ЛИНЕЙНАЯ СКОРОСТЬ СХОДИМОСТИ ОБОБЩЕННОГО MDM-МЕТОДА*

А. В. Плоткин
avplotkin@gmail.com

31 октября 2026 г.

В докладе предлагается обобщение MDM-метода для нахождения ближайшей к нулю точки выпуклой оболочки конечного множества точек. На каждой итерации выделяются две группы исходных точек. Для точек первой группы скалярные произведения с текущим приближением близки к максимуму, а для точек второй — к минимуму. Направление задается разностью произвольных точек выпуклых оболочек этих групп.

Доказана линейная сходимость представленного метода. В основе доказательства лежит оценка убывания целевой функции на блоке последовательных итераций. Получена верхняя оценка числа итераций, достаточного для достижения заданной точности. Приведены примеры на плоскости, демонстрирующие, как полученная свобода выбора направления позволяет избавиться от зигзагообразного поведения классического метода.

1°. Пусть в пространстве $\mathbb{R}^n$ заданы m точек,

$$H = \{a_i\}_{i=1}^m.$$

Выпуклую оболочку множества H обозначим G. Ставится экстремальная задача

$$\|v\|^2 \to \min_{v \in G}. \tag{1}$$

Очевидно, что задача (1) имеет единственное решение v_*.

Обозначим через A матрицу со столбцами $a_1, \ldots, a_m$. Тогда любой вектор v из выпуклой оболочки G множества H допускает представление $v = Ap$, где

$$p \geqslant \mathbb{O}, \quad \sum_{i=1}^m p[i] = 1. \tag{2}$$

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

Множество векторов p, удовлетворяющих условию (2), обозначим P. Введем носитель вектора коэффициентов p — индексное множество $M_+(p)$ вида

$$M_+(p) = \{i \in 1:m \mid p[i] > 0\}.$$

Отметим, что $M_+(p) \neq \varnothing$ для любого $p \in P$.

Величина

$$\Delta(p) = \max_{i \in M_+(p)} \langle a_i, v\rangle - \min_{i \in 1:m} \langle a_i, v\rangle,$$

где $v = Ap$, $p \in P$, называется *оценкой плана* v. Оценка плана всегда неотрицательна. Действительно,

$$\Delta(p) \geqslant \max_{i \in M_+(p)} \langle a_i, v\rangle - \min_{i \in M_+(p)} \langle a_i, v\rangle \geqslant 0.$$

Из результатов доклада [1] следуют приведенные ниже утверждения.

УТВЕРЖДЕНИЕ 1. *Для $p \in P$ равенство $\Delta(p) = 0$ выполняется тогда и только тогда, когда вектор $v = Ap$ является решением задачи* (1).

УТВЕРЖДЕНИЕ 2. *При любом $v = Ap$, $p \in P$, справедливо неравенство*

$$\|v - v_*\|^2 \leqslant \|v\|^2 - \|v_*\|^2. \tag{3}$$

УТВЕРЖДЕНИЕ 3. *При любом $v = Ap$, $p \in P$, справедливо неравенство*

$$\|v\|^2 - \|v_*\|^2 \leqslant 2\Delta(p). \tag{4}$$

В докладе [2] удалось усилить оценку (4) при малых значениях $\Delta(p)$. Приведем ее здесь без доказательства.

УТВЕРЖДЕНИЕ 4. *При любом $v = Ap$, $p \in P$, справедливо неравенство*

$$\|v\|^2 - \|v_*\|^2 \leqslant C\Delta(p)^2, \tag{5}$$

где C — положительная константа, зависящая только от множества H.

2°. Введем следующее обобщение MDM-метода для решения задачи (1). Зафиксируем параметры α и β, удовлетворяющие условиям

$$\alpha \geqslant 0, \quad \beta \geqslant 0, \quad \alpha + \beta < 1,$$

и обозначим $\gamma := 1 - \alpha - \beta > 0$.

Возьмем начальное приближение $v_0 = Ap_0$, $p_0 \in P$.

Пусть уже имеется k-е приближение $v_k = Ap_k$, $p_k \in P$. Опишем k-ю итерацию алгоритма, на которой происходит построение v_{k+1}.

