СИЛЬНАЯ СХОДИМОСТЬ SMO–АЛГОРИТМА*

В. Н. Малоземов — v.malozemov@spbu.ru

Н. А. Соловьева — 4vinyo@gmail.com

30 ноября 2023 г.

SMO — это сокращение от Sequential Minimal Optimization.

В докладе рассматривается вариант SMO-алгоритма, ориентированный на решение задачи строгого линейного отделения двух конечных множеств в евклидовом пространстве при условии, что полоса, разделяющая эти множества, имеет наибольшую ширину [1,2,3]. Дается углубленный математический анализ алгоритма, включающий его нестандартное построение, обоснование критерия оптимальности и, главное, доказательство его сильной сходимости. Предварительные результаты представлены в докладах [4,5].

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

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

где $s \in 1 : m-1$. Будем считать, что выпуклые оболочки C_1 и C_2 этих множеств не имеют общих точек. Положим $\xi_i = 1$ при $i \in 1 : s$, $\xi_i = -1$ при $i \in (s+1) : m$.

Рассмотрим задачу строгого линейного отделения множеств P_1 и P_2 с максимальной шириной разделяющей полосы. Она формализуется следующим образом [6]:

$$\begin{gathered} \tfrac{1}{2}\,\|w\|^2 \to \min, \\ \xi_i\big(\langle w, p_i\rangle + \beta\big) \geqslant 1, \quad i \in 1 : m. \end{gathered} \tag{1}$$

Задача (1) является задачей квадратичного программирования с неотрицательной целевой функцией. В силу условия $C_1 \cap C_2 = \varnothing$ множество ее планов непусто. В этом случае задача (1) имеет решение (w^*, β^*). Вид ее целевой функции гарантирует единственность нормали w^*. Отметим, что $w^* \neq \mathbb{O}$. В противном случае выполнялось бы неравенство $\xi_i\,\beta^* \geqslant 1$ при $i \in 1 : m$, что невозможно, так как $\xi_1 = 1$, $\xi_m = -1$.

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

Запишем двойственную задачу:

$$
\begin{gathered}
-\tfrac{1}{2}\,\|v\|^2 + \sum_{i=1}^{m} u[i] \to \max, \\
v = \sum_{i=1}^{m} u[i]\,\xi_i p_i, \\
\sum_{i=1}^{m} u[i]\,\xi_i = 0, \quad u[i] \geqslant 0 \ \text{ при } i \in 1:m.
\end{gathered}
\tag{2}
$$

По первой теореме двойственности в квадратичном программировании задача (2) также имеет решение (v^*, u^*). По второй теореме двойственности справедливы соотношения

$$
\begin{gathered}
w^* = v^*, \\
u^*[i]\big(\xi_i\big(\langle w^*, p_i\rangle + \beta^*\big) - 1\big) = 0, \quad i \in 1:m.
\end{gathered}
\tag{3}
$$

Переменную v в задаче (2) можно исключить. Для этого введем матрицу A со столбцами $\xi_1 p_1, \xi_2 p_2, \ldots, \xi_m p_m$ и перепишем первое ограничение в задаче (2) в виде $v = Au$. Поменяв знак у целевой функции, придем к эквивалентной задаче:

$$
\begin{gathered}
D(u) := \tfrac{1}{2}\langle A^T A u, u\rangle - \sum_{i=1}^{m} u[i] \to \min, \\
\sum_{i=1}^{m} u[i]\,\xi_i = 0, \quad u[i] \geqslant 0 \ \text{ при всех } i \in 1:m.
\end{gathered}
\tag{4}
$$

В силу эквивалентности, задача (4) имеет решение. Единственность решения не гарантируется.

Если u^* — решение задачи (4), то (v^*, u^*), где $v^* = Au^*$, — решение задачи (2). Согласно (3), $w^* = v^* = Au^*$. Так как $w^* \neq \mathbb{O}$, то и $u^* \neq \mathbb{O}$.

Существует более глубокая связь между векторами w^* и u^*.

ЛЕММА 1. *Справедливо равенство*

$$
\|w^*\|^2 = \sum_{i=1}^{m} u^*[i]. \tag{5}
$$

Доказательство. Обозначим $I^* = \big\{i \in 1:m \mid u^*[i] > 0\big\}$. Согласно второму ограничению задачи (2) имеем

$$
\sum_{i \in I^*} u^*[i]\,\xi_i = \sum_{i=1}^{m} u^*[i]\,\xi_i = 0. \tag{6}
$$

Кроме того, в силу (3) выполняются равенства

$$\xi_i\big(\langle w^*, p_i\rangle + \beta^*\big) = 1 \quad \text{при } i \in I^*. \tag{7}$$

С учетом соотношений (6) и (7) получаем

$$\|w^*\|^2 = \langle w^*, w^*\rangle = \Big\langle w^*, \sum_{i\in I^*} u^*[i]\,\xi_i p_i\Big\rangle =$$

$$= \sum_{i\in I^*} u^*[i]\,\xi_i\big(\langle w^*, p_i\rangle + \beta^*\big) = \sum_{i\in I^*} u^*[i] = \sum_{i=1}^{m} u^*[i].$$

Лемма доказана. □

Естественным дополнением к лемме 1 является следующее утверждение: *для вектора $v = Au$, где u — план задачи* (4), *справедливо неравенство*

$$\langle v, w^*\rangle \geqslant \sum_{i=1}^{m} u[i]. \tag{8}$$

Действительно, так как (w^*, β^*) — план задачи (1), то

$$\langle v, w^*\rangle = \Big\langle w^*, \sum_{i=1}^{m} u[i]\,\xi_i p_i\Big\rangle =$$

$$= \sum_{i=1}^{m} u[i]\,\xi_i\big(\langle w^*, p_i\rangle + \beta^*\big) \geqslant \sum_{i=1}^{m} u[i].$$

ЛЕММА 2. *Пусть $v = Au$, где u — план задачи* (4). *Тогда*

$$\|v - w^*\|^2 \leqslant 2\big(D(u) - D(u^*)\big).$$

Доказательство. Согласно (5) и равенствам $w^* = v^* = Au^*$ имеем

$$D(u^*) = \tfrac{1}{2}\,\|Au^*\|^2 - \sum_{i=1}^{m} u^*[i] =$$

$$= \tfrac{1}{2}\,\|v^*\|^2 - \|w^*\|^2 = -\tfrac{1}{2}\,\|w^*\|^2.$$

С учетом (8) получаем

$$\|v - w^*\|^2 = \|v\|^2 - 2\langle v, w^*\rangle + \|w^*\|^2 \leqslant$$

$$\leqslant \|v\|^2 - 2\sum_{i=1}^{m} u[i] + \|w^*\|^2 =$$

$$= 2\Big(\tfrac{1}{2}\,\|v\|^2 - \sum_{i=1}^{m} u[i]\Big) - 2\big(-\tfrac{1}{2}\,\|w^*\|^2\big) = 2\big(D(u) - D(u^*)\big).$$

Лемма доказана. □

2°. Оценка плана задачи (4). Обозначим через U множество планов задачи (4). Отметим, что U — выпуклый конус с вершиной в начале координат. Зафиксируем план u и рассмотрим его возмущение вида

$$\widetilde{u} = u + \delta' e_{i'} + \delta'' e_{i''},$$

где e_i — i-й орт в пространстве $\mathbb{R}^m$. Параметры δ', δ'' и индексы i', i'' из $1:m$ выберем так, чтобы $\widetilde{u} \in U$ и $D(\widetilde{u})$ было по возможности меньше $D(u)$.

