EAGE

47028

Neural Network Least Squares Migration

Zhaolun Liu* (King Abdullah University of Science and Technology), Gerard Schuster (King Abdullah University of Science and Technology)

Summary

Sparse least squares migration (SLSM) estimates the reflectivity distribution that honors a sparsity condition. This problem can be reformulated by finding both the sparse coefficients and basics functions from the data to predict the migration image. This is designated as neural network least squares migration (NLSM), which is a more general formulation of SLSM. This reformulation opens up new thinking for improving SLSM by adapting ideas from the machine learning community.

Introduction

The sparse least squares problem (SLSM) is defined as finding the reflectivity coefficients m_i in the $N \times 1$ vector m that minimize the objective function ε :

$$\varepsilon = \frac{1}{2}||\Gamma \mathbf{m} - \mathbf{m}^{mig}||_2^2 + \lambda S(\mathbf{m}), \tag{1}$$

where $\Gamma = \mathbf{L}^T\mathbf{L}$ represents the migration Green's function (Schuster and Hu, 2000), $\lambda > 0$ is a positive scalar, $\mathbf{m}^{mig} = \mathbf{L}^T\mathbf{d}$ is the migration image obtained by migrating the recorded data $\mathbf{d}$ with the migration operator $\mathbf{L}^T$, and $S(\mathbf{m})$ is a sparseness function. For example, the sparseness function might be $S(\mathbf{m}) = ||\mathbf{m}||_1$ or $S(\mathbf{m}) = log(1 + ||\mathbf{m}||_2^2)$.

The solution to equation 1 is

$$\mathbf{m}^* = \arg\min_{\mathbf{m}} \left[\frac{1}{2}||\Gamma \mathbf{m} - \mathbf{m}^{mig}||_2^2 + \lambda S(\mathbf{m})\right], \tag{2}$$

which can be approximated by an iterative gradient descent method:

$$\begin{aligned} m_i^{(k+1)} &= m_i^{(k)} - \alpha[\Gamma^T \overbrace{(\Gamma \mathbf{m} - \mathbf{m}^{mig})}^{\mathbf{r}=residual}]_i - \lambda S(\mathbf{m})'_i, \\ &= m_i^{(k)} - \alpha[\Gamma^T \mathbf{r}]_i - \lambda S(\mathbf{m})'_i. \end{aligned} \tag{3}$$

Here, $S(\mathbf{m})'_i$ is the derivative of the sparseness function with respect to the model parameter m_i and the step length is α. Vectors (matrices) are denoted by boldface lowercase (uppercase) letters.

As shown in Schuster and Hu (2000), the poststack migration (Yilmaz, 2001) image $m(\mathbf{x})^{mig}$ in the frequency domain is computed by weighting each reflectivity value $m(\mathbf{z})$ by $\Gamma(\mathbf{x}|\mathbf{z})$ and integrating over the model-space coordinates $\mathbf{z}$:

$$\text{for } \mathbf{x}, \mathbf{z} \in D_{model}: \quad m(\mathbf{x})^{mig} = \int_{D_{model}} d\mathbf{z} \overbrace{\int_{\mathbf{y} \in D_{data}} d\mathbf{y} \omega^4 G(\mathbf{x}|\mathbf{y})^{2*} G(\mathbf{y}|\mathbf{z})^2}^{\Gamma(\mathbf{x}|\mathbf{z})} \mathbf{m}(\mathbf{z}), \tag{4}$$

where the migration Green's function $\Gamma(\mathbf{x}|\mathbf{z})$ is given by

$$\text{for } \mathbf{x}, \mathbf{z} \in D_{model}: \quad \Gamma(\mathbf{x}|\mathbf{z}) = \int_{\mathbf{y} \in D_{data}} d\mathbf{y} \omega^4 G(\mathbf{x}|\mathbf{y})^{2*} G(\mathbf{y}|\mathbf{z})^2. \tag{5}$$

Here we implicitly assume a normalized source wavelet in the frequency domain, and D_{model} and D_{data} represent the sets of coordinates in, respectively, the model and data spaces. The term $G(\mathbf{x}'|\mathbf{x}) = e^{i\omega\tau_{xx'}}/||\mathbf{x} - \mathbf{x}'||$ is the Green's function for a source at $\mathbf{x}$ and a receiver at $\mathbf{x}'$ in a smoothly varying medium[1]. The traveltime $\tau_{xx'}$ is for a direct arrival to propagate from $\mathbf{x}$ to $\mathbf{x}'$.

