СВОЙСТВА ВАРИОГРАММЫ ВНУТРЕННЕ СТАЦИОНАРНЫХ СЛУЧАЙНЫХ ПРОЦЕССОВ

Т. В. Цеховая

Белорусский государственный университет
г. Минск, Беларусь
Tsekhavaya@bsu.by

В данной статье вводятся понятия внутренне стационарных случайных процессов, вариограммы. Вариограмма играет важную роль при анализе временных рядов в экономике, финансах, геологии, экологии и многих других областях человеческой деятельности. Возникает задача исследования ее свойств.

Ключевые слова: вариограмма, внутренне стационарный случайный процесс.

В середине XX в. появилось новое направление статистического анализа временных рядов (случайных процессов) – вариограммный анализ. Стало возможным более глубокое и полное изучение явлений, встречающихся в различных областях человеческой деятельности, поскольку современные методы вариограммного анализа используют информацию о внутренней структуре экспериментальных данных и, следовательно, охватывают всю сложность изучаемых процессов. При этом, как правило, основной моделью исследуемых событий являются стационарные в широком смысле и внутренне стационарные случайные процессы.

Случайный процесс $Y(s)$, $s \in Z = \{0, \pm 1, \pm 2, \ldots\}$ $(R = (-\infty, +\infty))$, называется внутренне стационарным, если справедливы следующие равенства:

$$M[Y(s_1) - Y(s_2)] = 0,$$

$$D[Y(s_1) - Y(s_2)] = 2\gamma(s_1 - s_2),$$

где $2\gamma(s_1 - s_2)$ – вариограмма рассматриваемого процесса, $s_1, s_2 \in Z(R)$.

Заметим, что функция $\gamma(s)$, $s \in Z(R)$, – называется семивариограммой. Нетрудно видеть, что $\gamma(s) \geq 0$, $\gamma(-s) = \gamma(s)$, $s \in Z(R)$, и $\gamma(0) = 0$.

Основоположником вариограммного анализа временных рядов является G. Matheron [1]. Дальнейшее развитие теоретических исследований по этой тематике продолжили ученые Д. Е. Мейерс [2], Дж. С. Дэвис [3], N. Cressie [4] и другие.

На практике часто оказывается, что процессы не обладают конечной дисперсией. В таких ситуациях делается предположение об их внутренней стационарности, и применяются методы вариограммного анализа, что позволяет существенно расширить круг решаемых проблем.

Вариограмма является основной характеристикой внутренне стационарных случайных процессов. Исследованию ее свойств посвящены работы А. Н. Колмогорова [5], А. С. Монина, А. М. Яглома [6], М. Давид [7], Н. Н. Труша, Т. В. Цеховой [8,9] и других. В данной статье приводятся некоторые свойства семивариограммы внутренне стационарных случайных процессов с дискретным и непрерывным временем.

Теорема 1. Для того чтобы функция $\gamma(s)$, $s\in Z(R)$, была семивариограммой некоторого случайного процесса $Y(s)$, $s \in Z(R)$, необходимо и достаточно, чтобы она являлась четной и условно отрицательно определенной функцией, т. е. для любого действительного ненулевого вектора $(a_1,...,a_n)$, такого, что $\sum_{i=1}^{n} a_i = 0$, было справедливо неравенство

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i - s_j) \le 0, \qquad (1)$$

$s_i \in Z(R)$, $i = \overline{1,n}$.

Доказательство. Докажем сначала необходимость условия (1). Из определения вариограммы очевидно, что функция $\gamma(s)$, $s\in Z(R)$, является четной.

