Released to public domain under Creative Commons license type BY (https://creativecommons.org/licenses/by/4.0).
Copyright (c) 2013 SLIM group @ The University of British Columbia.

Wave-equation based multi-parameter linearized inversion with joint-sparsity promotion

March 8, 2013

Abstract

The successful application of linearized inversion is affected by the prohibitive size of the data, computational resources required, and how accurately the model parameters reflects the real Earth properties. The issue of data size and computational resources can be addressed by combining ideas from sparsity promoting and stochastic optimization, which can allow us to invert model perturbation with a small subset of the data, yielding a few PDE solves for the inversion.

In this abstract, we are aiming at addressing the issue of accuracy of model parameters by inverting density and velocity simultaneously rather than only using velocity. As a matter of face, the effects of density and velocity variations towards the wavefield are very similar, which will cause energy leakage between density and velocity images. To overcome this issue, we proposed a incoherence enhanced method that can reduce the similarity between the effect of density and velocity. Moreover, the location of structural variations in velocity and density are often overlapped in geological setting, thus in this abstract, we also exploit this property with joint-sparsity promoting to further improve the imaging result.

Introduction

Wave-equation based linearized inversion is a procedure in which we minimize the least-square misfit as a function of the model perturbations. Within a defined tolerance level, we expect the misfit to decrease resulting in a model perturbation revealing subsurface geological structures. However, the successful application of linearized inversion is affected by the prohibitive size of the data, computational resources required, and how accurately the model reflects the real Earth properties.

In (Herrmann and Li, 2012), we addressed the issues of data size and computational resource requirements, where we proposed a sparsity-promoting algorithm using insights from the fields of compressive sensing and stochastic optimization to invert model perturbation, with a small random selected subset of the data. This approach can significantly reduce the number of PDE solves required for inversion, hence, we are able to compute a linearized inversion with computational cost roughly equal to computing a single RTM with *all* the data. Moreover, Curvelet domain sparsity promotion can help us suppress the source cross-talk interference caused by random subsampling yielding a better resolution for the result.

According to our previous experience (Herrmann and Li, 2012; Li et al., 2012), reasonable results can be extracted from the data by using velocity-only acoustic modeling, however, velocity modeling alone is insufficient to explain the data, which may be more accurately explained when combined with the density information.

In addition the location of structural variations in velocity model and density model usually overlap in geological settings. Thus, it allows us to exploit this property by joint-sparsity promotion. This approach helps improving the resolution according to the results enclosed in this abstract.

Theory

Joint-sparse recovery: In (van den Berg and Friedlander, 2009), a multiple-measurement-vector (MMV) problem can be formed as

$$\underset{\mathbf{X}}{\text{minimize}}\, \|\mathbf{X}\|_{p,q} \quad \text{subject to} \quad \mathscr{A}(\mathbf{X}) = \mathbf{b}, \tag{1}$$

where $\mathscr{A}$ is a sampling operator acting on the columns of $\mathbf{X}$ organized into a long vector. The vector $\mathbf{b}$ is the measurement. The $\ell_{p,q}$ norm of $\mathbf{X}$ is defined as $\|\mathbf{X}\|_{p,q} = (\sum_{j=1}^{n} \|\mathbf{X}_{j,:}\|_q^p)^{1/p}$, in which $\mathbf{X}_{j,:}$ is j^{th} row of $\mathbf{X}$. According to van den Berg and Friedlander (2009), joint sparsity ($\ell_{1,2}$ norm) can provides better recovery compared to sparsity promotion on $\mathbf{X}$ organized as a long vector (this corresponds to the $\ell_{1,1}$ norm), when $\mathbf{X}$ has nonzero entries in only a small number of rows (row-sparse). To test the improved recovery by joint sparsity promotion, we sample a sparse 512×2 matrix $\mathbf{X}$ with a sampling operator $\mathscr{A}$, which contains two random 120×512 Gaussian matrices. There are only 20 non-zeros in each column of $\mathbf{X}$, 16 non-zeros in these two columns are at the same location. We generate the measurement vector $\mathbf{b}$ by $\mathbf{b} = \mathscr{A}(\mathbf{X}) + \mathbf{e}$, where $\mathbf{e}$ is a random noise vector. From the recovered results in Figure 1, we can easily sense the advantage of joint-sparsity promotion when $\mathbf{X}$ is row-sparse.

