УДК 550.834

О корректности задачи нелинейной регрессии при мониторинге природных и рукотворных объектов

© *В. С. Мостовой, С. В. Мостовой, 2012*

Институт геофизики НАН Украины, Киев, Украина

Поступила 1 ноября 2011 г.

Представлено членом редколлегии В. Н. Шуманом

Розглянуто нелінійні моделі моніторингу, які побудовано на основі суперпозиції осциляторів з вільними параметрами, даними спостережень, як задачі нелінійної регресії. Для вибраних нелінійних математичних моделей з'ясовано питання, пов'язані з існуванням розв'язку, його єдиністю і стійкістю залежно від початкових даних. Остання обставина особливо важлива, оскільки алгоритми, побудовані на підставі цих моделей, орієнтовані на безпосередню обробку польових спостережень, що означає залежність від характеристик вимірювальної апаратури, помилок виміру і супутнього фону перешкод. Для побудови оптимальних оцінок параметрів моделі запропоновано відокремлення лінійних і нелінійних параметрів з метою оптимізації процесу обчислення. Під час пошуку квазіоптимальних розв'язків такий розподіл дає змогу використовувати тільки нелінійні вільні параметри для симуляції в методі Монте-Карло. Параметри, що входять лінійно, визначають розв'язком системи лінійних рівнянь. Таким чином, розмірність завдання пошуку оптимальних оцінок зменшується на розмірність вектора лінійних параметрів.

Under consideration there is a compliance with observed data and nonlinear models of monitoring. These models are based on superposition of oscillators with free parameters. Optimal estimation of free parameters of model which enter into model both linearly and nonlinearly, we shall consider as a problem of nonlinear regression. The optimality is understood in sense of a global minimum of an objective functional. The point in space of possible values of free parameters of model in which criterion has a global minimum is accepted as the optimal solution of a problem. For the chosen nonlinear mathematical models it is necessary to find out the questions connected with existence of the solution, its uniqueness, and stability of the solution depending on initial data. The last circumstance is especially important, as the algorithms constructed on the basis of these models, are concentrated on direct processing of field supervision. It means dependence on characteristics of the measuring equipment, errors of measurement and to accompanying by background noises. Separation of linear and nonlinear parameters with the purpose of calculation process optimization is offered for construction of optimal estimations model parameters. By search quasi-optimal solutions such division allows to use for the Monte-Carlo technique simulation only nonlinear parameters. Linearly entering parameters are defined by the solution of system of the linear equations. Thus, dimension of a search problem of optimal estimations is decreased on a size of a linear parameters vector dimension.

Вступление. Согласие нелинейных моделей мониторинга, основанных на суперпозиции осцилляторов со свободными параметрами, которые предложены в работах [Мостовой, 2008 а, б; Мостовой В. С., Мостовой С. В., 2008; 2011 а, б; Мостовой и др., 2011], с наблюденными данными, будем рассматривать как задачу нелинейной регрессии [Виноградов, 1977; Bethea et al., 1985]. Для выбранных нелинейных математических моделей необходимо выяснить вопросы, связанные с существованием решения, его единственностью и устойчивостью в зависимости от начальных данных. Последнее обстоятельство особенно важно, поскольку алгоритмы, построенные на основании этих моделей, ориентированы на непосредственную обработку полевых наблюдений, а это означает зависимость от характеристик измерительной аппаратуры, ошибок измерения и сопутствующего фона помех.

Для построения оптимальных оценок параметров модели предлагается разделение линейных и нелинейных параметров с целью оптимизации процесса вычисления. При поиске квазиоптимальных решений такое разделение позволяет использовать только нелинейные свободные параметры для симулирования в методе Монте-Карло. Линейно входящие параметры определяются решением системы линейных уравнений. Таким образом, размерность задачи поиска оптимальных оценок уменьшается на размерность вектора линейных параметров.

Существование решения задачи регрессии. Предположим, что A — компактное множество в $\Re^n$, B — произвольное множество в $\Re^m$, где n и m — натуральные числа; $F(\cdot,\cdot)$ — непрерывная функция $A\times B\to\Re$, где $A\times B$ — прямое произведение множеств A и B. Для произвольного $y\in B$ рассмотрим следующую оптимизационную задачу:

$$\min_{x\in A} F(x,y)\,. \tag{1}$$

Лемма 1. *Для произвольного $y\in B$ существует точка $\hat{x}(y)$, минимизирующая $F(\cdot,y)$ на множестве A.*

Доказательство. По теореме о поведении непрерывной функции, заданной на компактном множестве, функция $F(\cdot,y)$ достигает точной нижней грани на множестве A. Следовательно, существует точка $\hat{x}(y)$, минимизирующая $F(\cdot,y)$ на множестве A.

