Journal of the Korean Statistical Society (2001), **30**: *4, pp 573–583*

On Asymptotic Property of Matheron's Spatial Variogram Estimators

Yoon-Dong Lee[1] and Eun-Kyung Lee[2]

ABSTRACT

A condition in which the covariances of Matheron's variogram estimators are expressed in a simple form is reviewed. An asymptotic property of the covariances of the variogram estimators is examined, and a sufficient condition that guaranties the finiteness of the asymptotic variance of the normalized variogram estimators is provided.

Keywords: Spatial Variogram; Variogram estimators; Asymptotic covariance

1. INTRODUCTION

For an intrinsically stationary random field, $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in I\!R^d\}$, the *variogram* $2\gamma(\boldsymbol{h})$ which is defined as $2\gamma(\boldsymbol{h}) = Var(Z(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s}))$, for all $\boldsymbol{h} \in I\!R^d$ is a measure of spatial dependence of the process $Z(\cdot)$. Gaussian model, (piecewise) linear model, etc. are commonly used model variograms. Linear variogam model is a special case of the power variogram model, $2\gamma(\boldsymbol{h}) = \|\boldsymbol{h}\|^\theta$ with $\theta = 1$. To be a valid variogram model, the well-known two conditions, *conditional negative definite condition* and *order-2 stationary condition* should be satisfied. The conditional negative definite condition is directly derived from nonnegative definiteness of Fourier transform of spectral density (*ref.* Cressie, 1993, p. 60), and the order-2 stationary condition, derived by Armstrong and Diamond (1984), is stated as

$$\frac{2\gamma(\boldsymbol{h})}{\|\boldsymbol{h}\|^2} \to 0, \qquad \|\boldsymbol{h}\| \to \infty \,. \tag{1.1}$$

In weak stationary case, the order-2 stationary condition is automatically satisfied.

We consider the sample $\boldsymbol{Z}$ is the observed values of $Z(\boldsymbol{s})$ at the points contained in a set, $D \in I\!R$. That is, $\boldsymbol{Z} = \{Z(\boldsymbol{s}) : \boldsymbol{s} \in D\}$ and the set D is the set

[1]StatSoft, Inc., U.S.A./ KangNam-Ku NonHyun-Dong DongHyun APT 1-805
[2]Department of Statistics,Iowa State University, Ames, IA 50011, U.S.A.

of sample points. Matheron (1962) suggested the method of moment estimator, $2\hat{\gamma}(\boldsymbol{h})$, of the variogram $2\gamma(\boldsymbol{h})$ as follows;

$$2\hat{\gamma}(\boldsymbol{h}) = \frac{1}{N(\boldsymbol{h})} \sum_{\boldsymbol{s}\in D(\boldsymbol{h})} (Z(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s}))^2, \tag{1.2}$$

where $D(\boldsymbol{h}) = \{\boldsymbol{s} \in D : \boldsymbol{s}+\boldsymbol{h} \in D\}$ and $N(\boldsymbol{h})$ is the number of the elements in the set $D(\boldsymbol{h})$, *i.e.* $|D(\boldsymbol{h})|$. The set of sample points D is conveniently assumed to be a subset of $\mathbb{Z}^d$. When we deal with non-regular lattice data in practice, the definition of the set D and $D(\boldsymbol{h})$ needs to be slightly changed to give tolerance between exact lattice points and sample points which are not on the exact lattice points.

Cressie (1985,1993) obtained the exact expression of the covariance of the variogram estimators, $Cov\big(2\hat{\gamma}(\boldsymbol{h}_1), 2\hat{\gamma}(\boldsymbol{h}_2)\big)$, by assuming Gaussianity. He also obtained the asymptotic approximation of the variance of the variogram estimator,

$$Var(2\hat{\gamma}(\boldsymbol{h})) \simeq \frac{2}{N(\boldsymbol{h})}(2\gamma(\boldsymbol{h};\boldsymbol{\theta}))^2, \tag{1.3}$$

when $2\gamma(\boldsymbol{h};\boldsymbol{\theta})$ is the true variogram model. Asymptotic expression (1.3) was used to justify and to explain the good performance of his weighted least estimators (WLS) of the parameter $\boldsymbol{\theta}$ of the model variogram.

