Coupling Modeling and Migration for Seismic Imaging

Hervé Chauris and Daniela Donno
Centre de Géosciences, Mines ParisTech, UMR Sisyphé 7619 France

1. Introduction

Seismic imaging consists of retrieving the Earth’s properties, typically velocity and density models, from seismic measurements at the surface. It can be formulated as an inverse problem (Bamberger et al., 1982; Beylkin, 1985; Lailly, 1983; Tarantola, 1987). The resolution of the inverse problem involves two seismic operators: the modeling and the migration operators.

The modeling operator \( M \) applies to a given velocity model \( m(x) \), where \( x \) denotes the spatial coordinates, and indicates how to generate the corresponding shot gathers at any position in the model, usually at the surface. It consists of solving the wave equation for given velocity and density parameters. Fig. 1 and 2 illustrate the acoustic wave propagation for different travel times in two different velocity models. A point source generates a roughly circular wavefront for short travel times. The wavefront is then largely distorted due to the heterogeneous aspect of the velocity model. In simple models, it is easy to derive which part of the wave energy is diffracted, reflected, transmitted or refracted (Fig. 1). In more complex models, the wave modeling is obtained by numerically solving the wave equation, here with a finite difference scheme in the time domain (Fig. 2). In our definition, the migration operator is the adjoint \( M' \) of the modeling operator. It is related to kinematic migration, in the sense that the adjoint operator does not necessarily consider proper amplitudes. Equivalently, the modeling operator is also known to be the demigration operator.

Both the modeling and the migration operators can be very complicated. They provide the link between the time/data domain (shot, receiver and time) and the space/model domain (x positions). For example, a homogeneous model with a local density anomaly will create a data gather containing the direct arrival and a diffraction curve. For more complex models, the corresponding data gather is complicated, even under the Born approximation. Fig. 3 illustrates the fact that a single input trace extracted from a shot gather contributes to a large portion of the migrated image, whatever the type of migration used for implementation. The same conclusion holds for elastic modeling or more sophisticated wave equations.

We analyze in this work the combination of the modeling and the migration operators, with the objective of showing that the coupling of these operators can provide a large number of benefits for seismic imaging purposes. In particular, we consider the following general operator \( H = M'[m + \delta m] W M[m] \). Here, we call \( H \) the generalized Hessian. The operator \( W \) is typically a weighting or filtering matrix. The modeling and migration operators may be defined in two different models \( m \) and \( m + \delta m \), with \( \delta m \) representing a model perturbation.
Fig. 1. Snapshots of the acoustic wave propagation in a simple model for different travel time values (from left to right and top to bottom). The velocity model consists of two different homogeneous layers and a diffraction point (white point). The source position is indicated by a star.

The exact descriptions of $W$ and $\delta m$ are given in the following sections, depending on the applications. In the strict definition of the Hessian, the model perturbation $\delta m$ is equal to zero and $W$ is an identity matrix. The classical Hessian, also known as normal operator $M' M$, naturally appears in the solution of the imaging problem as proposed by Tarantola (1987). It makes the link between images defined in the same domains, here the space domain.

Compared to the effects of $M$ or $M'$, we would like to demonstrate through different examples from the literature that the application of operator $H$ has several advantages. In section 2 of this work, we present the exact expression of the Hessian and we show why this operator naturally appears in the resolution of the inverse problems as in the migration case. Then, we concentrate on three main seismic imaging tasks: pre-processing steps for reducing migration artifacts (section 3), true-amplitude imaging processes (section 4), and image sensitivity to model parameters (section 5). We also discuss the difficulties to construct the Hessian (very large matrix) and to invert for it (ill-conditioned matrix), and present several strategies to avoid its computation by sequentially applying the modeling and migration operators. We review the different approaches proposed within the geophysical community, mainly during
Fig. 2. Snapshots of the acoustic wave propagation in a complex model for different travel time values (from left to right and top to bottom). The velocity model is displayed in the image background. The source position is indicated by a star.

the last decade, to deal with these problems. Finally, in section 6, we conclude by suggesting new possible research directions, mainly along the estimation of unknown model parameters, where the coupling of modeling and migration could be useful.

