

С.И. Еременко

СПЕКТРАЛЬНАЯ ФИЛЬТРАЦИЯ ПОЛОЖИТЕЛЬНО ОПРЕДЕЛЕННЫХ ФЛУКТУИРУЮЩИХ СИГНАЛОВ

Предложен метод «положительной спектральной фильтрации» (ПСФ) – математический подход к анализу неотрицательных случайных сигналов, позволяющий разделить составляющие с разными характерными временами флуктуаций. Метод эффективен для сильно флуктуирующих сигналов, в том числе для флуктуаций концентрации примесей в атмосфере. Приведен пример разделения вкладов близких и удаленных источников загрязнения атмосферы города.

1. Введение

Анализ газовых и аэрозольных примесей в атмосфере представляет собой частный случай довольно широкого класса задач, где исследуемый сигнал по своей природе представляет собой неотрицательную случайную функцию времени (часто многомерную) и образован наложением нескольких неотрицательных компонент, обусловленных разными источниками. Одна из задач анализа состоит в разделении этих компонент.

В случае многомерных сигналов для этой цели используют статистические методы, основанные на анализе ковариации многомерных сигналов, такие как факторный анализ или метод главных компонент. Однако эти методы не вполне адекватны задаче, так как не учитывают положительной определенности сигнала, и поэтому корректны лишь в случаях, когда флуктуации малы по сравнению со средним значением, в то время как для многих задач этого класса, и в частности для атмосферы, типично противоположное.

В последние годы предпринимаются попытки построить математические методы, адекватные таким задачам. Стоит упомянуть работы П. Паатеро [1, 2], посвященные факторизации положительно определенных многомерных сигналов. Предлагаемый им подход позволяет корректно обойтись с суперпозицией неотрицательных компонент и основан на том, что вместо ковариационной матрицы анализу подвергается вся выборка временного ряда многомерного сигнала и компоненты выделяются в результате решения экстремальной задачи с ограничениями. Однако и в этих работах никак не используется информация, заключенная во временном ходе сигнала, т. е. сигнал воспринимается как неупорядоченная выборка случайной величины. Ясно, что сигнал, измеренный как функция времени, в данной точке содержит больше информации, чем просто неупорядоченная выборка. Например, сигналы, обусловленные разными источниками, могут иметь различные спектры флуктуаций.

В нашем частном случае исследуется концентрация какой-либо примеси в одной точке. Она представляет собой реализацию случайной функции времени, причем характер турбулентного перемешивания в атмосфере таков, что масштаб пульсаций концентрации обычно сравним с ее средним значением, а во многих случаях в несколько раз больше.

Теория турбулентного переноса в настоящее время не дает однозначного ответа на вопрос о том, как флуктуации концентрации зависят от расстояния до источника или от времени переноса до приемника, однако совокупность теоретических и экспериментальных данных [3–5] позволяет сформулировать некоторые качественные закономерности.

1) Концентрация примеси от локальных источников обычно имеет характер пульсаций, спадающих временами до нуля (эффект перемежаемости).

2) Относительный размах флуктуаций концентрации не спадает или спадает очень медленно от времени (расстояния) переноса, хотя среднее значение быстро уменьшается.

3) Время автокорреляции флуктуаций растет со временем (расстоянием) переноса, т.е. отклик «размывается» по времени и его спектр смещается в сторону низких частот.

Можно надеяться, что спектральный анализ позволит выделить из сигнала компоненты, имеющие разную природу, например разделить глобальный фон и вклад местных источников. Однако такие стандартные методы, как, например, разложение Фурье, в нашем случае малоэффективны, так как не учитывают знакопостоянство сигнала. Ниже будет приведен пример, поясняющий это утверждение.

В настоящей работе предлагается подход, позволяющий разделять положительно определенные случайные сигналы на компоненты с разным характерным временем флуктуаций, т.е. метод положительной спектральной фильтрации (ПСФ). Метод основан на использовании априорной информации о сигналах и учитывает в явном виде их фундаментальное свойство знакопостоянства. Предлагаемый подход эффективен в случае, когда масштаб флуктуаций сравним со средним значением сигнала.

