MDM-АЛГОРИТМ И ЗАДАЧА СИЛЬВЕСТРА. I*

В. Н. Малоземов — v.malozemov@spbu.ru

Н. А. Соловьева — 4vinyo@gmail.com

Г. Ш. Тамасян — grigoriytamasjan@mail.ru

14 сентября 2023 г.

1°. Постановка задачи. В пространстве $\mathbb{R}^n$ с евклидовой нормой заданы m точек $a_1, a_2, \ldots, a_m$. Задача Сильвестра ставится так:

найти шар наименьшего объема, содержащий все точки a_i, $i \in 1:m$.

Запишем формализацию этой задачи

$$\max_{i\in 1:m}\left\{\tfrac{1}{2}\,\|a_i - x\|^2\right\} \to \min_{x\in\mathbb{R}^n}. \tag{1}$$

Здесь x — центр шара. Задача (1) имеет решение x^* и оно единственно.

Обозначим через A матрицу со столбцами $a_1, a_2, \ldots, a_m$ и через b — вектор из $\mathbb{R}^m$ с компонентами $b[i] = \frac{1}{2}\,\|a_i\|^2$. Рассмотрим вспомогательную задачу квадратичного программирования

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

Множество ее планов обозначим через U. Очевидно, что у задачи (2) существует оптимальный план.

Справедливо следующее утверждение (см., например, [1, с. 117–119] и [2]).

ТЕОРЕМА 1. *Если u^* — оптимальный план задачи* (2)*, то вектор $x^* = Au^*$ будет решением задачи* (1)*. При этом*

$$\max_{i\in 1:m}\left\{\tfrac{1}{2}\,\|a_i - x^*\|^2\right\} = -Q(u^*).$$

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

Таким образом, задача Сильвестра (1) сводится к задаче квадратичного программирования (2). Для решения задачи (2) существуют известные конечные методы (например, метод Данцига [3, с. 64–85]). Они хорошо работают при небольших m и n. Однако при больших m и n могут разваливаться вследствие накопления вычислительных погрешностей (например, при пересчете обратных базисных матриц). В связи с этим разрабатываются простые итеративные методы, бесконечные в общем случае, с гарантированной сходимостью. В данном докладе предлагается такой метод для решения задачи (2). Он аналогичен MDM-алгоритму[1], разработанному для нахождения точки из выпуклой оболочки конечного множества, ближайшей к началу координат [4, 5].

З а м е ч а н и е. Целевая функция $Q(u)$ задачи (2) на множестве планов U допускает представление

$$Q(u) = -\tfrac{1}{2} \sum_{i=1}^{m} u[i] \cdot \|a_i - x\|^2 , \tag{3}$$

где

$$x = Au = \sum_{i=1}^{m} u[i]\, a_i.$$

Действительно, в силу определения $b[i]$ имеем

$$\begin{aligned} \sum_{i=1}^{m} u[i] \cdot \|a_i - x\|^2 &= 2\langle b, u\rangle - 2\langle Au, x\rangle + \|x\|^2 = \\ &= -2\bigl(\tfrac{1}{2}\, \|x\|^2 - \langle b, u\rangle\bigr) = -2Q(u), \end{aligned}$$

что равносильно (3). Из (3) следует, что $Q(u) \leqslant 0$ при всех $u \in U$ и что на ортах e_j пространства $\mathbb{R}^m$ справедливы равенства

$$Q(e_j) = -\tfrac{1}{2}\, \|a_j - Ae_j\|^2 = 0, \quad j \in 1 : m.$$

2°. Критерий оптимальности. Получим нестандартный вариант критерия оптимальности для задачи (2).

ТЕОРЕМА 2. *План u задачи (2) будет оптимальным тогда и только тогда, когда для вектора $x = Au$ при некотором вещественном μ выполняются соотношения*

$$\langle a_i, x\rangle - b[i] = \mu, \quad \text{\textit{если} } u[i] > 0, \tag{4}$$

$$\langle a_i, x\rangle - b[i] \geqslant \mu, \quad \text{\textit{если} } u[i] = 0. \tag{5}$$

[1] MDM — сокращение от Mitchell-Demyanov-Malozemov.

