

МЕТОДЫ УСТОЙЧИВОГО ПОСТРОЕНИЯ ЛИНЕЙНЫХ МОДЕЛЕЙ НА ОСНОВЕ СПУСКА ПО УЗЛОВЫМ ПРЯМЫМ¹

А. Н. Тырсин, А. А. Азарян

RESISTANT LINEAR MODEL FITTING METHODS BASED ON THE DESCENT THROUGH THE NODAL STRAIGHT LINES

A. N. Tyrsin, A. A. Azarian

Аннотация. *Актуальность и цели.* При построении линейных моделей во многих случаях приходится сталкиваться со стохастической неоднородностью экспериментальных данных. Это проявляется в нарушении условий теоремы Гаусса–Маркова, в частности, наблюдения могут быть засорены грубыми ошибками. В этих условиях оценивание параметров моделей требуется выполнять с помощью устойчивых методов. К их числу относят метод наименьших модулей и обобщенный метод наименьших модулей. Однако известные алгоритмы их реализации являются достаточно эффективными лишь для малых размерностей моделей и ограниченного объема выборок. Цель данного исследования – разработка эффективных вычислительных алгоритмов реализации данных методов, не имеющих ограничений на порядок моделей и объем экспериментальных данных. *Материалы и методы.* Реализация задач была достигнута за счет организации спуска к искомому решению по узловым прямым. Для снижения вычислительных затрат использована особенность узловых прямых – все расположенные на каждой такой прямой узловые точки являются пересечением набора гиперплоскостей, из которых отличается только одна гиперплоскость. *Результаты.* Дан обзор известных методов реализации метода наименьших модулей и обобщенного метода наименьших модулей. Описаны алгоритмы спуска по узловым прямым, позволяющие существенно снизить вычислительные затраты при реализации этих методов. Доказано достижение минимума за конечное число шагов при реализации метода наименьших модулей с помощью спуска по узловым прямым. Реализация обобщенного метода наименьших модулей с помощью спуска по узловым прямым позволяет найти глобальный минимум или близкое к нему решение. *Выводы.* Предложены эффективные алгоритмы реализации метода наименьших модулей и обобщенного метода наименьших модулей при оценивании параметров линейных моделей, основанные на спуске по узловым прямым. Вычислительная сложность этих алгоритмов делает возможной их практическую реализацию для анализа экспериментальных данных и построения многомерных линейных моделей.

Ключевые слова: метод наименьших модулей, обобщенный метод, линейная модель, регрессия, алгоритм, узловая точка, узловая прямая, гиперплоскость.

Abstract. *Background.* When fitting linear models, in many cases one has to deal with stochastic inhomogeneity of the experimental data. This is manifested in violation of the assumptions of the Gauss-Markov theorem, in particular, observations can contain outliers. Under the circumstances the estimation of the parameters of models is required to perform using resistant methods. Among them are the least absolute deviations method and the generalized method of the least absolute deviations. However, the known algorithms for

¹ Работа выполнена при финансовой поддержке гранта РФФИ, проект № 16-06-00048а.

their implementation are sufficiently effective only for small dimensions of models and a limited sample size. The purpose of this study is the development of effective computational algorithms for the implementation of these methods, which have no limitations on the order of the models and the amount of experimental data. *Materials and methods.* The implementation of the tasks was achieved by the descent to the solution through the nodal straight lines. To reduce computational costs, the feature of nodal straight lines is used – all nodes located on each such straight line are intersections of a set of hyperplanes, of which only one hyperplane is different. *Results.* A review of known methods for realizing the least absolute deviations method and the generalized method of the least absolute deviations is given. The algorithms of descent through the nodal straight lines are described, which allow to significantly reduce the computational costs when implementing these methods. The achievement of a minimum in a finite number of steps when implementing the method of the least absolute deviations based on descent through the nodal straight lines is proved. The implementation of the generalized method of the least absolute deviations by descent through the nodal straight lines allows us to find a global minimum or a solution close to it. *Conclusions.* Effective algorithms for realization of the least absolute deviations method and the generalized method of the least absolute deviations when estimating the parameters of linear models based on descent through the nodal straight lines are proposed. The computational complexity of these algorithms makes possible their practical implementation for the analysis of experimental data and the construction of multidimensional linear models.

Key words: least absolute deviations method, generalized method, linear model, regression, algorithm, nodal point, nodal straight line, hyperplane.

Введение

При исследовании стохастических систем часто используют линейные модели. К этому классу моделей относят линейные регрессионные зависимости, динамические модели авторегрессии и распределенного лага, а также линейные стохастические модели временных рядов (модели авторегрессии, авторегрессии – скользящего среднего).

Модель линейной регрессии имеет вид [1]

$$\mathbf{y} = \mathbf{X}\mathbf{a} + \boldsymbol{\varepsilon}, \quad (1)$$

где $\mathbf{X} = \{x_{ij}\}_{n \times m} = \begin{pmatrix} 1 & x_{12} & \dots & x_{1m} \\ 1 & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n2} & \dots & x_{nm} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \dots \\ \mathbf{x}_n \end{pmatrix}$ – заданная матрица;

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{pmatrix} \in \mathbf{R}^m, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \in \mathbf{R}^n, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix} \in \mathbf{R}^n$$
 – векторы коэффициентов регрес-

сии, измерений и случайных ошибок, соответственно.

