

# Homework Assignment 2

Lecturer: Kyunghyun Cho

January 23, 2017

1. 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?

2. When the distances to the nearest positive and negative examples are defined as  $d^+$  and  $d^-$ , the margin is

$$\gamma = \frac{1}{2}(d^+ + d^-).$$

Show that minimizing the norm of the weight vector of a support vector machine is equivalent to maximizing the margin.

## 3. PROGRAMMING ASSIGNMENT