ТЕОРІЯ ОПТИМАЛЬНИХ РІШЕНЬ

УДК 519.85

П.І. СТЕЦЮК, О.М. ХОМ'ЯК

МІНІМАЛЬНІ ЗА L_p-НОРМОЮ ЛІНІЙНІ СПЛАЙНИ

Вивчаються задачі безумовної мінімізації опуклих функцій для знаходження мінімальних за Lp-нормою лінійних сплайнів для випадків $p \geq 1$ *та* $1 < p \leq 2$. *Якщо* $p \geq 1$, *то використовується негладка функція, а якщо* $1 < p \leq 2$ *– гладка функція. Показано, що при певному виборі параметра* p *оптимізаційні задачі породжують відомі методи – метод найменших квадратів, метод найменших модулів та мінімаксний чебишевський метод. Наведено властивості розв'язків задачі для випадку* $1 < p \leq 2$.

Вступ. За допомогою лінійних сплайнів можна досить точно наблизити довільну неперервну криву. Це означає, що задачу знаходження невідомої кривої, яка найкращим чином апроксимує набір експериментальних кривих, можна замінити задачею пошуку невідомого лінійного сплайна, який найкращим чином апроксимує певний набір заданих лінійних сплайнів. Для оцінки невідомих величин за результатами проведених вимірювань використовують різноманітні методи регресійного аналізу. Серед них метод найменших модулів [1], використання якого доцільно в тих випадках, коли розподіл помилок вимірювань підпорядкований закону Лапласа, та метод найменших квадратів [2], теоретичною основою якого є закон нормального розподілу. Меншого поширення набули методи, в яких мінімізується L_p-норма вектора $v = (v_1, \ldots, v_m)^T$, яка визначена як: $\| v \|_p = (\sum_{i=1}^{m} | v_i |^p)^{1/p}$, де параметр $p \geq 1$.

У статті розглянемо оптимізаційні задачі для знаходження мінімального за L_p-нормою лінійного сплайна. Покажемо, що при певному виборі параметра p оптимізаційні задачі породжують відомі методи обробки експериментальних даних, а саме: методу найменших модулів відповідає задача мінімізації негладкої опуклої функції, а методу найменших квадратів – задача мінімізації квадратичної опуклої функції. Метод найменших модулів є робастним [3 – 5] у тому випадку, якщо в заданому наборі сплайнів є аномальні результати вимірювань.

1. Формулювання задачі. Побудуємо оптимізаційну задачу для пошуку лінійного сплайна, який за L_p-нормою ($p \geq 1$) мінімально відрізняється від m лінійних сплайнів $y^1, \ldots, y^m$, які визначені значеннями $y_1^i, \ldots, y_n^i$, $i = 1, \ldots, m$ в одних і тих же базових точках $x_1 < \ldots < x_n$ інтервалу $[x_1, x_n]$. Невідомими у задачах будуть значення $y = (y_1, \ldots, y_n)^T$ в базових точках $x_1 < \ldots < x_n$, а мінімальний за L_p-нормою лінійний сплайн будемо позначати $y_p^* = (y_{1p}^*, \ldots, y_{np}^*)^T$.

Якщо $p \geq 1$, то знаходженню лінійного сплайна y_p^* буде відповідати задача безумовної мінімізації опуклої негладкої функції: знайти

$$y_p^* = \arg\min_{y \in R^n} \left\{ f_p(y) = \left(\sum_{j=1}^{n} \sum_{i=1}^{m} \left| y_j - y_j^i \right|^p \right)^{1/p} \right\}. \quad (1)$$

Тут умова $p \geq 1$ гарантує опуклість функції $f_p(y_1, \ldots, y_n)$ від n змінних.

Задача (1) має розв'язок $y_p^* = (y_1^*, \ldots, y_n^*)^T$, де $\min_{j=1,\ldots,m} \{y_j^i\} \leq y_j^* \leq \max_{j=1,\ldots,m} \{y_j^i\}$, $j = 1, \ldots, n$, але не обов'язково цей розв'язок є єдиним. Так, наприклад, якщо $p = 1$, $n = 3$ та $m = 2$, то задача (1) є задачею мінімізації функції $f_1(y_1, y_2, y_3) = |y_1^1 - y_1| + |y_1^2 - y_1| + |y_2^1 - y_2| + |y_2^2 - y_2| + |y_3^1 - y_3| + |y_3^2 - y_3|$. Її оптимальним розв'язком є 3-вимірний вектор $y^* = (y_1^*, y_2^*, y_3^*)^T$ з компонентами $y_1^* = \lambda_1 y_1^1 + (1 - \lambda_1) y_1^2$, $y_2^* = \lambda_2 y_2^1 + (1 - \lambda_2) y_2^2$, $y_3^* = \lambda_3 y_3^1 + (1 - \lambda_3) y_3^2$, де $0 \leq \lambda_1 \leq 1$, $0 \leq \lambda_2 \leq 1$, $0 \leq \lambda_3 \leq 1$, а y_1^1, y_1^2, y_2^1, y_2^2, y_3^1 та y_3^2 – довільні числа. При цьому мінімальне значення функції буде рівним $f_1(y^*) = |y_1^2 - y_1^1| + |y_2^2 - y_2^1| + |y_3^2 - y_3^1|$. Цей випадок для $y_1^1 = 3$, $y_1^2 = 5$, $y_2^1 = 3.5$, $y_2^2 = 5.7$, $y_3^1 = 3.8$ та $y_3^2 = 6$ показано на рисунку, де обидва сплайни y^1 та y^2 наведені суцільними лініями.