Although the approximation (1.3) is very simple and persuasive in explaining WLS, the true variance of the variogram estimator is not well approximated by (1.3). Even in some cases, the value of $N(\boldsymbol{h}) \cdot Var(2\hat{\gamma}(\boldsymbol{h}))$ goes to infinity, but the term $\{2\gamma(\boldsymbol{h};\boldsymbol{\theta})\}^2$ in RHS of (1.3) may take only finite value. Inaccuracy of the approximation (1.3) is also noted in Zhang *et al.* (1995). In the following sections, we will generally consider the properties of the covariances of Matheron's variogram estimators. First, we will test the conditions under which the term $Cov\big(2\hat{\gamma}(\boldsymbol{h}_1), 2\hat{\gamma}(\boldsymbol{h}_2)\big)$ is expressed in a simple form as Cressie (1985) suggested. Next, we will develope more precise approximation of the variances and covariances of the variogram estimators than approximation (1.3) in asymptotic view point. Finally, we will suggest a sufficient condition that guarantees the finiteness of the asymptotic variances of the (normalized) variogram estimators.

2. COVARIANCES OF MATHERON'S VARIOGRAM ESTIMATORS

Generally covariances of variogram estimators are functions of the fourth order moment terms of the underlying process. To get the the covariance formula of variogram estimators in terms of the second order moment terms, we need to define the fourth order cumulant function of the spatial process. Assume that the intrinsically stationary spatial process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in I\!R^d\}$ has the characteristic function $\phi_Z(\cdot; \boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4)$ of the marginal distribution of $(Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_4))'$, at four arbitrary points of spatial indices $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4\}$. If the quantity

$$\xi_Z(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_4) \equiv \left[\frac{\partial^4}{\partial t_1 \partial t_2 \partial t_3 \partial t_4} \log \phi_Z(\boldsymbol{t}; \boldsymbol{s}_1, \ldots, \boldsymbol{s}_4) \right]_{\boldsymbol{t}=\boldsymbol{0}} \tag{2.1}$$

is the same constant regardless of the choice of $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_4 \in I\!R^d$, then the fourth moment terms of the process $Z(\boldsymbol{s})$ are expressed with their second moment terms.

Lemma 2.1. *For an intrinsically stationary process* $\{Z(\boldsymbol{s})\}$, *if the condition,*

$$\xi_Z(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_4) = \xi_Z \quad \text{for all} \quad \boldsymbol{s}_1, \ldots, \boldsymbol{s}_4 \in I\!R^d$$

is satisfied, then the covariance of the variogram estimators is shown as,

$$Cov\big(2\hat{\gamma}(\boldsymbol{h}_1), 2\hat{\gamma}(\boldsymbol{h}_2)\big) = \frac{2}{N(\boldsymbol{h}_1)N(\boldsymbol{h}_2)} \sum_{\boldsymbol{u} \in D(\boldsymbol{h}_1)} \sum_{\boldsymbol{v} \in D(\boldsymbol{h}_2)} \mathrm{q}(\boldsymbol{u} - \boldsymbol{v}, \boldsymbol{h}_1, \boldsymbol{h}_2) , \tag{2.2}$$

where

$$\mathrm{q}(\boldsymbol{u}, \boldsymbol{h}_1, \boldsymbol{h}_2) = \big\{\gamma(\boldsymbol{u}) - \gamma(\boldsymbol{u} + \boldsymbol{h}_1) - \gamma(\boldsymbol{u} - \boldsymbol{h}_2) + \gamma(\boldsymbol{u} + \boldsymbol{h}_1 - \boldsymbol{h}_2)\big\}^2 .$$

The proof is given in Appendix.

The form of covariance function (2.2) implies the correlation formula derived by Cressie (1985). We need to note that there is no restriction on the sampling points and sampling region in proving the lemma. That is, even for nonlattice sampling the lemma is also applicable.

If the random sequence $\{Z(i)\}_{i=1}^{\infty}$ is an i.i.d, ξ_Z should be 0 to satisfy the condition in lemma 2.1. The cumulants generating function $\log \phi_Z(\boldsymbol{t})$ has the form $\sum f(t_i)$ for a function $f(\cdot)$ for i.i.d. case, and the partial derivatives w.r.t. two different variables t_i and t_j make all terms 0. Generally, the *intrinsically stationary* condition in lemma 2.1 is not necessary. Without the condition, we just need to replace the terms of $\gamma(\cdot)$ in (2.2) with the terms of $\mu(\cdot)$ defined in (A.3). Since $\xi_Z = 0$ for all Gaussian processes $\{Z(\boldsymbol{s})\}$, all Gaussian processes having constant means autometically satisfy the condition of the lemma.

