Stanford Exploration Project, Report 92, November 12, 1997, pages 235–244

Nonlinear least squares and regularization

James G. Berryman[1]

keywords: *modeling, least squares, nonlinear, linear*

ABSTRACT

I present and discuss some general ideas about iterative nonlinear output least-squares methods. The main result is that, if it is possible to do forward modeling on a physical problem in a way that permits the output (*i.e.*, the predicted values of some physical parameter that could be measured) and the first derivative of the same output with respect to the model parameters (whatever they may be) to be calculated numerically, then it is possible (at least in principle) to solve the inverse problem using the method described. The main trick learned in this analysis comes from the realization that the steps in the model updates may have to be quite small in some cases for the implied guarantees of convergence to be realized.

INTRODUCTION

A problem frequently encountered in the earth sciences, and in other physical and biomedical sciences as well, requires deducing physical parameters of the system of interest from measurements of some other (hopefully) closely related physical quantity. The obvious example in seismology (either surface reflection seismology or crosswell seismic tomography) is the use of measurements of sound wave traveltime to deduce wavespeed distribution in the earth and then subsequently to infer the values of other physical quantities of interest such as porosity, water or oil saturation, permeability, etc.

Many of the problems of interest can be formulated in a way such that the measured quantities (often called the "output" for reasons that will become clear in a moment) may be compared to predicted values of those same quantities and the resulting observed discrepancies then used to make "improvements" in the system model parameters of real interest. Predicted values are obtained by forward modeling based on some assumed model of the physical quantities of interest, and these predicted values are clearly the "output" of such a forward modeling code. Comparisons be-

[1]**email:** berryman@sep.stanford.edu

tween predicted output and measured output parameters may be done in a variety of ways, but a common method is the output least-squares method: discrepancies are squared and summed, and some numerical procedure is established to reduce the overall least-squares error in the output quantity.

The trick in all this is to find a method that actually does guarantee convergence of the output least-squares functional to zero, or at least to a small number whose size is comparable to that expected from the sums of squares of the measurement errors. The purpose of this paper is to show that such a procedure can essentially always be constructed as long as one additional condition is present: If the nonlinear functional of the model parameters used to compute the outputs can be differentiated once with the respect to each of the model parameters of interest (whether this derivative is taken analytically or numerically does not appear to be important), then an evolution equation can be found that will essentially guarantee that a sequence of models gradually improving the agreement between the measured data and the predicted data can be found in a systematic way. The only caveat is that the step size of the improvement from one iteration to the next will be problem dependent and may be rather small in some applications of interest.

NONLINEAR OUTPUT LEAST-SQUARES METHOD

In an abstract setting, let $\{d_i\}$ for $i = 1, \dots, m$ be a set of m measurements. This data may be organized into a data m-vector $\mathbf{d}$. Let $\mathbf{s}$ be an n-vector containing the n parameters for the model/solution space to be constructed. Then, let $N_i(\mathbf{s})$ be a (possibly nonlinear) function of the model parameter vector $\mathbf{s}$ such that $N_i(\mathbf{s}) = d_i$ in infinite precision forward modeling of the physical problem of interest. If the data d_i and the functions N_i are known, then the inversion problem is to find an approximate $\mathbf{s}$ that satisfies the data (*i. e.*, has as small a data misfit as possible) in some sense.

If life were simple, the operators N_i would be analytically invertible, so one might think it could be possible to write the solution to the inverse problem as $\mathbf{s} = N_i^{-1}(d_i)$. However, the result is certainly a strange looking formula: While it is perfectly sensible that the domain of each function N_i is a set of n model parameters, and that the range is a single datum, the postulated inverse function with multiple model outputs from single datum inputs is clearly not sensible. In most cases, there will simply not be sufficient information contained in a single datum to determine multiple model parameters.

For strictly linear problems, there might exist an appropriate operator X acting on the full data m-vector such that $X(\mathbf{d}) = \mathbf{s}$ produces the model n-vector. However, even for linear problems, the existence of such an operator is not necessarily guaranteed since the sizes m and n of the two vector spaces will generally differ. This difficult situation leads to the introduction of pseudoinverses in linear problems or to the methods to be described now in nonlinear problems.

