ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 25 МАТЕМАТИКИ Выпуск 1
2018

А. Н. Тырсин, **А. А. Азарян** (Екатеринбург, УрФУ). **Оценивание нелинейных регрессионных зависимостей на основе обобщенного метода наименьших модулей.**

Рассмотрим задачу оценивания параметров некоторых нелинейных регрессионных зависимостей с помощью обобщенного метода наименьших модулей [1]

$$(\widehat{a}_1, \ldots, \widehat{a}_m) = \arg\min_{\mathbf{a}\in\mathbf{R}^m} Q(\mathbf{a}), \tag{1}$$

где $Q(\mathbf{a}) = \sum_{i=1}^n \rho(|y_i - f(\mathbf{x}_i, \mathbf{a})|)$; $\mathbf{x}_i = (x_{i1}, \ldots, x_{im})$ — i-я строка регрессионной матрицы $\mathbf{X}$; $y = (y_1, , y_n)$ — значения зависимой переменной; $\mathbf{a} = (a_1, \ldots, a_m)$ — искомый вектор параметров; $\rho(\cdot)$ — некоторая монотонно возрастающая, всюду дважды непрерывно-дифференцируемая на положительной полуоси функция, $\rho(0) = 0$, $\rho''(x) < 0$; $f(\mathbf{x}, \mathbf{a})$ — определенная на $\mathbf{R}^m$ действительная функция, дважды непрерывно-дифференцируемая по параметрам a_j на области определения вектора $\mathbf{x}$.

Утверждение 1. *Пусть функция $f(\mathbf{x}, \mathbf{a})$ является линейной по параметрам. Тогда все локальные минимумы задачи (1) могут находиться только в узловых точках, которые являются решениями систем уравнений*

$$\begin{cases} f(\mathbf{x}_{i_1}, \mathbf{a}) = y_{i_1}, \\ \quad\cdots \\ f(\mathbf{x}_{i_m}, \mathbf{a}) = y_{i_m}, \end{cases} \tag{2}$$

где $\{i_1, \ldots, i_m\}$ — набор индексов, заданный на множестве $\{1, \ldots, n\}$, такой что $\forall\, k \neq j \;\; i_k \neq i_j$.

Доказательство. С учетом линейности $f(\mathbf{x}, \mathbf{a})$ по параметрам

$$f(\mathbf{x}_i, \mathbf{a}) = \sum_{j=1}^m a_j f_j(\mathbf{x}_i) = \sum_{j=1}^m a_j z_{ij} = \langle \mathbf{a}, \mathbf{z}_i \rangle,$$

т. е. она приняла линейный вид. Отсюда $Q(\mathbf{a}) = \sum_{i=1}^n \rho\left(|y_i - \langle \mathbf{a}, \mathbf{z}_i \rangle|\right)$, следовательно, как показано в [2], все локальные минимумы $Q(\mathbf{a})$ будут находиться в узловых точках пересечения гиперплоскостей $\Omega_i = \Omega(\mathbf{a}, \mathbf{z}_i, y_i)$, задаваемых как

$$y_i - \langle \mathbf{a}, \mathbf{z}_i \rangle = 0.$$

Эти узловые точки находим как

$$\mathbf{u}_{(k_1,\ldots,k_m)} = \bigcap_{i=k_1}^{k_m} \Omega_i, 1 \leqslant k_1 < k_2 < \cdots < k_m \leqslant n, \quad k_s \in N. \tag{3}$$

Очевидно, что множество узловых точек (3) эквивалентно множеству решений систем (2).

Далее добавим к функции $f(\mathbf{x}, \mathbf{a})$ свободный член a_0.

Утверждение 2. *Пусть имеем задачу*

$$(\widehat{a}_0, \ldots, \widehat{a}_m) = \arg \min_{\mathbf{a} \in \mathbf{R}^{m+1}} \sum_{i=1}^{n} \rho \left(|y_i - f(\mathbf{x}_i; a_1, \ldots, a_m) - a_0| \right), \tag{4}$$

где $f(\mathbf{x}, \mathbf{a})$ *имеет те же свойства, что и в утверждении* 1.

Тогда все локальные минимумы задачи (4) могут находиться только в узловых точках, которые являются решениями систем уравнений

$$\begin{cases} a_0 + f(\mathbf{x}_{i_0}; a_1, \ldots, a_m) = y_{i_0}, \\ \qquad\qquad \ldots \\ a_0 + f(\mathbf{x}_{i_m}; a_1, \ldots, a_m) = y_{i_m}, \end{cases}$$

где $\{i_0, \ldots, i_m\}$ — набор индексов, заданный на множестве $\{1, \ldots, n\}$, такой что $\forall\, k \neq j \;\; i_k \neq i_j$.

Д о к а з а т е л ь с т в о. Обозначим

$$Q(\mathbf{a}) = \sum_{i=1}^{n} \rho \left(|y_i - f(\mathbf{x}_i;\, a_1,\, \ldots,\, a_m) - a_0| \right). \tag{5}$$

Рассмотрим произвольную стационарную точку $\mathbf{a}^\circ = (a_0^\circ,\, \ldots,\, a_m^\circ)$, для которой $\nabla Q(\mathbf{a}^\circ) = 0$. Предположим, что $\mathbf{a}^\circ$ является точкой локального минимума. Поскольку эта точка не является особой, то функция (5) в ней дважды дифференцируема. Тогда матрица Гессе $\mathbf{H}(\mathbf{a}^\circ)$ должна быть неотрицательно определенной, т. е. все ее угловые миноры не отрицательны. Определим знак первого углового минора Δ_1. В силу строгой вогнутости функции ρ для всех неособых точек

$$\Delta_1 = \frac{\partial^2 Q(\mathbf{a}^\circ)}{\partial \left(a_0^\circ\right)^2} = \sum_{i=1}^{n} \rho'' \left(|y_i - f(\mathbf{x}_i;\, a_1^\circ,\, \ldots,\, a_m^\circ) - a_0^\circ| \right) < 0.$$

Получено противоречие, следовательно, все стационарные точки функции $Q(a)$ не являются локальными минимумами. Поэтому, локальные минимумы, если они существуют, могут находиться только в особых точках. Применив рассуждения, аналогичные как и при доказательстве теоремы 1 в [2], получаем, что локальные минимумы не могут находиться в неузловых особых точках.

Работа выполнена при финансовой поддержке гранта РФФИ № 16-06-00048а.

СПИСОК ЛИТЕРАТУРЫ

1. *Тырсин А. Н.* Робастное построение регрессионных зависимостей на основе обобщенного метода наименьших модулей. — Записки научных семинаров ПОМИ, 2005, т. 328, с. 236–250.
2. *Тырсин А. Н., Соколов Л. А.* Оценивание линейной регрессии на основе обобщенного метода наименьших модулей. — Вестник Самарского государственного технического университета. Сер. Физ.-матем. науки, 2010, № 5(21), с. 134–142.