A visualization of the relationship between Kirchhoff migration and seismic inversion

John C. Bancroft - CREWES, University of Calgary

Abstract

The inversion of seismic data may be approximated by a Kirchhoff migration process. The kinematics of this process is illustrated using cartoon images of 2D data.

Introduction

This paper represents an abbreviated form of a paper that appeared recently in the December 2001 issue of the Recorder.

The term inversion is used in a number of different applications in geophysics such as

1. Estimating source and receiver statics from cross-correlation statics (trims)
2. Deconvolution
3. Converting seismic data to a geological subsurface image

These applications often use the mathematics of least squares inversion in which many observations are processed to extract the best estimate of a number of parameters, i.e. over-determined where we have many more equations than variables.

The last two applications may be combined for a definition of geophysical inversion that produces "models of the earths physical properties", (Lines 1992) from observations collected at the earth's surface or in a borehole. These processes include:

- Lindseth's method (1979) of estimating impedance from combining an "integrated" seismic trace with a low frequency velocity model,
- Line's method (1988) of estimating a geological model of horizons and interval velocities from seismic and potential field data, or a
- Kirchhoff type migration that attempt to produce an impedance image of the subsurface from seismic data.

Figure 1a illustrates the forward modelling process in which seismic data is created from reflectivity and a wavelet. Figure 1b illustrates the reverse process of inversion in which geological data is estimated from the seismic data.

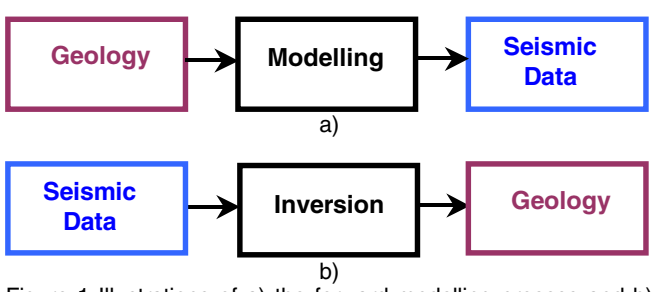

Figure 1 Illustrations of a) the forward modelling process and b) the reverse inversion process

Note the use of the term "estimate" in the inversion process. The forward modelling process uses established principles of physics to create exact seismic data. However, the reverse process is more difficult, usually because we don't know the wavelet, and some of the operations are arithmetically inverted and we get division by zeros. Consequently, approximations are included in inversion process to stabilize the results. Some of these approximations are illustrated in this paper using cartoon drawings to illustrate how Kirchhoff migration is an approximation to the inversion process.

In contrast to inversion, I think of seismic migration as a process that uses the wave equation to reposition or focus seismic energy at a location that represents the reflectors. The term imaging may be used to include both migrations and inversions.

Kirchhoff migration is based on the integral solution to the wave equation and has a kinematic and amplitude parts to the solution. Every migrated sample results from the amplitude weighting and summing of energy along a diffraction shape that is defined by the location of the migrated sample. The kinematics define the traveltimes of a diffraction shape, while the amplitude part of the solution provides an amplitude weighting that is applied along the diffraction. Modern inversion techniques (Bleistein 2001) also result with the same Kirchhoff type solution but requires "true amplitude" type processing and may apply different amplitudes weightings along the diffraction.

Convolution model and geophysical inversion

I will start with the convolutional model for a one dimensional trace to illustrate the inversion process and then proceed to the 2D model of seismic data and its inversion. The seismic data *s(t)* is formed by convolving the reflectivity *r* with a wavelet *w(t)*. Ignoring for the moment the spatial dimensions of *r*, i.e.,

$$s(t) = r * w(t) \qquad (1)$$

If equation (1) is expressed in the frequency domain, the convolution becomes a product, i.e.,

$$S(f) = R \bullet W(f) \qquad (2)$$

This linear form of the equation enables a simple estimation of the reflectivity at each frequency from

$$R = \frac{S(f)}{W(f)} \qquad (3)$$

The actual reflectivity could then be found by summing all the frequency components using the Fourier transform. The problem with this procedure is that *D(f)* may go to zero at some frequency, and that we can't perform the division in equation (3) when there are zeros in *D(f)*. With inversion, as in deconvolution, the procedure becomes an estimation problem that is solved with a variety of methods

