1 Exercises on elements of convex analysis

Exercise 1:

Theorem 1.1 (Weierstrass). *Given f a real continuous function on a compact set $K \subset \mathbb{R}^n$, then the optimization problem*

$$\begin{cases} \textit{Minimize} & f(x), \\ & x \in K \end{cases}$$

has an optimal solution $x^* \in K$.

Proof. We may prove that there exists $x^* \in K$ such that $f(x) \geq f(x^*)$, $\forall x \in K$. Suppose K is a non empty space, otherwise, as established, $x^* = -\infty$ and the problem is not considered.
Then, given it is a compact set, by definition it is closed and bounded in $\mathbb{R}^n$. Therefore, since it is bounded, there exists a lower bound $\alpha = Inf\{f(x) : x \in K\}$, i.e. for all $x \in K, \alpha \leq f(x)$. Let's consider the sequence $K_j = \{x \in K : \alpha \leq f(x) \leq \alpha + \varepsilon^j\}$, $\forall j \in \mathbb{N}$ and for a fixed $\varepsilon \in (0, 1)$. Note that, by definition, $K_j \neq \emptyset$ for all j.
Let's consider now a sequence given by $y^j \in K_j$, $\{y^j\}$. Since K is a compact space, we know that every sequence has a subsequence that converges, that we may denote as $\{x_j\}$. Therefore, we have a subsequence such that

$$\alpha \leq f(x_j) \leq \alpha + \varepsilon^j \tag{1}$$

that converges to $\hat{x}$.
Since K is closed and $\hat{x}$ is a limit point, then $\hat{x} \in K$, verifying $x_j \to \hat{x}$ when $j \to \infty$. Given the continuity of f, we also have $f(x_j) \to f(\hat{x})$, when $j \to \infty$. Then, taking limits into the expression (1) we get $\alpha \leq lim_{j\to\infty} f(x_j) = f(\hat{x}) \leq \alpha \iff \alpha = f(\hat{x}) \leq f(x)$ for all $x \in K$, and therefore $\hat{x} \in K$ is an optimal solution of the initial minimization problem.

□

Exercise 2:

Corollary 1.2. *If f is a real continuous function defined on all $\mathbb{R}^n$ such that $f(x) \to \infty$ when $||x|| \to \infty$, then the problem*

$$\begin{cases} \textit{Minimize} & f(x), \\ & x \in \mathbb{R}^n \end{cases}$$

has an optimal solution.

Proof. Let's consider $y \in \mathbb{R}^n$. Note, that by hypothesis, there exists $M > 0$ such that $||x|| \geq M \Rightarrow f(x) \geq f(y)$. Then, considering the closed ball $B = \{x \in \mathbb{R}^n : ||x|| \leq M\}$ we get a compact set in which we can apply the previous theorem to see that there exists $x^* \in \mathbb{R}^n$ solving the minimizing problem, as we wanted to prove. □

Exercise 3: *Let S be a convex subset of $\mathbb{R}^n$, and let λ_1, λ_2 be positive scalars.*

(a) *Show that $(\lambda_1 + \lambda_2)S = \lambda_1 S + \lambda_2 S$.*

Proof. We may prove it by double inclusion.

$\boxed{\subseteq}$ Given $x \in (\lambda_1 + \lambda_2)S$, by definition we have $x = (\lambda_1 + \lambda_2)y = \lambda_1 y + \lambda_2 y \in \lambda_1 S + \lambda_2 S$, with $y \in S$. Therefore, $(\lambda_1 + \lambda_2)S \subseteq \lambda_1 S + \lambda_2 S$.

