1973 г. ОПТИМИЗАЦИЯ Выпуск 10(27)

УДК 519.8

# НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ МИНИМАКСА

В.Н. Малоземов

В работе описывается новая методика получения необходимых и достаточных условий минимакса.

1°. Пусть $\Omega$ -выпуклое множество некоторого линейного нормированного пространства, а $\mathcal{D}$ - метрический компакт. Задана функция $\mathcal{F}(X,Y)$, непрерывная по $X\in\Omega$ равномерно относительно $Y\in\mathcal{D}$, непрерывная по $Y\in\mathcal{D}$ при любом фиксированном $X\in\Omega$ и выпуклая по $X$ на $\Omega$ при каждом $Y\in\mathcal{D}$. Рассматривается минимаксная задача

$$\max_{Y\in\mathcal{D}} \mathcal{F}(X,Y) \longrightarrow \min_{X\in\Omega}, \qquad (1)$$

причем считается, что решение этой задачи существует. При сделанных предположениях справедлива

ТЕОРЕМА 1. Для того чтобы элемент $X^*\in\Omega$ был решением задачи (1), необходимо и достаточно, чтобы выполнялось равенство

$$[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)]=0, \qquad (2)$$

где $R(X^*)=\{Y\in\mathcal{D}\,|\,\mathcal{F}(X^*,Y)=\max\limits_{Z\in\mathcal{D}}\mathcal{F}(X^*,Z)\}$.

ДОКАЗАТЕЛЬСТВО. НЕОБХОДИМОСТЬ. Пусть $X^*\in\Omega$ – решение задачи (1) и $X$ – произвольный элемент из $\Omega$. Тогда при любых $\alpha\in(0,1)$ и $Y\in\mathcal{D}$ в силу выпуклости $\mathcal{F}(X,Y)$ по $X$ будет

$$\mathcal{F}(X^*+\alpha(X-X^*),Y)\le\mathcal{F}(X^*,Y)+\alpha[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)]. \qquad (3)$$

Введем обозначения

$$\varphi(X)=\max_{Y\in\mathcal{D}}\mathcal{F}(X,Y),$$

$$R_\varepsilon(X)=\{Y\in\mathcal{D}\,|\,\varphi(X)-\mathcal{F}(X,Y)\le\varepsilon\}.$$

Из (3) следует, что при любом фиксированном $\varepsilon>0$

$$\max_{Y\in R_\varepsilon(X^*)}\mathcal{F}(X^*+\alpha(X-X^*),Y)\le$$
$$\le\varphi(X^*)+\alpha\cdot\max_{Y\in R_\varepsilon(X^*)}[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)]$$

При малых $\alpha>0$ выполняется неравенство

$$\max_{Y\in\mathcal{D}}|\mathcal{F}(X^*+\alpha(X-X^*),Y)-\mathcal{F}(X^*,Y)|<\frac{\varepsilon}{2},$$

поэтому

$$R(X^*+\alpha(X-X^*))\subset R_\varepsilon(X^*).$$

Теперь имеем

$$\varphi(X^*+\alpha(X-X^*))\leq\varphi(X^*)+\alpha\cdot\max_{Y\in R_\varepsilon(X^*)}\left[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)\right].$$

Поскольку $X^*$ – точка минимума $\varphi(X)$ на $\Omega$, то из последнего неравенства следует, что

$$\max_{Y\in R_\varepsilon(X^*)}\left[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)\right]\geq 0. \qquad (4)$$

Отсюда

$$\max_{Y\in R(X^*)}\left[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)\right]\geq 0,\qquad \forall X\in\Omega. \qquad (5)$$

Действительно, возьмем последовательность положительных чисел $\{\varepsilon_\nu\}$, стремящуюся к нулю. В силу (4) найдутся $Y_\nu\in R_{\varepsilon_\nu}(X^*)$ такие, что

$$\mathcal{F}(X,Y_\nu)-\mathcal{F}(X^*,Y_\nu)\geq 0.$$

Так как все $Y_\nu$ принадлежат компакту $\mathcal{D}$, то из последовательности $\{Y_\nu\}$ может быть выделена сходящаяся подпоследовательность. Будем считать, что вся последовательность $\{Y_\nu\}$ сходится к элементу $Y^*\in\mathcal{D}$. Тогда очевидно, что $Y^*\in R(X^*)$ и

$$\mathcal{F}(X, Y^*) - \mathcal{F}(X^*, Y^*) \geqslant 0.$$

Тем более,

$$\max_{Y \in R(X^*)} [\mathcal{F}(X, Y) - \mathcal{F}(X^*, Y)] \geqslant 0.$$

Неравенство (5) доказано. Осталось заметить, что (5) эквивалентно (2).