Linearization of convolution using matricies

We had linearized the convolution process above by using the frequency domain. We can also linearize convolution with matrix theory. I will start with a one dimensional model where the reflectivity, wavelet, and trace are functions of time as defined by the convolution equation

$$s(t) = \int_{-\infty}^{\infty} r(\tau) * w(t-\tau) d\tau \,, \qquad (4)$$

where we integrate over τ to compute one sample on the seismic trace *s(t)*, as illustrated in Figure 2. Corresponding time samples in *r* and w_r are first multiplied then summed to get the single n^{th} sample, i.e. the dot product. We then compute all values of *t* on the trace to complete the convolution.

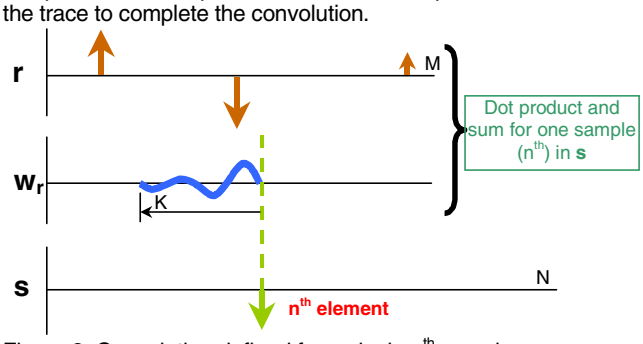

Figure 2 Convolution defined for a single n^{th} sample.

All elements of the seismic vector can be computed by repeating the process for all value of *n*, as illustrated in Figure 3 that shows numerous wavelets with varying delays. In this cartoon figure, (and those that follow), the sample interval is much finer along the rows to define the wavelet than between the rows for illustration purposes, but the intent is to imply that the reversed wavelet increases one sample to the right when progressing to the next row, i.e. at the nth row, the right side of the wavelet is at the nth sample. The sample by sample product of the reflectivity vector with the wavelet at the nth row, (containing the enlarged wavelet in blue), is summed then stored at the nth location in the seismic vector.

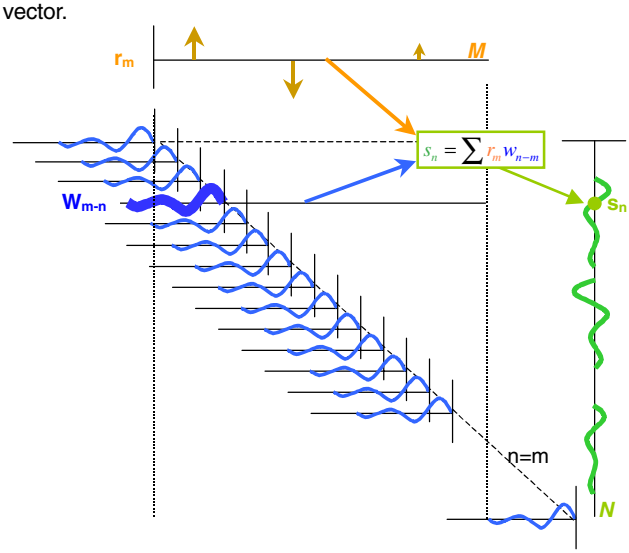

Figure 3 Computing other samples in convolutional model.

The wavelet data illustrated in Figure 3 is in a convenient form to be defined by a two dimensional matrix **W** that has *N* rows and *M* columns, with the number of columns defined to match the number of elements in the reflectivity vector. Each row in the matrix defines at least a portion of the delayed and time reversed wavelet w_{m-n}, where *n* represents the row number and *m* the column number. All the elements on a given diagonal of **W** will be the same, i.e. **W** is a Toeplitz matrix. To be consistent with matrix algebra, we define the reflectivity as a column vector **r**, and the seismic trace as a column vector **s**. The matrix equation for the convolution process then becomes a linear process

$$\underset{[N\times M]}{\mathbf{W}}\ \underset{[M]}{\mathbf{r}} = \underset{[N]}{\mathbf{s}} \qquad (5)$$

where, for convenience, the dimensions are shown in square brackets below the matrix and vectors. Equation (5) is visualized in Figure 4 similar to that in Margrave 1998.