$\boxed{\supseteq}$ Given now $x \in \lambda_1 S + \lambda_2 S$, we have $x = \lambda_1 y + \lambda_2 z$, for $y, z \in S$. Since $\lambda_1, \lambda_2 > 0$, we can consider the equivalent equality

$$\frac{1}{\lambda_1 + \lambda_2} x = \frac{\lambda_1}{\lambda_1 + \lambda_2} y + \frac{\lambda_2}{\lambda_1 + \lambda_2} z.$$

Note that $\frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\lambda_2}{\lambda_1+\lambda_2} = 1$, and $\frac{\lambda_1}{\lambda_1+\lambda_2}, \frac{\lambda_2}{\lambda_1+\lambda_2} \in (0,1)$. Therefore, by convexity we get

$$\frac{1}{\lambda_1 + \lambda_2} x \in S \iff x \in (\lambda_1 + \lambda_2) S,$$

and, thus $\lambda_1 S + \lambda_2 S \subseteq (\lambda_1 + \lambda_2) S$, as we wanted to prove[1].

$\square$

(b) *Give an example that shows that this does not need to be true when S is not convex.*

Consider the noncovex subset of $\mathbb{R}^2$ given by the circumference of radius 1 and centered at the origin, i.e. $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Note that if we consider $(x, y) \in (\lambda_1 + \lambda_2) S$, we get $(x, y) = (\lambda_1 + \lambda_2)(u, v)$, with $(u, v) \in S$. Therefore, it is direct to see that (x, y) verifies $x^2 + y^2 = (\lambda_1 + \lambda_2)^2 (u^2 + v^2) = (\lambda_1 + \lambda_2)^2$. Thus, $(x, y) \in (\lambda_1 + \lambda_2) S$ is contained in the circumference of radius $(\lambda_1 + \lambda_2)$. In fact, we have $(\lambda_1 + \lambda_2) S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = (\lambda_1 + \lambda_2)^2\}$.

Following the equivalent calculations for the case $(x, y) \in \lambda_1 S + \lambda_2 S \iff (x, y) = \lambda_1 (u, v) + \lambda_2 (w, z)$, with $(u, v), (w, z) \in S$, we get that (x, y) verifies the condition

$$x^2 + y^2 = \lambda_1^2 + \lambda_2^2 + 2\lambda_1 \lambda_2 (uv + wz) \notin (\lambda_1 + \lambda_2) S.$$

Exercise 4: *Let S be a nonempty closed convex set in $\mathbb{R}^n$, not containing the origin. Show that there exists a hyperplane that strictly separates S and the origin.*

Proof. We want to prove that there exists a hyperplane that strictly separates S and the origin, $\mathbf{0}$. More specifically, we want to show that there exists $a \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $b \in \mathbb{R}$ such that $a^T x > b$ for all $x \in S$, and $a^T x < b$ for $x = \mathbf{0}$. We may denote this hyperplane by $H(a, b)$.

Note that there exists a point $y \in S$ that is the closest to $\mathbf{0}$, i. e. $d(y, \mathbf{0}) \leq d(z, \mathbf{0})\ \forall z \in S$. This comes from the fact that S is closed. More explicitly, consider a general point $z \in S$ and let $r = d(z, \mathbf{0})$, then since $z \neq \mathbf{0}$ we have that $r > 0$. Thus, we can define the closed ball of radius r centered on the origin as $\overline{B} = cl(B(\mathbf{0}, r))$. Note that, since it is closed, $z \in \overline{B}$. In addition $\overline{B}$ is compact, and therefore $\overline{B} \cap S$ is nonempty (since $z \in \overline{B}$) and compact (because S is closed). Note now that the function $d(\cdot, \mathbf{0})$ is continuous, and therefore, from *Exercise 1* (*Weierstrass theorem)* it attains its minimum on the compact set $\overline{B} \cap S$. Let's denote this minimum point $y \in \overline{B} \cap S$. Then, suppose there is a point closer to $\mathbf{0}$, $y' \in S$, then we would have $d(y', \mathbf{0}) < d(y, \mathbf{0}) \leq r$, so we have $y' \in \overline{B}$ and therefore $y' \in \overline{B} \cap S$ with $d(y', \mathbf{0}) < d(y, \mathbf{0})$, contradicting the fact that y minimizes $d(x, \mathbf{0})$ on $\overline{B} \cap S$.

