Fast high resolution parabolic Radon transform

*Mauricio D. Sacchi *, University of Alberta, and Milton Porsani, Universidade Federal Da Bahia*

Summary

The high resolution Parabolic Radon transform proposed by Sacchi and Ulrych (1995) entails the utilization of a regularization technique that leads to an operator that does not exhibit a Toeplitz structure. In the original formulation of the high resolution Radon transform the operator is inverted using Cholesky decomposition. This is quite expensive compared to the classical least squares Radon transform that uses the Levinson recursion to invert a Toeplitz form.

We propose a method to achieve high resolution at a computational cost of the order of the conventional parabolic least squares Radon transform. This feature makes our new algorithm quite attractive to process large data sets.

Introduction

The Parabolic Radon transform is a widely accepted technique for multiple removal (Hampson, 1986). The technique can be implemented in the frequency domain via a fast algorithm that exploits the Toeplitz structure of the least squares Radon operator (Kostov, 1990; Darche, 1990). Recently, Sacchi and Ulrych (1995) proposed a high resolution algorithm to increment the ability of the transform to distinguish events with similar moveout curves. This algorithm is based on a procedure that attempts to find a sparse representation of the reflections in the parabolic Radon domain. A similar algorithm has been proposed by Cary (1998). In this case the Radon panel is constrained to be sparse in both the Radon parameter and the intercept time.

The high resolution parabolic Radon transform can be used to isolate multiples interferences with a few milliseconds of residual moveout at far offset. This is a problem frequently encountered when dealing with short period multiple reflections generated by carbonate targets in the Western Canadian Basin (Hunt et al., 1996).

One of the advantages of the high resolution parabolic Radon transform is that the focusing power of the transform is considerably increased with respect to the classical least squares parabolic Radon transform. Unfortunately, the high resolution parabolic Radon transform leads to the inversion of an operator that is Hermitian but does not exhibit a Toeplitz structure. The resulting Hermitian operator is inverted using Cholesky decomposition. The Cholesky method for solving Hermitian linear systems of equations requires a number of operations that is proportional to M^3, where M is the dimension of the Hermitian operator.

In this paper we present a strategy to reduce the computational cost of the high resolution Radon transform. Our technique is based on the inversion of the Hermitian operator via the method of conjugate gradients with the addition of a fast matrix times vector multiplication using circulant matrices.

Least squares Parabolic Radon transform

Common mid point (CMP) gathers after normal moveout (NMO) correction can be modeled as a superposition of events with parabolic moveout:

$$d(x_j,t) = \sum_{k=1}^{M} m(q_k, \tau = t - q_k\, x_j^2)\,, j = 1, N\,, \qquad (1)$$

where $d(x_j,t)$ denotes the CMP gather, x_j the offset, $m(q_k,\tau)$ is the Radon panel, q_k the discrete Radon parameter and τ the intercept time. The data consist of N seismic traces which do not need to be regularly sampled. The Radon parameter is uniformly discretized according to $q_k = q_0 + \Delta q\,(k-1)$, $k = 1,\ldots,M$.

Equation (1) is essentially a decomposition of the CMP gather in terms of parabolic events distributed in the plane τ, q. It is computational more convenient to rewrite the last equation in the frequency-offset domain. Taking Fourier transform with respect to the temporal variable t we arrive to the following expression

$$d(x_j,f) = \sum_{k=1}^{M} m(q_k,f)\, e^{i2\pi f\, q_k x_j^2}\,, j = 1,\ldots,N\,. \qquad (2)$$

The calculations can be carried out independently for each frequency f. Equation (2) can be written in matrix form as follows:

$$\mathbf{d}(f) = \underset{\sim}{\mathbf{L}}(f)\,\mathbf{m}(f)\,. \qquad (3)$$

To avoid notational clutter we will drop the frequency dependency in equation (3) and write $\mathbf{d} = \underset{\sim}{\mathbf{L}}\,\mathbf{m}$.

