Stanford Exploration Project, Report 108, April 29, 2001, pages 1–??

Model-space vs. data-space normalization for recursive depth migration

James Rickett[1]

ABSTRACT

Illumination problems caused by finite-recording aperture and lateral velocity lensing can lead to amplitude fluctuations in migrated images. I calculate both model and data-space weighting functions that compensate for these illumination problems in recursive depth migration results based on downward-contination. These weighting functions can either be applied directly with migration to mitigate the effects of poor subsurface illumination, or used as preconditioning operators in iterative least-squares ($L2$) migrations. Computational shortcuts allow the weighting functions to be computed at about the cost of a single migration. Results indicate that model-space normalization can significantly reduce amplitude fluctuations due to illumination problems. However, for the examples presented here, data-space normalization proved susceptible to coherent noise contamination.

INTRODUCTION

Migration is the adjoint of a linear forward modeling operator rather than the inverse [e.g. Claerbout (1995)]. This means that, although migration treats kinematics correctly, the amplitudes of migrated images do not accurately represent seismic reflectivity.

Geophysical inverse theory provides a rigorous framework for estimating earth models that are consistent with some observed data. Typically the matrices involved in industrial-scale geophysical inverse problems are too large to invert directly, and we depend on iterative gradient-based linear solvers to estimate solutions. However, operators such as prestack depth migration are so expensive to apply that we can only afford to iterate a handful of times, at best.

In this paper I compute diagonal weighting functions that can be applied directly to migrated images to compensate for the inadequacies of the adjoint with respect to seismic amplitudes. Furthermore, these weighting functions can be applied as preconditioning operators that speed the convergence of iterative linear solvers, facilitating least-squares recursive depth migration.

As well as looking at model-space weights, I also consider data-space weighting functions derived from the operator $\mathbf{A}\mathbf{A}'$ (where $\mathbf{A}$ is our linear forward modeling operator), and develop

[1]**email:** james@sep.stanford.edu

a framework for computing and applying both model and data-space weights simultaneously.

MODEL-SPACE WEIGHTING FUNCTIONS

For an over-determined system of equations, the inverse problem can be summarized as follows - given a linear forward modeling operator $\mathbf{A}$, and some recorded data $\mathbf{d}$, estimate a model $\mathbf{m}$ such that $\mathbf{A}\mathbf{m} \approx \mathbf{d}$. The model that minimizes the expected error in predicted data is given by:

$$\mathbf{m}_{L2} = (\mathbf{A}'\mathbf{A})^{-1}\,\mathbf{A}'\mathbf{d}. \tag{1}$$

To attempt to speed convergence, we can always change model-space variables from $\mathbf{m}$ to $\mathbf{x}$ through a linear operator $\mathbf{P}$, and solve the following new system for $\mathbf{x}$,

$$\mathbf{d} = \mathbf{A}\mathbf{P}\mathbf{x} = \mathbf{B}\mathbf{x}. \tag{2}$$

When we find a solution, we can then recover the model estimate, $\mathbf{m}_{L2} = \mathbf{P}\mathbf{x}$. If we choose the operator $\mathbf{P}$ such that $\mathbf{B}'\mathbf{B} \approx \mathbf{I}$, then even simply applying the adjoint ($\mathbf{B}'$) will yield a good model estimate; furthermore, gradient-based solvers should converge to a solution of the new system rapidly in only a few iterations. The problem then becomes: what is a good choice of $\mathbf{P}$?

Rather than trying to solve the full inverse problem given by equation (1), I look for a diagonal operator $\mathbf{W}_{\rm m}$ such that

$$\mathbf{W}_{\rm m}^2\,\mathbf{A}'\mathbf{d} \approx \mathbf{m}_{L2}. \tag{3}$$

$\mathbf{W}_{\rm m}$ can be applied to the migrated (adjoint) image with equation (3); however, in their review of $L2$ migration, Ronen and Liner (2000) observe that normalized migration is only a good substitute for full (iterative) $L2$ migration in areas of high signal-to-noise. In these cases, $\mathbf{W}_{\rm m}$ can be used as a model-space preconditioner to the full $L2$ problem, as described in the introduction.

