КАРТИНА МАСЛОМ*

В. Н. Малоземов Н. А. Соловьева

v.malozemov@spbu.ru 4vinyo@gmail.com

11 мая 2023 г.

В пространстве $\mathbb{R}^n$ заданы два конечных множества

$$P_1 = \{p_j\}_{j=1}^{s} \quad \text{и} \quad P_2 = \{p_j\}_{j=s+1}^{m},$$

где $s \in 1 : m-1$. Будем считать, что выпуклые оболочки C_1 и C_2 этих множеств не имеют общих точек. Введем вектор $\xi \in \mathbb{R}^m$ с компонентами $\xi_j = 1$ при $j \in 1 : s$ и $\xi_j = -1$ при $j \in (s+1) : m$. Обозначим через A матрицу, составленную из столбцов $\xi_1 p_1, \xi_2 p_2, \ldots, \xi_m p_m$.

Как известно [1], задача строгого линейного отделения множеств P_1 и P_2, имеющего наибольшую ширину разделяющей полосы, сводится к задаче квадратичного программирования следующего вида

$$\begin{gathered} Q(u) := \tfrac{1}{2}\langle A^T A u, u\rangle - \sum_{j=1}^{m} u_j \to \min, \\ \sum_{j=1}^{m} \xi_j u_j = 0, \\ u_j \geqslant 0 \ \text{ при всех } j \in 1 : m. \end{gathered} \tag{1}$$

Множество планов этой задачи обозначим через U. По теореме Куна-Таккера план $u \in U$ будет оптимальным тогда и только тогда, когда найдутся числа $\lambda, t_1, \ldots, t_m$, такие, что

$$\begin{gathered} A^T A u - e = \lambda \xi + \sum_{j=1}^{m} t_j e_j, \\ t_j u_j = 0, \quad t_j \geqslant 0 \ \text{ при всех } j \in 1 : m. \end{gathered} \tag{2}$$

Здесь e — вектор из $\mathbb{R}^m$, все компоненты которого равны единице, и e_j — j-й орт в $\mathbb{R}^m$.

---

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» `http://oml.cmlaboratory.com/`

В описании SMO-алгоритма решения задачи (1) используется другая форма критерия оптимальности [2]. Чтобы представить ее, введем индексные множества

$$
\begin{aligned}
&I_1'(u) = \bigl\{i \in 1:s \mid u_i = 0\bigr\}, && I_2'(u) = \bigl\{i \in 1:s \mid u_i > 0\bigr\},\\
&I_1''(u) = \bigl\{i \in (s+1):m \mid u_i = 0\bigr\}, && I_2''(u) = \bigl\{i \in (s+1):m \mid u_i > 0\bigr\},
\end{aligned}
$$

$$
\begin{aligned}
I'(u) &= I_1'(u) \cup I_2'(u) \cup I_2''(u),\\
I''(u) &= I_1''(u) \cup I_2''(u) \cup I_2'(u).
\end{aligned}
$$

Пусть $u \in U$ и $v = Au$. Положим

$$
\Delta(u) = \max_{i \in I''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} - \min_{i \in I'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\}.
$$

Нетрудно проверить, что $\Delta(u) \geqslant 0$ при всех $u \in U$. Действительно, вектор $u = \mathbf{0}$ является планом задачи (1). Ему соответствуют $v = \mathbf{0}$, $I'(\mathbf{0}) = 1 : s$, $I''(\mathbf{0}) = (s+1) : m$ и

$$
\Delta(\mathbf{0}) = \max_{i \in (s+1):m} \bigl\{-\xi_i\bigr\} - \min_{i \in 1:s} \bigl\{-\xi_i\bigr\} = 2 > 0. \tag{3}
$$

Возьмем ненулевой план u задачи (1). В силу условий $u_i \geqslant 0$ при всех $i \in 1:m$ и

$$
\sum_{i=1}^{s} u_i = \sum_{i=s+1}^{m} u_i
$$

имеем $I_2'(u) \neq \varnothing$ и $I_2''(u) \neq \varnothing$. Запишем

$$
\Delta(u) \geqslant \max_{i \in I_2'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} - \min_{i \in I_2'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} \geqslant 0. \tag{4}
$$

Получили, что $\Delta(u) \geqslant 0$ при всех $u \in U$.

ТЕОРЕМА. *План $u \in U$ будет оптимальным тогда и только тогда, когда $\Delta(u) = 0$.*

Доказательство основано на том, что условия (2) и $\Delta(u) = 0$ эквивалентны.

Предварительно перепишем условия (2) в развернутом виде:

$$
\begin{aligned}
\langle v, p_i\rangle - 1 &= \lambda + t_i && \text{при } i \in I_1'(u),\\
\langle v, p_i\rangle - 1 &= \lambda && \text{при } i \in I_2'(u),\\
-\langle v, p_i\rangle - 1 &= -\lambda + t_i && \text{при } i \in I_1''(u),\\
-\langle v, p_i\rangle - 1 &= -\lambda && \text{при } i \in I_2''(u),
\end{aligned}
$$

$$
t_i \geqslant 0 \quad \text{при } i \in 1:m.
$$

Это равносильно соотношениям

$$\begin{aligned}
&\langle v, p_i\rangle - \xi_i \geqslant \lambda \quad \text{при } i \in I_1'(u),\\
&\langle v, p_i\rangle - \xi_i = \lambda \quad \text{при } i \in I_2'(u),\\
&\langle v, p_i\rangle - \xi_i \leqslant \lambda \quad \text{при } i \in I_1''(u),\\
&\langle v, p_i\rangle - \xi_i = \lambda \quad \text{при } i \in I_2''(u).
\end{aligned} \tag{5}$$

Таким образом, план $u \in U$ задачи (1) будет оптимальным тогда и только тогда, когда найдется число λ, такое, что выполняются условия (5).

