ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

УДК 519.22

СОСТОЯТЕЛЬНОСТЬ ОЦЕНКИ РИСКА ПРИ МНОЖЕСТВЕННОЙ ПРОВЕРКЕ ГИПОТЕЗ С FDR-ПОРОГОМ

Заспа А.Ю.[*], **Шестаков О.В.**[*,**]
[*] МГУ им. М.В. Ломоносова, г. Москва
[**] Институт проблем информатики ФИЦ ИУ РАН, г. Москва

Поступила в редакцию 02.02.2017, после переработки 14.03.2017.

В работе рассматривается метод удаления шума в разреженных сигналах, основанный на процедуре множественной проверке нулевых гипотез с использованием FDR-порога. В модели с белым гауссовским шумом доказывается состоятельность несмещенной оценки среднеквадратичного риска.

Ключевые слова: множественная проверка гипотез, пороговая обработка, оценка риска, состоятельность.

Вестник ТвГУ. Серия: Прикладная математика. 2017. № 1. С. 5–16.

Введение

Во многих прикладных областях возникает потребность в обработке цифровых сигналов самых разных видов и происхождения. Одной из самых распространенных задач при этом является фильтрация шума, которая чаще всего становится отправной точкой для последующей обработки сигнала.

Для фильтрации сигнала обычно сначала осуществляется преобразование, приводящее к его экономному (разреженному) представлению, т.е. фактически разложение функции сигнала по некоторому базису, зависящему от конкретного класса рассматриваемых сигналов. Затем осуществляется пороговая обработка, обнуляющая часть коэффициентов разложения, которые считаются шумом. Фактически эта операция эквивалентна процедуре множественной проверки нулевых гипотез для набора коэффициентов сигнала.

Одной из основных проблем в этом подходе является выбор стратегии поиска порогового значения, для которого существуют различные методы вычисления, в основном ориентированные на минимизацию среднеквадратичной погрешности [1–4]. В данной работе рассматривается пороговая фильтрация с FDR-порогом для класса сигналов, вектор значений которых является разреженным. Такого рода сигналы часто возникают, например, после осуществления вейвлет-преобразования исходного сигнала. Доказывается, что в этом классе при мягкой

пороговой обработке с FDR-порогом оценка среднеквадратичного риска является состоятельной. Статистические свойства такой оценки при других стратегиях выбора порога изучены в работах [5–8].

1. Модель данных и пороговая обработка

Рассмотрим следующую модель отсчетов сигнала с шумом:

$$x_i = \mu_i + \varepsilon_i,\ i = 1, \ldots, n, \tag{1}$$

где μ_i – точные значения сигнала в узлах равномерной сетки, а ε_i – случайные величины, соответствующая шумовой компоненте измерения μ_i. Величины x_i будем называть зашумленными значениями исходной функции. Предполагается, что шум является белым гауссовским, т.е. ε_i независимы и имеют нормальное распределение с нулевым средним и дисперсией σ^2. Далее не ограничивая общности будем считать, что $\sigma^2 = 1$.

Введем обозначение для пространства разреженных векторов, которому принадлежат значения полезного сигнала μ_i:

$$l_0[\eta] = \{\mu \in \mathbb{R}^n : \|\mu\|_0 \leq \eta n\},$$

где $\|x\|_0 = \#\{i : x_i \neq 0\}$.

Для удаления шумовой компоненты воспользуемся алгоритмом мягкой пороговой обработки. Мягкой пороговой обработкой называется процедура оценивания сигнала, при которой к каждому отсчету x_i применяется пороговая функция, которая задается следующим образом:

$$\rho_t(x) = \begin{cases} x - t, & \text{если } x > t; \\ x + t, & \text{если } x < -t; \\ 0, & \text{если } |x| \leqslant t. \end{cases}$$

Риском обработки сигнала с порогом t назовем функцию

$$r(\mu, t) = \frac{1}{n} \sum_{i=1}^{n} r_i(\mu_i, t),$$

где $r_i(\mu_i, t)$ – индивидуальный риск (риск обработки отдельного отсчета). Индивидуальный риск в случае использования мягкой пороговой функции выглядит следующим образом:

$$r_i(\mu_i, t) = \mathsf{E}(\rho_t(x_i) - \mu_i)^2.$$

Помимо мягкой пороговой обработки часто используется жесткая пороговая обработка [9], при которой отсчет x_i обнуляется, если его абсолютное значение не превосходит порога, и остается неизменным в противном случае. При такой пороговой обработке используется разрывная пороговая функция, что приводит к появлению дополнительных артефактов в обрабатываемом сигнале. Обычно жесткая пороговая обработка используется для компрессии сигналов с потерями.

