P026

Least-Squares Signal Estimation with Complicated Mathematical Models of Seismic Data

Y. Tyapkin* (Ukrainian State Geological Prospecting Institute), B. Ursin (Norwegian University of Science and Technology), Y. Roganov (Ukrainian State Geological Prospecting Institute), I. Nekrasov (Pridneprovskaya Geophysical Prospecting Expedition) & G. Shumlyanskaya (Ukrainian State Geological Prospecting Institute)

SUMMARY

We present a least-squares method to estimate signals with a more realistic compared to the conventional one, mathematical model of seismic record comprising random noise and an arbitrary number of coherent noise wavetrains. Both kinds of noise are assumed to be stationary stochastic processes independent of each other and the signal. The signal and coherent noise wavetrains bear individual trace-independent waveforms, while their amplitudes and arrival times vary in space in any manner. Being uncorrelated in space, the random noise is supposed to have identical to within a scale factor, the variance, spectral functions on different traces. Under certain conditions, the method may be reduced to successively subtracting all the coherent noise wavetrains followed by estimating the signal on the residual record. Both procedures use optimum weighted stacking, with reference to the related arrival times. If the power of random noise is rather stable in space, the method permits further simplification by substituting singular value decomposition for optimum weighted stacking. The method is demonstrated with synthetic and field data sets contaminated by severe coherent noise. Emphasis is put on the comparison of the method with f-k filtering combined with straight stacking.

INTRODUCTION

When searching for and prospecting of hydrocarbon traps in sedimentary basins, geophysicists are often faced with the problem of optimally recovering signal from seismic data contaminated by spatially coherent and random noise. Among both kinds of noise, it is coherent noise that is the most persistent problem in seismic imaging, and a number of techniques have been developed to attenuate it: radial trace filtering (Zhu *et al.* 2004), *f-k* filtering (March and Bailey 1983), *τ-p* transform (Kelamis and Mitchell 1989; Sacchi and Ulrych 1995), spectral matrix filtering (Gounon *et al.* 1998). Various modifications of the singular value decomposition (SVD), whish is also known in the literature either as the Karhunen-Loeve or as the principal component transformation, operating in the time or frequency domain (Ulrych *et al.* 1999; Vrabie *et al.* 2004, Tyapkin *et al.* 2004, 2006), are also popular tools for removing coherent and random noise. In the 1960s-1970s, in order to solve the problem of optimally estimating the seismic signal imbedded in coherent and random noise, the majority of publications exploited multichannel Wiener filters (e.g., Cassano and Rocca 1973). Then the situation changed and the interest of geophysicists in these methods gradually was lost because of their insufficient effectiveness. Their place was taken by non-optimum methods implying simplified mathematical models of seismic data, with *f-k* filtering and Radon (*τ-p*) transform filtering being best known among them. Even though these filters are usually faster and more cost-effective, the most undesirable aspect of them is the mixing of the data. In our opinion, the optimum methods are often less effective because, first, they exploit imperfect mathematical models of the record and, second, they are not supplied with reliable estimates of the parameters required to operate. For this reason, in this paper, we make an attempt to rehabilitate these methods and to reanimate the interest of geophysicists in them. With this purpose in mind, we utilize a more complicated and adequate mathematical model of the record.

SOLUTION OF THE PROBLEM AND ITS ANALYSIS

Suppose that the *i*th trace of a record that consists of *N* traces may be written as:

$$u_i(t) = a_i s\left(t - \tau_{(s)i}\right) + \sum_{l=1}^{L} b_{il} r_l\left(t - \tau_{(r)il}\right) + n_i(t), \qquad i = 1, ..., N \ . \tag{1}$$

Here the signal component is described by the first term on the right-hand side of equation (1). It is assumed to have an identical waveform *s*(*t*) on each trace, with arbitrary trace-dependent amplitudes a_i and time delays $\tau_{(s)i}$. The second term represents a superposition of coherent noise wavetrains with individual waveforms $r_l(t),\ l = 1, ..., L$. Each of the wavetrains, as well as the signal, bears arbitrary trace-dependent amplitudes b_{il} and time delays $\tau_{(r)il}$. Random noise is expressed by the third term. It is supposed to be a stationary zero-mean Gaussian stochastic process independent of both the signal and the random noise on any other trace. It has identical to within a scale factor, the variance σ_i^2, autocorrelations on different traces. In turn, all the coherent noise waveforms are also assumed to be stationary zero-mean Gaussian stochastic processes independent of each other, the signal and random noise. Due to these assumptions, the cross-spectrum of the entire noise between channels *i* and *j* may be expressed as:

