Sample title

Brandon Hosley

University of Illinois - Springfield

September 3, 2020

Overview

1. Q1: Bias–variance tradeoff
2. Q2: Hastie and Tibshirani

Bias-Variance Trade-Off

Bias Error

- Also called 'Overfitting'

Variance

Bias-Variance Trade-Off

Bias Error

- Also called 'Overfitting'

Variance

- Also called 'Underfitting'

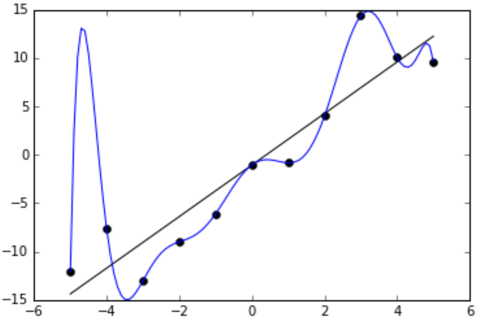
Bias-Variance Trade-Off

Bias Error

- Also called 'Overfitting'

Variance

- Also called 'Underfitting'



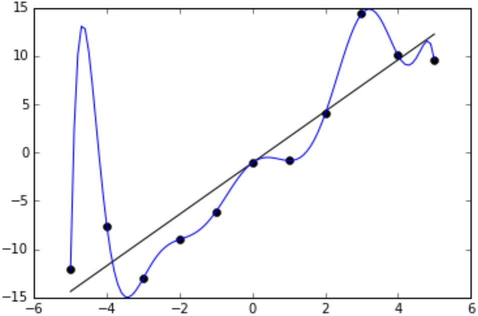
Bias-Variance Trade-Off

Bias Error

- Also called 'Overfitting'
- Predicts test data too well

Variance

- Also called 'Underfitting'



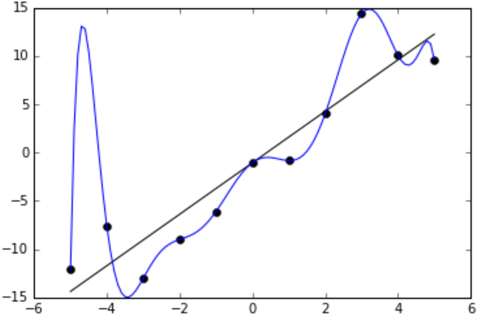
Bias-Variance Trade-Off

Bias Error

- Also called 'Overfitting'
- Predicts test data too well

Variance

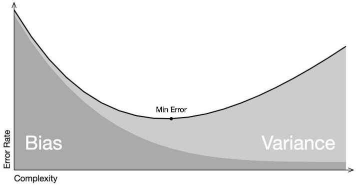
- Also called 'Underfitting'
- Generalizes too well



Bias-Variance Trade-Off

- Aiming for the lowest total error typically means finding a "middle-ground"
- A common technique for this is determining the minimum *mean squared error.*

$$\mathrm{MSE} = \left(E\left[\hat{f}(x)\right] - f(x)\right)^2 + E\left[\left(\hat{f}(x) - E\left[\hat{f}(x)\right]\right)^2\right] + \sigma_e^2$$
$$= \mathrm{Bias}^2 + \mathrm{Variance} + \text{Irreducible Error}$$



Hastie Lectures

- Statistical Learning and Regression
- Dimensionality and Parametric Models
- Assessing Model Accuracy and Bias-Variance Trade-Off
- Classification Problems and K-Nearest Neighbors

Statistical Learning and Regression

- Predicting values through various *means*

Statistical Learning and Regression

- Predicting values through various *means*
- The simplest method of regression may be to use:

$$\hat{f}(x_n) = \bar{x}_n$$

Statistical Learning and Regression

- Predicting values through various *means*
- The simplest method of regression may be to use:

$$\hat{f}(x_n) = \bar{x}_n$$

- This type of model may be improved by extending the values contributing to the average to:

$$\hat{f}(x_n) = \bar{x}_{n \pm i}$$

Statistical Learning and Regression

- Predicting values through various *means*
- The simplest method of regression may be to use:

$$\hat{f}(x_n) = \bar{x}_n$$

- This type of model may be improved by extending the values contributing to the average to:

$$\hat{f}(x_n) = \bar{x}_{n \pm i}$$

- Or the range may be adjusted such that it includes the k-nearest neighbors.

Statistical Learning and Regression

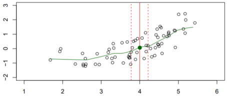
- Predicting values through various *means*
- The simplest method of regression may be to use:

$$\hat{f}(x_n) = \bar{x}_n$$

- This type of model may be improved by extending the values contributing to the average to:

$$\hat{f}(x_n) = \bar{x}_{n\pm i}$$

- Or the range may be adjusted such that it includes the k-nearest neighbors.

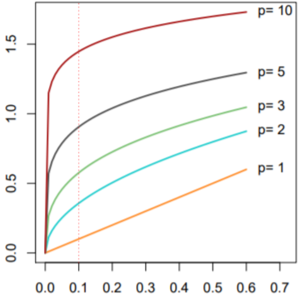


Dimensionality and Parametric Models

- Approximate proportion of a data's range needed to capture a certain fraction of the data scales very quickly with added dimensions

Dimensionality and Parametric Models

- Approximate proportion of a data's range needed to capture a certain fraction of the data scales very quickly with added dimensions



Dimensionality and Parametric Models

- Approximate proportion of a data's range needed to capture a certain fraction of the data scales very quickly with added dimensions
- Structured models may help address these scaling problems

Dimensionality and Parametric Models

- Approximate proportion of a data's range needed to capture a certain fraction of the data scales very quickly with added dimensions
- Structured models may help address these scaling problems

  The simplest structured model is a linear model:

$$f_L(X) = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n$$

Dimensionality and Parametric Models

- Approximate proportion of a data's range needed to capture a certain fraction of the data scales very quickly with added dimensions
- Structured models may help address these scaling problems

  The simplest structured model is a linear model:

  $$f_L(X) = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n$$

- Structured models allowed allow flexible adjustment to obtain preferred level of fitting

Assessing Model Accuracy and Bias-Variance Trade-Off

Classification Problems and K-Nearest Neighbors