ELSEVIER

Computational Statistics & Data Analysis 41 (2002) 47–57

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Nonparametric variogram and covariogram estimation with Fourier–Bessel matrices

Marc G. Genton[a,*], David J. Gorsich[b]

[a]Department of Statistics, 209-D, North Carolina State University, Raleigh, NC 27695-8203, USA
[b]National Automotive Center, U.S. Army TACOM, AMSTA-TR-N (MS 289), Warren, MI 48397-5000, USA

Abstract

The nonparametric estimation of variograms and covariograms for isotropic stationary spatial stochastic processes is considered. It is shown that Fourier–Bessel matrices play an important role in this context because they provide an orthogonal discretization of the spectral representation of positive definite functions. Their properties are investigated and an example from a simulated two-dimensional spatial process is provided. It is shown that this approach provides a smooth and positive definite nonparametric estimator in the continuum, whereas previous methods typically suffer from spurious oscillations. A practical example from Astronomy is used for illustration.

Keywords: Discretization; Fourier–Bessel expansion; Kriging; Nonnegative least squares; Orthogonality; Positive definiteness; Spatial statistics

1. Introduction

Kriging is a widely used method of spatial interpolation (e.g. Chilès and Delfiner, 1999), particularly in earth and environmental sciences. It is based on the variogram or the covariogram, two functions which describe the spatial dependence. A reliable estimation of the latter is therefore crucial. A traditional approach, under stationarity, consists of estimating the variogram or covariogram at several spatial lags and fitting a valid parametric model to those estimates. The main disadvantage with this method is that the practitioner is required to select a model, an often rather subjective

* Corresponding author. Tel.: +1-919-515-1936; fax: +1-919-515-1169.
E-mail address: genton@stat.ncsu.edu (M.G. Genton).
URL: http://www.stat.ncsu.edu/∼mggenton/

PII: S0167-9473(02)00062-2

task (Gorsich and Genton, 2000). Moreover, this choice typically implies a particular smoothness of the spatial process, determined by the behavior of the variogram or covariogram at the origin, for example such as linear or parabolic. In the last decade, there have been several attempts at eliminating the selection of variogram or covariogram models and their associated smoothness, see Shapiro and Botha (1991), Hall et al. (1994), Lele (1995), Cherry et al. (1996), Cherry (1997), Ecker and Gelfand (1997). These nonparametric methods are based on the isotropic spectral representation of positive definite functions derived by Yaglom (1957) from Bochner's (1955) theorem. In this paper, we present an optimal discretization of this spectral representation using Fourier–Bessel matrices. Our method provides smooth and positive definite nonparametric estimators in the continuum, whereas previous methods, such as the one first introduced by Shapiro and Botha (1991), typically suffer from spurious oscillations when no additional constraints on derivatives are imposed. Moreover, the orthogonality or near-orthogonality of our discretization improves the approximation power of the nonparametric variogram estimator.

Consider a spatial stochastic process $\{Z(\mathbf{x}) \mid \mathbf{x} \in D\}$, where the spatial domain D is a fixed subset of $\mathbb{R}^d$, $d \geqslant 1$. Assume that this process is second-order stationary with isotropic variogram $2\gamma(h)$ and covariogram $C(h)$, i.e. they depend only on the distance $h = \|\mathbf{h}\|$ and not on the direction of the lag vector $\mathbf{h}$. Let $Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n)$ be a realization of the spatial stochastic process. The classical estimator of the variogram proposed by Matheron (1962), based on the method of moments, is

$$2\hat{\gamma}(h) = \frac{1}{N_h} \sum_{N(h)} (Z(\mathbf{x}_i) - Z(\mathbf{x}_j))^2, \tag{1}$$

where the set $N(h) = \{(\mathbf{x}_i, \mathbf{x}_j) \colon \|\mathbf{x}_i - \mathbf{x}_j\| \in T(h)\}$ has cardinality N_h, and $T(h)$ is some tolerance neighborhood at distance h when the spatial locations are irregularly positioned in the domain D. Similarly, we estimate the covariogram by

$$\hat{C}(h) = \frac{1}{n} \sum_{N(h)} (Z(\mathbf{x}_i) - \bar{Z})(Z(\mathbf{x}_j) - \bar{Z}), \tag{2}$$

where $\bar{Z} = (1/n) \sum_{i=1}^{n} Z(\mathbf{x}_i)$ is the sample mean of the observations. Note that the estimator $n\hat{C}(h)/N_h$ is often found in the literature, but we prefer (2) because it provides a set of positive definite estimates in $\mathbb{R}^1$ (Yaglom, 1987). From now on, we focus on the estimation of the covariogram, since variogram estimation can easily be derived from the relation $\gamma(h) = C(0) - C(h)$.

It is well known that a valid covariogram has to be positive definite. Therefore, the key idea behind a nonparametric estimator for the covariogram is Bochner's (1955) theorem, which gives the spectral representation of positive definite functions. In particular, a covariogram $C(\mathbf{h})$ is positive definite if and only if it has the form:

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} \cos(\mathbf{u}^{\mathrm{T}}\mathbf{h}) F(\mathrm{d}\mathbf{u}), \tag{3}$$

where $F(\mathrm{d}\mathbf{u})$ is a positive bounded symmetric measure. If $C(\mathbf{h})$ is isotropic, Bochner's theorem can be written as (Yaglom, 1957)

$$C(h) = \int_0^\infty \Omega_d(th)F(\mathrm{d}t), \tag{4}$$

where

$$\Omega_d(x) = \left(\frac{2}{x}\right)^{(d-2)/2} \Gamma\left(\frac{d}{2}\right) J_{(d-2)/2}(x) \tag{5}$$

form a basis for functions in $\mathbb{R}^d$. Here $F(t)$ is any nondecreasing bounded function, $\Gamma(d/2)$ is the gamma function, and J_v is the Bessel function of the first kind of order v. Some familiar examples of Ω_d are $\Omega_1(x)=\cos(x)$, $\Omega_2(x)=J_0(x)$, and $\Omega_3(x)=\sin(x)/x$.

In order to obtain a nonparametric covariogram estimator from (4), we choose a vector $\mathbf{t}\in\mathbb{R}^m$, which represents the locations of the jump points, or nodes, in a discretization of $F(t)$. Specifically, let $F(t)=\sum_{j=1}^m p_j\Delta(t-t_j)$ where Δ is the step function

$$\Delta(t-t_j) = \begin{cases} 1 & \text{if } t \geqslant t_j, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

We know that $F(t)$ is nondecreasing and bounded, so the coefficients $\mathbf{p}=(p_1,\ldots,p_m)^{\mathrm{T}} \in\mathbb{R}^m$ are required to be nonnegative. Let $\mathbf{c}=(\hat{C}(h_1),\ldots,\hat{C}(h_l))^{\mathrm{T}}\in\mathbb{R}^l$ be the vector of l covariogram estimates obtained from (2). Given the nodes $\mathbf{t}=(t_1,\ldots,t_m)^{\mathrm{T}}$, $t_j\geqslant 0$, the nonparametric covariogram estimator has the form:

$$\tilde{C}(h) = \sum_{j=1}^m p_j\Omega_d(t_jh). \tag{7}$$

The nonnegative coefficients $\mathbf{p}$ are obtained by minimizing, under the constraint $\mathbf{p}\geqslant\mathbf{0}$, the objective function

$$S[\mathbf{p}] = (\mathbf{c}-M\mathbf{p})^{\mathrm{T}}W(\mathbf{c}-M\mathbf{p}), \tag{8}$$

where the $l\times m$ matrix M has entries $M_{ij}=\Omega_d(t_jh_i)$. Here W is an $l\times l$ weighting matrix that approximates the covariance structure of the covariogram estimates $\mathbf{c}$, see Genton (1999), and for the variogram, Genton (1998b, 2000), Genton et al. (2001), Gorsich et al. (2002). For simplicity, we set $W=I_l$, the identity matrix of size $l\times l$, in the sequel. Note that it is possible to include a nugget effect in (7) and use an iterative method for the minimization of (8), see Cherry et al. (1996, p. 448). It is now natural to ask how the nodes $\mathbf{t}$ should be chosen? Shapiro and Botha (1991), Cherry et al. (1996), Cherry (1997) all choose the nodes in an ad hoc way, e.g. about 200 equispaced nodes. However, in order to avoid arbitrary oscillations, the number of nodes m should be kept smaller than or equal to l, the number of covariogram estimates. In this paper, we show that the nodes should be chosen as zeros of Bessel functions in order to provide an orthogonal series in (7). With some freedom to select lags, near-orthogonal Fourier–Bessel matrices can be used. These choices ensure a smooth and positive definite nonparametric covariogram estimator in the continuum. We call the matrix M, with such choice of nodes, a Fourier–Bessel matrix and investigate its properties in the next section. In Section 3, we provide an example of nonparametric covariogram estimation from a simulated two-dimensional spatial process. An application to the characterization of variable stars in Astronomy is used for illustration in the last section.

