УДК 550.834

В. С. Мостовой

О корректности задачи нелинейной регрессии и сходимости алгоритма поиска глобального минимума в моделях мониторинга

(Представлено академиком НАН Украины В. И. Старостенко)

Рассматривается согласие нелинейных моделей мониторинга с наблюденными данными нелинейных моделей мониторинга. Эти модели основаны на суперпозиции осцилляторов со свободными параметрами. Оптимальную оценку свободных параметров модели, которые входят в модель как линейно, так и нелинейно, будем рассматривать как задачу нелинейной регрессии. Оптимальность понимается в смысле глобального минимума целевого функционала. Точка в пространстве возможных значений свободных параметров модели, в которой критерий имеет глобальный минимум, принимается как оптимальное решение. Для выбранных нелинейных математических моделей нужно выяснить вопросы, связанные с существованием решения, его единственностью и устойчивостью решения в зависимости от начальных данных. Последнее обстоятельство особенно важно, поскольку алгоритмы, построенные на основании этих моделей, ориентированы на непосредственную обработку полевых наблюдений, а значит, на зависимость от характеристик измерительной аппаратуры, ошибок измерения и сопутствующего фона помех.

Согласие нелинейных моделей мониторинга, основанных на суперпозиции осцилляторов со свободными параметрами [1–6], с наблюденными данными будем рассматривать как задачу нелинейной регрессии [7, 8]. Для выбранных нелинейных математических моделей нужно выяснить вопросы, связанные с существованием решения, его единственностью и устойчивостью решения в зависимости от начальных данных. Последнее обстоятельство особенно важно, поскольку алгоритмы, построенные на основании этих моделей, ориентированы на непосредственную обработку полевых наблюдений, а значит, зависимость от характеристик измерительной аппаратуры, ошибок измерения и сопутствующему фону помех.

1. Существование решения задачи регрессии. Предположим, что A — компактное множество в $\mathbb{R}^n$, B — произвольное множество в $\mathbb{R}^m$, где n и m — натуральные числа; $F(\cdot,\cdot)$ — непрерывная функция $A \times B \to \mathbb{R}$, 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 $\widehat{x}(y)$, 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 $\widehat{x}(y)$, минимизирующая $F(\cdot, y)$ на множестве A.

2. Единственности 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 мы предположим, что результаты экспериментов — случайные величины с любым непрерывным распределением, например Гауссовым или равномерным (на подмножестве $\mathbb{R}^m$), тогда 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 преставление функции $F(\cdot,\cdot)$ в виде полинома, отличного от константы. Заметим, что единственность решения оптимизационных задач вида (1), рассмотренных в работах [1–6], подтверждена практическими их исследованиями в численном эксперименте и обработке полевых наблюдений.

3. О непрерывной зависимости от начальных данных.

Лемма 2. *Пусть* $C = \{y \in B\colon$ *решение задачи* (1) *не единственно*$\}$. *Если дополнение множества* C, $C^c = B \setminus C$ *— открытое множество в* B, *то на множестве* C^c *решение задачи (1),* $\widehat{x} = \widehat{x}(y)$, *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 $\widehat{y} \in B_\delta$ (т. е. при $\|y - \widehat{y}\| < \delta$) мы получим: решение задачи (1), $\widehat{x}(y)$ удовлетворяет: $\|\widehat{x}(y) - \widehat{x}(\widehat{y})\| < \varepsilon$ для некоторого $\varepsilon > 0$. Следовательно, для любого $y \in C^c$ мы получим, что для любого $\varepsilon > 0$ существует $\delta > 0$, такое, что для любого $\widehat{y}$: $\|\widehat{y} - y\| < \delta$ мы имеем $\|\widehat{x}(y) - \widehat{x}(\widehat{y})\| < \varepsilon$. Заключаем, что решение задачи (1) — непрерывная функция от второй компоненты функции $F(\cdot,\cdot)$ на множестве C^c.

Отметим, открытость множества C^c выполняется для функций, используемых в постановке задачи регрессии (1), на практике. Как указано выше, для практических задач, рассмотренных в работах [1–6], $L(C) = 0$, где $L(C)$ — мера Лебега множества C. С вероятностной точки зрения Лемма 2 показывает, что (при некоторых условиях на $F(\cdot,)$ решение задачи регрессии непрерывно зависит от начальных данных с вероятностью (1).

Таким образом, задачи вида (1) для функций $F(\cdot,\cdot)$, рассмотренных в работах [1–6], являются корректными [9] с практической точки зрения. При этом строго мы можем 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 $F(\cdot,\cdot)$, рассмотренных в работах [1–6].

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

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

В данной части мы исследуем следующий алгоритм поиска приближенного решения задачи (2):

1. На множестве A выбирается некоторая вероятностная мера P, такая, что для любого множества $C \subseteq A$ с положительной мерой Лебега ($L(C) > 0$) выполняется следующее условие: $P(C) > 0$.

2. Выбрасывается 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.

3. Для каждого $n = \overline{1, N}$ используя алгоритм Левенберга–Марквардта [10] с начальной точкой x_n находим точку локального минимума $\widehat{x}_n$.

