

# Data Plotting and Dimensionality in Regression

Raghda Al Taei

2024

## 1 Introduction

In data analysis and machine learning, visualizing and manipulating data is crucial. This article discusses a dataset of 25 one-dimensional points, where the first column represents the x-values and the second column corresponds to the y-values. We will explore how to plot this data and the importance of adding a new dimension to the x-values.

## 2 Plotting Data with $y$ as a Function of $x$

The first step is to visualize the relationship between the x and y values. A simple scatter plot is effective. Below is the MATLAB code:

The code loads the data, extracts x and y values, and plots them.

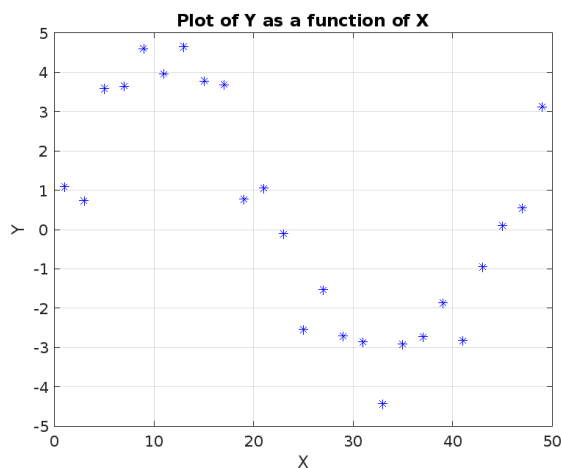


Figure 1: Scatter Plot of Data Points

### 3 Adding a New Dimension to the X Data

Next, we enhance the x-values by adding a new dimension where all values are equal to 1:

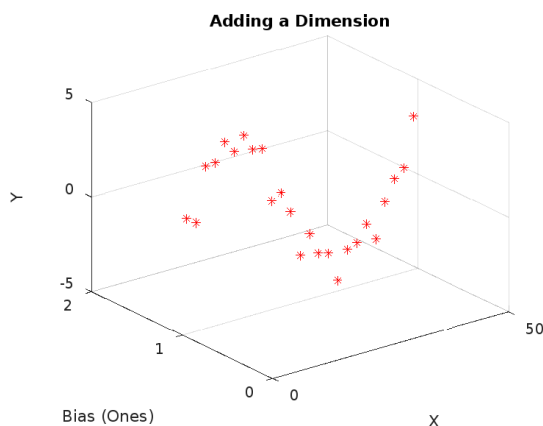


Figure 2: Adding a New Dimension to the X Data

#### 3.1 Explanation of the Effect

Adding a column of ones introduces a bias term, which is crucial in regression models:

- **Bias Term:** This allows the model to learn an intercept in addition to the slope when fitting a line to the data. Without this term, the model assumes the line passes through the origin, which may not be accurate.
- **Effect on Plot:** This transformation does not change the scatter plot visually, but it improves the model's ability to capture the underlying relationship by accounting for a possible intercept.

### 4 Repository Code

You can find the complete code and further analysis in my GitHub repository: [GitHub Repository](#).

### 5 In Summary

Adding a column of ones enables better data fitting by considering a constant bias alongside the input values, which is essential for improving prediction ac-

curacy.