Exercise 0.1. *Prove the extreme value theorem of Weierstrass: If f is a real continuous on a compact set $K \subset \mathbb{R}^n$, then the problem of optimization*

$$\begin{cases} \text{Minimize } f(x) \\ x \in K \end{cases}$$

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

Proof. Let m=$\inf_{x \in K}\{f(x)\}$. Then, there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ of elements of K such that $f(x_n) \xrightarrow[n\to\infty]{} m$($m$ may a priori have the value $-\infty$).
Since K is compact, there exists a sub-sequence $\{x_{n_k}\}_{n_k\in\mathbb{N}}$ which converge to $x^* \in K$. Since f is continuous, we have $f(x_{n_k}) \to f(x^*)$ and

$$m = \lim_{n\to\infty} f(x_n) = \lim_{n_k\to\infty} f(x_{n_k}) = f(x^*).$$

Since $x^* \in K$, and K is compact, $f(x^*) > -\infty$ and we have that $m > -\infty$. Hence, for all $x \in K$, $f(x^*) = m \leqslant f(x)$, so that $x^* \in K$ is an optimal solution of the given problem. □

Exercise 0.2. *Let f be a real continous function on $\mathbb{R}^n$ satisfying that $f(x) \to +\infty$ when $\|x\| \to +\infty$. Show that the problem of optimization*

$$\begin{cases} \text{Minimize } f(x) \\ x \in K \end{cases}$$

has an optimal solution $x^* \in \mathbb{R}^n$.

Proof. Let x_1 be any point in $\mathbb{R}^n$. As $f(x) \to +\infty$ when $\|x\| \to +\infty$, there exists $M > 0$ such that

$$\|x\| \geq \text{M} \Longrightarrow f(x) \geq f(x_1).$$

Therefore, the problem reduces to an optimization problem in the closed ball $B(0, M) := \{x \in \mathbb{R}^n : \|x\| \leqslant M\}$ which is compact and Weierstrass theorem applies. □

Exercise 0.3. ***(optional)*** *Let S be a convex subset of $\mathbb{R}^n$, and let λ_1 and λ_2 be positive scalars.*

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

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

Proof. For the first part of the exercise, we are going to proceed by double inclusion. The inclusion from left to right is easy. Let $p \in (\lambda_1 + \lambda_2)S$, then, there exists $x \in S$ such that $p = (\lambda_1 + \lambda_2)x$ and by the distributive property of $\mathbb{R}^n$, since $S \subset \mathbb{R}^n$, $p = \lambda_1 x + \lambda_2 x$. Therefore, $p \in \lambda_1 S + \lambda_2 S$.

For the other inclusion, we take $p \in \lambda_1 S + \lambda_2 S$. Therefore, $p = \lambda_1 x + \lambda_2 y$ for some $x, y \in S$. Now we are going to discuss different cases:
If $\lambda_1 + \lambda_2 = 0$, as $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 = \lambda_2 = 0$ and $(\lambda_1 + \lambda_2)S = \lambda_1 S + \lambda_2 S = \emptyset$.
If $\lambda_1 + \lambda_2 \neq 0$, in that case

$$p = \lambda_1 x + \lambda_2 y = (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x + \frac{\lambda_2}{\lambda_1 + \lambda_2} y\right).$$

As $\frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\lambda_2}{\lambda_1+\lambda_2} = 1$ and S is convex, $\frac{\lambda_1}{\lambda_1+\lambda_2}x + \frac{\lambda_2}{\lambda_1+\lambda_2}y \in S$ and $p \in (\lambda_1 + \lambda_2)S$.