The least squares Radon operator is estimated by minimizing the following cost function.

$$J = ||\mathbf{d} - \underset{\sim}{\mathbf{L}}\,\mathbf{m}\,||^2 + \mu||\mathbf{m}||^2\,. \qquad (4)$$

The regularization term $\mu||\mathbf{m}||^2$ is used to control the roughness of the solution. It can be shown that this term is one of the major sources of amplitude smearing in the Radon panel (Sacchi and Ulrych, 1995).

Taking derivatives of J with respect to $\mathbf{m}$ and equating them to zero yields

$$\begin{aligned}(\underset{\sim}{\mathbf{L}}^H \underset{\sim}{\mathbf{L}} + \mu \underset{\sim}{\mathbf{I}})\mathbf{m} &= \underset{\sim}{\mathbf{L}}^H \mathbf{d} \\ &= \mathbf{m}_{adj}\,. \end{aligned} \tag{5}$$

In the last equation $\mathbf{m}_{adj}$ denotes the low resolution Radon transform obtained using the adjoint or transpose operator $\mathbf{L}^H$. The least squares solution becomes

$$\begin{aligned}\mathbf{m} &= (\underset{\sim}{\mathbf{L}}^H \underset{\sim}{\mathbf{L}} + \mu \underset{\sim}{\mathbf{I}})^{-1}\mathbf{m}_{adj} \\ &= (\underset{\sim}{\mathbf{R}} + \mu \underset{\sim}{\mathbf{I}})^{-1}\mathbf{m}_{adj}\,. \end{aligned} \tag{6}$$

At this point some observations are in order. First it is clear that $\underset{\sim}{\mathbf{R}} = \underset{\sim}{\mathbf{L}}^H \underset{\sim}{\mathbf{L}} + \mu \underset{\sim}{\mathbf{I}}$ is a Toeplitz form (Kostov, 1990), with elements given by

$$\{\underset{\sim}{\mathbf{R}} + \mu \underset{\sim}{\mathbf{I}}\}_{l,m} = \sum_{k=1}^{N} e^{-i2\pi f \Delta q (l-m) x_k^2} + \mu \delta_{l,m}\,. \tag{7}$$

Solving this equation using the Levinson recursion requires approximately $4M^2 + 7M$ operations, and storage of only the first row of the Toeplitz matrix (Marple, 1987). This feature yields to a very efficient algorithm to compute the parabolic Radon transform.

High resolution parabolic Radon transform

In the high resolution parabolic Radon transform the vector $\mathbf{m}$ is retrieved by solving the following equation:

$$(\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{W}}^H \underset{\sim}{\mathbf{W}})\mathbf{m} = \mathbf{m}_{adj}\,. \tag{8}$$

The matrix $\underset{\sim}{\mathbf{W}}$ is a diagonal matrix with elements that depend on $\mathbf{m}$ (Sacchi and Ulrych, 1995). This leads to an iterative algorithm where $\underset{\sim}{\mathbf{W}}$ is bootstrapped from the result of a previous iteration. In general, the iterative procedure is not required if we are able to design $\underset{\sim}{\mathbf{W}}$ from a priori information. The matrix of weights $\underset{\sim}{\mathbf{W}}$ is a diagonal matrix with elements given by

$$\{\underset{\sim}{\mathbf{W}}\}_{l,m} = w_l\, \delta_{l,m}, \quad l, m = 1, \ldots, M\,. \tag{9}$$

The elements of the diagonal form $\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{W}}^H \underset{\sim}{\mathbf{W}}$ become:

$$\{\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{W}}^H \underset{\sim}{\mathbf{W}}\}_{l,m} = \sum_{k=1}^{N} e^{-i2\pi f \Delta q (l-m) x_k^2} + w_l^2 \delta_{l,m}\,. \tag{10}$$

It is clear that the addition of a diagonal matrix with non-constant elements has destroyed the Toeplitz structure of the operator. The above matrix can be inverted by the Cholesky method in a number of operations proportional to M^3.

