ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

УДК 519.22

ОЦЕНКА ОПТИМАЛЬНОГО ПОРЯДКА РИСКА ОБРАБОТКИ ВЕЙВЛЕТ-КОЭФФИЦИЕНТОВ, ОСНОВАННОГО НА ВЕРОЯТНОСТЯХ ОШИБОК[1]

Кудрявцев А.А., Шестаков О.В.
МГУ им. М.В. Ломоносова, г. Москва

Поступила в редакцию 15.02.2016, после переработки 20.02.2016.

В работе рассматривается метод построения оценки функции сигнала по зашумленным данным, основанный на минимизации вероятностей ошибок вычисления вейвлет-коэффициентов. В модели с аддитивным гауссовским шумом вычисляются оптимальные параметры метода и оценивается порядок риска.

Ключевые слова: вейвлеты, пороговая обработка, риск.

Вестник ТвГУ. Серия: Прикладная математика. 2016. № 1. С. 5–12.

Введение

Одной из наиболее актуальных задач обработки сигналов является подавление шума. В последнее время для решения этой задачи все чаще применяются методы, основанные на вейвлет-разложениях функции сигнала, поскольку вейвлет-разложение обеспечивает экономное представление полезного сигнала. Для подавления шума весьма популярными стали методы пороговой обработки, ориентированные на минимизацию среднеквадратичного риска или оценки этого риска (см. [1–3]). Статистические свойства таких методов и оценок среднеквадратичного риска подробно исследованы в работах [4–6]. В работе [7] был предложен альтернативный подход, заключающийся в масштабировании вейвлет-коэффициентов и минимизации функции потерь, основанной на вероятностях ошибок их вычисления. В данной работе оценивается порядок риска, выраженного через вероятности ошибок вычисления вейвлет-коэффициентов, и вычисляются оптимальные параметры масштабирования.

1. Вейвлет-разложение и модель данных

Пусть функция f, описывающая сигнал, задана на конечном отрезке $[a, b]$ и равномерно регулярна по Липшицу с некоторым показателем $\gamma > 0$: $f \in \mathrm{Lip}(\gamma)$.

[1]Работа выполнена при финансовой поддержке РФФИ (проект №16-07-00736).

Пусть задана вейвлет-функция ψ, которая M раз непрерывно дифференцируема ($M \geq \gamma$), имеет M нулевых моментов, то есть

$$\int\limits_{-\infty}^{\infty} x^k \psi(x) dx = 0, \quad k = 0, \ldots, M-1,$$

и быстро убывает на бесконечности вместе со своими производными, т. е. для всех $0 \leq k \leq M$ и любого $m \in \mathbf{N}$ найдется константа C_m, такая что при всех $x \in \mathbf{R}$

$$\left|\psi^{(k)}(x)\right| \leq \frac{C_m}{1 + |x|^m}.$$

При выполнении некоторых дополнительных условий (см. [8]) семейство $\{\psi_{j,k}\}_{j,k\in\mathbf{Z}}$, где $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, образует ортонормированный базис в $L^2(\mathbf{R})$. Индекс j называется масштабом, а индекс k – сдвигом. Вейвлет-разложение представляет собой ряд

$$f = \sum_{j,k\in\mathbf{Z}} \langle f, \psi_{j,k}\rangle \psi_{j,k}.$$

Известно [8], что при выполнении указанных выше условий найдется такая константа $C_f > 0$, что

$$|\langle f, \psi_{j,k}\rangle| \leq \frac{C_f}{2^{j(\gamma+1/2)}}. \tag{1}$$

При цифровой обработке сигналов функция f задана в дискретных отсчетах. Для применения дискретного вейвлет-разложения число этих отсчетов должно равняться 2^J для некоторого $J > 0$. Дискретное вейвлет-разложение представляет собой умножение вектора значений функции f на ортогональную матрицу, определяемую вейвлет-функцией ψ. При этом дискретные вейвлет-коэффициенты задаются соотношением $\mu_{j,k} = 2^{J/2}\langle f, \psi_{j,k}\rangle$ (см., например, [8]).