For the second part of the exercise, we can think about the set

$$S := \{(x,y) \in \mathbb{R}^2 | x = y, x \leqslant \frac{1}{2}\} \cup \{(x,y) \in \mathbb{R}^2 | y = 0, x \leqslant 1\}.$$

This set is a triangle of vertices in $(0,0), (1,0)$ and $(\frac{1}{2}, \frac{1}{2})$ without the edge joining $(1,0)$ and $(\frac{1}{2}, \frac{1}{2})$. If we choose $\lambda_1 = 1$ and $\lambda_2 = 2$, we can see that $S + 2S \neq 3S$. For example,

$$(1,0) + 2(\frac{1}{2}, \frac{1}{2}) = (2,1) \notin 3S.$$

□

Exercise 0.4. ***(optional)*** *Let S be a nonempty closed convex set in $\mathbb{R}^n$, not containing the origin. Show that there exist a hyperplane that strictly separates S and the origin.*

I have not done this one
Notice that I change the order of exercise 5 and 6!!

Exercise 0.5. *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. Let us first prove the direction from left to right. Consider $f : (a,b) \to \mathbb{R}$ a convex function. We are going to see, that if f is a convex function, then, for $x, y, z \in (a,b)$ such that $x < y < z$, we have

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(z) - f(y)}{z - y}. \tag{0.1}$$

As $y \in (x,z)$, there exist a unique $\lambda \in [0,1]$ such that $y = (1-\lambda)x + \lambda z$. Therefore,

$$\frac{f(y) - f(x)}{y - x} = \frac{f((1-\lambda)x + \lambda z) - f(x)}{(1-\lambda)x + \lambda z - x},$$

and by convexity,

$$\frac{f((1-\lambda)x + \lambda z) - f(x)}{(1-\lambda)x + \lambda z - x} \leqslant \frac{(1-\lambda)f(x) + \lambda f(z) - f(x)}{\lambda(z - x)} = \frac{f(z) - f(x)}{z - x}.$$

Similarly we obtain that

$$\frac{f(z) - f(x)}{z - x} \leqslant \frac{f(z) - f(y)}{z - y},$$

and (0.1) is proved. Now we are going to see that if we have a convex function $f : (a,b) \to \mathbb{R}$, then f' must increase. To see that, we choose $x, y, z, t \in (a,b)$ such that $x < y < z < t$. Using the convexity and what we have just proved, we can see that

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(z) - f(y)}{z - y} \leqslant \frac{f(t) - f(z)}{t - z}.$$

Ignoring the middle term and let $y \to x^+$ and $t \to z^+$, we obtain that $f'(x) \leqslant f'(z)$. So f' increase and that implies $f''(x) \geq 0$ for all $x \in (a,b)$.

To see the other direction, we use the Taylor expansion of the function, for all $x \in (a, b)$

$$f(y) = f(x) + f'(x)(y - x) + \frac{f''(\zeta)}{2}(y - x)^2,$$

with ζ between x and y. Since $f''(z) \geq 0$ for all $z \in (a, b)$, we have that

$$f(y) \geq f(x) + f'(x)(y - x).$$

This proves that f is convex. □

Exercise 0.6. ***(optional)*** *Show that a convex function $f : (a, b) \to \mathbb{R}$ is continuous.*

Assume $f : (a, b) \to \mathbb{R}$ a convex function. Using the same arguments as the previous exercise, we notice that if $x, y, z, t \in (a, b)$ such that $x < y < z < t$, we have

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(z) - f(y)}{z - y} \leqslant \frac{f(t) - f(z)}{t - z}.$$

The previous inequalities can be written as

$$f(y) + \frac{f(y) - f(x)}{y - x}(z - y) \leqslant f(z) \leqslant f(y) + \frac{f(t) - f(z)}{t - z}(z - y).$$

Then, $\lim_{z \to y} f(z) = f(y)$ and the function is continuous in (a, b) since z is arbitrary.