Claerbout and Nichols (1994) noticed that if we model and remigrate a reference image, the ratio between the reference image and the modeled/remigrated image will be a weighting function with the correct physical units. For example, the weighting function, $\mathbf{W}_{\rm m}$, whose square is given by

$$\mathbf{W}_{\rm m}^2 = \frac{\mathbf{diag}(\mathbf{m}_{\rm ref})}{\mathbf{diag}(\mathbf{A}'\,\mathbf{A}\,\mathbf{m}_{\rm ref})} \approx \left(\mathbf{A}'\mathbf{A}\right)^{-1}, \tag{4}$$

will have the same units as $\mathbf{A}^{-1}$. Furthermore, $\mathbf{W}_{\rm m}^2$ will be the *ideal* weighting function if the reference model equals the true model and we have the correct modeling/migration operator.

Equation (4) with forms the basis for the first part of this paper. However, when following this approach, there are two important practical considerations to take into account: firstly, the choice of reference image, and secondly, the problem of dealing with zeros in the denominator.

Similar normalization schemes [e.g. Slawson et al. (1995); Chemingui (1999); Duquet et al. (2000)] have been proposed for Kirchhoff migration operators. In fact, both Nemeth et al. (1999) and Duquet et al. (2000) report success with using diagonal model-space weighting functions as preconditioners for Kirchhoff $L2$ migrations. However, normalization schemes that work for Kirchhoff migrations are not computationally feasible for recursive migration algorithms based on downward-continuation.

Three choices of reference image

The ideal reference image would be the true subsurface model. However, since we do not know what that is, we have to substitute an alternative model. I experiment with three practical alternatives, which I will denote $\mathbf{m}_1$, $\mathbf{m}_2$, and $\mathbf{m}_3$.

Claerbout and Nichols (1994) attribute to Symes the idea of using the adjoint (migrated) image as the reference model. The rationale for this is that migration provides a robust estimate of the true model. As the first alternative I take Symes' suggestion, so that $\mathbf{m}_1 = \mathbf{A}'\mathbf{d}$. The second alternative is to try an reference image of purely random numbers: $\mathbf{m}_2 = \mathbf{r}$, where $\mathbf{r}$ is a random vector. This is has the advantage of not being influenced by the data, but has the disadvantage that different realizations of $\mathbf{r}$ may produce different weighting functions. The third alternative that I consider is a monochromatic reference image ($\mathbf{m}_3$) consisting of purely flat events: literally flat-event calibration.

Stabilizing the denominator

To avoid division by zero, Claerbout and Nichols (1994) suggest multiplying both the numerator and denominator in equation (4) by $\mathbf{diag}(\mathbf{A}'\,\mathbf{A}\,\mathbf{m}_{\rm ref})$, and stabilizing the division by adding a small positive number to the denominator:

$$\mathbf{W}_{\rm m}^2 = \frac{\mathbf{diag}(\mathbf{m}_{\rm ref})\cdot\mathbf{diag}(\mathbf{A}'\,\mathbf{A}\,\mathbf{m}_{\rm ref})}{|\mathbf{diag}(\mathbf{A}'\,\mathbf{A}\,\mathbf{m}_{\rm ref})|^2+\epsilon^2\mathbf{I}}, \tag{5}$$

Although this does solve the problem of division by zero, the numerator and denominator will still oscillate rapidly in amplitude with the phase of the image.

Illumination, however, should be independent of the wavefield's phase. Therefore, I calculate weighting functions from the ratio of the smoothed analytic signal envelopes (denoted by $<>$) of the model-space images:

$$\mathbf{W}_{\rm m}^2 = \frac{\mathbf{diag}(<\mathbf{m}_{\rm ref}>)}{\mathbf{diag}(<\mathbf{A}'\,\mathbf{A}\,\mathbf{m}_{\rm ref}>)+\epsilon^2\mathbf{I}}, \tag{6}$$

where ϵ is a damping parameter that is related to the signal-to-noise level.

Numerical comparison

The Amoco 2.5-D synthetic dataset (Etgen and Regone, 1998) provides an excellent test for the weighting functions discussed above.

