Advanced Mathematics for Scientific Challenges

EXERCISES 1.1

1) Prove the extreme value theorem of Weierstrass:
If f is a real continuous function 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

$f: K \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, K compact, f continuous

Since f is a continuous function on a compact set K, we have, by the boundedness theorem, that f is bounded on K.

Which means $\text{Image}(f) \subset \left(\inf_{x \in K} f(x), \sup_{x \in K} f(x) \right)$.

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 $f(x_n) \xrightarrow[n \to +\infty]{} m$

Since K is compact, there exists a sub-sequence $\{x_{u_n}\}_{u_n \in \mathbb{N}}$ which converges to $x^* \in K$.

Since f is continuous, we have $f(x_{u_n}) \to f(x^*)$

and $m = \lim_{n \to +\infty} f(x_n) = \lim_{u_n \to +\infty} f(x_{u_n}) = f(x^*)$

Therefore we have $x^* \in K : f(x^*) = m \leq f(x) \; \forall x \in K$ and $m > -\infty$ (limited), so x^* is a solution of the given problem. ■

2) Let f be a real continuous function on $\mathbb{R}^n$ satisfying that $f(x) \to +\infty$ when $\|x\| \to +\infty$. Show that the problem of optimization $\begin{cases} \text{Minimize } f(x) \\ x \in \mathbb{R}^n \end{cases}$ has an optimal solution $x^* \in K$

Proof

We have that $\lim_{\|x\| \to +\infty} f(x) = +\infty$, f continuous, $f: \mathbb{R}^n \to \mathbb{R}$.

Given a point $M \in \mathbb{R}^n$ such that $\|M\|$ is large enough, we have that $f(x) < f(M)$ $\forall x: \|x\| < \|M\|$, since f is continuous (no vertical asymptotes) and $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

Therefore we have a set $K = \{x \in \mathbb{R}^n : \|x\| < \|M\|\}$ which is compact and includes points for which $f(x) < f(M)$ is satisfied, in which f is continuous. On this set K the extreme value theorem of Weierstrass can be applied, as previously shown, so the problem $\begin{cases} \text{Minimize } f(x) \\ x \in \mathbb{R}^n \end{cases}$ has an optimal solution $x^* \in K$. ■

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

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

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

Proof

(a) By double inclusion, let $p \in (\lambda_1 + \lambda_2) S$, then, $\exists x \in S$ such that $p = (\lambda_1 + \lambda_2) x$ and by the distributive property of $\mathbb{R}^n$, $p = \lambda_1 x + \lambda_2 x$, therefore $p \in \lambda_1 S + \lambda_2 S$.
To show the opposite inclusion, let $p \in \lambda_1 S + \lambda_2 S$, then, $\exists x, y \in S$ such that $p = \lambda_1 x + \lambda_2 y$.
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 $\lambda_1 + \lambda_2 \neq 0$, then $p = \lambda_1 x + \lambda_2 y = (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x + \frac{\lambda_2}{\lambda_1 + \lambda_2} y\right)$
Since $\frac{\lambda_1}{\lambda_1 + \lambda_2} + \frac{\lambda_2}{\lambda_1 + \lambda_2} = 1$ and S is convex, $\frac{\lambda_1}{\lambda_1 + \lambda_2} x + \frac{\lambda_2}{\lambda_1 + \lambda_2} y \in S$, and $p \in (\lambda_1 + \lambda_2) S$

(b) Consider the set $S := \{(x,y) \in \mathbb{R}^2 : x = y, x \leq \frac{1}{2}\} \cup \{(x,y) \in \mathbb{R}^2 : y = 0, x \leq 1\}$
This set is a triangle of vertices in $(0,0)$, $(1,0)$ and $(1/2, 1/2)$, without the edge joining $(1,0)$ and $(1/2, 1/2)$. If we choose $\lambda_1 = 1$ and $\lambda_2 = 2$, we can see that $S + 2S \neq 3S$. For example: $(1,0) + 2(1/2, 1/2) = (2,1) \notin 3S$ ■

4) (OPTIONAL) 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.

<u>Proof</u>

Since S is a closed set we have that its complement S^c is an open set. (depending on the definition considered, it is either true by definition or it can be proven)

We have $0 \notin S \subset \mathbb{R}^n \Rightarrow 0 \in S^c$ open set in $\mathbb{R}^n$.