From (2.2), the variances of the variogram estimator are directly derived as

$$Var(2\hat{\gamma}(\boldsymbol{h})) = \frac{2}{(N(\boldsymbol{h}))^2} \sum_{\boldsymbol{u}\in D(\boldsymbol{h}_1)} \sum_{\boldsymbol{v}\in D(\boldsymbol{h}_2)} \{2\gamma(\boldsymbol{u}-\boldsymbol{v})-\gamma(\boldsymbol{u}-\boldsymbol{v}+\boldsymbol{h})-\gamma(\boldsymbol{u}-\boldsymbol{v}-\boldsymbol{h})\}^2 \tag{2.3}$$

If the process $\{Z(\boldsymbol{s})\}$ has varying mean, say $E\{Z(\boldsymbol{s})\} = \mu(\boldsymbol{s})$, from (A.1) we get

$$\begin{aligned}
&Cov\big((Z(\boldsymbol{u}) - Z(\boldsymbol{u}+\boldsymbol{h}_1))^2, (Z(\boldsymbol{v}) - Z(\boldsymbol{v}+\boldsymbol{h}_2))^2\big) \\
&\quad = Cov\big((\tilde{Z}(\boldsymbol{u}) - \tilde{Z}(\boldsymbol{u}+\boldsymbol{h}_1))^2, (\tilde{Z}(\boldsymbol{v}) - \tilde{Z}(\boldsymbol{v}+\boldsymbol{h}_2))^2\big) \\
&\quad +2\, d_\mu(\boldsymbol{u},\boldsymbol{h}_1)\cdot E\Big[\{\tilde{Z}(\boldsymbol{u}) - \tilde{Z}(\boldsymbol{u}+\boldsymbol{h}_1)\}\{\tilde{Z}(\boldsymbol{v}) - \tilde{Z}(\boldsymbol{v}+\boldsymbol{h}_2)\}^2\Big] \\
&\quad +2\, d_\mu(\boldsymbol{v},\boldsymbol{h}_2)\cdot E\Big[\{\tilde{Z}(\boldsymbol{v}) - \tilde{Z}(\boldsymbol{v}+\boldsymbol{h}_2)\}\{\tilde{Z}(\boldsymbol{u}) - \tilde{Z}(\boldsymbol{u}+\boldsymbol{h}_1)\}^2\Big] \\
&\quad +4\, d_\mu(\boldsymbol{u},\boldsymbol{h}_1)\, d_\mu(\boldsymbol{v},\boldsymbol{h}_2) E\Big[\{\tilde{Z}(\boldsymbol{u}) - \tilde{Z}(\boldsymbol{u}+\boldsymbol{h}_1)\}\{\tilde{Z}(\boldsymbol{v}) - \tilde{Z}(\boldsymbol{v}+\boldsymbol{h}_2)\}\Big]
\end{aligned}$$

for $\tilde{Z}(\boldsymbol{s}) = Z(\boldsymbol{s}) - \mu(\boldsymbol{s})$ and $d_\mu(\boldsymbol{u},\boldsymbol{h}) = \mu(\boldsymbol{u}) - \mu(\boldsymbol{u}+\boldsymbol{h})$. Thus, when the mean of the process is not constant, the third moment term of the process $\{Z(\boldsymbol{s})\}$ appears in $Cov(2\hat{\gamma}(\boldsymbol{h}_1), 2\hat{\gamma}(\boldsymbol{h}_2))$.

As an example we can consider the stationary process $\{Z(\boldsymbol{s})\}$ having the Gaussian variogram model ; $2\gamma(\boldsymbol{h}) = 2\theta_1 \cdot (1 - \exp(-(\theta_2|h_1|^2 + \theta_3|h_2|^2)))$, with $\theta_1 = 1.0$, $\theta_2 = 0.2$, and $\theta_3 = 0.3$. Suppose that we choose the three lag vectors, $\boldsymbol{h}_1 = (0,1), \boldsymbol{h}_2 = (1,0), \boldsymbol{h}_3 = (1,1)$. Assume $2\hat{\gamma} = (2\hat{\gamma}(\boldsymbol{h}_1), 2\hat{\gamma}(\boldsymbol{h}_2), 2\hat{\gamma}(\boldsymbol{h}_3))$, $D = \{(i,j) : i,j = -2,\ldots,3\}$, the integer-lattice points in the sampling region $R = (-3,3] \times (-3,3]$. Then, from lemma 2.1, we have

$$Cov(2\hat{\gamma}, 2\hat{\gamma}) = \frac{1}{36}\cdot \begin{pmatrix} 1.0793722 & 0.4032465 & 1.404173 \\ 0.4032465 & 2.3014128 & 2.635349 \\ 1.4041734 & 2.6353489 & 6.135458 \end{pmatrix}.$$