$$R_{ij}(\omega) = \sum_{l=1}^{L} b_{il} b_{jl} R_{(r)l}(\omega) \exp\left[i\omega\left(\tau_{(r)jl} - \tau_{(r)il}\right)\right] + \sigma_i^2 R_{(n)}(\omega)\delta_{ij} \ , \tag{2}$$

where $R_{(r)l}(\omega)$ and $R_{(n)}(\omega)$ are the power spectra of $r_l(t)$ and any $n_i(t)$, respectively, at an angular frequency ω, while δ_{ij} signifies the Kronecker delta.

Given this model, let us state the problem to obtain the best in least-squares sense estimate of the signal shape, *s*(*t*). With this purpose in mind, we represent the signal component as **f***s*, where $\mathbf{f} = \left\{a_1 \exp\left(i\omega\tau_{(s)1}\right), ..., a_N \exp\left(i\omega\tau_{(s)N}\right)\right\}^H$, the scalar *s* is the Fourier spectrum of *s*(*t*), and the superscript *H* denotes complex conjugate (Hermitian) transpose.

Then the least-squares estimate of *s* is (Helstrom 1968)

$$s = \left(\mathbf{f}^H \mathbf{R}^{-1} \mathbf{f}\right)^{-1} \mathbf{f}^H \mathbf{R}^{-1} \mathbf{u} \ , \tag{3}$$

where $\mathbf{u} = \left\{U_1^*, ..., U_N^*\right\}^H$ is the column vector whose elements U_i, $i = 1,…,N$, are the Fourier spectra of the traces, **R** is the matrix whose elements are defined by equation (2), -1 signifies matrix inverse, and the superscripted asterisk stands for complex conjugation. For shot, here and in the following the functional dependence on frequency is dropped.

After representing **R** in the form

$$\mathbf{R} = \mathbf{G}\mathbf{B}\mathbf{G}^H + R_{(n)}\mathbf{D}, \qquad (4)$$

where $\mathbf{G} = \{\mathbf{g}_j\} = \{b_{ij} \exp(i\omega\tau_{(r)ij})\}$, $i = 1, ..., N$, $j = 1, ..., L$, is a $N \times L$ matrix that consists of the column vectors $\mathbf{g}_j = \{b_{1j} \exp(i\omega\tau_{(r)1j}), ..., b_{Nj} \exp(i\omega\tau_{(r)Nj})\}^H$, whereas $\mathbf{B} = \mathrm{diag}\{R_{(r)1}, ..., R_{(r)L}\}$ and $\mathbf{D} = \mathrm{diag}\{\sigma_1^2, ..., \sigma_N^2\}$ are $L \times L$ and $N \times N$ diagonal matrices, respectively, we can take advantage of the method described by Horn and Johnson (1986) and obtain

$$\mathbf{R}^{-1} = R_{(n)}^{-1}\left(\mathbf{I} - R_{(n)}^{-1}\mathbf{D}^{-1}\mathbf{G}\mathbf{V}^{-1}\mathbf{G}^H\right)\mathbf{D}^{-1}, \qquad (5)$$

where **I** is an identity matrix and $\mathbf{V} = \mathbf{B}^{-1} + R_{(n)}^{-1}\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G}$. If the conditions

$$R_{(r)l}R_{(n)}^{-1}c_{ll} >> 1, \qquad c_{ii}c_{jj} >> \left|c_{ij}\right|^2 \quad \text{for } i \neq j, \qquad \text{and} \qquad c_s >> \sum_{l=1}^{L} c_{ll}^{-1}\left|c_{sl}\right|^2, \qquad (6)$$

where

$$c_{ij} = \mathbf{g}_i^H\mathbf{D}^{-1}\mathbf{g}_j = \tilde{\mathbf{g}}_i^H\tilde{\mathbf{g}}_j = \sum_{k=1}^{N}\frac{b_{ki}b_{kj}}{\sigma_k^2}\exp\left[i\omega\left(\tau_{(r)ki} - \tau_{(r)kj}\right)\right], \qquad \tilde{\mathbf{g}}_i = \mathbf{D}^{-1/2}\mathbf{g}_i, \qquad i, j = 1, ..., L,$$

