficiently accurate solution to the (nprojected) normal equations is obtained. The scheme is quite similar to the original Lanczos decomposition, although a small linear system must be solved at each step to couple equations for all Tx.

RESULTS

As an illustration of these methods we consider a very simple toy marine CSEM synthetic dataset. The model, and transmitter and receiver configuration is shown in Figure 1. The model is very simple, a local resistive body in a layered background with flat bottom bathymetry (depth of 1000 m). There are 6 Tx locations and threwe frequencies, so the total number of transmitters is 18.

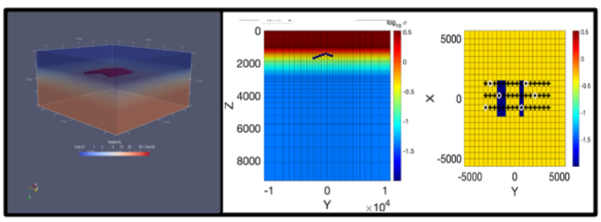

Figure 1: model and data configuration for tests.

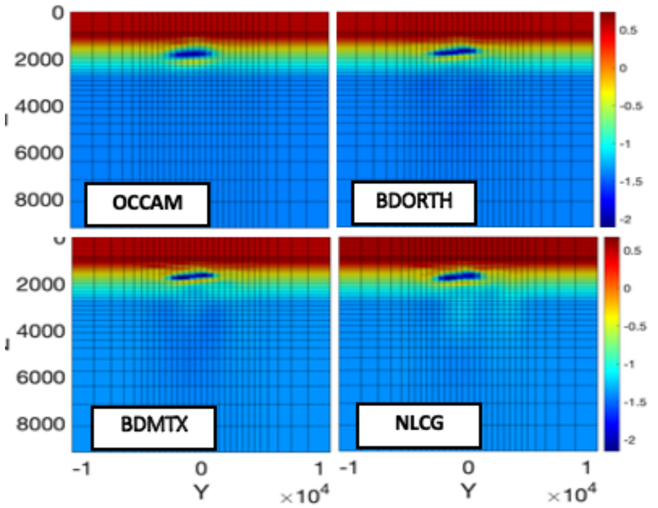

Figure 2: Inversion results for four algorithms: OCCAM, based on full Jacobian computation, BDORTH, the basic hybris scheme based on saving sensitivity computations from the Lanczos bidiagonalization, BDMTX, the multi-transmitter extension, and a the NLCG algorithm, as implemented in the ModEM code of Kelbert et al (2014; extended to CSEM inversion), unpublished). All algortihms converged to similar results for this simple problem. However the OCCAM scheme required 9270 adjoint/forward solves – making this impractical for more realistic problems. Other approaches required only 333 (NLCG), 354 (BDORTH) and 152 (BDMTX).

Results from the new algorithms proposed here, along with from OCCAM (full sensitivity calculation, and a more commonly used gradient search scheme (NLCG), are shown in Figure 2. Number of forward and adjoint solves required are summarized in the figure caption.

Both BDORTH and BDMTX solve the data space normal equations iteratively, generating an approximation to the Jacobian along the way. Convergence is significantly more rapid, and apparently smoother and more stable with the BDMTX approach, as illustrated in Figure 3.

Discussion

We have presented two hybrid schemes, which generate an approximate Jacobian through iterative solution of the data-space normal equations for a Gauss-Newton inversion algorithm. Both are about as efficient (in terms of number of forward and adjoint solves) as a gradient search method such as NLCG, and allow use of an OCCAM approach to choice of regularization parameter. The approximate Jacobian could also be used for linearized resolution analysis, something that is not possible with NLCG.

These hybrid schemes may also be useful for joint inversion of, for example, CSEM and MT data. With two different data types the appropriate balance between fitting each type often requires multiple runs with different relative weights. This exploration could be done quite efficiently using the computed approximate Jacobian, especially with the BDMTX scheme.

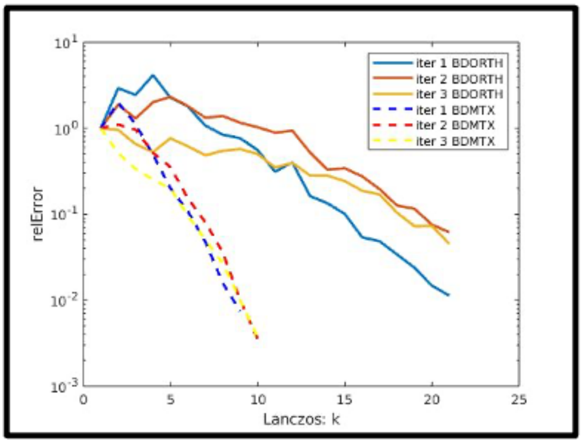

Figure 3: comparison of convergence of the inner loop (iterative solution to normal equations) for 3 outer loop steps. Dashed lines are for the multi-transmitter scheme, which reduces required number of iterations by a factor of roughly 2.5, and shows smoother and more stable convergence behaviour.

References

Constable, Steven C., Robert L. Parker, and Catherine G. Constable. "Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data." *Geophysics* 52, no. 3 (1987): 289-300.

Egbert, G.D., 2012. Hybrid conjugate gradient-Occam algorithms for inversion of multifrequency and multitransmitter EM data. *Geophysical Journal International*, *190*(1), pp.255-266.

Siripunvaraporn, Weerachai, Gary Egbert, Yongwimon Lenbury, and Makoto Uyeshima. "Three-dimensional magnetotelluric inversion: data-space method." *Physics of the Earth and planetary interiors* 150, no. 1-3 (2005): 3-14.