Обозначим

$$x_k = \max_{i \in M_+(p_k)} \langle a_i, v_k \rangle, \qquad y_k = \min_{i \in 1:m} \langle a_i, v_k \rangle.$$

Тогда оценка плана $v_k = Ap_k$ может быть записана в виде

$$\Delta_k := \Delta(p_k) = x_k - y_k.$$

Если $\Delta_k = 0$, то, согласно утверждению 1, вектор v_k является решением задачи (1). Процесс закончен.

Пусть $\Delta_k > 0$. Введем множества *околоэкстремальных* индексов

$$I_k^{\max} = \big\{ i \in M_+(p_k) \mid \langle a_i, v_k \rangle \geqslant x_k - \alpha\Delta_k \big\},$$
$$I_k^{\min} = \big\{ i \in 1:m \quad \mid \langle a_i, v_k \rangle \leqslant y_k + \beta\Delta_k \big\}.$$

Оба множества непусты и не пересекаются.

Рассмотрим выпуклые оболочки точек с индексами из этих множеств:

$$G_k^{\max} = \operatorname{conv}\{a_i \mid i \in I_k^{\max}\}, \qquad G_k^{\min} = \operatorname{conv}\{a_j \mid j \in I_k^{\min}\}.$$

Произвольно выберем

$$b_k^{\max} \in G_k^{\max}, \qquad b_k^{\min} \in G_k^{\min}$$

и зафиксируем их выпуклые представления

$$u_k \in P, \qquad M_+(u_k) \subset I_k^{\max}, \qquad b_k^{\max} = Au_k;$$
$$w_k \in P, \qquad M_+(w_k) \subset I_k^{\min}, \qquad b_k^{\min} = Aw_k.$$

При построении вектора p_{k+1} из p_k элементы с индексами из $M_+(u_k)$ будут уменьшены, а элементы с индексами из $M_+(w_k)$ — увеличены. Вычислим максимальный *вес*, который мы можем перенести при выбранных представлениях u_k, w_k:

$$\rho_k := \min_{i \in M_+(u_k)} \frac{p_k[i]}{u_k[i]} > 0.$$

Введем направление (см. рис. 1)

$$d_k = b_k^{\max} - b_k^{\min} = A(u_k - w_k).$$

Это направление ненулевое, и для него справедлива оценка

$$\gamma\Delta_k \leqslant \langle d_k, v_k \rangle \leqslant \Delta_k. \tag{6}$$

