geoconvention
Calgary • Canada • May 15-19 2017

Least squares Kirchhoff depth migration: important details

Daniel Trad
CREWES-University of Calgary

Summary

Least squares migration has been an important research topic in the academia for about two decades, but only recently it has attracted interest from the industry. The main reason is that from a practical point of view its ratio of benefit/cost has not been sufficient for its use in seismic exploration. Another problem is that these benefits are mixed with effects from the filtering techniques used to regularize the inversion, which are computationally much cheaper. In this paper, I discuss some challenges with least squares Kirchhoff depth migration. This algorithm, although less precise than the more popular least squares reverse time migration, has the advantage of being fast enough to be applied in a production environment, and flexible enough to be applied without data regularization. This last characteristic makes it a good candidate to understand benefits in terms of footprint acquisition and aliasing. In addition, its limitations in terms of modelling/imaging accuracy make more evident some problems that exist but are often ignored when using reverse time migration with synthetic data.

Introduction

Despite the existence of more accurate imaging algorithms, the most common migration algorithm for land seismic data is still Kirchhoff migration, mostly because of its flexibility to adapt to irregular sampling and its efficiency. Kirchhoff migration is basically a weighted mapping between data space (the original acquired data), and the image space (migrated section), where samples are weighted and shifted temporally and spatially. These weights come from the mathematics of integration and wave propagation, usually assuming infinite aperture, and taking only partially into account the exact locations of shots and receivers. Hard work over many decades has led to weights that produce a good approximation to true amplitude reflectivity. However, these weights have limitations because they do not contain information about the sampling operator and acquisition illumination.

Least squares migration (LSMIG) is a different approach to migration where, rather than mapping and weighting to find the model, a modeling operator is inverted by optimization algorithms (Nemeth et al., 1999; Kuhl and Sacchi, 2003). The inverse of this operator applied to seismic data produces an image that by design can predict the acquired data in a least squares sense. Although the inverse of this modeling operator is closely related to the imaging (mapping) operator, it represents a different approach to the problem and it takes into account the actual geometry (sampling and aperture) of the data. By inverting a modeling operator that contains the acquisition pattern, we can in principle remove acquisition and illumination effects.

Probably the most attractive characteristic of least squares migration is that it falls into a wide range of algorithmic techniques used not only in seismic but in physics in general. The mathematics and physics of LSMIG allows one to include concepts like prior information, in the form of mathematical constraints that help to obtain better images. In practice, however, LSMIG is a difficult process because of the high cost of applying the modeling/adjoint operator multiple times, and the challenging goal of creating an image that not only looks reasonable but also can predict the data. Although stacking samples (mapping) is a very robust process, inverting a multidimensional operator is usually an ill-conditioned problem.

From the practical point of view the advantages of LSMIG have not been shown in the industry except for specific examples, mostly 2D synthetic examples. Its application to 3D seismic processing is only starting to be attempted as an alternative to traditional methods and it remains as an elusive application of academic research. This is particularly true for land data sets, where noise is complex and correct signal amplitudes are very difficult to predict correctly. However, many possible advantages over conventional migration keep motivating researchers to solve these difficulties. Among these possibilities, we can mention (1) elimination of footprint noise caused by poor sampling and (2) better amplitudes in the image by compensation of illumination problems. In addition, the availability of the LSMIG engine provides a powerful and physical tool for interpolation, and de-noising (Trad, 2015).

Theory

The LSMIG implementation in this work uses a general philosophy widely used to solve for geophysical problems that originates from the Stanford Exploration Project and is nicely explained in Claerbout (1992). Given a transformation of some model **m** to some data **d** through some mathematical operator **L**,

$$\mathbf{d} = \mathbf{Lm}. \quad (1)$$

we can approximately solve the inverse problem of recovering the model that produced the data by inverting the operator in a least squares sense. The meaning of the different components of equation (1) depends on the details of the operator and are often associated with well-known transformations. For the case of LSMIG, **L** is a synthesis or modeling operator, that generates data from reflectivity using Green functions as basis functions. Any transformation could be used to generate the same data by this procedure, independently of whether it is physical or not. What is different and requires some consideration is the meaning we assign to the model. Even for very similar operators, small variations can change what the model represents. For example, by changing the weights in the operator we can switch the model from reflectivity (Kirchhoff) to velocity perturbations (Born).