Д о к а з а т е л ь с т в о. Напомним стандартный критерий оптимальности, использующий условия Куна–Таккера: для того, чтобы план u задачи (2) был оптимальным, необходимо и достаточно, чтобы нашлись вещественные числа μ, $t_1, \ldots, t_m$, такие, что

$$A^T Au - b = \mu e + \sum_{i=1}^{m} \lambda_i e_i, \tag{6}$$

$$\lambda_i \, u[i] = 0 \quad \text{и} \quad \lambda_i \geqslant 0 \;\text{ при всех } i \in 1:m. \tag{7}$$

Здесь e — вектор из $\mathbb{R}^m$, все компоненты которого равны единице и e_i — i-й орт в $\mathbb{R}^m$. Покажем, что условия (4), (5) и (6), (7) эквивалентны.

Предположим, что выполняются условия (6), (7). Обозначим $x = Au$ и распишем равенство (6) покомпонентно:

$$\langle a_i, x\rangle - b[i] = \mu + \lambda_i, \quad i \in 1:m. \tag{8}$$

Если $u[i] > 0$, то, согласно (7), $\lambda_i = 0$. В этом случае из (8) следует (4). При $u[i] = 0$ воспользуемся тем, что $\lambda_i \geqslant 0$. Тогда из (8) будет следовать (5).

Проверим обратное утверждение. Пусть выполнены условия (4), (5). Обозначим

$$\lambda_i = \langle a_i, x\rangle - b[i] - \mu, \quad i \in 1:m. \tag{9}$$

Формула (9) равносильна формуле (8), которая в свою очередь представляет собой покомпонентную запись равенства (6). Далее, по определению, $\lambda_i = 0$ при $u[i] > 0$ и $\lambda_i \geqslant 0$ при $u[i] = 0$. Это гарантирует выполнение условий (7).

Теорема доказана. $\square$

З а м е ч а н и е. Для константы μ можно получить явное представление

$$\mu = \|x\|^2 - \langle b, u\rangle.$$

Покажем это. Согласно (4), (5) и определению $b[i]$ имеем

$$\begin{aligned} &\|x\|^2 - \|a_i - x\|^2 = 2\big(\langle a_i, x\rangle - b[i]\big) = 2\mu \quad \text{при } u[i] > 0, \\ &\|x\|^2 - \|a_i - x\|^2 \geqslant 2\mu \quad \text{при } u[i] = 0. \end{aligned}$$

Отсюда следует, что

$$\begin{aligned} \|a_i - x\|^2 &= \|x\|^2 - 2\mu \quad \text{при } u[i] > 0, \\ \|a_i - x\|^2 &\leqslant \|x\|^2 - 2\mu \quad \text{при } u[i] = 0. \end{aligned}$$

Значит,

$$R^2 := \max_{i \in 1:m} \|a_i - x\|^2 = \|x\|^2 - 2\mu.$$

Вместе с тем, по теореме 1, $R^2 = -2Q(u)$, так что

$$\mu = \frac{1}{2}\|x\|^2 + Q(u) = \|x\|^2 - \langle b, u\rangle.$$

Критерий оптимальности (4), (5) можно еще упростить. Для этого введем *оценку плана* u — величину

$$\Delta(u) = \max_{i\in I^+(u)}\bigl\{\langle a_i, x\rangle - b[i]\bigr\} - \min_{i\in 1:m}\bigl\{\langle a_i, x\rangle - b[i]\bigr\}. \tag{10}$$

Здесь $x = Au$ и $I^+(u) = \{i \in 1:m \mid u[i] > 0\}$ — носитель плана u. Величина $\Delta(u)$ при всех $u \in U$ неотрицательна. Действительно,

$$\Delta(u) \geqslant \max_{i\in I^+(u)}\bigl\{\langle a_i, x\rangle - b[i]\bigr\} - \min_{i\in I^+(u)}\bigl\{\langle a_i, x\rangle - b[i]\bigr\} \geqslant 0. \tag{11}$$

ТЕОРЕМА 3. *План u задачи* (2) *будет оптимальным тогда и только тогда, когда* $\Delta(u) = 0$.

