A Note on Nonparametric Variogram Fitting

Chunfeng Huang[a], and Haimeng Zhang[b]

July 20, 2009

Technical Report 09-02

Department of Statistics

Indiana University

Bloomington, IN 47405

[a]Departments of Statistics, Indiana University. huang48@indiana.edu

[b]Department of Mathematics and Statistics, Mississippi State University. hzhang@math.msstate.edu

Abstract

A simplified version of the nonparametric variogram fitting method proposed by Shapiro and Botha (1991) is developed in this note. Instead of the Bessel functions of various orders, the same cosine function is used regardless of the dimensions of the space where the data was collected. A simulation study is presented to illustrate this approach.

Key words: Conditional Negative Definite, Isotropy, Quadratic Programming, Spectral Density

1. Introduction

Variogram estimation has received much attention in geostatistics. The variogram models the spatial dependency, and it plays an essential role in spatial prediction/kriging (e.g. Cressie, 1993 and Chilès and Delfiner, 1999). One crucial property of the variogram is the conditionally negative definiteness. Armstrong and Jabin (1981) present examples to show that negative variance can occur when the variogram is not valid. In practice, parametric models are developed to ensure this property. For nonparametric method proposed by Shapiro and Botha (1991), Bochner's theorem (Bochner, 1955) is utilized and a conditionally negative definite estimator is obtained through quadratic programming. In their method, when data are collected from the space of different dimensions, the Bessel function of according order is acquired. This approach has been extended in various directions (Barry and Ver Hoef, 1996, Cherry et al. 1996, Lele, 1995, Ecker and Gelfond, 1997, Genton and Gorsich, 2002, and Gorsich and Genton, 2004).

In this note, we present a simplified version of Shapiro and Botha (1991), where the Bessel function is replaced by a cosine function regardless of dimensions. See details in Section 2. In Section 3, a numerical study is provided to illustrate the approach.

2. Nonparametric Variogram Fitting

Let $\{Z(t) : t \in \mathbb{R}^d\}$ be a spatial random process with a constant mean. If the process is assumed to be intrinsically stationary and isotropic, the variogram is defined as

$$2\gamma(h) = \text{var}[Z(s) - Z(t)], \qquad \text{for } s, t \in \mathbb{R}^d,$$

where $h = ||s - t||$. For a given set of data observations $\{Z(t_i) : i = 1, 2, \cdots, n\}$, the method of moment variogram estimator is commonly used and it is given by (Matheron, 1963)

$$\hat{\gamma}(h) = \frac{1}{|N(h)|} \sum_{N(h)} (Z(t_i) - Z(t_j))^2,$$

where $h = ||t_i - t_j||$ is a possible spatial lag, $N(h) = \{(t_i, t_j), ||t_i - t_j|| = h\}$ with $|N(h)|$ representing the cardinality of $N(h)$. When data are not observed on grids, a tolerance region method can be used (Journel and Huijbregts, 1978). More discussion on variogram estimation can be found in Cressie (1993). Throughout this note, we assume that $\{\hat{\gamma}(h_1), \ldots, \hat{\gamma}(h_n)\}$ are the estimated semi-variogram values at lags $h_1, h_2, \cdots, h_n$. One would now be interested in estimating the variogram function $\gamma(\cdot)$.

The variogram is conditionally negative definite and has the spectral representation (Yaglom, 1987, Section 25)

$$2\gamma(h) = \int_0^\infty (1 - \Omega_d(uh))dF(u), \quad (1)$$

where

$$\Omega_d(x) = (2/x)^{(d-2)/2}\Gamma(d/2)J_{(d-2)/2}(x), \quad (2)$$

$J_\nu(\cdot)$ is the Bessel function of the first kind of order ν, and $F(\cdot)$ is a nondecreasing function on the half-line $(0, \infty)$, satisfying $\int_0^\infty \frac{u^2 dF(u)}{1+u^2} < \infty$. From this representation, a nonparametric estimation of variogram is formulated in Shapiro and Botha (1991), where a vector $\mathbf{u} = (u_1, u_2, \cdots, u_m)^T$ representing the locations of the jump points is used in a discretization of $F(u)$ (Genton and Gorsich, 2002). More specifically, let

$$F(u) = \sum_{j=1}^{m} y_j \Delta(u - u_j),$$

where Δ is the step function

$$\Delta(u - u_j) = \begin{cases} 1 & \text{if } u \geq u_j, \\ 0 & \text{otherwise.} \end{cases}$$

