Using the Levenberg-Marquardt Method for the Solution of Inverse Transport Problems

Jeffrey A. Favorite* and David I. Ketcheson**

*X-4, MS F663, Los Alamos National Laboratory, Los Alamos, NM 87545 USA, fave@lanl.gov
** Dept. of Applied Mathematics, U. of Washington, Seattle, WA 98105 USA, ketch@amath.washington.edu

INTRODUCTION

The inverse transport problem of determining unknown interface locations in a multilayer radioactive source/shield system based on gamma-ray flux measurements was recently attacked using a novel solution method dubbed the Schwinger inverse method [1], which was also used to determine an unknown source composition [1], identify an unknown material in a radiation shield [2], and identify unknown mass densities in the source and/or shield [3]. The ability to solve for these types of unknowns is useful in characterizing many types of radioactive systems for environmental or homeland security applications.

In this paper a standard, robust optimization algorithm, the Levenberg-Marquardt (or just Marquardt) method [4], is applied to these types of problems. The Marquardt method computes a set of unknown parameters using gradients of an error function with respect to each of the unknown parameters. The error function is typically taken to be the χ^2 difference between a set of measured data and a set of model (calculated) data. In this paper, the data of interest are the unscattered leakages of a set of decay gamma-ray lines. The gradients are calculating using an adjoint-based method [5-7].

In this paper, the Marquardt method is applied to the problem of determining unknown interface locations in a source/shield system [1]. The two questions addressed are 1) How well are the gradients calculated using adjoint-based methods?, and 2) Given these adjoint-based gradients, how well does the Marquardt method work for these types of problems?

COMPUTING THE GRADIENTS

Consider a system that includes some volumetric source of neutral particles and some shield. Both the source and the shield may be multilayered but, for simplicity, only homogeneous layers are considered. The angular flux of particles is given by the Boltzmann transport equation,

$$\hat{\mathbf{\Omega}} \cdot \vec{\nabla} \psi + \Sigma_t \psi - S(\Sigma_s \psi) = q, \tag{1}$$

where S is the usual scattering source.

Suppose a set of D measurements is taken at a detector. These may be, for example, peaks in a gamma-ray spectrum from which the flux or leakage of gamma-ray lines is obtained. The quantities of interest are

$$M_d = \left\langle \Sigma_d \psi \right\rangle, \; d = 1, \ldots, D, \tag{2}$$

where the detector response function $\Sigma_d(r, E, \hat{\mathbf{\Omega}})$ is defined as zero outside the detector volume, angle, and energy region of interest to detector d.

Now suppose the system is perturbed in some way. The perturbed cross sections and inhomogeneous source are denoted with a prime; they are related to the unperturbed terms by $\Sigma' = \Sigma + \delta\Sigma$ and $q' = q + \delta q$. The angular flux in the perturbed system, ψ', satisfies

$$\hat{\mathbf{\Omega}} \cdot \vec{\nabla} \psi' + \Sigma_t' \psi' - S(\Sigma_s' \psi') = q' \tag{3}$$

and is related to the angular flux in the unperturbed system by $\psi' = \psi + \delta\psi$. The exact value of the perturbed quantity of interest for detector d is

$$M_d' = \left\langle \Sigma_d \psi' \right\rangle, \tag{4}$$

where it is assumed that the detector is not perturbed. The variation in the quantity of interest is

$$\delta M_d = M_d' - M_d = \left\langle \Sigma_d \delta\psi \right\rangle. \tag{5}$$

Ignoring second-order terms, Eq. (3) becomes

$$\hat{\mathbf{\Omega}} \cdot \vec{\nabla} \delta\psi + \Sigma_t \delta\psi - S(\Sigma_s \delta\psi) = \delta q - \delta\Sigma_t \psi + \delta\mathbb{S}(\psi), \tag{6}$$

where $\delta\mathbb{S}(\psi) \equiv S(\delta\Sigma_s \psi)$.

