МДМ-МЕТОД ДЛЯ РЕШЕНИЯ ОБЩЕЙ КВАДРАТИЧНОЙ ЗАДАЧИ МАТЕМАТИЧЕСКОЙ ДИАГНОСТИКИ*

В. Н. Малоземов

Н. А. Соловьева

v.malozemov@spbu.ru

4vinyo@gmail.com

1 июня 2022 г.

МДМ-метод был разработан в 1971 г. для поиска направления наискорейшего спуска в минимаксных задачах [1]. В дальнейшем он нашел применение в машинном обучении [2, 3]. В основе МДМ-метода лежит алгоритм ортогонального проектирования произвольной точки на выпуклую оболочку конечного числа заданных точек. В данном докладе рассматривается обобщение МДМметода на случай отделения двух выпуклых оболочек. Обобщенный МДМметод ориентирован на решение общей квадратичной задачи математической диагностики [4].

 1° . Напомним постановку общей квадратичной задачи математической диагностики [4]. В пространстве \mathbb{R}^n заданы два конечных множества

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

где $s \in 1: m-1$. Обозначим через C_1 и C_2 выпуклые оболочки множеств P_1 и P_2 соответственно. Требуется решить экстремальную задачу

$$\frac{1}{2} \|w_1 - w_2\|^2 \to \min_{w_1 \in C_1, w_2 \in C_2}.$$
 (1)

Очевидно, что задача (1) имеет решение (w_1^*, w_2^*) .

По определению выпуклой оболочки точки $w_1 \in C_1$ и $w_2 \in C_2$ допускают представления

$$w_1 = \sum_{j=1}^{s} u_j p_j, \quad w_2 = \sum_{j=s+1}^{m} u_j p_j,$$

 $^{^*}$ Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» http://www.apmath.spbu.ru/oml/

где

$$\sum_{j=1}^{s} u_{j} = 1, \quad u_{j} \geqslant 0 \text{ при всех } j \in 1:s,$$

$$\sum_{j=s+1}^{m} u_{j} = 1, \quad u_{j} \geqslant 0 \text{ при всех } j \in (s+1):m.$$
(2)

Введем вектор $\xi=(\xi_1,\ldots,\xi_m)$ с компонентами $\xi_j=1$ при $j\in 1:s$ и $\xi_j=-1$ при $j\in (s+1):m$. Обозначим через A матрицу со столбцами ξ_1p_1,\ldots,ξ_mp_m . Тогда для любого плана (w_1,w_2) задачи (1) справедлива формула

$$w_1 - w_2 = \sum_{j=1}^m u_j \xi_j p_j = Au.$$

Здесь вектор u удовлетворяет условиям (2). Множество таких векторов обозначим через U. Задачу (1) можно переписать в следующем виде:

$$\frac{1}{2} \|w\|^2 \to \min,$$

$$w = Au, \quad u \in U.$$
(3)

2° . Отметим одно свойство планов задачи (1).

ЛЕММА 1. Пусть $w^* = w_1^* - w_2^* - pешение задачи (1)$. Для любого плана $w = w_1 - w_2$ этой задачи справедливо неравенство

$$\langle w, w^* \rangle \geqslant \langle w^*, w^* \rangle. \tag{4}$$

Доказательство. В силу выпуклости множества C_1 точка $w_1^* + t(w_1 - w_1^*)$ при всех $t \in (0,1)$ принадлежит C_1 . Точка w_2^* принадлежит C_2 . Значит, планом задачи (1) является точка

$$(w_1^* + t(w_1 - w_1^*)) - w_2^* = w^* + t(w_1 - w_1^*).$$

Воспользуемся оптимальностью w^* . Получим

$$\|w^*\|^2 \le \|w^* + t(w_1 - w_1^*)\|^2 = \|w^*\|^2 + 2t\langle w^*, w_1 - w_1^* \rangle + t^2 \|w_1 - w_1^*\|^2.$$

Отсюда следует, что

$$\langle w^*, w_1 - w_1^* \rangle + \frac{1}{2} t \| w_1 - w_1^* \|^2 \geqslant 0.$$

