УДК 519.8

РАЗЫСКАНИЕ НАПРАВЛЕНИЯ НАИСКОРЕЙШЕГО СПУСКА В МИНИМАКСНЫХ ЗАДАЧАХ СО СВЯЗАННЫМИ ОГРАНИЧЕНИЯМИ

Е.Ф. Войтон, В.Ф.Демьянов

Функции вида

$$\varphi(X) = \max_{Y \in G(X)} f(X,Y),$$

где $G(X) \subset E_m$ - множество, зависящее от $X \in E_n$, изучались в [1 - 4].

Пусть

$$G(X) = \{Y \in E_m \mid h_j(X,Y) \le 0 \quad \forall j \in J\}, \qquad (1)$$

где $J = [1:N]$, функции f и h_j непрерывно дифференцируемы по обеим переменным, функции $h_j(X,Y)$ выпуклы по Y при любом фиксированном X.

Предположим, что для $G(X)$ выполнено условие Слейтера, т.е. существует точка $Y_0 \in G(X)$ такая, что

$$h_j(X,Y_0) < 0 \qquad \forall j \in J. \qquad (2)$$

Пусть $$R(X)=\{Y\in G(X) \mid f(X,Y)=\varphi(X)\}.$$
Допустим, что $f(X,Y)$ вогнута по Y в окрестности множества $R(X)$. Тогда, как показано в [2,3], для того чтобы в точке $X^*\in E_n$ функция $\varphi(X)$ достигала своего минимального значения, необходимо, чтобы

$$\inf_{\|g\|=1}\ \sup_{Y\in R(X^*)}\ \sup_{V\in\Gamma(X^*,g,Y)}\left[\left(\frac{\partial f(X^*,Y)}{\partial Y},V\right)+\left(\frac{\partial f(X^*,Y)}{\partial X},g\right)\right]\geqslant 0, \quad (3)$$

где
$$\Gamma(X,g,Y)=\left\{V\in E_m \,\middle|\, \left(\frac{\partial h_j(X,Y)}{\partial Y},V\right)+\left(\frac{\partial h_j(X,Y)}{\partial X},g\right)\leqslant 0,\right.$$
$$\left.\forall j\in Q(X,Y)\right\},$$
$$Q(X,Y)=\{j\in J \mid h_j(X,Y)=0\}.$$

Точка X^*, удовлетворяющая (3), называется стационарной точкой функции $\varphi(X)$. Для эффективного нахождения стационарных точек надо уметь вычислять правую часть в (3) для любого $X\in E_n$.

Если инфимум в (3) отрицателен, то вектор g, на котором он достигается, является направлением наискорейшего спуска функции φ в точке X. Ниже и исследуется задача нахождения направления наискорейшего спуска.

Рассмотрим случай, когда множество $R(X)$ состоит из конечного числа точек. Обозначим их через $\{Y_i\}$. Положим $R=\{i \mid Y_i\in R(X)\}$. Тогда имеем задачу: Найти

$$\inf_{\|g\|=1}\ \max_{i\in R}\left[(D_i,g)+\sup_{V\in T_i(g)}(A_i,g)\right]\stackrel{df}{=}\inf_{\|g\|=1}\mathcal{H}(g), \quad (4)$$

где

$$T_i(g) \overset{df}{=} \Gamma(X,g,Y_i)=\{V\in E_m|(B_{ij},V)+(C_{ij},g)\leqslant 0 \quad \forall j\in J_i\},$$

$$A_i=\frac{\partial f(X,Y_i)}{\partial Y}, \quad D_i=\frac{\partial f(X,Y_i)}{\partial X}, \quad B_{ij}=\frac{\partial h_j(X,Y_i)}{\partial Y},$$

$$C_{ij}=\frac{\partial h_j(X,Y_i)}{\partial X}, \quad J_i=\{j\in J | h_j(X,Y_i)=0\}, \quad Y_i\in R(X),$$

$$\mathcal{H}(g) = \max_{i\in R} [(D_i,g) + \sup_{V\in T_i(g)} (A_i,V)].$$

Вначале покажем, что $\sup\limits_{V\in T_i(g)} (A_i,V)$ конечен.

