ЗАДАЧА h-ОТДЕЛЕНИЯ ДВУХ МНОЖЕСТВ И ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ*

Е. К. Чернэуцану

katerinache@yandex.ru

6 февраля 2014 г.

Аннотация. Рассматривается задача наилучшего отделения выпуклой оболочки одного конечного множества от другого конечного множества с помощью h гиперплоскостей. Показывается, что эта задача сводится к конечному числу задач линейного программирования.

1°. Пусть в $\mathbb{R}^n$ заданы два конечных множества

$$A = \{a_i\}_{i=1}^m \quad \text{и} \quad B = \{b_j\}_{j=1}^k.$$

Следуя [1], назовём выпуклую оболочку множества A и множество B *строго h-отделимыми*, если существует h гиперплоскостей вида

$$H_s = \big\{x \in \mathbb{R}^n \mid \langle w^s,\ x\rangle = \gamma_s\big\}, \quad w^s \neq \mathbb{O}, \quad s \in 1:h,$$

таких, что выполняются неравенства

$$\begin{array}{lllll} \langle w^s,\ a_i\rangle < \gamma_s & \text{при всех} & i \in 1:m & \text{и всех} & s \in 1:h, \\ \langle w^s,\ b_j\rangle > \gamma_s & \text{при каждом} & j \in 1:k & \text{и некотором} & s \in 1:h. \end{array} \tag{1}$$

На рис. 1 приведён пример строгого 2-отделения.

Введём функцию

$$F(G) = \frac{1}{m}\sum_{i=1}^m \max_{s\in 1:h}\big[\langle w^s,\ a_i\rangle - \gamma_s + c\big]_+ + \frac{1}{k}\sum_{j=1}^k \min_{s\in 1:h}\big[-\langle w^s,\ b_j\rangle + \gamma_s + c\big]_+.$$

Здесь G — матрица размера $h \times (n+1)$ со строками

$$g^s = (w^s,\ \gamma_s), \quad s \in 1:h;$$

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

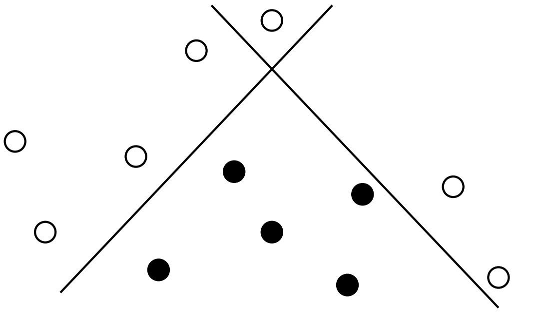

Рис. 1. Строго 2-отделимые множества

$c > 0$ — параметр и $[u]_+ = \max\{0, u\}$ — плюсиковая функция. Матрицу G указанного вида будем называть *подходящей*, если у неё все элементы w^s ненулевые ($w^s \neq \mathbb{O}$, $s \in 1 : h$). Ясно, что $F(G) \geq 0$ при всех G.

ТЕОРЕМА 1. *Выпуклая оболочка множества A и множество B строго h-отделимы тогда и только тогда, когда существует подходящая матрица G_*, такая, что $F(G_*) = 0$.*

Доказательство. Необходимость. Пусть выполнены соотношения (1) и $w^s \neq \mathbb{O}$ при всех $s \in 1 : h$. Обозначим

$$\delta := \min_{i \in 1:m,\ s \in 1:h} \left[-\langle w^s,\ a_i \rangle + \gamma_s \right] > 0.$$

Каждому $j \in 1 : k$ соответствует индекс $s_j \in 1 : h$, такой, что

$$\delta_j := \langle w^{s_j},\ b_j \rangle - \gamma_{s_j} > 0.$$

При $\delta_* = \min\{\delta,\ \delta_1, \ldots,\ \delta_k\}$, $\delta_* > 0$, выполняются неравенства

$$\begin{aligned} \delta_* &\leqslant -\langle w^s,\ a_i \rangle + \gamma_s, \quad i \in 1 : m,\ s \in 1 : h, \\ \delta_* &\leqslant \langle w^{s_j},\ b_j \rangle - \gamma_{s_j}, \quad j \in 1 : k. \end{aligned}$$