Доказательство. Равенство $\Delta(u) = 0$ непосредственно следует из условий (4), (5). Обратно, пусть $\Delta(u) = 0$. Согласно (11) имеем

$$\max_{i\in I^+(u)}\bigl\{\langle a_i, x\rangle - b[i]\bigr\} = \min_{i\in I^+(u)}\bigl\{\langle a_i, x\rangle - b[i]\bigr\}.$$

Значит,

$$\langle a_i, x\rangle - b[i] =: \mu \quad \text{при} \;\; i \in I^+(u). \tag{12}$$

Еще раз воспользуемся условием $\Delta(u) = 0$. Получим

$$\mu = \max_{i\in I^+(u)}\bigl\{\langle a_i, x\rangle - b[i]\bigr\} = \min_{i\in 1:m}\bigl\{\langle a_i, x\rangle - b[i]\bigr\},$$

так что

$$\langle a_i, x\rangle - b[i] \geqslant \mu \quad \text{при} \;\; i \in 1:m. \tag{13}$$

Условия (12) и (13) гарантируют оптимальность плана u.

Теорема доказана. □

С помощью $\Delta(u)$ можно оценить близость вектора $x = Au$ к решению x^* задачи (1).

ТЕОРЕМА 4. *Пусть u — план задачи* (2) *и $x = Au$. Тогда*

$$\|x - x^*\|^2 \leqslant \Delta(u). \tag{14}$$

Доказательство. По теореме 1 справедливо равенство $x^* = Au^*$, где u^* — оптимальный план задачи (2). Согласно теореме 3, $\Delta(u^*) = 0$.

Имеем

$$\|x - x^*\|^2 = \langle x, x - x^*\rangle - \langle x^*, x - x^*\rangle.$$

Воспользуемся ограничениями на план задачи (2). Получим

$$\begin{aligned}
&\langle x, x - x^*\rangle = \langle u, A^T(x - x^*)\rangle = \\
&= \sum_{i=1}^m u[i] \times \Big(\big(\langle a_i, x\rangle - b[i]\big) - \big(\langle a_i, x^*\rangle - b[i]\big) \Big) \leqslant \\
&\leqslant \max_{i \in I^+(u)} \big\{ \langle a_i, x\rangle - b[i] \big\} - \min_{i \in 1:m} \big\{ \langle a_i, x^*\rangle - b[i] \big\}.
\end{aligned} \tag{15}$$

Аналогично

$$\begin{aligned}
\langle x^*, x - x^*\rangle &= \sum_{i=1}^m u^*[i] \times \Big(\big(\langle a_i, x\rangle - b[i]\big) - \big(\langle a_i, x^*\rangle - b[i]\big) \Big) \geqslant \\
&\geqslant \min_{i \in 1:m} \big\{ \langle a_i, x\rangle - b[i] \big\} - \max_{i \in I^+(u^*)} \big\{ \langle a_i, x^*\rangle - b[i] \big\}.
\end{aligned} \tag{16}$$

Вычтем неравенство (16) из неравенства (15). Так как $\Delta(u^*) = 0$, то придем к требуемому неравенству (14).

Теорема доказана. □

Оценке $\Delta(u)$ при $u \in U$ можно дать эквивалентное определение в терминах исходной задачи (1). А именно,

$$\Delta(u) = \tfrac{1}{2}\big(\max_{i \in 1:m} \|a_i - x\|^2 - \min_{i \in I^+(u)} \|a_i - x\|^2 \big), \tag{17}$$

где $x = Au$. Действительно, так как $b[i] = \frac{1}{2}\|a_i\|^2$, то

$$\langle a_i, x\rangle - b[i] = \tfrac{1}{2}\big(\|x\|^2 - \|a_i - x\|^2 \big).$$

