PREDICTIVE DECONVOLUTION FROM THE POINT OF VIEW OF KRIGING

MICHAEL K. BROADHEAD

Saudi Aramco, P.O. Box 1372, Dhahran 31311, Saudi Arabia.

(Received November 8, 2008; accepted March 9, 2009)

ABSTRACT

Broadhead, M.K., 2009. Predictive deconvolution from the point of view of kriging. *Journal of Seismic Exploration*, 18: 239-247.

We review the geostatistical method of simple kriging and consider its application to the predictive deconvolution problem in seismology. We find that when kriging is applied to the 1D time-series prediction problem, it can be used to obtain the usual Wiener-Levinson system of equations that are normally arrived at with the methods of stationary time-series analysis. The kriging weights can then be interpreted in terms of the prediction filter coefficients. This connection between predictive decon and geostatistics does not appear to be known. Perhaps a better understanding and exploitation of this type of connection can lead to productive synergies between signal processing and geostatistics.

KEYWORDS: deconvolution, geostatistics, kriging, prediction filter, Wiener-Levinson, covariance.

INTRODUCTION

Geostatistics is well known for its usefulness in the oil industry for wireline borehole and seismic data integration in order to populate reservoir simulation models with porosity and permeability. Less well known is that geostatistical methods can be used for signal processing. The application of kriging to image processing, for example, is discussed by Ruiz-Alzola et al. (2005). In the reflection seismic world, geostatistical methods have been hybridized successfully with seismic inversion in several schemes (Dubrule, 2003). A recent advance in this area is due to Hansen et al. (2006), who extend

0963-0651/09/$5.00

the geostatistical methodology of sequential simulation to linear and weakly nonlinear inverse problems. In addition, some investigators have applied geostatistical methods to seismic data filtering (Piazza et al., 1997).

In this paper, we explore the kriging interpretation of one particular inverse problem that arises in geophysical signal analysis, namely, the linear prediction of time series as applied to the deconvolution problem. Predictive deconvolution (also called prediction error filtering) finds widespread application in exploration geophysics for multiple attenuation and resolution enhancement. It is based on the theory of stationary time series analysis developed by both Kolmogrov and Wiener, and further extended by Levinson, Robinson, Treitel and others (Robinson and Treitel, 2000). In predictive decon, a prediction filter is designed from the autocorrelation function using the Wiener-Levinson linear system of equations, and then a prediction error filter is formed and applied to the data.

We apply the method of simple kriging to the linear prediction filter design problem for a 1D time series. We find that we can obtain the Wiener-Levinson linear system of equations where the kriging weights can be interpreted in terms of the prediction filter coefficients. This result is interesting in that it reveals an unexpected connection between geostatistics and signal processing. This realization can be immediately exploited by anyone understanding either of these subjects in order to gain insight into (and perhaps, new ideas about) the other subject.

SIMPLE KRIGING THEORY

Kriging (named in honor of South African mining engineer Daniel Gerhardus Krige) refers to a set of statistical tools that were developed largely by Matheron (1963), which in turn was based on earlier work by Krige (1951). This subject has been thoroughly covered elsewhere (Goovaerts, 1997) and hence we present only a brief review for completeness and for establishing notation. Kriging refers to a collection of techniques for spatial interpolation of a set of data values $Z(\mathbf{u}_i)$, where a least-squares regression approach is central to all of the particular methods ($\mathbf{u}_i$ is the position vector for the i-th data point). Also common to all kriging methods is a random function model for the data, where $Z(\mathbf{u})$ is decomposed into a residual component, $R(\mathbf{u})$, and a trend component, $m(\mathbf{u})$, where

$$Z(\mathbf{u}) = R(\mathbf{u}) + m(\mathbf{u}) \quad . \tag{1}$$

$R(\mathbf{u})$ is treated as a stationary random variable with zero mean. The expected value of $Z(\mathbf{u})$ is given by the trend, i.e.,

