Exercise #1

January 17, 2023

Problem 1.

a) Prove that the real-valued function $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is lower semi-continuous at $\overline{x} \in \mathbb{R}^n$ if and only if for any $\lambda < f(\overline{x})$ there exists $\delta > 0$ such that $\lambda < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$, where $\mathbb{B}(\overline{x}, \delta)$ is the open ball with center at $\overline{x}$ and radius δ.

b) Prove that the real-valued function $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is lower semi-continuous at $\overline{x}$ if and only if for any $\epsilon > 0$ there exists $\delta > 0$ such that $f(\overline{x}) - \epsilon < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$. (**Hint:** use the condition given in a).)

c) Prove that the real-valued function $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is lower semi-continuous in $\mathbb{R}^n$ if and only if for any $\lambda \in \mathbb{R}$ the set $L_\lambda = \{x \in \mathbb{R}^n : f(x) > \lambda\}$ is open. (**Hint:** use the condition given in a) or b).)

d) Prove that the real-valued function $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is lower semi-continuous in $\mathbb{R}^n$ if and only if epi f is closed, where epi f is the epigraph of the function f, defined as epi $f = \{(x, p) \in \mathbb{R}^n \times \mathbb{R} : p \geq f(x)\}$. (**Hint:** use the conditions given in a) and c).)

Solution.

a) Assume that the function f is lower semi-continuous at $\overline{x}$. In order to prove our assertion *a*) of Problem 1, we suppose to the contrary that there exists $\lambda < f(\overline{x})$ such that for any $\delta > 0$ there exists $y_\delta \in \mathbb{B}(\overline{x}, \delta)$ with $\lambda \geq f(y_\delta)$. Thus, by setting $\delta = \frac{1}{k}$ we obtain a sequence a sequence $x_k \in \mathbb{B}(\overline{x}, \frac{1}{k})$ such that $\lambda \geq f(x_k)$. Therefore, $x_k \longrightarrow \overline{x}$ (i.e., x_k converges to $\overline{x}$) and

$$\liminf_{k\to\infty} f(x_k) \leq \lambda < f(\overline{x}).$$

Thus, we obtain

$$\liminf_{k\to\infty} f(x_k) \leq f(\overline{x}),$$

which yields a contradiction to the fact that f is lower semi-continuous function at $\overline{x}$. Consequently, the assertion *a*) of Problem 1 is validated.

Conversely, we suppose that assertion *a*) of Problem 1 is true. Let $\{x_k\}_{k\in\mathbb{N}}$ be any sequence in $\mathbb{R}^n$ such that $x_k \longrightarrow \overline{x}$. Further, we consider for arbitrary $\lambda < f(\overline{x})$ there exists $\delta > 0$ such that $\lambda < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$.

Since $x_k \longrightarrow \overline{x}$, there exists $n \in \mathbb{N}$ such that $x_k \in \mathbb{B}(\overline{x}, \delta)$ for all $k \geq n$. Consequently, $\lambda < f(x_k)$ for all $k \geq n$, which yields

$$\lambda \leq \liminf_{k\to\infty} f(x_k),$$

and since this is true for all $\lambda < f(\overline{x})$, we have that

$$f(\overline{x}) \leq \liminf_{k\to\infty} f(x_k).$$

Therefore, f is lower semi-continuous function at $\overline{x}$.

b) Let f be the lower semi-continuous function at $\overline{x}$. We consider for arbitrary fixed $\epsilon > 0$, $\lambda = f(\overline{x}) - \epsilon < f(\overline{x})$. Further, lower semi-continuity of the function f implies the assertion a) of the Problem 1. It follows that there exists $\delta > 0$ such that $\lambda = f(\overline{x}) - \epsilon < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$. Since this is true for any $\epsilon > 0$, assertion b) holds.