3. FINITE ASYMPTOTIC VARIANCE

The variances and covariances of variogram estimators are functions of variogram model and sample size. In spatial situation, *increasing domain asymptotics* (IDA), *infill asymptotics* (IFA), and mixed versions of IDA and IFA are asymptotic frameworks that can be considered. IDA assumes that one observes the spatial stochastic process at an increasing number of sites such that any two sites are at least a fixed distance apart. In this case, the sampling region eventually becomes unbounded, as the sample size tends to infinity. In IFA, the sampling region is necessarily bounded, and more and more samples are taken from the given region. As a result, the minimum distance between the data-sites tends to zero as the sample size tends to infinity (*ref.* Lahiri, 1996). In the followings, we will only consider the case of IDA which is the most basic asymptotic framework, for the simplicity of arguments. The related topics to the mixed version of IFA and IDA are reviewed in Lahiri *et al.* (2001).

To set the asymptotic structure, we assume that the set of sample points, D, is the set of lattice points in the sampling region, R_n, which is defined to be an inflated set of a set R_0, by a scaling factor λ_n, *i.e.* $D = R_n \cap \mathbb{Z}^d$ and $R_n = \lambda_n \cdot R_0$, where λ_n is a sequence of real numbers going to infinity as $n \to \infty$, and R_0 is an open subset of $(-\frac{1}{2}, \frac{1}{2}]^d$ containing the origin. R_0 is uaually assumed to have good properties such as *regular boundary condition* (*ref.* Lahiri, 1996) and *star-shape property* (*ref.* Sherman and Carlstein, 1994). Since the set of sample points D depends on the parameter n, all other symbols related to D also depend on n. To denote this, the subscript n will be appended to all related symbols when it is needed, as in the cases of D_n, N_n and $\hat{\gamma}_n(\boldsymbol{h})$.

Under the standard weak dependence assumptions on the random field and related asymptotics shown in Ibragimov and Linnik (1971), we derive the asymptotic quantity of (2.3) in the following theorem.

Theorem 3.1. *With the two sets* $D(\boldsymbol{h},\boldsymbol{u}) = \{\boldsymbol{s} \in D(\boldsymbol{h}) | \boldsymbol{s}+\boldsymbol{u} \in D(\boldsymbol{h})\}$, *and* $D^c(\boldsymbol{h},\boldsymbol{u}) = (D(\boldsymbol{h},\boldsymbol{u}))^c \cap D(\boldsymbol{h})$, *the set* $D(\boldsymbol{h})$ *is devided into two disjoint parts. When* $N_n^c(\boldsymbol{h},\boldsymbol{u})$ *denotes the number of the points in* $D^c(\boldsymbol{h},\boldsymbol{u})$, *if the condition,*

$$\sum_{\boldsymbol{u} \in D_n(\boldsymbol{h})} N_n^c(\boldsymbol{h},\boldsymbol{u}) \Big\{ 2\gamma(\boldsymbol{u}) - \gamma(\boldsymbol{u}-\boldsymbol{h}) - \gamma(\boldsymbol{u}+\boldsymbol{h}) \Big\}^2 = o(N_n(\boldsymbol{h})) \tag{3.1}$$

is satisfied, in asymptotic view points of IDA, the variance of the variogram given in (2.3) is approximated by

$$Var(2\hat{\gamma}(\boldsymbol{h})) \simeq \frac{2}{N(\boldsymbol{h})} \sum_{\boldsymbol{u} \in D(\boldsymbol{h})} \Big\{ 2\gamma(\boldsymbol{u}) - \gamma(\boldsymbol{u}-\boldsymbol{h}) - \gamma(\boldsymbol{u}+\boldsymbol{h}) \Big\}^2. \tag{3.2}$$

The proof is given in appendix.

The approximation (1.3) suggested by Cressie (1985) also gives a good approximation consistently with (3.2), if

$$\frac{\gamma(\boldsymbol{u}+\boldsymbol{h}) + \gamma(\boldsymbol{u}-\boldsymbol{h})}{2} - \gamma(\boldsymbol{u}) \simeq 0, \tag{3.3}$$

for all other points of $\boldsymbol{u}$ except few points near the origin $\mathbf{0}$. Approximation (1.3) underestimates the true asymptotic variance most of cases. In the example which will be stated at the end of this section, the value of $N(\boldsymbol{h})Var(2\hat{\gamma}(\boldsymbol{h}))$ obtained from (3.2) takes ∞, but $\{2\gamma(\boldsymbol{h})\}^2$ in (1.3) takes only finite values for all $\boldsymbol{h}$.