ДОСТАТОЧНОСТЬ. Пусть выполнено (2) или, что то же самое, (5). Поскольку $\mathcal{F}(X^*, Y)$ при всех $Y \in R(X^*)$ равно $\varphi(X^*)$, то (5) можно переписать в другом виде

$$\max_{Y \in R(X^*)} \mathcal{F}(X, Y) \geqslant \varphi(X^*), \quad \forall X \in \Omega.$$

Учитывая, что $\varphi(X) \geqslant \max\limits_{Y \in R(X^*)} \mathcal{F}(X, Y)$, получаем $\varphi(X) \geqslant \varphi(X^*)$, $\forall X \in \Omega$, а это и означает, что $X^*$ – точка минимума функции $\varphi(X)$ на $\Omega$. Теорема доказана.

ЗАМЕЧАНИЕ. При доказательстве достаточности не использовалась ни выпуклость множества $\Omega$, ни выпуклость функции $\mathcal{F}(X, Y)$ по $X$.

$2^0$. Рассмотрим задачу наилучшего приближения непрерывной функции обобщенными полиномами:

$$\max_{Y \in \mathcal{D}} \left| f(Y) - \sum_{i=1}^{n} a_i u_i(Y) \right| \longrightarrow \min_{A \in \Omega}, \qquad (6)$$

где $\mathcal{D}$ – метрический компакт, $A = (a_1, \ldots, a_n)$, $\Omega \subset E^n$ –

замкнутое выпуклое множество. Функции $f(Y), u_1(Y), \ldots, u_n(Y)$ непрерывны на $\mathscr{D}$. Покажем, что из теоремы I следует известный критерий А.Н. Колмогорова для полинома наилучшего приближения.

Предварительно заметим, что задача (6) имеет решение. Введем обозначения

$$\Phi(A,Y)=\sum_{i=1}^{n} a_i u_i(Y),$$

$$\Delta(A,Y)=f(Y)-\Phi(A,Y),$$

$$\mathscr{F}(A,Y)=|\Delta(A,Y)|.$$

ТЕОРЕМА А.Н.КОЛМОГОРОВА. *Для того чтобы вектор $A^*\in\Omega$ был решением задачи (6), необходимо и достаточно, чтобы*

$$\inf_{A\in\Omega}\ \max_{Y\in R(A^*)}\left\{\left[\Phi(A^*,Y)-\Phi(A,Y)\right]\cdot \operatorname{sign}\Delta(A^*,Y)\right\}=0. \qquad (7)$$

ДОКАЗАТЕЛЬСТВО. Поскольку для функции $\mathscr{F}(A,Y)$ выполняются все условия, сформулированные в начале пункта I, то в силу теоремы I достаточно установить эквивалентность соотношений (7) и

$$\inf_{A\in\Omega}\ \max_{Y\in R(A^*)}\left[\mathscr{F}(A,Y)-\mathscr{F}(A^*,Y)\right]=0. \qquad (8)$$

Покажем вначале, что из (7) следует (8).
Для любых $A\in\Omega$ и $Y\in\mathscr{D}$ имеем

$$\mathcal{F}(A,y)-\mathcal{F}(A^*,y)=|\Delta(A,y)|-|\Delta(A^*,y)|\geq$$

$$\geq(\Delta(A,y)-\Delta(A^*,y))\cdot \operatorname{sign}\Delta(A^*,y)=$$

$$=(\Phi(A^*,y)-\Phi(A,y))\cdot \operatorname{sign}\Delta(A^*,y).$$

Отсюда и из (7) следует (8).
Докажем обратное. Прежде всего, заметим, что если $|a|\geq|b|$, то

$$|a+b|=|a|+b\cdot \operatorname{sign} a. \qquad (9)$$

При всех $y\in R(A^*)$ величина $|\Delta(A^*,y)|$ имеет одно и то же значение, которое мы обозначим через $\mu$. Если $\mu=0$, то равенство (7), очевидно, выполняется. Поэтому будем считать, что $\mu>0$. Зафиксируем $A\in\Omega$ и выберем $\alpha\in(0,1)$ так, чтобы

$$\alpha\cdot\max_{y\in R(A^*)}|\Phi(A-A^*,y)|\leq\mu.$$

Тогда в силу (9) при всех $y\in R(A^*)$ будет

$$|\Delta(A^*+\alpha(A-A^*),y)|=|\Delta(A^*,y)+\alpha\Phi(A^*-A,y)|=$$

$$=|\Delta(A^*,y)|+\alpha\cdot\Phi(A^*-A,y)\cdot \operatorname{sign}\Delta(A^*,y).$$