The physical interpretation of the kernel $\Gamma(\mathbf{x}'|\mathbf{x})$ is that it is the migration operator's[2] response at $\mathbf{x}'$ to a point scatterer at $\mathbf{x}$, otherwise known as the MGF or the migration Green's function (Schuster and Hu, 2000). It is analogous to the point spread function (PSF) of an optical lens for a point light source at $\mathbf{x}$ in front of the lens and its optical image at $\mathbf{x}'$ behind the lens on the image plane. In discrete form, the modeling term $[\Gamma \mathbf{m}]_i$ in equation 4 can be expressed as

$$[\Gamma \mathbf{m}]_i = \sum_j \Gamma(\mathbf{x}_i|\mathbf{z}_j) m_j. \tag{6}$$

[1] If the source and receiver are coincident at $\mathbf{x}$ then the zero-offset trace is represented by the squared Green's function $G(\mathbf{x}|\mathbf{x}')^2$.

[2] This assumes that the zero-offset trace is generated with an impulsive point source with a smoothly varying background velocity model, and then migrated by a poststack migration operation. It is always assumed that the direct arrival is muted and there are no multiples.

with the physical interpretation that $[\Gamma\mathbf{m}]_i$ is the migration Green's function response at $\mathbf{x}_i$. An alternative interpretation is that $[\Gamma\mathbf{m}]_i$ is the weighted sum of basis functions $\Gamma(\mathbf{x}_i|\mathbf{z}_j)$ where the weights are the reflection coefficients m_j and the summation is over the j index. We will now consider this last interpretation and redefine the problem as finding both the weights m_j and the basis functions $\Gamma(\mathbf{x}_i|\mathbf{z}_j)$. This will be shown to be equivalent to the problem of a fully connected (FC) neural network.

Theory of Neural Network LSM

The neural network least squares migration (NNLSM) algorithm in the image domain is defined as solving for *both* the basis functions $\tilde{\Gamma}(\mathbf{x}_i|\mathbf{x}_j)$ and $\tilde{m}_j$ that minimize the objective function defined in equation 1. In contrast, SLSM only finds the least squares migration image in the image domain and uses the pre-computed migration Green's functions that solve the wave equation.

The NNLSM solution is defined as

$$(\tilde{\mathbf{m}}^*, \tilde{\Gamma}^*) \quad = \quad \arg\min_{\tilde{\mathbf{m}},\, \tilde{\Gamma}} \left[\frac{1}{2}||\tilde{\Gamma}\tilde{\mathbf{m}} - \mathbf{m}^{mig}||_2^2 + \lambda S(\tilde{\mathbf{m}})\right], \tag{7}$$

where now both $\tilde{\Gamma}$ and $\tilde{\mathbf{m}}$ are to be found. The functions with tilde's are mathematical constructs that are not necessarily identical to those based on the physics of wave propagation in equation 1.

The explicit matrix-vector form of the objective function in equation 7 is given by

$$\varepsilon \quad = \quad \frac{1}{2}\sum_i \left[\sum_j \tilde{\Gamma}(\mathbf{x}_i|\mathbf{z}_j)\tilde{m}_j - m_i^{mig}\right]^2 + \lambda S(\tilde{\mathbf{m}}). \tag{8}$$

and its Fréchet derivative with respect to $\tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_{j'})$ is given by

$$\frac{\partial \varepsilon}{\partial \tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_{j'})} \quad = \quad \sum_j (\tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_j)\tilde{m}_j - \tilde{m}_{i'}^{mig})\tilde{m}_{j'}. \tag{9}$$

The iterative solution of equation 7 is given in two steps (Olshausen and Field, 1996).

1. Iteratively estimate $\tilde{m}_i$ by the gradient descent formula used with SLSM:

$$\tilde{m}_i^{(k+1)} \quad = \quad \tilde{m}_i^{(k)} - \alpha[\tilde{\Gamma}^T(\tilde{\Gamma}\tilde{\mathbf{m}} - \mathbf{m}^{mig})]_i - \lambda S(\tilde{\mathbf{m}})'_i. \tag{10}$$

   However, one migration image $\mathbf{m}^{mig}$ is insufficient to find so many unknowns. In this case the original migration image is broken up into many small pieces so that there are many migration images to form examples from a large training set. For prestack migration, there will be many examples of prestack migration images, one for each shot.

