

СЕМИНАР ПО ОПТИМИЗАЦИИ,
МАШИННОМУ ОБУЧЕНИЮ И ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ
(<http://oml.cmlaboratory.com/>)

ТАМАСЯН Григорий Шаликович

ОБОБЩЕНИЕ МДМ–МЕТОДА

г. Санкт-Петербург
2 октября 2025 г.

ПОСТАНОВКА ЗАДАЧИ

В пространстве $\mathbb{R}^n$ заданы $s \geq 2$ конечных множеств

$$P_k = \{A_{k1}, \dots, A_{kd_k}\}, \quad k \in 1 : s,$$

где $A_{kj} \in \mathbb{R}^n$ — элементы соответствующих множеств, $d_k \geq 1$.

Обозначим через C_k выпуклую оболочку множества P_k , $k \in 1 : s$.
Рассмотрим выпуклое множество C равное сумме по Минковскому множеств C_k , $k \in 1 : s$.

Требуется решить экстремальную задачу

$$\frac{1}{2} \|v\|^2 \rightarrow \min_{v \in C}. \quad (1)$$

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

БАЗОВЫЕ СВЕДЕНИЯ НЕГЛАДКОГО АНАЛИЗА

Определение 1. Говорят, что функция f , заданная и конечная на U , гиподифференцируема в точке $x \in U$, если существует такой выпуклый компакт $df(x) \subset \mathbb{R}^n \times \mathbb{R}$, что справедливы разложение в точке x вида

$$f(x + \Delta) = f(x) + \max_{(v, \alpha) \in df(x)} [v^\top \Delta + \alpha] + o_x(\Delta),$$

и равенство

$$\max_{(v, \alpha) \in df(x)} \alpha = 0.$$

Здесь $\alpha \in \mathbb{R}$, $v \in \mathbb{R}^n$, $\frac{o_x(\gamma\Delta)}{\gamma} \rightarrow 0$ при $\gamma \rightarrow 0$ для всех $\Delta \in \mathbb{R}^n$.

Множество $df(x)$ называется гиподифференциалом функции f в точке x . Элементы множества $df(x)$ называются *гипоградиентами**.

Классы субдифференцируемых и гиподифференцируемых функций совпадают. Более того, справедливо соотношение

$$\partial f(x) = \left\{ v \in \mathbb{R}^n \mid \begin{pmatrix} v \\ 0 \end{pmatrix} \in df(x) \right\}.$$

* Демьянов В. Ф., Рубинов А. М. Основы негладкого анализа и квазидифференциальное исчисление. М.: Наука, 1990. 432 с.

ИСЧИСЛЕНИЕ ГИПОДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ

Пусть функции $f_i(x)$, $i \in 1 : s$ гиподифференцируемы в точке x .

1. $d(f_1 + f_2)(x) = df_1(x) + df_2(x)$.
2. $d(\lambda \cdot f_1)(x) = \lambda \cdot df_1(x)$ для всех $\lambda > 0$.
3. Тогда функция $f(x) = \max_{i \in 1:s} f_i(x)$ гиподифф. в этой точке:

$$df(x) = \text{conv} \left\{ \begin{pmatrix} \mathbf{0}_n \\ f_i(x) - f(x) \end{pmatrix} + df_i(x) \mid 1 \leq i \leq s \right\}.$$

Здесь над гиподифференциалами определены операции *сложения по Минковскому* и *умножения на константу*, а именно

$$\begin{aligned} W + V &= \{w + v \mid w \in W, v \in V\}, \\ \lambda V &= \{\lambda v \mid v \in V\}, \end{aligned}$$

где W, V — выпуклые компакты в $\mathbb{R}^n$, λ — вещественное число.

ПРИМЕРЫ

Пример 1. Для произвольной непрерывно дифференцируемой на всем пространстве функции f в качестве непрерывного гиподифференциала* можно взять

$$df(x) = \left\{ \begin{pmatrix} f'(x) \\ 0 \end{pmatrix} \right\}.$$

Пример 2. Пусть $h(x)$ непрерывно дифференцируемая функция. Найдем гиподифференциалы функций $f(x) = |h(x)|$ и $\varphi(x) = \max\{0, h(x)\}$. Заметим, что $f(x) = \max\{-h(x), h(x)\}$. Тогда

$$df(x) = \text{conv} \left\{ \begin{pmatrix} -h'(x) \\ -h(x) - f(x) \end{pmatrix}, \begin{pmatrix} h'(x) \\ h(x) - f(x) \end{pmatrix} \right\},$$
$$d\varphi(x) = \text{conv} \left\{ \begin{pmatrix} 0_n \\ -\varphi(x) \end{pmatrix}, \begin{pmatrix} h'(x) \\ h(x) - \varphi(x) \end{pmatrix} \right\}.$$