Минимальный риск при жесткой пороговой обработке равен [9]

$$r_{inf}(\mu) = \frac{1}{n}\sum_{i=1}^{n}\min(\mu_i^2, 1). \tag{2}$$

Значение $r_{inf}(\mu)$ и порог, при котором достигается это значение, вычислить на практике нельзя, поскольку они зависят от ненаблюдаемых чистых отсчетов сигнала.

Риск мягкой пороговой обработки $r(\mu, t)$ и порог t_{min}, минимизирующий этот риск, также зависят от неизвестных значений полезного сигнала, и на практике используют оценку риска при мягкой пороговой обработке, называемую $SURE$ (англ. Stein Unbiased Risk Estimator). Эта оценка является несмещенной и равна

$$\widehat{r}(x, t) = \frac{1}{n}\sum_{i=1}^{n}\Phi(x_i^2),$$

где

$$\Phi(x) = \begin{cases} x - 1, & \text{если } x \leqslant t^2; \\ 1 + t^2, & \text{если } x > t^2. \end{cases}$$

2. Универсальный порог

В работе [1] Донохо и Джонстон предложили использовать порог, называемый универсальным, который равен $t_U = \sqrt{2\ln n}$. Этот порог практически удаляет весь шум, поскольку

$$\mathsf{P}(\max_{1\leqslant i\leqslant n}|\varepsilon_i| > \sqrt{2\ln n}) \to 0 \text{ при } n \to \infty. \tag{3}$$

При этом значения отсчетов полезного сигнала затрагиваются не сильно.

Риск мягкой и жесткой пороговой обработки при использовании порога t_U близок к минимальному риску $r_{inf}(\mu)$ (2), и при $n \geq 4$ выполняется неравенство [9]

$$r(\mu, t_U) \leqslant (2\ln n + 1)\left(\frac{1}{n} + r_{inf}(\mu)\right). \tag{4}$$

Универсальный порог можно использовать практически в любых ситуациях, однако в некоторых случаях его значение оказывается слишком большим, и оценка сигнала получается слишком сглаженной.

3. FDR-порог

Одним из популярных методов построения порогового значения является использование алгоритмов множественной проверки гипотез. Так как x_i являются реализациями некоторых случайных величин, то можно каждой такой случайной величине $\mathsf{x_i}$, $i = 1, \ldots, n$ сопоставить нулевые гипотезы $\{H_{0_i},\ i = 1, \ldots, n\}$. Каждая нулевая гипотеза включает в себя предположение о распределении этой

Таблица 1: Результаты множественной проверки гипотез

	Число принятых гипотез	**Число отвергнутых гипотез**	**Всего**
Число верных гипотез	u	v	m_0
Число неверных гипотез	w	s	m_1
Всего	$n-k$	k	n

случайной величины и о том, что данное значение сигнала является полностью шумовым. Далее для имеющихся реализаций случайных величин используется некоторый алгоритм множественной проверки гипотез.

Таким образом, набору нулевых гипотез H_0 и реализациям случайных величин $X = \{x_1, x_2, \ldots, x_n\}$ ставится в соответствие набор p-значений $p = \{p_1, p_2, \ldots, p_n\}$. Так как на основе этих p-значений каждая нулевая гипотеза либо принимается, либо отвергается, то могут быть допущены ошибки второго и первого рода, соответственно. Число отклоненных нулевых гипотез k и число принятых $n-k$ являются наблюдаемыми случайными величинами, тогда как число отвергнутых неверных s, число отвергнутых верных v, число принятых неверных w, число принятых верных u из Таблицы 1 являются ненаблюдаемыми.

