РОССИЙСКАЯ АКАДЕМИЯ НАУК

СИБИРСКОЕ ОТДЕЛЕНИЕ

АВТОМЕТРИЯ

№ 6 2001

УДК 519.234

В. Г. Алексеев

(Звенигород Московской обл.)

ОЦЕНКА СПЕКТРАЛЬНОЙ ПЛОТНОСТИ ТИПА УЭЛЧА. СЛУЧАЙ ДИСКРЕТНОГО АРГУМЕНТА

Предложена новая модификация оценки спектральной плотности типа Уэлча (оценки, получаемой осреднением по сдвигу во времени) для стационарного случайного процесса с дискретным временем. В качестве сглаживающих окон данных при построении оценки типа Уэлча используются разложения в ряд Фурье полиномиальных тригонометрических ядер типа Джексона до пятого порядка включительно. Обсуждаются статистические свойства рассматриваемой оценки спектральной плотности, формулируются рекомендации по выбору ее параметров.

1. Данная работа, как и [1], посвящена прикладному спектральному анализу стационарных случайных процессов (ССП). Но в отличие от [1] в настоящей работе нас интересует случай дискретного временно́го аргумента. При этом мы не рассматривали проблему спектрального оценивания для случайных процессов с дискретным временем в ее полном объеме, а остановились лишь на оценке, известной под названием оценки типа Уэлча. Одним из наиболее привлекательных достоинств оценки типа Уэлча является ее высокая помехозащищенность, т. е. слабая зависимость от возмущений и всплесков на удаленных частотах.

Итак, пусть $\{X(k),\ k=0,\pm1,\pm2,\ldots\}$ – стационарный в широком смысле случайный процесс со средним $\langle X(k)\rangle\equiv 0$, корреляционной функцией $r(k)=$ $=\langle X(j)X(j+k)\rangle$ и спектральной плотностью

$$f(\omega)=(2\pi)^{-1}\sum_{k=-\infty}^{\infty}e^{-ik\omega}r(k),\qquad \omega\in\Pi=[-\pi,\pi]. \tag{1}$$

В целях упрощения выкладок мы полагаем в дальнейшем, что ССП $X(k)$ гауссов. Наконец, относительно спектральной плотности $f(\omega)$ предполагаем, что она, будучи продолженной периодическим образом за пределы отрезка Π, по крайней мере дважды дифференцируема, причем

$$\sup_{\omega}|f''(\omega)|<K<\infty. \tag{2}$$

Говоря об оценке спектральной плотности типа Уэлча, мы имеем в виду вычислительный алгоритм, включающий в себя разбиение всего интервала наблюдения длины N на конечное число неперекрывающихся или частично перекрывающихся сегментов длины $M < N$, вычисление по каждому из них периодограммы и ее последующее осреднение по числу сегментов длины M. При этом, говоря о периодограмме, вычисляемой по каждому из сегментов длины M, мы имеем в виду ее модифицированную версию, определяемую соотношением

$$I_M^{\{B\}}(\omega) = \left| \sum_{k=1}^{M} b(k) X(k) \mathrm{e}^{ik\omega} \right|^2 \left[2\pi \sum_{k=1}^{M} b^2(k) \right]^{-1}.$$

Здесь $B = \{b(k),\ k = \overline{1, M}\}$ – так называемое окно данных, используемое для домножения (неравномерного взвешивания) отрезка реализации $\{X(k), k = \overline{1, M}\}$. Последовательность $b(k)$ чаще всего плавно убывает от середины отрезка реализации к его краям, чем достигается сглаживание краев реализации и в конечном счете уменьшение смещения (систематической ошибки оценивания) периодограммы. Большое число разнообразных окон данных B можно найти в [2–4]. В настоящей работе в качестве сглаживающих окон данных B при построении оценки типа Уэлча используем разложения в ряд Фурье полиномиальных тригонометрических ядер типа Джексона.