The velocity model (Figure 1) contains significant structural complexity in the upper 3.8 km, and a flat reflector of uniform amplitude at about 3.9 km depth. Since the entire velocity model ("Canadian foothills overthrusting onto the North Sea") is somewhat pathological, I restricted my experiments to the North Sea section of the dataset ($x > 10$ km). The data were generated by 3-D acoustic finite-difference modeling of the 2.5-D velocity model. However, making the test more difficult is the fact that the 2-D linear one-way recursive extrapolators (Ristow and Ruhl, 1994) that I use for modeling and migration do not accurately predict the 3-D geometric spreading and multiple reflections that are present in this dataset.

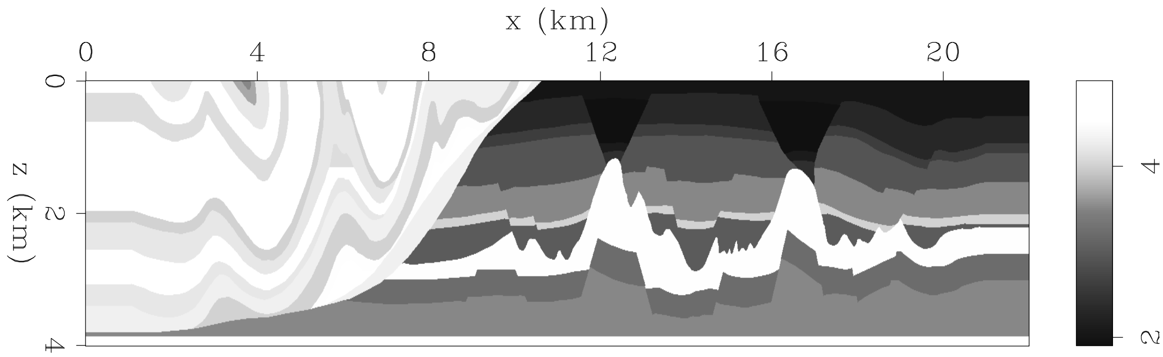

Figure 1: Velocity (in km/s) model used to generate the synthetic Amoco 2.5-D dataset. james3-amocovel [CR]

Figure 2 compares the migrated image ($\mathbf{m}_1$) with the results of remodeling and remigrating the three reference images described above. The imprint of the recording geometry is clearly visible on the three remigrations in Figures 2 (b-d).

Figure 3 compares the illumination calculated from the three reference images with the raw shot illumination. Noticably, the shot-only weighting function [panel (a)] does not take into account the off-end (as opposed to split-spread) receiver geometry. Panel (b), the weighting function derived from model $\mathbf{m}_1$, appears slightly noisy. However, in well-imaged areas (e.g. along the target reflector), the weighting function is well-behaved. Panel (c) shows the weighting function derived from the random reference image ($\mathbf{m}_2$). Despite the smoothing, this weighting function clearly bears the stamp of the random number field. A feature of white noise is that no amount of smoothing will be able to remove the effect of the random numbers completely. The final panel (d) shows the flat-event illumination weighting function, derived from $\mathbf{m}_3$. This is noise-free and very well-behaved since it depends only on the velocity model and recording geometry, not the data.

For a quantitative comparison, I picked the maximum amplitude of the 3.9 s reflection event on the calibrated images. The normalized standard deviation (NSD) of these amplitudes is shown in Table 1, where

$$\mathrm{NSD} = \sqrt{\sum_{i_x} \left(\frac{a_{i_x}}{\bar{a}} - 1 \right)^2}. \tag{7}$$

