

Методы оптимизации. Семинар 7

Субдифференциальное исчисление.

Корнилов Никита Максимович

Московский физико-технический институт

16 октября 2025г

Субдифференциал

Обобщение свойств градиента на точки недифференцируемости.

Definition (Субградиент)

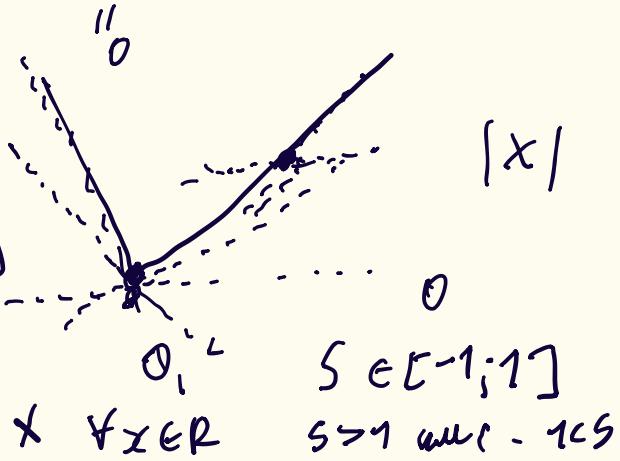
Пусть дана функция $f : U \rightarrow \overline{\mathbb{R}}$ и точка $x_0 \in \text{dom } f$. Элемент $s \in U$ называется *субградиентом функции f в точке x_0* , если

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

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



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



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

$$\partial f(0) = \{1; 1\}$$

$$|x| \geq s \cdot x \quad \forall x \in \mathbb{R}$$

$$s \in [-1; 1]$$

$$s > 1 \text{ und } -1 < s$$

Посчитаем по определению

Example

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

Example

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

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

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

$$f[X] = \lambda_{\max}(X) \leq \max_{\|\alpha\|_2=1} \alpha^T X \alpha$$

$$\vartheta_{\max} = \lambda_{\max}(X)$$

$$(\vartheta_m^T \in \partial f(x))$$

$$\forall Y \in S^d \quad \lambda_m(Y) = \max_{\|\alpha\|_2=1} \alpha^T Y \alpha \geq \vartheta_m^T Y \vartheta_m$$

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

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

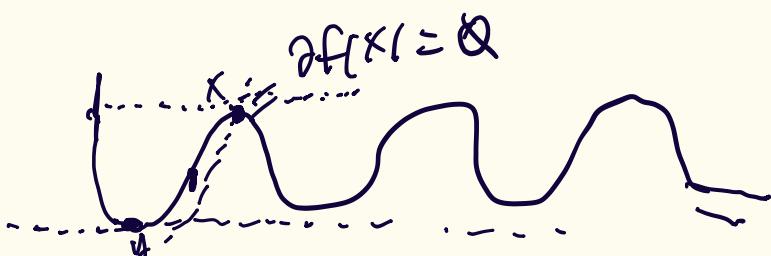
"

$$\text{Tr}(\vartheta_m \vartheta_m^T (Y-X)) =$$

$$\forall Y \in S^d \quad \langle \vartheta_m \vartheta_m^T, Y-X \rangle$$

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

$$f(y) \quad f(x) \quad \langle S, y-x \rangle$$



Свойства субдифференциала для выпуклых функций

- ① Субдифференциал — это выпуклое замкнутое множество (возможно, пустое)

$$\partial f(x_0) = \bigcap_{x \in \text{dom } f} \{s \in U : f(x) \geq f(x_0) + \langle s, x - x_0 \rangle\}$$

- ② Пусть $x_0 \in \text{int}(\text{dom } f)$ и функция f выпуклая. Тогда $\partial f(x_0)$ не пуст и является выпуклым компактом (f - субдиф на $\text{int}(\text{dom } f)$).

- ③ Пусть $x_0 \in \text{int}(\text{dom } f)$ и функция f выпуклая. Тогда функция дифференцируемая в точке x_0 тогда и только тогда, когда субдифференциал состоит из одного элемента $\underline{\partial f(x_0) = \{\nabla f(x_0)\}}$.

