

ИСПОЛЬЗОВАНИЕ НЕПРЕРЫВНОГО ВСПЛЕСК-ПРЕОБРАЗОВАНИЯ ДЛЯ ПАРАМЕТРИЧЕСКОЙ СПЕКТРАЛЬНОЙ ОБРАБОТКИ НЕСТАЦИОНАРНЫХ СИГНАЛОВ

Бочкарев В.В., Кацевман М.М., Петрова И.Р., Теплов В.Ю.

Казанский государственный университет

Введение. В последнее время все более широкое применение находят методы спектрального анализа, основанные на использовании всплеск-преобразования. Однако, разрешающая способность всплеск-преобразования, так же как и преобразования Фурье, зависит от длины анализируемого ряда данных. Она ограничена величиной $1/T$ (T - длина ряда) и недостаточна для коротких рядов. Физическая природа сигналов и особенности проведения исследований часто не позволяют получить серии данных достаточной длины. Применение параметрических методов может существенно увеличить разрешающую способность. В этой работе мы делаем попытку объединить достоинства всплеск- преобразований и параметрических методов высокого разрешения.

Используемая модель. Для построения параметрической модели принимается предположение, что исследуемый сигнал удовлетворяет некоторому подходящему линейному обыкновенному дифференциальному уравнению с постоянными коэффициентами:

$$y^{(r)} + a_1 \cdot y^{(r-1)} + \dots + a_r \cdot y = \xi(x) \quad (1)$$

Если записать входящие сюда производные в разностной форме, в зависимости от принимаемых предположений о свойствах процесса $\xi(x)$ получим модели, используемые в других параметрических методах. В предположении, что $\xi(x)$ – белый шум, получим модель авторегрессионного метода, предположив, что это импульс, получим полиэкспоненциальную модель, используемую в методе Прони [1]. В данной работе рассматривается второй случай. Использование континуальной модели процесса позволяет анализировать ряды с неэквидистантными отсчетами, что важно для широкого круга геофизических применений, а так же при обработке сигналов в системах измерения параметров релаксации и диффузии методом ядерного магнитного резонанса.

Для определения коэффициентов модельного дифференциального уравнения мы предлагаем использовать всплеск-преобразование, основанное на серии непрерывных всплесков. Применение непрерывного преобразования также обусловлено необходимостью применения к неэквидистантным рядам данных. Выберем систему всплесков ψ_m , обладающую следующим свойством:

$$(W_k f')_{\tau, \lambda} = \sum_m C_{k,m} \cdot (W_m f)_{\tau, \lambda}$$

т.е., образ производной функции должен линейно выражаться через образы самой функции (здесь и ниже приняты обозначения ψ_m для m -ого всплеска серии и $(W_m f)_{\tau, \lambda}$ для соответствующего ему образа функции f при сдвиге τ и масштабе λ). В этом случае задача определения коэффициентов модельного уравнения сводится к линейной регрессии. Простейший способ построения такой системы всплесков состоит в выборе некоторой базовой функции, быстро убывающей $x \rightarrow \pm\infty$, и определению

$$\psi_m(x) \sim \frac{\partial^m}{\partial x^m} \varphi(x) \quad (2)$$

Известна хорошо изученная последовательность всплесков, получающаяся при выборе в качестве базовой функции гауссоиды $\exp\{-x^2/2\}$. Второй всплеск последовательности известен как mexihat-всплеск. Использование гауссоиды дает очень хорошую пространственную и масштабную локализацию, что в нашем случае позволяет избежать появления возможных краевых эффектов, поэтому остановимся на таком выборе. Считая, что функция задана на интервале $x \in [A, B]$, интегрированием по частям устанавливаем для всплеск-образов серии вида (2) соотношение:

$$(W_n f')_{\tau, \lambda} = -\frac{1}{\lambda} (W_{n+1} f)_{\tau, \lambda} + \frac{1}{\sqrt{\lambda}} \left\{ f(B) \psi_n \left(\frac{B-\tau}{\lambda} \right) - f(A) \psi_n \left(\frac{A-\tau}{\lambda} \right) \right\} \quad (3)$$

