OPTIMIZACIÓN II
Práctica 4

Francisco J. Aragón Artacho
Dpto. Matemáticas, Universidad de Alicante

24-25 de marzo de 2022

En esta práctica aplicaremos el método de Gauss–Newton para buscar una solución del sistema de ecuaciones:

$$\left\{\begin{array}{c} x_1^2 + x_2x_3x_4 - 2 = 0 \\ x_1 + x_1x_2 - x_3x_4 - 1 = 0 \\ -x_1 + x_1^2 + 2x_2 - x_3^3 + x_4^2 - 2 = 0 \\ x_1 e^{x_2+x_3^2-x_4} - e = 0 \\ x_2\cos(\pi x_1) - x_3\sin(\pi x_4) + 1 = 0 \end{array}\right\}$$

Así, nuestro objetivo será minimizar la función

$$f(x) := \frac{1}{2}\|g(x)\|^2 = \frac{1}{2}\sum_{i=1}^{5} g_i(x)^2,$$

donde $g = (g_1, g_2, g_3, g_4, g_5)^T : \mathbb{R}^4 \to \mathbb{R}^5$*, siendo* $g_i : \mathbb{R}^4 \to \mathbb{R}$ *cada una de las funciones que define el sistema de ecuaciones de arriba.*

Crea un programa en Python llamado `Pr4.m` que realice lo siguiente:

1. Define la función `f` y su gradiente `grad_f`.

2. Tomando como punto inicial $(0, 0.2, 0.7, 0.5)^T$, crea un bucle que ejecute el método de Gauss–Newton puro el número de iteraciones que sea necesario (hasta un máximo de 1000) para que la norma de `grad_f` sea menor o igual que 10^{-7}. Muestra solo en pantalla el número de iteraciones, la solución obtenida, el valor de la función f y la norma de su gradiente.

3. Tomando el mismo punto inicial, crea un bucle que ejecute el método de Gauss–Newton con búsqueda lineal y tamaño de paso satisfaciendo las condiciones de Wolfe[1] el número de iteraciones que sea necesario (hasta un máximo de 1000) para que la norma de `grad_f` sea menor o igual que 10^{-7}. Muestra solo en pantalla el número de iteraciones, la solución obtenida, el valor de la función f y la norma de su gradiente. Compara con el resultado obtenido en el apartado anterior.

4. Repite el experimento del apartado anterior usando esta vez la dirección del método del descenso más rápido en vez de la de Gauss–Newton y compara el resultado.

5. Repite los experimentos de los apartados 2 y 3 usando $(-0.3, 1.5, -0.2, 0.8)^T$ como punto inicial. ¿Qué metodo obtiene una mejor solución?

[1] Haz las modificaciones necesarias en la función Wolfe de la Práctica 3.

6. (OPCIONAL) Explica por qué fallan los métodos anteriores si tomamos como punto inicial $(0,0,0,0)^T$. Programa el método de Lebenberg–Marquardt tomando $\Delta_k := \mu_k I$, donde I es la matriz identidad y $\mu_k := \|\nabla f(x_k)\|^2$, con tamaño de paso satisfaciendo las condiciones de Wolfe y usando la mismas reglas de parada de los apartados 2 y 3. Aplícalo al punto inicial $(0,0,0,0)^T$.