2. Update the basis functions $\tilde{\Gamma}(\mathbf{x}_i|\mathbf{z}_j)$ by inserting equation 9 into the gradient descent formula to get

$$\begin{aligned} \tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_{j'})^{(k+1)} &= \tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_{j'})^{(k+1)} - \alpha\frac{\partial \varepsilon}{\partial \tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_{j'})}, \\ &= \tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_{j'})^{(k+1)} - \alpha(\left\{\sum_j \tilde{\Gamma}(\mathbf{x}_{i'}|\mathbf{z}_j)\tilde{m}_j\right\} - m_{i'}^{mig})\tilde{m}_{j'}. \end{aligned} \tag{11}$$

   It is tempting to think of $\tilde{\Gamma}(\mathbf{x}|\mathbf{x}')$ as the migration Green's function and $\tilde{m}_i$ as the component of reflectivity. However, there is yet no justification to submit to this temptation and so we must consider, unlike in the SLSM algorithm, that $\tilde{\Gamma}(\mathbf{x}|\mathbf{x}')$ is a sparse basis function and $\tilde{m}_i$ is its coefficient. To get the true reflectivity then we should equate equation 6 to $\sum_j \tilde{\Gamma}(\mathbf{x}_i, \mathbf{x}_j)\tilde{m}_j$ and solve for m_j.

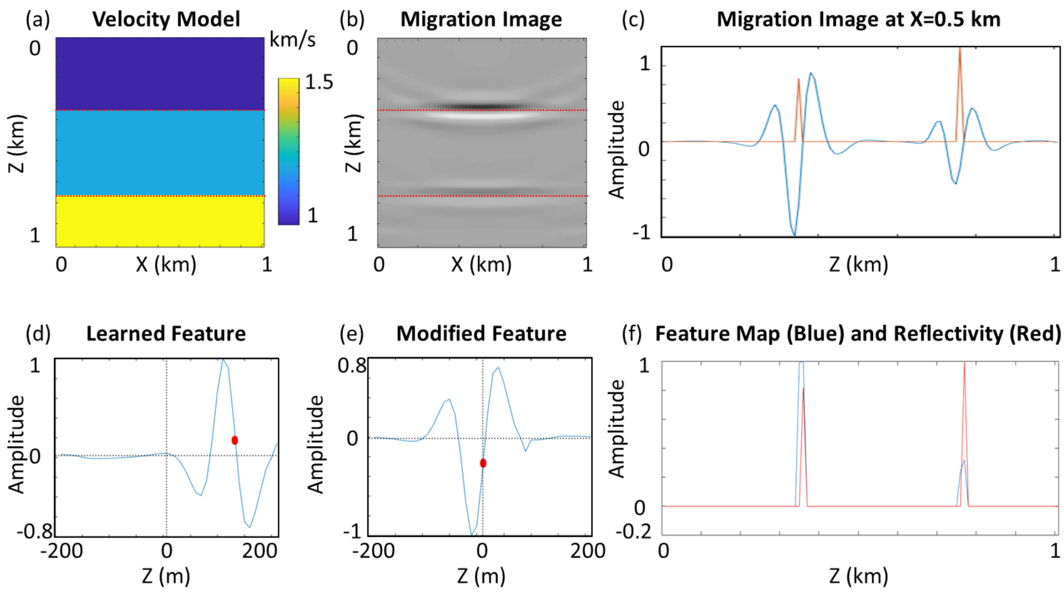

***Figure 1** a) Three-layer velocity model, b) RTM image, and c) migration image (blue curve) at X=0.5 km, where the red curve is the normalized reflectivity model. d) Basis function (also known as a feature), e) modified feature, and f) calculated coefficients (also known as a feature map) $\tilde{\mathrm{m}}$(blue) and reflectivity model (m).*

Numerical Examples

We will test the feasibility of estimating the basis functions and coefficients for NNLSM in the image domain. No particular application is addressed, except the success of the following tests should point the way towards using unsupervised learning methods for extracting features that can be inverted by the wave equation (Schuster, 2018).

The three-layer velocity model shown in Figure 1a is used to test NNLSM by the unsupervised feature learning method of convolutional sparse coding (Liu et al., 2018). The grid size of the model is 101 by 101. The grid interval is 10 m in both the x and z directions. There are 26 shots evenly spaced at a distance of 40 m on the surface. Each shot is recorded by 101 receivers with a sampling interval of 10 m. Figure 1b shows the reverse time migration (RTM) image.

The first test is for a 1D model where we extract the image located at X= 0.5 km, which is displayed as the blue curve in Figure 1c. The red curve in Figure 1c is the reflectivity model. Assume that there is only one basis function (feature) $\tilde{\Gamma}$ and it extends over the depth of 400 m (41 grid points). Equation 7 is solved for both $\tilde{\mathrm{m}}$ and $\tilde{\Gamma}$ by the two-step iterative procedure denoted as the alternating descent method. The computed basis function is shown in Figure 1d where the phase of the basis function $\tilde{\Gamma}$ is nonzero. However, the "physical" Γ is zero phase according to equation 5. If we use a nonzero-phase basis function $\tilde{\Gamma}$ to calculate its coefficient vector $\tilde{\mathrm{m}}$, the phases of the coefficient vector $\tilde{\mathrm{m}}$(also known as the feature map) and the true reflectivity m will be different. So, we need to modify the phase and polarity of the basis function $\tilde{\Gamma}$. The modified basis function is shown in Figure 1e, and its coefficients are displayed as the blue curve in Figure 1e. Compared with the true reflectivity m (red curve in Figure 1f), the feature map can give the correct positions but wrong values of the reflectivity distribution.

