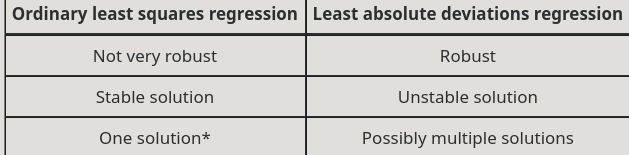
LAD

Least absolute deviation is an alternative to Ordinary Least Squares, whereLAD minimises the sum of absolute residuals, and OLS minimises the sum of squares of residuals.

LAD is more computationally expensive, and hence OLS was used more often, despite LAD predating OLS.

LAD is more robust than OLS (David Birkes and Yadolah Dodge. *Alternative methods of regression*. John Wiley & Sons, 2011.).

Summary: In OLS, a large residual for a single data point makes an outsized contribution to the loss function. This may cause predictions to be skewed.



Note: use same graph as https://colab.research.google.com/drive/1DjI3eDcwo0P-whG92qTTLHCdH1\_ImiYr

Gradient boosting

**https://en.wikipedia.org/wiki/Gradient\_boosting**

Note: tree docs

https://github.com/catboost/tutorials/blob/master/model\_analysis/visualize\_decision\_trees\_tutorial.ipynb