GeoConvention 2015
GEOSCIENCE
new horizons
Telus Convention Centre • 4-8 May 2015 • Calgary, AB Canada

Least Squares Kirchhoff Depth Migration: potentials, challenges and its relation to interpolation

Daniel Trad
CGG

Summary

Least Squares Migration (LSM), like interpolation, has the potential to address sampling issues on migrated images; it also generates images with better amplitudes than standard migration. Although both techniques share the same goal and often the same formulation, they differ in the nature of the forward model that is used to predict the data. While most interpolation algorithms use short linear events in windows, LSM is a powerful decomposition of data on physical basis functions, capable of using much more prior information than interpolation. However, in practice, the cost of the migration/modelling operator and its limited representation of reality limit its potential. Because of the size of a real scale production problem, we can only afford a handful of iterations to invert a large migration/modelling operator. Therefore, an early stop in the iterations could compromise the final result to the point of questioning the advantages of LSM over standard migration. In this abstract, I discuss some of these limitations and possible ways to overcome them. I also analyze similarities and differences with respect to seismic data interpolation. Finally, I present results using a Least Squares implementation of Kirchhoff Pre-Stack Depth Migration.

Introduction

It is well known that poor sampling and truncation in seismic acquisitions cause aliasing problems in migration. A common approach to address this issue is either to apply data interpolation before migration or to employ some weighting and filtering scheme during migration. Interpolation techniques use simple approximations of continuity of events on windows and sparsity of plane waves. They are relatively simple and computationally cheap, and create new traces containing geophysical information that is consistent with the acquired data. Migration algorithms benefit from these additional seismic data, which cancels aliasing effects induced by insufficient sampling of field acquisitions to produce more consistent Common Image Gathers (CIGs) and final stacks.

For almost two decades many researchers have proposed least squares migration (LSM) as an alternative solution to the incomplete acquisition problem (Nemeth et al. 1999, Kuehl and Sacchi, 2003). Although many examples from academia showed encouraging results, LSM has struggled to be adopted by the industry mainly because of its high computational cost in comparison with interpolation. Only recently has the industry become interested in using LSM (Dong et al, 2012, Zhang et al. 2013). In principle interpolation and LSM techniques rely on a similar concept, the minimization of the error between observed and predicted seismic data in a least squares sense. The main difference is that LSM attempts data prediction by using a physical model (the wave equation) whereas interpolation uses a non-physical model (localized plane-waves). Furthermore, interpolation mitigates poor sampling by removing aliasing in the model space (filtering). Conversely, the goal of LSM is to create a realistic image by inverting the migration operator. A second important difference is that interpolation divides data in small time-space windows which can be approximated by simpler models. The nature of migration algorithms does not allow one to divide the data into small windows, because seismic propagation involves large volumes of the earth. Finally interpolation uses linear operators (for example the Fourier transform) for which we know exactly how to recover proper amplitudes. LSM instead uses operators to predict seismic data whose amplitudes depend on complex physical phenomena like attenuation, scattering, velocity and anisotropy. These are modelled and estimated with some degree of inaccuracy. Since LSM methods rely on amplitude matching, this last factor proves to be a very challenging point for LSM compared to interpolation.

After these initial considerations we can wonder why it is desirable to use a more expensive and complicated procedure like LSM instead of interpolation. Two important reasons are:

- Although interpolation has become a mature technique and keeps improving, it is intrinsically a limited approach. Prior information has little participation in interpolation, and mostly comes in the form of normal moveout and static corrections. Events that have complicated shapes in data space can only partially be mapped into simple models. LSM, on the other hand, uses the velocity model, anisotropy information and horizons, as prior information. This feature is a blessing because it addresses complex structures, but it also a curse since it relies on the correctness of this prior information.
- Because interpolation works on windows, it is critical how the data are grouped. Most interpolations calculate transforms using midpoint locations as two of the four possible spatial coordinates. In complex environments the concept of midpoint is meaningless; even more for converted wave or ocean bottom surveys. In principle, it is possible to interpolate along acquisition coordinates, like shot and receiver positions instead of midpoints and offset vectors, but the line sampling intervals become so coarse as to violate the sampling principle for interpolation. Migration, on the other hand, works well on acquisition coordinates.

Least Squares Migration expectations

Although at first glance LSM appears able to mitigate the acquisition footprint and compensate the geometrical spreading in the image, it is important to realize that a proper inversion of the migration operator is never achieved for several reasons. LSM is generally halted early since each iteration cost is very high (equivalent to two migrations). Furthermore, a migration/modelling operator is never perfect, so differences between original and predicted data (residuals) are not necessarily the correct direction for inversion (to attain convergence). LSM may attempt to change the image to fit data events that cannot be predicted by the given operator. Finally, unmigrated data suffer from many alterations through preprocessing steps which can severely hamper the ability of a theoretically exact operator to predict observed seismic data.

