Доклады Академии Наук СССР
1940. Том XXVIII, № 3

МАТЕМАТИКА

Л. В. КАНТОРОВИЧ

ОБ ОДНОМ ЭФФЕКТИВНОМ МЕТОДЕ РЕШЕНИЯ НЕКОТОРЫХ КЛАССОВ ЭКСТРЕМАЛЬНЫХ ПРОБЛЕМ

(Представлено академиком С. Н. Бернштейном 8 V 1940)

Во многих экстремальных проблемах точка экстремума оказывается непременно не внутри, а на границе основной области, рассматриваемой в данном эвклидовом или функциональном пространстве. Фактическое решение таких задач часто оказывается затруднительным. Так, например, простейшая из задач такого рода об экстремуме линейной функции в многограннике имеет весьма простое теоретическое решение— для нахождения экстремума достаточно сравнить значения этой функции в вершинах многогранника; однако, если число этих вершин очень велико, задача оказывается практически неразрешимой.

Мы хотим здесь указать один метод решения задач о краевом экстремуме, имеющий известное сходство с методом Лагранжа разыскания относительных экстремумов. Метод этот, хотя и построен на весьма простых соображениях, дает для многих задач гораздо более эффективный прием нахождения решения, чем обычные методы.

В § 1 мы излагаем этот метод в общем виде, а в § 2 приводим примеры задач, к которым этот метод может быть применен.

§ 1. Теорема 1. *Пусть E линейное нормированное пространство, $A \subset E$ выпуклое и слабо-компактное множество. $F(x)$ функционал, определенный и конечный в A и обладающий следующими свойствами: 1) F слабо-полунепрерывен сверху: $x_n \dashrightarrow x_n$ влечет $\overline{\lim} F(x_n) \leqslant F(x)$, 2) $F(x)$ не достигает максимума внутри A, т. е. если x внутренняя точка A, в любой ее окрестности имеется точка x', для которой $F(x') > F(x)$.*

При этих условиях существует точка x_0 на границе A, в которой $F(x)$ достигает максимума, и существует линейный функционал $f_0(x)$, который достигает максимума в A также в точке x_0.

Доказательство. Пусть M точная верхняя граница $F(x)$ на A. Найдутся такие x_n, что $F(x_n) > M - \frac{1}{n}$. Из последовательности $\{x_n\}$ можно выбрать частичную слабо-сходящуюся $x_{n_k} \dashrightarrow x_0$. Вследствие условия 1) $F(x_0) = M$ и вследствие 2), x_0 граничная точка A. По теореме S. Mazur'a [1] существует линейный функционал $f_0(x)$ такой, что множество A лежит целиком по одну сторону от плоскости $f_0(x) = f_0(x_0)$. Очевидно, выбором знака у f_0 можно добиться того, чтобы для $x \in A$ выполнялось именно неравенство $f_0(x) \leqslant f_0(x_0)$. Тогда точка x_0 и линейный функционал f_0 и суть те, существование которых утверждается в теореме.

Основное следствие. Пусть f линейный функционал. Обозначим через H_f множество тех точек A, где он достигает максимума, и через $p(f)$ — максимум $F(x)$ на множестве H_f, тогда

$$M = \max_{x \in A} F(x) = \max_{f} \{p(f)\}. \tag{1}$$

Действительно, с одной стороны, всегда $p(f) \leqslant M$, с другой — имеем, согласно теореме 1, $p(f_0) = F(x_0) = M$.

Равенство (1) и служит исходным для предлагаемого метода нахождения краевого экстремума. Именно, рекомендуется поступать следующим образом: выбирая некоторый линейный функционал f, определяем множество H_f и затем $p(f)$. Далее, варьируем f так, чтобы значение $p(f)$ увеличивалось, и так постепенно приближаемся к максимуму. Существенное преимущество этого пути, по сравнению с непосредственным варьированием точки x, состоит в том, что здесь обеспечено движение по границе области A. Кроме того, при этом способе часто удается легче определить правильное направление варьирования.

Возникает, прежде всего, вопрос о том, как проверить, решая задачу по этому методу, то, что мы действительно пришли к максимуму. Такой критерий дает для многих случаев следующая теорема.