In asymptotic view points, to adjust sample size effect, the normalized variogram estimator $\sqrt{N(\boldsymbol{h})}\cdot 2\hat{\gamma}(\boldsymbol{h})$ is considered. When sampling sites lie on the integer grid D_n, the asymptotic covariance of normalized variogram estimators is given as

$$N_n \cdot Cov(2\hat{\gamma}_n(\boldsymbol{h}_1), 2\hat{\gamma}_n(\boldsymbol{h}_2)) \to \sum_{\boldsymbol{u}\in\mathbb{Z}^d} 2\cdot\{\gamma(\boldsymbol{u}) - \gamma(\boldsymbol{u}+\boldsymbol{h}_1) - \gamma(\boldsymbol{u}-\boldsymbol{h}_2) + \gamma(\boldsymbol{u}+\boldsymbol{h}_1-\boldsymbol{h}_2)\}^2 \tag{3.4}$$

as $n \to \infty$ provided the conditions of lemma 2.1 and the conditions on the sampling region R_n in the IDA framework are satisfied. We will call the RHS of (3.4) *asymptotic covariance matrix* (ACV), in short. From (3.4), each variance term is obtained as

$$N_n \cdot Var(2\hat{\gamma}_n(\boldsymbol{h}_1), 2\hat{\gamma}_n(\boldsymbol{h}_2)) \to \sum_{\boldsymbol{u}\in\mathbb{Z}^d} 2\cdot\{2\gamma(\boldsymbol{u}) - \gamma(\boldsymbol{u}-\boldsymbol{h}) - \gamma(\boldsymbol{u}-\boldsymbol{h})\}^2 \tag{3.5}$$

which depends on the model variogram itself.

In the Gaussian variogram model mentioned as an example in the previous section, the ACV of the model is

$$\begin{pmatrix} 1.3074411 & 0.6800292 & 1.805354 \\ 0.6800292 & 2.7185115 & 3.139840 \\ 1.8053538 & 3.1398401 & 6.481423 \end{pmatrix}.$$

Since the summation in (3.4) and (3.5) cover the whole region of $\mathbb{Z}^d$, the RHS of (3.4) and (3.5) may not be finite depending on the model variogram. Thus we need to devise a criterion to check whether the variogram model have finite asymptotic variance. The next theorem provides a sufficient condition that the ACV (3.4) will be finite, for 1-dimensional spatial process.

Theorem 3.2. *For an 1-dimensional intrinsic stationary process $\{Z(s) : s \in \mathbb{R}\}$ having the variogram model, $2\gamma_0(|h|)$, assume that there exists a sequence $\{C_k\}$ of constants satisfying the following conditions.*

i) $\limsup_{k\to\infty} \left|\frac{C_k}{k!}\right|^{1/k} = 0.$

ii) There exists a positive constant M, such that $\gamma_0(\cdot)$ is analytic on $[M, \infty)$ and

$$\int_{t>M} \left(\frac{\partial^k}{\partial t^k} \gamma_0(t) \right)^2 dt < C_k^2, \quad \forall \ k \geq 2 .$$

Then, the normalized variogram estimator $\sqrt{N_n(h)} \cdot 2\hat{\gamma}_n(h)$ has finite asymptotic variance,

$$\sum_{u \in \mathbb{Z}} 2 \cdot \{2\gamma_0(|u|) - \gamma_0(|u+h|) - \gamma_0(|u-h|)\}^2.$$

The proof is given in Appendix.

For the variogram model $2\gamma(h) = |h|^\theta$ of the 1-dimensional fractional isotropic Brownian motion in $\mathbb{R}$, because of order-2 stationary condition, (1.1), θ should be a value in $(0, 2)$. The conditions in theorem 3.2 require $\theta < 1.5$ to have finite asymptotic variance. When $\theta = 1.5$,

$$(2\gamma(u) - \gamma(u+1) - \gamma(u-1))^2 \simeq \frac{9}{16} \cdot \frac{1}{u} + O(\frac{1}{u^3})$$

for sufficiently large value of u. Hence the asymptotic variance cannot be finite for $\theta = 1.5$. Although theorem 3.2 says the condition only in the direction of sufficiency, the condition looks very tight and nearly necessary.

