1 Exercises on multi-dimensional optimization

Exercise 1. (Optional) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function such that -f is also a convex function. Prove that there exists $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $f(x) = a^T x + c$.*

Proof. Recall the definition of convexity of a function:

Definition 1.1. *$f : \mathbb{R}^n \to \mathbb{R}$, $dom(f) \subset \mathbb{R}^n$, is a convex function if*

1. *dom(f) is convex*
2. *$\forall x \in dom(f)$, $\forall y \in dom(f)$, $\forall \lambda \in [0, 1]$, then $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. If the inequality is strict, i.e. $<$, then f is strictly convex.*

Thus, we have that for any $x, y \in \mathbb{R}^n$, $\forall \lambda \in [0, 1]$ then, since f is convex

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

and given that $-f$ is convex

$$-f(\lambda x + (1 - \lambda)y) \leq -\lambda f(x) - (1 - \lambda)f(y) \iff f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

Consequently,

$$f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda)f(y)$$

and we have that the function $f|_L$ is linear, this is that f restricted to the line between x and y, denoted as L is linear, and this holds for any two points from $\mathbb{R}^n$. Then we have that f is linear, and consequently, there exist $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $f(x) = a^T x + c$ as wanted to prove. ☐

Exercise 2. *Use the Kuhn-Tucker conditions to solve the following problems*

(a)

$$(P) \begin{cases} Min & f(x) = x_1 x_2 \\ s.t. & x_1 + x_2 \geq 2 \\ & x_2 \geq x_1 \end{cases} \iff \begin{cases} Min & f(x) = x_1 x_2 \\ s.t. & g_1(x) = 2 - x_1 - x_2 \leq 0 \\ & g_2(x) = x_1 - x_2 \leq 0 \end{cases}$$

According to the KKT theorem, since clearly the (CG) condition holds, then $x^0 \in X = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i \in I\}$ is a local minimum of (P) if there exist $\lambda_i \geq 0, i \in I$ called KKT multipliers such that

$$\begin{cases} \nabla f(x^0) + \sum_{i \in I} \lambda_i \nabla g_i(x^0) = 0 \\ \lambda_i g_i(x^0) = 0, \ \forall i \in I \end{cases}$$

Computing the corresponding gradients we have

$$\begin{aligned} \nabla f(x) &= (x_2, x_1) \\ \nabla g_1(x) &= (-1, -1) \\ \nabla g_2(x) &= (1, -1) \end{aligned}$$

thus, the KKT conditions are given as follows:

$$\begin{cases} x_2 - \lambda_1 + \lambda_2 = 0 \\ x_1 - \lambda_1 - \lambda_2 = 0 \\ \lambda_1(2 - x_1 - x_2) = 0 \\ \lambda_2(x_1 - x_2) = 0 \end{cases}$$

From the two first linear equations, we have $x_2 = \lambda_1 - \lambda_2$ and $x_1 = \lambda_1 + \lambda_2$. Now, regarding the last two equations, the nonlinear ones, we can see that they hold if

$$\lambda_1(2 - x_1 - x_2) = 0 \iff 2\lambda_1(1 - \lambda_1) = 0 \iff \begin{cases} \lambda_1 = 0 \\ \lambda_1 = 1 \end{cases}$$
$$\lambda_2(x_1 - x_2) = 0 \iff -\lambda_2^2 = 0 \iff \lambda_2 = 0$$

Thus, if $\lambda_2 = 0$, we have $x = (0, 0)$, but note that this point does not verify the restriction $g_1(x) \leq 0$. Therefore, consider the case $\lambda_1 = 1$, and we have $x = (1, 1)$, which in fact satisfies $g_1(x) = 2 - 1 - 1 = 0 \leq 0$. Then, if the optimization problem (P) has a minimum it is in $\hat{x} = (1, 1)$.

