Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Virginia Estellers
Exercises: Emanuel Laude
Winter Semester 2017/18

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 7

Room: 02.09.023
Friday, 15.12.2017, 09:15-11:00
Extended Submission deadline: Wednesday, 13.12.2017, 11:59 p.m.

Coding: Consensus Primal-Dual for Sparse SVMs (24 Points)

Exercise 1 (24 Points). In this exercise sheet you are asked to implement the consensus Primal-Dual (PDHG) for sparse binary SVM training, that you have derived last week. The problem is phrased as an optimization problem of the form:

$$\min_{w \in \mathbb{R}^d, b \in \mathbb{R}} \sum_{i=1}^{N} \ell(w, b; x_i, y_i) + \lambda \|w\|_1, \tag{1}$$

where $\ell(\cdot, \cdot; x_i, y_i)$ is the hinge loss defined according to

$$\ell(w, b; x_i, y_i) := \max\{0, 1 - (\langle x_i, w \rangle + b) y_i\}. \tag{2}$$

More precisely, you are asked to do the following:

1. Implement a MATLAB function, that takes as an input argument the feature matrix $X \in \mathbb{R}^{N \times d}$ and a vector $y \in \{-1, 1\}^N$ of binary class labels and returns the classifier $(w, b) \in \mathbb{R}^{d+1}$. Here, N is the number of training examples and d is the feature dimension.

2. Use the code template from the logistic regression task and classify the toy data. Since the SVM-model is suited for two classes only, you are asked to train for each class an individual classifier in a one-vs-rest fashion. More precisely, you have to call the function c times for each class $1 \leq i \leq c$, where all examples belonging to class i are labeled as 1 and the rest is labeled as -1.

3. Stack the individual classifiers into a classifier matrix $W \in \mathbb{R}^{d \times c}$, $B \in \mathbb{R}^c$ and use the logistic regression code template to visualize the classifier.