В пределе при $t \to +0$ придем к неравенству

$$\langle w^*, w_1 - w_1^* \rangle \geqslant 0. \tag{5}$$

Далее, в силу выпуклости множества C_2 точка $w_2^* + t(w_2 - w_2^*)$ при всех $t \in (0,1)$ принадлежит C_2 . Точка w_1^* принадлежит C_1 . Значит, планом задачи (1) является точка

$$w_1^* - (w_2^* + t(w_2 - w_2^*)) = w^* - t(w_2 - w_2^*).$$

Воспользуемся оптимальностью w^* . Получим

$$\|w^*\|^2 \le \|w^* - t(w_2 - w_2^*)\|^2 = \|w^*\|^2 - 2t\langle w^*, w_2 - w_2^* \rangle + t^2 \|w_2 - w_2^*\|^2$$
.

Отсюда следует, что

$$-\langle w^*, w_2 - w_2^* \rangle + \frac{1}{2} t \|w_2 - w_2^*\|^2 \geqslant 0.$$

В пределе при $t \to +0$ придем к неравенству

$$-\langle w^*, w_2 - w_2^* \rangle \geqslant 0. \tag{6}$$

Требуемое неравенство (4) есть сумма неравенств (5) и (6). Лемма доказана. \Box

Отметим, что неравенство (4) равносильно следующему неравенству:

$$\|w - w^*\|^2 \leqslant \|w\|^2 - \|w^*\|^2. \tag{7}$$

 3° . Наша цель — распространить МДМ-метод на решение задачи (1), (3). Классический вариант МДМ-метода описан в работе [1]. Проведем предварительный анализ ситуации.

Пусть w = Au, $u \in U$. Обозначим

$$M_1^+(u) = \big\{ j \in 1 : s \mid u[j] > 0 \big\}, \quad M_2^+(u) = \big\{ j \in (s+1) : m \mid u[j] > 0 \big\}.$$

Введем величины

$$\Delta_1(u) = \max_{j \in M_1^+(u)} \langle p_j, w \rangle - \min_{j \in 1:s} \langle p_j, w \rangle,$$

$$\Delta_2(u) = \max_{j \in M_2^+(u)} \langle p_j, -w \rangle - \min_{j \in (s+1):m} \langle p_j, -w \rangle.$$

Положим

$$\Delta(u) = \max\{\Delta_1(u), \Delta_2(u)\}.$$

Очевидно, что

$$\Delta_1(u) \geqslant 0, \quad \Delta_2(u) \geqslant 0, \quad \Delta(u) \geqslant 0.$$
(8)

ЛЕММА 2. Возъмем произвольный план w = Au задачи (3) и её оптимальный план $w^* = Au^*$. Справедливо неравенство

$$\|w - w^*\|^2 \leqslant 2\Delta(u). \tag{9}$$

Доказательство. С учетом (4) запишем

$$\begin{split} \|w - w^*\|^2 &= \|w\|^2 - 2\langle w, w^* \rangle + \|w^*\|^2 \leqslant \|w\|^2 - \langle w, w^* \rangle = \\ &= \sum_{j=1}^m \xi_j u_j \langle p_j, w \rangle - \sum_{j=1}^m \xi_j u_j^* \langle p_j, w \rangle = \\ &= \sum_{j=1}^s u_j \langle p_j, w \rangle + \sum_{j=s+1}^m u_j \langle p_j, -w \rangle - \sum_{j=1}^s u_j^* \langle p_j, w \rangle - \sum_{j=s+1}^m u_j^* \langle p_j, -w \rangle. \end{split}$$

Объединив первую и третью, вторую и четвертую суммы, придем к неравенству

$$\|w - w^*\|^2 \leqslant \max_{j \in M_1^+(u)} \langle p_j, w \rangle - \min_{j \in 1:s} \langle p_j, w \rangle +$$

$$+ \max_{j \in M_2^+(u)} \langle p_j, -w \rangle - \min_{j \in (s+1):m} \langle p_j, -w \rangle = \Delta_1(u) + \Delta_2(u) \leqslant 2\Delta(u).$$

Лемма доказана.

TEOPEMA 1. Равенство $\Delta(u) = 0$ выполняется тогда и только тогда, когда вектор w = Au является решением задачи (1), (3).

