

UNIVERSIDAD GALILEO

Postgrado en Análisis y Predicción de Datos

Curso: Algoritmos en la Ciencia de Datos Horario: lunes: 18:00 – 21:00

Tutor: PhD. Alberth Alvarado



Laboratorio No. 3

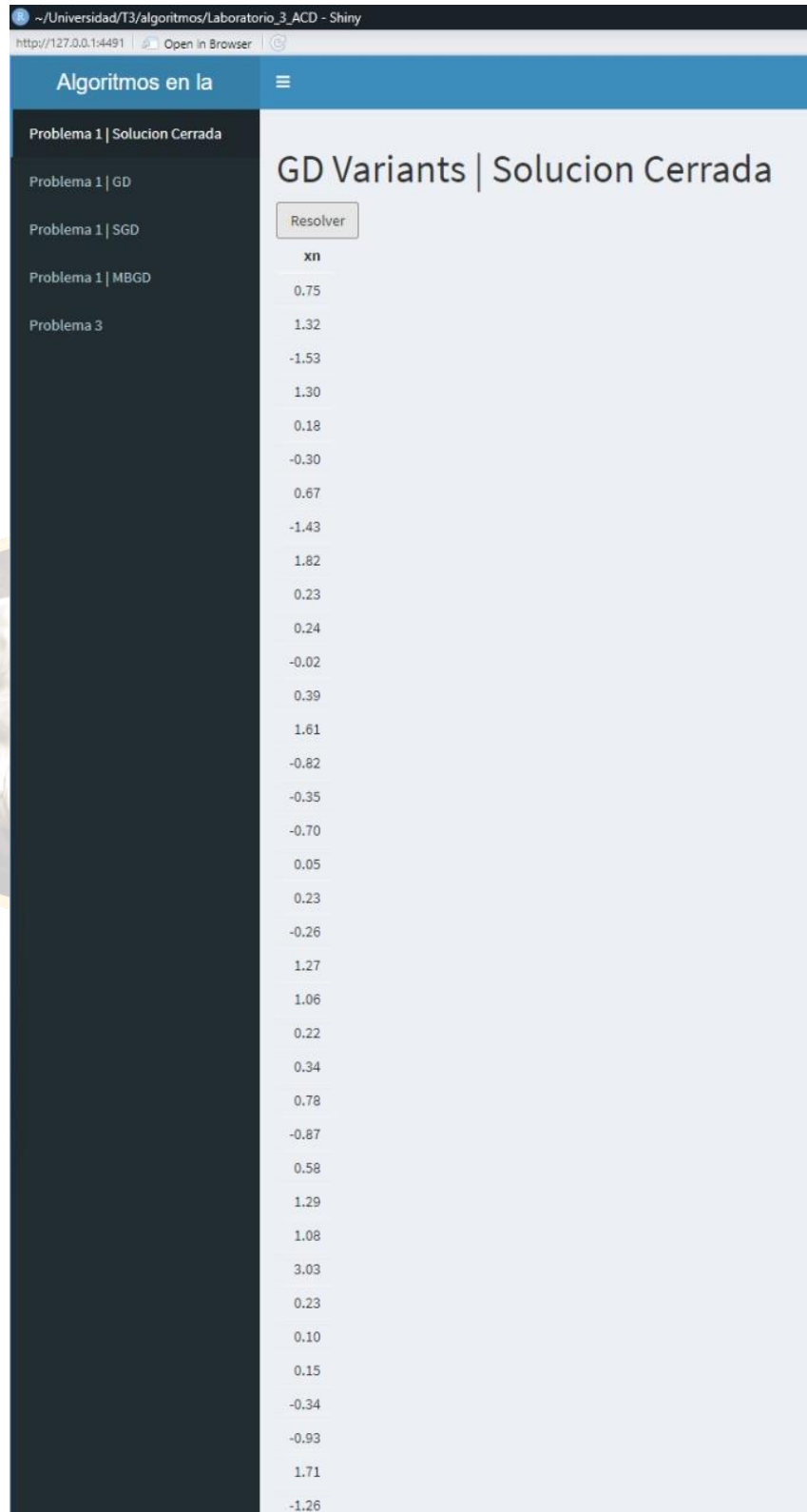
Galileo
UNIVERSIDAD

Edgar Geovany Ocaña Orozco – 24010004
Andrea María Hernández Marroquín – 240110074