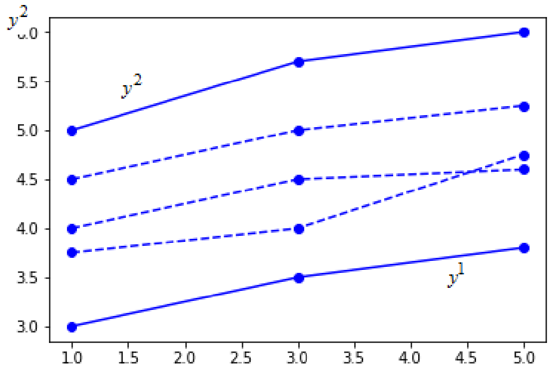

РИСУНОК. Структура оптимальних розв'язків задачі (1) при $p = 1$, $n = 3$ та $m = 2$

Тут оптимальні розв'язки задачі (1) зображено пунктирними лініями. Їм відповідає мінімальне значення функції $f_1^* = f_1(y^*) = 6.4$.

Залежно від значень параметра p із задачі (1) витікають три її часткові випадки, що тісно пов'язані з відомими методами обробки експериментальних даних. Якщо $p=1$, то отримуємо функцію

$$f_1(y) = \left|y_1 - y_1^1\right| + \ldots + \left|y_1 - y_1^m\right| + \ldots + \left|y_n - y_n^1\right| + \ldots + \left|y_n - y_n^m\right|, \tag{2}$$

якій відповідає метод найменших модулів (МНМ), тобто у задачі (1) мінімізуються сумарні абсолютні величини відхилень компонент невідомого сплайна y від компонент заданих сплайнів $y^1,\ldots,y^m$. Якщо $p=2$, то отримаємо функцію

$$f_2(a) = \sqrt{\left(y_1 - y_1^1\right)^2 + \ldots + \left(y_1 - y_1^m\right)^2 + \ldots + \left(y_n - y_n^1\right)^2 + \ldots + \left(y_n - y_n^m\right)^2}, \tag{3}$$

якій з точністю до кореня квадратного відповідає метод найменших квадратів (МНК), тобто у задачі (1) мінімізуються сумарні квадрати відхилень компонент сплайна y від компонент сплайнів $y^1,\ldots,y^m$. Якщо $p=\infty$, то отримаємо таку функцію

$$f_\infty(a) = \max_{j=1,\ldots,n}\left\{\left|y_j - y_j^1\right|,\ldots,\left|y_j - y_j^m\right|\right\}, \tag{4}$$

якій відповідає мінімаксний (чебишевський) метод, тобто у задачі (1) мінімізується максимальна серед абсолютних величин $m \times n$ відхилень компонент невідомого сплайна y від компонент заданих сплайнів $y^1,\ldots,y^m$.

Отже, кожному з наведених значень параметра p відповідає свій метод розв'язання задачі (1) – МНМ ($p=1$), МНК ($p=2$), або мінімаксний (чебишевський) метод ($p=\infty$). Зазначимо, що МНМ та мінімаксний метод можна замінити задачами лінійного програмування, враховуючи, що ці методи використовують кусочно лінійні функції вигляду (2) та вигляду (4) відповідно.

2. Спрощена задача та її властивості. Для випадку $1 \le p \le 2$ задачу (1) можна спростити, опустивши знак степеня $1/p$. В результаті для знаходження мінімального за L_p-нормою лінійного сплайна y_p^* отримаємо таку задачу мінімізації опуклої функції: знайти

$$y_p^* = \arg\min_{y \in R^n}\left\{F_p(y) = \sum_{j=1}^{n}\sum_{i=1}^{m}\left|y_j - y_j^i\right|^p\right\}, \tag{5}$$

де p – скалярний параметр, такий, що $1 \le p \le 2$. Тут $F_p(y)$ – сепарабельна опукла функція, яка є негладкою тільки при $p=1$. При цьому мінімальні значення функції $f_p^*(y_p^*)$ в оптимізаційній задачі (1) та функції $F_p^*(y_p^*)$ у оптимізаційній задачі (5) зв'язані співвідношенням $F_p^*(y_p^*) = \left(f_p^*(y_p^*) \right)^p$.

