

Week 6 Lecture Notes

ML: Advice for Applying Machine Learning

Deciding What to Try Next

Errors in your predictions can be broken down into:

- Getting more training examples
- Trying smaller sets of features
- Trying additional features
- Trying polynomial features
- Increasing or decreasing λ

Don't put too much of either structural constraints, like "Simpson's diagnostic" techniques for checking one of the above solutions in the following sections.

Evaluating a Hypothesis

A hypothesis may have low error for the training examples, but still be inaccurate (be outside of our fittings).

With a good dataset of training examples, we can split up the data into two sets: a **training set** and a **test set**.

The new procedure using the new set notation:

1. Learn off-and minimize $J_{\text{train}}(\theta)$ using the training set
2. Compute the test set error $J_{\text{test}}(\theta)$

The test set error:

1. For linear regression $J_{\text{test}}(\theta) = \frac{1}{2m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} (\theta_0 + \theta_1 x_{i,1}^{(2)} - y_{i,1}^{(2)})^2$
2. For classification - Misclassification error (also 1/1 misclassification error)

$$\text{misclassification error} = \frac{1}{m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} \begin{cases} 1 & \text{if } \text{hypo}(x) \neq y \text{ and } y \text{ is linear hypothesis is 0.0 and } y \neq 1 \\ 0 & \text{otherwise} \end{cases}$$

Regression is always for Linear model based on misclassification.

The average test error for the test set is:

$$\text{Test Error} = \frac{1}{m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} \text{err}(\text{hypo}(x_{i,1}^{(2)})) = \frac{1}{m_{\text{test}}} J_{\text{test}}(\theta)$$

Regression is the approach toward the test data that was motivated.

Model Selection and Train/Validation/Test Sets

- Just because a learning algorithm finds training set early, that does not mean it's a good hypothesis.
- The error of your hypothesis is measurement that the set with which you trained the parameters will be lower than any other data set.

In order to choose the model of your hypothesis, you can minimize degree of polynomial and look at the error result.

Without the validation set (note: this is a bad method - it's not used)

1. Optimize the parameters in θ using the training set for each polynomial degree.
2. Find the polynomial degree with the least error using the test set.
3. Estimate the generalization error choosing the test set with $J_{\text{test}}(\theta^{(2)})$ (data from polynomial with least error).

In this case, we have trained on one variable, d , the degree of the polynomial, using the test set. This will choose the error value to be greater for any other set of data.

Use of the CV set

To solve this, we can introduce a third set, the **Cross Validation Set**, to serve as an intermediate set that we can train it with. Then our test set will give us accurate, more optimal error.

Our example splits based on error divided into the three sets:

- Training set: 55%
- Cross validation set: 20%
- Test set: 25%

We can now calculate three separate error values for the three different sets.

With the validation set (note: this method procedure we do not often use the CV set for regression)

1. Optimize the parameters in θ using the training set for each polynomial degree.
2. Find the polynomial degree with the least error using the cross validation set.
3. Estimate the generalization error using the test set with $J_{\text{test}}(\theta^{(2)})$ (data from polynomial with least error).

However, the degree of the polynomial has not been minimizing the test set.

(Note: note for parameter using the "Cross Validation" means that we can choose it for the validation cross process of getting the optimal value).

Diagnosing Bias vs. Variance

In this section we examine the relationship between the degree of the polynomial and the underfitting or overfitting of our hypothesis.

- The reason to distinguish **underfitting** or **variance** is the problem contributing to bad predictions.
- High bias is underfitting and high variance is overfitting. The goal is to find a golden mean between the two.

The training error will tend to **decrease** as we increase the degree d of the polynomial.

As the same time, the cross validation error will **decrease** as we increase d up to a point, and then it will **increase** as we increase, forming a curve shape.

High bias (underfitting): both $J_{\text{train}}(\theta)$ and $J_{\text{CV}}(\theta)$ will be high, also $J_{\text{CV}}(\theta) \approx J_{\text{train}}(\theta)$.

High variance (overfitting): $J_{\text{train}}(\theta)$ will be low and $J_{\text{CV}}(\theta)$ will be much greater than $J_{\text{train}}(\theta)$.