Motivated by the observation in previous example (equation 1), we are able to exploit the coincidence of

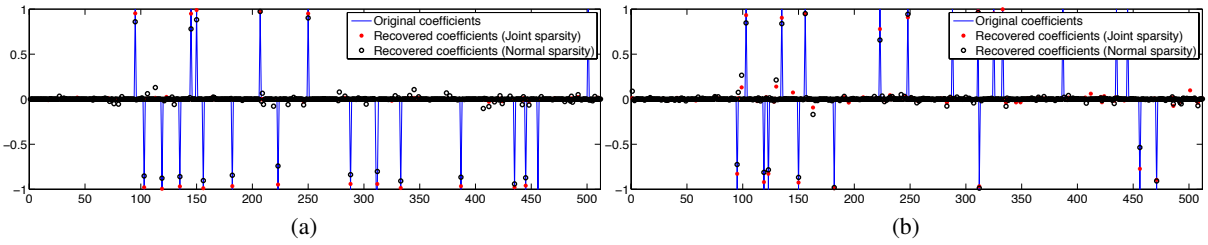

Figure 1: Multiple measurement vector recovery. **(a:)** First row of $\mathbf{X}$. **(b:)** Second row of $\mathbf{X}$.

density and velocity changes in relation to Geological structural or stratigraphic variation. This similarity also yields similar supports for sparsifying transform domain coefficients of these parameters as shown in figure 2. Figure 2a and Figure 2b are a synthetic geological model generated with constraint from real

well-log information. Figure 2c and Figure 2d are Curvelet synthesis of Figure 2a and Figure 2b. We observed from the experiment that about 93% non-zeros in each coefficients are overlapped. This can allow us to invert density and velocity together with joint-sparsity promotion.

Linearized inversion with joint-sparsity promotion: Similar to our earlier work (Herrmann and

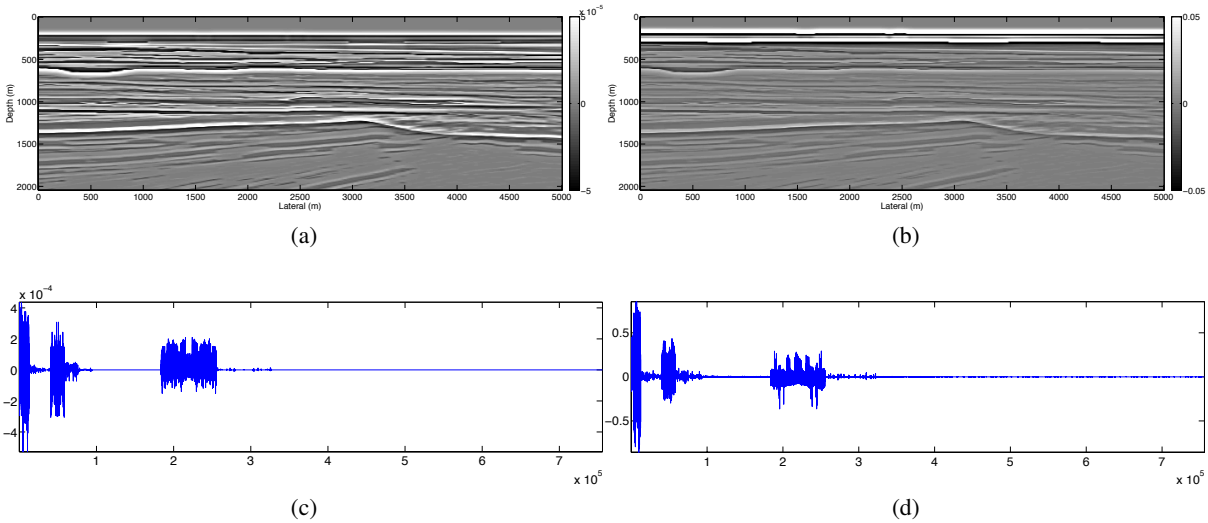

Figure 2: Curvelet domain sparsity. **(a:)** True velocity perturbation. **(b:)** True density perturbation. **(c:)** Curvelet synthesis of **a**. **(d:)** Curvelet synthesis of **b**.

