Tu_R04_09

Neural Network Least Squares Migration

Z. Liu[1]*, G. Schuster[1]

[1] 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.

Sparse Least Squares Migration

The sparse least squares problem (SLSM) is defined as finding the reflectivity coefficients m_i in the $N \times 1$ vector $\mathbf{m}$ 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 $\mathbf{z}$:

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

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

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

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 (Hu and Schuster, 1998). 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 $\mathbf{x}_i$ for the reflectivity model $\mathbf{m}$. 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 LSM problem as finding both the weights m_j and the basis functions $\Gamma(\mathbf{x}_i|\mathbf{z}_j)$ that minimize the objective function in equation 2. This can be shown to be equivalent to the problem of a fully connected (FC) neural network (Schuster and Liu, 2019).

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

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

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

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

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

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

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:

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

   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, and the compressive sensing technique denoted as VISTA (Ahmad et al., 2015) is used for the calculations.

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

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

   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.

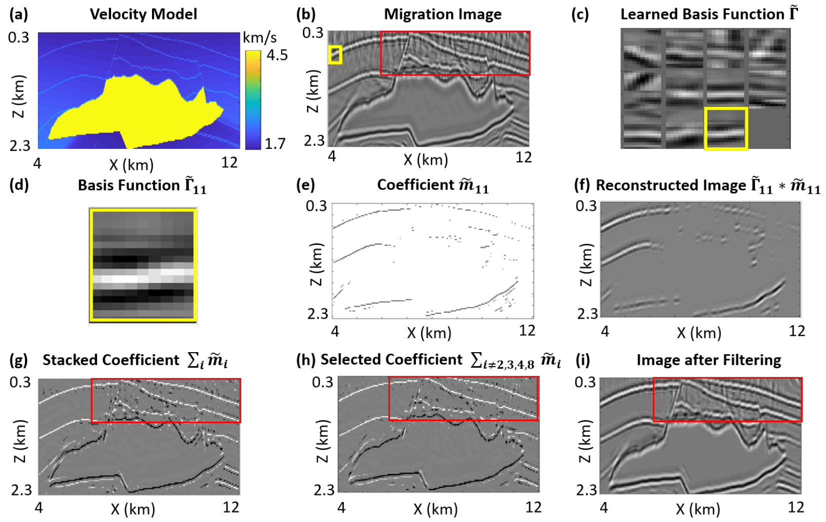

Figure 1 *a) 2D SEG/EAGE salt model, b) RTM image, and c) learned basis functions, d) basis function No. 11, e) coefficient No. 11, and f) reconstructed image by the eleventh basis function and its coefficient, g) stacked coefficients, h) stacked the selected coefficients, and i) migration image after filtering.*

Numerical Examples

We will test the feasibility of estimating the basis functions and coefficients for NNLSM in the image domain with synthetic and field data examples including data associated with 1) the 2D SEG/EAGE salt model and 2) the F3 block of data recorded offshore in the Netherlands.

The 2D SEG/EAGE salt 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 100 grid points in the z-direction and 401 grid points in the x-direction. The grid interval is 20 m in both the x and z directions. Figure 1b shows the reverse time migration (RTM) image. Assume that there are 11 basis functions $\tilde{\Gamma}$ and each of them has the size of 17 grid points in the z-direction (0.32 km, about one and a half wavelength) and nine grid points (0.16 km) in the x-direction marked as the yellow box in Figure 1b. Equation 7 is solved for both $\tilde{\mathbf{m}}$ and $\tilde{\Gamma}$ by the two-step iterative procedure denoted as the alternating descent method. The computed basis functions are shown in Figure 1c, where the eleventh basis function $\tilde{\Gamma}_{11}$ is marked by the yellow box which is displayed in Figure 1d in a zoom view. The coefficient $\tilde{\mathbf{m}}_{11}$ is shown in Figure 1e. The convolution result of the basis function $\tilde{\Gamma}_{11}$ and the coefficient $\tilde{\mathbf{m}}_{11}$ is displayed in Figure 1f. Here the basis function $\tilde{\Gamma}_{11}$ can be seen as the blurring filter which blurs the coefficient $\tilde{\mathbf{m}}_{11}$ to give the blurred image $\tilde{\Gamma}_{11} * \tilde{\mathbf{m}}_{11}$. If we stack all the coefficients, we can get the reflectivity distribution shown in Figure 1g. Because there are migration artifacts in the migration image as shown in the red box in Figure 1b, some of the basis functions may be learned from those artifacts, such as the basis function No. 2, 3, 4 and 8. Thus, we need to exclude their corresponding coefficients and stack the remaining coefficients to get the filtered reflectivity distribution as shown in Figure 1h. It is evident that there is less noise in the red box of Figure 1h after excluding the coefficients from the migration artifacts compared to that in Figure 1g. The reconstructed migration image after filtering is shown in Figure 1, and we can see that there are fewer migration artifacts in the red box.