Применяя понятие условно отрицательно определенной функции [4, с. 60], определение вариограммы, свойства математического ожидания, имеем

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i - s_j) = M\sum_{i,j=1}^{n} a_i a_j \frac{1}{2}(Y(s_i) - Y(s_j))^2 =$$

$$= \frac{1}{2}M\left[\sum_{i=1}^{n} a_i Y^2(s_i)\sum_{j=1}^{n} a_j - 2\sum_{i=1}^{n} a_i Y(s_i)\sum_{j=1}^{n} a_j Y(s_j) + \sum_{i=1}^{n} a_i \sum_{j=1}^{n} a_j Y^2(s_j)\right].$$

Поскольку $\sum_{i=1}^{n} a_i = 0$, то

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i - s_j) = -M\left[\sum_{i=1}^{n} a_i Y(s_i)\right]^2 \le 0,$$

что и требовалось доказать.

Обратно, пусть выполняется неравенство (1). Достаточность утверждения означает, что существует такой случайный процесс $Y(s)$, $s\in Z(R)$, что $\gamma(s_i - s_j) = \frac{1}{2}M(Y(s_i) - Y(s_j))^2$, $s_i \in Z(R)$, $i = \overline{1,n}$.

Обозначим

$$Q(s_i - s_j) = -\sum_{i,j=1}^{n} a_i a_j \gamma(s_i - s_j) \ge 0.$$

Из монографии М. Лоэва [10, с. 488] следует, что функция $e^{-Q(s_i - s_j)/2}$ является характеристической функцией n гауссовских случайных величин с нулевым математическим ожиданием и ковариацией $MY(s_i)Y(s_j) = -\gamma(s_i - s_j)$, $s_i \in Z(R)$, $i = \overline{1,n}$. Тогда, учитывая условие $\sum_{i=1}^{n} a_i = 0$ и определение вариограммы, запишем

$$\frac{1}{2}\sum_{i=1}^{n} a_i \sum_{j=1}^{n} a_j MY^2(s_j) - \sum_{i,j=1}^{n} a_i a_j MY(s_i)Y(s_j) + \frac{1}{2}\sum_{j=1}^{n} a_j \sum_{i=1}^{n} a_i MY^2(s_i) =$$

$$= \sum_{i,j=1}^{n} a_i a_j \frac{1}{2}M(Y(s_i) - Y(s_j))^2 = \sum_{i,j=1}^{n} a_i a_j \gamma(s_i - s_j).$$

Таким образом, характеристическая функция $e^{-Q(s_i-s_j)/2}$ полностью определяет закон распределения гауссовского случайного процесса $Y(s)$ для которого $\gamma(s)$ является семивариограммой. Теорема доказана.

Теорема 2. Пусть $\gamma_1(s)$, $\gamma_2(s)$, $s\in Z(R)$, – семивариограммы стационарных случайных процессов $Y_1(s)$, $Y_2(s)$, $s\in Z(R)$, соответственно. Тогда функция $\gamma(s)=$ $=\gamma_1(s)+\gamma_2(s)$, $s\in Z(R)$, также является семивариограммой некоторого стационарного случайного процесса.

Доказательство. Для доказательства теоремы 2 достаточно показать, что функция $\gamma(s)$, $s\in Z(R)$, является четной и условно отрицательно определенной.

Очевидно, что $\gamma(s)=\gamma(-s)$, $s\in Z(R)$. Используя определение вариограммы, свойства математического ожидания, имеем

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i-s_j)=\frac{1}{2}\sum_{i,j=1}^{n} a_i a_j\left[M(Y_1(s_i)-Y_1(s_j))^2+M(Y_2(s_i)-Y_2(s_j))^2\right]=$$

$$=-M\left[\sum_{i=1}^{n} a_i Y_1(s_i)\right]^2-M\left[\sum_{i=1}^{n} a_i Y_2(s_i)\right]^2\le 0.$$

Теорема доказана.

Теорема 3. Пусть $\gamma(s)$, $s\in Z(R)$,– семивариограмма стационарного случайного процесса $Y(s)$, $s\in Z(R)$. Тогда для любого $b>0$ функция $b\gamma(s)$, $s\in Z(R)$, является семивариограммой стационарного случайного процесса $bY(s)$, $s\in Z(R)$.