As in Refs. 5 and 6, we seek to remove the flux variation $\delta\psi$ from Eq. (5). To accomplish this, let the adjoint source be the response function for detector d, so that the unperturbed adjoint equation is

$$-\hat{\mathbf{\Omega}} \cdot \vec{\nabla} \psi_d^* + \Sigma_t \psi_d^* - S^*(\Sigma_s \psi_d^*) = \Sigma_d. \tag{7}$$

Using Eq. (7) in Eq. (5) and then using Eq. (6) yields

$$\delta M_d = \left\langle \psi_d^*, \delta q \right\rangle - \left\langle \psi_d^*, \delta\Sigma_t \psi \right\rangle + \left\langle \psi_d^*, \delta\mathbb{S}(\psi) \right\rangle. \tag{8}$$

Equation (8) will now be applied to determine the gradient of the set of measurements M_d of Eq. (2) with respect to a set of interface locations r_n, $n = 1,\ldots,N$, in a multilayer source/shield system. Using the same procedure as in Ref. 6, it can be shown that the total cross section variation due to variations δr_n in any of the interface locations is

$$\delta\Sigma_t = \sum_{n=1}^{N} \delta(r - r_n)\Delta\Sigma_{t,n}\delta r_n, \tag{9}$$

the total inhomogeneous source variation is

$$\delta q = \sum_{n=1}^{N} \delta(r - r_n)\Delta q_n \delta r_n, \tag{10}$$

and the total scattering source variation is

$$\delta S(\psi) = \sum_{n=1}^{N} \delta(r - r_n)\Delta S_n(\psi)\delta r_n. \tag{11}$$

In Eqs. (9)-(11), the ΔX_n quantities are the difference between the quantity in region n and region $n + 1$ and $\Delta S_n(\psi) \equiv S(\Delta\Sigma_{s,n}\psi)$.

Using Eqs. (9)-(11) in Eq. (8), the total variation in the quantity of interest M_d due to interface-location variations is

$$\begin{aligned}\delta M_d = \sum_{n=1}^{N} \Big[&\left\langle \psi_d^*, \delta(r - r_n)\Delta q_n \right\rangle - \left\langle \psi_d^*, \delta(r - r_n)\Delta\Sigma_{t,n}\psi \right\rangle \\ &+ \left\langle \psi_d^*, \delta(r - r_n)\Delta S_n(\psi) \right\rangle \Big] \delta r_n.\end{aligned} \tag{12}$$

Following Norton [5], using $\delta M_d = \sum_{n=1}^{N} \left(\partial M_d / \partial r_n\right)\delta r_n$ results in the desired equation for the gradient of M_d:

$$\begin{aligned}\frac{\partial M_d}{\partial r_n} = &\left\langle \psi_d^*, \delta(r - r_n)\Delta q_n \right\rangle - \left\langle \psi_d^*, \delta(r - r_n)\Delta\Sigma_{t,n}\psi \right\rangle \\ &+ \left\langle \psi_d^*, \delta(r - r_n)\Delta S_n \right\rangle.\end{aligned} \tag{13}$$

Note that the Dirac delta functions in Eq. (13) convert the volume integrals to surface integrals evaluated on the unperturbed surfaces.

Equations similar to Eq. (13) have also been worked out to calculate derivatives for the same types of problems dealt with in Refs. 1-3.

Although both first and second derivatives of the error functional are required to minimize it, in the standard implementation of the Marquardt method [4], only the first derivatives of the quantities of interest are used.

NUMERICAL RESULTS

The method was tested on the same one-dimensional spherical shielded Godiva model that was used in Refs. 1-3 and 6. The materials and outer radii in the model are: high-enriched uranium (HEU), $r_1 = 8.741$ cm; void, $r_2 = 12.4$ cm; lead, $r_3 = 12.9$ cm; and aluminum, $r_4 = 13.2$ cm. The densities of the HEU, lead, and aluminum are 18.74, 11.4, and 2.7 g/cm^3, respectively.

For this problem, the quantities of interest were the leakages from the system of four passive decay lines from uranium: 144, 186, 766, and 1001 keV. The forward and adjoint line fluxes were calculated using the multigroup discrete-ordinates code PARTISN [8] with S_{32} quadrature. Because passive decay lines were used, scattering was ignored under the assumption that a scattering event cannot land a photon in the narrow energy range of a decay line. With this approximation, the line leakages are independent, and, since only surface integrals are required, a semi-analytic ray-trace algorithm (for calculating the leakages and the derivatives) was also used for comparison [9]. Implementing the method to include scattering is a work in progress.

Accuracy of the Computed Derivatives

The adjoint-based derivatives of Eq. (13) were compared with a central-difference approximation obtained by varying each interface independently by ± 0.001 cm for both S_{32} and ray-trace calculations. The results are presented in Table I. Clearly, Eq. (13) is extremely accurate for shield-only interfaces but has some difficulty with source interfaces when discrete-ordinates fluxes are used. The reason is that when $\Delta q \neq 0$, the equation involves a subtraction of terms of similar size, so the inner products must be extremely accurate. Also, the low-energy, easily attenuated lines are more difficult to handle. It will be most interesting to add scattering to this problem.