$$E\{Z(\mathbf{u})\} = m(\mathbf{u}) \quad . \tag{2}$$

We concentrate on a particular form of kriging called "simple kriging", where the surface trend $m(\mathbf{u})$ is taken to be a constant, m, which is assumed to be known. Then, the linear estimator of some value Z, (e.g., porosity, depth, mineral concentration) to be interpolated at spatial location $\mathbf{u}$ is given by

$$Z^*_{SK}(\mathbf{u}) = m + \sum_{i=1}^{n(\mathbf{u})} \lambda_i^{SK}(\mathbf{u}) R(\mathbf{u}_i) \quad , \tag{3}$$

which constitutes an unbiased estimator. Note, the subscript or superscript SK refers to simple kriging and $n(\mathbf{u})$ is the number of data points in some neighborhood about the point of investigation. The kriging weights, λ_i^{SK}, are found by minimizing the variance expression given by

$$\sigma_E^2(\mathbf{u}) = \mathrm{Var}[Z^*_{SK}(\mathbf{u}) - Z(\mathbf{u})] \quad , \tag{4}$$

which leads to the linear system to be solved, namely

$$\sum_{i=1}^{n(\mathbf{u})} \lambda_j^{SK}(\mathbf{u}) C(\mathbf{u}_i - \mathbf{u}_j) = C(\mathbf{u}_i - \mathbf{u}) \quad , \tag{5}$$

where $i = 1, \ldots, n$. The quantity C represents the covariance function, which, under the assumption of stationarity, is only a function of coordinate separation (or lag), $\mathbf{h}$. The stationary expression for covariance is given by

$$C(\mathbf{h}) = E\{Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})\} - E\{Z(\mathbf{u})\}E\{Z(\mathbf{u} + \mathbf{h})\} \quad . \tag{6}$$

Note that $\mathbf{h}$ is a vector with magnitude h. The case where C depends only on h is termed "isotropic". In geostatistics, C is often indirectly estimated using the semivariogram γ, which is given by

$$2\gamma(\mathbf{h}) = E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2\} \quad . \tag{7}$$

The covariance function can then be recovered from the semivariogram using

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \quad . \tag{8}$$

However, this detour through the variogram, while traditional, is unnecessary. One can directly compute the covariance function. Also, in geostatistics the empirical covariance (or variogram) is not typically used directly in the kriging equations, but rather, an analytical representation is fitted to the empirical result

and used instead. This step would be, of course, omitted in our application to predictive deconvolution, where the use of the empirical covariance is made.

Now, let us adapt the covariance function to the time series case. If **u** is replaced with time, t, and **h** is replaced with time lag τ, and $E\{Z(\mathbf{u})\} = 0$ (zero mean) and Z is stationary, then $C(\mathbf{h})$ becomes

$$C(\tau) = r(\tau) = E\{Z(t)Z(t + \tau)\} \quad , \tag{9}$$

which is well known as the autocorrelation function $r(\tau)$ (Robinson, 1976). Then, using eq. (8), the connection between the semi-variogram and the autocorrelation function is given by

$$\gamma(\tau) = r(0) - r(\tau). \tag{10}$$

SIMPLE KRIGING APPLIED TO 1D TIME SERIES PREDICTION

Let's make the notation above more explicit by writing out the matrix form of the linear system, where we use a simple 3×3 case without loss of generality. We will only explicitly treat the unit time step prediction problem. The generalization of our treatment below to an arbitrary prediction distance is straightforward. Then, for the isotropic case Equation (5) becomes

$$\begin{bmatrix} C(0) & C(h_{12}) & C(h_{13}) \\ C(h_{21}) & C(0) & C(h_{23}) \\ C(h_{31}) & C(h_{32}) & C(0) \end{bmatrix} \begin{bmatrix} \lambda_1(\mathbf{u}) \\ \lambda_2(\mathbf{u}) \\ \lambda_3(\mathbf{u}) \end{bmatrix} = \begin{bmatrix} C(h_1) \\ C(h_2) \\ C(h_3) \end{bmatrix} \quad , \tag{11}$$