2. Hessian and linearized migration

Non-linear seismic inversion consists of minimizing the differences between the observed data $d^{\text{obs}}$ recorded at the surface and the computed data $d(m)$ generated in a given velocity model $m$, such that the objective function in the least-squares sense (Tarantola, 1987) is written as

$$ J(m) = \frac{1}{2} \|d(m) - d^{\text{obs}}\|^2. \quad (1) $$

The definition of the Hessian is given by the second derivative of the objective function with respect to the velocity model

$$ H(x, y) = \frac{\partial^2 J(m)}{\partial m^2}, \quad (2) $$

where $x$ and $y$ denote two spatial positions.
In the case of linear least-squares inversion, where \( d(m) = Mm \), then the Hessian is:

\[
H = M^T M,
\]

where \( M \) and \( M' \) represent the modeling and the migration operators, respectively. Under the Born approximation (single scattering), the velocity model is decomposed into two parts: \( m = m_0 + \delta m \), where \( m_0 \) is referred to the background model and \( \delta m \) to a velocity perturbation associated to the reflectivity (Fig. 4). The background model \( m_0 \) should contain the large wavelengths (low frequencies) of the velocity model and is classically obtained by travel time tomography (Bishop et al., 1985) or by migration velocity analysis techniques (Chauris et al., 2002; Mulder & ten Kroode, 2002; Shen & Symes, 2008; Symes, 2008b). Migration aims at finding the reflectivity model \( \delta m \), assuming a known smooth background model. In the linear case, the solution of equation 1 using the Hessian gives us the migration image as

\[
\delta m = -H^{-1} M' \left( d - d^{\text{obs}} \right).
\]

If migration is only obtained as the gradient of equation 1 with respect to the model \( m \)

\[
(\nabla J(m) = \frac{\partial J(m)}{\partial m}),
\]

then only the kinematic part of the migration is retrieved (Lailly, 1983; Tarantola, 1987):

\[
\delta m = -K \nabla J(m) = -KM' \left( d - d^{\text{obs}} \right),
\]

where \( K \) is a positive matrix. The application of the inverse of the Hessian yields better migration estimates, by getting a balance between amplitudes at shallow and deeper depths. As it will be discussed in section 4, the Hessian matrix is mainly diagonally banded. For simple models, its scaling properties are contained in the diagonal terms of the Hessian, while the non-diagonal terms take into account the limited-bandwidth of the data. The application of the Hessian in the inversion can be seen as a deconvolution process. Moreover, as indicated by Pratt et al. (1988), the Hessian can also potentially deal with multiscattering effects, such as multiples.
Fig. 4. Exact velocity model (top left), smoothed velocity model (top right), difference between the exact and the smooth velocity models (bottom left) and filtered version of the model difference to get a reflectivity model by taking into account the finite-frequency behavior of the migration result (bottom right).

The exact expression of the Hessian in the linear inversion case can be written using the Green’s functions (Plessix & Mulder, 2004; Pratt et al., 1988)

\[
H(x, y) = \sum_\omega \omega^4 |S(\omega)|^2 \sum_s G^*(s, y, \omega)G(s, x, \omega) \cdot \sum_r G^*(r, y, \omega)G(r, x, \omega),
\]

where \(s\) and \(r\) correspond to the source and receiver coordinates, \(S\) to the source term and \(\omega\) to the angular frequency. The star symbol denotes the complex conjugate. The diagonal term is

\[
H(x, x) = \sum_\omega \omega^4 |S(\omega)|^2 \sum_s |G(s, x, \omega)|^2 \cdot \sum_r |G(r, x, \omega)|^2.
\]

The physical meaning of the Hessian is presented in (Pratt et al., 1988; Ravaut et al., 2004; Virieux & Operto, 2009). Applied to a Dirac velocity perturbation, it provides the resolution operator. Fig. 5 displays the Hessian in a 1-D homogeneous model. For a delta-type source, the Hessian is diagonal and constant along the diagonal as there is not decay in amplitudes in a 1-D propagation case. The band-limited source, here a Ricker with a maximum frequency of 30 Hz, introduces non-zero terms off the main diagonal (Fig. 5, right).
Fig. 5. Hessian matrices for a 1-D homogeneous model with a delta-type source (left) and with a Ricker source with frequencies up to 30 Hz (right).

