

Proyecto Final: Algoritmos de Optimización

1. En este problema utilizaremos el [Boston Housing Dataset](#) para predecir el precio de las viviendas a partir de características tales como la cantidad de habitaciones, distancia al centro laboral, entre otras características disponibles en el dataset. Con este fin, nuestro objetivo es minimizar la función de mínimos cuadrados

$$\underset{\beta \in \mathbb{R}^p}{\text{minimize}} \left\| \vec{y} - X\vec{\beta} \right\|_2^2,$$

donde la matriz de datos $X \in \mathbb{R}^{506 \times 13}$ contiene las características, $\vec{y} \in \mathbb{R}^{506}$ contiene el valor medio de las viviendas y $\vec{\beta}$ contiene el vector de parámetros del modelo.

Utilice cualquier lenguaje programación de su elección (Matlab, Python, R, etc) para minimizar la función de mínimos cuadrados y, consecuentemente, estimar los parámetros del modelo. Para esto haga uso de los siguientes algoritmos de optimización:

- a) Descenso de gradiente con condición de Armijo
- b) Descenso de gradiente estocástico (con diferentes tamaños de Batch)
- c) Método de Newton

Envíe su código y un gráfico de la convergencia de los diferentes métodos durante las primeras 100 iteraciones. Comenta brevemente tus resultados.

2. En este problema utilizaremos el [Smarket Dataset](#) para predecir la dirección de la variación de precios del mercado (Up o Down) en funciones de características tales como de los retornos desfasados (Lags), volúmenes de negociación, entre otras características disponibles en el dataset. Por tanto, nuestro objetivo es maximizar la función de log-verosimilitud de la regresión logística

$$\underset{\vec{\beta} \in \mathbb{R}^p}{\text{maximize}} \sum_{i=1}^{1250} y^{(i)} \log g\left(\langle \vec{x}^{(i)}, \vec{\beta} \rangle\right) + (1 - y^{(i)}) \log \left(1 - g\left(\langle \vec{x}^{(i)}, \vec{\beta} \rangle\right)\right),$$

donde $g(t) = (1 - \exp^{-t})^{-1}$ es la función logística, los labels $y^{(i)}$, $1 \leq i \leq 1250$, son 1 (Up) o 0 (Down), $\vec{\beta}$ contiene los parámetros del modelo, y las características $\vec{x}^{(i)}$ son vectores en \mathbb{R}^7 .

Análogo al problema anterior, estime los parámetros de este modelo a través de la maximización de la log-verosimilitud, usando los siguientes algoritmos de optimización:

- a) Descenso de gradiente con condición de Armijo
- b) Descenso de gradiente estocástico (con diferentes tamaños de Batch)
- c) Método de Newton

Envíe su código y un gráfico de la convergencia de los diferentes métodos durante las primeras 100 iteraciones. Comenta brevemente tus resultados.