УДК 535.214.4

БЫСТРОЕ ПОМЕХОУСТОЙЧИВОЕ ОЦЕНИВАНИЕ УРОВНЯ ШУМ-ДОРОЖКИ РАДИОМЕТРА С ПОМОЩЬЮ АБСОЛЮТНОГО МЕДИАННОГО ОТКЛОНЕНИЯ

Б. Л. Ерухимов

Рассматривается метод помехоустойчивого оценивания уровня σ радиометра в реальном времени. Использован известный алгоритм абсолютного медианного отклонения. Приводится сравнение предлагаемого метода с классическим на записях шума, загрязненных типичными для радиоастрономических наблюдений помехами, а также эффективность используемой оценки.

A real time method of antijamming estimation of the σ level of the radiometer is regarded. A known algorithm of the absolute median deviation is used. A comparison of the proposed method with a classical one is given on noise records contaminated by jamming typical for radio-astronomical observations. An efficiency of the used estimation is shown.

Контроль уровня дорожки радиометра является традиционной процедурой в методике радиоастрономических наблюдений. При этом оценивается шумовая температура системы, характеризующая потенциальную точность результатов радиоастрономических исследований. Оперативный контроль уровня шума особенно важен при использовании радиометров предельной чувствительности, достигаемой, как правило, в обмен на некоторое ухудшение стабильности параметров.

На практике выходной сигнал радиометра представляет собой смесь нормального шумового процесса, уровень σ которого принято называть шум-дорожкой, и аномальных помех, наиболее характерными среди которых являются короткие мощные импульсы и медленные уходы уровня (тренды). Ясно, что если по «загрязненной» таким образом выборке оценивать σ с помощью стандартной, скорректированной для отсутствия смещения оценки максимального правдоподобия вида

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j\right)\right]^2}, \qquad (1)$$

то результат может быть весьма далеким от истинного значения. Это обстоятельство связано с отсутствием у оценки (1) свойства робастности, т. е. устойчивости к аномальным значениям выборки.

Существует достаточное число способов реализации хорошей оценки σ нормального шума по загрязненной выборке. Наиболее просто снимается проблема трендов, для чего используются различные дифференциальные методы, например алгоритм трехточечной или американской дисперсии, а также подсчет дисперсии по отдельным кускам выборки, где величиной тренда можно пренебречь, с последующим усреднением результатов. Несколько сложнее обстоит дело с импульсными помехами. Для получения хорошей оценки σ или, как ее принято называть в статистике, оценки масштаба задачу разбивают на два этапа. На первом этапе реализуется быстрый робастный алгоритм, дающий оценку масштаба с не очень высокой эффективностью (под эффективностью оценки мы здесь и далее будем понимать принятую в статистике величину, равную отношению дисперсии оптимальной и исследуемой оценок). Полученная таким образом оценка применяется в качестве нулевого приближения в некоторой итеративной процедуре, используемой на втором этапе для получения результирующей оценки с высокой эффективностью.

Однако на практике задача оценивания масштаба в чистом виде возникает достаточно редко [1]. Обычно параметр играет роль мешающего (уровень шум-дорожки) при оценке других параметров, что является типичным для задач радиоастрономии. Здесь существенно более важным оказывается получение устойчивой несмещенной оценки масштаба, нежели достижение высокой эффективности. В этом случае П. Хьюбер [1] рекомендует ограничиваться типовым нулевым приближением оценки масштаба, представляемой величиной абсолютного медианного отклонения (АМО) $\hat{\sigma}m$. Эта величина определяется как медиана * абсолютных отклонений от медианы:

$$\hat{\sigma}m = \operatorname{med}\{|x_i - M_n|\}, \tag{2}$$

где M_n — медиана выборки x_i размером n.

Очевидно, что эта оценка обладает хорошей робастностью, простой анализ показывает, что она «выдерживает» загрязнение помехами 50 % отсчетов выборки. Ее можно легко реализовать в реальном времени на типовых ЭВМ сбора данных с помощью очевидного алгоритма, состоящего в последовательном вычислении $\hat{\sigma}m$ по нескольким временным интервалам с дальнейшим усреднением результата. Как было указано выше, при таком подходе снимается проблема трендов. Кроме того, при разумном выборе размера интервала легко реализуется быстрая оценка $\hat{\sigma}m$, сводящаяся к двукратной сортировке n чисел.

