# Centro de Investigación en Matemáticas, A.C.

CIMAT

# Maestría en Ciencias con especialidad en Matemáticas Aplicadas

## Optimización

Tarea 3

Profesor:

Dr. Oscar S. Dalmau Cedeño

Alumno:

Gabriel Tirado Pruneda

16 de febrero de 2021

**Exercise 1** *Is the set* $S = \{a \in \mathbb{R}^k | p(0) = 1, |p(t)| \leq 1 \text{ for } t \in [\alpha, \beta]\}$, *where* $p(t) = a_1 + a_2 t + \cdots + a_k t^{k-1}$ *convex?*

**Solution.** *Para* $a = [a_1, \ldots, a_k]^T \in \mathbb{R}^k$, *definimos*

$$p_a(t) := a_1 + a_2 t + \cdots + a_k t^{k-1}.$$

*Sean* $x, y \in S$ *entonces para* $t \in [\alpha, \beta]$ *se cumple*

$$p_x(0) = x_1 = 1 \qquad y \qquad p_y(0) = y_1 = 1$$

*además,*

$$|p_x(t)| \leq 1, \qquad |p_y(t)| \leq 1.$$

*Ahora, sea* $\gamma\colon [0,1] \to \mathbb{R}^k$ *el segmento de recta entre* $x$ *y* $y$ *definido por*

$$\gamma(\lambda) = (1-\lambda)x + \lambda y,$$

*y tomemos* $z \in \gamma([0,1])$. *Entonces, existe* $\lambda^* \in [0,1]$ *tal que*

$$\gamma(\lambda^*) = (1-\lambda^*)x + \lambda^* y = z$$

*Luego,*

$$p_z(t) = (1-\lambda^*)x_1 + \lambda^* y_1 + [(1-\lambda^*)x_2 + \lambda^* y_2]\, t + \cdots + [(1-\lambda^*)x_k + \lambda^* y_k]\, t^{k-1}$$

*con lo que*

$$p_z(0) = (1-\lambda^*)x_1 + \lambda^* y_1 = 1 - \lambda^* + \lambda^* = 1.$$

*Por otro lado,*

$$\begin{aligned}
|p_z(t)| &= |(1-\lambda^*)x_1 + \lambda^* y_1 + [(1-\lambda^*)x_2 + \lambda^* y_2]\, t + \cdots + [(1-\lambda^*)x_k + \lambda^* y_k]\, t^{k-1}| \\
&= |(1-\lambda^*)(x_1 + x_2 t + \cdots + x_k t^{k-1}) + \lambda^*(y_1 + y_2 t + \cdots + y_k t^{k-1})| \\
&\leq |(1-\lambda^*)(x_1 + x_2 t + \cdots + x_k t^{k-1})| + |\lambda^*(y_1 + y_2 t + \cdots + y_k t^{k-1})| \\
&= (1-\lambda^*)|x_1 + x_2 t + \cdots + x_k t^{k-1}| + \lambda^* |y_1 + y_2 t + \cdots + y_k t^{k-1}| \\
&= (1-\lambda^*)|p_x(t)| + \lambda^* |p_y(t)| \\
&\leq 1 - \lambda^* + \lambda^* \\
&= 1.
\end{aligned}$$

*Así, si* $t \in [\alpha, \beta]$, *entonces* $z \in S$ *y por lo tanto* $S$ *es convexo.*

**Exercise 2** *Suppose $f$ is convex, $\lambda_1 > 0$ and $\lambda_2 \leq 0$ with $\lambda_1 + \lambda_2 = 1$, and let $x_1, x_2 \in \operatorname{dom} f$. Show that the inequality*

$$f(\lambda_1 x_1 + \lambda_2 x_2) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

*always holds.*

**Solution.** *Sean $x_1, x_2 \in \operatorname{dom} f$, expresamos a $x_1$ como:*

$$x_1 = \mu_1 (\lambda_1 x_1 + \lambda_2 x_2) + \mu_2 x_2,$$

*donde*

$$\mu_1 = \frac{1}{\lambda_1}, \qquad \mu_2 = -\frac{\lambda_2}{\lambda_1}.$$

*Claramente, $\mu_1 > 0$ y $\mu_2 \geq 0$ y, además*

$$\mu_1 + \mu_2 = \frac{1}{\lambda_1} - \frac{\lambda_2}{\lambda_1} = \frac{1 - \lambda_2}{\lambda_1} = \frac{\lambda_1}{\lambda_1} = 1.$$

