Преподавание математики и компьютерных наук в классическом университете: материалы 2-й научно-методической конференции преподавателей математического факультета и факультета информатики и вычислительной техники Ярославского государственного университета им. П. Г. Демидова / Отв. ред. М. В. Невский; Яросл. гос. ун-т им. П. Г. Демидова. – Ярославль, 2007. – 175 с.
ISBN 5-8397-0517-9 (978-5-8397-0517-3)

стр. 81–86.

Правило множителей Лагранжа в курсе математического анализа

В.С. Климов, А.Ю. Ухалов

В работе обсуждается методика изложения темы "Правило множителей Лагранжа"в курсе математического анализа. Приводится доказательство теоремы Лагранжа, не опирающееся на теорему о неявной функции.

Библиография: 1 название.

В курсе математического анализа теория функций одной переменной, как правило, излагается с достаточной полнотой. При переходе же к функциям многих переменных изложение существенно осложняется громоздкими выкладками. Даже когда основная идея утверждения проста, полное изложение доказательства часто оказывается технически сложным и его трудно привести целиком на лекции.

Лекторы вынуждены идти на упрощение курса - опускать важные теоремы или приводить их без доказательства. Одно такое упрощение влечет за собой другие и в результате страдает целостность курса.

В частности, часто сокращению подвергается изложение круга вопросов, связанных с теоремой о неявной функии. Сама теорема о неявной функции для многомерного случая, как правило, приводится без доказательства. В результате, оказываются плохо обоснованными и другие важные темы, например, теорема о правиле множителей Лагранжа.

Доказательство этой теоремы, приводимое обычно в учебниках и лекционных курсах, существенно опирается на теорему о неявной функции. Логичным становится опустить и доказательство теоремы Лагранжа об условном экстремуме, т.к. из под фундамента ее доказательства уже вынут главный камень.

Оказывается, однако, что для многих утверждений многомерного анализа существуют доказательства, использующие минимальное количество фактов из предшествующего материала.

В настоящей заметке приводится элементарное доказательство теоремы о правиле множителей Лагранжа, опирающееся только на теорему Вейерштрасса.

Пусть на открытом множестве U пространства R^n определены два конечных семейства функций $f_i : U \to R$ $(i \in I)$ и $f_i : U \to R$ $(i \in I_0)$ (I,I_0 - конечные непересекающиеся множества). Пусть множество

$$E = \{x \in U | f_i(x) \leq 0 \quad (i \in I), \quad f_i(x) = 0 \quad (i \in I_0)\}$$

непусто. Рассматривается задача

$$f_0(x) \to \min, \tag{1}$$

$$f_i(x) \leq 0 \quad (i \in I), \qquad f_i(x) = 0 \quad (i \in I_0), \qquad x \in U. \tag{2}$$

Здесь $f_0 : U \to R$ — действительная функция на множестве U. Элемент $\hat{x}$ из E называют локальным решением задачи (1), (2), если найдется такое $\delta > 0$, что $f_0(x) \geq f_0(\hat{x})$ $\forall x \in E \cap B(\hat{x}, \delta)$. Здесь и далее $B(\hat{x}, R) = \{x \in R^n, \|x - \hat{x}\| \leq R\}$.

Введем в рассмотрение функцию

$$g(t) = \begin{cases} \frac{t^2}{2}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Функция $g(t)$ дифференцируема и

$$g'(t) = \begin{cases} t, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Производная $g'(t)$ непрерывна на всей оси. Положим

$$\Phi(x) = \frac{1}{2} \sum_{i \in I_0} f_i^2(x) + \sum_{i \in I} g[f_i(x)].$$

Как нетрудно видеть, равенство $\Phi(x) = 0$ эквивалентно включению $x \in E$. Действительно, если $x \in E$, то $f_i(x) = 0 (i \in I_0)$, $g[f_i(x)] = 0 (i \in I)$, поэтому $\Phi(x) = 0$. Верно и обратное. Если функции f_i $(i \in I_1 = I \cup$

I_0) непрерывно дифференцируемы, то функция $\Phi(x)$ также непрерывно дифференцируема.

