УДК 512.25/26

К ХАРАКТЕРИСТИКЕ ШТРАФНЫХ ФУНКЦИЙ

А.А.Каплан

В работе исследуется вопрос о построении последовательности функций, с помощью которой может быть реализован метод штрафов для решения задач выпуклого программирования. Приводится описание структурных свойств функций штрафа, охватывающее широкий класс выпуклых функций.

Для иллюстрации полученных результатов рассматриваются функции штрафа, существенно отличающиеся от обычно применяемых на практике как по способу своего задания, так и по способу реализации градиентных методов отыскания экстремума.

Одним из универсальных методов решения экстремальных задач с ограничениями является так называемый метод штрафов, в котором исходная задача сводится к последовательности задач на безусловный экстремум путем введения штрафа за нарушение ограничений.

Схема метода применительно к задаче минимизации функции f при ограничениях

$$g^j(x) \leq 0, \quad j = 1, 2, \ldots, m,$$

(здесь и в дальнейшем все функции заданы в n -мерном арифметическом пространстве R^n) состоит в следующем.

Начиная процесс из некоторой фиксированной точки x^0, на $(k-1)$-м шаге мы определяем точку x^{k-1} и на k-м шаге с учетом x^{k-1} или независимо строим функцию штрафа

$$\Phi_k(x) = \varphi_k(g^1(x), g^2(x), \ldots, g^m(x)).$$

Далее, выбирая в качестве исходной точки x^{k-1}, находим x^k из условия,[ж)] что

$$F_k(x^k) = \min_{x \in R^n} F_k(x),$$

где $F_k(x) = f(x) + \Phi_k(x)$. Выбор функций Φ_k должен гарантировать существование у последовательности $\{x^k\}$ предельных точек, каждая из которых доставляет решение исходной задачи (в предположении, что таковое существует).

Следует отметить, что в некоторых работах рассматриваются штрафные функции Φ_k, заданные лишь в допустимой области, но обладающие тем свойством, что решение каждой вспомогательной задачи лежит строго внутри допустимой области и для его отыскания с некоторыми усложнениями могут быть использованы обычные методы отыскания безусловного экстремума [2], [3].

В настоящей статье метод штрафов исследуется применительно к задаче выпуклого программирования. Функции f и g^j, $j = 1, 2, \dots, m$, предполагаются выпуклыми, множество $\Omega = \{x : g^j(x) \le 0, j = 1, 2, \dots, m\}$ ограниченным и содержащим слейтеровскую точку. В этом случае естественно рассматривать выпуклые функции Φ_k: при их использовании существенно упрощается процесс решения вспомогательных задач.

Весьма общую характеристику функций штрафа для задач выпуклого программирования дает следующая теорема[жж)].

ТЕОРЕМА I. *Пусть Φ_k - выпуклые функции.*

1) $\lim_{k \to \infty} \Phi_k(x) = 0$; если $x \in \operatorname{int} \Omega$;

2) $\lim_{k \to \infty} \Phi_k(x) = +\infty$, если $x \notin \Omega$.

Тогда, начиная с некоторого K, функции $F_k(x) = f(x) + \Phi_k(x)$ достигают своего безусловного минимума. Последовательность точек минимума функций F_k ($k \ge K$) имеет предель-

---

ж) Предположение о достижимости минимума в каждой вспомогательной задаче не является, вообще говоря, обязательным (см., например, [I]).

жж) В работе [I] аналогичное утверждение доказано при существенно иных требованиях относительно функций f и Φ_k.

ные точки; любая предельная точка этой последовательности принадлежит множеству Ω и доставляет минимум f на Ω.

ДОКАЗАТЕЛЬСТВО. Пусть S - телесный многогранник с вершинами $y^1, y^2, \dots, y^\ell$; $S \subset int\,\Omega$. Нетрудно видеть, что при фиксированном $\varepsilon > 0$, начиная с некоторого K,

$$\max_{x \in S} \Phi_K(x) < \varepsilon.$$

Действительно, ввиду выпуклости функций Φ_K достаточно лишь проверить, что, начиная с некоторого K, справедливы неравенства

$$\Phi_K(y^i) < \varepsilon, \quad i = 1, 2, \dots, \ell.$$

Последнее же очевидно ввиду условия I).