На відміну від того, що для задачі (1) обов'язково використовувати тільки методи мінімізації негладких опуклих функцій, для розв'язання задачі (5) підійдуть методи мінімізації гладких опуклих функцій. Окрім того, якщо $p > 1$, то функція $F_p(y)$ є строго опуклою, тобто для довільних $a \in R^n$ та $b \in R^n$ виконується умова

$$F_p(\lambda a + (1-\lambda)b) < \lambda F_p(a) + (1-\lambda)F_p(b), \quad \forall a \neq b, 0 < \lambda < 1. \qquad (6)$$

Якщо $p = 1$, то задача (5) відповідає МНМ та зводиться до задачі лінійного програмування, враховуючи що функція (2) є кусково-лінійною. Якщо $p = 2$, то задача (5) відповідає МНК та є задачею мінімізації квадратичної функції

$$F_2(y) = \left(y_1 - y_1^1 \right)^2 + \ldots + \left(y_1 - y_1^m \right)^2 + \ldots + \left(y_n - y_n^1 \right)^2 + \ldots + \left(y_n - y_n^m \right)^2. \qquad (7)$$

При цьому оптимальні значення функцій в задачах (1) та (5) зв'язані співвідношенням $F_2^*(y_2^*) = \left(f_2^*(y_2^*) \right)^2$.

Лема. Якщо p – таке, що $1 < p \leq 2$, то задача (5) має єдиний розв'язок.

Доведення. Проведемо його методом від супротивного. Нехай $a^* \in R^n$ та $a^{**} \in R^n$ – два неспівпадаючі розв'язки задачі (5), яким відповідає оптимальне значення цільової функції $F_p^* = F_p(a^*) = F_p(a^{**})$.

Функція $F_p(y)$ при $1 < p \leq 2$ є строго опуклою, тому, враховуючи (6), для $F_p(a^{***})$ – значення функції $F_p(a)$ в точці $a^{***} = \lambda a^* + (1-\lambda)a^{**}$ − справедливі співвідношення

$$F_p(a^{***}) = F_p(\lambda a^* + (1-\lambda)a^{**}) <$$
$$< \lambda F_p(a^*) + (1-\lambda)F_p(a^{**}) = \lambda F_p^* + (1-\lambda)F_p^* = F_p^*,$$

з яких випливає нерівність $F_p(a^{***}) < F_p^*$. Вона суперечить тому, що a^* та a^{**} розв'язки задачі (5), так як в точці a^{***} значення функції $F_p(a^{***})$ є меншим за мінімальне значення F_p^*. Лема доведена.

Зазначимо, що при $p = 2$ задача (5) має аналітичний розв'язок $y_2^* = (y^1 + \ldots + y^m)/m$ – середнє арифметичне m лінійних сплайнів $y^1, \ldots, y^m$, який випливає із мінімізації квадратичної функції (7).

Висновки. Для знаходження мінімальних за L_p-нормою лінійних сплайнів у роботі розглянуто задачі (1) та (5) – задачі безумовної мінімізації опуклих функцій. У випадку $p \geq 1$ для задачі (1) використовується негладка функція, а у випадку $1 < p \leq 2$ для задачі (5) використовується гладка функція. Якщо задачі

(1) та (5) доповнити обмеженнями на властивості сплайна (монотонність, опуклість, увігнутість, кривина) на окремих ділянках інтервалу, то можна автоматизувати вибір сплайн функцій для інтерполяції чи апроксимації кривих ліній гладкими функціями необхідної степені гладкості. Це може бути використано для побудови аеродинамічних профілів із заданими ізогеометричними властивостями [6].

Розглянемо два приклади оптимізаційних задач з обмеженнями на властивості лінійного сплайна на двох окремих ділянках інтервалу $[x_1, x_n]$ – на ділянці $x_1 \leq \ldots \leq x_{n_1}$ та ділянці $x_{n_1} \leq \ldots \leq x_n$, де $1 < n_1 < n$. Так, наприклад, якщо $p \geq 1$ та невідомий лінійний сплайн повинен бути монотонно неспадним на першій ділянці інтервалу і монотонно незростаючим на другій ділянці інтервалу, то тоді достатньо вибрати таку задачу опуклого програмування: знайти

$$y_p^* = \arg\min_{y \in R^n} \left\{ f_p(y) = \left(\sum_{j=1}^{n} \sum_{i=1}^{m} \left| y_j - y_j^i \right|^p \right)^{1/p} \right\} \tag{7}$$

