УДК 519.8
УДК 519.8

Ю.П. ЛАПТИН

ИСПОЛЬЗОВАНИЕ ШТРАФНЫХ ФУНКЦИЙ ДЛЯ РЕШЕНИЯ НЕКОТОРЫХ ЗАДАЧ ОПТИМИЗАЦИИ

Рассматриваются проблемы использования точных штрафных функций при решении оптимизационных задач с ограничениями. Предложены подходы, позволяющие определять штрафные коэффициенты по ходу работы оптимизационного алгоритма. Сформулированы достаточные условия, при которых решение вспомогательной задачи является решением исходной задачи.

Введение. В настоящее время метод точных штрафных функций широко используется при решении задач оптимизации с ограничениями [1–5]. Однако применение этого метода связано с определенными проблемами – отсутствуют простые методики вычисления приемлемых значений штрафных коэффициентов, необходимо проверять непростые условия регулярности, гарантирующие получение решения исходной задачи. Для выпуклых задач подходы, позволяющие определять штрафные коэффициенты по ходу работы оптимизационного алгоритма, рассматривались в [6]. В данной работе эти подходы уточняются и обобщаются на случай невыпуклых задач. Полученные результаты полезны при построении оптимизационных алгоритмов.

Рассматривается задача: найти

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

где $\Omega = \left\{ x \in R^n : h(x) \le 0 \right\}$, $f, h : R^n \to R$ – либо выпуклые функции, либо являются суперпозициями выпуклой и непрерывно дифференцируемых функций и принимают конечные значения при любых x, Ω – компактное множество,

и задача: найти

$$F_\lambda^* = \min\left\{ F_\lambda(x) : x \in R^n \right\}, \qquad (2)$$

где $F_\lambda(x)$ – штрафная функция вида

$$F_\lambda(x) = f(x) + \lambda \cdot h^+(x), \qquad (3)$$

здесь $\lambda \in R,\ \lambda \ge 0,\ h^+ = \max\{0, h\}$.

Рассматриваемые задачи могут быть невыпуклыми и многоэкстремальными. При определенных условиях локальные решения задачи (2) являются также локальными решениями задачи (1). Это позволяет для поиска решений задачи (1) использовать решение задачи (2) и применять для этого развитые алгоритмы безусловной оптимизации [7]. При этом возникает ряд проблем, среди которых – определение приемлемых значений штрафного коэффициента λ и правила отсева локальных решений задачи (2), не являющихся решениями задачи (1).

Обозначим $F_\lambda^{'}(x,p)$, $f^{'}(x,p)$, $h^{'}(x,p)$ – односторонние производные функций F_λ, f, h в точке $x \in R^n$ по направлению p. Определим $h^{\downarrow}(x)$, скорость наискорейшего спуска функции h в точке x,

$$h^{\downarrow}(x) = \min\left\{h'(x,p) : \|p\| = 1\right\},$$

Следуя [4], рассмотрим множество $\Omega_\delta = \left\{x \in R^n : h(x) \le \delta\right\}$, где $\delta > 0$. Пусть при некотором $a > 0$ для точек из Ω_δ выполняется условие

$$h^{\downarrow}(x) \le -a,\ \forall x \in \Omega_\delta \setminus \Omega. \tag{4}$$

Тогда, если функция f липшицева на $\Omega_\delta \setminus \Omega$, то существует λ^* такое, что при $\lambda > \lambda^*$ все точки локального минимума задачи (2), принадлежащие множеству Ω_δ, являются и точками локального минимума задачи (1). Можно показать, что при $\lambda > \lambda^*$ для $\forall x \in \Omega_\delta \setminus \Omega$ выполняется

$$\min\left\{F_\lambda^{'}(x,p) : \|p\| = 1\right\} < 0. \tag{5}$$

Если при решении задачи (2) некоторым сходящимся алгоритмом все порождаемые точки попадают в множество $\Omega_\delta \setminus \Omega$, то условие (5) можно бы использовать для подбора значения λ^* – если это условие не выполняется, текущее значение коэффициента λ следует увеличить.

Однако при таком подходе возникает две проблемы:

- трудоемкость решения задачи (5) может быть слишком высокой;
- условие (5) может выполняться для любой точки последовательности, формируемой оптимизационным алгоритмом, но не выполняться для предельной точки.

Для преодоления этих проблем рассмотрим несколько другой подход.

