О. Заболотний, І.Яворський*

Фізико-механічний інститут ім. Г. Карпенка НАН України м. Львів

*Інститут телекомунікації Технічно-сільськогосподарської академії м.Бидгощ, Польща.

УДК 621,77(088,8)

ВПЛИВ ДИСКРЕТИЗАЦІЇ НА ЯКІСТЬ ОЦІНКИ НАЙМЕНШИХ КВАДРАТІВ КОРЕЛЯЦІЙНОЇ ФУНКЦІЇ

Досліджено імовірнісні характеристики оцінки найменших квадратів кореляційної функції поліритмічного часового ряду. Отримано аналітичні вирази для її зміщення та дисперсії. Досліджено їхню поведінку залежно від величини кроку дискретизації. Зроблено рекомендації щодо вибору частоти дискретизації при дослідженні стохастичної складової поліритмічного сигналу методами майже періодично корельованих випадкових процесів.

The probabilistic characteristics of the least squares estimation of polyrhythmic signal correlation function are observed. The according analytical expressions of bias and variance are obtained. Their sample rate dependence is observed. The recommendations for sample rate selection in investigation of signal stochastic component by methods of almost periodically correlate random processes are given.

Вступ

Багатьом сигналам як штучного, так і природного походження властива стохастичність і періодична повторюваність [1,2]. Використання методів, що ґрунтуються на детермінованих чи стаціонарних моделях, для дослідження процесів, що породжують такі сигнали, є недостатньо обгрунтованим. Зазначенні властивості сигналів було покладено в основу розроблених моделей періодично корельованих випадкових процесів (ПКВП) та їх узагальнень майже ПКВП (МПКВП). Для оцінювання параметрів цих моделей використовують когерентний і компонентний методи та метод найменших квадратів [2,3]. Отже, стають актуальними дослідження достовірності оцінок отриманих за допомогою цих методів.

Постановка задачі

У роботі [4] досліджено властивості оцінок найменших квадратів (ОНК) МПКВП-імовірнісних характеристик, які вважалися неперервними реалізаціями. Проте, коли мова іде про МПКВП - аналіз сигналу, то у більшості випадків йдеться про обробку сигналу за допомогою комп'ютера. Тобто, реально ми маємо справу із часовою послідовністю відліків $\left\{\xi_k = \xi(kh),\ k = -\overline{N, N}\right\}$, отриманою шляхом дискретизації випадкового процесу $\xi(t)$ на часовому інтервалі $t \in [-\theta, \theta]$, $\theta = Nh$. h – крок дискретизації. У [3] наведено результат досліджень достовірності дискретної ОНК математичного сподівання, і показано, що коли крок дискретизації такий, що не виконується умова

$$\forall\ \omega_{p_1},\omega_{p_2}\in\{\omega_p\}:\left|\omega_{p_1}\pm\omega_{p_2}\right|h\neq 2\pi d,\qquad d\in\mathbf{N},$$

де $\{\omega_p\}$ – множина частот коливань гармонічних складових математичного сподівання (так званий ефект елайзингу 1-го роду), то дискретна ОНК математичного сподівання не задовольняє умови слушності і асимптотичної незміщеності. Так само, у випадку невиконання умови

$$\forall\ \omega_{p_1},\omega_{p_2}\in\{\omega_p\}\quad \text{і}\quad \forall\ \nu_q\in\{\nu_q\}:\left|\omega_{p_1}\pm\omega_{p_2}\pm\nu_q\right|h\neq 2\pi d,\quad d\in\mathbf{N},\tag{1}$$

де $\{v_q\}$ множина частот коливань гармонічних складових кореляційної функції, спостерігається ефект елайзингу 2-го роду, який проявляється у погіршенні дисперсії дискретної ОНК математичного сподівання зі збільшенням довжини реалізації процесу.