2. Данный раздел условно назовем вспомогательным. В нем приведем все необходимые сведения, касающиеся ядер типа Джексона до пятого порядка включительно. Предварительно условимся, что далее интеграл без указания пределов будет обозначать интегрирование по отрезку $\Pi = [-\pi, \pi]$.

Говоря о полиномиальных ядрах типа Джексона, мы имеем в виду 2π-периодические функции (конечные тригонометрические полиномы) $J_{l,n}(\mu)$, определенные для произвольных натуральных чисел l и n соотношением

$$J_{l,n}(\mu) = C_{l,n} \left[\frac{\sin(n\mu/2)}{\sin(\mu/2)} \right]^{2l}, \qquad (3)$$

где $C_{l,n}$ – нормирующий множитель, обеспечивающий выполнение равенства $\int J_{l,n}(\mu) d\mu = 2\pi$.

Нормирующие множители $C_{l,n}$ для всех $l \le 5$, а также разложения в ряд Фурье ядер $J_{l,n}(\mu)$ для $l \le 4$ приведены в [5]. И лишь разложение в ряд Фурье ядра пятого порядка $J_{5,n}(\mu)$ ранее не публиковалось, что и обязывает нас привести здесь это разложение полностью. Величины $b(k)$ в формуле

$$J_{5,n}(\mu) = \sum_{k=-5(n-1)}^{5(n-1)} b(k) \mathrm{e}^{ik\mu}$$

определяются для $|k| \le n, n \le |k| \le 2n, 2n \le |k| \le 3n, 3n \le |k| \le 4n$ и $4n \le |k| \le 5n$ соотношениями:

$$b(k) = 1 - 42U^{-1}[1728|k| + 50(42n^7 + 57n^5 + 65n^3 + 82n)k^2 -$$

$$-2460|k|^3 - 5(114n^5 + 250n^3 + 455n)k^4 + 819|k|^5 +$$

$$+50(2n^3+7n)k^6-90|k|^7-15nk^8+3|k|^9],$$

$$b(k)=2U^{-1}[10(7799n^9+3990n^7-273n^5+10660n^3-4032n)+$$
$$+63(15n^8-350n^6+2275n^4-4100n^2+384)|k|-$$
$$-420(114n^7-15n^5+845n^3-410n)k^2+$$
$$+210(42n^6-525n^4+1365n^2-164)|k|^3-$$
$$-210(6n^5-650n^3+455n)k^4+882(15n^4-75n^2+13)|k|^5-$$
$$-420(26n^3-35n)k^6+1260(3n^2-1)|k|^7-630nk^8+42|k|^9],$$

$$b(k)=2U^{-1}[10(10871n^9-19050n^7+52143n^5-28700n^3+2880n)-$$
$$-27(5085n^8-29050n^6+43225n^4-12300n^2+384)|k|+$$
$$+300(762n^7-4011n^5+3185n^3-410n)k^2-$$
$$-90(3486n^6-9975n^4+4095n^2-164)|k|^3+$$
$$+210(1146n^5-1750n^3+325n)k^4-$$
$$-378(285n^4-225n^2+13)|k|^5+2100(14n^3-5n)k^6-$$
$$-540(9n^2-1)|k|^7+450nk^8-18|k|^9],$$

$$b(k)=U^{-1}[-5(133663n^9-514290n^7+388479n^5-84460n^3+4032n)+$$
$$+27(88245n^8-197050n^6+97825n^4-12300n^2+192)|k|-$$
$$-210(14694n^7-21345n^5+6695n^3-410n)k^2+$$
$$+90(23646n^6-22575n^4+4095n^2-82)|k|^3-$$
$$-105(8538n^5-5150n^3+455n)k^4+$$
$$+189(1290n^4-450n^2+13)|k|^5-210(206n^3-35n)k^6+$$
$$+270(18n^2-1)|k|^7-315nk^8+9|k|^9]$$

и соответственно

$$b(k)=(5n-|k|)U^{-1}\prod_{r=1}^{4}[(5n-|k|)^2-r^2],$$

где $U = 10(15619n^9 + 7350n^7 + 5187n^5 + 4100n^3 + 4032n)$.