4. CONCLUSION

In theoretical view point, we reviewed the conditions in which the covariance of Matheron's variogram estimators would be expressed in the simple form as Cressie (1985) suggested. The better asymptotic variance of the variogram estimator than the previously suggested approximation was obtained in IDA asymptotic framework. With the new asymptotic variance, We obtained sufficient conditions that guarantee the finiteness of the asymptotic variance. As shown in the example, since not all valid variogram models give finite variances of (normalized) variogram estimators, when we assume the finiteness of the asymptotic variances of the variogram estimators, we need to give careful attention in choosing the model variogram.

APPENDIX

Proof of lemma 2.1 Obviously,

$$Cov\big(2\hat{\gamma}(\boldsymbol{h}_1), 2\hat{\gamma}(\boldsymbol{h}_2)\big) = \frac{2}{N(\boldsymbol{h}_1)N(\boldsymbol{h}_2)} \sum_{\boldsymbol{u}\in D(\boldsymbol{h}_1)} \sum_{\boldsymbol{v}\in D(\boldsymbol{h}_2)} \mathbf{C}_{\mathbf{v}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{h}_1,\boldsymbol{h}_2)\ , \quad \text{(A.1)}$$

where

$$\mathbf{C}_{\mathbf{v}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{h}_1,\boldsymbol{h}_2) = Cov\big((Z(\boldsymbol{u}) - Z(\boldsymbol{u}+\boldsymbol{h}_1))^2, (Z(\boldsymbol{v}) - Z(\boldsymbol{v}+\boldsymbol{h}_2))^2\big)\ .$$

To represent covariance term of RHS in (A.1) with second moment terms of the underlying process, we need the followings. Define centralized moments of order k,

$$\mu(\boldsymbol{s}_1,\cdots,\boldsymbol{s}_k) = E\Big\{\prod_{i=1}^{k}\big(Z(\boldsymbol{s}_i) - \mu(\boldsymbol{s}_i)\big)\Big\},$$

and $m(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4) = \mu(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4) - \mu(\boldsymbol{s}_1,\boldsymbol{s}_2)\mu(\boldsymbol{s}_3,\boldsymbol{s}_4)$.

Since intrinsic stationary precesses are assumed to have constant mean, we may assume $E\{Z(\boldsymbol{s})\} = 0$ without loss of generality. Fix $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_4 \in \mathbb{R}^d$. For simplicity, assume that $(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2), Z(\boldsymbol{s}_3), Z(\boldsymbol{s}_4))$ has a characteristic function, which we again denote by $\phi_Z(\boldsymbol{t})$, *i.e.*

$$\phi_Z(\boldsymbol{t}) = E\Big\{\exp\big(i\,t_1\ Z(\boldsymbol{s}_1) + i\,t_2\ Z(\boldsymbol{s}_2) + i\,t_3\ Z(\boldsymbol{s}_3) + i\,t_4\ Z(\boldsymbol{s}_4)\big)\Big\}.$$

Define $P(\boldsymbol{t}) = \log\phi_Z(\boldsymbol{t})$, and

$$\frac{\partial^k}{\partial t_1\cdots\partial t_k}P(\boldsymbol{t}) = P_{t_1,\cdots,t_k}(\boldsymbol{t}) = P_{t_1,\cdots,t_k}.$$

Then,

$$\frac{\partial^2}{\partial t_1\partial t_2}\phi_Z(\boldsymbol{t}) = \big(P_{t_1,t_2} + P_{t_1}\cdot P_{t_2}\big)\phi_Z(\boldsymbol{t})\ .$$

Since $\mu = 0$, $P_{t_i}(\mathbf{0}) = 0\ \ \forall i$, and $P_{t_i,t_j}(\mathbf{0}) = \mu(\boldsymbol{s}_i,\boldsymbol{s}_j)$,

$$\Big[\frac{\partial^4}{\partial t_1\partial t_2\partial t_3\partial t_4}\phi_Z(\boldsymbol{t})\Big]_{\boldsymbol{t}=\mathbf{0}} =$$
$$\phi_Z(\mathbf{0})\cdot\Big[P_{t_1,t_2}(\mathbf{0})P_{t_3,t_4}(\mathbf{0}) + P_{t_1,t_3}(\mathbf{0})P_{t_2,t_4}(\mathbf{0}) + P_{t_1,t_4}(\mathbf{0})P_{t_2,t_3}(\mathbf{0}) + P_{t_1,t_2,t_3,t_4}(\mathbf{0})\Big].$$

If the condition of the 2.1 is satisfied,

$$\mu(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4) = \mu(\boldsymbol{s}_1,\boldsymbol{s}_2)\mu(\boldsymbol{s}_3,\boldsymbol{s}_4) + \mu(\boldsymbol{s}_1,\boldsymbol{s}_3)\mu(\boldsymbol{s}_2,\boldsymbol{s}_4) + \mu(\boldsymbol{s}_1,\boldsymbol{s}_4)\mu(\boldsymbol{s}_2,\boldsymbol{s}_3) + \xi_Z.$$