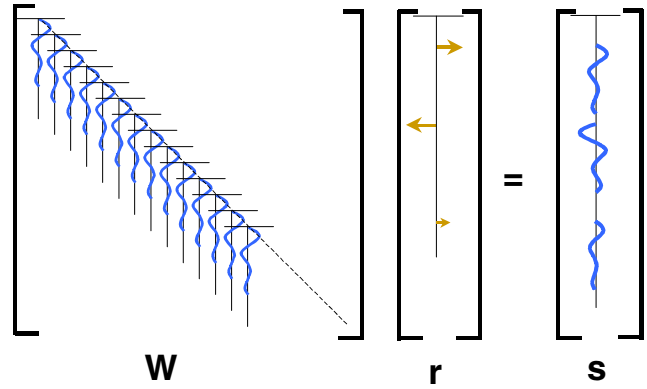

Figure 4 Matrix form of the convolutional model

Reversing the modelling process

It was easy to do the forward modelling where the seismic trace was computed from a reflectivity and wavelet using equation (4). We now "reverse" this modelling process to estimate the reflectivity from a given seismic trace and known wavelet. Lets start by trying to take the inverse of **W** in equation (5) to get

$$\underset{[M]}{\mathbf{r}} = ?\ \underset{[M\times N?]}{\mathbf{W}^{-1}}\ \underset{[N]}{\mathbf{s}}\ . \qquad (6)$$

That probably won't work because we can't take the inverse of a matrix that is not square, and we would have to make *M* = *N*. That could be solved, but the biggest problem is that **W** may contain zeros, which, when inverted, would become infinite. There are a number of ways to modify **W** to make it invertable such as adding small numbers to eliminate the zeros (such as prewhitening), however, in this paper I will only consider the linear process that leads to the "least squares" solution.

Least squares solution

One solution to this problem is to multiply both sides of equation (5) by the transpose of **W**,

$$\underset{[M\times N]}{\mathbf{W}^T}\ \underset{[N\times M]}{\mathbf{W}}\ \underset{[M]}{\mathbf{r}} = \underset{[M\times N]}{\mathbf{W}^T}\ \underset{[N]}{\mathbf{s}}\ , \qquad (7)$$

and then inverting $\mathbf{W}^T\mathbf{W}$ giving the well known solution (Lines 1984).

$$\underset{[M]}{\mathbf{r}} = \left(\underset{[M\times N]}{\mathbf{W}^T}\ \underset{[N\times M]}{\mathbf{W}}\right)^{-1}\ \underset{[M\times N]}{\mathbf{W}^T}\ \underset{[N]}{\mathbf{s}}\ . \qquad (8)$$

The product $\mathbf{A} = \mathbf{W}^T\mathbf{W}$ produces a square matrix which can be inverted only when there are no zeros in the spectrum of the auto-correlation wavelet.

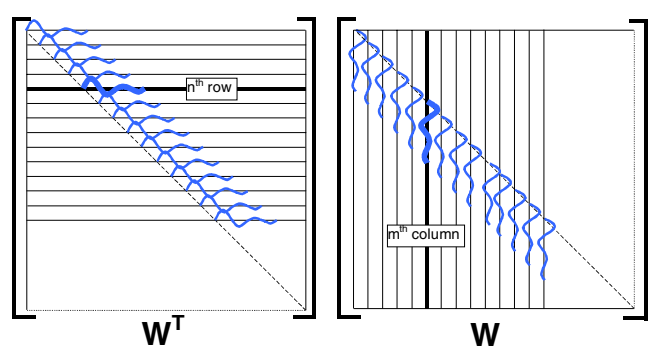

Figure 5 Illustration of $\mathbf{W}^T$ and **W** with corresponding rows and columns displayed for taking the product of the two matrices.

The square matrix **A** is shown in Figure 6 with vertical traces that illustrate that all the new wavelets are the delayed auto correlation of the original wavelet. Plotting the data with horizontal traces should produce the same image (with adequate column sampling).