Доказательство. Неравенство (9) гарантирует оптимальность вектора w при $\Delta(u)=0$. Проверим справедливость обратного утверждения. Возьмём решение $w^*=Au^*$ задачи (3) и покажем, что $\Delta(u^*)=0$. Напомним, что $w^*=w_1^*-w_2^*$, где $w_1^*\in C_1,\,w_2^*\in C_2$.

Вначале рассмотрим случай, когда $\Delta(u^*) = \Delta_1(u^*)$. Пусть

$$\Delta_1(u^*) = \langle p_{j'} - p_{j''}, w^* \rangle. \tag{10}$$

Индексы j' и j'' принадлежат множествам $M_1^+(u^*)$ и 1:s соответственно. Введем вектор

$$\widetilde{w}_1^* = w_1^* - u_{j'}^* (p_{j'} - p_{j''})$$

(коэффициент при $p_{j'}$ передали вектору $p_{j''}$). Очевидно, что $\widetilde{w}_1^* \in C_1$. Обозначим

$$\widetilde{w}^* = \widetilde{w}_1^* - w_2^* = w^* - u_{j'}^* (p_{j'} - p_{j''}).$$

Умножим \widetilde{w}^* скалярно на w^* . С учетом (10) получим

$$\langle \widetilde{w}^*, w^* \rangle = \langle w^*, w^* \rangle - u_{j'}^* \Delta_1(u^*).$$

Согласно лемме 1 справедливо неравенство $\langle \widetilde{w}^*, w^* \rangle \geqslant \langle w^*, w^* \rangle$, поэтому $-u_{j'}^* \Delta_1(u^*) \geqslant 0$. Условие $j' \in M_1^+(u^*)$ гарантирует, что $u_{j'}^* > 0$. Приходим к неравенству $\Delta_1(u^*) \leqslant 0$, которое вместе с обратным неравенством $\Delta_1(u^*) \geqslant 0$ (см. (8)) обеспечивает равенство $\Delta_1(u^*) = 0$.

Рассмотрим второй случай, когда $\Delta(u^*) = \Delta_2(u^*)$. Пусть

$$\Delta_2(u^*) = \langle p_{l'} - p_{l''}, -w^* \rangle. \tag{11}$$

Здесь индексы l' и l'' принадлежат множествам $M_2^+(u^*)$ и (s+1):m соответственно. Введем вектор

$$\widetilde{w}_{2}^{*} = w_{2}^{*} - u_{l'}^{*} (p_{l'} - p_{l''}).$$

Очевидно, что $\widetilde{w}_2^* \in C_2$. Обозначим

$$\widetilde{w}^* = w_1^* - \widetilde{w}_2^* = w^* + u_{l'}^* (p_{l'} - p_{l''}).$$

Умножим \widetilde{w}^* скалярно на w^* . С учетом (11) получим

$$\langle \widetilde{w}^*, w^* \rangle = \langle w^*, w^* \rangle - u_{l'}^* \langle p_{l'} - p_{l''}, -w^* \rangle = \langle w^*, w^* \rangle - u_{l'}^* \Delta_2(u^*).$$

Отсюда, как и в предыдущем случае, следует неравенство $\Delta_2(u^*) \leq 0$, которое вместе с обратным неравенством $\Delta_2(u^*) \geq 0$ (см. (8)) обеспечивает равенство $\Delta_2(u^*) = 0$.

4°. Обратимся к описанию обобщенного МДМ-метода.

Возьмем начальное приближение $w_0 = Au_0$. Если $\Delta(u_0) = 0$, то по теореме 1 вектор w_0 является решением задачи (1). В противном случае переходим к следующей итерации.

