STUDIA INFORMATICA
Nr 1(5) Systemy i technologie informacyjne 2005

PIERWSZE DWA MOMENTY OCENY WARIOGRAMU

Mikołaj Trusz[1], Iwona Arcimowicz[2]

Abstract. In this article we deal with theoretical bases of geostatistic which have employment in research of time series. Let we have stationary random process, defined on R, with expected value equal 0, given covariance function and spectral density. For this process we define variogram function and prove two theorems about variogram estimator property

Key words: variogram function, spatially variable data, modelling semivariograms, geostatistical methods

Wstęp

W połowie XX wieku pojawił się nowy rodzaj statystycznej analizy szeregów czasowych – analiza wariogramu. Stało się możliwe głębsze i pełniejsze badanie zjawisk spotykanych w różnych dziedzinach życia człowieka, między innymi geologii i meteorologii. Metody analizy wariogramu umożliwiają otrzymywanie pełniejszej informacji o wewnętrznej strukturze danych eksperymentalnych i określają złożoność badanych procesów. W tym podstawowym modelem okazują się stacjonarne i wewnętrznie stacjonarne procesy losowe.

Twórcą analizy wariogramu szeregów czasowych jest G. Materon. Dalej teorię tą i przełożenie jej na praktyczne zastosowania rozwijali D.E. Meiers, S. Devis, M. David, L.E. Borgman, N. Cressie, D.M. Hawkins i inni.

Problemy badania statystycznych własności ocen wartości oczekiwanej wydają się dobrze zbadane. Zadania analizy własności podstawowych charakterystyk drugiego rzędu procesów stochastycznych są jednak nadal aktualne. W pierwszej kolejności chodzi tu o obliczenie i badanie oceny wariogramu, wzajemnego wariogramu, funkcji kowariancyjnej, wzajemnej kowariancji, które są podstawowymi miarami zależności danych eksperymentalnych i pomagają w rozwiązywaniu wielu problemów.

Własności oceny wariogramu

Rozpatrujemy stacjonarny w szerokim znaczeniu proces losowy X(s), $s \in R$ z wartością oczekiwaną równą zeru, funkcją kowariancyjną R(s), $s \in R$ o gęstości spektralnej $f(\lambda)$, $\lambda \in R$.

[1] Instytut Informatyki Akademii Podlaskiej, Siedlce
[2] Instytut Informatyki Akademii Podlaskiej, Siedlce

Wariogramem procesu X(s) nazywamy funkcję:

$$2\gamma(s) = V[X(t+s) - X(t)], t, s \in R$$

Zakładamy, że jest to funkcja nie znana i potrzebujemy dla realizacji procesu X(s), $s \in R$ utworzyć ocenę funkcji $2\gamma(s)$, $s \in R$.

Będziemy budować ocenę wariogramu jako statystykę postaci:

$$2\tilde{\gamma}(h) = \frac{1}{T-h}\int_0^{T-h} (X(s+h) - X(s))^2 ds, \tag{1}$$

gdzie, $h \in [0,T)$ oraz zakładamy, że $\tilde{\gamma}(-h) = \tilde{\gamma}(h), h \in [0,T)$ i $\tilde{\gamma}(h) = 0$ dla $|h| \geq T$.

Podobne oceny dla procesów dyskretnych były tworzone i badane w pracach [1], [2] i [3]. W tej pracy będziemy badać właściwości statystyki (1).

Twierdzenie 1.

Dla oceny $2\tilde{\gamma}(s)$ zadanej wzorem (1) mamy

$$E(2\tilde{\gamma}(s)) = 2\gamma(s)$$

$$\operatorname{cov}\{2\tilde{\gamma}(s_1), 2\tilde{\gamma}(s_2)\} = \frac{2\pi}{T-s_2}\int_R \Psi_{T-s_1}(u)(g^{s_1s_2}(u) + h^{s_1s_2}(u))du +$$

$$+\frac{1}{(T-s_2)(T-s_2)}\int_R \Delta_{T-s_1}(u)\Delta_{s_1-s_2}(u)(g^{s_1s_2}(u) + h^{s_1s_2}(u))du + \tag{2}$$