Модель авторегрессии и распределенного лага, обозначаемая как $ADL(p, q_1, \dots, q_m)$, записывается в виде [2]

$$y_i = a_0 + \sum_{k=1}^p a_k y_{i-k} + \sum_{l=1}^m \sum_{j=1}^{q_l} b_{lj} x_{l,i-j} + \varepsilon_i. \quad (2)$$

Если в (2) $p = 0$, то получим модель распределенного лага $DL(q_1, \dots, q_m)$

$$y_i = a_0 + \sum_{l=1}^m \sum_{j=1}^{q_l} b_{lj} x_{l,i-j} + \varepsilon_i. \quad (3)$$

При $m = 0$, $a_0 = 0$ имеем модель авторегрессии $AR(p)$

$$y_i = \sum_{k=1}^p a_k y_{i-k} + \varepsilon_i, \quad (4)$$

которая при введении линейной комбинации лагов случайной компоненты будет моделью авторегрессии–скользящего среднего $ARMA(p, q)$:

$$y_i = \sum_{k=1}^p a_k y_{i-k} + \varepsilon_i + \sum_{j=1}^q b_j \varepsilon_{i-j}. \quad (5)$$

Построение конкретной математической модели по имеющимся наблюдениям реализуется с помощью статистических методов оценки ее параметров. Многие задачи, связанные с обработкой статистических данных, решаются в предположении существования достаточной информации об изучаемых объектах, процессах, явлениях и о свойствах, действующих на них возмущений. Для широкого класса задач разработаны методы эффективного оценивания неизвестных параметров с использованием классических методов максимального правдоподобия [3, 4]. В частности, в предположении, что случайные ошибки нормально распределены, методом максимального правдоподобия является метод наименьших квадратов (МНК) [1].

На основе МНК создана целостная система статистической обработки. С учетом простоты реализации он является наиболее распространенным статистическим методом построения линейных моделей. Его использование для оценивания параметров модели (1) требует выполнения ряда предпосылок, называемых условиями Гаусса–Маркова [5]. При их выполнении МНК-оценки параметров модели (1) являются состоятельными и несмещенными. Кроме того, если случайные ошибки имеют нормальный закон распределения, то МНК-оценки модели (1) становятся эффективными. У динамических моделей (2)–(5) условие некоррелированности значений случайной компоненты для различных лагов не выполняется, поэтому МНК-оценки остаются состоятельными и несмещенными.

Использование МНК при нарушении условий Гаусса–Маркова может привести к значительным ошибкам при оценивании параметров, а в случае присутствия в измерениях больших выбросов – даже к несостоятельности оценок [5]. В этом случае устойчивость оценок может быть достигнута за счет использования иных методов, самым популярным из которых является метод наименьших модулей (МНМ) [6–9]. МНМ для задачи (1) имеет вид

$$Q(\mathbf{a}) = \sum_{i=1}^n |y_i - \langle \mathbf{a}, \mathbf{x}_i \rangle| \rightarrow \min_{\mathbf{a} \in \mathbf{R}^m}. \quad (6)$$

Если выбросы в целом имеют симметричный вид, то для оценивания моделей (1), (3), не использующих в качестве входных переменных лагов вы-

ходной величины, МНМ-оценки обеспечивают приемлемые результаты. Однако при одностороннем характере выбросов наблюдается смещение оценок [10]. Для оценивания моделей (2), (4), (5) в условиях выбросов МНМ уже не обеспечивает состоятельности оценок [8]. В качестве альтернативы в [11] предложен обобщенный метод наименьших модулей (ОМНМ). Без потери общности для удобства будем рассматривать далее модель (1), так как остальные линейные модели формально можно свести к выражению (1).

ОМНМ-оценки для задачи (1) находят как решение задачи

$$W(\mathbf{a}) = \sum_{i=1}^n \rho_i(|y_i - \mathbf{x}_i \mathbf{a}|) \rightarrow \min_{\mathbf{a} \in \mathbb{R}^m}, \quad (7)$$

где $\rho(\cdot)$ – некоторая монотонно возрастающая, дважды непрерывно-дифференцируемая на положительной полуоси функция, причем $\rho(0) = 0$, $\forall t > 0 \quad 0 < \rho'(t) < \infty$, $-\infty < \rho''(t) < 0$.

МНМ и ОМНМ объединяет общее свойство – в обоих случаях решение находится в узловых точках [9, 12].

Введем гиперплоскости $\Omega_i = \Omega(\mathbf{a}, \mathbf{x}_i, y_i)$ в виде уравнений

$$y_i - \langle \mathbf{a}, \mathbf{x}_i \rangle = 0, \quad (i = 1, 2, \dots, n). \quad (8)$$

Зададим также узловые точки пересечения гиперплоскостей (8):

$$\mathbf{a}_{(k_1, \dots, k_m)} = \bigcap_{i=k_1}^{k_m} \Omega_i, \quad 1 \leq k_1 < k_2 < \dots < k_m \leq n, \quad k_s \in N. \quad (9)$$

Обозначим как $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_n\}$ – множество всех гиперплоскостей, U – множество всех узловых точек (9). Поскольку минимум целевых функций $Q(\mathbf{a})$ и $W(\mathbf{a})$ всегда принадлежит множеству U , поэтому точное решение $\mathbf{a}^*$ задач (6) и (7) может быть получено путем перебора всех узловых точек и выбора в качестве решения той, которая обеспечивает минимум целевой функции.

Перебор всех узловых точек требует решения C_n^m систем линейных уравнений порядка m и вычисления значения целевой функции в каждой узловой точке $\mathbf{a}_{(k_1, \dots, k_m)}$. Это означает, что с ростом n и m наблюдается экспоненциальный рост вычислительных затрат. Фактически практическое применение переборного алгоритма ограничено объемом выборки $n < 150$ и числом коэффициентов регрессии $m < 4$.

В настоящее время предложен ряд других методов решения задач (6) и (7). Однако они также для получения точного или близкого к точному решения требуют больших вычислительных затрат.