Вдали от границ области ($B-x, x-A \ll \lambda$) второе слагаемое в формуле стремится к нулю, что значительно упрощает вывод. Возьмем преобразование по базису ψ_n от обеих частей уравнения (1), r -кратным применением соотношения (3) получим

$$\left(-\frac{1}{\lambda}\right)^r \cdot (W_{n+r} f)_{\tau, \lambda} + a_1 \cdot \left(-\frac{1}{\lambda}\right)^{r-1} \cdot (W_{n+r-1} f)_{\tau, \lambda} + \dots + a_r \cdot (W_n f)_{\tau, \lambda} = (W_n \xi)_{\tau, \lambda}$$

Записывая это соотношение для различных τ, λ, n , и выбрав модель процесса ξ , получаем переопределенную систему линейных уравнений для оценки a_i . Для полиэкспоненциальной модели отсюда по критерию наименьших квадратов следует $A_{ij} a_j = b_i$,

$$A_{i,j} = \sum_{\tau,\lambda,n} (-\lambda)^{i+j-2r} \cdot P_{\tau,\lambda,n} \cdot (\overline{W}_{n+r-i} f)_{\tau,\lambda} \cdot (W_{n+r-j} f)_{\tau,\lambda}, \quad (4)$$

$$b_i = \sum_{\tau,\lambda,n} (-\lambda)^{i-r} \cdot P_{\tau,\lambda,n} \cdot (\overline{W}_{n+r-i} f)_{\tau,\lambda} \cdot (W_{n+r} f)_{\tau,\lambda}$$

где $P_{\tau,\lambda,n}$ – некоторая весовая функция.

Полученная таким образом модель может быть использована для оценки спектральной плотности мощности, для аппроксимации экспериментальных зависимостей, в задаче прогнозирования.

Выбор рабочей области. Значения τ , λ могут выбираться в области достоверности ряда. Наименьший масштаб, который может быть использован для определения модельных коэффициентов, определяется погрешностью, возникающей при вычислении непрерывного всплеск- преобразования по дискретному ряду отсчетов. Если функция ограничена вместе со своими первой и второй производными, причем $M_f^{(k)} = \max(f^{(k)}(x))$, имеем следующую оценку для погрешности при использовании формулы трапеций:

$$\delta(W_n f)_{a,\lambda} = \frac{h^2}{24} \cdot \sqrt{\lambda} \cdot \left\{ M_f^{(2)} \cdot \mu_n + \frac{2}{\lambda} \cdot M_f^{(1)} \cdot \mu_{n+1} + \frac{1}{\lambda^2} \cdot M_f^{(0)} \cdot \mu_{n+2} \right\},$$

$$\text{где } \mu_m = \int_{-\infty}^{+\infty} |\Psi_m(x)| dx$$

Таким образом, получается треугольная область на плоскости (τ,λ) .

Влияние аддитивного шума. Предположим, что сигнал $z = y + \zeta$ состоит из полезного сигнала, определяемого моделью (1), и аддитивного белого шума с дисперсией σ^2 . Присутствие аддитивного шума приводит, так же как и для других линейных параметрических методов, к появлению среднего смещения коэффициентов уравнений, и, следовательно, к наличию регулярной ошибки. В принятом приближении нетрудно подсчитать смещение математического ожидания коэффициентов определяющей системы уравнений:

$$\delta\langle A_{i,j} \rangle = \sum_{\tau,\lambda} (-\lambda)^{i+j-2r} \cdot P_{\tau,\lambda} \cdot \int_{-\infty}^{+\infty} \overline{\Psi}_{n+r-i}(x) \cdot \Psi_{n+r-j}(x) dx, \quad 1 \leq i, j \leq r-1$$

$$\delta\langle b_i \rangle = -\delta\langle A_{i,0} \rangle$$

Для выбранной системы последнее выражение принимает вид

$$(-1)^{\frac{i-j}{2}} \cdot \sqrt{\pi} \cdot \frac{[2n+2r-i-j-1]!!}{2^{n+r-\frac{i+j}{2}}} \cdot \sum_{\tau,\lambda} \lambda^{i+j-2r} \cdot P_{\tau,\lambda}$$

для случая четных $i+j$, и 0 для нечетных. Если известен уровень шума σ , несложно внести соответствующие поправки к коэффициентам формул (4), и получить несмещенную оценку спектральных параметров. Аналогично находятся поправки для небелого шума с известной корреляционной функцией.