From the computational point of view it is more convenient to compute the Radon transform using a constant diagonal regularization (equation (5)). However, if we want to estimate a high resolution Radon operator, the regularization term must be a diagonal form with non-constant elements (equation (9)). The elements of $\underset{\sim}{\mathbf{W}}$ are used to emphasize the Radon parameters q_k that need to be constrained to be zero. In general, the matrix $\underset{\sim}{\mathbf{W}}$ is bootstrapped from the data in an iterative manner. The aforementioned procedure is described in Sacchi and Ulrych (1995).

In our synthetic example, the elements of the diagonal matrix $\underset{\sim}{\mathbf{W}}^H \underset{\sim}{\mathbf{W}}$ are given by

$$w_k^2 = \begin{cases} 100. & if \;\; q_k \notin Q \\ 0.0001 & if \;\; q_k \in Q\,, \end{cases} \tag{11}$$

where Q indicates the set of parameters q_k where the reflections are localized. These weights can be interpreted as the inverse of a variance in model space. If w_l^2 is large, $1/w_l^2$ is small and therefore, the algorithm will constraint the areas of no reflections in the τ, q space to be zero. It is clear that the resolution is enhanced by inhibiting the creation of smearing in the Radon panel.

Conjugate gradients and circulant matrices

To solve equation (8) we adopt the method of conjugate gradients (see for instance, Strang, 1996), which is summarized below.

We want to solve $(\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{D}})\mathbf{m} = \mathbf{m}_{adj}$, where $\underset{\sim}{\mathbf{D}} = \underset{\sim}{\mathbf{W}}^H \underset{\sim}{\mathbf{W}}$.

Start with an initial solution $\mathbf{m}_0$, set $\mathbf{p}_0 = \mathbf{r}_0 = \mathbf{m}_{adj} - (\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{D}})\mathbf{m}_0$,

$$\alpha_{i+1} = (\mathbf{r}_i, \mathbf{r}_i)/(\mathbf{p}_i, (\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{D}})\mathbf{p}_i) \tag{12a}$$

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \alpha_{i+1}\mathbf{p}_i \tag{12b}$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_{i+1}(\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{D}})\mathbf{p}_i \tag{12c}$$

$$\beta_{i+1} = (\mathbf{r}_{i+1}, \mathbf{r}_{i+1})/(\mathbf{r}_i, \mathbf{r}_i) \tag{12d}$$

$$\mathbf{p}_{i+1} = \mathbf{r}_{i+1} + \beta_{i+1}\mathbf{p}_i \tag{12e}$$

where $i = 0, 1, 2, \ldots K$ denotes the iteration number.

The cost of the conjugate gradients algorithm is dominated by the cost of multiplying a matrix by a vector (12a). In general, matrix times vector multiplication is an $O(M^2)$ process. In our problem we will use the Toeplitz structure of $\underset{\sim}{\mathbf{R}}$ to find a fast manner to compute the aforementioned operation.

The product $(\underset{\sim}{\mathbf{R}} + \underset{\sim}{\mathbf{D}})\mathbf{x}$ can be decomposed into two products: $\underset{\sim}{\mathbf{R}}\mathbf{x} + \underset{\sim}{\mathbf{D}}\mathbf{x}$. The first product can be efficiently computed using the Fast Fourier Transform (FFT), the second product involves only $2M$ operations (M products plus M additions) and does not substantially increase the computational cost of the inversion.