Задача множественной проверки гипотез состоит в том, чтобы выбрать такой метод, который минимизирует число ложных отклонений v и ложных принятий w. Существует несколько мер оценивающих число произошедших ошибок первого рода при использовании определенного алгоритма проверки гипотез. Одной из таких мер является **FDR**. FDR (англ. False Discovery Rate) – средняя доля ошибок первого рода среди всех отклоненных гипотез:

$$\mathbf{FDR} = \mathsf{E}\frac{v}{k}, \text{ при } k > 0.$$

Бенджамини и Хочберг разработали алгоритм множественной проверки гипотез [10], который позволяет ограничить сверху значение FDR. Ими была доказана теорема, что для независимых тестовых статистик $X = \{x_1, x_2, \ldots, x_n\}$ и для произвольной конфигурации неверных нулевых гипотез, описанная ими процедура позволяет контролировать FDR на уровне q, т.е.

$$\mathsf{E}\frac{v}{k} \leqslant q,$$

где v – число ошибок первого рода среди отклоненных гипотез.

Опишем саму процедуру:

– упорядочим по возрастанию p-значения;

– возьмем за k такое максимальное i, для которого выполняется

$$p_{(i)} \leqslant \frac{i}{n}\, q; \tag{5}$$

– затем отвергнем все нулевые гипотезы $H_{0_{(i)}}$ с номерами $i = 1, 2, \ldots, k$.

В качестве нулевых гипотез возьмем гипотезы о том, что коэффициенты сигнала x_i являются полностью шумовыми и, в соответствии с используемой моделью шума, имеют нормальное распределение с нулевым математическим ожиданием и единичной дисперсией.

В рамках принятой модели p-значения выражаются следующим образом:

$$p_i = 2\left[1 - \Phi\left(|x_i|\right)\right],$$

где $\Phi(x)$ – функция стандартного нормального распределения.

Далее из этих p-значений формируется вариационный ряд, для которого ищется k из условия (5). Сам порог выражается следующим образом:

$$t_F = \begin{cases} \Phi^{-1}\left(1 - \dfrac{p_{(k+1)}}{2}\right), & \text{если } k = 1, 2, \ldots, n-1; \\ 0, & \text{если } k = n; \\ \Phi^{-1}\left(1 - \dfrac{p_{(1)}}{2}\right), & \text{если нет } k, \text{ для которых выполнено (5)}, \end{cases}$$

где $\Phi^{-1}(x)$ – обратная функция к функции стандартного нормального распределения.

Модифицируем алгоритм получения FDR-порога, исходя из рассуждений, справедливых для универсального порога. Так как справедливо утверждение теоремы (3), то можно сделать вывод, что брать порог $t > \sqrt{2\ln n}$ нецелесообразно, поэтому под FDR-порогом далее будем понимать $\min(t_F, t_U)$ и обозначать его также через t_F.

Для случайных величин, вектор математических ожиданий которых принадлежит $l_0[\eta]$, в [11] получен важный результат, описывающий свойства FDR-порога.

Теорема 1. *Пусть* $\mathrm{x_i} \sim N(\mu_i, 1)$, $i = 1, \ldots, n$, $\mu \in l_0[\eta_n]$ *и* $\eta_n \in [n^{-1}(\ln n)^5, n^{-\delta}]$, *где* $0 < \delta < 1$. *Тогда существует такая константа* $c > 0$, *что для FDR-порога* t_F *с управляющим параметром* $q_n \xrightarrow{n} 0$, *построенного для реализаций этих случайных величин, при достаточно больших* n *выполняется неравенство*

$$\sup_{\mu \in l_0[\eta_n]} \mathsf{P}(t_F < t_1) \le 2n \exp\{-c\, q_n \kappa_n \gamma_n^2\},$$

где $t_1 \sim (2\ln \eta_n^{-1})^{1/2}$, $\gamma_n = \frac{1}{\ln\ln n}$, $\kappa_n = \frac{n\,\eta_n}{1 - q_n - \gamma_n}$.

4. Состоятельность оценки риска при использовании FDR-порога

Следующая теорема показывает, что при выборе FDR-порога оценка риска по вероятности сходится к минимальному риску мягкой пороговой обработки.