2. Fourier–Bessel matrices

In this section, we investigate the properties of Fourier–Bessel matrices and their use for nonparametric covariogram estimation. We call an $l \times m$ matrix M a Fourier–Bessel matrix if its entries are defined by $M_{ij} = \Omega_d(t_j h_i)$, where $t_1, \ldots, t_m$ are zeros of Bessel functions, i.e. $J_{(d-2)/2}(t_j) = 0$, and $h_1, \ldots, h_l$ are a set of spatial lag distances. These matrices play a crucial role in nonparametric covariogram estimation because the choice of the nodes $\mathbf{t}$ significantly affects the solution of the minimization problem (8). If $\mathbf{t}$ is chosen in the right manner, the approximation error can be greatly reduced. In addition, if there is some freedom in the lag selection, near-orthogonal matrices can be used. In fact, for valid covariogram estimates (Gorsich and Genton, 2001) and the correct choice of nodes $\mathbf{t}$, a nonlinear algorithm is no longer needed to find the coefficients $\mathbf{p}$. It is important to notice that the series in Eq. (7) can be made orthogonal. Thus, the matrix M is nonsingular, and performance error is reduced (Björck, 1996). If the coefficients are positive for a given node selection, then nonnegative least squares becomes ordinary least squares, and covariogram estimates can be fitted to machine precision.

It is well known in physics that if the nodes t_j are chosen as the roots of Bessel functions, the series (7) becomes very similar to a Fourier–Bessel series (Bowman, 1958; Watson, 1958; Tolstov, 1962). Consider a Fourier–Bessel series of order v of $h^v C(h)$ given by

$$\sum_{j=1}^{\infty} \tilde{p}_j J_v(t_j h), \tag{9}$$

where the t_j are the roots of $J_v(t) = 0$, $0 < h < 1$, and $\tilde{p}_j = \Gamma(d/2) 2^v p_j / t_j^v$. Here, $v = (d-2)/2$ and d is the dimension of the spatial domain. This choice of nodes $\mathbf{t}$ makes the basis orthogonal, and the proof can be found in Bowman (1958). Because the series is orthogonal, multiplying by $hJ_v(t_k h)$ and integrating from 0 to 1 gives

$$\tilde{p}_k = \frac{2}{J_{v+1}^2(t_k)} \int_0^1 h^{v+1} C(h) J_v(t_k h)\, \mathrm{d}h \tag{10}$$

for the Fourier–Bessel expansion of the covariogram. For the variogram, orthogonality cannot be achieved and we cannot get such a simple form for p_k since $h^v = 2\sum_{j=1}^{\infty} J_v(t_j h)/(t_j J_{v+1}(t_j))$. For example, when $d = 1$, i.e. $v = -\frac{1}{2}$, the basis is $1 - \cos(t_k h)$, which is not orthogonal to $1 - \cos(t_j h)$, whereas $\cos(t_k h)$ and $\cos(t_j h)$ are orthogonal. Therefore, we focus on estimating p_j from the covariogram, and then derive the variogram from $\gamma(h) = C(0) - C(h)$.

We start by deriving the orthogonal matrices associated with the orthogonal Bessel functions, i.e. we want orthogonality for discrete h as well as for continuous h. Then, with orthogonality in the discrete case, we find a linear equation for $\mathbf{p}$ if the covariogram estimates are valid (Gorsich and Genton, 2001). If the nearly-orthogonal matrices can be used, then $\mathbf{p}$ can be solved for without inverting M. Although the function $F(t)$ has been discretized by a vector of nodes $\mathbf{t}$, the lags of $C(h)$ have not been. Given the continuous expansion

$$h^v C(h) = \sum_{j=1}^{m} \tilde{p}_j J_v(t_j h), \tag{11}$$

where $0 \leqslant h < 1$, we want the discrete expansion

$$h_i^v C(h_i) = \sum_{j=1}^{m} \tilde{p}_j J_v(t_j h_i) \tag{12}$$

to be orthogonal, but this is not automatic. Ensuring orthogonality for Fourier–Bessel matrices is not as simple as ensuring orthogonality for discrete cosine or Fourier series expansions. Consider the matrix B_v defined by $(B_v)_{ij} = J_v(t_j h_i / h_{\max})$. For the matrix B_v to be orthogonal (or nearly-orthogonal), given the nodes t_j, there must be a restriction on the choice of the lags h_i. This is because the Fourier–Bessel transform is its own inverse (Lemoine, 1994). Therefore, the transform matrix must be symmetric. As noted by Lemoine (1994), this leaves the following choice for an orthogonal (or nearly-orthogonal) matrix K_v of B_v