Теорема 1. Пусть заданы числа $\delta > 0, \beta > 0$, $\lambda > 0$, последовательность $x^k \in R^n$, $k = 0,1,\ldots$, сходится к локальному минимуму $\tilde{x}$ задачи (2), каждой точке x^k, $x^k \notin \Omega$ поставлен в соответствие вектор $p^k \in R^n$, $\left\|p^k\right\| = 1$ и выполняются условия

$$\max\left\{F_\lambda'(x^k+tp^k,p^k):t\in[0,\delta],\ h(x^k+tp^k)>0\right\}\le-\beta, \tag{6}$$

Тогда $\tilde{x}$ является локальным минимумом задачи (1).

Доказательство. Если $\tilde{x}\in\Omega$, то, очевидно, $\tilde{x}$ является локальным минимумом задачи (1). Предположим $\tilde{x}\notin\Omega$. Без ограничения общности можно считать, что последовательность векторов p^k, $k=0,1,\ldots$, сходится к некоторому вектору $\tilde{p}$. Обозначим

$$\tilde{t}=\min\{t:t>0,\ \tilde{x}+t\tilde{p}\in\Omega\}. \tag{7}$$

Точка $\tilde{x}+\tilde{t}\cdot\tilde{p}$ лежит на границе множества Ω. Положим $\tilde{\delta}=\min(\delta,\tilde{t})$.

Покажем, что существует номер K такой, что $h(x^k+tp^k)>0$, при $t\in\left[0,\frac{\tilde{\delta}}{2}\right]$, $k>K$. Предположим противное, т. е. для любого i найдется $k_i>i$, для которого существует точка $t_{k_i}\in\left[0,\frac{\tilde{\delta}}{2}\right]$ такая, что $h(x^{k_i}+t_{k_i}p^{k_i})\le0$. Без ограничения общности можно считать, что последовательность t_{k_i}, $i=1,\ldots$, сходится. Обозначим $\bar{t}=\lim\limits_{i\to\infty}t_{k_i}$. Имеем $\bar{t}\le\frac{\tilde{\delta}}{2}<\tilde{t}$, $h(\tilde{x}+\bar{t}\cdot\tilde{p})\le0$, т. е. $\tilde{x}+\bar{t}\cdot\tilde{p}\in\Omega$ что противоречит (7).

В силу (6) можно показать, имеет место

$$F_\lambda(x^k+tp^k)\le F_\lambda(x^k)-\beta t+o_1(t),\ t\in\left[0,\frac{\tilde{\delta}}{2}\right],\ k>K. \tag{8}$$

Покажем, что $F_\lambda'(\tilde{x},\tilde{p})\le-\beta$. Предположим противное, т. е. $F_\lambda'(\tilde{x},\tilde{p})>-\beta$. Выберем число $\gamma>0$ так, что $F_\lambda'(\tilde{x},\tilde{p})=-\beta+\gamma$. В силу дифференцируемости по направлению для $t\ge0$ имеет место разложение

$$F_\lambda(\tilde{x}+t\tilde{p})=F_\lambda(\tilde{x})+F_\lambda'(\tilde{x},\tilde{p})t+o_2(t)=F_\lambda(\tilde{x})-\beta t+\gamma t+o_2(t).$$

В силу $\frac{o_1(t)}{t}\to0$, $\frac{o_2(t)}{t}\to0$ при $t\to0$ можно выбрать $\tau\in\left[0,\frac{\tilde{\delta}}{2}\right]$ так, что $\left|\frac{o_1(\tau)}{\tau}\right|\le\frac{\gamma}{3}$, $\left|\frac{o_2(\tau)}{\tau}\right|\le\frac{\gamma}{3}$. Откуда $F_\lambda(x^k+\tau p^k)\le F_\lambda(x^k)-\beta\tau+\frac{\gamma}{3}\tau$, $F_\lambda(\tilde{x}+\tau\tilde{p})\ge F_\lambda(\tilde{x})-\beta\tau+\gamma\tau-\frac{\gamma}{3}\tau=F_\lambda(\tilde{x})-\beta\tau+\frac{2\gamma}{3}\tau$ и $F_\lambda(\tilde{x}+\tau\tilde{p})\ne$ $\ne\lim\limits_{k\to\infty}F_\lambda(x^k+\tau p^k)$, что противоречит условию непрерывности функции F_λ.

Таким образом, показано, что $F'_{\lambda}(\tilde{x}, \tilde{p}) \le -\beta$. Но, тогда точка $\tilde{x}$ не является локальным минимумом задачи (2), т. е. предположение $\tilde{x} \notin \Omega$ неверно. Теорема доказана.