Определение амплитуд спектральных составляющих. Рассмотрим соотношения между (W_{mf}) в пределах угла влияния начальной точки ряда A . Используя (3) и сгруппировав слагаемые при $f^{(k)}(A)$, получим выражение

$$\frac{1}{\sqrt{\lambda}} \cdot \sum_{k=0}^{r-1} \left\{ \sum_{m=0}^{r-1-k} (-\lambda)^{-m} \cdot \Psi_{n+m} \left(\frac{A-\tau}{\lambda} \right) \cdot a_{r-1-k} \right\} \cdot f^{(k)}(A) = \sum_{m=0}^r (-\lambda)^{-r} \cdot (W_{n+r} f)_{\tau,\lambda} \cdot a_{r-m}$$

Если a_i уже определены, то $f^{(k)}(A)$ находится из этой системы, после чего определить амплитуды не представляет труда. Достоинством данного способа определения амплитуд (по сравнению с тем, как это делается в методе Прони) является то, что при этом используется информация о структуре модели, т.е., о поведении аппроксимируемой функции на всем интервале. Это уменьшает вероятность появления ложных максимумов в оценке спектральной плотности мощности.

Пример применения. На рис.1. представлен спектр сигнала ЯМР, измеренный методом ССИ с использованием спектрометра высокого разрешения TESLA-BS587A [2]. В качестве тестового образца использовался бутанол, имеющий 4 спектральные линии. Для сравнения представлены результаты обработки данным параметрическим методом по 56 первым точкам и классическим Фурье-анализом с окном Кайзера по 4096 точкам. Сокращение длины исходного ряда необходимого для получения приемлемого разрешения позволяет исследовать образцы с малыми временами релаксации и сокращать время затрачиваемое на проведение диффузионных измерений. Ранее для решения подобных задач нами применялись методы Прони и Куликова [3]. По сравнению с ними применение разработанного метода приводит к получению более качественных и достоверных результатов.

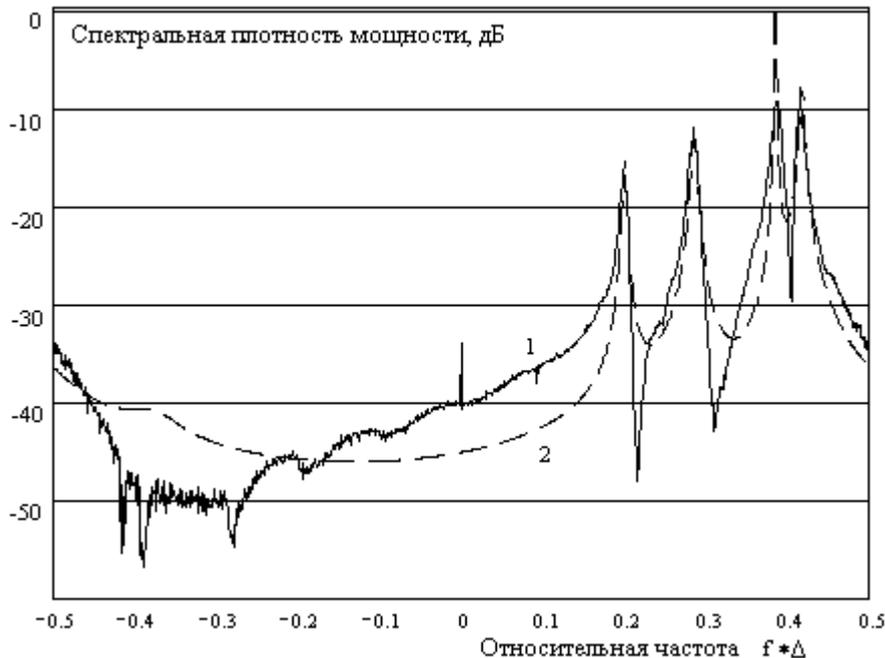


Рис.1. Спектр ЯМР бутанола.

