УДК 519.85

Д. А. Марковцев

Московский физико-технический институт (государственный университет)

Условия сходимости итерационного процесса решения задач параметрического программирования методом гладких штрафных функций

Рассматривается итерационный процесс решения задач параметрического программирования с использованием метода гладких штрафных функций. Приводятся условия сходимости этого процесса.

Ключевые слова: параметрическое программирование, условия сходимости, метод гладких штрафных функций, итерационный процесс.

1. Введение

Одним из возможных способов расширения множества задач математического программирования, поддающихся решению численными методами, является *параметризация* их постановки. Параметризация заключается в разделении множества переменных исходной задачи:

$$\begin{aligned}&\text{найти } \min_{x,u} f(x,u) \\ &\text{при условиях } g_s(x,u) \geqslant 0, \quad s=[1,m],\end{aligned} \tag{1}$$

на два подмножества: собственно *переменных* $x \in E^n$ и *параметров* $u \in E^l$, значения которых при необходимости могут быть зафиксированы. Такое разделение позволяет свести решение исходной задачи к двухуровневой системе задач [2, 4]:

$$\begin{aligned}&\text{найти } \min_{x} f(x,u) \\ &\text{при условиях } g_s(x,u) \geqslant 0, \quad s=[1,m],\end{aligned} \tag{2}$$

$$\begin{aligned}&\text{найти } \min_{u} f(x^*(u),u) \\ &\text{при условиях } g_s(x^*(u),u) \geqslant 0, \quad s=[1,m],\end{aligned} \tag{3}$$

где $x^*(u)$ есть решение задачи (2) при фиксированном векторе параметров u. Обозначим локальное решение задачи (1) как $\{x^*, u^*\}$. Условимся также в дальнейшем задачу (3) называть задачей *верхнего уровня*, а задачу (2) будем называть задачей *нижнего уровня*. Пару задач *верхнего и нижнего уровней* будем называть *двухуровневой системой задач*, или *задачей параметрического программирования*.

В этом случае упрощение процедуры решения исходной задачи (1) может быть обусловлено как очевидно меньшей размерностью задач (2) и (3) (по сравнению с размерностью задачи (1)), так и возможным упрощением их постановок. Примером такого упрощения является случай, когда за счет подходящего выбора параметризуемых переменных задача нижнего уровня оказывается линейной.

Проблемы и подход к решению двухуровневой системы задач с помощью метода штрафных функций рассматриваются в [4]. Для решения задачи используется достаточно гладкая по своим аргументам штрафная функция $P(r,\alpha)$ такая, что

$$\lim_{r\to+0} P(r,\alpha) = \begin{cases} +\infty, & \alpha < 0, \\ 0, & \alpha \geqslant 0, \end{cases} \tag{4}$$

и удовлетворяющая неравенствам

$$\frac{\partial P}{\partial \alpha} \leqslant 0 \text{ и } \frac{\partial^2 P}{\partial \alpha^2} \geqslant 0 \; \forall \; \alpha. \tag{5}$$

При использовании метода штрафных функций задача нижнего уровня (2) является стандартной задачей безусловной минимизации в E^n:

$$x^+(r, u) = \arg \min_{x \in E^n} \varepsilon(r, x, u), \tag{6}$$

где вспомогательная функция метода гладких штрафных функций имеет вид [3]:

$$\varepsilon(r, x, u) = f(x, u) + \sum_{s=1}^{m} P[r, g_s(x, u)]. \tag{7}$$

Итерационный процесс решения этой задачи рассматривается также в [3].

На верхнем уровне выполняется оптимизация функции:

$$E(r, u) = \varepsilon(r, x^+(r, u), u), \quad u \in E^l. \tag{8}$$

Обозначим

$$\tilde{u}(r) = \arg \min_{u \in E^l} E(r, u). \tag{9}$$

Для практического решения задач верхнего уровня может быть использована стандартная [1] итеративная схема построения последовательных приближений к искомому решению $\tilde{u}(r)$ в пространстве E^l, k-й шаг которой состоит из следующего набора операторов:

1) для некоторого исходного приближения u_k находится решение задачи (6): $x^+(r, u_k)$;
2) затем вычисляются улучшающее направление w_k в пространстве параметров E^l и величина шага σ_k по этому направлению;
3) выполняется переход к точке u_{k+1} по формуле
$$u_{k+1} = u_k + \sigma_k w_k; \tag{10}$$
4) если в u_{k+1} условие окончания итерационного процесса не выполнено, то u_k присваивается значение u_{k+1} и делается переход к пункту 1.

