УДК 519.21

Ю. В. Козаченко, Т. О. Яковенко (КНУ імені Тараса Шевченка)

КРИТЕРІЙ ПЕРЕВІРКИ ГІПОТЕЗИ ПРО ВИГЛЯД КОВАРІАЦІЙНОЇ ФУНКЦІЇ СТАЦІОНАРНОЇ ГАУСОВОЇ ВИПАДКОВОЇ ПОСЛІДОВНОСТІ

In this paper the criterion for checking a hypotheses on the covariance function of the Gaussian and stationary in wide sense random sequence was constructed. The critical apparatus of the theory of square-Gaussian random variables was used for the construction.

В роботі був побудований критерій для перевірки гіпотези про вигляд коваріаційної функції гаусової, стаціонарної в широкому сенсі випадкової послідовності. Для цього було використано математичний апарат теорії квадратично гаусових випадкових величин.

1. Вступ

Задача дослідження властивостей випадкових послідовностей є дуже важливою з точки зору практичних застосувань. Часто деякі явища ми можемо спостерігати лише в певні моменти часу, а при спостереженні неперервних величин, таких як, наприклад, температура повітря чи напруга, їх значення фіксуються в дискретні моменти часу. І навіть в тих випадках, коли значення можна записувати неперервно, для комп'ютерної обробки ми можемо використовувати лише дані, що відповідають дискретній множині моментів часу. Тому на практиці ми в основному працюємо саме з випадковими послідовностями.

В даній роботі досліджено центровані гаусові стаціонарні в широкому сенсі випадкові послідовності. Для них побудовано критерій для перевірки гіпотези про вигляд коваріаційної функції. Статистичний аналіз випадкових послідовностей (числових рядів) наведено в книгах [1], [2]. Для побудови критерію було використано теорію квадратично гаусових випадкових величин. Простір квадратично гаусових випадкових величин було вперше введено в [3]. В книзі [4] було досліджено властивості цих просторів та встановлено їх зв'язок з просторами Орліча випадкових величин.

2. Квадратично гаусові випадкові величини

Нехай $(\Omega, \mathfrak{F}, \mathsf{P})$ – стандартний імовірнісний простір, на якому визначені сумісно гаусові центровані вектори $\overrightarrow{\gamma}^{\mathrm{T}} = (\gamma_1, ..., \gamma_r)$ довільної розмірності $r \geq 1$. Розглянемо сім'ю $SG_{(0)}(\Omega)$ випадкових величин, які можна придставити у вигляді $\overrightarrow{\gamma}^{\mathrm{T}} A \overrightarrow{\gamma}$, де A – це будь-яка дійснозначна матриця розмірності $r \times r$, причому r пробігає множину натуральних чисел. Як було показано в [4] сім'я $SG_{(0)}(\Omega)$ являє собою лінійний підпростір в просторі $L_2(\Omega)$. Тоді лінійне замикання цієї сім'ї в середньому квадратичному буде підпростором в просторі $L_2(\Omega)$.

Означення 1. [4] Простір $SG(\Omega)$ будемо називати простором квадратично гаусових випадкових величин якщо будь-який елемент $\xi \in SG(\Omega)$ можна подати у вигляді
$$\xi = \overrightarrow{\gamma}^{\mathrm{T}} A \overrightarrow{\gamma} - \mathsf{E}\overrightarrow{\gamma}^{\mathrm{T}} A \overrightarrow{\gamma}, \tag{1}$$
де $\overrightarrow{\gamma}^{\mathrm{T}} = (\gamma_1, ..., \gamma_r)$ – вектор сумісно гаусових центрованих випадкових величин, A – дійснозначна матриця; або елемент $\xi \in SG(\Omega)$ є границею у середньому

квадратичному послідовності випадкових величин $\{\xi_n, n \geq 1\}$ виду (1)

$$\xi = \mathsf{l}.i.m_{n\to\infty}\xi_n.$$

Для квадратично гаусових випадкових величин справедливі наступні теореми.

