О МАТЕМАТИЧЕСКОЙ ДИАГНОСТИКЕ (ЛИНЕЙНАЯ МОДЕЛЬ)*

В. Н. Малозёмов — malv@math.spbu.ru

Е. К. Чернэуцану — katerinache@yandex.ru

17 апреля 2010 г.

В докладе анализируются основные результаты известной работы [1].

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

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

Эти множества называются *строго отделимыми*, если существует вектор $v \in \in \mathbb{R}^n$, такой, что

$$\min_{i \in 1:m} \langle v, a_i \rangle > \max_{j \in 1:k} \langle v, b_j \rangle. \tag{1}$$

ЛЕММА 1. *Множества $\mathcal{A}$ и $\mathcal{B}$ строго отделимы тогда и только тогда, когда найдутся вектор $w \in \mathbb{R}^n$ и число γ со свойствами*

$$\begin{aligned} \langle w, a_i \rangle &\geqslant \gamma + 1 \qquad \text{при всех} \quad i \in 1:m, \\ \langle w, b_j \rangle &\leqslant \gamma - 1 \qquad \text{при всех} \quad j \in 1:k. \end{aligned} \tag{2}$$

Доказательство. То, что из (2) следует (1) с $v = w$, очевидно. Проверим обратное заключение. Обозначим

$$d = \min_i \langle v, a_i \rangle - \max_j \langle v, b_j \rangle, \quad w = \big(2/d\big) v,$$
$$\gamma = \tfrac{1}{2} \big[\min_i \langle w, a_i \rangle + \max_j \langle w, b_j \rangle \big].$$

Тогда

$$\begin{gathered} \min_i \langle w, a_i \rangle - \max_j \langle w, b_j \rangle = 2, \\ \min_i \langle w, a_i \rangle = 2\gamma - \max_j \langle w, b_j \rangle = 2\gamma + \big(2 - \min_i \langle w, a_i \rangle\big), \\ \max_j \langle w, b_j \rangle = 2\gamma - \min_i \langle w, a_i \rangle = 2\gamma - \big(2 + \max_j \langle w, b_j \rangle\big). \end{gathered}$$

Последние два соотношения равносильны (2). □

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

2°. Рассмотрим экстремальную задачу

$$\tfrac{1}{m}\sum_{i=1}^{m}\bigl(-\langle w,a_i\rangle+\gamma+1\bigr)_+ + \tfrac{1}{k}\sum_{j=1}^{k}\bigl(\langle w,b_j\rangle-\gamma+1\bigr)_+ \to \min_{w,\gamma}. \tag{3}$$

Здесь $(u)_+ = \max\{0,u\}$. Задача (3) эквивалентна следующей задаче линейного программирования

$$\begin{gathered}
\tfrac{1}{m}\sum_{i=1}^{m} y_i + \tfrac{1}{k}\sum_{j=1}^{k} z_j \to \min,\\
\langle w,a_i\rangle-\gamma+y_i \geqslant 1, \quad i\in 1:m;\\
-\langle w,b_j\rangle+\gamma+z_j \geqslant 1, \quad j\in 1:k;\\
y_i\geqslant 0,\ \ i\in 1:m; \quad z_j\geqslant 0,\ \ j\in 1:k.
\end{gathered} \tag{4}$$

Множество планов задачи (4) непусто (можно взять, например, $w=\mathbb{O}$, $\gamma=0$, $y_i\equiv 1$, $z_j\equiv 1$) и целевая функция ограничена снизу нулём. Значит, задача (4) имеет решение. По эквивалентности и у задачи (3) существует решение.

Обозначим через μ минимальное значение целевых функций в задачах (3) и (4). Очевидно, что $\mu\geqslant 0$. Из определений следует такое утверждение.

ПРЕДЛОЖЕНИЕ 1. *Равенство $\mu=0$ выполняется тогда и только тогда, когда найдутся вектор $w\in\mathbb{R}^n$ и число γ со свойствами* (2).

При выполнении соотношений (2) говорят, что гиперплоскость

$$\langle w,x\rangle=\gamma \tag{5}$$

строго разделяет множества $\mathcal{A}$ и $\mathcal{B}$.

3°. Обратимся к случаю $\mu>0$.

ПРЕДЛОЖЕНИЕ 2. *При $\mu>0$ у задачи* (3) *существует решение $\{w,\gamma\}$ с $w\neq\mathbb{O}$.*

По решению задачи (3) формируется гиперплоскость (5), которая *приближённо* разделяет множества $\mathcal{A}$ и $\mathcal{B}$.

Мы будем доказывать предложение 2 с помощью эквивалентной задачи (4).

ЛЕММА 2. *Для того чтобы задача* (4) *имела решение с $w=\mathbb{O}$, необходимо и достаточно, чтобы совпадали средние точки множеств $\mathcal{A}$ и $\mathcal{B}$, то есть чтобы*

$$\tfrac{1}{m}\sum_{i=1}^{m} a_i = \tfrac{1}{k}\sum_{j=1}^{k} b_j\,. \tag{6}$$

Доказательство. Необходимость. При $w = \mathbb{O}$ целевая функция задачи (3) принимает вид

$$(1+\gamma)_+ + (1-\gamma)_+ .$$

Её минимальное значение равно 2 и достигается при всех $\gamma \in [-1, 1]$. Значит, $\mu = 2$.

