Submission code: PBRQVP74X9

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Nonlinear elastic seismic inversion using Levenberg-Marquardt with trust region.

SERGIO OLIVEIRA (UENF)

Nonlinear elastic seismic inversion using Levenberg-Marquardt with trust region.

Although the trust region method offers a mathematically natural and rigorous alternative to control the stability factor of the Levenberg-Marquardt nonlinear optimization algorithm, empirical criteria are commonly used for this task. In this work, we propose and test, on real data, an implementation of the full waveform prestack elastic seismic inversion using the Levenberg-Marquardt algorithm with trust region and compare it's results with those obtained when the most common empirical criterion is used to control the stability factor of this algorithm.

Introduction

Nonlinear optimization methods are the mathematical engine of full waveform seismic inversion. Among the most widely used methods are those based on the information of the first derivative of the objective function (or error function), such as the gradient method, conjugate gradient and BFGS, and methods that use the information of the second derivative, such as the Newton or Gauss-Newton method (Fletcher, 1987), the latter being more robust and converging more efficiently. A comparison of such methods, specifically applied to the waveform inversion problem, can be found in Pratt et al. (1998). In general, such methods work by taking a certain step in a certain direction in the model space at each iteration. What changes is the strategy for defining the search direction and the criterion for defining the step size. Conventionally, the search direction is established first and the step size is obtained later. Thus, for example, in the gradient method the search direction is the opposite direction to the gradient vector of the objective function and the step size is normally established by searching for the minimum of this function along this direction. In the Newton algorithm the search direction involves the inverse of the Hessian matrix and in the Gauss-Newton algorithm an approximation of the Hessian matrix is used (Nocedal and Wright, 1999).

The trust region method searches, at each iteration, for the minimum within a subdomain in the model space where the objective function is evaluated using an approximation function. If this function is able to adequately reproduce the values of the objective function within this subdomain, then the region is expanded; conversely, if the approximation is poor, then the region is contracted, since an approximation function is "trustworthy" only in the region where it provides a reasonable approximation. In this method, the radius of the trust region is first estimated and then a step direction that leads to the minimum of the approximation function within this region is determined (Sorensen, 1982). This method is also known as the constrained step method (Fletcher, 1987).

The trust region method can be naturally implemented in the classical Levenberg-Marquadt algorithm, since there is a direct relationship between the stabilization factor λ and the radius of the trust region of the quadratic approximation of the objective function, on which this optimization algorithm is based. In this work, we propose and test a version of the presack elastic inversion method AVAFWI (Oliveira et al, 2018) that incorporates a version of the LM algorithm with trust region. This inversion algorithm is used to invert a real seismic data and the result is confronted with that obtained by the conventional implementation of the LM method, which is based on an empirical criterion to dimension the stability factor.

Method

In full waveform elastic inversion, all propagation effects are taken into account, however it is necessary to assume a non-linear mathematical relationship between the data and the model parameters, given by:

$$\mathbf{d} = \mathbf{S}(\mathbf{m}) \tag{1}$$

Where $\boldsymbol{S}$ represents a mathematical operator that connects the data vector d to the model parameter vector $\mathbf{m}$. In the inverse problem in question, the subsurface is represented as a 1-D geological medium discretized into elementary layers, and the parameters to be determined are the P-wave velocity, S-wave velocity, and density in each elementary layer. The vector $\mathbf{d}$ represents a CMP set after undergoing seismic migration and being reordered by incidence angle. Once the observed data is known and $\mathbf{s(m)}$ is calculated, which is the seismic response of the earth as a function of the parameters, the inversion is done via nonlinear regression, that is, the model parameters are defined by fitting the data. Mathematically, this means that the solution consists of a vector $\mathbf{m}$ that minimizes an error function that measures the distance between the observed and calculated data:

$$\mathbf{m} = \min_{\mathbf{m}}\{\, E(\mathbf{m}) = \Delta\mathbf{d}^{\mathbf{T}}\Delta\mathbf{d}\,\} \tag{2}$$

Where $\mathbf{\Delta d}$ is a vector whose elements are the difference between the calculated and the observed data: $\mathbf{\Delta d=(d\text{-}S(m))}$ This error function uses L2 norm to measure the distance between the calculated and observed data (Menke, 1989). The objective function (2) can be expanded around $\mathbf{m}$ using Taylor series. This expansion can be truncated from the second-order terms which gives rise to a quadratic approximation for the error function around $\mathbf{m}$, which will here be given as a function of $\boldsymbol{\delta}\mathbf{m}$:

$$E_Q(\delta\mathbf{m}) = E(\mathbf{m}) + \mathbf{g}^T\boldsymbol{\delta}\mathbf{m} + \frac{1}{2}\boldsymbol{\delta}\mathbf{m}^{\mathbf{T}}\mathbf{H}\boldsymbol{\delta}\mathbf{m} \tag{3}$$

Where $\mathbf{g}$ is the gradient vector that contains the derivatives of the objective function with respect to the model parameters: $g_i=\partial E/\partial m_i$. The matrix $\mathbf{H}$ is known as the Hessian and contains the second-order derivatives of the objective function with respect to the parameters:

$$H_{ij} = \frac{\partial^2 E}{\partial m_i \partial m_j} \tag{4}$$

The gradient vector and an approximation to the Hessian matrix can be calculated as a function of the sensitivity matrix, also known as the Jacobian:

$$\mathbf{g} = -\mathbf{2J}^T\Delta\mathbf{d},\ \mathbf{H} \cong \mathbf{2J}^T\mathbf{J} \tag{5}$$

The Jacobian is obtained from differential seismograms; $J_{ik}=\partial S_i/\partial m_k$, for more details see Oliveira et all (2018). The Gauss-Newton method is obtained by seeking the minimum of E_Q, that is, the point where $\partial E_Q/\partial\delta\mathbf{m}=0$. Note that $\mathbf{J^TJ}$ must be positive definite, otherwise the Gauss-Newton method becomes unstable. To get around this problem, Levenberg proposed to damp the absolute value of $\boldsymbol{\delta}\mathbf{m}$, minimizing $E_L(\delta\mathbf{m}) = E_Q(\boldsymbol{\delta}\mathbf{m}) + \lambda\boldsymbol{\delta}\mathbf{m}^T\boldsymbol{\delta}\mathbf{m}$, which is done by the following iterative scheme:

$$\begin{aligned}(\mathbf{J}^T\mathbf{J} + \lambda\mathbf{I})\boldsymbol{\delta}\mathbf{m} &= \mathbf{J}^T\Delta\mathbf{d}\\ \mathbf{m}_{\mathbf{i+1}} &= \mathbf{m}_{\mathbf{i}} + \boldsymbol{\delta}\mathbf{m}\end{aligned} \tag{6}$$

The condition number of system (6) can be controlled by λ, since if the $\mathbf{J^T J}$ matrix has any negative eigenvalues, the stabilization factor λ must be greater than the absolute value of the smallest eigenvalue of this matrix in order to guarantee that $(\mathbf{J^T J} + \lambda \mathbf{I})$ will be positive definite. Calculating the $\mathbf{J^T J}$ eigenvalues is not a simple task, especially in problems with many parameters, due to this an empirical scheme is usually adopted to control λ at each iteration. In these schemes, the value of λ is usually decreased at each iteration by dividing it by a decay rate value: $\lambda_{k+1} = \lambda_k / c$, where $|c|>1$. The problem is that a low value for c can slow down convergence, while a high value can make it unstable, and there is no universal criterion for selecting an optimal value for this. The problem of dimensioning the stability factor of the LM method can be treated in a mathematically rigorous way according to the principle of the trust region. For this we take into account the following results: the value of $\boldsymbol{\delta}\mathbf{m}$ obtained by solving equation (6) minimizes the error function $E(\mathbf{m})$ (eq. 2) on the sphere whose radius $|\delta|^2$ satisfies $|\boldsymbol{\delta}\mathbf{m}|^2 = |\delta|^2$. It is then possible to demonstrate that $\boldsymbol{\delta}\mathbf{m}$ is a decreasing function of λ, since

$$|\boldsymbol{\delta}\mathbf{m}|^2 = \sum_k \frac{u_k^2}{(\gamma_k - \lambda)^2} \tag{7}$$

Where u_k represents the elements of the vector $\mathbf{u} = \mathbf{U^T J^T \delta d}$, $\mathbf{U}$ is the matrix whose columns are the eigenvectors of $\mathbf{J^T J}$ and γ_k represents the eigenvalues of this same matrix; Marquadt (1963), Pujol (2007). Thus, λ can be used naturally to control the radius of the trust region in the following way: let r be the ratio between the true decay of the error function and the decay predicted by the quadratic approximation: $r = \Delta E / \Delta E_Q$, where $\delta E = E(\mathbf{m} + \boldsymbol{\delta}\mathbf{m}) - E(\mathbf{m})$ and $\delta E_Q = E_Q(\boldsymbol{\delta}\mathbf{m}) - E(\mathbf{m})$. The central idea is to modify λ at each iteration in order to keep the ratio r within certain limits. For this we present the following version of a seismic inversion algorithm using LM with trust region. Given $\mathbf{m}, \mathbf{d}, \lambda$ and E_{min};

(1) *calculate* $\mathbf{S(m)}$
(2) $\mathbf{\Delta d = d - S}$
(3) $E = \mathbf{\Delta d^T \Delta d}$
(4) *calculate* $\mathbf{J(m)}$
(5) *solve* $(\mathbf{J^T J} + \lambda \mathrm{I}) \boldsymbol{\delta}\mathbf{m} = \mathbf{J^T \Delta d}$
(6) $\mathbf{m_2 = m} + \boldsymbol{\delta}\mathbf{m}$
(7) *calculate* $\mathbf{S_2(m_2)}$
(8) $\mathbf{\Delta d_2 = d - S_2}$
(9) $E_2 = (\mathbf{\Delta d_2})^T \mathbf{\Delta d_2}$
(10) $\Delta E_Q = \boldsymbol{\delta}\mathbf{m}^T \mathbf{J^T J} \boldsymbol{\delta}\mathbf{m} - 2 \mathbf{\Delta d^T J} \boldsymbol{\delta}\mathbf{m}$
(11) $\Delta E = E_2 - E$
(12) $r = \Delta E / \Delta E_Q$
(13) *if* $r < 0$ *then*
$\lambda = 4\lambda$
return to **(5)**
else
$E = E_2$, $\mathbf{m} = \mathbf{m_2}$
$\mathbf{\Delta d = \Delta d}$
$\mathbf{S = S_2}$
if $r < 1/4$ *then* $\lambda = 4\lambda$
If $r > 3/4$ *then* $\lambda = \lambda/2$
(14) *if* $E < E_{min}$ *then stop*
else return to **(4)**

Note that in this algorithm λ is modified in order to control the size of $\boldsymbol{\delta}\mathbf{m}$ in each iteration, in order to meet the trust region criterion. Thus, the LMTR algorithm (Levenberg-Marquardt with trust region) arrives at a solution that, in addition to minimizing the data error, will also meet the condition of controlling $\boldsymbol{\delta}\mathbf{m}$, what prevents the final solution from deviating too far from the initial model, thus implicitly acting as a regularization.

Results

For this research, the AVAFWI method was implemented using the LMTR algorithm; Levenberg-Marquardt with trust region and LMC; conventional Levenberg-Marquardt. For details on the calculation of the synthetic seismograms, Jacobian and solution of system (6) the reader is referred to (Oliveira et al, 2018). The two versions were then applied to invert the same data and the final results, as well as the convergence curves, were subjected to analysis. This test consists of the inversion of a marine data. In addition to the LMTR, this test was performed using the LMC algorithm with a conservative decay rate to slowly reduce the stability factor over the iterations ($c=2$) and a more aggressive decay rate ($c=10$) was also tested to try to accelerate convergence.

The input data consisted of an inline containing angle gathers with eight traces from 5 to 40 degrees and five degree spacing. The results of the inversion of an angle gather from this data are shown in figure (1). In these figures the results of Vp, Vs and Rho are presented together with the initial model (red dashed line) and with the data from a neighboring well filtered to the same frequency band of the seismic (black line). The results of the elastic inversion using LMTR and using the LMC method with $c=2$ are shown in figure 1a,b,c and figure1d,c,e respectively. The results of the elastic inversion using LMC with $c=10$ are shown in figure 1g,h,i. Figure 1j and 1k shows the real angle gather and the modeled angle gather from the LMTR Vp, Vs and Rho result. The three tests were performed using the same initial value for the stability factor (λ=5) and the same number of twelve iterations. Figure 2c shows the convergence curves of the LMTR and LMC methods with c=2 and c=10. In these, it is possible to observe the behavior of the normalized error throughout the iterations. The normalized error E_N is the value of the objective function in the current iteration divided by the value of this function for the initial model: $E_N=E(\mathbf{m_k})/E(\mathbf{m_0})$. Figure 2a shows the density section obtained using LMTR inversion and Figure 2b shows the density section obtained using LMC with c=10.

Conclusions

For the LMC method, choosing an adequate value for λ and also for c is a critical point for the correct convergence in a reasonable number of iterations. Large values for c made the LMC algorithm converge quickly to a low value of the error function. However, this early convergence ended up arriving at a model that, despite explaining the data well, did not reproduce the values of the elastic parameters satisfactorily and generated instability in the final inversion image. This problem can be solved by choosing a small value for c. This conservative choice tends to avoid these instabilities but at the cost of significantly increasing the number of iterations. The seismic inversion problem has a non-unique solution, so one way to restrict the number of possible solutions is to use some regularization criterion. It is interesting to note that the trust region method ends up playing the role of an implicit regularization in the seismic inversion, since it does not allow exaggerated steps, thus preventing the final solution from deviating too far from the initial model. LMTR is capable of quickly adjusting an adequate value for the stability factor λ, even if a very large or small initial value has been specified for it, what makes it converges in an optimized number of steps

