**Name:** _____

**MATH 361**
**Spring 2022**
**Written HW 7: Due 05/12**

*"I don't think that everyone should become a mathematician, but I do believe that many students don't give mathematics a real chance."*
                                                                    *— Maryam Mirzakhani*

**Problem 1. (Least Square Regression Line)** Linear models are of particular use for understanding data and making predictions because of their simplicity. However, perfectly linear datasets do not exist 'in nature.' But we have seen that even when a dataset is not *perfectly* linear, it is still possible to "approximate" the data using a linear model. Consider the following dataset:

| $x$  | $y$      |
|------|----------|
| 0.3  | 89.3     |
| 3.5  | 35.8     |
| 7.9  | 5.7      |
| 10.2 | $-55.4$  |
| 14.9 | $-109.2$ |

(a) Compute the least square regression line for this dataset.

(b) Compute the prediction and residual for $x = 7.9$.

(c) Is a linear model "appropriate" for this dataset? Explain.

(d) Compute the total "error" in the model using the following error measurements:

$$\text{Maximum:} \quad \max_i |y_i - \widehat{y_i}|$$

$$\text{Average Error:} \quad \frac{1}{n} \sum_i |y_i - \widehat{y_i}|$$

$$\text{Root Mean-Square Error:} \quad \frac{1}{n} \sum_i |y_i - \widehat{y_i}|^2$$

Which of the above error measurements is the closest used to compute the linear regression and how does this 'define' the regression line?

**Problem 2. (Least Square Parabola)** Although the most commonly used regression is a linear regression, we can fit any polynomial $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ to a given dataset. Just as with linear regressions, we choose the model that minimizes the squares of the error, i.e. the least square polynomial. This procedure allowed us to find the optimal choice of coefficients $\{a_i\}_{i=0}^n$. Consider the following dataset:

| $x$ | $y$ |
|---|---|
| $-5.5$ | $219.7$ |
| $-2.7$ | $47.7$ |
| $0.3$ | $1.4$ |
| $3.8$ | $-33.7$ |
| $6.9$ | $136.9$ |

(a) Use the 'normal equations' to find the least square parabola for the data given above.

(b) Compute the total square error for this model.

(c) Given a model $f(x) = ax^2 + bx + c$ for the data above, what is the smallest possible 'error' (measured by square error)? Explain.

**Problem 3. (Linearizing Data)** We have seen how to find models that best fit a dataset using analytic techniques that produced exact solutions. However, not all choices of models had the property that we could exactly find the conditions that minimized the square error. In some cases, we could perform a change of variables to linearize the data and then apply standard least square regression techniques to find the model.

(a) For each of the following models, find a chance of variables that 'linearizes' the data and write the model in linearized form:

(i) $y = Ax^2 + C$, $x \geq 0$.

(ii) $y = \dfrac{A}{x} + C$, $x \neq 0$

(iii) $y = Ab^{Cx}$

(iv) $y = \dfrac{1}{(Ax + C)^2}$, $x \neq -\frac{C}{A}$

(b) Suppose one chooses a model $y = 2^x$. Given the dataset below, find the linearized dataset.

| $x$ | $y$ |
|-----|------|
| 0.1 | 1.53 |
| 3.7 | 14.01 |
| 5.4 | 46.18 |
| 6.6 | 94.22 |
| 8.0 | 261.4 |

**Problem 4. (Gradient Descent)** One of the most commonly used minimization techniques in applied mathematics is gradient descent or one of its variants. The method can be rather slow, each iteration typically requires few flops; moreover, the method is simple to understand and implement. Consider the case where $f(x, y) = 8xy + \dfrac{1}{x} + \dfrac{1}{y}$.

(a) Show that $f(x, y)$ has a minimum at $(\frac{1}{2}, \frac{1}{2})$.

(b) Implement gradient descent using five iterations with step size $h = 0.1$. Use an initial value of $\mathbf{x}_0 = (1.0, 1.0)$.

(c) Find the error in your approximation from (b).

**Evaluation.**

Complete the following survey by rating each problem. Each area will be rated on a scale of 1 to 5. For interest, 1 is "mind-numbing" while a 5 is "mind-blowing." For difficulty, 1 is "trivial/routine" while 5 is "brutal." For learning, 1 means "nothing new" while 5 means "profound awakening." Then you to estimate the amount of time you spent on each problem (in minutes).

|  | Interest | Difficulty | Learning | Time Spent |
|---|---|---|---|---|
| Problem 1 |  |  |  |  |
| Problem 2 |  |  |  |  |
| Problem 3 |  |  |  |  |
| Problem 4 |  |  |  |  |

Finally, indicate whether you believe lectures were useful in completing this assignment and whether you believe the problems were useful enough/interesting enough to assign again to future students by checking the appropriate space.

|  | Lectures | | Assign Again | |
|---|---|---|---|---|
|  | Yes | No | Yes | No |
| Problem 1 |  |  |  |  |
| Problem 2 |  |  |  |  |
| Problem 3 |  |  |  |  |
| Problem 4 |  |  |  |  |