Постановка задачи

Целью статьи является исследование особенностей задач (6) и (7) для выяснения причин недостаточно высокой скорости сходимости приближенных методов их решения и проблем, связанных с получением точных реше-

ний, а также описание эффективных вычислительных алгоритмов нахождения МНМ- и ОМНМ-оценок линейных моделей.

Обзор методов нахождения МНМ- и ОМНМ-оценок линейных моделей

Эффективный алгоритм поиска минимума целевой функции должен учитывать ее свойства и особенности геометрии. Исследования показали, что у функции $Q(\mathbf{a})$, несмотря на кусочно-линейный вид, очень «плохая» геометрия – она имеет большое число «оврагов» (равно C_n^{m-1}) в виде прямых, причем эти овраги вытянуты, скрещены и многие из них почти параллельны и находятся на близком расстоянии друг от друга, образуя так называемый «пучок». При этом минимум находится внутри этого пучка. Стенки этих оврагов представляют собой выпуклые линейные гипергрani, которые постепенно при приближении к минимуму целевой функции также становятся почти параллельными. В качестве иллюстрации для случая для $m = 2$, $n = 128$ и нормально распределенных случайных ошибок на рис. 1 показаны прямые $y_i - a_1 - a_2 x_i = 0$, $i = 1, 2, \dots, n$, на которых расположены все овраги функции $Q(\mathbf{a})$.

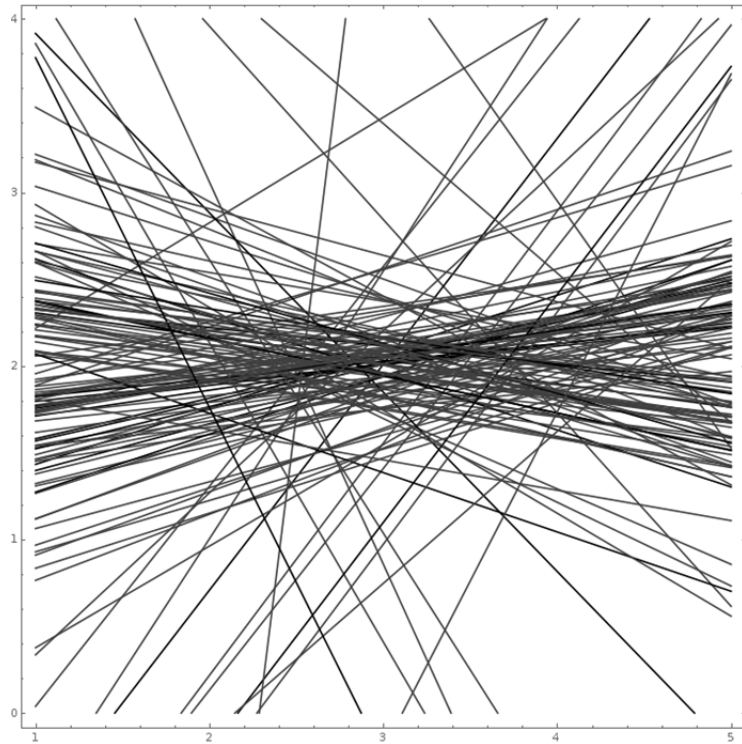


Рис. 1. Графики прямых, на которых расположены все овраги целевой функции $Q(\mathbf{a})$, для случая $m = 2$ и $n = 128$ и нормально распределенных случайных ошибок

Указанная специфика геометрии целевой функции приводит ко многим затруднениям при реализации разных алгоритмов нахождения точного решения задачи (6) на основе идей линейного программирования [7, 13–17]. Все эти алгоритмы основаны на симплекс-методе либо на различных его модифи-

кациях. Они осуществляют движение по узловым точкам, являющимся базисными решениями.

Но решение задачи линейного программирования ограничено случаем лишь нескольких десятков наблюдений. Это вызвано накоплением погрешностей из-за ошибок округлений. В результате этого процесс движения по базисным решениям может не дать точного решения задачи (6) или даже заикаться. Кроме того, требуется хранить в оперативной памяти большой массив данных в виде симплекс-таблиц. Использование техники безошибочных вычислений [18] при устранении роста вычислительных погрешностей приведет к существенному усложнению реализации и росту вычислительных затрат.

При переходе от одной к другой узловой точке указанные алгоритмы не учитывают эту «плохую» геометрию целевой функции.

Объясним, почему переходы по симплекс-таблицам не эффективны. С учетом специфики самой задачи, а именно того, что овраги целевой функции вытянуты, скрещены и многие из них почти параллельны и находятся на близком расстоянии друг от друга, такие переходы без учета геометрии целевой функции часто могут отдаляться от пучка этих гиперплоскостей вида (8), внутри которого лежит точное решение. После этого появляется потребность к возвращению к этому пучку гиперплоскостей, а конкретнее, к точному решению, и для этого осуществляются какие-то спиралевидные неэффективные переходы от одной узловой точки к другой. То есть процесс затягивается, кроме того, при этих спиралевидных переходах также накапливаются вычислительные погрешности. Все эти проблемы существенно усиливаются с увеличением размерности модели и объема выборки.

Известные алгоритмы недостаточно используют свойство ее выпуклости. А именно, ввиду выпуклости целевой функции спуск к точному решению можно осуществлять по различным направлениям. Поэтому нужно выбирать такие направления спуска, которые не позволяли бы отдаляться от пучка гиперплоскостей.