3. Будем считать вспомогательную часть настоящей статьи завершенной. Предполагая, что $\omega \in [0, \pi]$, определим предлагаемую нами модификацию оценки спектральной плотности типа Уэлча формулой

$$f_N(\omega) = (QT)^{-1} \sum_{t=-(T-1)}^{T-1} a(t) \left| \sum_{k=-(m-1)}^{m-1} b(k) X(k+tL) e^{ik\omega} \right|^2 . \qquad (4)$$

Здесь

$$a(t) = (2\pi)^{-1} \int e^{it\mu} J_{1,T}(\mu) d\mu = \begin{cases} 1 - |t|/T, & |t| \le T, \\ 0, & |t| > T; \end{cases}$$

$$b(k) = b_{l,n}(k) = (2\pi)^{-1} \int e^{ik\mu} J_{l,n}(\mu) d\mu; \qquad (5)$$

$$Q = Q_{l,n} = 2\pi \sum_{k=-(m-1)}^{m-1} b^2(k) = \int J_{l,n}^2(\mu) d\mu = (2\pi C_{l,n}^2)/C_{2l,n};$$

m, T, L, l и n – целочисленные параметры оценки (4); N – общий объем выборки, определяемый соотношением $N = 2m + 2(T-1)L - 1$.

Разумеется, параметры m, T, L, l и n не могут быть выбраны произвольно. Они связаны следующими соотношениями и условиями: $T > 1$, $1 \le L \le 2m - 1$, $n > 1$, $1 \le l \le 5$ и $m = l(n-1) + 1$. Последнее из приведенных выше соотношений обусловлено тем обстоятельством, что коэффициенты Фурье $b(k)$ ядра $J_{l,n}(\mu)$ тождественно равны нулю для всех $|k| > l(n-1)$. Что же касается неравенства $l \le 5$, то оно отражает тот факт, что разложениями в ряд Фурье ядер $J_{l,n}(\mu)$ для $l > 5$ мы пока еще не располагаем.

Вычисляя математическое ожидание оценки (4), получаем

$$\langle f_N(\omega) \rangle = \int J_{l,n}^2(\mu) f(\omega + \mu) d\mu \Big/ \int J_{l,n}^2(\mu) d\mu. \qquad (6)$$

Отсюда, разлагая функцию $f(\omega + \mu)$ в ряд Тейлора в окрестности точки ω и принимая во внимание формулу (3) и предположение (2), находим, что при любом l

$$\sup_{\omega \in [0,\pi]} \left| \langle f_N(\omega) \rangle - f(\omega) \right| \sim K/(2n^2), \qquad (7)$$

где символ $\sim$ обозначает пропорциональность двух величин. Можно показать также, что, по крайней мере при $l \le 5$, левая часть формулы (7) не превосходит $0{,}3K\pi^2/n^2$.

Переходя к вычислению дисперсии оценки (4) и принимая во внимание предположение о гауссовости ССП $X(k)$, находим

$$\mathrm{D} f_N(\omega) \equiv \langle f_N^2(\omega) \rangle - |\langle f_N(\omega) \rangle|^2 = (QT)^{-2} \sum_{t=-(T-1)}^{T-1} a(t) \sum_{s=-(T-1)}^{T-1} a(s) \sum_{j=-(m-1)}^{m-1} b(j) \times$$

$$\times \sum_{k=-(m-1)}^{m-1} b(k) \sum_{u=-(m-1)}^{m-1} b(u) \sum_{v=-(m-1)}^{m-1} b(v)\,\mathrm{e}^{i(j+u-k-v)\omega} \times$$

$$\times [\langle X(j+tL)X(v+sL)\rangle\langle X(k+tL)X(u+sL)\rangle +$$

$$+\langle X(j+tL)X(u+sL)\rangle\langle X(k+tL)X(v+sL)\rangle]. \qquad (8)$$

Таким образом, $\mathrm{D}f_N(\omega) = D_1 + D_2$, где величина D_j, $j = 1, 2$, возникает за счет j-го слагаемого в квадратных скобках в крайней правой части соотношения (8).