Теорема 1 *Пусть $\hat{x}$ - локальное решение задачи (1), (2) и функции $f_i (i \in \{0\} \cup I_1)$ непрерывно дифференцируемы в окрестности точки $\hat{x}$. Тогда найдутся такие числа $\hat{\lambda}_i$, не все равные нулю, что*

$$\hat{\lambda}_0 f_0'(\hat{x}) + \sum_{i \in I_1} \hat{\lambda}_i f_i'(\hat{x}) = 0, \tag{3}$$

$$\hat{\lambda}_i f_i'(\hat{x}) = 0 \quad (i \in I_1), \tag{4}$$

$$\hat{\lambda}_0 \geq 0, \hat{\lambda}_i \geq 0 \quad (i \in I). \tag{5}$$

Числа $\hat{\lambda}_i$ называют множителями Лагранжа, а саму теорему 1 — правилом множителей Лагранжа. Ясно, что вместе с набором $\hat{\lambda}_i$ соотношениям (3)-(5) удовлетворяет набор $t\hat{\lambda}_i$ $(t > 0)$ — набор множителей Лагранжа определяется неоднозначно. Наиболее важным необходимым условием оптимальности является равенство (3), называемое условием стационарности. Соотношение (4) называют условием дополнительной нежесткости. Если $f_i(\hat{x}) < 0$ для некоторого индекса i из I, то ограничение $f_i(x) \leq 0$ называют мягким в точке $\hat{x}$, в этом случае из (4) следует равенство $\hat{\lambda}_i = 0$.

Доказательство

Каждому натуральному числу N поставим в соответствие функцию

$$\Phi^N(x) = f_0(x) + N\Phi(x) + \frac{|x - \hat{x}|^2}{2}$$

и экстремальную задачу

$$\Phi^N(x) \to \min, \qquad x \in B(\hat{x}, R). \tag{6}$$

Число $R > 0$ считаем настолько малым, что $B(\hat{x}, R) \subset U$, $f_0(x) \geq f_0(\hat{x})$ $\forall x \in E \cap B(\hat{x}, R)$, функции f_i непрерывно дифференцируемы на шаре $B(\hat{x}, R)$.

Существование решения x^N задачи (6) следует из теоремы Вейерштрасса. Справедливо неравенство $\Phi^N(x^N) \leq \Phi^N(\hat{x})$ или, более подробно,

$$f_0(x^N) + N\Phi(x^N) + \frac{|x^N - \hat{x}|^2}{2} \leq f_0(\hat{x}). \tag{7}$$

В частности, из (7) вытекает оценка $N\Phi(x^N) \leq f_0(\hat{x}) - f_0(x^N)$. Последовательность $x^N \in B(\hat{x}, R)$. Если $\bar{x}$ - предельная точка этой последовательности, то $\Phi(\bar{x}) = 0$. Из (7) следует неравенство

$$f_0(\bar{x}) + \frac{|\bar{x} - \hat{x}|^2}{2} \leq f_0(\hat{x}).$$

С другой стороны, $\bar{x} \in E \cap B(\hat{x}, R)$, следовательно $f_0(\bar{x}) \geq f_0(\hat{x})$, поэтому $\bar{x} = \hat{x}$. Всякая предельная точка последовательности x^N совпадает с $\hat{x}$. Это означает, что $x^N \to \hat{x}$. При больших N элемент x^N есть внутренняя точка шара $B(\hat{x}, R)$.

Так как x^N — внутренняя точка шара $B(\hat{x}, R)$ и x^N минимизирует функцию Φ^N на этом шаре, то x^N — стационарная точка функции Φ^N: $(\Phi^N)'(x^N) = 0$ т.е.

$$f_0'(x^N) + N \sum_{i \in I_0} f_i(x^N) f_i'(x^N) + N \sum_{i \in I} g'(f_i(x^N)) f_i'(x^N) + (x^N - \hat{x})^T = 0$$

Разделим это равенство на положительное число K^N, так чтобы получилось соотношение

$$\lambda_0^N f_0'(x^N) + \sum_{i \in I_1} \lambda_i^N f_i'(x^N) + \frac{(x^N - \hat{x})^T}{K^N} = 0, \tag{8}$$

в котором

$$(\lambda_0^N)^2 + \sum_{i \in I_1} (\lambda_i^N)^2 = 1, \quad \forall N. \tag{9}$$