Положим $w_*^s = \frac{c}{\delta_*} w^s$, $\gamma_s^* = \frac{c}{\delta_*} \gamma_s$. Получим

$$\begin{gathered} \langle w_*^s,\ a_i \rangle - \gamma_s^* + c \leqslant 0, \quad i \in 1 : m,\ s \in 1 : h; \\ -\langle w_*^{s_j},\ b_{s_j} \rangle + \gamma_{s_j}^* + c \leqslant 0, \quad j \in 1 : k. \end{gathered}$$

Отсюда по определению плюсиковой функции следует, что на подходящей матрице G_* со строками $(w_*^s,\ \gamma_s^*)$, $s \in 1 : h$, выполняется равенство $F(G_*) = 0$. Достаточность. Если $F(G_*) = 0$ на некоторой подходящей матрице G_*, то

$$\begin{aligned} \max_{s \in 1:h} \left[\langle w_*^s,\ a_i \rangle - \gamma_s^* + c \right]_+ &= 0 \quad \text{при всех} \quad i \in 1 : m; \\ \min_{s \in 1:h} \left[-\langle w_*^s,\ b_j \rangle + \gamma_s^* + c \right]_+ &= 0 \quad \text{при всех} \quad j \in 1 : k. \end{aligned}$$

Это значит, что

$$\big[\langle w_*^s,\ a_i\rangle - \gamma_s^* + c\big]_+ = 0 \quad \text{при всех} \quad i \in 1:m \quad \text{и всех} \quad s \in 1:h;$$
$$\big[-\langle w_*^s,\ b_j\rangle + \gamma_s^* + c\big]_+ = 0 \quad \text{при каждом} \quad j \in 1:k \quad \text{и некотором} \quad s \in 1:h.$$

Перепишем две последние формулы в эквивалентном виде

$$\langle w_*^s,\ a_i\rangle \leqslant \gamma_s^* - c \quad \text{при всех} \quad i \in 1:m \quad \text{и всех} \quad s \in 1:h;$$
$$\langle w_*^s,\ b_j\rangle \geqslant \gamma_s^* + c \quad \text{при каждом} \quad j \in 1:k \quad \text{и некотором} \quad s \in 1:h.$$

Данные соотношения гарантируют выполнение условий (1) с $w^s = w_*^s$, $\gamma_s = \gamma_s^*$.

Теорема доказана. □

2°. В дальнейшем мы будем исследовать экстремальную задачу

$$F(G) \to \min, \tag{2}$$

где минимум берётся по всем матрицам G размера $h \times (n+1)$.

Задача (2) сводится к конечному числу задач линейного программирования. Доказательство опирается на лемму о сумме минимумов.

Пусть

$$M = \sum_{j=1}^{k} \min_{s\in 1:h_j} p(s,\ j).$$

На рис. 2 схематично представлено индексное множество, на котором определена функция $p(s,\ j)$.

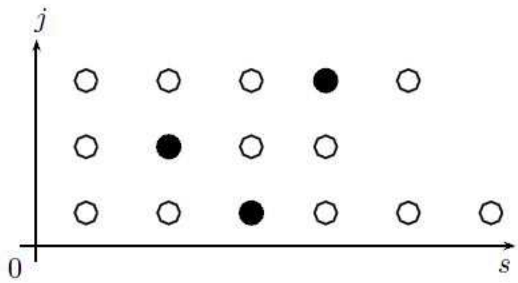

Рис. 2. Двухиндексное множество

Обозначим через Π множество всех цепочек $S = (s_1,\ s_2, \dots,\ s_k)$, где $s_j \in 1:h_j$ при каждом $j \in 1:k$. Выделим в Π цепочку $S^* = (s_1^*,\ s_2^*, \dots,\ s_k^*)$, элементы которой удовлетворяют условиям

$$s_j^* \in 1:h_j, \quad p(s_j^*,\ j) = \min_{s\in 1:h_j} p(s,\ j).$$

Введём функцию
$$P(S) = \sum_{j=1}^{k} p(s_j,\ j).$$
Ясно, что $M = P(S^*)$.