Once a meaningful operator is chosen, we can estimate the model by inverting the operator. To invert operators, we define a cost or objective function, which is a mathematical expression that measures the undesired characteristics of the model. The most common is to find a model that honors the data in a least error sense, measured on some norm, and has a minimum of information not required by the data. This statement of goals is commonly presented as

$$\text{minimize } ||\mathbf{W}_m\mathbf{m}||^p_p$$
$$\text{subject to } ||\mathbf{W}_d(\mathbf{d} - \mathbf{Lm})||= \varphi_d \quad (2)$$

where φ_d is some estimate of the noise level in the data plus a residual due to the failure of the proposed model to explain the data. p and q indicate that different norms can be applied to measure the norm of vectors. $\mathbf{W}_d$ could be a matrix or vector of data weights, often a diagonal matrix containing the inverse of the standard deviation of the data, but more generally it could be any kind of filter that leaves out bad data. $\mathbf{W}_m$ is an operator of model weights that can be customized to enhance our preferences regarding the model. For example, $\mathbf{W}_m$ might be a gradient (roughening) operator, then minimizing $||\mathbf{W}_m\mathbf{m}||$ will produce a smooth model. After absorving $\mathbf{W}_\mathrm{d}$ and $\mathbf{W}_\mathrm{m}$ into **L** by a change of variables (Trad et al., 2003), the solution to this problem results on the usual LS formulation:

$$\mathbf{m=(L^HL)^{-1}L^Hd} \quad (3)$$

were the ($\mathbf{L^HL}$) operator, known as the Hessian, is iteratively deconvolved from the migration result ($\mathbf{L^Hd}$). For LSMIG we usually need to favor, but not enforce, smooth models that lead to continuous structures. A typical choice for Kirchhoff type of migrations, is to use $\mathbf{W}_m$ to enforce smooth changes of reflectivity with offset or angle. This can be done as in Nemeth et al. (1999) by taking differences between consecutive reflectivities at different offsets, as in Kuhl and Sacchi (2003) by a sparse tau p transform of the reflectivity, or as in Moghaddam and Herrmann (2005) by a sparse curvelet transform of the reflectivity. A generic choice for $\mathbf{W}_m$ is the inverse of a multidimensional operator that takes a multioffset migrated volume and performs smoothing, for example a triangular filter, a dip filter, a fxy filter or a Radon transform. In Fomel (2007), this formulation is generalized with the name of shaping filter.

Weighting

It is not trivial to achieve the optimal weighting scheme for LS Kirchhoff migration. A common approach is to use standard migration weights with a migration operator, design its adjoint such that it passes the adjoint test (Claerbout, 1992), and use these two operators into a CG algorithm. Weights in migration algorithms often take the form of deconvolution (DC) weights, which come from the imaging principle

(Claerbout, 1971) as a ratio of scattered (u_s) over incident (u_i) wavefields. This leads to migration weights that are proportional to a ratio of shot and receiver wavefield amplitudes, or, after approximating for constant velocity, to a ratio of traveltime inverses. (Dellinger et al., 1999; Zhang et al., 2000). On the other hand, Nemeth et al. (1999) sets the operator $\mathbf{L}$ as a modelling rather than a migration (Casasanta, personal communication). This approach changes the imaging condition from deconvolution (migration) to cross-correlation (modelling). In the cross-correlation imaging condition (XC) the LS weights are proportional to the product of amplitudes, which leads, after using similar approximations, to a product of traveltimes instead of a ratio. Testing with simple flat synthetics supports this interpretation. Figure 1a-b illustrates that deconvolution weights produce the correct amplitudes for migration (a), but as we iterate during the inversion, the amplitudes get distorted (b). Figure 1c-d show that cross-correlation weights produce the wrong image for migration (c), but the amplitudes become correct after a few iterations (d).

Another related issue is whether the phase filter operator should be included inside the modelling - migration operator so it would act for every iteration, or should be only pre-applied as a data preconditioner, or not applied at all. The exact formulation for the phase shift operator depends on the dimensionality of the data acquisition, being equal to $i\omega$ for 3D, $\sqrt{i|\omega|}$ for 2.5D or $|\omega|$ for 2D (equations 2 in Bleistein and Gray (2001)). In this paper, I refer to this filter simply as ω filter. Figure 2 compares the amplitude spectra. These tests show that if the ω filter is not included inside the operator, then the spectrum of the image shift towards higher frequencies with iterations. Another factor to consider with data weights is whether to include or not velocity weights. In theory amplitude weights contain a power of velocity (v^{-2} for 2D, v^{-3} for 3D). Although in migration this factor can be replaced by a depth gain after migration, in LSMIG it has a different effect because by including the velocities inside the weights the convergence is focused to areas with higher velocity. This is one of many situations where simplifications taken in migration do not carry over to LSMIG.

