Machine Learning I: Foundations
Exercise Sheet 1

Prof. Marius Kloft TA: Billy Joe Franks

04.05.2021
Deadline: 11.05.2021

1) (MANDATORY) 10 Points

In this question we will be exploring eigenvectors and eigenvalues. Let $A \in \mathbb{R}^{d \times d}$. Recall the following definition from linear algebra: a vector $\mathbf{v} \in \mathbb{R}^d$ is an *eigenvector* of A if and only if there exists $\lambda \in \mathbb{R}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. We then call λ the *eigenvalue* of A associated with the vector $\mathbf{v}$. Note that, if $\mathbf{v}$ is an eigenvector of A, then, for any $a \in \mathbb{R}$, $a\mathbf{v}$ is also an eigenvector of A. Therefore, $\mathbf{v}$ and $a\mathbf{v}$ are not considered 'distinct' eigenvectors. Prove the following:

Proposition 1 *If A has a finite number of distinct eigenvectors then each eigenvector must have a unique eigenvalue.*

2) (MANDATORY) 10 Points

Recall the following definition from linear Algebra: a symmetric matrix $A \in \mathbb{R}^{d \times d}$ is called positive definite, if $\mathbf{x}^\top A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^d$ with $\mathbf{x} \neq \mathbf{0}$. Let A be symmetric.

a) Prove that, if all eigenvalues of A are positive, then A is positive definite.

b) Prove that, all eigenvalues of A are positive, if A is positive definite.

Now let $F : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & x^2 + 2y^2 + 4.97. \end{array}$

c) Find the critical point of F.

d) Compute the Hessian matrix H of F in any point $(x, y)^\top \in \mathbb{R}^2$.

e) Recall from multivariate calculus that, if H is positive definite in a critical point $(x, y)^\top$, then $(x, y)^\top$ is a local minimum. Show that the critical point of F is a local minimum. **Hint:** note that H is symmetric.

3) Find the global minima of the following functions $f : \mathbb{R} \to \mathbb{R}$ and $g, h, i : \mathbb{R}^2 \to \mathbb{R}$.

a) $f(w) := aw^2 + bw + c$

b) $g(\mathbf{w}) := \mathbf{w}^T A\mathbf{w} + \mathbf{b}^T\mathbf{w} + c$

c) $h(\mathbf{w}) := aw_1^2 + bw_2 + c$

d) $i(\mathbf{w}) := w_1^2 + w_2^2 + w_1^2 w_2$

Hint: Compute the gradient of the above functions and set it to zero. Do not forget to check the necessary conditions on global minima.

4) For a matrix $X \in \mathbb{R}^{m\times n}$ let $X_{i,:} = [X_{i,1}, \dots, X_{i,n}]$ be the i-th row vector and $X_{:,i} = [X_{1,i}, \dots, X_{m,i}]^T$ be the i-th column vector. For $X \in \mathbb{R}^{m\times n}$ and $Y \in \mathbb{R}^{n\times q}$ show that

$$XY = [X_{i,:}Y_{:,j}]_{i,j}$$

and

$$XY = \sum_{i=1}^{n} X_{:,i}Y_{i,:}.$$

Be sure to note the orientations of the vectors, some of these are row vectors and others are column vectors.