Из приближенных численных методов решения задачи (6) можно отметить метод вариационно-взвешенных квадратических приближений (называемый также алгоритмом Вейсфелда) [7, 19] и методы спуска нулевого порядка [9]. Метод вариационно-взвешенных квадратических приближений не обладает свойством сходимости к точному решению из-за неограниченного возрастания нескольких весовых коэффициентов, а предложенная регуляризация существенно ограничивает точность приближенного решения. Кроме того, с ростом m и n резко снижается скорость сходимости алгоритма в окрестности минимума целевой функции. Методы нулевого порядка просты и эффективны с вычислительной точки зрения. Однако они оказались чувствительны к «плохой» геометрии целевой функции, особенно при увеличении размерности модели m . В процессе перехода от одного приближенного решения к другому происходит перескок на другой овраг, который может проходить в стороне от точного решения, что с учетом близости к нулю градиента целевой функции в точках дифференцируемости часто приводит к несостоятельности оценок свободного члена, а значит, и вектора $\mathbf{a}$ в целом. Чтобы избежать этого, необходимо значительно повысить точность алгоритма, а это приведет к росту вычислительных затрат.

Анализ известных решений показал, что их основной недостаток состоит в неучете овражности и близости к нулю градиента целевой функции в точках дифференцируемости в достаточно большой окрестности ее минимума. Поскольку целевая функция выпуклая, то спуск у овражной функции наиболее эффективен по ее оврагам. Но этими оврагами являются отрезки прямых (назовем их узловыми прямыми), образованных пересечением гиперплоскостей (8). Таким образом, спускаясь по узловым прямым, мы устраняем все указанные выше проблемы известных методов решения задачи (6).

Целевая функция $W(\mathbf{a})$ также имеет овраги, расположенные вдоль узловых прямых. Кроме того, она является вогнутой (выпуклой вверх) и имеет множество локальных минимумов, расположенных в узловых точках, что существенно затрудняет нахождение точного решения задачи (7). Поэтому итерационный алгоритм [20, 21], основанный на идеях линейного программирования, весьма трудоемок и не гарантирует сходимости к точному решению задачи. В [22] для уменьшения вычислительных затрат предложена модификация ОМНМ, основанная на разбиении исходной выборки на подвыборки и формировании множества оценок по всем подвыборкам. В данном случае повышение точности оценивания достигается за счет осреднения полученных оценок и повторного случайного формирования подвыборок. Это приводит к росту вычислительных затрат. Данный алгоритм также не гарантирует получения точного решения задачи (7).

У целевой функции $W(\mathbf{a})$ глобальный минимум также находится в узловой точке, и она имеет вытянутые овраги, расположенные на узловых прямых. Поэтому поиск глобального минимума функции $W(\mathbf{a})$ также следует осуществлять по узловым прямым.

Реализация МНМ на основе спуска по узловым прямым

Обычный спуск. Данный алгоритм точного решения задачи (6) основан на спуске к точному решению, двигаясь вдоль узловых прямых $l_{(k_1, \dots, k_{m-1})}$, каждая из которых является пересечением $(m-1)$ различных гиперплоскостей Ω_i :

$$l_{(k_1, \dots, k_{m-1})} : \bigcap_{i=k_1}^{k_{m-1}} \Omega_i, \quad 1 \leq k_1 < k_2 < \dots < k_{m-1} \leq n, \quad k_s \in N.$$

В качестве начального приближения берется узловая точка $\mathbf{a}_{(k_1, \dots, k_m)}$, являющаяся пересечением m произвольных различных гиперплоскостей $\Omega_{k_1}, \dots, \Omega_{k_m}$. Исключив одну из гиперплоскостей, получим узловую прямую $l_{(k_1, \dots, k_{m-1})}$. Через любую узловую точку проходит m узловых прямых. Выберем из этих узловых прямых ту, вдоль которой целевая функция достигает наименьшего значения, которое всегда будет достигаться в одной из узловых точек. Найдя эту узловую точку, продолжим движение из нее по тому же принципу. В результате будет найдена узловая точка $\mathbf{a}_{(k_1, \dots, k_m)}^* = \mathbf{a}^*$, спуск из которой невозможен. Эта узловая точка будет являться точным решением задачи (6). В основе алгоритма лежат следующие теоремы.

Теорема 1. Рассмотрим модель (1), для которой имеется выборка наблюдений $(\mathbf{x}_i, y_i) = (x_{i1}, x_{i2}, \dots, x_{im}, y_i)$, $(i = 1, \dots, n)$, и пусть заданы функция $Q(\mathbf{a})$ задачи (6), гиперплоскости (3) и множество U всех узловых точек (9). Тогда функция (2) всегда имеет точку глобального минимума, эта точка либо единственна и принадлежит U , либо состоит из выпуклого линейного многогранника, вершины которого являются точками из U .

Доказательство теоремы 1. Известно, что выпуклая непрерывная кусочно-линейная функция либо имеет глобальный минимум, либо стремится к минус бесконечности. А поскольку функция $Q(\mathbf{a})$ является также ограниченной снизу ($Q(\mathbf{a}) \geq 0$, как сумма модулей) функцией, то она всегда имеет точку глобального минимума.

Пусть $\mathbf{a}^* = (a_1^*, a_2^*, \dots, a_m^*)^T$ – стационарная точка и является точкой минимума функции $Q(\mathbf{a})$. Тогда ее градиент в этой точке равен нулю. Поскольку $Q(\mathbf{a})$ – кусочно-линейная функция, то из равенства $\text{grad} Q(\mathbf{a}^*) = \mathbf{0}$ следует, что функция $Q(\mathbf{a})$ является постоянной функцией на выпуклом многограннике с вершинами $\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^l$, гранями которого являются гиперплоскости (8), имеющие вид [23]

$$A = \{\mathbf{a} : \mathbf{a} = \sum_{k=1}^l \lambda_k^2 \mathbf{a}^k, \sum_{k=1}^l \lambda_k^2 = 1\}.$$

Функция $Q(\mathbf{a})$ является постоянной до граничных точек многогранника A , лежащих на гиперплоскостях (3). Следовательно, вершины $\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^l$ этого многогранника являются узловыми точками и функция $Q(\mathbf{a})$ достигает минимума и в этих точках.

