Adv

Linear least squares regression has earned its place as the primary tool for process modeling because of its effectiveness and completeness.

Though there are types of data that are better described by functions that are nonlinear in the parameters, many processes in science and engineering are well-described by linear models. This is because either the processes are inherently linear or because, over short ranges, any process can be well-approximated by a linear model.

Dis

The main disadvantages of linear are limitations in the shapes that linear models can assume over long ranges, possibly poor extrapolation properties, and sensitivity to outliers.

Linear models with nonlinear terms in the predictor variables curve relatively slowly

while the method of normal equation often gives optimal estimates of the unknown parameters, it is very sensitive to the presence of unusual data points in the data used to fit a model. One or two outliers can sometimes seriously skew the results of a least squares analysis.that’s also the reason why we skewed the result when I fitted all features.

When n is low (n < 1000 or n < 10000) you can think of normal equations as the better option for calculation theta, however for greater values **Gradient Descent** is much more faster, so the only reason is the time :) n is the number of features.