Homework Assignment 2
Loss Functions and Support Vector Machines

Lecturer: Kyunghyun Cho

February 5, 2018

1. After replacing the label set from $\{0,1\}$ to $\{-1,1\}$, we introduced the log loss

$$D_{\log}(y,\mathbf{x};M) = \frac{1}{\log 2}\log(1+\exp(-s(y,\mathbf{x};M))),$$

as an alternative to the logistic regression distance function above. Show that these two are equivalent up to a constant multiplication for logistic regression.

2. Unlike the log loss, the hinge loss, defined below, is not differentiable everywhere:

$$D_{\text{hinge}}(y,\mathbf{x};M) = \max(0, 1 - s(y,\mathbf{x};M)).$$

Does it mean that we cannot use a gradient-based optimization algorithm for finding a solution that minimizes the hinge loss? If not, what can we do about it?

3. (Source: Koller) Recall that the formulation of SVM is

$$\frac{1}{N}\sum_{i=1}^{N}\max(0, 1 - y_i \mathbf{w}^T \mathbf{x}_i) + \lambda \|\mathbf{w}\|^2$$

Consider fitting an SVM with $\lambda > 0$ to a dataset that is linearly separable. Is the resulting decision boundary guaranteed to separate the classes? Please explain your answer.

4.