Единственность решения задачи (1). При дополнительных условиях на функцию $F(\cdot,\cdot)$ можно показать, что множество точек $y\in B$, таких что решение задачи (1) не единственно, имеет меру Лебега ноль. Физическая или, скорее, вероятностная интерпретация данного утверждения следующая: если предположить, что результаты экспериментов — случайные величины с любым непрерывным распределением, например гауссовым или равномерным (на подмножестве $\Re^m$), то вероятность того, что решение задачи (1) не единственно, равна нулю.

Таким дополнительным условием на функции вида $F(\cdot,\cdot)$ может быть представление $F(\cdot,\cdot)$ в виде композиции непрерывной (но не равной константе) функции и функции, заданной суперпозицией осцилляторов, а также представление функции $F(\cdot,\cdot)$ в виде полинома, отличного от константы. Заметим, что единственность решения оптимизационных задач вида (1), рассмотренных в настоящей статье, подтверждена практическими их исследованиями в численном эксперименте и обработке полевых наблюдений.

Непрерывная зависимость от начальных данных. ***Лемма 2.*** *Пусть $C=\{y\in B$: решение задачи (1) не единственно$\}$. Если дополнение C, $C^c=B\setminus C$ — открытое множество в B, то на множестве C^c решение задачи (1), $\hat{x}=\hat{x}(y)$, непрерывно зависит от второй компоненты функции $F(\cdot,\cdot)$.*

Доказательство. Из открытости C^c и непрерывности $F(\cdot,\cdot)$ следует, что для любого $y\in C^c$ существует некоторый шар (в евклидовом пространстве) B_δ такой, что для любой точки $\hat{y}\in B_\delta$ (т. е. при $\left\|y-\hat{y}\right\|<\delta$) получим решение задачи (1), $\hat{x}(y)$, удовлетворяющее $\left\|\hat{x}(y)-\hat{x}(\hat{y})\right\|<\varepsilon$ для некоторого $\varepsilon>0$. Следовательно, для любого $y\in C^c$ при $\varepsilon>0$ существует такое $\delta>0$, при котором $\left\|\hat{x}(y)-\hat{x}(\hat{y})\right\|<\varepsilon$ для любого $\hat{y}:\left\|\hat{y}-y\right\|<\delta$. Таким образом, решение задачи (1) – непрерывная функция от второй компоненты функции $F(\cdot,\cdot)$ на множестве C^c.

Отметим, что открытость множества C^c выполняется для функций, используемых в постановке задачи регрессии (1), на практике. Как было указано, для практических задач, рассмотренных в данной статье, $L(C)=0$, где $L(C)$ — мера Лебега множества C. С вероятностной точки зрения лемма 2 показывает, что (при некоторых условиях на $F(\cdot,\cdot)$) решение задачи регрессии непрерывно зависит от начальных данных с вероятностью 1.

Таким образом, задачи вида (1) для функций $F(\cdot,\cdot)$, рассмотренных в настоящей статье, являются корректными [Evans, 1998] с практической точки зрения. При этом строго можно показать лишь первое условие корректности — существование решения. Остальные два условия — единственность и непрерывная зависимость от начальных данных — выполняются при дополнительных условиях на функцию $F(\cdot,\cdot)$ и подтверждаются практическими исследованиями функций вида $F(\cdot,\cdot)$.

О сходимости алгоритма решения задачи регрессии. Предположим, что A — компактное подмножество $\Re^d$, где d — натуральное число; $F(\cdot)$ — непрерывная функция $A\to\Re$. Рассмотрим следующую оптимизационную задачу:

$$\min_{x\in A} F(x)\,. \tag{2}$$

Исследуем следующий алгоритм поиска приближенного решения задачи (2): на множестве A выберем некоторую вероятностную меру P такую, что для любого множества $C\subseteq A$ с положительной мерой Лебега ($L(C)>0$) будет выполняться условие $P(C)>0$.

Выбросим N случайных точек $x_n\in A$, $n=\overline{1,N}$, каждая из которых имеет распределение, удовлетворяющее условиям, описанным в предыдущем пункте, так что x_n, $n=\overline{1,N}$, — независимые одинаково распределенные случайные величины на A.

Для каждого $n=\overline{1,N}$, используя алгоритм Левенберга—Марквардта [Levenberg, 1944; Marquardt, 1963] с начальной точкой x_n, находим точку локального минимума $\hat{x}_n$.

Приближенным решением задачи (2) назовем такую точку $y_N \in \left\{\hat{x}_1,\ldots,\hat{x}_N\right\}$, что $y_N = \min\limits_{n=\overline{1,N}} F(\hat{x}_n)$.

