Non-Linear Regression Analysis

Subhadip Baidya, 221092

November 9, 2024

1 Introduction

In this project, a given dataset of paired observations $(t, y(t))$ is analyzed using three proposed non-linear models. The objective is to determine the model that best represents the data by estimating the unknown parameters through the least squares method.

2 Assumptions and Model Descriptions

Assumptions: The dataset $(t, y(t))$ includes an error term $\epsilon(t)$, assumed to be i.i.d. with mean zero and variance σ^2.

- **Model 1:**
$$y(t) = \alpha_0 + \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + \epsilon(t)$$
- **Model 2:**
$$y(t) = \frac{\alpha_0 + \alpha_1 t}{\beta_0 + \beta_1 t} + \epsilon(t)$$
- **Model 3:**
$$y(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \epsilon(t)$$

3 Question 1: Calculating LSE

Model 1: The least squares estimators (approx) are represented by

$$\hat{\theta} = [\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2] = [-170.6624, 3.3393, 1.3606, 174.0554, 0.0052]$$

Model 2: The least squares estimators (approx) are represented by

$$\hat{\theta} = [\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_0, \hat{\beta}_1] = [-1.9571, -0.8157, -0.2928, 0.1332]$$

Model 3: The least squares estimators (approx) are represented by

$$\hat{\theta} = [\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4] = [6.7069, 5.8999, 1.2776, 3.9131, -0.49]$$

4 Question 2: Initial Guess

The models given are non-linear in nature so to find the least square estimators for those unknown parameters we need to use numerical methods. For this project, the Gauss-Newton numerical method is used. For iterative approach like Gauss-Newton method to work well, choosing good initial guess is important as it may happen choosing a random initial guess the model converges to a local minima instead of the global one. To get the initial guess for each model the grid-search algorithm is used for a range. For guessing the initial range for the parameters Curvefit function of Spicy Library was used. For **Model 1**, the initial guess used is $\theta_0 = [-170, 3, 1.388, 174, 0.006]$ whereas for **Model 2** $\theta_0 = [-2, 0, -0.278, 0.166]$ and for **Model 3** $\theta_0 = [6.7, 5.889, 1.333, 3.88, -0.5398]$ are used.

5 Question 3: Best Fitted Model

For model 1 the Least Square error is 0.2735. For model 2 the Least Square error is 0.3171 wherever for model 3 the Least Square error is 0.2685. So comparing the Least Square Errors for all three models, model 3 is said to be the best fitted model as it has the least error.

6 Question 4: Variance Estimation

As model 3 is the best fitted model according the errors calculated so the variance(σ^2) is calculated using this model and the value found of σ^2 is 0.0026848. But if model 1 was considered then the value of σ^2 would be 0.00273 and for model 2 would be 0.00317.

7 Question 5: Confidence Intervals

Since we are assuming the residuals are normally distributed, we can use the Fisher Information Matrix to estimate the confidence intervals for the parameters in our model. In non-linear regression, when assuming normality, the Fisher Information Matrix $I(\theta)$ can be approximated by the product of the Jacobian transpose and the Jacobian matrices, denoted $J^\top J$.

To compute the confidence intervals, we follow these steps:

First, we calculate the Jacobian matrix J, which consists of the partial derivatives of the residuals with respect to each model parameter. The Fisher Information Matrix is then approximated as:

$$I(\theta) \approx J^\top J$$

The covariance matrix of the parameter estimates is given by:

$$\text{Cov}(\hat{\theta}) = \sigma^2 (J^\top J)^{-1}$$

where σ^2 is the estimated variance of the residuals. The standard errors for each parameter are obtained from the square root of the diagonal entries in the covariance matrix. Assuming a normal distribution, a 95% confidence interval for each parameter θ_i is given by:

$$\theta_i \pm 1.96 \times \mathrm{SE}(\theta_i)$$

where $\mathrm{SE}(\theta_i)$ is the standard error of θ_i.
The estimated parameter values and their 95% confidence intervals for each model are as follows:

Model 1		
Parameter	**Estimate**	**95% Confidence Interval**
α_0	-170.6624	(-243.1532, -98.1716)
α_1	3.3393	(1.5547, 5.1239)
β_1	1.3606	(1.0986, 1.6225)
α_2	174.0554	(101.5665, 246.5442)
β_2	0.0052	(-0.0054, 0.0159)

Table 1: Confidence Intervals for Model 1 Parameters

Model 2		
Parameter	**Estimate**	**95% Confidence Interval**
α_0	-1.9517	(-1009.2955, 1005.3812)
α_1	-0.8157	(-420.6575, 419.0261)
β_0	-0.2928	(-150.9877, 150.4021)
β_1	0.1332	(-68.4338, 68.7002)

Table 2: Confidence Intervals for Model 2 Parameters

Model 3		
Parameter	**Estimate**	**95% Confidence Interval**
β_0	6.7069	(6.6529, 6.7608)
β_1	5.8999	(5.1671, 6.6327)
β_2	1.2776	(-1.6521, 4.2073)
β_3	3.9131	(-0.4349, 8.2611)
β_4	-0.4900	(-2.6260, 1.6459)