Доказательство. Нетрудно видеть, что $b\gamma(s)=b\gamma(-s)$, $s\in Z(R)$, и справедливо неравенство

$$b\sum_{i,j=1}^{n} a_i a_j \gamma(s_i-s_j)\le 0$$

для любого действительного ненулевого вектора $(a_1,\ldots,a_n)$, такого, что $\sum_{i=1}^{n} a_i=0$, $s_i\in Z(R)$, $i=\overline{1,n}$. Тогда в силу теоремы 1 получаем требуемый результат.

В работе [6] показано, что семивариограмма $\gamma(s)$, $s\in Z(R)$, удовлетворяет соотношению

$$\gamma(s)=\int_{\Pi(R)}(1-\cos\lambda s)dF(\lambda), \tag{2}$$

где $F(\lambda), \lambda\in\Pi(R)$, $\Pi=[-\pi,\pi]$, – спектральная функция процесса $Y(s)$, $s\in Z(R)$, удовлетворяющая условиям:

$$F(\lambda)\ge 0, \qquad \int_{\Pi(R)}\lambda^2 dF(\lambda)<\infty. \tag{3}$$

Это спектральное представление будет использовано для доказательства следующего результата.

Теорема 4. Пусть $\gamma_1(s)$, $\gamma_2(s)$, $s\in Z(R)$ – соответственно семивариограммы стационарных случайных процессов $Y_1(s)$, $Y_2(s)$, $s \in Z(R)$. Тогда функция $\gamma(s)=\gamma_1(s)\cdot\gamma_2(s)$, $s\in Z(R)$, также является семивариограммой.

Доказательство. Легко заметить, что функция $\gamma(s)$, $s\in Z(R)$, является четной.

Покажем, что $\gamma(s)$, $s\in Z(R)$, удовлетворяет условию (1) для любого действительного ненулевого вектора $(a_1,...,a_n)$, такого, что $\sum\limits_{i=1}^{n} a_i=0$. Используя спектральное представление семивариограммы (2), запишем

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i-s_j)=$$

$$=\sum_{i,j=1}^{n} a_i a_j \iint\limits_{\Pi^2(R^2)} (1-\cos\lambda_1(s_i-s_j))(1-\cos\lambda_2(s_i-s_j))dF_1(\lambda_1)dF_2(\lambda_2),$$

где $F_1(\lambda_1)$, $F_2(\lambda_2)$ – спектральные функции процессов $Y_1(s)$, $Y_2(s)$, соответственно, удовлетворяющие условиям (3), $\lambda_1,\lambda_2\in\Pi(R)$, $s\in Z(R)$.

Применяя элементарное равенство $\cos\lambda=\dfrac{e^{i\lambda}+e^{-i\lambda}}{2}$ и учитывая, что $\sum\limits_{i=1}^{n} a_i=0$, получим

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i-s_j)=-\iint\limits_{\Pi^2(R^2)}\left|\sum_{i=1}^{n} a_i e^{i\lambda_1 s_i}\right|^2 dF_1(\lambda_1)dF_2(\lambda_2)-$$

$$-\iint\limits_{\Pi^2(R^2)}\left|\sum_{i=1}^{n} a_i e^{i\lambda_2 s_i}\right|^2 dF_1(\lambda_1)dF_2(\lambda_2)+\frac{1}{2}\iint\limits_{\Pi^2(R^2)}\left|\sum_{i=1}^{n} a_i e^{i(\lambda_1+\lambda_2)s_i}\right|^2 dF_1(\lambda_1)dF_2(\lambda_2)+$$

$$+\frac{1}{2}\iint\limits_{\Pi^2(R^2)}\left|\sum_{i=1}^{n} a_i e^{i(\lambda_1-\lambda_2)s_i}\right|^2 dF_1(\lambda_1)dF_2(\lambda_2).$$

В третьем слагаемом правой части последнего равенства сделаем замену переменных интегрирования $\lambda_1=\mu$, $\lambda_1+\lambda_2=\lambda$, а в четвертом – $\lambda_1=\mu$, $\lambda_1-\lambda_2=\lambda$. Тогда,

$$\sum_{i,j=1}^{n} a_i a_j \gamma(s_i-s_j)=-\iint\limits_{\Pi^2(R^2)}\left|\sum_{i=1}^{n} a_i e^{i\lambda_2 s_i}\right|^2 dF_1(\lambda_1)dF_2(\lambda_2)\le 0.$$

В силу теоремы 1 получаем требуемый результат.