Теорема 1. *[4] Нехай $\xi \in SG(\Omega)$ та $\mathsf{D}\xi > 0$, тоді наступна нерівність виконуеться для всіх $|s| < 1$*

$$\mathsf{E}\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{\xi}{\sqrt{\mathsf{D}\xi}}\right\} \leq \frac{1}{\sqrt{1-|s|}}\exp\left\{-\frac{|s|}{2}\right\} \tag{2}$$

Теорема 2. *Нехай $\{\xi_m, 1 \leq m \leq M\}$ – квадратично гаусові випадкові величини, такі, що $\max_{1\leq m\leq M}\mathsf{D}(\xi_m) > 0$. Тоді, для $M \geq 2$, $K \geq \frac{2}{\ln M}$ та $x \geq 2+\sqrt{2}$*

$$\mathsf{P}\left\{\max_{1\leq m\leq M}|\xi_m| > x\max_{1\leq m\leq M}\mathsf{D}(\xi_m)K\ln M\right\} \leq 2M^{1-K}(1+x\sqrt{2})^{\frac{1}{2}}\exp\left\{-\frac{x}{\sqrt{2}}\right\}. \tag{3}$$

Доведення. Введемо позначення: $\eta := \max_{1\leq m\leq M}|\xi_m|$, $a := \max_{1\leq m\leq M}\mathsf{D}(\xi_m)$ та $u_M := K\ln M$. Тоді для всіх $0 \leq s < 1$

$$\begin{aligned} I_M := \mathsf{P}\{\eta > xau_M\} &= \mathsf{E}\mathbf{1}\{\omega : \eta > xau_M\} \leq \\ &\leq \sum_{m=1}^{M}\mathsf{E}\mathbf{1}\{\eta = |\xi_m|\}\cdot\mathsf{E}\mathbf{1}\{\omega : |\xi_m| > xau_M\} \leq \\ &\leq M\max_{1\leq m\leq M}\mathsf{E}\mathbf{1}\{\omega : |\xi_m| > xau_M\} \leq \\ &\leq M^{1-K}M^K\max_{1\leq m\leq M}\mathsf{E}\mathbf{1}\{\omega : |\xi_m| > xau_M\}\frac{\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{au_M}\right\}}{\exp\left\{\frac{s}{\sqrt{2}}x\right\}}. \end{aligned}$$

Для тих s, для яких

$$\frac{s}{\sqrt{2}}x \geq 2 \tag{4}$$

маємо $\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{au_M} \geq \frac{s}{\sqrt{2}}x \geq 2$. Наступну нерівність отримуємо скориставшись тим, що $y+z \leq yz$ для $y, z \geq 2$:

$$M^K\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{au_M}\right\} = \exp\left\{K\ln M + \frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{au_M}\right\} \leq \exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{a}\right\}.$$

Тоді

$$\begin{aligned} I_M &\leq M^{1-K}\max_{1\leq m\leq M}\mathsf{E}\mathbf{1}\{\omega : |\xi_m| > xau_M\}\frac{\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{a}\right\}}{\exp\left\{\frac{s}{\sqrt{2}}x\right\}} \leq \\ &\leq M^{1-K}\exp\left\{-\frac{s}{\sqrt{2}}x\right\}\max_{1\leq m\leq M}\mathsf{E}\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{a}\right\}. \end{aligned}$$

З нерівності (2) випливає, що для будь-якого m та $0 \leq s < 1$

$$\mathsf{E}\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{a}\right\} \leq \mathsf{E}\exp\left\{\frac{s}{\sqrt{2}}\cdot\frac{|\xi_m|}{\sqrt{\mathsf{D}\xi_m}}\right\} \leq \frac{2}{\sqrt{1-s}}\exp\left\{-\frac{s}{2}\right\},$$