gdzie

$$\Psi_T(u) = \frac{1}{2\pi T}\Delta_T^2(u), \tag{3}$$

$$\Delta_T(u) = \frac{\sin\left(\frac{Tu}{2}\right)}{\frac{Tu}{2}}, \tag{4}$$

$$g^{s_1s_2}(u) = \iint_{R^2} f_4(x, u-z, z) g_1^{s_1s_2}(x,u,z)dxdz, \tag{5}$$

$$g_1^{s_1s_2}(x,u,z) = [e^{iu(s_1-s_2)} + e^{ius_1} - 2e^{i(us_1+zs_2)} + e^{-ius_2} + 1 - 2e^{izs_2} -$$

$$2e^{ix(s_1-s_2)-i(u-x)s_2} - 2e^{ixs_1} + 4e^{ixs_1+izs_2}]\cos\left(\frac{i(s_1-s_2)u}{2}\right)e^{\frac{i(s_1-s_2)u}{2}}, \tag{6}$$

$$h^{s_1s_2}(u)=\int_R f(x)f(u-x)h_1^{s_1s_2}(x,u)dx\,, \tag{7}$$

$$\begin{aligned}h_1^{s_1s_2}(u)=[&2e^{iu(s_1-s_2)}+2e^{ius_1}+2e^{-ius_2}+\\&+2-4e^{ix(s_1-s_2)+i(u-x)s_1}-4e^{-ixs_2}\\&-4e^{ix(s_1-s_2)+i(u-x)s_2}-4e^{ixs_1}+4e^{ix(s_1-s_2)}+\\&+4e^{ixs_1-i(u-x)s_2}]\cos\left(\frac{(T-s_1)u}{2}\right)e^{\frac{i(s_1-s_2)u}{2}}\,,\end{aligned} \tag{8}$$

oraz $s_1,s_2\in[0,T)$.

Dowód: Posługując się własnościami wartości oczekiwanej otrzymujemy:

$$E(2\bar{\gamma}(s))=\frac{1}{T-h}\int_0^{T-h}E(X(s+h)-X(s))^2ds=V2\gamma(s)\frac{1}{T-h}\int_0^{T-h}ds=2\gamma(s).$$

Pierwsza równość w Twierdzeniu 1 jest udowodniona.

Będziemy dowodzić równość (2).

Posługując się definicją kowariancji i jawną postacią oceny (1) otrzymujemy:

$$\operatorname{cov}\{2\bar{\gamma}(s_1),2\bar{\gamma}(s_2)\}=E2\bar{\gamma}(s_1)2\bar{\gamma}(s_2)-E\bar{\gamma}(s_1)E2\bar{\gamma}(s_2)=$$

$$=E\left[\frac{1}{T-s_1}\int_0^{T-s_1}(X(s+s_1)-X(s)]^2ds\frac{1}{T-s_2}\int_0^{T-s_2}(X(t+s_2)-X(t))^2dt\right]-$$

$$-E[\frac{1}{T-s_1}\int_0^{T-s_1}(X(s+s_1)-X(s))^2ds]E[\frac{1}{T-s_2}\int_0^{T-s_2}[X(t+s_2)-X(t)]^2dt]=$$

$$=\frac{1}{T-s_1}\frac{1}{T-s_2}\int_0^{T-s_1}\int_0^{T-s_2}[E(X(s+s_1)-X(s))^2(X(t+s_2)-X(t))^2-E(X(s+s_1)-X(s)^2\times$$

$$\times E(X(t+s_2)-X(s_2))^2]dsdt=\frac{1}{T-s_1}\frac{1}{T-s_2}\int_0^{T-s_1}\int_0^{T-s_2}[E(X(s+s_1))^2(X(t+s_2))^2+$$

$$+E(X(s+s_1))^2(X(t))^2-2E(X(s+s_1))^2X(t+s_2)X(t)+E(X(s))^2(X(t+s_2))^2+$$

$$+E((X(s))^2(X(t))^2-2EX(t+s_2)(X(s))^2X(t)-2EX(s+s_1)X(s)(X(t+s_2))^2-$$

$$-2EX(s+s_1)X(s)(X(t))^2+4EX(s+s_1)X(s)X(t)X(t+s_2)-E(X(s+s_1))^2E(X(t+s_2))^2$$

$$-E(X(s+s_1))^2E(X(t))^2+2E(X(s+s_1))^2EX(t+s_2)X(t)-E(X(s))^2E(X(t+s_2))^2-$$
$$-E(X(s))^2E(X(t))^2+2E(X(s))^2EX(t+s_2)X(t)+2EX(s+s_1)X(s)E(X(t+s_2))^2+$$
$$+2EX(s+s_1)X(s)E(X(t))^2-4EX(s+s_1)X(s)EX(t+s_2)X(t)]dsdt\,.$$