$$c_{si} = \mathbf{g}_i^H\mathbf{D}^{-1}\mathbf{f} = \tilde{\mathbf{g}}_i^H\tilde{\mathbf{f}} = \sum_{k=1}^{N}\frac{a_k b_{ki}}{\sigma_k^2}\exp\left[i\omega\left(\tau_{(r)ki} - \tau_{(s)k}\right)\right], \qquad c_s = \mathbf{f}^H\mathbf{D}^{-1}\mathbf{f} = \tilde{\mathbf{f}}^H\tilde{\mathbf{f}} = \sum_{k=1}^{N}\frac{a_k^2}{\sigma_k^2}, \qquad \tilde{\mathbf{f}} = \mathbf{D}^{-1/2}\mathbf{f},$$

are valid, **V** reduces to

$$\mathbf{V} = R_{(n)}^{-1}\mathrm{diag}\{c_{ii}\}, \quad i = 1, ..., L. \qquad (7)$$

Inequalities (6) imply that, first, any coherent noise prevails significantly over random noise at the output from related optimum weighted stacking (Tyapkin and Ursin 2005) and, second, all the vectors $\tilde{\mathbf{g}}_i$ and the vector $\tilde{\mathbf{f}}$ are almost mutually orthogonal.

Combining equation (7) with (3) and (5) yields

$$s = c_s^{-1}\mathbf{f}^H\left(\mathbf{I} - \sum_{i=1}^{L}\frac{\mathbf{D}^{-1}\mathbf{g}_i\mathbf{g}_i^H}{c_{ii}}\right)\mathbf{D}^{-1}\mathbf{u} = \mathbf{h}^H\mathbf{u}, \qquad (8)$$

where **h** is the spectral characteristic of the least-squares signal estimator. Note that unlike the strict solution from equations (3) and (5), this filter is independent of the spectra of all the noise components.

Equation (8) allows us to suitably interpret the sequence of operations needed to implement this kind of filtering. First, for each of the coherent noise wavetrains, two operations are carried our, namely:

- optimum weighted stacking (Tyapkin and Ursin 2005) of all the traces in order to estimate the waveform of the coherent noise; prior to this operation, the related arrival times should be cancelled out to align the coherent noise and to remove its time delays,
- subtracting the coherent noise estimate from all the traces with regard for its arrival time, amplitude, and waveform.

Once the entire coherent noise has been removed, optimum weighted stacking of the residual record is performed to estimate the signal waveform. As in the previous case, the related time shifts are cancelled out prior to stacking.

The effectiveness of the method developed is highly dependent on the accuracy of the necessary signal and noise parameter determination. Therefore, for the method to be feasible, we propose algorithms described in (Tyapkin and Ursin 2005), which are based on the same record model. When random noise is rather stable in space, the method permits further simplification by the substitution of SVD for optimum weighted stacking (Tyapkin and Ursin 2005). In such a case, with the record treated as a rectangular matrix formed with regard for the corresponding arrival times, any coherent noise can be estimated as the dominant term of the SVD. After subtracting the entire coherent noise, the signal waveform is estimated as the left singular vector associated with the dominant singular value of the residual record. This technique was used to process synthetic data demonstrated below.

SYNTHETIC DATA EXAMPLES

The proposed method is demonstrated with a set of synthetic shot gathers contaminated by a linear coherent noise wavetrain of various relative, in regard to the signal component, amplitudes (0.5, 1, 2, 4, 8 and 16) and dips (1, 2, 4, 8 and 16 sample intervals per trace), with the sample interval being 2 ms. Emphasis is put on the comparison of its performance with that of the conventional combination of *f-k* filtering and straight stacking. The waveforms of both the signal and coherent noise were generated by convolution of independent stochastic processes with a 20 Hz Ricker wavelet. The signal bears root-mean-square (rms) amplitudes decreasing linearly between 1500 and 500 with offset. White noise with an rms amplitude 100 is always superimposed on each trace to simulate additive random noise.

