ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2006 МАТЕМАТИКА № 12 (535)

УДК 519.853

З.Р. ГАБИДУЛЛИНА

ТЕОРЕМА ОТДЕЛИМОСТИ ВЫПУКЛОГО МНОГОГРАННИКА ОТ НУЛЯ ПРОСТРАНСТВА И ЕЕ ПРИЛОЖЕНИЯ В ОПТИМИЗАЦИИ

В теории экстремальных задач важную роль играют теоремы, называемые теоремами отделимости. Обычно доказательство теорем отделимости выпуклого множества $D \subset R_n$ и точки $x_0 \notin D$ (напр., [1], с.200; [2]–[4]) сводится к построению одной отделяющей их гиперплоскости вида

$$\langle d, x - x_0\rangle \geq \|d\|^2 \quad \forall x \in D,$$

где $d = \mathbf{P}_D(x_0) - x_0$, $\mathbf{P}_D(x_0)$ — проекция точки x_0 на множество D. В данной работе предлагается новый способ построения множества нормальных векторов гиперплоскостей, каждая из которых сильно отделяет выпуклую комбинацию конечного числа точек от нуля пространства. В работе также дано приложение теоремы отделимости, относящееся к теории недифференцируемой оптимизации, в частности к задаче нахождения проекции нуля пространства на выпуклый многогранник конечными методами.

1^0. Теорема отделимости

Пусть $J = \{1, 2, \dots, m\}$, $z_i \in R_n$, $i \in J$ и

$$L = \Big\{z \in R_n : z = \sum_{i\in J} \alpha_i z_i,\ \alpha_i \geq 0,\ \sum_{i\in J}\alpha_i = 1\Big\}.$$

Очевидно, L является непустым замкнутым ограниченным выпуклым множеством. Положим

$$\Omega_i = \{y \in R_n : \langle z_i, y\rangle \geq \|y\|^2\}, \quad i \in J.$$

Лемма 1. *Множество $\Omega_i \subset R_n$, $i \in J$, является n-мерным шаром радиуса $R_i = \|z_i\|/2$ с центром в точке $O_i = z_i/2$.*

Доказательство. Прибавив к обеим частям неравенства, определяющего множество Ω_i, $\|z_i\|^2/4$, получим $\|y - z_i/2\|^2 \leq \|z_i\|^2/4$. $\square$

Так как множество Ω_i, $i \in J$, является выпуклым, замкнутым и ограниченным, то множество

$$\Omega = \bigcap_{i\in J} \Omega_i = \{y \in R_n : \langle z_i, y\rangle \geq \|y\|^2 \ \forall i \in J\}$$

является выпуклым компактом.

Определение 1 ([1], с. 198). Пусть D — множество из R_n. Говорят, что гиперплоскость $\langle c, u\rangle = a$ с нормальным вектором $c \neq \mathbf{0}$ отделяет множество D от точки x_0, если $\langle c, d\rangle \geq a$ при всех $d \in D$ и $\langle c, x_0\rangle \leq a$, т. е. если выполняются неравенства $\langle c, x_0\rangle \leq a \leq \inf\limits_{d\in D}\langle c, d\rangle$. Если $\langle c, x_0\rangle < \inf\limits_{d\in D}\langle c, d\rangle$, то говорят, что множество D и точка x_0 сильно отделимы.

Определение 2 ([1], с. 200). Гиперплоскость $\langle c, d\rangle = a$ называют опорной к множеству в точке D ($\overline{D}$ — замыкание множества D), если $\langle c, d\rangle \geq a$ при всех $d \in D$ и $\langle c, y\rangle = a$, $y \in \overline{D}$.

Теорема 1. *Пусть* $\Pi = \{z \in R_n : \langle c, z\rangle = a\}$, $a > 0$, *и* $\langle c, z\rangle \geq a$ *для всех* $z \in L$. *Тогда гиперплоскость* Π *сильно отделяет множество* L *от нуля пространства* R_n *и* $y = ac/\|c\|^2 \in \Omega$.

