

Métodos numéricos y Optimización - primer semestre 2024

Trabajo Práctico 4 - Optimización

Fecha de entrega: Domingo 23 Junio 23h59

Escribir un informe reportando los resultados de los siguientes experimentos numéricos. El informe debe contar con una introducción, descripción de los métodos numéricos, análisis de los resultados y conclusiones. El informe puede tener hasta 15 páginas sin contar las referencias y deberá venir junto a los códigos .py (referenciados en el informe).

1 Cuadrados mínimos mediante descenso por gradiente

El objetivo de este trabajo es aplicar el algoritmo de gradiente descendente al problema de encontrar la solución del sistema

$$Ax = b, \quad (1)$$

donde $A \in \mathbb{R}^{n \times d}$, $x \in \mathbb{R}^d$ y $b \in \mathbb{R}^n$. Para hacer esto primero definimos la función de costo

$$F(x) = (Ax - b)^T(Ax - b). \quad (2)$$

El algoritmo de gradiente descendente busca encontrar x^* , la solución que minimiza F , mediante el proceso iterativo

$$x_{k+1} = x_k - s \nabla F(x_k), \quad (3)$$

donde s es el paso utilizado.

Cuando el problema tiene más incógnitas que ecuaciones se suelen aplicar *regularizaciones* que inducen alguna propiedad deseada en la solución obtenida. Una variante típica es agregarle un término que dependa de la norma-2 del vector al cuadrado a la función de costo, lo que resulta en

$$F_2(x) = F(x) + \delta_2 \|x\|_2^2, \quad (4)$$

donde δ_2 es un nuevo parámetro a elegir. Esto se conoce como regularización $L2$. Algunas definiciones antes de pasar a las consignas: σ son los valores principales de A , λ son los autovalores de H , donde H es el Hessiano de F (no de F_2).

Objetivos:

- Tomando $n = 5$ y $d = 100$ genere matrices A y vectores b aleatorios y resuelva el problema minimizando F y F_2 . Tome $\delta_2 = 10^{-2} \sigma_{\max}$ cuando trabaje con F_2 . En todos los casos utilice $s = 1/\lambda_{\max}$, una condición inicial aleatoria y realice 1000 iteraciones. Estudiar como evoluciona la solución del gradiente descendente iterativamente. Compare con la solución obtenida mediante SVD. Analice los resultados. ¿Por qué se elige este valor de s ? ¿Qué sucede si se varían los valores de δ_2 ?