УДК 519.254

НЕЛИНЕЙНАЯ ФИЛЬТРАЦИЯ ВРЕМЕННЫХ РЯДОВ ГЕО- И ГЕЛИОФИЗИЧЕСКИХ ПАРАМЕТРОВ

И.А. Белашова, В.В. Бочкарев

NONLINEAR FILTERING OF TIME SERIES OF GEO- AND HELIOPHYSICAL PARAMETERS

I.A. Belashova, V.V. Bochkarev

В данной работе рассматриваются методы нелинейной фильтрации гео- и гелиофизических рядов. В качестве альтернативы критерию квадратичной ошибки, применяемому в стандартных методах фильтрации, к которым относится и вейвлет-трешолдинг, используется критерий максимального правдоподобия, как более общий метод оценки, не зависящий от закона распределения. Вид функции правдоподобия может сложным образом зависеть от параметров, поэтому применялись различные методы оптимизации (генетические алгоритмы, метод имитации отжига, методы детерминированного поиска). В качестве примеров рассматриваются фильтрация рядов среднемесячного числа солнечных пятен, критических частот ионосферных слоев, допплеровского сдвига частоты сигнала наклонного зондирования ионосферы.

In this paper considered methods of nonlinear filtering and geo heliophysical series. Alternatively squared error criterion, employed in standard filtering techniques, which include wavelet tresholding, the criterion of maximum likelihood, as a more general method of evaluation does not depend on the law of distribution. View of the likelihood function may depend in a complicated way on the parameters, so used different methods of optimization (genetic algorithms, simulated annealing method, deterministic search methods). As examples, filtering rows monthly sunspot numbers, critical frequency of the ionospheric layers Doppler shift signal oblique ionospheric sounding.

Вейвлет-фильтрация рядов

В настоящее время все большее распространение получил такой метод анализа и обработки сигналов, как вейвлет-анализ. Он широко применяется в различных областях науки, техники, медицины. Такую популярность вейвлет-анализ получил благодаря своим свойствам: частотно-временная локализация, возможность исследования фрактальных функций, высокая эффективность. Преимущество вейвлет-анализа перед спектральным анализом Фурье колоссально, поскольку последний, как известно, предполагает рассмотрение бесконечно длинного по времени сигнала, который, в практических целях, конечно, не рассматривается.

Развитие вычислительной техники привело к широкому распространению дискретного вейвлет-анализа. Одним из его интереснейших приложений является фильтрация данных, основной принцип которой заключен в следующем: при фильтрации оставляются коэффициенты, содержащие наибольший объем информации о фильтруемом сигнале. В большинстве стандартных методов это приводит к тому, что оставляются самые большие коэффициенты. Алгоритмы и методы вейвлет-фильтрации известны и хорошо изучены и имеют широкое применение в акустике, сейсмологии, медицине и других областях.

Использование критерия квадратичной ошибки как функции качества дает очень хорошие результаты, когда речь идет о рядах с гауссовским законом распределения вероятности. Однако во многих случаях приходится сталкиваться с рядами, флуктуации которых описываются законом, отличным от гауссовского. В частности, во всех случаях, когда фиксируется количество случаев редких событий (например, ряд числа солнечных пятен), мы будем иметь дело с законом Пуассона. Применение стандартных методов фильтрации, предполагающих, что мы оставляем наибольшие по величине коэффициенты разложения, может в этих случаях быть не корректным.

Методы нелинейной фильтрации

Для решения задачи фильтрации пуассоновских рядов можно использовать другой критерий выбора ненулевых коэффициентов и их значений, а именно критерий максимального правдоподобия:

$$\hat{\theta}_{мп} = \arg\max_{\theta\in\Theta} P(\bar{y}/\theta), \qquad (1)$$

где θ – параметр.

Так же часто рассматривается логарифмическая функция правдоподобия $\ln P(\frac{\bar{y}}{\theta})$. Для ряда $y_1, y_2 \ldots y_N$, подчиняющегося пуассоновскому распределению, условная вероятность дается выражением

$$P = \prod_k \frac{\lambda_k^{y_k e^{-\lambda_k}}}{y_k},$$

где λ_k – математическое ожидание процесса в момент времени k. Мы хотим представить оценки для ряда λ_k в форме $\hat{y}_k = \sum_{i=1}^{M} c_i \psi_{ik}$, $k = 1:N$, $M \ll N$.