Отсюда следует, что

$$\max_{y\in R(A^*)}\left[\mathcal{F}(A^*+\alpha(A-A^*),y)-\mathcal{F}(A^*,y)\right]=$$

$$=\alpha\cdot\max_{y\in R(A^*)}\left\{[\Phi(A^*,y)-\Phi(A,y)]\cdot \operatorname{sign}\Delta(A^*,y)\right\}.$$

Учитывая (8) и то, что $\alpha > 0$, получаем

$$\max_{y \in R(A^*)} \left\{ \left[ \Phi(A^*, y) - \Phi(A, y) \right] \cdot \operatorname{sign} \Delta(A^*, y) \right\} \geq 0, \quad \forall A \in \Omega,$$

что равносильно (7). Теорема доказана.

3°. Рассмотрим задачу дробно-рациональной аппроксимации :

$$\max_{y \in \mathscr{D}} \left| f(y) - \frac{\sum_{i=1}^{n} a_i u_i(y)}{\sum_{i=1}^{m} b_i v_i(y)} \right| \longrightarrow \min_{X \in \Omega}, \tag{10}$$

где $\mathscr{D}$ – метрический компакт, $X = (a_1, \dots, a_n, b_1, \dots, b_m)$. Функции $f(y), u_1(y), \dots, u_n(y), v_1(y), \dots, v_m(y)$ непрерывны на $\mathscr{D}$. Если ввести обозначения

$$A = (a_1, \dots, a_n), \quad B = (b_1, \dots, b_m),$$

$$P(A, y) = \sum_{i=1}^{n} a_i u_i(y), \quad Q(B, y) = \sum_{i=1}^{m} b_i v_i(y),$$

$$H(X, y) = \frac{P(A, y)}{Q(B, y)}, \quad \delta(X, y) = f(y) - H(X, y),$$

то (10) можно переписать в следующем виде :

$$\max_{y \in \mathscr{D}} \left| \delta(X, y) \right| \longrightarrow \min_{X \in \Omega}. \tag{11}$$

Здесь

$$\Omega = \left\{ X \in E^{n+m} \mid Q(B, y) > 0, \ \forall y \in \mathscr{D} \right\} \cap \Omega_1,$$

где $\Omega_1 \subset E^{n+m}$ – выпуклое множество. Допустим, что задача (11) имеет решение. В этом случае для элемента наилуч-

шего приближения имеет место критерий, аналогичный критерию А.Н. Колмогорова.

ТЕОРЕМА 2. Для того чтобы вектор $X^* \in \Omega$ был решением задачи (11), необходимо и достаточно, чтобы выполнялось равенство

$$\inf_{X\in\Omega} \max_{Y\in R(X^*)} \left\{ \left[ H(X^*,Y) - H(X,Y) \right] \cdot \operatorname{sign} \delta(X^*,Y) \right\} = 0. \qquad (12)$$

ЗАМЕЧАНИЕ. Поскольку функция $|\delta(X,Y)|$ не является выпуклой по $X$, то теоремой 1 непосредственно воспользоваться нельзя. Однако идея доказательства теоремы 1 оказывается плодотворной и в данной ситуации.

ДОКАЗАТЕЛЬСТВО. НЕОБХОДИМОСТЬ. Пусть $X^* \in \Omega$ -решение задачи (11) и $X$ -любой вектор из $\Omega$. Тогда при малых $\alpha > 0$

$$X^* + \alpha(X - X^*) \in \Omega$$

и

$$\delta(X^* + \alpha(X - X^*), Y) = \delta(X^*, Y) +$$

$$+ \frac{P(A^*,Y)}{Q(B^*,Y)} - \frac{P(\alpha A + (1-\alpha)A^*, Y)}{Q(\alpha B + (1-\alpha)B^*, Y)} =$$

$$= \delta(X^*,Y) + \alpha \cdot \frac{P(A^*,Y)Q(B,Y) - P(A,Y)Q(B^*,Y)}{Q(B^*,Y)Q(B^* + \alpha(B - B^*), Y)} . \qquad (13)$$