*Долгополик М. В. О непрерывной кодифференцируемости // Семинар «CNSA & NDO». Избранные доклады. 12 ноября 2019 г.
(<http://cnsa.cmlaboratory.com/rep19.shtml#1119>)

Пример 3. Пусть $f_k(x)$, $k \in 1 : s$, непрерывно дифференцируемые функции. Найдем гиподифференциалы функций

$$\varphi(x) = \max_{k \in 1:s} f_k(x), \quad f(x) = \sum_{k=1}^s |f_k(x)|.$$

Имеем

$$\begin{aligned} d\varphi(x) &= \text{conv} \left\{ \begin{pmatrix} f'_1(x) \\ f_1(x) - \varphi(x) \end{pmatrix}, \dots, \begin{pmatrix} f'_s(x) \\ f_s(x) - \varphi(x) \end{pmatrix} \right\}, \\ df(x) &= \sum_{k=1}^s \text{conv} \left\{ \begin{pmatrix} -f'_k(x) \\ -f_k(x) - |f_k(x)| \end{pmatrix}, \begin{pmatrix} f'_k(x) \\ f_k(x) - |f_k(x)| \end{pmatrix} \right\}. \end{aligned}$$

УСЛОВИЯ ОПТИМАЛЬНОСТИ*

Пусть функция f определена на открытом множестве $U \subset \mathbb{R}^n$. Положим $S = \{g \in \mathbb{R}^n \mid \|g\| = 1\}$.

Определение 2. Говорят, что функция f дифференцируема в точке $x \in U$ по направлению $g \in S$, если существует конечный предел

$$f'(x, g) = \lim_{\alpha \rightarrow +0} \frac{f(x + \alpha g) - f(x)}{\alpha}.$$

Величина $f'(x, g)$ называется производной функции f в точке x по направлению g . Для гладкой функции справедливо $f'(x, g) = \langle f'(x), g \rangle$.

ТЕОРЕМА 1. Для того чтобы точка x_* была точкой минимума функции $f(x)$ на $\mathbb{R}^n$, необходимо, а в случае выпуклости f и достаточно, чтобы выполнялось неравенство

$$f'(x_*, g) \geq 0 \quad \forall g \in S. \quad (2)$$

Точка x_* , для которой выполняется неравенство (2), называется стационарной точкой функции $f(x)$ на $\mathbb{R}^n$.

ТЕОРЕМА 2. Условие стационарности равносильно включению

$$0_{n+1} \in df(x_*).$$

*Малозёмов В. Н., Тамасян Г. Ш. О направлении наискорейшего спуска // Вестн. СПбГУ. Сер. 10. Прикл. матем. Информ. Проц. упр., 15:4 (2019), 489–501.

МЕТОД ГИПОДИФФЕРЕНЦИАЛЬНОГО СПУСКА

Пусть уже имеется $x^{(\ell)} \in \mathbb{R}^n$. Переход от ℓ -го приближения к $(\ell + 1)$ -му осуществляется следующим образом ($\ell \geq 0$):

1. Вычисление направления спуска:

$$(v_\ell, \alpha_\ell) = \arg \min \left\{ \|v\|^2 + \alpha^2 \mid (v, \alpha) \in df(x^{(\ell)}) \right\}.$$

Если $\|(v_\ell, \alpha_\ell)\| = 0$, то $x^{(\ell)}$ — решение задачи. Иначе $\|v_\ell\| > 0$ и переходим к следующему шагу.

2. Линейный поиск. Найдём $\gamma_\ell \geq 0$ такое, что

$$\min_{\gamma \geq 0} f(x^{(\ell)} - \gamma v_\ell) = f(x^{(\ell)} - \gamma_\ell v_\ell).$$

3. Очередное приближение: $x^{(\ell+1)} = x^{(\ell)} - \gamma_\ell v_\ell$.

Нахождение направления гиподифференциального спуска*

Рассмотрим функции $\varphi(x) = \max_{k \in 1:s} f_k(x)$, $f(x) = \sum_{k=1}^s |f_k(x)|$. Имеем

$$d\varphi(x) = \text{conv} \left\{ \begin{pmatrix} f'_1(x) \\ f_1(x) - \varphi(x) \end{pmatrix}, \dots, \begin{pmatrix} f'_s(x) \\ f_s(x) - \varphi(x) \end{pmatrix} \right\}, \quad (3)$$

$$df(x) = \sum_{k=1}^s \text{conv} \left\{ \begin{pmatrix} -f'_k(x) \\ -f_k(x) - |f_k(x)| \end{pmatrix}, \begin{pmatrix} f'_k(x) \\ f_k(x) - |f_k(x)| \end{pmatrix} \right\}. \quad (4)$$

Поиск направления спуска сводится к задаче:

$$\frac{1}{2}\mu^\top G(x)\mu - \mu^\top q(x) \longrightarrow \min_{\mu \in \Omega},$$

где $G(x)$ — матрица Грама системы векторов $\begin{pmatrix} f'_k(x) \\ f_k(x) \end{pmatrix}$, $k = 1 : s$,

$q(x) = (f_1(x), \dots, f_s(x))^\top f(x)$. Множество Ω в задаче (3) является стандартным симплексом, а в задаче (4) — n -мерный параллелепипед

вида $-1 \leq \mu_k \leq 1$, $k = 1 : s$. Направление спуска $g = \sum_{k=1}^s f'_k(x)\mu_k^*$.

*Полякова Л. Н. Непрерывные методы безусловной минимизации гиподифференцируемых функций // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2007. № 3. С. 71–81.

ОБОБЩЕННЫЙ МДМ–МЕТОД*

В пространстве $\mathbb{R}^n$ заданы $s \geq 2$ конечных множеств

$$C_k = \text{conv}\{P_{k1}, P_{k2}, \dots, P_{kd_k}\}, \quad k \in 1 : s.$$

Рассмотрим выпуклое множество C равное сумме по Минковскому множеств C_k , $k \in 1 : s$. Требуется решить экстремальную задачу

$$\frac{1}{2} \|v\|^2 \rightarrow \min_{v \in C}. \quad (5)$$

Далее рассмотрим случай, когда все $d_k = 2^\dagger$, т.е.

$$C_k = \text{conv}\{P_{k1}, P_{k2}\}, \quad k \in 1 : s.$$

***1.** Малоземов В. Н. МДМ-методу — 50 лет // Семинар «CNSA & NDO». Избранные доклады. 10 ноября 2021 г. (<http://cnsa.cmlaboratory.com/rep21.shtml#1110>)

2. Малоземов В. Н., Соловьева Н. А. МДМ-метод для решения общей квадратичной задачи математической диагностики // Семинар «O & ML». Избранные доклады. 1 июня 2022 г. (<http://oml.cmlaboratory.com/rep22.shtml#0601>)

[†]Общий случай рассмотрен в работе: Тамасян Г.Ш. Обобщение МДМ-метода // Семинар «O & ML». Избранные доклады. 2 октября 2025 г. (<http://oml.cmlaboratory.com/rep25.shtml#1002>)

По определению выпуклой оболочки точка $v_k \in C_k$, $k \in 1 : s$, допускает представление

$$v_k = \lambda_{k1} \cdot P_{k1} + \lambda_{k2} \cdot P_{k2}, \quad \lambda_{k1} + \lambda_{k2} = 1, \quad \lambda_{k1} \geq 0, \lambda_{k2} \geq 0.$$

Для $v \in C$ справедливо

$$v = v_1 + \dots + v_s = A\lambda,$$

где

$$A = [P_{11}, P_{12}, \dots, P_{s1}, P_{s2}],$$

$$\lambda^\top = [\lambda_{11}, \lambda_{12}, \dots, \lambda_{s1}, \lambda_{s2}].$$

Положим $\Lambda = \left\{ \lambda \in \mathbb{R}^{2s} \mid \lambda_{k1} + \lambda_{k2} = 1, \quad \lambda_{k1} \geq 0, \lambda_{k2} \geq 0, \quad k \in 1 : s \right\}$.

Задачу (5) можно переписать в следующем виде:

$$\begin{aligned} \frac{1}{2} \|v\|^2 &\rightarrow \min, \\ v &= A\lambda, \quad \lambda \in \Lambda. \end{aligned} \tag{6}$$

$$\Lambda : \quad \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \end{pmatrix} \lambda = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \lambda \geq \mathbf{0}_{2s}.$$

ОСНОВНЫЕ УТВЕРЖДЕНИЯ

Пусть $v = A\lambda$, $\lambda \in \Lambda$. Обозначим

$$M_k^+(\lambda) = \{j \in 1 : 2 \mid \lambda_{kj} > 0\}, \quad k \in 1 : s.$$

Введем величины

$$\Delta_k(\lambda) = \max_{j \in M_k^+(\lambda)} \langle A_{kj}, v \rangle - \min \{ \langle A_{k1}, v \rangle, \langle A_{k2}, v \rangle \}, \quad k \in 1 : s,$$

$$\Delta(\lambda) = \max_{k \in 1:s} \Delta_k(\lambda).$$

Очевидно, что для каждого $k \in 1 : s$ имеем $\Delta_k(\lambda) \geq 0$, а следовательно $\Delta(\lambda) \geq 0 \quad \forall \lambda \in \Lambda$.