Fig. 6. Homogeneous (left) and heterogeneous (right) models used for the Hessian operator computations. Two points are selected and marked in the models.

We study the Hessian in 2-D in two different models, an homogeneous model at 1.9 km/s and the same model with a velocity perturbation of 1 km/s in the central part (Fig. 6). The maximum frequency of the data is 30 Hz. The Hessian remains mainly diagonal (Fig. 7), with an amplitude decaying with depth due to the geometrical spreading of energy and to the acquisition at the surface. The same structure is also observed in (Pratt et al., 1988; Ravaut et al., 2004; Virieux & Operto, 2009). Non-diagonal terms are present due to the band-limited data (up to 30 Hz) and to the heterogeneity of the model.

Finally, \( H(x, x_0) \) is represented for fixed \( x_0 \) at either positions (250, 150) or (350, 350) meters (Fig. 6). The resolution degrades with depth and is function of the velocity model used to compute the Green’s functions (Fig. 8). These results are consistent with those obtained by Ren et al. (2011). From these illustrations, it appears that a good approximation of the Hessian, as for example proposed by Plessix & Mulder (2004), should take into account...
3. Pre-processing for reducing migration artifacts

The quality of a migrated image is strongly influenced by uneven or partial illumination of the subsurface, which creates distortions in the migrated image. Such partial illumination can be caused by the complexity of the velocity model in the overburden, as well as by limited or irregular acquisition geometries. In the case of complex overburdens, the partial illumination is due to strong velocity variations that prevent the seismic energy either from reaching the...
reflectors or from propagating back to the surface, where it is recorded. The irregularity of the data spatial sampling is instead mainly due to practical acquisition constraints, such as truncated recording aperture, coarse source-receiver distributions, holes due to surface obstacles or cable feathering in marine acquisitions. In both cases, with complex overburden and with poorly sampled data, the resulting effect is that strong artifacts degrade the migrated images (Nemeth et al., 1999; Salomons et al., 2009). Fig. 9 shows an example of acquisition artifacts in a common-offset migrated section where 50 input traces are missing in the central part (Fig. 9, right). Compared with the migrated section with all traces (Fig. 9, left), we note that artifacts are localized around different positions, as a function of the reflectivity and the model used for the migration. In milder cases, when distortions are limited to the image amplitudes, true-amplitude imaging processes can be employed (see section 4 for more details). Otherwise, data need to be regularized prior to imaging. In this section, we consider the data interpolation methods that combine the migration $M'$ and demigration $M$ operators. Seismic reflection data acquired on an irregular grid are migrated (using a velocity model as accurate as possible) and then demigrated with the same model back into the data space onto a regular grid. In this case, the application of $WMM'$ is required, where $W$ is a (diagonal) weighting matrix with zero weights for dead traces and non-zero weights for live traces according to their noise level, i.e. the inverse of the standard deviation of the noise in the data (Kühl & Sacchi, 2003; Trad, 2003). Note that this combined operator, defined in the data domain, can also be interpreted as the Hessian of an alternative objective function (Ferguson, 2006). Although expensive, this process provides a good data interpolation (and even extrapolation) technique because it accounts more correctly for the propagation effects in the reflector overburden (Santos et al., 2000). With the use of the same model $m$ for the modeling and migration parts, the kinematic aspect of the wave propagation is preserved (Bleistein, 1987). Moreover, interpolation using migration followed by demigration allows to model only those events that the migration operator can image: the demigration result is thus free of multiples (Duquet et al., 2000).

Several approaches have been proposed in the literature for implementing these techniques. We can distinguish between methods that rely on a direct inversion of the combination of migration and demigration (Ferguson, 2006; Stolt, 2002) and methods that consecutively

![Fig. 9. Example of a migrated section with all input traces (left) and with 50 missing traces in the central part (right). The main differences are underlined with red circles.](image-url)
apply the migration and modeling operators to reconstruct the data at the new locations. The latter methods can be separated into algorithms based on partial prestack migration (Chemingui & Biondi, 2002; Ronen, 1987) and those based on full prestack migration, either with a Kirchhoff operator (Duquet et al., 2000; Nemeth et al., 1999; Santos et al., 2000) or with a wavefield-continuation operator (Kaplan et al., 2010; Kühl & Sacchi, 2003; Trad, 2003).

