Métodos Numéricos II 2022

Lista 03

29.septiembre.2022

1. Implementar los siguientes métodos de descenso gradiente (naïve = tamaño de paso α constante):

   - descenso gradiente naïve con dirección de descenso aleatoria,
   - descenso máximo naïve (*steepest descent*).
   - descenso gradiente con búsqueda en línea *Backtracking* (α no constante).
   - descenso gradiente con dirección de Newton (Hessiano exacto).
   - descenso gradiente con dirección de Newton, con Hessiano aproximado.

   En cada uno de los métodos, su función debe recibir los siguientes argumentos: la función objetivo f, el gradiente de la función objetivo df, $d^2 f$ (cuando sea necesario), un punto inicial $\mathbf{x}_0 \in \mathbb{R}^n$, el tamaño de paso $\alpha > 0$, el número máximo de iteraciones $maxIter$, la tolerancia ε, así como un criterio de paro.
   En los casos que se use Backtracking, $\alpha = \alpha_0$ corresponde al inicio de la búsqueda lineal, y se debe incorporar un parámetro $\rho > 0$ de tasa de decaimiento. En el caso del Backtracking, usted decide si usar las condiciones de Wolfe o de Goldstein.

   Como resultado, sus algoritmos deben devolver: la mejor solución encontrada $best\ \mathbf{x}$ (la última de las aproximaciones calculadas); la secuencia de iteraciones $\mathbf{x}_k$; la secuencia de valores $f(\mathbf{x}_k)$; la secuencia de errores en cada paso (según el error de su criterio de paro).
   Además, es deseable indicar el número de iteraciones efectuadas por el algoritmo, y si se obtuvo o no convergencia del método.

2. Testar los algoritmos del Ejercicio 1 con las siguientes funciones:

   a) La función de Rosembrock 2-dimensional $f : \mathbb{R}^2 \to \mathbb{R}$, dada por
   $$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$
   Punto inicial: $\mathbf{x}_0 = (-1.2, 1)^T$. Óptimo: $\mathbf{x}^* = (1, 1)^T$, $f(\mathbf{x}^*) = 0$.

   b) La función de Wood $f : \mathbb{R}^4 \to \mathbb{R}$, dada por
   $$f(x, y) = 100(x_1^2 - x_2) + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4).$$
   Punto inicial: $\mathbf{x}_0 = (-3, 1, -3, 1)^T$. Óptimo: $\mathbf{x}^* = (1, 1, 1, 1)^T$, $f(\mathbf{x}^*) = 0$.

   c) La función de Rosembrock 100-dimensional $f : \mathbb{R}^{100} \to \mathbb{R}$, dada por
   $$f(\mathbf{x}) = \sum_{i=1}^{99} \left[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\right].$$
   Punto inicial: $\mathbf{x}_0 = (-1.2, 1, 1, \ldots, 1, -1.2, 1)^T$. Óptimo: $\mathbf{x}^* = (1, 1, \ldots, 1)^T$, $f(\mathbf{x}^*) = 0$.

   En cada uno de los casos, hallar un tamaño de paso α que garantice la convergencia de los métodos, y elabore una tabla con las primeras 3 y las últimas 3 aproximaciones $\mathbf{x}_k$ obtenidas.

   Para este tamaño de paso, comparar:

- la solución aproximada obtenida
- el error de aproximación
- la norma del gradiente en la solución

Elabore gráficas que muestren el error de aproximación, en función del número de iteración, y muestre la comparación de la evolución de la convergencia en sus tres métodos. A partir de estas gráficas, discuta cuál de los métodos es más efectivo, en cada caso.

3. Construya una función "suma de gaussianas" 2-dimensional, en la forma

$$f(\mathbf{x}) = -\sum_{i=1}^{k} \exp\big(-\tfrac{1}{2\sigma}||\mathbf{x}-\mathbf{x}_k||_2^2\big),$$

donde $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ son puntos en el rectángulo $[0,8]\times[0,8]$ elegidos de forma aleatoria (distribución uniforme). Use $k=8$, Aquí, $\sigma > 0$ es un parámetro de escala definido por el usuario (que indica qué tan suave se desea la función).

Aplique varias veces el método de descenso gradiente a la función f, con inicializaciones $\mathbf{x}_0$ distintas, de forma que se puedan obtener los diferentes mínimos locales de la función.

Muestre visualizaciones de diferentes secuencias de aproximaciones $\{\mathbf{x}_k\}$ convergiendo a cada uno de los mínimos locales de su función.