Math Homework 1

CS540

May 30, 2019

1 Instruction

Please submit your answers on Canvas → Assignments → M1. Late submission will not be accepted. Please add a file named "comments.txt", and in the first line of the file, grade yourself: $1, 1.5, 2$ (for the entire homework, not for individual questions). In your submission, please do not write your name if you do not want other students to see it (in the case it is posted as a sample solution).

Grade	Meaning
1	You attempted something but mostly incorrect.
1.5	You attempted something but there are mistakes.
2	You have the correct answers + permission to post as a sample solution.

2 Questions

2.1 Question 1

2017 May Final Exam Q3

Consider a LTU perceptron with initial weights $w = \begin{bmatrix} 0.2 \\ 0.7 \\ 0.9 \end{bmatrix}$ and bias $b = -0.7$. Given a new input $x_i = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $y_i = 0$. Let the learning rate by 0.2, compute the updated weights.

2.2 Question 2

Show that the following cost (objective) functions are equivalent if $a_i = \mathbb{1}_{\{w^T x_i + b \geqslant 0\}}$ and $y_i \in \{0, 1\}$.

Squared cost: $\dfrac{1}{2} \displaystyle\sum_{i=1}^{n} (a_i - y_i)^2$

Zero-one cost: $\displaystyle\sum_{i=1}^{n} \mathbb{1}_{\{a_i \neq y_i\}}$

Absolute value cost: $\displaystyle\sum_{i=1}^{n} |a_i - y_i|$

Hinge cost: $\displaystyle\sum_{i=1}^{n} \max\{0, 1 - (2 \cdot a_i - 1)(2 \cdot y_i - 1)\}$

2.3 Question 3

2015 Fall Midterm Exam Q5

Use gradient descent on $f(w) = w^4$ with learning rate 0.1.

Show that the algorithm converges to the correct global minimum when the initial guess is $w = 1$.

Show that the algorithm diverges when the initial guess is $w = 3$.

2.4 Question 4

2011 Fall Midterm Exam Q11

Consider a linear model (not LTU perceptron)$a_i = w^T x_i + b$ with squared error cost function $\sum_{i=1}^{n} (y_i - a_i)^2$.

Given a single training data, $x_1 = \begin{bmatrix} 1 \\ ... \\ 1 \end{bmatrix} \in \mathbb{R}^d, y_1 = 100$. Compute the gradient at $w = \begin{bmatrix} 3 \\ ... \\ 3 \end{bmatrix} \in \mathbb{R}^d$ and $b = 3$.

2.5 Question 5

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, show that the function $f(x) = \frac{1}{2} x^T A x, x \in \mathbb{R}^2$ is not convex: find the Hessian matrix and show that it is not positive semidefinite.