Polynomial Regression

Polynomial Regression is a form of **linear regression** where the relationship between the **independent variable** x and the **dependent variable** y is modeled as an n-degree polynomial.

It is used when data shows a **non-linear** relationship that a straight line (simple linear regression) cannot fit well.

Mathematical Representation

A **degree-n polynomial regression** is:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \cdots + \beta_n x^n + \epsilon$$

- $\beta_0, \beta_1, \ldots, \beta_n$: model coefficients

  Although the equation is nonlinear in terms of x, it is **linear in parameters** βi, hence it is still considered a linear model.

🔹 Why Use Polynomial Regression?

- To capture **curved trends** in the data.
- Useful when plotting residuals of linear regression shows **non-random patterns**.
- Helps model **complex relationships** with more flexibility.

🔹 Visual Intuition

Degree	Fit Shape	Use Case
1	Line	Linear relationship
2	Parabola	Single curve (e.g., U-shape trends)
3	S-curve	More complex patterns

4+	Wiggles/Waves	Even more complexity, but beware overfitting

Choosing the Right Degree

- **Too low** → underfitting (high bias)
- **Too high** → overfitting (high variance)

Use techniques like:

- **Cross-validation**
- **Plotting learning curves**
- **Regularization (Ridge/Lasso with polynomial features)**

🔹 Advantages

✅ Captures non-linear relationships
✅ Easy to implement
✅ Flexible model behavior

🔹 Disadvantages

❌ Sensitive to outliers
❌ Can overfit with high degree
❌ Model interpretability decreases with degree

🔹 Real-world Applications

- Predicting population growth
- Modeling economic trends
- Fitting curves in physics/engineering experiments

🔹 Summary

Concept	Description
Type	Supervised, Regression
Model Shape	Polynomial curve
Use case	When data shows non-linear patterns
Key Concern	Overfitting with high-degree polynomials