Then, let's denote this minimum by $y \in S$, that we already have proven that exists. Let's define $b = y^T y = ||y||^2$, and then consider the hyperplane $H(y, b)$. Note that clearly, $y^T \cdot \mathbf{0} = 0 < b$, since we assume $y \neq \mathbf{0} \notin S$.
Consider now any general point $x \in S$, we have to see that $y^t x \geq b = y^T y$. Let's assume, by way of contradiction, that $y^T x < y^T y$. From convexity, we have that for all $\lambda \in (0, 1)$, $z_\lambda = \lambda x + (1 - \lambda) y \in S$. We

[1] Note that, by definition, it is direct to see that $\frac{1}{\lambda_1+\lambda_2} x \in S \iff \frac{1}{\lambda_1+\lambda_2} x = y \in S \iff x = (\lambda_1 + \lambda_2) y \in (\lambda_1 + \lambda_2) S$

may see that for small values of λ, we have $d(z_\lambda, \mathbf{0}) < d(y, \mathbf{0})$, contradicting the fact that y minimizes $d(x, \mathbf{0})$ on S. Let's consider then the corresponding distance:

$$\begin{aligned} d(z_\lambda, \mathbf{0}) &= z_\lambda^T z = \lambda^2 x^T x + (1-\lambda)^2 y^T y + 2\lambda(1-\lambda) x^T y \\ &= y^T y + \lambda[\lambda(x^T x + y^T y - 2x^T y) + 2(x^T y - y^T y)] \\ &= d(y, \mathbf{0}) + \lambda\underbrace{[\lambda||x-y||^2 + 2(x^T y - y^T y)]}_{g(\lambda)} \end{aligned}$$

Note that by assumption $y^T x < y^T y \Rightarrow lim_{\lambda\to 0} g(\lambda) < 0$. Then for a λ small enough, i.e. inside a neighborhood of y it exists a point over the segment $\lambda x + (1-\lambda)y$ with $d(z_\lambda, \mathbf{0}) < d(y, \mathbf{0})$, that contradicts the fact that y minimizes this function in S. Thus, we have $y^t x \geq b = y^T y$, and consequently the hyperplane $H(y, y^T y)$ strictly separates S and the origin. □

Exercise 5: *Show that a convex function $f : (a, b) \to \mathbb{R}$ is continuous.*

Proof. Assume that f is a convex function, and we want to prove that it is continuous, i.e. that given $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ if $|x - y| < \delta$, for all $x, y \in (a, b)$.

Firstly, note that for any $x \in (a, b)$ there exist $u, v \in (a, b)$ such that $u < x < v$. Consider $\frac{x-u}{v-u} \in (0, 1)$, so we can express x in terms of u, v as follows:

$$x = \left(\frac{x-u}{v-u}\right) v + \left[1 - \left(\frac{x-u}{v-u}\right)\right] u.$$

Since f is convex, we have

$$f(x) \leq \left(\frac{x-u}{v-u}\right) f(v) + \left[1 - \left(\frac{x-u}{v-u}\right)\right] f(u). \tag{2}$$

Rearranging the terms from the inequality we get

$$f(x) - f(u) \leq \frac{x-u}{v-u}[f(v) - f(u)] \iff \frac{f(x) - f(u)}{x-u} \leq \frac{f(v) - f(u)}{v-u}. \tag{3}$$

Rewriting in a different way the inequality (2), we can also see that it verifies that

$$(f(v) - f(u))\left[1 - \left(\frac{x-u}{v-u}\right)\right] \leq f(v) - f(x) \iff \frac{f(v) - f(u)}{v-u} \leq \frac{f(v) - f(x)}{v-x} \tag{4}$$