Согласно определению (10) имеем

$$\Delta(u) = \tfrac{1}{2}\Big(\max_{i \in I^+(u)} \big\{ \|x\|^2 - \|a_i - x\|^2 \big\} - \min_{i \in 1:m} \big\{ \|x\|^2 - \|a_i - x\|^2 \big\} \Big).$$

Воспользуемся следующими свойствами функций максимума и минимума:

$$\begin{aligned}
&\max_{i \in 1:m} \big\{ \alpha_i + c \big\} = c + \max_{i \in 1:m} \big\{ \alpha_i \big\}, &\qquad &\min_{i \in 1:m} \big\{ \alpha_i + c \big\} = c + \min_{i \in 1:m} \big\{ \alpha_i \big\}, \\
&\max_{i \in 1:m} \big\{ -\alpha_i \big\} = -\min_{i \in 1:m} \big\{ \alpha_i \big\}, & &\min_{i \in 1:m} \big\{ -\alpha_i \big\} = -\max_{i \in 1:m} \big\{ \alpha_i \big\}.
\end{aligned}$$

Получим

$$\Delta(u) = \tfrac{1}{2}\Big(\max_{i\in I^+(u)}\big\{-\|a_i - x\|^2\big\} - \min_{i\in 1:m}\big\{-\|a_i - x\|^2\big\}\Big) =$$
$$= \tfrac{1}{2}\big(\max_{i\in 1:m}\|a_i - x\|^2 - \min_{i\in I^+(u)}\|a_i - x\|^2\big),$$

что соответствует формуле (17).

ТЕОРЕМА 5. *План u задачи* (2) *будет оптимальным тогда и только тогда, когда для вектора $x = Au$ при всех $s \in I^+(u)$ выполняются равенства*

$$\|a_s - x\| = \max_{i\in 1:m}\|a_i - x\|. \tag{18}$$

Доказательство. Согласно (17) критерий оптимальности $\Delta(u) = 0$ равносилен соотношению

$$\max_{i\in 1:m}\|a_i - x\| = \min_{i\in I^+(u)}\|a_i - x\|. \tag{19}$$

Если u — оптимальный план, то согласно (19) при всех $s \in I^+(u)$ имеем

$$\|a_s - x\| \geqslant \min_{i\in I^+(u)}\|a_i - x\| = \max_{i\in 1:m}\|a_i - x\|.$$

Обратное неравенство очевидно. Значит, справедливы равенства (18).

Наоборот, если выполнено условие (18), то справедливо соотношение (19). Оно эквивалентно равенству $\Delta(u) = 0$, гарантирующему оптимальность плана u.

Теорема доказана. □

Теорему 5 можно переформулировать в геометрических терминах:
план u задачи (2) *будет оптимальным тогда и только тогда, когда точки a_s при всех $s \in I^+(u)$ лежат на поверхности шара с центром в точке $x = Au$ и радиусом R, равным* $\max\limits_{i\in 1:m}\|a_i - x\|$.

Такие точки a_s будем называть *опорными*.

3°. Вариация плана задачи (2). Зафиксируем план u задачи (2), отличный от оптимального. Согласно теореме 2, его оценка $\Delta(u)$ положительна. Рассмотрим вариацию $\widetilde{u}$ плана u следующего вида:

$$\widetilde{u} = u + \delta\big(e_{i'} - e_{i''}\big). \tag{20}$$

Распишем равенство векторов (20) покомпонентно:

$$\widetilde{u}[i] = \begin{cases} u[i'] + \delta \quad \text{при } i = i', \\ u[i''] - \delta \quad \text{при } i = i'', \\ u[i] \quad \text{при остальных } i \in 1:m. \end{cases}$$

Будем считать, что выполнено условие

$$0 < \delta \leqslant u[i'']. \tag{21}$$

Представление (20) и условие (21) гарантируют принадлежность вектора $\widetilde{u}$ множеству планов U задачи (2). Выбор параметра δ и индексов i', i'' из множества $1 : m$ должен обеспечить максимально возможное уменьшение значения $Q(\widetilde{u})$ целевой функции по сравнению со значением $Q(u)$.

