Polynomial regression is a type of regression analysis in which the relationship between the independent variable x and the dependent variable y is modeled as an nth degree polynomial. It is used to model nonlinear relationships between variables.

The history of polynomial regression can be traced back to the 19th century when mathematicians like **Adrien-Marie Legendre and Carl Friedrich Gauss developed methods to fit polynomial curves to data.** However, the specific person credited as the "founder" of polynomial regression is not clear, as the concept has been developed and refined by many mathematicians and statisticians over the years.

One of the earliest known uses of polynomial regression was by **Sir Francis Galton** in 1877, who used it to analyze the relationship between parents' heights and their children's heights. Since then, the technique has been widely used in many fields, including physics, engineering, economics, and social sciences.

Polynomial regression is a type of linear regression where the relationship between the independent variable(s) and the dependent variable is modeled as an nth-degree polynomial function.

The general formula for polynomial regression with a single independent variable (i.e., simple polynomial regression) is:

y = b0 + b1*x + b2*x^2 + ... + bn\*x^n + e

where:

* y is the dependent variable (i.e., the target variable you want to predict)
* x is the independent variable
* b0, b1, b2, ..., bn are the coefficients of the polynomial function, which are estimated from the data
* n is the degree of the polynomial
* e is the error term or residual, which represents the unexplained variation in the dependent variable that is not accounted for by the independent variable(s)

The formula for polynomial regression with multiple independent variables (i.e., multiple polynomial regression) is similar, except that there are multiple independent variables and the polynomial function includes cross-terms (i.e., interactions) between them.

To perform polynomial regression in practice, you can use various software packages and libraries such as scikit-learn in Python.

* Let the quadratic polynomial regression model be

y = a0 + a1x + a2x^2

where:

* y is the dependent variable
* x is the independent variable
* a0, a1, and a2 are coefficients that determine the shape of the curve

This equation represents a parabolic curve, which can have a concave-up or concave-down shape depending on the sign of the coefficient a2. If a2 is positive, the parabola opens upwards, while if a2 is negative, the parabola opens downwards.

This type of equation is commonly used in mathematics and physics to model a wide range of phenomena, such as the trajectory of a projectile, the shape of a lens, or the behavior of an electrical circuit.

a1 and a2 are called the linear effect parameter and quadratic effect

* The values of **, and** are calculated using the following system of equations:

| x | Y |
| --- | --- |
| 3 | 2.5 |
| 4 | 3.2 |
| 5 | 3.8 |
| 6 | 6.5 |
| 7 | 11.5 |

| x | y |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 3 | 2.5 | 9 | 27 | 81 | 7.5 | 22.5 |
| 4 | 3.2 | 16 | 64 | 256 | 12.8 | 51.2 |
| 5 | 3.8 | 25 | 125 | 625 | 19 | 95 |
| 6 | 6.5 | 36 | 216 | 1296 | 39 | 234 |
| 7 | 12 | 49 | 343 | 2401 | 80.5 | 563.5 |
| **25** | **27.5** | **135** | **775** | **4659** | **158.8** | **966.2** |

Using the given data we

Solving this system of equations we get

| x | y (predicted)(  ) | Y pred -Y |
| --- | --- | --- |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |
| 6 |  |  |
| 7 |  |  |
|  |  |  |
|  |  |  |

R^2= ∑(y p -ẏ)2 / ∑(y- ẏ) 2

| x | y (predicted)(  ) | **y p -ẏ** | **(y p -ẏ)2** | **y- ẏ** | **(y- ẏ) 2** |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

In the above program, the degree 2 in the PolynomialFeatures function is the degree of the polynomial regression model that we want to fit.

When we set the degree to 2, we are creating a quadratic polynomial regression model, which means that the model will fit a curve to the data that is a quadratic equation of the form y = a + bx + cx^2.