Для справедливости включения $\widetilde{u} \in U$ необходимо, в частности, чтобы сумма $\sum_{i=1}^m \widetilde{u}[i]\,\xi_i$ равнялась нулю. Это приводит к условию

$$\delta'\xi_{i'} + \delta''\xi_{i''} = 0,$$

согласно которому

$$\delta'' = -\delta'\xi_{i'}\xi_{i''}.$$

Обозначим $\delta = \delta'$. Тогда

$$\widetilde{u} = u + \delta\big(e_{i'} - \xi_{i'}\xi_{i''}e_{i''}\big) = u + \delta\,\xi_{i'}\big(\xi_{i'}e_{i'} - \xi_{i''}e_{i''}\big). \tag{9}$$

Обратимся к целевой функции задачи (4). Имеем

$$D(u) = \tfrac{1}{2}\,\|Au\|^2 - \sum_{i=1}^m u[i].$$

Вычислим $D(\widetilde{u})$, где $\widetilde{u}$ имеет вид (9). Положим $v = Au$. Отметим, что $Ae_i = \xi_i p_i$. Так как

$$\tfrac{1}{2}\,\|A\widetilde{u}\|^2 = \tfrac{1}{2}\,\big\|v + \delta\,\xi_{i'}\big(p_{i'} - p_{i''}\big)\big\|^2 = \tfrac{1}{2}\,\|v\|^2 + \delta\,\xi_{i'}\langle v, p_{i'} - p_{i''}\rangle + \tfrac{1}{2}\delta^2\,\|p_{i'} - p_{i''}\|^2,$$

$$\sum_{i=1}^m \widetilde{u}[i] = \sum_{i=1}^m u[i] + \delta\,\xi_{i'}\big(\xi_{i'} - \xi_{i''}\big),$$

то

$$D(\widetilde{u}) = D(u) + \delta\,\xi_{i'}\Big(\langle v, p_{i'} - p_{i''}\rangle - \big(\xi_{i'} - \xi_{i''}\big)\Big) + \tfrac{1}{2}\delta^2\,\|p_{i'} - p_{i''}\|^2. \tag{10}$$

Введем обозначение (индексы i' и i'' меняются местами)

$$\Delta(i', i'') = \langle v, p_{i''} - p_{i'}\rangle - \big(\xi_{i''} - \xi_{i'}\big) = \big(\langle v, p_{i''}\rangle - \xi_{i''}\big) - \big(\langle v, p_{i'}\rangle - \xi_{i'}\big).$$

Оно позволяет упростить запись формулы (10):

$$\Phi(\delta) := D(\widetilde{u}) = D(u) - \delta\,\xi_{i'}\Delta(i', i'') + \tfrac{1}{2}\delta^2\,\|p_{i'} - p_{i''}\|^2. \tag{11}$$

Функция $\Phi(\delta)$ является квадратным трехчленом относительно δ. Желая уменьшить значение $D(\widetilde{u})$, найдем точку минимума функции $\Phi(\delta)$. Из уравнения $\Phi'(\delta)=0$ получаем

$$\delta = \xi_{i'} \frac{\Delta(i',i'')}{\|p_{i'}-p_{i''}\|^2}\,.$$

Обозначим

$$\widetilde{\lambda} = \delta\,\xi_{i'} = \frac{\Delta(i',i'')}{\|p_{i'}-p_{i''}\|^2}\,. \tag{12}$$

При выбранном δ согласно (9) имеем

$$\widetilde{u} = u + \widetilde{\lambda}\big(\xi_{i'}e_{i'} - \xi_{i''}e_{i''}\big). \tag{13}$$

При этом в силу (11) справедливо разложение

$$D(\widetilde{u}) = D(u) - \tfrac{1}{2}\,\widetilde{\lambda}\,\Delta(i',i''). \tag{14}$$

Теперь займемся выбором индексов i' и i'' из $1:m$, обеспечивающих неотрицательность величины $\Delta(i',i'')$ и возможно большее ее значение. Введем индексные множества

$$I_0'(u) = \big\{i\in 1:s \mid u[i]=0\big\}, \qquad I_1'(u) = \big\{i\in 1:s \mid u[i]>0\big\},$$
$$I_0''(u) = \big\{i\in (s+1):m \mid u[i]=0\big\}, \quad I_1''(u) = \big\{i\in (s+1):m \mid u[i]>0\big\},$$

$$I'(u) = I_0'(u)\cup I_1'(u)\cup I_1''(u) = \big(1:s\big)\cup I_1''(u),$$
$$I''(u) = I_0''(u)\cup I_1''(u)\cup I_1'(u) = \big((s+1):m\big)\cup I_1'(u).$$

Отметим, что как $I'(u)$, так и $I''(u)$ содержат множество $I_1(u)=I_1'(u)\cup I_1''(u)$. Индексы i' и i'' выберем следующим образом: $i'\in I'(u)$, $i''\in I''(u)$,

$$\begin{aligned}\langle v,p_{i'}\rangle - \xi_{i'} &= \min_{i\in I'(u)}\big\{\langle v,p_i\rangle - \xi_i\big\},\\ \langle v,p_{i''}\rangle - \xi_{i''} &= \max_{i\in I''(u)}\big\{\langle v,p_i\rangle - \xi_i\big\}.\end{aligned} \tag{15}$$

В этом случае

$$\Delta(i',i'') = \max_{i\in I''(u)}\big\{\langle v,p_i\rangle - \xi_i\big\} - \min_{i\in I'(u)}\big\{\langle v,p_i\rangle - \xi_i\big\}. \tag{16}$$

Индексы i' и i'' определяются планом u. В дальнейшем вместо $\Delta(i',i'')$ будем писать $\Delta(u)$. Величину $\Delta(u)$ назовем *оценкой плана u*.

