

Matemática para ciencias de los datos:

Trabajo práctico 2

M. Sc. Saúl Calderón Ramírez
Instituto Tecnológico de Costa Rica,
Escuela de Computación, bachillerato en Ingeniería en Computación,
PAttern Recongition and MACHine Learning Group (PARMA-Group)

4 de octubre de 2020

Fecha de entrega: Martes 6 de Octubre

Entrega: Un archivo .zip con el pdf generado con lyx o latex, y un jupyter notebook en Pytorch, debidamente documentado, con una función definida por ejercicio que lo necesite. A través del TEC-digital.

Resumen

En el presente trabajo práctico se repasarán aspectos básicos del álgebra lineal, relacionados con los conceptos a desarrollar a lo largo del curso, mezclando aspectos teóricos y prácticos.

1. Para la siguiente función: $f_1(x_1, x_2) = (x_1 - 0, 7)^2 + (x_2 - 0, 5)^2$ con $x_1, x_2 \in [-4, 4]$. Realice lo siguiente:

R/para los cálculos siguientes del trabajo usaremos:

$$\nabla_{\vec{x}} f = \begin{bmatrix} 2(x_1 - 0, 7) \\ 2(x_2 - 0, 5) \end{bmatrix}$$

2. **(10 puntos)** Según tales gráficas, grafique las funciones usando la función meshgrid, y distinga si las funciones son convexas o no, y los puntos mínimos y regiones o puntos silla. **(ver gráfica en colab)**

R/La función f_1 es convexa, no tiene puntos silla y tiene un mínimo **aproximadamente** en $x = [0, 5, 0, 5]$

- a) **(50 puntos)** Implemente el algoritmo del descenso del gradiente, para cada función:

- 1) Escoja un coeficiente de aprendizaje que permita la convergencia y reporte los resultados para 10 corridas:

a' la cantidad de iteraciones necesarias para converger,

b' el punto de convergencia.

c' Reporte si convergió al punto correcto.

d' Escoga una de las corridas y en una gráfica muestre los puntos probados por el algoritmo.

R/ (Ver colab)

- 2) Escoja un α relativamente grande respecto al valor seleccionado ¿Qué sucede? ¿Permite un α muy grande la convergencia?

a' ¿Qué sucede si escoge un muy pequeño?

R/ Si el valor de α es muy alto el algoritmo da pasos muy grandes por lo que no logra aproximar el mínimo. Si el α es muy pequeño el algoritmo puede aproximar el mínimo pero requiere de una gran cantidad de iteraciones. (ver colab)

- b) (40 puntos) Implemente el algoritmo de Newton-Raphson, para cada función:

- 1) Calcule la matriz Hessiana demostrando cada paso intermedio.

R/ Derivadas para la matriz Hessiana:

$$\begin{aligned}\frac{\partial^2}{\partial x_1^2} ((x_1 - 0, 7)^2 + (x_2 - 0, 5)^2) &= \frac{\partial}{\partial x_1} (2(x_1 - 0, 7)) = 2 \\ \frac{\partial^2}{\partial x_1 x_2} ((x_1 - 0, 7)^2 + (x_2 - 0, 5)^2) &= \frac{\partial}{\partial x_2} (2(x_1 - 0, 7)) = 0 \\ \frac{\partial^2}{\partial x_2 x_1} ((x_1 - 0, 7)^2 + (x_2 - 0, 5)^2) &= \frac{\partial}{\partial x_1} (2(x_2 - 0, 5)) = 0 \\ \frac{\partial^2}{\partial x_2^2} ((x_1 - 0, 7)^2 + (x_2 - 0, 5)^2) &= \frac{\partial}{\partial x_2} (2(x_2 - 0, 5)) = 2\end{aligned}$$

Matriz Hessiana:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

- 2) Reporte los resultados para 10 corridas:

a' La cantidad de iteraciones necesarias para converger.

b' El punto de convergencia y reporte si convergió al punto correcto.

c' Escoga una de las corridas y en una gráfica muestre los puntos probados por el algoritmo, usando una gráfica de las curvas de nivel para la función optimizada.

R/ (ver colab)