ЛЕММА (о сумме минимумов)**.** *Справедливо равенство*
$$M = \min_{S \in \Pi} P(S). \tag{3}$$

Доказательство. При каждом $S \in \Pi$ имеем
$$P(S) = \sum_{j=1}^{k} p(s_j,\ j) \geq \sum_{j=1}^{k} \min_{s \in 1:h_j} p(s,\ j) = M,$$
так что $P(S) \geq M$ при всех $S \in \Pi$. Вместе с тем, как отмечалось, $P(S^*) = M$, причём $S^* \in \Pi$. Приходим к равенству (3). □

Перепишем формулу (3) в развёрнутом виде
$$\sum_{j=1}^{k} \min_{s \in 1:h_j} p(s,\ j) = \min_{S \in \Pi} \sum_{j=1}^{k} p(s_j,\ j).$$
Таким образом, лемма показывает, как правильно переставлять местами знаки "$\sum$" и "min".

3°. Вернёмся к задаче (2). Её подробная запись выглядит так:
$$F(G) := \frac{1}{m} \sum_{i=1}^{m} \varphi_i(G) + \frac{1}{k} \sum_{j=1}^{k} \psi_j(G) \to \min, \tag{4}$$
где
$$\varphi_i(G) = \max_{s \in 1:h} \big[\langle w^s,\ a_i\rangle - \gamma_s + c\big]_+,$$
$$\psi_j(G) = \min_{s \in 1:h} \big[-\langle w^s,\ b_j\rangle + \gamma_s + c\big]_+.$$
Напомним обозначение $\Pi = \big\{S = (s_1, \ldots,\ s_k) \mid s_j \in 1 : h \text{ при всех } j \in 1 : k\big\}$.

ТЕОРЕМА 2. *Справедливо равенство*
$$\inf_G F(G) = \min_{S \in \Pi} \inf_G \Big\{\frac{1}{m} \sum_{i=1}^{m} \varphi_i(G) + \frac{1}{k} \sum_{j=1}^{k} \big[-\langle w^{s_j},\ b_j\rangle + \gamma_{s_j} + c\big]_+\Big\}. \tag{5}$$

Д о к а з а т е л ь с т в о. По лемме о сумме минимумов

$$\sum_{j=1}^{k} \psi_j(G) = \min_{S\in\Pi} \sum_{j=1}^{k} \big[-\langle w^{s_j},\ b_j\rangle + \gamma_{s_j} + c\big]_+.$$

Отсюда и из определения функции $F(G)$ следует, что

$$F(G) = \min_{S\in\Pi} \Big\{ \frac{1}{m} \sum_{i=1}^{m} \varphi_i(G) + \frac{1}{k} \sum_{j=1}^{k} \big[-\langle w^{s_j},\ b_j\rangle + \gamma_{s_j} + c\big]_+ \Big\}.$$

Значит,

$$\inf_G F(G) = \inf_G \min_{S\in\Pi} \Big\{ \frac{1}{m} \sum_{i=1}^{m} \varphi_i(G) + \frac{1}{k} \sum_{j=1}^{k} \big[-\langle w^{s_j},\ b_j\rangle + \gamma_{s_j} + c\big]_+ \Big\}. \tag{6}$$

Остаётся в правой части (6) поменять местами инфимум по G и минимум по $S \in \Pi$. $\square$

Теорема 2 показывает, что задача (2) эквивалентна конечному числу экстремальных задач вида

$$\frac{1}{m} \sum_{i=1}^{m} \max_{s\in 1:h} \big[\langle w^{s},\ a_i\rangle - \gamma_s + c\big]_+ + \frac{1}{k} \sum_{j=1}^{k} \big[-\langle w^{s_j},\ b_j\rangle + \gamma_{s_j} + c\big]_+ \to \inf_G, \tag{7}$$

