Домашнее задание 6

Авласов Владислав

Задание 1

Постройте пример того, что 0-1 функция потерь может обладать локальным минимумом, не являющимся глобальным.

Функция 0-1 является ступенчатой функцией, т.е. нам достаточно взять отрезок на "ступени"со значением 1. Все точки на этом отрезке будут равны и будут являться локальными минимумами. Но все они больше, чем любые точки, принадлежащие "ступени"со значением 0.
Для примера возьмём выборку из одного вектора x=(0,1). И возьмём w=(0, -1). Он лежит левее нашей точки x и неправильно её классифицирует. В окрестности $\epsilon = 0.1$ все вектора w$'$ тоже неправильно классифицируют нашу точку, т.е. они локальные минимумы. Если же взять w*=(0,1), то теперь классификация происходит верно и ошибка равна 0, т.е. мы нашли глобальный минимум, отличающийся от локального.

Задание 2

Покажите, что задача одновременно и выпукло-липшицево-ограничена, и выпукло-гладко-ограничена.
Приведите соответствующие параметры липшицевости и гладкости.

Заданное множество является ограниченным, а функция логарифма log(1 + exp(x)) является выпуклой по критерию выпуклости, т.к. $f''(x) = \frac{e^x}{(e^x+1)^2} \geq 0$.
Тогда наша функция потерь тоже является выпуклой, по лемме о композиции выпуклой и линейной функций.
Липшицевость для log(1 + exp(x)) следует из того свойства, что $|f'(x)| = \frac{\exp(x)}{1+\exp(x)} = \frac{1}{\exp(-x)+1} \leq 1$.
Т.е. функция является 1-липшецовой. Воспользовавшись свойствами о композиции липшецевых функций и утверждением из лекции, что $f(w) = \langle w, v \rangle + b$ является $\|v\|$ -липшецовой, получим, что наша функция потерь является B-липшецовой.
Через взятие второй производной на лекции мы показали, что функция log(1 + exp(x)) является также 1/4-гладкой.
По лемме о композиции гладкой и линейной функции $f(w) = \log(1 + \exp(-y\langle w, x \rangle))$ является $(\|x\|^2/4)$-гладкой, т.е. для нашей задачи ограничим сверху и получим $B^2/4$-гладкость.
Т.е. наша задача является B-липшецовой и $B^2/4$-гладкой.
ЧТД.

Задание 3

Покажите, что эта функция потерь является также R-липшицевой.

Докажем по определению, т.е. покажем, что $\|f(w_1) - f(w_2)\| \leqslant R\|w_1 - w_2\|$. Зафиксируем некоторые точки (x, y) и выберем два вектора w_1,$w_2 \in \mathbb{R}^d$. И пусть l_1, l_2 будут значениями функции потерь в этих точках.
Т.е. хотим показать $|\ell_1 - \ell_2| \leq R\|w_1 - w_2\|$.
Переберём возможные варианты.
Если оба значения $y\langle \mathbf{w}_1, \mathbf{x}\rangle \geq 1$, $y\langle \mathbf{w}_2, \mathbf{x}\rangle \geq 1$, то $l_1 = l_2 = 0$ и получаем тривиальный ответ: $0 \leq R\|\mathbf{w}_1 - \mathbf{w}_2\|$.
Пусть хотя бы одно из произведений $y\langle \mathbf{w}_i, x\rangle < 1$. Пусть это будет $y\langle \mathbf{w}_1, \mathbf{x}\rangle < y\langle \mathbf{w}_2, \mathbf{x}\rangle$. Тогда раскроем модуль: $|\ell_1 - \ell_2| = \ell_1 - \ell_2$.

$$\begin{aligned}\ell_1 - \ell_2 = 1 - y\langle \mathrm{w}_1, \mathrm{x}\rangle - \max\{0, 1 - y\langle \mathrm{w}_2, \mathrm{x}\rangle\} \leq 1 - y\langle \mathrm{w}_1, \mathrm{x}\rangle - 1 + y\langle \mathrm{w}_2, \mathrm{x}\rangle = \\ = y\langle \mathrm{w}_2, \mathrm{x}\rangle - y\langle \mathrm{w}_1, \mathrm{x}\rangle = y\langle \mathrm{w}_2 - \mathrm{w}_1, \mathrm{x}\rangle \leq \|\mathrm{w}_1 - \mathrm{w}_2\|\|\mathbf{x}\| \leq R\|\mathrm{w}_1 - \mathrm{w}_2\|\end{aligned}$$

Таким образом, мы рассмотрели все варианты и доказали R-липшицевость функции потерь.
ЧТД.