По условию задача (4) имеет решение, и мы только что установили, что минимальное значение её целевой функции равно 2. Воспользуемся критерием оптимальности для задачи линейного программирования (см., например, [2, с. 27]), согласно которому совместна система

$$\sum_{i=1}^{m} u_i + \sum_{j=1}^{k} v_j = 2; \tag{7}$$

$$\sum_{i=1}^{m} u_i\, a_i - \sum_{j=1}^{k} v_j\, b_j = \mathbb{O}; \tag{8}$$

$$-\sum_{i=1}^{m} u_i + \sum_{j=1}^{k} v_j = 0; \tag{9}$$

$$0 \leqslant u_i \leqslant \tfrac{1}{m}\,, \ i \in 1:m; \quad 0 \leqslant v_j \leqslant \tfrac{1}{k}\,, \ j \in 1:k. \tag{10}$$

Из (7) и (9) следует, что

$$\sum_{i=1}^{m} u_i = 1, \qquad \sum_{j=1}^{k} v_j = 1.$$

Принимая во внимание (10), заключаем, что все u_i равны $\frac{1}{m}$ и все v_j равны $\frac{1}{k}$. Теперь формула (8) становится эквивалентной (6).

Достаточность. Запишем задачу, двойственную к (4):

$$\sum_{i=1}^{m} u_i + \sum_{j=1}^{k} v_j \to \max$$

при ограничениях (8)–(10). В силу (6) набор $u_i \equiv \frac{1}{m}$, $v_j \equiv \frac{1}{k}$ является планом этой задачи. Значение целевой функции на нём равно 2.

Вместе с тем, на плане

$$w = \mathbb{O}, \quad \gamma = 0, \quad y_i \equiv 1, \quad z_j \equiv 1 \tag{11}$$

задачи (4) значение целевой функции также равно 2. Отсюда следует, что план (11) задачи (4) с $w = \mathbb{O}$ является оптимальным.

Лемма доказана. $\square$

4°. Переходим к доказательству предложения 2.

Возьмём решение задачи (4), и пусть оказалось, что у него $w = \mathbb{O}$. В этом случае $\mu = 2$ и выполняется соотношение (6). Покажем, что задача (4) имеет другое решение с $w \neq \mathbb{O}$.

Зафиксируем ненулевой вектор $h \in \mathbb{R}^n$ и рассмотрим систему линейных соотношений

$$\begin{gathered} -\tfrac{1}{m}\sum_{i=1}^{m} y_i - \tfrac{1}{k}\sum_{j=1}^{k} z_j = -2; \\ \langle w, a_i\rangle - \gamma + y_i \geqslant 1, \quad i \in 1:m; \\ -\langle w, b_j\rangle + \gamma + z_j \geqslant 1, \quad j \in 1:k; \\ y_i \geqslant 0,\ \ i \in 1:m; \quad z_j \geqslant 0,\ \ j \in 1:k; \\ -\langle h, w\rangle > 0. \end{gathered} \tag{12}$$

Предложение 2 будет доказано, если удастся установить, что система (12) совместна.

Допустим противное. Несовместность системы (12) означает, что у задачи линейного программирования

$$\langle h, w\rangle \to \min,$$

ограничения которой составляют условия (12) без строгого неравенства, минимальное значение целевой функции равно нулю. По критерию оптимальности для задачи линейного программирования совместна система

$$\sum_{i=1}^{m} u_i + \sum_{j=1}^{k} v_j - 2\zeta = 0; \tag{13}$$

$$\sum_{i=1}^{m} u_i\, a_i - \sum_{j=1}^{k} v_j\, b_j = h; \tag{14}$$

$$-\sum_{i=1}^{m} u_i + \sum_{j=1}^{k} v_j = 0; \tag{15}$$

$$0 \leqslant u_i \leqslant \tfrac{1}{m}\,\zeta,\ \ i \in 1:m; \quad 0 \leqslant v_j \leqslant \tfrac{1}{k}\,\zeta,\ \ j \in 1:k. \tag{16}$$

Но данная система совместна только при $h = \mathbb{O}$. Это очевидно в случае $\zeta = 0$. Если же $\zeta > 0$, то в силу (13) и (15)

$$\sum_{i=1}^{m} u_i = \zeta, \qquad \sum_{j=1}^{k} v_j = \zeta.$$

На основании (16) заключаем, что $u_i \equiv \frac{1}{m}\zeta$, $v_j \equiv \frac{1}{k}\zeta$. Согласно (6) формула (14) принимает вид

$$h = \zeta\Big(\tfrac{1}{m}\sum_{i=1}^{m} a_i - \tfrac{1}{k}\sum_{j=1}^{k} b_j\Big) = \mathbb{O}.$$

Таким образом, система (13)–(16) совместна только при $h = \mathbb{O}$, а должна быть совместной при фиксированном ненулевом векторе h. Полученное противоречие завершает доказательство предложения 2. □

Отметим, что в качестве h можно брать произвольный ненулевой вектор из $\mathbb{R}^n$.

5°. Допустим, что мы нашли решение $\{w, \gamma\}$ задачи (3) с $w \neq \mathbb{O}$. При фиксированном w целевая функция этой задачи представляет собой выпуклую ломаную $\varphi(\gamma)$. Минимальное значение ломаной может достигаться на целом отрезке $[\gamma_*, \gamma^*]$. Чтобы определить такой отрезок, достаточно вычислить значения $\varphi(\lambda)$ в узлах $\gamma = \langle w, a_i\rangle - 1$, $i \in 1:m$, и $\gamma = -\langle w, b_j\rangle - 1$, $j \in 1:k$. Некоторая свобода при выборе параметра γ позволяет получить приближённое разделение множеств $\mathcal{A}$ и $\mathcal{B}$ с дополнительными свойствами.

ЛИТЕРАТУРА

1. Bennett K. P., Mangasarian O. L. *Robust linear programming discrimination of two linearly inseparable sets* // Optimization Methods and Software. 1992. Vol. 1. P. 23–34.

2. Гаврин М. К., Малозёмов В. Н. *Экстремальные задачи с линейными ограничениями.* Л.: Изд-во ЛГУ, 1984. 176 с.