Будем в качестве функции качества рассматривать функцию правдоподобия с обратным знаком:

$$Err = -\sum_{k=1}^{N} (y_k \ln \sum_{i=1}^{M} c_i \psi_{ik} - \sum_{i=1}^{M} c_i \psi_{ik}). \qquad (2)$$

Следует обратить внимание, что мы пытаемся минимизировать функцию качества при очень малом числе ненулевых коэффициентов.

Функция качества может сложным образом зависеть от параметра, поэтому для нахождения ее минимума предлагается применение различных методов оптимизации: генетических алгоритмов, метода имитации отжига, методов детерминированного поиска.

Применение генетических алгоритмов при фильтрации пуассоновских рядов

Прежде всего, хочется отметить, что применение аппарата генетических алгоритмов является весьма

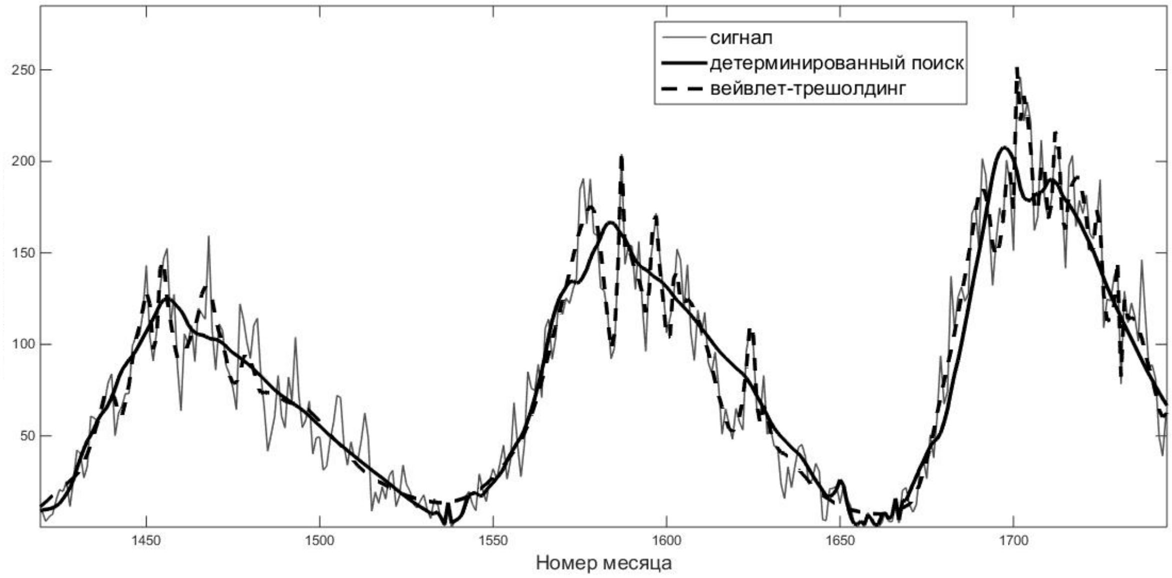

Рис. 1. Фильтрация среднемесячного числа солнечных пятен.

интересной и творческой работой. В подтверждение этому можно привести большое число параметров, изменение которых приводит к изменению величины оптимизируемой функции, такие как: численность популяции, число родителей в начальной популяции, число поколений, вероятность кроссинговера и мутации, число «элитных» потомков в новой популяции и так далее.

Метод фильтрация при помощи генетических алгоритмов можно разделить на два ключевых блока:

- Определение нового набора ненулевых коэффициентов;
- Варьирование значений ненулевых коэффициентов с целью минимизации функции качества – функции максимального правдоподобия.