- ④ Критерий глобального минимума. Точка x_0 — глобальный минимум функции $f : U \rightarrow \mathbb{R}$ тогда и только тогда, когда $0 \in \partial f(x_0)$.

$$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle \quad \forall x \in U$$

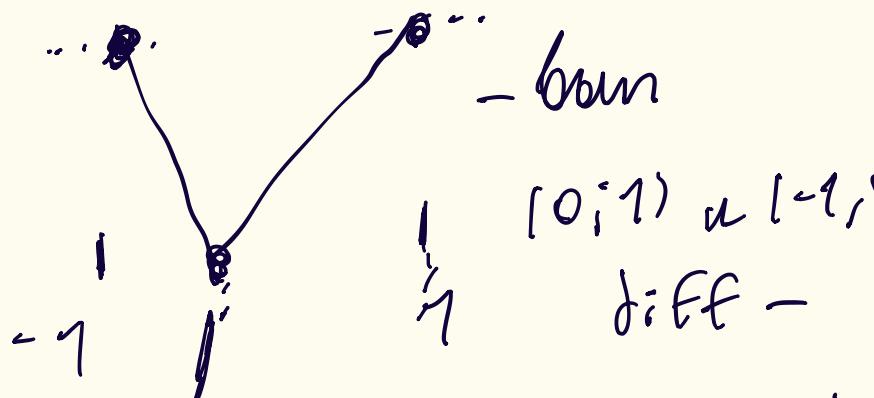
- ⑤ Критерий условного минимума. Точка $x_0 \in Q$ — условный минимум функции $f : U \rightarrow \mathbb{R}$ на $Q \subset U$ тогда и только тогда, когда $\exists g \in \partial f(x_0) : \langle g, x - x_0 \rangle \geq 0, \forall x \in Q$.



int dom f - бетега енсүү сүдүрүүлүп

орында барып
у заман

$$\partial F(x) = 3\sqrt{f(x)^2}$$



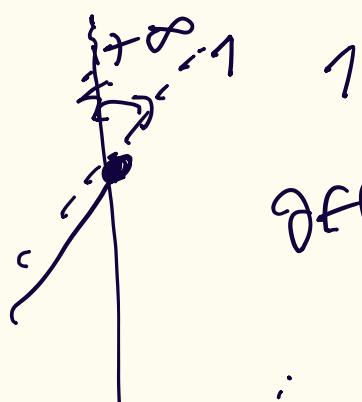
$$\partial F(x) = 3f'(x)^{\frac{1}{2}}$$

$$= 3, \quad x > 0$$

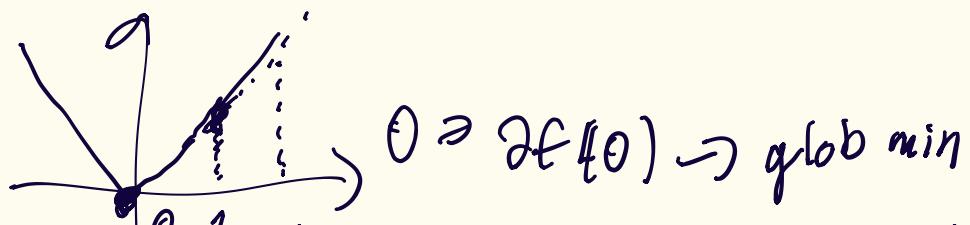
$$-1, \quad x < 0$$

$$\Theta \subset \text{not diff} = [-1; 1]$$

1 - ырынчы



$$\partial F(x) = [-1; +\infty) \leftarrow \text{ре көмүнүк}$$



min на $[-\frac{1}{2}; 1]$ но $\frac{1}{2}$ ын мин

$$\partial F(x) = 3x^2 \quad \langle 1, x - \frac{1}{2} \rangle \geq 0 \quad \forall x \in \frac{1}{2}$$

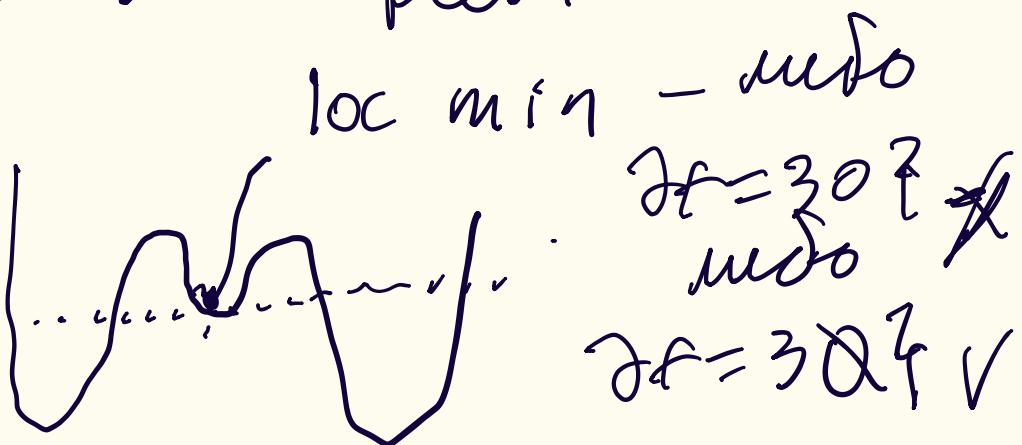
Свойства субдифференциала для произвольных функций