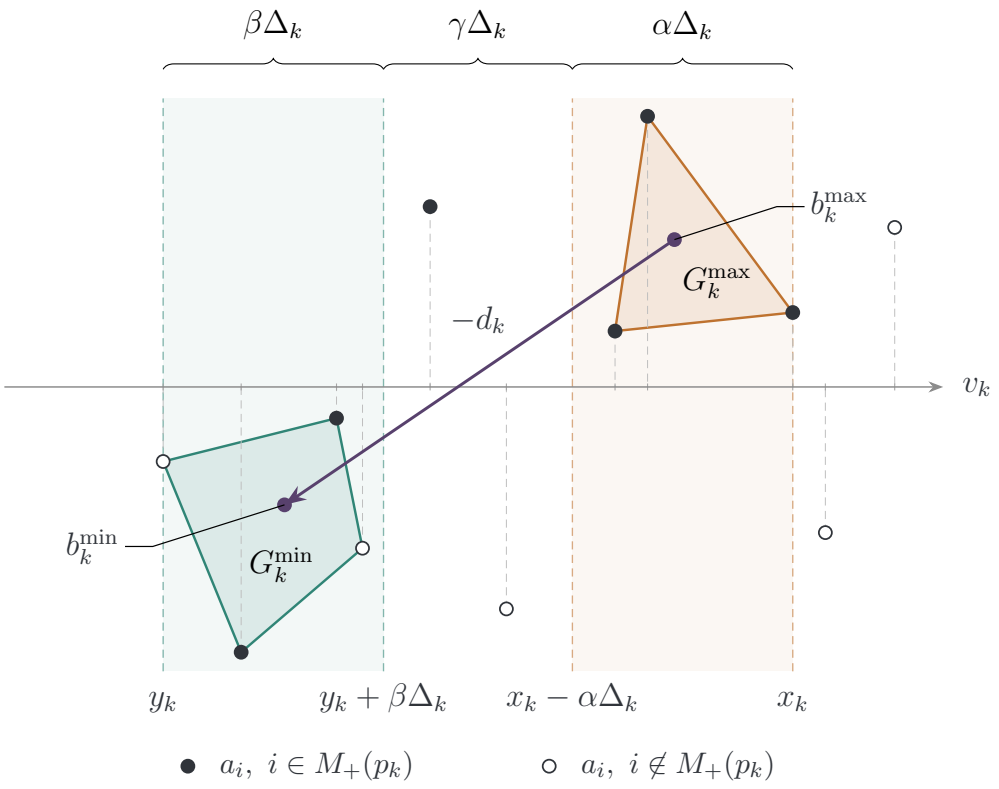

Рис. 1. Геометрия выбора направления.
Отметки на оси соответствуют значениям $\langle v_k, a_i \rangle$.

Рассмотрим вектор

$$\hat{v}_k = v_k - \rho_k d_k.$$

Определение ρ_k гарантирует, что $\hat{v}_k \in G$. Рассмотрим отрезок

$$v_k(t) = v_k + t(\hat{v}_k - v_k) = v_k - t\rho_k d_k, \quad t \in [0, 1].$$

В силу выпуклости множества G все точки $v_k(t)$ этого отрезка принадлежат G. Выберем $t_k \in [0, 1]$ из условия

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

Положим $v_{k+1} = v_k(t_k)$.

Коэффициенты разложения вектора v_{k+1} по столбцам матрицы A выглядят так:

$$p_{k+1}[i] = \begin{cases} p_k[i] & \text{при } i \notin M_+(u_k) \cup M_+(w_k), \\ p_k[i] - t_k\rho_k u_k[i] & \text{при } i \in M_+(u_k), \\ p_k[i] + t_k\rho_k w_k[i] & \text{при } i \in M_+(w_k). \end{cases}$$

Укажем явную формулу для t_k. Абсолютный минимум квадратичной функции одной переменной $\|v_k(t)\|^2$ на $\mathbb{R}$ достигается в точке

$$\hat{t}_k = \frac{\langle d_k, v_k \rangle}{\rho_k \|d_k\|^2} > 0.$$

Значит,

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

Если $\hat{t}_k \leqslant 1$, то $t_k = \hat{t}_k$; если $\hat{t}_k > 1$, то $t_k = 1$. В первом случае k-я итерация называется *неусеченной*, во втором — *усеченной*. Для неусеченной итерации имеем

$$v_{k+1} = v_k(\hat{t}_k) = v_k - \frac{\langle d_k, v_k \rangle}{\|d_k\|^2} d_k,$$

а для усеченной

$$v_{k+1} = v_k(1) = \hat{v}_k.$$

Описание обобщенного MDM-метода завершено.

Представленный алгоритм строит последовательность $v_0, v_1, \ldots$ точек из G. Если она конечна, то последний ее элемент является решением задачи (1). Вообще говоря, последовательность $\{v_k\}$ бесконечна. Такая ситуация возникает, когда

$$\Delta_k > 0 \quad \text{при всех} \quad k = 0, 1, \ldots$$

Далее будем рассматривать случай бесконечной последовательности $\{v_k\}$.

3°. Отметим важное свойство монотонности алгоритма. Функция $\|v_k(t)\|^2$ строго выпукла, а $t_k > 0$ является точкой ее минимума на $[0, 1]$, поэтому

$$\|v_{k+1}\|^2 = \|v_k(t_k)\|^2 < \|v_k(0)\|^2 = \|v_k\|^2 .$$

То есть последовательность $\left\{\|v_k\|^2\right\}$ является строго убывающей.

Введем обозначение $Q_k = \|v_k\|^2 - \|v_*\|^2$ для оценки скорости сходимости метода. Последовательность $\{Q_k\}$ также строго убывает. Из неравенства (3) для всех k следует

$$\|v_k - v_*\|^2 \leqslant Q_k.$$

Через D обозначим диаметр множества G, а через $\tau_k = t_k\rho_k$ — вес, переносимый на k-й итерации. Оценим уменьшение Q_k за одну итерацию.