В случае сходимости данной процедуры будет найдено $\{x^+(r, \tilde{u}), \tilde{u}(r)\}$ – приближенное локальное решение задачи (1). При этом будут справедливы предельные соотношения [5]:

$$\lim_{r \to +0} \tilde{u}(r) = u^*,$$
$$\lim_{r \to +0} x^+(r, \tilde{u}(r)) = x^*.$$

2. Условия сходимости итерационного процесса

Теперь рассмотрим вопрос о сходимости итерационного процесса (10). Предварительно докажем ряд вспомогательных утверждений.

Лемма 1. *Пусть функции $\varphi(u)$ и $\psi(u)$ удовлетворяют условию Липшица на компакте $M \subset E^l$:*

$$\|\varphi(u_1) - \varphi(u_2)\| \leqslant L_\varphi \|u_1 - u_2\| \quad \forall \; u_1, u_2 \in M,$$
$$\|\psi(u_1) - \psi(u_2)\| \leqslant L_\psi \|u_1 - u_2\| \quad \forall \; u_1, u_2 \in M.$$

Тогда функции

1) $\varphi(u) + \psi(u)$,

2) $\varphi(u) \cdot \psi(u)$

также удовлетворяют условию Липшица на множестве M.

Доказательство.

Рассмотрим функцию $\varphi(u)+\psi(u)$. Для нее будет справедливо

$$\begin{aligned}&\|\ \varphi(u_1)+\psi(u_1)-\varphi(u_2)-\psi(u_2)\| \leqslant \|\ \varphi(u_1)-\varphi(u_2)\|+\|\ \psi(u_1)-\psi(u_2)\| \leqslant\\ &\quad \leqslant L_\varphi\|\ u_1-u_2\|+L_\psi\|\ u_1-u_2\|=(L_\varphi+L_\psi)\|\ u_1-u_2\|.\end{aligned}$$

Следовательно, она удовлетворяет условию Липшица на M. Теперь рассмотрим произведение $\varphi(u)\cdot\psi(u)$:

$$\begin{aligned}&\|\ \varphi(u_1)\psi(u_1)-\varphi(u_2)\psi(u_2)\| =\\ &\quad=\|\ \varphi(u_1)\psi(u_1)-\varphi(u_1)\psi(u_2)+\varphi(u_1)\psi(u_2)-\varphi(u_2)\psi(u_2)\| =\\ &\quad=\|\ \varphi(u_1)(\psi(u_1)-\psi(u_2))+\psi(u_2)(\varphi(u_1)-\varphi(u_2))\| \leqslant\\ &\quad\leqslant \|\ \varphi(u_1)\|\cdot\|\ \psi(u_1)-\psi(u_2)\|+\|\psi(u_2)\|\cdot\|\ \varphi(u_1)-\varphi(u_2)\|.\end{aligned} \tag{11}$$

Если функция удовлетворяет условию Липшица, то она непрерывна. Так как M — компакт, то $\varphi(u)$ и $\psi(u)$ ограничены на M. Следовательно, существуют такие числа C_φ и C_ψ, что

$$\begin{aligned}&\|\ \varphi(u)\| \leqslant C_\varphi \quad \forall\ u\in M,\\ &\|\ \psi(u)\| \leqslant C_\psi \quad \forall\ u\in M.\end{aligned} \tag{12}$$

Из (11) и (12) получаем

$$\|\ \varphi(u_1)\psi(u_1)-\varphi(u_2)\psi(u_2)\| \leqslant (C_\varphi L_\psi+C_\psi L_\varphi)\|\ u_1-u_2\|.$$

■

Лемма 2. *Пусть функция* $\varphi(u)$

1) *непрерывна на отрезке* $[a,b]\in R$,

2) *дифференцируема на отрезке* $[a,b]$, *за исключением точек* $x_1,x_2,...,x_n\in[a,b]$,

3) $\exists\ \varphi'(x_k\pm 0)\in R$.

Тогда $\varphi(u)$ *удовлетворяет условию Липшица на отрезке* $[a,b]$:

$$|\ \varphi(u_1)-\varphi(u_2)| \leqslant L|u_1-u_2| \quad \forall\ u_1,u_2\in[a,b], \tag{13}$$

и

$$|\ \varphi'(u\pm 0)| \leqslant L \quad \forall\ u\in[a,b]. \tag{14}$$

Доказательство.

Из условий (2) и (3) леммы следует ограниченность лево- и правосторонних производных функции $\varphi(u)$ на $[a,b]$, значит, существует константа L, удовлетворяющая (14). Докажем (13). Без ограничения общности можно считать, что $u_1<x_1<x_2<...<x_n<u_2$. Применим формулу конечных приращений Лагранжа на отрезках $[u_1,x_1]$, $[x_1,x_2]$, ..., $[x_n,u_2]$:

$$\begin{aligned}&|\ \varphi(u_1)-\varphi(u_2)|=|\ [\varphi(u_1)-\varphi(x_1)]+[\varphi(x_1)-\varphi(x_2)]+...+[\varphi(x_n)-\varphi(u_2)]\ | \leqslant\\ &\leqslant|\ \varphi(u_1)-\varphi(x_1)|+|\ \varphi(x_1)-\varphi(x_2)|+...+|\ [\varphi(x_n)-\varphi(u_2)|=\\ &=|\ \varphi'(c_1)\ |\cdot|\ u_1-x_1\ |+|\ \varphi'(c_2)\ |\cdot|\ x_1-x_2\ |+...+|\ \varphi'(c_{n+1})\ |\cdot|\ x_n-u_2\ | \leqslant\\ &\qquad\leqslant L|\ u_1-x_1\ |+L|\ x_1-x_2\ |+...+L|\ x_n-u_2\ |=\\ &\qquad\quad=L|\ u_1-u_2\ |.\end{aligned}$$

■

Следствие 1. *Если в условии леммы 2 предположить, что количество точек разрыва производной счетно и* $\lim\limits_{n\to+\infty} x_n = u_2$, *то утверждение леммы останется верным. Это следует из непрерывности функции* $\varphi(u)$ *на отрезке* $[a,b]$ *и свойств пределов. Действительно, из*

$$| \varphi(u_1) - \varphi(x_n)| \leqslant L| u_1 - x_n |$$

при стремлении n *к бесконечности следует*

$$| \varphi(u_1) - \varphi(u_2)| \leqslant L| u_1 - u_2 |.$$

Лемма 3. *Пусть функция* $g(u)$ *непрерывно дифференцируема в выпуклой области* $G \subset E^l$. *Рассмотрим функцию*

$$h(u) = \Theta(-g(u))g(u),$$

где $\Theta(\alpha)$ *– стандартная функция Хевисайда. Если для любых* $u_1, u_2 \in G$ *граница множества, являющегося пересечением множества нулей функции* $g(u)$ *и отрезка, соединяющего точки* u_1 *и* u_2, *состоит из счетного количества точек, то функция* $h(u)$ *удовлетворяет условию Липшица в области* G:

$$\| h(u_1) - h(u_2)\| \leqslant L\|u_1 - u_2\| \quad \forall\, u_1, u_2 \in G.$$

Доказательство.

Отметим, что функция $g(u)$ в условии леммы — это некоторая $g_s(x^+(r,u),u)$ из задачи (3). Из условий леммы следует, что функция $h(u)$ непрерывна в области G, а производная $h'(u)$ существует всюду, кроме точек, являющихся нулями функции $g(u)$, и ограничена:

$$\exists\, L : \| h'(u)\| \leqslant L$$

для всех u, где производная определена.

Рассмотрим пересечение множества нулей функции $g(u)$ и отрезка, соединяющего точки u_1 и u_2. Занумеруем точки, принадлежащие границе этого пересечения, в порядке движения от u_1 к u_2: $x_1, x_2, \ldots, x_n, \ldots$ Без ограничения общности можно считать, что $\lim\limits_{n\to+\infty} x_n = u_2$.

Теперь применим к функции одной переменной

$$\varphi(t) = h(u_1 + t(u_2 - u_1)), \quad 0 \leqslant t \leqslant 1,$$

на отрезке $[0,1]$ следствие из леммы 2:

$$\| h(u_1) - h(u_2)\| \leqslant L\|u_1 - u_2\| \quad \forall\, u_1, u_2 \in G.$$

■

Теперь рассмотрим условия сходимости итерационного процесса (10) при решении задачи верхнего уровня (задачи безусловной минимизации (8)). В общем случае минимизируемая функция верхнего уровня (8) является невыпуклой, поэтому представляют интерес теоремы сходимости для минимизации как выпуклых, так и невыпуклых функций (см. [1]). В этих теоремах при условии существования локального минимума предъявляются следующие требования:

1) функция (8) должна быть дважды непрерывно дифференцируема, и матрица её вторых производных локально положительно определена (при использовании методов 2-го порядка, например, метода Ньютона [3]);

2) градиент функции (8) должен удовлетворять условию Липшица (при использовании методов 1-го порядка, например, градиентного поиска);

3) величина шага определяется методом дробления или одномерной минимизации [1].

В первую очередь отметим, что функция (8) является дважды непрерывно дифференцируемой, если дважды непрерывно дифференцируемыми по своим аргументам являются функции $f(x,u)$, $g_s(x,u)$, $s=[1,m]$ и $P(r,\alpha)$ [4], и является также локально выпуклой в окрестности решения (9) [5]. Сформулируем теперь условия, которым должны удовлетворять функции $f(x,u)$, $g_s(x,u)$, $s=[1,m]$ и $P(r,\alpha)$, чтобы функция (8) удовлетворяла условию Липшица.

Теорема 1. *Пусть функции $f(x,u), g_s(x,u), s=[1,m]$ и $P(r,\alpha)$ дважды дифференцируемы на компакте $M \subset E^l$. Тогда градиент функции (8) удовлетворяет условию Липшица на множестве M.*

Доказательство.

Утверждение теоремы следует из леммы 1, а также того, что градиент дважды дифференцируемой функции удовлетворяет условию Липшица на множестве M и дифференцируемости функции $x^+(r,u)$ [4].

■

Теорема 2. *Пусть функции $f(x,u), g_s(x,u), s=[1,m]$ дважды непрерывно дифференцируемы в замыкании ограниченной выпуклой области $G \subset E^l$, $P(r,\alpha)$ – квадратичная штрафная функция:*

$$P(r,\alpha) = \begin{cases} \frac{\alpha^2}{2r}, & \alpha < 0, \\ 0, & \alpha \geqslant 0. \end{cases}$$

Множество нулей функции $g_s(x^+(u),u)$ удовлетворяет условиям леммы 3 для любого s. Тогда градиент функции (8) удовлетворяет условию Липшица на множестве G.

Доказательство.

Утверждение теоремы следует из леммы 1 и леммы 3.

■

3. Заключение

Таким образом, рассмотрен итерационный процесс решения задачи параметрического программирования (2) – (3). Обоснованы условия сходимости двух классов численных методов безусловной минимизации для целевой функции (8) в пространстве параметров.

Литература

1. *Сухарев А.Г., Тимохов А.В., Федоров В.В.* Курс методов оптимизации. — М.: Физматлит, 2005.
2. *Соколов А.В., Токарев В.В.* Методы оптимальных решений. Т. 1. — М.: Физматлит, 2011.
3. *Фиакко А., Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной оптимизации. — М.: Мир, 1972.
4. *Марковцев Д.А., Умнов Е.А.* Использование метода штрафных функций для решения задач параметрического программирования // Динамика линейных и нелинейных систем. — М.: ИСА РАН, 2006.
5. *Марковцев Д.А.* Об обосновании метода параметризации в задачах математического программирования // Современные проблемы фундаментальной и прикладной математики. — М.: МФТИ, 2012.

Поступила в редакцию 06.02.2012.