Доказательство. Так как $a > 0$, то первое утверждение очевидно. Поскольку $z_i \in L$ для всех $i \in J$, то $\langle c, z_i\rangle - a \geq 0$ для всех $i \in J$. Тогда $\langle z_i, y\rangle - \|y\|^2 = \dfrac{a}{\|c\|^2}(\langle c, z_i\rangle - a) \geq 0 \quad \forall i \in J$, т. е. $y \in \Omega$. $\square$

Замечание 1. Очевидно, гиперплоскость $\Pi_0 = \{z \in R_n : \langle y, z\rangle = b\}$ совпадает с Π при $y = ac/\|c\|^2$, $b = a^2/\|c\|^2$.

Поскольку $\varphi(z) = \|z\|^2$ строго выпукла, L — выпуклое замкнутое множество, то существует единственная точка $\rho \in L$ такая, что $\min\limits_{z \in L} \|z\|^2 = \|\rho\|^2$. Точку ρ называют проекцией нуля пространства R_n на множество L, поэтому будем употреблять далее обозначение $\mathbf{P}_L(\mathbf{0}) = \rho$.

Из следующих двух теорем получим, что гиперплоскость $\langle c, z\rangle = a$ при $c = \rho$, $a = \|\rho\|^2$ сильно отделяет множество L от точки $\mathbf{0}$. Введем обозначение $\Gamma\wp\,\Omega$ — граница множества Ω.

Теорема 2. *Точка* $\mathbf{P}_L(\mathbf{0}) \in \Gamma\wp\,\Omega$.

Доказательство. По известному свойству проекции (напр., [1], с. 193–194) выполняется $\langle \rho - \mathbf{0}, z - \rho\rangle \geq 0 \quad \forall z \in L$. Тогда $\langle z, \rho\rangle \geq \|\rho\|^2 \quad \forall z \in L$.

Поскольку $z_i \in L \quad \forall i \in J$, то $\langle z_i, \rho\rangle \geq \|\rho\|^2 \quad \forall i \in J$, т. е. $\rho \in \Omega$. Известно, что $\min\limits_{i \in J} \langle z_i, \rho\rangle = \|\rho\|^2$ ([2], с. 337). Тогда множество индексов $J(\rho) = \{i \in J : \langle z_i, \rho\rangle = \min\limits_{i \in J}\langle z_i, \rho\rangle = \|\rho\|^2\}$ не пусто и вектор $\rho \in \Gamma\wp\,\Omega$. $\square$

Лемма 2. *Если* $\mathbf{0} \notin L$, *то* $\Omega \neq \{\mathbf{0}\}$.

Доказательство. Поскольку $\mathbf{0} \notin L$, то $\mathbf{P}_L(\mathbf{0}) \neq \mathbf{0}$. Отсюда с учетом теоремы 2 имеем $\Omega \neq \{\mathbf{0}\}$. $\square$

Обобщением известных теорем отделимости выпуклого многогранника и нуля пространства является

Теорема 3. *Если* $\mathbf{0} \notin L$, *то* $\forall c \in \Omega$, $c \neq \mathbf{0}$, *гиперплоскость* $\langle c, z\rangle = \|c\|^2$ *сильно отделяет множество* L *от нуля пространства* R_n.

Доказательство. Согласно лемме 2 множество $\Omega \neq \{\mathbf{0}\}$. В силу определения множества Ω

$$\langle z_i, c\rangle \geq \|c\|^2 \quad \forall i \in J, \quad \forall c \in \Omega.$$

Отсюда для любой точки $z \in L$ выполняется соотношение $\langle c, z\rangle = \left\langle c, \sum\limits_{i \in J} \alpha_i z_i\right\rangle \geq \|c\|^2 \sum\limits_{i \in J} \alpha_i = \|c\|^2 > 0$, т. к. $\alpha_i \geq 0$, $\sum\limits_{i \in J} \alpha_i = 1$. Поэтому гиперплоскость $\langle c, z\rangle = \|c\|^2$ разделяет множество L и точку $\mathbf{0}$. Поскольку $\inf\limits_{z \in L}\langle c, z\rangle \geq \|c\|^2 > 0 = \langle c, \mathbf{0}\rangle$, то множество L и нуль пространства R_n сильно отделимы. $\square$