So, instead of using some hypothetical analytical approach, the desired solution to the inverse problem will make simultaneous use of all (or at least most) of the data while determining some sort of optimal fit to a chosen set of model parameters. This situation is common in data analysis and often leads to formulation of least-squares methods for this type of inversion problem.

An approach to this problem based on nonlinear least-squares inversion considers the nonlinear function

$$F(\mathbf{s}) = \sum_{i=1}^{m} [d_i - N_i(\mathbf{s})]^2, \tag{1}$$

which is simply the sum of the squares of the data residuals $r_i = d_i - N_i$. I formulate an "evolution" equation for the model vector $\mathbf{s} = \mathbf{s}(\eta)$, where η is an evolution parameter – treated as a continuous generalization of an iteration number. In particular, I must ultimately discretize this parameter in my final numerical method, in which case the particular values of $\eta = 0, 1, 2, \ldots$ will be exactly the iteration numbers. But for purposes of explaining the method, it will prove useful to treat η initially as a continuous parameter.

The thrust of the nonlinear least-squares method is to produce a model $\mathbf{s}$ that minimizes the least-squares error function $F(\mathbf{s})$. One way to guarantee [see, for example, Jeffrey and Rosner (1986) and Lu and Berryman (1991)] that a local minimum is achieved is to pose the evolution equation for $\mathbf{s}(\eta)$ in a way that guarantees the value of $F(\mathbf{s})$ decreases monotonically as η increases (or at each iteration step for the discretized problem). Taking the derivative of $F(\mathbf{s})$ with respect to the evolution parameter, the chain rule gives

$$\frac{\partial F(\mathbf{s})}{\partial \eta} = -2 \sum_{j=1}^{n} \sum_{i=1}^{m} [d_i - N_i(\mathbf{s})] \frac{\partial N_i}{\partial s_j} \frac{\partial s_j}{\partial \eta}. \tag{2}$$

It is desired that the evolution equation for $\mathbf{s}$ be chosen to guarantee that (2) is ≤ 0. It is easy to see that (2) will always be negative or zero if the evolution equations for the model parameters s_j are chosen to be

$$\frac{\partial s_j}{\partial \eta} = \gamma(\eta) \sum_{i=1}^{m} [d_i - N_i(\mathbf{s})] \frac{\partial N_i}{\partial s_j}, \tag{3}$$

where $\gamma(\eta)$ is a positive parameter to be determined. Then,

$$\frac{\partial F(\mathbf{s})}{\partial \eta} = -2\gamma(\eta) \sum_{j=1}^{n} \left[\sum_{i=1}^{m} [d_i - N_i(\mathbf{s})] \frac{\partial N_i}{\partial s_j} \right]^2, \tag{4}$$

so each term in the sum over j in (4) is a square quantity and must be positive or vanish identically. The choice (3) of evolution then clearly guarantees the desired monotonically decreasing behavior of the total squared error functional as $\eta \to \infty$.

To implement this procedure, a choice of discretization must be made together with a choice for $\gamma(\eta)$. The obvious choice of discretization for the evolution equation

(3) is

$$s_j(k+1) - s_j(k) \equiv \Delta\mathbf{s}_j = \gamma(\eta)[d_i - N_i(\mathbf{s}(k))]\frac{\partial N_i(\mathbf{s})}{\partial s_j}|_{\mathbf{s}=\mathbf{s}(k)}, \tag{5}$$

where I have taken the finite step size for the evolution to be $\Delta\eta = 1$. In the continuous evolution problem, the infinitesimal changes in the evolution parameter guarantee similarly infinitesimal changes in the model vector $\mathbf{s}(\eta)$ and therefore this makes the choice of $\gamma(\eta)$ largely arbitrary. In contrast, for the discretized problem, the evolution of $\mathbf{s}$ is finite at each step and care must be taken not to violate the desired condition that the least-squares functional should decrease at each step. Such violations may occur if the step size is taken too large.

