Departamento de Ciencia de la Computación e Inteligencia Artificial

Alumno:	
Grupo teoría:	
DNI:	

Ejercicio RESUELTO teoría Bloque Resolución Ecuaciones. Matemáticas II, 2019/20 GRUPO ?

Instrucciones generales:
Todos los ejercicios deben estar bien explicados. Entregar todas las hojas firmadas y con DNI.

EJERCICIO 1. Resolver el siguiente sistema no-lineal

$$x^2 + y^2 = 0$$

$$x - y = 0$$

Su solución obvia es $x^* = 0$, $y^* = 0$. Pero ¿cómo la encontraría el método de descenso por gradiente con **punto inicial** $x_0 = 2$, $y_0 = 2$ y **constante de ajuste** $\gamma = 0.1$? Realizar 3 iteraciones del método. Guardar cuatro decimales en cada iteración.

Para ello seguiremos los siguientes pasos

1. **Formular la función de error:** Como hay dos ecuaciones, la función de error tendrá dos sumandos

$$E(x, y) = \frac{1}{2}[(x^2 + y^2 - 0)^2 + (x - y - 0)^2] = \frac{1}{2}[(x^2 + y^2)^2 + (x - y)^2]$$

2. **Calcular el gradiente de la función de error:** $\nabla E(x, y) = [\frac{\partial E}{\partial x} \quad \frac{\partial E}{\partial y}]^T$ que viene dado por

$$\frac{\partial E}{\partial x} = (x^2 + y^2)2x + (x - y)$$
$$\frac{\partial E}{\partial y} = (x^2 + y^2)2y - (x - y)$$

3. **Hacer** $t = 0$ **e iterar en la dirección opuesta al gradiente** según la ecuación:

$$\begin{bmatrix} x_{t+1} \\ y_{y+1} \end{bmatrix} = \begin{bmatrix} x_t \\ y_y \end{bmatrix} - \gamma \nabla E(x_t, y_t)$$

Que para nuestro problema queda así:

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \end{bmatrix} - \gamma \begin{bmatrix} (x_t^2 + y_t^2)2x_t + (x_t - y_t) \\ (x_t^2 + y_t^2)2y_t - (x_t - y_t) \end{bmatrix}$$

Iteramos para $t = 1$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - 0.1 \begin{bmatrix} (2^2 + 2^2)4 + (2 - 2) \\ (2^2 + 2^2)4 - (2 - 2) \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - 0.1 \begin{bmatrix} 32 \\ 32 \end{bmatrix} = \begin{bmatrix} -1.2 \\ -1.2 \end{bmatrix}$$

Vemos que nos aparta del óptimo (es debido al elevado valor de $\gamma = 0.1$). Veamos si podemos converger.

Iteramos para $t = 2$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1.2 \\ -1.2 \end{bmatrix} - 0.1 \begin{bmatrix} ((-1.2)^2 + (-1.2)^2)2(-1.2) + (-1.2 + 1.2) \\ ((-1.2)^2 + (-1.2)^2)2(-1.2) - (-1.2 + 1.2) \end{bmatrix} = \begin{bmatrix} -1.2 \\ -1.2 \end{bmatrix} - 0.1 \begin{bmatrix} -6.912 \\ -6.912 \end{bmatrix} = \begin{bmatrix} -0.5088 \\ -0.5088 \end{bmatrix}$$

Nos volvemos a acercar al óptimo.

Iteramos para $t = 3$

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} -0.5088 \\ -0.5088 \end{bmatrix} - 0.1 \nabla E(-0.5088, -0.5088) = \begin{bmatrix} -0.5088 \\ -0.5088 \end{bmatrix} - 0.1 \begin{bmatrix} -0.5268 \\ -0.5268 \end{bmatrix} = \begin{bmatrix} -0.4561 \\ -0.4561 \end{bmatrix}$$

Después de la tercera iteración vemos que hemos recuperado terreno y volvemos a ser menos negativos. Si fijamos un error $\epsilon = 10^{-4}$ de manera que paremos cuando $\sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2} \leq \epsilon$ entonces necesitaríamos 1836 iteraciones y llegaríamos a [-0.026 -0.026].

4. EN FORMA DE TABLA

Una forma resumida de poner estos datos es cumplimentar una tabla con tres columnas. En la primera ponemos el x,y de la iteración anterior, en la segunda el gradiente Grad(x,y) y en la tercera el nuevo x',y' tras multiplicar el Grad(x,y) por gamma y restar el resultado a la primera columna:

	x,y	Grad(x,y)	x',y'
1	(2,2)	(32,32)	(-1.2,-1.2)
2	(-1.2,-1.2)	(-6.192,-6.192)	(-0.5088,-0.5088)
3	(-0.5088,-0.5088)	(-0.5268,-0.5268)	(-0.4561,-0.4561)
4	(-0.4561,-0.4561)		

5. INTERPRETACIÓN GRÁFICA

UTILIZANDO MATLAB es fácil ver que tenemos una función de error convexa (eso se puede preguntar en el examen y entonces habría que evaluar la Hessiana en 0,0 para demostrar que es un mínimo). Visto el recorrido del descenso por gradiente desde arriba (fcontour) tenemos, para 1836 iteraciones, lo siguiente: