Subspace method for nonlinear acoustic multi-parameter FWI

Yu Geng*, Kris Innanen and Wenyong Pan
yu.geng@ucalgary.ca

Abstract

Full waveform inversion (FWI) is a powerful tool to reconstruct subsurface parameters. This highly nonlinear inverse problem is normally solved iteratively to minimize a misfit function, which is usually defined as the distance between the observed and predicted data, by gradient-based method or Newton type method. Incorporating more nonlinearity within each update in FWI, especially for multiparameter reconstruction, may have very important consequences for convergence rates and discrimination of different parameter classes. In this study, we focus on acoustic media with variable density, and the goal is to simultaneously update velocity and density, other parameterization is also discussed. We start from the physical interpretation of both the gradient and the Hessian of the misfit function, and derive one approach from the Newton method, to include the additional term of the Hessian, which contains the second-order partial derivative of the wavefield and related to the second-order scattering, into the gradient, to construct a new descent direction, as well as a set of basis vectors of a subspace that step length can be searched separately for different parameter classes or perturbations.

Theory and Method

FWI method is to seek the high resolution estimation of the subsurface model parameters by solving a nonlinear least-squares minimization problem. The misfit function is defined as the norm of the data residuals

$$\phi(\mathbf{m}) = \frac{1}{2}\sum_{ns}\sum_{n\omega}\left\|\mathbf{d}_{obs}(\mathbf{x}_s,\omega) - \mathbf{d}_{syn}(\mathbf{m},\mathbf{x}_s,\omega)\right\|^2 = \frac{1}{2}\sum_{ns}\sum_{n\omega}\delta\mathbf{d}^T\delta\mathbf{d}^*$$

with $\mathbf{d}_{syn}(\mathbf{m}, \mathbf{x}_s, \omega)=\mathbf{R}\mathbf{u}(\mathbf{m}, \mathbf{x}_s, \omega)$ is the synthetic data generated using the current model $\mathbf{m}$, $\mathbf{R}$ is the sampling matrix that sampling the wavefield from the whole space to the receiver's locations, and $\mathbf{d}_{obs}(\mathbf{x}_s, \omega)$ is the observed data, ^{T} is the transpose operator and * is the conjugate operator. Here, we suppose the wavefield can be obtained by the frequency domain acoustic wave equation

$$\mathbf{A}(\mathbf{m},\omega)\mathbf{u}(\mathbf{m},\mathbf{x}_s,\omega) = \mathbf{f}(\mathbf{x}_s,\omega)$$

where vector $\mathbf{m}$ is the model parameters with different classes, e.g., bulk modulus κ and density ρ, or velocity v and density ρ and so on, $\mathbf{A}(\mathbf{m}, \omega)$ is the impedance matrix, and it is a sparse banded matrix, as the number of non-zero diagonals are related to the finite-difference scheme, e.g., in this study, we use a five-point finite difference scheme, so the impedance matrix has five non-zero diagonals, and $\mathbf{u}(\mathbf{m}, \mathbf{x}_s, \omega)$ is the pressure wavefield, generated by a point source $\mathbf{f}(\mathbf{x}_s, \omega)$ located at $\mathbf{x}_s$.

By expanding the misfit function as a Taylor series up to the second order

$$\phi(\mathbf{m}+\delta\mathbf{m}) = \phi(\mathbf{m}) + \delta\mathbf{m}^T\mathbf{g} + \frac{1}{2}\delta\mathbf{m}^T\mathbf{H}(\mathbf{m})\delta\mathbf{m} + O(|\delta\mathbf{m}|^3)$$

Then a perturbation $\delta\mathbf{m}$ can be found to minimum of the misfit function under the quadratic approximation

$$\delta\mathbf{m} = -\mathbf{H}^{-1}(\mathbf{m})\mathbf{g}$$

The Hessian $\mathbf{H}(\mathbf{m})$ contains two terms, the first term $\mathbf{H}_1(\mathbf{m})$ contains the second-order derivatives of the wavefield, and when it is neglected, the second term $\mathbf{H}_2(\mathbf{m})$ becomes the Gauss-Newton approximation of the Hessian operator.

Rewrite the full Newton inversion equation (Pratt, 1998)

$$(\mathbf{H}_1 + \mathbf{H}_2)\delta\mathbf{m} = -\mathbf{g}$$

$$(\mathbf{H}_2^{-1}\mathbf{H}_1 + \mathbf{I})\delta\mathbf{m} = -\mathbf{H}_2^{-1}\mathbf{g}$$

$$\delta\mathbf{m} = -(\mathbf{H}_2^{-1}\mathbf{H}_1 + \mathbf{I})^{-1}\mathbf{H}_2^{-1}\mathbf{g}$$

$$\delta\mathbf{m} = -\left(\mathbf{I} - \mathbf{H}_2^{-1}\mathbf{H}_1 + (\mathbf{H}_2^{-1}\mathbf{H}_1)^2 - \ldots\right)\mathbf{H}_2^{-1}\mathbf{g}$$

1. Taking only the first two terms, we can get an approximate version of the full Newton method

$$\delta\mathbf{m} = -\mathbf{H}_2^{-1}(\mathbf{g} - \mathbf{H}_1\mathbf{H}_2^{-1}\mathbf{g})$$

When adding nonlinearity into the descent direction

$$\delta\mathbf{m} = -\mathbf{H}_2^{-1}(\mathbf{g} + \mathbf{H}_1\delta\mathbf{m}_1)$$

Which can be considered as another linearized inverse problem with a new descent direction $\mathbf{g}+\mathbf{H}_1\delta\mathbf{m}_1$, or to say including the second-order scattering caused by the perturbation from a linearized inverse problem in the original gradient. In the monoparameter case, we can find out that it is consistent with the nonlinear descent direction we have studied based on the scattering theory in last year's report.

2. Two perturbation can be inverted using this approximate version of the full Newton method, which are related to the standard Gauss-Newton method result and second-order terms

$$\delta\mathbf{m}_1 = -\mathbf{H}_2^{-1}\mathbf{g}$$

$$\delta\mathbf{m}_2 = \mathbf{H}_2^{-1}\mathbf{H}_1\mathbf{H}_2^{-1}\mathbf{g} = -\mathbf{H}_2^{-1}\mathbf{H}_1\delta\mathbf{m}_1$$

Then $\delta\mathbf{m}_1$ and $\delta\mathbf{m}_2$ can be used to form a set of basis vectors $\{\mathbf{a}_j\}$, so that the true perturbation can be found as

$$\delta\mathbf{m} = \sum_{j=1}^{n}\mu_j\mathbf{a}_j$$

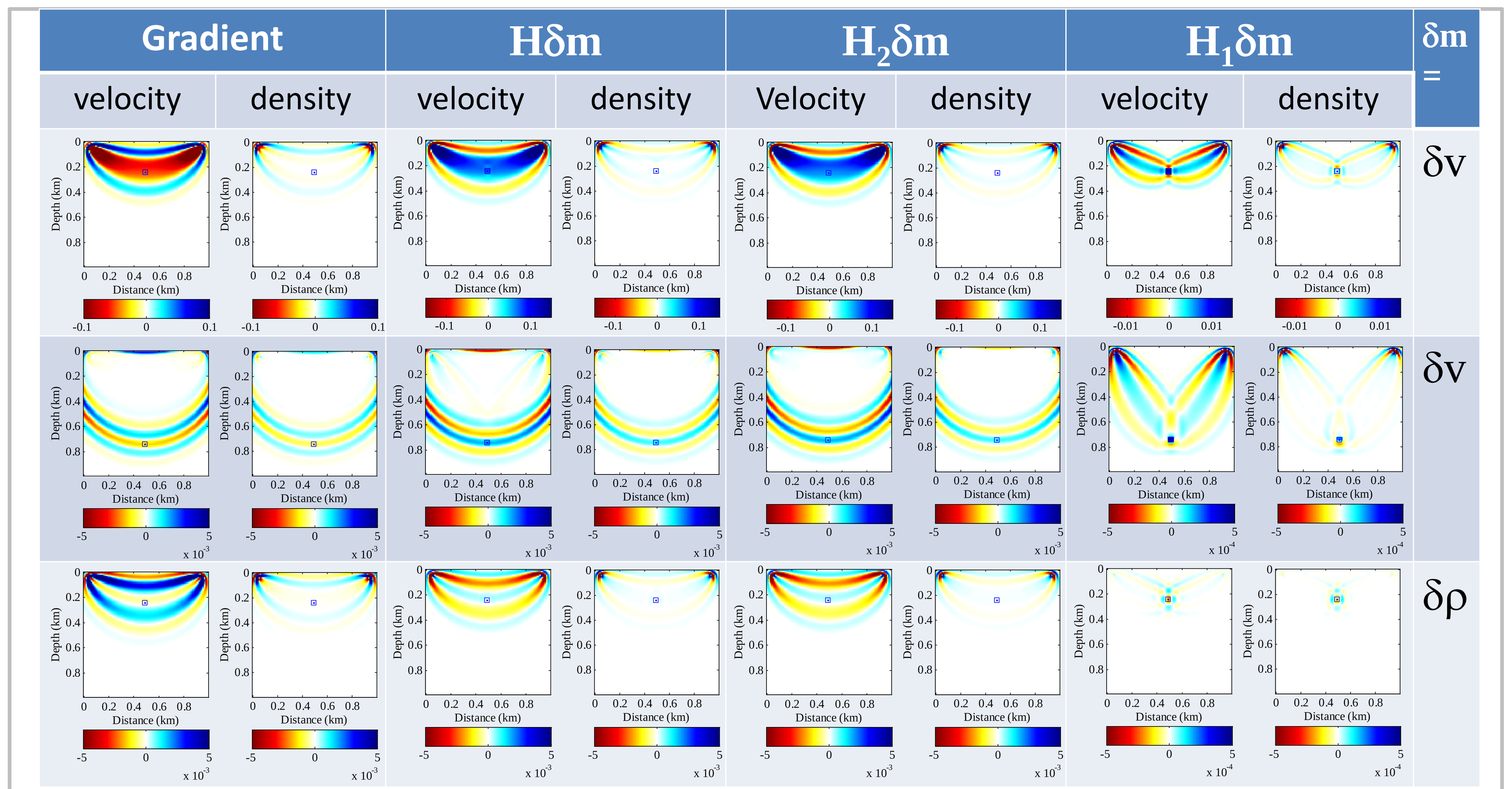
The scaler μ_j can be found by solving

$$\langle\mathbf{g},\mathbf{a}_j\rangle + \sum_{i=1}^{n}\mu_j\langle\mathbf{H}\mathbf{a}_j,\mathbf{a}_i\rangle \approx 0$$

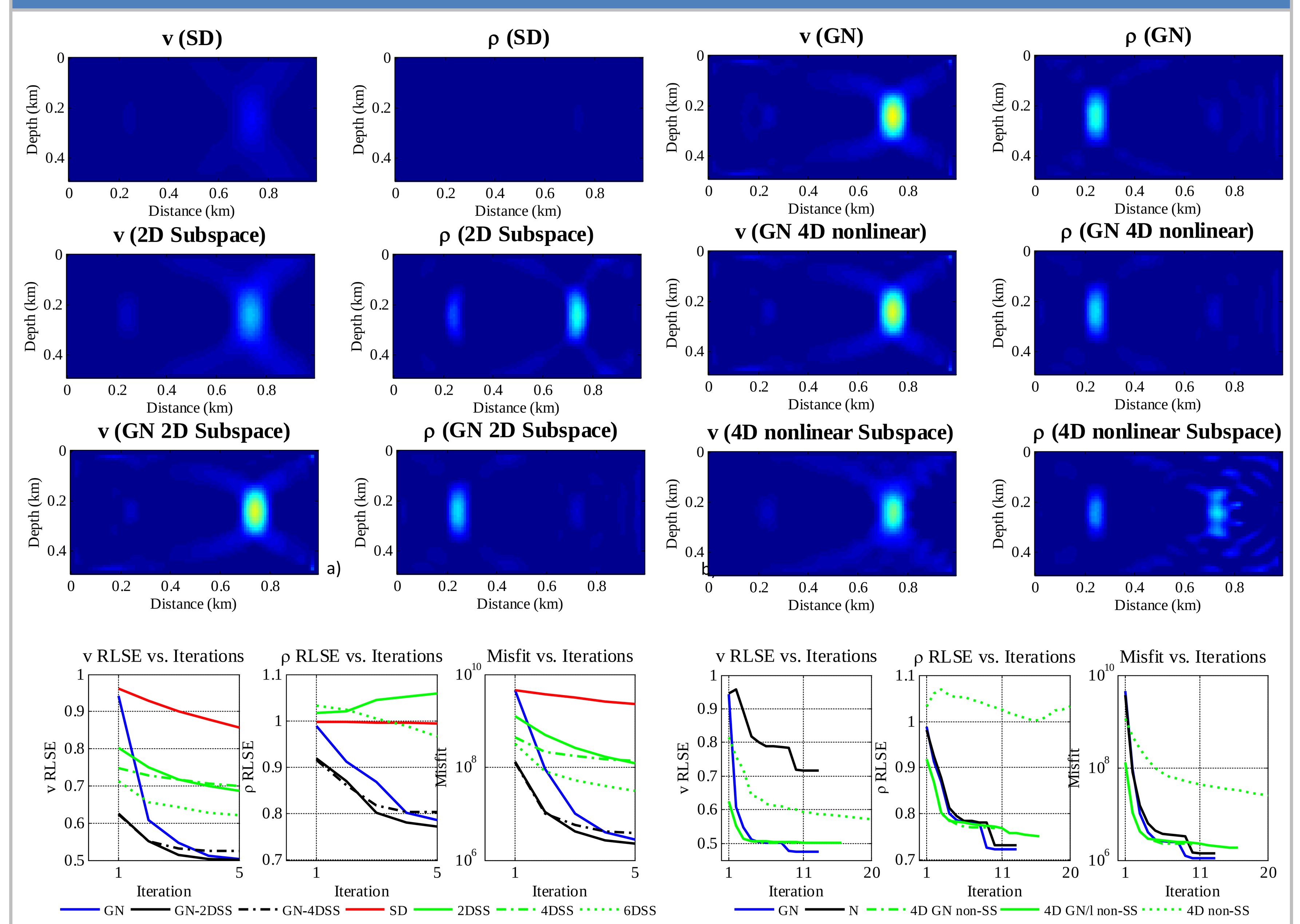
Or in matrix form

$$\boldsymbol{\mu} = -(\mathbf{A}^T\mathbf{H}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{g}$$

Explicitly calculating both terms of the Hessian operator in this proposed method is still not possible for the large size of FWI problem, especially in the multiparameter case. However, the calculation of a product between the Hessian $\mathbf{H}_1(\mathbf{m})$ and a vector is needed to first calculate the second-order scattering related perturbation, and the inverse of $\mathbf{H}_2(\mathbf{m})$ can be added to the perturbation iteratively by a truncated Gauss-Newton method, with the help of a Gauss-Newton Hessian-vector product.



Subspace method combined with nonlinear updates



Conclusion

In this study, we first derive a nonlinear descent direction from the Newton method to perform a multiparameter FWI, which alters the gradient with the second-order scattering, and can be used to predict the model perturbation using a truncated Gauss-Newton method. We then use this nonlinear descent direction as a set of basis vectors to construct a subspace, in which different step lengths can be obtained for different parameter class, as well as linear/nonlinear perturbations for each parameter class. Gauss-Newton Hessian product with a vector is involved to find the local minimum in the spanned space. The behavior of the subspace methods for both linear updates and nonlinear updates are compared with traditional FWI methods. The subspace methods have better convergence rate, as well as better reconstruction of the velocity model. The reconstruction of density model, however, could still be effected by the cross-talk artifacts, when Hessian is not considered in the inversion.

We thank the sponsors of CREWES for support. This work was funded by CREWES and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 379744-08.

CREWES www.crewes.org NSERC CRSNG UNIVERSITY OF CALGARY FACULTY OF SCIENCE Department of Geoscience