Despite these negative factors, we can still expect some benefits from LSM. First, LSM iterations perform a deconvolution (Yu et. al, 2006) whose effect is to increase resolution. This effect becomes enhanced when attenuation and/or de-ghosting are taken into account inside the operator. Second, in integral techniques like Kirchhoff, by fitting the data LSM adds back into the image energy which has been removed by anti-alias filters. This also has a gap filling effect, since energy from far offsets, normally muted by antialiasing, contributes to illuminate the area under the gap. Finally, we expect a further improvement of amplitudes in all directions, since the inversion of the migration operator performs illumination compensation (although this is only partially achieved because of the small number of iterations).

On the other hand, it is not clear that we should expect LSM to eliminate sampling artefacts produced by acquisition holes. In fact there is nothing inherent in the data fitting procedure that would lead to elimination of artefacts produced by limited wave-front interference. Although sampling artefact mitigation is always advocated in the literature, it is not a consequence of data fitting, but of filters applied during the inversion. Nevertheless, regularization is legitimate and necessary to produce a sensible solution out of ill conditioned inverse problems. What is less understood is whether the same sampling artefact attenuation can be attained by a simple filtering of the common image gathers after standard migration. However, this last approach can easily distort the amplitude on the migrated image.

Method

Many variants of LSM are possible but in this work I implemented a similar approach to the one described by Nemeth et al., 1999. A Kirchhoff modelling/migration (or forward/adjoint) operator pair is included in a conjugate gradient scheme to minimize the following cost function:

$$J = \|\mathbf{W}_d\mathbf{d} - \mathbf{W}_d\mathbf{L}\mathbf{W}_m\mathbf{m}\|^2 + \lambda\|\mathbf{m}\|^2. \qquad (1)$$

In this equation, $\mathbf{d}$ is the input data, $\mathbf{m}$ is the migrated image, $\mathbf{L}$ is the Kirchhoff modeling operator, $\mathbf{W}_d$ is a data weight function to eliminate bad data from the system of equations, $\mathbf{W}_m$ is an image weighting function that emphasizes particular features of the image. A solution to (1) can be expressed as

$$\mathbf{m} = (\mathbf{W}_m^H \mathbf{L}^H \mathbf{W}_d^H \mathbf{W}_d \mathbf{L} \mathbf{W}_m + \lambda \mathbf{I})^{-1} \mathbf{W}_m^H \mathbf{L}^H \mathbf{W}_d^H \mathbf{d} \ , \qquad (2)$$

where superscript H means adjoint operator ($\mathbf{L}^H$ is the Kirchhoff pre-stack depth-migration operator). Equation 2 can be considered as a standard migration further modified by the de-convolution of the modeling operator, which is done iteratively. At each iteration, the algorithm maps back to the image space the part of the data that has not been predicted properly. Because the physics of the operator has limited accuracy, there is always a large portion of the residuals dominated by events that cannot be predicted. This limits considerably the convergence of the method. It is possible through the operator $\mathbf{W_d}$ to remove from the residuals the non-predictable part of the data. Similarly, it is possible to use the operator $\mathbf{W_m}$ as a filter (or chain of filters) that attenuates the undesirable part of the data, like sampling artefacts, for example through some smoothing across spatial dimensions. $\mathbf{W_m}$ can also be used to increase resolution by using a correlation with a weight function proportional to the amplitude of the data, in a way similar to other high resolution transforms. However, since in migration we are not looking for spiky models, an additional transformation is required to map the subsurface geology to a much sparser domain (see for example Herrmann and Li, 2012).

At this point it is useful to make a comparison between the interpolation and LSM equations. The LSM algorithm presented here is very similar to a 5D interpolation algorithm (Trad, 2009). In the interpolation framework, the filter $\mathbf{W_m}$ removes artifacts related to sampling, but is applied in a domain where low-amplitude artifacts have no physical significance, such as evanescent energy in the Fourier space. Conversely, a filter applied in the physical image space can be dangerous because the geological complexity of the subsurface is a-priori unknown. Interpolation and LSM have also different computational costs: solving the cost function (1) in interpolation, for example using the Fourier operator in small windows, allows hundreds of iterations. In LSM just a handful, usually less than 10 iterations, are practically feasible today.