$$(K_v)_{ij} = \frac{2}{t_l} (N_v)_{ii} J_v(t_i t_j / t_l)(N_v)_{jj}, \tag{13}$$

where now $h_i = t_i/t_l$, $(N_v)_{ij} = \delta(i-j)/|J_{v+1}(t_i)|$, and $\delta(i-j)$ is the Kronecker delta, i.e. is one when $i = j$ and zero otherwise. The matrix K_v becomes a discrete sine transform for $v = \frac{1}{2}$ (but with different quadrature weights). Now $(B_v)_{ij} = J_v(t_i t_j / t_l)$, and K_v is a nearly-orthogonal matrix since $K_v^2 \approx I$, the identity matrix. Using an overlap matrix (discussed below), K_v can be made exactly orthogonal, and the difference of the overlap matrix from the identity matrix is as small as $2E-10$ for a 100×100 matrix.

Using the nearly-orthogonal matrix (without the diagonal N_v matrices) and rewriting Eq. (12) in matrix form gives the following expansion for the covariogram

$$\mathbf{c} = b_v D_v B_v \tilde{D}_v \mathbf{p}, \tag{14}$$

where $(\tilde{D}_v)_{ij} = \delta(i-j)/t_j^v$, $(D_v)_{ij} = \delta(i-j) t_l^v / t_i^v$, and $b_v = 2^v \Gamma(v+1)$. Thus, the Fourier–Bessel matrix M for the covariogram, $M_{ij} = \Omega_v(t_j h_i)$, is the same as $b_v D_v B_v \tilde{D}_v$ except that the lag values are chosen to be the node values divided by the lth node value. This normalizes the lag values to range between 0 and 1. Now using Eq. (14) there is a discrete expression for $\mathbf{p}$ without having to deal with approximating the integrals in (10) by exploiting near-orthogonality. The following proposition gives a linear expression for the coefficients $\mathbf{p}$.

Proposition 1. *Let the nodes t_j satisfy $J_v(t_j) = 0$ and the lags be related to the nodes by $h_i = t_i/t_l$, so $0 < h_i \leqslant 1$, and $v \geqslant -\frac{1}{2}$. Then the coefficients $\mathbf{p}$ of the expansion in (14) are given by*

$$\mathbf{p} = b_v^{-1} N_v^2 \tilde{D}_v^{-1} B_v N_v S_v^2 D_v^{-1} N_v \mathbf{c}, \tag{15}$$

where $S_v = (K_v^{\mathrm{T}} K_v)^{-1/2}$ is an overlap matrix.

Proof. Let $\mathbf{c} = b_v D_v B_v \tilde{D}_v \mathbf{p}$ and set $\mathbf{p} = N_v \mathbf{p}^*$. All the entries of N_v are positive. Then $N_v \mathbf{c} = b_v N_v D_v B_v \tilde{D}_v N_v \mathbf{p}^*$. Since diagonal matrices commute, we can interchange the N matrices so that there is a nearly symmetric orthogonal matrix in the middle. This gives $\mathbf{p}^* \approx b_v^{-1} N_v \tilde{D}_v^{-1} N_v B_v N_v D_v^{-1} N_v \mathbf{c}$. Now using N_v again, we arrive at a very close approximation to $\mathbf{p}$ without taking an inverse of a matrix

$$\mathbf{p} \approx b_v^{-1} N_v^2 \tilde{D}_v^{-1} B_v D_v^{-1} N_v^2 \mathbf{c}. \tag{16}$$

For l as large as a few hundred, this approximation is within machine precision. Otherwise equality can be reached by using an overlap matrix (Lemoine, 1994) that is nearly the identity. The overlap matrix is $S_v = (K_v^{\mathrm{T}} K_v)^{-1/2}$, so that $S_v K_v$ is truly orthogonal and

$$\mathbf{p} = b_v^{-1} N_v^2 \tilde{D}_v^{-1} B_v N_v S_v^2 D_v^{-1} N_v \mathbf{c}. \tag{17}$$

In most cases, Eq. (16) can be used and no inverse is needed because the noise in the covariogram estimates is much greater than the entries of the overlap matrix. This improves the performance of the optimization problem (8), should nonnegative least squares be needed. This completes the proof. □

Proposition 1 gives a simple linear equation for the coefficients $\mathbf{p}$. Now the question is whether these coefficients will be positive? If they are, then nonnegative least squares is no longer needed for (8). This depends essentially on the covariogram estimates $\mathbf{c}$. Further discussions on this topic can be found in Gorsich (2000), and Gorsich and Genton (2001). Note that in Proposition 1, we have avoided evaluating the integral by using a discretization scheme that nearly preserves orthogonality.

