PH3120 - Computational Physics Laboratory 1
CPL101 - Curve Fitting
Section 1 & 2

1. **Interpolation vs Regression**

   I. Identify the fundamental difference between interpolation and regression.

   II. Explain the advantages and disadvantages of the interpolation.

   III. Explain the advantages and disadvantages of the regression.

   IV. Assume that you have some data points ($y(t)$) within a given time (t) interval. You need to predict the y values beyond that time interval. What is the technique you use (interpolation or regression)?

   V. How do outliers and noise in the data affect interpolation?

   VI. How do outliers and noise in the data affect regression?

2. **Least Squares Regression**

   I. Answer the following questions based on the concepts of Least square regression.

      (a) What is Least Squares Regression, and what is its primary objective?

      (b) How does Least Squares Regression find the line of best fit in a linear regression model?

      (c) In Least Squares Regression, how are the coefficients (slope and intercept) of a straight line estimated?

      (d) What are the assumptions of Least Squares Regression?

      (e) What is the significance of the residual sum of squares (RSS) in Least Squares Regression?

      (f) What is "mean absolute error"?

      (g) How is the quality of a Least Squares Regression model evaluated using mean absolute error?

   II. Consider the simple harmonic motion of an object attached to a spring. The time period (T) of the motion is given by the following equation.

$$T = 2\pi\sqrt{\frac{m}{k}}$$

Here, m is the mass of the object and k is the spring constant.

(a) Rewrite the above expression in linear form ($y = bx$).
(b) Find the spin constant (k) using Numpy and least square fit based on the data in Table 1.

m (kg)	T (s)
0.1	1.4
0.3	2.4
0.5	3.1
0.7	3.7
0.9	4.2
1.1	4.6
1.3	5.1
1.5	5.4
1.7	5.7
1.9	6.14

Table 1: Time period (T) of a simple harmonic motion for several masses (m).

3. Polynomial Regression

I. Answer the following questions based on the concepts of Polynomial Regression.

(a) What is Polynomial Regression, and how does it differ from Least Squares Regression?

(b) How does Polynomial Regression model the relationship between variables?

(c) What is the degree of a polynomial in Polynomial Regression, and what does it represent?

(d) How are the coefficients of a polynomial equation estimated in Polynomial Regression?

(e) What is mean squared error?

(f) How is the quality of a Polynomial Regression model evaluated using mean squared error?

(g) How does the choice of polynomial degree impact the model's performance and complexity?

(h) How can overfitting be addressed in Polynomial Regression?

II. In a particle accelerator experiment, the momentum (p) of accelerated particles is measured at different velocities (v) in relevant units. The following data is collected:

$$u = [0.2,\ 0.4,\ 0.6,\ 0.8,\ 1.0,\ 1.2,\ 1.4,\ 1.6,\ 1.8,\ 2.0]$$
$$p = [1.3,\ 2.7,\ 4.9,\ 8.1,\ 12.5,\ 18.9,\ 27.3,\ 37.7,\ 50.1,\ 64.5]$$

Based on particle physics, the relativistic-momentum equation is given by,

$$p = \gamma m u, \text{ where } \gamma = \frac{1}{\sqrt{1-\frac{u^2}{v^2}}}$$

Here, u is the velocity of the particle, c is the speed of light, and m is the rest mass of an electron.

(a) Use polynomial regression and mean squared error to find the best degree of polynomial for the above data.

(b) By plotting the original data and polynomial curves with different degrees, prove your answer for part (a) is correct.

(c) Estimate the coefficients and write down the equation of the best-fit polynomial.

(d) Find p when $u = 0.9c$.

4. Lagrange Polynomial interpolation

I. Answer the following questions based on the concepts of Lagrange Polynomial Interpolation.

(a) What is the main idea behind Lagrange Polynomial Interpolation?

(b) How is the Lagrange interpolating polynomial constructed using the given data points?

(c) What is the Lagrange basis function, and how does it contribute to the construction of the interpolating polynomial?

(d) How does Lagrange Polynomial Interpolation ensure that the interpolating polynomial passes through all the given data points?

(e) What is the role of the Lagrange interpolating polynomial in estimating values at points between the given data?

(f) What are the limitations or potential issues associated with Lagrange Polynomial Interpolation?

II. Use the **‘lagrange’** function from the **spicy.interpolate** module to solve the following problem.

A ball is shot vertically upwards, and its height *(h)* in meters above the ground is recorded at different time intervals *(t)* in seconds. The recorded data is as follows:

t = [0, 1, 2, 3, 4, 5]

h = [0, 5, 20, 45, 80, 125]

(a) Construct the Lagrange interpolating polynomial based on the given data points.

(b) Print the coefficients of the above polynomial.

(c) Calculate the mean squared error for the above polynomial.

(d) Plot the original data and polynomial curve found against the time intervals.

(e) Estimate the velocity of the ball at time $t = 2.5$ seconds.

5. Linear Interpolation

I. Answer the following questions based on the concepts of Linear Interpolation.

(a) What is the underlying assumption behind linear interpolation?

(b) How does linear interpolation estimate values between two known data points?

(c) Can linear interpolation be used to estimate values outside the range of the given data points? Why or why not?

(d) How does the distance between the target position and the known data points affect the accuracy of linear interpolation?

(e) What are the limitations of linear interpolation in capturing complex or nonlinear relationships in the data?

(f) How can you determine whether linear interpolation is appropriate for a given set of data?

II. Use the **interp1d** function from the **spicy.interpolate** module to solve the problem in question 5-II. Compare your answers from Lagrange and linear interpolations.