https://www.coursera.org/learn/ml-regression/lecture/TIGJ5/symptoms-of-overfitting-in-polynomial-regression

Symptoms of overfitting in polynomial regression

review: bias - variance tradeoff.

ridge regression

when a model is overfit - the estimated co-efficients become very large.

https://www.coursera.org/learn/ml-regression/lecture/38mGi/overfitting-for-more-general-multiple-regression-models

Overfitting for more general multiple regression models

slide 9 : number of inputs affects overfitting.

Lots of features(inputs) = very flexible model – can easily overfit

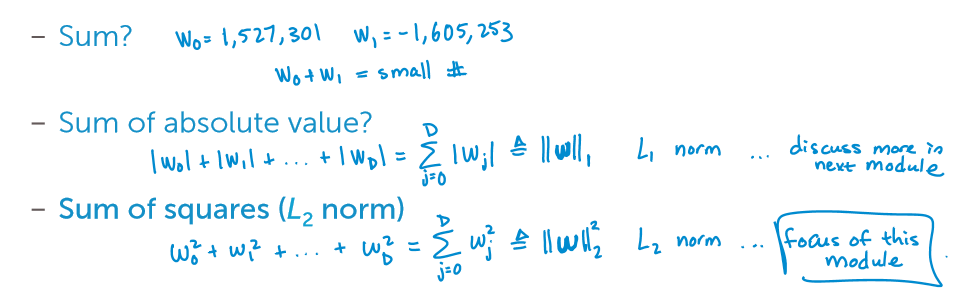
to avoid overfitting with single feature – must have data points for every possible value of that feature.

For multiple features it become more difficult – need data point for each combination of feature values.

Adding term – cost of fit to prefer small coeﬃcients – slide 11

we need to balance – fit of function to data vs magnitude of coefficients.

Formula on slide 15





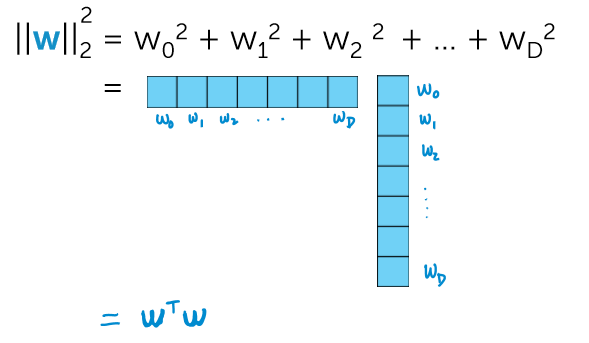
<https://www.coursera.org/learn/ml-regression/lecture/b1fbX/balancing-fit-and-magnitude-of-coefficients>

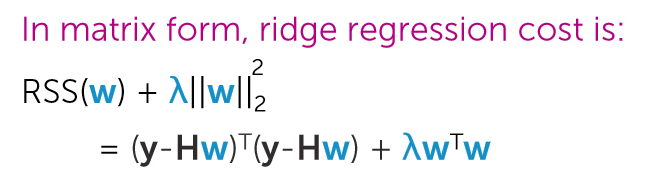
Balancing fit and magnitude of coefficients

https://www.coursera.org/learn/ml-regression/lecture/Ki05k/the-ridge-coefficient-path

The ridge coefficient path

||w||^2 (sub 2) calculation method – slide 28

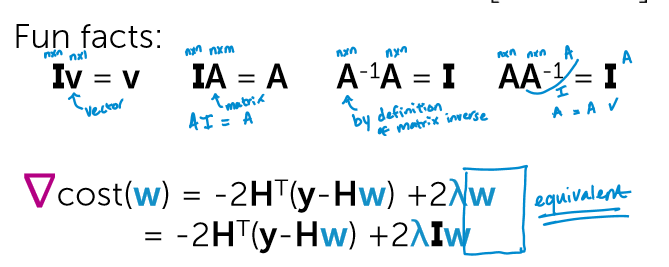




^ from slide 29

<https://www.coursera.org/learn/ml-regression/lecture/uYzai/approach-1-closed-form-solution>

Approach 1: closed-form solution



https://www.coursera.org/learn/ml-regression/lecture/SlFBr/discussing-the-closed-form-solution

Discussing the closed-form solution

slide 35 onwards

https://www.coursera.org/learn/ml-regression/lecture/UEH0h/selecting-tuning-parameters-via-cross-validation

Selecting tuning parameters via cross validation

starts @ slide 45

split available data > training + validation + test

which subset of data should I use as the validation set?

Average performance over validation sets drawn from all over the original data.

https://www.coursera.org/learn/ml-regression/lecture/FJcUw/k-fold-cross-validation

K-fold cross validation

starts at slide 52

K-fold cross validation

preprocessing – randomly assign data to k groups.

For each k group of data – treat the selected block as the validation set.

Compute error on validation blocks

^ is slide 53.

computer average error CV(lambda)

slide 59

best approximation occurs for validation sets of size 1. ie : K = number of data points.

This is computationally intensive.

Requires computing N fits of model per lambda.

NB: choosing the tuning parameters is an important step in the machine learning workflow.

<https://www.coursera.org/learn/ml-regression/lecture/3KZiN/how-to-handle-the-intercept>

How to handle the intercept