Exercise 0.7. *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,*

$$Q(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) = \nabla^2 f(x)$$

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

Proof. Let us first prove the direction from left to right. Assume f is convex and choose an arbitrary point $a \in \mathbb{R}^n$. We define a new function $g : S \subset \mathbb{R}^n \to \mathbb{R}$ as

$$g(x) = f(x) - [\nabla f(a)]^T (x - a).$$

Because $-[\nabla f(a)]^T (x - a)$ is a linear function, thus it is convex and as g is the summation of two convex function, is a convex function. The first order and second order partial derivatives of g are

$$\begin{aligned} \nabla g(x) &= \nabla f(x) - \nabla f(a), \\ \nabla^2 g(x) &= \nabla^2 f(x). \end{aligned}$$

It is easy to see that $\nabla g(a) = 0$, thus a is a global minimum of g. To see that, according to the sufficient and necessary condition for convex functions, we have

$$g(x) \geq g(a) + [\nabla g(a)]^T (x - a) = g(a).$$

Because x is arbitrary, a is a global minimum of g. Then, $\nabla^2 g(a)$ is positive semi-definite. As $\nabla^2 g(a) = \nabla^2 f(a)$, $\nabla^2 f(a)$ is also positive semi-definite. Because $a \in \mathbb{R}^n$ was arbitrary, the Hessian matrix $\nabla^2 f(x)$ is positive semi-definite for all $x \in S$.

For the other direction, assume that the Hessian matrix of f is positive semi-definite for all $x \in S$. According to the Taylor's theorem, we have

$$f(y) = f(x) + [\nabla f(x)]^T(y-x) + \frac{1}{2}(y-x)^T\nabla^2 f(\zeta)(y-x),$$

with ζ between x and y. As the Hessian matrix of f is positive semi-definite, we have that

$$f(y) \geq f(x) + [\nabla f(x)]^T(y-x).$$

Therefore, f is convex. $\square$

Exercise 0.8. *Assume that $S \subset \mathbb{R}^n$ is a convex set and that $g : S \to \mathbb{R}$. Show that the set $g(x) \leqslant 0$ is convex if g is convex. What about the opposite implication?*

Proof. We have to see that if g is convex, then the set $\Omega := \{x \in S \mid g(x) \leqslant 0\}$ is convex. In order to prove that, we choose $x, y \in \Omega$ and we want to see that for all $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in \Omega$. As S is convex, $\lambda x + (1-\lambda)y \in S$ and by the convexity of g, we have

$$g(\lambda x + (1-\lambda)y) \leqslant \lambda g(x) + (1-\lambda)g(y).$$

Since $x, y \in \Omega$, $g(x) \leqslant 0$ and $g(y) \leqslant 0$, thus

$$g(\lambda x + (1-\lambda)y) \leqslant \lambda g(x) + (1-\lambda)g(y) \leqslant 0.$$

Hence, $\lambda x + (1-\lambda)y \in \Omega$ as we wanted to see.
The other implication is false. For example, choose a convex subset of $\mathbb{R}$, $S := (-2,2)$ and $g : S \to \mathbb{R}$ given by $g(x) = -x^2 + 4$. Then, the set $\Omega := \{x \in S \mid g(x) \leqslant 0\}$ is the empty set, which is trivially convex, but g is not a convex function, because $g \in C^2$ and $g''(x) < 0$ for all $x \in S$. $\square$

Exercise 0.9. ***(optional)*** *A linear combination with positive coefficients of convex functions is a convex function*

Proof. Let μ_1 and μ_2 be positive coefficients and $f, g : \mathbb{R}^n \to \mathbb{R}$ two functions with the same domain. We want to prove that $h(x) := \mu_1 f(x) + \mu_2 g(x)$ is a convex function. We have to see that

$$h(\lambda x + (1-\lambda)y) \leqslant \lambda h(x) + (1-\lambda)h(y)$$

for all $\lambda \in [0,1]$ and for all $x, y \in$ dom(h)=dom(f)=dom(g). We notice that

$$h(\lambda x + (1-\lambda)y) = \mu_1 f(\lambda x + (1-\lambda)y) + \mu_2 g(\lambda x + (1-\lambda)y).$$

By the convexity of f and g, we have that

$$h(\lambda x + (1-\lambda)y) \leqslant \mu_1\lambda f(x) + \mu_1(1-\lambda)f(y) + \mu_2\lambda g(x) + \mu_2(1-\lambda)g(y).$$

Grouping the terms properly, we can see what we want to prove, i.e:

$$h(\lambda x + (1-\lambda)y) \leqslant \lambda h(x) + (1-\lambda)h(y).$$

$\square$