Покажем, что $\Delta(u)\geqslant 0$ при всех $u\in U$. Действительно, вектор $u=\mathbb{O}$ принадлежит U. Ему соответствуют $v=\mathbb{O}$, $I'(\mathbb{O})=1:s$, $I''(\mathbb{O})=(s+1):m$ и

$$\Delta(\mathbb{O}) = \max_{i\in(s+1):m}\big\{-\xi_i\big\} - \min_{i\in 1:s}\big\{-\xi_i\big\} = 2 > 0. \tag{17}$$

Возьмем ненулевой план u задачи (4). 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 $I_1'(u) \neq \varnothing$ и $I_1''(u) \neq \varnothing$. Сославшись на формулу (16), получим

$$\Delta(u) \geqslant \max_{i \in I_1(u)} \big\{ \langle v, p_i \rangle - \xi_i \big\} - \min_{i \in I_1(u)} \big\{ \langle v, p_i \rangle - \xi_i \big\} \geqslant 0. \tag{18}$$

Неотрицательность оценки $\Delta(u)$ при всех $u \in U$ установлена. Согласно (12) неотрицательным будет и шаг $\widetilde{\lambda}$.

Осталось проверить условие $\widetilde{u}[i] \geqslant 0$ при всех $i \in 1 : m$, чтобы гарантировать, что $\widetilde{u} \in U$. Согласно (13) дело сводится к анализу неравенств

$$\begin{aligned} u[i'] \,+\, \xi_{i'}\widetilde{\lambda} &\geqslant 0, \\ u[i''] - \xi_{i''}\widetilde{\lambda} &\geqslant 0. \end{aligned} \tag{19}$$

Придется рассмотреть четыре случая.

1) $i' \in 1 : s$, $i'' \in (s+1) : m$. Согласно определению ξ_i неравенства (19) примут вид

$$\begin{aligned} u[i'] \,+\, \widetilde{\lambda} &\geqslant 0, \\ u[i''] + \widetilde{\lambda} &\geqslant 0. \end{aligned}$$

В силу неотрицательности $\widetilde{\lambda}$ эти неравенства справедливы.

2) $i' \in 1 : s$, $i'' \in I_1'(u)$. Неравенства (19) принимают вид

$$\begin{aligned} u[i'] \,+\, \widetilde{\lambda} &\geqslant 0, \\ u[i''] - \widetilde{\lambda} &\geqslant 0. \end{aligned}$$

Второе неравенство может нарушаться. Оно будет выполняться, если $\widetilde{\lambda}$ заменить на $\lambda = \min\{\widetilde{\lambda}, u[i'']\}$.

3) $i' \in I_1''(u)$, $i'' \in (s+1) : m$. Неравенства (19) принимают вид

$$\begin{aligned} u[i'] \,-\, \widetilde{\lambda} &\geqslant 0, \\ u[i''] + \widetilde{\lambda} &\geqslant 0. \end{aligned}$$

Следует заменить $\widetilde{\lambda}$ на $\lambda = \min\{\widetilde{\lambda}, u[i']\}$.

4) $i' \in I_1''(u)$, $i'' \in I_1'(u)$. Неравенства (19) принимают вид

$$u[i'] - \widetilde{\lambda} \geqslant 0,$$
$$u[i''] - \widetilde{\lambda} \geqslant 0.$$

Это равносильно одному неравенству $\widetilde{\lambda} \leqslant \min\{u[i'], u[i'']\}$. Оно будет выполняться, если $\widetilde{\lambda}$ заменить на следующую величину

$$\lambda = \min\bigl\{\widetilde{\lambda}, \min\{u[i'], u[i'']\}\bigr\} = \min\bigl\{\widetilde{\lambda}, u[i'], u[i'']\bigr\}.$$

Во всех случаях $\lambda \leqslant \widetilde{\lambda}$.

Таким образом, для справедливости включения $\widetilde{u} \in U$ нужно в формуле (13) заменить $\widetilde{\lambda}$ на λ, где

$$\lambda = \begin{cases} \widetilde{\lambda}, \quad \text{если } i' \in 1:s,\ i'' \in (s+1):m, \\ \min\{\widetilde{\lambda}, u[i'']\}, \quad \text{если } i' \in 1:s,\ i'' \in I_1'(u), \\ \min\{\widetilde{\lambda}, u[i']\}, \quad \text{если } i' \in I_1''(u),\ i'' \in (s+1):m, \\ \min\{\widetilde{\lambda}, u[i'], u[i'']\}, \quad \text{если } i' \in I_1''(u),\ i'' \in I_1'(u). \end{cases} \tag{20}$$

В результате получим

$$\widetilde{u} = u + \lambda\bigl(\xi_{i'} e_{i'} - \xi_{i''} e_{i''}\bigr). \tag{21}$$

Это базовая формула. На ее основе строится SMO-алгоритм решения задачи (1). Индексы i' и i'' определяются из условий (15), в которых $v = Au$. В (21) неявно присутствует оценка $\Delta(u)$ плана u,

$$\Delta(u) = \Delta(i', i'') = \langle v, p_{i''} - p_{i'} \rangle - \bigl(\xi_{i''} - \xi_{i'}\bigr). \tag{22}$$

Действительно, нужно иметь в виду, что $\widetilde{\lambda} = \Delta(u) / \|p_{i'} - p_{i''}\|^2$ (см. (12)) и что шаг λ вычисляется по формуле (20).

3°. Дальнейшие свойства оценки плана задачи (4). В предыдущем разделе было доказано, что оценка $\Delta(u)$ неотрицательна при всех $u \in U$. Продолжим изучение оценки $\Delta(u)$ вида (22).

ТЕОРЕМА 1. *План $u \in U$ задачи* (4) *будет оптимальным тогда и только тогда, когда $\Delta(u) = 0$.*

Доказательство. По теореме Куна–Таккера для оптимальности плана u задачи (4) необходимо и достаточно, чтобы нашлись числа $\mu, t_1, \ldots, t_m$, такие, что

$$\begin{gathered} A^T A u - e = \mu \xi + \sum_{i=1}^{m} t_i e_i, \\ t_i u[i] = 0, \quad t_i \geqslant 0 \ \text{ при всех } i \in 1:m. \end{gathered} \tag{23}$$

Здесь $\xi = (\xi_1, \xi_2, \ldots, \xi_m)^T$, $e \in \mathbb{R}^m$ — вектор, все компоненты которого равны единице, и e_i — i-й орт в $\mathbb{R}^m$. Доказательство теоремы 1 основано на том, что условия (23) и $\Delta(u) = 0$ эквивалентны.