Сопоставим теперь произвольному $\delta > 0$ выпуклое множество $\Omega_\delta = \{x : \rho(x, \Omega) \le \delta\}$, где $\rho(x, \Omega)$ - расстояние от точки x до множества Ω, и рассмотрим множество $\Omega_{\delta_2} \setminus \Omega_{\delta_1}$, считая, что $0 < \delta_1 < \delta_2$. Пусть $x^1 \in int\,S$, $u \in Q = \overline{\Omega_{\delta_2} \setminus \Omega_{\delta_1}}$ (черта означает замыкание). Точку $x^2 \in int(\Omega_{\delta_1} \setminus \Omega)$ выберем на отрезке с концами x^1 и u. Очевидно, множество

$$S' = \{z = x^2 + \lambda(x^2 - x) : \lambda > 0,\ x \in int\,S\}$$

содержит точку u с некоторой окрестностью U.

Задаваясь $\varepsilon > 0$ и $M > \varepsilon$, можно указать номер K^1, начиная с которого

$$\max_{x \in S} \Phi_K(x) < \varepsilon,$$

$$\Phi_K(x^2) > M.$$

Ввиду выпуклости функций Φ_K отсюда следует, что при $K \geq K^1$

$\Phi_K(x) > M$ для $x \in U$.

А так как u - произвольная точка множества Q, можно утверждать, что при любом $M > 0$ и любом $u \in Q$ найдется номер $K(u)$ такой, что при $K \geq K(u)$ неравенства $\Phi_K(x) > M$ будут одновременно выполняться в некоторой открытой окрестности точки u.

На основании компактности множества Q отсюда следует, что можно построить конечное покрытие

$$\{S^i\},\ i=1,2,\dots,\nu;\ \bigcup_{i=1}^{\nu} S^i \supset Q,$$

и при этом, начиная с некоторого L^i , для $x\in S^i$ имеем $\Phi_K(x) > M$. Тем самым при $K \geq L$, где $L=\max_{1\le i\le\nu} L^i$, неравенство $\Phi_K(x)>M$ должно выполняться на всем множестве Q .

Ввиду ограниченности функции f на Ω_{δ_2} и произвольности ε и M теперь ясно, что, начиная с некоторого $\widetilde{K}$, каждая из функций F_K должна достигать своего минимума, причем $\min_{x\in R^n} F_K(x) < F_K(y)$, если $y\in R^n\setminus\Omega_{\delta_1}$. Следовательно, из $\{x^K\}$ можно выделить подпоследовательность $\{x^{K_j}\}$, сходящуюся к некоторой точке $\bar{x}$. При этом, ввиду произвола в выборе δ_1 и δ_2 , точка $\bar{x}$ должна принадлежать Ω .

Остается показать, что $f(\bar{x})=\min_{x\in\Omega} f(x)$. Если это не так, то непременно найдется точка $\bar{\bar{x}}\in \operatorname{int}\Omega$, в которой имеет место неравенство $f(\bar{\bar{x}})<f(\bar{x})$ $(f(\bar{\bar{x}})=f(\bar{x})-a,\ a>0)$. Тогда можно указать телесный многогранник $G\subset \operatorname{int}\Omega$ $(\bar{\bar{x}}\notin G,\ \{\bar{x}+\lambda(\bar{\bar{x}}-\bar{x}):\lambda>0\}\cap \operatorname{int} G\neq\emptyset)$ и число J так, что при $j>J$ для $x\in G$ имеют место неравенства

$$f(x)<f(x^{K_j})-\frac{a}{2}$$

и, следовательно, неравенства

$$\Phi_{K_j}(x)>\Phi_{K_j}(x^{K_j})+\frac{a}{2}. \tag{I}$$

С другой стороны, так как G - многогранник, найдутся числа $\mu(K_j)$ такие, что для $x\in G$

$$\Phi_{K_j}(x)<\frac{a}{\mu(K_j)}, \tag{2}$$

причем $\mu(K_j)\to\infty$ при $j\to\infty$.

Замечая теперь, что, как бы ни было мало $\delta>0$, при достаточно больших j мы имеем

$$\Phi_{K_j}(\bar{\bar{x}})>-\delta \tag{3}$$

и

$$\{\bar{\bar{x}}+\lambda(\bar{\bar{x}}-x^{K_j}):\lambda>0\}\cap G\neq\emptyset,$$

нетрудно показать, что соотношения (I), (2) и (3) при больших j противоречивы ввиду выпуклости функций Φ_K .

Как видно из доказательства теоремы I, требование, чтобы функции Φ_K были выпуклы на всем R^n, может быть несколько ослаблено. Именно, справедливо следующее утверждение.