The is represented in the figure below:



Regularization and Bias/Variance

Instead of looking at the degree d contributing to bias/variance, now we will look at the regularization parameter λ .

- Large λ : high bias (underfitting)
- Intermediate λ : just right
- Small λ : high variance (overfitting)

A large lambda automatically penalizes all the θ parameters, which greatly simplifies the resulting function, so reduces underfitting.

The relationship of bias to training and the variance set is as follows:

Low λ : $J_{\text{train}}(\theta)$ is low and $J_{\text{CV}}(\theta)$ is high (high variance/overfitting).

Intermediate λ : $J_{\text{train}}(\theta)$ and $J_{\text{CV}}(\theta)$ are somewhat low and $J_{\text{train}}(\theta) \approx J_{\text{CV}}(\theta)$.

Large λ : both $J_{\text{train}}(\theta)$ and $J_{\text{CV}}(\theta)$ will be high (underfitting, high bias)

The figure below illustrates the relationship between lambda and the hypothesis.



In order to choose the model and the regularization λ , we need:

1. Create a list of lambda's (e.g. 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8.0, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10.0, 10.1, 10.2, 10.3, 10.4, 10.5, 10.6, 10.7, 10.8, 10.9, 11.0, 11.1, 11.2, 11.3, 11.4, 11.5, 11.6, 11.7, 11.8, 11.9, 12.0, 12.1, 12.2, 12.3, 12.4, 12.5, 12.6, 12.7, 12.8, 12.9, 13.0, 13.1, 13.2, 13.3, 13.4, 13.5, 13.6, 13.7, 13.8, 13.9, 14.0, 14.1, 14.2, 14.3, 14.4, 14.5, 14.6, 14.7, 14.8, 14.9, 15.0, 15.1, 15.2, 15.3, 15.4, 15.5, 15.6, 15.7, 15.8, 15.9, 16.0, 16.1, 16.2, 16.3, 16.4, 16.5, 16.6, 16.7, 16.8, 16.9, 17.0, 17.1, 17.2, 17.3, 17.4, 17.5, 17.6, 17.7, 17.8, 17.9, 18.0, 18.1, 18.2, 18.3, 18.4, 18.5, 18.6, 18.7, 18.8, 18.9, 19.0, 19.1, 19.2, 19.3, 19.4, 19.5, 19.6, 19.7, 19.8, 19.9, 20.0, 20.1, 20.2, 20.3, 20.4, 20.5, 20.6, 20.7, 20.8, 20.9, 21.0, 21.1, 21.2, 21.3, 21.4, 21.5, 21.6, 21.7, 21.8, 21.9, 22.0, 22.1, 22.2, 22.3, 22.4, 22.5, 22.6, 22.7, 22.8, 22.9, 23.0, 23.1, 23.2, 23.3, 23.4, 23.5, 23.6, 23.7, 23.8, 23.9, 24.0, 24.1, 24.2, 24.3, 24.4, 24.5, 24.6, 24.7, 24.8, 24.9, 25.0, 25.1, 25.2, 25.3, 25.4, 25.5, 25.6, 25.7, 25.8, 25.9, 26.0, 26.1, 26.2, 26.3, 26.4, 26.5, 26.6, 26.7, 26.8, 26.9, 27.0, 27.1, 27.2, 27.3, 27.4, 27.5, 27.6, 27.7, 