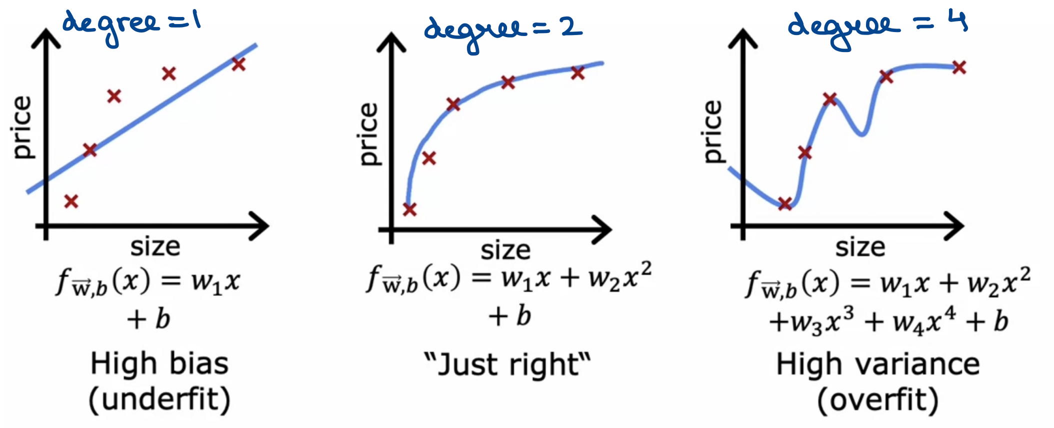
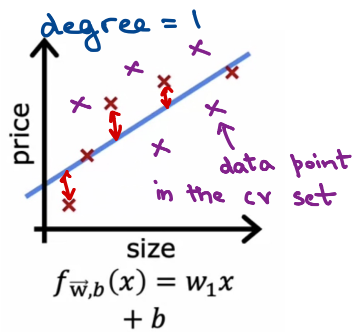


When there is less data, it is easy to determine from the graph whether a model has high bias, is just right or underfits.

In case where a lot of data is present, we measure J_{train}, J_{cv} and J_{test} error to determine which state is our model in.

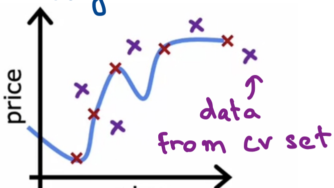


J_{train} will be high because the straight line doesn't accomodate majority of the features

Even J_{cv} set will be high

Usually if J_{train} is high then it has high bias.

degree = 4



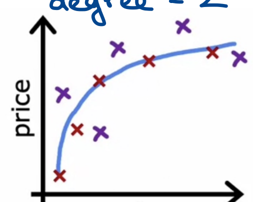
$$f_{\vec{w},b}(x) = w_1x + w_2x^2 + w_3x^3 + w_4x^4 + b$$

J_{train} will be low because it perfectly fits the graph

J_{cv} will be high because the graph is not aligned with new data because it hugs the training data and doesn't generalize

Since J_{cv} is much higher than J_{train} the data has high variance.

degree = 2

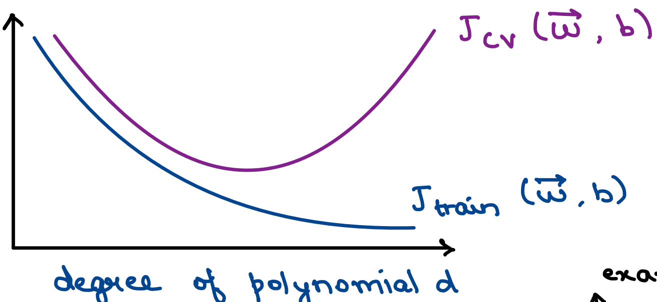


$$f_{\vec{w},b}(x) = w_1x + w_2x^2 + b$$

J_{train} is low because it is somewhat close to each training data

J_{cv} is also low because it generalizes to new data

Since J_{train} and J_{cv} are relatively close. Neither one's too high nor too low, the graph is just right.

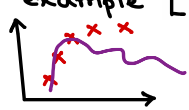


High bias (underfit)
J_{train} will be high
($J_{train} \approx J_{cv}$)

High variance (overfit)
$J_{cv} \gg J_{train}$
(J_{train} may be low)
$\gg$ means much greater

High bias and high variance
J_{train} will be high
and $J_{cv} \gg J_{train}$

example



fits perfectly for a part of data then becomes underfit.