Розглянемо задачу оцінювання методом найменших квадратів кореляційної функції $\{\xi_k\}$ часового ряду.

$$b(t,u)=\mathbf{B}(u)^T\,\mathbf{e}^{\nu}(t),\tag{2}$$

$$\mathbf{B}(u)=\left(B_0(u),B_1^c(u),...,B_{L_2}^c(u),B_1^s(u),...,B_{L_2}^s(u)\right)^T,$$

$$\mathbf{e}^{\nu}(t)=\left(1,\cos\nu_1 t,...,\cos\nu_{L_2}t,\sin\nu_1 t,...,\sin\nu_{L_2}t\right)^T.$$

Оцінку математичного сподівання і кореляційної функції шукають у вигляді [1–3]

$$\hat{m}(t)=\hat{\mathbf{m}}^T\mathrm{e}^{\omega}(t);\qquad \hat{b}(t)_s=\hat{\mathbf{B}}_s^T\mathbf{e}^{\nu}(t),\tag{3}$$

де $\mathbf{m}=\left(m_0,m_1^c,...,m_{L_1}^c,m_1^s,...,m_{L_1}^s\right)^T$, $\mathbf{e}^{\omega}(t)=\left(1,\cos\omega_1 t,...,\cos\omega_{L_1}t,\sin\omega_1 t,...,\sin\omega_{L_1}t\right)^T$; L_1, L_2 – кількість ненульових гармонік математичного сподівання і кореляційної функції відповідно; $\left\{\omega_k,\ k=\overline{1,L_1}\right\}$ – множина частот коливань гармонічних складових математичного сподівання, $\left\{\omega_k,\ k=\overline{1,L_1}\right\}$ – множина частот коливань гармонічних складових кореляційної функції.

Результати наукових досліджень

Мінімізуючи функціонал

$$\mathbf{F}[\mathbf{B}_u]=\frac{1}{2N+1}\sum_{k=-N}^{N}\left[\eta_{k,u}-\hat{b}_{k,u}\right]^2,$$

де $\eta_{ku}=\mathring{\xi}_k\,\mathring{\xi}_{k+u}$, $\mathring{\hat{\xi}}_k=\xi_k-\hat{m}(kh)$, отримуємо розв'язок:

$$\hat{\mathbf{B}}_u=2\tilde{\mathbf{M}}^{-1}\overline{\mathbf{B}}_u,$$

де $\tilde{\mathbf{M}}^{\nu}=\frac{1}{2N+1}\sum\limits_{k=-N}^{N}\mathbf{e}_k^{\nu}\mathbf{e}_k^{\nu T}$, $\mathbf{e}_k^{\nu}=\mathbf{e}^{\nu}(kh)$, $\bar{\mathbf{B}}_u=\frac{1}{2N+1}\sum\limits_{k=-N}^{N}\eta_{ku}\mathbf{e}_k^{\nu}$.

Знайдений вектор $\hat{\mathbf{B}}_u$ мінімізує функціонал $F[\mathbf{B}_u]$. Це випливає із того, що квадратична форма, побудована на основі його других частинних похідних, додатньо-визначена.

Матриця $\tilde{\mathbf{M}}^{\nu}$ симетрична, і при $Nh\to\infty$ прямує до діагональної:

$$\lim_{Nh\to\infty}\tilde{M}_{pq}=\frac{1}{2}\begin{cases}2, & p=q=0;\\ \delta_{pq}, & pq=\overline{1,L_2},\end{cases}$$

окрім випадку, коли

$$\left|\nu_{q_1}\pm\nu_{q_2}\right|h=2\pi d,\quad d\in\mathbf{N},\quad \nu_{q_1}\ \text{і}\ \nu_{q2}\in\left\{\nu_q\right\}. \tag{4}$$

Вибір такого кроку дискретизації, коли існує хоча б дві такі частоти, для яких справедливе (4), призводить до того, що $\tilde{M}^{\nu}_{q_1q_1}=\tilde{M}^{\nu}_{q_2q_2}=\tilde{M}^{\nu}_{q_2q_1}=\tilde{M}^{\nu}_{q_1q_2}=\frac{1}{2}$; матриця $\tilde{\mathbf{M}}^{\nu}$ має однакові рядки і є виродженою. У цьому випадку неможливо обчислити обернену матрицю $\tilde{\mathbf{M}}^{\nu-1}$, а разом із тим і знайти дискретну ОНК кореляційної функції, отже, це ефект елайзингу 1-го роду.

