Ахмаджонов Мумтозбек Б05-927б

Домашнее задание №4

Задача 1.

Найти субдифференциал функции $\mathbf{ReLU}(x) = \max\{0, x\}$.

Решение:

По теореме о субдифференциале поточечного максимума,

$$\partial(\mathbf{ReLU}(x_0)) = \begin{cases} 0 & x_0 < 0 \text{ - очевидно} \\ 1 & x_0 > 0 \text{ - очевидно} \\ \mathbf{conv}\{\partial 0 \cup \partial x_0\} = \mathbf{conv}\{0, 1\} = [0, 1] & x_0 = 0 \end{cases}$$

Ответ:

$$\partial(\mathbf{ReLU}(x_0)) = \begin{cases} 0 & x_0 < 0 \\ 1 & x_0 > 0 \\ [0, 1] & x_0 = 0 \end{cases}$$

Задача 2.
Найти субдифференциал функции $f_p(x) = ||x||_p$ при $p = 1, 2, \infty$.
Решение:
Без ограничения общности, пусть $f_p : \mathbb{R}^n \to \mathbb{R}$.

- $p = 1$

$$f(x) = ||x||_1 = |x_1| + ... + |x_n| = \max\left\{g_k(x) = e_k^T x \ : \ e_k^i \in \{-1, 1\} \ \forall i \in \overline{1n}\right\}.$$

т.е. в $f(x) = \max\{g_1(x), \ ..., g_k(x)\} = e^T x, \ \ e^i = 1$, если $x_i > 0$, и $e^i = -1$, если $x_i < 0$. И -1 либо 1 в нуле. Т.к. $\forall k \ g_k(x) = e_k^T x$ дифференцирумы, то их субдифференциалы содержат только один субградиент:

$$\partial g_k(x) = e_k$$
$$e_k^i = \begin{cases} 1 & x_i > 0 \\ -1 & x_i < 0 \\ [-1, 1] & x_i = 0 \end{cases}$$

Таким образом, выпуклая оболочка всех таких субградиентов описанных выше имеет вид:

$$\partial f(x) = \left\{e \mid e \in [-1, 1]^n \wedge e^T x = ||x||_1\right\}.$$

- $p = 2$

$$f(x) = ||x||_2 = \sqrt{x_1^2 + ... + x_n^2}$$

Достаточно очевидно, что эта функция дифференцируема везде кроме 0. Тогда:

$$\left\{\nabla f(x_0) = \frac{x_0}{||x_0||_2}, \ \forall x_0 \in \mathbb{R}^n \,/\, \{0\}\right\} \subset \partial f(x)$$

Поисследуем, точку $x_0 = 0$. Расммотрим субградиент g в этой точке: Из неравенства Гельдера слудет, что:

$$\forall y \in \mathbb{R}^n \ f(y) - f(0) \geqslant \langle g, y - 0\rangle$$
$$\forall y \in \mathbb{R}^n \ ||y||_2 \geqslant \langle g, y\rangle$$

Из неравенства Гельдера имеем:

$$\forall y \in \mathbb{R}^n \ \forall g \in \mathbb{R}^n \ ||y||_2 \cdot ||g||_2 \geqslant \langle g, y\rangle$$

Значит, очевидно $||g||_2 \leqslant 1$. Получаем:

$$\partial f(x) = \begin{cases} \dfrac{x}{||x||_2} & x \neq 0 \\ \forall g \ ||g||_2 \leqslant 1 & x = 0 \end{cases}.$$

- $p = \infty$

$$f(x) = \max_i \{|x_i|\} = \max\left\{g_k(x) = e_k^T x \ : \ e_k^k = \begin{bmatrix} 1 & x_k > 0 \\ -1 & x_k < 0 \\ \{-1, 1\} & x_k = 0 \end{bmatrix} \ \forall i \neq k \ e_k^i = 0\right\}$$

т.е. у векторов e_k кроме одной координаты все остальные нулевые. Очевидно, что все $g_k(x)$ выпуклы и дифференцируемы всюду. Значит:

$$\partial g_k(x) = e_k$$

$$
e_k^1 = 0,\ \dots, e_k^{k-1} = 0, e_k^k = \begin{cases} 1 & x_i > 0 \\ -1 & x_i < 0 \\ [-1, 1] & x_i = 0 \end{cases}, e_k^{k+1} = 0,\ \dots, e_k^n = 0
$$

$$
\partial f(x) = \mathbf{conv}\{\partial g_k(x) \mid g_k(x) = f(x)\} = \left\{ e \mid ||e||_\infty \leqslant 1 \land e^T x = ||x||_\infty \right\}.
$$

Ответ:

1. $$\partial f(x) = \left\{ e \mid e \in [-1, 1]^n \land e^T x = ||x||_1 \right\}$$

2. $$\partial f(x) = \begin{cases} \dfrac{x}{||x||_2} & x \neq 0 \\ \forall g \ ||g||_2 \leqslant 1 & x = 0 \end{cases}$$

3. $$\partial f(x) = \left\{ e \mid ||e||_\infty \leqslant 1 \land e^T x = ||x||_\infty \right\}$$

Задача 3.

Найти субдифференциал функции $f_p(x) = ||Ax - b||_p$ при $p = 1, 2, \infty$.

Решение:

Без ограничения общности, пусть $f_p : \mathbb{R}^n \to \mathbb{R}$.

Воспользуемся фактом из учебника: $f(x)$ - выпукла, и $h(x) = f(Ax + b)$, тогда

$$\partial h(x) = A^T \partial f(Ax + b).$$

Получается, можно воспользоваться результатами полученными из предыдущей задачи.

Пусть:

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix}, a_i \in \mathbb{R}^n.$$

- $p = 1$

$$f(Ax - b) = ||Ax - b||_1 = |a_1^T x - b_1| + ... + |a_n^T x - b_n| =$$
$$= \max\left\{ g_k(x) = e_k^T(Ax - b) \ : \ e_k^i \in \{-1, 1\} \ \forall i \in \overline{1n} \right\}.$$

т.е. в $f(Ax - b) = \max\{g_1(Ax - b), \dots, g_k(Ax - b)\} = e^T(Ax - b), \ e^i = 1$, если $a_i^T x_i > b_i$, и $e^i = \ -1$, если $a_i^T x_i < b_i$, и -1 либо 1 при $a_i^T x_i = b_i$. Т.к. $\forall k \ g_k(Ax - b) = e_k^T(Ax - b)$ дифференцирумы, то их субдифференциалы содержат только один субградиент:

$$\partial g_k(Ax - b) = e_k$$
$$e_k^i = \begin{cases} 1 & a_i^T x_i > b_i \\ -1 & a_i^T x_i < b_i \\ [-1, 1] & a_i^T x_i = b_i \end{cases}$$

Таким образом, выпуклая оболочка всех таких субградиентов описанных выше имеет вид:

$$\partial f(Ax - b) = \left\{ e \mid e \in [-1, 1]^n \wedge e^T(Ax - b) = ||Ax - b||_1 \right\} \Longrightarrow$$
$$\partial h(x) = A^T \partial f(Ax - b) = \left\{ A^T e \mid e \in [-1, 1]^n \wedge e^T(Ax - b) = ||Ax - b||_1 \right\}.$$

- $p = 2$

$$f(Ax - b) = ||Ax - b||_2 = \sqrt{\left(a_1^T x - b_1\right)^2 + ... + \left(a_n^T x - b_n\right)^2}$$

Достаточно очевидно, что эта функция дифференцируема везде кроме 0. Тогда:

$$\left\{ \nabla f(Ax_0 - b) = \frac{Ax_0 - b}{||Ax_0 - b||_2}, \ \forall (Ax_0 - b) \in \mathbb{R}^n \ / \ \{0\} \right\} \subset \partial f(Ax - b)$$

Поисследуем, точку $Ax_0 = b$. Рассмотрим субградиент g в этой точке: Из неравенства Гельдера слудет, что:

$$\forall y \in \mathbb{R}^n \ \ f(Ay - b) - f(0) \geqslant \langle g, Ay - b - 0 \rangle$$
$$\forall y \in \mathbb{R}^n \ ||Ay - b||_2 \geqslant \langle g, Ay - b \rangle$$

Из неравенства Гельдера имеем:

$$\forall y \in \mathbb{R}^n \ \forall g \in \mathbb{R}^n \ ||y||_2 \cdot ||g||_2 \geqslant \langle g, y \rangle$$

Значит, очевидно $||g||_2 \leqslant 1$. Получаем:

$$\partial f(Ax-b) = \begin{cases} \dfrac{Ax-b}{||Ax-b||_2} & Ax \neq b \\ \forall g\ ||g||_2 \leqslant 1 & Ax = b \end{cases} \Longrightarrow$$

$$\partial h(x) = A^T \partial f(Ax-b) = \begin{cases} \dfrac{A^TAx - A^Tb}{||Ax-b||_2} & Ax \neq b \\ \left\{ A^T g \mid ||g||_2 \leqslant 1 \right\} & Ax = b \end{cases}.$$

- $p = \infty$

$$f(Ax-b) = \max_i \left\{ |a_i^T x_i - b_i| \right\} = \max\left\{ g_k(Ax-b) = e_k^T(Ax-b) \right\}$$

$$e_k^k = \left[\begin{array}{ll} 1 & a_k^T x > b_i \\ -1 & a_k^T x < b_i \\ \{-1, 1\} & a_k^T x = b_i \end{array} \right. \quad \forall i \neq k\ e_k^i = 0$$

т.е. у векторов e_k кроме одной координаты все остальные нулевые. Очевидно, что все $g_k(Ax-b)$ выпуклы и дифференцируемы всюду. Значит:

$$\partial g_k(Ax-b) = e_k$$

$$e_k^1 = 0,\ \dots, e_k^{k-1} = 0, e_k^k = \begin{cases} 1 & a_k^T x > b_i \\ -1 & a_k^T x < b_i \\ [-1, 1] & a_k^T x = b_i \end{cases}, e_k^{k+1} = 0,\ \dots, e_k^n = 0$$

$$\partial f(Ax-b) = \mathbf{conv}\{\partial g_k(Ax-b) \mid g_k(Ax-b) = f(Ax-b)\} =$$

$$= \left\{ e \mid ||e||_\infty \leqslant 1 \wedge e^T(Ax-b) = ||Ax-b||_\infty \right\} \Longrightarrow$$

$$\partial h(x) = A^T \partial f(Ax-b) = \left\{ A^T e \mid ||e||_\infty \leqslant 1 \wedge e^T(Ax-b) = ||Ax-b||_\infty \right\}.$$

Ответ:

1. $$\partial ||Ax-b||_1 = \left\{ A^T e \mid e \in [-1, 1]^n \wedge e^T(Ax-b) = ||Ax-b||_1 \right\}$$

2. $$\partial ||Ax-b||_2 = \begin{cases} \dfrac{A^TAx - A^Tb}{||Ax-b||_2} & Ax \neq b \\ \left\{ A^T g \mid ||g||_2 \leqslant 1 \right\} & Ax = b \end{cases}$$

3. $$\partial ||Ax-b||_\infty = \left\{ A^T e \mid ||e||_\infty \leqslant 1 \wedge e^T(Ax-b) = ||Ax-b||_\infty \right\}$$

Задача 4.

Найти субдифференциал функции $f_p(x) = e^{||x||_p}$ при $p = 1, 2, \infty$.

Решение:

Будем пользоваться теоремой о субдифференциале сложной функции:

$$\partial e^{g(x_0)} = \frac{de^u}{du}(u_0) \cdot \partial g(x_0), u = g(x), u_0 = g(x_0) \Longrightarrow$$
$$\partial e^{||x||_p} = e^{||x||_p} \cdot \partial(||x||_p).$$

Ответ:

1. $$\partial e^{||x||_1} = \left\{ e^{||x||_1} g \mid g \in [-1, 1]^n \wedge g^T x = ||x||_1 \right\}$$

2. $$\partial e^{||x||_2} = \begin{cases} \dfrac{e^{||x||_2} x}{||x||_2} & x \neq 0 \\ \left\{ e^{||x||_2} g \mid ||g||_2 \leqslant 1 \right\} & x = 0 \end{cases}$$

3. $$\partial e^{||x||_\infty} = \left\{ e^{||x||_\infty} \cdot g \mid ||g||_\infty \leqslant 1 \wedge g^T x = ||x||_\infty \right\}$$

Задача 5.

Покажите, что x_0 - минимум выпуклой функции f тогда и только тогда, когда $\overline{0} \in \partial f(x_0)$.

Решение:

1) x_0 - минимум выпуклой функции f. Покажем, что $\overline{0}$ является субградиентом этой функции в точке x_0. Итак, раз это минимум, то $\forall y \ f(x) \geqslant f(x_0)$, значит:

$$\forall y \in \mathbf{dom} f \ : \ f(y) - f(x_0) \geqslant 0 = \langle \overline{0}, y - x_0 \rangle \Longrightarrow \overline{0} \in \partial f(x_0).$$

2) Покажем, что $\overline{0} \in \partial f(x_0) \Longrightarrow \forall y \in \mathbf{dom} f \ : \ f(y) \geqslant f(x_0)$:

$$\overline{0} \in \partial f(x_0) \Longrightarrow \forall y \in \mathbf{dom} f \ : \ f(y) - f(x_0) \geqslant \langle \overline{0}, y - x_0 \rangle = 0 \Longrightarrow f(y) \geqslant f(x_0). \qquad \square$$