Spatial Predictions Based on Skew-normal Models

Previsioni spaziali basate su modelli normali asimmetrici

Cinzia Franceschini
e-mail: cinziafranceschini@msn.com

Nicola Loperfido
Istituto di Scienze Economiche, Università degli Studi di Urbino “Carlo Bo”
Via Saffi, 42 – 61029 (PU), e-mail: nicola@econ.uniurb.it

Riassunto: Le predizioni spaziali basate sul kriging possono essere distorte se il modello non è specificato correttamente. Il presente lavoro considera il caso in cui erroneamente si ipotizzi che i dati provengano da un processo normale stazionario, quando provengono dal modello normale asimmetrico proposto da Kim e Mallick (2004). Si dimostra che l’errore non pregiudica le proprietà del kriging se le posizioni dei dati e del valore da predire sono disposte sui vertici di un poligono regolare.

Keywords: Skew-Normal, Isotropy, Unbiased Predictor, Sampling Design.

1. Introduction

Kriging is the prevalent technique for spatial prediction (Cressie, 1993). It is particularly useful when data are generated from a gaussian random field. However, assuming normality does not seem appropriate when the data are skewed, as it often happens with rainfall data. Kim and Mallick (2004) and Kim, Ha and Mallick (2004) deal with the problem by assuming that data came from a skew-normal process, i.e. their probability density function (pdf, hereafter) is of the form

\[ f(x; \xi, \Omega, \alpha) = 2\phi_p(x; \xi, \Omega)\Phi(\alpha^T(x-\xi)) \]  

(1)

where \( \phi_p(y; \Omega) \) and \( \Phi(y) \) are the pdf and the cumulative distribution function of \( N_p(0, \Omega) \) and \( N(0,1) \) respectively. A random vector whose pdf is (1) is said to have a p-dimensional skew-normal distribution with scale parameter \( \xi \), location parameter \( \Omega \) and shape parameter \( \alpha \), denoted with \( SN_p(\xi, \Omega, \alpha) \). The normal distribution is a skew-normal distribution whose shape parameter is the null vector: \( N_p(\xi, \Omega) \) is equivalent to \( SN_p(\xi, \Omega, 0) \). When observation’s pdf is \( SN_p(\xi, \Omega, \alpha) \) but it is wrongly assumed to be \( N_p(\xi, \Omega) \), statistical predictions based on kriging can be misleading. More precisely, a kriging predictor which is optimal among all unbiased ones under the model \( N_p(\xi, \Omega) \) is, in general, biased and inadmissible under the model \( SN_p(\xi, \Omega, \alpha) \). Moreover, Genton (2004) showed that the sample variogram might not be able to detect the presence of skewness. This paper gives conditions for ordinary kriging to be the best unbiased predictor when data come from a skew-normal process.
2. The model

This section describes the skew-normal model used by Kim et al. (2004) and Kim and Mallick (2004) for spatial predictions. We shall consider several locations \( x_1, \ldots, x_n \), all belonging to a region \( D \) of the \( d \)-dimensional real space. We shall use the vector of observations \( z=(Z(x_1),\ldots,Z(x_n))^T \) to predict the unobserved value \( Z(x_0) \). The joint distribution of \( Z(x_0) \) and \( z \) is often modelled through the assumptions of normality and stationarity, that is:

\[
\begin{pmatrix}
Z(x_0) \\
z
\end{pmatrix} \sim N_{n+1}(\xi_{n+1}, \Omega)
\]

for \( i, j=0, \ldots, n \). The \( i,j \)-th element of the matrix \( \Omega \) only depends on the Euclidean distance between \( x_{i-1} \) and \( x_j \), with \( i=1, \ldots, n+1 \) and \( j=0, 1, \ldots, n \). Kim and Mallick (2004) and Kim et al. (2004) modelled skewness generalizing (2) as follows:

\[
\begin{pmatrix}
Z(x_0) \\
z
\end{pmatrix} \sim SN_{n+1}(\xi_{n+1}, \Omega, \psi_{n+1})
\]

where \( \psi \) is a scalar quantifying skewness. They also considered a more general model, where the location parameter is the product of a design matrix \( F \) and a vector \( \beta \) of unknown coefficients, but statistical analysis found that none of the covariates were relevant. Model (2) satisfies first and second order stationarity assumptions, while model (3), in general, does not. It follows that model (2) is isotropic (correlations only depend on distances between \( x_{i}, x_{j} \)) while model (3) is in general anisotropic (correlations depend on relative orientation of \( x_{0}, x_{1}, \ldots, x_{n} \) as well as on their distances). It is an immediate consequence of expectation and variance of skew-normal distributions not being in general equal to the location and scale parameter, respectively. This poses a problem, from the modelling point of view, when assuming skewness is considered reasonable but assuming anisotropy is not.
3. Prediction

Ordinary kriging (Cressie, 1993) is defined as $\lambda^T z$, where:

$$
\lambda^T = \left( \gamma + \frac{1 - 1_n \Gamma^{-1} \gamma}{1_n \Gamma^{-1} 1_n} \right)^T \Gamma^{-1} \gamma = \begin{bmatrix}
\gamma(x_0 - x_i) \\
\vdots \\
\gamma(x_0 - x_n)
\end{bmatrix}
$$

$$
\Gamma = \begin{bmatrix}
\gamma(x_h - x_k) \\
\vdots \\
\gamma(x_i - x_j)
\end{bmatrix} \qquad V(Z(x) - Z(x_j)) = \gamma
$$

for $h, k = 1, \ldots, n$ and $i, j = 0, \ldots, n$. It easily follows that $\lambda^T z$ is an unbiased predictor for $Z(x_0)$ under model (2), but not necessarily under model (3). It is a direct consequence of expectation and variance of $(Z(x_0), z)^T$ not being in general equal to $\xi 1_{n+1}$ and $\Omega$, respectively. The following theorem states a condition on the design scheme which guarantees unbiasedness and optimality for kriging under model (3).

**THEOREM 1:** Let model (3) hold and let the locations $x_0, x_1, \ldots, x_n$ be located on the vertices of a regular polygon. Then $\lambda^T z$ is the optimal predictor, under squared loss, among all unbiased ones which are linear functions of $Z(x_1), \ldots, Z(x_n)$.

**PROOF:** By assumption, the joint distribution of $Z(x_0)$ and $z$ is $SN_{n+1}(\xi 1_{n+1}, \Omega, \psi 1_{n+1})$. Hence the corresponding expectation is (Dalla Valle, 2004):

$$
E\left( \frac{Z(x_0)}{z} \right) = \frac{2}{\pi} \frac{\psi \Omega_{1_{n+1}}}{\sqrt{1 + \psi^2 1_{n+1}^T \Omega 1_{n+1}}}
$$

By assumption, the value $\omega_{ij}$ of the matrix $\Omega$ only depends on the Euclidean distance between locations $x_i$ and $x_j$ ($i, j = 0, 1, \ldots, n$). Since locations $x_0, x_1, \ldots, x_n$ are located on the vertices of a regular polygon, the scale matrix $\Omega$ is symmetric circulant (Olkin and Press, 1969). Hence the unit vector $1_{n+1}/(n+1)^{0.5}$ is a normalized eigenvector of $\Omega$. Let $\varepsilon/(n+1)^{0.5}$ be the corresponding eigenvalue. Hence we can write

$$
E\left( \frac{Z(x_0)}{z} \right) = \xi 1_{n+1} + \frac{2}{\pi} \frac{\psi \varepsilon 1_{n+1}}{\sqrt{1 + \varepsilon^2 (n+1)}} = \left( \xi + \frac{2}{\pi} \frac{\psi \varepsilon}{\sqrt{1 + \varepsilon^2 (n+1)}} \right) 1_{n+1}
$$

It follows that expectations of $Z(x_0), Z(x_1), \ldots, Z(x_n)$ are all equal and that ordinary kriging $\lambda^T z$ is an unbiased predictor for $Z(x_0)$, due to the constraint $\lambda^T = 1$. Efficiency of $\lambda^T z$ among all predictors of the form $l^T z$, where $l^T 1_{n+1} = 1$ easily follow from basic properties of ordinary kriging (Cressie, 1993). The proof is then complete.

The following theorem shows that taking observations on the vertices of a regular polygon implies isotropy, under model (3).
THEOREM 2: Under the assumptions of the above theorem, the joint distribution of \( Z(x_0) \) and \( z \) is isotropic.

PROOF: Under model (3), the variance matrix of \( Z(x_0) \) and \( z \) is (Dalla Valle, 2004):

\[
V\left( \begin{array}{c} Z(x_0) \\ z \end{array} \right) = \Omega - \frac{2 \psi^2 \varepsilon^2 1_{n+1} \Omega \Omega_{n+1}^T}{\pi 1 + \psi^2 1_{n+1}^T \Omega 1_{n+1}}
\]

From the proof of theorem 1, we know that the unit vector \( 1_{n+1}/(n+1)^{0.5} \) is a normalized eigenvector of \( \Omega \) and that \( \varepsilon (n+1)^{0.5} \) is the corresponding eigenvalue. Hence

\[
V\left( \begin{array}{c} Z(x_0) \\ z \end{array} \right) = \Omega - \frac{2 \psi^2 \varepsilon^2 1_{n+1} \Omega_{n+1}^T}{\pi 1 + (n+1)\varepsilon\psi^2}
\]

By assumption, The scale parameter is a matrix whose \( i,j \)-th element only depends on the Euclidean distance between \( x_{i-1} \) and \( x_j \), with \( i=1,...,n+1 \) and \( j=0,1,...,n \). It follows that \( C[Z(x_i),Z(x_j)]=w(||x_i-x_j||)-(2/\pi)(\psi\varepsilon)^2/[1+(n+1)\varepsilon\psi^2] \) for \( i,j=0,1,...,n \). Hence covariances between \( Z(x_i) \) and \( Z(x_j) \) only depend on the Euclidean distance between \( x_i \) and \( x_j \). The model is then isotropic and this completes the proof.

References