Все гиперплоскости (9), и только они, являются особыми точками функции $Q(\mathbf{a})$, поскольку только в них она не дифференцируема. Если взять $(m-1)$ произвольных невырожденных соотношений вида $l_i = y_i - \langle \mathbf{a}, \mathbf{x}_i \rangle = 0$, то они в совокупности определяют прямую

$$l_{(k_1, \dots, k_{m-1})} : \bigcap_{i=k_1}^{k_{m-1}} \Omega_i, \quad k_i \in \{1, 2, \dots, n\} \quad (10)$$

в пространстве $(a_1, a_2, \dots, a_m)$ и вместе с тем плоскость $P_{(Q, k_1, \dots, k_{m-1})}$, параллельную оси Q в пространстве $(Q, a_1, a_2, \dots, a_m)$. Присоединяя к системе (10) выражение (6) и рассматривая их совместно, найдем уравнение ломаной M , полученной в результате пересечения поверхности (6) плоскостью $P_{(Q, k_1, \dots, k_{m-1})}$.

Если с помощью уравнений, входящих в систему (10), выразить $(m-1)$ неизвестных $a_1, a_2, \dots, a_{m-1}$ через оставшееся неизвестное и подставить в выражение (6), то получим уравнение проекции ломаной M на плоскость (Q, a_m) (рис. 2). Точки $T_{i_1}, T_{i_2}, \dots, T_{i_{n-m+1}}$ являются проекциями на эту плоскость точек пересечения прямой $l_{(k_1, \dots, k_{m-1})}$ гиперплоскостями (8), не вошедшими в (10).

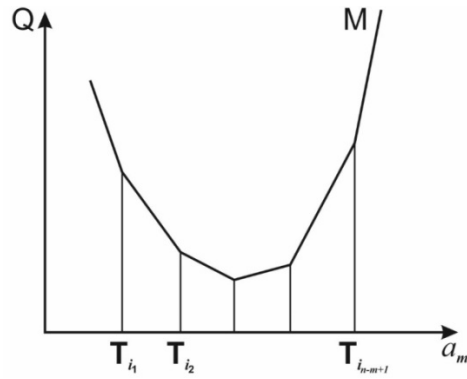


Рис. 2. Вид функции $Q(a_m)$

Функция $Q(a_m)$ на плоскости (Q, a_m) – выпуклая и кусочно-линейная с особыми точками $T_{i_1}, T_{i_2}, \dots, T_{i_{n-m+1}}$, которые являются узловыми. Отсюда следует, что функция $Q(\mathbf{a})$ достигает минимума в узловой точке.

Теорема 2. Алгоритм спуска вдоль узловых прямых для нахождения решения задачи (6) сходится к точному решению за конечное число шагов.

Доказательство теоремы 2. На каждом шаге алгоритма мы находим узловую точку с все меньшим значением целевой функции, а поскольку количество узловых точек конечно, алгоритм выполняется за конечное число шагов.

Докажем, что алгоритм всегда будет достигать минимума целевой функции (т.е. решения). Допустим, что мы находимся на k -м шаге алгоритма, в какой-то узловой точке и пытаемся осуществить спуск вдоль одной из m узловых прямых, проходящих через эту точку, к другой узловой точке с меньшим значением целевой функции. Покажем, что либо мы сможем найти такую точку, либо текущая узловая точка является минимальной.

Целевая функция $Q(\mathbf{a})$ является выпуклой функцией в пространстве $\mathbf{R}^m$. Поэтому, если взять произвольную точку $\tilde{\mathbf{x}}$ и m линейно независимых прямых $\tilde{l}_j$ ($j=1, 2, \dots, m$), проходящих через нее, либо найдется такая окрестность этой точки, в которой целевая функция будет убывать хотя бы вдоль одного из направлений, либо в текущей узловой точке достигается минимум целевой функции.

Поскольку вероятность того, что случайная величина (случайные погрешности) более одного раза примет одно и то же значение равна нулю, то узловые прямые, проходящие через узловую точку, являются линейно независимыми. То есть вдоль одной из них целевая функция убывает, причем она будет убывать вплоть до следующей узловой точки на этой узловой прямой, поскольку между двумя соседними узловыми точками целевая функция (2) будет линейна, так как все ее подмодульные выражения не будут менять знак. Действительно, смена знака означает пересечение одной из образующих гиперплоскостей, что приведет к образованию узловой точки, а мы рассматриваем часть узловой прямой между двумя соседними узловыми точками.

На практике из-за конечной точности измерений гипотетически могут возникнуть ситуации, когда узловые прямые, проходящие через узловую точ-

ку, окажутся параллельными. В этом случае узловой точки не существует, и переходим к рассмотрению другой узловой прямой.

Таким образом, либо произвольная узловая точка является минимумом целевой функции, либо хотя бы по одной из проходящих через нее узловых прямых можно сделать переход к узловой точке с меньшим значением целевой функции. Теорема доказана.

Отметим, что вычислительные затраты данного алгоритма можно значительно сократить, если использовать информацию о предыдущей найденной узловой точке.

Спуск с использованием разреженных матриц. Двигаясь вдоль прямой $l_{(k_1, \dots, k_{m-1})}$ для нахождения узловых точек, принадлежащих этой прямой, нужно для каждой точки решать систему линейных алгебраических уравнений (СЛАУ) порядка m :

$$\begin{cases} a_1 + a_2 x_{k_1,2} + a_3 x_{k_1,3} + \dots + a_m x_{k_1,m} = y_{k_1}, \\ a_1 + a_2 x_{k_2,2} + a_3 x_{k_2,3} + \dots + a_m x_{k_2,m} = y_{k_2}, \\ \dots \\ a_1 + a_2 x_{k_{m-1},2} + a_3 x_{k_{m-1},3} + \dots + a_m x_{k_{m-1},m} = y_{k_{m-1}}, \\ a_1 + a_2 x_{i,2} + a_3 x_{i,3} + \dots + a_m x_{i,m} = y_i, \end{cases} \quad (11)$$

где $1 \leq k_1 < k_2 < \dots < k_{m-1} \leq n$, $i \in \{1, 2, \dots, n\}$, $i \notin \{k_1, k_2, \dots, k_{m-1}\}$.