Because there is a simple form for $\mathbf{p}$ using the covariogram, there also must be a simple form for the variogram. Moreover, the values of the coefficients provided by the two linear equations must coincide. Let $\boldsymbol{\gamma} = (\hat{\gamma}(h_1), \ldots, \hat{\gamma}(h_l))^{\mathrm{T}} \in \mathbb{R}^l$ be the vector of l variogram estimates obtained from (1). We have the following proposition.

Proposition 2. *Let the nodes t_j satisfy $J_v(t_j) = 0$ and the lags be related to the nodes by $h_i = t_i/t_l$, so $0 < h_i \leqslant 1$, and $v \geqslant -\frac{1}{2}$. Then the coefficients $\mathbf{p}$ of the variogram expansion are given by*

$$\mathbf{p} = -N_v(A + A\mathbf{n}_v \mathbf{n}_v^{\mathrm{T}} A(1 - \mathbf{n}_v^{\mathrm{T}} A \mathbf{n}_v)^{-1}) N_v \boldsymbol{\gamma}, \tag{18}$$

where $A = (b_v D_v N_v B_v N_v \tilde{D}_v)^{-1}$ and the vector $\mathbf{n}_v = \mathrm{diag}(N_v)$.

Proof. The values of the coefficients coincide for both the variogram and covariogram but the equation for the variogram $\boldsymbol{\gamma}$ takes the form

$$\boldsymbol{\gamma} = (\mathbf{1}\mathbf{1}^{\mathrm{T}} - b_v D_v B_v \tilde{D}_v)\mathbf{p}. \tag{19}$$

As before, setting $\mathbf{p} = N_v \mathbf{p}^*$ and $\boldsymbol{\gamma} = N_v^{-1} \boldsymbol{\gamma}^*$ and using the relation

$$(A - \mathbf{1}\mathbf{1}^{\mathrm{T}})^{-1} = A^{-1} + A^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}} A^{-1}(1 - \mathbf{1}^{\mathrm{T}} A^{-1} \mathbf{1})^{-1} \tag{20}$$

we get the desired result. Again, the inverse of $N_v B_v N_v$ can be found approximately by taking the transpose, or the overlap matrix can be used as in Proposition 1 for exact equality. □

The Fourier–Bessel matrices in this section exploit near-orthogonality (or orthogonality for $d=1$ or 3, see Strang, 1999) and find approximate expressions for $\mathbf{p}$ without the need to take an inverse. In some cases, $\mathbf{p}$ is found to be positive and therefore the nonparametric covariogram estimator is positive definite. Otherwise, nonnegative least squares (8) is used with improved performance (no inverse of a matrix is needed) and with greater approximation power. As a result, fewer nodes are necessary and there are no spurious oscillations. An example of this fact is discussed in the next section.

3. A simulated example

In the previous section, we discussed the nearly orthogonal matrices that are found when the nodes are chosen as zeros of Bessel functions. These matrices improve the performance of the nonnegative least squares algorithm, and provide a nonparametric estimator of variograms and covariograms that uses a small number of nodes under reasonable sampling conditions (see Gorsich, 2000). With a small number of nodes, the covariogram estimates can be fit well at a finite number of estimates, and in the continuum, i.e. the nonparametric covariogram estimator is smooth between the covariogram estimates. With other node selections, this is not true as discussed in Gorsich and Genton (2001).

In this section, we simulate for illustration a Gaussian spatial process Z on a two dimensional grid of size 20×20. More simulations can be found in Gorsich (2000) and Gorsich and Genton (2001). The process has mean zero and covariogram defined by a Bessel function of order zero. We window the covariogram with $(1-h)^2$ so that it is valid in $\mathbb{R}^2$ and is zero for $h>1$. The resulting covariogram is plotted in Fig. 1. Covariogram estimates are computed with the estimator (2), thus guaranteeing positive definiteness in $\mathbb{R}^1$. This positive definiteness does not extend to $\mathbb{R}^d$, which means that the estimated points are not positive definite in higher dimensions, and nonnegative least squares must be used. The Fourier–Bessel matrix M, as well as a matrix M defined by equally spaced nodes, are used to find the coefficients $\mathbf{p}$. As done in earlier papers (Shapiro and Botha, 1991; Cherry et al., 1996; Cherry, 1997), the number of nodes is increased beyond the number of covariogram estimates available so that the estimates can be fitted closely. Thus, using $\tilde{C}(h)$ defined by (7), a nonparametric covariogram estimator can be found that gives values for the covariogram in the continuum. For the equally spaced nodes, we have to go up to 80 nodes to get a good fit to the covariogram estimates. With the nodes chosen as zeros of Bessel functions, i.e. with Fourier–Bessel matrices, we can use 80 nodes, but only twenty are really needed here because the values of p_j for $j>20$ are effectively zero. This type of result is based on the orthogonality and the improved approximation power of the series. Fig. 1 shows the two nonparametric estimators of the covariogram, as well as the true underlying covariogram (dashed line). The stared points are the covariogram estimates given by Eq. (2). We view only the first nine of the 20 for clarity. The two curves represent the