By including higher degrees in the polynomial regression model, we can fit more complex curves to the data. However, it's important to note that increasing the degree of the polynomial can also lead to overfitting, where the model fits the noise in the data instead of the underlying pattern. So, it's important to choose an appropriate degree for the polynomial regression model based on the complexity of the data and the desired accuracy of the model.

fits a third-degree polynomial curve to the data points (x,y). The function numpy.polyfit() computes the coefficients of the polynomial curve that best fits the data, and numpy.poly1d() creates a polynomial object that can be used to evaluate the curve at any point. The resulting mymodel object is a third-degree polynomial function that can be used to predict the value of y for any given value of x.

The line yp = mymodel(11) uses the polynomial function mymodel to predict the value of y for x=11. The predicted value is then printed with the print(yp) statement.

A third-degree polynomial curve, also known as a cubic polynomial, is a mathematical function of the form:

y = ax^3 + bx^2 + cx + d

where x is the independent variable, y is the dependent variable, and a, b, c, and d are constants that determine the shape and position of the curve.

A third-degree polynomial curve is characterized by having one "hump" or "valley" in the curve, which means it can capture more complex and non-linear relationships between variables than a simple linear regression model.

In the context of polynomial regression, fitting a third-degree polynomial curve to a set of data means finding the values of a, b, c, and d that minimize the sum of the squared errors between the predicted values of y and the actual values of y for the given values of x.

A second-degree polynomial curve, also known as a quadratic polynomial, is a mathematical function of the form:

y = ax^2 + bx + c

where x is the independent variable, y is the dependent variable, and a, b, and c are constants that determine the shape and position of the curve.

A second-degree polynomial curve is characterized by having one "peak" or "valley" in the curve, which means it can capture non-linear relationships between variables better than a simple linear regression model but less complex than a third-degree polynomial.

In the context of polynomial regression, fitting a second-degree polynomial curve to a set of data means finding the values of a, b, and c that minimize the sum of the squared errors between the predicted values of y and the actual values of y for the given values of x.

the equation y = a0 + a1x + a2x^2 + c is a **second-order e**quation because it contains the second-order term a2x^2. The term "order" in mathematics generally refers to the highest power of the derivative or variable that appears in the equation. In this case, the highest power of the independent variable x is 2, making the equation a second-order equation.

The second-order term a2x^2 determines the curvature of the quadratic curve, which can be concave-up or concave-down depending on the sign of a2. When a2 is positive, the curve opens upwards and is concave-up, while when a2 is negative, the curve opens downwards and is concave-down.

Polynomial regression is a type of regression analysis in which the relationship between the independent variable x and the dependent variable y is modeled as an nth degree polynomial. The third-order polynomial regression model is expressed as:

y = b0 + b1*x + b2*x^2 + b3\*x^3 + e

where:

* y is the dependent variable
* x is the independent variable
* b0, b1, b2, and b3 are the coefficients of the third-order polynomial model, which determine the shape of the curve.
* x^2 and x^3 represent the second and third powers of x, respectively.
* e is the error term, representing the random error in the model.

The equation can be rewritten as:

y = a0 + a1*x + a2*x*x +a3*x*x*x + e

The goal of polynomial regression is to find the values of the coefficients b0, b1, b2, and b3 that minimize the sum of the squared errors between the predicted values of y and the actual values of y. This is typically done using a method called least squares regression, which involves finding the line of best fit that minimizes the sum of the squared distances between the data points and the line.

Polynomial regression is a type of regression analysis in which the relationship between the independent variable x and the dependent variable y is modeled as an nth degree polynomial. The fourth-order polynomial regression model is expressed as:

y = b0 + b1*x + b2*x^2 + b3*x^3 + b4*x^4 + e

where:

* y is the dependent variable
* x is the independent variable
* b0, b1, b2, b3, and b4 are the coefficients of the fourth-order polynomial model, which determine the shape of the curve.
* x^2, x^3, and x^4 represent the second, third, and fourth powers of x, respectively.
* e is the error term, representing the random error in the model.

The equation can be rewritten as:

y = b0 + b1*x + b2*x*x + b3*x*x*x + b4*x*x*x*x + e

The goal of polynomial regression is to find the values of the coefficients b0, b1, b2, b3, and b4 that minimize the sum of the squared errors between the predicted values of y and the actual values of y. This is typically done using a method called least squares regression,