а тому

$$I_M \leq 2M^{1-K} \exp\left\{-\frac{s}{\sqrt{2}}x\right\} \frac{1}{\sqrt{1-s}} \exp\left\{-\frac{s}{2}\right\} = \frac{2M^{1-K}}{\sqrt{1-s}} \exp\left\{-\frac{s}{\sqrt{2}}x - \frac{s}{2}\right\}. \quad (5)$$

Права частина нерівності (5) набуває свого мінімального значення

$$2M^{1-K}(1+x\sqrt{2})^{\frac{1}{2}} \exp\left\{-\frac{x}{\sqrt{2}}\right\}$$

при $s_{min} = 1 - \frac{1}{1+\sqrt{2}x}$.

Для значення s_{min} виконується умова (4) коли $x \geq 2 + \sqrt{2}$.

Наслідок 1. *Якщо при виконанні умов Теореми 2 у (3) вибрати деяке число ε, так що $\varepsilon \geq (2+\sqrt{2})aK \ln M$, де $a = \max_{1\leq m\leq M} \mathsf{D}(\xi_m)$, то*

$$\mathsf{P}\left\{\max_{1\leq m\leq M} |\xi_m| > \varepsilon\right\} \leq 2M^{1-K}\left(1 + \frac{\sqrt{2}\varepsilon}{aK\ln M}\right)^{\frac{1}{2}} \exp\left\{-\frac{\varepsilon}{\sqrt{2}aK\ln M}\right\}. \quad (6)$$

Доведення. Якщо покласти $\varepsilon = xaK\ln M$, тоді з того, що $x = \frac{\varepsilon}{aK\ln M} \geq 2 + \sqrt{2}$ випливає (6).

3. Коваріаційна функція стаціонарної гаусової послідовності

Розглянемо стаціонарну гаусову послідовність $\{\gamma_i, i \geq 1\}$ таку, що для всіх i $\mathsf{E}\gamma_i = 0$ та $\mathsf{E}\gamma_i\gamma_{i+m} = B(m)$, $m \geq 0$ – коваріаційна функція цієї послідовності.

В якості оцінки коваріаційної функції $B(m)$ розглянемо статистику

$$\widehat{B}_N(m) = \frac{1}{N-m}\sum_{i=1}^{N-m} \gamma_i\gamma_{i+m}, \quad N \geq 1.$$

Оцінка $\widehat{B}_N(m)$ є незміщеною:

$$\mathsf{E}\widehat{B}_N(m) = \frac{1}{N-m}\sum_{i=1}^{N-m} \mathsf{E}\gamma_i\gamma_{i+m} = B(m).$$

Тоді величина $\Delta_N(m) = \widehat{B}_N(m) - B(m)$ – буде квадратично гаусовою випадковою величиною, оскільки $\widehat{B}_N(m)$ можна записати як $(\gamma_1, ..., \gamma_{N-m})^{\mathrm{T}} A(\gamma_m, ..., \gamma_N)$, де матриця

$$A = \begin{pmatrix} \frac{1}{N-m} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{N-m} \end{pmatrix}.$$

Математичне сподівання $\Delta_N(m)$ дорівнює нулю та дисперсія

$$\begin{aligned} \mathsf{D}\Delta_N(m) &= \mathsf{E}(\widehat{B}_N(m) - B(m))^2 = \mathsf{E}\left[\frac{1}{N-m}\sum_{i=1}^{N-m}(\gamma_i\gamma_{i+m} - B(m))\right]^2 = \\ &= \frac{1}{(N-m)^2}\mathsf{E}\left[\sum_{i=1}^{N-m}\sum_{j=1}^{N-m}(\gamma_i\gamma_{i+m} - B(m))(\gamma_j\gamma_{j+m} - B(m))\right] = \\ &= \frac{1}{(N-m)^2}\sum_{i=1}^{N-m}\sum_{j=1}^{N-m}\mathsf{E}\gamma_i\gamma_{i+m}\gamma_j\gamma_{j+m} - B^2(m). \end{aligned}$$