Предположим, что δ-критерий остановки в алгоритме Левенберга—Маркварда, т. е. алгоритм Левенберга—Маркварда, выполняется (и утверждаем, что локальный минимум найден), если уменьшение значения функции $F(\cdot)$ при двух последовательных итерациях алгоритма не превосходит δ.

Лемма 3. *Алгоритм поиска глобального минимума функции $F(\cdot)$ сходится к решению задачи (2) с точностью δ, т. е.*

$$\lim_{N\to\infty} P\left[F(y_N)-\min_{x\in A}F(x)>\delta\right]=0.$$

Доказательство. Пусть z — точка, минимизирующая функцию $F(\cdot)$ на множестве A, т. е. $F(z)=\min\limits_{x\in A}F(x)$. Заметим, что существование z следует из леммы 1. Так как функция $F(\cdot)$ непрерывна, найдем такое число $\rho(z)>0$, что $B_{\rho(z)}(z)=\left\{y\in A:\|y-z\|\le\rho(z)\right\}$, шар радиуса $\rho(z)$ с центром в точке z, пересеченный с A, содержится в A, и для любого $z_1\in B_{\rho(z)}$ имеем $F(z_1)\ge F(z)$. Из условия (1) в описании алгоритма следует, что при выбросе случайной точки x выполняется следующее условие: $P\left[x\in B_{\rho(z)}(z)\right]>0$. Следовательно, используя независимость случайных величин x_n, $n=\overline{1,N}$, получаем, что при выбросе N точек (и соответственных N запусках алгоритма Левенберга—Макварда) приближенное решение задачи (2) y_N удовлетворяет следующему условию: $P[F(y_N)-F(z)\le\delta]\ge 1-\left(P\left[x\in B_{\rho(z)}(z)\right]\right)^N$. Так как $P\left[x\in B_{\rho(z)}(z)\right]>0$, заключаем, что $\lim\limits_{N\to\infty}P[F(y_N)-F(z)\le\delta]=1$.

Так как лемма 3 выполняется для любого критерия остановки $\delta>0$, то в частности из леммы 3 следует, что алгоритм поиска решения задачи (2) сходится по вероятности. Также следует отметить, что для задач, рассмотренных в настоящей статье, описанный алгоритм позволяет эффективно решать задачи вида (2). В частности, для этих задач скорость сходимости не является принципиальным вопросом, так как время поиска решения составляет считанные секунды. Тем не менее, скорость сходимости может быть определена при дополнительных условиях на функцию $F(\cdot)$ и множество A в задаче (2).

Выбор модели регрессии. В качестве моделей в мониторинге объектов, отражающих их динамику, выберем суперпозицию осцилляторов. Зарегистрированные колебания объекта $y(t)$ представим в виде

$$y(t)=M(t,\mathbf{h},\boldsymbol{\alpha})+n(t), \tag{3}$$

где $M(t,\ \mathbf{h},\ \boldsymbol{\alpha})$ — модель регрессии, $\mathbf{h}$ — вектор линейно входящих в модель параметров, $\boldsymbol{\alpha}$ — матрица нелинейно входящих параметров, $\mathbf{h}$ и $\boldsymbol{\alpha}$ — свободные параметры модели, $n(t)$ — аддитивный шум;

$$M_N(t,\mathbf{h},\boldsymbol{\alpha}) = \sum_{k=1}^{N} h_k\,\eta\left(t-\alpha_{k,1}\right)\exp\left(-\alpha_{k,2}\left(t-\alpha_{k,1}\right)\right)\times$$
$$\times\sin\left(\alpha_{k,2}\left(t-\alpha_{k,1}\right)\right). \tag{4}$$

В уравнении (4) матрица нелинейных параметров $\boldsymbol{\alpha}=\{\alpha_{k,s}\}$, $k=\overline{1,N}$; $s=\overline{1,3}$; $\eta(t)$ — единичная функция Хэвисайда, N — количество осцилляторов (параметр, также подлежащий оптимизации и фиксирующий модель). В рамках модели с номером N ищется оптимальное решение (1). Для множества моделей M_N, $N\in\overline{1,\mathfrak{M}}$ ($\mathfrak{M}$ — множество моделей данного класса), ищется модель $M_{\hat{N}}$, где $\hat{N}$ — номер модели, для которого (1) имеет глобальный экстремум.

Задача заключается в оптимальном выборе параметров $\mathbf{h}$ и $\boldsymbol{\alpha}$ для каждой модели M_N и переборе по N моделей M_N. Для простоты рассмотрим в качестве критерия оптимальности минимум расстояния между моделью $M(t,\mathbf{h},\boldsymbol{\alpha})$ и $y(t)$ в метрике $L_2\left[0,T\right]$, где $t\in\left[0,T\right]$ — область регистрации данных во времени. Такой выбор критерия эквивалентен максимуму правдоподобия при гипотезе о том, что $n(t)$ —белый шум.