ЛЕММА 1. *На каждой итерации k справедлива оценка*

$$Q_k - Q_{k+1} \geqslant \gamma\tau_k\Delta_k. \tag{7}$$

На каждой неусеченной итерации k дополнительно справедливо неравенство

$$Q_k - Q_{k+1} \geqslant \frac{\gamma^2\Delta_k^2}{D^2}. \tag{8}$$

Доказательство. Из определения t_k получим

$$\tau_k = \min\left\{\rho_k, \frac{\langle d_k, v_k\rangle}{\|d_k\|^2}\right\} \leqslant \frac{\langle d_k, v_k\rangle}{\|d_k\|^2}.$$

Тогда по построению v_{k+1} и полученному неравенству для τ_k имеем

$$\begin{gathered} Q_k - Q_{k+1} = \|v_k\|^2 - \|v_{k+1}\|^2 = 2\tau_k\langle v_k, d_k\rangle - \tau_k^2\|d_k\|^2 = \\ = \tau_k\big(2\langle d_k, v_k\rangle - \tau_k\|d_k\|^2\big) \geqslant \tau_k\langle d_k, v_k\rangle \geqslant \gamma\tau_k\Delta_k. \end{gathered}$$

Получили неравенство (7). Для неусеченных итераций имеем $\tau_k = \frac{\langle d_k, v_k\rangle}{\|d_k\|^2}$. Подставляя в неравенство (7), получим

$$Q_k - Q_{k+1} \geqslant \gamma\tau_k\Delta_k \geqslant \frac{\gamma^2\Delta_k^2}{\|d_k\|^2} \geqslant \frac{\gamma^2\Delta_k^2}{D^2}.$$

Для последнего перехода мы пользуемся фактом того, что d_k является разностью двух точек из G. Неравенство (8) установлено. □

Теперь докажем ключевую оценку убывания Q_k на непрерывном блоке из m итераций.

ТЕОРЕМА 1. *Рассмотрим блок из m последовательных итераций с номерами $s, \ldots, s+m-1$. Найдется номер $k_0 \in s : s+m-1$, для которого выполняется*

$$Q_s - Q_{s+m} \geqslant \frac{\gamma^2\Delta_{k_0}^2}{D^2}. \tag{9}$$

Доказательство. Если в таком блоке имеется неусеченная итерация, то в качестве k_0 достаточно взять ее номер из-за монотонности Q_k и оценки (8).

Далее будем считать, что все итерации блока являются усеченными. На каждой усеченной итерации как минимум один индекс покидает носитель.

Носитель не может стать пустым, поэтому для рассматриваемого блока итераций размера m найдется индекс $i_0 \in 1:m$, который покидает носитель на итерации l и возвращается в носитель на итерации r. То есть

$$i_0 \in I_l^{\max} \cap I_r^{\min}, \qquad s \leqslant l < r \leqslant s+m-1.$$

Оценим снизу суммарный вес, перенесенный на итерациях l, …, $r-1$. Введем величины

$$\widehat{h}_k = x_k - \langle v_k, a_{i_0}\rangle, \qquad \check{h}_k = \langle v_k, a_{i_0}\rangle - y_k, \qquad k \in l:r.$$

На концах участка $k = l, r$ проекция $\langle v_k, a_{i_0}\rangle$ лежит в отрезке $[y_k, x_k]$, поэтому $\widehat{h}_k$ и $\check{h}_k$ равны расстояниям до его верхней и нижней границ соответственно. На промежуточных итерациях $\widehat{h}_k$ может быть отрицательным. При этом для всех $k \in l:r$ справедливо равенство

$$\Delta_k = \widehat{h}_k + \check{h}_k.$$

Оценим изменения $\widehat{h}_k$ и $\check{h}_k$. Зафиксируем $k \in l : r-1$. Для каждого $j \in 1:m$ имеем

$$\langle v_{k+1} - v_k, a_j - a_{i_0}\rangle = -\tau_k \langle d_k, a_j - a_{i_0}\rangle \leqslant D^2\tau_k.$$

Новые индексы носителя при прежнем v_k имели проекции не выше x_k, поскольку они находились в $I_k^{\min}$. Поэтому, переходя к максимуму по новому носителю и минимуму по всем индексам, получаем

$$\begin{aligned} \widehat{h}_{k+1} - \widehat{h}_k &\leqslant D^2\tau_k, \\ \check{h}_k - \check{h}_{k+1} &\leqslant D^2\tau_k. \end{aligned}$$