Conversely, we assume that the claim b) of Problem 1 is true. Therefore, for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$f(\overline{x}) - \epsilon < f(x), \ \forall\, x \in \mathbb{B}(\overline{x}, \delta). \tag{1}$$

Let the arbitrary $\lambda < f(\overline{x})$. Now, we can consider a $\overline{\epsilon} > 0$ such that

$$\lambda < f(\overline{x}) - \overline{\epsilon}. \tag{2}$$

By using inequality (1) in (2), we can write there exists $\delta > 0$ such that

$$f(\overline{x}) - \overline{\epsilon} < f(x), \ \forall\, x \in \mathbb{B}(\overline{x}, \delta). \tag{3}$$

By inequalities (2) and (3), we can write for any $\lambda < f(\overline{x})$ there exists $\delta > 0$ such that $\lambda < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$. Now, the assertion a) implies that the function f is lower semi-continuous at $\overline{x}$.

c) Assume that the function f is lower semi-continuous and fix $\lambda \in \mathbb{R}$. Let $y \in L_\lambda$ be arbitrary, it follows that $\lambda < f(y)$. Since f is lower semi-continuous function, by the assertion a) we can write there exists $\delta > 0$ such that $\lambda < f(x)$ for all $x \in \mathbb{B}(y, \delta)$. Therefore, $x \in L_\lambda$ and then $\mathbb{B}(y, \delta) \subset L_\lambda$. Moreover, this is true for any $y \in L_\lambda$. That means for any $y \in L_\lambda$ there exists an open ball $\mathbb{B}(y, \delta)$ centered at y with radius δ, such that $\mathbb{B}(y, \delta) \subset L_\lambda$. Consequently, L_λ is open set for any $\lambda \in \mathbb{R}$.

Conversely, we suppose that assertion c) holds. Let arbitrary $\lambda < f(\overline{x})$ for any $\overline{x} \in \mathbb{R}^n$. It follows that $\overline{x} \in L_\lambda$. Since L_λ is open set, there exists $\delta > 0$ such that

$$B(\overline{x}, \delta) \subset L_\lambda. \tag{4}$$

Let $x \in B(\overline{x}, \delta)$ be arbitrary. Then, inequality (4) implies $x \in L_\lambda$. It follows that $\lambda < f(x)$. Since x is arbitrary, we can write that $\lambda < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$. Now, the assertion a) implies f is lower semi-continuity function at $\overline{x}$, moreover $\overline{x} \in \mathbb{R}^n$ is arbitrary that turns into lower semi-continuity of the function f in $\mathbb{R}^n$.

d) Let f be lower semi-continuous function at any $\overline{x} \in \mathbb{R}^n$. Now we shall show that epi f is closed by showing that $(\text{epi } f)^c$ (complement of epi f) is open. Take any $(\overline{x}, \lambda) \in (\text{epi } f)^c$, which yields $\lambda < f(\overline{x})$. Choose a small enough $\epsilon > 0$ such that $\lambda + \epsilon < f(\overline{x})$. Since f is lower semi-continuous function at $\overline{x}$, the assertion a) implies that there exists $\delta > 0$ such that $\lambda + \epsilon < f(x)$ for all $x \in \mathbb{B}(\overline{x}, \delta)$. Then it can be easily seen that $\mathbb{B}(\overline{x}, \delta) \times (\lambda - \epsilon, \lambda + \epsilon) \subset (\text{epi } f)^c$. Since such an open ball exists for any $(\overline{x}, \lambda) \in (\text{epi } f)^c$ which gives that $(\text{epi } f)^c$ is open. Consequently, epi f is closed.

Conversely, let epi f be closed. Taking any $\lambda \in \mathbb{R}$, we shall show that the set $A = \{x \in \mathbb{R}^n : f(x) \leq \lambda\}$ is closed. Then assertion c) implies that the function f is lower semi-continuous in $\mathbb{R}^n$. To this end, let $\{x_k\}_{k \in \mathbb{N}}$ be a sequence in the set A that converges to $\overline{x} \in \mathbb{R}^n$. Since $x_k \in A$, $f(x_k) \leq \lambda$. Therefore, $(x^k, \lambda) \in \text{epi } f$, and we also have $(x_k, \lambda) \longrightarrow (\overline{x}, \lambda)$. Moreover, epi f is closed. Consequently, $(\overline{x}, \lambda) \in \text{epi } f$, which implies $\lambda \geq f(\overline{x})$. Thus, $\overline{x} \in A$. Eventually, we get that A is closed, as asserted.