Table 3: Confidence Intervals for Model 3 Parameters

8 Question 6: Residual Plots

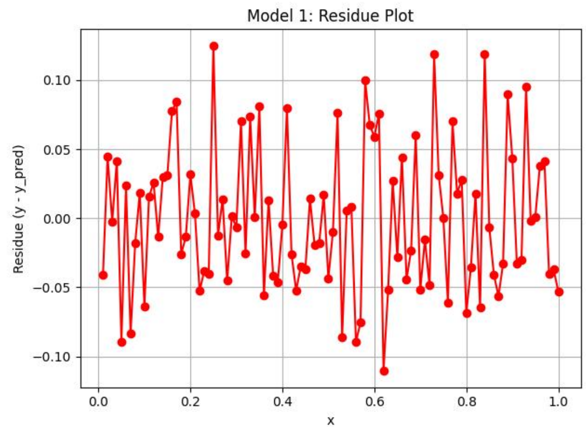

Figure 1: Residual plot for Model 1

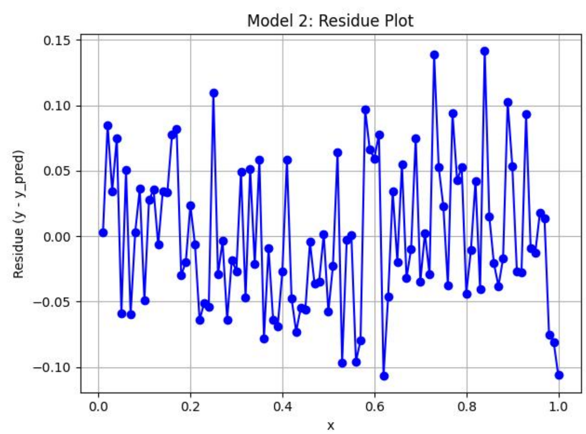

Figure 2: Residual plot for Model 2

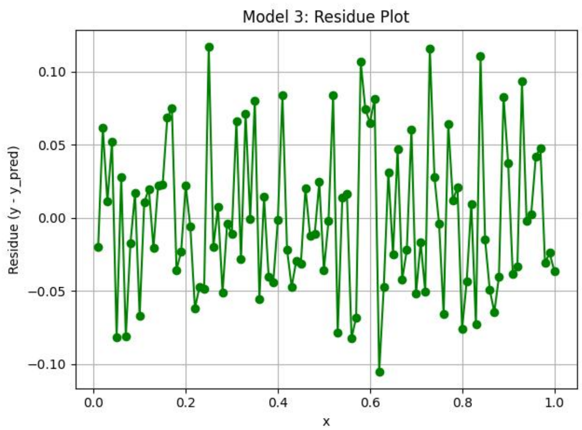

Figure 3: Residual plot for Model 3

9 Question 7: Normality Assumption

For the test of normality assumption for non-linear models Shapiro-Wilk test was conducted and the p-value for each model was calculated for the assumption. If p-value is greater than 0.05 then we said that the normality assumptions are satisfied. For model-1 the p-value is: 0.0595 , model-2 p-value is: 0.079, model-3 p-value is: 0.02581. As the p-value is greater than 0.05 we can surely say that the assumption of normality is satisfied for model 1 and 2.

10 Question 8: Data and Model Fit Plots

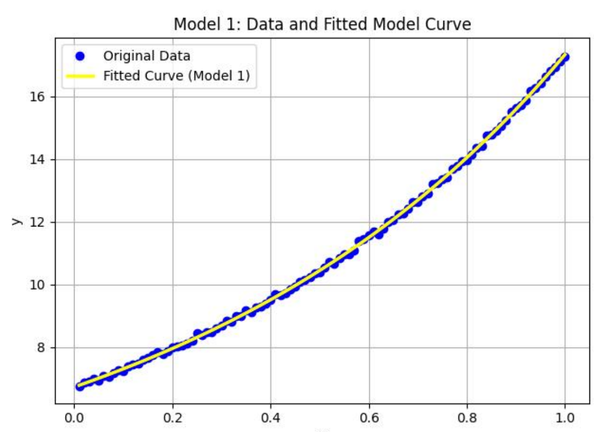

Figure 4: Observed Data and Fitted Curve for Model 1

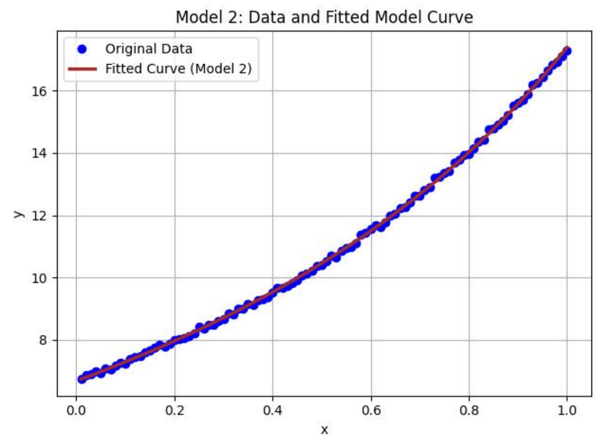

Figure 5: Observed Data and Fitted Curve for Model 2

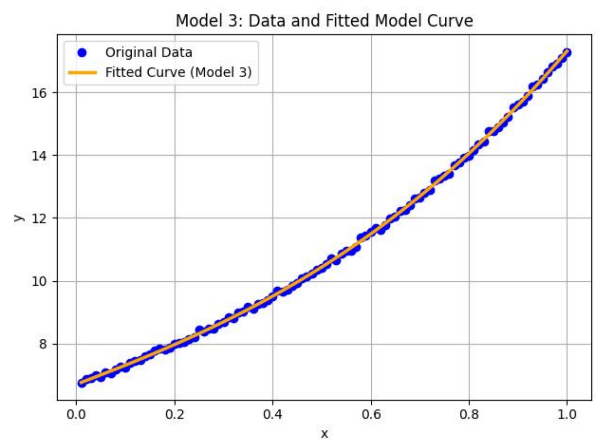

Figure 6: Observed Data and Fitted Curve for Model 3