Обозначим суммарный перенесенный вес через

$$T = \sum_{k=l}^{r-1} \tau_k.$$

Суммируя полученные неравенства по $k = l, \ldots, r-1$, имеем

$$\begin{aligned} \widehat{h}_r - \widehat{h}_l &\leqslant D^2 T, \\ \check{h}_l - \check{h}_r &\leqslant D^2 T. \end{aligned}$$

Заметим, что обе разности получаются вычитанием одной и той же суммы из соответствующей длины отрезка:

$$\begin{aligned} \widehat{h}_r - \widehat{h}_l &= \Delta_r - (\widehat{h}_l + \check{h}_r), \\ \check{h}_l - \check{h}_r &= \Delta_l - (\widehat{h}_l + \check{h}_r). \end{aligned}$$

Оценим сумму $\widehat{h}_l + \check{h}_r$, входящую в оба равенства. Из выбора i_0 следует

$$\widehat{h}_l \leqslant \alpha\Delta_l, \qquad \check{h}_r \leqslant \beta\Delta_r.$$

Тогда

$$\widehat{h}_l + \check{h}_r \leqslant \alpha\Delta_l + \beta\Delta_r \leqslant (\alpha+\beta)\max\{\Delta_l, \Delta_r\}.$$

Поэтому

$$D^2 T \geqslant \max\{\widehat{h}_r - \widehat{h}_l, \check{h}_l - \check{h}_r\} = \max\{\Delta_l, \Delta_r\} - (\widehat{h}_l + \check{h}_r) \geqslant \gamma\max\{\Delta_l, \Delta_r\}.$$

Получили оценку снизу на суммарный перенесенный вес:

$$T \geqslant \frac{\gamma\max\{\Delta_l, \Delta_r\}}{D^2}. \tag{10}$$

Среди итераций с номерами $l, \ldots, r-1$ выберем итерацию k_0 с минимальной оценкой плана Δ_{k_0}. Применяя лемму 1 и оценку (10), получим

$$Q_s - Q_{s+m} \geqslant Q_l - Q_r = \sum_{k=l}^{r-1}\bigl(Q_k - Q_{k+1}\bigr) \geqslant \gamma\sum_{k=l}^{r-1}\tau_k\Delta_k \geqslant$$
$$\geqslant \gamma\Delta_{k_0}T \geqslant \frac{\gamma^2\Delta_{k_0}\max\{\Delta_l, \Delta_r\}}{D^2} \geqslant \frac{\gamma^2\Delta_{k_0}^2}{D^2}.$$

Теорема доказана. □

СЛЕДСТВИЕ 1. *Для любого номера $s \geqslant 0$ справедливо*

$$Q_{s+m} \leqslant \sigma Q_s, \tag{11}$$

где $\sigma \in (0,1)$ — константа, зависящая только от множества H и параметра γ.

Доказательство. Из утверждения 4, теоремы 1 и монотонности $\{Q_k\}$ следует

$$Q_s - Q_{s+m} \geqslant \frac{\gamma^2\Delta_{k_0}^2}{D^2} \geqslant \frac{\gamma^2 Q_{k_0}}{CD^2} \geqslant \frac{\gamma^2 Q_{s+m}}{CD^2}.$$

Отсюда получим требуемое неравенство (11) с константой

$$\sigma = \frac{CD^2}{\gamma^2 + CD^2} \in (0,1).$$

□

Докажем линейную скорость сходимости обобщенного MDM-метода.

СЛЕДСТВИЕ 2. *Бесконечная последовательность приближений $\{v_k\}$, построенная обобщенным MDM-методом, сходится к решению v_* с линейной скоростью.*

Доказательство. Применяя следствие 1 последовательно к блокам из m итераций и используя монотонность последовательности $\{Q_k\}$, получим

$$\|v_k - v_*\|^2 \leqslant Q_k \leqslant \sigma^{\lfloor \frac{k}{m} \rfloor} Q_0 \leqslant (\sigma^{\frac{1}{m}})^k \frac{Q_0}{\sigma}.$$

Извлекая корень и полагая $B = \sqrt{Q_0/\sigma}$, $q = \sigma^{\frac{1}{2m}} \in (0,1)$, приходим к оценке

$$\|v_k - v_*\| \leqslant Bq^k,$$

которая свидетельствует о линейной скорости сходимости. □

Оценим число итераций, достаточное для достижения заданной точности.