Далее рассмотрим некоторые примеры вариограмм $2\gamma(s)$, $s\in R_+=[0,+\infty)$, и исследуем их поведение при $s\to 0$ для процессов $Y(s)$, $s\in R_+$.

Степенная функция $2\gamma(s)=\omega^2 s^\lambda$, $0\le\lambda\le 2$, $s\in R_+$, определяет важный класс вариограмм случайных процессов, для которых $\omega^2=M(Y(s+1)-Y(s))^2$, $s\in R_+$. Если $\lambda=0$, то соответствующий случайный процесс – нормированный процесс «белый шум», если $\lambda=1$ – винеровский процесс. Случайному процессу вида $Y(s)=Y_1 s+Y_0$, где $s\in R_+$, Y_1, Y_0 – некоторые случайные величины [6], соответствует $\lambda=2$. Случай дробного параметра λ, $0<\lambda<2$, рассматривался Мониным, Ягломом [6] для анализа локальной изменчивости изучаемого процесса.

По поведению вариограммы в окрестности нуля можно судить о свойствах рассматриваемого процесса.

Л е м м а. Пусть вариограмма $2\gamma(s)$, $s\in R_+$, стационарного случайного процесса $Y(s)$, $s\in R_+$, $Y(0)=0$, убывает при $s\to 0$, как s^{λ}, $\lambda \geq 0$. Тогда параметр λ определяет характер траектории процесса:

1) при $\lambda = 0$ рассматриваемый процесс всюду разрывный;

2) при $0 < \lambda < 2$ процесс $Y(s)$ непрерывный недифференцируемый в среднеквадратическом смысле;

3) при $\lambda = 2$ рассматриваемый процесс непрерывный дифференцируемый в среднеквадратическом смысле;

4) при $\lambda > 2$ процесс не существует.

Доказательство. Пусть t_1, t_2 – длины двух непересекающихся отрезков, расположенных на расстоянии ε друг от друга. Предположим далее, что левый конец интервала длиной t_1 совмещен с точкой $s=0$.

Рассмотрим ковариацию приращений процесса на этих отрезках.

$$\operatorname{cov}\{Y(t_1)-Y(0), Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon)\} =$$
$$= M\{(Y(t_1)-Y(0))-M(Y(t_1)-Y(0))\}\times$$
$$\times\{(Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon))-M(Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon))\}.$$

Учитывая стационарность рассматриваемого процесса, правую часть последнего равенства запишем

$$M\{(Y(t_1)-Y(0))\}\{(Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon))\} =$$
$$= M\{Y(t_1)(Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon))\}.$$

Найдем коэффициент корреляции приращений процесса:

$$\rho(t_1,t_2)=\frac{\operatorname{cov}\{Y(t_1)-Y(0), Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon)\}}{\sqrt{M(Y(t_1)-Y(0))^2}\sqrt{M(Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon))^2}}.$$

Используя выражение для ковариации приращений, полученное выше, определение вариограммы, имеем

$$\rho(t_1,t_2)=\frac{M\,Y(t_1)[Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon)]}{\sqrt{M[Y(t_1)]^2\,M[Y(t_1+\varepsilon+t_2)-Y(t_1+\varepsilon)]^2}}=$$
$$=\frac{\operatorname{cov}\{Y(t_1),Y(t_1+\varepsilon+t_2)\}-\operatorname{cov}\{Y(t_1),Y(t_1+\varepsilon)\}}{2\sqrt{\gamma(t_1)\gamma(t_2)}}=$$
$$=\frac{\gamma(\varepsilon)-\gamma(t_1+\varepsilon)-\gamma(t_2+\varepsilon)+\gamma(t_1+\varepsilon+t_2)}{2\sqrt{\gamma(t_1)\gamma(t_2)}}.$$

