Prace Naukowe Wyższej Szkoły Pedagogicznej w Częstochowie
Matematyka IX, Częstochowa 2003

Asymptotic Distribution of the Mutual Variogram Estimate

Nikolai Troush, Tatiana Tsekhovaya

The article deals with the problem of a statistical analysis of time series connected with the estimation of mutual variogram, a measure of spatial correlation. G. Matheron [1] has coined the term variogram, although earlier appearances of this function can be found in the scientific literature [2, 3].

The problem of estimating the mutual variogram and examination the statistical properties of this statistics has been considered in [1, 4, 5]. We present the limiting expressions of the first two moments and the higher order cumulants of the mutual variogram estimate of the second-order-stationary stochastic process with discrete time. These expressions are then used to prove the theorem concerning the asymptotic distribution of the mutual variogram estimate. The approach is similar to the approach taken in the time series literature, and the reader is referred to D. Brillinger [6] for theorems regarding the asymptotic distribution of the spectral density estimate of a time series.

Consider a random process

$$Y^r(s) = \{Y_a(s), a = \overline{1,r}\}, s \in Z = \{0, \pm 1, \pm 2, ...\}, r \geq 1.$$

Suppose further that $Y^r(s)$, $s \in Z$, is a zero-mean stochastic process with unknown mutual variogram

$$2\,\gamma_{ab}(h) = cov(Y_a(s+h) - Y_a(s),\ Y_b(s+h) - Y_b(s)),$$

$s, h \in Z,\ a, b = \overline{1,r}.$

The mutual variogram estimate $2\,\widetilde{\gamma}_{ab}\,(h)$ in terms of sequence of observations, $Y_a(1), Y_a(2), ..., Y_a(n)$, is defined as

$$2\,\widetilde{\gamma}_{ab}\,(h) = \frac{1}{n-h}\sum_{s=1}^{n-h}(Y_a(s+h) - Y_a(s))(Y_b(s+h) - Y_b(s)),$$

$\widetilde{\gamma}_{ab}\,(-h) = \widetilde{\gamma}_{ab}\,(h)$, $h = \overline{0, n-1}$, and $\widetilde{\gamma}_{ab}\,(h) = 0$ for $|h| \geq n$, $a, b = \overline{1,r}$.

We calculate the first two moments of the examined estimate. The mutual variogram estimate is unbiased for $2\,\gamma_{ab}(h)$.

Theorem 1. *For the mutual variogram estimate* $2\,\widetilde{\gamma}_{ab}\,(h)$, $h=\overline{0,n-1}$, $a,b=\overline{1,r}$, *the expression*

$$(n-h^-)cov\{2\,\widetilde{\gamma}_{a_1b_1}\,(h_1),2\,\widetilde{\gamma}_{a_2b_2}\,(h_2)\}=$$
$$=\sum_{k=-(n-h_2-1)}^{n-h_1-1}\Big(1+\frac{m(k)}{n-h^+}\Big)H^{h_1h_2}_{a_1b_1a_2b_2}(k),$$