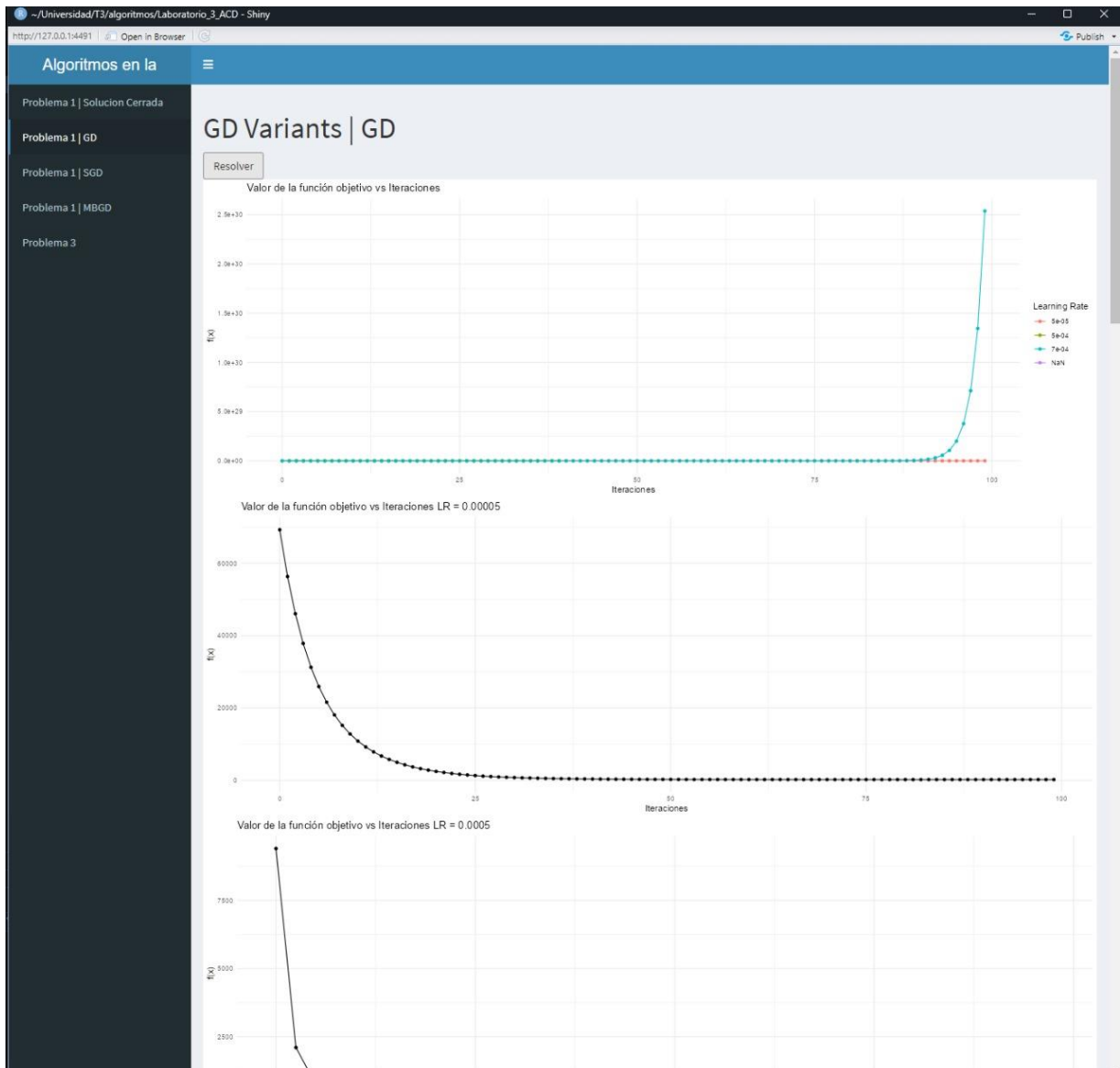
Guatemala, 15 de septiembre de 2024

2. SECCIÓN DOCUMENTAL DE EXPERIMENTACIÓN

1) Solución Cerrada