Table 1 presents the correlation coefficients between the 'pure' signal and its estimates. It is clearly seen that the SVD-based method always yields superior results. This advantage over *f-k* filtering combined with straight stacking is most pronounced when the relative dip of the coherent noise is either too small (1 sample interval per trace) or too large (16 sample intervals per trace), with the relative amplitude being large enough (16). In such critical situations, *f-k* filtering is incapable of separating the signal and coherent noise or suffers from spatial aliasing of the noise, respectively. The related correlation coefficients are seen in Table 1 on a yellow background. The first case is shown in Figure 1. Here it is obvious that the coherent noise is so strong that before the noise is removed, there is almost no signal that can be identified on the panel (c). The SVD-based method removes the coherent noise almost perfectly and then yields the signal estimate that can hardly be distinguished visually from the original. As opposed to this method, *f-k* filtering produces large artifacts not allowing the signal to be identified. For this reason, the subsequent straight stack differs considerably from the signal waveform. For a better visual analysis, the signal and both estimates obtained are represented in Figure 2 on a large scale.

The method developed also was tested successfully with field data sets (Tyapkin *et al*. 2006). Since in that case the data basically were corrupted by ground roll, which had a divergent, fan-like character, a special preliminary transformation (Tyapkin *et al.* 2004) was utilized to favour the subsequent application of SVD.

CONCLUSIONS

The method developed can be a valuable tool in processing seismic data. The results of its testing on various synthetic and field data sets, some of which are demonstrated, indicate that in many circumstances when estimating signal contaminated by severe coherent noise our method can significantly outperform *f-k* filtering combined with straight stacking and may therefore be prescribed as a better choice than the conventional process.

REFERENCES

Cassano, E. and Rocca, F. [1973] Multichannel linear filters for optimal rejection of multiple reflections. *Geophysics* **38**, 1053-1061.

Gounon, P.Y., Mars, J.I. and Goncalves, D. [1998] Wideband spectral matrix filtering. 68th Meeting, SEG, Expanded Abstracts, 1150-1153.

Helstrom, C.W. [1968] *Statistical Theory of Signal Detection*. Pergamon Press.

Horn, R.A. and Jonhson, C.R. [1986] *Matrix Analysis*. Cambridge University Press.

Kelamis, P.G. and Mitchell, A.R. [1989] Slant-stack processing. *First Break* **7 (**2), 43-54.

March, D.W. and Bailey, A.D. [1983] A review of the two-dimensional transform and its use in seismic processing. *First Break* **1** (1), 9-21.

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

Tyapkin, Yu., Marmalyevskyy, N. and Gornyak, Z. [2004] Suppression of source-generated noise using the singular value decomposition. 66th Conference, EAGE, Extended Abstracts, D038.

Tyapkin, Yu. and Ursin, B. [2005] Optimum stacking of seismic record with irregular noise. *Journal of Geophysics and Engineering* **2**, 177-187.

Tyapkin, Yu., Ursin, B., Roganov, Yu. and Nekrasov, I. [2006] Optimum seismic signal estimation with complicated models of coherent and random noise. 76th Meeting, SEG, Expanded Abstracts, 2867-2871.

Ulrych, T.J., Sacchi, M.D. and Graul, J.M. [1999] Signal and noise separation: Art and science. *Geophysics* **64,** 1648-1672.

Vrabie, V.D., Mars, J.I. and Lacoume, J.L. [2004] Modified singular value decomposition by means of independent component analysis. *Signal Processing* **84**, 645-652.

Zhu, W., Kelamis, P.G. and Liu, Q. [2004] Acquisition/Processing - Linear noise attenuation using local radial trace median filtering. *The Leading Edge* **23,** 728-737.

Table 1. Correlation coefficients of the 'pure' signal with its estimates by SVD-based method (in numerators) and by *f-k* filtering combined with straight stacking (in denominators).