Величина $|\delta(X^*,y)|$ при всех $y\in R(X^*)$ имеет одинаковое значение, которое мы обозначим через $\mu$. Если $\mu=0$, то (12), очевидно, выполняется. Пусть $\mu>0$. Зафиксируем $\varepsilon\in(0,\mu)$. Тогда в силу (13) и (9) найдется такое $\delta>0$, что для $\alpha\in(0,\delta)$ равномерно по $y\in R_\varepsilon(X^*)$ будет

$$|\delta(X^*+\alpha(X-X^*),y)|=|\delta(X^*,y)|+$$
$$+\alpha\cdot\frac{P(A^*,y)Q(B,y)-P(A,y)Q(B^*,y)}{Q(B^*,y)Q(B^*+\alpha(B-B^*),y)}\cdot \operatorname{sign}\delta(X^*,y).$$

Так же, как при доказательстве теоремы 1, отсюда получаем при малых $\alpha>0$

$$\max_{y\in R_\varepsilon(X^*)}\left\{\frac{P(A^*,y)Q(B,y)-P(A,y)Q(B^*,y)}{Q(B^*,y)Q(B^*+\alpha(B-B^*),y)}\times\right.$$
$$\left.\times \operatorname{sign}\delta(X^*,y)\right\}\geq 0.$$

В последнем неравенстве заменим $Q(B^*+\alpha(B-B^*),y)$ на $Q(B,y)$, учитывая, что обе эти функции положительны на $\mathcal{D}$. Получим

$$\max_{y\in R_\varepsilon(X^*)}\left\{[H(X^*,y)-H(X,y)]\times\right.$$
$$\left.\times \operatorname{sign}\delta(X^*,y)\right\}\geq 0,\quad \forall X\in\Omega.$$

Остается воспользоваться тем, что $\varepsilon$ – любое число из $(0,\mu)$, и так же, как при доказательстве теоремы 1, вывести (12).

ДОСТАТОЧНОСТЬ. Пусть имеет место (12). Тогда для любого $X\in\Omega$ будет

$$\max_{y\in\mathcal{D}} |\delta(x,y)| \geq \max_{y\in R(x^*)} \{\delta(x,y)\cdot \operatorname{sign}\delta(x^*,y)\} \geq$$

$$\geq \max_{y\in R(x^*)} \{\delta(x^*,y)\cdot \operatorname{sign}\delta(x^*,y)\} = \max_{y\in\mathcal{D}} |\delta(x^*,y)|.$$

Значит, $x^*$ является точкой минимума функции $\max\limits_{y\in\mathcal{D}} |\delta(x,y)|$ на $\Omega$. Теорема доказана.

4°. В заключение покажем, как на основе теоремы I может быть получена теорема из [I], стр. 237.

Пусть $\mathcal{F}(x,y)$ –функция, непрерывная вместе с $\frac{\partial \mathcal{F}(x,y)}{\partial x}$ на $G\times\mathcal{D}$, где $G\subset E^n$ – открытое множество, а $\mathcal{D}\subset E^m$ – компакт. Рассмотрим задачу

$$\varphi(x) \stackrel{def}{=} \max_{y\in\mathcal{D}} \mathcal{F}(x,y) \longrightarrow \min_{x\in\Omega}. \tag{14}$$

Здесь $\Omega\subset G$ – выпуклое множество. Предполагается, что задача (14) имеет решение.

ТЕОРЕМА 3. Для того чтобы вектор $x^*\in\Omega$ был решением задачи (14), необходимо (а в случае выпуклости $\varphi(x)$ на $\Omega$ и достаточно), чтобы выполнялось равенство

$$\inf_{x\in\Omega} \max_{y\in R(x^*)} \left(\frac{\partial \mathcal{F}(x^*,y)}{\partial x}, x-x^*\right) = 0. \tag{15}$$

ДОКАЗАТЕЛЬСТВО. НЕОБХОДИМОСТЬ. Пусть $x^*\in\Omega$ – решение задачи (14), а $x$ – произвольный вектор из $\Omega$. При $\alpha\in(0,1)$ имеем

$$\mathcal{F}(X^*+\alpha(X-X^*),\ y)=\mathcal{F}(X^*,\ y)+$$

$$+\alpha\left(\frac{\partial\mathcal{F}(X^*,y)}{\partial X},\ X-X^*\right)+o(y;\alpha), \qquad (16)$$