соответствующих различным $S \in \Pi$. В свою очередь, задача (7) эквивалентна задаче линейного программирования

$$\begin{gathered} \frac{1}{m} \sum_{i=1}^{m} p_i + \frac{1}{k} \sum_{j=1}^{k} q_j \to \inf, \qquad (8) \\ -\langle a_i,\ w^s\rangle + \gamma_s + p_i \geqslant c, \quad i \in 1:m,\ s \in 1:h; \\ \langle b_j,\ w^{s_j}\rangle - \gamma_{s_j} + q_j \geqslant c, \quad j \in 1:k; \\ p_i \geqslant 0, \quad i \in 1:m; \quad q_j \geqslant 0, \quad j \in 1:k. \end{gathered}$$

Множество планов задачи (8) непусто (планом является, например, набор $w^s = \mathbb{O}$ при всех $s = 1:h$, $\gamma_s \equiv 0$, $p_i \equiv c$, $q_j \equiv c$) и целевая функция ограничена снизу нулём. Значит, все задачи (8) имеют решение. По эквивалентности задача (7) при всех $S \in \Pi$ также имеет решение. На основании теоремы 2 заключаем, что существует решение и у задачи (2).

4°. Пусть G_* — какое-нибудь решение задачи (2). Возможны следующие случаи:

1) $F(G_*) = 0$ и G_* — подходящая матрица. Тогда строки матрицы G_* являются решением задачи строгого h-отделения.

2) $F(G_*) = 0$ и G_* — не является подходящей матрицей. При этом не все w^s равны нулевому вектору (иначе $F(G_*) \geqslant 2c > 0$). Если $w_*^s = \mathbb{O}$ на множестве $J \subset 1 : h$, то $\mathrm{co}(A)$ и B строго $\big(h - |J|\big)$-отделимы (см. [1, 2]).

3) $F(G_*) > 0$. Тогда по по теореме 1 строгое h-отделение невозможно. Если соответствующая матрица G_* подходящая, то будем говорить, что она обеспечивает *наилучшее приближённое h-отделение* множеств $\mathrm{co}(A)$ и B.

5°. Рассмотрим пример. Пусть на плоскости заданы множества A и B, состоящие соответсвенно из точек

$$a_1 = (-2, 0),\ a_2 = (2, 0),\ a_3 = (0, 2),\ a_4 = (0, 1);$$
$$b_1 = (0, 3),\ b_2 = (3, 0),\ b_3 = (-3, 0).$$

Очевидно, что $\mathrm{co}(A) \cap B = \varnothing$ (см. рис. 3).

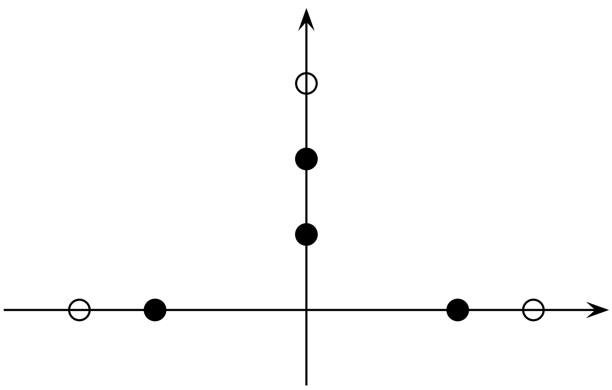

Рис. 3

Решим задачу строгой 2-отделимости. В данном случае

$$n = 2,\ m = 4,\ k = 3,\ h = 2.$$

Выясним, как выглядит задача (8) при $S = (1, 1, 2)$. Выпишем вектор неизвестных

$$z = (w_1^1,\, w_2^1,\, \gamma_1,\, w_1^2,\, w_2^2,\, \gamma_2,\, p_1,\, p_2,\, p_3,\, p_4,\, q_1,\, q_2,\, q_3)$$

и матрицу ограничений

$$D = \begin{bmatrix}
2 & 0 & 1 & & & & 1 & & & & & & \\
-2 & 0 & 1 & & & & & 1 & & & & & \\
0 & -2 & 1 & & & & & & 1 & & & & \\
0 & -1 & 1 & & & & & & & 1 & & & \\
 & & & 2 & 0 & 1 & 1 & & & & & & \\
 & & & -2 & 0 & 1 & & 1 & & & & & \\
 & & & 0 & -2 & 1 & & & 1 & & & & \\
 & & & 0 & -1 & 1 & & & & 1 & & & \\
0 & 3 & -1 & & & & & & & & 1 & & \\
3 & 0 & -1 & & & & & & & & & 1 & \\
 & & & -3 & 0 & -1 & & & & & & & 1
\end{bmatrix}.$$