За формулою Іссерліса для центрованих гаусових випадкових величин

$$\begin{aligned}\mathsf{E}\gamma_i\gamma_{i+m}\gamma_j\gamma_{j+m} &= \mathsf{E}\gamma_i\gamma_{i+m}\mathsf{E}\gamma_j\gamma_{j+m} + \mathsf{E}\gamma_i\gamma_j\mathsf{E}\gamma_{i+m}\gamma_{j+m} + \mathsf{E}\gamma_i\gamma_{j+m}\mathsf{E}\gamma_{i+m}\gamma_j = \\ &= B^2(m) + B^2(j-i) + B(j-i+m)B(j-i-m).\end{aligned}$$

Тоді

$$\begin{aligned}\mathsf{D}\Delta_N(m) &= \frac{1}{(N-m)^2}\sum_{i=1}^{N-m}\sum_{j=1}^{N-m}\left[B^2(j-i) + B(j-i+m)B(j-i-m)\right] = \\ &= \frac{2}{N-m}\sum_{k=0}^{N-m}\left(1-\frac{k}{N-m}\right)\left[B^2(k) + B(k+m)B(k-m)\right]\end{aligned}$$

Приклад 1. Для центрованої гаусової послідовності з коваріаційною функцією $B(m) = e^{-\lambda m}$, $m \geq 0$, $\lambda > 0$

$$\begin{aligned}\mathsf{D}\Delta_N(m) &= \frac{2}{N-m}\sum_{k=0}^{N-m}\left(1-\frac{k}{N-m}\right)\left[e^{-2\lambda k} + e^{-\lambda(k+m)}e^{-\lambda(k-m)}\right] = \\ &= \frac{4}{N-m}\sum_{k=0}^{N-m}\left(1-\frac{k}{N-m}\right)e^{-2\lambda k} = \\ &= \frac{4}{N-m}\left[\sum_{k=0}^{N-m}e^{-2\lambda k} - \frac{e^{-2\lambda}}{N-m}\sum_{k=1}^{N-m}ke^{-2\lambda(k-1)}\right] = \\ &= \frac{4}{N-m}\left[\frac{1-e^{-2\lambda(N-m)}}{1-e^{-2\lambda}} - \right. \\ &- \left.\frac{e^{-2\lambda}(1-(N-m)e^{-2\lambda(N-m-1)} + (N-m-1)e^{-2\lambda(N-m)})}{(N-m)(1-e^{-2\lambda})^2}\right] = \\ &= \frac{4}{(N-m)^2(1-e^{-2\lambda})^2}\left[(N-m) - (N-m+1)e^{-2\lambda} + e^{-2\lambda(N-m+1)}\right] = \\ &= O\left(\frac{1}{N-m}\right), \quad \text{при} \quad N \to \infty.\end{aligned}$$

4. Критерій перевірки гіпотези про вигляд коваріаційної функції

Припустимо, що ми маємо $\gamma_1, ..., \gamma_N$ – N послідовних спостережень деякої випадкової послідовності $\{\gamma_i, i \geq 1\}$ і маємо підстави припускати, що ця послідовність є центрованою, стаціонарною та гаусовою з коваріаційною функцією $B(m)$.

Критерій 1. *Нехай основна гіпотеза H_0 полягає в тому, що $B(m)$, $m \geq 0$ є коваріаційною функцією послідовності $\{\gamma_i, i \geq 1\}$, а альтернативна гіпотеза H_a – в тому, що $B(m)$ не є коваріаційною функцією цієї послідовності. Якщо для рівня значущості α та деяких M та K, таких, що $2 \leq M < N$, $K \geq \frac{2}{\ln M}$*

$$\max_{1\leq m\leq M}\Delta_N(m) > \varepsilon_\alpha,$$

де ε_α *– це значення, при якому*

$$2M^{1-K}\left(1+\frac{\sqrt{2}\varepsilon_\alpha}{\max\limits_{1\le m\le M}\mathsf{D}\Delta_N(m)K\ln M}\right)^{\frac{1}{2}}\exp\left\{-\frac{\varepsilon_\alpha}{\sqrt{2}\max\limits_{1\le m\le M}\mathsf{D}\Delta_N(m)K\ln M}\right\}=\alpha,$$

тоді гіпотеза H_0 *відхиляється, та не відхиляється в іншому випадку.*

Зауваження 1. Помилка першого роду для критерію 1, тобто, ймовірність того, що буде зроблено висновок про те, що $B(m)$ не може бути коваріаційною функцією послідовності $\{\gamma_i, i\ge 1\}$, хоча насправді це так, менша або дорівнює α.

Приклад 2. Припустимо, що залежність між елементами стаціонарної гаусової послідовності $\{\gamma_i, i\ge 1\}$ може бути добре описана за допомогою моделі ковзного середнього

$$\gamma_i=\sum_{k=0}^{\infty}\beta_k\zeta_{i-k},\quad i\ge 1$$

де $\beta_k=e^{-\lambda k}$, $k\ge 0$, $\lambda>0$ та $\{\zeta_k, k\in\mathbb{Z}\}$ – послідовність незалежних випадкових величин, таких, що для всіх k $\mathsf{E}\zeta_k=0$, $\mathsf{E}\zeta_k^2=1$. В цьому випадку $\mathsf{E}\gamma_i=0$ для всіх i та $B(m)=\mathsf{E}\gamma_i\gamma_{i+m}=\sum_{k=0}^{\infty}\beta_k\beta_{k+m}=\sum_{k=0}^{\infty}e^{-\lambda k}e^{-\lambda(k+m)}=\frac{e^{-\lambda m}}{1-e^{-2\lambda}}$. Тоді з прикладу 1 маємо, що

$$\begin{aligned}\max_{1\le m\le M}\mathsf{D}\Delta_N(m) &= \max_{1\le m\le M}\frac{4\left[(N-m)-(N-m+1)e^{-2\lambda}+e^{-2\lambda(N-m+1)}\right]}{(N-m)^2(1-e^{-2\lambda})^4}\le\\ &\le \max_{1\le m\le M}\frac{4}{(N-m)(1-e^{-2\lambda})^4}=\frac{4}{(N-M)(1-e^{-2\lambda})^4}.\end{aligned}$$

Якщо $\lambda=1$, $N=200$, $M=10$, тоді можна вибрати $K=1$. При рівні значущості $\alpha=0,1$ значення $\varepsilon_\alpha=0,8127$, а при рівні значущості $\alpha=0,05$ значення $\varepsilon_\alpha=0,9038$.

5. Висновки

В роботі було оцінено розподіл максимуму відхилень емпіричної оцінки коваріаційної функції стаціонарної гаусової послідовності від теоретичної коваріаційної функції. Це дало змогу побудувати критерій перевірки гіпотези про вигляд коваріаційної функції для цих послідовностей. Було розглянуто випадок, коли послідовність може бути зображена у вигляді ковзного середнього.

1. *Anderson, T. W.* The statistical analysis of time series. – New York, John Wiley & Sons, 1971. – 704 p.
2. *Brockwell,P. J., Davis,R. A.* Time Series: Theory and Methods. – New York, Springer Series in Statistics, Springer-Verlag, Second Edition, 2009. – 586 p.
3. *Kozachenko, Yu. V., Stus, O. V.* Square-Gaussian random processes and estimators of covariance functions// Math. Communications – 1998. – **3**, №1. – P. 83–94.
4. *Buldygin, V. V., Kozachenko, Yu. V.* Metric characterization of random variables and random processes. – Amer. Math. Soc., Providence, R I, 2000. – 257 p.

Одержано ..2010