Hoja de Trabajo 1 - Introducción a Optimización para Ciencia de Datos

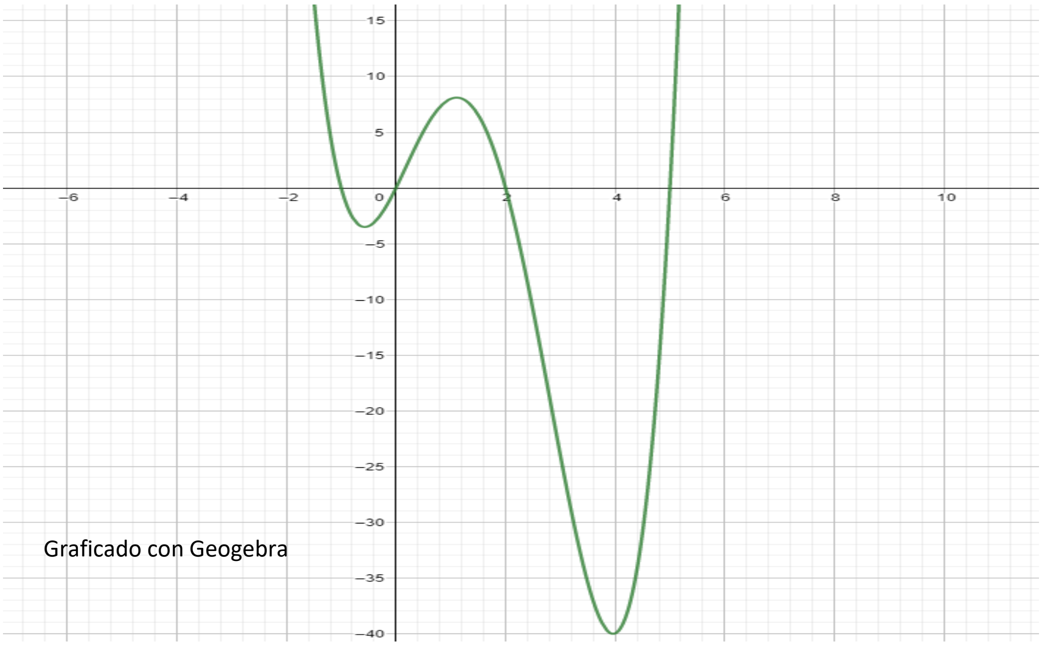
Eduardo Andrés Santizo Olivet (21007361)