Задача (8) примет вид

$$\frac{1}{4}\sum_{i=1}^{4} p_i + \frac{1}{3}\sum_{j=1}^{3} q_j \to \inf, \tag{9}$$
$$Dz \geqslant c\, e,$$
$$p_i \geqslant 0, \quad i \in 1:4; \qquad q_j \geqslant 0, \quad j \in 1:3.$$

где e — вектор, все компоненты которого равны единице. При $c = 1$ решением данной задачи является вектор

$$z_* = (111.2210,\ 112.0012,\ 272.9097,\ -78.1474,\ 27.3511,\ 192.1601,$$
$$0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 0.0000).$$

Минимальное значение целевой функции равно нулю. Строгое 2-отделение множеств A и B при $S = (1,\, 1,\, 2)$ выглядит так, как показано на рис. 4.

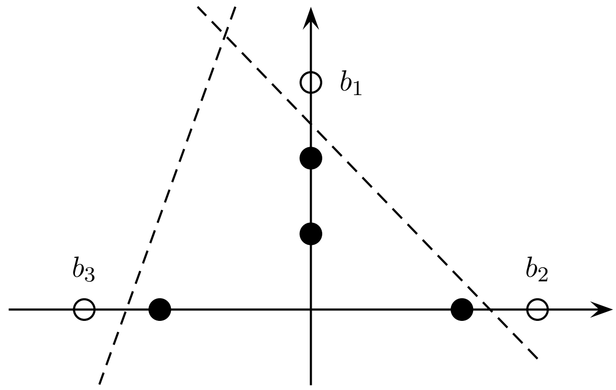

Рис. 4

Задание вектора индексов $S = (1,\ 1,\ 2)$ соответсвует разбиению множества B на два подмножества $\{b_1,\ b_2\} \cup \{b_3\}$. Эти два подмножества согласовано отделяются от $\mathrm{co}(A)$ с помощью двух прямых

$$\langle w^1,\ x\rangle = \gamma_1 \quad \text{и} \quad \langle w^2,\ x\rangle = \gamma_2.$$

Существуют ещё два разбиения множества B на два подмножества:

$$\{b_1,\ b_3\} \cup \{b_2\} \quad \text{и} \quad \{b_2,\ b_3\} \cup \{b_1\}.$$

Им соответствуют векторы $S = (1,\ 2,\ 1)$ и $S = (2,\ 1,\ 1)$.

Результат строгого 2-отделения при $S = (1,\ 2,\ 1)$ показан на рис. 5.

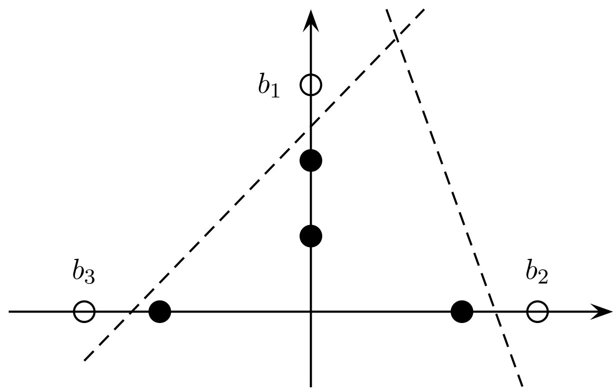

Рис. 5

Этот случай симметричен случаю $S = (1,\ 1,\ 2)$.

При $S = (2,\ 1,\ 1)$ строгой 2-отделимости нет (см. рис. 6).