TABLE I. Difference Between Finite-Difference Derivatives and Adjoint-Based Derivatives.

S_{32}	r_1[a]	r_2	r_3	r_4
144 keV	–71.415%	–0.309%	–0.439%	0.000%
186 keV	–28.410%	–0.051%	–0.133%	0.000%
766 keV	–5.180%	0.013%	–0.001%	0.000%
1001 keV	–1.922%	0.011%	0.000%	0.000%
Ray-trace				
144 keV	0.002%	–0.011%	–0.011%	0.000%
186 keV	0.004%	–0.003%	–0.003%	0.000%
766 keV	0.000%	0.000%	0.000%	0.000%
1001 keV	0.000%	0.000%	0.000%	0.000%

[a] Note that r_1 is the source radius.

Convergence of the Marquardt Method

The convergence of the method for problems involving unknown interface locations is merely demonstrated with two test problems. A true statistical analysis of the convergence properties is beyond the scope of this brief summary. Results from the adjoint-based differentiation of Eq. (13) are compared with those with semi-analytic ray-trace differentiation, which is extremely accurate (see Table I).

In this application of the Marquardt method, the error functional to minimize was the standard χ^2:

$$\chi^2 \equiv \sum_{d=1}^{D} \left(\frac{M_{d,o} - M_d(\mathbf{r})}{\sigma_{d,o}} \right)^2, \qquad (14)$$

where $M_{d,o}$ and $\sigma_{d,o}$ are the observed value of the quantity of interest and its associated uncertainty, respectively, and $\mathbf{r} = \{r_n, n = 1,\ldots,N\}$. The "observed" values $M_{d,o}$ were actually the computed values for the known true configuration and the associated uncertainties were arbitrarily set to 1%. (The iteration histories presented below would change if, as in the real world, different lines were measured with different precision.)

For the first problem, the source radius was known and the initially guessed radii for the shields were $\mathbf{r}$ = {12.3, 12.6, 15.0}. The root-mean-squared (RMS) error for the calculated lines for this model was ~3900% when all four lines were used but ~390% when the 144-keV line was not used. (These RMS values suggest the extreme sensitivity of the leakages, particularly for the low-energy lines, to the interface locations.) The iteration history (χ^2 vs. iteration number) for these two cases is shown in Fig. 1. With ray-trace differentiation, the method converged exactly to the correct solution, $\mathbf{r}$ = {12.4, 12.9, 13.2}. With S_{32} fluxes, the method converged exactly to the correct solution when all four lines were used. When the

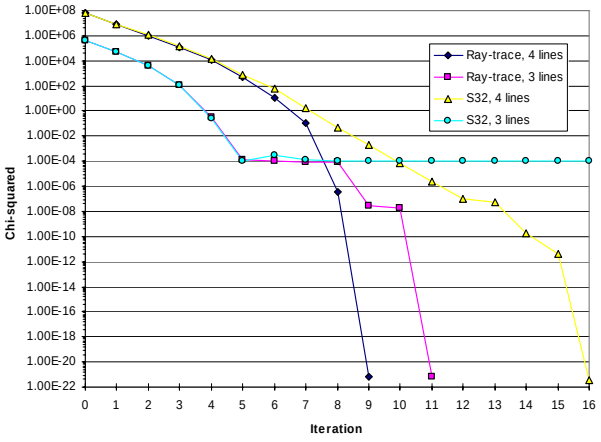

Fig. 1. Convergence for the problem when the source radius was known.

144-keV line was not used with S_{32} fluxes, the method got stuck too early at a model that was quite different from the true one, $\mathbf{r}$ = {12.2066, 12.7050, 12.9968}, which is about 0.2 cm too small for each interface location. This result indicates that although the derivatives are less accurate for the 144-keV line (see Table I), there is important information contained in them.