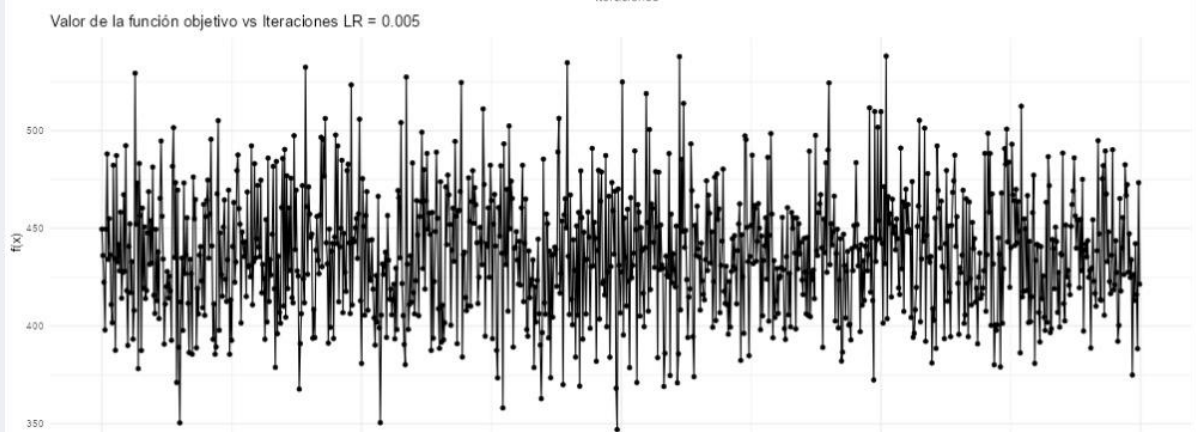
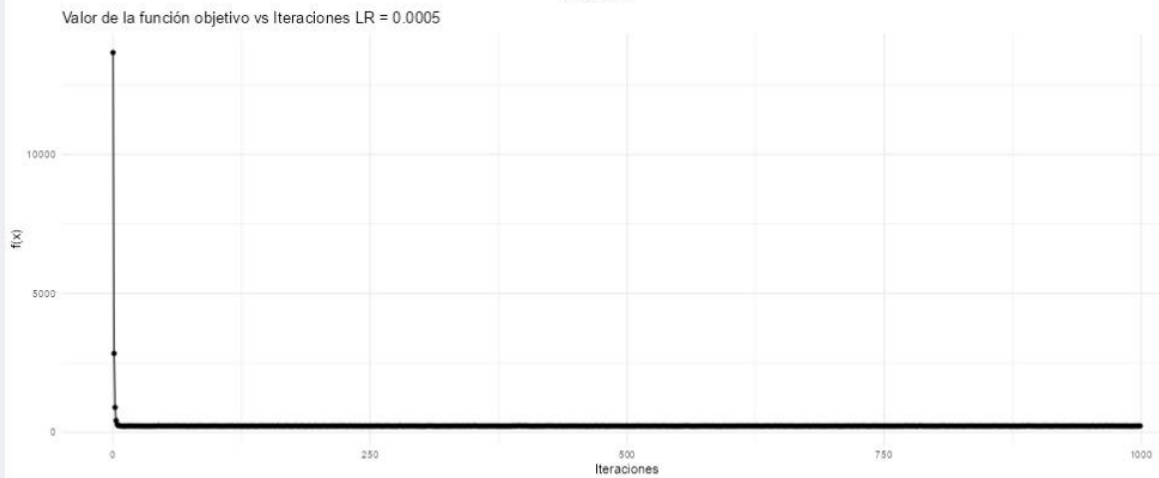
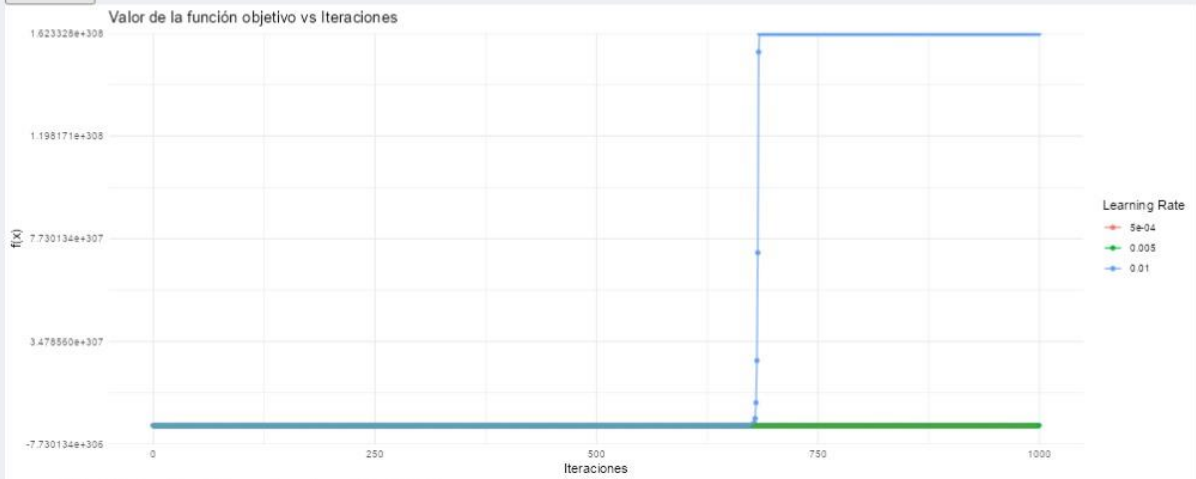
2) GD



