Методы оптимизации. Семинар 7
Субдифференциальное исчисление.

Корнилов Никита Максимович

Московский физико-технический институт

16 октября 2025г

Субдифференциал

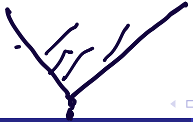
Обощение свойств градиента на точки недифференцируемости.

Definition (Субградиент)

Пусть дана функция $f : U \to \overline{\mathbb{R}}$ и точка $x_0 \in \operatorname{dom} f$. Элемент $s \in U$ называется *субградиентом функции f в точке x_0*, если

$$f(x) \geq f(x_0) + \langle s, x - x_0 \rangle, \quad \forall x \in \operatorname{dom} f.$$

- Множество всех субград. в точке $x_0 \in \operatorname{dom} f$ называется *субдифференциалом f в точке x_0* и обозначается как $\partial f(x_0)$. Для $x_0 \notin \operatorname{dom} f$ считаем $\partial f(x_0) = \emptyset$.
- Если $\partial f(x_0) \neq \emptyset$, то *функция f называется субдифференцируемой в точке x_0*.

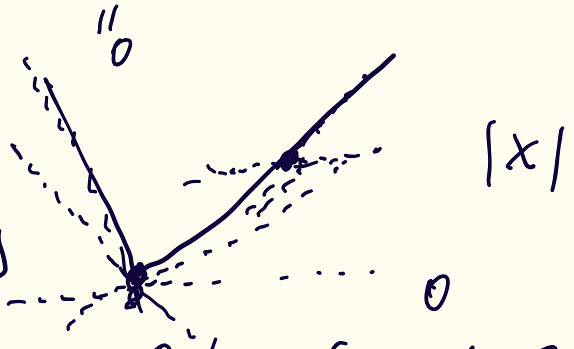


$-|x|$

$f(x) \geq f(y) + \langle s, x-y \rangle$

$\partial f(1) = \{1\}$

$\partial f(0) = [-1;1]$



$0,$ $\quad s \in [-1;1]$

$|x| \geq s \cdot x \quad \forall x \in R \qquad s>1$ или $-1<s$

Посчитаем по определению

Example

Пусть $f : \mathbb{R} \to \mathbb{R}$ функция модуля $f(x) = |x|$. Посчитайте субдифференциал $\partial f(0)$ в точке 0.

Example

Пусть $f : \mathbb{S}^n \to \mathbb{R}$ задаётся как $f(X) = \lambda_{\max}(X)$. Покажите, что $uu^\top \in \partial f(X)$, где u — нормированный собственный вектор, соответствующий наибольшему собственному числу $\lambda_{\max}(X)$.

Example

Пусть даны функции $f : \mathbb{R}^n \to \mathbb{R}$ и $h : \mathbb{R}^n \to \mathbb{R}^d$. По определению функция $g : \mathbb{R}^d \to \mathbb{R}$ задана как

$$g(\lambda) = -\min{}_{x \in \mathbb{R}^n} \left\{ f(x) + \langle \lambda, h(x) \rangle \right\}.$$

Покажите, что $-h(x_0) \in \partial g(\lambda_0)$ для $\lambda_0 \in \operatorname{dom} g$, где x_0 — точка, в которой достигается min в $g(\lambda_0)$.

$$f(X) = \lambda_{max}(X) = \max_{\|\omega\|=1} \omega^T X \omega$$

$$\omega_{max} - \lambda_{max}(X)$$

$$\omega_m \omega_m^T \in \partial f(X)$$

$$\forall Y \in S^d \quad \lambda_m(Y) = \max_{\omega} \omega^T Y \omega \geq \omega_m^T Y \omega_m$$

$$= \omega_m^T Y \omega_m + \underbrace{\omega_m^T X \omega_m}_{\lambda_{max}(X)} - \omega_m^T X \omega_m$$

$$= \lambda_{max}(X) + \underbrace{\omega_m^T (Y-X) \omega_m}_{\| \ \mathrm{Tr}(\omega_m^T (Y-X) \omega_m)}$$

$$\| \quad \mathrm{Tr}(\omega_m \omega_m^T (Y-X)) = \langle \omega_m \omega_m^T, Y-X \rangle$$