Замечание 2. Согласно теореме 3

$$L \subset \{z \in R_n : \langle c, z\rangle \geq \|c\|^2\}, \quad \mathbf{0} \in \{z \in R_n : \langle c, z\rangle < \|c\|^2\}.$$

2^0. Приложения теоремы отделимости

Для формулировки основных результатов поставим следующие две задачи нелинейного программирования:

$$\max \|y\|^2, \tag{1}$$
$$\langle z_i, y\rangle \geq \|y\|^2, \quad i \in J, \tag{2}$$
$$\min \|y\|^2, \tag{3}$$
$$\langle z_i, y\rangle \geq 1, \quad i \in J. \tag{4}$$

Следующие две леммы доказывают существование взаимозависимостей между допустимыми и оптимальными решениями задач (1)–(2) и (3)–(4).

Лемма 3. *Пусть $y^* \neq \mathbf{0}$ — решение системы неравенств (2), $\overline{y}$ — решение системы неравенств (4). Тогда*
1) *вектор $\widehat{y} = \overline{y}/\|\overline{y}\|^2$ удовлетворяет системе неравенств (2),*
2) *вектор $\widetilde{y} = y^*/\|y^*\|^2$ удовлетворяет системе неравенств (4).*

Лемма доказывается путем подстановки.

Лемма 4. *Пусть $y^* \neq \mathbf{0}$ — решение задачи (1)–(2), $\overline{y}$ — решение задачи (3)–(4). Тогда*
1) *вектор $\widehat{y} = \overline{y}/\|\overline{y}\|^2$ является решением задачи (1)–(2),*
2) *вектор $\widetilde{y} = y^*/\|y^*\|^2$ — решение задачи (3)–(4).*

Доказательство. 1. Очевидно, $\overline{y} \neq \mathbf{0}$. По лемме 3 вектор $\widehat{y}$ является допустимым решением задачи (1)–(2). Предположим, что y' — допустимое решение задачи (1)–(2) такое, что $\|y'\|^2 > \|\widehat{y}\|^2 = 1/\|\overline{y}\|^2$. Тогда согласно лемме 3 вектор $\breve{y} = y'/\|y'\|^2$ удовлетворяет условиям задачи (3)–(4). Но $\|\breve{y}\|^2 = 1/\|y'\|^2 < 1/\|\widehat{y}\|^2 = \|\overline{y}\|^2$. Получаем противоречие с тем, что $\overline{y}$ — решение задачи (3)–(4).

2. По лемме 3 вектор $\widetilde{y}$ — допустимое решение задачи (3)–(4). Предположим, что $\ddot{y}$ — допустимое решение задачи (3)–(4) такое, что выполняется $\|\ddot{y}\|^2 < \|\widetilde{y}\|^2 = 1/\|y^*\|^2$. Но тогда по лемме 3 вектор $\dot{y} = \ddot{y}/\|\ddot{y}\|^2$ является допустимым решением задачи (1)–(2), для которого $\|\dot{y}\|^2 = 1/\|\ddot{y}\|^2 > 1/\|\widetilde{y}\|^2 = \|y^*\|^2$. Это противоречит условию леммы. □

Два следствия легко доказываются методом от противного с помощью лемм 3, 4.

Следствие 1. Если решение задачи (1)–(2) $y^* = \mathbf{0}$, то задача (3)–(4) решений не имеет.

Следствие 2. Если область допустимых решений задачи (3)–(4) пуста, то решение задачи (1)–(2) $y^* = \mathbf{0}$.

Множество, заданное системой линейных неравенств типа (4), выпукло и замкнуто. Если это множество непусто, то решение задачи минимизации сильно выпуклой функции на множестве, заданном системой (4), существует и единственно ([5], с. 41). Для решения выпуклой задачи квадратичного программирования такого типа существуют различные конечные методы ([4], с. 232), например, выпуклый симплексный метод, метод Лемке описаны в ([6], сс. 15, 134).

Замечание 3. Из единственности решения задачи (3)–(4) по утверждению леммы 4 следует единственность решения задачи (1)–(2).

Обозначим через $\Gamma(y) = \{z \in R_n : \langle y, z\rangle \geq \|y\|^2\} \quad \forall y \in \Omega$ замкнутое полупространство, определяемое гиперплоскостью $\langle y, z\rangle = \|y\|^2$. Заметим, что $\mathbf{0} \notin \Gamma(y)$ при $y \neq \mathbf{0}$, т. к. $\langle y, \mathbf{0}\rangle < \|y\|^2$. Следущие две теоремы дают обоснование возможности сведения решения задачи нахождения проекции нуля на выпуклый многогранник к решению задачи (3)–(4).

Теорема 4. *Если y^* — решение задачи (1)–(2), то $\rho = y^*$.*

Доказательство. Очевидно, что для всех $y \in \Omega$ задача

$$\min_{z \in \Gamma(y)} \|z\|^2$$

эквивалентна задаче нахождения проекции нуля на полупространство $\Gamma(y)$, решение которой определяется по явной формуле ([1], с. 196)

$$\mathbf{P}_{\Gamma(y)}(\mathbf{0}) = \mathbf{0} + (\|y\|^2 - \langle y, \mathbf{0} \rangle) \frac{y}{\|y\|^2} = y \quad \forall y \in \Omega. \tag{5}$$

Очевидно, что $L \subset \Gamma(\mathbf{0}) = R_n$. Согласно теореме 3 $\forall y \in \Omega$, $y \neq \mathbf{0}$ гиперплоскость $\langle y, z \rangle = \|y\|^2$ отделяет множество L и нуль пространства R_n. Тогда $\forall z \in L$ выполняется $\langle y, z \rangle \geq \|y\|^2$, т. е. $L \subset \Gamma(y) \quad \forall y \in \Omega$. Следовательно, $\min\limits_{z \in L} \|z\|^2 \geq \min\limits_{z \in \Gamma(y)} \|z\|^2 \quad \forall y \in \Omega$, т. е. в силу (5) выполняется соотношение

$$\|\rho\|^2 \geq \|y\|^2 \quad \forall y \in \Omega. \tag{6}$$

Поскольку по теореме 2 $\rho \in \Omega$, то из неравенства (6) следует

$$\max_{y \in \Omega} \|y\|^2 = \|\rho\|^2.$$

Отсюда с учетом замечания 3 получим, что для решения задачи (1)–(2) выполняется равенство $y^* = \rho$. $\square$

Теорема 5. 1. *Если в задаче* (3)–(4) *область допустимых решений пуста, то* $\mathbf{0} \in L$.
2. *Если* $\overline{y}$ *— решение задачи* (3)–(4), *то* $\mathbf{P}_L(\mathbf{0}) = \overline{y}/\|\overline{y}\|^2$.

Доказательство. 1. В силу следствия 2 решением задачи (1)–(2) будет $y^* = \mathbf{0}$. В этом случае согласно теореме 4 $\|\rho\|^2 = \max\limits_{y \in \Omega} \|y\|^2 = 0$, т. е. $\mathbf{0} \in L$.