Предварительно перепишем условия (23) в развернутом виде:

$$\begin{aligned} \langle v, p_i\rangle - 1 &= \mu + t_i &&\text{при } i \in I_0'(u), \\ \langle v, p_i\rangle - 1 &= \mu &&\text{при } i \in I_1'(u), \\ -\langle v, p_i\rangle - 1 &= -\mu + t_i &&\text{при } i \in I_0''(u), \\ -\langle v, p_i\rangle - 1 &= -\mu &&\text{при } i \in I_1''(u), \end{aligned}$$

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

Это равносильно соотношениям

$$\begin{aligned} \langle v, p_i\rangle - \xi_i &\geqslant \mu &&\text{при } i \in I_0'(u), \\ \langle v, p_i\rangle - \xi_i &= \mu &&\text{при } i \in I_1'(u), \\ \langle v, p_i\rangle - \xi_i &\leqslant \mu &&\text{при } i \in I_0''(u), \\ \langle v, p_i\rangle - \xi_i &= \mu &&\text{при } i \in I_1''(u). \end{aligned} \tag{24}$$

Таким образом, план $u \in U$ задачи (4) будет оптимальным тогда и только тогда, когда найдется число μ, такое, что выполняются условия (24).

Теперь предположим, что $\Delta(u) = 0$ и покажем, что при некотором μ выполняются условия (24). В силу (17) вектор u отличен от нулевого. Из (18) следует, что

$$\max_{i \in I_1(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} = \min_{i \in I_1(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} =: \mu.$$

Значит,

$$\langle v, p_i\rangle - \xi_i = \mu \quad \text{при всех } i \in I_1(u).$$

Напомним, что $I_1(u) = I_1' \cup I_1''$.

Осталось проверить первое и третье условия из (24). Имеем

$$\begin{aligned} \mu = \max_{i \in I_1'(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} \leqslant \max_{i \in I''(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} = \\ = \min_{i \in I'(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} \leqslant \min_{i \in I_0'(u)} \big\{\langle v, p_i\rangle - \xi_i\big\}. \end{aligned}$$

Отсюда следует первое условие из (24). Аналогично

$$\begin{aligned} \mu = \min_{i \in I_1''(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} \geqslant \min_{i \in I'(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} = \\ = \max_{i \in I''(u)} \big\{\langle v, p_i\rangle - \xi_i\big\} \geqslant \max_{i \in I_0''(u)} \big\{\langle v, p_i\rangle - \xi_i\big\}. \end{aligned}$$

Отсюда следует третье условие из (24).

Таким образом, можно считать доказанным следующее утверждение: если $\Delta(u) = 0$ при некотором $u \in U$, то найдется число μ, такое, что выполняются соотношения (24). Это гарантирует оптимальность плана u.

Справедливо и обратное утверждение: $\Delta(u) = 0$, если u — решение задачи (4). Проверим его.

По критерию оптимальности при некотором μ выполняются соотношения (24). В разд. 1 отмечалось, что решение задачи (4) является ненулевым вектором. Поэтому, как показано в разд. 2, индексные множества $I_1'(u)$ и $I_1''(u)$ будут непустыми. Опираясь на эти факты и определение индексных множеств $I'(u)$ и $I''(u)$, приходим к заключению, что

$$\max_{i \in I''(u)} \big\{ \langle v, p_i \rangle - \xi_i \big\} = \mu \qquad \text{и} \qquad \min_{i \in I'(u)} \big\{ \langle v, p_i \rangle - \xi_i \big\} = \mu.$$

Отсюда очевидным образом следует равенство $\Delta(u) = 0$.

Теорема доказана. $\square$

З а м е ч а н и е. Появление нестандартных индексных множеств $I'(u)$ и $I''(u)$ в определении оценки $\Delta(u)$ диктуется теоремой Куна–Таккера, точнее, условиями Куна–Таккера в форме (24). Скалярное равенство $\Delta(u) = 0$ представляет собой свернутый вариант условий Куна–Таккера (23).

Теперь покажем, как с помощью $\Delta(u)$ оценить близость вектора $v = Au$ при $u \in U$ к нормали w^* (см. разд. 1).

ТЕОРЕМА 2. *Справедливо неравенство*

$$\|v - w^*\|^2 \leqslant \tfrac{1}{2}\, \Delta(u) \Big(\sum_{i=1}^{m} u[i] + \sum_{i=1}^{m} u^*[i] \Big). \tag{25}$$

Д о к а з а т е л ь с т в о. С помощью (5) и (8) получаем

$$\begin{aligned} \|v - w^*\|^2 &\leqslant \|v\|^2 - \langle v, w^* \rangle - \sum_{i=1}^{m} u[i] + \sum_{i=1}^{m} u^*[i] = \\ &= \Big(\langle v, v \rangle - \sum_{i=1}^{m} u[i] \Big) - \Big(\langle v, w^* \rangle - \sum_{i=1}^{m} u^*[i] \Big) =: S_1 - S_2. \end{aligned} \tag{26}$$

Оценим S_1. Запишем

$$S_1 = \sum_{i=1}^{m} u[i]\,\xi_i\langle v, p_i\rangle - \sum_{i=1}^{m} u[i]\,\xi_i^2 = \sum_{i=1}^{m} u[i]\,\xi_i\big(\langle v, p_i\rangle - \xi_i\big) =$$
$$= \sum_{i\in I_1'(u)} u[i]\big(\langle v, p_i\rangle - \xi_i\big) - \sum_{i\in I_1''(u)} u[i]\big(\langle v, p_i\rangle - \xi_i\big) \leqslant$$
$$\leqslant \max_{i\in I''(u)}\big\{\langle v, p_i\rangle - \xi_i\big\} \sum_{i=1}^{s} u[i] - \min_{i\in I'(u)}\big\{\langle v, p_i\rangle - \xi_i\big\} \sum_{i=s+1}^{m} u[i].$$

Так как

$$\sum_{i=1}^{s} u[i] = \sum_{i=s+1}^{m} u[i] = \tfrac{1}{2}\sum_{i=1}^{m} u[i],$$

то

$$S_1 \leqslant \tfrac{1}{2}\,\Delta(u)\sum_{i=1}^{m} u[i]. \tag{27}$$

Аналогично оценивается $(-S_2)$. Запишем

$$-S_2 = -\sum_{i=1}^{m} u^*[i]\,\xi_i\langle v, p_i\rangle + \sum_{i=1}^{m} u^*[i]\,\xi_i^2 = -\sum_{i=1}^{m} u^*[i]\,\xi_i\big(\langle v, p_i\rangle - \xi_i\big) =$$
$$= \sum_{i=s+1}^{m} u^*[i]\big(\langle v, p_i\rangle - \xi_i\big) - \sum_{i=1}^{s} u^*[i]\big(\langle v, p_i\rangle - \xi_i\big) \leqslant$$
$$\leqslant \max_{i\in I''(u)}\big\{\langle v, p_i\rangle - \xi_i\big\} \sum_{i=s+1}^{m} u^*[i] - \min_{i\in I'(u)}\big\{\langle v, p_i\rangle - \xi_i\big\} \sum_{i=1}^{s} u^*[i].$$

Так как

$$\sum_{i=s+1}^{m} u^*[i] = \sum_{i=1}^{s} u^*[i] = \tfrac{1}{2}\sum_{i=1}^{m} u^*[i],$$

то

$$-S_2 \leqslant \tfrac{1}{2}\,\Delta(u)\sum_{i=1}^{m} u^*[i]. \tag{28}$$

Объединив (26), (27) и (28), придем к требуемому неравенству (25). Теорема доказана. □

4°. Описание SMO-алгоритма. SMO-алгоритм ориентирован на одновременное решение задач (4) и (1). Опишем шаг алгоритма.