- ① Субдифференциал — это всё ещё выпуклое замкнутое множество (возможно, пустое).
- ② Пусть $x_0 \in \text{int}(\text{dom } f)$ 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 $f : U \rightarrow \mathbb{R}$ тогда и только тогда, когда $0 \in \partial f(x_0)$.



$$\partial f = 304 = 3\partial^2 f(x)^2 - \text{why?}$$

bowl - loc min num
flat bowl



$$\begin{aligned} \partial f &= 304 - \cancel{\text{why}} \\ \text{min} &\quad \cancel{\text{why}} \\ \partial f &= 304 \checkmark \end{aligned}$$

Примеры

Example

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

Example

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

Многомерный пример

Example (Норма)

Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 -белзор - көзүнүштөм

$$\forall y \in V$$

$$\begin{aligned} f(y) - f(x) - \langle g, y - x \rangle &= \|y\| - \|x\| - \langle g, y - x \rangle \\ &= \|y\| - \langle g, y \rangle + (\|x\| - \langle g, x \rangle) \geq 0 \end{aligned}$$

\Downarrow

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

\Downarrow

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

Көрсүк g -иншік берилгенде

$$\langle g, y \rangle \leq \|g\| \|y\| \quad \|g\|_* = \sup_{\|y\| \leq 1} \langle g, y \rangle$$

$$\begin{aligned} \bullet \|g\|_* \leq 1 &\Rightarrow \langle g, y \rangle - \|y\| \leq \|g\| (\|g\|_* - 1) \leq 0 \\ \|y\| \leq \|y\|_* &\Leftrightarrow \sup_y \{\langle g, y \rangle - \|y\|\} \leq 0 \end{aligned}$$

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

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

$$\exists g \mid \|g\|_* \leq 1, \langle g, x \rangle = \|x\|^2 = 2f(x)$$

\mathcal{C} - негизгиси

$$\|g\|_* > 1 \Rightarrow \|g\|_* = \sup_{\|y\|_1 \leq 1} \langle g, y \rangle$$

$$\exists \vec{y} \quad \langle g, \vec{y} \rangle > 1 \quad \|g\|_* \leq 1$$

$$\langle g, \vec{y} \rangle - \|g\|_* \geq 0 \quad \vec{y}_t = \vec{y} \cdot t \quad t > 0$$

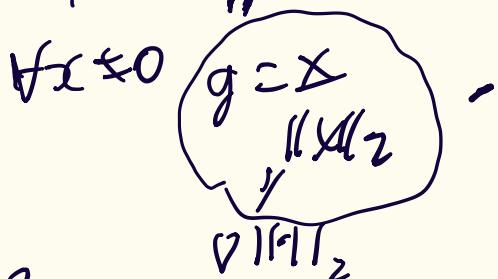
$$\langle g, \vec{y}_t \rangle - \|g\|_* \geq t(\langle g, \vec{y} \rangle - \|g\|_*)$$

$t \rightarrow +\infty \quad \sup_t \{ \} = +\infty$

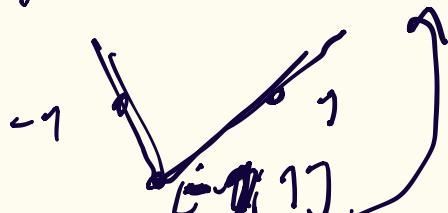
$$\langle g, x \rangle - \|x\|_* \geq +\infty$$

$$\mathcal{F}(x) = \{g \mid \|g\|_* \leq 1 \quad \langle g, x \rangle = \|x\|_*\}$$

$$\mathcal{F} \cdot \| \cdot \|_2(x) = \{g \mid \|g\|_2 \leq 1, \langle g, x \rangle = \|x\|_2\}$$



$$\mathcal{F} \cdot \| \cdot \|_2(0) = \{g \mid \|g\|_2 \leq 1\}^2 =$$



(Суб)Дифференцируемость и выпуклость

на всем \mathbb{R}^n

- Если функция f имеет открытый $\text{dom } f$, то субдифференцируемость эквивалентна выпуклости.
- В общем случае для неоткрытых $\text{dom } f$ 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 $\text{int dom } f$.



