Теорія оптимальних рішень

УДК 519.8

Ю.П. ЛАПТИН

НЕКОТОРЫЕ ВОПРОСЫ ОПРЕДЕЛЕНИЯ КОЭФФИЦИЕНТОВ НЕГЛАДКИХ ШТРАФНЫХ ФУНКЦИЙ

Рассматриваются процедуры автоматического определения значений штрафных коэффициентов по ходу работы оптимизационного алгоритма. Устанавливается связь с известными результатами.

Точные штрафные функции были впервые предложены в работах [1, 2]. В дальнейшем этому направлению было посвящено большое количество публикаций. Укажем лишь некоторые [3-5]. В настоящее время метод штрафных функций является основным средством при использовании *r*-алгоритма Н.З. Шора для решения оптимизационных задач с ограничениями. Однако использование точных штрафных функций связано с определенными проблемами – отсутствуют простые методики вычисления приемлемых значений штрафных коэффициентов. Выбор значений коэффициентов, обычно, возлагается на пользователя, что приводит либо к завышению используемых значений, либо к необходимости многократного решения одной и той же задачи для удовлетворительного подбора штрафных коэффициентов. В работе [6] рассматривался подход, позволяющий построить процедуру автоматического определения значений штрафных коэффициентов по ходу работы оптимизационного алгоритма. В настоящей работе предлагается обобщение этого подхода и приводится ряд полезных результатов.

Рассматривается задача: найти

$$f^* = \min\{f(x) : x \in C\}, \qquad (1)$$

где $C = \left\{x \in R^n : h(x) \le 0\right\}$, $f, h : R^n \to R$ – выпуклые функции, принимающие конечные значения при любых x, C – компактное множество. Будем рассматривать штрафные функции вида

$$S(x,s) = f(x) + s \cdot h^{+}(x), \ s \in R, \ s \geq 0, \tag{2}$$

где $h^{+} = \max\{0, h\}$, и выпуклую задачу: найти

$$S^{*}(s) = \min\left\{S(x,s) : x \in R^{n}\right\}. \tag{3}$$

Штрафная функция $S(x,s)$ называется точной при заданных значениях штрафных коэффициентов s, если $S^{*}(s) = f^{*}$ и решения задач (1) и (3) совпадают.

Пусть $S^{'}(x,s,p)$, $f^{'}(x,p)$, $h^{'}(x,p)$ – производные функций S, f, h в точке $x \in R^{n}$ по направлению p при фиксированном значении s, $p(x,y) = (y-x)/\|y-x\|, \ y \neq x$.

Лемма 1. Пусть для любого $x \notin C$ найдется $y \in C$ такое, что

$$S^{'}(x,s,p(x,y)) < 0, \tag{4}$$

тогда $S(x,s)$ – точная штрафная функция.

Доказательство очевидно, поскольку если оптимальное решение $\tilde{x}$ задачи (3) не принадлежит допустимому множеству C, то для $S(x,s)$ всегда найдется направление убывания $p(\tilde{x}, y)$ в этой точке, что противоречит предположению об оптимальности точки $\tilde{x}$.

Лемма 1 позволяет формулировать простые условия, которые должны выполняться на каждом шаге оптимизационных алгоритмов при решении задачи (3), если $S(x,s)$ точная штрафная функция. Для этого должно быть определено правило, по которому для текущей точки x^{k}, сгенерированной на итерации k алгоритма, ставится в соответствие точка $y^{k} \in C$, если $x^{k} \notin C$. Такие правила можно задавать различным образом, например, всегда полагать $y^{k} = y^{0}$, где y^{0}, такая что $y^{0} : h(y^{0}) < 0$ – начальная допустимая точка, или y^{k} выбирать среди допустимых точек, сгенерированных на предыдущих итерациях.

Однако, такие необходимые условия оказываются недостаточными – условия могут выполняться для всех точек, генерируемых оптимизационным алгоритмом при решении задачи (3), но не выполняться для предельной точки. Достаточные условия будут приведены в теореме 1.

Рассмотрим эти вопросы более подробно. Пусть $\tilde{x}$ – решение задачи (3), заданы сходящиеся последовательности $x^k \in R^n$, $y^k \in C$, $k = 0,1, \ldots$, $x^k \to \tilde{x}$ при $k \to \infty$, Обозначим $\tilde{y} = \lim\limits_{k\to\infty} y^k$. Пусть значение s зафиксировано и для каждого x^k, такого что $x^k \notin C$, выполняется

$$S'(x^k, s, p(x^k, y^k)) < 0. \tag{5}$$

Следующий пример одномерной задачи демонстрирует, что в предельной точке $\tilde{x}$ неравенство (5) может не выполняться:

$f(x) = -x$, $h(x) = \max\{h_1(x), h_2(x)\}$, где $h_1(x) = |x| - 1$, $h_2(x) = |2x| - 3$.