Examples

Figure 1 shows some standard and LSM migration results using a visco-elastic synthetic data set (courtesy of Chevron), which mimics a seafloor acquisition using autonomous Ocean Bottom Nodes (OBN) Figure 1a shows the original velocity model. Figure 1b shows the migration obtained using a densely sampled dataset (100m separation between the nodes). We can notice some attenuation with depth due to the seismic absorption. Figure 1c shows the migration of a decimated version of the same dataset: one in three receivers are kept resulting in 300m spacing between nodes. In figure 1c-d, the velocity model has been heavily smoothed to make it more similar to the standard resolution that would be obtained from velocity analysis on a real data set: in both figures we observe a loss of resolution and also sampling artefacts. Figure 1d shows the LSM results after 20 iterations, using the heavily decimated dataset. Three effects are observed: first, amplitudes are enhanced at depth thanks to LSM intrinsic illumination compensation. Second, the image has a resolution comparable to (b) which was obtained with three times more data and a better velocity model. By fitting the data, LSM is forcing the image to be more precise and therefore it enhances the resolution. Third, the sampling artefacts were mitigated. On the other hand, there is no additional energy in the very shallow section between the nodes. The LSM data fitting cannot infill that energy; in fact, shallow smiles from missing nodes fit the observed data. To remove the migration smiles it is necessary to include additional data, either by acquisition interpolation or by applying some filtering in the model space.

This reasoning suggests a possible improvement of LSM by creating data during the iterative process. For this interpolation-migration approach, new nodes are predicted in empty locations and added during the inversion. In that fashion the method resembles the Projection Onto Convex Sets (POCS) interpolation method (Abma and Kabir, 2006). In both cases the amplitude of the predicted data is weaker than it should be, but unlike POCS, for the interpolation-migration approach the amplitudes cannot be enhanced by simple filtering in the model space. A more sophisticated approach like matching filtering between new data and original data in the neighborhood of the new data is required.

Figure 2a show common image gathers (CIGs) for the decimated migration (Figure 1c), and Figure 2b shows the LS CIGs (corresponding to Figure 1d). The gathers show the benefits of LSM. Far offsets have been populated with energy that is normally removed by the antialiasing filter in standard migration. The final stack image would not normally benefit much from that energy because far offset traces tend to be noisier and stretched and therefore muted. However velocity analysis, AVO and AVAz could benefit from this extra information.

Conclusions

LSM is a powerful technology that can enhance the quality of migration. A comparison with interpolation reveals some difficulties but also a number of valid reasons why LSM is worthwhile to pursue: Improvements in resolution and amplitudes, and recovery of far offset energy. The use of filters, weights and domain transformation allows LSM to attenuate sampling artefacts, although the cost of LSM is much higher than that of interpolation.

Acknowledgements

I thank Chevron for the synthetic data set and my colleagues Sam Gray and Lorenzo Casasanta for their advice and useful discussions.