Noise control

One of the most important problems in LSMIG is the accumulation of noise with iterations. This noise comes from broken contributions to the model as the algorithm tries to predict small details on the data that are not properly mapped by the migration operator, in other words, when the operator does not approximate the physics correctly. For example, this effect is accentuated with discontinuities or gaps in Kirchhoff depth migration traveltime tables, which make the operator a poor representation of the physics. Conversely, this noise disappears if the data are created by the same numerical approximation used in the migration engine. For this reason, LSMIG noise is more obvious with a Kirchhoff migration than with reverse time migration (RTM). However, RTM may not show this problem if tested with finite difference data, but the noise may appear in real data processing as the numerical approximation deviate from the actual physics that produced the data.

As usual in ill-conditioned inversions, this problem requires some form of noise control by means of numerical regularization. Regularization in this work is performed by a multidimensional triangular filter (the $\mathbf{W}_m$ operator in equations 2). The filter acts on the inline, crossline and offset dimensions. The length of the filter, which is independent across dimensions, permits one to control the noise effectively, but it is limited by the structural complexity of the image. In Figure 3a we see one iteration of LSMIG (equivalent to migration) for the Marmousi data set, Figure 3b we see the same result after 10 iterations when no noise control is applied, and Figure 3c shows the same with a triangle filter along crosslines (cdps) and offsets. The length of the filter was very mild to preserve structure (length along inline was 5, length along offset was 5). Increasing the length would make the result cleaner, but some faults would be smoothed. We need to emphasize that improvements in the image we seek should not be a pure consequence of the regularization filter but of the inversion of the Hessian operator $\mathbf{L}^{\mathbf{H}}\mathbf{L}$.

Convergence

A critical aspect for a successful application of LSMIG in an industrial setting is to improve its convergence. One factor that has a large effect on convergence is the dimensionality of the model. When the model is defined in terms of offset gathers instead of their summation, convergence improves dramatically. This is understandable because of less mixing, which makes prediction easier, and therefore convergence faster. Using offset gathers has the effect of subdividing the problem into smaller sub-inversions. Similarly, convergence for a part of the data is much faster than for the whole data. The trade-off, however, is that by using part of the data inversion is only partially accomplished, even if data fitting is better.

Conclusions

Although straightforward in principle, there are many options related to the operator design whose choice can have significant influence in results and efficiency, and are not obvious from the theoretical point of view. In particular it is important to understand the operator weights in terms of phase and amplitude as well as regularization operators to control the noise during iterations.

Acknowledgements

My gratitude to Lorenzo Casasanta and Sam Gray for many enlightful discussions about migration during my work at CGG. I also thank CREWES sponsors for contributing to this seismic research. I also gratefully acknowledge support from NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13.

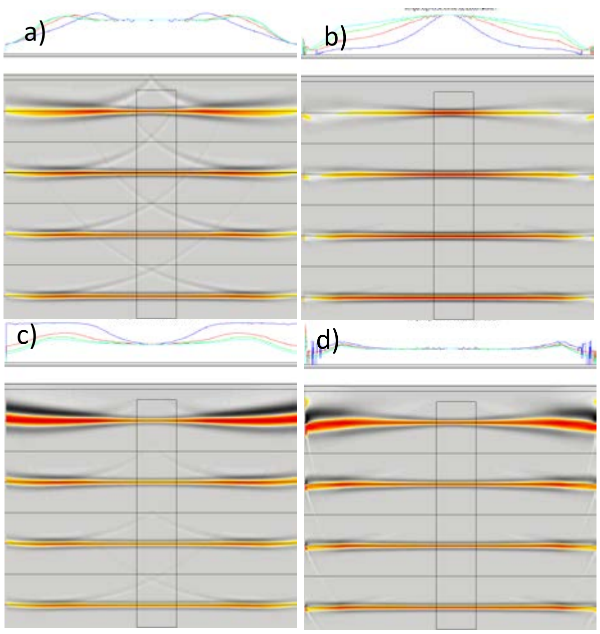

Figure 1: Top DC weights, a) migration, b) LSMIG, bottom XC weights, c) migration d) LSMIG

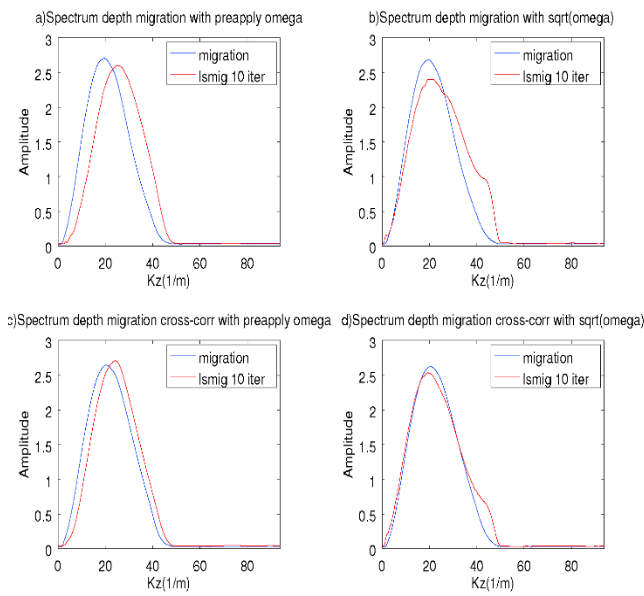

Figure 2: Effect of phase filter. Top DC weights, a) pre-applied b) inside operator, bottom XC weights, c) pre-applied d) inside operator

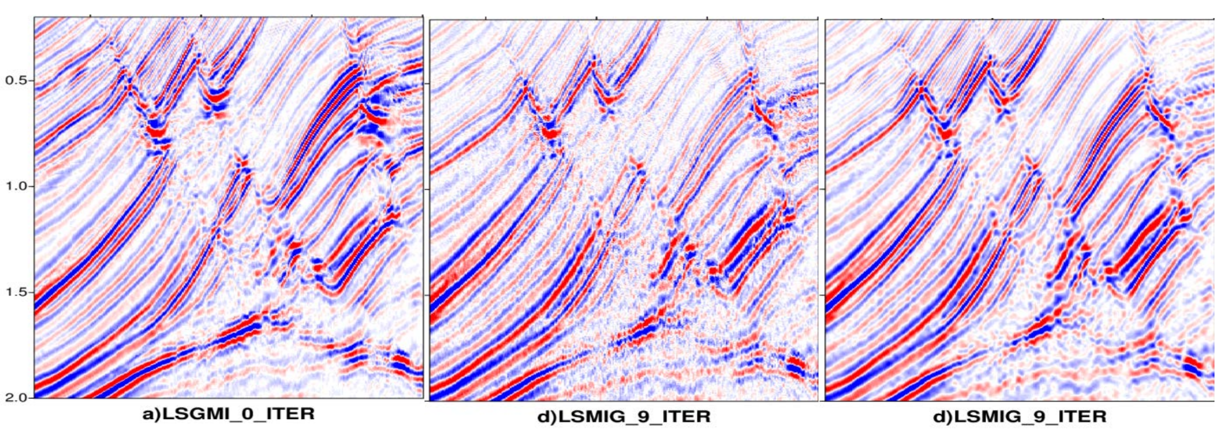

Figure 2: Noise with iterations. a) 1 iteration b) 9 iterations without noise control c) 9 iterations with filter inside operator.

References

Bleistein, N., and Gray, S. H., 2001, From the hagedoorn imaging technique to kirchhoff migration and inversion: Geophysical Prospecting, 49, No. 6, 629–643.

Claerbout, J., 1992, Earth sounding analysis, Processing versus inversion: Blackwell Scientific Publications. Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: GEOPHYSICS, 36, No. 3, 467–481.

Dellinger, J., Murphy, G., Etgen, J., Fei, T., and Gray, S., 1999, Efficient 2.5-d trueâA˘ Ramplitude migration:˘ The Leading Edge, 18, No. 8, 946–949.

Docherty, P., 1991, A brief comparison of some kirchhoff integral formulas for migration and inversion: GEOPHYSICS, 56, No. 8, 1164–1169.

Fomel, S., 2007, Shaping regularization in geophysical-estimation problems: GEOPHYSICS, 72, No. 2, R29–R36.

Kuhl, H., and Sacchi, M. D., 2003, Least squares wave equation migration for avp/ava inversion: GEOPHYSICS, 68, No. 1, 262–273.

Lumley, D. E., Claerbout, J. F., and Bevc, D., 1994, AntiâA˘ Raliased kirchhoff 3â˘ A˘ RD migration, 1282–1285.˘ Moghaddam, P. P., and Herrmann, F. J., 2005, Migration preconditioning with Curvelets, 2204–2207.

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

Trad, D., 2015, Least squares kirchhoff depth migration: implementation, challenges, and opportunities, in 2015 SEG Annual Meeting, Society of Exploration Geophysicists, 4238–4242.

Trad, D., Ulrych, T., and Sacchi, M., 2003, Latest views of the sparse radon transform: GEOPHYSICS, **68**, No. 1, 386–399.

Zhang, Y., Gray, S., and Young, J., 2000, Exact and approximate weights for Kirchhoff migration, 1036– 1039.