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

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

Лемма 3. *Алгоритм поиска глобального минимума функции $F(\cdot)$, описанный выше, сходится к решению задачи* (2) *с точностью δ, т. е* $\lim\limits_{N\to\infty} P[F(y_N) - \min\limits_{x\in A} F(x) > \delta] = 0$.

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, мы найдем такое число $\rho(z) > 0$, что $B_{\rho(z)}(z) = \{y \in A\colon \|y - z\| \leqslant \rho(z)\}$, шар радиуса $\rho(z)$ с центром в точке z, пересеченный с A, содержится в A, и для любого $z_1 \in B_{\rho(z)}(z)$ мы имеем: $F(z_1) \geqslant F(z)$. Из условия (1) в описании алгоритма следует, что при выбросе случайной точки x выполняется следующее условие: $P[x \in B_{\rho(z)}(z)] > 0$. Следовательно, используя независимость случайных величин x_n, $n = \overline{1, N}$, мы получим, что при выбросе N точек (и соответственных N запусках алгоритма Левенберга–Макварда) приближенное решение задачи (2) y_N удовлетворяет следующему условию: $P[F(y_N) - F(z) \leqslant \delta] \geqslant 1 - (P[x \in B_{\rho(z)}(z)])^N$. Так как $P[x \in B_{\rho(z)}(z)] > 0$, мы заключаем, что $\lim\limits_{N\to\infty} P[F(y_N) - F(z) \leqslant \delta] = 1$.

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 работах [1–6], описанный выше 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).

1. *Мостовой В. С., Мостовой С. В.* Математическое моделирование оценки старения природных и техногенных объектов в системах мониторинга // Доп. НАН України. – 2011. – № 7. – С. 114–118.
2. *Мостовой В. С., Мостовой С. В.* Оптимальные оценки нелинейных параметров в моделях сейсмоакустического мониторинга // Там само. – 2011. – № 8. – С. 103–107.
3. *Мостовой В. С, Мостовой С. В., Кондра С. М., Страшко Ж. С.* Оценка информативных параметров состояния строительных конструкций в режиме мониторинга / Пром. стр-во и инж. сооружения. – 2011. – № 1. – С. 24–29.
4. *Мостовой В. С.* Оптимальное обнаружение сигналов на фоне микросейсмического шума // Доп. НАН України. – 2008. – № 1. – С. 106–110.
5. *Мостовой В. С.* Математическая модель накопления сейсмических сигналов при активном мониторинге // Там само. – 2008. – № 4. – С. 132–136.
6. *Мостовой В. С., Мостовой С. В.* Вариационный подход к решению обратной задачи при накоплении сейсмических сигналов в активном мониторинге // Там само. – 2008. – № 8. – С. 113–116.
7. *Математическая* энциклопедия. Т. 4. – Москва: Сов. энцикл., 1977. – С. 742.
8. *Bethea R. M., Duran B. S., Boullion T. L.* Statistical methods for engineers and scientists. – New York: Marcel Dekker, 1985. – 105 p.

9. *Evans L. C.* Partial differential equations: methods and applications. – Providence, RI: Amer. Math. Soc., 1998. – 668 p.
10. *Marquardt D.* An algorithm for least-squares estimation of nonlinear parameters // SIAM J. Appl. Math. – 1963. – **11**. – P. 431–441.

Институт геофизики им. С. И. Субботина НАН Украины, Киев
Поступило в редакцию 09.12.2011

В. С. Мостовий

Про коректність задач нелінійної регресії і збіжності алгоритму пошуку глобального мінімуму в моделях моніторингу

Розглянуто згоду нелінійних моделей моніторингу із спостереженими даними нелінійних моделей моніторингу. Ці моделі грунтуються на суперпозиції осциляторів з вільними параметрами. Оптимальну оцінку вільних параметрів моделі, які входять у модель як лінійно, так і нелінійно, розглядатимемо як задачу нелінійної регресії. Оптимальність розуміється в сенсі глобального мінімуму цільового функціонала. Точка в просторі можливих значень вільних параметрів моделі, в якій критерій має глобальний мінімум, приймається як оптимальне рішення. Для вибраних нелінійних математичних моделей треба з'ясувати питання, що пов'язані з існуванням рішення, його єдиністю і стійкістю рішення залежно від початкових даних. Остання обставина особлива важливо, оскільки алгоритми, що побудовані на підставі цих моделей, орієнтовані на безпосередню обробку польових спостережень, а це означає: залежність від характеристик вимірювальної апаратури, помилок виміру і супутньому фону перешкод.

V. S. Mostovyi

About the correctness of a nonlinear problem of regression and convergence of an algorithm of search for a global minimum in models of monitoring

A compliance of observed data and nonlinear models of monitoring is considered. These models are based on a superposition of oscillators with free parameters. Optimal estimation of free parameters of a model, which enter into the model both linearly and nonlinearly, is considered as a problem of nonlinear regression. The optimality is understood in the sense of the global minimum of an objective functional. A point in the space of free parameters of the model, at which the criterion has a global minimum, is accepted as the optimal solution of the problem. For the chosen nonlinear mathematical models, it is necessary to find out the questions connected with the existence of a solution and its uniqueness and stability depending on initial data. Last circumstance is especially important, as the algorithms constructed on the basis of these models are oriented on the direct processing of field data. This means the dependence on characteristics of a measuring equipment, errors of measurement, and accompanying background noises.