Therefore, for any $x, y, \in [u, v]$, assuming without loss of generality that $x < y$ and since (a, b) is open, there exist $u_0, v_1 \in (a, b)$ such that $u_0 < u < y < x < v < v_1$. Thus, combining (3) applied to $x \leq y < v_1$ and (4) applied to $x < v < v_1$ we get:

$$\frac{f(y) - f(x)}{y-x} \leq \frac{f(v_1) - f(x)}{v_1 - x} \leq \frac{f(v_1) - f(v)}{v_1 - v}$$

Similarly, applying now (3) to the set $u_0 < u < y$ and then (4) applied to $u_0 < x < y$ we have:

$$\begin{gathered} \frac{f(u) - f(u_0)}{u - u_0} \leq \frac{f(y) - f(u)}{y-u} \leq \frac{f(y) - f(x)}{y-x} \iff \\ \frac{f(x) - f(y)}{y-x} \leq \frac{f(u) - f(y)}{y-u} \leq \frac{f(u_0) - f(u)}{u - u_0} \end{gathered}$$

Note that this is equivalent to affirm that:

$$\frac{|f(y)-f(x)|}{y-x} \leq C = max\left\{\frac{|f(u_0)-f(u)|}{u-u_0}, \frac{|f(v_1)-f(v)|}{v_1-v}\right\}.$$

Consequently, given $\epsilon > 0$, $x \in [u,v]$, let $\delta = min\{\frac{\epsilon}{C}, \frac{v-u}{2}\} > 0$, then for any $y \in (x-\delta, x+\delta) \subset [u,v]$, then $|f(y)-f(x)| \leq C|y-x| \leq C\frac{\epsilon}{C} = \epsilon$. Thus, we have f continuous, by definition in the interval (a,b). $\square$

Exercise extra:

Theorem 1.3. *A linear combination with positive coefficients of convex functions is a convex function.*

Proof. Let $f_1, \ldots, f_n$ be convex functions. Suppose we have a general linear combination $g = \lambda_1 f_1 + \cdots + \lambda_n f_n$, with $\lambda_1, \ldots, \lambda_n > 0$. Let's assume $dom\ f_i = dom\ g$, $\forall i = 1, \ldots, n$. Then, considering general $x, y \in dom\ g, \alpha \in (0,1)$ we have

$$\begin{aligned} g(\alpha x + (1-\alpha)y) &= \lambda_1 f_1(\alpha x + (1-\alpha)y) + \cdots + \lambda_n f_n(\alpha x + (1-\alpha)y)) \\ &\leq \lambda_1(\alpha f_1(x) + (1-\alpha)f_1(y)) + \cdots + \lambda_n(\alpha f_n(x) + (1-\alpha)f_n(y)) \\ &= \alpha(\lambda_1 f_1(x) + \cdots + \lambda_n f_n(x)) + (1-\alpha)(\lambda_1 f_1(y) + \cdots + \lambda_n f_n(y)) \\ &= \alpha g(x) + (1-\alpha)g(y) \end{aligned}$$

$\square$

Exercise 6: *Consider a function $f : (a,b) \to \mathbb{R}$ of class C^2. Show that f is convex if and only if $f''(x) \geq 0$ for all $x \in (a,b)$.*

Proof. $\boxed{\Rightarrow}$ Consider $c \in (a,b)$ and x close to c, i.e. let $x \in V(c)$, where $V(c)$ denotes a neighbourhood of c. Then we can compute the Taylor expansion of order 2 around c as

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \mathcal{O}(|x-c|^3)$$

Note that $f(c) + f'(c)(x-c)$ represents the tangent line of f at point c. By convexity, we know that the tangent line will remain bellow the function, therefore we have $f(x) - f(c) + f'(c)(x-c) \geq 0$. Therefore, given that $\mathcal{O}(|x-c|^3)$ is given by the error added by terms of greater order on the expansion, with x close enough to c we can neglect those terms so we have $f(x) - f(c) + f'(c)(x-c) \approx \frac{f''(x)}{2}(x-c)^2 \geq 0 \iff f''(c) \geq 0$.