27.8, 27.9, 28.0, 28.1, 28.2, 28.3, 28.4, 28.5, 28.6, 28.7, 28.8, 28.9, 29.0, 29.1, 29.2, 29.3, 29.4, 29.5, 29.6, 29.7, 29.8, 29.9, 30.0, 30.1, 30.2, 30.3, 30.4, 30.5, 30.6, 30.7, 30.8, 30.9, 31.0, 31.1, 31.2, 31.3, 31.4, 31.5, 31.6, 31.7, 31.8, 31.9, 32.0, 32.1, 32.2, 32.3, 32.4, 32.5, 32.6, 32.7, 32.8, 32.9, 33.0, 33.1, 33.2, 33.3, 33.4, 33.5, 33.6, 33.7, 33.8, 33.9, 34.0, 34.1, 34.2, 34.3, 34.4, 34.5, 34.6, 34.7, 34.8, 34.9, 35.0, 35.1, 35.2, 35.3, 35.4, 35.5, 35.6, 35.7, 35.8, 35.9, 36.0, 36.1, 36.2, 36.3, 36.4, 36.5, 36.6, 36.7, 36.8, 36.9, 37.0, 37.1, 37.2, 37.3, 37.4, 37.5, 37.6, 37.7, 37.8, 37.9, 38.0, 38.1, 38.2, 38.3, 38.4, 38.5, 38.6, 38.7, 38.8, 38.9, 39.0, 39.1, 39.2, 39.3, 39.4, 39.5, 39.6, 39.7, 39.8, 39.9, 40.0, 40.1, 40.2, 40.3, 40.4, 40.5, 40.6, 40.7, 40.8, 40.9, 41.0, 41.1, 41.2, 41.3, 41.4, 41.5, 41.6, 41.7, 41.8, 41.9, 42.0, 42.1, 42.2, 42.3, 42.4, 42.5, 42.6, 42.7, 42.8, 42.9, 43.0, 43.1, 43.2, 43.3, 43.4, 43.5, 43.6, 43.7, 43.8, 43.9, 44.0, 44.1, 44.2, 44.3, 44.4, 44.5, 44.6, 44.7, 44.8, 44.9, 45.0, 45.1, 45.2, 45.3, 45.4, 45.5, 45.6, 45.7, 45.8, 45.9, 46.0, 46.1, 46.2, 46.3, 46.4, 46.5, 46.6, 46.7, 46.8, 46.9, 47.0, 47.1, 47.2, 47.3, 47.4, 47.5, 47.6, 47.7, 47.8, 47.9, 48.0, 48.1, 48.2, 48.3, 48.4, 48.5, 48.6, 48.7, 48.8, 48.9, 49.0, 49.1, 49.2, 49.3, 49.4, 49.5, 49.6, 49.7, 49.8, 49.9, 50.0, 50.1, 50.2, 50.3, 50.4, 50.5, 50.6, 50.7, 50.8, 50.9, 51.0, 51.1, 51.2, 51.3, 51.4, 51.5, 51.6, 51.7, 51.8, 51.9, 52.0, 52.1, 52.2, 52.3, 52.4, 52.5, 52.6, 52.7, 52.8, 52.9, 53.0, 53.1, 53.2, 53.3, 53.4, 53.5, 53.6, 53.7, 53.8, 53.9, 54.0, 54.1, 54.2, 54.3, 54.4, 54.5, 54.6, 54.7, 54.8, 54.9, 55.0, 55.1, 55.2, 55.3, 55.4, 55.5, 55.6, 55.7, 55.8, 55.9, 56.0, 56.1, 56.2, 56.3, 56.4, 56.5, 56.6, 56.7, 56.8, 56.9, 57.0, 57.1, 57.2, 57.3, 57.4, 57.5, 57.6, 57.7, 57.8, 57.9, 58.0, 58.1, 58.2, 58.3, 58.4, 58.5, 58.6, 58.7, 58.8, 58.9, 59.0, 59.1, 59.2, 59.3, 59.4, 59.5, 59.6, 59.7, 59.8, 59.9, 60.0, 60.1, 60.2, 60.3, 60.4, 60.5, 60.6, 60.7, 60.8, 60.9, 61.0, 61.1, 61.2, 61.3, 61.4, 61.5, 61.6, 61.7, 61.8, 61.9, 62.0, 62.1, 62.2, 62.3, 62.4, 62.5, 62.6, 62.7, 62.8, 62.9, 63.0, 63.1, 63.2, 63.3, 63.4, 63.5, 63.6, 63.7, 63.8, 63.9, 64.0, 64.1, 64.2, 64.3, 64.4, 64.5, 64.6, 64.7, 64.8, 64.9, 65.0, 65.1, 65.2, 