Iteraciones								
Iter_1	f_x_1	Learning_rate_1	Iter_2	f_x_2	Learning_rate_2	Iter_3	f_x_3	Learning_rate_3
0.00	69237.18	0.00	0.00	9399.80	0.00	0.00	33228.93	0.00
1.00	56315.09	0.00	1.00	2102.69	0.00	1.00	32162.55	0.00
2.00	46045.15	0.00	2.00	742.52	0.00	2.00	38604.90	0.00
3.00	37839.86	0.00	3.00	402.52	0.00	3.00	52009.98	0.00
4.00	31249.83	0.00	4.00	299.01	0.00	4.00	75915.71	0.00
5.00	25929.81	0.00	5.00	262.62	0.00	5.00	117601.02	0.00
6.00	21613.39	0.00	6.00	248.50	0.00	6.00	190415.43	0.00
7.00	18094.08	0.00	7.00	242.65	0.00	7.00	318494.87	0.00
8.00	15211.07	0.00	8.00	240.13	0.00	8.00	545508.57	0.00
9.00	12838.53	0.00	9.00	239.02	0.00	9.00	950729.18	0.00
10.00	10877.55	0.00	10.00	238.52	0.00	10.00	1678562.06	0.00
11.00	9249.98	0.00	11.00	238.29	0.00	11.00	2992891.90	0.00
12.00	7893.78	0.00	12.00	238.18	0.00	12.00	5377270.53	0.00
13.00	6759.47	0.00	13.00	238.13	0.00	13.00	9719890.31	0.00
14.00	5807.40	0.00	14.00	238.11	0.00	14.00	17655544.92	0.00
15.00	5005.63	0.00	15.00	238.10	0.00	15.00	32198542.22	0.00
16.00	4328.32	0.00	16.00	238.09	0.00	16.00	58915305.09	0.00
17.00	3754.48	0.00	17.00	238.09	0.00	17.00	108098787.66	0.00
18.00	3266.96	0.00	18.00	238.09	0.00	18.00	198803791.00	0.00
19.00	2851.71	0.00	19.00	238.09	0.00	19.00	366341034.30	0.00
20.00	2497.18	0.00	20.00	238.09	0.00	20.00	676203360.89	0.00
21.00	2193.81	0.00	21.00	238.09	0.00	21.00	1249959599.03	0.00
22.00	1933.66	0.00	22.00	238.09	0.00	22.00	2313425855.13	0.00
23.00	1710.15	0.00	23.00	238.09	0.00	23.00	4286323581.39	0.00
24.00	1517.76	0.00	24.00	238.09	0.00	24.00	7949229033.02	0.00
25.00	1351.87	0.00	25.00	238.09	0.00	25.00	14754574895.85	0.00
26.00	1208.60	0.00	26.00	238.09	0.00	26.00	27406228150.17	0.00
27.00	1084.68	0.00	27.00	238.09	0.00	27.00	50939986985.22	0.00
28.00	977.34	0.00	28.00	238.09	0.00	28.00	94738652011.70	0.00
29.00	884.24	0.00	29.00	238.09	0.00	29.00	176291318077.42	0.00
30.00	803.39	0.00	30.00	238.09	0.00	30.00	328208977654.30	0.00
31.00	733.09	0.00	31.00	238.09	0.00	31.00	611321245766.81	0.00
32.00	671.89	0.00	32.00	238.09	0.00	32.00	1139132900800.88	0.00
33.00	618.57	0.00	33.00	238.09	0.00	33.00	2123507289222.33	0.00
34.00	572.05	0.00	34.00	238.09	0.00	34.00	3960025436687.47	0.00
35.00	531.44	0.00	35.00	238.09	0.00	35.00	7387524912049.54	0.00
36.00	495.95	0.00	36.00	238.09	0.00	36.00	13786367214398.90	0.00
37.00	464.90	0.00	37.00	238.09	0.00	37.00	25736220923239.18	0.00
38.00	437.72	0.00	38.00	238.09	0.00	38.00	48059443389001.19	0.00
39.00	413.90	0.00	39.00	238.09	0.00	39.00	89773338759838.66	0.00

GD Variants | SGD

Resolver



Iter_1	f_x_1	Learning_rate_1	Iter_2	f_x_2	Learning_rate_2	Iter_3
0.00	13674.84	0.00	0.00	449.61	0.00	0.00
1.00	2836.16	0.00	1.00	436.07	0.00	1.00
2.00	896.16	0.00	2.00	422.41	0.00	2.00
3.00	423.21	0.00	3.00	397.96	0.00	3.00
4.00	285.93	0.00	4.00	449.58	0.00	4.00
5.00	243.51	0.00	5.00	488.08	0.00	5.00
6.00	232.40	0.00	6.00	434.13	0.00	6.00
7.00	229.19	0.00	7.00	454.99	0.00	7.00
8.00	229.36	0.00	8.00	436.37	0.00	8.00
9.00	223.52	0.00	9.00	410.95	0.00	9.00
10.00	226.14	0.00	10.00	401.66	0.00	10.00
11.00	226.15	0.00	11.00	482.27	0.00	11.00
12.00	225.72	0.00	12.00	434.58	0.00	12.00
13.00	227.79	0.00	13.00	387.59	0.00	13.00
14.00	227.92	0.00	14.00	487.22	0.00	14.00
15.00	227.24	0.00	15.00	432.81	0.00	15.00
16.00	227.37	0.00	16.00	441.71	0.00	16.00
17.00	229.19	0.00	17.00	428.23	0.00	17.00
18.00	230.82	0.00	18.00	458.26	0.00	18.00
19.00	226.77	0.00	19.00	414.23	0.00	19.00
20.00	227.22	0.00	20.00	427.40	0.00	20.00
21.00	225.20	0.00	21.00	467.24	0.00	21.00
22.00	228.64	0.00	22.00	427.99	0.00	22.00
23.00	226.67	0.00	23.00	492.32	0.00	23.00
24.00	226.62	0.00	24.00	418.10	0.00	24.00
25.00	228.62	0.00	25.00	390.06	0.00	25.00
26.00	228.28	0.00	26.00	440.78	0.00	26.00
27.00	225.08	0.00	27.00	452.26	0.00	27.00
28.00	227.26	0.00	28.00	417.05	0.00	28.00
29.00	226.15	0.00	29.00	432.96	0.00	29.00

