УДК 519.234

ОЦЕНКА СПЕКТРАЛЬНОЙ ПЛОТНОСТИ ТИПА УЭЛЧА. СЛУЧАЙ НЕПРЕРЫВНОГО АРГУМЕНТА

В. Г. Алексеев, В. А. Суходоев

Институт физики атмосферы им. А. М. Обухова РАН,
119017, Москва, пер. Пыжевский, 3
E-mail: aleks.v.g@mail.ru

Рассмотрена оценка типа Уэлча (оценка, получаемая осреднением по сдвигу во времени) спектральной плотности стационарного случайного процесса с непрерывным временны́м аргументом. Исследуются смещение и дисперсия рассматриваемой статистической оценки, формулируются рекомендации по выбору ее параметров.

Ключевые слова: стационарный случайный процесс с непрерывным временем, оценка спектральной плотности типа Уэлча, B-сплайны Шенберга нечетных порядков.

Введение. Оценка спектральной плотности типа Уэлча была рассмотрена в [1, 2] для случая дискретного временно́го аргумента. В данной работе рассмотрим оценку типа Уэлча для непрерывного (пробегающего все вещественные значения) временно́го аргумента. При этом вместо разложений в ряд Фурье ядер типа Джексона, сыгравших ключевую роль в построении и исследовании оценки типа Уэлча в [1, 2], здесь используем B-сплайны Шенберга нечетных порядков, аргумент которых, как и аргумент исследуемого случайного процесса $X(t)$, пробегает все вещественные значения. Оценка типа Уэлча для случайного процесса $X(t)$ с непрерывным временем изучалась в монографиях [3, 4].

Цель данной работы — исследовать статистические свойства оценки типа Уэлча и сформулировать рекомендации по выбору ее параметров, сводящих к минимуму средний квадрат ошибки оценивания спектральной плотности.

1. Некоторые предварительные сведения. Говоря о B-сплайнах Шенберга нечетных порядков, имеем в виду функции $a_l(t)$, определяемые соотношением

$$a_l(t) = (2\pi)^{-1} \int \Phi_l(x)\, \mathrm{e}^{itx}\, dx, \qquad l = 1, 2, \ldots, \tag{1}$$

где

$$\Phi_l(x) = \Big[\frac{\sin(x/2)}{x/2}\Big]^{2l}. \tag{2}$$

В формуле (1) (и далее по тексту) интеграл без указания пределов означает интегрирование в пределах от $-\infty$ до ∞.

Все функции $a_l(t)$ финитны (носителем функции $a_l(t)$ является интервал $(-l, l)$), каждая из них есть сплайн порядка $2l-1$. Последнее означает, что функция $a_l(t)$ «склеена» из многочленов степени $2l-1$.

Точные выражения для функций $a_l(t)$ можно найти в работе [5] для $l = 1, 2, \ldots, 6$ и в работе [6] для $l = 7$. Формулы для функций $a_l(t)$, $l = 1$, $l = 2$, занимают одну и соответственно две строки, однако с каждым переходом от l к $l+1$ они становятся все более объемными. Воспроизводить их в предлагаемой работе не будем.

Далее для $l = 1, 2, \ldots, 14$ потребуются величины $\gamma_l = a_l(0)$. Они могут быть представлены следующими соотношениями:

$$\gamma_1 = 1, \quad \gamma_2 = \frac{2}{3} \approx 0{,}6667, \quad \gamma_3 = \frac{11}{20} = 0{,}5500, \quad \gamma_4 = \frac{151}{315} \approx 0{,}4794, \quad \gamma_5 = \frac{15\,619}{36\,288} \approx 0{,}4304,$$

$$\gamma_6 = \frac{655\,177}{1\,663\,200} \approx 0{,}3939, \quad \gamma_7 = \frac{27\,085\,381}{74\,131\,200} \approx 0{,}3654, \quad \gamma_8 = \frac{2\,330\,931\,341}{6\,810\,804\,000} \approx 0{,}3422,$$

$$\gamma_9 = \frac{12\,157\,712\,239}{37\,638\,881\,280} \approx 0{,}3230, \quad \gamma_{10} = \frac{37\,307\,713\,155\,613}{121\,645\,100\,408\,832} \approx 0{,}3067,$$