Лемма. Возьмем произвольный план $v = A\lambda$ задачи (6) и её оптимальный план $v^* = A\lambda^*$. Справедливы неравенства

$$\|v - v^*\|^2 \leq \sum_{k=1}^s \Delta_k(\lambda) \leq s \cdot \Delta(\lambda). \quad (7)$$

ТЕОРЕМА. Равенство $\Delta(\lambda) = 0$ выполняется тогда и только тогда, когда вектор $v = A\lambda$ является решением задачи (6).

Линейная скорость сходимости*.

*Соловьева Н. А. Линейная скорость сходимости МДМ-метода // Семинар «CNSA & NDO». Избранные доклады. 29 сентября 2016 г. (<http://cnsa.cmlaboratory.com/rep16.shtml#0929>)

Рабочая схема № 1 (экстремального проектирования)

Общий шаг. Пусть уже имеется ℓ -е приближение $v^\ell = A\lambda^\ell$.

1. Для каждого $k \in 1 : s$ находим индексы $j'_k \in M_k^+(\lambda^\ell)$, $j''_k \in 1 : 2$, такие, что

$$\max_{j \in M_k^+(\lambda^\ell)} \langle A_{kj}, v^\ell \rangle = \langle A_{kj'_k}, v^\ell \rangle =: \bar{\delta}_k, \quad \min_{j \in 1:2} \langle A_{kj}, v^\ell \rangle = \langle A_{kj''_k}, v^\ell \rangle =: \underline{\delta}_k.$$

2. Вычисляем $\Delta(\lambda^\ell) = \max_{k \in 1:s} \Delta_k(\lambda^\ell)$, где $\Delta_k(\lambda^\ell) = \bar{\delta}_k - \underline{\delta}_k$.

Если $\Delta(\lambda^\ell) = 0$, то процесс закончен и v^ℓ — решение задачи (6).

3. Пусть при некотором $k \in 1 : s$ имеем $\Delta(\lambda^\ell) = \Delta_k(\lambda^\ell) > 0$. Вычисляем $t_\ell^* = \min \left\{ \lambda_{kj'_k}^\ell, \frac{\Delta(\lambda^\ell)}{\|A_{kj'_k} - A_{kj''_k}\|^2} \right\}$ и получаем очередное приближение $v^{\ell+1} = v^\ell - t_\ell^* (A_{kj'_k} - A_{kj''_k})$.

4. Коэффициенты вектора $\lambda^{\ell+1}$ пересчитываем по правилам:

$$\lambda_{ij}^{\ell+1} = \begin{cases} \lambda_{kj'_k}^\ell - t_\ell^*, & \text{при } i = k, j = j'_k, \\ \lambda_{kj''_k}^\ell + t_\ell^*, & \text{при } i = k, j = j''_k, \\ \lambda_{ij}^\ell, & \text{иначе.} \end{cases}$$

Замечание 1. *Возможность независимого вычисления*

$$\Delta_k(\lambda) = \max_{j \in M_k^+(\lambda)} \langle A_{kj}, v \rangle - \min \{ \langle A_{k1}, v \rangle, \langle A_{k2}, v \rangle \}, \quad k \in 1 : s$$

существенно ускоряет процесс поиска $\Delta(\lambda)$ при больших значениях s .

Замечание 2. *Очередное приближение вектора λ отличается от предыдущего всего двумя компонентами.*

Замечание 3. *Очевидно, что очередное приближение вектора $v^{\ell+1}$ целесообразно вычислять по соотношению из пункта 3), нежели по формуле $v^{\ell+1} = A\lambda^{\ell+1}$.*

Рабочая схема № 2 (циклического проектирования)

Общий шаг. Пусть уже имеется ℓ -е приближение $v^\ell = A\lambda^\ell$.

