Machine Learning
V. Adamchik CSCI 567 Spring 2019
Discussion 4 University of Southern California

Gradient Descent

Perceptron

Logistic Regression

Problem 1

Why is the Hessian of logistic loss positive semidefinite.

Problem 2

Apply Newton's method to perceptron to minimize classification error.

$$F(w) = \sum_{n=1}^{N} \max(0, -y_n \mathbf{w}^T \mathbf{x}_n)$$

Problem 3

Which of the following surrogate losses is not an upper bound of the 0-1 loss?

(A) perceptron loss max{0, -z}
(B) hinge loss max{0, 1-z}
(C) logistic loss log(1 + exp(-z))
(D) exponential loss exp(-z)

Problem 4

The following table shows a binary classification training set and the number of times each point is misclassified during a run of the perceptron algorithm. Which of the following is the final output of the algorithm?

$\boldsymbol{x}$	y	Times misclassified
(-3, 2)	+1	5
(-1, 1)	-1	5
(5, 2)	+1	3
(2, 2)	-1	4
(1, -2)	+1	3

Problem 5

Suppose we obtain a hyperplane w via logistic regression and are going to make a randomized prediction on the label y of a new point x based on the sigmoid model. What is the probability of predicting y = +1?

$(a)\ \boldsymbol{e}^{-\boldsymbol{w}^T\boldsymbol{x}}$

$(b)\ \dfrac{1}{1+\boldsymbol{e}^{-\boldsymbol{w}^T\boldsymbol{x}}}$

$(c)\ \dfrac{1}{1+\boldsymbol{e}^{\boldsymbol{w}^T\boldsymbol{x}}}$

$(d)\ \mathbb{I}[\boldsymbol{w}^T\boldsymbol{x} \geq 0]$

Problem 6

Assume we have a training set (x_1, y_1), …, (x_N, y_N), the probability of seeing out come *y* is given by

$$P(y|x_n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y - w^T x_n)^2}{2\,\sigma^2})\right)$$

a) Assume σ is given, find the MLE for *w*

b) Assume σ is a parameter, find the MLE for σ.