2. Неотрицательное спектральное разложение в общем виде

В общем виде можно представить задачу следующим образом. Пусть имеется измеренный сигнал $y(t) \geq 0$ и задано из каких-то априорных соображений нормированное неотрицательное ядро $g(\tau, \omega)$ линейного оператора свертки

$$\mathbf{G}_\omega \psi(t) = \int_{-\infty}^{\infty} \psi(t - \tau) g(\tau, \omega) d\tau. \quad (1)$$

Предполагается, что сигнал $y(t)$ можно представить как результат интегрального преобразования свертки функции $\psi(t, \omega)$, которую будем условно называть «спектральной»:

$$y(t) = \int_0^{\infty} \mathbf{G}_\omega \psi(t, \omega) d\omega. \quad (2)$$

По аналогии с разложением Фурье параметр ω назовем частотой. Очевидно, что обратная задача (нахождение ψ из $y(t)$) не может иметь единственного решения.

3. Ядро спектрального разложения

В качестве ядра для разложения сигнала обычно применяется набор функций, главные свойства которых – полнота и ортогональность (или хотя бы линейная независимость). Первое необходимо для того, чтобы можно было разложить любой сигнал, второе – чтобы такое разложение было однозначным. В данной работе предлагается использовать набор положительно определенных функций, которые выбираются исходя из априорного знания о природе сигнала. Такой набор полон для исследуемого класса сигналов в той мере, в какой верны априорные посылки, исходя из которых он выбран. Что же касается линейной независимости, то в ней нет необходимости, если однозначность достигается другими средствами. Множество возможных разложений сигнала можно упорядочить, используя какие-то критерии предпочтения, и тогда выбор единственного разложения осуществляется решением экстремальной задачи на этом множестве.

Итак, вид функций ядра $g(t, \omega)$ следует строить, основываясь на модели изучаемого процесса. Если, например, этот процесс – изотропная диффузия, то в качестве $g(t, \omega)$ естественно взять набор функций Гаусса, а в качестве параметра ω – их обратную ширину. В любом случае будем полагать, что функции ядра обладают следующими свойствами:

$$g(t, \omega) \geq 0; \quad g(t, \omega) \rightarrow 0 \text{ при } t \rightarrow \pm\infty; \quad |g(t, \omega)| \equiv \int_{-\infty}^{\infty} g(t, \omega) dt = 1. \quad (3)$$

Символ $||$ здесь и далее будет обозначать интегральную меру функции. Далее, для всяких ω_1 и ω_2 , если $\omega_1 > \omega_2$, то существует $\varphi(t) \geq 0$ такое, что

$$g(t, \omega_2) = \mathbf{G}_{\omega_1} \varphi(t), \quad (4)$$

т.е. более широкая функция может быть представлена как свертка более узких. Последнее свойство присуще функциям переноса и, в частности, упомянутой функции Гаусса.

4. Формулировка задачи

Пусть сигнал может быть представлен сверткой с известным ядром G и неизвестной «спектральной» функцией $\psi(t, \omega) \geq 0$. Мы хотели бы использовать такое разложение для того, чтобы разделить, если это возможно, исходный сигнал на составляющие, соответствующие разным значениям ω . Простейшая задача такого рода – выделение нижних частот, т. е. выделение из $y(t)$ «низкочастотной» составляющей, которая целиком может быть представлена спектром с частотами не выше заданного значения Ω :

$$y(t) = \int_0^{\Omega} G_{\omega} \psi(t, \omega) d\omega + f'(t), \quad f'(t) \geq 0, \quad (5)$$

если это возможно при $|\psi(t, \omega)| > 0$, то будем говорить, что $y(t)$ содержит частоты ниже и равные Ω , и если при этом остаточная функция $f'(t)$ таких частот не содержит, то скажем, что выражение вида (5) решает задачу фильтрации нижних частот сигнала $y(t)$.