4. True-amplitude migration schemes

The application of the inverse of the classical Hessian can be seen as a deconvolution step applied to the migration result (Aoki & Schuster, 2009). It corrects for uneven subsurface illumination (due to energy spreading and heterogeneous velocity models), takes into account the limited and non-regular acquisition geometry, and potentially increases the resolution. An extremely rich literature is available on this subject. We cite in this section the key references and detail some of them.

We distinguish between approaches based on the high frequency approximation (Beylkin & Burridge, 1990; Lecomte, 2008; Operto et al., 2000), and on the wave-equation approximation (Ayeni & Biondi, 2010; Gherasim et al., 2010; Valenciano et al., 2006; Wang & Yang, 2010; Zhang et al., 2007). Recent extensions have been proposed for sub-surface offsets (Valenciano et al., 2009).

An interesting approach has been developed in Jin et al. (1992); Operto et al. (2000), where the authors have proposed to modify the original objective function

$$ J_Q(m) = \frac{1}{2} \| Q \left( d(m) - d_{\text{obs}} \right) \|^2, \quad (8) $$

such that the Hessian becomes diagonal. This is possible by choosing a correct weighting factor $Q$ in the context of ray theory. The estimation of the Hessian reduced to a diagonal term is the way to correct for illumination, but it is valid only under the high frequency approximation and an infinite acquisition geometry (Lecomte, 2008). For band-limited data, other non-diagonal terms should be considered (Chavent & Plessix, 1999; Symes, 2008a; Virieux & Operto, 2009). Different strategies have been developed for estimating the non-diagonal terms (Kiyashchenko et al., 2007; Operto et al., 2006; Plessix & Mulder, 2004; Pratt et al., 1988; Ren et al., 2011; Shin et al., 2001; Yu et al., 2006), among them the mass lumping technique (Chavent & Plessix, 1999) and the phase encoding (Tang, 2009). In practice, the estimation of the pseudo-inverse of the Hessian remains a difficult task, as the operator is large and ill-conditioned. Alternatives have been proposed to avoid the computation of the Hessian. A first approach consists of iteratively minimizing equation 1 using a gradient approach, as done in equation 5. An example is given in Fig. 10. Starting from a homogeneous model close to the exact model, $J$ is minimized with a simple non-linear steepest descent algorithm. The model is laterally invariant, with a velocity perturbation around 400 m depth. A single shot with a maximum offset of 2 km was used. After a single iteration (Fig. 10, middle), the position of the top interface is correctly retrieved. This corresponds to the kinematic migration. After 100 iterations (Fig. 10, right), the velocity jump at the top interface is also well retrieved (+100 m/s). Since all frequencies up to 30 Hz were used at the same time, it is not possible to fully update the smooth part of the velocity model. For that reason, the second interface around 500 m is positioned at about 10 m above the exact location. More importantly here, the velocity jump is under-estimated, because no Hessian has been applied to correctly balance amplitudes (Fig. 10, right). A quasi-Newton approach (Pratt et al., 1988)
would have been more suited in that respect, since the preconditioning of the gradient by an approximated inverse Hessian yields improved convergence rates in iterative methods.