Пусть уже имеется k-е приближение $w_k = Au_k$. Опишем переход к w_{k+1} . Найдем индексы $j_k' \in M_1^+(u_k), j_k'' \in 1: s$ и $l_k' \in M_2^+(u_k), l_k'' \in (s+1): m$, такие,

$$\begin{split} \max_{i \in M_1^+(u)} \langle p_i, w_k \rangle &= \langle p_{j_k'}, w_k \rangle, & \min_{i \in 1:s} \langle p_i, w_k \rangle &= \langle p_{j_k''}, w_k \rangle, \\ \max_{i \in M_2^+(u)} \langle p_i, -w_k \rangle &= \langle p_{l_k'}, -w_k \rangle, & \min_{i \in (s+1):m} \langle p_i, -w_k \rangle &= \langle p_{l_k''}, -w_k \rangle. \end{split}$$

Вычислим

$$\Delta_1(u_k) = \langle p_{j_k'} - p_{j_k''}, w_k \rangle, \quad \Delta_2(u_k) = \langle p_{l_k'} - p_{l_k''}, -w_k \rangle$$
$$\Delta(u_k) = \max\{\Delta_1(u_k), \Delta_2(u_k)\}.$$

Вначале рассмотрим случай, когда $\Delta(u_k) = \Delta_1(u_k)$. Для простоты будем использовать обозначения

$$p'_k := p_{j'_k}, \quad p''_k := p_{j''_k},$$

так что

$$\Delta(u_k) = \langle p_k' - p_k'', w_k \rangle. \tag{12}$$

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

Пусть $\Delta(u_k) > 0$. Запишем $w_k = x_k - y_k$, где $x_k \in C_1$, $y_k \in C_2$. Введем вектор

$$\widehat{x}_k = x_k - u_k' \big(p_k' - p_k'' \big),$$

где $u_k' = u_k[j_k']$ (коэффициент при p_k' передали вектору p_k''). Очевидно, что $\widehat{x}_k \in C_1$. Рассмотрим отрезок

$$x_k(t) = x_k + t(\hat{x}_k - x_k) = x_k - tu'_k(p'_k - p''_k), \quad t \in [0, 1].$$

В силу выпуклости множества C_1 все точки $x_k(t)$ при $t\in[0,1]$ принадлежат C_1 . Обозначим

$$w_k(t) = x_k(t) - y_k = w_k - tu'_k (p'_k - p''_k).$$
(13)

При всех $t \in [0,1]$ вектор $w_k(t)$ является планом задачи (1). Выберем $t_k \in [0,1]$ из условия

$$||w_k(t_k)||^2 = \min_{t \in [0,1]} ||w_k(t)||^2$$
.

Положим $w_{k+1} = w_k(t_k)$.

Нетрудно понять, учитывая (13), что при $j \in 1: s$ справедливо представление

$$u_{k+1}[j] = \begin{cases} u_k[j], & \text{если } j \neq j_k' \text{ и } j \neq j_k'', \\ (1 - t_k)u_k[j_k'], & \text{если } j = j_k', \\ u_k[j_k''] + t_k u_k[j_k'], & \text{если } j = j_k''. \end{cases}$$

При этом $u_{k+1}[j] = u_k[j]$, если $j \in (s+1) : m$.

Для t_k можно указать явную формулу. В силу (13) и (12) имеем

$$||w_k(t)||^2 = ||w_k||^2 - 2tu_k'\langle p_k' - p_k'', w_k \rangle + t^2(u_k')^2 ||p_{k'} - p_{k''}||^2 =$$

$$= ||w_k||^2 - 2tu_k'\Delta(u_k) + t^2(u_k')^2 ||p_{k'} - p_{k''}||^2.$$

Абсолютный минимум $\|w_k(t)\|^2$ на \mathbb{R} достигается в точке

$$\widehat{t}_k = \frac{\Delta(u_k)}{u_k' \|p_k' - p_k''\|^2}.$$

Ясно, что $\widehat{t}_k>0,$ поэтому для точки минимума на [0,1] справедлива формула

$$t_k = \min\{1, \widehat{t}_k\}.$$

Теперь рассмотрим случай, когда

$$\Delta(u_k) = \Delta_2(u_k) = \langle p_{l'_k} - p_{l''_k}, -w_k \rangle.$$

Здесь $l_k' \in M_2^+(u_k), \, l_k'' \in (s+1): m.$ Для простоты будем использовать обозначения

$$p'_k := p_{l'_k}, \quad p''_k := p_{l''_k},$$

так что

$$\Delta(u_k) = \langle p_k' - p_k'', -w_k \rangle. \tag{14}$$

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

Пусть $\Delta(u_k) > 0$. Запишем $w_k = x_k - y_k$, где $x_k \in C_1$, $y_k \in C_2$. Введем вектор

$$\widehat{y}_k = y_k - u_k' (p_k' - p_k''),$$

где $u_k' = u_k[l_k']$ (коэффициент от p_k' передали p_k''). Очевидно, что $\widehat{y}_k \in C_2$. Рассмотрим отрезок

$$y_k(t) = y_k + t(\hat{y}_k - y_k) = y_k - tu'_k(p'_k - p''_k), \quad t \in [0, 1].$$