$\boxed{\Leftarrow}$ Similarly, let's consider any point of the interval, $c \in (a,b)$, then we can compute the Taylor expansion around this point as previously, so the fact that $f''(x) \geq 0$ for all $x \in (a,b)$ implies that

$$0 \leq \frac{f''(c)}{2}(x-c)^2 \approx f(x) - f(c) + f'(c)(x-c).$$

Therefore, for any point of the considered interval we have that the tangent line is bellow the function's image, so by definition, the function is convex.

$\square$

Exercise 7: *Let f be a real valued function on an open convex set $S \subset \mathbb{R}^n$, of class C^2. Show that f is convex on S if and only if its Hessian matrix,*

$$Q(x) = H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right), \tag{5}$$

is positive semi-definite for all $x \in S$.

Proof. $\boxed{\Rightarrow}$ Assume f is a convex function. By hypothesis, $f \in C^2$, in particular it is C^1, therefore, by a previous theorem we have

$$f(y) \geq f(x) + \nabla f(x)^T(y-x),\ \forall x, y \in dom\ f.$$

Similarly to *Exercise 6*, let's consider Taylor's expansion of the function f around a general point $c \in S$ of degree 2:

$$f(x) = f(c) + \nabla f(c)^T(x-c) + \frac{1}{2}(x-c)^T H_f(c)(x-c) + \mathcal{O}(|x-c|^3).$$

By convexity, we have

$$0 \leq f(x) - f(c) + \nabla f(c)^T(x-c) \approx \frac{1}{2}(x-c)^T H_f(c)(x-c)$$

Since this is true for any $c \in S$, $x \in V(c)$, and we are on a convex open set we have $x - c \in S$. Then, we have that $(x-c)^T H_f(c)(x-c) \geq 0$, thus by definition of positive semi-definite, we have that the Hessian is positive semi-definite for any point of S (since they all are convex combinations of points of C) as we wanted to prove.

$\boxed{\Leftarrow}$ Similarly, assume that the Hessian matrix is positive semi-definite, $y^T H_f(x) y \geq 0,\ \forall x, y \in S$. Applying the Taylor's theorem, for a fixed point $c \in S$ and any $x \in S$ we have:

$$f(x) = f(c) + \nabla f(c)^T(x-c) + \frac{1}{2}(x-c)^T H_f(a)(x-c),$$

for some a in the segment between x and c, which belongs to S because of convexity. Then using that H_f is positive semi-definite we have

$$0 \leq \frac{1}{2}(x-c)^T H_f(a)(x-c) \approx f(x) - f(c) + \nabla f(c)^T(x-c) \iff f(x) \geq f(c) + \nabla f(c)^T(x-c),$$

which in virtue of a theorem previously proven, implies that the function f is convex, as we wanted to show.

$\square$

Note that this proof could also been demonstrated using the fact that the convexity of f on S is equivalent to the convexity of f restricted to each segment of S, and using therefore the previous exercise.

Exercise 8: *Assume $S \subset \mathbb{R}^n$ is a convex set and that $g : S \to \mathbb{R}$. Show that the set $g(x) \leq 0$ is convex if g is convex. What about the opposite implication?*

Proof. Assume g is a convex function, therefore, as stated in a previous proposition, the epigraph set

$$epi(g) = \{(\mu, x) : g(x) \leq \mu, x \in \mathbb{R}^n, \mu \in \mathbb{R}\}$$