The first product, $\mathbf{y} = \underset{\sim}{\mathbf{R}}\mathbf{x}$, is evaluated by augmenting the system as follows:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}' \end{bmatrix} = \underset{\sim}{\mathbf{R}}_{aug} \begin{bmatrix} \mathbf{x} \\ \mathbf{0} \end{bmatrix}, \qquad (13)$$

where $\underset{\sim}{\mathbf{R}}_{aug}$ is the original Toeplitz matrix after being properly folded to become a circulant matrix (Strang, 1986; Schonewille and Duijndam, 1998). The right hand side can be computed by multiplying the Fourier transform of the first row of $\underset{\sim}{\mathbf{R}}_{aug}$ by the Fourier transform of vector $[\mathbf{x}, \mathbf{0}]^T$, and taking the inverse Fourier transform of this product. Now our matrix times vector operation takes $O(M' \log M')$ operations where M' is the size of augmented matrix ($M' = 2M$). We have found that the conjugate gradients algorithm convergences after a few iterations ($K \approx M/5$). Therefore, the inversion becomes an $O(K\, M'\, Log(M'))$ process. This is more efficient than the direct inversion of equation (8) by the Cholesky method.

Example

In Table 1 we present a comparison of CPU times in seconds for 3 different algorithms. The times in Table 1 correspond to the total computational cost for 512 frequencies. These simulations were performed on a SGI Origin 2000.

In both cases we have 4 parabolic events which were mapped to the Radon domain using the following algorithms:

1. **Lev**: Classical least squares parabolic Radon transform implemented via the Levinson recursion (equation (5)).
2. **Chol**: High resolution Radon transform implemented via the Cholesky decomposition (equations (8) and (9)).
3. **CG+FFT**: High resolution parabolic Radon transform implemented via conjugate gradients plus matrix times vector multiplication using the FFT.

Is is clear that the new algorithm can achieved high resolution at a computational cost comparable to the one of the classical least squares Radon transform computed with the Levinson recursive solution.

In Figure 1 we portray the results obtained for the 256×256 simulation. Note that the differences between the high resolution Radon transform computed with the Cholesky decomposition and the proposed algorithm are minimal.

Conclusion

We have presented a new algorithm to compute the high resolution parabolic Radon transform. This algorithm operates at a speed that is comparable to the least squares Radon operator obtained by the Levinson recursion.

We have also shown the importance of using a non-constant diagonal regularization matrix to enhance the focusing power of the parabolic Radon transform. The efficiency of the high resolution parabolic Radon transform is improved by an order of magnitude with respect to the original algorithm based on a direct inversion using the Cholesky decomposition.

Acknowledgments

This research has been partially funded by Pan Canadian Petroleum, Geo-X Ltd., and a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Cary, P., W., 1998, The simplest discrete Radon transform, 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1999-2002.

Darche, G., 1990, Spatial interpolation using a fast parabolic transform: 60th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1647-1650.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: J. Can. Soc. Expl. Geophys., **22**, 1, 44-55.

Hunt, L., Cary, P., and Upham, W., 1996, An improved Radon Transform for short period multiple attenuation: CSEG 23rd Annual Mtg., Expanded Abstracts, 58-59.

Kostov, C., 1990, Toeplitz structure in slant-stack inversion: 60th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1618-1621.

Marple, S., L., 1987, Digital spectral analysis with applications: Prentice-Hall, Inc.

Sacchi, M.. D. and Ulrych, T., J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 4, 1169-1177.

Schonewille, M., A., and Duijndam, A., J., W., 1998, Efficient nonuniform Fourier and Radon filtering and reconstruction: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1692-1695.

Strang, G., 1986, Introduction to applied mathematics: Wellesley-Cambridge Press.

$N \times M$	**Lev**	**Chol**	**CG+FFT**
128×128	2	6	3
256×256	8	42	12

Table 1: CPU times in seconds for the 3 algorithms tested in this study. N denotes the number of traces and M the number of q parameters.