Очевидно, что СЛАУ двух различных узловых точек, принадлежащих этой прямой одной прямой, отличаются лишь одним (последним) уравнением. Следовательно, вычислительная эффективность алгоритма спуска существенно повысится, если для нахождения узловых точек, которые лежат на прямой $l_{(k_1, \dots, k_{m-1})}$, первые $(m-1)$ строк расширенной матрицы, соответствующей СЛАУ (11), предварительно преобразуем с помощью элементарных преобразований к ступенчатому виду. Расширенная матрица СЛУ прямой $l_{(k_1, \dots, k_{m-1})}$ имеет вид

$$\mathbf{A}_{(k_1, \dots, k_{m-1})} = \begin{pmatrix} 1 & x_{k_1,2} & x_{k_1,3} & \dots & x_{k_1,m-1} & x_{k_1,m} & y_{k_1} \\ 1 & x_{k_2,2} & x_{k_2,3} & \dots & x_{k_2,m-1} & x_{k_2,m} & y_{k_2} \\ 1 & x_{k_3,2} & x_{k_3,3} & \dots & x_{k_3,m-1} & x_{k_3,m} & y_{k_3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{k_{m-1},2} & x_{k_{m-1},3} & \dots & x_{k_{m-1},m-1} & x_{k_{m-1},m} & y_{k_{m-1}} \end{pmatrix}.$$

Применив алгоритм прямого хода метода Гаусса, преобразуем матрицу $\mathbf{A}_{(k_1, \dots, k_{m-1})}$ к ступенчатому виду:

$$\mathbf{A}'_{(k_1, \dots, k_{m-1})} = \begin{pmatrix} 1 & x_{k_1,2} & x_{k_1,3} & \dots & x_{k_1,m-1} & x_{k_1,m} & y_{k_1} \\ 0 & 1 & x'_{k_2,3} & \dots & x'_{k_2,m-1} & x'_{k_2,m} & y'_{k_2} \\ 0 & 0 & 1 & \dots & x'_{k_3,m-1} & x'_{k_3,m} & y'_{k_3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & x'_{k_{m-1},m} & y'_{k_{m-1}} \end{pmatrix}.$$

Используя ступенчатую матрицу $\mathbf{A}'_{(k_1, \dots, k_{m-1})}$, можно значительно сократить вычислительные затраты на нахождение всех узловых точек, лежащих на прямой $l_{(k_1, \dots, k_{m-1})}$. Действительно для каждой искомой узловой точки имеем расширенную матрицу

$$\mathbf{A}_{(k_1, \dots, k_{m-1}, i)} = \begin{pmatrix} 1 & x_{k_1, 2} & x_{k_1, 3} & \dots & x_{k_1, m-1} & x_{k_1, m} & y_{k_1} \\ 0 & 1 & x'_{k_2, 3} & \dots & x'_{k_2, m-1} & x'_{k_2, m} & y'_{k_2} \\ 0 & 0 & 1 & \dots & x'_{k_3, m-1} & x'_{k_3, m} & y'_{k_3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & x'_{k_{m-1}, m} & y'_{k_{m-1}} \\ 1 & x_{i, 2} & x_{i, 3} & \dots & x_{i, m-1} & x_{i, m} & y_i \end{pmatrix}. \quad (12)$$

Варьируя номер i в (12), найдем все узловые точки, лежащие на прямой $l_{(k_1, \dots, k_{m-1})}$.

Вычислительную эффективность алгоритма спуска можно повысить, если учитывать направление спуска.

Спуск с использованием разреженных матриц и с учетом направления спуска. Используя ступенчатую матрицу $\mathbf{A}'_{(k_1, \dots, k_{m-1})}$ и решив СЛАУ, соответствующую расширенной матрице (12), находим значение m -го коэффициента $a_m^{(k_1, k_2, \dots, k_{m-1}, i)}$, $1 \leq k_1 < k_2 < \dots < k_{m-1} \leq n$, $i \in \{1, 2, \dots, n\}$, $i \notin \{k_1, k_2, \dots, k_{m-1}\}$ для каждой узловой точки. После этого по возрастанию $a_m^{(k_1, k_2, \dots, k_{m-1}, i)}$ упорядочиваем все узловые точки, которые лежат на прямой $l_{(k_1, \dots, k_{m-1})}$, и выполняем описанный выше алгоритм спуска, но с учетом направления. Если при непосредственном переходе от одной узловой точки к другой значение целевой функции увеличивается, то в этом направлении значение целевой функции будет увеличиваться во всех узловых точках (вытекает из выпуклости целевой функции). Назовем такое направление «плохим». Для осуществления спуска до вычисления значения целевой функции в очередной узловой точке рассматриваем направление спуска. Если оно «плохое», то переходим к следующей точке, не вычисляя в данной узловой точке ни значения целевой функции, ни значения коэффициентов $a_j^{(k_1, k_2, \dots, k_{m-1}, i)}$, $j = 1, 2, \dots, m-1$.

В [24] доказано, что алгоритм спуска по узловым прямым имеет вычислительную сложность $W = O(m^2 n^2 + m^4 n \ln n + m^2 n \ln^2 n)$. Отметим, что для $n > \max(\ln^2 n; m^2 \ln n)$ вычислительная сложность спуска по узловым прямым $W = O(m^2 n^2)$.