*Luego, como $f$ es convexa, por la desigualdad de Jensen*

$$\begin{aligned}
f(x_1) &= f(\mu_1 (\lambda_1 x_1 + \lambda_2 x_2) + \mu_2 x_2) \\
&\leq \mu_1 f(\lambda_1 x_1 + \lambda_2 x_2) + \mu_2 f(x_2) \\
&\leq \frac{1}{\lambda_1} f(\lambda_1 x_1 + \lambda_2 x_2) - \frac{\lambda_2}{\lambda_1} f(x_2) \\
\lambda_1 f(x_1) &\leq f(\lambda_1 x_1 + \lambda_2 x_2) - \lambda_2 f(x_2) \\
\lambda_1 f(x_1) + \lambda_2 f(x_2) &\leq f(\lambda_1 x_1 + \lambda_2 x_2).
\end{aligned}$$

**Exercise 3** *Show that the following function $f : \mathbb{R}^n \to \mathbb{R}$ is convex.*

$$f(x) = -\exp(-g(x))$$

*where $g : \mathbb{R}^n \to \mathbb{R}$ has a convex domain and satisfies*

$$\begin{bmatrix} \nabla^2 g(x) & \nabla g(x) \\ \nabla^T g(x) & 1 \end{bmatrix} \succeq 0$$

*for $x \in \operatorname{dom} g$.*

**Solution.**

$$\frac{\partial}{\partial x_i} f(x) = \exp(-g(x)) \frac{\partial}{\partial x_i} g(x)$$

$$
\begin{aligned}
\frac{\partial^2}{\partial x_j \partial x_i} &= -\exp(-g(x))\left[\frac{\partial}{\partial x_j} g(x)\right]\left[\frac{\partial}{\partial x_i} g(x)\right] + \exp(-g(x))\frac{\partial^2}{\partial x_j \partial x_i} g(x) \\
&= \exp(-g(x))\left[\frac{\partial^2}{\partial x_j \partial x_i} g(x) - \left[\frac{\partial}{\partial x_j} g(x)\right]\left[\frac{\partial}{\partial x_i} g(x)\right]\right]
\end{aligned}
$$

*Luego,*

$$
\nabla^2 f(x) = \exp(-g(x))\left[\nabla^2 g(x) - \nabla g(x) \nabla g(x)^T\right]
$$

*dado que* $\exp(-g(x))$ *siempre es positivo, resta verificar que la matriz*

$$
\nabla^2 g(x) - \nabla g(x) \nabla g(x)^T
$$

*es semidefinida positiva. Ahora, sea* $x \in \mathbb{R}^n$ *entonces, convenientemente, elegimos el vector* $z = (x^T, -x^T \nabla g) \in \mathbb{R}^{n+1}$, *por hipótesis*

$$
\begin{aligned}
0 &\leq \left[x^T, -x^T \nabla g\right] \begin{bmatrix} \nabla^2 g & \nabla g \\ \nabla g^T & 1 \end{bmatrix} \begin{bmatrix} x \\ -\nabla g^T x \end{bmatrix} \\
&= \left[x^T \nabla^2 g - x^T \nabla g \nabla g^T, x^T \nabla g - x^T \nabla g\right] \begin{bmatrix} x \\ -\nabla g^T x \end{bmatrix} \\
&= x^T \nabla^2 g x - x^T \nabla g \nabla g^T x \\
&= x^T \left(\nabla^2 g(x) - \nabla g(x) \nabla g(x)^T\right) x
\end{aligned}
$$

*por lo que* $\nabla^2 g(x) - \nabla g(x) \nabla g(x)^T$ *es semidefinida positiva, y por consiguiente,* $\nabla^2 f(x)$ *es semidefinida positiva, luego* $f$ *es convexa.*

**Exercise 4** *Show that* $f(x, y) = x^2/y, y > 0$ *is convex.*

**Solution.** *Comenzamos calculando el gradiente y el hessiano de* $f$*:*

$$
\nabla f(x, y) = \left[\frac{2x}{y}, -\frac{x^2}{y^2}\right]
$$

$$
\nabla^2 f(x, y) = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}
$$

*Los menores del hessiano son*

$$
\frac{2}{y} \qquad y \qquad \frac{4x^2}{y^4} - \frac{4x^2}{y^4} = 0
$$

*los cuales son no negativos en* $\mathbb{R} \times \mathbb{R}^+$, *con lo que el hessiano es semidefinido positivo y por consiguiente,* $f$ *es convexa en* $\mathbb{R} \times \mathbb{R}^+$.

**Exercise 5** *Consider the function $f(x_1, x_2) = (x_1 + x_2^2)^2$. At the point $x^T = [1, 0]$ we consider the search direction $p^T = [-1, 1]$. Show that $p$ is a descent direction and find all minimizers of the function.*

**Solution.**

$$f(1, 0) = 1$$

*Sea $y = x + \varepsilon p = [1 - \varepsilon, \varepsilon]^T$, entonces*

$$f(y) = (1 - \varepsilon + \varepsilon^2)^2$$

*así, buscamos $\varepsilon > 0$ tal que*

$$(1 - \varepsilon + \varepsilon^2)^2 < 1$$

*o equivalentemente*

$$-\varepsilon(1 - \varepsilon) < 0$$

*lo cual se cumple si $\varepsilon < 1$. Por lo tanto, en el punto $x^T = [1, 0]$, $p^T = [-1, 1]$ es una dirección de descenso para un tamaño de paso $\varepsilon \in (0, 1)$.*

*Ahora, calculamos el gradiente de $f$ y buscamos puntos estacionarios*

$$\nabla f(x_1, x_2) = [2(x_1 + x_2^2), 4x_2(x_1 + x_2^2)]$$

$$\nabla f(x_1, x_2) = 0 \qquad \Rightarrow \qquad \begin{cases} 2(x_1 + x_2^2) = 0 & \Rightarrow \qquad x_1 = -x_2^2 \\ 4x_2(x_1 + x_2^2) = 0 & \Rightarrow \qquad x_1 = -x_2^2 \end{cases}$$

*Así, $(x_1, x_2)$ es un punto estacionario si $x_1 = -x_2^2$. Luego, el hessiano de $f$ es*

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 4x_2 \\ 4x_2 & 4x_1 + 12x_2^2 \end{bmatrix}$$

*Si evaluamos en $x_1 = -x_2^2$ obtenemos*

$$\nabla^2 f(-x_2^2, x_2) = \begin{bmatrix} 2 & 4x_2 \\ 4x_2 & 8x_2^2 \end{bmatrix}$$

*los menores del hessiano son $2$ y $0$, al ser no negativos, se cumple que el hessiano es semidefinido positivo, con lo que no se puede concluir si los puntos $(-x_2^2, x_2)$ son mínimos.*

**Exercise 6** *Find all the values of the parameter a such that $[1, 0]^T$ is the minimizer or maximizer of the function $f(x_1, x_2) = a^3 x_1 e^{x_2} + 2a^2 \log(x_1 + x_2) - (a + 2)x_1 +$*

$8ax_2 + 16x_1x_2$

**Solution.** *El gradiente de $f$ es:*

$$\nabla f\left(x_1, x_2\right) = \left[a^3 e^{x_2} + \frac{2a^2}{x_1 + x_2} - a + 16x_2 - 2, a^3 x_1 e^{x_2} + \frac{2a^2}{x_1 + x_2} + 8a + 16x_1\right]$$

*Evaluando en $[x_1, x_2] = [1, 0]$ e igualando a 0*

$$\nabla f(1,0) = 0 \qquad \Rightarrow \qquad \begin{cases} a^3 + 2a^2 - a - 2 = 0 & \Rightarrow \quad (a-1)(a+1)(a+2) = 0 \\ a^3 + 2a^2 + 8a + 16 = 0 & \Rightarrow \quad (a+2)(a^2+8) = 0 \end{cases}$$

*Por lo tanto, para que $\nabla f(1,0) = 0$ necesariamente $a = -2$. Así,*

$$f\left(x_1, x_2\right) = -8x_1 e^{x_2} + 8 \log\left(x_1 + x_2\right) - 16x_2 + 16x_1x_2$$

*y*

$$\nabla f\left(x_1, x_2\right) = \left[-8e^{x_2} + \frac{8}{x_1 + x_2} + 16x_2, -8x_1 e^{x_2} + \frac{8}{x_1 + x_2} - 16 + 16x_1\right]$$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} -\frac{8}{(x_1+x_2)^2} & -8e^{x_2} - \frac{8}{(x_1+x_2)^2} + 16 \\ -8e^{x_2} - \frac{8}{(x_1+x_2)^2} + 16 & -8x_1 e^{x_2} - \frac{8}{(x_1+x_2)^2} \end{bmatrix}$$

*Así, evaluando el hessiano en $[1, 0]^T$*

$$\nabla^2 f(1,0) = \begin{bmatrix} -8 & 0 \\ 0 & -16 \end{bmatrix}$$

*se obtiene una matriz negativa definida, y por consiguiente, $[1, 0]^T$ es un máximo de $f$.*

**Exercise 7** *Consider the sequence $x_k = 1 + 1/k!, k = 0, 1, \cdots$. Does this sequence converge linearly to 1? Justify your response.*

**Solution.** *Primero, comenzamos calculando $\beta$ para $p = 1$:*

$$
\begin{aligned}
\beta = \lim_{k\to\infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^p} &= \lim_{k\to\infty} \frac{\left|\left(1 + \frac{1}{(k+1)!}\right) - 1\right|}{\left|\left(1 + \frac{1}{k!}\right) - 1\right|} \\
&= \lim_{k\to\infty} \frac{\left|\frac{1}{(k+1)!}\right|}{\left|\frac{1}{k!}\right|} \\
&= \lim_{k\to\infty} \frac{k!}{(k+1)!} \\
&= \lim_{k\to\infty} \frac{1}{(k+1)} \\
&= 0.
\end{aligned}
$$

*Dado que $\beta = 0$, la sucesión $\{x_k\}_{k\in N\cup\{0\}}$ converge superlinealmente a $1$.*

**Exercise 8** *Show that $f(x) = \log \sum_{i=1}^{n} \exp(x_i)$ is convex.*

**Solution.** *Primero calculamos el hessiano de $f$*

$$
\frac{\partial}{\partial x_j} f(x) = \frac{\exp(x_j)}{\sum_{i=1}^{n} \exp(x_i)}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial x_k \partial x_j} f(x) &= -\frac{\exp(x_j + x_k)}{(\sum_{i=1}^{n} \exp(x_i))^2} \\
\frac{\partial^2}{\partial x_j \partial x_j} f(x) &= \frac{\exp(x_j)\sum_{i=1}^{n} \exp(x_i) - \exp(x_j)\exp(x_j)}{(\sum_{i=1}^{n} \exp(x_i))^2} \\
&= \frac{\exp(x_j)}{\sum_{i=1}^{n} \exp(x_i)} - \frac{\exp(2x_j)}{(\sum_{i=1}^{n} \exp(x_i))^2}
\end{aligned}
$$

*O de forma equivalente*

$$
\frac{\partial^2}{\partial x_k \partial x_j} f(x) = \delta_{kj} \frac{\exp(x_j)}{\sum_{i=1}^{n} \exp(x_i)} - \frac{\exp(x_j + x_k)}{(\sum_{i=1}^{n} \exp(x_i))^2} = [H(x)]_{k,j}
$$

*donde $\delta_{kj}$ es la delta de Kronecker. Ahora, para $y \in \mathbb{R}^n$ tenemos*

$$
y^T H(x) y = \frac{\left(\sum_i \exp(x_i) y_i^2\right)\left(\sum_i \exp(x_i)\right) - \left(\sum_i y_i \exp(x_i)\right)^2}{\left(\sum_i \exp(x_i)\right)^2} \tag{1}
$$

*Por la desigualdad de Cauchy-Schwarz para el numerador de (1):*

$$
\left(\sum_i y_i \exp(x_i)\right)^2 \leq \left(\sum_i \exp(x_i) y_i^2\right)\left(\sum_i \exp(x_i)\right)
$$

*Además, como $(\sum_i \exp(x_i))^2 > 0$, se cumple que $y^T H(x) y \geq 0$. Así, $H(x)$ es semipositiva definida y finalmente $f(x)$ es convexa.*

**Exercise 9** *Show that $f(x) = \log \sum_{i=1}^{n} \exp\left(g_i(x)\right) : \mathbb{R} \to \mathbb{R}$ is convex if $g_i : \mathbb{R} \to \mathbb{R}$ are convex*

**Solution.**

**Exercise 10** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Show that $f$ is convex over a nonempty convex set $C$ if and only if*

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0, \forall x, y \in C$$

*Note: the proof we have is only for the case $(\Rightarrow)$*

**Solution.** *Sabemos que $f$ es convexa si y solo si se satisface*

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

*Ahora, dado que $x$ y $y$ son arbitrarias, también se cumple*

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

*sumando ambas desigualdades*

$$\begin{aligned}
f(y) + f(x) &\geq f(x) + \nabla f(x)^T (y - x) + f(y) + \nabla f(y)^T (x - y) \\
0 &\geq \nabla f(x)^T (y - x) + \nabla f(y)^T (x - y) \\
0 &\leq \nabla f(x)^T (x - y) - \nabla f(y)^T (x - y) \\
0 &\leq \left(\nabla f(x) - \nabla f(y)\right)^T (x - y).
\end{aligned}$$