From (1), Shapiro and Botha (1991) then look for an $(m+1) \times 1$ vector $\mathbf{y} = (y_1, \cdots, y_m, c_0)^T$ that minimizes the quadratic function

$$Q(\mathbf{y}) = \sum_{i=1}^{n} w_i \left(\hat{\gamma}(h_i) - c_0 + \sum_{j=1}^{m} \Omega_d(h_i u_j) y_j \right)^2 ,$$

subject to

$$y_j \geq 0, j = 1, \cdots, m, \quad \text{and} \quad c_0 - \sum_{j=1}^{m} y_j \geq 0.$$

Here the linear constraints of $y_j \geq 0$ ensure the resulting variogram estimate is conditionally negative definite and w_i's are the weights to be specified (see Remark 5). Hence the variogram function is estimated by

$$\tilde{\gamma}(h) = \tilde{c}_0 - \sum_{j=1}^{m} \Omega_d(h u_j) \tilde{y}_j,$$

where $\tilde{\mathbf{y}} = (\tilde{y}_1, \cdots, \tilde{y}_m, \tilde{c}_0)^T$ is the solution of above quadratic programming.

Clearly, for each dimension d, $\Omega_d(\cdot)$ takes a different form. For example, $\Omega_1(x) = \cos(x), \Omega_2(x) = J_0(x)$, and $\Omega_3(x) = \frac{\sin(x)}{x}$, corresponding to $d = 1, 2, 3$, respectively. Hence, a corresponding quadratic programming procedure for each dimension d has to be developed. In geostatistics, the most interesting case is $d = 2$, which acquires the Bessel function of order 0. Note that, instead of the equation (1), a different spectral representation of an isotropic variogram in $\mathbb{R}^d$ can be found (Yaglom, 1987, Section 25.3):

$$2\gamma(h) = \int_0^\infty (1 - \cos(uh)) dF_1(u), \tag{3}$$

where $F_1(\cdot)$ is a nondecreasing function on the half-line $(0, \infty)$, satisfying $\int_0^\infty \frac{u^2 dF_1(u)}{1+u^2} < \infty$. Here the Bessel function Ω_d in (2) is replaced by the same cosine function regardless of the dimension d. Hence, the Shapiro and Botha (1991)'s approach can be simplified. In particular, one is to find an $(m+1) \times 1$ vector $\mathbf{y} = (y_1, \cdots, y_m, c_0)^T$ to minimize

$$Q(\mathbf{y}) = \sum_{i=1}^{n} w_i \left(\hat{\gamma}(h_i) - c_0 + \sum_{j=1}^{m} \cos(u_j h_i) y_j \right)^2 , \tag{4}$$

subject to

$$y_j \geq 0, j = 1, \ldots, m, \quad \text{and} \quad c_0 - \sum_{j=1}^{m} y_j \geq 0.$$

The case of $c_0 - \sum_{j=1}^{m} y_j > 0$ will respond to the nugget effects (Fernandez-Casal *et al.*, 2003). The nonparametric variogram estimate is then given by,

$$\tilde{\gamma}(h) = \tilde{c}_0 - \sum_{j=1}^{m} \cos(hu_j)\tilde{y}_j,$$

where $\tilde{\mathbf{y}} = (\tilde{y}_1, \cdots, \tilde{y}_m, \tilde{c}_0)^T$ is the solution of quadratic programming.