Сразу ясно, что задача (5) не может иметь однозначного решения, хотя бы в силу свойства (4), однако это же свойство позволяет записать ее иначе. Действительно, в силу свойства (4) для любого $\psi(t, \omega)$ и Ω можно найти $\varphi(t) \geq 0$ такое, что

$$\int_0^{\Omega} G_{\omega} \psi(t, \omega) d\omega = G_{\Omega} \varphi(t), \quad (6)$$

т. е. свести все компоненты с частотами меньше Ω к частоте Ω . Это позволяет записать задачу фильтрации нижних частот следующим образом:

$$y(t) = G_{\Omega} \varphi(t) + f'(t) \equiv f(t, \Omega) + f'(t); \quad (7)$$

$$f(t, \Omega) \geq 0, \quad f'(t) \geq 0.$$

Назовем $f(t, \Omega)$ интегральной низкочастотной (НЧ) ПСФ-составляющей сигнала для частоты разделения Ω . Идеальным решением задачи было бы построение алгоритма, позволяющего на основании каких-то разумных критериев однозначно получить для $y(t)$ монотонно возрастающую по ω и дифференцируемую $f(t, \omega)$. Это позволило бы представить наш сигнал как

$$y(t) = \int_0^{\infty} \frac{\partial f(t, \omega)}{\partial \omega} d\omega \quad (8)$$

и получить составляющую для любого интервала частот (ω_1, ω_2) как интеграл в этих пределах. Как это сделать, пока неясно. Но, во всяком случае, можно сформулировать экстремальные критерии, позволяющие корректно решить задачу (7).

5. Верхняя оценка низкочастотной компоненты и ее спектр

Назовем *верхней по среднему значению* оценкой соответствующей НЧ-компоненты функцию $f^*(t, \omega)$, минимизирующую $|f^*|$ в выражении (7) при заданном ω . Очевидно, минимум $|f^*|$ соответствует максимуму $|\varphi|$ и $|f|$. Введем сокращенные обозначения $f_{\omega}^* = f^*(t, \omega) = G_{\omega} \varphi_{\omega}$, $f'_{\omega} = y(t) - f_{\omega}^*$ и сформулируем основные свойства такой оценки.

а) Из (3) следует, что $|\varphi_{\omega}(t)| \leq |f_{\omega}(t)|$.

б) Из (4) и (6) можно показать, что если $a \geq b$, то $|\varphi_a| \geq |\varphi_b|$ и соответственно $|f_b^*| \geq |f_a^*|$. К сожалению, нельзя утверждать, что при тех же условиях $\varphi_a(t) \geq \varphi_b(t)$ для любого t , т. е. оценка, верхняя по среднему значению, строго говоря, не является верхней локально. Тем не менее при разумных требованиях к гладкости $y(t)$ можно показать, что

в) $f'_{\omega}(t) \rightarrow 0$ и $|f'_{\omega}| \rightarrow 0$ при $\omega \rightarrow 0$.

Задача отыскания максимальной по модулю f_{ω}^* для любого ω представляет собой стандартную задачу линейного программирования и обычно имеет единственное решение. Таким образом, задача фильтрации нижних частот может быть решена корректно.

Указанная оценка еще не позволяет корректно разделить сигнал на низко- и высокочастотную составляющую, но позволяет оценить сверху низкочастотную составляющую и получить оценку спектра частот сигнала.

Зададим отрезок $[\omega_1, \omega_2]$, на котором нас интересует ПСФ-спектр сигнала $y(t)$, и найдем $f_\omega(t)$ для каждого значения ω на этом отрезке. Полученную зависимость $S(\omega) = |f_\omega(t)|$ назовем верхней по среднему значению оценкой интегрального спектра функции $y(t)$. Из вышеприведенных свойств следует, что $S(\omega)$ – монотонно возрастающая функция, причем $S(\omega) \rightarrow |y(t)|$ при $\omega \rightarrow \infty$. Полученный таким образом спектр можно использовать для качественного анализа сигнала. Покажем это на предельно упрощенном примере.

6. Демонстрационный пример

Пусть сигнал представлен суммой двух случайных сигналов от разных источников, находящихся на разном расстоянии от точки измерения. Отклик мгновенного выброса источника в точке измерения определяется диффузионным размытием, и характерное время отклика зависит от расстояния до источника. Флуктуирующий сигнал от каждого из источников можно представить как суперпозицию случайного набора таких откликов. Пусть элементарный отклик имеет вид функции Гаусса, ширина которых для каждого источника своя. На рис. 1, *a, б* приведен пример симуляции такого сигнала из двух составляющих (подробности под рисунком). Целью анализа в идеале является восстановление исходных составляющих.