$$\text{dom } f = [0; +\infty)$$
$$(0; +\infty) - \frac{1}{2}tx$$

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

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

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

$$\partial(f + g)(x_0) = \partial f(x_0) + \partial g(x_0). - \text{Минимизация}$$

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

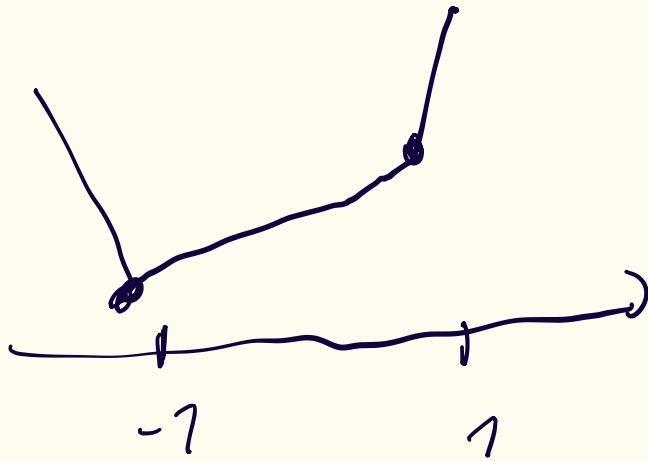
$$\partial(f + g)(x_0) \supseteq \partial f(x_0) + \partial g(x_0).$$

Примеры

Example

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

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



$$\begin{cases} 2x, & x > -1 \\ -1, & x < -1 \\ [-1; 1], & x = -1 \end{cases}$$

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

$$\begin{cases} 2x, & x > -1 \\ -2, & x < -1 \\ [-2; 1], & x = -1 \end{cases}$$

$$\begin{cases} 2x-1, & x > 1 \\ -1, & x < 1 \\ [-1; 1], & x = 1 \end{cases}$$

$$\begin{aligned} x < -1 & \quad 2[-2; 1](x) + 2[-1; 1](x) \\ & = -2 - 1 = -3 \end{aligned}$$

$$\begin{aligned} x < -1 & \quad = [-2; 1] + [-1; 1] = \\ & = [-3; 1] \end{aligned}$$

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

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

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

Субдифференциальное исчисление II

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

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

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

$$\frac{\partial f}{\partial x} = \frac{\partial \phi}{\partial q_1} \frac{\partial q_1}{\partial x} + \frac{\partial \phi}{\partial q_2} \frac{\partial q_2}{\partial x} + \dots$$
$$\partial f(x_0) = \sum_{i=1}^m \frac{\partial \phi}{\partial u_i} \partial g_i(x_0).$$

Примеры

Example

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

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

беск.
норма

компьютер с гипер

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

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

$$\partial (\| \cdot \|_2)(y) = \begin{cases} \frac{y}{\|y\|_2}, & y \neq 0 \\ \underbrace{\{y \mid \|y\|_2 \leq 1\}}, & y = 0 \end{cases}$$

$$= \begin{cases} \frac{A^T(Ax + b)}{\|Ax + b\|_2}, & Ax + b \neq 0 \end{cases}$$

$$\begin{cases} \{A^Tg \mid \|g\|_2 \leq 1\}, & Ax + b = 0 \end{cases}$$

Субдифференциальное исчисление III

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

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

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

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

Субдифференциальное исчисление IV

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

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

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

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

Примеры

Example (Модуль через max)

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

Example (Норма через max)

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

Example (ℓ_1 -норма)