Remarks:

1. Our approach can be used for nonparametric variogram estimation regardless of the dimension d of the space where the data was collected. The method by Shaprio and Botha (1991) is based on the Bessel function (2). In this note, the same cosine function is used instead, which simplifies the programming in the estimation procedure. For example, when $d = 2$, the Bessel function of order zero $J_0(\cdot)$ is replaced by $\cos(\cdot)$.

2. If we further assume both $F(\cdot)$ and $F_1(\cdot)$ in (1) and (3) are absolutely continuous with spectral densities $f(\cdot)$ and $f_1(\cdot)$, we have

   $$2\gamma(h) = \int_0^\infty (1 - \Omega_d(uh))f(u)du = \int_0^\infty (1 - \cos(uh))f_1(u)du.$$

   The relationship between $f(\cdot)$ and $f_1(\cdot)$ is expressed through the following (Yaglom, 1987, equation 4.120)

   $$f_1(u_1) = \frac{2\pi^{(d-1)/2}}{\Gamma((d-1)/2)} \int_{u_1}^\infty f(u)(u^2 - u_1^2)^{(d-3)/2} u du. \tag{5}$$

   For example, consider the following variogram in $\mathbb{R}^2$,

   $$2\gamma(h) = \frac{1}{a^2} - \frac{a}{(a^2 + h^2)^{3/2}}, \qquad h > 0,$$

where $a > 0$. It can be verified to have the following spectral representation

$$2\gamma(h) = \int_0^\infty (1 - J_0(uh))ue^{-au} du,$$

with the spectral density $f(u)$ given by $f(u) = ue^{-au}, u > 0$. Meanwhile, the variogram can also be represented by

$$2\gamma(h) = \int_0^\infty (1 - \cos(uh))\frac{1}{\pi}uK_1(uh) du,$$

with $f_1(u) = \frac{1}{\pi}uK_1(u)$. Here $K_1(u)$ is the modified Bessel functions of the second kind of order 1. Direct calculation shows that (5) holds for $f(u)$ and $f_1(u)$. Note that the function $f(u) = ue^{-au}$ can be viewed as a two-dimensional isotropic spectral density, while $f_1(u) = \frac{1}{\pi}uK_1(u)$ is the one-dimensional spectral density.

3. A stationary process is always intrinsically stationary. Therefore, its covariance function $C(h)$ also has both spectral densities $f(u)$ and $f_1(u)$

$$C(h) = \int_0^\infty \Omega_d(uh) f(u) du = \int_0^\infty \cos(uh) f_1(u) du.$$

The variogram fitting procedure based on the cosine function discussed above can be applied to obtain a nonparametric covariance fitting. In particular, letting $\hat{C}(h_1), \cdots, \hat{C}(h_n)$ be experimental covariance estimates, one is to find the vector $\mathbf{y} = (y_1, y_2, \cdots, y_m)^T$ through minimizing

$$Q(\mathbf{y}) = \sum_i w_i \left(\hat{C}(h_i) - \sum_{j=1}^m \cos(u_j h_i) y_j \right)^2,$$

subject to $y_j \geq 0, j = 1, 2, \cdots, m$. Hence, the covariance estimate is given by

$$\tilde{C}(h) = \sum_{j=1}^m \cos(u_j h) \tilde{y}_j,$$

with $(\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_n)^T$ is the solution of the quadratic programming. Genton and Gorsich (2002) and Gorsich and Genton (2004) provide an orthogonal discretization of nonparametric covariogram estimator through the use of Fourier-Bessel matrices where the roots of Bessel functions are used, which is not a simple task for applications. This complication of ensuring orthogonality for Fouries-Bessel matrices in their approach can be reduced if the discrete cosine series expansion is used.

4. It is noticed that the resulting estimate $\tilde{\gamma}(\cdot)$ may change rapidly (Shapiro and Botha, 1991, Gorsich and Genton, 2000). One may impose smoothness conditions, for example, a bound on the first derivative. These conditions can be formulated in the quadratic programming procedure as proposed in Shapiro and Botha (1991). Such approach can also be adapted in our estimation method as well.