Reconsidering (1), I find that, by keeping only those terms proportional to the first and second powers of the components of $\Delta\mathbf{s}$, I have

$$F(\mathbf{s}(k+1)) = \sum_{i=1}^{m}[d_i - N_i(\mathbf{s}(k)+\Delta\mathbf{s})]^2 \simeq \sum_{i=1}^{m}[d_i - N_i(\mathbf{s}(k))]^2$$
$$-2\sum_{j=1}^{n}\sum_{i=1}^{m}[d_i - N_i(\mathbf{s})]\frac{\partial N_i}{\partial s_j}\Delta\mathbf{s}_j + \sum_{i=1}^{m}\left[\sum_{j=1}^{n}\frac{\partial N_i}{\partial s_j}\Delta\mathbf{s}_j\right]^2. \tag{6}$$

After substituting (5), the parameter $\gamma(\eta)$ can now be chosen so that the right-hand side of (6) decreases as much as possible at each step of the iteration scheme. The optimum choice is easily shown to be

$$\gamma(k+1) = \frac{\sum_{j=1}^{n}\left[\sum_{i=1}^{m}[d_i - N_i(\mathbf{s}(k))]\frac{\partial N_i}{\partial s_j}\right]^2}{\sum_{j,j'}\left[\sum_{i'=1}^{m}[d_{i'} - N_{i'}(\mathbf{s}(k))]\frac{\partial N_{i'}}{\partial s_j}\sum_{i=1}^{m}\frac{\partial N_i}{\partial s_j}\frac{\partial N_i}{\partial s_{j'}}\sum_{i''=1}^{m}[d_{i''} - N_{i''}(\mathbf{s}(k))]\frac{\partial N_{i''}}{\partial s_{j'}}\right]}, \tag{7}$$

since this minimizes the right-hand side of (6).

It will prove enlightening to compare the procedure just presented with the well-known method of conjugate gradients (Hestenes and Stiefel, 1952; Fomel, 1996) for a linear operator such that $N_i(\mathbf{s}) = (\mathbf{Ls})_i$. Then, model updates are obtained using

$$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} + \gamma^{(k+1)}\mathbf{u}^{(k+1)}, \tag{8}$$

where, for the discrete iterative problem, I put the iteration numbers in superscripts. The new vector $\mathbf{u}^{(k+1)}$ is the (somehow) known update direction in the model space and $\gamma^{(k+1)}$ is a parameter used to optimize the step size. The updated residual vector is again given by

$$\mathbf{r}^{(k+1)} = \mathbf{d} - \mathbf{Ls}^{(k+1)} = \mathbf{r}^{(k)} - \gamma^{(k+1)}\mathbf{Lu}^{(k+1)}. \tag{9}$$

The magnitude of the residual vector is easily shown to decrease most at each step of the iteration sequence if the optimization parameter satisfies

$$\gamma^{(k+1)} = \frac{(\mathbf{r}^{(k)}, \mathbf{Lu}^{(k+1)})}{||\mathbf{Lu}^{(k+1)}||^2}. \tag{10}$$

Equation (10) is exactly analogous to the formula (7) obtained in the nonlinear least-squares problem if the components of the residual are given by

$$r_i^{(k)} = d_i - N_i(\mathbf{s}^{(k)}), \tag{11}$$

the matrix elements of the linear operator are

$$(\mathbf{L})_{ij} = L_{ij} = \frac{\partial N_i}{\partial s_j}\Big|_{\mathbf{s}=\mathbf{s}^{(k)}}, \tag{12}$$

and the update direction vector satisfies

$$\gamma^{(k+1)}\mathbf{u}^{(k+1)} = \frac{\partial \mathbf{s}}{\partial \eta}, \tag{13}$$

where the components of the right-hand side of (13) are given by (3) evaluated at $\mathbf{s}^{(k)}$. With this identification, it becomes clear that the nonlinear least-squares method outlined above is one natural generalization of the well-known conjugate gradients technique.