при лінійних обмеженнях

$$y_j \leq y_{j+1}, \quad j = 1, \ldots, n_1, \tag{8}$$

$$y_j \geq y_{j+1}, \quad j = n_1, \ldots, n. \tag{9}$$

Якщо при $1 \leq p \leq 2$ невідомий лінійний сплайн повинен бути увігнутим на першій ділянці інтервалу та опуклим на другій ділянці інтервалу, то тоді можна вибрати таку задачу опуклого програмування: знайти

$$y_p^* = \arg\min_{y \in R^n} \left\{ F_p(y) = \sum_{j=1}^{n} \sum_{i=1}^{m} \left| y_j - y_j^i \right|^p \right\} \tag{10}$$

при лінійних обмеженнях

$$\frac{y_{j+1} - y_j}{x_{j+1} - x_j} \geq \frac{y_{j+2} - y_{j+1}}{x_{j+2} - x_{j+1}}, \quad j = 1, \ldots, n_1 - 2, \tag{11}$$

$$\frac{y_{j+1} - y_j}{x_{j+1} - x_j} \leq \frac{y_{j+2} - y_{j+1}}{x_{j+2} - x_{j+1}}, \quad j = n_1, \ldots, n - 2. \tag{12}$$

Оптимізаційні задачі (7) – (9) та (10) – (12) можна ефективно розв'язувати за допомогою сучасних модифікацій r-алгоритмів – субградієнтних методів з розтягом простору в напрямку різниці двох послідовних субградієнтів [7 – 9]. При цьому кількість ітерацій для знаходження мінімального значення f^* з точністю ε для опуклих функцій від n змінних емпірично оцінюється як $N = O\left(n \log\left(1/\varepsilon\right)\right)$. Відмітимо, що використання методів мінімізації негладких функцій дає можливість будувати негладкі цільові функції, що значно розширює набір критеріїв оптимальності аеродинамічних профілів, та дозволяє включати негладкі функції в обмеження оптимізаційної задачі, що значно розширює можливості на вимоги до сплайнів на окремих ділянках інтервалу.

Робота виконана за фінансової підтримки НАН України (проект № 0118U005227) та Volkswagen Foundation (грант No 90 306).

П.И. Стецюк, О.Н. Хомяк

МИНИМАЛЬНЫЕ В L_p-НОРМЕ ЛИНЕЙНЫЕ СПЛАЙНЫ

Изучаются задачи безусловной минимизации выпуклых функций для нахождения минимальных в L_p-норме линейных сплайнов для случаев $p \geq 1$ и $1 \leq p \leq 2$. Если $p \geq 1$, то используется негладкая функция, а если $1 < p \leq 2$ – гладкая функция. Показано, что при определенном выборе параметра p оптимизационные задачи порождают известные методы – метод наименьших квадратов, метод наименьших модулей и минимаксный чебышевский метод. Приведены свойства решений задачи для случая $1 < p \leq 2$.

P.I. Stetsyuk, O.M. Khomiak

THE MINIMAL LINEAR SPLINES IN L_p-NORM

Problems of unconstrained minimization of convex functions for finding the minimal linear splines in L_p-norm for cases $p \geq 1$ and $1 \leq p \leq 2$ are investigated. If $p \geq 1$, then the non-smooth function is used, and if $1 < p \leq 2$ then the smooth function is used. It is shown, that with a certain choice of parameter p, the optimization problems generate the known methods: the method of least squares, the method of least absolute deviations, and the Chebyshev minimax method. The properties of solutions of problems with $1 < p \leq 2$ are given.

Список літератури

1. Мудров В.И., Кушко В.Л. Метод наименьших модулей. М.: Знание, 1971. 64 с.
2. Зоркальцев В.И. Метод наименьших квадратов: геометрические свойства, альтернативные подходы, приложения. Новосибирск: Наука, 1995. 220 с.
3. Хьюбер П. Робастность в статистике. М.: Мир, 1984. 304 с.
4. Стецюк П.И., Колесник Ю.С., Лейбович М.М. О робастности метода наименьших модулей. *Компьютерная математика*. 2002. Вып. 2. С. 114 – 123.
5. Кларк Ф. Оптимизация и негладкий анализ. М.: Наука, 1988. 280 с.
6. Квасов Б.И. Методы изогеометрической аппроксимации сплайнами. М.: ФИЗМАТЛИТ, 2006. 360 с.
7. Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. Киев: Наук. думка, 1979. 200 с.
8. Стецюк П.И. Теория и программные реализации r-алгоритмов Шора. *Кибернетика и системный анализ*. 2017. № 5. С. 43 – 57.
9. Стецюк П.І. Комп'ютерна програма "Octave-програма ralgb5a: $r(\alpha)$-алгоритм з адаптивним кроком". Свідоцтво про реєстрацію авторського права на твір № 85010. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. Дата реєстрації 29.01.2017.

Одержано 19.03.2019