(b)

$$(P) \begin{cases} Min & f(x) = (x_1 - 1)^2 + x_2 - 2 \\ s.t. & x_2 - x_1 = 1 \\ & x_1 + x_2 - 2 \leq 0 \end{cases}$$

Recalling the KKT theorem for optimization problems that have equality constraints $h_i(x) = 0$, we have that a necessary condition for x^0 to be a local minimum of (P) is that there exist $\lambda_i \geq 0, (i \in I)$ and $\mu_j > 0, (j \in L)$ such that

$$\begin{aligned} &\nabla f(x^0) + \sum_{i \in I} \lambda_i \nabla g_i(x^0) + \sum_{j \in L} \mu_j h_j(x^0) = 0 \\ &\lambda_i g_i(x^0) = 0, \ \forall i \in I \end{aligned}$$

Computing the corresponding gradients as follows

$$\begin{aligned} \nabla f(x) &= (2(x_1 - 1), 1) \\ \nabla g(x) &= (1, 1) \\ \nabla h(x) &= (-1, 1) \end{aligned}$$

so the KKT conditions correspond to:

$$\begin{cases} 2(x_1 - 1) + \lambda - \mu = 0 \\ 1 + \lambda + \mu = 0 \\ \lambda(x_1 + x_2 - 2) = 0 \end{cases}$$

Following the same procedure as before, note that from the linear equations we get $\lambda = -1 - \mu$ and $x_1 = 1 + \frac{\mu - \lambda}{2} = \frac{3}{2} + \mu$. Thus, by considering the last equation we have:

$$\lambda(x_1 + x_2 - 2) = 0 \iff \begin{cases} \lambda = 0 \\ (x_1 + x_2 - 2) = 0 \end{cases}$$

Note that if $\lambda = 0$, then $\mu = -1$, $x_1 = \frac{1}{2}$. Therefore, from the constraint $h(x) = 0$ we get $x_2 = \frac{3}{2}$, which also verifies $g(x) = 0 \leq 0$. Then, $\hat{x} = (\frac{1}{2}, \frac{3}{2})$ is an optimal solution of (P).

Note that by considering the other condition we obtain the same result, since we get $(x_1+x_2-2) = 0 \iff (x_1, x_2) = (\frac{3}{2} + \mu, \frac{1}{2} - \mu)$, and when restricting this solution to the remaining constraint we get $h(x^0) = 0 \iff \mu = -1$ obtaining again the point $\hat{x}$.

(c)

$$(P)\begin{cases} Min & f(x) = x_1^2 + 2x_2^2 + 3x_3^2 \\ s.t. & x_1 - x_2 - 2x_3 - 12 \leq 0 \\ & x_1 + 2x_2 - 3x_3 - 8 \leq 0 \end{cases}$$

Consider the corresponding gradients:

$$\nabla f(x) = (2x_1, 4x_2, 6x_3)$$
$$\nabla g_1(x) = (1, -1, -2)$$
$$\nabla g_2(x) = (1, 2, -3)$$

Then, the KKT conditions are given as follows:

$$\begin{cases} 2x_1 + \lambda_1 + \lambda_2 = 0 \\ 4x_2 - \lambda_1 + 2\lambda_2 = 0 \\ 6x_3 - 2\lambda_1 - 3\lambda_3 = 0 \\ \lambda_1(x_1 - x_2 - 2x_3 - 12) = 0 \\ \lambda_2(x_1 + 2x_2 - 3x_3 - 8) = 0 \end{cases}$$

Following again the same procedure as previously, from the linear equations we have

$$x_1 = -\frac{\lambda_1 + \lambda_2}{2}, \quad x_2 = \frac{\lambda_1 - 2\lambda_2}{4}, \quad x_3 = \frac{2\lambda_1 + 3\lambda_2}{3}$$

Consider now the non-linear part of the KKT system. Note that we have 4 possibilities:

$$\lambda_1(x_1 - x_2 - 2x_3 - 12) = 0 \iff \begin{cases} \lambda_1 = 0 \\ (x_1 - x_2 - 2x_3 - 12) = 0 \end{cases}$$

$$\lambda_2(x_1 + 2x_2 - 3x_3 - 8) = 0 \iff \begin{cases} \lambda_2 = 0 \\ (x_1 + 2x_2 - 3x_3 - 8) = 0 \end{cases}$$

Consider the case in which $\lambda_1 = 0 = \lambda_2$. Note that for this case we have $\hat{x} = (0, 0, 0)$ which actually verifies all the constraints $g_i(\hat{x}) \leq 0$, and in fact gives the minimum possible value for the optimization function, since $f(x) \geq 0$, $\forall x$, and is therefore the optimal solution of (P).