A. Multiple Linear Regression

Saturday, 30 August 2025 9:33 PM

1. Introduction

Multiple Linear Regression is Linear Regression with Multi Features/Variables whereas Linear Regression with One Feature/Variable is called Univariate Regression.

Let's take an example. Earlier we had Linear Regression with one variable (size in feet^2 (x)) that would predict (Price ($) in 1000's (y)). But now, what if we have multiple features such as the size of the house, number of bedrooms, number of floors, age of house in years. This would give us a much better context to predict the house price.

Multiple features (variables)

	Size in feet^2	Number of bedrooms	Number of floors	Age of home in years	Price ($) in $1000's
	X_1	X_2	X_3	X_4	
	2104	5	1	45	460
$i=2$	1416	3	2	40	232
	1534	3	2	30	315
	852	2	1	36	178
	...	...	...	...	...

$j = 1 \ldots 4$

$n = 4$

x_j = j^{th} feature

n = number of features

$\vec{x}^{(i)}$ = features of i^{th} training example

$x_j^{(i)}$ = value of feature j in i^{th} training example

$\vec{x}^{(2)} = [1416 \ 3 \ 2 \ 40]$

$x_3^{(2)} = 2$

Now that we have multiple features, let's look at what our model would look like.

Previously for LR with One Feature, we had,

$$f_{w,b}(x) = wx + b$$

Now, for LR with Multi Features, the model will look like this:

$$f_{w,b}(X) = w_1 X_1 + w_2 X_2 + w_3 X_3 + w_4 X_4 + b$$

For the housing-price example, let's say we have a model:

$$f_{w,b}(X) = 0.1 X_1 + 4 X_2 + 10 X_3 + -2 X_4 + 80$$

X_1: size, X_2: #bedrooms, X_3: #floors, X_4: years, 80: base price

Here, **b = 80** that means a house, assuming it has no size, no bedrooms, no floors, no age in years, it would have a base price of **80,000 $.**
$0.1x_1$ means that for increase in every additional square foot, the price of house will go up by **0.1(1000 $) = 100 $**
$4x_2$ means the addition in house price will be **4*1000 $ = 4000 $** for increase by 1 bedroom.
$10x_3$ means, with the increase of 1 floor, the price increment would be **10 * 1000 $ = 10000 $**
$-2x_4$ means, with the increase in house age in years by 1 year, the price will decrement by **2 * 1000 $ = 2000 $**

In general, if we have n features, the Multiple LR model will be:

$$f_{w,b}(x) = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b$$

Let's simplify our formula:
We'll define: w as a list of numbers for parameters [w_1, w_2, … w_n]. This is a row vector.
x as a list of all features [x_1, x_2, x_3, … x_n]. This is a row vector.
Now,

$$f_{\vec{w},b}(\vec{x}) = \vec{w} \cdot \vec{x} + b = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n + b$$

dot product