Условие $\lim\limits_{t_1,t_2\to 0}\rho(t_1,t_2)\to 1$ является необходимым и достаточным для дифференцируемости процесса в среднеквадратическом смысле [1]. Полагая $\varepsilon = 0$ и учитывая тот факт, что $2\gamma(s)\sim s^{\lambda}$ при малых значениях аргумента, т. е. $\lim\limits_{s\to 0}(2\gamma(s)-s^{\lambda})=0$, получим

$$\lim_{t_1,t_2\to 0}\frac{(t_1+t_2)^{\lambda}-t_1^{\lambda}-t_2^{\lambda}}{2\sqrt{t_1^{\lambda}t_2^{\lambda}}}\to 1.$$

Пусть $t_1 = t_2 = t$, тогда $\lim\limits_{t\to 0}\rho(t,t) = 2^{\lambda-1} - 1$. Этот предел будет равен единице при $\lambda = 2$, откуда следует условие дифференцируемости процесса.

Поскольку $|\rho(t,t)| \le 1$, то должно выполняться неравенство $|2^{\lambda-1} - 1| \le 1$. Оно справедливо при $0 < \lambda \le 2$. Отсюда вытекает невозможность существования случайного процесса при $\lambda > 2$.

Рассмотрим $\lim\limits_{s_2\to s_1} 2\gamma(s_1 - s_2)$ для произвольных $s_1, s_2 \in R_+$ и параметра λ, $0 \le \lambda \le 2$. Этот предел равен нулю при $0 < \lambda \le 2$, т. е. $Y(s)$, $s \in R_+$, – непрерывный в среднеквадратическом смысле процесс. Если $\lambda = 0$, то предел больше нуля и рассматриваемый процесс является всюду разрывным. Лемма доказана.

ЛИТЕРАТУРА

1. *Matheron G.* Principles of Geostatistics // Economic Geology. 1963. Vol. 58. P. 1246–1266.
2. *Мейерс Д. Е.* Перекрестное подтверждение и оценивание вариограмм // Теория вероятностей и ее применение. 1992. Т. 37(2). С. 377–380.
3. *Дэвис Дж. С.* Статистический анализ данных в геологии. М., 1990. Т. 1. 318 с.
4. *Cressie N.* Statistics for Spatial Data. N. Y., Wiley. 1991. 900 p.
5. *Колмогоров А. Н.* Локальная структура турбулентности в несжимаемой вязкой жидкости при очень больших числах Рейнольдса // Докл. АН СССР. 1941. Т. 30(4). С. 299–303.
6. *Монин А. С., Яглом А. М.* Статистическая гидромеханика. Механика турбулентности: В 2 т. М., 1967. Т. 2. 720 с.
7. *Давид М.* Геостатистические методы при оценке запасов руд. Л., 1980. 360 с.
8. *Труш Н. Н., Цеховая Т. В.* Исследование статистических свойств оценок вариограммы и ковариационной функции // Весці НАН Беларусі . Сер. фіз.-мат. навук. 2001. № 2. С. 24–29.
9. *Troush N. N., Tsekhovaya T. V.* Estimation of the variogram // Remote sensing techniques for subsurface resources exploration and management: 13-th International Symposium, Damascus, Siria, 9–12 december, 2002. The international and arab cooperation in remote sensing and space sciences. Damascus, 2002.
10. *Лоэв М.* Теория вероятностей. М., 1962. 719 с.
11. *Леви П.* Стохастические процессы и броуновское движение. М., 1972. 375 с.