Теорема 2. *Пусть в модели (1) для вектора чистых отсчетов сигнала выполнено* $\mu \in l_0[\eta_n]$. *Пусть FDR-порог* t_F *вычисляется с управляющим параметром* q_n. *Если выполняется* $\eta_n \in [n^{-1}(\ln n)^5, n^{-\delta}]$, *где* $0 < \delta < 1$, *и* $q_n \xrightarrow{n} 0$ *так, что* $\frac{q_n \kappa_n \gamma_n^2}{\ln n} \to \infty$, *тогда*

$$\mathsf{P}(|\widehat{r}(x, t_F) - r(\mu, t_{min})| > \varepsilon) \to 0 \text{ при } n \to \infty.$$

Доказательство. Справедлива оценка

$$\mathsf{P}(|\widehat{r}(x,t_F)-r(\mu,t_{min})|>\varepsilon)\le$$
$$\le \mathsf{P}\left(|\widehat{r}(x,t_F)-r(\mu,t_F)|>\frac{\varepsilon}{2}\right)+\mathsf{P}\left(|r(\mu,t_F)-r(\mu,t_{min})|>\frac{\varepsilon}{2}\right).$$

Оценим сначала первую вероятность. Для удобства заменим $\varepsilon/2$ на ε. Тогда

$$\mathsf{P}(|\widehat{r}(x,t_F)-r(\mu,t_F)|>\varepsilon)\le$$
$$\le \mathsf{P}(|\widehat{r}(x,t_F)-r(\mu,t_F)|>\varepsilon, t_F\in[t_1,t_U])+\mathsf{P}(t_F\notin[t_1,t_U]).$$

По теореме 1 $\mathsf{P}(t_F\notin[t_1,t_U])\xrightarrow{n}0$, а первое слагаемое не превосходит

$$\mathsf{P}(\sup_{t\in[t_1,t_U]}|\widehat{r}(x,t)-r(\mu,t)|>\varepsilon). \tag{6}$$

Обозначим $\mathrm{Z}(t)=\widehat{r}(x,t)-r(\mu,t)$ и представим $\mathrm{Z}(t)$ в виде $\mathrm{Z}(t)=\frac{1}{n}\sum\limits_{i=1}^{n}\mathrm{Y}_i(t)$, где $\mathrm{Y}_i(t)$ – независимые случайные величины с нулевым математическим ожиданием. Несложно показать, что $|\mathrm{Y}_i(t)|\le 2+t^2$. Применяя неравенство Хёффдинга [12] для фиксированного t и произвольного $r_d>1$, имеем

$$\mathsf{P}\left(|\mathrm{Z}(t)|\ge r_d\,\frac{1}{\sqrt{n}}\right)\le 2\exp\left(-\frac{r_d^2}{2(t^2+2)^2}\right). \tag{7}$$

Возьмём $t'>t$ и обозначим $N(t,t')=\#\{i: t<|x_i|\le t'\}$. Для следующего шага оценим приращение $\widehat{r}(x,t)$:

$$|\widehat{r}(x,t)-\widehat{r}(x,t')|\le$$
$$\le\left|\frac{2}{n}\sum_{i=1}^{n}\mathbf{1}(t^2<x_i^2\le t'^2)\right|+\left|\frac{1}{n}\sum_{i=1}^{n}(x_i^2\wedge t^2-x_i^2\wedge t'^2)\right|\le$$
$$\le\frac{2}{n}N(t,t')+(t'^2-t^2). \tag{8}$$

Также можно оценить приращение $r(\mu,t)-r(\mu,t')$. Используя оценку для модуля производной

$$\left|\frac{\partial r(\mu,t)}{\partial t}\right|\le 4\,t_U,$$

справедливую при $t<t_U$ и достаточно больших n, и ограничив изменение аргумента $|t-t'|<\delta_n$, можно получить неравенство

$$|r(\mu,t)-r(\mu,t')|\le 4\,t_U\delta_n. \tag{9}$$

Объединив оценки (8) и (9), получим:

$$|Z(t)-Z(t')|\le\frac{2}{n}N(t,t')+(t'-t)(t'+t)+4t_U\delta_n\le\frac{2}{n}N(t,t')+6t_U\delta_n.$$

Теперь разобьем отрезок $[t_1, t_U]$ на равные части точками $t_j = t_1 + j\delta_n \in [t_1, t_U]$. Очевидно, выполняется

$$A_n = \left\{ \sup_{[t_1,t_U]} |Z(t)| \geq 3r_d \frac{1}{\sqrt{n}} \right\} \subset D_n \cup E_n \,,$$