$$\gamma_{11} = \frac{339\,781\,108\,897\,078\,469}{1\,161\,157\,776\,629\,760\,000} \approx 0{,}2926, \quad \gamma_{12} = \frac{75\,489\,558\,096\,433\,522\,049}{269\,291\,841\,030\,051\,840\,000} \approx 0{,}2803,$$

$$\gamma_{13} = \frac{11\,482\,547\,005\,345\,338\,463\,969}{42\,613\,214\,404\,755\,456\,000\,000} \approx 0{,}2695,$$

$$\gamma_{14} = \frac{3\,607\,856\,726\,470\,666\,022\,715\,979}{13\,888\,864\,094\,921\,367\,552\,000\,000} \approx 0{,}2598.$$

Очевидно, что для $l = 1, 2, \ldots, 7$ вычисление каждой из величин γ_l каких-либо трудностей не представляет: достаточно положить $t = 0$ в формуле для функции $a_l(t)$. Что же касается величин γ_l, $l = 8, 9, \ldots, 14$, то каждая из них была вычислена из выражения

$$\gamma_l = a_l(0) = \int_{-m}^{m} a_7(t) a_m(t)\, dt, \qquad m = l - 7.$$

Кроме того, для всех четных $l \leqslant 14$ потребуются величины

$$a_l''(0) = -\frac{1}{2\pi} \int \Phi_l(x) x^2\, dx.$$

Приведем их приближенные значения:

$$a_2''(0) = -2, \quad a_4''(0) \approx -0{,}6667, \quad a_6''(0) \approx -0{,}3745, \quad a_8''(0) \approx -0{,}2472,$$

$$a_{10}''(0) \approx -0{,}1785, \quad a_{12}''(0) \approx -0{,}1367, \quad a_{14}''(0) \approx -0{,}1090.$$

Наконец, наряду с B-сплайнами Шенберга, будут полезны ядра Фейера и Джексона (первые два члена последовательности ядер типа Джексона), определяемые соответственно для всех натуральных n соотношениями

$$F_n(x) = \frac{1}{n} \left[\frac{\sin(nx/2)}{\sin(x/2)} \right]^2, \qquad J_n(x) = \frac{3}{2n^3 + n} \left[\frac{\sin(nx/2)}{\sin(x/2)} \right]^4.$$

2. Основные результаты. Оценка спектральной плотности типа Уэлча. Итак, пусть $\{X(t),\ -\infty < t < \infty\}$ — стационарный в широком смысле случайный процесс со средним $\langle X(t)\rangle \equiv 0$, корреляционной функцией $r(t) = \langle X(s)X(s+t)\rangle$ и спектральной плотностью

$$f(\omega) = \frac{1}{2\pi}\int \mathrm{e}^{-it\omega}\, r(t)\, dt, \qquad \omega \in \mathbf{R} = (-\infty, \infty). \tag{3}$$

В целях упрощения выкладок полагаем далее случайный процесс $X(t)$ гауссовым. Относительно спектральной плотности $f(\omega)$ будем предполагать, что она, по крайней мере дважды, дифференцируема, причем

$$\sup_{\omega} |f''(\omega)| < K < \infty. \tag{4}$$

Говоря об оценке спектральной плотности типа Уэлча, имеем в виду вычислительный алгоритм, включающий в себя разбиение всего интервала наблюдения длины M на конечное число неперекрывающихся или частично перекрывающихся сегментов длины $H < M$, вычисление периодограммы по каждому из них и ее последующее осреднение по числу интервалов длины H. При этом под периодограммой, вычисляемой по каждому сегменту длины H, понимается ее модифицированная версия, определяемая соотношением

$$I_H(\omega) = \Big|\int_{-H/2}^{H/2} a(t)X(t)\,\mathrm{e}^{it\omega}\ dt\Big|^2 \Big/ \Big(2\pi\int_{-H/2}^{H/2} a^2(t)\,dt\Big).$$

