Optimization and Computational Linear Algebra for Data Science
Homework 9: Convex functions

Due on November 19, 2019

- Unless otherwise stated, all answers must be mathematically justified.
- Partial answers will be graded.
- You can work in groups but each student must write his/her own solution based on his/her own understanding of the problem. Please list on your submission the students you work with for the homework (this will not affect your grade).
- Problems with a ($\star$) are extra credit, they will not (directly) contribute to your score of this homework. However, for every 4 extra credit questions successfully answered your lowest homework score get replaced by a perfect score.
- If you have any questions, feel free to contact me (`lm4271@nyu.edu`) or to stop at the office hours.

Problem 9.1 (2 points)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. We assume that the minimum $m \stackrel{\text{def}}{=} \min_{x \in \mathbb{R}^n} f(x)$ of f on $\mathbb{R}^n$ is finite, and that the set of minimizers of f*

$$\mathcal{M} \stackrel{\text{def}}{=} \{v \in \mathbb{R}^n \,|\, f(v) = m\}$$

is non-empty.

(a) *Show that $\mathcal{M}$ is a convex set.*

(b) *Show that if f is strictly convex, then $\mathcal{M}$ has only one element.*

Problem 9.2 (2 points)**.** *Let $M \in \mathbb{R}^{n \times n}$ be a symmetric matrix, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. For $x \in \mathbb{R}^n$ we define*

$$f(x) = x^{\mathsf{T}} M x + \langle x, b \rangle + c.$$

f is called a quadratic function.

(a) *Compute the gradient $\nabla f(x)$ and the Hessian $H_f(x)$ at all $x \in \mathbb{R}^m$. Show that f is convex if and only if M is positive semi-definite.*

(b) *In this question, we assume M to be positive semi-definite. Show that f admits a minimizer if and only if $b \in \mathrm{Im}(M)$.*

Problem 9.3 (3 points)**.** *We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex if there exists $\alpha > 0$ such that the function $x \mapsto f(x) - \frac{\alpha}{2}\|x\|^2$ is convex. In other words, f is strongly convex if there exists $\alpha > 0$ and a convex function $g : \mathbb{R}^n \to \mathbb{R}$ such that*

$$f(x) = g(x) + \frac{\alpha}{2}\|x\|^2.$$

- **(a)** *Show that a strongly convex function is strictly convex. (Hint: start by showing that $x \mapsto \|x\|^2$ is strictly convex).*
- **(b)** *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. Show that φ is strongly convex if and only if there exists $\alpha > 0$ such that for all $x \in \mathbb{R}^n$ the eigenvalues of $H_\varphi(x)$ are greater or equal than α.*

Problem 9.4 (3 points)**.** *Let $A \in \mathbb{R}^{n \times m}$ and $y \in \mathbb{R}^n$. For $x \in \mathbb{R}^m$ we define*

$$f(x) = \|Ax - y\|^2.$$

- **(a)** *Compute the gradient $\nabla f(x)$ and the Hessian $H_f(x)$ at all $x \in \mathbb{R}^m$. Show that f is convex.*
- **(b)** *Show that if* $\operatorname{rank}(A) < m$*, then f is not strictly convex.*
- **(c)** *Show that is* $\operatorname{rank}(A) = m$*, then f is strongly convex (use the definition and results of Problem 9.3).*

Problem 9.5 ($\star$)**.** ***Notation:*** *For a symmetric matrix $M \in \mathbb{R}^{n \times n}$ we denote respectively $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ the smallest and largest eigenvalue of M.*

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function that is twice continuously differentiable. We assume that

$$\gamma \stackrel{\text{def}}{=} \inf_{x \in \mathbb{R}^n} \lambda_{\min}(H_f(x)) \qquad \textit{and} \qquad L \stackrel{\text{def}}{=} \sup_{x \in \mathbb{R}^n} \lambda_{\max}(H_f(x))$$

are both finite. Show that for all $x, h \in \mathbb{R}^n$:

$$f(x) + \langle \nabla f(x), h \rangle + \frac{\gamma}{2}\|h\|^2 \le f(x+h) \le f(x) + \langle \nabla f(x), h \rangle + \frac{L}{2}\|h\|^2.$$