Li, 2012), linearized inversion for velocity and density with joint-sparsity promotion can be formed as follows,

$$\underset{\mathbf{X}}{\text{minimize}} \|\mathbf{X}\|_{p,q} \quad \text{subject to} \quad \|\delta\mathbf{d} - \mathscr{A}(\mathbf{X})\|_2 \leq \sigma, \tag{2}$$

where the vector $\delta\mathbf{d}$ is the data residual, while the matrix $\mathbf{X}$ has two columns that contain the vectorized density and velocity perturbations ($\delta\mathbf{v}$ and $\delta\rho$ throughout the abstract). The action of the operator $\mathscr{A}$ on the matrix $\mathbf{X}$ is given by $\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]\delta\mathbf{v} + \nabla\mathscr{F}[\rho,\mathbf{q}]\delta\rho$. The operator $\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]$ represents the Jacobian of the velocity, while $\nabla\mathscr{F}[\rho,\mathbf{q}]$ is the Jacobian of the density.

To test the properties of those two Jacobians, we carried out a simple experiment shown in Figure 3 with a constant background model. In this case, we migrate one single event (center image in Figure 3), which is caused by a spike in density perturbation (right bottom image in Figure 3). The left images in Figure 3 are the reverse-time migration (RTM) results of velocity (top) and density (bottom), which are given by the normal equation

$$-\begin{bmatrix}\mathbf{g_v}\\ \mathbf{g}_\rho\end{bmatrix} = \mathbf{H}\begin{bmatrix}\delta\mathbf{v}\\ \delta\rho\end{bmatrix}, \tag{3}$$

where $\mathbf{H}$ is the Gauss-Newton Hessian. We can easily observe that these two RTM results are numerically similar to each other eventhough there is no perturbation in the velocity. From this test we learn that these two Jacobians are highly coherent, which means they have similar actions on wavefield.

To understand the consequence of the coherence between these two Jacobians, we use two coherent Gaussian matrices ($\mathbf{G_v}$ and $\mathbf{G}_\rho$, while $\mathbf{G}_\rho = \mathbf{G_v} + \delta\mathbf{G}$ and $\delta\mathbf{G}$ is a small perturbation) to represent Jacobians of density and velocity. The corresponding Hessian is shown in Figure 4a, which has three dominant diagonals due to the similarity in $\mathbf{G_v}$ and $\mathbf{G}_\rho$. The effect of this Hessian is energy leakage between the two model parameters as shown in Figure 4a, in which we have only two spikes in density and velocity but 4 spikes in the gradient, two of them are energy leakage which correspond to the two off-center diagonals in the Hessian matrix. This leakage will easily lead to ill-condition of $\mathscr{A}$ if we iterate the model with this gradient. To overcome this issue, we need to come up with a method to reduce the coherence between these two Jacobians.

Incoherence enhancement: From the observation in the previous section, the coherence of density and

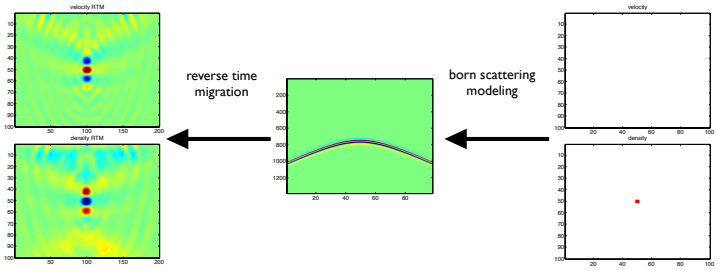

Figure 3: Gradient test. **(Top left:)** RTM of velocity. **(Bottom left:)** RTM of density. **(Center:)** Data residual. **(Top right:)** Velocity perturbation. **(Bottom right:)** Density perturbation.