Здесь $a(t)$, $|t| \leqslant H/2$, — так называемое окно данных, используемое для домножения (неравномерного взвешивания) отрезка реализации $\{X(t),\ |t| \leqslant H/2\}$. Это окно чаще всего плавно убывает от середины отрезка реализации (т. е. от точки $t = 0$) к его краям, чем достигается сглаживание краев реализации и в конечном счете уменьшение смещения (систематической ошибки оценивания) периодограммы. Рекомендации предлагаемой работы, касающиеся выбора окна данных $a(t)$, сводятся к использованию B-сплайнов Шенберга $a_l(t)$. Основное достоинство B-сплайнов $a_l(t)$ состоит в том, что их преобразования Фурье $\Phi_l(x)$ имеют низкие боковые лепестки и быстро убывают с ростом абсолютной величины аргумента x (см. формулу (2)).

Пусть $\omega \geqslant 0$. Предлагаемая модификация оценки спектральной плотности типа Уэлча описывается соотношением

$$f_M(\omega) = \frac{1}{hQn}\sum_{t=-(n-1)}^{n-1}\Big(1 - \Big|\frac{t}{n}\Big|\Big)\ \Big|\int_{-H/2}^{H/2} a_l\Big(\frac{u}{h}\Big)X(u+tL)\,\mathrm{e}^{iu\omega}\ du\Big|^2. \tag{5}$$

Здесь H, h, n, L и l — параметры оценки (5); M — общая длина реализации, определяемая величиной $M = H + 2(n-1)L$;

$$Q = Q(l) = 2\pi\int_{-l}^{l} a_l^2(t)\,dt = \int \Phi_l^2(x)\,dx = \int \Phi_{2l}(x)\,dx\ = 2\pi\gamma_{2l}.$$

Разумеется, параметры H, h, n, L и l не могут быть выбраны произвольным образом. Они связаны следующими соотношениями и условиями: параметры n и l целочисленны,

$n > 1$, $1 \leqslant l \leqslant 7$, $0 < L \leqslant H = 2lh$. При этом неравенство $l \leqslant 7$ обусловлено тем обстоятельством, что для $l > 7$ B-сплайны $a_l(t)$ пока еще неизвестны.

Вычисляя математическое ожидание оценки (5) и принимая во внимание, что

$$\sum_{t=-(n-1)}^{n-1} \Bigl(1 - \Bigl|\frac{t}{n}\Bigr|\Bigr) = n,$$

приходим к соотношению

$$\langle f_M(\omega)\rangle = \int \Phi_{2l}(x) f\Bigl(\omega + \frac{x}{h}\Bigr)\, dx \Bigm/ \int \Phi_{2l}(x)\, dx.$$

Отсюда, применяя формулу Тейлора с остаточным членом в форме Лагранжа, получаем

$$\langle f_M(\omega)\rangle - f(\omega) = \int \Phi_{2l}(x) x^2 f''\Bigl(\omega + \frac{\theta x}{h}\Bigr)\, dx \Bigm/ \Bigl(2h^2 \int \Phi_{2l}(x)\, dx\Bigr),$$

где $0 < \theta = \theta(x) < 1$. Наконец, пользуясь предположением (4), приходим к неравенству

$$|\langle f_M(\omega)\rangle - f(\omega)| \leqslant K \int \Phi_{2l}(x)\, x^2\, dx \Bigm/ \Bigl(2h^2 \int \Phi_{2l}(x)\, dx\Bigr) = -\frac{K a''_{2l}(0)}{2h^2 \gamma_{2l}}. \tag{6}$$

Пусть $\varepsilon_l = -a''_{2l}(0)/\gamma_{2l}$. Используя результаты разд. 1, находим

$$\varepsilon_1 = 3, \quad \varepsilon_2 \approx 1{,}3907, \quad \varepsilon_3 \approx 0{,}9507, \quad \varepsilon_4 \approx 0{,}7224,$$

$$\varepsilon_5 \approx 0{,}5820, \quad \varepsilon_6 \approx 0{,}4877, \quad \varepsilon_7 \approx 0{,}4196.$$

Это значит, что при фиксированном значении параметра h полученная оценка сверху для смещения оценки $f_M(\omega)$ существенно убывает с ростом l.