где

$$D_n = \left\{ \sup_j |Z(t_j)| \geq r_d \frac{1}{\sqrt{n}} \right\}$$

и

$$E_n = \left\{ \sup_j \sup_{t\in[t_j,t_j+\delta_n)} |Z(t) - Z(t_j)| \geq 2r_d \frac{1}{\sqrt{n}} \right\}.$$

Выберем последовательность δ_n такую, чтобы $t_U\delta_n = o\left(\frac{1}{\sqrt{n}}\right)$. Тогда при достаточно больших n

$$E_n \subset E'_n = \left\{ \sup_j \frac{2}{n} N(t_j, t_j + \delta_n) \geq r_d \frac{1}{\sqrt{n}} \right\}, \text{ так как}$$

$$\left\{ \sup_j \sup_{t\in[t_j,t_j+\delta_n)} |Z(t) - Z(t_j)| \geq 2r_d \frac{1}{\sqrt{n}} \right\} \subset$$
$$\subset \left\{ \sup_j \frac{2}{n} N(t_j, t_j + \delta_n) \geq 2r_d \frac{1}{\sqrt{n}} - o\left(\frac{1}{\sqrt{n}}\right) \right\}.$$

Заметим, что $N(t,t') = \#\{i : t < |x_i| \leq t'\}$ – не что иное, как сумма индикаторов того, что случайная величина попала в заданный интервал: $N(t,t') = \sum\limits_{i=1}^{n} \mathbf{1}(t < |x_i| \leq t')$. Нетрудно также показать, что $\mathsf{E}\, N(t', t' + \delta_n) \leq \frac{2}{\sqrt{2\pi}} n\, \delta_n$, поэтому справедливо $\mathsf{E}\, N(t_j, t_j + \delta_n) = o(r_d\sqrt{n})$. Для события E'_n выполняется

$$E'_n \subset \left\{ \sup_j \left(N(t_j, t_j + \delta_n) - \mathsf{E}\, N(t_j, t_j + \delta_n)\right) \geq \frac{1}{2} r_d \sqrt{n} - o(r_d\sqrt{n}) \right\}.$$

Заметим, что, начиная с некоторого n, верно $\frac{1}{2}\, r_d\sqrt{n} - o(r_d\sqrt{n}) \geq \frac{1}{3}\, r_d\sqrt{n}$, поэтому можно продолжить цепочку вложенности событий

$$\left\{ \sup_j \left(N(t_j, t_j + \delta_n) - \mathsf{E}\, N(t_j, t_j + \delta_n)\right) \geq \frac{1}{2} r_d \sqrt{n} - o(r_d\sqrt{n}) \right\} \subset$$
$$\subset \left\{ \sup_j \frac{1}{n} \left|\left(N(t_j, t_j + \delta_n) - \mathsf{E}\, N(t_j, t_j + \delta_n)\right|\right. \geq \frac{1}{3\sqrt{n}} r_d \right\} = E''_n.$$

Применяя неравенство Хёффдинга, получаем

$$\mathsf{P}(E''_j) = \mathsf{P}\left(\frac{1}{n} \left|\left(N(t_j, t_j + \delta_n) - \mathsf{E}\, N(t_j, t_j + \delta_n)\right|\right. \geq \frac{1}{3\sqrt{n}} r_d \right) \leq 2e^{-\frac{2}{9} r_d^2},$$

и далее для события E_n''

$$\mathsf{P}(E_n'') \le 2\,\frac{t_U}{\delta_n}\,e^{-\frac{2}{9}r_d^2}.$$

Теперь оценим вероятность события D_n. Для каждой случайной величины $\mathrm{Y}_i(t_j)$ выполняется $|\mathrm{Y}_i(t_j)| \le 2 + t_U^2$, поэтому, используя неравенство (7), по аналогии с предыдущим неравенством получаем

$$\mathsf{P}(D_n) \le \sum_j \mathsf{P}\left(|Z_n(t_j)| \ge \frac{1}{\sqrt{n}} r_d\right) \le 2\,\frac{t_U}{\delta_n}\,\exp\left(-\frac{r_d^2}{2(t_U^2+2)^2}\right).$$

