Polynomial Regression: Bias-Variance Trade-off

March 21, 2023

1 Overview

This repository contains code for exploring the bias-variance trade-off in polynomial regression. The goal is to understand how the bias and variance of a model change as the degree of the polynomial is increased, and how this affects the overall performance of the model.

2 Requirements

To run the code, you will need the following Python packages:

- NumPy
- Matplotlib
- scikit-learn

You can install these packages using pip:

```
pip install numpy matplotlib scikit-learn
```

3 Usage

To run the code, simply clone this repository and run the following command:

```
python polynomial_regression.py
```

This will generate a plot showing the bias, variance, training error, and test error for polynomial models of different degrees. You can modify the range of degrees to try and the size of the test set in the code.

4 Results

The plot shows that as the degree of the polynomial is increased, the bias decreases and the variance increases. This is because the model becomes more flexible and can fit the training data more closely, but also becomes more sensitive to noise and overfits the training data.

As a result, there is a trade-off between bias and variance, and the overall error of the model depends on finding the right balance between the two. This can be achieved by choosing a model with an appropriate degree of complexity that fits the data well without overfitting.

Polynomial regression is a popular method for modeling the relationship between a dependent variable and one or more independent variables. It involves fitting a polynomial function to a set of input-output data pairs. However, choosing the right degree of polynomial is crucial, as selecting a degree that is too high can lead to overfitting, while a degree that is too low can lead to underfitting. This is where the concept of bias-variance trade-off comes in.

5 Bias-Variance Trade-off

The bias-variance trade-off is a fundamental concept in machine learning that aims to find the right balance between two types of errors that can occur in a model: bias and variance. Bias is the error that arises from approximating a real problem with a simpler model. On the other hand, variance is the error that arises from sensitivity to small fluctuations in the training set.

In polynomial regression, bias can be thought of as how well the model captures the underlying relationship between the input and output variables. In other words, bias measures how much the model's predictions deviate from the true relationship between the variables. A high bias means that the model is too simple and cannot capture the complexity of the data, leading to underfitting.

Variance, on the other hand, measures how much the model's predictions vary depending on the training data. A high variance means that the model is too sensitive to the training data and does not generalize well to new, unseen data, leading to overfitting.

The goal of the bias-variance trade-off is to find a model that achieves low bias and low variance, while maintaining reasonable overall error. The trade-off is illustrated by the following formula for the mean squared error (MSE) of a model:

$$\text{MSE} = \text{Bias}^2 + \text{Variance} + \text{Irreducible Error} \tag{1}$$

The MSE measures the expected squared difference between the predictions of the model and the true values of the output variable. The irreducible error represents the noise in the data that cannot be explained by any model.

5.1 Derivation of Bias-Variance Tradeoff

Let D be a dataset consisting of input-output pairs (x, y), where x is an input feature vector and y is the corresponding target variable. Our goal is to learn a function $f(x)$ that approximates the true function $y = g(x)$ that maps inputs to outputs. We can think of $f(x)$ as a machine learning model that we train on the dataset D.

The expected prediction error (EPE) of $f(x)$ on a new input x' can be decomposed into three components: the squared bias, the variance, and the irreducible error: Suppose we have a set of data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, where each x_i is a vector of features and y_i is the corresponding target value. We want to learn a function f that can predict y given new input x. We assume that the relationship between x and y can be modeled by a probabilistic model, such as a linear regression model, and that the true relationship is given by $y = g(x) + \epsilon$, where $g(x)$ is the true function that maps x to y and ϵ is the irreducible error, which represents the noise in the data that cannot be explained by the model.

We can use a learning algorithm to learn a function $\hat{f}$ that approximates the true function g. However, because the training data is not a perfect representation of the true relationship between x and y, the learned function will not be an exact match for the true function. Instead, there will be some error between the predicted values $\hat{f}(x_i)$ and the true values y_i. We can measure the quality of the learned function using the expected prediction error (EPE):

$$EPE = E[(y - \hat{f}(x))^2],$$

where the expectation is taken over all possible training sets that we might encounter in the future.

We can decompose the EPE into three components: bias, variance, and irreducible error. The bias of an estimator is the difference between the expected value of the estimator and the true value of the function being estimated. In the context of supervised learning, bias measures how much the learned function differs from the true function, on average, over all possible training sets. The variance of an estimator measures how much the estimator varies as a function of different training sets. In the context of supervised learning, variance measures how much the learned function would vary if we trained it on different training sets. The irreducible error represents the noise in the data that cannot be explained by the model.

Using these definitions, we can derive the bias-variance tradeoff as follows:

$$EPE = E[(y - \hat{f}(x))^2] \quad = E[(g(x) + \epsilon - \hat{f}(x))^2] \ = E[(g(x) - E[\hat{f}(x)])^2] + E[(\epsilon)^2] + E[(\hat{f}(x) - E[\hat{f}(x)])^2]$$

The first line is simply the definition of the EPE. The second line substitutes $y = g(x) + \epsilon$ for the true values y. The third line expands the squared term and separates out the three components of the EPE. The fourth line uses the law of total expectation to decompose the covariance term into bias and variance components. The bias is the square of the difference between the true function and the expected value of the learned function, while the variance is the expected value of the squared difference between

where

- $\text{Bias}^2 = E[(g(x') - E[f(x')])^2]$ is the squared bias, which measures the difference between the expected prediction of $f(x)$ and the true value of $g(x)$.
- $\text{Variance} = E[(E[f(x')] - f(x'))^2]$ is the variance, which measures the amount that $f(x)$'s predictions vary for different training sets.
- $\text{Irreducible Error} = E[\epsilon^2]$ is the irreducible error, which represents the noise inherent in the data.

6 Additional information

6.1 Mean Squared Error and Variance

Before delving further into the bias-variance trade-off, let's first define two important concepts: mean squared error (MSE) and variance.

6.2 Mean Squared Error (MSE)

The mean squared error is a measure of the average squared difference between the predicted values of a model and the actual values. The formula for MSE is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2 \tag{2}$$

where n is the number of samples, y_i is the actual value for the i-th sample, and $\hat{y_i}$ is the predicted value for the i-th sample.

6.3 Variance

Variance is a measure of how spread out a set of data is. It is calculated as the average of the squared differences from the mean. The formula for variance is:

$$\text{Var}(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{3}$$

where n is the number of samples, X_i is the i-th value in the data set, and $\bar{X}$ is the mean of the data set.

In the context of the bias-variance trade-off, the MSE can be decomposed into three components: the squared bias, the variance, and an irreducible error term.

The squared bias measures how well the model fits the true relationship between the inputs and outputs, while the variance measures how much the model's predictions vary depending on the specific training data used.

The irreducible error term represents the inherent noise or variability in the data that cannot be reduced by any model.

The bias-variance trade-off is the problem of finding a model that balances these two sources of error to achieve low overall MSE on new, unseen data [1] [2].

References

[1] Andriy Burkov. *The Hundred-Page Machine Learning Book*.

[2] Trevor Hastie et al. *An Introduction to Statistical Learning*.