Вычислим $Q(\widetilde{u})$. Обозначим $x = Au$. Отметим, что $Ae_i = a_i$. На основании (20) имеем

$$\begin{gathered}\tfrac{1}{2}\langle A^T A\widetilde{u}, \widetilde{u}\rangle = \tfrac{1}{2}\,\|A\widetilde{u}\|^2 = \tfrac{1}{2}\,\big\|\,x + \delta\big(a_{i'} - a_{i''}\big)\big\|^2 = \\ = \tfrac{1}{2}\,\|x\|^2 - \delta\langle a_{i''} - a_{i'}, x\rangle + \tfrac{1}{2}\delta^2\,\|a_{i'} - a_{i''}\|^2\,, \\ \langle b, \widetilde{u}\rangle = \langle b, u\rangle - \delta\big(b[i''] - b[i']\big).\end{gathered}$$

Вычитая второе равенство из первого, получаем

$$Q(\widetilde{u}) = Q(u) - \delta\big[\big(\langle a_{i''}, x\rangle - b[i'']\big) - \big(\langle a_{i'}, x\rangle - b[i']\big)\big] + \tfrac{1}{2}\delta^2\,\|a_{i'} - a_{i''}\|^2\,. \tag{22}$$

Теперь выберем индексы i' и i'' из следующих соображений: $i' \in 1 : m$, $i'' \in I^+(u)$ и

$$\begin{gathered}\langle a_{i''}, x\rangle - b[i''] = \max_{i \in I^+(u)}\big\{\langle a_i, x\rangle - b[i]\big\}, \\ \langle a_{i'}, x\rangle - b[i'] = \min_{i \in 1:m}\big\{\langle a_i, x\rangle - b[i]\big\}.\end{gathered} \tag{23}$$

Тогда выражение в квадратных скобках из правой части формулы (22) будет равно $\Delta(u)$ (см. определение (10)). Сама формула (22) примет вид

$$Q(\widetilde{u}) = Q(u) - \delta\Delta(u) + \tfrac{1}{2}\delta^2\,\|a_{i'} - a_{i''}\|^2\,. \tag{24}$$

Условие $i'' \in I^+(u)$ означает, что $u[i''] > 0$. Это важное для дальнейшего неравенство. Оно, в частности, согласуется с условием (21).

Выражение $Q(\widetilde{u})$ как функция от δ является квадратным трехчленом. Его минимум достигается при $\delta = \widetilde{\delta}$, где

$$\widetilde{\delta} = \frac{\Delta(u)}{\|a_{i'} - a_{i''}\|^2}\,. \tag{25}$$

Чтобы удовлетворить условию (21), придется подкорректировать δ. Положим

$$\delta = \begin{cases}\widetilde{\delta}, & \text{если } \widetilde{\delta} \leqslant u[i''], \\ u[i''], & \text{если } \widetilde{\delta} > u[i''].\end{cases} \tag{26}$$

Ясно, что $0 < \delta \leqslant \widetilde{\delta}$.

Все параметры в формуле (20) для $\widetilde{u}$ определены. Индексы i' и i'' находятся по формуле (23), а параметр δ — по формуле (26). При этом значение $Q(\widetilde{u})$ вычисляется по формуле (24).

Отметим, что при $\delta = \widetilde{\delta}$ равенство (24) упрощается:

$$Q(\widetilde{u}) = Q(u) - \tfrac{1}{2}\widetilde{\delta}\Delta(u).$$

Отсюда в силу неравенства $\delta \leqslant \widetilde{\delta}$ следует, что в общем случае

$$Q(\widetilde{u}) \leqslant Q(u) - \tfrac{1}{2}\delta\Delta(u). \tag{27}$$

4°. MDM-алгоритм. Опишем, как применяется MDM-алгоритм к решению задачи Сильвестра.

В качестве начального приближения возьмем произвольный план u_0 задачи (2). Ему соответствует вектор $x_0 = Au_0$.