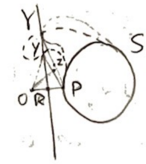
This allows us to identify a neighborhood of points $x \notin S$, $x \in B_{\varepsilon>0}(0) \subset S^c$

To show this is true is enough to consider the projection p of the origin on the set S. $p \in S$, since S is closed, which means that a point $(0+\varepsilon) \in S^c$, $\varepsilon > 0$ can be surely found on the projection ray, allowing for $B_{\varepsilon>0}(0) \subset S^c$ to exist. Let $v = [v_1, \dots, v_n]$ be the projection ray.

Consider once again the point $R = (0+\varepsilon) \in S^c$, this lies on v which, being a projection ray, is the shortest path from 0 to the frontier of S.

Consider now the hyperplane γ passing through $R = (0+\varepsilon)$ and orthogonal to v. γ strictly separates 0 from part of $\mathbb{R}^n$

ex. in 2 dim.



Suppose γ intersects S in y.

Since S is convex, $\forall \lambda \in [0,1]$ $\lambda p + (1-\lambda) y \in S$.

We can find z point of such convex combination such that $0z < 0p$, which means that either p is not the projection of 0 on S, or γ cannot intersect S. Therefore γ is a hyperplane that strictly separates S and the origin. ■

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

<u>Proof</u>

$f: (a,b) \to \mathbb{R}$ convex. Given $x, y, z, t \in (a,b)$ such that $x < y < z < t$, the following inequality holds:

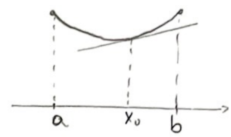
$$\frac{f(y)-f(x)}{y-x} \le \frac{f(z)-f(y)}{z-y} \le \frac{f(t)-f(z)}{t-z}$$

which can be written as:

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

and when $z \to y$, $f(z) \to f(y)$. Therefore we have that $\lim_{z \to y} f(z) = f(y)$ for any point $z \in (a,b)$, which means f is continuous on (a,b) ■

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



<u>Proof</u>

(f convex $\Rightarrow f''(x)\ge 0$)

We have already shown that:

$$f \text{ convex} \Rightarrow f(y)\ge f(x)+f'(x)(y-x) \quad \forall x,y\in(a,b) \quad (1)$$

Since $f\in C^2$ we can write the Taylor approximation of order 2 in x_0, $\forall x_0\in(a,b)$: $f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2+o((x-x_0)^2)$ $\quad x\approx x_0$

Let $x=y,\ x_0=x$: $f(y)=f(x)+f'(x)(y-x)+\frac{1}{2}f''(x)(y-x)^2+o((y-x)^2)$

Rearranging: $\underbrace{\tfrac{1}{2}}_{>0}f''(x)\underbrace{(y-x)^2}_{\ge 0}=\underbrace{f(y)-f(x)-f'(x)(y-x)}_{\text{from (1) we have } f(y)-f(x)-f'(x)(y-x)\ge 0}-o((y-x)^2)$ → error over the order of 2, its sum will not change the sign

Therefore we conclude that $f''(x)\ge 0\ \forall x\in(a,b)$

($f''(x)\ge 0 \Rightarrow f$ convex)

Similarly, from the previous equation, we have:

$$f(y)-f(x)-f'(x)(y-x)=\underbrace{\tfrac{1}{2}f''(x)(y-x)^2}_{\ge 0}+o((y-x)^2)\ge 0 \quad y\approx x$$

(*) We have show the result for $y\approx x$, which generalizes to any combination of $x,y\in(a,b)$

$\Rightarrow f(y)-f(x)-f'(x)(y-x)\ge 0$

$f(y)\ge f(x)+f'(x)(y-x) \overset{(*)}{\Longleftrightarrow} f$ is convex (as previously shown) ∎

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, $Q(x)=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right)$ is positive semi-definite for all $x\in S$.

<u>Proof</u>

We have already shown that f convex $\Leftrightarrow f(y)\ge f(x)+[\nabla f(x)]^T(y-x)$ (1) $\forall x,y\in dom(f)\cap\mathbb{R}^n$

Similary to the previous proof, consider the Taylor approximation of order 2 of $f\in C^2$ in $x\in S$:

$$f(y)=f(x)+[\nabla f(x)]^T(y-x)+(y-x)^T Q_f(x)(y-x)+o((y-x)^2)$$

$\underbrace{(y-x)^T Q_f(y-x)}_{\Downarrow}=\underbrace{f(y)-f(x)-[\nabla f(x)]^T(y-x)}_{\text{from (1) we have } f(y)-f(x)-[\nabla f(x)]^T(y-x)\ge 0}-o((y-x)^2)$ → error over the order of 2, its sum will not change the sign

Must, therefore, be ≥ 0; which means $v^TQ_fv\ge 0$ $\Rightarrow$

$\Rightarrow Q_f$ is positive semi-definite $\forall x\in S$

(This result holds in both directions, since it relies on (1) which is a relation of necessary and sufficient conditions; the previous consideration, (*), still applies) ∎

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

Proof

$S \subset \mathbb{R}^n$ convex: $\forall x \in S$, $\forall y \in S$, $\forall \lambda \in [0,1]$ $\quad \lambda x + (1-\lambda) y \in S$

Let $g: S \to \mathbb{R}$ convex: $\forall x \in S$, $\forall y \in S$, $\forall \lambda \in [0,1]$ $\quad g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda) g(y)$

Consider the set $C = \{x \in S : g(x) \leq 0\}$, $C \subseteq S \subset \mathbb{R}^n$

Let $x, y \in C$, meaning $g(x) \leq 0, g(y) \leq 0$. Since $x, y \in C$, we have $x, y \in S$.

Therefore, $\forall \lambda \in [0,1]$ $\quad g(\lambda x + (1-\lambda)y) \leq \underbrace{\underbrace{\lambda}_{\in[0,1]} \underbrace{g(x)}_{\leq 0}}_{\leq 0} + \underbrace{\underbrace{(1-\lambda)}_{\in[0,1]} \underbrace{g(x)}_{\leq 0}}_{\leq 0} \leq 0$

$\Rightarrow \forall x \in C$, $\forall y \in C$, $\forall \lambda \in [0,1]$ $\quad \lambda x + (1-\lambda) y \in C$ $\quad \Rightarrow$ C is convex

Consider now the opposite implication: suppose $C = \{x \in S : g(x) \leq 0\}$ convex

$\forall x \in C$, $\forall y \in C$, $\forall \lambda \in [0,1]$ $\quad \lambda x + (1-\lambda) y \in C$ which in terms of g means: $g(x) \leq 0$, $g(y) \leq 0$, $\lambda \in [0,1]$ $\Rightarrow g(\lambda x + (1-\lambda)y) \leq 0$

For g to be convex on C we would need $\underbrace{g(\lambda x + (1-\lambda)y)}_{\leq 0} \leq \lambda \underbrace{g(x)}_{\leq 0} + (1-\lambda) \underbrace{g(y)}_{\leq 0}$

This is not guaranteed by the convexity of C.

In fact, consider the function $g: C \to \mathbb{R}$, $C \subseteq \mathbb{R}^n$, $g(x) = -(x_1^2 + \dots + x_n^2)$, $\forall x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in C \subset \mathbb{R}^n$ and restrict, for example, C to the set in which $\|x\| \leq \sqrt{n}$

Let $x = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in C \subset \mathbb{R}^n$, $y = \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} \in C \subset \mathbb{R}^n$, $\lambda = 1/2$, we have the following:

$\lambda x + (1-\lambda) y = \frac{1}{2} x + \frac{1}{2} y = 0 \in C \subset \mathbb{R}^n$, $\quad g(\lambda x + (1-\lambda) y) = g(0) = 0$

$\lambda g(x) + (1-\lambda) g(y) = \frac{1}{2} \underset{\overset{\parallel}{-n}}{g(1)} + \frac{1}{2} \underset{\overset{\parallel}{-n}}{g(-1)} = -n$

$\Rightarrow g(\lambda x + (1-\lambda)y) \not\leq \lambda g(x) + (1-\lambda) g(y)$

$\Rightarrow g$ is not convex on $C \subset S \subset \mathbb{R}^n$ $\Rightarrow$ the opposite implication does not hold true ■