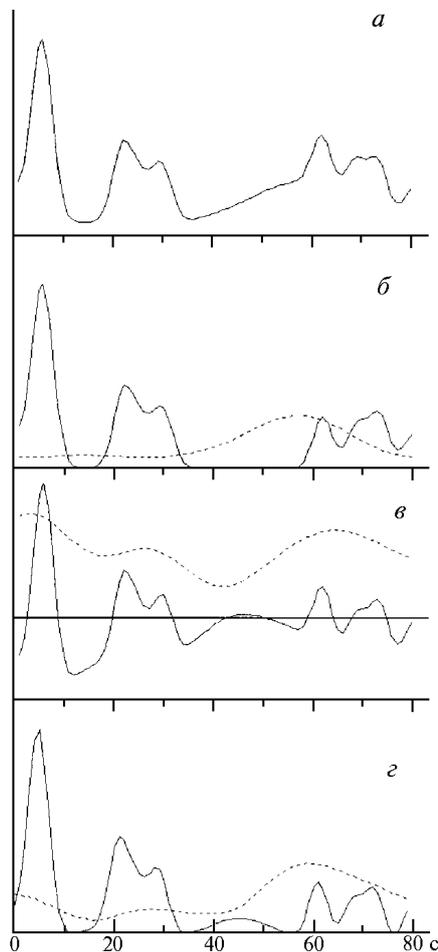


Рис. 1. Демонстрационный пример сигнала, набранного из гауссовых пиков разной ширины. Сигнал составлен из 5 гауссовых пиков с шириной 20 с («медленный») и 10 пиков с шириной 4 с («быстрый»). Пики имели случайную амплитуду и положение по оси абсцисс. Граничные условия периодические. *a* – суммарный сигнал; *б* – быстрая и медленная компоненты; *в* – результат разделения методом Фурье; *г* – результат разделения методом ПСФ

На рис. 1, в показан результат стандартного спектрального разделения методом Фурье. Видно, что метод не позволяет получить более одной положительно определенной компоненты и что целиком теряется информация о среднем значении высокочастотной компоненты.

Теперь посмотрим, какие возможности содержит предлагаемый метод ПСФ. Во-первых, получим, как описано раньше, верхнюю оценку интегрального спектра функции. На рис. 2 приведен спектр $S(\omega)$ для сигнала из примера на рис. 1. На спектре хорошо видно плато в области низких частот около $\omega = 0,05 \text{ с}^{-1}$. Это дает естественную границу для разделения сигнала на две частотные компоненты, f_ω и f'_ω . Эти компоненты приведены на рис. 1, з. Видно, что они неплохо аппроксимируют исходные составляющие.

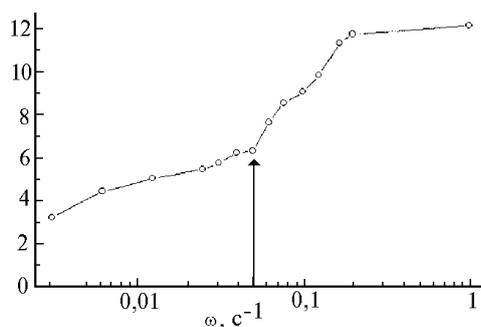


Рис. 2. Спектр верхней оценки НЧ-составляющей ПСФ сигнала рис. 1. Стрелкой помечена частота разделения спектра

7. Применение метода ПСФ для анализа реальных данных

Данные для анализа взяты из результатов, полученных в ходе исследования загрязнения атмосферы г. Новосибирска при помощи передвижной лаборатории ИХКиГ [6]. Измерения проводились в центральной части города в декабре 1995 г. Для анализа выбраны данные по концентрации газовых примесей SO_2 и NO_x в ночь с 3 на 4 декабря. Исходные данные приведены на рис. 3.