Наконец, оценим вероятность события A_n:

$$\mathsf{P}(A_n) \le \mathsf{P}(D_n) + \mathsf{P}(E_n'') \le 2\,\frac{t_U}{\delta_n}\left(\exp\left(-\frac{r_d^2}{2(t_U^2+2)^2}\right) + \exp\left(-\frac{2}{9}r_d^2\right)\right).$$

Возьмем в качестве r_d последовательность $t_U(t_U^2+2) = \sqrt{2\ln n}\,(t_U^2+2)$, члены которой, начиная с некоторого n больше 1. Тогда неравенство преобразуется к виду $\mathsf{P}(A_n) \le \frac{4t_U}{\delta_n n}$. Уточним длину отрезков разбиения $\delta_n = t_U/n^{\frac{3}{4}}$, при этом ранее заявленное условие $t_U\delta_n = o\left(\frac{1}{\sqrt{n}}\right)$ выполняется. Получаем

$$\mathsf{P}(A_n) \le \mathsf{P}\left(\sup_{[t_1,t_U]} |Z(t)| \ge \frac{6\sqrt{2\ln n}\,(\ln n + 1)}{\sqrt{n}}\right) = O\left(n^{-\frac{1}{4}}\right).$$

Таким образом вероятность (6) стремится к нулю при $n \to \infty$.

Оценим теперь вероятность $\mathsf{P}(|r(\mu,t_F) - r(\mu,t_{min})| > \varepsilon)$ по аналогии с (6). Исходя из этого, достаточно оценить

$$\mathsf{P}(\sup_{t\in[t_1,t_U]} |r(\mu,t) - r(\mu,t_{min})| > \varepsilon),$$

причем знак модуля можно убрать, так как любой риск больше минимального. Далее нам понадобится соотношение $0 \le r_i(\mu_i,t) - r_i(0,t) \le \mu_i^2$ [1]. Объединяя это неравенство с неравенством $r_i(\mu_i,t) \le 1 + t^2$, получаем

$$r_i(\mu_i,t) \le \min(r_i(0,t) + \mu_i^2, 1+t^2) \le r_i(0,t) + \min(\mu_i^2, 1+t^2).$$

Теперь для $\sup\limits_{t\in[t_1,t_U]} (r(\mu,t) - r(\mu,t_{min}))$ справедлива оценка

$$\sup_{t\in[t_1,t_U]} (r(\mu,t) - r(\mu,t_{min})) \le$$

$$\le \sup_{t\in[t_1,t_U]} \left(\frac{1}{n}\sum_{i=1}^{n} \left(r_i(0,t) + \min(\mu_i^2, 1+t^2)\right) - r(\mu,t_{min})\right).$$

Поскольку $\frac{\partial r_i(0,t)}{\partial t} \le 0$, максимум $r_i(0,t)$ достигается на левой границе отрезка при $t = t_1$, а выражение $\min(\mu_i^2, 1+t^2)$ принимает максимальное значение на правой границе $t = t_U$. С учетом этого

$$\sup_{t\in[t_1,t_U]} (r(\mu,t) - r(\mu,t_{min})) \le \frac{1}{n}\sum_{i=1}^{n} (\min(\mu_i^2, 1+t_U^2) + r_i(0,t_1)) - r(\mu,t_{min}). \quad (10)$$

Можно показать, что для любого $t > 0$ выполняется неравенство

$$r_i(0,t) \le \sqrt{\frac{2}{\pi}}\,\frac{1}{t}\,e^{-\frac{t^2}{2}}. \tag{11}$$

Подставляя в (10) значения t_1 и t_U, а также используя оценку (11), ограничение на η_n и тот факт, что неравенство (4) справедливо не только для t_U, но и для t_{min}, получаем, что правая часть (10) стремится к нулю при $n \to \infty$. Таким образом, случайная величина $|r(\mu, t_F) - r(\mu, t_{min})|$ сходится почти наверно к 0, и, следовательно, $\mathsf{P}(|r(\mu, t_F) - r(\mu, t_{min})| > \varepsilon) \xrightarrow{n} 0$. Теорема доказана. □

Заключение

Рассмотрен метод удаления шума в разреженном сигнале, основанный на процедуре множественной проверке нулевых гипотез. В модели с белым гауссовским шумом доказана состоятельность оценки среднеквадратичного риска при выборе FDR-порога.