Несмотря на сделанное выше замечание о малой важности получения высокоэффективной оценки масштаба в случае, если она является вспомогательной, вопрос об ее эффективности обязательно возникает. При использовании полученной оценки уровня шум-дорожки в дальнейшей обработке данных или в качественных выводах наблюдателю необходимо знать ее точность, т. е. дисперсию. Известно, что для нормального шума оценка (1) является оптимальной по минимуму дисперсии (эффективной). Естественно определить эффективность предлагаемой оценки, с тем чтобы наблюдатель мог знать, на сколько процентов он должен увеличить суммарное число отсчетов (в обмен на свойство робастности) для получения оценки с точностью, равной точности оптимальной оценки (1). В этих целях полезно аналитически определить дисперсии оценок $\hat{\sigma}$ и $\hat{\sigma}m$, что и приводится ниже.**

Определим асимптотическую дисперсию D оценки σ, получаемую классическим методом максимального правдоподобия (1). Для этого рассмотрим случайную величину η, определяемую выражением

$$\eta = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\,, \tag{3}$$

где x_i — одинаково распределенные нормальные случайные величины с нулевым средним и дисперсией σ^2. В соответствии с [2] плотность распределения этой величины равна

$$W(\eta) = \begin{cases} 0 & \eta < 0 \\ \dfrac{2\eta^{n-1} l^{-\frac{\eta^2}{2\sigma^2}}}{(2\sigma^2)^{\frac{n}{2}}\,\Gamma(n/2)} & \eta \geqslant 0 \end{cases}$$

---

* Медиана — средний член вариационного ряда, т. е. выборки экспериментальных данных, упорядоченных по возрастанию величины.

** Автор ни в коей мере не претендует на новизну полученных выводов с точки зрения математической статистики. Первоначальной целью этой работы была иллюстрация с успехом используемой нами оценки $\hat{\sigma}m$ с указанием ссылки на вывод выражения ее эффективности. Однако подобного вывода не удалось обнаружить в доступных автору работах, что, вероятно, не исключает наличия такого. В этой связи автор считает целесообразным привести здесь собственный вывод эффективности оценки $\hat{\sigma}m$, хотя бы потому, что это явилось бы необходимым условием для использования предложенной оценки самим автором, если бы он был в роли читателя.

При $n=2K+1$ моменты этого распределения равны:

$$m_\eta=\frac{\sigma\sqrt{2}\left(\frac{n-1}{2}\right)!}{\Gamma\left(\frac{n}{2}\right)}; \tag{4}$$

$$D_\eta=\sigma^2\left[n-\frac{2\Gamma^2\left(\frac{n+1}{2}\right)}{\Gamma(n/2)}\right]. \tag{5}$$

Очевидно, что выражения (1) и (3) эквивалентны с точностью до некоторого масштабного множителя C. Таким образом, искомая дисперсия равна

$$D=C^2D_\eta. \tag{6}$$

Для нахождения величины C приравняем математические ожидания величин $\hat\sigma$ и η. По условию математическое ожидание оценки $\hat\sigma$ равно σ. Отсюда

$$C=\frac{\Gamma(n/2)}{\sqrt{2}\left(\frac{n-1}{2}\right)!}. \tag{7}$$

Полагая n достаточно большим ($n>10$) и воспользовавшись формулами удвоения гамма-функции, а также приближением Стирлинга для факториала, получим:

$$C=\frac{1}{\sqrt{n-1}};$$

$$D=\frac{\sigma^2}{n-1}. \tag{8}$$

Теперь определим приближенно асимптотическую дисперсию K оценки $\hat\sigma m$. По условию для получения оценки $\hat\sigma m$ надо сначала оценить медиану исходной нормальной выборки x_i, далее построить вариационный ряд, состоящий из модулей отклонений величины x от этой оценки. Медиана этого ряда, взятая с соответствующим коэффициентом, и берется в качестве несмещенной и состоятельной оценки величины σ.

В соответствии с [3] искомая дисперсия K представляется выражением

$$K=\frac{1}{4nf^2(tm)}, \tag{9}$$

где n — число элементов выборки; tm — значение медианы вышеупомянутого ряда абсолютных отклонений; $f(tm)$ — значение плотности распределения вероятности абсолютных отклонений в точке медианы.