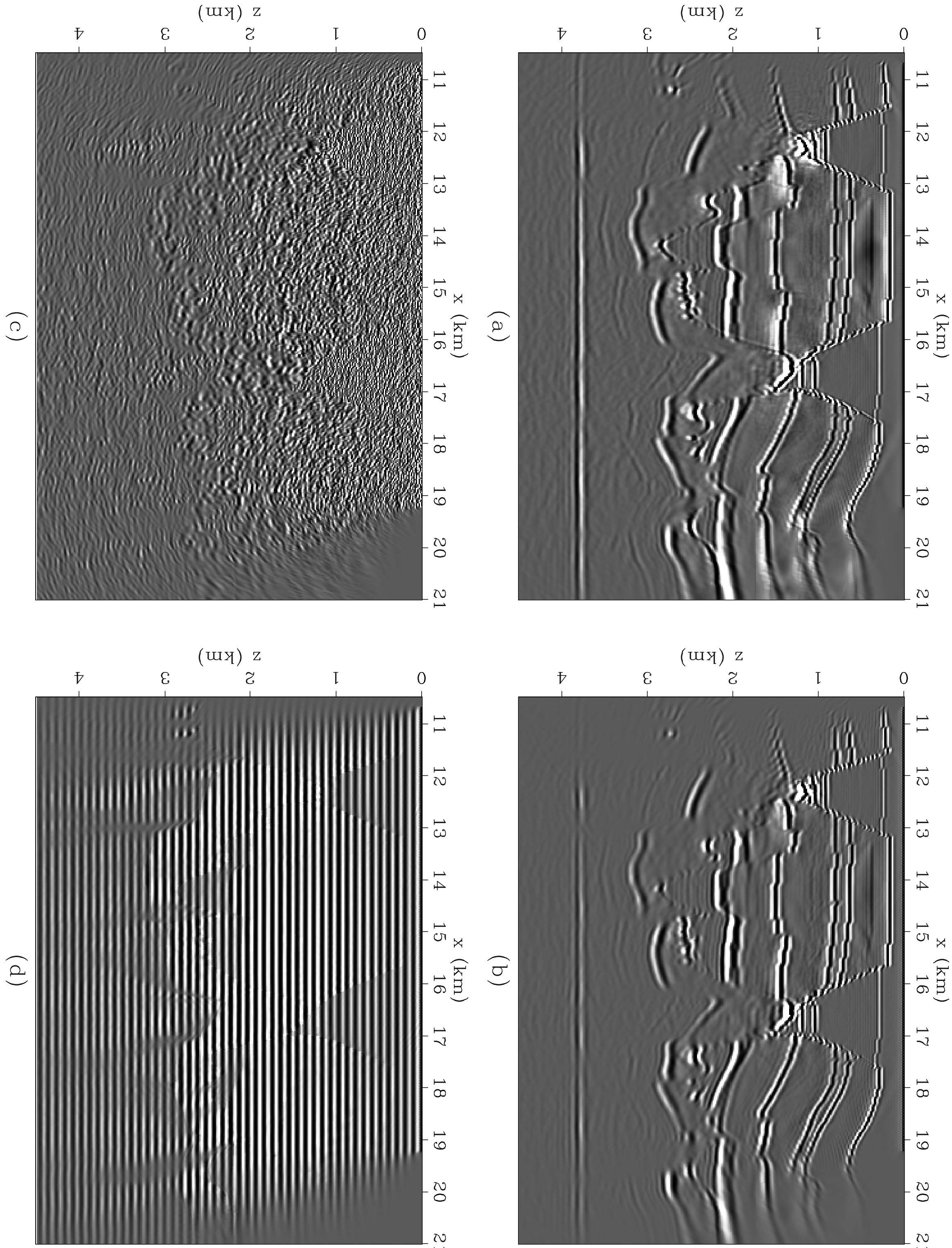

Figure 2: Comparison of calibration images: (a) original migration, (b) original migration after modeling and migration, (c) random image after modeling and migration, and (d) flat event image after modeling and migration. james3-amocomigs [CR,M]