Пусть уже имеется k-е приближение $u_k \in U$. Ему соответствует вектор $x_k = Au_k$. Проверим план u_k на оптимальность. Для этого выберем индексы $i'_k \in 1:m$ и $i''_k \in I^+(u_k)$ из условий

$$\begin{aligned} \langle a_{i'_k}, x_k\rangle - b[i'_k] &= \min_{i\in 1:m}\bigl\{\langle a_i, x_k\rangle - b[i]\bigr\}, \\ \langle a_{i''_k}, x_k\rangle - b[i''_k] &= \max_{i\in I^+(u_k)}\bigl\{\langle a_i, x_k\rangle - b[i]\bigr\}. \end{aligned}$$

Вычислим оценку $\Delta(u_k)$ плана u_k:

$$\Delta(u_k) = \langle a_{i''_k} - a_{i'_k}, x_k\rangle - \bigl(b[i''_k] - b[i'_k]\bigr).$$

Если $\Delta(u_k) = 0$, то u_k — решение задачи (2), а x_k — решение задачи (1). Процесс закончен.

Пусть $\Delta(u_k) > 0$. В этом случае будем говорить, что план u_k порождает k-ю итерацию. Вычислим

$$\widetilde{\delta}_k = \frac{\Delta(u_k)}{\bigl\|a_{i'_k} - a_{i''_k}\bigr\|^2} \tag{28}$$

и положим

$$\delta_k = \begin{cases} \widetilde{\delta}_k, & \text{если } \widetilde{\delta}_k \leqslant u_k[i''_k], \\ u_k[i''_k], & \text{если } \widetilde{\delta}_k > u_k[i''_k]. \end{cases} \tag{29}$$

Очевидно, что $0 < \delta_k \leqslant \widetilde{\delta}_k$. Если $\delta_k = \widetilde{\delta}_k$, то k-ю итерацию будем называть *неусеченной*. При $\delta_k < \widetilde{\delta}_k$ получаем *усеченную итерацию*. Формулы

$$\begin{aligned} u_{k+1} &= u_k + \delta_k\bigl(e_{i'_k} - e_{i''_k}\bigr), \\ x_{k+1} &= x_k + \delta_k\bigl(a_{i'_k} - a_{i''_k}\bigr) \end{aligned} \tag{30}$$

определяют очередное $(k+1)$-е приближение. Описание алгоритма завершено.

С помощью предложенного алгоритма строятся две последовательности

$$\begin{aligned} &u_0, u_1, \ldots, u_k, \ldots, \\ &x_0, x_1, \ldots, x_k, \ldots \end{aligned} \tag{31}$$

Они конечны, если при некотором k выполнится равенство $\Delta(u_k) = 0$. Они бесконечны, когда $\Delta(u_k) > 0$ при всех $k = 0, 1, \ldots$ Нам предстоит изучить вопрос о сходимости этих последовательностей.

ЛИТЕРАТУРА

1. Гавурин М. К., Малоземов В. Н. *Экстремальные задачи с линейными ограничениями.* Л.: Изд-во ЛГУ, 1984. 176 с.

2. Малоземов В. Н., Плоткин А. В. *Двойственность в квадратичном программировании. Задача Сильвестра* // Семинар «CNSA & NDO». Избранные доклады. 8 декабря 2021 г.
(http://cnsa.cmlaboratory.com/reps21.shtml#1208)

3. Даугавет В. А. *Численные методы квадратичного программирования.* СПб.: Изд-во СПбГУ, 2004. 128 с.

4. Митчелл Б. Ф., Демьянов В. Ф., Малоземов В. Н. *Нахождение ближайшей к началу координат точки многогранника* // Вестник ЛГУ. 1971. № 19. С. 38–45.

5. Малоземов В. Н. *MDM-методу — 50 лет* // Семинар «CNSA & NDO». Избранные доклады. 10 ноября 2021 г.
(http://cnsa.cmlaboratory.com/reps21.shtml#1110)