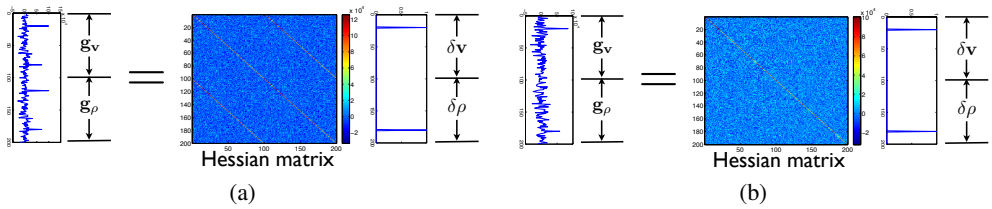

Figure 4: **(a)** Normal equation based on two random coherent Gaussian matrices. **(b)** Incoherence enhanced normal equation.

velocity Jacobians is caused by the common part between them. In order to reduce the coherence, we can simply subtract these two Jacobians to get rid of the common part. Consequently, the combination of model parameters also needs to be modified to maintain the system in equation 4. Notice that $\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]\delta\mathbf{v}+\nabla\mathscr{F}[\rho,\mathbf{q}]\delta\rho$ equals to $(\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]-\nabla\mathscr{F}[\rho,\mathbf{q}])\delta\mathbf{v}+\nabla\mathscr{F}[\rho,\mathbf{q}](\delta\rho+\delta\mathbf{v})$, hence we can invert for $\delta\mathbf{v}$ and $\delta\mathbf{z}$ by solving

$$\text{minimize}\,\|\mathbf{X}\|_{p,q} \quad \text{subject to} \quad \|\delta\mathbf{d}-\mathscr{A}(\mathbf{X})\|_2 \leq \sigma, \tag{4}$$

where we redefined $\mathscr{A}$ and $\mathbf{X}$. The vector $\delta\mathbf{z}$ represents $\frac{1}{\beta}(\delta\rho+\delta\mathbf{v})$ (van Wijngaarden, 1998; Hennenfent and Herrmann, 2004), while β is a number given by $\frac{\|(\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]-\nabla\mathscr{F}[\rho,\mathbf{q}])\mathbf{x}\|_2}{\|\nabla\mathscr{F}[\rho,\mathbf{q}]\mathbf{x}\|_2}$ that can approximately scale the nuclear norm of $(\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]-\nabla\mathscr{F}[\rho,\mathbf{q}])$ and $\nabla\mathscr{F}[\rho,\mathbf{q}]$, $\mathbf{x}$ is a random Gaussian vector. Columns of $\mathbf{X}$ now represent $\delta\mathbf{v}$ and $\delta\mathbf{z}$. The action of the new operator $\mathscr{A}$ on $\mathbf{X}$ is given by $(\nabla\mathscr{F}[\mathbf{v},\mathbf{q}]-\nabla\mathscr{F}[\rho,\mathbf{q}])\delta\mathbf{v}+(\beta\nabla\mathscr{F}[\rho,\mathbf{q}])\frac{1}{\beta}(\delta\rho+\delta\mathbf{v})$. With this linear combination we can suppress the two off-center Hessian diagonals shown in Figure 4a, yielding a new Hessian with only one dominant diagonal as shown in Figure 4b. With the new Hessian, the energy leakage is suppressed as depicted in the Figure 4b. In this example, the gradient of two spikes are still two spikes but with random interference, this random interference can be easily suppressed with sparse regularized method (Herrmann and Li, 2012).

Examples

To demonstrate the performance of incoherence enhanced linearized inversion with joint sparsity promotion. We generate data with a synthetic model as shown in Figure 2a and Figure 2b which is constrained by real well log information. For simplicity, we simulated 251 shots with interval 20 meters by using linearized born scattering modeling. All shots share the same 330 receivers placed with interval of 15 meters. We neglect surface-related multiple by using an absorbing boundary condition at the surface. Inversion is carried out simultaneously with 8 randomly selected frequencies from 20 to 30 Hz and 20 randomly selected simultaneous shots. In which we use 100 iterations of linear solver (SPGℓ_1) (Berg and Friedlander, 2008).

For comparison, We also carried out an experiment with the same setting but using the normal linearized inversion (equation 4), whose results are depicted in Figure 5a and Figure 5b. As we reported in the text, the coherence of Jacobian operators may cause energy leakage and ill-conditioning, that is the rea-

son why we are not able to invert the result with correct amplitude in Figure 5a and Figure 5b. On the other hand, Figure 5c and Figure 5d illustrate the behavior of our new method. With a limited amount of computation cost (roughly equivalent to 1/12 full linearized inversion problem), the new method is able to capture most structures in the original density and velocity model with the right amplitude. In addition, we observe turning waves caused by the increasing velocity in Figure 5a and Figure 5c. Based on the findings in this experiment, it is easy to scatter this process to the whole FWI problem.