СЛЕДСТВИЕ 3. *Пусть $0 < \varepsilon < \sqrt{Q_0}$. Для достижения неравенства $\|v_k - v_*\| \leqslant \varepsilon$ требуется не более*

$$N(\varepsilon) = m \left\lceil \frac{\ln(Q_0/\varepsilon^2)}{\ln(1/\sigma)} \right\rceil \tag{12}$$

итераций обобщенного MDM-метода.

Доказательство. Пусть b — число блоков из m итераций. После mb итераций по следствию 1 имеем

$$\|v_{mb} - v_*\|^2 \leqslant Q_{mb} \leqslant \sigma^b Q_0.$$

Для достижения точности ε достаточно выполнения условия

$$\sigma^b Q_0 \leqslant \varepsilon^2 \quad \Longleftrightarrow \quad (1/\sigma)^b \geqslant Q_0/\varepsilon^2.$$

Поскольку $\ln(1/\sigma) > 0$, это равносильно неравенству

$$b \geqslant \frac{\ln(Q_0/\varepsilon^2)}{\ln(1/\sigma)}.$$

Выбирая наименьшее целое b, удовлетворяющее этому условию, и умножая его на m, получим оценку (12). □

4°. В представленном обобщении метода параметры α, β задают множество допустимых направлений, а правило выбора определяет, какое из них будет использовано на каждой итерации. Классический MDM-метод получается при $\alpha = \beta = 0$ и выборе $b_k^{\max} = a_i$, $i \in I_k^{\max}$, $b_k^{\min} = a_j$, $j \in I_k^{\min}$.

Известен эффект зигзагообразного поведения классического метода. Покажем, что полученная свобода выбора направления может помочь сгладить такое поведение.

Рассмотрим пример на плоскости (см. рис. 2): треугольник с вершинами

$$a_1 = (0, 4), \qquad a_2 = (10, 2), \qquad a_3 = (-10, 2).$$

Начинать вычисления будем с $v_0 = a_1$.

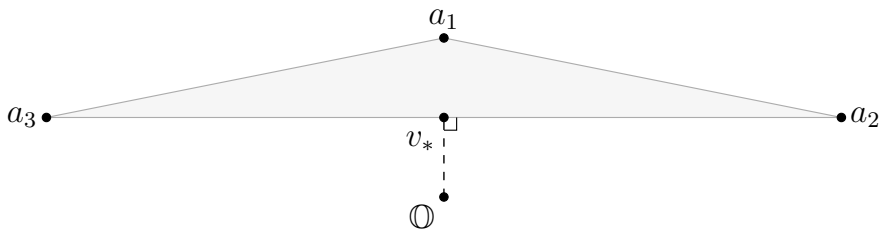

Рис. 2. Исходный треугольник.

Классический MDM-метод, выбирающий при равенстве экстремальных значений наименьший индекс, приходит к решению $v_* = (0, 2)$ за 11 итераций, двигаясь зигзагами.

Для обобщенного метода зафиксируем следующее правило выбора направления. Точки $b_k^{\max}$ и $b_k^{\min}$ будем выбирать как ближайшую пару точек множеств $G_k^{\max}$ и $G_k^{\min}$ соответственно. Тогда при любых допустимых α, β обобщенный метод достигает решения за один шаг (см. рис. 3).