Переходя к рассмотрению дисперсии оценки (5) и принимая во внимание предположение о гауссовости случайного процесса $X(t)$, получаем

$$\mathbf{D} f_M(\omega) \equiv \langle f_M^2(\omega)\rangle - |\langle f_M(\omega)\rangle|^2 = \frac{1}{(hQn)^2} \sum_{t=-(n-1)}^{n-1} \Bigl(1 - \Bigl|\frac{t}{n}\Bigr|\Bigr) \sum_{s=-(n-1)}^{n-1} \Bigl(1 - \Bigl|\frac{s}{n}\Bigr|\Bigr) \times$$

$$\times \int_{-H/2}^{H/2} a_l\Bigl(\frac{u}{h}\Bigr)\, du \int_{-H/2}^{H/2} a_l\Bigl(\frac{p}{h}\Bigr) \mathrm{e}^{i\omega(u-p)}\, dp \int_{-H/2}^{H/2} a_l\Bigl(\frac{v}{h}\Bigr)\, dv \int_{-H/2}^{H/2} a_l\Bigl(\frac{q}{h}\Bigr) \mathrm{e}^{i\omega(q-v)} \times$$

$$\times \bigl[\langle X(u+tL)X(v+sL)\rangle \langle X(p+tL)X(q+sL)\rangle +$$

$$+ \langle X(u+tL)X(q+sL)\rangle \langle X(p+tL)X(v+sL)\rangle\bigr]\, dq = D_1 + D_2, \tag{7}$$

где величина D_i, $i = 1, 2$, возникает за счет i-го слагаемого в квадратных скобках под знаком интеграла.

В соответствии с формулой

$$r(t) = \langle X(s)X(s+t)\rangle = \int \mathrm{e}^{it\omega}\, f(\omega)\, d\omega,$$

полученной в качестве обращения (3), каждый из сомножителей $\langle X(u+tL)X(v+sL)\rangle$ и $\langle X(p+tL)X(q+sL)\rangle$ в (7) может быть выражен через спектральную плотность $f(\omega)$. Это позволяет представить величину D_1 (первое из двух слагаемых в формуле (7)) в следующем виде:

$$D_1 = \frac{1}{(hQn)^2}\int f(\mu)\, d\mu \int\limits_{-H/2}^{H/2} a_l\Big(\frac{u}{h}\Big)\, du \int\limits_{-H/2}^{H/2} a_l\Big(\frac{v}{h}\Big)\, \mathrm{e}^{i(\mu-\omega)(v-u)}\, dv \times$$

$$\times \int f(\theta)\, d\theta \int\limits_{-H/2}^{H/2} a_l\Big(\frac{p}{h}\Big)\, dp \int\limits_{-H/2}^{H/2} a_l\Big(\frac{q}{h}\Big)\, \mathrm{e}^{i(\omega+\theta)(q-p)} \times$$

$$\times \Big[\sum_{t=-(n-1)}^{n-1}\Big(1-\Big|\frac{t}{n}\Big|\Big)\sum_{s=-(n-1)}^{n-1}\Big(1-\Big|\frac{s}{n}\Big|\Big)\, \mathrm{e}^{i(s-t)L(\mu+\theta)}\Big]\, dq. \qquad (8)$$

Заметим, что

$$\int\limits_{-H/2}^{H/2} a_l\Big(\frac{u}{h}\Big)\, \mathrm{e}^{i(\omega-\mu)u}\, du = \int\limits_{-H/2}^{H/2} a_l\Big(\frac{v}{h}\Big)\, \mathrm{e}^{i(\mu-\omega)v}\, dv = h\Phi_l(h(\mu-\omega)),$$

$$\int\limits_{-H/2}^{H/2} a_l\Big(\frac{p}{h}\Big)\, \mathrm{e}^{-i(\theta+\omega)p}\, dp = \int\limits_{-H/2}^{H/2} a_l\Big(\frac{q}{h}\Big)\, \mathrm{e}^{i(\theta+\omega)q}\, dq = h\Phi_l(h(\theta+\omega)),$$

$$\sum_{s=-(n-1)}^{n-1} \mathrm{e}^{isL(\mu+\theta)}\Big(1-\Big|\frac{s}{n}\Big|\Big) = \frac{1}{n}\Big\{\frac{\sin[n(L/2)(\mu+\theta)]}{\sin[(L/2)(\mu+\theta)]}\Big\}^2 = F_n(L(\mu+\theta))$$

(см., например, [7, упр. 1.7.12]),

$$F_n^2(L(\mu+\theta)) = \frac{2n^2+1}{3n}\, J_n(L(\mu+\theta))$$

и $\Phi_l^2(x) = \Phi_{2l}(x)$ для всех x.