Надалі вважатимемо, що рівність (4) не виконується.

Так само, як і у неперервному випадку, дискретна ОНК коефіцієнтів Фур'є кореляційної функції має зміщення, зумовлене центруванням сигналу на оцінку його математичного сподівання:

$$E\hat{\mathbf{B}}_u=\tilde{\mathbf{M}}^{\nu-1}E\bar{\mathbf{B}}_u=\mathbf{B}_u+\mathbf{e}^m_{Bu},$$

де

$$\mathbf{e}^m_{Bu}=\sum_{i=1}^{3}\mathbf{e}^m_{B_iu},\ \mathbf{e}^m_{B_iu}=\frac{1}{2N+1}\sum_{k=-N}^{N}\varepsilon_{\eta_i ku}\mathbf{e}^{\nu}(kh),\ \varepsilon_{\eta_1 ku}=E\left[\overset{\circ}{\xi}_k\ \overset{\circ}{\hat{m}}\left(\{k+u\}h\right)\right],$$

$$\varepsilon_{\eta_2 ku}=E\left[\overset{\circ}{\xi}_{k+u}\ \overset{\circ}{\hat{m}}(kh)\right],\ \varepsilon_{\eta_3 ku}=-E\left[\overset{\circ}{\hat{m}}(kh)\overset{\circ}{\hat{m}}\left(\{k+u\}h\right)\right],$$

$$\overset{\circ}{\hat{m}}(t)=\mathbf{e}^{\omega}(t)^T\ \tilde{\mathbf{M}}^{\omega-1}\ \overset{\circ}{\bar{\mathbf{m}}},\ \overset{\circ}{\bar{\mathbf{m}}}=\frac{1}{2N+1}\sum_{k=-N}^{N}\overset{\circ}{\xi}_k\ \mathbf{e}_k^{\omega}.$$

Для $\mathbf{e}^m_{B_3u}$ можна записати

$$\varepsilon_{\eta_3 ku}=-E\left[\overset{\circ}{\hat{m}}(kh)\overset{\circ}{\hat{m}}\left(\{k+u\}h\right)\right]=-\mathbf{e}_k^{\omega T}\tilde{\mathbf{M}}^{\omega-1}E\left[\overset{\circ}{\bar{\mathbf{m}}}\overset{\circ}{\bar{\mathbf{m}}}^T\right]\tilde{\mathbf{M}}^{\omega-1}\mathbf{e}_k^{\omega},$$

$$\mathrm{E}\left[\overset{\circ}{\bar{\mathbf{m}}}\overset{\circ}{\bar{\mathbf{m}}}^T\right]=\frac{1}{(2N+1)^2}\sum_{k,l=-N}^{N}E\left[\overset{\circ}{\xi}_k\ \overset{\circ}{\xi}_l\right]\mathbf{e}_k^{\omega}\mathbf{e}_l^{\omega T}=\bar{\mathbf{R}}_m,$$

де $\tilde{\mathbf{R}}_m = \dfrac{1}{(2N+1)^2} \sum\limits_{k,l=-N}^{N} b_{k,k-l}\mathbf{e}_k^{\omega}\mathbf{e}_l^{\omega T}$.

Отже

$$\varepsilon_{\eta_3 ku} = -\mathbf{e}_k^{\omega T}\tilde{\mathbf{M}}^{\omega-1}\tilde{\mathbf{R}}_m\tilde{\mathbf{M}}^{\omega-1}\mathbf{e}_{k+u}^{\omega}, \qquad \mathbf{e}_{B_3 u}^{m} = -\frac{1}{2N+1}\sum_{k=-N}^{N}\mathbf{e}_k^{\nu}\ \mathbf{e}_k^{\omega T}\tilde{\mathbf{M}}^{\omega-1}\tilde{\mathbf{R}}_m\tilde{\mathbf{M}}^{\omega-1}\mathbf{e}_{k+u}^{\omega}.$$

У роботі [3] досліджено властивості матриці $\tilde{\mathbf{R}}_m$, де показано, що

$$\lim_{Nh\to\infty}\tilde{\mathbf{R}}_m = 0, \tag{5}$$

якщо компоненти кореляційної функції задовольняють умову

$$\int_{-\theta}^{\theta}\mathbf{B}(x)dx < \infty, \quad \theta = Nh, \tag{6}$$

а невиконання умови відсутності елайзингу 2-го роду (1) призводить до погіршання збіжності (5).