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

$$|x| \in \mathbb{R}$$

$$|x| = \max(-x, x)$$

$$\partial I_{+}(x) = \partial \max(-x, x)$$

$$x < 0 \quad \max(-x, x) = -x$$

bogus - X generalisiert

$$\partial I_{+}(x) = \text{conv} \left(\bigcup_{i=1}^{\infty} \partial f_i \right) = \text{conv} \left(\begin{array}{c} \{1\} \\ \dots \\ \{3-1\} \end{array} \right)$$

$$x > 0 \quad \partial I_{+}(x) = \text{conv} \left(\begin{array}{c} \{1\} \\ \dots \\ \{3\} \end{array} \right) = \{1\}$$

$$x = 0 \quad \partial I_{+}(x) = \text{conv} \left(\partial I_{+}(0) \cup \{\partial I_{-}(0)\} \right)$$

$$= \text{conv} \left(\{1\} \cup \{3-1\} = [-1; 1] \right)$$

$$\underline{-1} \qquad \overline{1}$$

$$\partial ||\cdot||_p(x) = \max_{\|g\|_p \leq 1} \langle g, x \rangle$$

$$\langle g, x \rangle$$

$$\partial \max_{\|g\|_p \leq 1} \langle g, x \rangle = \text{conv} \left(\bigcup_{\|x\|=1} \langle g, \cdot \rangle \right)$$

$$\langle g, x \rangle$$

$$\|g\|_p \leq 1$$

$$\text{conv} \left(\begin{array}{c} \{g \mid \|g\|_p \leq 1\} \\ \langle g, x \rangle = \|x\| \end{array} \right) = \{g \mid \|g\|_p \leq 1 \wedge \langle g, x \rangle = \|x\|\}$$

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

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

$$\sum_{i \in I_x} \text{sing}(x_i) \ell_i + \sum_{i \in I_x^c} [-\ell_i, \ell_i]$$

$$I_x = \{ i ; |x_i| \neq 0 \}$$

$$I_x^c = \{ i ; |x_i| = 0 \}$$

В калесонији је овој броју $\ell_i \text{sing}(x_i)$
из броја $[-\ell_i, \ell_i]$ где $i \in I_x^c$

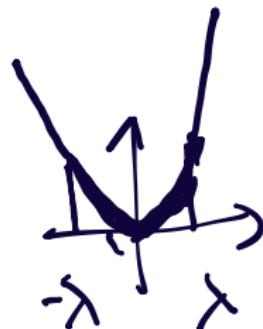
$$\Rightarrow \text{sing}(x) \in \partial \| \cdot \|_1(x)$$

Moreau-Yosida Envelope

- $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $|x| : \mathbb{R} \rightarrow \mathbb{R}$



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

$$\min_u |u| + \frac{1}{2\lambda} \|u-x\|^2 = g(u)$$

$$\partial \left(\frac{1}{2\lambda} (u-x)^2 \right) = \frac{1}{\lambda} (u-x)$$

$$\partial|u| = \begin{cases} 1, & u>0 \\ -1, & u<0 \\ [-1;1], & u=0 \end{cases}$$

$$\partial g(u) = \begin{cases} 1 + \frac{1}{\lambda} (u-x), & u>0 \\ -1 + \frac{1}{\lambda} (u-x), & u<0 \\ [-1;1] + \frac{1}{\lambda} (u-x), & u=0 \end{cases}$$

$$0 \in \text{werte } u>0 \quad 1 + \frac{1}{\lambda} (u-x) = 0$$

$$\cancel{u} - x - \cancel{\lambda} = 0 \quad x > \cancel{x}$$

$$g(u) = |x-u| + \frac{1}{2\lambda} (x-u-x)^2 = x-\lambda + \frac{1}{2\lambda} x^2 = x + \frac{\lambda}{2}$$

$$0 \in \partial g(u) \text{ werte } u<0$$

$$-1 + \frac{1}{\lambda} (u-x) = 0 \Leftrightarrow u = x + \lambda < 0 \quad x < -\lambda$$

o

$$g(\bar{x}) = \underbrace{|x+\lambda|}_{-x-\lambda} + \frac{|x+\lambda-x|^2}{2\lambda} =$$

$$-x + \frac{\lambda}{2}$$

$$0 \in \partial g(0)$$

$$\underbrace{[-1; 1]}_{\gamma} + \frac{1}{\lambda}(0 - x) = 0$$

$$x \in [-x; \lambda] \Rightarrow \bar{x} = 0$$

$$g(\bar{x}) = 0 + \frac{x^2}{2\lambda} \text{ and } x \in [-\lambda; \lambda]$$

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.

♥

$$g(x) = \sup_y f(x, y)$$

$$\nabla g(x) = \nabla_x f(x, \hat{y}) \rightarrow \sup \text{дальнейшее}$$