Пусть имеется k-е приближение $u_k \in U$. Обозначим

$$I'(u_k) = \big(1:s\big) \cup \big\{i \in (s+1):m \mid u_k[i] > 0\big\},$$
$$I''(u_k) = \big\{i \in 1:s \mid u_k[i] > 0\big\} \cup \big((s+1):m\big).$$

Положим $v_k = Au_k$ и выберем индексы $i'_k \in I'(u_k)$, $i''_k \in I''(u_k)$ из условий

$$\langle v_k, p_{i'_k}\rangle - \xi_{i'_k} = \min_{i \in I'(u_k)} \big\{\langle v_k, p_i\rangle - \xi_i\big\},$$
$$\langle v_k, p_{i''_k}\rangle - \xi_{i''_k} = \max_{i \in I''(u_k)} \big\{\langle v_k, p_i\rangle - \xi_i\big\}.$$

Вычислим оценку плана u_k:

$$\Delta(u_k) = \langle v_k, p_{i''_k} - p_{i'_k}\rangle - \big(\xi_{i''_k} - \xi_{i'_k}\big).$$

Если $\Delta(u_k) = 0$, то по теореме 1 план u_k — решение задачи (4).

Пусть $\Delta(u_k) > 0$ (в частности, $i'_k \neq i''_k$). Введем предварительный шаг

$$\widetilde{\lambda}_k = \frac{\Delta(u_k)}{\big\|p_{i'_k} - p_{i''_k}\big\|^2}\,.$$

Уточним его:

$$\lambda_k = \begin{cases} \widetilde{\lambda}_k, \quad \text{если } i'_k \in 1:s,\ i''_k \in (s+1):m; \\ \min\{\widetilde{\lambda}_k, u_k[i''_k]\}, \quad \text{если } i'_k \in 1:s,\ i''_k \in 1:s; \\ \min\{\widetilde{\lambda}_k, u_k[i'_k]\}, \quad \text{если } i'_k \in (s+1):m,\ i''_k \in (s+1):m; \\ \min\{\widetilde{\lambda}_k, u_k[i'_k], u[i''_k]\}, \quad \text{если } i'_k \in (s+1):m,\ i''_k \in 1:s. \end{cases}$$

Положим

$$u_{k+1} = u_k + \lambda_k\big(\xi_{i'_k} e_{i'_k} - \xi_{i''_k} e_{i''_k}\big). \tag{29}$$

Очевидно, что $\widetilde{\lambda}_k$ и λ_k положительны, причем $\lambda_k \leqslant \widetilde{\lambda}_k$. При $\lambda_k = \widetilde{\lambda}_k$ итерацию с номером k будем называть *неусеченной*, при $\lambda_k < \widetilde{\lambda}_k$ — *усеченной*. Подчеркнем, что на k-й итерации SMO-алгоритма вычисляется u_{k+1}.

Найдем значение целевой функции задачи (4) на плане u_{k+1}. Так же, как при выводе формулы (11), получим

$$D(u_{k+1}) = D(u_k) - \lambda_k \Delta(u_k) + \tfrac{1}{2}\,\lambda_k^2\, \big\|p_{i'_k} - p_{i''_k}\big\|^2\,.$$

Перепишем эту формулу в виде

$$D(u_k) - D(u_{k+1}) = \lambda_k \Delta(u_k) - \tfrac{1}{2}\,\lambda_k^2\, \big\|p_{i'_k} - p_{i''_k}\big\|^2\,. \tag{30}$$

ЛЕММА 3. *Предположим, что* $\Delta(u_k) > 0$. *Тогда* $D(u_{k+1}) < D(u_k)$.

Доказательство. Напомним, что

$$\lambda_k \leqslant \widetilde{\lambda}_k = \frac{\Delta(u_k)}{\left\|p_{i'_k} - p_{i''_k}\right\|^2}.$$

Отсюда следует, что

$$\lambda_k \left\|p_{i'_k} - p_{i''_k}\right\|^2 \leqslant \Delta(u_k). \tag{31}$$

На основании (30) и (31) получаем

$$D(u_k) - D(u_{k+1}) \geqslant \tfrac{1}{2}\,\lambda_k \Delta(u_k).$$

Учитывая положительность $\Delta(u_k)$ и λ_k, приходим к заключению леммы. $\square$

Обозначим через d диаметр объединения $P_1 \cup P_2$ исходных конечных множеств P_1 и P_2.

ЛЕММА 4. *Предположим, что* $\Delta(u_k) > 0$. *Если итерация с номером* k *неусеченная, то верно неравенство*

$$D(u_k) - D(u_{k+1}) \geqslant \frac{\Delta^2(u_k)}{2d^2}.$$

Доказательство. В данном случае $\lambda_k = \widetilde{\lambda}_k$. Согласно (30) и определению $\widetilde{\lambda}_k$ имеем

$$\begin{aligned} D(u_k) - D(u_{k+1}) &= \widetilde{\lambda}_k \Delta(u_k) - \tfrac{1}{2}\,\widetilde{\lambda}_k^2 \left\|p_{i'_k} - p_{i''_k}\right\|^2 = \\ &= \tfrac{1}{2}\,\widetilde{\lambda}_k \Delta(u_k) = \frac{\Delta^2(u_k)}{2\left\|p_{i'_k} - p_{i''_k}\right\|^2} \geqslant \frac{\Delta^2(u_k)}{2d^2}. \end{aligned}$$

Лемма доказана. $\square$

Рекуррентная формула (29) позволяет построить последовательность планов задачи (4)

$$u_0, u_1, \ldots, u_k, \ldots, \tag{32}$$

в которой $u_0 \in U$ — произвольное начальное приближение. Эта последовательность будет конечной, если при некотором k_0 выполнится условие $\Delta(u_{k_0}) = 0$. Получим, что u_{k_0} — решение задачи (4) и $v_{k_0} = Au_{k_0}$ — решение задачи (1).

Бесконечная последовательность (32) характеризуется тем, что

$$\Delta(u_k) > 0 \quad \text{при всех } k = 0, 1, \ldots \tag{33}$$

В дальнейшем рассматриваются только бесконечные последовательности (32).

5°. Ограниченность последовательности $\{u_k\}$. Имеется в виду следующее утверждение.

ТЕОРЕМА 3. *Пусть выполняется условие* (33) *и* $D(u_0) \leqslant 0$. *Тогда найдется константа* B, *такая, что*

$$\|u_k\| \leqslant B \quad \text{при всех } k = 0, 1, \ldots$$

Нам понадобятся две алгебраические леммы.