Hence,

$$m(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4) = \mu(\boldsymbol{s}_1, \boldsymbol{s}_3)\mu(\boldsymbol{s}_2, \boldsymbol{s}_4) + \mu(\boldsymbol{s}_1, \boldsymbol{s}_4)\mu(\boldsymbol{s}_2, \boldsymbol{s}_3) + \xi_Z. \tag{A.2}$$

Since $Cov\big(Z(\boldsymbol{u}_1)Z(\boldsymbol{u}_2), Z(\boldsymbol{v}_1)Z(\boldsymbol{v}_2)\big) = m(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}_1, \boldsymbol{v}_2)$,

$$\begin{aligned}
&Cov\Big\{\big(Z(\boldsymbol{u}+\boldsymbol{h}_1) - Z(\boldsymbol{u})\big)^2, \big(Z(\boldsymbol{v}+\boldsymbol{h}_2) - Z(\boldsymbol{v})\big)^2\Big\} \\
&\quad = m(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}) + m(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}+\boldsymbol{h}_2,\boldsymbol{v}+\boldsymbol{h}_2) - 2m(\boldsymbol{u},\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v},\boldsymbol{v}) \\
&\quad + m(\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v},\boldsymbol{v}) - 2m(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}+\boldsymbol{h}_2)4m(\boldsymbol{u},\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v},\boldsymbol{v}+\boldsymbol{h}_2) \\
&\quad + m(\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v}+\boldsymbol{h}_2,\boldsymbol{v}+\boldsymbol{h}_2) - 2m(\boldsymbol{u},\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v}+\boldsymbol{h}_2,\boldsymbol{v}+\boldsymbol{h}_2) \\
&\quad - 2m(\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v},\boldsymbol{v}+\boldsymbol{h}_2).
\end{aligned}$$

By substitututing (A.2), we get

$$\begin{aligned}
&Cov\Big\{\big(Z(\boldsymbol{u}+\boldsymbol{h}_1) - Z(\boldsymbol{u})\big)^2, \big(Z(\boldsymbol{v}+\boldsymbol{h}_2) - Z(\boldsymbol{v})\big)^2\Big\} \\
&\quad = 2\big\{\mu(\boldsymbol{u},\boldsymbol{v}) - \mu(\boldsymbol{u},\boldsymbol{v}+\boldsymbol{h}_2) - \mu(\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v}) + \mu(\boldsymbol{u}+\boldsymbol{h}_1,\boldsymbol{v}+\boldsymbol{h}_2)\big\}^2,
\end{aligned} \tag{A.3}$$

and by plugging this into (A.1), the result follows.

Proof of lemma 3.1 With the $\mathrm{q}(\boldsymbol{u}, \boldsymbol{h_1}, \boldsymbol{h_2})$ defined in lemma 2.1,

$$\begin{aligned}
N(\boldsymbol{h}) \cdot Var\big(2\hat{\gamma}(\boldsymbol{h})\big) &= \frac{2}{N(\boldsymbol{h})} \sum_{\boldsymbol{u}\in D(\boldsymbol{h}_1)} \sum_{\boldsymbol{v}\in D(\boldsymbol{h}_2)} \mathrm{q}(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{h},\boldsymbol{h}) \\
&= \frac{2}{N(\boldsymbol{h})} \sum_{\boldsymbol{w}\in D(\boldsymbol{h},\boldsymbol{u}-\boldsymbol{v})} \sum_{\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{w}} \mathrm{q}(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{h},\boldsymbol{h}) \\
&= 2 \sum_{\boldsymbol{w}\in D(\boldsymbol{h})} \{\frac{N(\boldsymbol{h}) - N^c(\boldsymbol{h},\boldsymbol{w})}{N(\boldsymbol{h})}\}\mathrm{q}(\boldsymbol{w},\boldsymbol{h},\boldsymbol{h}).
\end{aligned}$$

With the condition (3.1) the result directly follows.