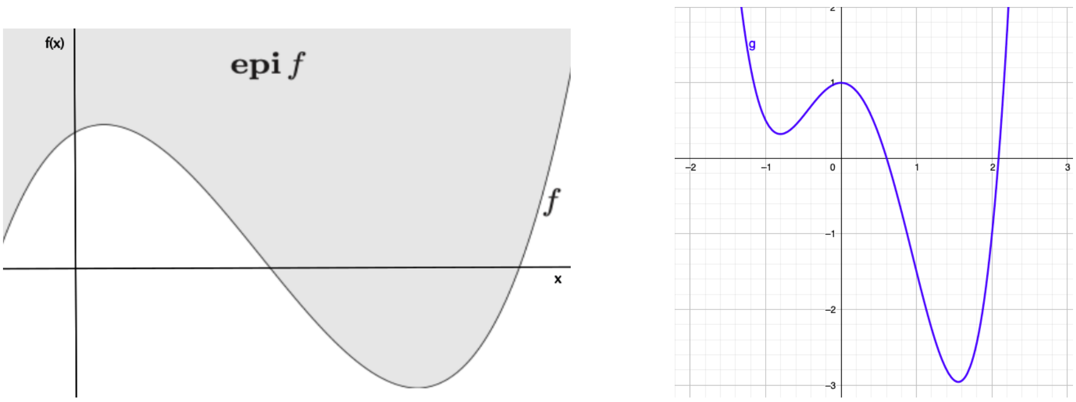
is convex. Consider now the subset $L = \{z = (z_0, z_1, \ldots, z_n) \in \mathbb{R}^{n+1} : z_0 \leq 0\} \subset \mathbb{R}^{n+1}$, which is a convex set. Consider the finite intersection of the two convex sets, which is, in fact, also convex:

$$epi(g)|_{\mu \leq 0} = epi(g) \cap L = \{(\mu, x) : g(x) \leq \mu, x \in \mathbb{R}^n, \mu \in \mathbb{R}, \mu \leq 0\}$$

Note that, geometrically, $epi(g)|_{\mu \leq 0}$ represents the epigraph of g restricted to the values where the function has negative values, i.e. $g(x) \leq 0$. Then, it is direct to note that if the restricted epigraph is convex, then the function bellow, $\{g(x) \leq 0\}$ is convex.
This can be trivially deduced from the proof of the proposition that states that g is convex if and

only if $epi(g)$ is convex. Explicitly, assume $epi(g)|_{\mu\leq 0}$ is a convex set, and suppose that $\{g(x) \leq 0\}$ is not convex. Then, there would exist $x, y \in \{z \in dom(g) : g(z) \leq 0\}$, $\lambda \in [0,1]$ such that $g(\lambda x + (1-\lambda)y) > \lambda g(x) + (1-\lambda)g(y)$. Thus, picking $\hat{\mathbf{x}} = (x, g(x)), \hat{\mathbf{y}} = (y, g(y)) \in epi(g)|_{\mu\leq 0}$, we would have $\lambda\hat{\mathbf{x}} + (1-\lambda)\hat{\mathbf{y}} = (\lambda x + (1-\lambda)y, \lambda g(x) + (1-\lambda)g(y)) \notin epi(g)|_{\mu\leq 0}$ contradicting the fact that $epi(g)|_{\mu\leq 0}$ is convex. $\square$

Let' take a closer look to the opposite implication. Assume that the set $g(x) \leq 0$ is convex and we want to know whether it can be affirmed that g is convex. Note that the hypothesis is a local affirmation, and we could have some regions of the whole epigraph of g on the positive hyperplane of $\mathbb{R}^{n+1}$ where the set is not globally convex. As a contra-example, consider the function $f : \mathbb{R} \to \mathbb{R}$ as shown on Figure 1a. More explicitly, we could consider the function $g(x) = x^4 - x^3 - \frac{5}{2}x + 1$, as shown on the Figure 1b.



(a) Illustrative contra-example of a function with $f(x) \leq 0$ convex but globally non convex.

(b) Graphical representation of $g(x) = x^4 - x^3 - \frac{5}{2}x + 1$.