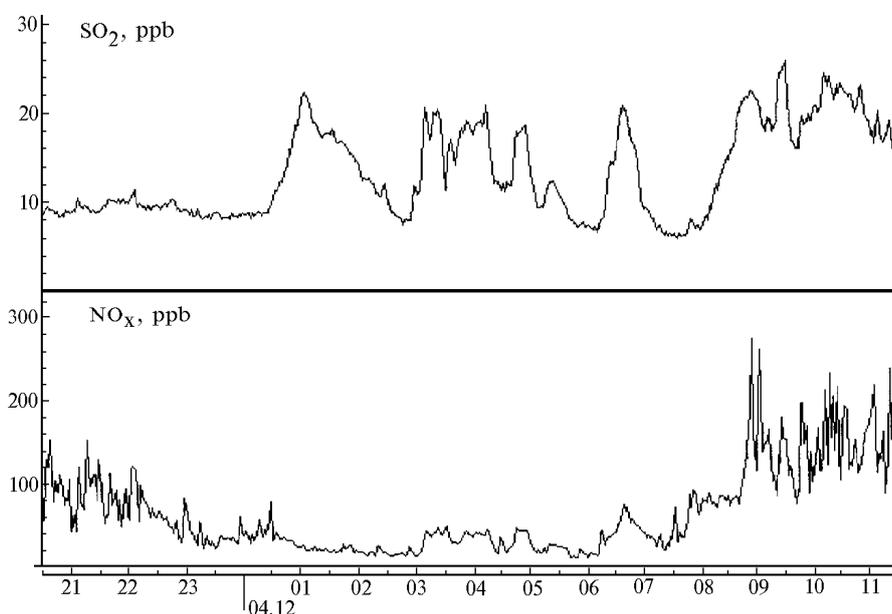


Рис. 3. Временная зависимость концентрации SO_2 и NO_x в центральной части г. Новосибирска в ночь с 3 на 4 декабря 1995 г.

Из рис. 3 видно, что концентрация примесей падает ночью и что на фоне медленных изменений присутствуют быстрые (длительностью в несколько минут) пики, обусловленные, по-

видимому, проходящим поблизости транспортом. Можно выделить период (примерно с полуночи до 7:30), когда вклад транспорта относительно невелик.

Для спектрального разделения сигналов были выбраны два характерных времени: 60 мин для отделения медленного дрейфа и 5 мин для отсечения быстрых флуктуаций. Вначале выделялась низкочастотная компонента (с $\omega^{-1} = 60$ мин), а затем остаток подвергался новому разделению с $\omega^{-1} = 5$ мин. Таким образом, были получены низко-, средне- и высокочастотная компоненты сигналов, приведенные на рис. 4.