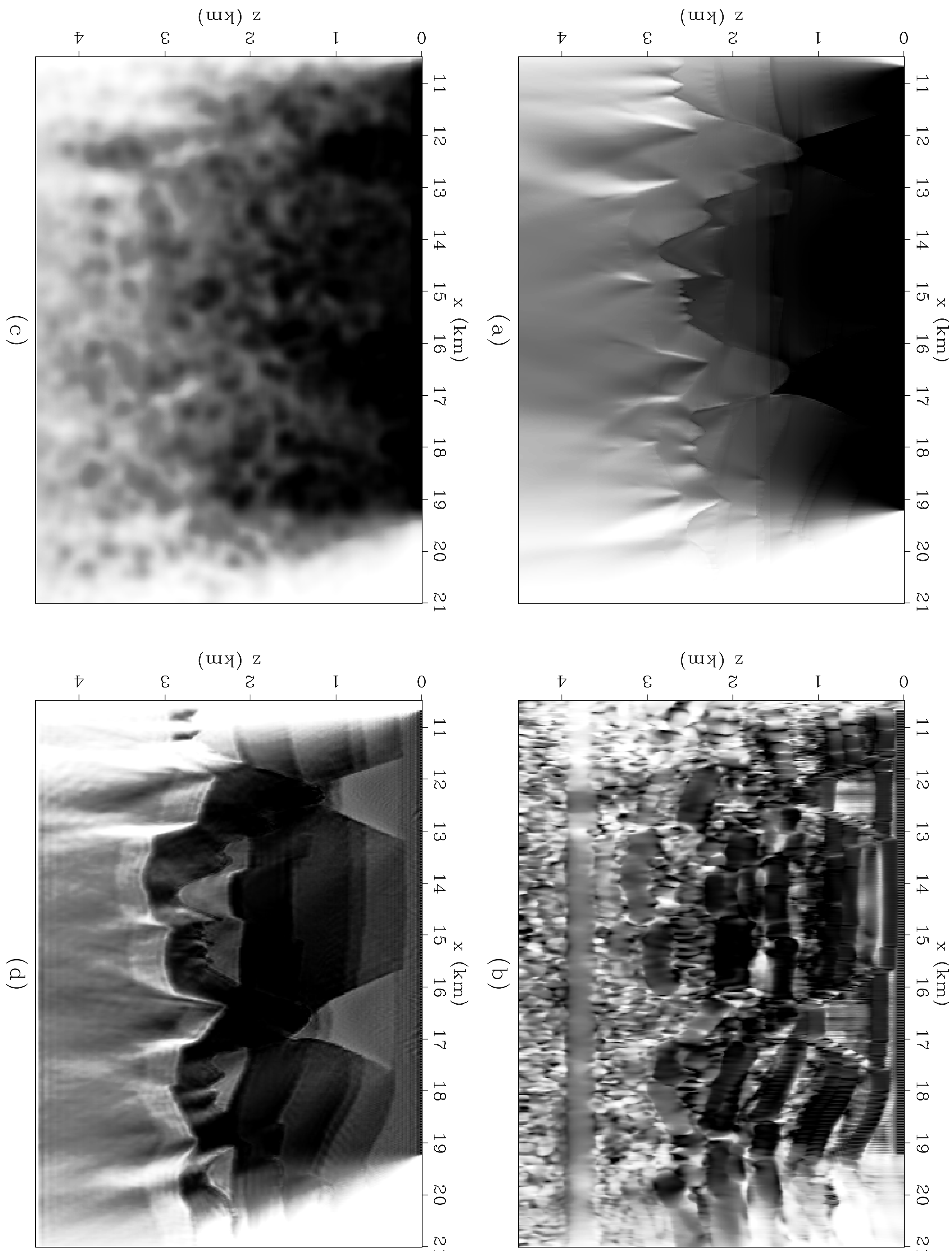

Figure 3: Comparison of weighting functions: (a) original migration, (b) original migration after modeling and migration, (c) random image after modeling and migration, and (d) flat event image after modeling and migration. james3-amocowghts [CR,M]

Weighting function:	NSD:
No weighting function	0.229
Shot illumination	0.251
$\mathbf{m}_{\text{ref}} = \mathbf{m}_1$ (migrated image)	0.148
$\mathbf{m}_{\text{ref}} = \mathbf{m}_2$ (random image)	0.195
$\mathbf{m}_{\text{ref}} = \mathbf{m}_3$ (flat events)	0.140

Table 1: Comparison of the reflector strength for different choices of illumination-based weighting function.

Table 1, therefore, provides a measure of how well the various weighting function compensate for illumination difficulties. The amplitudes of the raw migration, and the migration after flat-event normalization are shown in Figure 4. This illustrates the numerical results from Table 1: for this model the normalization procedure improves amplitude reliability by almost a factor of two.