Действительно, поскольку $f(X,Y_i)=\max\limits_{Y\in G(X)} f(X,Y)$, то

по необходимому условию максимума

$$(A_i,V) \overset{df}{=} \left(\frac{\partial f(X,Y_i)}{\partial Y}, V\right)\leqslant 0 \tag{5}$$

для всех $V\in\Gamma_1(Y_i)$, где $\Gamma_1(Y_i)$ – конус допустимых направлений множества $G(X$ (X фиксировано)в точке Y_i. т.е.

$$\Gamma_1(Y_i)=\{V\in E_m | (V, \frac{\partial h_j(X,Y_i)}{\partial Y})\leqslant 0 \quad \forall j\in J_i\}$$

Из условия Слейтера (2) и из свойств выпуклых функций для

$$j\in J: \quad 0>h_j(X,Y_0)=h_j(X,Y_i+(Y_0-Y_i))\geqslant h_j(X,Y_i)+$$

$$+\left(\frac{\partial h_j(X,Y_i)}{\partial Y}, Y_0-Y_i\right)=\left(\frac{\partial h_j(X,Y_i)}{\partial Y}, Y_0-Y_i\right)$$

Теперь ясно, что при достаточно большом $\mu > 0$ будет

$$\left(\frac{\partial h_j(X,Y_i)}{\partial Y},\ \mu(Y_0-Y_i)\right) < -\left(\frac{\partial h_j(X,Y_i)}{\partial X}, g\right),$$

или, что то же,

$$(B_{ij},\ \mu(Y_0-Y_i)) < -(C_{ij}, g), \quad j \in J_i \tag{6}$$

Допустим, что $\sup\limits_{V \in T_i(g)} (A_i, V_\kappa)$ не является конечным.
Тогда существует последовательность $\{V\}$ такая, что

$$V_\kappa \in T_i(g), \quad (A_i, V_\kappa) \xrightarrow[\kappa\to\infty]{} +\infty.$$

Для $W_\kappa = V_\kappa + \mu(Y_0 - Y_i)$ имеем

$$(A_i, W_\kappa) \xrightarrow[\kappa\to\infty]{} +\infty. \tag{7}$$

С другой стороны, для всех $j \in J_i$ из (6) и определения $T_i(g)$ будет

$$(B_{ij}, W_\kappa) = \left(\frac{\partial h_j(X,Y_i)}{\partial Y},\ W_\kappa\right) =$$

$$= (B_{ij}, V_\kappa) + \mu(B_{ij}, Y_0 - Y_i) < 0,$$

т.е. $W_\kappa \in \Gamma_i(Y_i)$. Но для $\{W_\kappa\}$ выполнено (7), что противоречит (5). Конечность доказана.

Теперь из теоремы разрешимости задач линейного программирования заключаем, что этот супремум достигается, т.е. существует $V_i \in T_i(g)$ такое, что

$$(A_i, V_i) = \sup_{V \in T_i(g)} (A_i, V)$$

Теперь стандартными рассуждениями можно показать, что функция $\mathcal{H}(g)$ непрерывна.

Рассмотрим теперь более подрооно задачу разыскания

$$\sup_{V\in T_i(g)} (A_i, V).$$

Выше уже доказано, что супремум достигается.

Для удобства изложения индекс i временно опустим.

Имеем следующую задачу: Н а й т и

$$\max_{V\in T(g)} (A, V).$$

Зафиксируем произвольное $g_0 \in E_n$, $\|g_0\| = 1$.

Пусть

$$(A, V_0) = \max_{V\in T(g_0)} (A, V),$$

где

$$T(g_0) = \{V \in E_m \mid (B_j, V) + (C_j, g_0) \leq 0 \quad \forall j \in J\}.$$

Из необходимых условий максимума заключаем, что существуют подмножество $J(g_0)$ индексного множества J и положительные коэффициенты α_j такие, что

$$A = \sum_{j\in J(g_0)} \alpha_j B_j, \tag{8}$$

причем векторы B_j, $j \in J(g_0)$, линейно независимы и

$$(B_j, V_0) + (C_j, g_0) = 0 \quad \text{для} \quad j \in J(g_0).$$

Пусть $V_0 = V_1 + V_2$, где V_1 принадлежит линейному множеству, натянутому на векторы $\{B_j \mid j \in J(g_0)\}$, а V_2 – его ортогональному дополнению, т.е. $(B_j, V_2) = 0$ для всех $j \in J(g_0)$.

