1 Exercises: Introduction to optimization

Jaime Sánchez Salazar

Exercise 1

(i) In order to prove that there exists a global minimum, we define the following function: $f(x) := \sum_{i=1}^{m} w_i||x - y_i||$. Since norms are continuous and the sum of continuous functions are also continous, we have that f is continous. Now, let x_1 be any point in $\mathbb{R}^2$. As $f(x) \to +\infty$ when $||x|| \to +\infty$, there exists $M > 0$ such that $||x|| \geq M \implies f(x) \geq f(x_1)$.

Therefore, the problem to find the minimum in $\mathbb{R}^2$ reduces to find the minimum in the closed ball $B(0, M) := \{x \in \mathbb{R}^n : ||x|| \leq M\}$ which is compact and by Weierstrass theorem we can conclude that there exist a global minimum for this function.

We want to know the point at which the minimum is achieved. That is, minimize $f(x)$. Let x^* be the minimum point, as $f(x) \subseteq \mathbb{R}^2$ we will to satisfy that $\nabla f(x^*) = 0$. That is to say, $\nabla(\sum_{i=1}^{m} w_i||x - y_i||) = \sum_{i=1}^{m} w_i \nabla(||x - y_i||) = 0$, Which means that the weighted vectors $\sum_{i=1}^{m} w_i(||x - y_i||)$ must be zero. And, solving for x we easily obtain that $x = \frac{\sum_{i=1}^{m} w_i y_i}{\sum_{i=1}^{m} w_i}$.

(ii) If we have a look at the previous formula we can easily see that the optimal solution won't be always unique. In fact, it will be only unique if $w_i = w \in \mathbb{R}^+ \ \forall i$, more precisely, $\mathbf{x}^* = \frac{\sum_{i=1}^{m} y_i}{m}$. Otherwise, if $w_i \neq w \ \forall i$ we will obtain different optimal solutions.

I also tried another mathematical argument that contradicts this argument. As I don't really see the mistake, at least I will share the mathematical idea.

We want to prove that the global minimum is unique. For that, suppose that the function f has at least one global minimum. If we are able to prove that f is a convex function we are done, since a global minimum of a convex function in a convex set must be unique. Let's start proving that, in fact, f is a convex function.

Given $g : \mathbb{R}^2 \to \mathbb{R}$ by $g(x) := ||x||$ we want to prove that g is convex. We have to see that for all $x, y \in \mathbb{R}^2$ and for all $\lambda \in [0, 1]$ satisfies $g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y)$. This comes from the triangle inequality of the norms. So, $g(\lambda x + (1 - \lambda)y) = ||\lambda x + (1 - \lambda)y|| \leq ||\lambda x|| + ||(1 - \lambda)y|| = \lambda||x|| + (1 - \lambda)||y||$, since $1 - \lambda \geq 0$. Then, the norm is a convex function and since the finite sum of convex functions is also a convex function, we have that f is a convex function.

To see that, assume f, g convex functions in $\mathbb{R}^2$, let us prove that $f + g$ is also a convex function. For all $x, y \in \mathbb{R}^2$ and for all $\lambda \in [0, 1]$ we have $(f + g)(\lambda x + (1 - \lambda)y) = f(\lambda x + (1 - \lambda)y) + g(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(x) + \lambda g(x) + (1 - \lambda)g(x) = \lambda(f + g)(x) + (1 - \lambda)(f + g)(x)$.

Finally we are going to prove that if we have a convex function defined in a convex set, if it has a global minimum, it's unique. Notice that this concludes the prove, since our function is convex and the ball $B(0, M)$ is a convex set. In order to prove it, we may assume that we have two global minimum, x_1, x_2. Without loss of generality, assume $x1 < x2, f(x1) = f(x2) and f(x) > f(x1) = f(x2)$ for all x in the domain. Since the domain is convex, for all $\lambda \in [0,1], \lambda x_1 + (1-\lambda)x_2$ belongs to the domain and $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x1) + (1-\lambda)f(x_2) = \lambda f(x_1) + (1-\lambda)f(x_1) = f(x_1)$, which is a contradiction with $f(x) > f(x_1) = f(x_2)$ for all x in the domain.

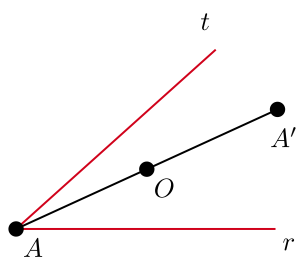
(iii) In the mechanical model shown in the figure, we can think each point $y_i \in \mathbb{R}^2$ as a weight with a mass proportional to its weight w_i. Therefore, the problem can be translated as to find the equilibrium point for a multiple weights attached to the plane. In the same way, the optimal solution x^* can be seen as the point that keeps the system in equilibrium. That is to say, if all the weights were attached to a "body" located at $x*$, all the the forces exerted by the weights would balance and the system would remain in equilibrium.

(iv) It's known that the potential energy of a gravitational field is given by $U = mgh$ where m is the mass of the object, g is the gravitational acceleration and h is the height of the object above some reference level. So, if we consider each weight at the position y_i, then we have that the potential energy associated with each weights is equal to $gw_i h$, where we have substituted $m \to w_i$ and $h \to h_i$ where h_i is the distance between the weight and the reference level. That means that the potential energy of the sistem is: $U = g\sum_{i=1}^{m} w_i h_i$.

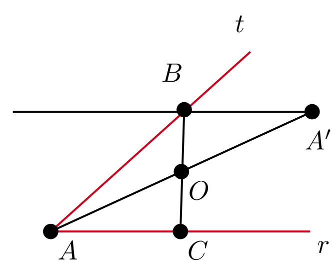
Now, in order to minimize U we want to minimize that h_i, but if we imagine the distance from x to y_i as a rope joining them, the shorter the rope the shorter the distance will be. Thus, to minimize U is the same as minimize $g\sum_{i=1}^{m} w_i(||x - y_i||)$, which actually is the previous question, given the fact that the constant g doesn't modify the solution.

Exercise 2

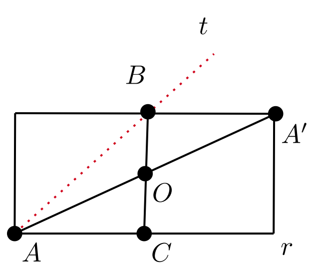
We will answer this question using geometric arguments. To start with, let's consider the central symmetry with respect to the point O.



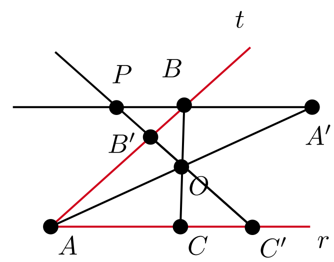
In particular, we draw a parallel line to r through A', that is r'. Let $B := r' \cap t$ and C to be the intersection point between r and a line going through B and O.



Now, if we have a look at the currently situation we can see that there's a rectangle, whose centre is precisely the point O. In fact, this centre satisfies that every single line that goes through it cuts the rectangle at two point equidistant from the centre, which actually defines a symmetry with respect to a point for all the points of the rectangle. In other words, the distance $||\vec{BO}||$ and $||\vec{CO}||$ are the same.



We consider now an arbitrary line different from $l := B'C'$ that goes through the point O and intersect t and r lines. Let's call this points B' and C' respectively and $P := r' \cup l$.



Note that, using the same argument as before, $||\vec{PO}||$ and $||\vec{C'O}||$. Let f be the area function, then we have that

$f(AB'C') = f(AB'OC) + f(OC'C) = f(AB'OC) + f(PBB') > f(AB'OC) + f(B'BO) = f(ABC)$

(The case when C' is located between A and C is treated in the same way.) Therefore, as B' and C' are random points, the line $l' := BC$ cuts off a triangle of minimal area from the given angle.