Решением задачи (1) в этом случае есть $x^* = 1$, $f^* = -1$. Пусть $x^k = 2 + 1/k$, $k = 0,1, \ldots$, $y^k = 0$, $k = 0,1, \ldots$, штрафной коэффициент $s = 0,75$. Нетрудно видеть, что $\tilde{x} = 2$ есть решение задачи (3), условия (5) выполняются, но $\tilde{x} = 2$ не является решением задачи (1).

Пусть $x^k \notin C$. Обозначим $\pi_C(x^k, y^k)$ точку пересечения отрезка $\left[x^k, y^k\right]$ с границей множества C, $\overline{x}^k = \pi_C(x^k, y^k)$.

Теорема 1. Пусть заданы $\varepsilon > 0$, $s > 0$ такие, что для каждого x^k, $x^k \notin C$, $k = 0,1, \ldots$ выполняется

$$S'(\overline{x}^k, s, p(\overline{x}^k, x^k)) \geq \varepsilon. \tag{6}$$

тогда $\tilde{x}$ является решением задачи (1), т.е. $S(x, s)$ – точная штрафная функция.

Доказательство. Понятно, что если $\tilde{x} \in C$, то $\tilde{x}$ – решение задачи (1).

Предположим, что $\tilde{x} \notin C$. В силу сходимости последовательностей x^k, y^k, $k = 0,1, \ldots$ последовательность $\overline{x}^k$, $k = 0,1, \ldots$ сходится к точке $\tilde{z} = \pi_C(\tilde{x}, \tilde{y})$. В силу выпуклости функции $S(x, s)$ имеем

$$-S'(\tilde{x}, s, p(\tilde{x}, \tilde{z})) \geq \frac{S(\tilde{x}, s) - S(\tilde{z}, s)}{\|\tilde{x} - \tilde{z}\|}, \tag{7}$$

$$\frac{S(x^k, s) - S(\overline{x}^k, s)}{\left\|x^k - \overline{x}^k\right\|} \geq S'(\overline{x}^k, s, p(\overline{x}^k, x^k)). \tag{8}$$

В силу непрерывности функции $S(x, s)$ получаем $\dfrac{S(\tilde{x}, s) - S(\tilde{z}, s)}{\|\tilde{x} - \tilde{z}\|} = \lim\limits_{k\to\infty} \dfrac{S(x^k, s) - S(\overline{x}^k, s)}{\left\|x^k - \overline{x}^k\right\|}$ и, с учетом (6) и (8),

$\dfrac{S(\tilde{x},s)-S(\tilde{z},s)}{\|\tilde{x}-\tilde{z}\|} \ge \dfrac{\varepsilon}{2}$. Откуда $S^{'}(\tilde{x},s,p(\tilde{x},\tilde{z})) \le -\dfrac{\varepsilon}{2}$. Т.е. в точке $\tilde{x}$ существует направление убывания $p(\tilde{x},\tilde{z}))$, что противоречит предположению об оптимальности точки $\tilde{x}$. ■

Для решения задачи (3) при использовании данного подхода может использоваться любой глобально сходящийся алгоритм выпуклой безусловной оптимизации. Последовательность $x^k \in R^n$, $k = 0,1,\ldots$ в этом случае генерируется оптимизационным алгоритмом. Последовательность $y^k \in C$, $k = 0,1,\ldots$ должна задаваться некоторым специальным образом. На каждом шаге оптимизационного алгоритма необходимо проверять условие (6), что требует решения одномерной задачи поиска точки пересечения отрезка $\left[x^k, y^k\right]$ с границей множества C.

В случае, когда неравенство (6) на некоторой итерации алгоритма нарушается, будем увеличивать (на этой итерации) штрафной коэффициент s так, чтобы неравенство (6) выполнилось. При этом увеличение штрафного коэффициента будем производить на величину не менее B, где $B > 0$ – заданный параметр. Легко видеть, что если существует конечное $\overline{s}$ такое, что при $s > \overline{s}$ неравенство (6) выполняется на всех итерациях алгоритма, то количество таких увеличений штрафного коэффициента по ходу работы оптимизационного алгоритма будет конечно. Условия существования конечного $\overline{s}$ будут сформулированы в теореме 2.

