Mathematical Geology, Vol. 32, No. 1, 2000

The Correlation Structure of Matheron's Classical Variogram Estimator Under Elliptically Contoured Distributions[1]

Marc G. Genton[2]

The classical variogram estimator proposed by Matheron can be written as a quadratic form of the observations. When data have an elliptically contoured distribution with constant mean, the correlation between the classical variogram estimator at two different lags is a function of the spatial design matrix, the covariance matrix, and the kurtosis. Several specific cases are studied closely. A subclass of elliptically contoured distributions with a particular family of covariance matrices is shown to possess exactly the same correlation structure for the classical variogram estimator as the multivariate independent Gaussian distribution. The consequences on variogram fitting by generalized least squares are discussed.

KEY WORDS: variogram estimation, quadratic form, kurtosis, variogram fitting, generalized least squares.

INTRODUCTION

Variogram estimation is a crucial stage of spatial prediction, because it determines the kriging weights. Today, the most widely used variogram estimator is certainly the one proposed by Matheron (1962), although it is highly nonrobust to outliers in the data (Cressie, 1993; Genton, 1998a, 1998c). The main reasons for this popularity are its simple appealing formulation and unbiasedness property. If $\{Z(\mathbf{x}) : \mathbf{x} \in D \subset \mathbb{R}^d\}, d \geq 1$, is a spatial stochastic process, ergodic, and intrinsically stationary, Matheron's classical variogram estimator, based on the method-of-moments, is

$$2\hat{\gamma}(\mathbf{h}) = \frac{1}{N_{\mathbf{h}}} \sum_{N(\mathbf{h})} (Z(\mathbf{x}_i) - Z(\mathbf{x}_j))^2, \qquad \mathbf{h} \in \mathbb{R}^d, \tag{1}$$

[1]Received July 1998; accepted 24 February 1999.

[2]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA. e-mail: genton@math.mit.edu

where $N(\mathbf{h}) = \{(\mathbf{x}_i, \mathbf{x}_j) : \mathbf{x}_i - \mathbf{x}_j = \mathbf{h}\}$ and $N_{\mathbf{h}}$ is the cardinality of $N(\mathbf{h})$. The simple formulation of this estimator allows (1) to be written as a quadratic form. In fact, if $\mathbf{z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))^T$ is the data vector and $A(\mathbf{h})$ is the spatial design matrix of the data at lag $\mathbf{h}$, then

$$2\hat{\gamma}(\mathbf{h}) = \mathbf{z}^T A(\mathbf{h})\mathbf{z}. \tag{2}$$

The spatial design matrix $A(\mathbf{h})$ is a symmetric matrix of size $n \times n$, derived from Equation (1). For regularly spaced data in $\mathbb{R}^1$, $A(h)$ has three possible forms depending on h ($h < n/2$, $h = n/2$, and $h > n/2$). For instance, if $h < n/2$, then

$$A(h) = \frac{1}{n-h}\left(\begin{array}{ccc|ccc|ccc}
1 & & & -1 & & & & & \\
 & \ddots & & & \ddots & & & O & \\
 & & 1 & & & -1 & & & \\
\hline
-1 & & & 2 & & & -1 & & \\
 & \ddots & & & \ddots & & & \ddots & \\
 & & -1 & & & 2 & & & -1 \\
\hline
 & & & -1 & & & 1 & & \\
 & O & & & \ddots & & & \ddots & \\
 & & & & & -1 & & & 1
\end{array}\right).$$

For data on a regularly spaced multidimensional grid in $\mathbb{R}^d$, $d > 1$, the spatial design matrix $A(\mathbf{h})$ can be split along each axis of the grid and described by Kronecker products of matrices (Genton, 1998b). An important issue is to understand the statistical properties of the variogram estimator (2). For data with a Gaussian distribution, the mean and variance (Cressie, 1993) of $2\hat{\gamma}(\mathbf{h})$, as well as its correlation structure (Genton, 1998b), are easily computed. The next theorem summarizes their results.

Theorem 1. *Let* $\mathbf{z}$ *be a random vector with expectation* $E(\mathbf{z}) = \mu\mathbf{1}_n$ *and covariance matrix* $Var(\mathbf{z}) = \Sigma$. *Then, the expectation of Matheron's classical variogram estimator is*

(a) $E(2\hat{\gamma}(\mathbf{h})) = tr[A(\mathbf{h})\Sigma]$.

Moreover, if $\mathbf{z}$ *is Gaussian, then*

(b) $Var(2\hat{\gamma}(\mathbf{h})) = 2\,tr[A(\mathbf{h})\Sigma A(\mathbf{h})\Sigma]$,
(c) $Cov(2\hat{\gamma}(\mathbf{h}_1), 2\hat{\gamma}(\mathbf{h}_2)) = 2\,tr[A(\mathbf{h}_1)\Sigma A(\mathbf{h}_2)\Sigma]$,
(d) $Corr(2\hat{\gamma}(\mathbf{h}_1), 2\hat{\gamma}(\mathbf{h}_2)) = \dfrac{\{tr[A(\mathbf{h}_1)\Sigma A(\mathbf{h}_2)\Sigma]\}}{\{\sqrt{tr[A(\mathbf{h}_1)\Sigma A(\mathbf{h}_1)\Sigma]\, tr[A(\mathbf{h}_2)\Sigma A(\mathbf{h}_2)\Sigma]}\}}$,

where $tr[\cdot]$ *is the trace operator.*

Although gaussianity is a nice assumption that makes many mathematical problems tractable, it is not always met in practice, or sometimes only approximately. This problem is addressed in this paper by considering elliptically contoured distributions, a wide class of multivariate distributions, which generalizes the Gaussian one. The correlation structure of Matheron's classical variogram estimator is computed under elliptically contoured distributions in the next section. It is shown to depend on the spatial design matrix $A(\mathbf{h})$, the covariance matrix Σ of the data, and kurtosis parameter κ. The second section presents a subclass of elliptically contoured distributions, with a particular family of covariance matrices Σ, that yields exactly the same correlation structure as for the multivariate independent Gaussian distribution. In the last section, the correlation structure is computed and depicted with graphics for a covariance matrix Σ based on a spherical underlying variogram. The effects of the range of this variogram and of the kurtosis parameter κ on the correlation structure are pointed out. The consequences on the generalized least squares method with an explicit covariance structure (GLSE) for variogram fitting (Genton, 1998b) are discussed.

ELLIPTICALLY CONTOURED DISTRIBUTIONS

Recall some concepts on elliptically contoured distributions (Fang, Kotz, and Ng, 1989; Fang and Zhang, 1990; Fang and Anderson, 1990). A random vector $\mathbf{z} \in \mathbb{R}^n$ is said to have an elliptically contoured distribution $EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$ if its characteristic function has the form

$$e^{i\mathbf{t}^T\boldsymbol{\mu}}\phi(\mathbf{t}^T\Sigma^*\mathbf{t}), \tag{3}$$

where $i = \sqrt{-1}$, $\mathbf{t} \in \mathbb{R}^n$, $\boldsymbol{\mu} \in \mathbb{R}^n$, $\Sigma^* \in \mathbb{R}^{n\times n}$ is a positive definite matrix, and ϕ a real function such that (3) be a characteristic function. The expectation of $\mathbf{z}$ is $E(\mathbf{z}) = \boldsymbol{\mu}$ and its covariance matrix is Var($\mathbf{z}$)$= \Sigma = -2\phi'(0)\Sigma^*$. This is a general class of distributions whose contours of equal density have the same elliptical shape as the multivariate Gaussian, but that contains long-tailed and short-tailed distributions. Some important subclasses of elliptically contoured distributions are the Kotz type, Pearson type, multivariate t, multivariate Cauchy, multivariate Bessel, logistic, scale mixture, and of course the multivariate Gaussian with $\phi(u) = e^{-u/2}$. If $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma^* = I_n$, then $\mathbf{z}$ has a spherical distribution, and the contours of equal density have a circular shape. Recently, some Q–Q probability plots were proposed (Li, Fang, and Zhu, 1997) to test spherical and elliptical symmetry in the data. Muirhead (1982) defines the kurtosis parameter κ of an elliptically contoured distribution $EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$ as

$$\kappa = \frac{\phi''(0)}{(\phi'(0))^2} - 1, \tag{4}$$

where $\phi'(0)$ and $\phi''(0)$ are the first and the second derivatives of ϕ, evaluated at zero. In particular, the kurtosis parameter is equal to zero for the multivariate Gaussian distribution. Subsequently, it will be shown that Matheron's classical variogram estimator (2) has the same correlation structure for elliptically contoured distributions with kurtosis parameter $\kappa = 0$ as for the multivariate independent Gaussian distribution. First, the general correlation structure is established in the following theorem.

Theorem 2. *Let* $\mathbf{z} \sim EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$ *be a random vector with elliptically contoured distribution, where* $\boldsymbol{\mu} = \mu \mathbf{1}_n$ *and the kurtosis parameter is* κ. *Then, the correlation structure of Matheron's classical variogram estimator* $2\hat{\gamma}(\mathbf{h}) = \mathbf{z}^T A(\mathbf{h})\mathbf{z}$ *is*

$$Corr(2\hat{\gamma}(\mathbf{h}_1), 2\hat{\gamma}(\mathbf{h}_2)) = \frac{T(\kappa, \mathbf{h}_1, \mathbf{h}_2, \Sigma^*)}{\sqrt{T(\kappa, \mathbf{h}_1, \mathbf{h}_1, \Sigma^*) T(\kappa, \mathbf{h}_2, \mathbf{h}_2, \Sigma^*)}},$$

where

$$T(\kappa, \mathbf{h}_1, \mathbf{h}_2, \Sigma^*) = \kappa\, tr[A(\mathbf{h}_1)\Sigma^*]\, tr[A(\mathbf{h}_2)\Sigma^*] + 2(\kappa + 1)\, tr[A(\mathbf{h}_1)\Sigma^* A(\mathbf{h}_2)\Sigma^*].$$

Proof. This result is derived from multivariate analysis of quadratic forms for elliptically contoured distributions (Li, 1987):

$$\begin{aligned} &\mathrm{Cov}(\mathbf{z}^T C_1 \mathbf{z}, \mathbf{z}^T C_2 \mathbf{z}) \\ &\quad = 4(\phi''(0) - \phi'(0)^2)\, \mathrm{tr}[C_1 \Sigma^*]\, \mathrm{tr}[C_2 \Sigma^*] + 8\phi''(0)\, \mathrm{tr}[C_1 \Sigma^* C_2 \Sigma^*] \\ &\qquad - 2\phi'(0) \boldsymbol{\mu}^T (C_1 \Sigma^* C_2 + 3 C_2 \Sigma^* C_1) \boldsymbol{\mu}, \end{aligned}$$

where C_1 and C_2 are real symmetric matrices. Setting $C_1 = A(\mathbf{h}_1)$, $C_2 = A(\mathbf{h}_2)$, and using the property $A(\mathbf{h})\boldsymbol{\mu} = \mathbf{0}$, $\forall \mathbf{h}$, as well as the definition of the kurtosis parameter κ, yields:

$$\begin{aligned} &\mathrm{Cov}(\mathbf{z}^T A(\mathbf{h}_1)\mathbf{z}, \mathbf{z}^T A(\mathbf{h}_2)\mathbf{z}) \\ &\quad = 4\kappa\phi'(0)^2\, \mathrm{tr}[A(\mathbf{h}_1)\Sigma^*]\, \mathrm{tr}[A(\mathbf{h}_2)\Sigma^*] + 8(\kappa + 1)\phi'(0)^2\, \mathrm{tr}[A(\mathbf{h}_1)\Sigma^* A(\mathbf{h}_2)\Sigma^*], \end{aligned}$$

and thus proves this theorem after simplification by $4\phi'(0)^2$ in the formula for the correlation. ■

Corollary 2.1. *Let* $\mathbf{z} \sim EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$ *be a random vector with elliptically contoured distribution, where* $\boldsymbol{\mu} = \mu \mathbf{1}_n$ *and the kurtosis parameter is* $\kappa = 0$. *Then, the correlation structure of Matheron's classical variogram estimator*

$2\hat{\gamma}(\mathbf{h}) = \mathbf{z}^T A(\mathbf{h})\mathbf{z}$ *is*

$$Corr(2\hat{\gamma}(\mathbf{h}_1), 2\hat{\gamma}(\mathbf{h}_2)) = \frac{tr[A(\mathbf{h}_1)\Sigma^* A(\mathbf{h}_2)\Sigma^*]}{\sqrt{tr[A(\mathbf{h}_1)\Sigma^* A(\mathbf{h}_1)\Sigma^*]\, tr[A(\mathbf{h}_2)\Sigma^* A(\mathbf{h}_2)\Sigma^*]}}.$$

Theorem 2 gives the general formula for the correlation structure, which depends on the spatial design matrix $A(\mathbf{h})$, the covariance matrix Σ of the data, and the kurtosis parameter κ. Positive values of κ correspond to long-tailed distributions, whereas negative values correspond to short-tailed ones. However, note that κ has a greatest lower bound (Bentler and Berkane, 1986) given by $\kappa > -2/(n+2)$ in order to ensure positive definiteness of the covariance matrix Σ. Corollary 2.1 presents the correlation structure when the kurtosis parameter is $\kappa = 0$. In particular, this is true for the multivariate Gaussian distribution. Note furthermore that the matrix Σ^* in the correlation formula of Theorem 2 and Corollary 2.1 can be replaced by Σ without any change, thus being in agreement with Theorem 1.

A PARTICULAR CLASS OF COVARIANCE MATRICES

Suppose that the covariance matrix Σ of the data belongs to the particular family $\mathcal{S}$ of matrices:

$$\mathcal{S} = \left\{\Sigma \,\middle|\, \Sigma = \alpha I_n + \mathbf{1}_n\mathbf{a}^T + \mathbf{a}\mathbf{1}_n^T\right\}, \tag{5}$$

where $\alpha \in \mathbb{R}$ and $\mathbf{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$ are defined in such a way that Σ is positive definite. For instance, simple computations show that the eigenvalues of a covariance matrix $\Sigma \in \mathcal{S}$ are α with multiplicity $n-2$ and

$$\alpha + \sum_{i=1}^n a_i \pm \left(n\sum_{i=1}^n a_i^2\right)^{1/2}.$$

Therefore, for any vector $\mathbf{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$, one can choose $\alpha > 0$ such that

$$\alpha > \left(n\sum_{i=1}^n a_i^2\right)^{1/2} - \sum_{i=1}^n a_i, \tag{6}$$

in order to ensure positive definiteness of Σ. For this particular family $\mathcal{S}$ of matrices, the formula of Theorem 2 for the correlation reduces to the expression given in the next theorem, which does not depend on the matrix Σ (i.e., does not depend on the vector $\mathbf{a}$, nor on the constant α).

Theorem 3. *Let* $\mathbf{z} \sim EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$ *be a random vector with elliptically contoured distribution, where* $\boldsymbol{\mu} = \mu \mathbf{1}_n$, $\Sigma = -2\phi'(0)\Sigma^* \in \mathcal{S}$ *and the kurtosis parameter is* κ. *Then, the correlation structure of Matheron's classical variogram estimator* $2\hat{\gamma}(\mathbf{h}) = \mathbf{z}^T A(\mathbf{h})\mathbf{z}$ *is*

$$Corr(2\hat{\gamma}(\mathbf{h}_1), 2\hat{\gamma}(\mathbf{h}_2)) = \frac{2\kappa + (\kappa+1)\, tr[A(\mathbf{h}_1)A(\mathbf{h}_2)]}{\sqrt{(2\kappa + (\kappa+1)\, tr[A^2(\mathbf{h}_1)])(2\kappa + (\kappa+1)\, tr[A^2(\mathbf{h}_2)])}}.$$

Proof. First compute

$$\begin{aligned}
&\mathrm{tr}[A(\mathbf{h}_1)\Sigma^* A(\mathbf{h}_2)\Sigma^*] \\
&\quad= \frac{1}{4\phi'(0)^2}\mathrm{tr}\big[A(\mathbf{h}_1)\big(\alpha I_n + \mathbf{1}_n\mathbf{a}^T + \mathbf{a}\mathbf{1}_n^T\big)A(\mathbf{h}_2)\big(\alpha I_n + \mathbf{1}_n\mathbf{a}^T + \mathbf{a}\mathbf{1}_n^T\big)\big] \\
&\quad= \frac{1}{4\phi'(0)^2}\mathrm{tr}\big[\big(\alpha A(\mathbf{h}_1) + A(\mathbf{h}_1)\mathbf{1}_n\mathbf{a}^T + A(\mathbf{h}_1)\mathbf{a}\mathbf{1}_n^T\big)\big(\alpha A(\mathbf{h}_2) + A(\mathbf{h}_2)\mathbf{1}_n\mathbf{a}^T \\
&\qquad + A(\mathbf{h}_2)\mathbf{a}\mathbf{1}_n^T\big)\big] \\
&\quad= \frac{1}{4\phi'(0)^2}\mathrm{tr}\big[\big(\alpha A(\mathbf{h}_1) + A(\mathbf{h}_1)\mathbf{a}\mathbf{1}_n^T\big)\big(\alpha A(\mathbf{h}_2) + A(\mathbf{h}_2)\mathbf{a}\mathbf{1}_n^T\big)\big] \\
&\quad= \frac{\alpha^2}{4\phi'(0)^2}\mathrm{tr}[A(\mathbf{h}_1)A(\mathbf{h}_2)],
\end{aligned}$$

by using the fact that $A(\mathbf{h})\mathbf{1}_n = \mathbf{0}$, $\forall \mathbf{h}$. Second, notice that as a special case, we have

$$\mathrm{tr}[A(\mathbf{h})\Sigma^*] = \frac{\alpha}{-2\phi'(0)}\mathrm{tr}[A(\mathbf{h})].$$

Combining these two formulas yields:

$$T(\kappa, \mathbf{h}_1, \mathbf{h}_2, \Sigma^*) = \frac{\alpha^2}{4\phi'(0)^2}[\kappa\, \mathrm{tr}[A(\mathbf{h}_1)]\, \mathrm{tr}[A(\mathbf{h}_2)] + 2(\kappa+1)\, \mathrm{tr}[A(\mathbf{h}_1)A(\mathbf{h}_2)]],$$

from which the formula for the correlation in the theorem's statement follows, using $\mathrm{tr}[A(\mathbf{h})] = 2$, $\forall \mathbf{h}$. ■

Corollary 3.1. *Let* $\mathbf{z} \sim EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$ *be a random vector with elliptically contoured distribution, where* $\boldsymbol{\mu} = \mu\mathbf{1}_n$, $\Sigma = -2\phi'(0)\Sigma^* \in \mathcal{S}$ *and the kurtosis* *parameter is* $\kappa = 0$. *Then, the correlation structure of Matheron's classical variogram estimator* $2\hat{\gamma}(\mathbf{h}) = \mathbf{z}^T A(\mathbf{h})\mathbf{z}$ *is*

$$Corr(2\hat{\gamma}(\mathbf{h}_1), 2\hat{\gamma}(\mathbf{h}_2)) = \frac{tr[A(\mathbf{h}_1)A(\mathbf{h}_2)]}{\sqrt{tr[A^2(\mathbf{h}_1)]\, tr[A^2(\mathbf{h}_2)]}}.$$

The correlation in Theorem 3 depends on the spatial design matrix $A(\mathbf{h})$ and the kurtosis parameter κ, but not on the covariance matrix $\Sigma \in \mathcal{S}$. Corollary 3.1 is a special case with $\kappa = 0$, yielding a correlation that depends only on the spatial design matrix $A(\mathbf{h})$. Genton (1998b) has a simple explicit formula for this latter case with data in $\mathbb{R}^1$, as well as in $\mathbb{R}^d$, $d > 1$, by using Kronecker products of matrices (Fang and Zhang, 1990).

The family $\mathcal{S}$ contains several interesting structures. In particular, the uncorrelated case is obtained by letting $\alpha = \sigma^2$ and $\mathbf{a} = \mathbf{0}$, thus yielding $\Sigma = \sigma^2 I_n$. Note that "uncorrelated" becomes "independent" for the multivariate Gaussian distribution. The equicorrelation case is also a member of $\mathcal{S}$ obtained by letting $\alpha = 1 - \rho$ and $\mathbf{a} = (\rho/2)\mathbf{1}_n$, thus yielding

$$\Sigma = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}.$$

Other choices of $\mathbf{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}^n$ and $\alpha > 0$ satisfying (6) yield a wide range of dependency structures. In all these cases, the correlation structure of Matheron's classical variogram estimator depends only on the spatial design matrix $A(\mathbf{h})$, but not on the vector $\mathbf{a}$, nor on the constant α.

THE CORRELATION STRUCTURE

We start by describing the correlation structure of Matheron's classical variogram estimator, and afterwards discuss its effect on variogram fitting.

Let $\mathbf{z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))^T$ be the data vector, with unidimensional and regular support for simplicity, i.e., $\mathbf{x}_i \in \mathbb{R}^1$, $i = 1, \ldots, n$. The case when $\mathbf{x}_i \in \mathbb{R}^d$, $i = 1, \ldots, n$, $d > 1$, is very similar as shown by Genton (1998b), using Kronecker products of matrices (Fang and Zhang, 1990) to describe the spatial design matrix $A(\mathbf{h})$. Consider a spherical variogram model given by

$$\gamma(h, a, b, c) = \begin{cases} 0 & \text{if } h = 0, \\ a + b\left(\frac{3}{2}\left(\frac{h}{c}\right) - \frac{1}{2}\left(\frac{h}{c}\right)^3\right) & \text{if } 0 < h \le c, \\ a + b & \text{if } h > c, \end{cases} \tag{7}$$

where a, b, and c are nonnegative real numbers. Suppose the data vector $\mathbf{z}$ has an elliptically contoured distribution $EC_n(\boldsymbol{\mu}, \Sigma^*, \phi)$, with $\boldsymbol{\mu} = \mu\mathbf{1}_n$, and covariance matrix Σ based on the spherical variogram (7), i.e., $\Sigma_{ij} = a + b - \gamma(|i - j|)$. The correlation between Matheron's classical variogram estimator at two different lags is given by the formula in Theorem 2, and represented in Figure 1 for a sample size $n = 50$. The first column corresponds to $c = 1$ (pure nugget), and the second one to a range $c = 10$, both with a sill of $a + b = 1$. The three rows correspond to a kurtosis parameter κ of 0, 0.5, and 1, respectively.

The differences between the two columns describe the effect of the range on the correlation structure, and the differences between the three rows describe the effect of the kurtosis parameter. When κ increases, the correlation becomes larger, i.e., the variogram estimates become more and more dependent. Note that $\kappa > 0$ indicates more heavy-tailed distributions. Negative values of κ have the lower bound $-2/(n + 2)$ and therefore tend to zero when the sample size n increases, leading to plots similar to the case $\kappa = 0$. The use of other variogram models did not show significant differences from the plots in Figure 1.

Variogram fitting is the second crucial stage of spatial prediction because it determines the kriging weights. Because variogram estimates at different spatial lags are correlated, variogram fitting by ordinary least squares is not satisfactory. Genton (1998b) describes variogram fitting by generalized least squares with an explicit formula for the covariance structure (GLSE). It consists in finding the estimator $\hat{\boldsymbol{\theta}}$ which minimizes:

$$G(\boldsymbol{\theta}) = (2\hat{\boldsymbol{\gamma}} - 2\boldsymbol{\gamma}(\boldsymbol{\theta}))^T \Omega^{-1} (2\hat{\boldsymbol{\gamma}} - 2\boldsymbol{\gamma}(\boldsymbol{\theta})), \tag{8}$$

where $2\hat{\boldsymbol{\gamma}} = (2\hat{\gamma}(\mathbf{h}_1), \dots, 2\hat{\gamma}(\mathbf{h}_k))^T \in \mathbb{R}^k$ is the random vector with covariance matrix $\mathrm{Var}(2\hat{\boldsymbol{\gamma}}) = \Omega$, with $\mathbf{h}_i = i\mathbf{h}/\|\mathbf{h}\|$, $i = 1, \dots, k$, and $2\boldsymbol{\gamma}(\boldsymbol{\theta}) = (2\gamma(\mathbf{h}_1, \boldsymbol{\theta}), \dots, 2\gamma(\mathbf{h}_k, \boldsymbol{\theta}))^T \in \mathbb{R}^k$ is the vector of a valid parametric variogram. Journel and Huijbregts (1978) suggest using $k = K/2$, where K is the maximal possible distance between data in the direction $\mathbf{h}$. Genton (1998b) proposes using the matrix Ω defined by

$$\Omega_{ij} = \mathrm{Corr}(2\hat{\gamma}(\mathbf{h}_i), 2\hat{\gamma}(\mathbf{h}_j))\gamma(\mathbf{h}_i, \boldsymbol{\theta})\gamma(\mathbf{h}_j, \boldsymbol{\theta})/\sqrt{N_i N_j}, \tag{9}$$

where N_i is the number of differences at lag $\mathbf{h}_i$, and the correlation $\mathrm{Corr}(2\hat{\gamma}(\mathbf{h}_i), 2\hat{\gamma}(\mathbf{h}_j))$ is computed with an explicit formula in the multivariate independent Gaussian case, i.e., with $\Sigma = \sigma^2 I_n$. Corollary 3.1 extends the validity of this formula, showing that it is still true for elliptically contoured distributions with kurtosis parameter $\kappa = 0$, and covariance matrix $\Sigma \in \mathcal{S}$. In particular, the formula is valid for all multivariate Gaussian distributions with covariance matrix $\Sigma \in \mathcal{S}$.

Figure 1 shows that the correlation structure of Matheron's classical variogram estimator under elliptically contoured distributions still depends on the kurtosis

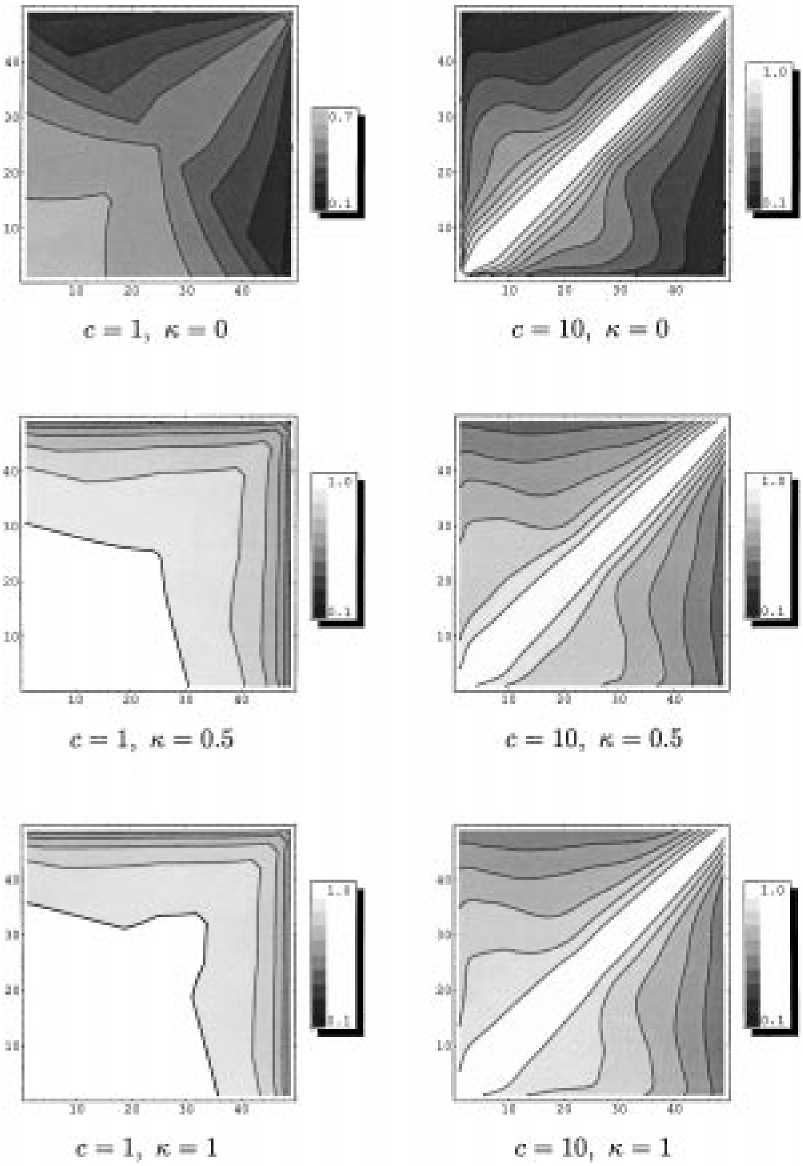

Figure 1. These plots show the dependence of the correlation structure $\mathrm{Corr}(2\hat{\gamma}(h_1), 2\hat{\gamma}(h_2))$ of Matheron's classical variogram estimator on the lags h_1 (horizontal axis) and h_2 (vertical axis), for data with an elliptically contoured distribution and sample size $n = 50$. The covariance matrix Σ is based on a spherical variogram with range $c = 1$ (pure nugget), $c = 10$, and the kurtosis parameter is $\kappa = 0, 0.5, 1$. The plots show the contour lines for correlations between 0 and 0.1, between 0.1 and 0.2, etc.

parameter κ. Therefore, the formula for the correlation $\mathrm{Corr}(2\hat{\gamma}(\mathbf{h}_i), 2\hat{\gamma}(\mathbf{h}_j))$ in Equation (9) can be improved by using Theorem 3 and estimating κ from the data. For instance, a consistent estimator $\hat{\kappa}$ can be found in Waternaux (1976), or Muirhead and Waternaux (1980). Note that in this situation, the covariance matrix Ω is positive definite because the correlation matrix defined by $\mathrm{Corr}(2\hat{\gamma}(\mathbf{h}_i), 2\hat{\gamma}(\mathbf{h}_j))$ is positive definite. This is a direct consequence of Schur's theorem (Horn and Johnson, 1991), as is shown by Genton (1998b) for the multivariate Gaussian case.

For other variogram estimators that cannot be expressed as a quadratic form of the data, like the one proposed by Cressie and Hawkins (1980), or the highly robust one proposed by Genton (1998a), closed forms of the correlation structure are not available. However, simulations like the ones in Genton (1998b) show that these correlation structures are not very different from the one of Matheron's classical variogram estimator. This is especially true for the highly robust variogram estimator. Therefore, one can still use the results of the previous theorems and corollaries as approximations for the correlation structure of other variogram estimators.

CONCLUSIONS

In this paper, the correlation structure of Matheron's classical variogram estimator has been computed when data have an elliptically contoured distribution. In general, this correlation structure depends on the spatial design matrix, the covariance matrix, and the kurtosis. However, for a subclass of elliptically contoured distributions with a particular family of covariance matrices, the correlation structure depends only on the spatial design matrix. Moreover, it is exactly the same as for the multivariate independent Gaussian distribution. This result allows one to extend the validity of the method of variogram fitting proposed by Genton (1998b), which is based on generalized least squares with an explicit formula for the covariance structure (GLSE).

REFERENCES

Bentler, P. M., and Berkane, M., 1986, Greatest lower bound to the elliptical theory kurtosis parameter: Biometrika, v. 73, no. 1, p. 240–241.

Cressie, N., and Hawkins, D. M., 1980, Robust estimation of the variogram, I: Math. Geology, v. 12, no. 2, p. 115–125.

Cressie, N., 1993, Statistics for spatial data: John Wiley & Sons, New York, 900 p.

Fang, K.-T., Kotz, S., and Ng, K.-W., 1989, Symmetric multivariate and related distributions: Chapman & Hall, London, 200 p.

Fang, K.-T., and Zhang, Y.-T., 1990, Generalized Multivariate Analysis: Springer, Berlin, 220 p.

Fang, K.-T., and Anderson, T. W., 1990, Statistical Inference in Elliptically Contoured and Related Distributions: Allerton Press, New York, 498 p.

Genton, M. G., 1998a, Highly robust variogram estimation: Math. Geology, v. 30, no. 2, p. 213–221.

Genton, M. G., 1998b, Variogram fitting by generalized least squares using an explicit formula for the covariance structure: Math. Geology, v. 30, no. 4, p. 323–345.

Genton, M. G., 1998c, Spatial breakdown point of variogram estimators: Math. Geology, v. 30, no. 7, p. 853–871.

Horn, R. A., and Johnson, C. R., 1991, Topics in matrix analysis: Cambridge University Press, Cambridge, 607 p.

Journel, A. G., and Huijbregts, Ch. J., 1978, Mining geostatistics: Academic Press, London, 600 p.

Li, G., 1987, Moments of a random vector and its quadratic forms: Jour. Stat. Appl. Prob., v. 2, p. 219–229.

Li, R.-Z., Fang, K.-T., and Zhu, L.-X., 1997, Some Q-Q probability plots to test spherical and elliptical symmetry: Jour. Comp. Graph. Stat., v. 6, p. 435–450.

Matheron, G., 1962, Traité de géostatistique appliquée, Tome I: Mémoires du Bureau de Recherches Géologiques et Minières, no. 14, Editions Technip, Paris, 333 p.

Muirhead, R. J., 1982, Aspects of multivariate statistical theory: Wiley, New York, 673 p.

Muirhead, R. J., and Waternaux, C. M., 1980, Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations: Biometrika, v. 67, p. 31–43.

Waternaux, C. M., 1976, Asymptotic distribution of the sample roots for a nonnormal population: Biometrika, v. 63, p. 639–646.