Galileo
UNIVERSIDAD

Problema 1 | Solucion Cerrada

Problema 1 | GD

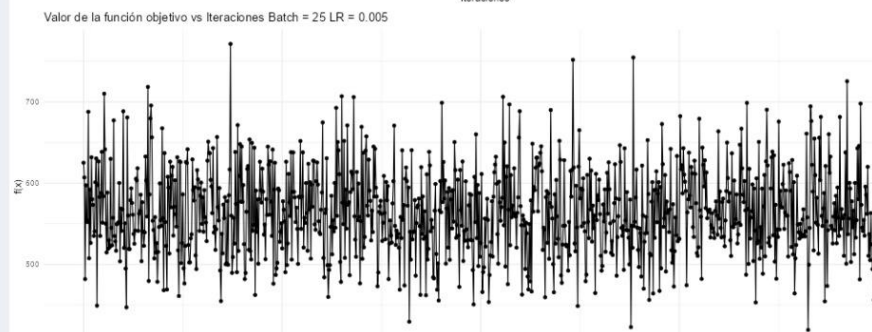
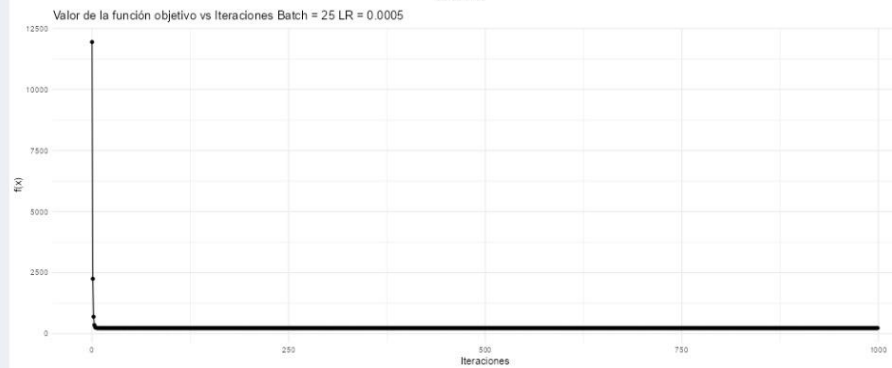
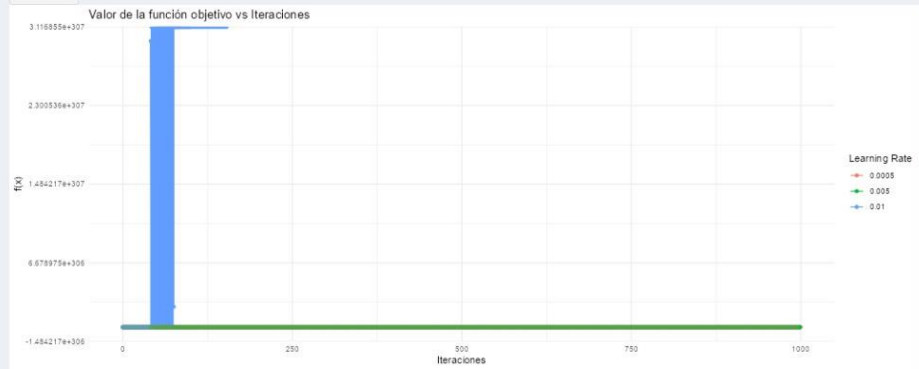
Problema 1 | SGD