Для формирования нового набора ненулевых коэффициентов составляется хромосома из коэффициентов вейвлет-разложения, подвергшихся вейвлет-фильтрации. Она представляет из себя строку из нулей и единиц, где «1» соответствует оставленному при фильтрации коэффициенту, а «0»-занулившемуся. При помощи генетических алгоритмов происходит варьирование положения «1» в хромосоме, и находится такой набор, при котором величина функции правдоподобия минимальна.

Важнейшим условием, которое следует проверять на протяжении всей работы минимизации, является постоянство числа ненулевых коэффициентов, определенного при вейвлет-фильтрации. Это условие необходимо учитывать при разработке функции кроссинговера и мутации, а так же при создании начальной популяции.

Можно рассматривать различные разновидности кроссинговера (одноточечный, многоточечный, однородный, перетасовочный), главное, следует накладывать условие постоянства числа ненулевых коэффициентов. В данной работе применялся однородный кроссинговер.

В результате мутации задается произвольное число генов, которые должны измениться, при этом так же контролируется число ненулевых коэффициентов: если в хромосоме «0» изменился на «1», то в каком-то другом локусе происходит обратная операция.

На этапе варьирования значений ненулевых коэффициентов применяется, так называемый, непрерывный генетический алгоритм. Здесь хромосомой выступает уже не битовая строка, а вектор ненулевых коэффициентов. Применяя смешанный кроссинговер и мутацию, так же необходимо следить за постоянством числа ненулевых коэффициентов.

Применение других методов оптимизации при фильтрации пуассоновских рядов

Структура фильтрации методом отжига так же состоит из двух частей:

- Определение наилучшего набора ненулевых коэффициентов;
- Варьирование значений ненулевых коэффициентов.

Идея выбора нового набора ненулевых коэффициентов аналогична генетическим алгоритмам, различие заключается лишь в способе поиска новых коэффициентов. Однако метод имитации отжига для случая фильтрации ряда числа солнечных пятен дает не очень хороший результат – величина функции правдоподобия больше, чем при вейвлет-фильтрации.

Для использования наиболее эффективных методов детерминированного поиска – квазиньютоновских алгоритмов оптимизируемая функция должна быть непрерывной. Это ограничение не позволяет нам решать задачу определения наилучшего набора ненулевых коэффициентов, поскольку позиции ненулевых коэффициентов – целые числа. Таким образом, метод оптимизации будет проходить в один этап, а именно, будут варьироваться ненулевые значения коэффициентов, определенных вейвлет-трешолдингом.

Мы рассмотрели применение одного из наиболее распространенных алгоритмов метода доверительных областей – алгоритм «собачья нога», который состоит в решении системы линейных уравнений для нахождения направления поиска.

Эффективность метода «собачья нога» заключается в том, что он требует только одно линейное вычисление на итерации, когда вычисляется шаг Ньютона–Гаусса. Кроме того, данный метод является более надежным по сравнению с методом линейного поиска Ньютона–Гаусса.

Пример фильтрации ряда среднемесячного числа солнечных пятен

На рисунке изображены результаты стандартной вейвлет-фильтрации и фильтрации по критерию максимального правдоподобия (использовался квазиньютоновский алгоритм) ряда среднемесячного числа солнечных пятен. Из рисунка видно, что при стандартной вейвлет-фильтрации наиболее детально выделены флуктуации на вершинах, в то время как при фильтрации по критерию наибольшего правдоподобия более детально оказались описано поведение ряда в минимумах солнечной активности.

Для распределения Пуассона математическое ожидание равно λ, а среднеквадратическое отклонение $\sigma = \sqrt{\lambda}$. Отсюда следует, что при одинаковой величине флуктуаций более значимы флуктуации в областях с малым λ. Таким образом, можно сделать вывод, что критерий максимального правдоподобия дал более адекватные результаты, чем стандартный метод вейвлет-фильтрации.

СПИСОК ЛИТЕРАТУРЫ

Малла С. Вейвлеты в обработке сигналов. М.: Мир, 2005. 671 с.

Нурминский Е. Методы оптимизации. Курс лекций ДВГУ–ДВФУ.

Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы / Пер. с польск. И.Д. Рудинского. М.: Горячая линия – телеком, 2006. 452 с.

Казанский (Приволжский) федеральный университет, Казань, Россия