Список литературы

[1] Donoho D., Johnstone I.M. Ideal spatial adaptation via wavelet shrinkage // Biometrika. 1994. Vol. 81, № 3. Pp. 425–455.

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

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

[4] Jansen M. Noise Reduction by Wavelet Thresholding. New York: Springer, 2001. Lecture notes in Statistics. Vol. 161. doi:10.1007/978-1-4613-0145-5

[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] Шестаков О.В. Центральная предельная теорема для функции обобщенной кросс-валидации при пороговой обработке вейвлет-коэффициентов // Информатика и ее применения. 2013. Т. 7, № 2. С. 40–49.

[8] Shestakov O.V. On the strong consistency of the adaptive risk estimator for wavelet thresholding // Journal of Mathematical Sciences. 2016. Vol. 214, № 1. Pp. 115–118.

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

[10] Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing // Journal of the Royal Statistical Society: Series B (Statistical Methodology). 1995. Vol. 57. Pp. 289–300.

[11] Abramovich F., Benjamini Y., Donoho D. Johnstone I. Adapting to unknown sparsity by controlling the false discovery rate // Annals of Statistics. 2006. Vol. 34, № 2. Pp. 584–653.

[12] Hoeffding W. Probability inequalities for sums of bounded random variables // Journal of the American Statistical Association. 1963. Vol. 58, № 301. Pp. 13–30.

Библиографическая ссылка

Заспа А.Ю., Шестаков О.В. Состоятельность оценки риска при множественной проверке гипотез с FDR-порогом // Вестник ТвГУ. Серия: Прикладная математика. 2017. № 1. С. 5–16.

Сведения об авторах

1. **Заспа Андрей Юрьевич**
студент факультета вычислительной математики и кибернетики МГУ имени М.В. Ломоносова.

Россия, 119992, г. Москва, ГСП-1, Воробьевы горы, МГУ им. М.В. Ломоносова. E-mail: zacpa@ya.ru.

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. E-mail: oshestakov@cs.msu.su.

CONSISTENCY OF THE RISK ESTIMATE OF THE MULTIPLE HYPOTHESIS TESTING WITH THE FDR THRESHOLD

Zaspa Andrey Yurievich
student of Computational Mathematics and Cybernetics faculty,
Lomonosov Moscow State University.
Russia, 119992, Moscow, GSP-1, Vorobievy gory, Lomonosov MSU.
E-mail: zacpa@ya.ru

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.
E-mail: oshestakov@cs.msu.su

Received 02.02.2017, revised 14.03.2017.

We consider the method of denoising the sparse signals based on the multiple hypothesis testing with the use of the FDR threshold. In the model with a white Gaussian noise we prove the consistency of the unbiased mean-square risk estimator.

Keywords: multiple hypothesis testing, thresholding, risk estimate, consistency.

Bibliographic citation

Zaspa A.Yu., Shestakov O.V. Consistency of the risk estimate of the multiple hypothesis testing with the FDR threshold. *Vestnik TvGU. Seriya: Prikladnaya Matematika* [Herald of Tver State University. Series: Applied Mathematics], 2017, no. 1, pp. 5–16. (in Russian)

References

[1] Donoho D., Johnstone I.M. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 1994, vol. 81(3), pp. 425–455.

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

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

[4] Jansen M. *Noise Reduction by Wavelet Thresholding*. Springer, New York, 2001. Lecture notes in Statistics. Vol. 161. doi:10.1007/978-1-4613-0145-5

[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] Shestakov O.V. Central limit theorem for generalized cross-validation function in wavelet thresholding method. *Informatika i ee primeneniya* [Informatics and its Applications], 2013, vol. 7(2), pp. 40–49. (in Russian)

[8] Shestakov O.V. On the strong consistency of the adaptive risk estimator for wavelet thresholding. *Journal of Mathematical Sciences*, 2016, vol. 214(1), pp. 115–118.

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

[10] Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 1995, vol. 57, pp. 289–300.

[11] Abramovich F., Benjamini Y., Donoho D. Johnstone I. Adapting to unknown sparsity by controlling the false discovery rate. *Annals of Statistics*, 2006, vol. 34(2), pp. 584–653.

[12] Hoeffding W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 1963, vol. 58(301), pp. 13–30.