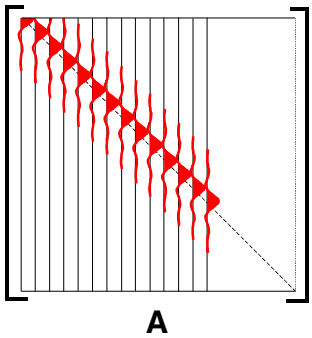

Figure 6 The result of **A** = $\mathbf{W}^T\mathbf{W}$, the product of a matrix and its transpose.

All elements on any diagonal in **A** are equal (also Toeplitz) with the largest value on the diagonal. This is very significant as **A** is similar to the identity matrix **I** in which the data is zero, except for unit values on the diagonal, as illustrated in Figure 7.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Figure 7 An identity matrix.

A feature of the identity matrix is that the inverse is also an identity matrix, i.e. $\mathbf{I}^{-1} = \mathbf{I}$. If we assume that the auto-correlation matrix **A** is an approximation to the identity matrix **I** (except for a scale factor), we can now assume

$$\underset{[M\times N]}{\mathbf{W}^T}\underset{[N\times M]}{\mathbf{W}} = \underset{[M\times M]}{\mathbf{A}} \approx \underset{[M\times M]}{\mathbf{I}}, \qquad (9)$$

and also that

$$\left(\underset{[M\times N]}{\mathbf{W}^T}\underset{[N\times M]}{\mathbf{W}}\right)^{-1} \approx \underset{[M\times M]}{\mathbf{I}}. \qquad (10)$$

This simplification allows equation (8) to be reduced to a very simple form

$$\underset{[M]}{\mathbf{r}} \approx \underset{[M\times N]}{\mathbf{W}^T}\underset{[N]}{\mathbf{s}}, \qquad (11),$$

which allows us to define an estimate of the reflectivity $\hat{\mathbf{r}}$ by

$$\underset{[M]}{\hat{\mathbf{r}}} = \underset{[M\times N]}{\mathbf{W}^T}\underset{[N]}{\mathbf{s}}. \qquad (12)$$

The reflectivity can be estimated from the product of the seismic vector **s** with the transpose of the wavelet matrix $\mathbf{W}^T$. Our simplifications have assumed that the inverse to a matrix $\mathbf{W}^{-1}$ can be approximated by a very simple transpose $\mathbf{W}^T$, i.e.,

$$\underset{[M\times N]}{\mathbf{W}^T} \approx \underset{[M\times N?]}{\mathbf{W}^{-1}}. \qquad (13)$$

The simplifications leading to equation (13) produced a band-limited form of the reflectivity, and is identical to that obtained with a matched-filter.

Kirchhoff migration as a transpose process

There are many "reverse" processes that use the concept of approximating the inverse with a transpose. What about seismic modelling and migration? In seismic modelling we place a diffraction at every scatterpoint. When using Kirchhoff migration, we define a kinematic diffraction shape for a scatterpoint, then weight and sum the input energy defined by this shape, and insert the energy at the scatterpoint location. This is in essence, a two-dimensional cross correlation, and will produce a peak of energy when the model diffraction matches a diffraction in the input data, i.e. matched filtering, or by approximating the inverse process with a transpose process. We are now in a position to visualize and evaluate the limitations of our migration algorithms with true inversion.

Visualizing modelling data with diffractions

We will start by modelling a gather of reflectivity traces with a wavelet matrix **W** in which the wavelets are time varying to get a gather of traces **s** as illustrated in Figure 8

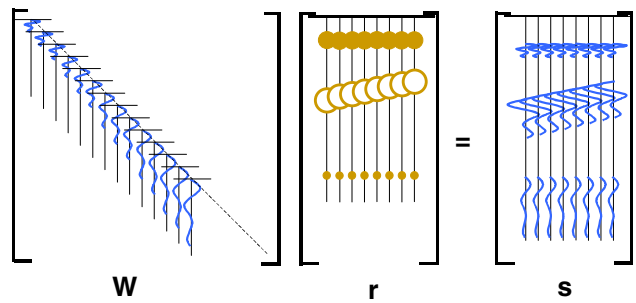

Figure 8 Matrix form of the convolutional model with time varying wavelets.

