AM 221: Advanced Optimization	Spring 2016
Prof. Yaron Singer	*Problem Set 1 — Due Wed, Feb. 3rd at 23:59*

Instructions: All your solutions should be prepared in LaTeX and the PDF and .tex should be submitted to Canvas. For each question, the best and correct answers will be selected as sample solutions for the entire class to enjoy. If you prefer that we do not use your solutions, please indicate this clearly on the first page of your assignment.

The programming parts can be written in the programming language of your choice and the code should be submitted alongside your solutions.

1. Sequences, Limits, Functions

a. Remember that the field $\mathbb{R}$ is characterized (among ordered fields containing $\mathbb{Q}$) by the least upper bound property: every non-empty bounded set has a least upper bound. Use this property to show that any non-decreasing upper-bounded sequence of real numbers is convergent.

b. Let $u = (u_n)_{n \geq 0}$ and $v = (v_n)_{n \geq 0}$ be two sequences of real numbers such that:

- u is non-increasing and v is non-decreasing
- $\lim_{n\to\infty} u_n - v_n = 0$

Show that both u and v are convergent and that they have the same limit.

c. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $f(0) > 0$. Show that there exists $\varepsilon > 0$ such that:

$$|x| < \varepsilon \Rightarrow f(x) > 0.$$

2. Linear Algebra

A is a matrix in $\mathbb{R}^{m\times n}$ with $m \leq n$.

a. Give the definition of the rank of A. What is the largest possible rank of A?

b. Let us denote by $\mathbf{a}_1, \dots, \mathbf{a}_m$ the rows of A, *i.e* $\mathbf{a}_i \in \mathbb{R}^n$ and $A = [\mathbf{a}_1 \ \dots \ \mathbf{a}_m]^\intercal$. Show that:

$$\text{rank}(A) < m \Leftrightarrow \exists \mathbf{x} \in \mathbb{R}^m \setminus \{0\}, \ \sum_{i=1}^{m} \mathbf{a}_i x_i = 0$$

3. Inner product, norm For $\mathbf{x}$ and $\mathbf{y}$ two vectors of $\mathbb{R}^d$, we write $\mathbf{x}^\intercal\mathbf{y} = \sum_{i=1}^d x_i y_i$ their inner product. $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\intercal\mathbf{x}}$ denotes the Euclidean norm of $\mathbf{x}$.

a. Let $\mathbf{a} \in \mathbb{R}^d$. Show that:

$$\mathbf{a} = 0 \iff \forall \mathbf{x} \in \mathbb{R}^d,\ \mathbf{a}^\intercal\mathbf{x} = 0$$
$$\mathbf{a} \geq 0 \iff \forall \mathbf{x} \geq 0,\ \mathbf{a}^\intercal\mathbf{x} \geq 0$$

b. Let $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{R}^d$. Show that:

$$\mathbf{x}^\intercal\mathbf{y} = \frac{\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2}{2} = \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2}{2}$$

c. Deduce the parallelogram law:

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$$

d. Let us denote by $B_2(0,1)$ the unit ball of $\mathbb{R}^d$, $B_2(0,1) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x}\| \leq 1\}$ and let us consider $\mathbf{v} \in \mathbb{R}^d$. Show that:

$$\max_{\mathbf{x} \in B_2(0,1)} \mathbf{v}^\intercal\mathbf{x} = \|\mathbf{v}\|$$

4. Multivariate Calculus

a. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function twice. For $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{d} \in \mathbb{R}^n$, we define the function $f_{\mathbf{x},\mathbf{d}} : \mathbb{R} \to \mathbb{R}$ by:

$$f_{\mathbf{x},\mathbf{d}}(\lambda) = f(\mathbf{x} + \lambda\mathbf{d})$$

Express the first and second derivative of $f_{\mathbf{x},\mathbf{d}}$ in terms of the gradient and Hessian of f.

b. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and let $\mathbf{x}$ be a local minimum of f, *i.e* there exists $\varepsilon > 0$ such that:

$$\|\mathbf{y} - \mathbf{x}\| \leq \varepsilon \Rightarrow f(\mathbf{y}) \geq f(\mathbf{x})$$

show that $\nabla f(\mathbf{x}) = 0$. **Hint:** remember the Taylor expansion of f at $\mathbf{x}$:

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{h}^\intercal\nabla f(\mathbf{x}) + o(\|\mathbf{h}\|)$$

c. Let $\mathbf{a} \in \mathbb{R}^d$ and $M \in \mathbb{R}^{d\times d}$. What are the gradients of $f(\mathbf{x}) = \mathbf{a}^\intercal\mathbf{x}$, $g(\mathbf{x}) = \|\mathbf{x}\|^2$ and $h(\mathbf{x}) = \mathbf{x}^\intercal M\mathbf{x}$?

5. Programming Download the file at `http://thibaut.horel.org/access.log`. This file is a server log, each line has the following format:

```
<time>\t<ip-adress>
```

i.e it contains a time and the IP address which accessed the server at that time; the time and the IP address are separated by a tab character. Using the programming language of your choice, write a program to find the list of the ten IP addresses who accessed the server the most (in decreasing order). Report both the list you obtained and the code you used.