Decomposition by Markov Processes

Jon F. Claerbout

Given the sound of several speakers added together into a single "seismogram", or given a VSP containing events of several dips, or a profile with several velocities, or a section with both geologic layering and fault edge diffractions, how can we decompose such data into its components?

Let d_j represent the j^{th} component of such a decomposition. The sum of the components of the data should give the data itself. Let $d_j^{(n)}$ represent our n^{th} estimate of the j^{th} component. For each estimate n the sum of the components should give the data. Markov matrices are matrices whose columns sum to unity — to preserve the unity sum of probability. Actually, multiplication by a Markov matrix preserves the sum of the components of any vector. So we can use Markov matrices to take us from any decomposition $d^{(n)}$ to the next $d^{(n+1)}$. Initially we could simply divide the data equally among the components, or we could put all the data in the first component.

First Markov matrix

Consider the following Markov transition:

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}^{(n+1)} = \begin{bmatrix} 1-\alpha_{11}B_1 & \alpha_{12}B_2 & \alpha_{13}B_3 \\ \alpha_{21}B_1 & 1-\alpha_{22}B_2 & \alpha_{23}B_3 \\ \alpha_{31}B_1 & \alpha_{32}B_2 & 1-\alpha_{33}B_3 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}^{(n)} \tag{1}$$

The α_{ij} contain the cross coupling. To preserve the data being the sum of its components, we must have the column sums being unity, namely, $\sum_i \alpha_{ij} = 0$. At convergence each component $d_i^{(n+1)}$ equals $d_i^{(n)}$ so

$$\alpha_{jj}\ B_j\ d_j \quad = \sum_{j\,,\ j \neq i} \alpha_{ij}\ B_j\ d_j \tag{2}$$

Prediction error

Consider a two component prediction error filtering problem with $\alpha_{ij} = 1/2$ and two given prediction-error filters B_1 and B_2. Equation (2) says that at convergence, the prediction error coming out of the first filter B_1 matches the prediction error coming out of the second filter B_2. This is the problem that led me to thinking about decomposition; my two prediction operators are each velocity dependent filter pairs.

Markov matrices ordinarily have real positive elements, but notice that in the Fourier domain, we are extending their use to complex elements. I don't think this creates any problems.

Vertical seismic profile (VSP)

Next consider the example of a VSP. Here the goal is to divide the VSP into components with various dips. It is helpful to think of the operator B_j as a "badpass" operator. It passes with unit gain, all that does not fit some model. That is,

$$B_j \quad = \quad 1 - G_j \tag{3}$$

where G_j tries to predict the j^{th} component of the data. In multidimensional problems, G_j is usually some kind of a spatial low-pass filter. For VSP's the filter G_j is a shift to flatten the j^{th} event followed by a spatial lowpass filter which might be implemented by a tridiagonal scheme, followed by the reversed shift.

Velocity decomposition of a CDP gather

Consider

$$G_j \quad = \quad NMO^{-1}\, SpatialLowPass \;\; NMO \tag{4}$$

Various j-values refer to various NMO velocities and/or various filter cutoffs. I think the tridiagonal *NMO* inversion scheme I worked out a few weeks ago should work nicely. Of course *NMO* does not have an exact inverse. A pseudoinverse should do, as long as the the eigenvalues of B_j are less than unit magnitude. I expect instability should any of the eigenvalues of G be negative. For *NMO* by linear interpolation, I suspect the transpose NMO^T would work too if its column sums were normalized to unity. This is worth trying. It should be fast. The decomposition should be useful. Who knows what difficulties lurk in the side boundaries ...

Spectral decomposition

Let k_j, $k_j < k_{j+1}$ be a set of cutoff frequencies in the spatial spectral domain. A family of lowpass filters is

$$L_j \quad = \quad \frac{1}{1 \ + \ T/k_j^2} \tag{5}$$

where T is the familiar tridiagonal matrix with the negative of the second difference operator on its diagonal. Positive bandpass filters, less than unity in magnitude, are given by $G_j = L_{j+1} - L_j$.

Mixed voices

If you thought I was going to solve the mixed voices problem, you are a dreamer. For that problem I suspect that the α_{ij} are not given apriori but must be deduced via envelope functions of the data and the data components. The mathematics of problems of this kind should be buried in the literature on adaptive antenna theory.

Convergence procedure

The convergence of the iteration is something that needs to be studied. In the deconvolution problem, I would iterate for a while, then stop and re-estimate the deconvolution filters from the improved data components. That is, initially my first two filter pairs would depend on velocity and the original data d. After a decomposition of $d = d_1 + d_2$, one new pair would be based on v_1 and d_1, while the other pair would be based on v_2 and d_2.

Missing data

Since the output of each B_j is a prediction error, naturally it would be set to zero where the predicted data was not recorded.

Another Markov matrix

In the interests of computational efficiency, particularly while starting out the iteration, we might throw all the energy out from one channel into the next channel processed.

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}^{(n+1)} = \begin{bmatrix} 1-\alpha_1 B_1 & 0 & \alpha_3 B_3 \\ \alpha_1 B_1 & 1-\alpha_2 B_2 & 0 \\ 0 & \alpha_2 B_2 & 1-\alpha_3 B_3 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}^{(n)} \tag{6}$$

Later, as convergence approaches, the more symmetrical arrangement (1) might be preferable.

Formulation as an Optimization Problem

Finally I should point out that data can be decomposed by optimization with the constraint $d = d_1 + d_2 + \cdots$. It is possible that the Markov sequence is a numerical solution to some such optimization problem. Maybe it is this:

$$\min = \sum R_1 d_1^2 + R_2 d_2^2 + R_3 d_3^2 + \cdots \tag{7}$$

where $R = B^T B$.