Exercise Sheet 8

Due: 23.01.2019, 09:00

Download the following files

- Train data: https://pjreddie.com/media/files/mnist_train.csv
- Test data: https://pjreddie.com/media/files/mnist_test.csv
- Jupyter notebook: SoftmaxRegression.ipynb from ISIS

The training and test datasets are csv-versions of the MNIST dataset: http://yann.lecun.com/exdb/mnist

Exercise 8.1 (1 point for each a) and b), 1 bonus point for c))

The goal of this exercise is to derive the softmax function

$$h_k(x) = P(y = k|x) = \frac{\exp\left(w_k^{\mathsf{T}} x\right)}{\sum_{l=1}^{c} \exp\left(w_l^{\mathsf{T}} x\right)} \quad , \; k = 1, \dots, c$$

by modeling the Posterior with a categorical distribution. For this, complete the sketch of the probabilistic interpretation of softmax regression. In particular,

a) Show that
   - $p_k(x) = \frac{\exp(w_k^T x)}{1+\sum_{l=1}^{c-1} \exp(w_l^T x)}, \; k = 1, \dots, c-1,$
   - $p_c(x) = \frac{1}{1+\sum_{l=1}^{c-1} \exp(w_l^T x)}$

b) Find $w_1', \dots, w_c'$ such that $p_k(x) = \frac{\exp(w_k'^T x)}{\sum_{l=1}^{c} \exp(w_l'^T x)}$, for all $k = 1, \dots, c$ and all $x \in \mathbb{R}^{n+1}$.

c) Present the whole probabilistic derivation.

Exercise 8.2

Implement softmax regression in matrix formulation. The fit method should return a training loss curve 'LOSS_TRAIN' and an evaluation loss curve 'LOSS_EVAL', both containing loss-values after each k iterations. For this, split the given training data into two splits, train and eval.

Exercise 8.3

Train a softmax regression model on the whole MNIST train dataset (all 10 classes). Plot the train and eval loss curves. Evaluate the model on the test set. Plot each of the learned weight vectors $w_1, \dots, w_c$ as image. Discuss your results.

Using Scikit-Learn

Scikit-Learn should only be used if indicated by the task. You can use Scikit-Learn in another task, A say, in order to be able to solve further tasks building upon task A. But then you'll not receive a point in task A.