Теорема 2. *Пусть выполнены условия теоремы 1 и, кроме того, известно, что множество $F(x) \geqslant C$ выпукло для любого C. Тогда, если найдены точка x_0 и линейный функционал f_0 такие, что* 1) $f_0(x) \leqslant f_0(x_0)$ *для $x \in A$ и* 2) *на плоскости $f_0(x) = f_0(x_0)$ справедливо неравенство $F(x) \leqslant F(x_0)$, либо* 2′) *на поверхности $F(x) = F(x_0)$ справедливо неравенство $f_0(x) \geqslant f_0(x_0)$, то функционал $F(x)$ достигает максимума на A в точке x_0. При этом существование требуемых здесь f_0 и x_0 всегда обеспечено.*

Доказательство. В самом деле, из условий теоремы вытекает, что плоскость $f_0(x) = f_0(x_0)$ разделяет выпуклые множества A и $\mathcal{E}(F(x) \geqslant F(x_0))$, а потому ясно, что во всех точках множества A будет справедливо неравенство $F(x) \leqslant F(x_0)$, т. е. x_0—точка максимума для $F(x)$. Наконец, существование f_0 и x_0 всегда обеспечено, так как, по теореме Eidelheit'a([2]), существует плоскость, отделяющая два выпуклых множества.

Отметим, что в условиях теоремы 2 можно уточнить то, что было сказано о направлении вариации f. Именно, если функционал f_0 и точка x_0 в некотором приближении еще не искомые, то плоскость $f_0(x) = f_0(x_0)$ будет касательной для A, но не для множества $\mathcal{E}(F(x) \geqslant F(x_0))$. Проведем через x_0 касательную плоскость к этому множеству $f_1(x) = f_1(x_0)$. Тогда при нахождении следующего приближения нужно функционал f_0 заменить на $f_0 + \varepsilon f_1$.

Сделаем теперь несколько замечаний, показывающих, что данный метод применим не только в случаях, предусмотренных теоремами 1 и 2.

Замечание 1. Если существование конечного экстремума для F известно из других соображений, то требования слабой компактности A и слабой полунепрерывности F являются излишними.

Замечание 2. Условие выпуклости множества A может быть также ослаблено. Именно, достаточно потребовать, чтобы для каждой граничной точки x множества A существовал такой функционал f, что для него

$$\overline{\lim_{\substack{x' \to x \\ x' \in A}}} \frac{f(x' - x)}{\|x' - x\|} = 0. \tag{2}$$

В этом случае надлежит только за множество H_f брать совокупность точек x, в которых выполняется условие (2).

§ 2. Приведем теперь примеры задач, к которым данный метод приложим.

I. Функции $\alpha_i(t)$ $(i=1, 2, \dots, m)$ определены и интегрируемы (L) в промежутке $[a, b]$. Найти измеримые функции $h_i(t)$ $(i=1, 2, \dots, m)$ так, чтобы 1) $h_i(t) \geqslant 0$ $(i=1, 2, \dots, m)$, 2) $h_1(t)+h_2(t)+\dots+h_m(t)=1$, 3) $\min\limits_i \int\limits_a^b \alpha_i(t) h_i(t)\,dt$ имел максимальное возможное значение.

Теоремы 1 и 2 могут быть применены, если за A принять множество систем чисел $(z_1, z_2, \dots, z_m)$:

$$z_i = \int\limits_a^b \alpha_i(t) h_i(t)\,dt,$$

где $h_1, h_2, \dots, h_m$—любые функции, удовлетворяющие условиям 1) и 2), и положить

$$F(z_1, z_2, \dots, z_m) = \min(z_1, z_2, \dots, z_m).$$

Воспользовавшись теоремами 1 и 2, можно установить, что существуют множители $\lambda_1, \lambda_2, \dots, \lambda_m$ такого рода, что линейный функционал

$$\sum_{i=1}^{m} \lambda_i z_i = \int\limits_a^b \sum_{i=1}^{m} \lambda_i \alpha_i(t) h_i(t)\,dt \tag{3}$$

достигает максимума одновременно с F. Но относительно функционала (3) ясно, что, если система функций $h_1, h_2, \dots, h_m$ доставляет максимум ему, то почти везде должно быть выполнено условие:

$$\left.\begin{aligned} &h_i(t)=0, \\ \text{если}\quad &\lambda_i\alpha_i(t) h_i(t) < \max_j \lambda_j \alpha_j(t) h_j(t). \end{aligned}\right\} \tag{4}$$

Если λ_i известны, то из последнего условия $h_i(t)$ могут быть легко определены и притом так, чтобы $h_i(t)=0; 1$ и чтобы

$$\int\limits_a^b \alpha_1(t) h_1(t)\,dt = \dots = \int\limits_a^b \alpha_m(t) h_m(t)\,dt. \tag{5}$$

Наоборот, если числа λ_i таковы, что можно определить функции $h_i(t)$, удовлетворяющие условиям (4) и (5), то функции $h_i(t)$ доставляют экстремум функционалу F. Таким образом, разыскание системы функций заменяется разысканием системы чисел $\lambda_1, \lambda_2, \dots, \lambda_m$. Что касается последних, то для их нахождения может быть указан удобный метод последовательных приближений.