Нахождение ОМНМ-оценок линейных моделей на основе спуска по узловым прямым

Для реализации идеи неполного перебора узловых точек необходима априорная информация о примерном расположении точного решения $\mathbf{a}^*$

задачи (7). Многочисленные вычислительные эксперименты методом статистических испытаний Монте-Карло [25] показали, что вектор $\mathbf{a}^*$, как правило, мало отличается от решения задачи (7) методом наименьших модулей, равного

$$\mathbf{a}^{(0)} = \arg \min_{\mathbf{a} \in \mathbb{R}^m} \sum_{i=1}^n |y_i - \langle \mathbf{a}, \mathbf{x}_i \rangle|. \quad (13)$$

Описанный выше алгоритм реализации МНМ на основе спуска по узловым прямым имеет число операций порядка $O(m^2 n^2)$. Это пренебрежимо мало по сравнению с затратами на реализацию полного перебора узловых точек.

Алгоритм решения задачи (7) также основан на использовании узловых прямых и состоит в следующем.

1. Используем в качестве начального приближения $\mathbf{a}^{(1)}$ узловую точку решения $\mathbf{a}^{(0)}$ задачи (13) методом наименьших модулей, т.е. считаем $\mathbf{a}^{(1)} = \mathbf{a}^{(0)}$.

2. Уточняем решение путем спуска из узловой точки $\mathbf{a}^{(1)}$, поочередно двигаясь вдоль каждой из m узловых прямых, проходящих через $\mathbf{a}^{(1)}$. Выберем ту узловую точку, в которой целевая функция $W(\mathbf{a})$ достигает наименьшего значения. Найдя эту точку, продолжим движение из нее по тому же принципу. В результате будет найдена узловая точка $\mathbf{a}^{(2)}$, спуск из которой невозможен. При движении вдоль узловой прямой используем ступенчатые расширенные матрицы вида (12).

3. Для найденной узловой точки $\mathbf{a}^{(2)}$ определим множество гиперплоскостей $\Omega(\mathbf{a}^{(2)}) \subset \Omega$, состоящее из $\alpha \cdot 100\%$ гиперплоскостей, наименее удаленных от $\mathbf{a}^{(2)}$.

Каждая из гиперплоскостей (8) включает в себя случайную ошибку. Ошибки группируются относительно нуля; чем меньше ошибка, тем меньше гиперплоскости отличаются друг от друга. Решение обычно лежит в узловых точках, образованных гиперплоскостями с малыми величинами ошибок. Все эти гиперплоскости будут располагаться в пространстве очень «тесно».

4. С помощью перебора среди всех узловых точек, получаемых в результате пересечения гиперплоскостей из множества $\Omega(\mathbf{a}^{(2)})$, находим ту точку $\mathbf{a}^{(3)}$, в которой целевая функция принимает наименьшее значение.

5. Найденную узловую точку $\mathbf{a}^{(3)}$ сравниваем с начальным приближением $\mathbf{a}^{(1)}$. Если они совпадают, то работа алгоритма останавливается и в качестве решения задачи (2) берется полученная узловая точка. В противном случае в качестве начального приближения $\mathbf{a}^{(1)}$ берем полученную узловую точку, переходим к п. 2 и повторяем вышеописанные действия.

Пример. Для нескольких значений n и m проведем исследование описанного алгоритма. Считаем для определенности, что случайная ошибка распределена по стандартному нормальному закону распределения. Воспользуемся методом Монте-Карло. Число испытаний зададим равным 1000. В качестве функции потерь используем $\rho(|x|) = \arctg|x|$.

В табл. 1 приведено отношение вычислительных затрат переборного алгоритма к затратам предложенного алгоритма спуска по узловым прямым.

Таблица 1

Отношение вычислительных затрат переборного алгоритма к затратам предложенного алгоритма спуска по узловым прямым

n	$m = 3$	$m = 4$	$m = 5$
32	7,2	40,4	202,5
64	46,0	414,4	3961,6
128	158,4	1157,9	6203,2
256	222,1	1273,4	6875,0
512	225,4	1339,3	7363,6

Алгоритм спуска по узловым прямым показал высокую точность. В частности, для $m = 3$, $n = 128$ в 995 случаях из 1000 полученное решение совпало с точным решением по переборному алгоритму. В остальных пяти случаях средняя относительная погрешность по сравнению с точным решением составила 0,11 %.

Заключение

1. Предложены эффективные алгоритмы реализации метода наименьших модулей и обобщенного метода наименьших модулей при оценивании параметров линейных моделей, основанные на спуске по узловым прямым.
2. Данные алгоритмы позволяют значительно снизить вычислительные затраты по сравнению с переборным алгоритмом.
3. Вычислительная сложность спуска по узловым прямым позволяет на практике реализовать предложенные алгоритмы для анализа экспериментальных данных.

Библиографический список

1. Айвазян, С. А. Прикладная статистика: исследование зависимостей / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин. – М. : Финансы и статистика, 1985. – 488 с.
2. Сулов, В. И. Эконометрия / В. И. Сулов, Н. М. Ибрагимов, Л. П. Талышева, А. А. Цыплаков. – Новосибирск : Изд-во СО РАН, 2005. – 744 с.
3. Кендалл, М. Статистические выводы и связи : пер. с англ. / М. Кендалл, А. Стюарт. – М. : Наука : ФИЗМАТЛИТ, 1973. – 900 с.
4. Крамер, Г. Математические методы статистики / Г. Крамер ; пер. с англ. – М. : Мир, 1975. – 648 с.
5. Демиденко, Е. З. Линейная и нелинейная регрессия / Е. З. Демиденко. – М. : Финансы и статистика, 1981. – 302 с.
6. Bloomfield, P. Least absolute deviations: theory, applications, and algorithms / P. Bloomfield, W. L. Steiger. – Boston-Basel-Stuttgart : Birkhauser, 1983. – 349 p.
7. Мудров, В. И. Методы обработки измерений. Квазиравдоподобные оценки / В. И. Мудров, В. Л. Кушко. – М. : Радио и связь, 1983. – 304 с.
8. Болдин, М. В. Знаковый статистический анализ линейных моделей / М. В. Болдин, Г. И. Симонова, Ю. Н. Тюрин. – М. : Наука : ФИЗМАТЛИТ, 1997. – 288 с.