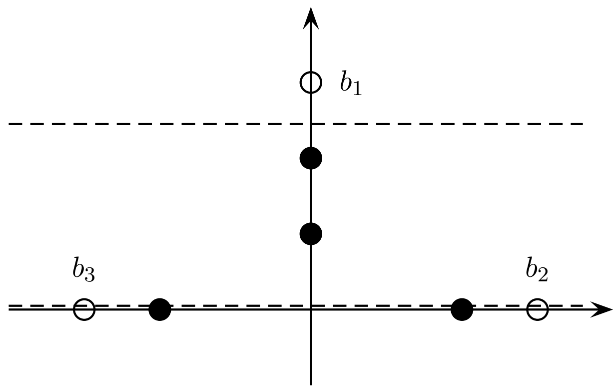

Рис. 6

Решением задачи, аналогичной (9), является вектор

$$z = (0.0000,\ -112.0230,\ -1.0000,\ 0.0000,\ 111.8673,\ 273.7824,\\ 2.0000,\ 2.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 0.0000).$$

Минимальное значение целевой функции равно единице.

6°. В общем случае задание вектора $S \in \Pi$ соответствует разбиению множества B, состоящего из k векторов, на h подмножеств. Число таких разбиений и определяет количество задач линейного программирования вида (8), к решению которых сводится решение задачи (2).

Если заранее известно, что множества A и B строго h-отделимы, то решение задачи (2) можно упростить. После разбиения множества B на h подмножеств следует *независимо* решать задачи линейного отделения каждого из этих подмножеств от множества A. В случае успешного отделения совокупность разделяющих гиперплоскостей образует решение задачи (2).

В рассмотренном выше примере при $S = (1,\, 1,\, 2)$ будем независимо решать задачи линейного отделения множеств $\{b_1,\, b_2\}$ и $\{b_3\}$ от A. Запишем соответствующие задачи линейного программирования (см. [3])

$$\begin{gathered}
\frac{1}{4}\sum_{i=1}^{4} p_i + \frac{1}{2}\sum_{j=1}^{2} q_j \to \inf,\\
-\langle a_i, w^1\rangle + \gamma_1 + p_i \geqslant 1, \quad i \in 1:4;\\
\langle b_j, w^1\rangle - \gamma_1 + q_j \geqslant 1, \quad j \in 1:2;\\
p_i \geqslant 0, \quad i \in 1:4; \qquad q_j \geqslant 0, \quad j \in 1:2,
\end{gathered}$$

и

$$\frac{1}{4}\sum_{i=1}^{4} p_i + q_3 \to \inf,$$
$$-\langle a_i, w^2\rangle + \gamma_2 + p_i \geqslant 1, \quad i \in 1:4;$$
$$\langle b_3, w^2\rangle - \gamma_2 + q_3 \geqslant 1;$$
$$p_i \geqslant 0, \quad i \in 1:4; \qquad q_3 \geqslant 0.$$

Их решения $\{w^1, \gamma_1\}$ и $\{w^2, \gamma_2\}$ определяют две прямые, строго отделяющие $\mathrm{co}(A)$ от B (см. рис. 7).

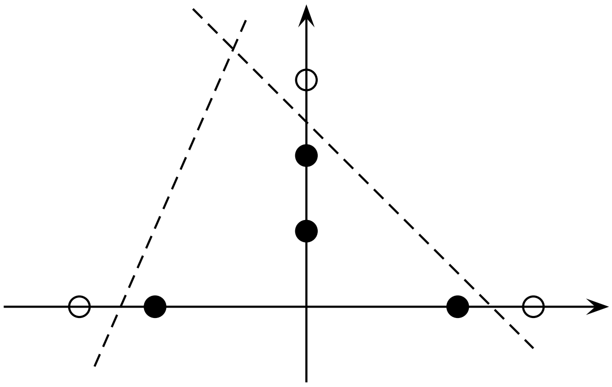

Рис. 7

ЛИТЕРАТУРА

1. Astorino A., Gaudioso M. *Polyhedral separability through successive LP* // JOTA. 2002. Vol. 112. No 2. P. 265–293.

2. Малозёмов В. Н., Чернэуцану Е. К. *Строгая h-отделимость двух множеств* // Семинар «DHA & CAGD». Избранные доклады. 18 декабря 2010 г. (`http://dha.spb.ru/reps10.shtml#1218`)

3. Малозёмов В. Н., Чернэуцану Е. К. *О математической диагностике (линейная модель)* // Семинар «DHA & CAGD». Избранные доклады. 17 апреля 2010 г. (`http://dha.spb.ru/reps10.shtml#0417`)