Таким образом,

$$\lambda_0^N = \frac{1}{K^N}, \quad \lambda_i^N = \frac{N}{K^N} f_i(x^N) \quad (i \in I_0), \quad \lambda_i^N = \frac{N}{K^N} g'(f_i(x^N)) \quad (i \in I), \tag{10}$$

$$K^N = \left(1 + N^2 \sum_{i \in I_0} f_i^2(x^N) + N^2 \sum_{i \in I} |g'(f_i(x^N))|^2 \right)^{\frac{1}{2}} \geq 1.$$

Каждая из последовательностей λ_i^N $(N = 1, 2, \dots)$ ограничена, поэтому, не уменьшая общности, можно считать, что $\lambda_i^N \to \hat{\lambda}_i$, $\forall i \in \{0\} \cup I_1$. Переходя в (9) к пределу, получаем равенство

$$\hat{\lambda}_0^2 + \sum_{i \in I_1} \hat{\lambda}_i^2 = 1.$$

Из этого равенства следует, что хотя бы одно из чисел $\hat{\lambda}_i$ отлично от нуля.

Полученный набор $\hat{\lambda}_i$ является искомым. Действительно, устремляя в (8) N к ∞, получаем условие стационарности (3) (используются соотношения $x^N \to \hat{x}$, $\lambda_i^N \to \hat{\lambda}_i$).

Если $f_i(\hat{x}) = 0$, то $\hat{\lambda}_i f_i(\hat{x}) = 0$ и условие дополнительной нежесткости выполнено. Если $f_i(\hat{x}) < 0$, то $f_i(x^N) < 0$ $\forall N >> 1$, следовательно,

$$\lambda_i^N = \frac{N}{K^N} g'(f(x^N)) = 0,$$

поэтому $\hat{\lambda}_i = 0$, т.е. (4) справедливо и в этом случае. В силу (10) $\lambda_0^N \geq 0$, $\lambda_i^N \geq 0$ $(i \in I)$, поэтому $\hat{\lambda}_0 \geq 0$, $\hat{\lambda}_i \geq 0$ $\forall i \in I$. Теорема доказана. $\square$

Варианты теоремы 1 сохраняются, если одно из множеств I_0, I пусто.

Теорема 2 *Пусть $\hat{x}$ - локальное решение задачи*

$$f_0(x) \to \min,$$

$$f_i(x) \leq 0 \quad (i \in I), \quad x \in U.$$

и функции f_i непрерывно дифференцируемы в окрестности точки $\hat{x}$. Тогда найдутся такие числа $\hat{\lambda}_i \geq 0$, не все равные 0, что выполнены условия (3), (4) с $I_1 = I$.

Теорема 3 *Пусть $\hat{x}$ - локальное решение задачи*

$$f_0(x) \to \min,$$

$$f_i(x) = 0 \quad (i \in I_0), \quad x \in U.$$

и функции f_i непрерывно дифференцируемы в окрестности точки $\hat{x}$. Тогда найдутся такие числа $\hat{\lambda}_i \geq 0$, не все равные 0 одновременно, что выполнено условие (3) с $I_1 = I_0$.

Теоремы 1–3 выражают правило множителей Лагранжа. Во всех теоремах $\hat{\lambda}_0 \geq 0$. Если $\hat{\lambda}_0 > 0$, то разделив на $\hat{\lambda}_0$, получаем снова набор множителей Лагранжа, для которого $\hat{\lambda}_0 = 1$.

Лагранж считал, что всегда можно взять $\hat{\lambda}_0 = 1$. Однако это законно, лишь если $\hat{\lambda}_0 > 0$. Любое предположение, гарантирующее неравенство $\hat{\lambda}_0 > 0$, называют условием регулярности. В экстремальной задаче

$$f_0(x) = x_2 \to \min, \quad f_1(x_1, x_2) = x_2^3 - x_1^2 = 0$$

решение $\hat{x} = (0, 0)^T$, условие регулярности не выполнено. Действительно,

$$f_0'(\hat{x}) = (0, 1), \quad f_1'(\hat{x}) = (0, 0)$$

и равенство $\hat{\lambda}_0 f_0'(\hat{x}) + \hat{\lambda}_1 f_1'(\hat{x}) = 0$ возможно лишь при $\hat{\lambda}_0 = 0$.

Основу доказательства теоремы 1 составляет широко используемый в оптимизации метод штрафных функций ([1], гл. 8). Там же приведены достаточно обозримые условия регулярности задачи (1), (2).

Список литературы

[1] *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.