For the second problem, all four interfaces were unknown and the initially guessed outer radii were $\mathbf{r}$ = {7.5, 12.3, 12.6, 15.0} (the initial shield interfaces were the same as in the previous problem). The initial RMS error was ~3700% when all four lines were used but ~340% when the 144-keV line was not used. The iteration histories for these two cases are shown in Fig. 2. In this problem, the iterations using the ray-trace derivatives were similar to those for the S_{32} fluxes, although the S_{32} derivatives are quite inaccurate for the source radius (see Table 1). Using all four lines prevented the method from making much progress. Not using the 144-keV line allowed the method to converge all lines to within 0.0001% but to a solution that was not exact: $\mathbf{r}$ = {8.736, 12.205, 12.703, 12.989} for both S_{32} calculations and ray-tracing. Note the similarity of this solution to the one obtained when three lines were used with S_{32} fluxes in the previous problem. Interestingly, although the shield interfaces were not very well calculated, the source radius was (8.736 cm vs. the true answer of 8.741 cm). When the 144-keV line was added back in to the model once χ^2 was down to 10^{-9}, the method with ray-trace derivatives converged to the exact solution, but not with S_{32} fluxes.

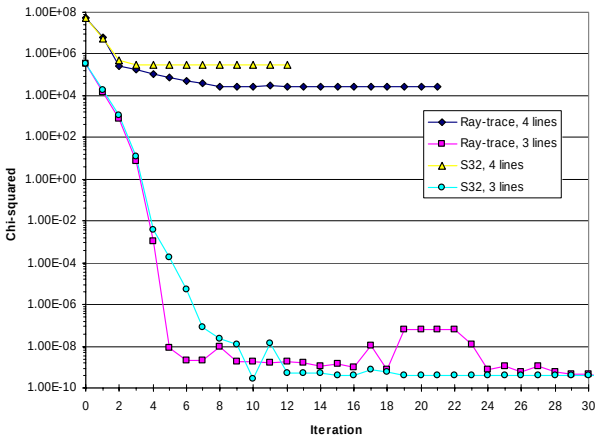

Fig. 2. Convergence for the problem when all four interfaces were unknown.

CONCLUSION

The Marquardt method has been used in an attempt to reconstruct the internal interfaces of an unknown radioactive object using the measured leakage of passive gamma-ray lines. Initial results, using unscattered

gamma-rays only, are promising but indicate that some operator intervention would be required for successful application. Thus, the method could be useful for characterizing static systems given well-measured data, rather than, for example, automatic characterization and discrimination at portal monitors.

The necessary gradients are obtained using an adjoint-based method in which the inner products simplify to surface integrals (for the problem of determining unknown interface locations). For problems that do not involve scattering, semi-analytic ray-trace derivatives , which are very accurate, can be used.

The results of this paper indicate that there are problems for which the use of adjoint-based derivatives, which may be inaccurate for source interfaces, result in the same iteration history and convergence as the use of extremely accurate derivatives. This result suggests that adjoint-based derivatives might be accurate enough to be used on problems with scattering. Important future work is to implement the method for problems that include scattering.

In addition, the method still needs to be tested against measurements of real systems or simulations that more realistically resemble measurements.

REFERENCES

1. J. A. FAVORITE, "Using the Schwinger Variational Functional for the Solution of Inverse Transport Problems," *Nucl. Sci. Eng.*, **146**, 51-70 (2004).
2. J. A. FAVORITE and K. C. BLEDSOE, "Identification of an Unknown Material in a Radiation Shield Using the Schwinger Inverse Method," *Nucl. Sci. Eng.*, **152**, 106-117 (2006).
3. J. A. FAVORITE, "Identification of Unknown Densities in a Source/Shield System Using the Schwinger Inverse Method," *Trans. Am. Nucl. Soc.*, **93**, 588-589 (2005).
4. W. H. PRESS et al., *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd Ed. (reprinted with corrections), Chap. 15, Cambridge University Press (1994).
5. S. J. NORTON, "A General Nonlinear Inverse Transport Algorithm Using Forward and Adjoint Flux Computations," *IEEE Trans. Nucl. Sci.*, **44**, 153-162 (1997).
6. J. A. FAVORITE and R. SANCHEZ, "An Inverse Method for Radiation Transport," 10th International Conference on Radiation Shielding/Radiation Protection and Shielding 2004, Funchal, Portugal, May 9-14; *Radiation Protection Dosimetry*, **116**, *1-4*, 482-485 (2005).
7. J. C. RAGUSA, "Application of Duality Principles to Solve Inverse Particle Transport Problems: A Framework," *Trans. Am. Nucl. Soc.*, **93**, 427-429 (2005).
8. R. E. ALCOUFFE et al., "PARTISN," LA-CC-98-62, Los Alamos National Laboratory (1997).
9. D. KETCHESON, "Inverse Solution Methods and Perturbation Estimates for Radiation Leakage Without Scattering," Research Note X-5-RN(U)-05-35, LA-UR-05-7048, Los Alamos National Laboratory (2005).