Next, we perform a 2D test where the input is the 2D migration image in Figure 1b. Three 35-by-35 (grid point) features are learned (see Figure 2a), where a feature is also denoted as a basis function. The modified features are shown in Figure 2b. The feature maps of these three features are displayed in Figures 3a, 3b and 3c. Figure 3d shows the sum of these three feature maps. It is evident that the stacked feature maps can estimate the correct locations of the reflectivity spikes.

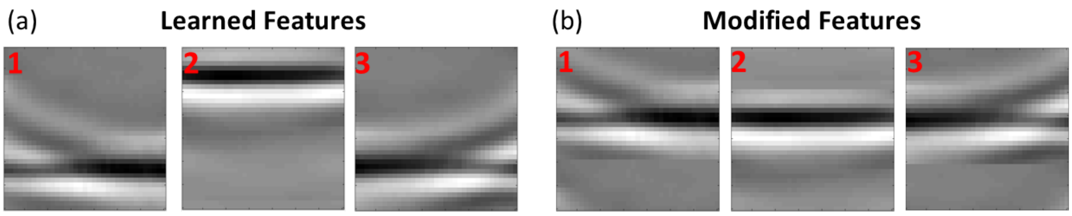

Figure 2 *a) Learned and b) modified features.*

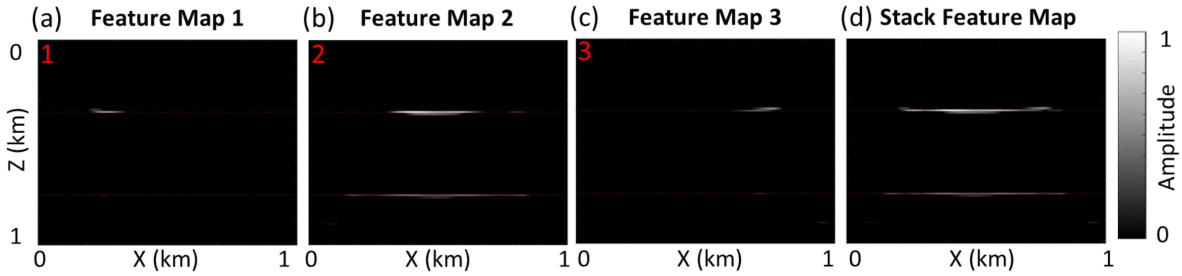

Figure 3 *Feature maps for the features a) 1, b) 2, and c) 3 shown in Figure 2. The stacked feature map is shown in d).*

Summary

Sparse least squares migration (SLSM) finds the optimal reflectivity distribution $m(\mathbf{x})$ that minimizes a sum of migration misfit and sparsity functions. This problem can be reformulated by finding both the sparse coefficients $\tilde{m}(\mathbf{z})$ and basis functions $\tilde{\Gamma}(\mathbf{x}|\mathbf{z})$ from the data to predict the migration image $m(\mathbf{x})^{mig} = \sum_z \tilde{\Gamma}(\mathbf{x}|\mathbf{z})\tilde{m}(\mathbf{z})$. This generalization of SLSM is designated as neural network least squares migration (NLSM). If the basis functions are required to be from the migration Green's functions then NLSM reduces to SLSM. One possibility is to invert the feature maps using the wave equation as described in Schuster (2018). NSLM provides new ideas from the technology of machine learning for improving least squares migration.

References

Liu, Z., Lu, K. and Ge, X. [2018] Convolutional Sparse Coding for Noise Attenuation of Seismic Data. In: *2018 SEG Beijing Workshop: Maximizing asset value through Artificial Intelligence and Machine Learning*.

Olshausen, B.A. and Field, D.J. [1996] Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, **381**(6583), 607.

Schuster, G.T. [2018] Machine learning and wave equation inversion of skeletonized data. In: *80th EAGE Conference and Exhibition 2018*.

Schuster, G.T. and Hu, J. [2000] Green's function for migration: Continuous recording geometry. *Geophysics*, **65**(1), 167–175.

Yilmaz, Ö. [2001] *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists.