**Lab #4**

**PROBLEMA 1**

**Parte 0 – Generación de datos**

Graphical user interface, text, application

Description automatically generated

**Parte 1 – Solución Cerrada**

En la práctica no se suele aplicar una solución cerrada para la resolución del problema de mínimos cuadrados ya que para altos volúmenes de datos el cálculo se vuelve computacionalmente demandante.

Graphical user interface, text, application, Word, email

Description automatically generated

Text

Description automatically generated with medium confidence

**\*Nota. La columna E\_n representa el error entre X\_true y X al computar *||x\_k – x\_true||***

**Parte 2 – GD**

Un step size muy pequeño puede ocasionar un proceso de aprendizaje lento, por lo cual la función objetivo a minimizar requiere de un mayor número de iteraciones para converger, mientras un valor muy grande puede ocasionar que diverja. En este ejercicio el primer valor 0.00005 es muy pequeño por lo cual en la gráfica se observa que su error disminuye más lento que los otros dos. El step size que mejor funciona en esta práctica es el step size mayor con un valor de 0.0007.

Graphical user interface, application

Description automatically generated

**Lr = 0.00005**

Graphical user interface, application

Description automatically generated

**Lr = 0.0005**

A screenshot of a computer

Description automatically generated with medium confidence

**Lr = 0.0007**

Graphical user interface, text, application

Description automatically generated

Chart, line chart

Description automatically generated

**Parte 3 – SGD**

En el caso de SGD el step size se comporta igual que en lo descrito para GD, sin embargo, es importante tomar en consideración que dado la actualización de los parámetros de forma recurrente que genera este algoritmo, la función objetivo fluctúa y no es monótona decreciente. En general el step size de 0.0005 genera valores más bajos que los demás dado que es menor y evita que la función diverja.

Graphical user interface, application

Description automatically generated

**Lr = 0.0005**

Graphical user interface, text, application, email

Description automatically generated

**Lr = 0.005**

Graphical user interface, text, application, email

Description automatically generated

**Lr = 0.01**

Graphical user interface, text, application, email

Description automatically generated

Chart, line chart, histogram

Description automatically generated

**Parte 4 – MBGD**

* Un mini batch size muy grande generar mayor error en el conjunto de datos, por lo cual el mini batch de 25 es el que presenta mejor desempeño
* Los mejores resultados se obtienen con un step size de 0.0005 y un mini batch size de 25, esto dado que presenta el menor error de todos los valores evaluados.

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Chart

Description automatically generated

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

**Parte 5 – Comparación**

Es posible observar que los algoritmos tienen un desempeño similar con respecto a la optimización de la función objetivo. Sin embargo, la diferencia en su comparación se resaltar a nivel de número de iteraciones entre los cuales la aplicación de la solución cerrada, así como el método clásico de GD obtienen el mejor desempeño en cuento a tiempo y resultado. A este análisis es necesario agregarle las virtudes que pueden tener los otros algoritmos ya que no utilizan todos los datos (MBGD) o bien pueden ser usados para online learning (SGD).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Método** | **LR** | **Mini Batch size** | **Valor óptimo (f\*)** | **No. Iteraciones** | **Error (x\_k, x\_true)** |
| Cerrado | - | - | 225.02 | 1 | 0.157 |
| GD | 0.0007 | 1000 | 225.02 | 58 | 0.157 |
| SGD | 0.0005 | 1 | 225.18 | 1000 | 0.156 |
| MB-GD | 0.0005 | 25 | 225.19 | 1000 | 0.158 |

**PROBLEMA 2**

**Parte 1 – GD con Backtracking Line Search**

|  |  |  |
| --- | --- | --- |
| **X0** | **Error Backtracking** | **Error LR constante** |
| (0, 0) | Inf | 0.69 |
| (0.6, 0.6) | 0.17 | 0.22 |
| (-0.5, 1) | 0.77 | 2.18 |
| (-1.2, 1) | 2.09 | 2.11 |

Usando un step size constante de 0.0005 es posible notar que backtracking funciona mejor para todos los puntos iniciales con excepción del punto (0, 0). Adicionalmente la selección del step size constante fue a través de experimentación y evaluación de resultados pues valores mayores al utilizado tienden a divergir, mientras que backtracking determina el step size de forma automática según el algoritmo.

Graphical user interface, application

Description automatically generated

**X0 = [0, 0]**

Graphical user interface, text, application, email

Description automatically generated

**X0 = [0.6, 0.6]**

Graphical user interface, text, application

Description automatically generated

**X0 = [-0.5, 1]**

A picture containing application

Description automatically generated

**X0 = [-1.2, 1]**

Graphical user interface, application

Description automatically generated

**Parte 2 – Método de Newton con Backtracking Line Search**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **X0** | **Error Backtracking** | **Iteraciones Backtracking** | **Error LR constante** | **Iteracions LR constante** |
| (0, 0) | 0 | 344 | 0 | 3 |
| (0.6, 0.6) | 4.03 | 3000 | 0 | 6 |
| (-0.5, 1) | 11.21 | 3000 | 0 | 6 |
| (-1.2, 1) | 53.83 | 3000 | 7 | 0 |

En el caso del método de Newton este algoritmo se beneficia de tener un step size constante en contraste de calcularlo a través de backtracking, pues el número de iteraciones es menor al tener un step size constante de 1 mientras con backtracking solo en el primer punto inicial logra converger a 0.

Graphical user interface, application

Description automatically generated

**X0 = [0, 0]**

Graphical user interface

Description automatically generated with medium confidence

**X0 = [0.6, 0.6]**

Graphical user interface, text, application

Description automatically generated

**X0 = [-0.5, 1]**

Graphical user interface, application

Description automatically generated

**X0 = [-1.2, 1]**

Graphical user interface, text, application

Description automatically generated