
DSC 140A - Discussion 03

Problem 1.

A subgradient of the absolute loss is:

$$\begin{cases} \text{Aug}(\vec{x}), & \text{if } \text{Aug}(\vec{x}) \cdot \vec{w} - y > 0, \\ -\text{Aug}(\vec{x}), & \text{if } \text{Aug}(\vec{x}) \cdot \vec{w} - y < 0, \\ \vec{0}, & \text{otherwise.} \end{cases}$$

Suppose you are running subgradient descent to minimize the risk with respect to the absolute loss in order to train a function $H(x) = w_0 + w_1x$ on the following data set:

x	y
1	3
2	5
3	7

Suppose that the initial weight vector is $\vec{w} = (0, 0)^T$ and that the learning rate $\eta = 1$. What will be the weight vector after one iteration of subgradient descent?

Problem 2.

Consider the function $f(\vec{z}) = f(z_1, z_2) = \max(z_1, z_2)$.

- a) Using the definition of the subgradient, check if $(1, 1)^T$ is a subgradient at the point $(2, 2)^T$.
- b) Show that $(1, 0)^T$ is a valid subgradient at $(2, 2)^T$.
- c) Show that $(\frac{1}{2}, \frac{1}{2})^T$ is a subgradient at $(2, 2)^T$.

Problem 3.

The absolute loss of a linear predictor is

$$\ell_{\text{abs}}(\text{Aug}(x) \cdot \vec{w}, y) = |\text{Aug}(x) \cdot \vec{w} - y|.$$

We can write this as a piecewise function:

$$\ell_{\text{abs}}(\text{Aug}(x) \cdot \vec{w}, y) = \begin{cases} \text{Aug}(x) \cdot \vec{w} - y, & \text{if } \text{Aug}(x) \cdot \vec{w} - y > 0, \\ y - \text{Aug}(x) \cdot \vec{w}, & \text{if } \text{Aug}(x) \cdot \vec{w} - y < 0, \\ 0, & \text{if } \text{Aug}(x) \cdot \vec{w} = y. \end{cases}$$

This loss function is not differentiable at $\text{Aug}(x) \cdot \vec{w} = y$, but has a well-defined gradient everywhere else.

- a) What is the gradient of the absolute loss with respect to \vec{w} when $\text{Aug}(x) \cdot \vec{w} - y > 0$?
- b) What is the gradient of the absolute loss with respect to \vec{w} when $\text{Aug}(x) \cdot \vec{w} - y < 0$?
- c) **Optional:** Show that $\vec{0}$ is a subgradient of the absolute loss at $\text{Aug}(x) \cdot \vec{w} = y$.