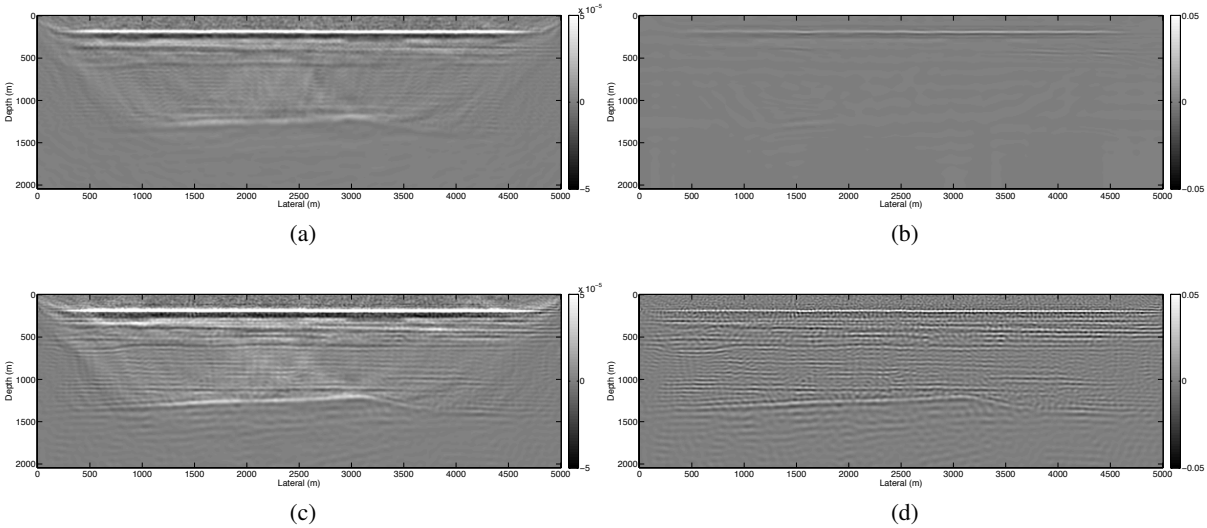

Figure 5: Recovery results. **(a-b)** Inverted velocity and density through normal Guass-Newton. **(c-d)** Inverted velocity and density through incoherence enhanced method with joint-sparsity.

Conclusions

In this abstract, we introduced a new method to break the coherence between Jacobian of density and Jacobian of velocity. By doing this, we are able to get rid of off-center diagonals of the Hessian for linearized inversion to avoid energy leakage between inversion result of velocity and density. Moreover, joint-sparsity approach further improved the quality of the final result when density and velocity share similar sparsity support.

Acknowledgements

We would like to thank the sponsors of the SINBAD project. This work was partly supported by a NSERC Discovery Grant (22R81254) and CRD Grant DNOISE II (375142-08).

References

Berg, E.V.D. and Friedlander, M.P. [2008] Probing the Pareto frontier for basis pursuit solutions. *Siam J. Sci. Comput.*, **31**(2), 890–912.

Hennenfent, G. and Herrmann, F.J. [2004] Three-term amplitude-versus-offset (avo) inversion revisited by curvelet and wavelet transform. *Expanded Abstracts*, Soc. Expl. Geophys., Tulsa.

Herrmann, F.J. and Li, X. [2012] Efficient least-squares imaging with sparsity promotion and compressive sensing. *Geophysical Prospecting*, **60**(4), 696–712.

Li, X., Aravkin, A.Y., van Leeuwen, T. and Herrmann, F.J. [2012] Fast randomized full-waveform inversion with compressive sensing. *Geophysics*, **77**(3), A13–A17.

van den Berg, E. and Friedlander, M.P. [2009] Joint-sparse recovery from multiple measurements. Tech. rep., Department of Computer Science.

van Wijngaarden, A. [1998] *Imaging and characterization of angle-dependend seismic reflection data*. Ph.D. thesis, Delft University of Technology.