DSC 140A - Homework 03

Due: Wednesday, January 31

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Unless otherwise noted by the problem's instructions, show your work or provide some justification for your answer. Homeworks are due via Gradescope at 11:59 PM.

Problem 1.

Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is convex and $g : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing. That is, if $a > b$, then $g(a) \geq g(b)$.

Show that the composition of these functions, $h(\vec{x}) = g(f(\vec{x}))$, is also convex.

Hint: you'll want to go back to the definition to show this is true. A similar problem was solved in discussion.

Problem 2.

Recall that the *hinge loss* is

$$L_{\text{hinge}}(\vec{w}, \vec{x}, y) = \max\{0, 1 - y\,\vec{w} \cdot \text{Aug}(\vec{x})\}$$

The Soft-SVM problem aims to minimize the regularized empirical risk:

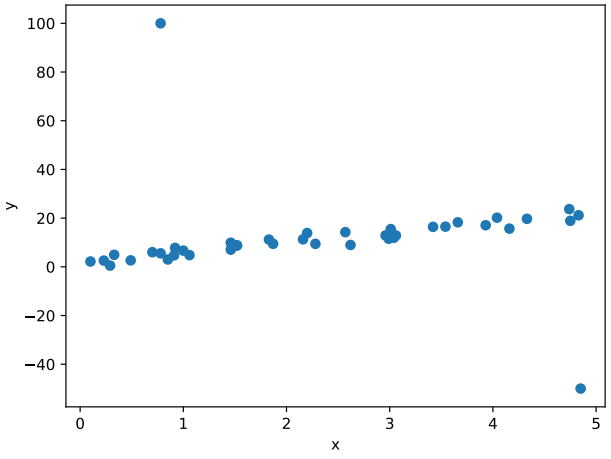
$$R(\vec{w}) = \frac{C}{n} \sum_{i=1}^{n} L_{\text{hinge}}(\vec{w}, \vec{x}^{(i)}, y_i) + \|\vec{w}\|^2$$

Show that $R(\vec{w})$ is a convex function of $\vec{w}$.

Hint: you will probably *not* want to use the formal definition of convexity here. Instead, you'll want to show that R is composed of simpler functions which themselves are convex.

Problem 3.

In the last homework, you fit a line to the data set below by minimizing the mean squared error:



Notably, the data contains outliers which may affect our regression.

Remember that you can download the data set at the link below:

`https://f000.backblazeb2.com/file/jeldridge-data/002-regression_outlier/data.csv`

a) Recall the absolute loss:

$$L_{\text{abs}}(H(x), y) = |H(x) - y|.$$

The absolute loss is not as sensitive to outliers as compared to the square loss. However, it is not differentiable, so unlike the mean squared error there is no closed form solution for the minimizer of the risk.

Implement subgradient descent and use it to fit a function of the form $w_0 + w_1 x$ by minimizing the risk with respect to the absolute loss. Report w_0 and w_1, and provide your code.

Hint: the subgradient of the absolute loss was considered in discussion.

b) Plot both of the fitted lines on top of a scatter plot of the data; that is, plot the least squares regression line you found in Homework 02 and the line fit by minimizing the risk of the absolute loss.

You should see that the two lines are quite different. Briefly explain why this is the case.