Практически всегда в сигнале присутствует шум. В данной работе рассматривается модель сигнала с аддитивным шумом:

$$X_i = f_i + w_i, \quad i = 1, \ldots, 2^J,$$

где f_i – отсчеты полезного сигнала, а w_i – независимые случайные величины, имеющие нормальное распределение с нулевым средним и дисперсией σ^2. В силу ортогональности матрицы дискретного вейвлет-разложения для эмпирических дискретных вейвлет-коэффициентов получается следующая модель:

$$Y_{j,k} = \mu_{j,k} + W_{j,k}, \;\; j = 0, \ldots, J-1, \; k = 0, \ldots, 2^j - 1,$$

где случайные величины $W_{j,k}$ также независимы и имеют нормальное распределение с нулевым средним и дисперсией σ^2.

2. Метод подавление шума

Для подавления шума обычно используется пороговая обработка, смысл которой заключается в обнулении коэффициентов, чьи абсолютные значения не превышают заданного порога. Среднеквадратичный риск таких методов имеет меньший порядок, чем у линейных методов подавления шума (см. [2]). Асимптотический порядок риска, статистические свойства его оценок, а также методы выбора порога подробно исследованы (см., например, [1–6]). В работе [7] предложен альтернативный метод масштабирования вейвлет-коэффициентов и экспериментально показано, что при минимизации вероятностей ошибок вычисления коэффициентов этот метод в некоторых случаях превосходит пороговую обработку по критерию отношения сигнал/шум.

Обозначим через $\hat{Y}_{j,k}$ оценку вейвлет-коэффициента, которая получается путем умножения на «масштабирующий» множитель $a_{j,k} \in [0,1]$: $\hat{Y}_{j,k} = a_{j,k}Y_{j,k}$.

Рассмотрим риск, основанный на вероятностях ошибок вычисления вейвлет-коэффициентов и определяемый для заданного $\varepsilon > 0$ как

$$r_J(f) = \frac{\sum\limits_{j=0}^{J-1}\sum\limits_{k=0}^{2^j-1} \mathsf{P}\left(\left|\hat{Y}_{j,k} - \mu_{j,k}\right| > \varepsilon\right)}{2^J}.$$

Такое определение риска является обобщением определения, предложенного в [7].

Целью данной работы является вычисление оптимальных множителей $a_{j,k}$ и оценивание максимального порядка риска r_J пороговой обработки наблюдаемого сигнала в классе функций $f \in \mathrm{Lip}(\gamma)$:

$$R_J = \sup_{f \in \mathrm{Lip}(\gamma)} r_J(f). \tag{2}$$

В большинстве следующих соотношений рассматривается только порядок величины по J. Поэтому для сокращения записи будем использовать обозначение C для некоторых положительных констант, которые могут зависеть от параметров модели, но не зависят от J. При этом в разных соотношениях одними и теми же буквами могут обозначаться различные константы.

3. Основные результаты

Пусть функция $g_1(J) > 0$ сколь угодно медленно убывает по J к нулю, $g_2(J) > 0$ сколь угодно медленно неограниченно возрастает по J.

Неравенство (1) дает возможность разбить множество индексов $\{0,\ldots,J-1\}$ на три класса в зависимости от величины $|\mu_{j,k}|$. Пусть индексы j_1 и j_2 ($j_1 < j_2$) таковы, что:

$$|\mu_{j,k}| \le C(g_1(J))^{-(\gamma+1/2)}, \quad j_1 \le j \le j_2 - 1;$$

$$|\mu_{j,k}| \le C(g_2(J))^{-(\gamma+1/2)}, \quad j_2 \le j \le J - 1.$$

При этом в силу (1)

$$j_i = \frac{J}{2\gamma+1} + \log_2 g_i(J) + C, \quad i = 1,2. \tag{3}$$

Разобьем сумму в числителе (2) на три составляющие:

$$\sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1}\mathsf{P}\left(\left|\hat{Y}_{j,k}-\mu_{j,k}\right|>\varepsilon\right)=\sum_{j=0}^{j_1-1}\sum_{k=0}^{2^j-1}\mathsf{P}\left(\left|\hat{Y}_{j,k}-\mu_{j,k}\right|>\varepsilon\right)+$$
$$+\sum_{j=j_1}^{j_2-1}\sum_{k=0}^{2^j-1}\mathsf{P}\left(\left|\hat{Y}_{j,k}-\mu_{j,k}\right|>\varepsilon\right)+\sum_{j=j_2}^{J-1}\sum_{k=0}^{2^j-1}\mathsf{P}\left(\left|\hat{Y}_{j,k}-\mu_{j,k}\right|>\varepsilon\right)\equiv$$
$$\equiv S_1+S_2+S_3. \tag{4}$$