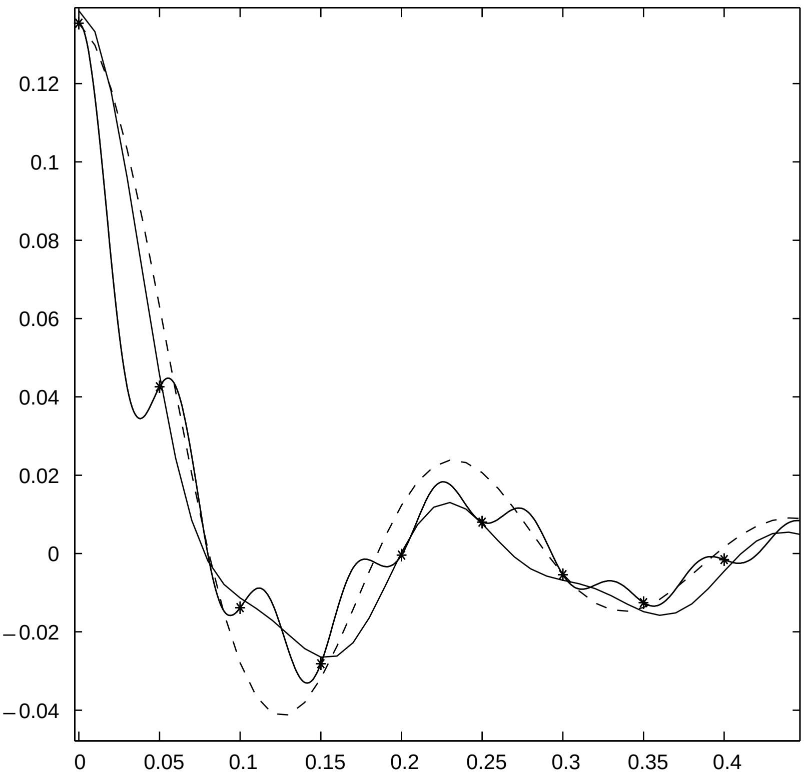

Fig. 1. Two nonparametric covariogram estimators fitting covariogram estimates (stars) of a two-dimensional spatial process. The dashed line represents the true underlying covariogram. The fit with equally spaced nodes needs 80 nodes and has oscillations that are not represented in the true underlying covariogram. The smooth fit uses nodes as zeros of Bessel functions, and the Fourier–Bessel matrix M as discussed in the paper. In this case, the last 60 values of $\mathbf{p}$ are effectively zero, so this estimator needs only twenty nodes, and therefore provides a much smoother fit in the continuum. This fit is closer to the underlying covariogram.

two nonparametric covariogram estimators, one with nodes as zeros of Bessel functions, and the other as equally spaced nodes. The estimator that uses equally spaced nodes requires more nodes to fit the covariogram estimates effectively, and therefore has spurious oscillations appearing that have nothing to do with the true underlying covariogram, a windowed Bessel function of order zero. With the nodes properly chosen, fewer are needed, and a smoother fit can be reached. The use of Fourier–Bessel matrices is thus highly recommended in order to avoid spurious oscillations in the continuum. If the nodes were chosen in another manner, or randomly, the fits can become much worse, and to fit covariogram estimates, prior authors chose up to 200 nodes (see e.g. Cherry et al., 1996). Of course, choosing a large number of nodes allows one to fit the covariogram estimates very closely, but causes even larger spurious oscillations to occur that have nothing to do with the underlying covariogram.

4. An example from Astronomy

In this section, we present an application of nonparametric variogram estimation to photometric data arising from Astronomy. In 1989, the European Space Agency launched a satellite, named Hipparcos (HIgh Precision PARallax COllecting Satellite), with the goal of collecting a huge number of measurements about stars. Actually, the Hipparcos satellite provided about 13,000,000 measurements of fluxes for 118,204 stars during a time interval of 3.3 years. A systematic search for periodic stars was carried out (Eyer, 1998), and a particular goal was also to characterize stars with pseudo-periodic light curves, i.e. approximately periodic signal, such as super giant, spotted, and semi-regular variables stars.

