

[LinearRegression.fit\(\)](#) is a python function from the sklearn library that is used to approximate the best fit for the given input and output. What it does is consider a variety of linear functions of the form :

$$y=a+\sum b_i x_i \text{ (Multivariable Regression) (} x_i \text{ is input}$$

and b_i is parameter coeff)

and returns the best fit for our training set. It tries to minimize the average error from a linear model to each of the points.

Degree	Variance	Bias	Irreducible_error	Total_error
1	21119.180670	573.923766	4.672529e-12	513590.841371
2	44627.267234	567.201418	4.779395e-11	514391.787545
3	51104.362752	47.088141	3.637979e-12	55317.540824
4	77044.471429	50.138500	-1.728040e-12	81202.430673
5	84405.232754	48.679513	4.547474e-14	88062.116134
6	111514.917441	49.397690	-3.637979e-13	115454.719361
7	131157.888689	55.930659	3.637979e-12	135383.778473
8	150611.600984	61.328224	4.138201e-12	155809.288671
9	182132.759061	64.287127	-2.819434e-12	188725.575680
10	225725.676308	65.005919	1.000444e-12	231890.729272
11	241536.909524	65.557320	2.150955e-11	248836.928029
12	274899.567057	84.327113	1.414264e-11	290473.506272
13	255429.820319	67.234747	6.002665e-12	266302.527880
14	287134.711130	106.906166	-7.366907e-12	318341.459811
15	294114.494602	74.580276	-3.501555e-12	309418.689731

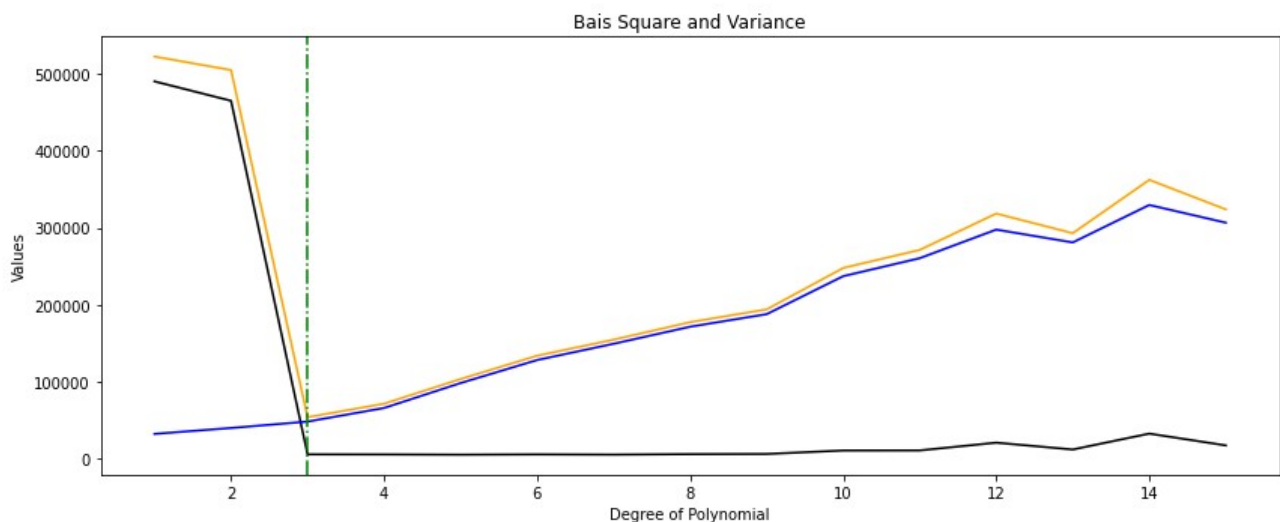
We see that the [Irreducible Error](#) is very small compared to bias and variance. From the above values it changes with the degree. This is to be expected because in some overfit and underfit predictions noise may get factored in as parameters or features.

Conclusion:

In these sets we can see the variance mostly increases with increasing degree whereas bias is very high till degree 2 and drops down at degree 3 maintaining a lower bias with a slight increase at the higher degrees.

Overall degree 3 seems to have a better balance of Variance and bias (irreducible error is very small) since its total error is the lowest among all of the degrees. It is the best fit since bias and variance are balanced too.

This can be visually seen in the graph below.



As we can see till degree 2 bias² is high and after degree 4 variance dominates and the least total error occurs at degree 3 as shown by the green dotted line.

Degree < 3 bias is dominant (underfit)

Degree > 3 variance is dominant (overfit)

Best fit is for degree 3 with least total error