

Homework 5

Qinyun Song

1 Logistic Regression

1.

$$g(w)' = -\frac{y_i x_i e^{-y_i w^T x_i}}{1 + e^{-y_i w^T x_i}} \quad (1)$$

2. From the problem we know that the function we aim to minimize is

$$J(w) = \sum_{i=1}^m \log(1 + e^{-y_i w^T x_i}) + \frac{1}{\sigma^2} w^T w \quad (2)$$

By using the SGD algorithm, for a simple example (x_i, y_i) , wthe gradient with respect to the weight vector is

$$J(w)' = \frac{-y_i x_i e^{-y_i w^T x_i}}{1 + e^{-y_i w^T x_i}} + \frac{2}{\sigma^2} w \quad (3)$$

Algorithm 1 Algorithm for SGD

3. 1: Define trainig set $S = \{(x_i, y_i)\}, x \in R^n$
2: Init $w^0 = 0 \in R^n$
3: **for** $t = 1, \dots, T$ **do**
4: Pick random example (x_i, y_i)
5: Treat the example as the whole training set with the object function

$$J(w) = \log(1 + e^{-y_i w^T x_i}) + \frac{1}{\sigma^2} w^T w \quad (4)$$

- 6: Update $w^t \leftarrow w^{t-1} - r^t \nabla J^t(w^{t-1})$ where r^t is the learning rate for this epoch
7: **end for**
8: **return** final w
-