Из выбора критерия следует, что для фиксированной матрицы $\boldsymbol{\alpha}$, полученной в результате симуляции методом Монте-Карло, для поиска вектора $\mathbf{h}$, оптимального для модели с нелинейными параметрами $\boldsymbol{\alpha}$, можно построить матрицу $\boldsymbol{\Psi}$ для системы линейных уравнений, решением которой будет $\mathbf{h}$:

$$\boldsymbol{\Psi}=\left\{\Psi_{k,s}\right\} \text{ и } \mathbf{I}=\left\{I_k\right\}, \tag{5}$$

$$\Psi_{k,q}=\int_0^T \eta\left(t-\alpha_{k,1}\right)\exp\left(-\alpha_{k,2}\left(t-\alpha_{k,1}\right)\right)\times$$
$$\times\sin\left(\alpha_{k,2}\left(t-\alpha_{k,1}\right)\right)\eta\left(t-\alpha_{q,1}\right)\exp\left(-\alpha_{q,2}\left(t-\alpha_{q,1}\right)\right)\times$$
$$\times\sin\left(\alpha_{q,2}\left(t-\alpha_{q,1}\right)\right)dt;\ k,q=\overline{1,N}, \tag{6}$$

$$I_k=\int_0^T \eta\left(t-\alpha_{k,1}\right)\exp\left(-\alpha_{k,2}\left(t-\alpha_{k,1}\right)\right)\times$$
$$\times\sin\left(\alpha_{k,2}\left(t-\alpha_{k,1}\right)\right)y(t)dt;\ k=\overline{1,N}. \tag{7}$$

Методом Монте-Карло генерируем псевдослучайную матрицу $\boldsymbol{\alpha}$ по априорным распределениям. Источником гипотез о параметрах распределения свободных параметров h_k, $\alpha_{k,s}$ служит априорный анализ преобразований Фурье и вейвлет.

Для каждой матрицы $\boldsymbol{\alpha}$, решая систему линейных уравнений, находим вектор $\mathbf{h}$, который является оптимальным для модели с матрицей нелинейных параметров $\boldsymbol{\alpha}$. Далее, используя эту точку ($\mathbf{h}$, $\boldsymbol{\alpha}$) (линейные и нелинейные параметры) в пространстве параметров методом Левенберга—Марквардта, находим ближайший локальный минимум. Процедура повторяется Q раз. На множестве Q экспериментов находятся оптимальные значения $\left(\hat{\mathbf{h}},\hat{\boldsymbol{\alpha}}\right)$ Как следует из леммы 3, с ростом Q обеспечивается сходимость по вероятности к оптимальному решению.

Список литературы

Виноградов И. М. Математическая энциклопедия. Т. 4. — Москва: Сов. энциклопедия, 1977. — 742 с.

Мостовой В. С. Математическая модель накопления сейсмических сигналов при активном мониторинге // Докл. НАН Украины. — 2008а. — № 4. — С. 132—136.

Мостовой В. С. Оптимальное обнаружение сигналов на фоне микросейсмического шума // Доп. НАН Украины. — 2008б. — № 1. — С. 106—110.

Мостовой В. С., Мостовой С. В. Вариационный подход к решению обратной задачи при накоплении сейсмических сигналов в активном мониторинге // Докл. НАН Украины. — 2008. — № 8. — С. 113—116.

Мостовой В. С., Мостовой С. В. Математическое моделирование оценки старения природных и техногенных объектов в системах мониторинга // Докл. НАН Украины. — 2011а. — № 7. — С. 114—118.

Мостовой В. С., Мостовой С. В. Оптимальные оценки нелинейных параметров в моделях сейсмоакустического мониторинга // Докл. НАН Украины. — 2011б. — № 8. — С. 103—107.

Мостовой В. С, Мостовой С. В., Кондра С. М., Страшко Ж. С. Оценка информативных параметров состояния строительных конструкций в режиме мониторинга // Промышленное строительство и инженерные сооружения. — 2011. — № 1. — С. 7—13.

Bethea R. M., Duran B. S., Boullion T. L. Statistical Methods for Engineers and Scientists. — New York: Marcel Dekker, 1985. — 447 p.

Evans L. C. Partial Differential Equations // Amer. Math. Soc. — 1998. — **19**. — 662 p.

Levenberg K. A. Method for the Solution of Certain Non-Linear Problems in Least Squares // Quart. Appl. Math. — 1944. — № 2. — P. 164—168.

Marquardt D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters // SIAM J. Appl. Math. — 1963. — **11**. — P. 431—441.