where $h_{ij} = |\mathbf{u}_i - \mathbf{u}_j|$ and $h_k = |\mathbf{u}_k - \mathbf{u}|$. Now, let us assume that our data to be interpolated or extrapolated (seismic reflection amplitudes) are in the form of an N point time series on a uniform grid, i.e.,

$$\mathbf{Z} = \{Z(t_0), Z(t_1), \ldots, Z(t_{N-1})\} \quad , \tag{12}$$

where $t_i = t_0 + i\Delta t$ and Δt is the sample rate. Then, we wish to cast the following unit time step prediction problem in terms of our geostatistical notation,

$$Z^*(t_3) = \lambda_0(t_3)Z(t_0) + \lambda_1(t_3)Z(t_1) + \lambda_2(t_3)Z(t_2) \quad . \tag{13}$$

The kriging weights are then given by

$$\begin{bmatrix} C_0 & C_1 & C_2 \\ C_1 & C_0 & C_1 \\ C_2 & C_1 & C_0 \end{bmatrix} \begin{bmatrix} \lambda_0(t_3) \\ \lambda_1(t_3) \\ \lambda_2(t_3) \end{bmatrix} = \begin{bmatrix} C_3 \\ C_2 \\ C_1 \end{bmatrix} , \tag{14}$$

where $C_i = C(i\Delta t)$ for $i = 0, \ldots, 3$. Note that we already have a data value measured at t_3, so our purpose in predicting it is different than is usual for geostatistical application. In the seismic case, we are actually interested in the prediction errors,

$$PE = Z(t) - Z^*(t) , \tag{15}$$

which is taken as an estimate of reflectivity. Note that the time we are predicting is still tied to all of the previous data points. Hence, when we advance in time, more data become available for use (since we are restricting ourselves to pure prediction). Hence, our kriging system could grow in size on the next time step for prediction, or

$$\begin{bmatrix} C_0 & C_1 & C_2 & C_3 \\ C_1 & C_0 & C_1 & C_2 \\ C_2 & C_1 & C_0 & C_1 \\ C_3 & C_2 & C_1 & C_0 \end{bmatrix} \begin{bmatrix} \lambda_0(t_4) \\ \lambda_1(t_4) \\ \lambda_2(t_4) \\ \lambda_3(t_4) \end{bmatrix} = \begin{bmatrix} C_4 \\ C_3 \\ C_2 \\ C_1 \end{bmatrix} , \tag{16}$$

where

$$Z^*(t_4) = \lambda_0(t_4)Z(t_0) + \lambda_1(t_4)Z(t_1) + \lambda_2(t_4)Z(t_2) + \lambda_3(t_4)Z(t_3) . \tag{17}$$

However, because we have many data points located on an uniformly sampled grid and are assuming stationarity, we can choose to decouple the kriging weights from time by leaving the matrix size fixed. This is done to be in accordance with how linear prediction is usually done. Then, the interpolator becomes

$$Z^*(t_n) = \lambda_0 Z(t_{n-3}) + \lambda_1 Z(t_{n-2}) + \lambda_2 Z(t_{n-1}) , \tag{18}$$

for fixed weights (see Figs. 1 and 2). Note that these choices are not prescribed or required by geostatistical theory, but are reasonable, and are done in order to assure we arrive at the Wiener-Levinson equations. This is why we can only claim that kriging can be used to arrive at that result, not that there is a strong equivalence between predictive decon and kriging. Now, recognizing that the covariance function on a uniform time grid is essentially the autocorrelation function for a zero mean random process, we can rewrite eq. (14) as

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} r_3 \\ r_2 \\ r_1 \end{bmatrix} . \tag{19}$$

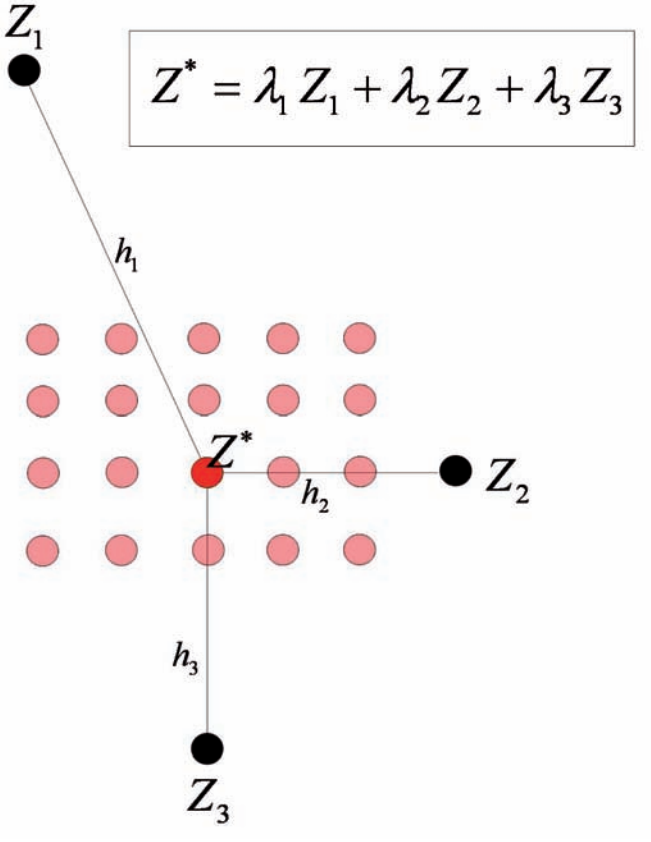

Fig. 1. Typical spatial data interpolation template. Black circles are measured data locations. Pink circles represent the desired grid. The red circle is the point under present consideration.

Using elementary row and column operations, we can rewrite this as

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} \lambda_2 \\ \lambda_1 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} . \tag{20}$$

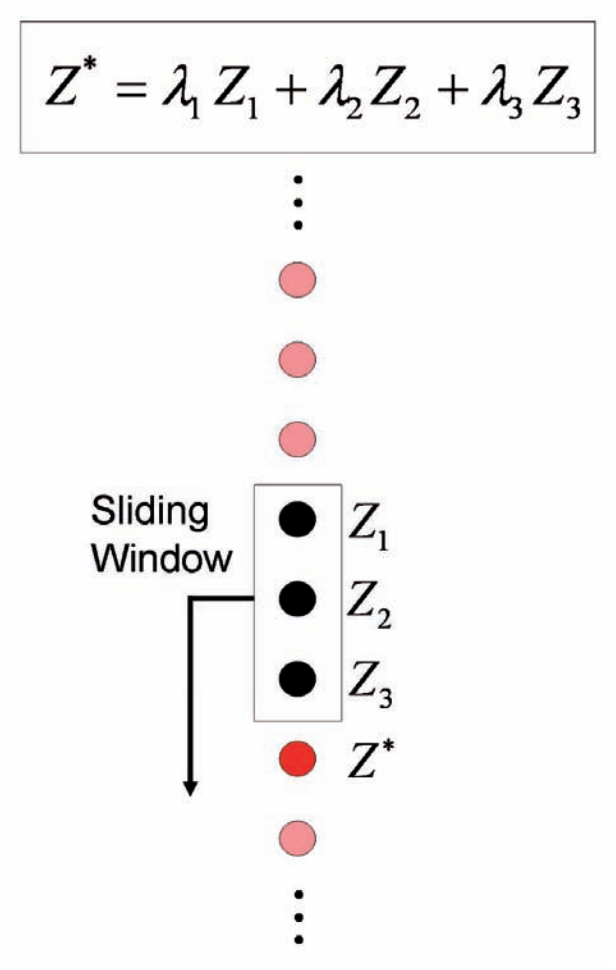

Fig. 2. Kriging template adapted to the prediction filter problem.