Other approaches for the estimation of the Hessian matrix (Guitton, 2004; Herrmann et al., 2009; Nemeth et al., 1999; Rickett, 2003; Symes, 2008a; Tygel et al., 1996) consist of migrating and demigrating a result several times and of computing optimal scaling and filtering operators. This is valid in the case of single scattering. A recent article exactly shows the type of scaling and filtering to apply (Symes, 2008a). The first step consists of performing a classical prestack migration with \(m_{\text{mig}} = M'd\) and, from this result, of regenerating data with the adjoint operator \(d_{\text{new}} = Mm_{\text{mig}}\). A second migration is run to obtain \(m_{\text{remig}} = M'd_{\text{new}}\). Then the inverse Laplace filter \(\text{Lap}^{-(n-1)/2}\) is applied \(m_{\text{filt}} = \text{Lap}^{-(n-1)/2}m_{\text{remig}}\), where \(n\) is the space dimension. In that case, \(m_{\text{filt}}\) and \(m_{\text{mig}}\) are very similar except for a scaling factor \(S\): \(Sm_{\text{filt}} = m_{\text{mig}}\). The final result is obtained as \(m_{\text{inv}} = Slap^{-(n-1)/2}m_{\text{mig}}\). According to Symes (2008a), this strategy is successful if the migrated result consists of nicely defined dips (see Fig. 11). For this reason, curvelet or space-phase domains are well suited for these types of applications (Herrmann et al., 2009). Curvelets can be seen as an extension of wavelets to multi-dimensional spaces and are characterized by elongated shapes (Candès et al., 2006; Chauris & Nguyen, 2008; Do, 2001; Herrmann et al., 2008). All curvelets can be deduced from a reference one (Fig. 12). For true-amplitude purposes, curvelet should be understood in a broad sense as being close to the representation of local plane waves. We refer to curvelets in the next section for other applications. To summarize the approach, the effect of the inverse of the Hessian can be obtained through two migration processes and a modeling step. The scaling part only is not sufficient. A Laplace operator needs also to be applied.
Coupling Modeling and Migration
for Seismic Imaging

Fig. 11. Seismic data gathers can be seen as a combination of local event “curvelets”, both in the unmigrated (left) and migrated (right) domains.

5. Image sensitivity

Starting from a reference migrated section, the objective of the image sensitivity techniques presented in this section is to predict the migrated section that would have been migrated with a different velocity model. For example, the migrated image in Fig. 13 (left) was obtained by using a smooth version of the exact velocity model, whereas the second migrated section (Fig. 13 right) was built with a homogeneous model. The two gathers clearly differ in terms of positioning and focusing.

In this section, we study the extended Hessian operator $H = M'[m + \delta m] M[m]$. This approach is an alternative to fully migrate the same input data for different velocity models, even though other efficient strategies have been proposed in that direction (Adler, 2002). An important aspect of the techniques based on the extended Hessian operator is that the kinematic of events remains the same through the migration/modeling operator for the same background velocity model $m$ (Bleistein, 1987). Original ideas were first developed in the case of time migration (Fomel, 2003b). The extended Hessian $H$ can be simplified in different ways, depending on the approximation behind the modeling and migration operators. For example, in the work of Chauris & Nguyen (2008), the operator $H$ has a very simple shape. For that, the authors use ray tracing (high frequency approximation) and decompose the reference migrated image into curvelets (Fig. 12). The application of $H$ to a curvelet is restricted to a shift, a rotation and a stretch of that curvelet. In practice, the model perturbation $\delta m$ should be small for this method to be valid. With this strategy and for a given velocity anomaly, it is possible to predict which part of the migrated section is affected (Fig. 14). As for the approaches proposed by Symes (2008a) and Herrmann et al. (2009), a key aspect is to decompose the migrated image as a combination of local events such as curvelets. Then each curvelet is potentially distorted, if the rays connecting the curvelet to the surface penetrate the velocity perturbation. The spatial position and the orientation of the curvelets are thus important. In that context, the objective is to derive the dependency of the migrated image with respect to a given velocity anomaly.
Fig. 12. Representation of different curvelets in the spatial domain. They all can be deduced from the reference curvelet (top left), either after translation/shift (top right), rotation (bottom left) or dilation/stretch (bottom right).

Fig. 13. Migrated images with the same input data but with two different velocity models, the correct smoothed model (left) and a constant velocity model at 3 km/s (right).
Fig. 14. Part of the migrated image unperturbed (left) and perturbed (right) by a velocity anomaly.

Finally, we refer to Chauris & Benjemaa (2010), where the authors extend the method of Chauris & Nguyen (2008) to heterogeneous models in a wave-equation approach. They propose an approximation of the Hessian that can be efficiently computed. In that case, the model perturbation $\delta m$ can also be large. An example of sub-salt imaging with synthetic data is presented (Fig. 15). The first step consists of migrating the data in a given velocity model (here a smooth model that does not contain the salt body) for a series of different time-delays. A time shift is introduced during the imaging condition (Chauris & Benjemaa, 2010; Sava & Fomel, 2006). These images are considered as new input data. It is then possible to predict the new migrated section obtained in a different velocity model, at least from a kinematic point of view and with a slight frequency lost. When the new model is a smooth model with the salt body, interfaces below salt become visible (Fig. 15). In practice, the migration information is preserved on different time-delay sections, except when the exact model is used: in that case, most of the energy is concentrated around small time-delay values.