Можна також показати, що при виконанні (6) і (1), $\lim\limits_{Nh\to\infty}\mathbf{e}_{B_3 u}^{m} = 0$. Так само поводять себе і решта доданків: $\mathbf{e}_{B_1 u}^{m}$ і $\mathbf{e}_{B_2 u}^{m}$. Отже, коли компоненти кореляційної функції задовольняють умову (6), то ОНК кореляційної функції асимптотично незміщена.

Знайдемо дисперсію ОНК кореляційної функції:

$$D\left[\hat{b}(t,uh)\right] = E\left[\hat{b}(t,uh) - E\hat{b}(t,uh)\right]^2 = E\left[\hat{b}(t,uh)\right]^2 - \left[E\hat{b}(t,uh)\right]^2 = E\left[\mathbf{e}^{\nu}(t)^T\hat{\mathbf{B}}_u\right]^2 -$$

$$-\left[\mathbf{e}^{\nu}(t)^T\mathbf{B}_u + \mathbf{e}^{\nu}(t)^T\mathbf{e}_{Bu}^{m}\right]^2 = \mathbf{e}^{\nu}(t)^T\tilde{\mathbf{M}}^{\nu-1}E\left[\bar{\mathbf{B}}_u\bar{\mathbf{B}}_u^T\right]\tilde{\mathbf{M}}^{\nu-1}\mathbf{e}^{\nu}(t) - \left[\mathbf{e}^{\nu}(t)^T\mathbf{e}_{Bu}^{m}\right]^2.$$

$$E\left[\bar{\mathbf{B}}_u\bar{\mathbf{B}}_u^T\right] = \frac{1}{(2N+1)^2}\sum_{k,n=-N}^{N}E\left[\hat{\eta}_{ku}\hat{\eta}_{nu}\right]\mathbf{e}^{\nu}(kh)\mathbf{e}^{\nu}(nh)^T,$$

відкидаючи члени вищого порядку малості, припускаючи гауссовість стохастичної складової процесу, можна записати

$$E\left[\hat{\eta}_{ku}\hat{\eta}_{nu}\right] = b_{k,u}b_{n,u} + f_{k,n,u}, \qquad f_{k,n,u} = b_{k,n-k}b_{k+u,n-k} + b_{k,n-k+u}b_{n,k-n+u}.$$

Ввівши позначення

$$\tilde{\mathbf{R}}_b = \frac{1}{(2N+1)^2}\sum_{k,n=-N}^{N} f_{k,n,u}\mathbf{e}^{\nu}(kh)\mathbf{e}^{\nu}(nh)^T, \tag{7}$$

запишемо

$$E\left[\bar{\mathbf{B}}_u\bar{\mathbf{B}}_u^T\right] = \tilde{\mathbf{R}}_b + \tilde{\mathbf{M}}^{\nu}\mathbf{B}_u\mathbf{B}_u^T\tilde{\mathbf{M}}^{\nu}.$$

Після підстановки для достатньо довгої довжини реалізації наближено можна записати

$$D\left[\tilde{b}(t,uh)\right] \approx \mathbf{e}^{\nu}(t)^{T}\ \tilde{\mathbf{M}}^{\nu-1}\tilde{\mathbf{R}}_{b}\tilde{\mathbf{M}}^{\nu-1}\mathbf{e}^{\nu}(t).$$