(1- оценка, полученная с помощью дискретного преобразования Фурье;
2. параметрическая оценка)

Литература

1. С.Л.Марпл-мл. Цифровой спектральный анализ и его приложения. М.: Мир, 1990 г. 584 С.
2. Идиятуллин Д.Ш., Идиятуллин З.Ш., Идиятуллин Б.З., Теплов В.Ю.
3. Теплов В.Ю., Бочкарев В.В. Алгоритмы определения параметров релаксации, диффузии и потоков методом ЯМР. //Структура и динамика молекулярных систем. Сб. статей, Ч.2 Йошкар-Ола – Москва, 1998 г. с. 28-32.

USAGE OF CONTINUOUS WAVELET - TRANSFORMATION FOR PARAMETRIC SPECTRAL PROCESSING OF NON-STEADY SIGNALS

Bochkarev V.V., Katsevman M.M., Petrova I.R., Teplov V.Yu.

Kazan State University

The methods of spectral analysis, founded on usage of wavelet - transformation are widely applied recently. However, the resolving power of wavelet - transformation, as well as Fourier transforms, depends on length of data series. It is limited to value $1/T$ (T - length of a series) and is poor for short series. The physical nature of signals and the features of researches often do not allow to receive series of the data of sufficient length. The application of parametric methods can essentially increase resolving power. In this paper we attempt to aggregate advantages wavelet - transformations and parametric methods of high-res.

For construction of a parametric model we receive the supposition, that the studied signal contents to some linear ordinary differential equation with constant factors:

$$y^{(r)} + a_1 \cdot y^{(r-1)} + \dots + a_r \cdot y = \xi(x)$$

If to record entering here derivative in the finite-differential form, depending on the received suppositions about properties of process $\xi(x)$, we receive models used in other parametric methods. If to suspect, that $\xi(x)$ - white noise, we shall receive model auto regression method. Having suspected, that it is impulse, we shall receive polyexponential model used in a Prony method. In this paper we consider the second case. Usage of continuum model of process allows to analyse series with not equidistant readouts, that is important for a wide areas of geophysical applications, and as at a signal processing in systems of measurement of parameters of a relaxation and diffusion by a method of nuclear magnetic resonance. For definition of coefficient of a model differential equation we offer to use wavelet - transformation, founded on a series of continuous wavelets. The application of continuous transformation also is conditioned by necessity of application to non-equidistant series of data. Let's select a system of wavelets ψ_m , having following property:

$$(W_k f')_{\tau, \lambda} = \sum_m C_{k,m} \cdot (W_m f)_{\tau, \lambda}$$

The image of a derivative of function should linearly be connected to images of the function. In this case problem of definition of factors of a model equation is reduced to linear regression. The elementary way of construction of such system of wavelets consists in selection of some base function fast decreasing $x \rightarrow \infty$, and definition

$$\psi_m(x) \sim \frac{\partial^m}{\partial x^m} \varphi(x)$$

The well studied series of wavelets receiving by selection as a base function gaussoid $\exp\{-x^2/2\}$ is known. The second splash of series is known as mexihat-wavelet. Usage of gaussoid gives very good time and scale localization, that in our case allows to avoid appearance of possible edge effects, therefore we shall stay on such selection.

