

# How can I perform least squares fitting in NumPy? Can you provide an example?

Authored by  
**stats writer**

June 27, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I perform least squares fitting in NumPy? Can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154674>

Least squares fitting is a statistical method used to find the best fitting line for a set of data points. This method is commonly used in data analysis and machine learning. In NumPy, the "polyfit" function can be used to perform least squares fitting by finding the optimal coefficients for a polynomial equation that best fits the given data. An example of using the "polyfit" function in NumPy is as follows:

```
import numpy as np
x = np.array()
y = np.array()

# fit a linear least squares line to the data
coefficients = np.polyfit(x, y, 1)

# print the optimal coefficients
print(coefficients)

# output:
```

This output indicates that the best fitting line for the given data is  $y = 2x + 1$ . In this way, NumPy provides a simple and efficient method for performing least squares fitting.

## Perform Least Squares Fitting in NumPy (With Example)

**The method of least squares is a method we can use to find the regression line that best fits a given dataset.**

**We can use the `linalg.lstsq()` function in NumPy to perform least squares fitting.**

**The following step-by-step example shows how to use this function in practice.**

### Step 1: Enter the Values for X and Y

**First, let's create the following NumPy arrays:**

```
import numpy as np

#define x and y arrays
x = np.array()

y = np.array()
```

**Step 2: Perform Least Squares Fitting**

**We can use the following code to perform least squares fitting and find the line that best "fits" the data:**

```
#perform least squares fitting
np.linalg.lstsq(np.vstack().T, y, rcond=None)

array()
```

**The result is an array that contains the slope and intercept values for the line of best fit.**

**From the output we can see:**

**Slope: 0.969 Intercept: 7.767**

**Using these two values, we can write the equation for**

**the line of best fit:**

$$y = 7.767 + 0.969x$$

**Step 3: Interpret the Results**

**Here's how to interpret the line of best fit:**

**When  $x$  is equal to 0, the average value for  $y$  is 7.767. For each one unit increase in  $x$ ,  $y$  increases by an average of .969.**

**For example, if  $x$  has a value of 10 then we predict that the value of  $y$  would be 17.457:**

$$y = 7.767 + 0.969x \quad y = 7.767 + 0.969(10) \quad y = 17.457$$

**Bonus: Video Explanation of Least Squares Fitting**

**Refer to the video below for a simple explanation of least squares fitting:**

**The following tutorials explain how to perform other common tasks in NumPy:**