$\forall Y \in S^d$

$$\lambda_{max}(Y) \geq \lambda_{max}(X) + \langle \omega_m \omega_m^T, Y-X \rangle$$

$$f(y) \qquad f(x) \qquad \langle s, y-x \rangle$$

$\partial f(x) = \emptyset$

$\partial f(y) = \{0\}$

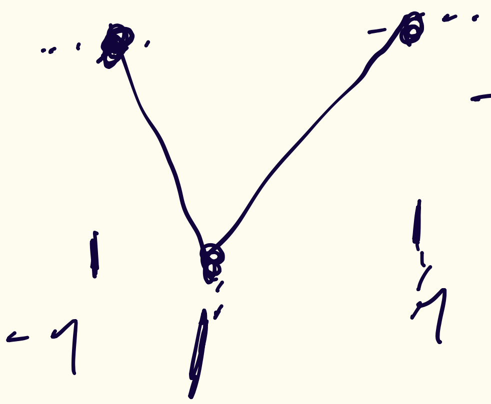
Свойства субдифференциала для **выпуклых функций**

1. Субдифференциал — это выпуклое замкнутое множество (возможно, пустое)
   $\partial f(x_0) = \cap_{x \in \mathrm{dom}\, f} \{s \in U : f(x) \geq f(x_0) + \langle s, x - x_0 \rangle\}$
2. Пусть $x_0 \in int(\mathrm{dom}\, f)$ и функция f выпуклая. Тогда $\partial f(x_0)$ не пуст и является выпуклым компактом (f - субдиф на $int(\mathrm{dom}\, f)$).
3. Пусть $x_0 \in int(\mathrm{dom}\, f)$ и функция f выпуклая. Тогда функция дифференцируемая в точке x_0 тогда и только тогда, когда субдифференциал состоит из одного элемента $\partial f(x_0) = \{\nabla f(x_0)\}$.
4. **Критерий глобального минимума**. Точка x_0 — глобальный минимум функции $f : U \to \mathbb{R}$ тогда и только тогда, когда $0 \in \partial f(x_0)$.
   $f(x) \geq f(x_0) + \langle g, x - x_0 \rangle \; \forall x \in dom$
5. **Критерий условного минимума**. Точка $x_0 \in Q$ — условный минимум функции $f : U \to \mathbb{R}$ на $Q \subset U$ тогда и только тогда, когда $\exists g \in \partial f(x_0) : \langle g, x - x_0 \rangle \geq 0, \forall x \in Q$.

int dom f – всегда есть субдифф

огр и вып и замк

$\partial f(x) = \{\nabla f(x)\}$



– вып

$[0;1)$ и $(-1,0]$ – diff –

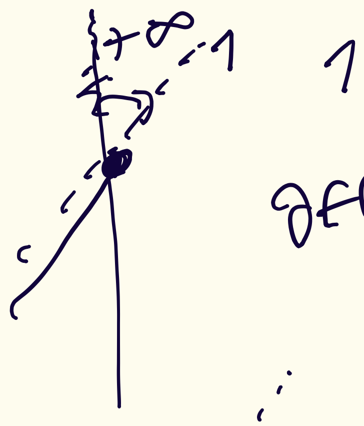
$\partial f(x) = \{f'(x)\}$

$= \{1, \ x>0$

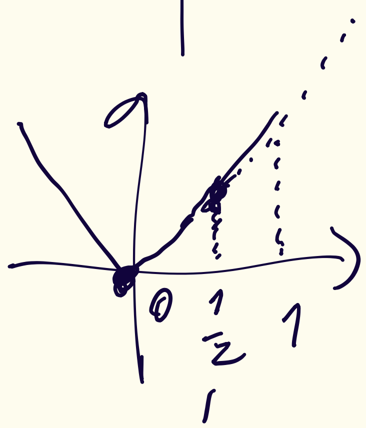
$-1, \ x<0$

0 – not diff – $[-1;1]$

1 – граница



$\partial f(x) = [-1; +\infty)$ – не компакт



$0 \in \partial f(0) \to$ glob min

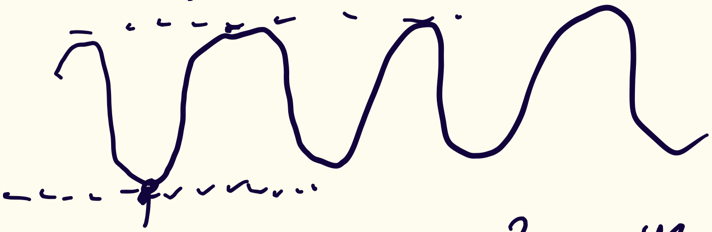
min на $[\frac{1}{2}; 1]$ но $\frac{1}{2}$ усл min