В силу выпуклости C_2 все точки $y_k(t)$ при $t \in [0,1]$ принадлежат этому множеству. Обозначим

$$w_k(t) = x_k - y_k(t) = w_k + tu'_k (p'_k - p''_k).$$
(15)

При всех $t \in [0,1]$ вектор $w_k(t)$ является планом задачи (1). Выберем $t_k \in [0,1]$ из условия

$$||w_k(t_k)||^2 = \min_{t \in [0,1]} ||w_k(t)||^2$$
.

Положим $w_{k+1} = w_k(t_k)$.

Нетрудно понять, опираясь на (15), что при $l \in (s+1)$: m справедливо представление

$$u_{k+1}[l] = \begin{cases} u_k[l], & \text{если } l \neq l_k' \text{ и } l \neq l_k'', \\ (1 - t_k) u_k[l_k'], & \text{если } l = l_k', \\ u_k[l_k''] + t_k u_k[l_k'], & \text{если } l = l_k''. \end{cases}$$

При этом $u_{k+1}[l] = u_k[l]$, если $l \in 1:s$.

Для t_k можно указать явную формулу. В силу (15) и (14) имеем

$$||w_k(t)||^2 = ||w_k||^2 - 2tu_k'\langle p_k' - p_k'', -w_k \rangle + t^2(u_k')^2 ||p_k' - p_k''||^2 =$$

$$= ||w_k||^2 - 2tu_k'\Delta(u_k) + t^2(u_k')^2 ||p_k' - p_k''||^2.$$

Абсолютный минимум $\|w_k(t)\|^2$ на \mathbb{R} достигается в точке

$$\widehat{t}_k = \frac{\Delta(u_k)}{u_k' \|p_k' - p_k''\|^2}.$$

Ясно, что $\widehat{t}_k>0,$ поэтому для точки минимума на [0,1] справедлива формула

$$t_k = \min\{1, \hat{t}_k\}.$$

Описание обобщенного МДМ-метода завершено.

Замечание. Обозначим $\Delta_k = \Delta(u_k)$. В обоих рассмотренных случаях справедливо равенство

$$||w_k(t)||^2 = ||w_k||^2 - 2tu_k'\Delta_k + t^2(u_k')^2||p_k' - p_k''||^2.$$
(16)

Одинаковы формулы и для \hat{t}_k , и для t_k :

$$\widehat{t}_k = \frac{\Delta_k}{u_k' ||p_k' - p_k''||^2}, \quad t_k = \min\{1, \widehat{t}_k\}.$$
(17)

Если $t_k=1$, то итерацию с номером k будем называть yсеченной, если $t_k=\widehat{t}_k<1$, то hеyсеченной.

 ${\bf 5}^{\circ}$. Рассмотрим неусеченную итерацию. Обозначим через D_1 диаметр множества P_1 , через D_2 — диаметр множества P_2 , и пусть $D=\max\{D_1,D_2\}$.

 Π EMMA 3. На неусеченной k-й итерации выполняется неравенство

$$\|w_k\|^2 - \|w_{k+1}\|^2 \geqslant \frac{\Delta_k^2}{D^2}$$
. (18)

Действительно, согласно (16) и (17) имеем

$$||w_{k+1}||^2 = ||w_k||^2 - 2\widehat{t}_k u_k' \Delta_k + \widehat{t}_k^2 (u_k')^2 ||p_k' - p_k''||^2 =$$

$$= ||w_k||^2 - 2\widehat{t}_k u_k' \Delta_k + \widehat{t}_k u_k' \Delta_k = ||w_k||^2 - \frac{\Delta_k^2}{||p_k' - p_k''||^2}.$$

Отсюда очевидным образом следует (18).