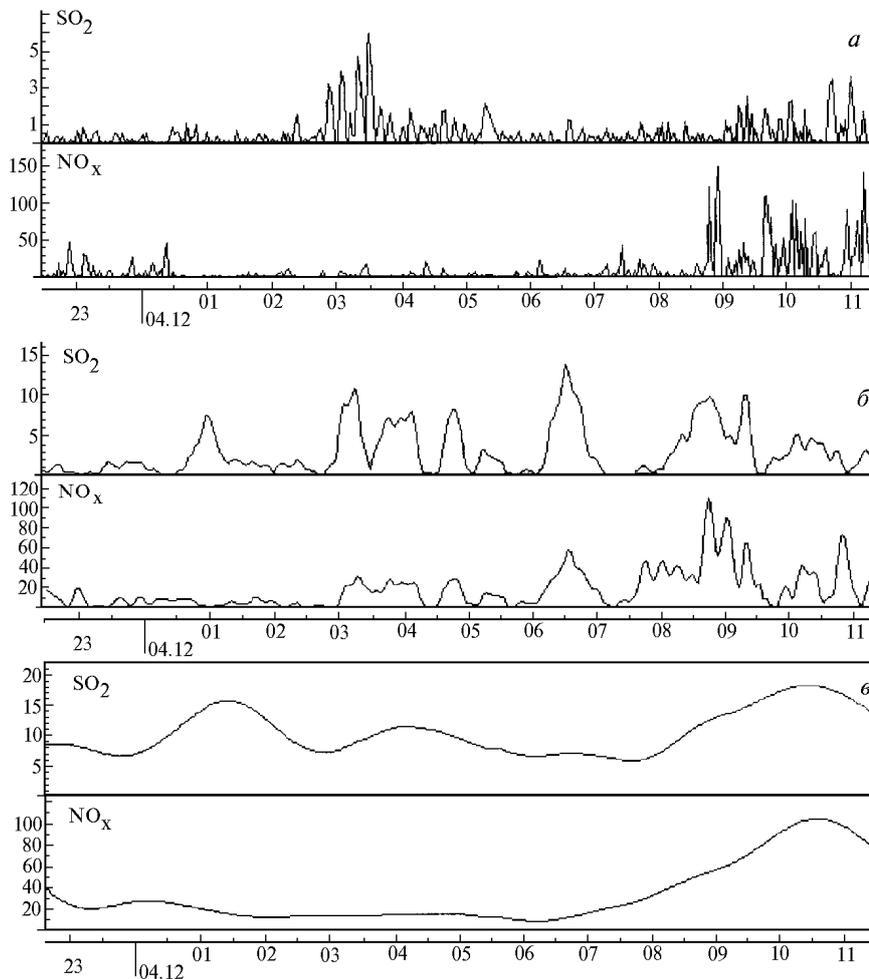


Рис. 4. Результат частотного разделения данных рис. 3 методом ПСФ: *а* – высокочастотная компонента, *б* – среднечастотная, *в* – низкочастотная. Характерные времена разделения 5 и 60 мин. Все концентрации в ppb

Из рис. 4 видно, что низкочастотные компоненты ведут себя симбатно утром, а среднечастотные хорошо коррелируют между собой ночью. Корреляционный анализ компонент дал следующие результаты (таблица).

Коэффициенты корреляции между SO₂ и NO_x

Компонента	Весь интервал	Ночь (0:00–7:30)
Низкочастотная	0,73	0,13
Среднечастотная	0,50	0,84
Высокочастотная	0,07	-0,17
Исходный сигнал	0,48	0,35

Не вдаваясь в более детальный анализ, можно сказать, что спектральное разделение позволило отделить сигнал от дымовых факелов от медленно меняющегося фона и быстрых флуктуаций, обусловленных близкими источниками. После выделения среднечастотной ком-

поненты в ней проявилась сильная корреляция между концентрациями SO_2 и NO_x (см. таблицу), что позволило определить газовый состав этой компоненты, а именно: установить соотношение $\text{SO}_2/\text{NO}_x = 0,074 \pm 0,006$ (линейная регрессия для среднечастотной компоненты с 0 до 7:30). Характерное время флуктуаций и газовый состав позволили предположительно идентифицировать этот сигнал с прохождением факела небольшой котельной.

На основе предложенного подхода создана компьютерная программа, которая используется в ИХКиГ СО РАН и полезна для изучения динамики газовых и аэрозольных примесей в атмосфере с характерными временами от минут до месяцев.

Из приведенных примеров видно, что предлагаемый подход при всей своей неточности может быть эффективен для частотной фильтрации положительно определенных сильно флуктуирующих сигналов и заслуживает дальнейшего развития. Главным недостатком его в настоящем виде является то, что оценка НЧ-составляющей f_ω принципиально односторонняя, и потому не вполне корректно трактовать остаточную функцию f как соответствующую высокочастотную составляющую. Это и указывает направление дальнейшего развития как поиск подходов, позволяющих сделать оценку двусторонней.

1. Paatero P., Tapper U. // *Environmetrics*. 1994. V. 5. P. 111–126.
2. Paatero P., Makela J.M., Jokinen V. // *J. Aeros. Sci.* 1995. V. 26. P. S269–S270.
3. Thomson D.J. // *J. Fluid Mech.* 1990. V. 210. P. 113–153.
4. Бородулин А.И., Майстренко Г.М., Чалдин Б.М. Статистическое описание распространения аэрозолей в атмосфере. Новосибирск: НГУ, 1992.
5. Wilson D.J., Robins A.G., Fackrell J.E. // *Atm. Environ.* 1982. V. 16. N 3. P. 497–504.
6. Rwyer S. et al. // *J. Aeros. Sci.* 1995. V. 26. P. 379–381.

Институт химической кинетики и горения СО РАН,
Новосибирск

Поступила в редакцию
15 января 1997 г.

S. I. Erementsko. **Spectral Filtration of Positive Definite Fluctuation Signals.**

The mathematical Positive Spectral Decomposition technique is developed to analyze non-negative random signals. The technique allows one to divide mixed signal into the components that differ in fluctuation characteristic time. The technique is most effective in application to strongly fluctuating signals, such as atmospheric pollution concentrations. The application to dividing of the city gas pollution signal into the components corresponding to the sources of different distances is presented.