При нормальном распределении исходной выборки с дисперсией σ^2 функция $f(t)$ есть плотность одностороннего нормального распределения вида

$$f(t)=\begin{cases}\frac{1}{\Delta}\sqrt{\frac{2}{\pi}}\,l^{-\frac{t^2}{2\Delta^2}}; & t\geqslant 0\\ 0, & t<0\end{cases} \tag{10}$$

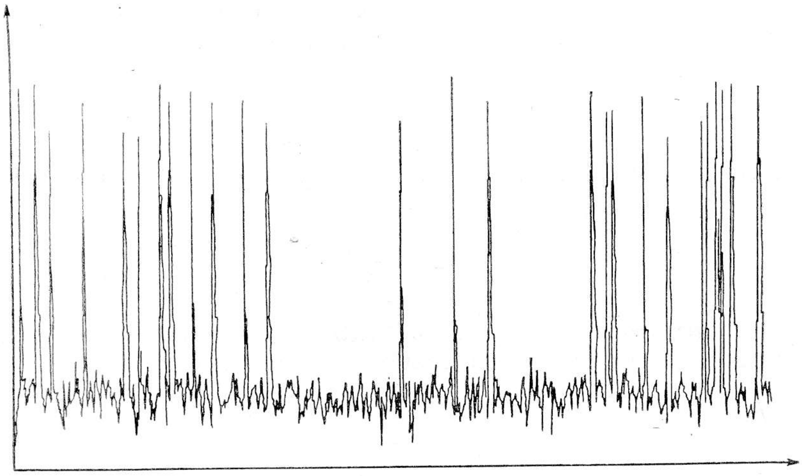
Величина Δ в выражении (10) несколько больше величины σ исходной выборки x_i, так как при формировании отклонений из каждого x_i вычитается оценка медианы x_i, имеющая ненулевую дисперсию. В соответствии с [3] для нормальной выборки размером n дисперсия оценки медианы равна $\sigma^2\pi/(2n)$. Отсюда имеем для Δ^2:

$$\Delta^2=\sigma^2[1+\pi/(2n)]. \tag{11}$$

Обозначив $z=t/\Delta$, запишем нормированную интегральную функцию распределения, соответствующую $f(t)$:

$$F(z) = \begin{cases} \sqrt{\frac{2}{\pi}} \int\limits_0^z l^{-\frac{z^2}{2}}; & z \geqslant 0 \\ 0. & z < 0 \end{cases}$$

При $F(z)=1/2$, что соответствует медиане, по таблицам находим $z=0.674$. Следовательно, $tm=\Delta \cdot 0.674$. Таким образом, чтобы оценка $\hat{\sigma}m$ была несмещенной, необходимо результат нашей процедуры (медиана ряда абсолютных откло-



Нормальный шум с параметром $\sigma=1$, загрязненный искусственно введенными импульсными помехами, амплитудой 20 и случайным временем появления, соответствующим распределению Пуассона с параметром $\lambda=0.07$.

Классическая оценка σ составляет 5.13, оценка АМО — 1.02.

нений) разделить на $0.674\ [1+\pi/(2n)]^{1/2}$. Для получения дисперсии этой оценки Dm необходимо значение K, полученной из (9), разделить на квадрат указанной величины. В результате получаем

$$Dm \approx 1.37\sigma^2/n, \qquad (12)$$

и асимптотическая эффективность АМО

$$E = D/Dm = 0.73.$$

Таким образом, оценка σ с помощью АМО требует примерно в 1.37 раз больше отсчетов, чем оптимальная оценка (1), для получения одинаковой точности. Рисунок иллюстрирует робастные свойства оценки АМО. На рисунке представлен стационарный гауссов шум ($\sigma=1$) с искусственно введенными импульсными помехами, амплитудой 20 и случайным временем появления, соответствующим распределению Пуассона с параметром $\lambda=0.07$. Классическая оценка $\hat{\sigma}$ загрязненного таким образом шума в соответствии с выражением (1) составляет —5.13, при этом оценка АМО составляет —1.02.

Литература

1. Хьюбер П. Робастность в статистике / Пер. с англ.; Под ред. И. Г. Журбенко. М Мир, 1984. 303 с.
2. Заездный А. М. Основы расчетов по статистической радиотехнике. М.: Связь, 1969. 447 с.
3. Юстуссон Б. И. Медианная фильтрация: статистические свойства // Быстрые алгоритмы в цифровой обработке изображений / Под ред. Т. С. Хуанга; Пер. с англ.; Под ред. Л. П. Ярославского. М.: Радио и связь, 1984. 221 с.

Поступила в редакцию
21 апреля 1989 г.