Поэтому формула (8) может быть переписана в виде

$$D_1 = \frac{h^2(2n^2+1)}{3Q^2n^3}\int f(\mu)\Phi_{2l}(h(\mu-\omega))\, d\mu \int f(\theta)\Phi_{2l}(h(\omega+\theta))J_n(L(\mu+\theta))\, d\theta.$$

Отсюда после ряда тождественных преобразований приходим к соотношению

$$D_1 = \frac{2\pi h^2(2n^2+1)}{3Q^2Ln^3} \int \Phi_{2l}(h\mu) f(\mu+\omega)\, d\mu \times$$

$$\times \int\limits_{-\pi}^{\pi} \frac{1}{2\pi} J_n(\nu) \Big[\sum_{k=-\infty}^{\infty} \Phi_{2l}\Big(h\Big(\frac{\nu+2k\pi}{L} - \mu\Big)\Big) f\Big(\frac{\nu+2k\pi}{L} - \mu - \omega\Big)\Big]\, d\nu.$$

Дальнейший анализ полученного выражения требует предположения, что $h \gg 1$ и $Ln \gg 12lh$. Пользуясь свойствами сингулярного интеграла Джексона (см. [8, гл. 4, § 2, формулы (96) и (99)] или [9]), нетрудно показать, что в этом случае

$$D_1 \approx \frac{2\pi h^2(2n^2+1)}{3Q^2Ln^3} \int \Phi_{2l}(h\mu) f(\mu+\omega) \Big[\sum_{k=-\infty}^{\infty} \Phi_{2l}\Big(h\Big(\mu - \frac{2k\pi}{L}\Big)\Big) f\Big(\mu + \omega - \frac{2k\pi}{L}\Big)\Big]\, d\mu. \quad (9)$$

Если, кроме того, $2L < h$, то главный лепесток функции $\Phi_{2l}(h\mu)$ не пересекается с главными лепестками функций $\Phi_{2l}(h(\mu - 2k\pi/L))$, $k \neq 0$, что позволяет ввести дальнейшее упрощение в формулу (9) для величины D_1. Сохраняя лишь нулевое слагаемое в квадратных скобках под знаком интеграла, приходим к соотношению

$$D_1 \approx \frac{2\pi h^2(2n^2+1)}{3Q^2Ln^3} \int \Phi_{2l}^2(h\mu) f^2(\omega+\mu)\, d\mu = \frac{2\pi h(2n^2+1)}{3Q^2Ln^3} \int \Phi_{4l}(\nu) f^2\Big(\omega + \frac{\nu}{h}\Big)\, d\nu,$$

из которого следует, что при достаточно больших h

$$D_1 \approx \frac{2\pi h(2n^2+1)}{3Q^2Ln^3} f^2(\omega) \int \Phi_{4l}(\nu)\, d\nu = \frac{h(2n^2+1)\gamma_{4l}}{3Ln^3\gamma_{2l}^2} f^2(\omega) \approx \frac{2h\gamma_{4l}}{3nL\gamma_{2l}^2} f^2(\omega).$$

Исследуя аналогичным образом величину D_2 (второе слагаемое в формуле (7)), находим, что $D_2 = D_1$, если $\omega = 0$, и $D_2 = o(h/nL)$ при $h/nL \to 0$, если $\omega \neq 0$.

Теперь можно утверждать, что

$$\mathbf{D} f_M(\omega) = D_1 + D_2 \approx \frac{2h(2 - \operatorname{sign}\omega)\gamma_{4l}}{3nL\gamma_{2l}^2} f^2(\omega), \quad (10)$$

если $h \gg 1$, $Ln \gg 12lh$ и $2L < h$ (напомним также, что по предположению $\omega \geqslant 0$).