Considering, that the function presets on an interval $x \in [A, B]$, by an integration by parts we set for wavelet - images of a series of a kind (2) a ratio:

$$(W_n f')_{\tau, \lambda} = -\frac{1}{\lambda} (W_{n+1} f)_{\tau, \lambda} + \frac{1}{\sqrt{\lambda}} \left\{ f(B) \psi_n \left(\frac{B-\tau}{\lambda} \right) - f(A) \psi_n \left(\frac{A-\tau}{\lambda} \right) \right\}$$

Far from edges of area ($B-x, x-A \ll 1$) the addend in the formula aims at zero point, that considerably simplifies a conclusion. Let's take transformation on basis ψ_n from both parts of an equation (1), by r -multiple application of a ratio (3) we shall receive

$$\left(-\frac{1}{\lambda}\right)^r \cdot (W_{n+r} f)_{\tau, \lambda} + a_1 \cdot \left(-\frac{1}{\lambda}\right)^{r-1} \cdot (W_{n+r-1} f)_{\tau, \lambda} + \dots + a_r \cdot (W_n f)_{\tau, \lambda} = (W_n \xi)_{\tau, \lambda}$$

Recording this ratio for different τ, λ, n , and having selected model of process ξ , we receive the reassigned system of simple equations for an estimation a_i . For polyexponential model from here on a least squares criterion follows $A_{i,j} a_j = b_i$,

$$A_{i,j} = \sum_{\tau, \lambda, n} (-\lambda)^{j+i-2r} \cdot P_{\tau, \lambda, n} \cdot (\overline{W}_{n+r-i} f)_{\tau, \lambda} \cdot (W_{n+r-j} f)_{\tau, \lambda},$$

$$b_i = \sum_{\tau, \lambda, n} (-\lambda)^{i-r} \cdot P_{\tau, \lambda, n} \cdot (\overline{W}_{n+r-i} f)_{\tau, \lambda} \cdot (W_{n+r} f)_{\tau, \lambda}$$

The obtained thus model can be utilised for an estimation of a spectral concentration of power, for approximation of experimental relations, in a problem of prediction.

Let's suspect, that the signal $z = y + \zeta$ consists of a desired signal instituted by model (1), and additive white noise with a dispersion σ^2 . The presence of an additive noise results in appearance of mean displacement of coefficients of equations, and, therefore, to availability of a regular error. In an adopted approaching it is uneasy to calculate displacement of average of distribution of coefficients of a determining equations set:

$$\delta \langle A_{i,j} \rangle = \sum_{\tau,\lambda} (-\lambda)^{i+j-2r} \cdot P_{\tau,\lambda} \cdot \int_{-\infty}^{+\infty} \bar{\Psi}_{n+r-i}(x) \cdot \Psi_{n+r-j}(x) dx, \quad 1 \leq i, j \leq r-1$$

$$\delta \langle b_i \rangle = -\delta \langle A_{i,0} \rangle$$

For a selected system last expression becomes

$$(-1)^{\frac{i-j}{2}} \cdot \sqrt{\pi} \cdot \frac{[2n+2r-i-j-1]!!}{2^{n+r-\frac{i+j}{2}}} \cdot \sum_{\tau,\lambda} \lambda^{i+j-2r} \cdot P_{\tau,\lambda}$$

For a case even $i+j$, and 0 for odd. If the noise level σ is known, it is possible to introduce the conforming corrections to coefficients of the formulas (4), and to receive a unbiased estimator of spectral parameters. Similarly we calculate the corrections for a non-white noise with a known correlation function.

Let's consider ratio between $(W_m f)$ within the limits of a angle of influencing of an first point of a series A.

Using (3) and having grouped together addends at $f^{(k)}(A)$, we shall receive expression

$$\frac{1}{\sqrt{\lambda}} \cdot \sum_{k=0}^{r-1} \left\{ \sum_{m=0}^{r-1-k} (-\lambda)^{-m} \cdot \Psi_{n+m} \left(\frac{A-\tau}{\lambda} \right) \cdot a_{r-1-k} \right\} \cdot f^{(k)}(A) = \sum_{m=0}^r (-\lambda)^{-r} \cdot (W_{n+r} f)_{\tau,\lambda} \cdot a_{r-m}$$

If a_i are already determined, that $f^{(k)}(A)$ is defined from this system then to define amplitudes presents no difficulty.

Usage of the information about structure of model, i.e., about behavior of an approximable function on all interval reduces probability of appearance of false maxima in an estimation of a power spectral density.