2. По лемме 4 вектор $\widehat{y} = \overline{y}/\|\overline{y}\|^2$ является решением задачи (1)–(2). Следовательно, по теореме 4 $\mathbf{P}_L(\mathbf{0}) = \widehat{y}$. $\square$

Теорема 6. *Если* $y \in \Gamma_{\wp}\, \Omega$, *то вектор* y *будет опорным вектором множества* L *в точках* $z \in L$, *удовлетворяющих соотношениям*

$$z = \sum_{i \in J(y)} \alpha_i z_i, \quad \alpha_i \geq 0 \quad \forall i \in J(y) = \{i \in J : \langle z_i, y \rangle = \|y\|^2\}, \qquad \sum_{i \in J(y)} \alpha_i = 1.$$

Доказательство. Так как $y \in \Gamma_{\wp}\, \Omega$, то $J(y) \neq \emptyset$. Поскольку $L \subset \Gamma(y) \quad \forall y \in \Omega$, то выполняется $\langle z, y \rangle \geq \|y\|^2 \quad \forall z \in L$, причем равенство достигается в точках $z = z_i$, $i \in J(y)$. Тогда гиперплоскость $\langle z, y \rangle = \|y\|^2$ является опорной к множеству L в точках $z = z_i$, $i \in J(y)$. Следовательно, вектор y, являющийся нормальным вектором этой гиперплоскости, будет опорным вектором множества L в точках $z = z_i$, $i \in J(y)$. Очевидно, вектор y является опорным также во всех точках $z \in L$ таких, что $z = \sum\limits_{i \in J(y)} \alpha_i z_i$, $\alpha_i \geq 0 \quad \forall i \in J(y)$, $\sum\limits_{i \in J(y)} \alpha_i = 1$. $\square$

Замечание 4. По теореме 2 $\rho \in \Gamma_{\wp}\, \Omega$. Тогда согласно теореме 6 вектор ρ является опорным вектором множества L во всех точках $z \in L$ таких, что

$$z = \sum_{i \in J(\rho)} \alpha_i z_i, \quad \alpha_i \geq 0 \quad \forall i \in J(\rho), \qquad \sum_{i \in J(\rho)} \alpha_i = 1.$$

Следующий алгоритм является одним из возможных приложений теоремы отделимости из данной работы.

Алгоритм 1 для нахождения проекции нуля на выпуклый многогранник

Шаг 1. Сформулировать задачу (3)–(4) и решить любым конечным методом.

Шаг 2. Если множество допустимых решений задачи (3)–(4)
$D = \{y \in R_n : \langle z_i, y\rangle \geq 1,\ i \in J\} = \emptyset$, то $\mathbf{P}_L(\mathbf{0}) = \mathbf{0}$. В противном случае полагаем $\mathbf{P}_L(\mathbf{0}) = y^*/\|y^*\|^2$, где y^* — решение задачи (3)–(4).

Перепишем ограничения задачи (3)–(4) в матрично-векторной форме

$$\min \|y\|^2,$$
$$Cy + e \leq 0,$$

где C — $m \times n$-матрица, строками которой являются векторы $-[z_i]^T$, $i \in J$; e — m-мерный вектор, $e^T = (1, \dots, 1)$.

Для решения задачи (3)–(4) построим функцию Лагранжа $\Lambda(y, u) = \|y\|^2 + \langle u, Cy + e\rangle$, $y \in R_n$, $u \geq 0$. Ее седловая точка (y^*, u^*) определяется соотношением $\Lambda(y^*, u^*) = \min\limits_{y \in R_n} \max\limits_{u \geq 0} \Lambda(y, u)$.

Строгая выпуклость и квадратичность по y функции $\Lambda(y, u)$ обеспечивает достижение ее минимума в точке $y(u)$ такой, что $\Lambda'_y(y, u) = 2y + C^T u = 0 \Rightarrow y(u) = -\dfrac{1}{2} C^T u$. Следовательно, $\Lambda(y(u), u) = -\dfrac{1}{4}\langle u, CC^T u\rangle + \langle e, u\rangle$. Тогда для нахождения y^* необходимо решить задачу минимизации квадратичной выпуклой функции на множестве простой структуры

$$\max_{u \geq 0} \Lambda(y(u), u) \quad \text{либо}$$
$$\min_{u \geq 0} \left(f(u) = \frac{1}{4}\langle u, CC^T u\rangle - \langle e, u\rangle \right). \tag{7}$$