Найдем точку минимума по $a_{j,k}$ слагаемого в (4), воспользовавшись равенством

$$\mathsf{P}\left(\left|\hat{Y}_{j,k}-\mu_{j,k}\right|>\varepsilon\right)=\Phi\left(\frac{-\varepsilon-(a_{j,k}-1)\mu_{j,k}}{a_{j,k}\sigma}\right)+\Phi\left(\frac{-\varepsilon+(a_{j,k}-1)\mu_{j,k}}{a_{j,k}\sigma}\right). \tag{5}$$

Имеем

$$\left(\mathsf{P}\left(\left|\hat{Y}_{j,k}-\mu_{j,k}\right|>\varepsilon\right)\right)'_{a_{j,k}}=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(a_{j,k}-1)^2\mu_{j,k}^2+\varepsilon^2}{2a_{j,k}^2\sigma^2}\right\}\times$$
$$\times\left[\frac{\varepsilon-\mu_{j,k}}{a_{j,k}^2\sigma}\exp\left\{-\frac{\varepsilon(a_{j,k}-1)\mu_{j,k}}{a_{j,k}^2\sigma^2}\right\}+\frac{\varepsilon+\mu_{j,k}}{a_{j,k}^2\sigma}\exp\left\{\frac{\varepsilon(a_{j,k}-1)\mu_{j,k}}{a_{j,k}^2\sigma^2}\right\}\right]. \tag{6}$$

Поэтому для существования нуля производной необходимо выполнение условия

$$\frac{\mu_{j,k}-\varepsilon}{\mu_{j,k}+\varepsilon}=\exp\left\{\frac{2\varepsilon(a_{j,k}-1)\mu_{j,k}}{a_{j,k}^2\sigma^2}\right\}. \tag{7}$$

Заметим, что при $|\mu_{j,k}|\le\varepsilon$ условие (7) не выполняется, поэтому минимум (5) достигается при $a_{j,k}=0$ и равен нулю.

При $|\mu_{j,k}|>\varepsilon$ и $a_{j,k}\ne 0$ (в точке $a_{j,k}=0$ в этом случае слагаемые в (4) принимают максимальные значения) условие (7) эквивалентно

$$p(a_{j,k})\equiv\sigma^2\ln\frac{\mu_{j,k}-\varepsilon}{\mu_{j,k}+\varepsilon}\cdot a_{j,k}^2-2\varepsilon\mu_{j,k}\cdot a_{j,k}+2\varepsilon\mu_{j,k}=0,$$

поэтому нули производной достигаются в точках

$$a_{j,k}^{(\pm)}=\frac{\varepsilon\mu_{j,k}\pm\sqrt{\varepsilon^2\mu_{j,k}^2-2\varepsilon\mu_{j,k}\sigma^2\ln\frac{\mu_{j,k}-\varepsilon}{\mu_{j,k}+\varepsilon}}}{\sigma^2\ln\frac{\mu_{j,k}-\varepsilon}{\mu_{j,k}+\varepsilon}}, \tag{8}$$

причем подкоренное выражение всегда положительно.

Заметим, что координата абсцисс вершины параболы $p(a_{j,k})$ всегда отрицательна, а $p(0)$ и $p(1)$ всегда имеют разные знаки. Из этого следует, что наименьший из корней (8) всегда отрицательный, а наибольший – всегда лежит строго между нулем и единицей и является точкой минимума, поскольку производная (6) переходит в этой точке из отрицательной области в положительную.