Relative amplitudes of coherent noise	Relative dips of coherent noise (in sample intervals per trace)				
	1	2	4	8	16
0.5	0.981 / 0.960	0.994 / 0.972	0.999 / 0.987	1 / 0.990	0.999 / 0.993
1	0.995 / 0.964	0.998 / 0.972	0.999 / 0.990	1 / 0.993	0.999 / 0.996
2	0.996 / 0.950	0.998 / 0.970	0.999 / 0.990	1 / 0.993	0.999 / 0.990
4	0.996 / 0.903	0.998 / 0.962	0.999 / 0.989	1 / 0.995	0.999 / 0.965
8	0.996 / 0.790	0.998 / 0.927	0.999 / 0.980	1 / 0.995	0.999 / 0.880
16	0.996 / 0.615	0.998 / 0.830	0.999 / 0.941	1 / 0.987	0.999 / 0.678

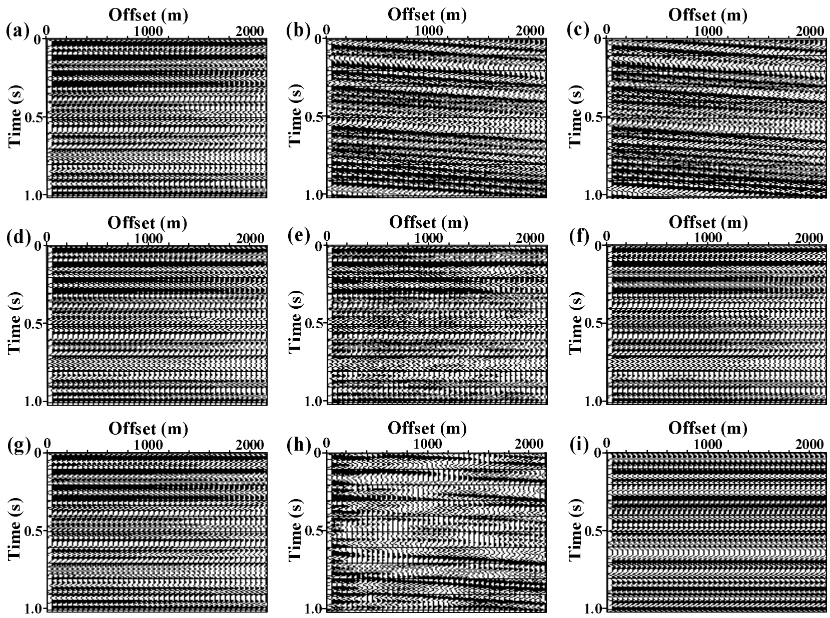

Figure 1. Synthetic data example. (a), (d), (g) 'Pure' signal. (b) Coherent noise with a relative amplitude 16 and dip 1 sample interval per trace. (c) Shot gather containing signal, coherent noise and random noise. Result obtained after (e) subtracting coherent noise by SVD, (f) approximating signal by SVD, (h) subtracting coherent noise by *f-k* filtering, and (i) straight stacking of (h) multiplied for a better visual analysis.

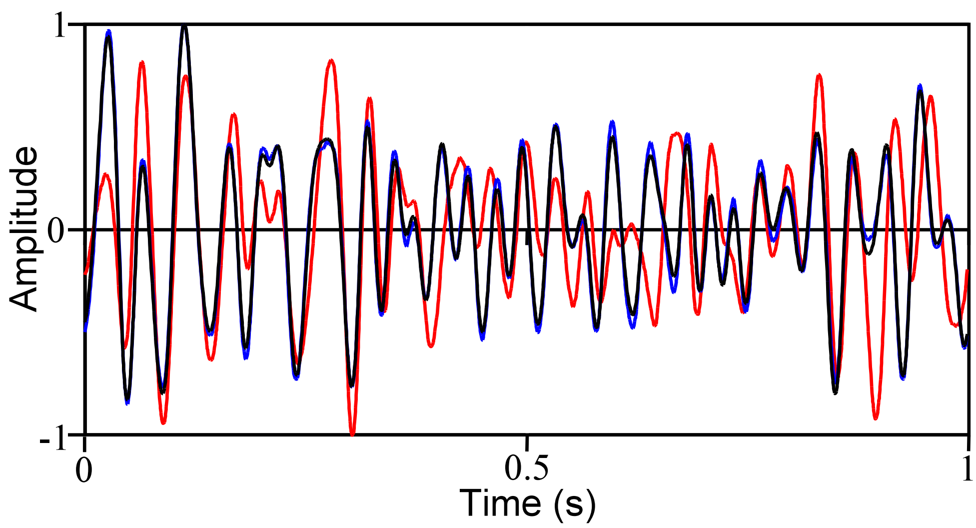

Figure 2. Signal (black) and its estimates by the SVD-based method (blue) and by *f-k* filtering combined with straight stacking (red) from Figure 1.