Полученное на основании формулы (1) спектральное разложение корреляционной функции $r(k) = \langle X(j)X(j+k)\rangle$ позволяет утверждать, что

$$D_1 = \int \left[\sum_{k=-(m-1)}^{m-1} b(k) \sum_{u=-(m-1)}^{m-1} b(u)\,\mathrm{e}^{i(k-u)(\mu-\omega)} \right] f(\mu)d\mu \times$$

$$\times \int \left[\sum_{j=-(m-1)}^{m-1} b(j) \sum_{v=-(m-1)}^{m-1} b(v)\,\mathrm{e}^{i(j-v)(\theta+\omega)} \right] f(\theta) \times$$

$$\times \left[(QT)^{-2} \sum_{t=-(T-1)}^{T-1} a(t) \sum_{s=-(T-1)}^{T-1} a(s)\,\mathrm{e}^{i(t-s)L(\mu+\theta)} \right] d\theta. \qquad (9)$$

С помощью соотношений

$$\sum_{k=-(m-1)}^{m-1} b(k)\mathrm{e}^{ik\mu} = J_{l,n}(\mu), \qquad \left| \sum_{t=-(T-1)}^{T-1} a(t)\mathrm{e}^{it\mu} \right|^2 = J_{1,T}^2(\mu) = \frac{2T^2+1}{3T} J_{2,T}(\mu)$$

формула (9) легко преобразуется к виду

$$D_1 = \frac{2T^2+1}{3Q^2T^3} \int J_{l,n}^2(\mu) f(\mu+\omega)d\mu \int J_{2,T}(L\theta) J_{l,n}^2(\theta-\mu) f(\theta-\mu-\omega)d\theta. \qquad (10)$$

Переходя к переменной $\lambda = L\theta$ под знаком внутреннего интеграла в правой части соотношения (10), находим

$$D_1 = \frac{2\pi(2T^2+1)}{3Q^2T^3L} \int J_{l,n}^2(\mu) f(\mu+\omega)d\mu \int (2\pi)^{-1} J_{2,T}(\lambda) \times$$

$$\times \left[\sum_{j=0}^{L-1} J_{l,n}^2\left(\frac{\lambda+2j\pi}{L} - \mu \right) f\left(\frac{\lambda+2j\pi}{L} - \mu - \omega \right) \right] d\lambda. \qquad (11)$$

Дальнейшее преобразование полученного нами выражения требует предположения, что $n \gg 2$ и $TL \gg 12nl$. Пользуясь свойствами сингулярного интеграла Джексона (см., например, [6, гл. IV, § 2, формулы (96) и (99)]), можно показать, что в этом случае формула (11) преобразуется в приближенное равенство

$$D_1 \approx \frac{2\pi(2T^2+1)}{3Q^2T^3L}\int J_{l,n}^2(\mu)f(\mu+\omega)\times$$
$$\times\left[\sum_{j=0}^{L-1} J_{l,n}^2\left(\mu-\frac{2j\pi}{L}\right)f\left(\mu+\omega-\frac{2j\pi}{L}\right)\right]d\mu. \quad (12)$$

