# Métodos Numéricos II 2025

Lista 05

30.octubre.2025

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

   - descenso máximo naïve
   - descenso gradiente con backtracking
   - descenso coordenado
   - descenso de Newton
   - descenso 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 $df$, el hessiano $ddf$ (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 $\varepsilon$, 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 sus algoritmos del Ejercicio 1 con las siguientes funciones:

   a) La función $f : \mathbb{R}^2 \to \mathbb{R}$, dada por
   $$f(x, y) = x^4 + y^4 - 4xy + \tfrac{1}{2}y + 1.$$

   b) La función de Rosembrock 2-dimensional $f : \mathbb{R}^2 \to \mathbb{R}$, dada por
   $$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

   c) La función de Rosembrock 10-dimensional $f : \mathbb{R}^{10} \to \mathbb{R}$, dada por
   $$f(\mathbf{x}) = \sum_{i=1}^{9} 100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2.$$

   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 $\alpha$ que garantice la convergencia de los métodos, y elabore una tabla con las primeras 4 y las últimas 4 aproximaciones $\mathbf{x}_k$ obtenidas.
   En el caso de la función de Rosembrock, puede utilizar como punto inicial el punto $\mathbf{x}_0 = (-1.2, 1)$.

   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.
No se olvide de fijar el `np.random.seed()` antes de correr los puntos aleatorios para siempre obtener la misma 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.

4. El archivo `BD2023.csv` contiene una base de datos relacionada con accidentes de tránsito en 2023. Asuma que los datos de la variable **edad_per** (edad persona) corresponden a una variable continua. Hacer lo siguiente:

   a) Hallar la distribución Gamma$(a, b)$ que mejor se ajusta a estos datos, implementando algún algoritmo de descenso gradiente para optimizar los parámetros de la distribución.
   b) Hallar la distribución LogNormal$(\mu, \sigma^2)$ que mejor se ajusta a estos datos, implementando algún algoritmo de descenso gradiente para optimizar los parámetros de la distribución.
   c) ¿Cuál de las dos distribuciones es mejor para modelar esta variable?

5. **(Mezcla de Gaussianas.)**
Siempre en la base de datos `BD2023.csv`, elabore una regresión con base de funciones gaussianas para determinar la función que mejor aproxime los datos de la variable **hora_ocu** (hora ocurrencia). Asuma que los datos de esta variable corresponden a una v.a. continua.

Para ello, implemente tres posibles aproximaciones:

   a) Fijar los centros $\mu_j$ y las desviaciones $\sigma_j$ de las gaussianas. Sólo optimizar los parámetros $\beta_j$ (alturas) en el modelo de regresión.
   b) Fijar las desviaciones $\sigma_j$ de las gaussianas, y optimizar los centros $\mu_j$ y los parámetros $\beta_j$ en el modelo de regresión.
   c) No fijar ningún parámetro, y optimizar todas las $\mu_j, \sigma_j, \beta_j$ en el modelo.
   (No olvide la restricción $\sigma_j > 0$).

Muestre gráficas contrastando cada resultado de los tres casos anteriores con la distribución de los datos, y elabore alguna tabla de métricas para determinar cuál de los modelos da mejores resultados.