MA2823: Introduction to Machine Learning **Homework 1**

Due October 06, 2017.

Turn in your homework
— as a PDF file
— named HW01_<LastName>_<FirstName>.pdf (no accents)
— at http://tinyurl.com/ma2823-2017-hw

1 Homework Problem

Question 1

Are the following problems machine learning problems? If yes, of what type? What are the design matrix and, if appropriate, the target vector?

(a) Given a car's owner manual and the price of gas at the nearest filling station, predict how much it will cost to fill up the car's tank.

Solution: The price (in €) will be the volume of the tank (in L), which can be found in the car's owner manual, multiplied by the price (in €/L). There is nothing to learn here!

(b) Compute a house's heating load[1] from its building plan and a civil engineer's records of plans and heating loads for hourses in the same neighborhood.

Solution: We do not know the formula to compute heating load from a house's plan. We can formulate this problem as a supervised (regression) machine learning problem. The design matrix X will be a $n \times p$ matrix where n is the number of houses in our records and p a number of characteristics we have extracted from the plans (such as surface area, windows surface area, wall surface area, etc.). The target vector y is a n-dimensional vector of heating loads.

Question 2

Are the following problems convex optimization problems? What technique could you apply to solve them?

(a)

$$\min_{\boldsymbol{w} \in \mathbb{R}^p} \sum_{i=1}^{n} \log\left(1 + e^{-y^i . \boldsymbol{x}^{i\top} \boldsymbol{w}}\right)$$

where $\boldsymbol{y} = (y^1, y^2, \dots, y^n)^\top \in \{-1, 1\}^n$ is a n-dimensional binary vector, and $X \in \mathbb{R}^{n \times p}$.

---

[1] i.e. the amount of energy that is needed to maintain the temperature.

Solution: $f : \mathbb{R}^p \to \mathbb{R}, \boldsymbol{w} \mapsto \sum_{i=1}^n \log\left(1 + e^{-y^i . \boldsymbol{x}^{i\top} \boldsymbol{w}}\right)$ is a convex function. You can prove this using the composition rules, and the fact that $u \mapsto \log\left(1 + e^{\alpha u}\right)$ is convex (independently from the sign of α) can be proven because it is twice differentiable and its second derivative is positive.
This is an unconstrained convex minimization problem. It can be solved by gradient descent and its variants.

(b)

$$\min_{\boldsymbol{w} \in \mathbb{R}^p} \sum_{j=1}^p w_j^2 \quad \text{subject to } w_j^3 \leq 0 \text{ for all } 0 \leq j \leq (p-1).$$

Solution: $\sum_{j=1}^p w_j^2$ is convex but w_j^3 is not convex. This is not a convex optimization problem on $\mathbb{R}^p$. However, $w_j^3 \leq 0$ can be replaced with the equivalent convex (indeed, affine) constraint $w \leq 0$. It also has a trivial solution $w_0 = w_1 = \cdots = w_{p-1} = 0$.