Приведем приближенные значения отношения $\delta_l = \gamma_{4l}/\gamma_{2l}^2$ для трех минимальных значений l:

$$\delta_1 \approx 1{,}0785, \qquad \delta_2 \approx 1{,}4890, \qquad \delta_3 \approx 1{,}8066.$$

Видно, что при фиксированном значении отношения h/nL величина $\mathbf{D} f_M(\omega)$ заметно возрастает с ростом l.

Формулу (10) для дисперсии оценки $f_M(\omega)$ следует рассматривать вместе с формулой (6), дающей оценку сверху для ее смещения. Заметим, что параметры n и L оценки $f_M(\omega)$ входят в (10) лишь в виде произведения nL, которое приближенно равно $M/2$, если $M \gg H$. Поэтому от выбора каждого из них в отдельности мало что зависит (до тех пор, пока остается в силе неравенство $2L < h$). Что же касается параметров h и l, то от них зависит очень многое. Как было установлено выше (см. формулу (6)), с ростом l существенно убывает смещение оценки $f_M(\omega)$. Параметр h следует по возможности выбирать таким образом, чтобы дисперсия оценки $f_M(\omega)$ (правая часть формулы (10)) и квадрат ее смещения были одного порядка малости. Тем самым при заданных значениях параметров M и l сводится к минимуму средний квадрат ошибки оценивания спектральной плотности $f(\omega)$. Наиболее общая закономерность может быть сформулирована следующим образом: с ростом длины реализации M при разумном выборе параметра h постепенно возрастает оптимальное (сводящее к минимуму средний квадрат ошибки оценивания функции $f(\omega)$) значение параметра l.

Кроме того, необходимо иметь в виду, что при фиксированном значении параметра H время счета по формуле (5) в первом приближении пропорционально n (т. е. обратно пропорционально параметру L). Поэтому не следует выбирать параметр L слишком малым, хотя неравенство $2L < h$ должно быть сохранено.

Заключение. Предложенная в данной работе оценка спектральной плотности $f(\omega)$ легко может быть реализована на персональном компьютере. Ее вычисление не требует больших затрат времени. Располагая достаточно длинной реализацией стационарного случайного процесса $X(t)$ и следуя сформулированным выше рекомендациям относительно выбора параметров l, h и L, исследователь получает возможность построить высококачественную оценку интересующей его спектральной плотности.

Нужно иметь в виду, что формула (10) для дисперсии оценки $f_M(\omega)$ получена в предположении, что исследуемый случайный процесс $X(t)$ гауссов. Если же последнее предположение не реализуется, могут возникнуть дополнительные слагаемые в правой части формулы (10), однако решающего значения они, скорее всего, иметь не будут. В любом случае дисперсия оценки $f_M(\omega)$ будет стремиться к нулю, если $h \gg 1$ и $h/nL \to 0$.

СПИСОК ЛИТЕРАТУРЫ

1. **Алексеев В. Г.** Оценка спектральной плотности типа Уэлча. Случай дискретного аргумента // Автометрия. 2001. № 6. С. 91–97.
2. **Алексеев В. Г., Суходоев В. А.** Уточненный метод оценивания спектральной плотности типа Уэлча // Автометрия. 2008. **44**, № 4. С. 23.
3. **Yaglom A. M.** Correlation Theory of Stationary and Related Random Functions. N. Y.: Springer-Verlag, 1987. Vols. I, II.
4. **Wirshing P. H., Paez T. L., Ortiz K.** Random Vibrations. Theory and Practice. N. Y.: John Wiley and Sons, 1995.
5. **Алексеев В. Г.** Новые непрерывные фильтры нижних частот // Радиотехника. 1998. № 4. С. 34.
6. **Алексеев В. Г.** Новый аналоговый линейный фильтр нижних частот // Радиотехника. 2005. № 10. С. 143.
7. **Бриллинджер Д.** Временны́е ряды. Обработка данных и теория. М.: Мир, 1980.
8. **Натансон И. П.** Конструктивная теория функций. М.—Л.: Гостехтеориздат, 1949.
9. **Ефимов А. В.** Джексона сингулярный интеграл // Математическая энциклопедия. Т. 2 /Под ред. И. М. Виноградова. М.: Сов. энц., 1979. С. 110.

Поступила в редакцию 14 октября 2008 г.