1. Последовательно для каждого $k = 1 : s$ производим:

- Находим индексы $j'_k \in M_k^+(\lambda^\ell)$, $j''_k \in 1 : d_k$, такие, что

$$\max_{j \in M_k^+(\lambda^\ell)} \langle A_{kj}, v^\ell \rangle = \langle A_{kj'_k}, v^\ell \rangle =: \bar{\delta}_k, \quad \min_{j \in 1:d_k} \langle A_{kj}, v^\ell \rangle = \langle A_{kj''_k}, v^\ell \rangle =: \underline{\delta}_k.$$

- Вычисляем $\Delta_k(\lambda^\ell) = \bar{\delta}_k - \underline{\delta}_k$. Если $\Delta_k(\lambda^\ell) = 0$, то переходим к очередному номеру k .
- Если $\Delta_k(\lambda^\ell) > 0$, то вычисляем $t^* = \min \left\{ \lambda_{kj'_k}^\ell, \frac{\Delta_k(\lambda^\ell)}{\|A_{kj'_k} - A_{kj''_k}\|^2} \right\}$ и переопределяем

$$v^\ell := v^\ell - t^*(A_{kj'_k} - A_{kj''_k}), \quad \lambda_{kj}^\ell := \begin{cases} \lambda_{kj'_k}^\ell - t^*, & \text{при } j = j'_k, \\ \lambda_{kj''_k}^\ell + t^*, & \text{при } j = j''_k, \\ \lambda_{kj}^\ell, & \text{иначе.} \end{cases}$$

2. Вычисляем $\Delta(\lambda^\ell) = \max_{k \in 1:s} \Delta_k(\lambda^\ell)$. Если $\Delta(\lambda^\ell) = 0$, то процесс закончен и v^ℓ — решение задачи (6).

Если $\Delta(\lambda^\ell) > 0$, то полагаем $v^{\ell+1} = v^\ell$, $\lambda^{\ell+1} = \lambda^\ell$.

Замечание 4. Для рабочей схемы № 2 также справедливы замечания 2 и 3.

Замечание 5. Окончание вычислений в рабочей схеме № 2 можно немного оптимизировать.

Действительно, пусть для некоторого $j \in 2 : s$ имеем

$$\begin{aligned}\Delta_k(\lambda^\ell) &> 0 \quad \text{при } k = j - 1, \\ \Delta_k(\lambda^\ell) &= 0 \quad \text{при } k = j : s, \\ \Delta_k(\lambda^{\ell+1}) &= 0 \quad \text{при } k = 1 : (j - 1),\end{aligned}$$

то $v^{\ell+1} = v^\ell$ — решение задачи. Проще говоря, для окончания работы алгоритма нужно, чтобы s раз подряд обращалась в ноль оценка плана Δ_k .

Замечание 6. Рабочая схема № 2 без существенных доработок функционирует при динамическом изменении как количества множеств C_k , так и количества элементов в их составе, т.е. при изменениях s и d_k .

ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

$$\sum_{k=1}^s |x^T a_k - b_k| \longrightarrow \min_{x \in \mathbb{R}^n}$$

МЕТОД ГЛОБАЛЬНОГО ГИПОДИФФЕРЕНЦИАЛЬНОГО СПУСКА *

Пусть уже имеется $x^{(\ell)} \in \mathbb{R}^n$. Переход от ℓ -го приближения к $(\ell+1)$ -му осуществляется следующим образом ($\ell \geq 0$):

1. Вычислим направление спуска и величину шага:

$$(v_\ell, \alpha_\ell) = \arg \min \left\{ \|v\|^2 + \alpha^2 \mid (v, \alpha) \in df(x^{(\ell)}) \right\}.$$

Если $\|(v_\ell, \alpha_\ell)\| = 0$, то $x^{(\ell)}$ — решение задачи. Иначе $\alpha_\ell < 0$ и переходим к следующему шагу.

2. Очередное приближение: $x^{(\ell+1)} = x^{(\ell)} + \frac{v_\ell}{\alpha_\ell}$.

*1. Долгополик М. В. Метод кодифференциального спуска для кусочно-аффинных функций // Семинар «CNSA & NDO». Избранные доклады. 16 мая 2019 г. (<http://cnsa.cmlaboratory.com/rep19.shtml#0516>)

БЛАГОДАРЮ ЗА ВНИМАНИЕ!*

Тамасян Григорий Шаликович, к.ф.-м.н., с.н.с.,
Военно-космическая академия имени А.Ф.Можайского,
Институт проблем машиноведения РАН,
g.tamasyan@ya.ru

*Работа выполнена в Институте проблем Машиноведения РАН при финансовой поддержке Российского научного фонда (проект № 23-41-00060).