Z określenia zmieszanych momentów drugiego i czwartego stopnia mamy poniższy związek:

$$m_4(s_1,s_2,s_3,s_4)=c_4(s_1,s_2,s_3,s_4)+c_2(s_1,s_2)c_2(s_3,s_4)+c_2(s_1,s_3)c_2(s_2,s_4)+$$
$$+c_2(s_1,s_4)c_2(s_2,s_3)$$

Biorąc dodatkowo pod uwagę stacjonarność procesu X(s) mamy:

$$\operatorname{cov}\{2\tilde{\gamma}(s_1),2\tilde{\gamma}(s_2)\}=$$
$$\frac{1}{T-s_1}\frac{1}{T-s_2}\int_0^{T-s_1}\int_0^{T-s_2}[c_4(s-t+s_1-s_2,s-t+s_1-s_2,0)+c_4(s-t+s_1,s-t+s_1,0)-$$
$$-2c_4(s-t+s_1,s-t+s_1,s_2)+c_4(s-t-s_2,s-t-s_2,0)+c_4(s-t,s-t,0)-$$
$$-2c_4(s-t,s-t,s_2)-2c_4(s-t+s_1-s_2,s-t-s_2,0)-2c_4(s-t+s_1,s-t,0)+$$
$$+4c_4(s-t+s_1,s-t,s_2)+2R^2(s-t+s_1-s_2)+2R^2(s-t+s_1)+2R^2(s-t-s_2)+$$
$$+2R^2(s-t)-4R(s-t+s_1-s_2)R(s-t+s_1)-4R(s-t-s_2)R(s-t)-$$
$$-4R(s-t+s_1-s_2)R(s-t-s_2)-4R(s-t+s_1)R(s-t)+4R(s-t+s_1-s_2)R(s-t)+$$
$$+4R(s-t+s_1)R(s-t-s_2)]dsdt.$$

Z pracy [3] wiedząc, że

$$c_4(a,b,c)=\iiint_{R^3}f_4(x,y,z)e^{i(xa+yb+zc)}dxdydz\,,$$

oraz

$$R(a)=\int_R f(x)e^{ixa}dx,\quad a,b,c\in R\,,$$

otrzymujemy:

$$\operatorname{cov}\{2\tilde{\gamma}(s_1),2\tilde{\gamma}(s_2)\}=$$
$$\frac{1}{T-s_1}\frac{1}{T-s_2}\int_0^{T-s_1}\int_0^{T-s_2}[\iiint_{R^3}f_4(x,y,z)[e^{i(x+y)(s-t+s_1-s_2)}+e^{i(x+y)(s-t+s_1)}-$$
$$-2e^{i(x+y)(s-t+s_1)+izs_2}+e^{i(x+y)(s-t-s_2)}+e^{i(x+y)(s-t)}-2e^{i(x+y)(s-t)+izs_2}-$$

$$2e^{ix(s-t+s_1-s_2)+iy(s-t-s_2)} - 2e^{ix(s-t+s_1)+iy(s-t)} + 4e^{ix(s-t+s_1)+iy(s-t)+izs_2}]dxdydz +$$

$$+ \iint_{R^2} f(x)f(y) + [2e^{i(x+y)(s-t+s_1-s_2)} + 2e^{i(x+y)(s-t+s_1)} + 2e^{i(x+y)(s-t-s_2)} +$$

$$+ 2e^{i(x+y)(s-t)} - 4e^{ix(s-t+s_1-s_2)+iy(s-t+s_1)} - 4e^{ix(s-t-s_2)+iy(s-t)} - 4e^{ix(s-t+s_1-s_2)+iy(s-t-s_2)} +$$

$$- 4e^{ix(s-t+s_1)+iy(s-t)} + 4e^{i(s-t+s_1-s_2)+iy(s-t)} + 4e^{ix(s-t+s_1)+iy(s-t-s_2)}]dxdy]dsdt$$