6. Discussion

In the case of single scattering, the Hessian has an explicit expression. We have reviewed different strategies to efficiently compute it or part of it, usually terms around the diagonal. However, for multiple scattering, e.g. in the case of multiples, the different approaches are not valid. Further work should be conducted along that direction. Pratt et al. (1988) indicated how to compute the Hessian without relying on the Born application. Alternatively, iterative processes for the resolution of the inverse problem potentially may deal with multiple scattering, but this should be further demonstrated.

The aim of this work is reviewing methods that combine the migration and modeling operators for seismic imaging purposes. However, it is worth noting that for seismic processing tasks several approaches exist that use a specific operator and its adjoint, particularly for data interpolation (Berkhout & Verschuur, 2006; Trad et al., 2002; van
Fig. 15. Exact migrated section (top left), exact velocity model (top right), migration result with the initial velocity model (bottom left), and migration result after demigration/migration (bottom right).

Groenestijn & Verschuur, 2009), for multiple prediction (Pica et al., 2005; van Dedem & Verschuur, 2005) and for signal/noise separation (Nemeth et al., 2000). Moreover, we refer to Fomel (2003a) for other applications (stacking, redatuming, offset continuation), for which a technique is proposed to obtain a unitary modeling operator in the context of high frequency approximation.

In the developments mentioned above, the background velocity model is supposed to be known. In the context of velocity model building, we think interesting research directions should be developed along that line. For example, full waveform inversion is a general technique to retrieve the Earth’s properties. However, the objective function is very oscillating and a gradient approach for the minimization leads to a local minimum. Alternative approaches have been proposed, among them Plessix et al. (1995). The Migration Based Travel Time (MBTT) method first migrates the data, and then uses the stack version to generate new data. The objective function consists of minimizing the differences between the new data and the observed data. As a benefit and compared to the classical method, it enlarges the attraction basin during the minimization process. We believe further work in that direction can deliver interesting results.
7. Conclusion

Extended research has been conducted around applying the Hessian operator in the context of pre-processing/interpolation techniques for reducing migration artifacts (section 3) and true-amplitude migration (section 4). In fact, with the use of Hessian, it is possible to correct for a limited acquisition, to provide more reliable amplitudes and to increase the resolution. However, we believe that the extended Hessian operator (refer to section 5) is a powerful tool for model estimation and that further research should be conducted along that line in the coming years.

8. Acknowledgements

The authors would like to thank a number of persons for fruitful discussions and new insights into seismic imaging. They are especially grateful to Henri Calandra (Total), Eric Dussaud (Total), Fons ten Kroode (Shell), Gilles Lambaré (CGGVeritas), Patrick Lailly (IFP), Wim Mulder (Shell), Mark Noble (Mines ParisTech), Stéphane Operto (Géoazur), René-Edouard Plessix (Shell), Bill Symes (Rice university), Jean Virieux (Grenoble university) and Sheng Xu (CGGVeritas).

9. References


The importance of seismic wave research lies not only in our ability to understand and predict earthquakes and tsunamis, it also reveals information on the Earth's composition and features in much the same way as it led to the discovery of Mohorovicic's discontinuity. As our theoretical understanding of the physics behind seismic waves has grown, physical and numerical modeling have greatly advanced and now augment applied seismology for better prediction and engineering practices. This has led to some novel applications such as using artificially-induced shocks for exploration of the Earth's subsurface and seismic stimulation for increasing the productivity of oil wells. This book demonstrates the latest techniques and advances in seismic wave analysis from theoretical approach, data acquisition and interpretation, to analyses and numerical simulations, as well as research applications. A review process was conducted in cooperation with sincere support by Drs. Hiroshi Takenaka, Yoshio Murai, Jun Matsushima, and Genti Toyokuni.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: