Block Kriging for Lognormal Spatial Processes

Noel Cressie

Lognormal spatial data are common in mining and soil-science applications. Modeling the underlying spatial process as normal on the log scale is sensible; point kriging allows the whole region of interest to be mapped. However, mining and precision agriculture is carried out selectively and is based on block averages of the process on the original scale. Finding spatial predictions of the blocks assuming a lognormal spatial process has a long history in geostatistics. In this article, we make the case that a particular method for block prediction, overlooked in past times of low computing power, deserves to be reconsidered. In fact, for known mean, it is optimal. We also consider the predictor based on the “law” of permanence of lognormality. Mean squared prediction errors of both are derived and compared both theoretically and via simulation; the predictor based on the permanence-of-lognormality assumption is seen to be less efficient. Our methodology is applied to block kriging of phosphorus to guide precision-agriculture treatment of soil on Broom’s Barn Farm, UK.

KEY WORDS: geostatistics, MSPE, permanence of lognormality, phosphorus, precision agriculture, spatial prediction.

INTRODUCTION

There have been quite a few publications in the past on geostatistics for lognormal data. The themes of these papers (Dowd, 1982; Journel, 1980; Marechal, 1974; Matheron, 1974; Rendu, 1979; Rivoirard, 1990; Roth, 1998) draw on the very best traditions of geostatistics: determine types of variogram models for lognormal data; decide whether to do inference on the original scale or the log scale; choose an optimality criterion for kriging; derive the kriging equations according to the optimality criterion; consider the cases of known or unknown mean (on the log scale); and consider whether knowing just the variogram (on the log scale) is enough to do kriging.

The purpose of this article is to take a fresh look at geostatistics for lognormal data, build on the results of the earlier papers, and develop new results in light of the statistical literature on linear models and transformations. We shall...
vigorously pursue two of the many possibilities, chosen based on the following principles:

- The original scale is for optimality criteria (including unbiasedness) but the log scale is for linear statistical analysis.
- Some form of stationarity is needed for estimation of spatial dependence but it is not needed for spatial prediction (i.e., kriging).
- Kriging is an empirical-Bayes methodology that requires efficient estimators of unknown parameters to be “plugged into” kriging equations.

Notice that “permanence of lognormality” (e.g., Rivoirard, 1990) is not one of our principles. In fact, one of the spatial predictors we consider is based on permanence and one is not, and in this article we give a set of recommendations (based on both geostatistical theory and a carefully designed simulation study) as to when assuming permanence is a reasonable thing to do. In a related paper, Cressie and Pavlicova (2005) find a way to correct the inherent bias in the predictor based on permanence, but still find it to be inefficient.

When discussing lognormality for geostatistical processes, two quite different issues have often arisen together. One has been the development and use of the de Wijsian variogram (Matheron, 1962) for modeling spatial dependence (the variogram has logarithmic shape), and the other has been lognormal kriging. de Wijs (1951) developed a simple model that split an orebody randomly into two halves, one with grade proportionately above average and the other with grade proportionately below average. If this is successively repeated and \( X_i \) is the random grade of one of the halves at the \( i \)th split, then after \( k \) splits, the grade of any one of the \( 2^k \) pieces is distributed as \( \prod_{i=1}^{k} X_i \). As \( k \to \infty \), the central limit theorem implies convergence in distribution to a lognormal random variable. This rather specialized model also gives rise to a variogram that is logarithmic in shape (Matheron, 1962), but it is clear that lognormal genesis can also happen in other ways (e.g., Brown and Sanders, 1981). It is the lognormality that we discuss in this article, and there is no requirement here that variograms be of de Wijsian form.

A very influential piece of writing on lognormal kriging has been the unpublished 43-page “note” by Matheron (1974). Matheron’s approach is to look at the problem from all sides, with many calculations drawn from Matheron (1962) but no definitive conclusions. His writing touches on all the geostatistical themes given earlier. At the time it was written, the statistical influences of linear models, efficient parameter estimation, prediction theory, and workstation computing were not yet felt in geostatistics.

Some notation is needed for what follows in this article. Let the process \( \{ Z(s) : s \in D \} \) denote a lognormal spatial process defined on a domain \( D \subset \mathbb{R}^d \) such that it has a positive \( d \)-dimensional volume \( |D| \). That is, the process:

\[
Y(s) \equiv \log Z(s); \quad s \in D,
\]