Отрицательным значениям $\mu_{j,k}$ соответствует точка минимума $a^{(+)}_{j,k}$, а положительным – точка минимума $a^{(-)}_{j,k}$. Очевидно, что

$$\lim_{\mu_{j,k}\to+\infty} a^{(-)}_{j,k} = 1 \text{ и } \lim_{\mu_{j,k}\to-\infty} a^{(+)}_{j,k} = 1.$$

Поэтому для любого $\delta \in (0,1)$ и достаточно больших $|\mu_{j,k}|$ выполняется неравенство $a^{(\pm)}_{j,k} > 1-\delta$, а следовательно

$$\mathsf{P}\left(\left|\hat{Y}_{j,k} - \mu_{j,k}\right| > \varepsilon\right) \geq \Phi\left(-\frac{\varepsilon}{(1-\delta)\sigma}\right). \tag{9}$$

Воспользуемся изложенными результатами для определения оптимального порядка риска (2). Рассмотрим S_3 из (4). Зафиксируем некоторое положительное число ε. Заметим, что для любого $\varepsilon > 0$ найдется такое $J_0 = J_0(\varepsilon)$, что $Cg_2(J)^{-(\gamma+1/2)} \leq \varepsilon$ для всех $J > J_0$. Следовательно имеет место соотношение $|\mu_{j,k}| \leq \varepsilon$ для $j_2 \leq j \leq J-1$. При этом минимизирующими компоненту риска S_3 коэффициенты $a_{j,k} = 0$ обеспечивают равенство $S_3 = 0$.

Предположив, что все слагаемые в суммах S_1 и S_2 из (4) отделены от нуля некоторой константой, получаем, что наихудшим порядком (верхней границей) в минимаксном смысле для риска (2) является значение порядка $(S_1+S_2)/2^J$, то есть 2^{j_2-J}, где j_2 определено в (3).

Теперь найдем нижнюю границу для риска (2). Заметим что для любой константы C_f из (1) найдется такая функция $f \in \mathrm{Lip}(\gamma)$, что в неравенстве (1) будет достигаться равенство для $0 \leq j \leq j_1-1$ (см. [8]). Следовательно, существует такое $J_1 > 0$, что для всех $\varepsilon > 0$, $\delta \in (0,1)$ и $J > J_1$ выполняется $|\mu_{j,k}| > \varepsilon$, $a^{(\pm)}_{j,k} > 1-\delta$ и соотношение (9) при $0 \leq j \leq j_1 - 1$. В этом случае порядок суммы S_1 в (4) равен числу слагаемых, то есть 2^{j_1}, где j_1 определено в (3). Поскольку сумма S_2 может лишь ухудшить порядок риска, истинное значение риска R_J имеет порядок не ниже 2^{j_1-J}, то есть рассматриваемый порядок является нижней оценкой для истинного порядка риска.

Приведенные рассуждения позволяют сформулировать следующее утверждение.

Теорема 1. *При выборе оптимальных «масштабирующих» множителей $a_{j,k}$ риск (2) удовлетворяет неравенствам*

$$C_m 2^{-\frac{2\gamma}{2\gamma+1}J} g_1(J) \leq R_J \leq C_M 2^{-\frac{2\gamma}{2\gamma+1}J} g_2(J),$$

где C_m и C_M – некоторые положительные константы.

Замечание 1. Выражение (2) и, следовательно, оптимальные множители $a_{j,k}$ зависят от неизвестных величин $\mu_{j,k}$. Для практического применения в определении риска (2) величины $\mu_{j,k}$ заменяют на оценки максимального правдоподобия (то есть $Y_{j,k}$) и вычисляют оптимальные множители $a_{j,k}$ для получившейся оценки риска (см. [7]).

Заключение

Рассмотрен метод построения оценки функции сигнала по зашумленным данным, основанный на минимизации вероятностей превышения ошибками вычисления вейвлет-коэффициентов заданного критического уровня. В модели с аддитивным гауссовским шумом вычисляются оптимальные «масштабирующие» множители и оценивается порядок риска.

Список литературы

[1] Donoho D., Johnstone I. Adapting to unknown smoothness via wavelet shrinkage // Journal of the American Statistical Association. 1995. Vol. 90. Pp. 1200–1224.

[2] Donoho D., Johnstone I.M. Minimax estimation via wavelet shrinkage // Annals of Statistics. 1998. Vol. 26, no. 3. Pp. 879–921.