Problem 2.

Check the properties of lower semi-continuity, coercivity and existence of a global minimizer for the following functions:

a) $\ell\colon \mathbb{R} \mapsto \mathbb{R}$ defined as $\ell(x) = 5x^{10} + 8x^7 - 9x^2 + x + c$, where $c \in \mathbb{R}$ is a constant.

b) $m\colon \mathbb{R} \mapsto \mathbb{R}$ defined as $m(x) = e^x - \frac{1}{1+x^2}$.

c) $p\colon \mathbb{R} \mapsto \mathbb{R}$ defined as $p(x) = x^4 - 20x^3 + \sup\limits_{k \in \mathbb{N}} \sin(k^2 x)$.

d) $q\colon \mathbb{R}^2 \mapsto \mathbb{R}$ defined as $q(x) = x_1^2(1 + x_2^3) + x_1^2$.

Solution.

a) The given function l is a polynomial and hence continuous. Therefore, l is the lower semi-continuous function. Moreover, it is coercive too because of the term x^{10}, as it dominates all other terms for large $|x|$. Consequently, a global minimizer exists (by the week 2 lecture note).

b) The given function m is continuous. Therefore, it is lower semi-continuous too. Moreover, it is not coercive because $\lim_{x\to-\infty} g(x) = 0 \neq +\infty$. On the other hand, the closed set $\Omega = \{x \in \mathbb{R} : g(x) \leq g(-1)\}$ is bounded (and therefore compact), since $g(-1) < 0 = \lim_{x\to-\infty} g(x)$ and $g(-1) < 0 < +\infty = \lim_{x\to+\infty} g(x)$. Therefore, the global minimum of g is attained over Ω.

c) Since $x^4 - 20x^3$ is a polynomial, $x^4 - 20x^3$ is continuous. Further $\sin(k^2x)$ is continuous, implies it's semi-continuity too, and then $\sup_{k\in\mathbb{N}} \sin(k^2x)$ is semi-continuous (by the week 2 lecture). Now, $p(x)$ is the sum of lower semi-continuous function and hence lower semi continuous. The function $p(x)$ is coercive, because of x^4, as it dominates all other terms for large $|x|$. Consequently, a global minimizer exists (by the week 2 lecture note).

d) The given function $q(x)$ is a polynomial and hence continuous, which implies its lower semi-continuity. Moreover, it is not coercive because for any fixed $x_1 \neq 0$ and $x_2 \longrightarrow -\infty$ we get $q(x) \longrightarrow -\infty$. Therefore, it does not attain global minimizers on $\mathbb{R}^2$.

Problem 3.

Find the gradient, Hessian, and local minimizers of the objective function f of the optimization problem $\min_{x,y} f(x, y)$, where $f: \mathbb{R}^2 \mapsto \mathbb{R}$ is defined as:

a) $f(x, y) = \frac{x^2}{2} + x\cos y$.

b) $f(x, y) = 2x^2 - 4xy + y^4 + 5y^2 - 10y$.

Solution.

a) (Gradient of f) $\nabla f(x, y) = (x + \cos y, -x\sin y)^T$.
(Hessian of f) $\nabla^2 f(x, y) = \begin{bmatrix} 1 & -\sin y \\ -\sin y & -x\cos y \end{bmatrix}$. The first-order necessary condition of optimality gives $\nabla f(x, y) = 0$. Now, we have two sets of critical points;
1. $x = 0$ and $y = (2n + 1)\frac{\pi}{2}$, for $n = 0, \pm1, \pm2, \pm3, \ldots$. In this case the Hessian matrix is

$$\nabla^2 f(x, y) = \begin{bmatrix} 1 & (-1)^{n+1} \\ (-1)^{n+1} & 0 \end{bmatrix}.$$