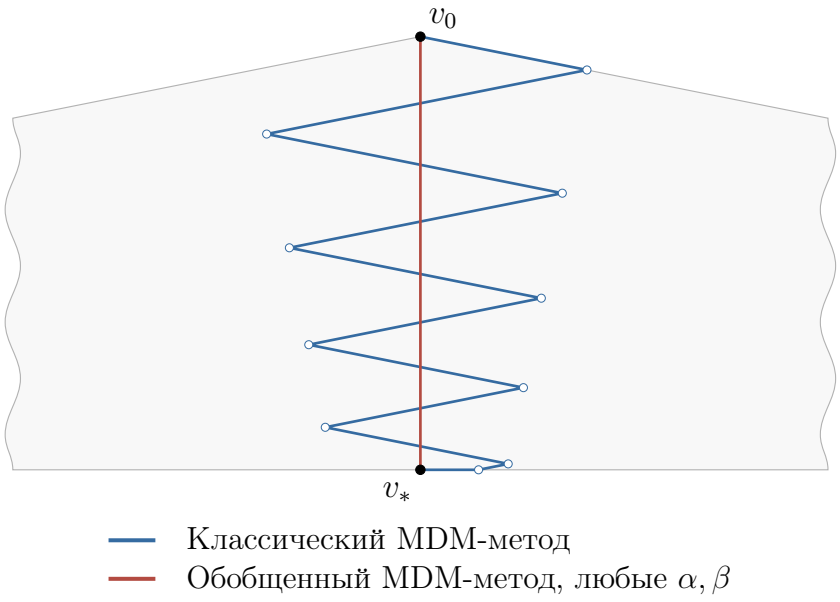

Рис. 3. Траектории методов для исходного треугольника.

Таким образом, только за счет изменения правила выбора направления удалось избежать зигзагообразного поведения. Однако в общем случае не менее важен выбор параметров α и β, позволяющий расширить множество допустимых направлений. Покажем это на примере.

Сдвинем все три вершины треугольника вправо на 0.1 и снова начнем с верхней вершины. Решение по-прежнему равно $v_* = (0, 2)$. Сравним два варианта обобщенного метода с тем же правилом выбора направления. При $\alpha = \beta = 0$ получается зигзагообразная траектория классического MDM-метода, и решение достигается за 11 итераций. При $\alpha = \beta = 1/3$ множество $G_0^{\min}$ совпадает с основанием, и обобщенный метод достигает решения за два шага:

$$(0.1, 4) \longrightarrow (0.1, 2) \longrightarrow (0, 2).$$

Здесь расширение околоэкстремальных индексных множеств позволяет выбирать дополнительные направления, вновь делая возможным движение к середине основания треугольника (см. рис. 4).

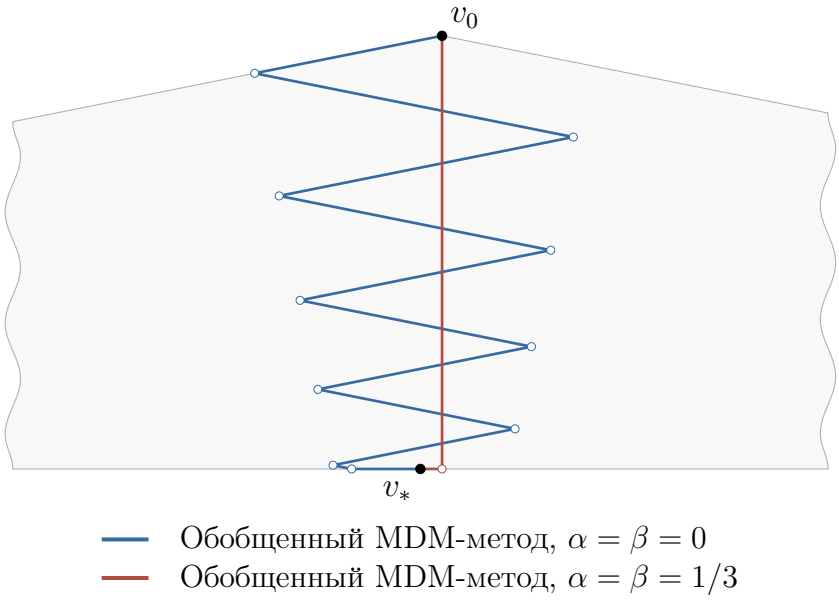

Рис. 4. Траектории методов после сдвига треугольника.

Обобщенный MDM-метод позволяет расширять множества точек, участвующих в перераспределении веса, и выбирать способ этого перераспределения. При этом линейная скорость сходимости гарантируется для любого допустимого правила выбора направления.

ЛИТЕРАТУРА

1. Малоземов В. Н. *MDM-методу — 50 лет* // Семинар «OML». Литература. Публикации участников семинара. (`http://oml.cmlaboratory.com/literatura.shtml`)

2. Плоткин А. В., Соловьева Н. А. *Линейная скорость сходимости MDM-метода* // Семинар «OML». Избранные доклады. 17 октября 2026 г. (`http://oml.cmlaboratory.com/reps26.shtml#1017`)