65.3, 65.4, 65.5, 65.6, 65.7, 65.8, 65.9, 66.0, 66.1, 66.2, 66.3, 66.4, 66.5, 66.6, 66.7, 66.8, 66.9, 67.0, 67.1, 67.2, 67.3, 67.4, 67.5, 67.6, 67.7, 67.8, 67.9, 68.0, 68.1, 68.2, 68.3, 68.4, 68.5, 68.6, 68.7, 68.8, 68.9, 69.0, 69.1, 69.2, 69.3, 69.4, 69.5, 69.6, 69.7, 69.8, 69.9, 70.0, 70.1, 70.2, 70.3, 70.4, 70.5, 70.6, 70.7, 70.8, 70.9, 71.0, 71.1, 71.2, 71.3, 71.4, 71.5, 71.6, 71.7, 71.8, 71.9, 72.0, 72.1, 72.2, 72.3, 72.4, 72.5, 72.6, 72.7, 72.8, 72.9, 73.0, 73.1, 73.2, 73.3, 73.4, 73.5, 73.6, 73.7, 73.8, 73.9, 74.0, 74.1, 74.2, 74.3, 74.4, 74.5, 74.6, 74.7, 74.8, 74.9, 75.0, 75.1, 75.2, 75.3, 75.4, 75.5, 75.6, 75.7, 75.8, 75.9, 76.0, 76.1, 76.2, 76.3, 76.4, 76.5, 76.6, 76.7, 76.8, 76.9, 77.0, 77.1, 77.2, 77.3, 77.4, 77.5, 77.6, 77.7, 77.8, 77.9, 78.0, 78.1, 78.2, 78.3, 78.4, 78.5, 78.6, 78.7, 78.8, 78.9, 79.0, 79.1, 79.2, 79.3, 79.4, 79.5, 79.6, 79.7, 79.8, 79.9, 80.0, 80.1, 80.2, 80.3, 80.4, 80.5, 80.6, 80.7, 80.8, 80.9, 81.0, 81.1, 81.2, 81.3, 81.4, 81.5, 81.6, 81.7, 81.8, 81.9, 82.0, 82.1, 82.2, 82.3, 82.4, 82.5, 82.6, 82.7, 82.8, 82.9, 83.0, 83.1, 83.2, 83.3, 83.4, 83.5, 83.6, 83.7, 83.8, 83.9, 84.0, 84.1, 84.2, 84.3, 84.4, 84.5, 84.6, 84.7, 84.8, 84.9, 85.0, 85.1, 85.2, 85.3, 85.4, 85.5, 85.6, 85.7, 85.8, 85.9, 86.0, 86.1, 86.2, 86.3, 86.4, 86.5, 86.6, 86.7, 86.8, 86.9, 87.0, 87.1, 87.2, 87.3, 87.4, 87.5, 87.6, 87.7, 87.8, 87.9, 88.0, 88.1, 88.2, 88.3, 88.4, 88.5, 88.6, 88.7, 88.8, 88.9, 89.0, 89.1, 89.2, 89.3, 89.4, 89.5, 89.6, 89.7, 89.8, 89.9, 90.0, 90.1, 90.2, 90.3, 90.4, 90.5, 90.6, 90.7, 90.8, 90.9, 91.0, 91.1, 91.2, 91.3, 91.4, 91.5, 91.6, 91.7, 91.8, 91.9, 92.0, 92.1, 92.2, 92.3, 92.4, 92.5, 92.6, 92.7, 92.8, 92.9, 93.0, 93.1, 93.2, 93.3, 93.4, 93.5, 93.6, 93.7, 93.8, 93.9, 94.0, 94.1, 94.2, 94.3, 94.4, 94.5, 94.6, 94.7, 94.8, 94.9, 95.0, 95.1, 95.2, 95.3, 95.4, 95.5, 95.6, 95.7, 95.8, 95.9, 96.0, 96.1, 96.2, 96.3, 96.4, 96.5, 96.6, 96.7, 96.8, 96.9, 97.0, 97.1, 97.2, 97.3, 97.4, 97.5, 97.6, 97.7, 97.8, 97.9, 98.0, 98.1, 98.2, 98.3, 98.4, 98.5, 98.6, 98.7, 98.8, 98.9, 99.0, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, 100.0, 100.1, 100.2, 100.3, 100.4, 100.5, 100.6, 100.7, 100.8, 100.9, 101.0, 101.1, 101.2, 101.3, 101.4, 101.5, 101.6, 101.7, 101.8, 101.9, 102.0, 102.1, 102.2, 102.3, 