Пусть задано число $\sigma > 0$ и выполняются условия теоремы 1. Предположим, что для каждой точки $x^k \in R^n$, $k = 0,1, \ldots,$ $x^k \notin \Omega$ также выполняется

$$\max\left\{h'(x^k + tp^k, p^k) : t \in [0, \delta],\, h(x^k + tp^k) > 0\right\} \le -\sigma. \qquad (9)$$

Тогда при выполнении условия (9), если условие (6) не выполняется на некотором шаге оптимизационного алгоритма, штрафной коэффициент λ может быть увеличен на конечную величину так, чтобы условие (6) выполнилось.

Наиболее сложным при использовании полученных результатов является выбор векторов p^k. В некоторых случаях это может быть сделано достаточно эффективно. Рассмотрим ситуацию, когда функции f, h – выпуклые. Пусть задана точка x^0 такая, что $h(x^0) < 0$. В качестве вектора p^k выберем вектор, порождающий прямую, исходящую из точки x^k и проходящую через x^0. Тогда, можно показать, что

1) существует $\sigma > 0$ такое, что условие (9) выполняется для любой точки $x^k \notin \Omega$;

2) существует $\lambda^* > 0$ такое, что при $\lambda > \lambda^*$ условие (6) выполняется для любой точки $x^k \notin \Omega$.

Таким образом, для задачи выпуклого программирования процедура подбора значения штрафного коэффициента λ достаточно простая

В невыпуклом случае ситуация существенно усложняется. В общем случае задача (1) многоэкстремальная, допустимое множество может быть многосвязным. Штрафная добавка $h^+(x)$ может иметь локальные минимумы в недопустимой области

Будем предполагать заданной допустимую точку $x^0 \in \Omega$ такую, что $h(x^0) < 0$. Обозначим Ω^0 односвязную компоненту допустимого множества задачи (1), содержащую точку x^0. Будем искать точку x^* локального минимума задачи (1) при дополнительном требовании $x^* \in \Omega^0$.

Естественно предполагать, что если точки последовательности, порождаемой при решении «штрафной» задачи (2), не слишком далеко удаляются от границы множества Ω^0, то условие (9) будет выполняться и предельная точка будет принадлежать множеству Ω^0. Для контроля за степенью удаления от границы Ω^0 введем вспомогательные понятия.

Траекторией $\theta(x,t)$, исходящей из точки x и проходящей через точку x^0, назовем непрерывно дифференцируемую по t вектор-функцию $\theta : R \to R^n$ такую, что $\theta(x,0) = x$ и $\theta(x,t_x^0) = x^0$ при некотором $t_x^0 > 0$. Предполагается, что $\left\|\theta_t^{'}(x,t)\right\| = 1$ для любого $t \geq 0$. Для непрерывной функции $\varphi : R^n \to R$, имеющей односторонние производные по направлениям, обозначим $\varphi'_\theta(x,t) = \lim\limits_{\Delta t \to +0} \dfrac{\varphi(\theta(x,t+\Delta t)) - \varphi(\theta(x,t))}{\Delta t}$.

Для контроля за степенью удаления от границы Ω^0 для каждой точки $x^k \in R^n$, $k = 0,1, ...,$ порождаемой при решении задачи (2), будем анализировать поведение функции h вдоль траектории $\theta(x^k,t)$, $t \geq 0$ – проверять условие

$$\max\left\{h'_\theta(x^k,t) : t \in \left[0, t^0_{x^k}\right], h(\theta(x^k, t)) > 0\right\} \leq -\sigma, \tag{10}$$

а для уточнения штрафного коэффициента λ – условие

$$\max\left\{F'_{\lambda\theta}(x^k,t) : t \in [0,\delta], h(\theta(x^k, t)) > 0\right\} \leq -\beta. \tag{11}$$

Здесь $\delta > 0$, $\beta > 0$, $\sigma > 0$.

Можно показать, что при выполнении этих условий для предельной точки $\tilde{x}$ (локального минимума задачи (2)) имеет место $\tilde{x} \in \Omega^0$.

Существенной проблемой является выбор траекторией $\theta(x,t)$. Простейшей траекторией является прямая, исходящая из точки x и проходящая через точку x^0. Понятно, что расширение класса возможных траекторий предоставляет дополнительные возможности.