$\partial f(x) = \{1\}$ $\langle 1, x - \frac{1}{2} \rangle \geq 0$ $\forall x \geq \frac{1}{2}$

Свойства субдифференциала для **произвольных функций**

1. Субдифференциал — это всё ещё выпуклое замкнутое множество (возможно, пустое).
2. Пусть $x_0 \in int(\text{dom } f)$ и f дифференцируема в x_0. Тогда либо $\partial f(x_0) = \emptyset$, либо $\partial f(x_0) = \{\nabla f(x_0)\}$.
   Только в случае выпуклых функций множество дифференцируемых точек является подмножеством субдифференцируемых!
3. **Критерий глобального минимума**. Точка x_0 — глобальный минимум функции $f : U \to \mathbb{R}$ тогда и только тогда, когда $0 \in \partial f(x_0)$.

$\partial f(x) = \varnothing$



— везде diff

$\partial f \ni 0 ? = \exists f'(x) ?$ — из 1

вып — loc min нет

loc min — либо

невып

$\partial f \ni 0 ?$ ✗

либо

$\partial f \ni \varnothing ?$ ✓

Примеры

Example

Пусть $f : \mathbb{R} \to \mathbb{R}$ функция $f(x) = |x|$. Найдите $\partial f(x)$.

Example

Пусть $f : \left[0, \frac{3\pi}{2}\right] \to \mathbb{R}$ функция $f(x) = \cos(x)$. Найдите $\partial f(x)$.

Example (Норма)

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ функция $f(x) = \|x\|$. Найдите $\partial f(x)$.
Hint. По определению, сопряженная норма $\|g\|_* := \sup_{\|x\| \leq 1} \langle g, x \rangle$.

$f(x) = \|x\|$

g - вектор - кандидат

$\forall y \in U$

$f(y) - f(x) - \langle g, y - x\rangle = \|y\| - \|x\| - \langle g, y - x\rangle$

$= \|y\| - \langle g, y\rangle - (\|x\| - \langle g, x\rangle) \geq 0$

$\Updownarrow$

$\langle g, x\rangle - \|x\| \geq \langle g, y\rangle - \|y\| \quad \forall y$

$\Updownarrow$

$\langle g, x\rangle - \|x\| \geq \sup\limits_y \{\langle g, y\rangle - \|y\|\}$

Какой g - чтобы выполнялось

$\langle g, y\rangle \leq \|y\| \|g\|_* \qquad \|g\|_* = \sup\limits_{\|y\| \leq 1} \langle y, g\rangle$

• $\|g\|_* \leq 1 \Rightarrow \langle g, y\rangle - \|y\| \leq \|y\| \|g\|_* - \|y\| \leq \|y\| - \|y\| \leq 0 \Rightarrow \sup \{\langle g, y\rangle - \|y\|\} \leq 0$

$\langle g, x\rangle - \|x\| \geq 0$

$\|g\|_* \|x\| \leq \|x\| \Rightarrow$ • $\langle g, x\rangle = \|x\|$

$\{g \mid \|g\|_* \leq 1, \langle g, x\rangle = \|x\|\} = \partial f(x)$

- показано

$\|g\|_* > 1 \Rightarrow \|g\|_* = \sup\limits_{\|y\| \le 1} \langle g, y \rangle$

$\exists \hat{y} \quad \langle g, \hat{y} \rangle > 1 \quad \|\hat{y}\| \le 1$

$\underset{>1}{\langle g, \hat{y} \rangle} - \underset{\le 1}{\|\hat{y}\|} > 0 \qquad \hat{y}_t = \hat{y} \cdot t \quad t > 0$

$\langle g, \hat{y}_t \rangle - \|\hat{y}_t\| = t \underset{>0}{(\langle g, \hat{y} \rangle - \|\hat{y}\|)}$

$t \to +\infty \quad \sup\limits_t \{ \; \} = +\infty$

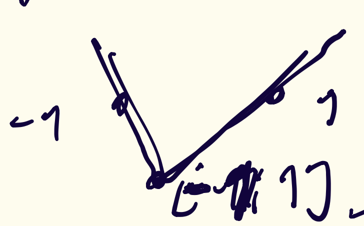
$\underset{\in \mathbb{R}}{\langle g, x \rangle - \|x\|} \ge +\infty \quad \times$