ТЕОРЕМА 2. Заключение теоремы I имеет место, если выполнены условия I) и 2), а функции Φ_K выпуклы в области $\Omega_\delta = \{x : \rho(x, \Omega) \le \delta\}$ при некотором $\delta > 0$ и квазивыпуклы на всем R^n.

ТЕОРЕМА 3. Пусть выполнены условия теоремы I или 2 и, кроме того, для граничных точек множества Ω

$$\Phi_K(x) \geq c > 0 \ , \quad K = 1, 2, \ldots$$

Тогда, начиная с некоторого K, точки x^K лежат внутри Ω.

ДОКАЗАТЕЛЬСТВО. Ввиду непрерывности функции f можно указать многогранник $S \subset int\,\Omega$ и компакт Ω' ($int\,\Omega' \supset \Omega$) такие, что

$$\min_{x \in \Omega'} f(x) > \min_{x \in S} f(x) - c/2 \ .$$

В силу условия I), начиная с некоторого K, будем иметь

$\Phi_K(x) < c/2$ для $x \in S$.

Отсюда ввиду квазивыпуклости функций Φ_K при $K \geq K$

$\Phi_K(x) \geq c$ для $x \in \overline{\Omega' \setminus \Omega}$,

а, следовательно,

$$\min_{x \in \Omega} \{ f(x) + \Phi_K(x) \} \le \min_{x \in S} \{ f(x) + \Phi_K(x) \} <$$

$$< \min_{x \in \Omega'} f(x) + c/2 + c/2 \le \min_{x \in \overline{\Omega' \setminus \Omega}} \{ f(x) + \Phi_K(x) \} ,$$

что и дает $x^K \in int\,\Omega$.

В приведенной ниже теореме 4 формулируются дополнительные требования относительно функций штрафа, при выполнении которых метод штрафов одновременно с решением исходной задачи, по существу, дает решение двойственной задачи. Отметим, что аналогичные результаты, но для штрафных функций специального вида (именно, $\Phi_k(x)=A_k\,G(g^1(x),g^2(x),\ldots,g^m(x))$, где в зависимости от свойств функции G либо $A_k\to\infty$, либо $A_k\to 0$) приведен в монографии [4] (см. также [3]). В сделанных выше предположениях относительно множества Ω эти результаты представляют собой частный случай*) теоремы 4.

Напомним, что в предположении о дифференцируемости функций $f, g^j, j=1,2,\ldots,m$, двойственная задача (см., например, [6]) состоит в отыскании максимума функции

$$f(x)+\sum_{j=1}^{m} y_j\, g^j(x)$$

при ограничениях

$$y_j \geqslant 0\ ,\quad j=1,2,\ldots,m,$$

$$\nabla f(x)+\sum_{j=1}^{m} y_j\, \nabla g^j(x)=\mathbb{0}\,.$$

Допустимое решение z исходной задачи и допустимое решение x, y ($y=(y_1,\ldots,y_m)'$) двойственной задачи связаны соотношением

$$f(z)\geqslant f(x)+\sum_{j=1}^{m} y_j\, g^j(x)\,,$$

и в наших предположениях относительно множества Ω в обеих задачах существуют оптимальные решения, при подстановке которых последнее неравенство превращается в равенство.

ТЕОРЕМА 4. Пусть функции $f, g^j, j=1,2,\ldots,m$, дифференцируемы, Φ_k, $k=1,2,\ldots$, дифференцируемы и удовлетворяют условиям теоремы I. Пусть также

$$\nabla\Phi_k(x)=\sum_{k=1}^{m} \psi_k^j(x)\,\nabla g^j(x)\,,$$

где функции ψ_k^j неотрицательны и непрерывны на R^n и при любом j последовательность $\{\psi_k^j(z^k)\}$ сходится к 0, если $\overline{\lim\limits_{k\to\infty}}\, g^j(z^k)<0$.

---

*) Следует иметь в виду, что в так называемых методах внутренней точки при каждом k функцию штрафа Φ_k, заданную лишь внутри Ω, можно считать продолженной с некоторого замкнутого выпуклого множества $\Omega_k\subset \operatorname{int}\Omega$ с сохранением выпуклости и дифференциальных свойств таким образом, что множества точек минимума во вспомогательных задачах при этом не изменяются [5].