Неравенство (4) перепишем в виде

$$f^{'}(x, p(x,y)) + s \cdot h^{'}(x, p(x,y)) < 0. \tag{9}$$

Заметим, что $h^{'}(x, p(x,y)) < 0$, если $x \notin C$ и $y \in C$.

Положим $r(x,y) = \max\left\{0, -\dfrac{f^{'}(x,p(x,y))}{h^{'}(x,p(x,y))}\right\}$,

$$r(y^0) = \sup\left\{r(\overline{x},x) : \overline{x} = \pi_C(x,y^0),\ x \notin C\right\},\ y^0 \in C \tag{10}$$

Лемма 2. Пусть задана точка $y^0 \in C$, такая что $h(y^0) < 0$, тогда $r(y^0) < +\infty$.

*Доказательство***.** В силу ограниченности множества C функция f липшицева, и существует $M > 0$, такое что $\left|f'(\overline{x}, p(\overline{x},x))\right| \le M$ для любых $x \notin C$,

$\overline{x} = \pi_C(x, y^0)$. Положим $m = \max\left\{ \left\| \overline{x} - y^0 \right\| : \overline{x} = \pi_C(x, y^0),\ x \notin C \right\}$. В силу выпуклости функции h имеем $h'(\overline{x}, p(\overline{x}, x)) \geq \dfrac{h(\overline{x}) - h(y^0)}{\left\| \overline{x} - y^0 \right\|}$. Откуда

$$h'(\overline{x}, p(\overline{x}, x)) \geq \left| h(y^0) \right| m^{-1}, \tag{11}$$

и $r(y^0) \leq \dfrac{mM}{\left| h(y^0) \right|}$. ■

Теорема 2. Пусть задана последовательность $x^k \in R^n$, $k = 0,1, \ldots$, сходящаяся к решению $\tilde{x}$ задачи (3), $y^k = y^0$, $k = 1, 2, \ldots$, $h(y^0) < 0$. Тогда существуют $\overline{s} < \infty$ такое, что при $s > \overline{s}$ выполняются условия теоремы 1.

Доказательство. Положим $\overline{s} = r(y^0)$. Нетрудно видеть, что

$$S'(\overline{x}^k, s, p(\overline{x}^k, x^k)) = f'(\overline{x}^k, p(\overline{x}^k, x^k)) + (s + \overline{s} - \overline{s}) \cdot h'(\overline{x}^k, p(\overline{x}^k, x^k)) =$$

$$= f'(\overline{x}^k, p(\overline{x}^k, x^k)) + \overline{s} \cdot h'(\overline{x}^k, p(\overline{x}^k, x^k)) + (s - \overline{s}) \cdot h'(\overline{x}^k, p(\overline{x}^k, x^k)) =$$

$$= h'(\overline{x}^k, p(\overline{x}^k, x^k)) \left[\overline{s} - \frac{-f'(\overline{x}^k, p(\overline{x}^k, x^k))}{h'(\overline{x}^k, p(\overline{x}^k, x^k))} \right] + (s - \overline{s}) \cdot h'(\overline{x}^k, p(\overline{x}^k, x^k)) \geq$$

$$\geq (s - \overline{s}) \cdot h'(\overline{x}^k, p(\overline{x}^k, x^k)).$$

Учитывая (11) и полагая $\varepsilon = (s - \overline{s}) \left| h(y^0) \right| m^{-1}$, получаем утверждение теоремы. ■

Приведенные результаты позволяют строить достаточно эффективные процедуры автоматического определения значений штрафных коэффициентов по ходу работы алгоритмов безусловной оптимизации. Простейшая процедура заключается в использовании начальной точки y^0, такой что $h(y^0) < 0$, на всех итерациях алгоритма для проверки условия (6). Однако при неудачном выборе точки y^0 значения штрафных коэффициентов могут устанавливаться достаточно большими. Уменьшить эти значения можно за счет специального выбора точек y^k на каждой итерации алгоритма. Вопросы оценки возможных значений штрафных коэффициентов рассматриваются далее.

Будем выбирать точки y^k на каждой итерации оптимизационного алгоритма так, чтобы неравенство

$$S'(\overline{x}^k, s, p(\overline{x}^k, x^k)) = f'(\overline{x}^k, p(\overline{x}^k, x^k)) + s \cdot h'(\overline{x}^k, p(\overline{x}^k, x^k)) > 0 \tag{12}$$

выполнялось при наименьшем значении штрафного коэффициента, здесь $\bar{x}^k = \pi_C(x^k, y^k)$. Обозначая, как и ранее, $r(x,y) = \max\left\{0, -\dfrac{f'(x, p(x,y))}{h'(x, p(x,y))}\right\}$, положим

$$r(x^k) = \min\left\{r(\bar{x}^k, x^k) : \bar{x}^k = \pi_C(x^k, y), y \in C\right\}, \tag{13}$$

$$y^k = \arg\min\left\{r(\bar{x}^k, x^k) : \bar{x}^k = \pi_C(x^k, y), y \in C\right\}, \tag{14}$$

$$r^* = \sup\left\{r(x) : x \notin C\right\}. \tag{15}$$

Используя для выбора y^k правило (14), и выбирая $s > r^*$, получим выполнение неравенств (12) на каждой итерации оптимизационного алгоритма. Таким образом, справедлива

Лемма 3. Пусть $s > r^*$, тогда $S(x,s)$ – точная штрафная функция.

Естественным является вопрос о том, как соотносятся полученные результаты с известными.

Пусть задача (1) является задачей линейного программирования: найти

$$f^* = \min\langle c, x\rangle, \tag{16}$$

$$\langle a_i, x\rangle + b_i \le 0,\ i = 1,\dots,m. \tag{17}$$

Для этой задачи $f(x) = \langle c, x\rangle$, $h(x) = \max\left\{h_i(x),\ i = 1,\dots,m\right\}$, где $h_i(x) = \langle a_i, x\rangle + b_i$, $i = 1,\dots,m$.

Теорема 3**.** Пусть задача (16), (17) имеет единственное решение x^*. Тогда $r^* = \sum_{i=1}^{m} u_i^*$, где u_i^*, $i = 1,\dots,m$ – оптимальные значения двойственных переменных.

Доказательство не приводится ввиду ограниченности объема статьи.

В работе рассмотрен подход, позволяющий строить достаточно эффективные процедуры автоматического определения значений штрафных коэффициентов по ходу работы алгоритмов безусловной оптимизации. При этом каждой точке x^k, генерируемой на итерации k алгоритма, должна ставиться в соответствие некоторая точка y^k из допустимого множества C исходной задачи. Для точек x^k и y^k выполняется проверка условия (6). Простейшая процедура заключается в использовании начальной точки y^0, такой что $h(y^0) < 0$, на всех итерациях алгоритма, т.е. $y^k = y^0$. Показано, что наилучший выбор точек y^k при решении задачи линейного программирования приводит к известному результату – для штрафного коэффициента r^* выполняется $r^* = \sum_{i=1}^{m} u_i^*$, где u_i^* – оптимальные значения двойственных переменных. Для повышения эффективности предлагаемого подхода необходимо разрабатывать правила (желательно простые) формирования допустимых точек y^k.

Ю.П. Лаптін

ДЕЯКІ ПИТАННЯ ВИЗНАЧЕННЯ КОЕФІЦІЕНТІВ НЕГЛАДКИХ ШТРАФНИХ ФУНКЦІЙ

Розгядаються процедури автоматичного визначення значень штрафних коефіцієнтів по ходу роботи оптімізаційного алгоритму. Аналізуються зв’язки з відомими результатами.

Yu.P.Laptin

SOME QUESTIONS FOR DETERMINING THE VALUES OF NONSMOOTH PENALTY FUNCTIONS

We consider an approach to construct an automatic procedure for determining the values of penalty coefficients in the process of optimization algorithm. Connections with known results are analized.

1. *Еремин И.И.* Метод "штрафов" в выпуклом про-граммировании // Докл. АН СССР. 1967. Т. 173, № 4. С. 748-751.
2. *Zangwill W.* Non-linear programming via penalty function // Manag. Sci. 1967. P. 344-358.
3. *Пшеничный Б.Н.* Метод линеаризации. – М.: Наука. – 1983. – 136 с.
4. *Shor N. Z.* Nondifferentiable Optimization and Polynomial Problems. – Amsterdam / Dordrecht / London: Kluwer Academic Publishers, 1998. – 381 p.
5. *Евтушенко Ю.Г., Жадан В.Г.* Точные вспомогательные функции в задачах оптимизации // ЖВМ и МФ, 1990. **30**. № 1. с. 43-57.
6. *Лаптин Ю.П.* Некоторые вопросы использования негладких штрафных функций // Теорія оптимальних рішень. 2011, № 10. с. 127 – 135.

Получено 14.05.2012