Рассмотрим задачу: найти

$$f^* = \min\left\{f(y) : y \in \Omega_y,\ y = \Phi(x),\ x \in \Omega_x\right\}, \tag{12}$$

где $\Omega_y = \left\{y \in Y : h^y(y) \leq 0\right\}$, $\Omega_x = \left\{x \in X : h^x(x) \leq 0\right\}$, $Y = R^{n_y}$, $X = R^{n_x}$, $f, h^y : Y \to R$, $h^x : X \to R$ – выпуклые функции, принимающие конечные значения при любых x, y, Ω_x, Ω_y – компактные множества. Будем предполагать, что $n_x = n_y$ и $\Phi : X \to Y$ – непрерывно дифференцируемое взаимнооднозначное отображение. В случае, когда Φ – линейное невырожденное отображение, задача (12) есть задача выпуклого программирования.

Обозначим $\Omega_y^{\Phi} = \{x : x \in X, \Phi(x) \in \Omega_y\}$ – прообраз множества Ω_y в пространстве X, Ω – множество допустимых точек задачи (12), $\Omega = \Omega_x \bigcap \Omega_y^F$. Будем предполагать заданной допустимую точку $x^0 \in \Omega_x$ такую, что $h^x(x^0) < 0$, $h^y(y^0) < 0$, где $y^0 = \Phi(x^0)$.

Для произвольной точки $x \notin \Omega$ возможными траекториями будем считать прямую $\theta_x(x,t)$ в пространстве X, соединяющую точки x и x^0, и прямую $\theta_y(y,t)$ в пространстве Y, соединяющую точки $y = F(x)$ и $y^0 = \Phi(x^0)$. Прообразом прямой $\theta_y(y,t)$ будет некоторая траектория (кривая) $\theta_y^{\Phi}(x,t)$ в пространстве X, соединяющая точки x и x^0.

Проверка условий (10), (11) может выполняться сначала для траектории $\theta_x(x,t)$ и, если эти условия не выполняются, производиться для траектории $\theta_y^{\Phi}(x,t)$. Как возможные траектории могут также рассматриваться линейные комбинации $\theta_x(x,t)$ и $\theta_y^{\Phi}(x,t)$.

Необходимо отметить, что проверка условий (10) является, в общем случае, достаточно трудоемкой. Также и построение последовательностей, удовлетворяющих условиям (10), (11), связано со значительными проблемами – могут потребоваться достаточно сложные процедуры поиска подходящей точки на текущей итерации оптимизационного алгоритма, а также процедуры выбора новой базовой точки x^0.

Ю.П. Лаптін

ВИКОРИСТАННЯ ШТРАФНИХ ФУНКЦІЙ
ДЛЯ РОЗВ'ЯЗАННЯ ДЕЯКИХ ЗАДАЧ ОПТИМІЗАЦІЇ

Розглядаються проблеми використання точних штрафних функцій при розв'язанні оптимізаційних задач з обмеженнями. Запропоновані підходи, що дозволяють визначати штрафні коефіцієнти по ходу роботи оптимізаційного алгоритму. Сформульовані достатні умови, за яких розв'язок допоміжної задачі є розв'язком вихідної задачі.

Yu.P. Laptin

USING PENALTY FUNCTIONS FOR SOLVING SOME OPTIMIZATION PROBLEMS

There are considered the problems of use of exact penalty functions for solving optimization problems with constraints. The approaches are proposed that allow to determine the coefficients of penalty in the course of an optimization algorithm. Sufficient conditions are formulated under which the solution of the auxiliary problem is the solution of the original problem.

1. *Пшеничный Б.Н.* Метод линеаризации. – М.: Наука. – 1983. – 136 с.
2. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – Amsterdam / Dordrecht / London: Kluwer Academic Publishers, 1998. – 381 p.
3. *Евтушенко Ю.Г., Жадан В.Г.* Точные вспомогательные функции в задачах оптимизации // ЖВМ и МФ. – 1990. – Т. 30. – № 1. – С. 43–57.
4. *Демьянов В.Ф.* Условия экстремума и вариационное исчисление. – М.: Высш. шк., 2005. – 335 с.
5. *Данилин Ю.М.* Линеаризация и штрафные функции // Кибернетика и системный анализ. – 2002. – № 5. – С. 65–78.
6. *Лаптин Ю.П.* Некоторые вопросы опредления коэффициентов негладких штрафных функций // Теорія оптимальних рішень. – 2012. – С. 73–79.
7. *Шор Н.З.* О классе почти-дифференцируемых функций и одном методе минимизации функций этого класса // Кибернетика. – 1972. – № 4. – С. 9 – 17.

Получено 19.03.2013