

Métodos numéricos y Optimización - primer cuatrimestre de 2023

Trabajo Práctico 4

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 10 páginas sin contar las referencias.

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 $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ y $b \in \mathbb{R}^m$. 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 $L2$. Otra variante usual es la regularización $L1$ donde el término que se agrega va como la norma-1 del vector, resultando en

$$F_1(x) = F(x) + \delta_1 \|x\|_1, \quad (5)$$

donde δ_1 es el nuevo parámetro.

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_1 ni F_2).

Objetivos:

- Tomando $m = 5$ y $n = 100$ genere matrices A y vectores b aleatorios y resuelva el problema minimizando F , F_2 y F_1 . Tome $\delta_2 = 10^{-2}\sigma_{\max}$ y $\delta_1 = 10^{-3}\sigma_{\max}$ cuando trabaje con F_2 y F_1 , respectivamente. En todos los casos utilice $s = 1/\lambda_{\max}$, una condición inicial aleatoria y realice 1000 iteraciones. 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_1 y δ_2 ?

- Tomando $m = 100$ y $n = 100$ genere matrices A aleatorias pero con número de condición dado y vectores b aleatorios y resuelva el problema minimizando F para distintos números de condición. Tome $s = 1/\lambda_{\max}$ y condiciones iniciales aleatorias en cada caso y corra el algoritmo hasta que el error sea de 10^{-2} . Calcule el número de iteraciones en función del número de condición de A y compare con la predicción teórica. Un tip: para generar matrices aleatorias con un número de condición dado puede generar una matriz aleatoria cualquiera y luego modificar sus valores singulares.