[3] Jansen M. Noise Reduction by Wavelet Thresholding. Springer New York, 2001. doi:10.1007/978-1-4613-0145-5

[4] Marron J.S., Adak S., Johnstone I.M., Neumann M.H., Patil P. Exact risk analysis of wavelet regression // Journal of Computational and Graphical Statistics. 1998. Vol. 7. Pp. 278–309.

[5] Маркин А.В., Шестаков О.В. О состоятельности оценки риска при пороговой обработке вейвлет-коэффициентов // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. 2010. № 1. С. 26–34.

[6] Шестаков О.В. Асимптотическая нормальность оценки риска пороговой обработки вейвлет-коэффициентов при выборе адаптивного порога // Доклады Академии наук. 2012. Т. 445, № 5. С. 513–515.

[7] Sadasivan J., Mukherjee S., Seelamantula C.S. An optimum shrinkage estimator based on minimum-probability-of-error criterion and application to signal denoising // Proceedings of the ICASSP. 2014. Pp. 4249–4253.

[8] Mallat S. A Wavelet Tour of Signal Processing. NY: Academic Press, 1999. 857 p.

Библиографическая ссылка

Кудрявцев А.А., Шестаков О.В. Оценка оптимального порядка риска обработки вейвлет-коэффициентов, основанного на вероятностях ошибок // Вестник ТвГУ. Серия: Прикладная математика. 2016. № 1. С. 5–12.

Сведения об авторах

1. **Кудрявцев Алексей Андреевич**
доцент кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М.В. Ломоносова.

Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В. Ломоносова.

2. **Шестаков Олег Владимирович**
доцент кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М.В. Ломоносова, старший научный сотрудник Института проблем информатики ФИЦ ИУ РАН.

Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В. Ломоносова.

ESTIMATION OF THE OPTIMAL RATE OF THE WAVELET THRESHOLDING RISK BASED ON THE ERROR PROBABILITIES

Kudryavtsev Alexey Andreevich
Associate Professor at Mathematical Statistics department, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University
Russia, 119992, Moscow, GSP-1, Vorobyovi gory, Lomonosov MSU.

Shestakov Oleg Vladimirovich
Associate Professor at Mathematical Statistics department, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University;
Senior Researcher at Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences
Russia, 119992, Moscow, GSP-1, Vorobyovi gory, Lomonosov MSU.

Received 15.02.2016, revised 20.02.2016.

We consider the method of estimating the signal function from the noised observations based on the minimization of the probabilities of errors in calculating the wavelet coefficients. In the model with additive Gaussian noise we find the optimal parameters and estimate the risk rate.

Keywords: wavelets, thresholding, risk.

Bibliographic citation

Kudryavtsev A.A., Shestakov O.V. Estimation of the optimal rate of the wavelet thresholding risk based on the error probabilities. *Vestnik TvGU. Seriya: Prikladnaya Matematika* [Herald of Tver State University. Series: Applied Mathematics], 2016, no. 1, pp. 5–12. (in Russian)

References

[1] Donoho D., Johnstone I. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 1995, vol. 90, pp. 1200–1224.

[2] Donoho D., Johnstone I.M. Minimax estimation via wavelet shrinkage. *Annals of Statistics*, 1998, vol. 26(3), pp. 879–921.

[3] Jansen M. *Noise Reduction by Wavelet Thresholding*. Springer New York, 2001. doi:10.1007/978-1-4613-0145-5

[4] Marron J.S., Adak S., Johnstone I.M., Neumann M.H., Patil P. Exact risk analysis of wavelet regression. *Journal of Computational and Graphical Statistics*, 1998, vol. 7, pp. 278–309.

[5] Markin A.V., Shestakov O.V. Consistency of risk estimation with thresholding of wavelet coefficients. *Moscow University Computational Mathematics and Cybernetics*, 2010, vol. 34(1), pp. 22–30.

[6] Shestakov O.V. Asymptotic normality of adaptive wavelet thresholding risk estimation. *Doklady Mathematics*, 2012, vol. 86(1), pp. 556–558.

[7] Sadasivan J., Mukherjee S., Seelamantula C.S. An optimum shrinkage estimator based on minimum-probability-of-error criterion and application to signal denoising. *Proceedings of the ICASSP*, 2014, pp. 4249–4253.

[8] Mallat S. *A Wavelet Tour of Signal Processing*. NY: Academic Press, 1999. 857 p.