Т о г д а ,   о б о з н а ч а я ж) $u_j^\kappa = \psi_\kappa^j(x^\kappa)$, м о ж н о   у т в е р ж д а т ь ,   ч т о

1) п о с л е д о в а т е л ь н о с т ь $\{u^\kappa\}$,   г д е $u^\kappa=(u_1^\kappa,\ldots,u_m^\kappa)$ и м е е т   п р е д е л ь н ы е   т о ч к и ;

2) т о ч к а $(\overline{x},\overline{u})$,   я в л я ю щ а я с я   п р е д е л о м   л ю б о й   с х о д я щ е й с я   п о д п о с л е д о в а т е л ь н о с т и   и з $\{(x^\kappa, u^\kappa)\}$,   д о с т а в л я е т   р е ш е н и е   д в о й с т в е н н о й   з а д а ч и .

ДОКАЗАТЕЛЬСТВО. В силу неотрицательности функций ψ_κ^j имеем $u_j^\kappa \geqslant 0$. Так как точка x^κ доставляет максимум функции F_κ , то

$$\mathbb{O} = \nabla F_\kappa(x^\kappa) = \nabla f(x^\kappa) + \sum_{j=1}^{m} \psi_\kappa^j(x^\kappa)\nabla g^j(x^\kappa) =$$

$$= \nabla f(x^\kappa) + \sum_{j=1}^{m} u_j^\kappa \nabla g^j(x^\kappa)$$

и, следовательно, (x^κ, u^κ) - допустимая точка двойственной задачи.

Возьмем произвольную фиксированную точку $x^* \in \operatorname{int}\Omega$. На основании выпуклости функций f и g^j , $j=1,2,\ldots,m$,

$$f(x^*) - f(x^\kappa) \geqslant (\nabla f(x^\kappa), x^* - x^\kappa) =$$

$$= \sum_{j=1}^{m} \psi_\kappa^j(x^\kappa)\,(\nabla g^j(x^\kappa), x^\kappa - x^*) \geqslant$$

$$\geqslant \sum_{j=1}^{m} \psi_\kappa^j(x^\kappa)\,(g^j(x^\kappa) - g^j(x^*)).$$

Пусть $\overline{x} = \lim_{\ell\to\infty} x^{\kappa_\ell}$ (см. теорему I), $J^1(\overline{x}) = \{j : g^j(\overline{x}) = 0\}$ и $J^2(\overline{x}) = \{j : g^j(\overline{x}) < 0\}$. В силу свойств функций ψ_κ^j ясно, что $\lim_{\ell\to\infty} u_j^{\kappa_\ell} = 0$ при $j \in J^2(\overline{x})$.

Далее, для достаточно больших ℓ при $j \in J^1$ имеем

$$g^j(x^{\kappa_\ell}) - g^j(x^*) > -\frac{g^j(x^*)}{2} > 0,$$

ж) Под x^κ по-прежнему понимается точка минимума функции F_κ . В силу теоремы I, начиная с некоторого K , функции F_κ достигают своего наименьшего значения. Соответственно, u^κ имеет смысл при $\kappa \geqslant K$.

и из соотношений

$$f(x^*)-f(x^{k_\ell})-\sum_{j\in J^2}\psi^j_{k_\ell}(x^{k_\ell})(g^j(x^{k_\ell})-g^j(x^*))>$$

$$>-\sum_{j\in J^1}\psi^j_{k_\ell}(x^{k_\ell})\cdot\frac{g^j(x^*)}{2}\geqslant-\max_{j\in J^1}\frac{g^j(x^*)}{2}\sum_{j\in J^1}u_j^{k_\ell}$$

видно, что $\sup_{j,\ell} u_j^{k_\ell}\leq c$, где c - некоторая константа. Ввиду допустимости точек (x^k, u^k) утверждение теоремы следует отсюда непосредственно.

ЗАМЕЧАНИЕ. Утверждение теоремы 4 сохраняет силу, если требование выпуклости функций Φ_k на всем R^n заменить соответствующим требованием из теоремы 2. Доказательство переносится на этот случай без сколь-нибудь существенных изменений.