is valid, where

$$\begin{gathered}H^{h_1h_2}_{a_1b_1a_2b_2}(k)=c_{a_1b_1a_2b_2}(k+h_1-h_2,k+h_1-h_2,0)+ \qquad (1)\\
+c_{a_1b_1a_2b_2}(k+h_1,k,h_2)+c_{a_1b_1a_2b_2}(k+h_1,k+h_1,0)-\\
-c_{a_1b_1a_2b_2}(k+h_1-h_2,k+h_1-h_2,-h_2)+c_{a_1b_1a_2b_2}(k+h_1-h_2,k-h_2,-h_2)-\\
-c_{a_1b_1a_2b_2}(k-h_2,k+h_1-h_2,0)+\\
+c_{a_1b_1a_2b_2}(k,k+h_1,h_2)-c_{a_1b_1a_2b_2}(k,k+h_1,0)-\\
-c_{a_1b_1a_2b_2}(k+h_1,k+h_1,h_2)+c_{a_1b_1a_2b_2}(k-h_2,k-h_2,0)+c_{a_1b_1a_2b_2}(k,k,0)+\\
+c_{a_1b_1a_2b_2}(k-h_2,k+h_1-h_2,-h_2)-c_{a_1b_1a_2b_2}(k-h_2,k-h_2,-h_2)-\\
-c_{a_1b_1a_2b_2}(k,k,h_2)-c_{a_1b_1a_2b_2}(k+h_1-h_2,k-h_2,0)-c_{a_1b_1a_2b_2}(k+h_1,k,0)+\\
+R_{a_1a_2}(k+h_1-h_2)R_{b_1b_2}(k+h_1-h_2)+R_{a_1b_2}(k+h_1-h_2)R_{b_1a_2}(k+h_1-h_2)-\\
-R_{a_1a_2}(k+h_1-h_2)R_{b_1b_2}(k+h_1)-R_{a_1b_2}(k+h_1)R_{b_1a_2}(k+h_1-h_2)-\\
-R_{a_1a_2}(k+h_1)R_{b_1b_2}(k+h_1-h_2)-R_{a_1b_2}(k+h_1-h_2)R_{b_1a_2}(k+h_1)+\\
+R_{a_1a_2}(k+h_1)R_{b_1b_2}(k+h_1)+R_{a_1b_2}(k+h_1)R_{b_1a_2}(k+h_1)-\\
-R_{a_1a_2}(k+h_1-h_2)R_{b_1b_2}(k-h_2)-R_{a_1b_2}(k+h_1-h_2)R_{b_1a_2}(k-h_2)+\\
+R_{a_1a_2}(k+h_1-h_2)R_{b_1b_2}(k)+\\
+R_{a_1b_2}(k+h_1)R_{b_1a_2}(k-h_2)-R_{a_1b_2}(k+h_1)R_{b_1a_2}(k)+\\
+R_{a_1a_2}(k+h_1)R_{b_1b_2}(k-h_2)+\\
+R_{a_1b_2}(k+h_1-h_2)R_{b_1a_2}(k)-R_{a_1a_2}(k+h_1)R_{b_1b_2}(k)-\\
-R_{a_1a_2}(k-h_2)R_{b_1b_2}(k+h_1-h_2)-R_{a_1b_2}(k-h_2)R_{b_1a_2}(k+h_1-h_2)+\\
+R_{a_1a_2}(k-h_2)R_{b_1b_2}(k+h_1)+R_{a_1b_2}(k)R_{b_1a_2}(k+h_1-h_2)+\\
+R_{a_1a_2}(k)R_{b_1b_2}(k+h_1-h_2)+R_{a_1b_2}(k-h_2)R_{b_1a_2}(k+h_1)-\\
-R_{a_1a_2}(k)R_{b_1b_2}(k+h_1)-R_{a_1b_2}(k)R_{b_1a_2}(k+h_1)+R_{a_1a_2}(k-h_2)R_{b_1b_2}(k-h_2)+\end{gathered}$$

$$+R_{a_1b_2}(k-h_2)R_{b_1a_2}(k-h_2)-R_{a_1a_2}(k-h_2)R_{b_1b_2}(k)-R_{a_1b_2}(k)R_{b_1a_2}(k-h_2)-$$
$$-R_{a_1a_2}(k)R_{b_1b_2}(k-h_2)-R_{a_1b_2}(k-h_2)R_{b_1a_2}(k)+$$
$$+R_{a_1a_2}(k)R_{b_1b_2}(k)+R_{a_1b_2}(k)R_{b_1a_2}(k),$$

$R_{ab}(k)$, $k\in Z$, $a,b=\overline{1,r}$, *is the mutual covariance function,* $c_{a_1a_2a_3a_4}(k_1,k_2,k_3)$, $k_1,k_2,k_3\in Z$, $a_i=\overline{1,r}$, $i=\overline{1,4}$, *is the fourth-order sample cumulants of the stochastic process* $Y^r(s), s\in Z$, $a_1,a_2,b_1,b_2=\overline{1,r}$,

$$m(k)=\begin{cases} m_1(k), & h_2>h_1,\\ m_2(k), & h_1>h_2,\end{cases}$$

$$m_1(k)=\begin{cases} k, & k=\overline{1-n+h_2,-1},\\ 0, & k=\overline{0,h_2-h_1},\\ h_2-h_1-k, & k=\overline{h_2-h_1+1,n-h_1-1},\end{cases}$$

$$m_2(k)=\begin{cases} k-h_2-h_1, & k=\overline{1-n+h_2,h_2-h_1-1},\\ 0, & k=\overline{h_2-h_1,0},\\ -k, & k=\overline{1,n-h_1-1},\end{cases}$$

$$h^-=\min\{h_1,h_2\},\quad h^+=\max\{h_1,h_2\}, \tag{2}$$

$h_1,h_2=\overline{0,n-1}$.