REGULARIZATION

Now it may happen, due to a combination of poor conditioning of the linear operator $\mathbf{L}$ and unfortunate choices of starting guess of the model vector $\mathbf{s}$, that the denominator of the right-hand side of (10) may be very small or vanish to numerical accuracy. Similar circumstances can arise in the nonlinear least-squares problem in cases of sparse or irregularly sampled data. When such circumstances arise in practice, it may be necessary to regularize the method by adding an additional constraint equation to the least-squares functional (1). Regularization is a well-known technique often associated with the names of Tikhonov and Arsénine (1976), Levenberg (1944), and Marquardt (1963, 1970), among others.

The regularization constraint usually takes the form of a quadratic functional of the model vector. One typical choice is $\mu(\mathbf{s} - \bar{\mathbf{s}})^T(\mathbf{s} - \bar{\mathbf{s}})$ where μ (often called the damping parameter) is some small positive constant and $\bar{\mathbf{s}}$ is some value of the model vector that the final solution should not deviate from too much. Other typical choices of regularization constraint are based on differentials of the model taking the form $(\mathbf{D}\mathbf{s})^T(\mathbf{D}\mathbf{s})$, where $\mathbf{D}\mathbf{s}$ might be either a simple gradient or a Laplacian of the model – assuming that $\mathbf{s}$ is some simple physical quantity. If $\mathbf{s}$ is a more complicated vector of model parameters which is not easily given a simple physical interpretation, then some other choice of regularization constraint might be needed.

It is preferable to avoid using regularization if possible, because such techniques tend to modify the entire spectrum of the operator to be inverted and therefore tend to degrade resolution.

For the linear least-squares problem, I can carry the analysis further to understand how the iteraton scheme modifies the model estimate at each step. Assuming that the linear operator $\mathbf{L}$ is a matrix of dimensions $m \times n$, where m and n are usually not equal, it is helpful to use $\mathbf{L}$ and its transpose $\mathbf{L}^T$ to form a square, symmetric matrix and the associated eigenvalue problem

$$\begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{L}^T & 0 \end{pmatrix} \begin{pmatrix} \psi_q \\ \phi_q \end{pmatrix} = \lambda_q \begin{pmatrix} \psi_q \\ \phi_q \end{pmatrix}. \tag{14}$$

Assuming that the m-vectors ψ_q and the n-vectors ϕ_q are normalized to unity, the singular value decomposition of the matrix $\mathbf{L}$ may now be written in terms of these eigenfunctions as

$$\mathbf{L} = \sum_{q=1}^{r} \psi_q \lambda_q \phi_q^T. \tag{15}$$

The sum is taken over r terms, where r is the rank of $\mathbf{L}$.

Now, each model estimate $\mathbf{s}^{(k)}$ may be expanded in terms of the appropriate eigenfunctions according to

$$\mathbf{s}^{(k)} = \sum_{q=1}^{r} \alpha_q^{(k)} \phi_q, \tag{16}$$

where the $\alpha_q^{(k)}$s are the expansion coefficients for the kth iteration. Similarly, the data vector may also be expanded as

$$\mathbf{d} = \sum_{q=1}^{r} \delta_q \psi_q, \tag{17}$$

where the δ_qs are constant coefficients. Both the model vector and the data vector expansion might in addition include a term from the null space of $\mathbf{L}$, but for simplicity I will ignore this possibility for the present purposes.

With these definitions of the various coefficients, I find that equation (10) becomes

$$\gamma^{(k+1)} = \frac{\sum_{q=1}^{r} \lambda_q^2 (\delta_q - \lambda_q \alpha_q^{(k)})^2}{\sum_{q=1}^{r} \lambda_q^4 (\delta_q - \lambda_q \alpha_q^{(k)})^2}, \tag{18}$$

and the update equation for the α_qs becomes

$$\alpha_q^{(k+1)} = \alpha_q^{(k)} + \gamma^{(k+1)} \lambda_q (\delta_q - \lambda_q \alpha_q^{(k)}). \tag{19}$$

It follows from (19) that the iteration process has converged when

$$\alpha_q^{(k)} = \frac{\delta_q}{\lambda_q}. \tag{20}$$

As the coefficients approach convergence, I discover that the denominator of (18) can become quite small. If I assume that the eigenvectors with the largest eigenvalues converge most quickly, then after some number of iterations the main contributions to the denominator will be from the terms associated with the smallest eigenvalues, and these contributions are proportional to the fourth power of these small eigenvalues. If this happens, some type of regularization may be required to obtain useful results.