102.4, 102.5, 102.6, 102.7, 102.8, 102.9, 103.0, 103.1, 103.2, 103.3, 103.4, 103.5, 103.6, 103.7, 103.8, 103.9, 104.0, 104.1, 104.2, 104.3, 104.4, 104.5, 104.6, 104.7, 104.8, 104.9, 105.0, 105.1, 105.2, 105.3, 105.4, 105.5, 105.6, 105.7, 105.8, 105.9, 106.0, 106.1, 106.2, 106.3, 106.4, 106.5, 106.6, 106.7, 106.8, 106.9, 107.0, 107.1, 107.2, 107.3, 107.4, 107.5, 107.6, 107.7, 107.8, 107.9, 108.0, 108.1, 108.2, 108.3, 108.4, 108.5, 108.6, 108.7, 108.8, 108.9, 109.0, 109.1, 109.2, 109.3, 109.4, 109.5, 109.6, 109.7, 109.8, 109.9, 110.0, 110.1, 110.2, 110.3, 110.4, 110.5, 110.6, 110.7, 110.8, 110.9, 111.0, 111.1, 111.2, 111.3, 111.4, 111.5, 111.6, 111.7, 111.8, 111.9, 112.0, 112.1, 112.2, 112.3, 112.4, 112.5, 112.6, 112.7, 112.8, 112.9, 113.0, 113.1, 113.2, 113.3, 113.4, 113.5, 113.6, 113.7, 113.8, 113.9, 114.0, 114.1, 114.2, 114.3, 114.4, 114.5, 114.6, 114.7, 114.8, 114.9, 115.0, 115.1, 115.2, 115.3, 115.4, 115.5, 115.6, 115.7, 115.8, 115.9, 116.0, 116.1, 116.2, 116.3, 116.4, 116.5, 116.6, 116.7, 116.8, 116.9, 117.0, 117.1, 117.2, 117.3, 117.4, 117.5, 117.6, 117.7, 117.8, 117.9, 118.0, 118.1, 118.2, 118.3, 118.4, 118.5, 118.6, 118.7, 118.8, 118.9, 119.0, 119.1, 119.2, 119.3, 119.4, 119.5, 119.6, 119.7, 119.8, 119.9, 120.0, 120.1, 120.2, 120.3, 120.4, 120.5, 120.6, 120.7, 120.8, 120.9, 121.0, 121.1, 121.2, 121.3, 121.4, 121.5, 121.6, 121.7, 121.8, 121.9, 122.0, 122.1, 122.2, 122.3, 122.4, 122.5, 122.6, 122.7, 122.8, 122.9, 123.0, 123.1, 123.2, 123.3, 123.4, 123.5, 123.6, 123.7, 123.8, 123.9, 124.0, 124.1, 124.2, 124.3, 124.4, 124.5, 124.6, 124.7, 124.8, 124.9, 125.0, 125.1, 125.2, 125.3, 125.4, 125.5, 125.6, 125.7, 125.8, 125.9, 126.0, 126.1, 126.2, 126.3, 126.4, 126.5, 126.6, 126.7, 126.8, 126.9, 127.0, 127.1, 127.2, 127.3, 127.4, 127.5, 127.6, 127.7, 127.8, 127.9, 128.0, 128.1, 128.2, 128.3, 128.4, 128.5, 128.6, 128.7, 128.8, 128.9, 129.0, 129.1, 129.2, 129.3, 129.4, 129.5, 129.6, 129.7, 129.8, 129.9, 130.0, 130.1, 130.2, 130.3, 130.4, 130.5, 130.6, 130.7, 130.8, 130.9, 131.0, 131.1, 131.2, 131.3, 131.4, 131.5, 131.6, 131.7, 131.8, 131.9, 132.0, 132.1, 132.2, 132.3, 132.4, 132.5, 132.6, 132.7, 132.8, 132.9, 133.0, 133.1, 133.2, 133.3, 133.4, 133.5, 133.6, 133.7, 133.8, 133.9, 134.0, 134.1, 134.2, 134.3, 134.4, 134.5