Для $j \in J(g_0)$ вектор B_j представим тоже в виде

$$B_j = B_j' + B_j''$$

где $(B_j'', B_k) = 0$ для $k \in J(g_0)$

Для вектора V_0 имеем $(A, V_0) = (A, V_1)$,

$$(B_j, V_0) + (C_j, g_0) = (B_j, V_1) + (C_j, g_0) = 0 \quad \forall j \in J(g_0),$$

$$(B_j, V_0) + (C_j, g_0) = (B_j' + B_j'', V_1 + V_2) + (C_j, g_0) =$$

$$= (B_j', V_1) + (B_j'', V_2) + (C_j, g_0) \leq 0 \qquad \forall j \in J(g_0).$$

Заметим, что на основании (8)

$$(A, V_1) = \max_{V \in T_1(g_0)} (A, V),$$

где

$$T_1(g_0) = \{V \in E_m \mid (B_j, V) + (C_j, g_0) \leq 0 \quad \forall j \in J(g_0)\}.$$

Рассмотрим задачу : Н а й т и

$$\max_{V \in T_2(g)} (A, V) \tag{9}$$

где $T_2(g) = \{V \in E_m \mid (B_j, V) + (C_j, g) \leq 0 \quad \forall j \in J(g_0)\}$

В силу (8) можно утверждать, что решение системы

$$(B_j, V) + (C_j, g) = 0, \qquad \forall j \in J(g_0) \tag{10}$$

является решением задачи (9). Решение системы (10) ищем в виде

$$V' = \sum_{j \in J(g_0)} \beta_j B_j \ . \qquad \text{(11)}$$

Пусть $J(g_0) = [1:s]$, $s \leqslant m$.

Подставляя (11) в (10) , получаем

$$\sum_{k=1}^{s} \beta_k (B_k, B_j) + (C_j, g) = 0 \qquad \forall_j \in [1:s]. \qquad \text{(12)}$$

Поскольку векторы B_j, $j \in [1:s]$, линейно-независимы, то система (12) имеет единственное решение

$$\beta \overset{df}{=} (\beta_1, \dots, \beta_s) = - \bar{B}^{-1} C^* g, \qquad \text{(13)}$$

где

$$\bar{B} = \begin{pmatrix} (B_1, B_1)(B_1, B_2) & \dots & (B_1, B_s) \\ (B_2, B_1)(B_2, B_2) & \dots & (B_2, B_s) \\ \dots & \dots & \dots \\ (B_s, B_1)(B_s, B_2) & \dots & (B_s, B_s) \end{pmatrix}, \qquad \text{(14)}$$

$$C = (C_1, C_2, \dots, C_s). \qquad \text{(15)}$$

Любое другое решение системы (10) может быть представлено в виде $V = V' + V''$,

где V' есть вектор из (11) (и β_j удовлетворяют (13) , а V'' принадлежит ортогональному дополнению множества векторов

$$\{ B_j \mid j \in [1:s] \overset{df}{=} J(g_0) \}.$$

Поскольку вектор A имеет вид (6), то вектор V'' не влияет на значение скалярного произведения и потому

$$\max_{V \in T_2(g)} (A, V) = (A, V') = (A, -B\bar{B}^{-1}C^* g) =$$

$$= -\left((B\bar{B}^{-1}C^*)^* A, g\right),$$

где

$$B = (B_1, \ldots, B_s). \tag{16}$$

Так как $$T(g) \subset T_2(g),$$

то $$\max_{V \in T(g)} (A, V) \leq \max_{V \in T_2(g)} (A, V) = (F, g),$$

где $F = -(B\bar{B}^{-1}C^*)^* A$. Заметим, что $F = F(g_0)$

Восстанавливая теперь опущенный ранее индекс i, получим

$$\mathcal{H}(g) \leq \max_{i \in R} (G_i, g) \overset{df}{=} \varphi(g, g_0), \tag{17}$$

где

$$G_i = G_i(g_0) = D_i + F_i(g_0), \quad F_i(g_0) = -(B_i \bar{B}_i^{-1} C_i^*)^* A_i .$$

Здесь B_i, $\bar{B}_i$ и C_i определяются по формулам (14), (15), (16) с учетом индекса i.

Ясно, что G_i от g не зависит.
Поэтому из (17) имеем

$$\min_{\|g\|=1} \mathcal{H}(g) \leq \min_{\|g\|=1} \max_{i \in R} (G_i, g) \overset{df}{=} \varphi(g_0) =$$

$$= \min_{\|g\|=1} \varphi_1(g, g_0) = \varphi_1(g_1, g_0). \tag{18}$$

Но минимакс в (I8) можно найти. Если $\eta(g_0 < 0)$, то

$$\min_{\|g\|=1} \max_{i\in R} (G_i, g) = -\|\bar Z\|,$$

где $\|\bar Z\| = \min_{Z\in L} \|Z\|$, $\quad L = L(g_0) \overset{df}{=} co\{G_i | i\in R\}$.

