

Linear Regression (Difficulty: 1/3)

Linear Regression is one of the simplest and most widely used machine learning algorithms. It helps us understand the relationship between two variables by fitting a straight line to the given data points.

1. What is Linear Regression?

Linear Regression tries to model the relationship between a dependent variable (y) and an independent variable (x) by fitting a line of the form:

$$y = m \cdot x + c$$

- **m** = slope of the line
- **c** = intercept (value of y when x = 0)

2. Why is it useful?

- Predicting future values
- Identifying trends
- Understanding how two variables are related

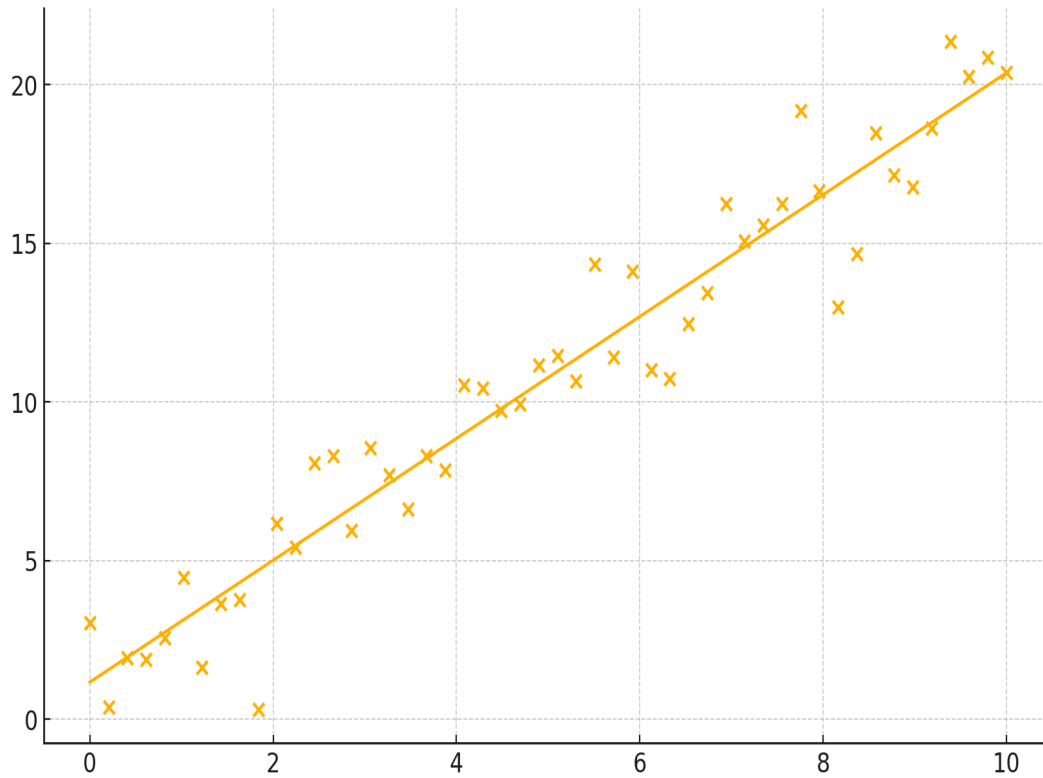
3. How does it work?

Linear Regression finds the best-fitting line by minimizing the error between the predicted values and actual values. This is usually done using the "Least Squares Method."

4. Example

If we want to predict a student's marks based on hours studied, Linear Regression can help discover a pattern and form an equation that predicts marks for any number of hours studied.

Below is a simple example of Linear Regression plotted on a graph:



5. Key Terms

- **Data Points:** The actual observed values.
- **Regression Line:** The line that best fits the data.
- **Error:** Difference between actual and predicted values.
- **Overfitting/Underfitting:** When the model learns too much noise or too little pattern.

6. Advantages

- Very simple to understand
- Fast to compute
- Works well for linearly related data

7. Limitations

- Doesn't work well for non-linear data
- Sensitive to outliers

Conclusion

Linear Regression is the foundation of many advanced machine learning models. Understanding it makes it easier to learn more complex algorithms later.

End of Notes