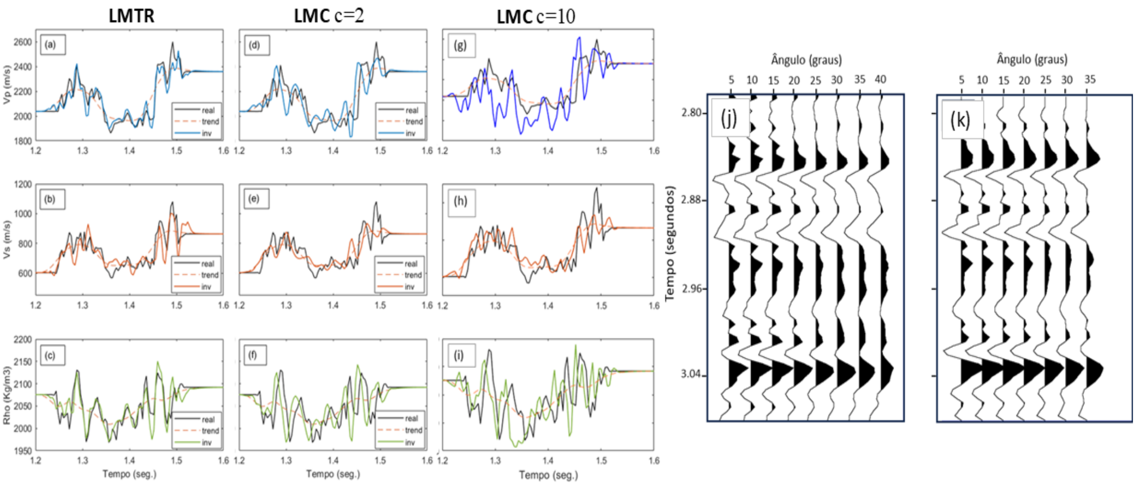

Figure 1: LMRC Inversion results: a) Vp, b) Vs and c) density. LMC Inversion results with c=2: d) Vp, e) Vs and f) density. LMC Inversion results with c=10: g) Vp, h) Vs and i) density. J) Real angle gather k) synthetic angle gather from LMTR inversion.

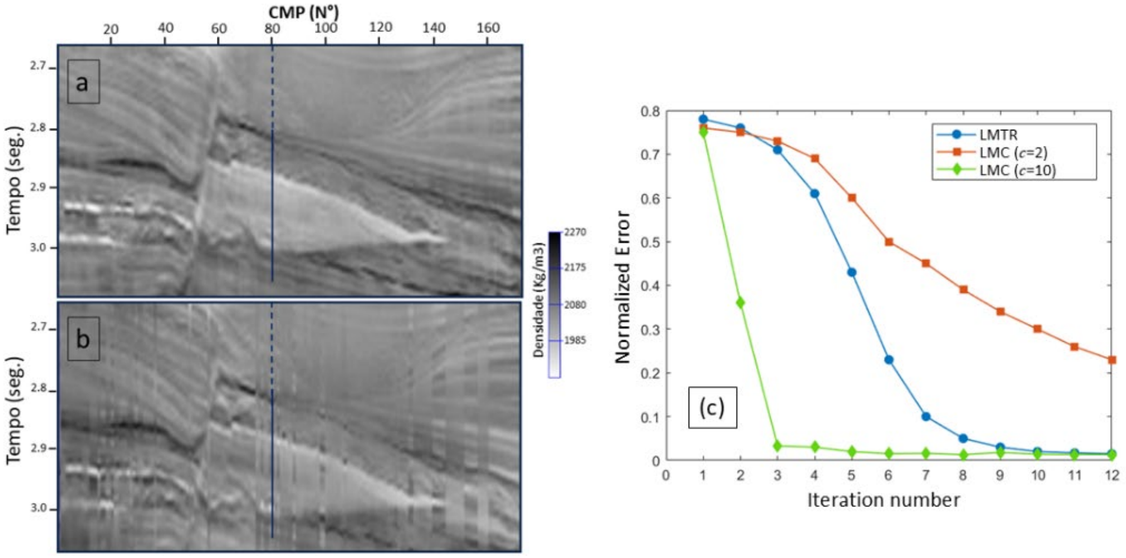

Figure 2: (a) Density section obtained using LMTR (b) Density section obtained using LMC with c=10 (c) Convergence curves for LMTR (blue) and LMC with c=2 (red) and c=10 (green).

Acknowledgments

Thanks to Petrobras and Invision Geophysics for the permission to publish this work.

References

Fletcher, R., 1987. Practical methods of optimization. Wiley International Science Publications.

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

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

Nocedal, J. e Wright, S., 1999. Numerical optimization. New York: Spring.

Oliveira, S. A. M., Braga, I. L. S., Lacerda, M. B., Ouverney, G. F. and Franco, A. W. P., 2018, Extending the useful angle range for elastic inversion through the amplitude-versus-angle full-waveform inversion method, Geophysics, 83(3), R213-R226.

Pratt, R. G., C. Shin, and G. Hicks, 1998. Gauss-Newton and full Newnton methods in frequency space seismic waveform inversion: Geophysical Journal International, 133, 341-362.

Pujol, S., 2007. The solution of nonlinear inverse problems and the Levenberg-Marquadt method: Geophysics, 72, W1-W16.

Sorensen, D. C., 1982. Newton's Method with a Model Trust Region Modification. *SIAM J. Numer. Anal*. **19** (2): 409–426.