Consider the simplest type of regularization involving a model vector constraint so that the modified objective function becomes:

$$F_\mu(\mathbf{s}) = \sum_{i=1}^{m}[d_i - N_i(\mathbf{s})]^2 + \mu\sum_{j=1}^{n}(s_j - \bar{s}_j)^2. \tag{21}$$

The expansion of the constraint vector is given by

$$\bar{\mathbf{s}} = \sum_{q=1}^{r}\bar{\alpha}_q\phi_q. \tag{22}$$

The equation for the modified model update is then given by

$$\alpha_q^{(k+1)} = \alpha_q^{(k)} + \gamma^{(k+1)}\left[\lambda_q\delta_q + \mu\bar{\alpha}_q - (\lambda_q^2 + \mu)\alpha_q^{(k)}\right]. \tag{23}$$

Clearly, the modified iteration sequence for α_q has converged when

$$\alpha_q^{(k)} = \frac{\lambda_q\delta_q + \mu\bar{\alpha}_q}{\lambda_q^2 + \mu}, \tag{24}$$

showing that the limiting relation is just a linear combination of the starting value and the result for pure least-squares as seen in (20). If the damping parameter μ is small, these two values will of course be quite close. But, in general, for any nonzero value of the damping parameter, it is inevitable that the resolution will suffer due to the fact that the coefficients cannot approach their optimum value (20) but are actually constrained away from it by the regularization procedure. This is why regularization should be avoided, or at least minimized, as much as possible.

SUMMARY AND SUGGESTIONS FOR FUTURE WORK

The main result obtained in this paper is this: If it is possible to do forward modeling on a physical problem in a way that permits the output (*i.e.*, the predicted values of some physical parameter that could be measured) and the first derivative of that same output with respect to model parameters (whatever they may be) to be calculated numerically, then (at least in principle) it is possible to solve the inverse problem using the method described. The main trick learned in this analysis comes from the realization that the steps in the model updates may have to be quite small in some cases for the implied guarantees of convergence to be realized.

We have concentrated on least-squares methods in the presentation in order to keep the analysis simple. However, it is clear that the general method presented could be applied as well to any norm of the data discrepancies that possesses the same necessary qualities, the most important of these being the existence of a first derivative with respect to the model parameters. One interesting example of such an alternative is Huber's hybrid of l_1 and l_2 [see Clearbout (1996)]. Other choices include special cases of the more general l_p norm that also have a first derivative. These other methods will be discussed in future work.

ACKNOWLEDGMENTS

I want to thank Liliana Borcea and George Papanicolaou for numerous discussions of inverse problems and the use of nonlinear output least-squares methods.

REFERENCES

Claerbout, J., 1996, Conjugate-direction Huber regression: **SEP-92**, 229–236.

Fomel, S., 1996, Least-square inversion with inexact adjoints. Method of conjugate directions: A tutorial: **SEP-92**, 253–266.

Hestenes, M. R., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: *J. Res. NBS* **49**, 409–436.

Jeffrey, W., and R. Rosner, 1986, Optimization algorithms: Simulated annealing and neural network processing: *Astrophys. J.* **310**, 473–481.

Levenberg, K., 1944, A method for the solution of certain non-linear problems in least squares: *Quart. Appl. Math.* **2**, 164–168.

Lu, S.-Y., and J. G. Berryman, 1991, Inverse scattering, seismic traveltime tomography, and neural networks: *Intern. J. Imaging Sys. Tech.* **2**, 112–118.

Marquardt, D. W., 1963, An algorithm for least-squares estimation of nonlinear parameters: *SIAM J. Appl. Math.* **11**, 431–441.

Marquardt, D. W., 1970, Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation: *Technometrics* **12**, 591–612.

Tikhonov, A., and V. Arsénine, *Méthodes de Résolution de Problèmes Mal Posés*, Éditions Mir, Moscow, 1976.