1. Dada la función $f : \mathbb{R} \to \mathbb{R}_\ni$

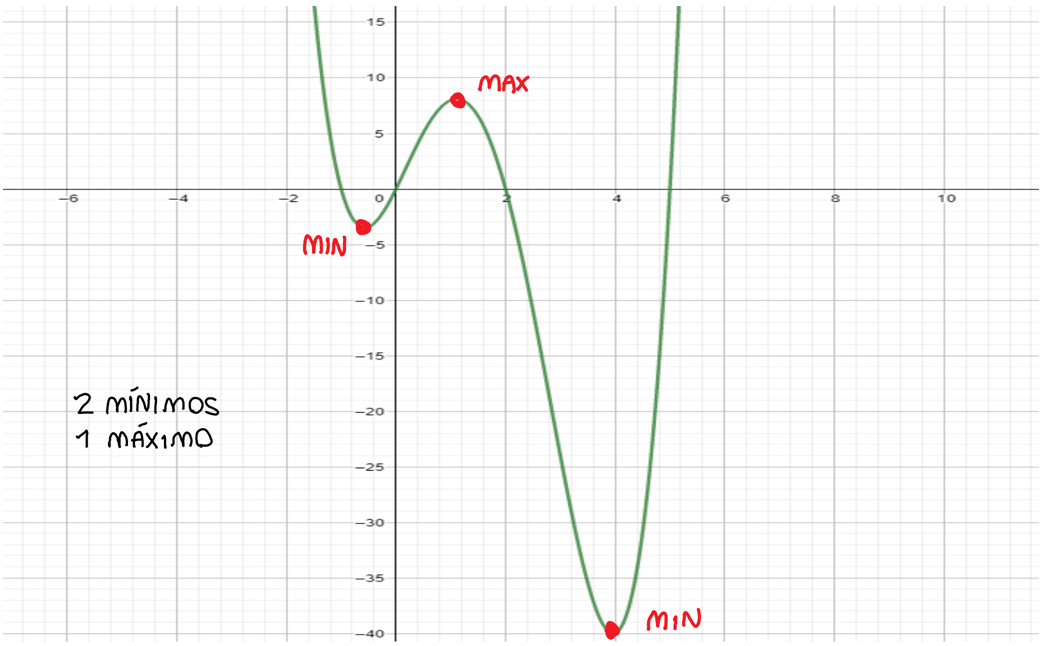
$$f(x) = x^4 - 6x^3 + 3x^2 + 10x.$$

   a) Utilice cualquier software para graficar dicha función.
   b) Localice los *máximos* y *mínimos locales*, si los hay.
   c) ¿Existe un mínimo global? ¿y un máximo global? justifique su respuesta.

A.



B.

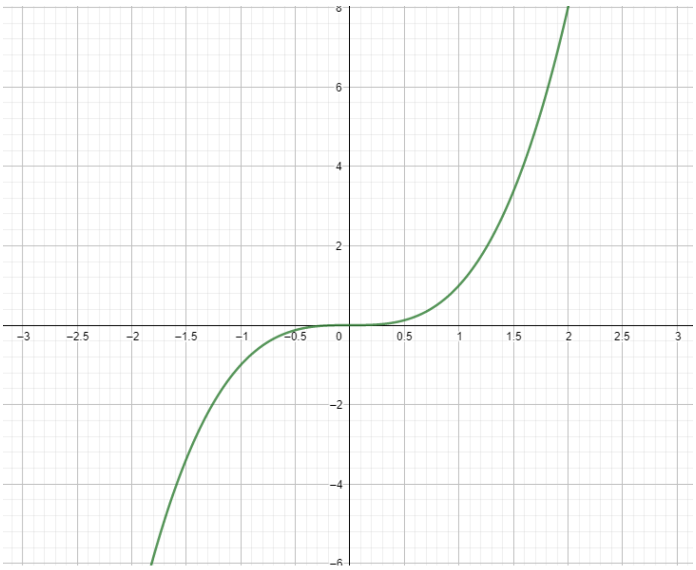


C. En la función existe un mínimo global ubicado en aproximadamente X = 4. En este valor de X se encuentra "la altura" más baja que adquiere la función sobre todo su dominio. Por otro lado, esta función no cuenta con un máximo global, ya que luego de pasar por sus dos mínimos (uno a la izquierda y otro a la derecha), la función se torna monotónica creciente, lo que implica que su valor de altura máximo es de infinito. El máximo que presenta la función en la cercanía a X = 1 es únicamente un máximo local.

2. Construya un ejemplo de una función (de una variable real) que *no* tenga mínimo global *ni* máximo global.

$f(x) = x^3$

Esta función es monotónica creciente cuando X > 0 y monotónica decreciente cuando X < 0. Esto implica que su valor máximo (si no se acota la función) está indefinido, al igual que su valor mínimo.



3. Considere el problema de optimización:

$$\begin{aligned} \text{mín} \quad & f(x,y) = x \\ \text{s.t.} \quad & x^2 + y^2 \leq 4 \\ & x^2 \geq 1. \end{aligned}$$

