

Optimización II

Hoja de ejercicios Parte II: Problemas sin restricciones

Francisco Javier Mercader Martínez

- 1) Consideramos la función $f(x_1, x_2) = 2x_1^2 + x_2^2 + x_1x_2 - x_1 - 3x_2$
 - a) Aplicar dos iteraciones del método del gradiente con búsquedas lineales exactas para minimizar f empezando desde el punto $x^{(0)} = (1, 1)^T$.
 - b) ¿Son ortogonales las direcciones de búsqueda empleadas en esas dos iteraciones?
 - c) Demuestra en general que las direcciones consecutivas de búsqueda en el método del gradiente con búsquedas lineales exactas son ortogonales.
 - d) A partir de dicho resultado, ¿cómo se comportará el método del gradiente en un problema de optimización sin restricciones en el que el número de condición de la función objetivo sea grande, siendo dicha función estrictamente convexa?
- 2) Resuelve el ejercicio anterior usando el método de gradiente conjugado para la misma inicialización $x^{(0)} = (1, 1)^T$.
- 3) **Ejercicio para entregar.** Consideremos el problema de minimizar la función cuadrática

$$f(x) = x^2.$$

Tomemos la sucesión de puntos $x_1 = 2, \quad x_n = 1 + \frac{1}{2^{n-1}}, \quad n = 2, 3, \dots$

- 1) Comprueba a mano que la sucesión x_n se obtiene usando un método de descenso del tipo

$$x_{n+1} = x_n + \alpha_n d_n, \quad \alpha_n = \frac{1}{2^n}, \quad d_n = -1$$

Un cálculo directo muestra que $x_n \rightarrow 1$. Sin embargo $x^* = 1$ no es el mínimo de f .

- 2) Ilustra esta situación gráficamente en Python.
- 3) Demuestra a mano que para la elección de α_n y el algoritmo anterior existe $0 < c_1 < 1$ que cumple la condición de Armijo

$$f(x_n + \alpha_n d_n) < f(x_n) + c_1 \alpha_n f'(x_n) d_n$$

- 4) Demuestra a mano que no existe $c_1 < c_2 < 1$ de modo que se cumpla la condición de curvatura de Wolfe

$$f'(x_n + \alpha_n d_n) d_n \geq c_2 f'(x_n) d_n$$

- 5) Implementa el algoritmo anterior usando un paso que cumpla las condiciones de Wolfe. Puedes usar ChatGPT, para lo cual te sugiero el prompt *May you please provide me with a code for the implementation of the Wolfe conditions?*

- 6) Implementa el algoritmo de descenso anterior, incluyendo paso que cumplen las condiciones de Wolfe, y para los datos del inicio.
- 7) Ya que estamos recibiendo la ayuda de ChatGPT para resolver problemas matemáticos, averigua cómo hemos de comunicarnos con ChatGPT cuando tengamos que transmitirle una fórmula en un prompt.
- 4) Calcula los grafos computaciones en modos forward y backward para las siguientes funciones y calcula, utilizando diferenciación automática, sus gradientes en los puntos que se indican:
 - a) $f(x_1, x_2) = x_1 x_2 + \sin(x_1)$ en el punto $(x_1, x_2) = (1, 1)$
 - b) $f(x_1, x_2) = \log(3x_1^2 + 5x_1 x_2)$ en el punto $(x_1, x_2) = (2, 3)$
 - c) $\omega_1, \omega_2 = (\sigma(2\omega_1 + 3\omega_2) - 1)^2$ en el punto $(\omega_1, \omega_2) = (1, 2)$ siendo $\sigma(s) = \tanh(s)$.

Comprueba los resultados obtenidos con `tensorflow` y `pytorch`.

- 5) Para cada una de las funciones del ejercicios anterior, calcula, usando `tensorflow` y `pytorch`, sus matrices hessianas en los puntos que se indican, e indica si dichas funciones son convexas, estrictamente convexas, o ni una cosa ni otra, en un entorno de dichos puntos. Recuerda que puedes calcular los valores propios de una matriz con el método `eig` de `numpy`. También has de dibujar dichas funciones en un entorno de dichos puntos.
- 6) Sea $L \in \mathbb{N}$ un número natural positivo. Consideramos la función

$$f_L(\omega) = g_L(g_{L-1}(\cdots g_1(\omega))).$$

Supongamos que para cada $1 \leq \ell \leq L$ se tiene que $g'_\ell(\omega) \approx a^\ell$ para todo $\omega \in \mathbb{R}$. Calcula el $\lim_{L \rightarrow \infty} f'_L(\omega)$ en los casos $0 < a < 1$ y $a > 1$. Explica qué le sucederá a un algoritmo de gradiente que utilizase los cálculos anteriores.

- 7) Sea $0 < b < 1$. Consideremos el problema

$$\text{Minimizar en } (x_1, x_2) : \quad f(x_1, x_2) = \frac{1}{2}(x_1^2 + bx_2^2)$$

Se pide:

- a) Tomando como inicialización $x^{(0)} = (2, 2)$ y para $b = 0.01$, realiza a manos dos iteraciones del método de gradiente con búsquedas lineales exactas. Realiza también 100 iteraciones con la implementación del algoritmo de gradiente que vimos en la práctica de Python.
- b) Realiza dos iteraciones con el método de gradiente conjugado.
- c) Tomando como inicialización un punto genérico $x^{(0)} = (x_1^0, x_2^0)$, realiza una iteración con el método de Newton.
- d) A tenor de los resultados obtenidos, razona sobre qué método es más conveniente usar para problemas cuadráticos.
- 8) Consideremos la función $f(x) = 7x - \ln x$. Se pide:
 - a) Realiza tres iteraciones del método de Newton partiendo de $x^{(0)} = 0.2$ e indica cuál crees que podría ser, aproximadamente, el mínimo buscado. Para ello es conveniente que dibujes la función $f(x) = 7x - \ln x$.
 - b) Realiza tres iteraciones del método de Newton partiendo de $x^{(0)} = 0.1$.
 - c) Explica los resultados que has obtenido en los dos apartados anteriores.