Zmieniając porządek całkowania oraz wyciągając wyrażenie $e^{is(x+y)}e^{-it(x+y)}$ przed nawias, otrzymujemy:

$$\mathrm{cov}\{2\hat{\gamma}(s_1), 2\hat{\gamma}(s_2)\} =$$

$$\frac{1}{T-s_1}\frac{1}{T-s_2}\Bigg[\iiint_{R^3} f_4(x,y,z)\Bigg[\ [e^{i(x+y)(s_1-s_2)} + e^{i(x+y)s_1} - 2e^{i((x+y)s_1+zs_2)} + e^{-i(x+y)s_2} +$$

$$+1 - 2e^{izs_2} - 2e^{ix(s_1-s_2)-iys_2} - 2e^{ixs_1} + 4e^{ixs_1+izs_2}\Bigg]\int_0^{T-s_1}\int_0^{T-s_2} e^{is(x+y)}e^{-it(x+y)}dsdt\ \Bigg]dxdydz +$$

$$+ \iint_{R^2} f(x)f(y)\Bigg[\ [2e^{i(x+y)(s_1-s_2)} + 2e^{i(x+y)s_1} + 2e^{-i(x+y)s_2} + 2 - 4e^{ix(s_1-s_2)+iys_1} - 4e^{-ixs_2} +$$

$$-4e^{ix(s_1-s_2)-iys_2} - 4e^{ixs_1} + 4e^{ix(s_1-s_2)} + 4e^{ixs_1-iys_2}]\int_0^{T-s_1}\int_0^{T-s_2} e^{is(x+y)}e^{-it(x+y)}dsdt\ \Bigg]dxdydz\ \Bigg]$$

Stosując podstawienie:

$$\begin{cases} x = x \\ x + y = u \\ z = z \end{cases}$$

a następnie oznaczając wyrażenia w małych nawiasach kwadratowych odpowiednio $g_2^{s_1,s_2}(x,u,z)$ i $h_2^{s_1,s_2}(x,u)$, otrzymujemy:

$$\mathrm{cov}\{2\hat{\gamma}(s_1), 2\hat{\gamma}(s_2)\} =$$

$$\frac{1}{T-s_1}\frac{1}{T-s_2}\Bigg[\iiint_{R^3} f_4(x,u-x,z)\Bigg[\ g_2^{s_1,s_2}(x,u,z)\int_0^{T-s_1}\int_0^{T-s_2} e^{isu}e^{-itu}dsdt\ \Bigg]dxdudz +$$

$$+ \iint_{R^2} f(x)f(u-x)\Bigg[\ h_2^{s_1,s_2}(x,u)\int_0^{T-s_1}\int_0^{T-s_2} e^{isu}e^{-itu}dsdt\ \Bigg]dxdu\ \Bigg]$$

Ponieważ:

$$\int_0^T e^{ixs}ds = \frac{\sin\frac{xT}{2}}{\frac{x}{2}} e^{\frac{ixT}{2}}$$

mamy

$$\int_0^{T-s_1}\int_0^{T-s_2} e^{ius}e^{-uti}dsdt = \Delta_{T-s_1}(u)\Delta_{T-s_2}(u)e^{\frac{i(s_2-s_1)u}{2}},$$

gdzie $\Delta_T(x)$ zadane jest wzorem (4).
Posługując się elementarnym wzorem, otrzymujemy:

$$\sin\frac{(T-s_2)u}{2} = \sin\frac{(T-s_1)u}{2}\cos\frac{(s_1-s_2)u}{2} + \cos\frac{(T-s_1)u}{2}\sin\frac{(s_1-s_2)u}{2},$$

a więc

$$\Delta_{T-s_1}(u)\Delta_{T-s_2}(u) = \Delta^2{}_{T-s_1}(u)\cos\frac{(s_1-s_2)u}{2} + \Delta_{T-s_1}(u)\Delta_{s_1-s_2}(u)\cos\frac{(T-s_1)u}{2}.$$

Stosując powyższe mamy:

$$\operatorname{cov}\{2\gamma(s_1),2\gamma(s_2)\} =$$

$$\frac{1}{T-s_1}\frac{1}{T-s_2}\Bigg[\iiint_{R^3} f_4(x,u-x,z)\Bigg[g_2^{s_1,s_2}(x,u,z)\cdot(\Delta^2{}_{T-s_1}(u)\cos\frac{(s_1-s_2)u}{2}e^{\frac{i(s_2-s_1)u}{2}} +$$

$$+\Delta_{T-s_1}(u)\Delta_{s_1-s_2}(u)\cos\frac{(T-s_1)u}{2}e^{\frac{i(s_2-s_1)u}{2}})\Bigg]dxdudz +$$

