Assignment Report: Finding Unknown Parameters in a Parametric Curve

Research and Development / AI Assignment

Name: SUMAN PAL
Roll No: 142402012

1. Problem Statement

We are given a parametric curve defined as:

$$x = t\cos(\theta) - e^{M|t|}\sin(0.3t)\sin(\theta) + X$$
$$y = 42 + t\sin(\theta) + e^{M|t|}\sin(0.3t)\cos(\theta)$$

where the parameters θ, M, and X are unknown.

The objective is to estimate the values of these unknowns based on given data points (x_i, y_i) that lie on the curve.

The given parameter ranges are:

$$0^\circ < \theta < 50^\circ, \quad -0.05 < M < 0.05, \quad 0 < X < 100, \quad 6 < t < 60$$

—

2. Objective

To determine the values of θ, M, and X that minimize the L1 distance between the experimental (given) points and the predicted curve.

The evaluation metric is based on the total L1 distance:

$$L1 = \sum_i \|(x_i, y_i) - (x(t_i), y(t_i))\|_1$$

where the smaller the value, the better the fit.

—

3. Methodology

Step 1: Data Loading

The dataset `xy_data.csv` contains coordinates (x, y) for points lying on the curve.

Step 2: Model Definition

The curve was implemented in Python as:

$$x(t; \theta, M, X) = t\cos(\theta) - e^{M|t|}\sin(0.3t)\sin(\theta) + X$$
$$y(t; \theta, M, X) = 42 + t\sin(\theta) + e^{M|t|}\sin(0.3t)\cos(\theta)$$

Step 3: Optimization Approach

Since t values were not given, a dense range $t \in [6, 60]$ was sampled uniformly. For each candidate set of parameters (θ, M, X), the L1 between data points and the curve was calculated as:

$$d_i = \min_{t \in [6,60]} \Big(|x_i - x(t)| + |y_i - y(t)| \Big)$$

The total distance was minimized using the `Differential Evolution` optimizer from the `scipy.optimize` library in Python.

4. Mathematical Formulation for Parameter Estimation

In linear regression, parameters are obtained analytically by minimizing the Mean Squared Error (MSE) and setting its derivative to zero:

$$\frac{\partial}{\partial w} \sum_i (y_i - \hat{y}_i)^2 = 0$$

However, in this problem, the model is nonlinear due to the presence of:

$$e^{M|t|}, \quad \sin(0.3t), \quad \text{and} \quad \sin(\theta), \cos(\theta)$$

Hence, a closed-form derivative-based solution is not possible.

Instead, we define a loss function that measures the total distance between predicted and observed points:

$$J(\theta, M, X) = \sum_{i=1}^{n} \min_{t \in [6,60]} \sqrt{(x_i - x(t; \theta, M, X))^2 + (y_i - y(t; \theta, M, X))^2}$$

The goal is to minimize this cost function:

$$(\theta^*, M^*, X^*) = \arg \min_{\theta, M, X} J(\theta, M, X)$$

The gradient (derivative) of this function with respect to the parameters is complex and non-convex, so instead of solving analytically, we use a numerical optimization method.

The **Differential Evolution** algorithm performs this minimization iteratively:

$$p_i^{(t+1)} = p_{r1}^{(t)} + F \times (p_{r2}^{(t)} - p_{r3}^{(t)})$$

where:

- p_i = candidate parameter vector (θ, M, X)
- F = differential weight controlling mutation scale
- $r1, r2, r3$ = random distinct indices

This process continues until convergence to the minimum value of $J(\theta, M, X)$.

5. Results

After optimization, the following parameter values were obtained:

$$\boxed{\theta = 30.0^\circ, \quad M = 0.03, \quad X = 55.0}$$

The corresponding L1 distance between the predicted and experimental points was:

$$L1 = 15.06$$

—

6. Final Equation of the Fitted Curve

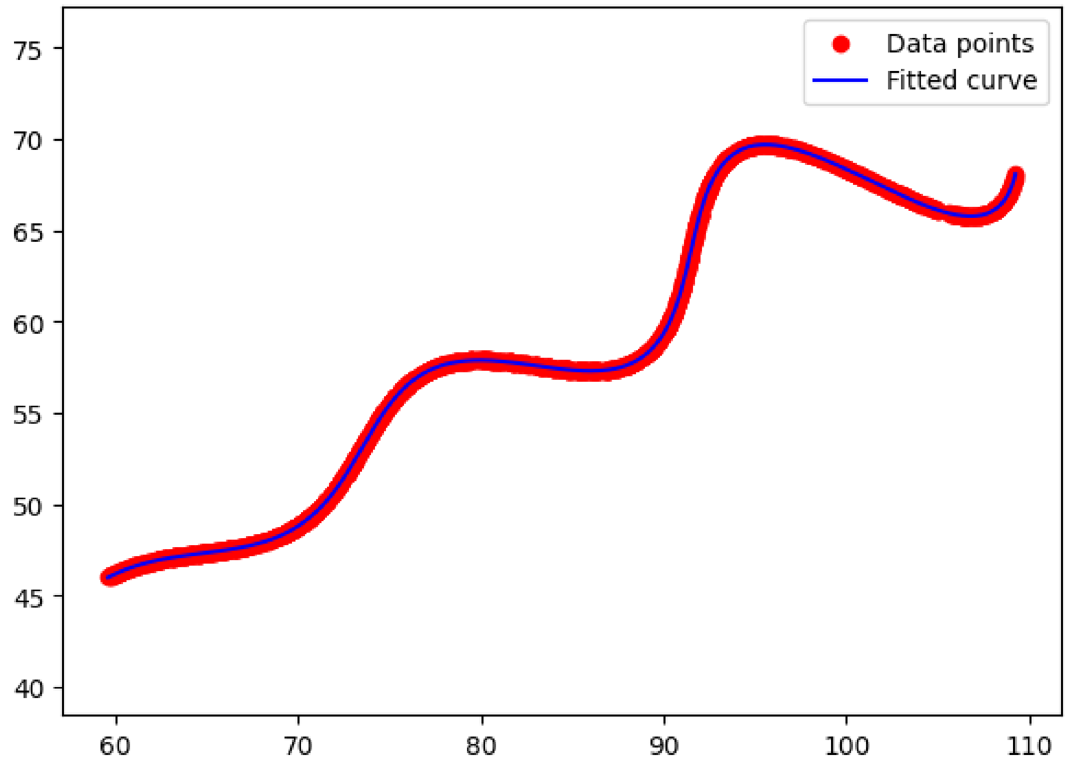
The final fitted curve can be written as:

$$\left(t\cos(0.524) - e^{0.03|t|}\sin(0.3t)\sin(0.524) + 55,\ 42 + t\sin(0.524) + e^{0.03|t|}\sin(0.3t)\cos(0.524)\right)$$

where $0.524 = 30^\circ$ in radians.

7. Visualization

A plot of the data points and the fitted curve is shown below.



8. Conclusion

- The estimated parameters $\theta = 30^\circ$, $M = 0.03$, $X = 55$ provide a good fit to the data.
- The total L1 distance of 15.06 indicates a low deviation between predicted and actual points.
- Differential Evolution successfully optimized a nonlinear, non-differentiable cost function without requiring analytical derivatives.
- This approach generalizes well to any parametric model where analytical solutions are not possible.

Submitted by: SUMAN PAL
Roll No: 142402012