We now replace the time varying wavelet with a time varying diffraction, and require an added dimension to account for the number of traces in the diffraction. A side view is shown in Figure 9a that shows the diffraction as a time varying "wavelet", similar to that in Figure 8. The full 3D diffraction model is illustrated in Figure 9b, which shows the kinematic shape of three diffractions. The **D** matrix is mainly composed of null space, with scaled values on a surface, which, when the velocities are constant, is a cone with a vertical axis.

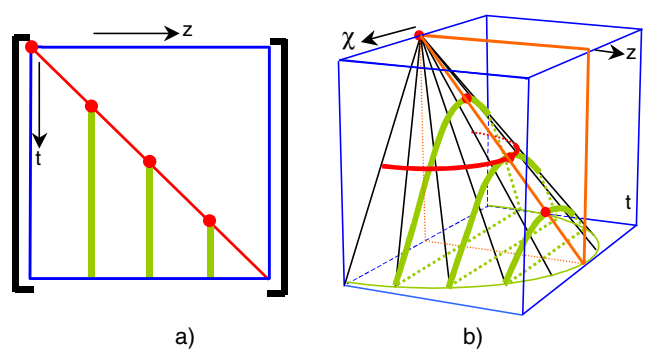

Figure 9 The diffraction matrix **D**, showing a) the (z, t) view and b) a perspective view showing the added dimension of χ.

Seismic modelling is illustrated in Figure 10 where one sample in **S** at $s(x_p, t_q)$ is formed when the 2D matrix **R** is dot multiplied by one plane of the 3D diffraction matrix **D** at constant time *t*. The matrix **R** is aligned such that the location of the migrated trace coincides with the location of $\chi = 0$, as illustrated in Figure 10 that

shows the **R** matrix above the diffraction matrix **D**. As different migration traces are selected, the matrix **R** above **D** is shifted in the corresponding χ direction. For example, when the red trace in **S** is being evaluated, the corresponding red trace in $\mathbf{R}^T$ is located above $\chi = 0$.

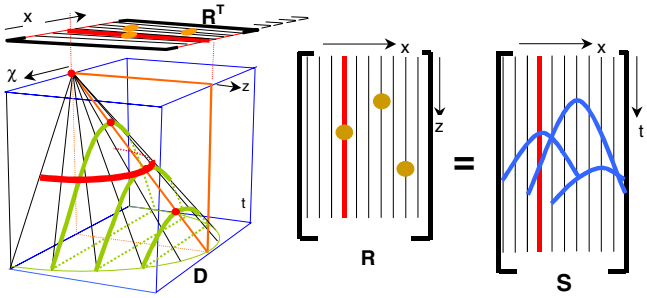

Figure 10 Matrix view to obtain a zero-offset section **S** from a reflectivity matrix **R** and a diffraction matrix **D**.

The time of a modelled sample t_q is selected by the horizontal plane within **D**, and the trace location x_p by the relative alignment of that trace above $\chi = 0$ as given by

$$s(x_p, t_q) = \sum_{\chi_i=-I}^{I} \sum_{z_j=1}^{J} d(\chi_i, z_j, t_q) r\left(x_p - \chi_i, z_j\right). \quad (14)$$

A constant time (slice) in **D** intersects the diffraction cone at a semi-circle, as identified by one red curve in the perspective view of Figure 10. This corresponds to the method of producing a seismic modelled section by summing reflector energy over a semi-circle, i.e. when creating a seismic section from modelling, we can either sum scatterpoints over a semi-circle to get one modelled sample, or spread energy from a scatterpoint along a hyperbolic diffraction.