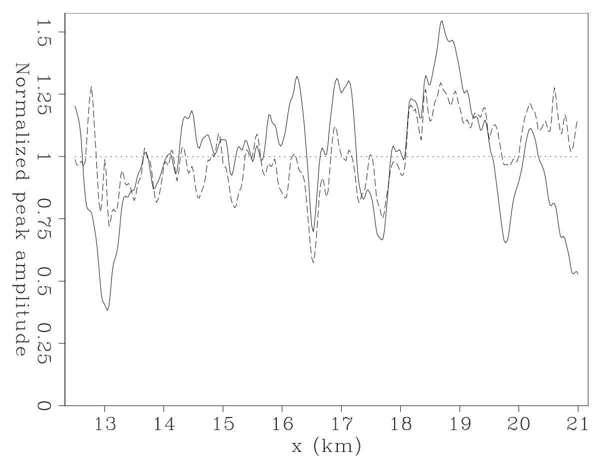

Figure 4: Normalized peak amplitude of 3.9 km reflector after migration (solid line), and then normalization by flat-event illumination (dashed line) derived with $\mathbf{m}_{\text{ref}} = \mathbf{m}_3$. The ideal result would be a constant amplitude of 1. james3-eventamp [CR]

Computational cost

As it stands, the cost of computing a weighting function of this kind is twice the cost of a single migration. Add the cost of the migration itself, and this approach is 25% cheaper than running two iterations of conjugate gradients, which costs two migrations per iteration.

However, the bandwidth of the weighting functions is much lower than that of the migrated images. This allows considerable computational savings, as modeling and remigrating a narrow frequency band around the central frequency produces similar weighting functions than the full bandwidth. Repeating the first experiment ($\mathbf{m}_{\text{ref}} = \mathbf{m}_1$) with half the frequencies gives a NSD $= 0.147$ - the same as before within the limits of numerical error.

DATA-SPACE WEIGHTING FUNCTIONS

If the system of equations, $\mathbf{d} = \mathbf{A}\,\mathbf{m}$, is underdetermined, then a standard approach is to find the solution with the minimum norm. For the $L2$ norm, this is the solution to

$$\hat{\mathbf{m}}_{L2} = \mathbf{A}' \left(\mathbf{A}\,\mathbf{A}'\right)^{-1} \mathbf{d}. \tag{8}$$

As a corollary to the the methodology outlined above for creating model-space weighting functions, Claerbout (1998) suggests constructing diagonal approximations to $\mathbf{A}\,\mathbf{A}'$ by probing the operator with a reference data vector, $\mathbf{d}_{\text{ref}}$. This gives data-space weighting functions of the form,

$$\mathbf{W}_{\text{d}}^2 = \frac{\mathbf{diag}(\mathbf{d}_{\text{ref}})}{\mathbf{diag}(\mathbf{A}\,\mathbf{A}'\,\mathbf{d}_{\text{ref}})} \approx \left(\mathbf{A}\,\mathbf{A}'\right)^{-1}, \tag{9}$$

which can be used to provide a direct approximation to the solution in equation (8),

$$\hat{\mathbf{m}}_{L2} \approx \mathbf{A}'\,\mathbf{W}_{\text{d}}^2\,\mathbf{d}. \tag{10}$$

Alternatively, we could use $\mathbf{W}_{\text{d}}$ as a data-space preconditioning operator to help speed up the convergence of an iterative solver:

$$\mathbf{W}_{\text{d}}\,\mathbf{d} = \mathbf{W}_{\text{d}}\,\mathbf{A}\,\mathbf{m}. \tag{11}$$

Combining weighting functions

With two possible preconditioning operators, $\mathbf{W}_{\text{m}}$ and $\mathbf{W}_{\text{d}}$, the question remains, what is the best strategy for combining them?