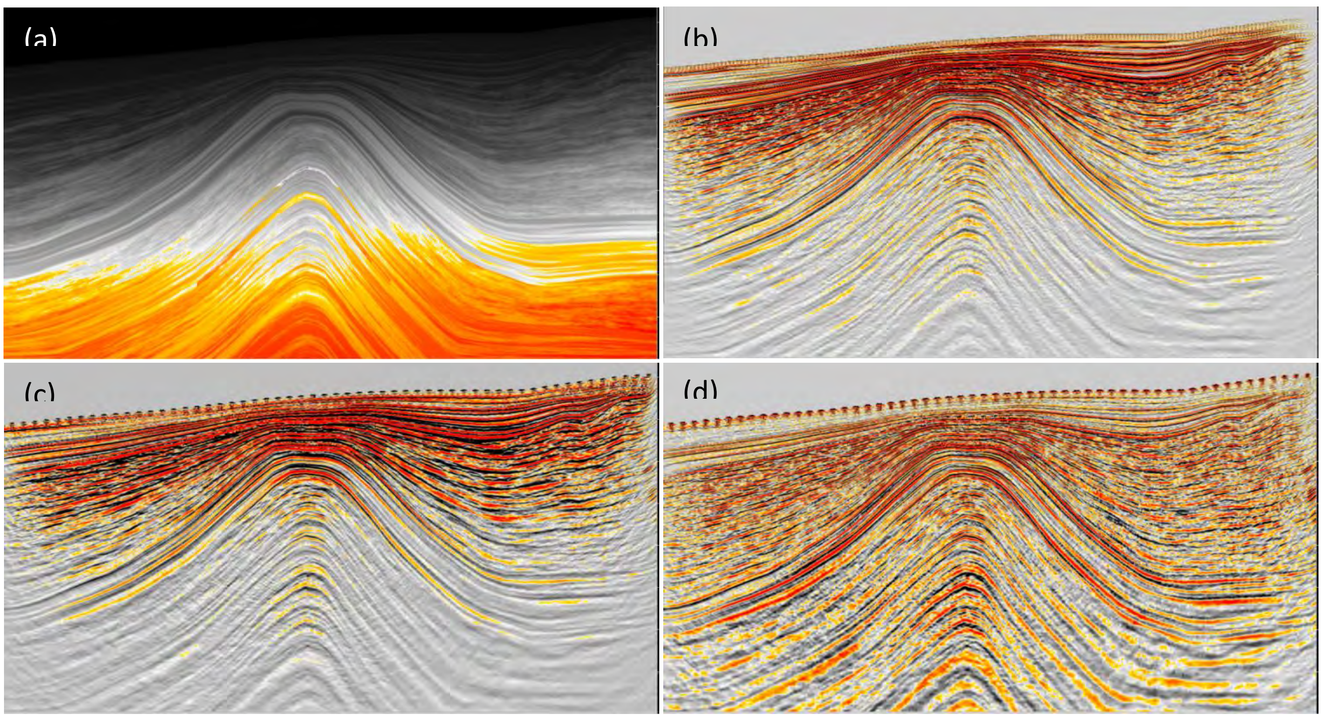

Figure 1(a) Velocity model (b) Migration from dense node acquisition (100m) and detailed velocity model. (c) Migration from sparse nodes (300m) and less detailed velocity model. (d) LSM(20 iterations) from the same data and model as in (c).

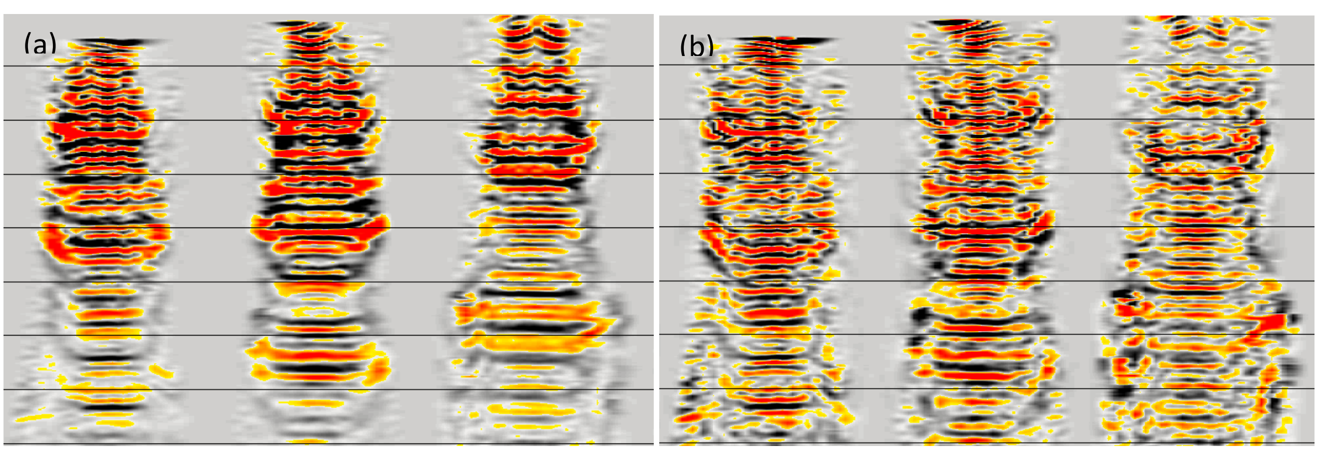

Figure 2 (a) CIGs for standard migration of decimated data. (b) LS CIGS for the same data.

References

Abma R. and N. Kabir, 2006, 3D interpolation of irregular data with POCS algorithm: Geophysics, 71, E91-E97

Dong S., J. Cai, M. Guo, S. Suh, Z. Zhang, B. Wang and Z. Li, (2012), Least-squares reverse time migration: towards true amplitude imaging and improving the resolution, SEG Technical Program Expanded abstracts 2012, pp 1-5

Herrmann F. and X. Li, 2012, Efficient least-squares imaging with sparsity promotion and compressive sensing, Geophysical Prospecting, vol. 60, p. 696-712.

Kuehl, H., and M. Sacchi, 2003, Least-squares wave-equation migration for AVP/AVA inversion: Geophysics, 68, 262–273.

Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, 64, No. 1, pp. 208-221.

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, 74, no. 6, V123–V132.

Yu, J., J. Hu, G. T. Schuster, and R. Estill, 2006, Prestack migration deconvolution: Geophysics, **71**, no. 2, S53–S62, http://dx.doi.org/10.1190/1.2187783

Zhang, Y.,Duan L., and Xie Y., 2013, A stable and practical implementation of least-squares reverse time migration, SEG Technical Program Expanded Abstracts 2013, 3716-3720.