5. Shapiro and Botha (1991) considered the equal weights $w_i = 1$ and the weights $w_i = |N(h_i)|/\gamma^2(h_i)$ in Cressie (1985). They noticed that the result differs little from each other. This is also observed in our simulation study when equation (4) is used.

3. Numerical Examples

In this section, a small simulation study is conducted. We consider two commonly used variogram models.

The Spherical Model with semivariogram equation

$$\gamma(h;\theta) = \begin{cases} 0, & h = 0 \\ n_0 + c\{3/2(h/a) - (1/2)(h/a)^3\}, & 0 < h \leq a, \\ n_0 + c, & h > a, \end{cases}$$

where $\theta = (a, n_0, c)$ with $a \geq 0, n_0 \geq 0, c \geq 0$.

The Exponential Model with semivariogram equation

$$\gamma(h;\theta) = \begin{cases} 0, & h = 0 \\ n_0 + c(1 - \exp(-h/a)), & h > 0 \end{cases}$$

where $\theta = (a, n_0, c)$ with $a \geq 0, n_0 \geq 0, c \geq 0$.

In the simulation, a 20×20 grid is used for data locations and 100 replications are conducted for each setting. For each simulated data, the nonparametric fittings through both cosine and the Bessel functions, along with the parametric fitting, are performed. The corresponding mean integrated squared errors (MISE)

$$\text{MISE} = \int (\gamma(h) - \tilde{\gamma}(h))^2 dh.$$

are computed and the results are presented in Table 1. It is clear from Table 1 that both cosine and Bessel approaches only differ little in all our simulations. Notice that, as expected, the MISE value (MISE(p)) based on the parametric fitting is the smallest. It is also worthy noting that the nonparametric methods perform well compared to the parametric fitting. Both capture the feature of true variogram function. In the simulation, we choose the nodes using $u_j = \delta \cdot j$ with δ predetermined (Shaprio and Botha, 1991, Cherry et al., 1996). It appears that the MISEs are not much affected by various choice of nodes.

Models	MISE(p)	MISE(cos)	MISE(J_0)
Spherical ($a = 5, n_0 = 0.32, c = 1$)	0.382(0.060)	0.385(0.061)	0.385(0.061)
Spherical ($a = 3, n_0 = 0.20, c = 1$)	0.205(0.040)	0.207(0.039)	0.207(0.039)
Exponential ($a = 0.5, n_0 = 0.20, c = 1$)	0.142(0.020)	0.151(0.020)	0.151(0.020)
Exponential ($a = 1, n_0 = 0.30, c = 1$)	0.118(0.020)	0.121(0.020)	0.121(0.020)

Table 1: MISE comparison for parametric and nonparametric methods

Figure 1 presents a simulation result from two nonparametric fitting (cosine and Bessel functions) and parametric fitting (three dashed lines), along with the true spherical variogram ($a = 5, n_0 = 0.32, c = 1$, solid line). They are very close to each other and fit well with the experimental variogram (dots). For the parametric fitting, the parameter estimates are given by $\tilde{a} = 5.73, \tilde{n}_0 = 0.31$, and $\tilde{c} = 1.08$, respectively. For nonparametric fitting, the nugget effect estimates are $\tilde{n}_0 = 0.32$, and 0.31 for cosine and Bessel functions fitting, respectively.

Acknowledgements: The work is motivated by the discussion with Professor Tailen Hsing at University of Michigan. Chunfeng Huang's research is supported by the National Science Foundation DMS 0808993.

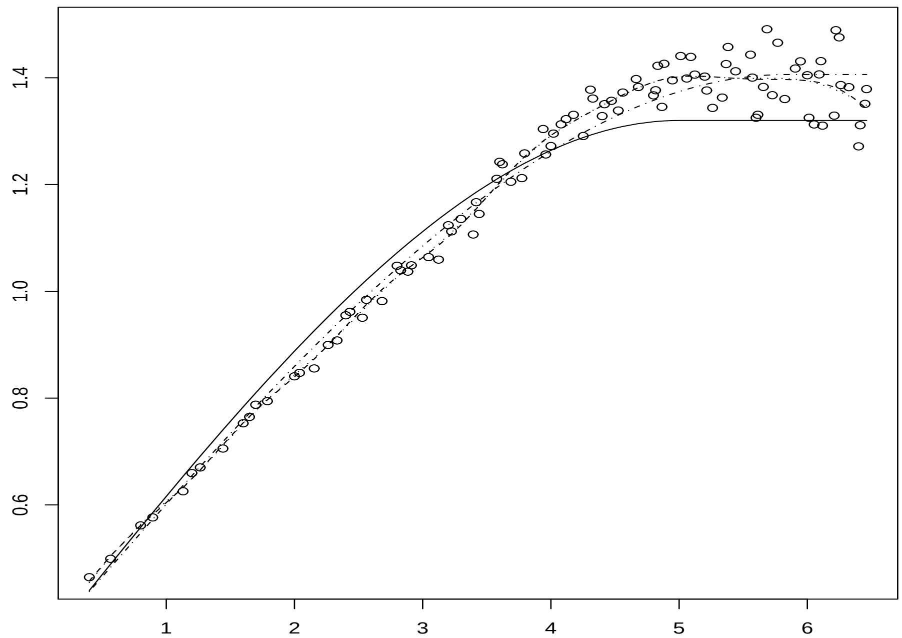

Figure 1: Two nonparametric fittings - cosine function (short dashed lines) and the Bessel function (mid-dashed lines); parametric fitting (long-dashed line); true variogram (solid line); experimental variogram (dots)

References

1. Armstrong, M., Jabin, R., 1981. Variogram models must be positive-definite. Journal of the International Association of Mathematical Geology, 13, 455 - 459.

2. Barry, R.P., Ver Hoef, J.M., 1996. Blackbox Kriging: spatial prediction without specifying variogram models. Journal of Agricultural, Biological and Environmental Statistics, 1, 297 - 322.

3. Bochner, S., 1955. Harmonic analysis and the theory of probability. University of California Press, Berkeley and Los Angeles, 176.

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

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

6. Cressis, N., 1985. Fitting variogram models by weighted least squares. Mathematical Geology, 17, 63-586.

7. Cressis, N., 1993. Statistics for spatial data. Revised edition, Wiley, New York.

8. Ecker, M.D., Gelfand, A.E., 1997. Bayesian variogram modeling for an isotropic spatial process. Journal of Agricultural, Biological and Environmental Statistics, 2, 347 - 369.

9. Fernández-Casal, R., González-Manteiga, W., Febrero-Bande, M., 2003. Flexible spatio-temporal stationary variogram models. Statis. Comput. 13, 127 - 136.

10. Genton, M.G., Gorsich, D.J., 2002. Nonparametric variogram and covariogram estimation with Fourier-Bessel matrices. Comput. Statist. Data Anal., 41, 47 - 57.

11. Gorsich, D.J., Genton, M.G., 2000. Variogram model selection via nonparametric derivative estimation. Mathematical geology, 32, 249 - 270.

12. Gorsich, D.J., Genton, M.G., 2004. On the discretization of nonparametric isotropic covariogram estimators. Statist Comput., 14, 99 - 108

13. Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London.

14. Lele, S., 1995. Inner product matrices, kringing, and nonparametric estimation of variogram. Math. Geol. 27, 673 - 692.

15. Matheron, G., 1963. Principles of geostatistics. Econometrics geology, 58, 1246 - 1266.

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

17. Yaglom, A.M., 1987. Correlation Theory of Stationary and Related Random Functions. Springer-Verlag, New York.