Теперь предположим, что $\Delta(u) = 0$ и покажем, что при некотором λ выполняются условия (5). В силу (3) вектор u отличен от нулевого. Из (4) следует, что

$$\max_{i\in I_2'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} = \min_{i\in I_2'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\}.$$

Значит,

$$\langle v, p_i\rangle - 1 = \nu \quad \text{при всех } i \in I_2'(u).$$

Аналогично получаем

$$0 = \Delta(u) \geqslant \max_{i\in I_2''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} - \min_{i\in I_2''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} \geqslant 0.$$

Значит,

$$\langle v, p_i\rangle + 1 = \eta \quad \text{при всех } i \in I_2''(u).$$

Покажем, что $\nu = \eta$. Допустив, что $\nu > \eta$, придем к противоречию с условием $\Delta(u) = 0$. Действительно,

$$\begin{aligned}
\max_{i\in I''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} &\geqslant \max_{i\in I_2'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} = \nu > \eta = \\
= \min_{i\in I_2''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} &\geqslant \min_{i\in I'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\}.
\end{aligned}$$

Аналогично придем к противоречию с условием $\Delta(u) = 0$, допустив, что $\eta > \nu$. Действительно,

$$\begin{aligned}
\max_{i\in I''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} &\geqslant \max_{i\in I_2''(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} = \eta > \nu = \\
= \min_{i\in I_2'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\} &\geqslant \min_{i\in I'(u)} \bigl\{\langle v, p_i\rangle - \xi_i\bigr\}.
\end{aligned}$$

Установлено, что

$$\nu = \eta =: \lambda.$$

При этом справедливы соотношения

$$\begin{aligned}
&\langle v, p_i\rangle - 1 = \lambda \quad \text{при } i \in I_2'(u),\\
&\langle v, p_i\rangle + 1 = \lambda \quad \text{при } i \in I_2''(u).
\end{aligned}$$

Неравенства

$$\begin{aligned} \langle v, p_i\rangle - 1 &\geqslant \lambda \quad \text{при } i \in I_1'(u), \\ \langle v, p_i\rangle + 1 &\leqslant \lambda \quad \text{при } i \in I_1''(u) \end{aligned}$$

доказываются от противного. Допустим, что

$$\min_{i \in I_1'(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} < \lambda.$$

Тогда получим

$$\begin{aligned} \max_{i \in I''(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} &\geqslant \max_{i \in I_2'(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} = \lambda > \\ > \min_{i \in I_1'(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} &\geqslant \min_{i \in I'(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\}. \end{aligned}$$

Это противоречит условию $\Delta(u) = 0$. Аналогично придем к противоречию с условием $\Delta(u) = 0$, допустив, что

$$\max_{i \in I_1''(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} > \lambda.$$

Действительно,

$$\begin{aligned} \max_{i \in I''(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} &\geqslant \max_{i \in I_1''(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} > \lambda = \\ = \min_{i \in I_2''(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} &\geqslant \min_{i \in I'(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\}. \end{aligned}$$

Таким образом, можно считать доказанным следующее утверждение: если $\Delta(u) = 0$ при некотором $u \in U$, то найдется число λ, такое, что выполняются соотношения (5). Это гарантирует оптимальность плана u.

Справедливо и обратное утверждение: если $u \in U$ — решение задачи (1), то $\Delta(u) = 0$. Проверим это.

Пусть $u \in U$ — оптимальный план задачи (1). Тогда найдется число λ, такое, что выполняются соотношения (5). Отметим, что $u \neq \mathbf{0}$. В противном случае соотношения (5) примут вид

$$\begin{aligned} -1 &\geqslant \lambda \quad \text{при } i \in 1 : s, \\ 1 &\leqslant \lambda \quad \text{при } i \in (s+1) : m. \end{aligned}$$

Но они несовместны. Из условия $u \neq \mathbf{0}$ следует, что $I_2'(u) \neq \varnothing$ и $I_2''(u) \neq \varnothing$. На основании соотношений (5) и определения индексных множеств $I'(u)$, $I''(u)$ получаем

$$\max_{i \in I''(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} = \lambda, \quad \min_{i \in I'(u)} \big\{ \langle v, p_i\rangle - \xi_i \big\} = \lambda.$$

Значит, $\Delta(u) = 0$.

Теорема доказана. $\square$

ЛИТЕРАТУРА

1. Малоземов В. Н., Плоткин А. В. *SVM-метод строгого линейного отделения двух конечных множеств* // Семинар «O&ML». Избранные доклады. 30 марта 2022 г. (`http://oml.cmlaboratory.com/reps22.shtml#0330`)

2. Малоземов В. Н., Соловьева Н. А. *SMO-алгоритм как обобщение МДМ-алгоритма* // Семинар «O&ML». Избранные доклады. 16 февраля 2023 г. (`http://oml.cmlaboratory.com/reps23.shtml#0216`)