Let's return to matrix theory for a moment, and see what happens when we take the transpose of **D** to estimate a migrated section from a zero-offset seismic section, i.e.,

$$\underset{[x,z]}{\hat{\mathbf{r}}} = \underset{[x,z,t]}{\mathbf{D}^{\mathbf{T}}} \underset{[x,t]}{\mathbf{s}} . \quad (15)$$

This result is illustrated in Figure 11 where $\mathbf{D}^T$ contains a cone with a horizontal axis. Now, the section **s** is dot multiplied with constant depth planes (slices) and the energy is summed where the horizontal plane intersects the cone, illustrated at one depth level by a red hyperbola. The equation for migration one sample $r(x_p, z_q)$ is

$$\hat{r}(x_p, z_q) = \sum_{\chi_i=-I}^{I} \sum_{t_j=1}^{J} d(\chi_i, z_q, t_j) s\left(x_p - \chi_i, t_j\right). \quad (16)$$

The dual property to hyperbolic summation in migration is the spreading the energy along a semi-circle, defined in Figure 11 by vertical planes in $\mathbf{D}^T$ at constant time.

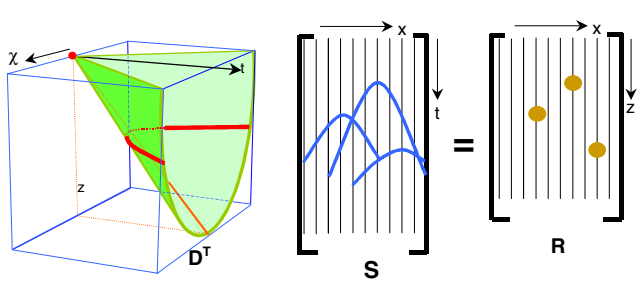

Figure 11 Matrix view to obtain a zero-offset section **S** from a reflectivity matrix **R** and a transpose diffraction matrix $\mathbf{D}^T$.

Comments and conclusions

The kinematics of diffraction stack migrations are straight forward, however the amplitude weightings used when summing the diffractions are still under investigation. Migrations that are derived from solution to the wave equation include weighting schemes. The integral solution of the wave equation changed the heuristic diffraction stack process to the mathematically deterministic Kirchhoff algorithm.

Inversion techniques provide an alternate solution when imaging the subsurface, and in some cases produce algorithms that are very similar to Kirchhoff migrations. Approximate linear algebra inversions were shown to be identical to matched filters, and the difference from an ideal inversion identified.

References

Bleistein, N., Cohen, J. K., and Stockwell, Jr., J. W., 2001, Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion, Springer

Cary, P. W., and Li, X., 2001, Some basic imaging problems with regularly-sampled seismic data, Exp. Abs., SEG Int. Nat. Exp., 981-984

Claerbout, J. F., 1992, "Earth Soundings Analysis: Processing versus inversion", Blackwell Scientific Publications

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping, Geophysics, 36, 467-481

Dellinger, J., Gray, S., Murphy, G., Etgen, J., and Fei, T., 1999, Efficient two and one-half dimension true-amplitude migration, Exp. Abs., SEG Nat. Conv., 69, 1540-1543

Gazdag, J., and Squazzero, P., 1984, Migration of seismic data, Proceedings of the IEEE, 72, 1302-1315

Geiger, H. D., 2001, Relative-amplitude preserving prestack time migration by the equivalent offset method, Ph.D. Thesis, Department of Geology and Geophysics, University of Calgary

Larner, K., and Hatton, L., 1990, Wave-equation migration: two approaches (1976), First Break, 8, 443-448

Lindseth, R. O., 1979, Synthetic sonic logs-a process for stratigraphic interpretation, Geophysics, 44, 3-26

Lines, L. R., and Levin, F. K., 1992 Inversion of Geophysical Data, Geophysical reprint series, No. 9, SEG

Lines, L. R., Schultz, A. K. and Treitel, S., 1988, Cooperative inversion of geophysical data: Geophysics, 53, 8-20.

Lines, L. R., and Treitel, S., 1984, Tutorial: a review of least-squares inversion and its application to geophysical problems, Geophysical Prospecting, 32, 159-186

Margrave, G. F., 1998, Theory of non-stationary linear filtering in the Fourier domain with applications to time-variant filtering, Geophysics, 63, 255-259

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions, Geophysics, 43, 49-76

Ziemer, R. E., and Tranter, W. H., 1995, Principles of Communications, Wiley