II. Задача I может быть решена и если имеется дополнительное ограничительное условие вида

$$R = \int\limits_a^b \sum_{i=1}^{m} \beta_i(t) h_i(t)\,dt \leqslant B.$$

Действительно, достаточно принять теперь

$$F(z_1, z_2, \dots, z_m, R) = \min_i z_i + \delta(R),$$

где $\delta(R)=0$, если $R \leqslant B$ и $\delta(R) = -\infty$, если $R > B$. Применение теорем 1 и 2 показывает здесь, что экстремум F достигается одновременно с экстремумом функционала $\sum \lambda_i z_i - \mu R$.

III. Функция $\alpha(x, y) > 0$ непрерывна в прямоугольнике $[a, b; c, e]$. Требуется определить функцию элементарного интервала $\Phi(d) \geqslant 0$ так, чтобы

$$1) \int\limits_a^b \Phi(d_x) = 1, \qquad 2) \int\limits_c^e \alpha(x, y)\, \Phi(d_y) = C,$$

и постоянная C имела максимальное возможное значение. (Условие 1) означает точнее, что для любых $c \leqslant \gamma < \delta \leqslant e$ должно быть $\Phi[(a, b; \gamma, \delta)] = \delta - \gamma$.) Если рассмотреть пространство функций $z(x) = \int\limits_c^e \alpha(x, y)\, \Phi(d_y)$, то решение задачи состоит в отыскании максимума функционала $F(z) = \min\limits_{a \leqslant x \leqslant e} z(x)$. Опять метод применим и позволяет провести исследование задачи.

Кроме целого ряда других задач, аналогичных задачам I, II, III, данный метод применим и в некоторых задачах более обычных типов.

Приведем два примера подобных задач из теории апроксимаций.

IV. Приближенное решение несовместных уравнений. Эта задача ставится так (см., например [3], статья 4). Определить числа $x_1, x_2, \ldots, x_n$ так, чтобы, полагая

$$\begin{gathered} a_{1,1} x_1 + \ldots + a_{1,n} x_n - c_1 = z_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m,1} x_1 + \ldots + a_{m,n} x_n - c_m = z_m, \end{gathered}$$

мы имели бы минимум $\Phi(z_1, z_2, \ldots, z_m) = [|z_1|^p + \ldots + |z_m|^p]^{\frac{1}{p}}$.

Если $p = \infty$, то речь идет о наилучшей чебышевской апроксимации; если $p = 2$, то этот вопрос встречается и решается достаточно просто в методе наименьших квадратов Гаусса.

Положим $z_i^2 = y_i$ и $F(y_1, \ldots, y_m) = \Phi(\sqrt{y_1}, \ldots, \sqrt{y_m})$. Тогда, если учесть замечание 2, наш метод может быть применен. Применив его, убеждаемся, что экстремум задачи достигается одновременно с экстремумом функционала

$$\sum \lambda_i y_i = \sum \lambda_i z_i^2.$$

Такого рода задача об экстремуме решается легко. Что касается нахождения λ_i, то опять можно воспользоваться способом последовательных приближений. В частности, в чебышевском случае $(p = \infty)$ λ_i нужно варьировать, исходя из того, что если $|z_{i_0}|$ больше прочих $|z_i|$, соответствующее λ_{i_0} следует увеличить.

V. Те же соображения, что и в предыдущем случае, могут быть применены для нахождения полинома данной степени, дающего наилучшую апроксимацию некоторой функции в смысле Чебышева или Джэксона. Именно можно искать такой вес $p(x)$, что полином, дающий наилучшее среднее квадратическое приближение по этому весу, и будет искомым.

Ленинградский
государственный университет

Поступило
11 V 1940

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] S. Mazur, Studia Math., IV. [2] Eidelheit, Studia Math., VI. [3] Н. Ахиезер и М. Крейн, О некоторых проблемах теории моментов (1938).