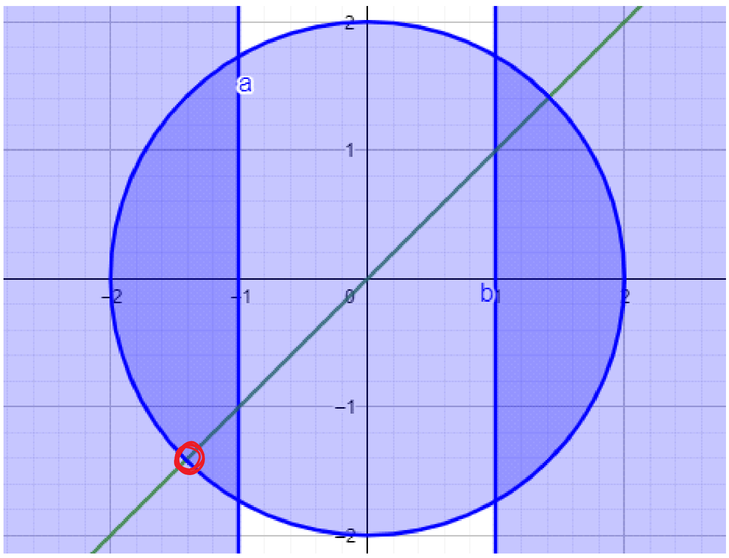
a) ¿Es este problema de optimización *lineal*? justifique su respuesta.

b) Grafique la *región factible* para dicho problema.

c) Utilice la gráfica del inciso (*a*) para determinar la *solución* del problema de optimización.

A. No se trata de un problema lineal ya que aunque la función que se está minimizando es lineal, las restricciones a las que está sujeta la función son no lineales.

B. La región factible consiste de las secciones coloreadas en azul oscuro. La función a minimizar consiste de la línea verde.



C. El valor mínimo de la función sujeta a la región factible dada, se encuentra en la intersección presente en el extremo inferior izquierdo del círculo con radio 2 (ver punto rojo arriba). Para encontrar el valor específico de la función se realiza lo siguiente:

en el extremo inferior izquierdo del círculo con radio 2 (ver punto rojo arriba). Para encontrar el valor específico de la función se realiza lo siguiente:

SE DESPEJA y

$$x^2 + y^2 = 4$$
$$y = \sqrt{4 - x^2}$$

SE IGUALA A f(x)

$$x = \sqrt{4 - x^2}$$
$$x^2 = 4 - x^2$$
$$2x^2 = 4$$
$$x = \sqrt{2}$$

SE OBTIENEN LAS COORDS. DE LA INTERSECCIÓN

$$x = \sqrt{2}$$
$$y = \sqrt{4 - (\sqrt{2})^2} = \sqrt{4-2} = \sqrt{2}$$

$\longrightarrow$ MÍNIMO $f(\sqrt{2}) = \sqrt{2}$

4. A continuación se le presentan un conjunto de datos, se desea construir un *modelo de regresión lineal* para predecir la Presión Arterial Sistólica en función del Peso y la Edad a partir de un conjunto de 7 personas seleccionadas aleatoriamente.

Observación	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Edad (en años)	16	25	39	45	49	64	70
Peso (en lb)	140	149	165	170	165	159	144
Presión Arterial Sistólica (en mm Hg)	125	129	127	150	161	144	132

a) Escriba el conjunto de datos en la forma $D = \{(\alpha_j, y_j) \mid j = 1, \ldots, m\}$, identificando claramente cada uno de sus elementos.

b) Plantee un *problema de optimización* que le permita resolver la situación planteada. Escriba la función objetivo L_D en forma matricial.

c) Utilizando cualquier software, *revuelva* el problema de optimización planteado en el inciso anterior. ¿Qué tipo de solución (global o local) encontró? justifique su respuesta.

d) Escriba la función $\phi(\alpha)$ que utilizará para realizar las predicciones.

A. $\alpha_j \; (A) = \begin{array}{c} \begin{array}{cc} \alpha_{j1} & \alpha_{j2} \end{array} \\ \begin{bmatrix} 16 & 140 \\ 25 & 149 \\ 39 & 165 \\ 45 & 170 \\ 49 & 165 \\ 64 & 159 \\ 70 & 144 \end{bmatrix} \end{array}$ $\qquad y_j \; (y) = \begin{bmatrix} 125 \\ 129 \\ 127 \\ 150 \\ 161 \\ 144 \\ 132 \end{bmatrix}$

Parámetros o features — Valor a predecir

B. $\boxed{\min_{x \in \mathbb{R}^2} L_D(x)}$

$$L_D(x) = \frac{1}{2m} \sum_{j=1}^{m} (\alpha_j^T x + B - y_j)^2 = \frac{1}{2m} \| Ax + B - y \|^2$$

$$= \frac{1}{2m} \left\| \underbrace{\begin{bmatrix} 16 & 140 \\ 25 & 149 \\ 39 & 165 \\ 45 & 170 \\ 49 & 165 \\ 64 & 159 \\ 70 & 144 \end{bmatrix}}_{A} \underbrace{[x_1 \; x_2]}_{x} + \underbrace{\begin{bmatrix} B_0 \\ \vdots \\ B_0 \\ \vdots \\ B_0 \end{bmatrix}}_{B_0} - \underbrace{\begin{bmatrix} 125 \\ 129 \\ 127 \\ 150 \\ 161 \\ 144 \\ 132 \end{bmatrix}}_{y} \right\|^2$$

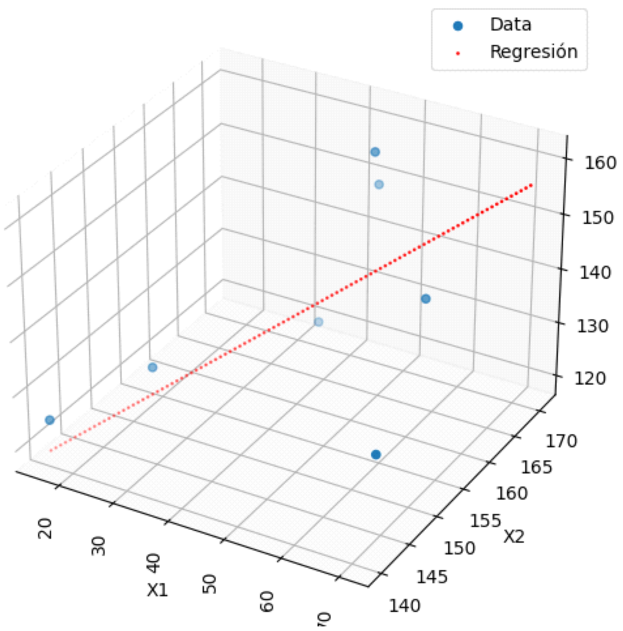
Se expanden las dimensiones copiando el valor de B0 en todas las filas

D. $\phi = \alpha^T x + B$

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad B = B_0$ $\Big\}$ MODELO

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad B = B_0 \quad \Big\} \text{ MODELO}$$

filas

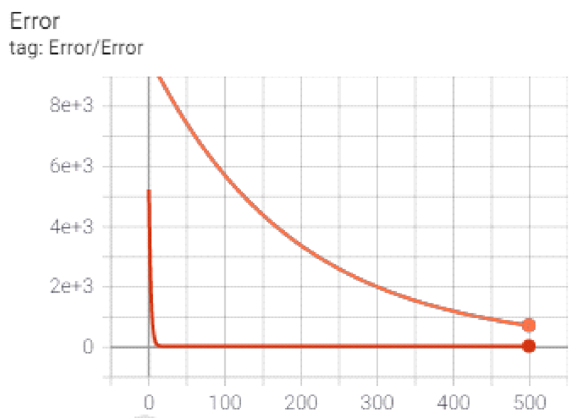


C. Luego de crear un modelo de regresión utilizando Tensorflow se encontró la siguiente solución o el siguiente modelo:

$$\phi = \alpha^T X + B$$

$$X = \begin{bmatrix} 0.199 \\ 0.830 \end{bmatrix} \qquad B = 0.006$$

Se puede argumentar que estos "mínimos" encontrados consisten de un mínimo global debido al carácter convexo de la función de costo L_D empleada. Por lo tanto, aunque en diferentes corridas del algoritmo se hayan obtenido valores ligeramente diferentes para los parámetros, técnicamente, bajo las condiciones correctas, todas las soluciones deberían de consistir de un valor muy cercano a los parámetros encontrados.



Esto puede ser parcialmente comprobado al observar la gráfica de error generada durante el entrenamiento, donde se observa que no importando la configuración empleada, la gráfica parece dirigirse a un mismo valor de costo mínimo.