Если, кроме того, $2L \le n$, то формула (12) может быть еще более упрощена:

$$D_1 \approx \frac{4\pi}{3Q^2TL}\int J_{l,n}^4(\mu)f^2(\mu+\omega)d\mu.$$

Замечая теперь, что $J_{l,n}^4(\mu)=C_{l,n}^4 C_{4l,n}^{-1} J_{4l,n}(\mu)$ и $C_{l,n} \sim n^{1-2l}$ [7, гл. II, § 3] при всех l, и принимая во внимание (5), находим, что при наших предположениях

$$D_1 \approx \frac{C_{2l,n}^2}{3\pi TLC_{4l,n}}\int J_{4l,n}(\mu)f^2(\mu+\omega)d\mu \sim \frac{2n}{3TL}f^2(\omega).$$

Нетрудно показать, что при тех же предположениях $D_2 = D_1$, если $\omega = 0, \pi$, и $D_2 = o(n/TL)$ при $n/(TL) \to 0$ в остальных случаях.

Мы можем утверждать теперь, что

$$\mathrm{D}f_N(\omega) = D_1 + D_2 \sim \frac{2n\{2-\mathrm{sign}[\omega(\pi-\omega)]\}}{3TL}f^2(\omega), \quad (13)$$

если $n \gg 2$, $TL \gg 12nl$ и $2L \le n$.

Формулами (7) и (13) описываются важнейшие статистические характеристики оценки (4). Видно, что смещение оценки (4) и ее дисперсия решающим образом зависят от выбора параметра n. Можно ожидать, в частности, что при малых n квадрат смещения оценки (4) будет существенно превосходить ее дисперсию. Однако с ростом n положение быстро меняется в обратную сторону. При достаточно больших n дисперсия оценки (4) будет уже заметно превосходить квадрат ее смещения. В том случае, когда дисперсия оценки (4) имеет тот же порядок малости, что и квадрат ее смещения, достигается минимум среднего квадрата ошибки оценивания.

Что же касается параметра l, то он ответствен, прежде всего, за помехозащищенность оценки (4). В соответствии с формулами (3) и (6) в наибольшей мере это качество оценки (4) достигается при l равном 4 и 5.

Замечание. Уже после того, как статья была отправлена в редакцию, автору удалось доказать следующее неравенство при любых натуральных l и n:

$$\frac{\int_\Pi \left[\frac{\sin(n\mu/2)}{\sin(\mu/2)}\right]^{4l+4}\mu^2 d\mu}{\int_\Pi \left[\frac{\sin(n\mu/2)}{\sin(\mu/2)}\right]^{4l+4} d\mu} \le \frac{\int_\Pi \left[\frac{\sin(n\mu/2)}{\sin(\mu/2)}\right]^{4l}\mu^2 d\mu}{\int_\Pi \left[\frac{\sin(n\mu/2)}{\sin(\mu/2)}\right]^{4l} d\mu}.$$

В соответствии с формулами (3) и (6) это означает, что при сделанных в разд. 1 предположениях полученная нами оценка сверху для смещения оценки (4) с ростом параметра l может лишь убывать. Поэтому переход от l к $l+1$ всегда желателен (если только объем выборки не слишком мал).

СПИСОК ЛИТЕРАТУРЫ

1. **Алексеев В. Г.** Непараметрический спектральный анализ стационарных случайных процессов // Автометрия. 2000. № 4. С. 131.
2. **Бриллинджер Д.** Временные ряды. Обработка данных и теория. М.: Мир, 1980.
3. **Журбенко И. Г.** Анализ стационарных и однородных случайных систем. М.: Изд-во МГУ, 1987.
4. **Kay S. M.** Modern Spectral Estimation: Theory and Application. Englewood Cliffs (NJ): Prentice-Hall, 1988.
5. **Алексеев В. Г.** Ядра типа Джексона и Джексона – Валле-Пуссена и их вероятностные применения // Теория вероятностей и ее применения. 1996. **41**, № 1. С. 170.
6. **Натансон И. П.** Конструктивная теория функций. М.–Л.: Гостехтеориздат, 1949.
7. **Дзядык В. К.** Введение в теорию равномерного приближения функций полиномами. М.: Наука, 1977.

Институт физики атмосферы РАН

Поступила в редакцию
5 октября 2000 г.