ЛЕММА 5. *Если* $Au = \mathbb{O}$ *при некотором* $u \in U$, *то необходимо* $u = \mathbb{O}$.

Доказательство. Допустим противное, что $Au = \mathbb{O}$ на ненулевом плане $u \in U$. Запишем равенство $Au = \mathbb{O}$ в виде

$$\sum_{i=1}^{m} u[i]\,\xi_i p_i = \mathbb{O}.$$

По определению ξ_i имеем

$$\sum_{i=1}^{s} u[i]\,p_i = \sum_{i=s+1}^{m} u[i]\,p_i. \tag{34}$$

Из определения множества U следует, что вектор $u \in U$ имеет неотрицательные компоненты и

$$\sum_{i=1}^{s} u[i] = \sum_{i=s+1}^{m} u[i] =: q.$$

Так как $u \neq \mathbb{O}$, то $q > 0$.

Введем коэффициенты $\widehat{u}[i] = u[i]/q$. Они неотрицательны и

$$\sum_{i=1}^{s} \widehat{u}[i] = 1, \qquad \sum_{i=s+1}^{m} \widehat{u}[i] = 1.$$

Поделив равенство (34) на q, получим

$$\sum_{i=1}^{s} \widehat{u}[i]\,p_i = \sum_{i=s+1}^{m} \widehat{u}[i]\,p_i. \tag{35}$$

Вектор, стоящий в левой части равенства (35), принадлежит выпуклой оболочке C_1 множества P_1, а вектор, стоящий в правой части равенства (35), принадлежит выпуклой оболочке C_2 множества P_2 (см. разд. 1). Само равенство (35) означает, что $C_1 \cap C_2 \neq \varnothing$. Но это противоречит основному условию $C_1 \cap C_2 = \varnothing$.

Лемма доказана. □

ЛЕММА 6. *Существует константа $\gamma > 0$ такая, что при всех $u \in U$ выполняется неравенство*
$$\|Au\| \geqslant \gamma \|u\| . \tag{36}$$
Доказательство. Пусть $S = \{u \in \mathbb{R}^m \mid \|u\| = 1\}$ — единичная сфера и
$$\gamma := \min_{u \in U \cap S} \|Au\| . \tag{37}$$
По определению, множество U является выпуклым конусом с вершиной в начале координат, поэтому пересечение $U \cap S$ непусто. Минимум γ достигается и неотрицателен. На самом деле, $\gamma > 0$. В противном случае найдется вектор $u_0 \in U \cap S$, на котором $\|Au_0\| = 0$. По лемме 5 имеем $u_0 = 0$, что противоречит условию $u_0 \in S$.

При $u = \mathbb{O}$ неравенство (36) тривиально. Пусть $u \in U$ и $u \neq \mathbb{O}$. Введем вектор $\widehat{u} = u / \|u\|$. Он принадлежит пересечению $U \cap S$. Согласно (37) получаем $\|A\widehat{u}\| \geqslant \gamma$. Отсюда очевидным образом следует требуемое неравенство (36).

Лемма доказана. □

Доказательство теоремы 3. По лемме 3 последовательность $\{D(u_k)\}$ строго убывает. Предполагается также, что $D(u_0) \leqslant 0$. Это гарантирует, что при $k = 1, 2, \ldots$ выполняется неравенство
$$D(u_k) := \tfrac{1}{2} \|v_k\|^2 - \sum_{i=1}^{m} u_k[i] < 0. \tag{38}$$
В частности, $u_k \neq \mathbb{O}$ при $k = 1, 2, \ldots$ Перепишем неравенство (38) в виде
$$\sum_{i=1}^{m} u_k[i] > \tfrac{1}{2} \|v_k\|^2 .$$
Из (36) следует, что $\|v_k\|^2 \geqslant \gamma^2 \|u_k\|^2$, поэтому
$$\sum_{i=1}^{m} u_k[i] > \tfrac{1}{2} \gamma^2 \|u_k\|^2 , \qquad k = 1, 2, \ldots$$
Вместе с тем, по неравенству Коши–Буняковского имеем
$$\sum_{i=1}^{m} u_k[i] \leqslant \sqrt{m} \|u_k\| , \qquad k = 1, 2, \ldots$$
Приходим к двойному неравенству
$$\tfrac{1}{2} \gamma^2 \|u_k\|^2 < \sum_{i=1}^{m} u_k[i] \leqslant \sqrt{m} \|u_k\| .$$
Поделив на $\|u_k\|$, получим требуемое неравенство с $B = \frac{2\sqrt{m}}{\gamma^2}$.

Теорема доказана. □

6°. Об усеченных итерациях. Если k-я итерация является усеченной ($\lambda_k < \widetilde{\lambda}_k$), то формула (29) для u_{k+1} упрощается.

Рассмотрим четыре случая.

1) $i'_k \in 1:s$, $i''_k \in 1:s$.
Имеем $\xi_{i'_k} = \xi_{i''_k} = 1$, $\lambda_k = u_k[i''_k]$ и

$$u_{k+1} = u_k + u_k[i''_k]\big(e_{i'_k} - e_{i''_k}\big).$$

Покомпонентно:

$$u_{k+1}[i] = \begin{cases} u_k[i'_k] + u_k[i''_k] \quad \text{при } i = i'_k, \\ 0 \quad \text{при } i = i''_k, \\ u_k[i] \quad \text{при остальных } i. \end{cases} \tag{39}$$

2) $i'_k \in (s+1):m$, $i''_k \in (s+1):m$.
Имеем $\xi_{i'_k} = \xi_{i''_k} = -1$, $\lambda_k = u_k[i'_k]$ и

$$u_{k+1} = u_k - u_k[i'_k]\big(e_{i'_k} - e_{i''_k}\big).$$

Покомпонентно:

$$u_{k+1}[i] = \begin{cases} 0 \quad \text{при } i = i'_k, \\ u_k[i''_k] + u_k[i'_k] \quad \text{при } i = i''_k, \\ u_k[i] \quad \text{при остальных } i. \end{cases} \tag{40}$$

3) $i'_k \in (s+1):m$, $i''_k \in 1:s$.
Имеем $\xi_{i'_k} = -1$, $\xi_{i'_k} = 1$, $\lambda_k = u_k[i'_k]$ (если $u_k[i'_k] \leqslant u_k[i''_k]$) и

$$u_{k+1} = u_k - u_k[i'_k]\big(e_{i'_k} + e_{i''_k}\big).$$

Покомпонентно:

$$u_{k+1}[i] = \begin{cases} 0 \quad \text{при } i = i'_k, \\ u_k[i''_k] - u_k[i'_k] \quad \text{при } i = i''_k, \\ u_k[i] \quad \text{при остальных } i. \end{cases} \tag{41}$$