$\partial f(x) = \{ g \mid \|g\|_* \le 1 \quad \langle g, x \rangle = \|x\| \}$

$\partial \|\cdot\|_2 (x) = \{ g \mid \|g\|_2 \le 1, \langle g, x \rangle = \|x\|_2 \}$

$\forall x \ne 0 \quad g = \frac{x}{\|x\|_2} = \nabla \|x\|_2$

$\partial \|\cdot\|_2 (0) = \{ g \mid \|g\|_2 \le 1 \} =$



на всем R^n

- Если функция f имеет открытый dom f, то субдифференцируемость эквивалентна выпуклости.
- В общем случае для неоткрытых dom f из выпуклости не следует субдифференцируемость на граничных точках ($f(x) = -\sqrt{x}$).
- Но из субдифференцируемости всегда следует выпуклость.
- Если функция f выпуклая, то она может быть не дифференцируема только в счетном числе точек из *int* dom f.

$f' \to +\infty$ $\partial f(0) = \emptyset$

dom $f = [0; +\infty]$

$(0; +\infty) - -\frac{1}{2\sqrt{x}}$

Общий алгоритм

Для выпуклых функций:

1. Для внутренних дифференцируемых точек считаем градиент, он и есть субдифференциал $\partial f(x_0) = \{\nabla f(x_0)\}$.
2. Для недифф, сложно дифференцируемых или граничных точек — смотрим по субдифф исчислению или по определению (или на график).

Для невыпуклых функций:

Внимательно изучаем каждую внутреннюю и граничную точку по определению или графику.

Субдифференциальное исчисление I

Для простоты изложения считаем, что dom всех функций это U.

- **Умножение на константу.** Пусть дана произвольная функция f, точка $x_0 \in U$ и положительный коеф $c \geq 0$, тогда

$$\partial \left[c \cdot f\right](x_0) = c \cdot \partial f(x_0).$$

- **Сумма.** Пусть даны *выпуклые* функции f и g на U, тогда

$$\partial(f+g)(x_0) = \partial f(x_0) + \partial g(x_0).$$ — Минковски

Обобщение на m выпуклых функций и $c_i \geq 0, i \in \overline{1, m}$

$$\partial \left(\sum_{i=1}^{m} c_i f_i \right)(x_0) = \sum_{i=1}^{m} c_i \partial f_i(x_0).$$

Если f, g — произвольные, то верно только

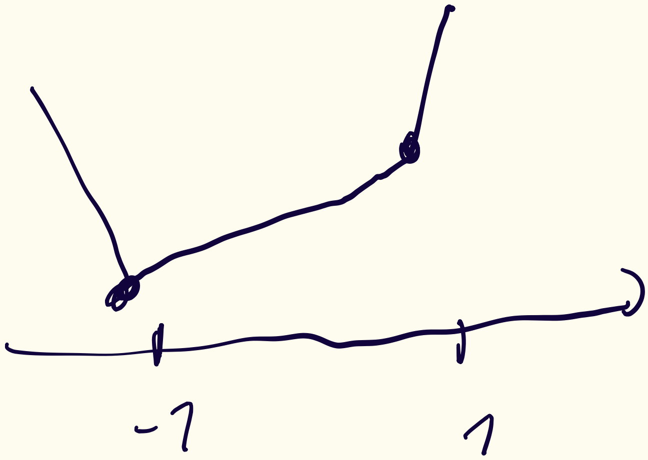
$$\partial(f+g)(x_0) \supseteq \partial f(x_0) + \partial g(x_0).$$

Примеры

Example

Посчитайте $\partial f(x)$ для $f(x) : \mathbb{R} \to \mathbb{R} = 2|x+1| + |x-1|$.

$f(x) = 2|x+1| + |x-1|$



$\partial|\cdot+1|(x) = \begin{cases} 1, & x>-1 \\ -1, & x<-1 \\ [-1;1], & x=-1 \end{cases}$

$\partial|\cdot-1|(x) = \begin{cases} 1, & x>1 \\ -1, & x<1 \\ [-1;1], & x=1 \end{cases}$

$2\cdot\partial|\cdot+1|(x) = \begin{cases} 2, & x>-1 \\ -2, & x<-1 \\ [-2;2], & x=-1 \end{cases}$

$x<-1$ $\quad 2\partial|\cdot+1|(x) + \partial|\cdot-1|(x)$

$= -2 - 1 = -3$

$x=-1$ $\quad = [-2;2] + (-1) =$

$[-3;1]$

$x \in (-1;1)$ $\quad = 2 - 1 = 1$

$x = 1$ $\quad = 2 + [-1;1] = [1;3]$

$x > 1$ $\quad$ $2 + 1 = 3$

Субдифференциальное исчисление II

- **Аффинное преобразование.** Аффинное преобразование $g(x) = Ax + b$ и *выпуклая* f в точке x_0 дают

$$\partial \left(f(Ax+b) \right)(x_0) = A^\top \partial f(Ax_0 + b).$$

- **Неубывающая композиция.** Пусть $g_i(x) : \mathbb{R}^n \to \mathbb{R}$ *выпуклые* функции для всех $i \in \overline{1, m}$, а $\phi(u) : \mathbb{R}^m \to \mathbb{R}$ *неубывающая выпуклая* функция. Тогда для композиции $f(x) : \mathbb{R}^n \to \mathbb{R} := \phi(g_1(x), \ldots, g_m(x))$ и точки x_0 верно

$$\partial f(x_0) = \bigcup_{p \in \partial \phi(g(x_0))} \sum_{i=1}^{m} p_i \partial g_i(x_0).$$

Если ϕ дифф, то

$$\frac{\partial f}{\partial x} = \frac{\partial \varphi}{\partial g_1}\frac{\partial g_1}{\partial x} + \frac{\partial \varphi}{\partial g_2}\frac{\partial g_2}{\partial x} + \ldots$$

$$\partial f(x_0) = \sum_{i=1}^{m} \frac{\partial \phi}{\partial u_i} \partial g_i(x_0).$$

Примеры

Example

Посчитайте $\partial f(x)$ для $f(x) : \mathbb{R}^n \to \mathbb{R} = \exp(\|Ax + b\|_2)$.

$\partial(f)$ $\quad f = \exp\left(\underbrace{\|Ax+b\|_2}_{\text{вып}}\right)$

вып неубыв

композиция с дифф

$$\partial f(x) = \exp(\|Ax+b\|_2)\,\partial(\|Ax+b\|_2)$$

$$\partial(\|Ax+b\|_2) = A^T\,\partial(\|\cdot\|_2)(Ax+b)$$

$$\partial(\|\cdot\|_2)(\cancel{x}\,y) = \begin{cases} \dfrac{y}{\|y\|_2}, & y \neq 0 \\ \underline{\{g \mid \|g\|_2 \leq 1\}}, & y = 0 \end{cases}$$

$$= \begin{cases} \dfrac{A^T(Ax+b)}{\|Ax+b\|_2}, & Ax+b \neq 0 \\ \{A^T g \mid \|g\|_2 \leq 1\}, & Ax+b = 0 \end{cases}$$

- **Конечный максимум.** Пусть даны *выпуклые функции* $f_i : U \to \mathbb{R}$ функции для $i = \overline{1, m}$ и $f(x) := \max_{i=\overline{1,m}} f_i(x)$. Тогда для $x_0 \in U$ верно

$$\partial f(x_0) = \mathrm{Conv}\left(\bigcup_{i \in I(x_0)} \partial f_i(x_0)\right),$$

где $I(x_0) := \{i \in \overline{1, m} : f_i(x_0) = f(x_0)\}$ – множество индексов, на которых достигается max.
Если f_i — произвольные, то верно только

$$\partial f(x_0) \supseteq \overline{\mathrm{Conv}}\left(\bigcup_{i \in I(x_0)} \partial f_i(x_0)\right).$$

Субдифференциальное исчисление IV

- **Бесконечный максимум.** Пусть даны *выпуклые* функции $f_i : U \to \mathbb{R}$ функции для $i \in I$, где I — произвольное *компактное* множество такое, что $i \to f_i(x)$ полунепрерывна сверху на I. Тогда для $f(x) := \max_{i \in I} f_i(x)$ и точки $x_0 \in U$ верно

$$\partial f(x_0) = \operatorname{Conv}\left(\bigcup_{i \in I(x_0)} \partial f_i(x_0)\right),$$

где $I(x_0) := \{i \in I : f_i(x_0) = f(x_0)\}$ – множество индексов, на которых достигается max.