При этом поскольку $\|\bar Z\| \neq 0$, то

$$\min_{\|g\|=1} \max_{i\in R} (G_i, g) = \max_{i\in R} (G_i, \bar g), \quad \bar g = -\frac{\bar Z}{\|\bar Z\|} .$$

Ясно, что $\bar g = \bar g(g_0)$.

Если же $\eta(g_0) \geqslant 0$, то $\eta(g_0) = \tau(g_0)$,

где $\tau(g_0)$ – радиус максимальной сферы, которая может быть вписана в многогранник $L(g_0)$.

Из изложенного выше вытекает следующий метод последовательных приближений для минимизации $\mathcal{H}(g)$.

Фиксируем произвольное g_0 . Как выше, находим все векторы $F_i(g_0)$. Находим $\eta(g_0) = \eta_1(g_1, g_0)$.

При этом если $\eta_1(g_1, g_0) < \eta_1(g_0, g_0)$, то в силу (I7)

$$\mathcal{H}(g_1) < \mathcal{H}(g_0) \text{ , ибо } \eta_1(g_0, g_0) = \mathcal{H}(g_0).$$

Исходя из g_1 , далее поступаем аналогично (строим векторы $F_i(g_1)$ и т.д.).

Если же оказалось $\eta_1(g_1, g_0) = \eta_1(g_0, g_0)$, то процесс прекращается.

Продолжая далее аналогично, строим последовательность конечную или нет) $\{g_k\}$ такую, что $\mathcal{H}(g_{k+1}) < \mathcal{H}(g_k)$.

К сожалению, нельзя утверждать, что предельная точка последовательности $\{g_k\}$ есть точка минимума функции $\mathcal{H}(g)$

на множестве $\{g \mid \|g\| = 1\}$

Это может произойти по двум причинам. Прежде всего, может оказаться, что

$$\eta_1(g_k, g_{k-1}) = \eta_1(g_{k-1}, g_{k-1}). \qquad (19)$$

Но при другом выборе множеств $J_i(g_{k-1})$, удовлетворяющих (8), могут получиться совсем другие векторы $\{F_i(g_{k-1})\}$, а потому и (19) может не иметь места.

Кроме того, функция $\mathcal{H}(g)$ может иметь локальные минимумы.

Справедлива следующая теорема.

ТЕОРЕМА. Существуют такие подмножества индексов $J_i^* \subset J_i$, что

$$\min_{\|g\|=1} \mathcal{H}(g) = \mathcal{H}(g^*),$$

где

$$\max_{i \in R} (G_i, g^*) = \min_{\|g\|=1} \max_{i \in R} (G_i, g),$$

$$G_i = D_i + F_i, \quad F_i = -(B_i \bar{B}_i^{-1} C_i^*) A_i.$$

Здесь B_i, $\bar{B}_i$, C_i определяются по формулам (14), (15), (16) с добавлением индекса i, причем в качестве $J(g_0)$ следует брать J^*

Доказательство немедленно следует из предыдущих рассмотрений, если положить $J_i^* = J_i(\tilde{g})$.

Действительно, по определению g^* (см. формулировку теоремы) имеем

$$\mathcal{H}(g^*) \leq \eta_1(g^*, \tilde{g}) \leq \eta_1(\tilde{g}, \tilde{g}) = \mathcal{H}(\tilde{g}) = \min_{\|g\|=1} \mathcal{H}(g).$$

Литература

I. Сотсков А.И. О необходимых условиях минимакса на связанных множествах."Кибернетика," № 4 (1970), 96 - 102.

2. Демьянов В.Ф. О задаче минимакса при связанных ограничениях. Ж. вычисл.матем. и матем.физ., 12 : 3 (1972).

3. Демьянов В.Ф., Певный А.Е. Вычисление первых и вторых маргинальных значений в задачах математического программирования. I.-"Оптимизация", 5 (22) (1972), 63 - 72.

4. Hogan W.W. Directional derivatives for extremal value functions. Western Management Science Institute, Working Paper No. 177, Los Angeles, 1971.

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