4) $i'_k \in (s+1):m$, $i''_k \in 1:s$.
Имеем $\xi_{i'_k} = -1$, $\xi_{i'_k} = 1$, $\lambda_k = u_k[i''_k]$ (если $u_k[i''_k] < u_k[i'_k]$) и

$$u_{k+1} = u_k - u_k[i''_k]\big(e_{i'_k} + e_{i''_k}\big).$$

Покомпонентно:

$$u_{k+1}[i] = \begin{cases} u_k[i'_k] - u_k[i''_k] \quad \text{при } i = i'_k, \\ 0 \quad \text{при } i = i''_k, \\ u_k[i] \quad \text{при остальных } i. \end{cases} \tag{42}$$

ЛЕММА 7. *В SMO-алгоритме любая последовательность идущих подряд усеченных итераций конечна.*

Доказательство. Допустим противное, что после усеченной r-й итерации идут подряд только усеченные итерации. Тогда все u_{k+1} при $k = r, r+1, \ldots$ вычисляются по одной из формул (39)–(42). Начнем анализ с формулы (42). В ней $u_k[i''_k] > 0$ (иначе $u_{k+1} = u_k$) и $u_k[i'_k] > 0$ (по условию $u_k[i''_k] < u_k[i'_k]$). Кроме того, $u_{k+1}[i'_k] > 0$ и $u_{k+1}[i''_k] = 0$. Получаем, что при переходе от вектора u_k к вектору u_{k+1} у вектора u_{k+1} появляется дополнительная нулевая компонента. Это основной мотив доказательства леммы.

В случае, когда u_{k+1} вычисляется по формуле (41), анализ аналогичен. Имеем $u_k[i'_k] > 0$ (иначе $u_{k+1} = u_k$) и $u_k[i''_k] > 0$ (по условию $u_k[i'_k] \leqslant u_k[i''_k]$). Кроме того, $u_{k+1}[i''_k] \geqslant 0$ и $u_{k+1}[i'_k] = 0$. И здесь у вектора u_{k+1} появляется дополнительная нулевая компонента (даже две, если $u_k[i'_k] = u_k[i''_k]$).

Обратимся к формуле (40). В ней $u_k[i'_k] > 0$. Если и $u_k[i''_k] > 0$, то у вектора u_{k+1} появляется дополнительная нулевая компонента. Однако при $u_k[i''_k] = 0$ количество нулевых компонент у u_{k+1} такое же, как у u_k. Действительно, нужно учесть, что

$$u_k[i'_k] > 0,\ u_k[i''_k] = 0 \quad \text{и} \quad u_{k+1}[i'_k] = 0,\ u_{k+1}[i''_k] > 0.$$

Аналогичное заключение можно сделать и о формуле (39), которая получается из (40) при замене i'_k на i''_k и i''_k на i'_k. При $u_k[i'_k] = 0$ количество нулевых компонент у u_{k+1} такое же, как у u_k.

Неблагоприятная ситуация возникает, когда на всех итерациях, начиная с r-й, компоненты вектора u_{k+1} вычисляются по формулам (39) или (40), причем $u_k[i'_k] = 0$ в формуле (39) и $u_k[i''_k] = 0$ в формуле (40). В этом случае при переходе от u_k к u_{k+1} дополнительная нулевая компонента не появляется. Допустим, что этот процесс бесконечен.

Здесь следует заметить, что последовательные преобразования вектора u_r по формулам (39), (40), когда одна компонента очередного вектора u_k «передается» другой компоненте, порождают лишь конечное число различных векторов. Поэтому в бесконечной последовательности $\{u_k\}$, $k = r, r+1, \ldots$, встретятся два одинаковых вектора. Но это противоречит лемме 3.

Таким образом, через конечное число итераций возникнет такая ситуация: либо $u_k[i'_k] > 0$ в (39), либо $u_k[i''_k] > 0$ в (40). У вектора u_{k+1} появится дополнительная нулевая компонента.

Подведем итог. В последовательности $\{u_k\}$ при $k = r, r+1, \ldots$, во всех случаях через конечное число итераций появляется вектор u_{k+1}, у которого имеется дополнительная нулевая компонента. Учитывая бесконечность последовательности $\{u_k\}$, описанный процесс можно продолжить. Тогда обнаружится бесконечная последовательность векторов u_{k+1}, у которых появляется дополнительная нулевая компонента. Но это невозможно, поскольку все векторы u_k ненулевые.

Лемма доказана. □

Следствие. SMO-алгоритм порождает бесконечное число неусеченных итераций. Действительно, после каждого блока усеченных итераций начинаются неусеченные итерации.

7°. Слабая сходимость последовательности $\{u_k\}$ и сильная сходимость последовательности $\{v_k\}$. Будем считать выполненными условия теоремы 3, гарантирующей ограниченность последовательности (32). Из этой последовательности выделим бесконечную подпоследовательность, соответствующую неусеченным итерациям SMO-алгоритма. Обозначим ее

$$u_{k_1}, u_{k_2}, \ldots, u_{k_j}, \ldots \tag{43}$$

Неусеченность итерации с номером k_j характеризуется свойством $\lambda_{k_j} = \widetilde{\lambda}_{k_j}$. При этом по лемме 4 справедливо неравенство

$$D(u_{k_j}) - D(u_{k_j+1}) \geqslant \frac{\Delta^2(u_{k_j})}{2d^2}\,. \tag{44}$$

Согласно лемме 3 последовательность $\{D(u_k)\}$ строго убывает. Кроме того, она ограничена снизу наименьшим значением функции $D(u)$ на множестве U. Значит, последовательность $\{D(u_k)\}$ имеет предел. Обозначим его через α. В силу (38) имеем $\alpha < 0$.

Левая часть неравенства (44) стремится к нулю при $j \to \infty$. Как следствие, получаем предельное соотношение

$$\lim_{j\to\infty} \Delta(u_{k_j}) = 0. \tag{45}$$

Далее, последовательность (43) ограничена. Из нее можно выделить сходящуюся подпоследовательность. Можно считать, что вся подпоследовательность (43) сходится к некоторому вектору $\check{u}$,

$$\lim_{j\to\infty} u_{k_j} = \check{u}. \tag{46}$$

Очевидно, что $\check{u} \in U$.

В силу непрерывности функции $D(u)$ имеем

$$\lim_{j\to\infty} D(u_{k_j}) = D(\check{u}). \tag{47}$$

Необходимо $D(\check{u}) = \alpha$. Так как $\alpha < 0$, то $\check{u} \neq \mathbb{O}$.

ЛЕММА 8. *План* $\check{u} \in U$ *задачи* (4) *будет оптимальным.*

Доказательство. Нужно проверить, что $\Delta(\check{u}) = 0$. Допустим противное, что $\Delta(\check{u}) > 0$. Поскольку $\check{u} \neq \mathbb{O}$, то $I_1'(\check{u}) \neq \varnothing$ и $I_1''(\check{u}) \neq \varnothing$.

