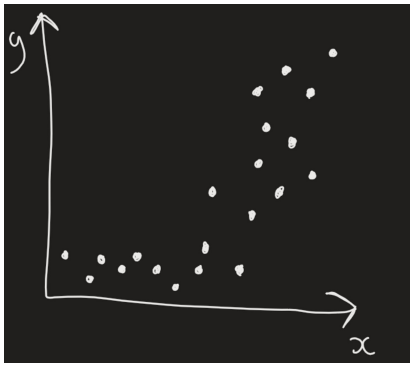
# Assignment question 1c i.

Tuesday, 31 March 2020 12:42 AM

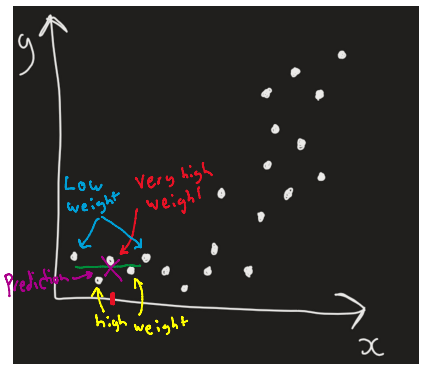
Question 1c i.
c. Now examine the behaviour of the training and test MSE, for a `loess` fit.
i. (1)Look up the `loess` model fit, and write a paragraph explaining how this fitting procedure works. In particular, explain what the `span` argument does. Add a (hand) sketch illustrating the method.

A loess model (Locally weighted Scatter-plot Smoother) is a non-parametric modelling approach (as linearity assumptions are relaxed) that fits a least squares regression to localised subsets. The value of the data points closest to each fitted point have a higher weight in determining the predicted value then those further away. The 'closeness' of the points around the fitted point can be controlled by using the 'span' argument in r. The value of span can be interpreted as the proportion of data that is considered 'close' and therefore influencing the fitted point.

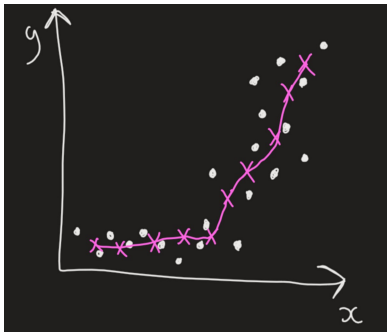
For example consider the following dataset



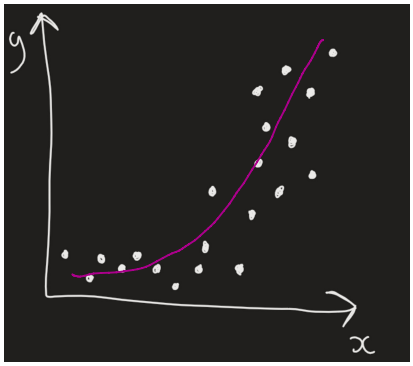

Using a loess model, y is predicted by observing the values of close data points.



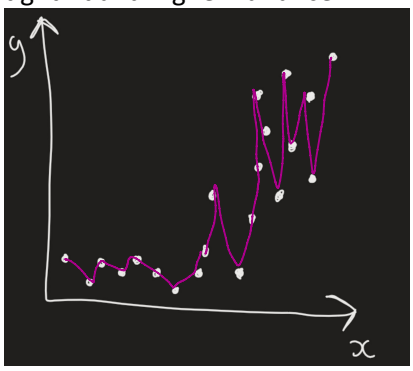

This process is then repeated to fit the whole dataset.



A high span means that a fitted value is influenced by a wider range of points which leads to a smooth fit with higher bias.




A low span means that each fitted value is influenced by small range or points which leads to a tight fit and higher variance.




References:

https://www.statsdirect.com/help/nonparametric_methods/loess.htm
http://r-statistics.co/Loess-Regression-With-R.html
StatQuest: Fiitting a curve to data, aka lowess, aka loess
R - LOWESS smoothing curve