 6° . Переходим к доказательству сходимости обобщенного МДМ-метода [5]. По описанным выше правилам строится последовательность планов w_0, w_1, \ldots задачи (1), (3). Эта последовательность конечна, если при некотором k_0 выполнится равенство $\Delta(u_{k_0}) = 0$. По теореме 1 план w_{k_0} будет оптимальным.

Предположим, что

$$\Delta(u_k) > 0$$
 при всех $k = 0, 1, \dots$ (19)

Условие (19) и формула (16) гарантируют, что бесконечная последовательность $\{\|w_k\|^2\}$ является строго убывающей.

ЛЕММА 4. Пусть выполняется условие (19). Тогда существует такое натуральное N, что количество подряд идущих усеченных итераций не превосходит N.

Доказательство. Усеченная k-я итерация характеризуется тем, что $t_k=1$. Как следствие, на множестве 1:s или (s+1):m справедливо представление

$$u_{k+1}[i] = \begin{cases} u_k[i], & \text{если } i \neq i'_k \text{ и } i \neq i''_k, \\ 0, & \text{если } i = i'_k, \\ u_k[i''_k] + u_k[i'_k], & \text{если } i = i''_k. \end{cases}$$
(20)

Предположим, что, начиная с r-й итерации, подряд идут усеченные итерации. Компоненты вектора u_{k+1} , определяющего w_{k+1} , при $k \geqslant r$ получаются путем перераспределения коэффициентов $u_r[i], i \in 1:m$, по формуле (20). Поэтому в цепочке последовательных усеченных итераций может присутствовать лишь конечное число различных векторов u_k . В силу условия (19) последовательность $\{\|w_k\|\}$ строго убывает. Значит, количество подряд идущих усеченных итераций ограничено сверху некоторым числом N, зависящим только от m и s.

TEOPEMA 2. При выполнении условия (19) справедливо предельное соотношение

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

Доказательство. Последовательность $\{\|w_k\|^2\}$ строго убывает и ограничена снизу нулем. Значит, она сходится. Как следствие,

$$\|w_k\|^2 - \|w_{k+1}\|^2 \to 0 \quad \text{при } k \to \infty.$$
 (21)

По лемме 4 существует бесконечная последовательность неусеченных итераций с номерами k_i . Согласно лемме 3 имеем

$$\|w_{k_i}\|^2 - \|w_{k_i+1}\|^2 \geqslant \frac{\Delta_{k_i}^2}{D^2}.$$

С учетом (21) заключаем, что $\Delta_{k_i} \to 0$ при $i \to \infty$. По лемме 2

$$\|w_{k_i} - w^*\|^2 \leqslant 2\Delta_{k_i},$$

так что $w_{k_i} \to w^*$ при $i \to \infty$. В частности, $||w_{k_i}|| \to ||w^*||$. Так как вся последовательность $\{||w_k||\}$ строго убывает, то $||w_k|| \to ||w^*||$ при $k \to \infty$. Остается сослаться на неравенство (7).

Теорема доказана.

ЛИТЕРАТУРА

- 1. Малоземов В. Н. MДM-методу 50 лет // Семинар «CNSA & NDO». Избранные доклады. 10 ноября 2021 г. (http://apmath.spbu.ru/cnsa/reps21.shtml#1110)
- Barbero A., Lopez J., Dorronsoro J. R. An Accelerated MDM Algorithm for SVM Training // European Symposium on Artificial Neural Networks

 Advances in Computational Intelligence and Learning. Bruges (Belgium).
 23–25 April 2008. Proceedings, pp. 421–426.
- 3. Ming Zeng, Yu Yang, Junsheng Cheng. A generalized Mitchell-Dem'yanov Malozemov algorithm for one-class support vector machine // Knowledge-Based Systems. 109 (2016), pp. 17–24.
- 4. Малоземов В. Н., Соловьева Н. А. *Общая квадратичная задача математи ческой диагностики* // Семинар «О & ML». Избранные доклады. 11 мая 2022 г. (http://apmath.spbu.ru/oml/reps22.shtml#0511)
- Lopez J., Dorronsoro J. R. A Common Framework for the Convergence of the GSK, MDM and SMO Algorithms. Lecturer Notes in Computer Science: Artificial Neural Networks, ICANN 2010. Vol. 6353. 2010. Pp. 82–87.