Наиболее полно изучены в литературе, по-видимому, следующие функции штрафа:

$$\Phi_k(x)=A_k\sum_{j=1}^{m}\max[0,g^j(x)],\ \text{где } A_k>0,\ A_k\to\infty \text{ при } k\to\infty\quad;$$

$$\Phi_k(x)=A_k\sum_{j=1}^{m}\{\max[0,g^j(x)]\}^2,\ \text{где } A_k>0,\ A_k\to\infty \text{ при } k\to\infty\quad.$$

Другие применяемые в практике непрерывные функции штрафа также в той или иной форме опираются на идею "срезки". При отыскании минимума таких функций по заданному направлению приходится предварительно определять точки пересечения луча с множествами $\{x: g^j(x)=0\}$, $j=1,2,\dots,m$, или с частью из этих множеств, что требует решения на каждом шаге некоторого (значительного при большом m) числа, вообще говоря, нелинейных уравнений.

Рассматриваемые ниже функции штрафа свободны от указанного недостатка.

1. $$\Phi_k(x)=\sum_{j=1}^{m} e^{A_k^j g^j(x)}, \qquad (4)$$

где $A_k^j>0$, $A_k^j\to\infty$ при $k\to\infty$; $e^{A_k^j g^j(x)}$ представляет суперпозицию выпуклой монотонно возрастающей функции e^z и выпуклой функции $z=A_k^j g^j(x)$ и поэтому выпукла.

2. $$\Phi_k(x)=\sum_{j=1}^{m}(\bar g^j(x)+1)^k; \qquad (5)$$

K берутся нечетными, $\bar{g}^j(x)=\frac{1}{M^j}g^j(x)$, где M^j - достаточно большое положительное число (существенно больше, нежели $-\min_{x\in\Omega} g^j(x)$). Отметим сразу, что, если $\inf_{x\in R^n} g^j(x)>-\infty$, достаточно просто взять $M^j\geq-\inf_{x\in R^n} g^j(x)$. . Если $M^j\geq-\inf_{x\in R^n} g^j(x)$ для всех j , то K может быть любым натуральным.

За счет выбора M^j функция $(\bar{g}^j(x)+1)^K$ является выпуклой в области $\{x: g^j(x)\leq M^j\}$ и квазивыпуклой на всем R^n (если же $M^j\geq-\inf_{x\in R^n} g^j(x)$, то $(\bar{g}^j(x)+1)^K$ выпукла на R^n).

Функции (4), как легко видеть, удовлетворяют всем условиям теорем I, 3 и 4, а функции (5) - условиям теорем 2, 3 и 4 (с учетом замечания к последней). Следует иметь в виду, что при реализации монотонных градиентных процессов минимизации в случае, когда функция (5) не является выпуклой, не возникает сколь-нибудь существенных осложнений. Дело в том, что, начиная с некоторого K , такой процесс может быть осуществлен внутри области, где Φ_K - выпуклые функции. Вместо точного K можно использовать некоторую оценку для K сверху.

Отметим, что идея использования функции (4) для решения систем линейных неравенств принадлежит Моцкину [7]. При построении функции (5), по существу, была использована известная в литературе функция штрафа

$$\Phi_K(x)=\sum_{j=1}^{m}\{\max[0,\, g^j(x)+1]\}^K.$$

Переход от g^j к $\bar{g}^j$ как раз и продиктован желанием избежать срезки, которая имеет место при условии, что $\inf_{x\in R^n} g^j(x)<-1$ для некоторого j , и особенно неприятна, если $\min_{x\in\Omega} g^j(x)<-1$.

Литература

1. ЛЕВИТИН Е.С., ПОЛЯК Б.Т. Методы минимизации при наличии ограничений.- Ж. вычисл. матем. и матем. физ., 1966,6, № 5.
2. FRISH R.The logarithmik potential method for solving linear programming problems.- Memorandum Univ, of Inst. econ., Oslo, 1955.
3. FIACCO A., McCORMIK. The sequential unconstrained minimization technique for nonlinear programming.- Man. Sci., 1964, v.IO, N 2

4. ФИАККО А., МАК-КОРМИК Г. Нелинейное программирование (перев. с английского). Изд-во "Мир", 1972.

5. КАПЛАН А.А. Численные методы решения задач выпуклого программирования. "Оптимальное планирование", 1970, № 17.

6. WOLFE Ph. A duality theorem for nonlinear programming. Quart. Appl. Math., 1961, v.19, N3.

7. MOTZKIN T.S. New techniques for linear inequalities and optimization. Project SCOOP, Symposium on Linear Inequalities and Programming, Planning Research Division, Director of Management Analisis Service, U.S.Air Force, Washington, D.C., 1952, v.10.

Поступила в ред.-изд. отд.
4.УII. 1972 г.