Theorem 2. *For the mutual variogram estimate* $2\stackrel{\sim}{\gamma}_{ab}(h)$, $h=\overline{0,n-1}$, $a,b=\overline{1,r}$, *the expression*

$$(n-h^-)cov\{2\stackrel{\sim}{\gamma}_{a_1b_1}(h_1),2\stackrel{\sim}{\gamma}_{a_2b_2}(h_2)\}=$$
$$=2\pi\int_\Pi \Phi_{n-h^+}(x)(G_1^{h_1-h_2}(x)+G_2^{h_1-h_2}(x))dx+$$
$$+\frac{1}{n-h^+}\int_\Pi \Delta_{n-h^+}(x)\Delta_{h^+-h^-}(x)\cos\frac{(n-h^+)x}{2}(G_1(x)+G_2(x))dx,$$

is valid, where

$$G_1^{h_1-h_2}(x)=\cos\frac{(h_1-h_2)x}{2}G_1(x), \tag{3}$$

$$G_2^{h_1-h_2}(x)=\cos\frac{(h_1-h_2)x}{2}G_2(x),$$

$$G_1(x) = \int\int_{\Pi^2} f_{a_1b_1a_2b_2}(x-y, y, z) q(x, y, z) dy dz,$$

$$G_2(x) = \int_{\Pi} (f_{a_1a_2}(x-y) f_{b_1b_2}(y) + f_{a_1b_2}(x-y) f_{b_1a_2}(y)) q(x, y) dy,$$

$$q(x,y,z) = e^{i\frac{x(h_2-h_1)}{2}}(1 + e^{ix(h_1-h_2)} e^{i(xh_1+zh_2)} - e^{ix(h_1-h_2)-izh_2} + e^{ixh_1} -$$
$$-e^{ix(h_1-h_2)-iyh_1} + e^{i(x-y)h_1+izh_2} + e^{i(x-y)(h_1-h_2)-iyh_2-izh_2} - e^{i(x-y)h_1} -$$
$$-e^{iyh_1-ixh_2} + e^{iyh_1+izh_2} + e^{iyh_1-i(x+z)h_2} - e^{iyh_1} - e^{izh_2} + e^{-ixh_2} - e^{-ih_2(x+z)}),$$

$$q(x,y) = e^{i\frac{x(h_2-h_1)}{2}}(1 + e^{ix(h_1-h_2)} - e^{ixh_1-i(x-y)h_2} - e^{ixh_1-iyh_2} + e^{ixh_1} -$$
$$-e^{ixh_2+i(x-y)h_1} + e^{i(x-y)(h_1-h_2)} + e^{i(x-y)h_1-iyh_2} - e^{i(x-y)h_1} - e^{iyh_1-ixh_2} +$$
$$+e^{iyh_1-i(x-y)h_2} + e^{iy(h_1-h_2)} - e^{iyh_1} + e^{-ixh_2} - e^{-i(x-y)h_2} - e^{-iyh_2}),$$

$f_{ab}(x)$, $x \in \Pi = [-\pi, \pi]$, *is the mutual spectral densities,* $f_{a_1b_1a_2b_2}(x_1, x_2, x_3)$, $x_i \in \Pi$, $i = \overline{1,3}$, *is the fourth-order cumulants spectral densities of stationary stochastic processes* $Y^r(s)$, $s \in Z$, $a, b, a_1, b_1, a_2, b_2 = \overline{1,r}$,

$$\Phi_n(x) = \frac{1}{2\pi n}\Delta_n^2(x), \; x \in \Pi,$$

$$\Delta_n^2(x) = \frac{\sin\frac{nx}{2}}{\sin\frac{x}{2}},$$

h^+, h^- *are (2),* $h_1, h_2 = \overline{0, n-1}$.

Theorem 3. *Let*

$$\sum_{k=-\infty}^{+\infty} |R_{ab}(k)| < \infty, \qquad \sum_{k_1,k_2,k_3=-\infty}^{+\infty} |c_{a_1b_1a_2b_2}(k_1,k_2,k_3)| < \infty.$$

Then

$$\lim_{n\to\infty}(n-h^-)cov\{2\widetilde{\gamma}_{a_1b_1}(h_1), 2\widetilde{\gamma}_{a_2b_2}(h_2)\} = \sum_{k=-\infty}^{+\infty} H^{h_1h_2}_{a_1b_1a_2b_2}(k), \quad (4)$$

where $H^{h_1h_2}_{a_1b_1a_2b_2}(k)$ *is (1),* h^- *is (2),* $h_1, h_2 = \overline{0, n-1}$, $a_1, b_1, a_2, b_2 = \overline{1,r}$.

Corollary 1. *Under the assumptions of Theorem 3,*

$$\lim_{n\to\infty} cov\{2\widetilde{\gamma}_{a_1b_1}(h_1), 2\widetilde{\gamma}_{a_2b_2}(h_2)\} = 0,$$

$$\lim_{n\to\infty} D\{2\,\widetilde{\gamma}_{ab}(h)\} = 0,$$

$h, h_1, h_2 = \overline{0, n-1}$, $a, b, a_1, b_1, a_2, b_2 = \overline{1, r}$.