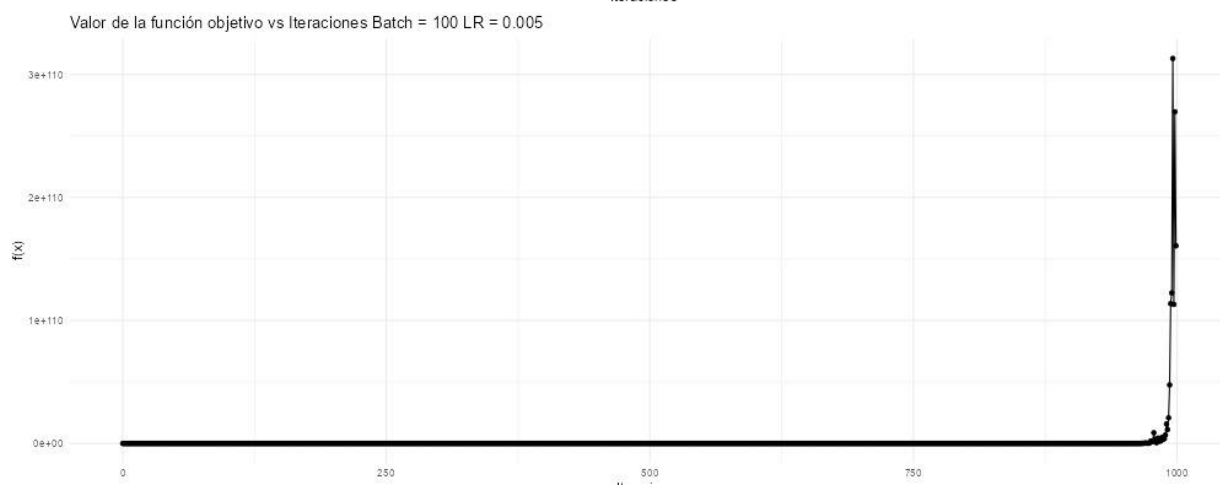
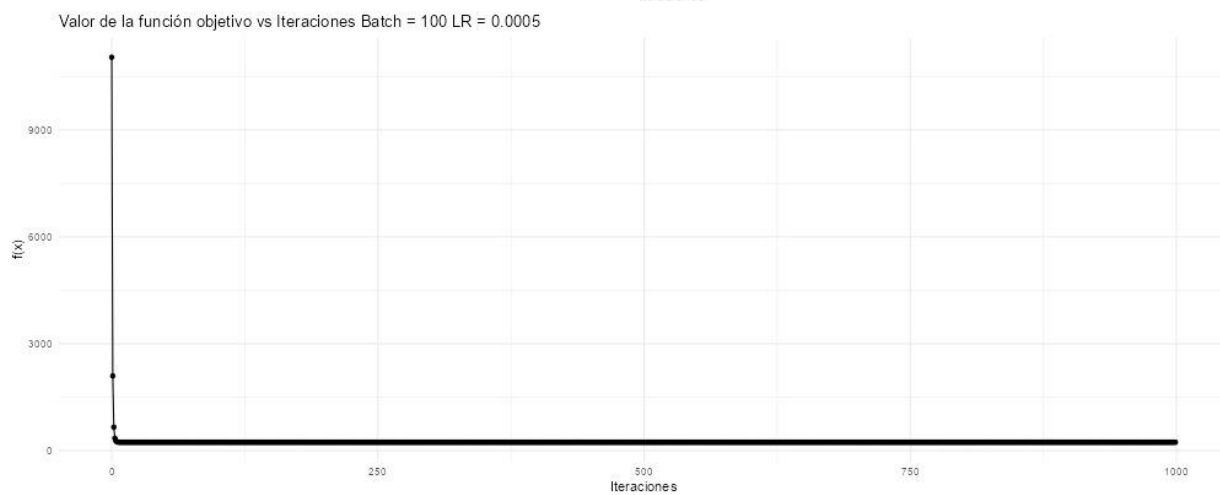
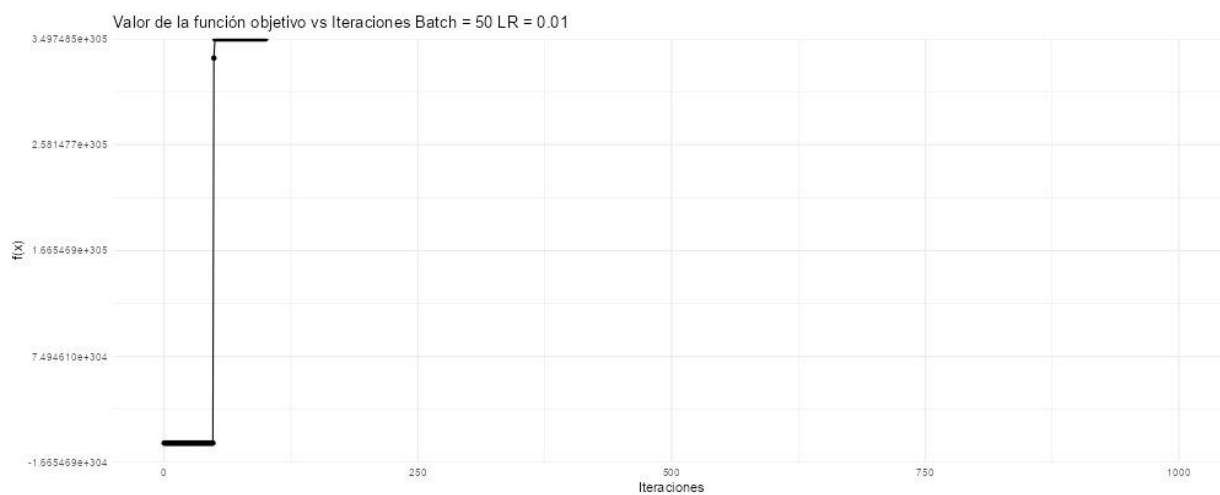
Problema 1 | MBGD

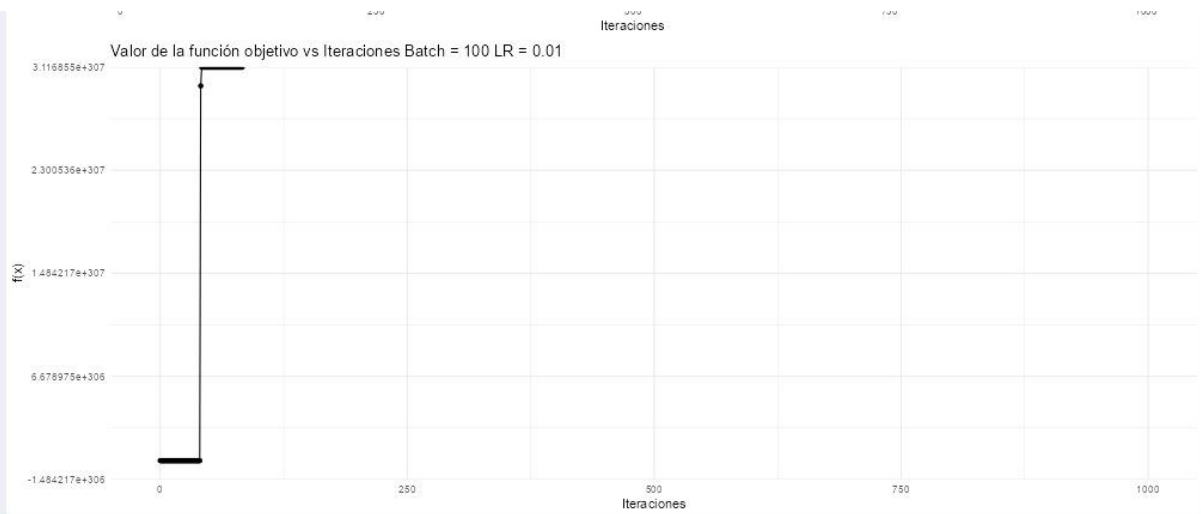
Problema 1 | Comparaciones

GD Variants | MBGD

Resolver







Iter_1	f_x_1	Learning_rate_1	Batch1	Iter_2	f_x_2	Learning_rate_2	Batch2	Iter_3
0.00	11943.42	0.0005	25.00	0.00	625.19	0.005	25.00	0.00
1.00	2248.46	0.0005	25.00	1.00	607.04	0.005	25.00	1.00
2.00	697.18	0.0005	25.00	2.00	481.99	0.005	25.00	2.00
3.00	355.73	0.0005	25.00	3.00	597.31	0.005	25.00	3.00
4.00	269.42	0.0005	25.00	4.00	554.81	0.005	25.00	4.00
5.00	245.23	0.0005	25.00	5.00	552.13	0.005	25.00	5.00
6.00	237.18	0.0005	25.00	6.00	687.84	0.005	25.00	6.00
7.00	235.21	0.0005	25.00	7.00	507.66	0.005	25.00	7.00
8.00	234.53	0.0005	25.00	8.00	592.00	0.005	25.00	8.00
9.00	233.79	0.0005	25.00	9.00	526.43	0.005	25.00	9.00
10.00	234.12	0.0005	25.00	10.00	631.89	0.005	25.00	10.00
11.00	234.30	0.0005	25.00	11.00	580.22	0.005	25.00	11.00
12.00	233.56	0.0005	25.00	12.00	573.17	0.005	25.00	12.00
13.00	233.45	0.0005	25.00	13.00	535.04	0.005	25.00	13.00
14.00	233.79	0.0005	25.00	14.00	540.72	0.005	25.00	14.00
15.00	233.87	0.0005	25.00	15.00	601.13	0.005	25.00	15.00
16.00	233.67	0.0005	25.00	16.00	630.47	0.005	25.00	16.00
17.00	233.62	0.0005	25.00	17.00	449.23	0.005	25.00	17.00
18.00	233.33	0.0005	25.00	18.00	597.77	0.005	25.00	18.00
19.00	233.74	0.0005	25.00	19.00	626.78	0.005	25.00	19.00
20.00	233.83	0.0005	25.00	20.00	583.05	0.005	25.00	20.00
21.00	233.98	0.0005	25.00	21.00	535.04	0.005	25.00	21.00
22.00	233.27	0.0005	25.00	22.00	587.98	0.005	25.00	22.00
23.00	234.10	0.0005	25.00	23.00	638.55	0.005	25.00	23.00
24.00	233.70	0.0005	25.00	24.00	551.41	0.005	25.00	24.00
25.00	233.35	0.0005	25.00	25.00	583.95	0.005	25.00	25.00

GD Variants | Comparaciones

Resolver							
Method	f(x)		Iterations	Error in x		Learning Rate	Batch Size
Solucion Cerrada	213.76		0.00	0.16		NA	NA
GD	232.34		45.00	0.16		0.0005	NA
GD	232.51		100.00	0.16		5e-05	NA
MBGD	233.60		1000.00	0.17		0.0005	25.00
MBGD	233.74		1000.00	0.16		0.0005	100.00
MBGD	233.77		1000.00	0.16		0.0005	25.00
MBGD	233.78		1000.00	0.17		0.0005	50.00
MBGD	233.98		1000.00	0.17		0.0005	50.00
MBGD	234.09		1000.00	0.17		0.0005	100.00
SGD	244.81		1000.00	0.19		0.0005	NA
SGD	443.72		1000.00	0.49		0.005	NA
MBGD	577.33		1000.00	0.60		0.005	25.00
MBGD	578.71		1000.00	0.61		0.005	25.00
MBGD	720.93		1000.00	0.70		0.005	50.00
MBGD	809.08		1000.00	0.74		0.005	50.00
GD	2016757198513827836068042828440640.00		100.00	1079943681556536.00		0.0007	NA
MBGD	16068513372927901335468080864846448604840462288628646480640284802448666060066666082084422082442628206022002886.00		1000.00	375563026304982210240884222260482020846626848888080400.00		0.005	100.00
MBGD	17261272019124192620408200262860000884002482242866042060640608082266224866444648044224600648442284484888208640442.00		1000.00	12384488219566878603240426264886846684200022680868246208.00		0.005	100.00
SGD			Inf	Inf		0.01	NA
MBGD			NA	NA		0.01	25.00
MBGD			NA	NA		0.01	50.00
MBGD			NA	NA		0.01	100.00
MBGD			NA	NA		0.01	25.00
MBGD			NA	NA		0.01	50.00
MBGD			NA	NA		0.01	100.00

3. SECCIÓN DOCUMENTAL DE CONCLUSION

1) ¿Por qué en la práctica no se utiliza el método de solución cerrada?

La solución cerrada, aunque teóricamente perfecta porque da la respuesta exacta en una operación, tiene limitaciones significativas en la práctica, especialmente en contextos de datos grandes:

- 1) Complejidad Computacional
- 2) Requerimientos de Memoria
- 3) Sensibilidad a Ruido y Outliers

Comente cómo el step size afecta la convergencia del algoritmo GD y con cuál step size constante obtuvo el “mejor” resultado.

El tamaño del paso de 0.0005 pareció proporcionar un buen equilibrio entre velocidad de convergencia y estabilidad, alcanzando valores bajos de la función objetivo sin mostrar signos de divergencia incontrolable.

Comente cómo el step size afecta la convergencia del algoritmo SGD y con cuál step size constante obtuvo el “mejor” resultado.

El step size de 0.005, aunque variable, parece ofrecer el mejor equilibrio entre velocidad y estabilidad, pero con una recomendación de ser ligeramente reducido para disminuir la variabilidad y las oscilaciones observadas. Es vital realizar más pruebas con tasas ligeramente menores para optimizar la convergencia.

Implementar técnicas adicionales como decaimiento adaptativo del step size o métodos de optimización avanzados (como Adam o RMSprop) que ajustan el step size de manera adaptativa podría ser beneficioso para controlar las oscilaciones y mejorar la estabilidad y velocidad de convergencia del SGD.

Comente cómo el step size afecta la convergencia del algoritmo MBGD, puede experimentar con otro step sizes para justificar su respuesta. ¿Con cuál step size constante obtuvo el “mejor” resultado?

Basado en los datos proporcionados, el step size de 0.0005 parece ofrecer el mejor equilibrio entre estabilidad y progreso hacia la convergencia, aunque para aplicaciones donde el tiempo de convergencia es crítico, un ajuste fino del step size en torno a 0.005 podría explorarse con precaución para mejorar la eficiencia sin comprometer demasiado la estabilidad.

Sería beneficioso implementar técnicas como **decaimiento del step size** o **ajustes adaptativos del step size** para manejar mejor la variabilidad y la convergencia, especialmente en el caso de step sizes mayores como 0.005.

Compare el desempeño de cada uno de los métodos implementados.

Solución Cerrada: Tiene un valor de función objetivo comparativamente bajo y un error en xxx de 0.16, pero no requiere iteraciones, lo que la hace ideal desde el punto de vista de la convergencia inmediata. Sin embargo, como discutido previamente, puede ser inviable para grandes dimensiones o datos correlacionados debido a problemas de escalabilidad y precisión.

GD: Varias entradas con diferentes tasas de aprendizaje. En general, parece que mientras más baja es la tasa de aprendizaje, más estables son los resultados, pero a costa de requerir más iteraciones para converger.

MBGD: Muestra una amplia gama de resultados, dependiendo del tamaño del lote y la tasa de aprendizaje. En general, los tamaños de lote más pequeños y las tasas de aprendizaje más **bajas**

parecen ofrecer un equilibrio entre estabilidad y eficiencia, aunque algunas configuraciones muestran una alta variabilidad en la función objetivo.

SGD: Presenta una variabilidad considerable en el valor de $f(x)$, lo que es típico para este método debido a su naturaleza estocástica. Las tasas de aprendizaje más altas parecen causar una mayor variabilidad e incluso divergencia en algunos casos.





Galileo
UNIVERSIDAD