Дослідимо матрицю $\tilde{\mathbf{R}}_b$. Зробимо в (7) підстановку $n = k - s$

$$\tilde{\mathbf{R}}_b = \mathbf{R}_{0u} + \mathbf{R}_{1u} + \mathbf{R}_{2u},$$

де

$$\mathbf{R}_{0u} = \frac{1}{(2N+1)^2}\sum_{k=-N}^{N} f_{k,k,u}\mathbf{e}_k^{\nu}\mathbf{e}_k^{\nu\,T},\quad \mathbf{R}_{1u} = \frac{1}{(2N+1)^2}\sum_{s=1}^{2N}\sum_{k=-N}^{N-s} f_{k,k+s,u}\mathbf{e}_k^{\nu}\mathbf{e}_{k+s}^{\nu\,T},$$

$$\mathbf{R}_{2u} = \frac{1}{(2N+1)^2}\sum_{s=-2N}^{-1}\sum_{k=-N-s}^{N} f_{k,k+s,u}\mathbf{e}_k^{\nu}\mathbf{e}_{k+s}^{\nu\,T}.$$

Шляхом тотожних перетворень можна показати, що $\mathbf{R}_u = \mathbf{R}_{1u} = \mathbf{R}_{2u}^T$. Отже

$$\tilde{\mathbf{R}}_b = \mathbf{R}_{0u} + \mathbf{R}_u + \mathbf{R}_u^T.$$

Перепишемо вираз для кореляційної функції (2) у вигляді

$$b_{k,u} = B_{0u} + \sum_{i=1}^{L_2} A_{iu}\cos\left[\omega_i kh + \phi_{iu}\right]$$

де $A_{iu} = \sqrt{B_{iu}^{c2} + B_{iu}^{s2}}$, а $\phi_{iu} = \mathrm{arctg}\,\frac{B_{iu}^s}{B_{iu}^c}$. Тоді для добутку $b_{n,r}b_{n+u,r}$ можна записати

$$b_{n,r}b_{n+u,r} = B_{0r}^2 + B_{0r}\sum_{i=1}^{L_2} A_{ir}\left(\cos\left[\omega_i nh + \phi_{ir}\right] + \cos\left[\omega_i(n+u)h + \phi_{ir}\right]\right) +$$

$$+\frac{1}{2}\sum_{i,j=1}^{L_2} A_{ir}A_{jr}\left[\cos\left(\left[\omega_j - \omega_i\right]nh + u\omega_j h + \phi_{jr} - \phi_{ir}\right) +\right.$$

$$\left.+\cos\left(\left[\omega_j + \omega_i\right]nh + u\omega_j h + \phi_{jr} + \phi_{ir}\right)\right].$$

Розглянемо один із доданків p-го, q-го ($1 \le p, q \le L_2$) елемента матриці

$$\rho_{pq} = \frac{1}{(2N+1)^2}\sum_{i,j=1}^{L_2}\sum_{r=1}^{2N} A_{ir}A_{jr}\sum_{n=-N}^{N-r}\cos\left(\left[\nu_j \pm \nu_i\right]nh + u\nu_j h + \phi_{jr} \pm \phi_{ir}\right)\cos(n+r)\nu_p h\cos n\nu_q h.$$

Виконуючи елементарні тригонометричні перетворення та підсумувавши за n, можна показати, що

$$\left|\rho_{pq}\right| \le \frac{1}{2N+1} \sum_{i,\,j=1}^{L} \sum_{r=1}^{2N} A_{ir} A_{jr} \frac{1}{4}\left(1 - \frac{r}{2N+1}\right) \times$$

$$\times \left(S_{\nu_j \pm \nu_i + \nu_p - \nu_q} + S_{\nu_j \pm \nu_i - \nu_p + \nu_q} + S_{\nu_j \pm \nu_i + \nu_p + \nu_q} + S_{\nu_j \pm \nu_i - \nu_p - \nu_q} \right), \tag{8}$$

де

$$S_\nu = \begin{cases} 1, & \nu = 0; \\ 1, & \nu h = 2\pi d,\ d \in \mathrm{Z}; \\ 0, & \text{в усіх інших випадках.} \end{cases}$$

Подібні співвідношення можна отримати для решти доданків p, q-го елемента матриці $\mathbf{R}_u$ та для усіх елементів матриці $\tilde{\mathbf{R}}_b$.