Теорема 7 ([1], с. 173). *Пусть U — выпуклое множество из R_n, $\operatorname{int} U \neq \emptyset$, $f(u) \in C^2(U)$. Тогда для выпуклости $f(u)$ на U необходимо и достаточно, чтобы $\langle f''(u)\xi, \xi\rangle \geq 0$ при всех $\xi = (\xi^1, \dots, \xi^n) \in R_n$ и всех $u \in U$.*

Вычислим $f'(u) = \dfrac{1}{2} CC^T u - e$, $f''(u) = \dfrac{1}{2} CC^T$. Отсюда имеем $\langle f''(u)\xi, \xi\rangle = \dfrac{1}{2}\langle CC^T \xi, \xi\rangle = = \dfrac{1}{2}\|C^T \xi\|^2 \geq 0$ при всех $\xi \in R_n$, $u \in R_m$. Тогда в силу теоремы 7 функция $f(u)$ в задаче (7) выпукла в R_m. Согласно теореме 4.6 из ([4], с. 181) если допустимое множество задачи (3)–(4) непусто, то задача (7) имеет решение u^*, а решение задачи (3)–(4) можно выписать в явном виде $y^* = -\dfrac{1}{2} C^T u^*$. Поскольку множество D в задаче (3)–(4) задано линейными ограничениями, то согласно теореме Куна–Таккера ([5], с. 61) необходимое и достаточное условие оптимальности точки $y^* \in D$ состоит в существовании такого $u^* \geq 0$, что пара (y^*, u^*) является седловой точкой функции Лагранжа на множестве $y \in R_n$, $u \geq 0$.

Алгоритм 2 для нахождения проекции нуля на выпуклый многогранник

Шаг 1. Сформулировать задачу (7) и решить ее любым конечным методом.

Шаг 2. Если целевая функция задачи (7) неограничена снизу на множестве $\{u \in R_m : u \geq 0\}$, то $\mathbf{P}_L(\mathbf{0}) = \mathbf{0}$. В противном случае полагаем $\mathbf{P}_L(\mathbf{0}) = y^*/\|y^*\|^2$, где $y^* = -\dfrac{1}{2} C^T u^*$, u^* — решение задачи (7).

На примере метода Лемке покажем, что для реализации одной итерации метода при решении задачи (7) требуется существенно меньше вычислительных затрат, чем при решении задачи (3)–(4). Отметим, что задача (7) поставлена в канонической для метода Лемке форме, а задачу (3)–(4) необходимо привести к канонической форме путем замены, например, каждой переменной разностью двух дополнительных неотрицательных переменных. Тогда матрица, пересчитываемая на каждой итерации метода Лемке, для задачи (3)–(4) будет иметь $m + 2n$ строк и

$2(m+2n+1)$ столбцов. Соответствующая матрица пересчета для задачи (7) будет иметь размерность $n \times 2(n+1)$. Следовательно, для реализации одной итерации метода Лемке при решении задачи (7) подсчитывается на $2(m+n)^2$ меньше величин, чем при решении задачи (3)–(4).

Литература

1. Васильев Ф.П. *Численные методы решения экстремальных задач.* – М.: Наука, 1980. – 518 с.
2. Демьянов В.Ф., Малоземов В.Н. *Введение в минимакс.* – М.: Наука, 1972. – 368 с.
3. Пшеничный Б.Н. *Выпуклый анализ и экстремальные задачи.* – М.: Наука, 1980. – 319 с.
4. Сухарев А.Г., Тимохов А.В., Федоров В.В. *Курс методов оптимизации.* – М.: Наука, 1986. – 328 с.
5. Карманов В.Г. *Математическое программирование.* – М: Наука, 1986. – 286 с.
6. Реклейтис Г., Рейнвиндран А., Регсдел К. *Оптимизация в технике.* Ч. II. – М.: Наука, 1986. – 320 с.

Казанский государственный университет

Поступила 26.06.2006