Homework Assignment 3

Lecturer: Kyunghyun Cho

February 25, 2018

1. Cross-validation is a useful strategy for model selection, especially when the training data is small. However, it cannot be used for early-stopping (in other words, you cannot pick the best fold). Why is this the case?

2. In multiclass classification, given the definitions in the lecture notes, derive the following distance function. defined as

$$\begin{aligned} D(y^*, M, \mathbf{x}) &= -\log p_{M^*(\mathbf{x})} \\ &= -a_{y^*} + \log \sum_{k=1}^{K} \exp(a_k), \end{aligned}$$

3. Given the definition of the distance function above, derive a learning rule step-by-step for each column vector $\mathbf{w}_c$ of the weight matrix $\mathbf{W}$ (Equation 1.28 in the lecture notes).

4. Multiclass Classification on MNIST Please download `https://github.com/nyu-dl/Intro_to_ML_Lecture_Note/blob/master/homeworks/hw2.ipynb` and follow its instructions.