$$+\iint_{R^2} f(x)f(u-x)\Bigg[h_2^{s_1,s_2}(x,u)\cdot(\Delta^2{}_{T-s_1}(u)\cos\frac{(s_1-s_2)u}{2}e^{\frac{i(s_2-s_1)u}{2}} +$$

$$+\Delta_{T-s_1}(u)\Delta_{s_1-s_2}(u)\cos\frac{(T-s_1)u}{2}e^{\frac{i(s_2-s_1)u}{2}})\Bigg]dxdu\Bigg] =$$

$$= \frac{1}{T-s_1}\frac{1}{T-s_2}\Bigg[\iiint_{R^3} f_4(x,u-x,z)\Bigg[g_2^{s_1,s_2}(x,u,z)\cdot$$

$$\cdot\Delta^2{}_{T-s_1}(u)\cos\frac{(s_1-s_2)u}{2}e^{\frac{i(s_2-s_1)u}{2}}\Bigg]dxdudz + \iint_{R^2} f(x)f(u-x)$$

$$+\Bigg[h_2^{s_1,s_2}(x,u)\cdot\Delta^2{}_{T-s_1}(u)\cos\frac{(s_1-s_2)u}{2}e^{\frac{i(s_2-s_1)u}{2}}\Bigg]dxdu +$$

$$+\iiint_{R^3} f_4(x,u-x,z)\left[\ g_2^{s_1,s_2}(x,u,z)\cdot\Delta_{T-s_1}(u)\Delta_{s_1-s_2}(u)\cos\frac{(T-s_1)u}{2}e^{\frac{i(s_2-s_1)u}{2}}\ \right]dxdudz+$$

$$+\iint_{R^2} f(x)f(u-x)\left[\ h_2^{s_1,s_2}(x,u)\cdot\Delta_{T-s_1}(u)\Delta_{s_1-s_2}(u)\cos\frac{(T-s_1)u}{2}e^{\frac{i(s_2-s_1)u}{2}}\ \right]dxdu\ \Bigg]$$

Wprowadzając oznaczenia (3)-(8) otrzymujemy wyrażenie (2) co kończy dowód.

Twierdzenie 2.

$$(T-s_1)V2\tilde{\gamma}(s_1)=4\pi\int_R \Psi_T(u)(g^0(u)+h^0(u))du\,, \tag{9}$$

gdzie

$$\begin{aligned} g^{s_1}(u) = &\iint_{R^2} f_4(x,u-x,z)[1+\cos(us_1)-e^{i(us_1+zs_1)}- \\ &-e^{izs_1}-e^{-i(u-x)s_1}-e^{ixs_1}+2e^{ixs_1+izs_1}]dxdy\,, \end{aligned} \tag{10}$$

$$\begin{aligned} h^{s_1}(u) = &\int_R f(x)f(u-x)[4+2\cos(us_1)-4\cos(xs_1)-2e^{i(u-x)s_1}- \\ &-2e^{-i(u-x)s_1}+2e^{ixs_1-i(u-x)s_1}]dx\quad. \end{aligned} \tag{11}$$

Dowód: Wzór (9) wynika z Twierdzenia 1, kiedy s_1=s_2.

Losowy proces $X(s), s\in R$ będziemy nazywać wewnętrznie stacjonarnym, jeśli $E(X(s_1)-X(s_2))=0$ oraz $V(X(s_1)-X(s_2))=2\gamma(s_1-s_2)$ gdzie $2\gamma(s_1-s_2)$ jest wariogramem procesu $X(s), s\in R$.
Własności wewnętrznie stacjonarnego procesu rozpatrywane są w pracy [5].

Bibliografia

1. Cressie N. *Statistics for Spatial Data*. New York 1991.
2. Н. Н. Труш, Т. Цеховая *Исследование статистических свойств оценок вариограммы и ковариационной функции* Вести НАН Беларуси 2001.
3. Н. Н. Труш *Асимтотические методы статистического анализа временных рядов*. Минск БГУ 1999.
4. Troush N. N., Tsekhovaya T.V. *Asymptotic Distribution of The Mutual Variogram Estimate*, Polish-Czech Mathematical School. Poraj, 2003.
5. Т. Цеховая, I. Arcimowicz *Анализ внутренные стационарных случайных процессов.* Брест, 2004.
6. Т. Андерсон *Статистический анализ временных рядов* Мир, 1976.