Послідовність $1 - \dfrac{r}{2N+1}$, $r = \overline{1, 2N}$ монотонно спадна і обмежена на області визначення. Із ознак абсолютної збіжності рядів необхідною і достатньою умовою того, щоб $\lim\limits_{Nh \to \infty} \tilde{\mathbf{R}}_b = 0$, є виконання умов

$$\forall\ i, j = \overline{1, L_2} : \lim_{Nh \to \infty} \sum_{r=1}^{2N} \sqrt{\left(B_{ir}^{c2} + B_{ir}^{s2}\right)\left(B_{jr}^{c2} + B_{jr}^{s2}\right)} < \infty, \tag{9}$$

і

$$\lim_{Nh \to \infty} \sum_{r=1}^{2N} B_{0r} < \infty. \tag{10}$$

Отже, якщо компоненти кореляційної функції задовольняють нерівності (9), (10), то ОНК кореляційної функції справедлива

$$\lim_{Nh \to \infty} D\left[b_{k,u} \right] = 0.$$

Однак, з нерівності (8) випливає, що якщо серед частотного набору кореляційної функції $\{\nu_k\}$ існують такі частоти $\left|\nu_j \pm \nu_i \pm \nu_p \pm \nu_q\right| h = 2\pi d,\ d \in \mathbf{Z}$, то це призводить до погіршення швидкості загасання дисперсії ОНК кореляційної функції до нуля, тобто спостерігаємо ефект елайзингу 3-го роду.

Висновок

Отже, узагальнюючи отримані результати досліджень впливу дискретизації на якість ОНК імовірнісних характеристик, в тому числі й результати, наведені у роботі [4], можна рекомендувати дослідникам вибирати частоту дискретизації $f_d = 1/h$ так, щоб $f_d > 3f_{\max}, f_{\max} = 2\pi\max[\{\omega_k\}, \{\nu_l\}]$, коли до уваги береться лише математичне сподівання. Якщо ж об'єктом дослідження є стохастична складова сигналу, то необхідно вибирати частоту дискретизації $f_d > 4f_{\max}$. У цих випадках можна гарантувати уникнення небажаних ефектів елайзингу і, відповідно, забезпечити оптимальну збіжність оцінок імовірнісних характеристик МПКВП.

1. Efstratios Nikolaidis, Honggang Wang, Akhilesh Jha, Dan M. Ghiocel. Fatigue reliability of cars under road-induced cyclostationary excitation // of Materials & Product Technology. - 2001. – 16, № 4/5. – p.404–416.
2. Драган Я.П., Рожков В.А., Яворский И.Н. Методы вероятностного анализа ритмики гидрометеорологических процесов. – Л.: Гидрометеоиздат, 1987. – 347с.
3. Заболотний О.В., Михайлишин В.Ю., Яворський І.М., Метод найменших квадратів при статистичному аналізі поліритміки // Доповіді Національної академії наук України - 2000. – №8. – С. 93-101.
4. Заболотний О.В, Михайлишин В.Ю. Оцінка математичного сподівання полі ритмічних часових рядів за методом найменших квадратів // Відбір і обробка інформації. – 2002. – №17(93). – С.65–70.

В.Шекета
Івано-Франківський національний технічний університет нафти і газу

УДК 681. 3:622. 276

АНАЛІЗ СЕМАНТИКИ ШАБЛОНІВ ВИКЛИКУ МОДИФІКАЦІЙНИХ ПРЕДИКАТНИХ ЗАПИТІВ ДЛЯ ІНФОРМАЦІЙНИХ СИСТЕМ НА ОСНОВІ БАЗ ДАНИХ І ЗНАНЬ

Запропоновано спосіб обчислення семантик шаблонів виклику модифікаційних предикатних запитів для інформаційних систем на основі баз даних і знань.

The method of calculation semantics for patterns of modification predicate queries calls is offered for the information systems on the basis of databases and knowledges.

Вступ

Під семантикою будемо розуміти відображення, яке кожній програмі, створеній заданою мовою програмування, дозволяє поставити у відповідність певну практичну