The first strategy that I propose is to calculate a model-space weighting function, $\mathbf{W}_{\text{m}}$, and use it to create a new preconditioned system with the form of

$$\mathbf{d} = \mathbf{A}\,\mathbf{W}_{\text{m}}\,\mathbf{x} = \mathbf{B}\,\mathbf{x}.$$

Now probe the composite operator, $\mathbf{B}$, for a data-space weighting function for the new system,

$$\tilde{\mathbf{W}}_{\text{d}}^2 = \frac{< \mathbf{diag}(\mathbf{d}_{\text{ref}}) >}{< \mathbf{diag}(\mathbf{B}\,\mathbf{B}'\,\mathbf{d}_{\text{ref}}) > + \epsilon_{\text{d}}\mathbf{I}} \approx \frac{1}{\mathbf{B}\,\mathbf{B}'}. \tag{12}$$

The new data-space weighting function is dimensionless, and can be applied in consort with the model-space operator. This leads to a new system of equations,

$$\begin{aligned} \tilde{\mathbf{W}}_{\text{d}}\,\mathbf{d} &= \tilde{\mathbf{W}}_{\text{d}}\,\mathbf{A}\,\mathbf{W}_{\text{m}}\,\mathbf{x} \\ \text{with} \quad \mathbf{m} &= \mathbf{W}_{\text{m}}\,\mathbf{x}, \end{aligned} \tag{13}$$

with appropriate model-space and data-space preconditioning operators. The adjoint solution to this system is given by

$$\mathbf{m} = \mathbf{W}_{\rm m}^2 \, \mathbf{A}' \, \tilde{\mathbf{W}}_{\rm d}^2 \, \mathbf{d}. \tag{14}$$

A second alternative strategy is the corollary of this: create a new system that is preconditioned by an appropriate data-space weighting function, and then calculate a model-space weighing function based on the new system.

Numerical comparisons

Again, the Amoco 2.5-D dataset provides an excellent test dataset for comparing flavors of weighting function. Unfortunately the data-space weights proved susceptible to coherent noise in the form of multiples not predicted by the modeling operator. While data-space weights did improve the signal in poorly illuminated areas, they also boosted up the noise level causing an *increase* in NSD. So further work will be required to make this approach useful.

DISCUSSION

While the methodologies described in this paper are valid for general linear operators, they have several fundamental limitations. Most importantly, they require an accurate forward modeling operator: both the physics of wave propagation and the true earth velocity must be accurately modeled. While the physics of wave-propagation is broadly understood, earth velocity models are never completely true-to-life.

Another important caveat is that the “wave-equation methods” outlined here require the data and models to be represented on a regular grid. While we can choose our model-space, prestack seismic data is never recorded on a perfectly regular grid. Before we can apply any wave-equation technique (such as those described here), we need to regularize the data. Chemingui (1999) and Fomel (2000) provide two different approaches to solving this problem.

CONCLUSION

Model-space weighting functions based on equation (6) provide a robust way to compensate for illumination problems during recursive depth migration based on downward-continuation. Data-space weights can be calculated either to work alone, or in consort with model-space weights. However they are less robust to errors caused by inadequate forward modeling.

ACKNOWLEDGEMENTS

I would like to thank BP Amoco for releasing this dataset to SEP, and Joe Dellinger for installing it in our data library.

REFERENCES

Chemingui, N., 1999, Imaging irregularly sampled 3D prestack data: Ph.D. thesis, Stanford University.

Claerbout, J., and Nichols, D., 1994, Spectral preconditioning: SEP–**82**, 183–186.

Claerbout, J. F., 1995, Basic earth imaging: Stanford Exploration Project.

Claerbout, J. F., 1998, Geophysical estimation by example: Stanford Exploration Project.

Duquet, B., Marfurt, K. J., and Dellinger, J. A., 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: Geophysics, **65**, no. 4, 1195–1209.

Etgen, J., and Regone, C., 1998, Strike shooting, dip shooting, widepatch shooting - Does prestack depth migration care? A model study.: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 66–69.

Fomel, S., 2000, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.

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

Ristow, D., and Ruhl, T., 1994, Fourier finite-difference migration: Geophysics, **59**, no. 12, 1882–1893.

Ronen, S., and Liner, C. L., 2000, Least-squares DMO and migration: Geophysics, **65**, no. 5, 1364–1371.

Slawson, S. E., Monk, D. J., and Sudhakar, V., 1995, DMO implications for 3-D survey design: 65th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 935–938.