Если f_i и I произвольные, то верно только

$$\partial f(x_0) \supseteq \overline{\operatorname{Conv}}\left(\bigcup_{i \in I(x_0)} \partial f_i(x_0)\right).$$

Примеры

Example (Модуль через max)

Посчитайте $\partial f(x)$ для $f(x) : \mathbb{R} \to \mathbb{R} = |x|$ через субдифф максимума.

Example (Норма через max)

Посчитайте $\partial f(x)$ для $f(x) : \mathbb{R}^n \to \mathbb{R} = \|x\|$ через субдифф максимума.

Example (ℓ_1-норма)

Посчитайте $\partial f(x)$ для $f(x) : \mathbb{R}^n \to \mathbb{R} = \|x\|_1$ через сумму.

$|x| \in R$

$|x| = \max\{-x; x\}$

$\partial|\cdot| = \partial \max\{-x, x\}$

$x<0$ max в $-x$

в одной $-x$ достигается

$$\partial|\cdot|(x) = \operatorname{conv}\Big(\bigcup_{i\in I} \partial f_i\Big) = \operatorname{conv}(\{-1\}) = \{-1\}$$

max X достигается

$$x>0 \quad \partial|\cdot|(x) = \operatorname{conv}(\{1\}) = 1$$

$$x=0 \quad \partial|\cdot|(x) = \operatorname{conv}\big(\partial(x)(0) \cup (\partial(-x))(0)\big)$$

$$= \operatorname{conv}(\{1\} \cup \{-1\}) = [-1;1]$$

-1 $\quad$ 1

$$\partial\|\cdot\|(x) = \max_{\|g\|_* \le 1} \langle g, x\rangle$$

$\partial\langle g, x\rangle$

$$\partial \max_{\|g\|_*\le 1} \langle g, x\rangle = \operatorname{conv}\Big(\bigcup_{\substack{\|x\| = \langle g,x\rangle \\ \|g\|_*\le 1}} g\Big)$$

выпуклое

$$= \operatorname{conv}\big(\{g \mid \|g\|_* \le 1,\ \langle g, x\rangle = \|x\|\}\big) = \{g \mid \|g\|_* \le 1\ \langle g, x\rangle = \|x\|\}$$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\partial \|\cdot\|_1 (x) = \sum_{i=1}^{n} \partial |x_i| =$$

$$\sum_{i \in I_{\neq}} \mathrm{sing}(x_i) e_i + \sum_{i \in I_{=}} [-e_i, e_i]$$

$$I_{\neq} = \{ i : |x_i| \neq 0 \}$$

$$I_{=} = \{ i : |x_i| = 0 \}$$

В частности можно брать $e_i \, \mathrm{sing}(x_i)$ из всех $[-e_i, e_i]$ для $i \in I_{=}$

$$\Rightarrow \mathrm{sing}(x) \in \partial \|\cdot\|_1 (x)$$

- $f(x) : \mathbb{R}^n \to \mathbb{R}$ выпуклая, но негладкая.
- Moreau-Yosida envelope($\lambda > 0$) позволяет сделать новую выпуклую гладкую функцию

$$M_{\lambda f}(x) = \inf_{u \in \mathbb{R}^n} \left(f(u) + \frac{1}{2\lambda} \|u - x\|_2^2 \right).$$

- Посчитаем $M_{\lambda f}$ для $|x|$

$$M_{\lambda|\cdot|}(x) = \begin{cases} \frac{x^2}{2\lambda}, & |x| \leq \lambda, \\ |x| - \lambda/2, & |x| \geq \lambda. \end{cases}$$

Свойства Moreau-Yosida Envelope

- $M_{\lambda f}(x)$ — выпуклая (infimal convolution) и $\frac{1}{\lambda}$-гладкая.
- Можно построить градиентные методы над $M_{\lambda f}$

$$\nabla_x M_{\lambda f}(x) = \frac{1}{\lambda}(x - u^*), \quad u^* = \arg\min \left(f(u) + \frac{1}{2\lambda}\|u - x\|_2^2 \right).$$

- **Множество минимумов f и $M_{\lambda f}$ совпадают.**