Letting $f_0 = \lambda_3$, $f_1 = \lambda_2$, and $f_2 = \lambda_1$, we then have

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} . \tag{21}$$

This last result is known in time-series analysis as the Wiener-Levinson Equations (Robinson and Treitel, 1980) for linear prediction in the special case of predicting one time sample into the future, where the length M filter **f** is called a prediction filter. In general,

$$\begin{bmatrix} r_0 & r_1 & \cdots & r_{M-1} \\ r_1 & r_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1 \\ r_{M-1} & \cdots & r_1 & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{M-1} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_M \end{bmatrix} , \tag{22}$$

and the predicted time series is given by

$$\begin{aligned} Z^*(t_n) &= f_0 Z(t_{n-1}) + \ldots + f_{M-2} Z(t_{n-M+1}) + f_{M-1} Z(t_{n-M}) \\ &= \lambda_{M-1} Z(t_{n-1}) + \ldots + \lambda_1 Z(t_{n-M+1}) + \lambda_0 Z(t_{n-M}) , \end{aligned} \tag{23}$$

where we have identified the kriging weights as being related to the prediction filter in the following manner,

$$f_i^{LP} = \lambda_{M-i-1}^{SK} , \quad i = 0,\ldots,M-1. \tag{24}$$

We see that, in this case, linear prediction filtering and kriging are the time reverse of one another. Both methods lead to the same result when applied to the prediction problem for a uniformly sampled time series. This analysis can easily be extended to the case where the prediction distance is greater than unity. Finally, the usual prediction error filter is formed before applying to the data. For a prediction distance of α samples, this would be written as

$$\mathbf{f}^{PE}[1,0,\ldots,0,-f_0,\ldots,-f_{M-1}] , \tag{25}$$

where there are $\alpha - 1$ zeros.

CONCLUSIONS

We have shown that linear prediction theory and simple kriging (assuming certain reasonable choices) both lead to the Wiener-Levinson equations for the prediction filter in the 1D time series case. This connection between two seemingly disparate subjects gives one the ability to use their intuition and experience with one subject for a better understanding of the other. As well, it might foster more cross-disciplinary research between the two fields.

ACKNOWLEDGMENTS

I would like to thank Jerry Schuster of the University of Utah and Chris Liner of the University of Houston (formerly of Saudi Aramco) for helpful reviews.

REFERENCES

Dubrule, O., 2003. Geostatistics for Seismic Data Integration in Earth Models. SEG, Tulsa, OK.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford Univ. Press, Oxford.

Hansen, T.M., Journel, A.G., Tarantola, A. and Mosegaard, K., 2006. Linear inverse Gaussian theory and geostatistics. Geophysics, 71: R101-R111.

Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. J. of the Chem., Metal. and Mining Soc. of South Africa, 52: 119-139.

Matheron, G., 1963. Principles of geostatistics. Economic Geol., 58: 1264-1266.

Piazza, J.L., Sandjivy, L. and Legeron, S., 1997. Use of geostatistics to improve seismic velocities: Case studies. Expanded Abstr., 67th Ann. Internat. SEG Mtg. Dallas: 1293-1296.

Robinson, E.A., 1976. Physical Applications of Stationary Time-Series: with special reference to digital data processing of seismic signals. Macmillan Inc., New York.

Robinson, E.A. and Treitel, S., 2000. Geophysical Signal Analysis. SEG, Tulsa, OK.

Ruiz-Alzola, J., Alberola-López, C. and Westin, C.F., 2005. Kriging filters for multidimensional signal processing. Signal Proces., 85: 413-439.