Next, we apply NNLSM to the field data collected in block F3 in the Southern North Sea (Schroot and Scüttenhelm, 2003). The time migration image is shown in Figure 2a. The number of the time samplings is 213 grid points, and the grid size in the x-direction is 301 grid points. Twenty-one 13-by-5 (grid point) basis functions are learned (see Figure 2b). The size of the basis function is marked as the yellow box in Figure 2a and 2b. The stacked coefficients (reflectivity distribution) are displayed in Figure 2c. It is evident that the stacked coefficients can provide a high-resolution migration image. After reconstruction from the learned basis functions and coefficients, the migration image is shown in Figure 2d with less noise.

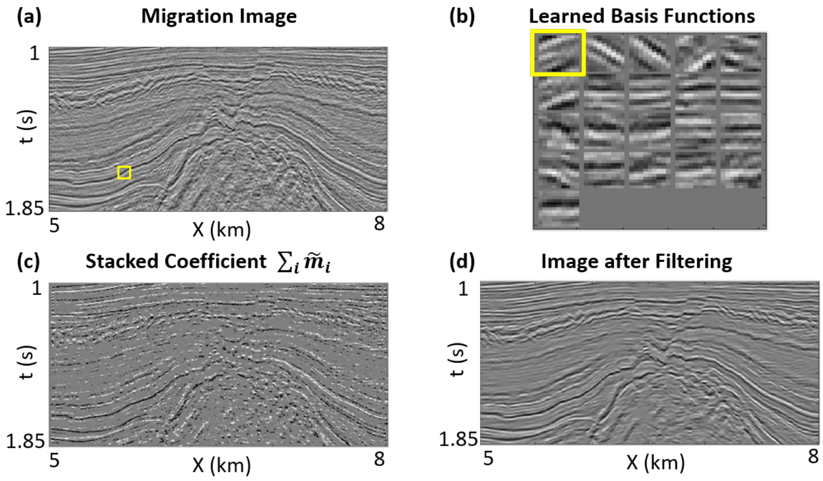

Figure 2 *a) Migration image computed from the F3 offshore block data, b) learned basis functions, c) stacked coefficients and d) migration image after filtering.*

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 the physics-based migration Green's functions then NLSM reduces to SLSM. As demonstrated in this paper, coherent artifacts in the migration image can be eliminated by excluding their associated basis functions in the filtered image. 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

Ahmad, R., Xue, H., Giri, S., Ding, Y., Craft, J. and Simonetti, O.P. [2015] Variable density incoherent spatiotemporal acquisition (VISTA) for highly accelerated cardiac MRI. *Magnetic resonance in medicine*, **74**(5), 1266–1278.

Hu, J. and Schuster, G.T. [1998] Migration deconvolution. *Mathematical Methods in Geophysical Imaging V*, **3453**, 118–125.

Liu, Z., Lu, K. and Ge, X. [2018] Convolutional sparse coding for noise attenuation of seismic data. In: *2018 Beijing SEG 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.

Schroot, B. and Scüttenhelm, R. [2003] Expressions of shallow gas in the Netherlands North Sea. *Netherlands Journal of Geosciences - Geologie en Mijnbouw*, **82**(1), 91–105.

Schuster, G. and Liu, Z. [2019] Geophysical sparse inversion=neural networks. *SEG Technical Program Expanded Abstracts 2019*, submitted.

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.