Eyer and Genton (1999) proposed to characterize variable stars by robust wave variograms. Indeed, the variograms of time series with periodic or pseudo-periodic signal exhibit negative correlations caused by the periodicity of the series, and are named wave or hole effect variograms (Chilès and Delfiner, 1999, p. 92). They show oscillations of decreasing amplitude around a horizontal plateau called the sill, which is the total variance of the observations. The first minimum of the wave variogram after the origin is considered as a measure of the period or pseudo-period of the time series, and its computation can be used to characterize variable stars. Unfortunately, only about 110 irregularly sampled observations are available for each star, among which a nonnegligible fraction of outliers can be found. As a consequence, Matheron's (1962) variogram estimator (1) is unreliable, nor is it enough to make simple modifications to (1) such as the ones proposed by Cressie and Hawkins (1980) in order to achieve robustness. The use of a highly robust variogram estimator (Genton, 1998a) has been advocated by Eyer and Genton (1999), who demonstrated the significant advantages of working with this estimator for determining pseudo-periodicities.

To date, this robust wave variogram methodology was applied only qualitatively. In order to automate the estimation of the first minimum, one could fit a valid parametric variogram model to the variogram estimates, for instance by a least square algorithm, and then determine the first minimum analytically. Because the time sampling obtained from the Hipparcos mission is very irregular, the variogram estimates have large variabilities and the previous approach typically fails. Moreover, the pseudo-periodicities characterizing variables stars tend to yield variogram oscillations of different shapes, thus preventing the use of a single parametric model of variogram. For these reasons, we herein propose to utilize a nonparametric variogram estimator, based on Fourier–Bessel matrices as described in Section 2, for the estimation of the first minimum. Fig. 2 depicts highly robust variogram estimates for two typical stars from the Hipparcos mission, HIP 052507 and HIP 023743, respectively, along with nonparametric variogram estimates in the continuum. The number l of highly robust variogram estimates are, respectively, 28 and 31 for the two stars, and the number of nodes is $m = l$. The Fourier–Bessel expansions are performed in dimension $d = 1$, although higher dimensions could be used if smoother fits are desired. The estimated pseudo-periods for the two variable stars are, respectively, 91.2 and 68.4 days. We are therefore able to characterize the variable stars from the Hipparcos mission by an automatic estimation of their pseudo-periodicities. If we would not use Fourier–

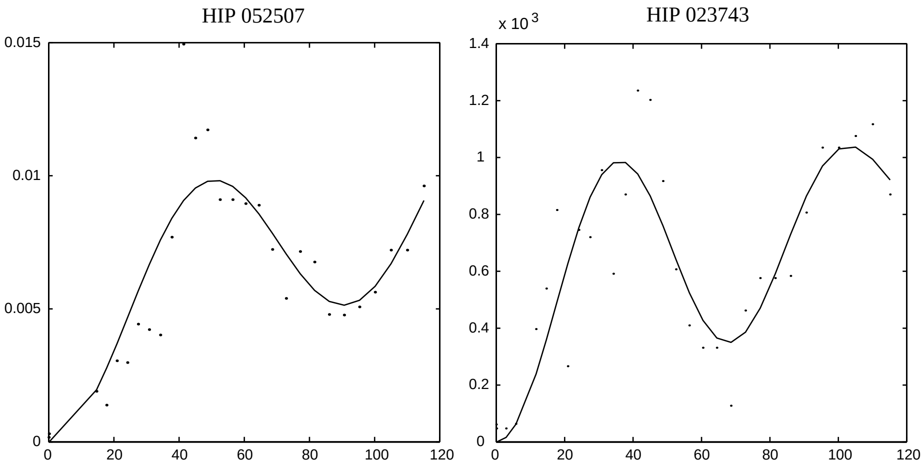

Fig. 2. Highly robust variogram estimates for two typical stars from the Hipparcos mission, HIP 052507 and HIP 023743, respectively, along with nonparametric variogram estimates in the continuum.

Bessel matrices, spurious oscillations might mask the first minimum of the wave variogram.

5. Conclusion

In this paper, we have shown that Fourier–Bessel matrices play an important role in the nonparametric estimation of variograms and covariograms. They provide the correct discretization of the spectral representation of positive definite functions, in the sense that the resulting covariogram estimators are guaranteed to be smooth and positive definite in the continuum. The issue of spurious oscillations was first raised by Hall et al. (1994), and our approach answers this point by advocating the use of Fourier–Bessel matrices. The main advantage with a nonparametric approach is that the choice of a particular model and the associated smoothness of the spatial process are avoided.

Acknowledgements

The authors thank the Editor and three anonymous reviewers for their comments that improved the manuscript, as well as Laurent Eyer, from the Geneva Observatory in Switzerland, for making data from the Hipparcos mission available. The first author's work was supported by the TACOM Research, Development and Engineering Center under the auspices of the U.S. Army Research Office Scientific Services Program administered by Battelle (Delivery Order 634, Contract No. DAAH04-96-C-0086, TCN 00-131).

References

Björck, A., 1996. Numerical Methods for Least Squares Problems. SIAM, Philadelphia.

Bochner, S., 1955. Harmonic Analysis and the Theory of Probability. University of California Press, Los Angeles.

Bowman, F., 1958. Introduction to Bessel Functions. Dover, New York.

Cherry, S., 1997. Non-parametric estimation of the sill in geostatistics. Environmetrics 8, 13–27.

Cherry, S., Banfield, J., Quimby, W.F., 1996. An evaluation of a non-parametric method of estimating semi-variograms of isotropic spatial processes. J. Appl. Statist. 23, 435–449.

Chilès, J.-P., Delfiner, P., 1999. Geostatistics: Modelling Spatial Uncertainty. Wiley, New York.

Cressie, N., Hawkins, D.M., 1980. Robust estimation of the variogram: I. Math. Geol. 12 115–125.

Ecker, M.D., Gelfand, A.E., 1997. Bayesian variogram modeling for an isotropic spatial process. J. Agri. Biol. Environ. Statist. 2, 347–368.

Eyer, L., 1998. Les étoiles variables de la mission Hipparcos. Ph.D. Thesis, Geneva Observatory.

Eyer, L., Genton, M.G., 1999. Characterization of variable stars by robust wave variograms: an application to Hipparcos mission. Astron and Astrophys. Ser. (Suppl) 136, 421–428.

Genton, M.G., 1998a. Highly robust variogram estimation. Math. Geol. 30, 213–221.

Genton, M.G., 1998b. Variogram fitting by generalized least squares using an explicit formula for the covariance structure. Math. Geol. 30, 323–345.

Genton, M.G., 1999. The correlation structure of the sample autocovariance function for a particular class of time series with elliptically contoured distribution. Statist. Probab. Lett. 41, 131–137.

Genton, M.G., 2000. The correlation structure of Matheron's classical variogram estimator under elliptically contoured distributions. Math. Geol. 32, 127–137.

Genton, M.G., He, L., Liu, X., 2001. Moments of skew-normal random vectors and their quadratic forms. Statist. Probab. Lett. 51, 319–325.

Gorsich, D.J., 2000. Nonparametric estimation of dependencies for spatial interpolation. Ph.D. Thesis, MIT.

Gorsich, D.J., Genton, M.G., 2000. Variogram model selection via nonparametric derivative estimation. Math. Geol. 32, 249–270.

Gorsich, D.J., Genton, M.G., 2001. On the discretization of nonparametric covariogram estimators. Submitted for publication.

Gorsich, D.J., Genton, M.G., Strang, G., 2002. Eigenstructures of spatial design matrices. J. Multivariate Anal. 80, 138–165.

Hall, P., Fisher, N., Hoffmann, B., 1994. On the nonparametric estimation of covariance functions. Ann. Statist. 22, 2115–2134.

Lele, S., 1995. Inner product matrices, kriging, and nonparametric estimation of variogram. Math. Geol. 27, 673–692.

Lemoine, D., 1994. The discrete Bessel transform algorithm. J. Chem. Phys. 101, 3936–3944.

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

Shapiro, A., Botha, J.D., 1991. Variogram fitting with a general class of conditionally nonnegative definite functions. Comput. Statist. Data Anal. 11, 87–96.

Strang, G., 1999. The discrete cosine transform. SIAM Rev. 41, 135–147.

Tolstov, G.P., 1962. Fourier Series. Prentice-Hall, Englewood Cliffs, NJ.

Watson, G.N., 1958. Treatise on the Theory of Bessel Functions. The University Press, Cambridge.

Yaglom, A.M., 1957. Some classes of random fields in *n*-dimensional space, related to stationary random processes. Theory Probab. Appl. 2, 273–320.

Yaglom, A.M., 1987. Correlation Theory of Stationary and Related Random Functions, Vol. I & II. Springer Series in Statistics, New York.