Proof of lemma 3.2 Under condition ii) $\gamma_0(t)$ is analytic for sufficiently large values of t, and here,

$$\gamma_0(t+h) = \sum_{k=0}^{\infty} \frac{{\gamma_o}^{(k)}(t)}{k!}h^k, \quad \forall\, t > M, \quad t+h > M \tag{A.4}$$

when ${\gamma_o}^{(k)}(\cdot)$ means k-th order derivative of $\gamma_o(\cdot)$. In the followings, we will use two new symbols a set A and a coeficient function c for notational convenience. They are defined as follows ;

$$A = \{(\alpha, \beta) : \alpha, \beta = 0, -h, h\}$$
$$c(\alpha, \beta) = \begin{cases} 4, & \text{if } \alpha = 0 \text{ and } \beta = 0, \\ 1, & \text{if } \alpha \neq 0 \text{ and } \beta \neq 0, \\ -2, & \text{otherwise.} \end{cases}$$

From (A.4),

$$\begin{aligned} &\left(2\gamma_0(t) - \gamma_0(t-h) - \gamma_0(t+h)\right)^2 \\ &= \sum_{(\alpha,\beta)\in A} c(\alpha,\beta)\left[\sum_{k=2}^{\infty}\sum_{l=2}^{\infty} \frac{{\gamma_o}^{(k)}(t){\gamma_o}^{(l)}(t)}{k!\ l!}\alpha^k\beta^l\right], \quad \text{for} \quad t,\, t+h,\, t-h \in [M, \infty). \end{aligned} \tag{A.5}$$

In asymptotic variance, we need to sum (A.5) over the region $\{s \ : \ |s| > M\}$. This summation is approximated by 1-dimmensional integeration because of the symetric property of $\gamma_0(\cdot)$. Thus the asymptotic variance is bounded by

$$d\,M \;+\; d\left|\sum_{(\alpha,\beta)\in A} c(\alpha,\beta)\left[\sum_{k=2}^{\infty}\sum_{l=2}^{\infty}\left\{\int_{t>M}\gamma_0^{(k)}(t)\gamma_0^{(l)}(t)dt\right\}\frac{\alpha^k\beta^l}{k!\ l!}\right]\right|, \tag{A.6}$$

for a suitable constant $d > 0$. The integral term in (A.6) has *Cauchy-Schwarz* bound;

$$\left|\int_{t>M}\gamma_0^{(k)}(t)\gamma_0^{(l)}(t)dt\right| \le \left(\int_{t>M}\left(\gamma_0^{(k)}(t)\right)^2dt\right)^{\frac{1}{2}}\left(\int_{t>M}\left(\gamma_0^{(l)}(t)\right)^2dt\right)^{\frac{1}{2}}. \tag{A.7}$$

Now, the summation with respect to k and l in (A.6) are separated and we have terms inside summation over the set A ;

$$\sum_{k=2}^{\infty}\left(\int_{t>M}\left(\gamma_0^{(k)}(t)\right)^2dt\right)^{\frac{1}{2}}\frac{\alpha^k}{k!} \qquad \text{and} \qquad \sum_{k=2}^{\infty}\left(\int_{t>M}\left(\gamma_0^{(l)}(t)\right)^2dt\right)^{\frac{1}{2}}\frac{\beta^l}{l!} \tag{A.8}$$

From the definition of the set A, we can see α and β have values depending on h. To guarantee that the series in (A.8) have finite values for all h, the series in (A.8) have to have infinite radii of convergence with respect to α and β which is ensured by i).

REFERENCES

Armstrong, M. and Diamond, P. (1984). Testing variograms for positive-definiteness, *Journal of the International Association for Mathematical Geology*, *16*, 407-421.

Cressie, N. (1985). Fitting variogram models by weighted least squares, *Journal of the International Association for Mathematical Geology*, *17*, 693-702.

Cressie, N. (1993). *Statistics for spatial data*, Wiley, New York.

Ibragimov, I. A. and Linnik, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*, Wolter-Noordhoff Publishing, Nethelands.

Lahiri, S. N. (1996). On Inconsistency of Estimators Based on Spatial Data Under Infill Asymptotics, *Sankhyā, Series A, Indian Journal of Statistics*, *58*, 403-417.

Lahiri, S., Lee, Y. and Cressie, N. (2001). On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters, *Journal of Statistical Planning and Inference*, will appear.

Matheron, G. (1962). Traite de Geostatistique Appliquee, Tome I, *Memoires du Bureau de Recherches Geologiques et Minieres,* No *14*. Editions Technip, Paris.

Sherman, M. and Carlstein, E. (1994). Nonparametric estimation of the moments of a general statistic computed from spatial data, *Journal of the American Statistical Association*, *89*, 496- 500.

Zhang, X., van Eijkeren, J., and Heemink, A. (1995). On the weighted least-squares method for fitting a semivariogram model, *Computers and Geosciences*, *21*,605 - 608.