Theorem 4. *Let the mutual spectral densities* $f_{ab}(x)$, $x \in \Pi$, *and the fourth-order cumulants spectral densities* $f_{a_1 b_1 a_2 b_2}(x_1, x_2, x_3)$, $x_i \in \Pi$, $i = \overline{1,3}$, *of the stationary stochastic processes* $Y^r(s)$, $s \in Z$, *are continuous on* Π, Π^3, *respectively. Then*

$$\lim_{n\to\infty} (n - h^-) cov\{2\,\widetilde{\gamma}_{a_1 b_1}(h_1), 2\,\widetilde{\gamma}_{a_2 b_2}(h_2)\} = 2\pi (G_1^{h_1-h_2}(0) + G_2^{h_1-h_2}(0)), \tag{5}$$

where $G_1^{h_1-h_2}(x)$, $G_2^{h_1-h_2}(x), x \in \Pi$, *are (3)*, h^- *is (2)*, $h_1, h_2 = \overline{0, n-1}$, $a_1, b_1, a_2, b_2 = \overline{1, r}$.

Corollary 2. *Under the assumptions of Theorem 4,*

$$\lim_{n\to\infty} cov\{2\,\widetilde{\gamma}_{a_1 b_1}(h_1), 2\,\widetilde{\gamma}_{a_2 b_2}(h_2)\} = 0,$$

$$\lim_{n\to\infty} D\{2\,\widetilde{\gamma}_{ab}(h)\} = 0,$$

$h, h_1, h_2 = \overline{0, n-1}$, $a, b, a_1, b_1, a_2, b_2 = \overline{1, r}$.

In order to found the asymptotic distribution of the mutual variogram estimate $2\,\widetilde{\gamma}_{ab}(h)$, $h = \overline{0, n-1}$, $a, b = \overline{1, r}$, *it is necessary to investigate an asymptotic behavior of the cumulant* $cum\{2\,\widetilde{\gamma}_{a_1 b_1}(h_1), ..., 2\,\widetilde{\gamma}_{a_p b_p}(h_p)\}$, $h_j = \overline{0, n-1}$, $a_j, b_j = \overline{1, r}$, $j = \overline{1, p}$.

Theorem 5. *Let*

$$\sum_{t_1, ..., t_{p-1} = -\infty}^{+\infty} |c_{a_1 ... a_p}(t_1, ..., t_{p-1})| < \infty,$$

where $c_{a_1 ... a_p}(t_1, ..., t_{p-1})$ *is the sample cumulant of the stationary stochastic processes* $Y^r(s)$, $s \in Z$, $t_j \in Z$, $j = \overline{1, p-1}$, $p > 2$. *Then*

$$\lim\nolimits_{n\to\infty} cum\{2\,\widetilde{\gamma}_{a_1 b_1}(h_1), ..., 2\,\widetilde{\gamma}_{a_p b_p}(h_p)\} = 0, \tag{6}$$

where $h_j = \overline{0, n-1}$, $a_j, b_j = \overline{1, r}$, $j = \overline{1, p}$, $n = 1, 2, 3, ...$

Theorem 6. *Let processes* $Y^r(s)$, $s \in Z$, *be second-order-stationary stochastic processes with continuous spectral densities up to order* s, $s = \overline{2, p}$, *for any* p, $p > 2$. *Then conclusion (6) of Theorem 5 holds.*

Theorem 7. *Let all the assumptions of Theorem 5 be satisfied. Then the mutual variogram estimate* $2\,\widetilde{\gamma}_{ab}(h)$, $h = \overline{0, n-1}$, $a, b = \overline{1, r}$, *is asymptotically normally distributed with mean* $2\gamma_{ab}(h)$, $h \in Z$, $a, b = \overline{1, r}$, *and asymptotic variance (4).*

Theorem 8. *Let all the assumptions of Theorem 6 be satisfied. Then the mutual variogram estimate* $2\,\widetilde{\widetilde{\gamma}}_{ab}\,(h)$, $h = \overline{0, n-1}$, $a, b = \overline{1, r}$, *is asymptotically normally distributed with mean* $2\gamma_{ab}(h)$, $h \in Z$, $a, b = \overline{1, r}$, *and asymptotic variance (5).*

References

[1] G. Matheron, *Principles of geostatistics,* Econ. Geol, **58**, 1246-1266, 1963.

[2] A. N. Kolmogorov, *The local structure of turbulence in an incompressible fluid at very large Reynolds numbers,* Doklady Akademii Nauk SSSR, **30**, 301-305, 1941.

[3] A. M. Yaglom, *Some classes of random fields in n - dimensional space, related to stationary random processes,* Theory of Probability and Its Applications, **2**, 273-320, 1957.

[4] A. G. Journel, C. J. Huijbregts, *Mining Geostatistics.* Academic Press, London, 1978.

[5] N. Cressie, *Statistics for Spatial Data.* John Wiley and Sons, New York. 1991.

[6] D. R. Brillinger, *Time Series, Data Analysis and Theory.* Holt, New York. 1975.

Academia Podlaska,
Siedlce,
Poland

Belarus State University
Minsk
Belarus