It is evident that the Hessian matrix is symmetric and it's eigen values are $\frac{1}{2}(1 \pm \sqrt{5})$, in which the value of one eigen value is negative. Therefore, the above mentioned critical points are not local minimum.
2. $x = (-1)^{n+1}$ and $y = n\pi$ for $n = 0, \pm1, \pm2, \pm3, \ldots$. In this case the Hessian matrix is

$$\nabla^2 f(x, y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

which is clearly positive definite matrix and hence the above mentioned critical points are strict local minimizers.
Note: Since Hessian matrix is positive definite, the function f is strictly convex. Therefore, the above mentioned critical point are unique global minimizer too.

b) (Gradient of f) $\nabla f(x,y) = (4x-4y, -4x+4y^3+10y-10)^T$.
(Hessian of f) $\nabla^2 f(x,y) = \begin{bmatrix} 4 & -4 \\ -4 & 12y^2+10 \end{bmatrix}$. The first-order necessary condition of optimality gives $\nabla f(x,y) = 0$. Now, we have the critical point $x = (1,1)$ where the Hessian matrix is

$$\nabla^2 f(x,y) = \begin{bmatrix} 4 & -4 \\ -4 & 22 \end{bmatrix}.$$

It is easy to show that $p^T \nabla^2 f(x,y) p > 0$ for all non zero vector p, or, we can find the eigen values of Hessian matrix which are approximately 22.8489 and 3.1512. Thus Hessian matrix is positive definite. Therefore, $(1,1)$ is strict local minimizer and hence local minimizer.

Problem 4.

Compute the gradient, Hessian and local minimizers of the Rosenbrock function $f\colon \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x) = 100(x_2-x_1^2)^2+(1-x_1)^2$.

Solution.

(Gradient of f) $\nabla f(x,y) = (-400x_1(x_2-x_1^2) - 2(1-x_1), 200(x_2-x_1^2))^T$.
(Hessian of f) $\nabla^2 f(x,y) = \begin{bmatrix} -400x_2+1200x_1^2+2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$. The first-order necessary condition of optimality gives $\nabla f(x,y) = 0$. Now, we have the solution is $x = (1,1)$ and the Hessian matrix is

$$\nabla^2 f(x,y) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}.$$

It is evident that the Hessian matrix is symmetric and it's eigen values are $501 \pm \sqrt{250601}$, in which the value of both eigen values are positive. Therefore, $(1,1)$ is strict local minimizer and hence the local minimizer.

Problem 5.

For a matrix $A \in \mathbb{R}^{d\times d}$, we denote by

$$\|A\|_F := \left(\sum_{i,j=1}^{d} a_{ij}^2\right)^{\frac{1}{2}}$$

its *Frobenius norm*. Show that the optimization problem

$$\min_{\substack{A\in\mathbb{R}^{d\times d},\\ \det A>0}} \left(\|A\|_F + \frac{1}{\det A}\right)$$

admits a global minimum.

Solution.

The objective function $f(A) = \|A\|_F + \frac{1}{\det A}$ is continuous because the Frobenius norm is continuous and the determinant is also continuous (because it is polynomial of matrix elements). Thus $f(A)$ is continuous over the feasible set $\Omega = \{A \in \mathbb{R}^{d\times d} : \det A > 0\}$, which is neither closed nor bounded. Now we aim to construct a compact set, say $S \subset \Omega$ in which the objective function f is continuous. Since I(identity matrix) $\in \Omega$ and $\det I > 0$ with $f(I) = d^{\frac{1}{2}} + 1$. Therefore, if global minima exists, then it should be in the set $\{A \in \Omega : f(A) \le f(I)\} = \{A \in \Omega : f(A) \le d^{\frac{1}{2}}+1\} \subset \{A \in \mathbb{R}^{d\times d} : \|A\|_F \le d^{\frac{1}{2}}+1\} \cap \{A \in \Omega : \det A \ge (d^{\frac{1}{2}}+1)^{-1}\}$. We say,

$$S = \{A \in \mathbb{R}^{d\times d} : \|A\|_F \le d^{\frac{1}{2}}+1\} \cap \{A \in \Omega : \det A \ge (d^{\frac{1}{2}}+1)^{-1}\}$$

which is bounded (the first set in the intersection is bounded) and closed (both intersected sets are closed, since the functions defining the inequality constraints are continuous). Therefore, the given optimization problem attains a global minimum in S and hence in Ω too.