DATA 220 Mathematical Methods for Data Analytics – Homework -4 | Deadline – 11.59 PM – 11/20/2023

Problem 1 (**2 pts):** Find the critical points of the multivariable function and check whether the critical points are local minima, local maxima, or saddle points.

a. $f(x, y) = x^3 - 12xy + 8y^3$

Problem 2 (**2 pts):** Find the minimum or maximum points using Lagrange multipliers

$$f(x, y) = x^2 + xy + y^2;$$

$$s.t., x^2 + y^2 = 8$$

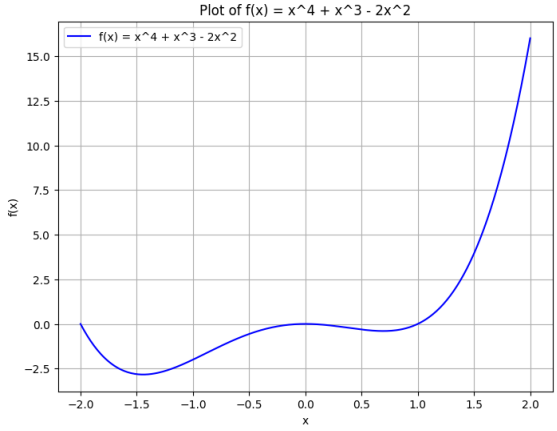
Problem 3 (**2 pts):** Use gradient descent to find the minimum value of the univariate function

$f(x) = x^4 + x^3 - 2x^2$

With initial guess $x_0 = 0.5$ and step size, $\alpha = 0.01$; show two iteration updates (x_1 *and* x_2)

Problem 4 (Coding) (**3 pts):** Use gradient descent to find the minimum value of the univariate function

$f(x) = x^4 + x^3 - 2x^2$ for minimum of 200 iterations. Consider, step size, $\alpha = 0.01$. Show the output in a table including columns: $n(iteration\ number)$,
x_n $(value\ of\ x\ for\ each\ iteration\ where\ x_0\ is\ the\ initial\ guess)$, and $\alpha * f'(x_n)$



a. With initial guess $x_0 = 1$;
b. With initial guess $x_0 = -1$
c. Discuss: Compare the convergence in terms of local and global minima for the above a and b.

Problem 5 (Coding) **11 pts**:

For the following tasks use the breast cancer dataset:

```
from sklearn.datasets import load_breast_cancer
cancer = load_breast_cancer()
```

Let's look at the first five observation of the data (X) that include 30 features and the target variable (y), respectively:

```
1 pd.DataFrame(cancer['data']).head()
```

	0	1	2	3	4	5	6	7	8	9	...
0	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419	0.07871	...
1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812	0.05667	...
2	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069	0.05999	...
3	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597	0.09744	...
4	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809	0.05883	...

5 rows × 30 columns

```
1 pd.DataFrame(cancer['target'])
```

	0
0	0
1	0
2	0
3	0
4	0
...	...
564	0
565	0
566	0
567	0
568	1

569 rows × 1 columns

Split the dataset into training and test sets of 80:20 ratio (use **random_seed = 2023**) and test_size = 0.20. You must train the linear regression model using the training data and compute MSE using the test dataset.

Perform Multiple linear regression using **gradient descent optimization technique**. You must **not** use any direct package for this. You should generate the parameters (intercepts & coefficients) initial values from standard normal distribution. Set initial learning rate and number of iterations to 0.001 and 100, respectively.

a. **(2 pts)** Mention: loss function, initial values of parameters (should be generated from Standard Normal Distribution: $X \sim N(0,1)$), hyperparameters, and partial derivatives of the loss function with respect to intercept and coefficients
b. **(6 pts)** Write a function that takes input (as arguments) of hyperparameters and provides output: best set of optimal parameters (intercept and coefficients), and the loss (cost).
c. (**1 pt**) Plot learning curve where x-axis will be number of iteration and y-axis will be loss for each iteration and explain the curve
d. (**1 pts**) Form the regression equation using the optimal parameters and find ($\hat{y}$) (predict for each datapoints of x_test) – show in DataFrame including two columns: y (y_actual) & $\hat{y}$ (y_predict)
e. (**1 pts**) Finally, for the test dataset: Find Mean Squared Error (MSE).

You are required to submit:

1. An MS/PDF/Scanned document:
    a. Include all the steps of your calculations.
    b. Attach screenshots of the code output.
2. Source code:
    a. Python (Jupyter Notebook)
    b. Ensure it is well-organized with comments and proper indentation.

- Failure to submit the source code will result in a deduction of 5 points.
- Format your filenames as follows: "your_last_name_HW1.pdf" for the document and "your_last_name_HW1_source_code.ipynb" for the source code.
- Before submitting the source code, please double-check that it runs without any errors.
- Must submit the files separately.
- Do not compress into a zip file.
- HW submitted more than 24 hours late will not be accepted for credit.