9. Тырсин, А. Н. Оценивание линейных регрессионных уравнений с помощью метода наименьших модулей / А. Н. Тырсин, К. Е. Максимов // Заводская лаборатория. Диагностика материалов. – 2012. – Т. 78, № 7. – С. 65–71.
10. Тырсин, А. Н. Робастное построение линейных регрессионных моделей по экспериментальным данным / А. Н. Тырсин // Заводская лаборатория. Диагностика материалов. – 2005. – Т. 71, № 11. – С. 53–57.
11. Тырсин, А. Н. Робастное построение регрессионных зависимостей на основе обобщенного метода наименьших модулей / А. Н. Тырсин // Записки научных семинаров ПОМИ. – 2005. – Т. 328. – С. 236–250.
12. Тырсин, А. Н. Оценивание линейной регрессии на основе обобщенного метода наименьших модулей / А. Н. Тырсин, Л. А. Соколов // Вестник Самарского государственного технического университета. Серия Физико-математические науки. – 2010. – № 5 (21). – С. 134–142.
13. Fisher, W. D. A note on curve fitting with minimum deviations by linear programming / W. D. Fisher // Journal of the American Statistical Association. – 1961. – Vol. 56. – P. 359–362.
14. Barrodale, I. An improved algorithm for discrete L1 linear approximation / I. Barrodale, F. D. K. Roberts // SIAM Journal on Numerical Analysis – 1973. – Vol. 10. – P. 839–848.
15. Narula, S. C. Algorithm AS108: Multiple linear regression with minimum sum of absolute errors / S. C. Narula, J. F. Wellington // Applied Statistics. – 1977. – Vol. 26. – P. 106–111.
16. Armstrong, R. D. Algorithm AS132: Least absolute value estimates for a simple linear regression problem / R. D. Armstrong, D. S. Kung // Applied Statistics. – 1978. – Vol. 27. – P. 363–366.
17. Wesolowsky, G. O. A new descent algorithm for the least absolute value regression problem / G. O. Wesolowsky // Communications in Statistics, Simulation and Computation. – 1981. – Vol. B10, № 5. – P. 479–491.
18. Грегори, Р. Безошибочные вычисления. Методы и приложения / Р. Грегори, Е. Кришнамурти. – М. : Мир, 1988. – 208 с.
19. Weiszfeld, E. On the point for which the sum of the distances to n given points is minimum / E. Weiszfeld // Annals of Operations Research. – 2008. – Vol. 167. – P. 7–41.
20. Панюков, А. В. Взаимосвязь взвешенного и обобщенного вариантов метода наименьших модулей / А. В. Панюков, А. Н. Тырсин // Известия Челябинского научного центра. – 2007. – № 1 (35). – С. 6–11.
21. Панюков, А. В. Об устойчивом оценивании параметров авторегрессионных моделей на основе обобщенного варианта метода наименьших модулей / А. В. Панюков // Вестник НГУЭУ. – 2015. – № 4. – С. 339–346.
22. Тырсин, А. Н. Об одном алгоритме реализации обобщенного метода наименьших модулей / А. Н. Тырсин, А. А. Азарян // Обзорение прикладной и промышленной математики. – 2017. – Т. 24, В. 4. – С. 375–376.
23. Рокафеллар, Р. Выпуклый анализ : пер. с англ. / Р. Рокафеллар. – М. : Мир, 1973. – 470 с.
24. Тырсин, А. Н. Точные алгоритмы реализации метода наименьших модулей на основе спуска по узловым прямым / А. Н. Тырсин, А. А. Азарян // Вестник Бурятского государственного университета. Математика, информатика. – 2017. – № 4. – С. 21–32.
25. Ермаков, С. М. Метод Монте-Карло и смежные вопросы / С. М. Ермаков. – 2-е изд., доп. – М. : Наука : ФИЗМАТЛИТ, 1975. – 472 с.

Тырсин Александр Николаевич

доктор технических наук, доцент,
заведующий кафедрой прикладной
математики,
Уральский энергетический институт,
Уральский федеральный университет
им. первого Президента России
Б. Н. Ельцина
E-mail: at2001@yandex.ru

Tyrsin Alexander Nikolayevich

doctor of technical sciences,
associate professor,
head of sub-department
of applied mathematics,
Ural power engineering Institute,
Ural Federal University named after
the first President of Russia B. N. Yeltsin

Азарян Алексан Артурович

аспирант,
Уральский энергетический институт,
Уральский федеральный университет
им. первого Президента России
Б. Н. Ельцина
E-mail: a.a.azaryan@gmail.com

Azaryan Alexan Arturovich

postgraduate student,
Ural power engineering Institute,
Ural Federal University named after
the first President of Russia B. N. Yeltsin

УДК 519.237.5:519.24

Тырсин, А. Н.

Методы устойчивого построения линейных моделей на основе спуска по узловым прямым / А. Н. Тырсин, А. А. Азарян // Модели, системы, сети в экономике, технике, природе и обществе. – 2018. – № 1 (25). – С. 188–202.