где $\frac{o(y;\alpha)}{\alpha}\xrightarrow[\alpha\to+0]{}0$ равномерно относительно $y\in\mathcal{D}$ .
Так же, как при доказательстве теоремы 1, отсюда получаем

$$\max_{y\in R_\varepsilon(X^*)}\left\{\left(\frac{\partial\mathcal{F}(X^*,y)}{\partial X},X-X^*\right)+\frac{o(y;\alpha)}{\alpha}\right\}\geqslant 0.$$

Значит,

$$\max_{y\in R_\varepsilon(X^*)}\left(\frac{\partial\mathcal{F}(X^*,y)}{\partial X},\ X-X^*\right)\geqslant 0,\quad \forall X\in\Omega.$$

Из последнего неравенства ввиду произвольности $\varepsilon>0$ следует (15).

ДОСТАТОЧНОСТЬ. Зафиксируем $X\in\Omega$ . По условию при $\alpha\in(0,1)$ выполняется неравенство

$$\varphi(X^*+\alpha(X-X^*))\leq\varphi(X^*)+\alpha\left[\varphi(X)-\varphi(X^*)\right]. \qquad (17)$$

Но в силу (16)

$$\varphi(X^*+\alpha(X-X^*))-\varphi(X^*)\geqslant$$

$$\geqslant\max_{y\in R(X^*)}\left[\mathcal{F}(X^*+\alpha(X-X^*),y)-\mathcal{F}(X^*,y)\right]=$$

$$=\alpha\cdot\max_{y\in R(X^*)}\left[\left(\frac{\partial\mathcal{F}(X^*,y)}{\partial X},\ X-X^*\right)+\frac{o(y;\alpha)}{\alpha}\right].$$

Отсюда и из (17) следует, что

$$\varphi(X)-\varphi(X^*)\geq \max_{Y\in R(X^*)}\left[\left(\frac{\partial\mathcal{F}(X^*,Y)}{\partial X},X-X^*\right)+\frac{o(Y;\alpha)}{\alpha}\right].$$

Переходя в последнем неравенстве к пределу при $\alpha\to+0$ и учитывая (15), получаем

$$\varphi(X)-\varphi(X^*)\geq 0,\qquad \forall X\in\Omega.$$

Значит, $X^*$ является точкой минимума функции $\varphi(X)$ на $\Omega$. Теорема доказана.

Из теоремы I следует также известная теорема о минимаксе (см., например, [2] , стр. 42-45 ).

ТЕОРЕМА 4. Пусть $\Omega$ и $\mathcal{D}$ - выпуклые компактные множества в $E^n$ и $E^m$ соответственно, а $\mathcal{F}(X,Y)$ - функция, заданная и непрерывная на $\Omega\times\mathcal{D}$, выпуклая по $X$ при каждом $Y$ и вогнутая по $Y$ при каждом $X$. Тогда

$$\min_{X\in\Omega}\max_{Y\in\mathcal{D}}\mathcal{F}(X,Y)=\max_{Y\in\mathcal{D}}\min_{X\in\Omega}\mathcal{F}(X,Y). \tag{18}$$

ДОКАЗАТЕЛЬСТВО. Можно считать, что $\mathcal{F}(X,Y)$ строго вогнута по $Y$ при каждом фиксированном $X$. Обозначим через $X^*$ точку минимума функции $\varphi(X)$ на $\Omega$. По теореме I

$$\max_{Y\in R(X^*)}\left[\mathcal{F}(X,Y)-\mathcal{F}(X^*,Y)\right]\geq 0,\qquad \forall X\in\Omega. \tag{19}$$

В силу строгой вогнутости $\mathcal{F}(X^*,Y)$ множество $R(X^*)$ сос -

тоит из единственной точки, которую мы обозначим через $y^*$. Тогда по определению $R(X^*)$

$$\mathcal{F}(X^*, y) \leq \mathcal{F}(X^*, y^*) \quad \forall y \in \mathcal{D}, \tag{20}$$

а по (19)

$$\mathcal{F}(X, y^*) \geq \mathcal{F}(X^*, y^*) \quad \forall X \in \Omega. \tag{21}$$

Из (20) и (21) очевидным образом следует (18). Теорема доказана.

## Литература


1. Демьянов В.Ф., Малоземов В.Н. Введение в минимакс. М., "Наука", 1972.
2. Карлин С. Математические методы в теории игр, программировании и экономике. М., "Мир," 1964.



Поступила в ред.-изд. отд.
6. XII. 1972 г.