Выберем столь большое J, чтобы при $j \geqslant J$ выполнялись неравенства

$$u_{k_j}[i] > 0 \quad \text{при } i \in I_1'(\check{u}) \cup I_1''(\check{u}), \tag{48}$$

$$\Big| \big(\langle v_{k_j}, p_i\rangle - \xi_i\big) - \big(\langle \check{v}, p_i\rangle - \xi_i\big) \Big| < \varepsilon \quad \text{при } i \in 1:m, \tag{49}$$

где $v_{k_j} = Au_{k_j}$, $\check{v} = A\check{u}$ и $\varepsilon = \frac{1}{4}\,\Delta(\check{u})$. В силу (48) имеем

$$I_1'(\check{u}) \subset I_1'(u_{k_j}), \qquad I_1''(\check{u}) \subset I_1''(u_{k_j}),$$

так что

$$I''(\check{u}) \subset I''(u_{k_j}), \qquad I'(\check{u}) \subset I'(u_{k_j}). \tag{50}$$

Принимая во внимание (50), (49) и определение ε, получаем

$$\begin{aligned}
\Delta(u_{k_j}) &= \max_{i\in I''(u_{k_j})} \big\{ \langle v_{k_j}, p_i\rangle - \xi_i \big\} - \min_{i\in I'(u_{k_j})} \big\{ \langle v_{k_j}, p_i\rangle - \xi_i \big\} \geqslant \\
&\geqslant \max_{i\in I''(\check{u})} \big\{ \langle v_{k_j}, p_i\rangle - \xi_i \big\} - \min_{i\in I'(\check{u})} \big\{ \langle v_{k_j}, p_i\rangle - \xi_i \big\} \geqslant \\
&\geqslant \max_{i\in I''(\check{u})} \big\{ \langle \check{v}, p_i\rangle - \xi_i - \varepsilon \big\} - \min_{i\in I'(\check{u})} \big\{ \langle \check{v}, p_i\rangle - \xi_i + \varepsilon \big\} = \\
&= \Delta(\check{u}) - 2\varepsilon = \tfrac{1}{2}\,\Delta(\check{u}) > 0.
\end{aligned}$$

Это противоречит предельному соотношению (45).

Лемма доказана. □

Таким образом, предельный вектор $\check{u}$ является решением задачи (4). Переобозначим его. Вместо $\check{u}$ будем писать u^*. Формулы (46) и (47) примут вид

$$\lim_{j\to\infty} u_{k_j} = u^*, \qquad \lim_{j\to\infty} D(u_{k_j}) = D(u^*).$$

ТЕОРЕМА 4 (о слабой сходимости последовательности $\{u_k\}$)**.** *Справедливо предельное соотношение*

$$\lim_{k\to\infty} D(u_k) = D(u^*).$$

Доказательство. Известно, что последовательность $\{D(u_k)\}$ сходится к числу α. Любая сходящаяся подпоследовательность имеет тот же предел. Значит, $\alpha = D(u^*)$. Это соответствует заключению теоремы. □

ТЕОРЕМА 5 (о сильной сходимости последовательности $\{v_k\}$). *Справедливо предельное соотношение*

$$\lim_{k\to\infty} v_k = w^*.$$

Результат следует из теоремы 4 и леммы 2. □

8°. Заключительные замечания.

- В докладе рассматривается задача (1) о строгом линейном отделении двух конечных множеств P_1 и P_2 при условии, что их выпуклые оболочки не имеют общих точек.

- Реально решается преобразованная двойственная задача (4), единственность оптимального плана которой не гарантируется. С помощью SMO-алгоритма строится минимизирующая последовательность $\{u_k\}$ с предельной точкой u^*, являющейся решением задачи (4).

- Оптимальная нормаль w^* в задаче (1) вычисляется по формуле $w^* = Au^*$. Константа β^* определяется из условий дополнительности (3), если записать их в виде

  $$\xi_i\big(\langle w^*, p_i\rangle + \beta^*\big) = 1 \quad \text{при } i \in I^*,$$

  где $I^* = \big\{i \in 1 : m \mid u^*[i] > 0\big\}$.

- Гиперплоскость, отделяющая множество P_1 от P_2, задается уравнением

  $$\langle w^*, x\rangle + \beta^* = 0.$$

  Границы разделяющей полосы определяются так:

  $$\langle w^*, x\rangle + \beta^* = 1 \quad \text{и} \quad \langle w^*, x\rangle + \beta^* = -1.$$

  Точки p_i при $i \in I^*$ называются *опорными*. Они лежат на гиперплоскостях, ограничивающих разделяющая полосу. С их помощью оптимальная нормаль w^* допускает представление

  $$w^* = \sum_{i\in I^*} u^*[i]\,\xi_i p_i.$$

- В качестве начального приближения для SMO-алгоритма можно взять нулевой вектор, $u_0 = \mathbb{O}$ (см. разд. 4). В этом случае $v_0 = \mathbb{O}$, $I'(u_0) = 1 : s$, $I''(u_0) = (s+1) : m$, i'_0 — любой индекс из множества $1 : s$, i''_0 — любой индекс из множества $(s+1) : m$. Согласно (17), $\Delta(u_0) = 2$. По определению λ_0 имеем

$$\lambda_0 = \widetilde{\lambda}_0 = \frac{2}{\left\|p_{i'_0} - p_{i''_0}\right\|^2}\,.$$

В силу (29) получаем

$$u_1 = \frac{2}{\left\|p_{i'_0} - p_{i''_0}\right\|^2}\left(e_{i'_0} + e_{i''_0}\right).$$

ЛИТЕРАТУРА

1. J. C. Platt. *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines.* Technical Report MSR-TR-98-14. April 21, 1998.

2. J. Lopez, J. R. Dorronsoro. *A Simple Proof of the Convergence of the SMO Algorithm for Linearly Separable Problems.* ICANN 2009, Part I, LNCS 5768, pp. 904–912, 2009.

3. J. Lopez, A. Barbero, and J. R. Dorronsoro. *On the Equivalence of the SMO and MDM Algorithms for SVM Training.* W. Daelemans et al. (Eds.): ECML PKDD 2008, Part I, LNAI 5211, pp. 288–300, 2008.

4. Малоземов В. Н., Соловьева Н. А. *SMO-алгоритм как обобщение MDM-алгоритма*// Семинар «OML». Избранные доклады. 16 февраля 2023 г. (http://oml.cmlaboratory.com/reps23.shtml#0216)

5. Малоземов В. Н., Соловьева Н. А. *Картина маслом* // Семинар «OML». Избранные доклады. 11 мая 2023 г. (http://oml.cmlaboratory.com/reps23.shtml#0511)

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