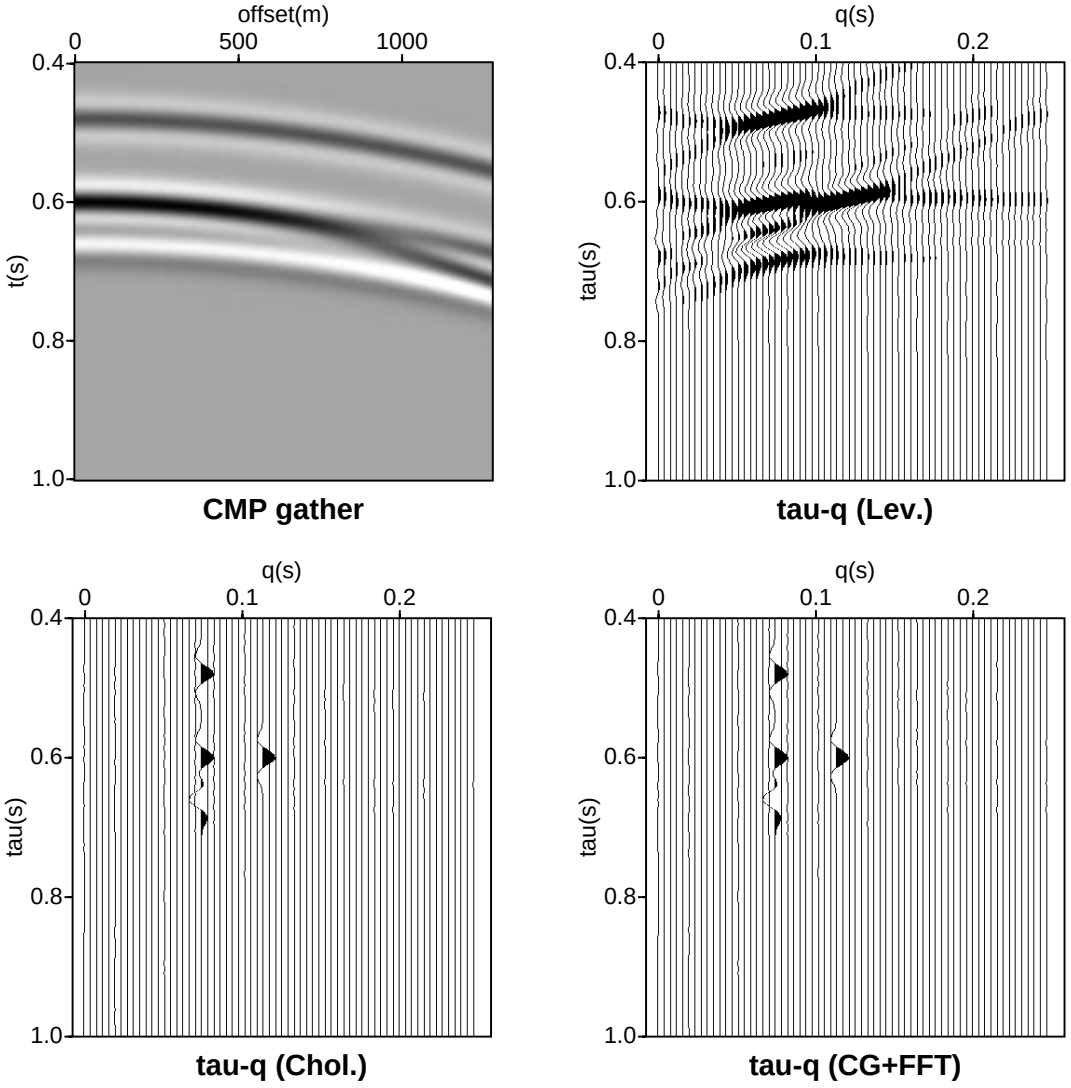

Fig. 1: A synthetic CMP gather composed of 4 parabolic events is used to test 3 different algorithms to compute the Radon transform. **Lev.** indicates the classical solution using least squares with a constant damping term (equation(6)); the Levinson algorithm is used to invert the resulting Toeplitz form. **Chol.** indicates the high resolution solution using non-constant damping (equation (8)), this solution is computed by means of the Cholesky decomposition. **CG+FFT** indicates the proposed fast algorithm to compute the high resolution Radon transform. In this example the size of the Radon operator is 256×256. CPU times in seconds are given in Table 1.