TECHNISCHE UNIVERSITÄT MÜNCHEN

Zentrum Mathematik

Prof. Dr. M. Wolf
J. Cuesta

Mathematical Foundations of Machine Learning

http://www-m5.ma.tum.de/Allgemeines/MA4801_2018S

Summer 2018
Blatt 11
(27.06.2018)

The following homework will be discussed on July 3rd and July 6th:

H11.1. **(Computational Graph)** Assume $f \in \mathcal{C}^2(\mathbb{R}^k)$ has a computational graph with $o(k)$ vertices. Sketch an argument so that for any $v, w \in \mathbb{R}^k$ and $i \in \{1, \dots, k\}$ the function

$$\sum_j^k \partial_i \partial_j f(w) v_j$$

can be computed with $o(k)$ steps.

H11.2. (**Linear Programming and Classification**) Linear programs are optimization problems that can be expressed as maximizing a linear function subject to linear inequalities, i.e.

$$\begin{aligned} \max_{w \in \mathbb{R}^d} \quad & \langle w, u \rangle \\ \text{subject to} \quad & Aw \geq v, \end{aligned}$$

where $A \in \mathbb{R}^{m \times d}$ and $v \in \mathbb{R}^m$. Let $S = (x_i, y_i)_{i=1}^n \in \left(\mathbb{R}^d \times \{-1, 1\}\right)$ be a training data set of size n. Assume that there is a linear half-space $\mathcal{H}$ (i.e. a half-space defined by a hyperplane that goes through the origin) such that $x_i \in \mathcal{H}$ if and only if $y_i = 1$. Show that finding such a $\mathcal{H}$ can be formulated as a linear program.