The Bias-Variance Tradeoff

METIS

METIS

Learning Goals

- Understand and correctly describe the bias-variance tradeoff, and its relationship with over/underfitting
- Precisely identify and differentiate the 3 sources of predictive model error
- Accurately describe model adjustments that impact complexity such as varying polynomial feature degrees

3 Sources of Model Error

Being wrong

Bias

Being unstable

Variance

Unavoidable randomness

Irreducible Error

3 Sources of Model Error in Detail

Bias

- Tendency of predictions to miss true values
- Worsened by missing information and simplifying assumptions about feature-target relationships
- Miss real patterns (underfit)

Variance

- Tendency of predictions to fluctuate
- Worsened by model's sensitivity to small changes in training data, often due to overly complex models
- Overfit to noise patterns

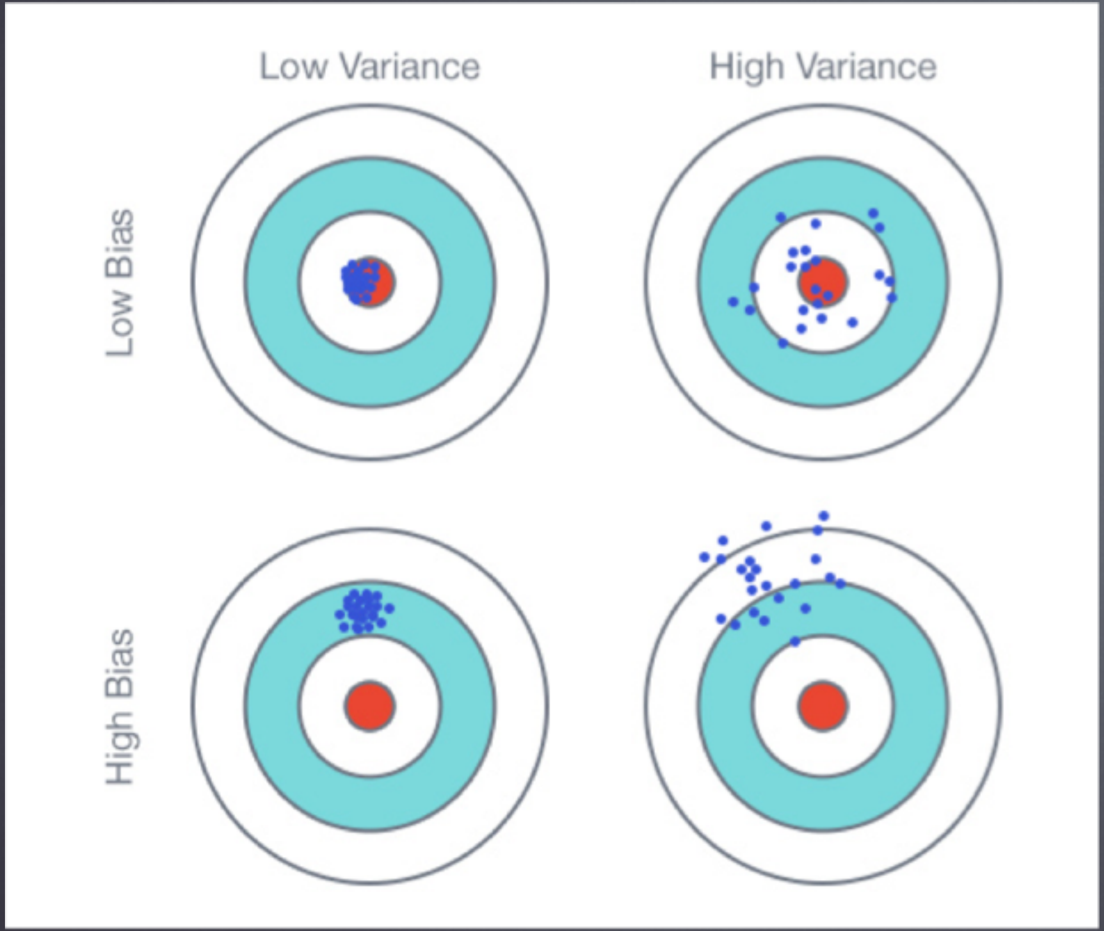
Irreducible Error

- Intrinsic uncertainty/randomness
- Present in even the best possible model

Tendency = expectation of out-of-sample behavior over many training set samples

Bias & Variance Intuition

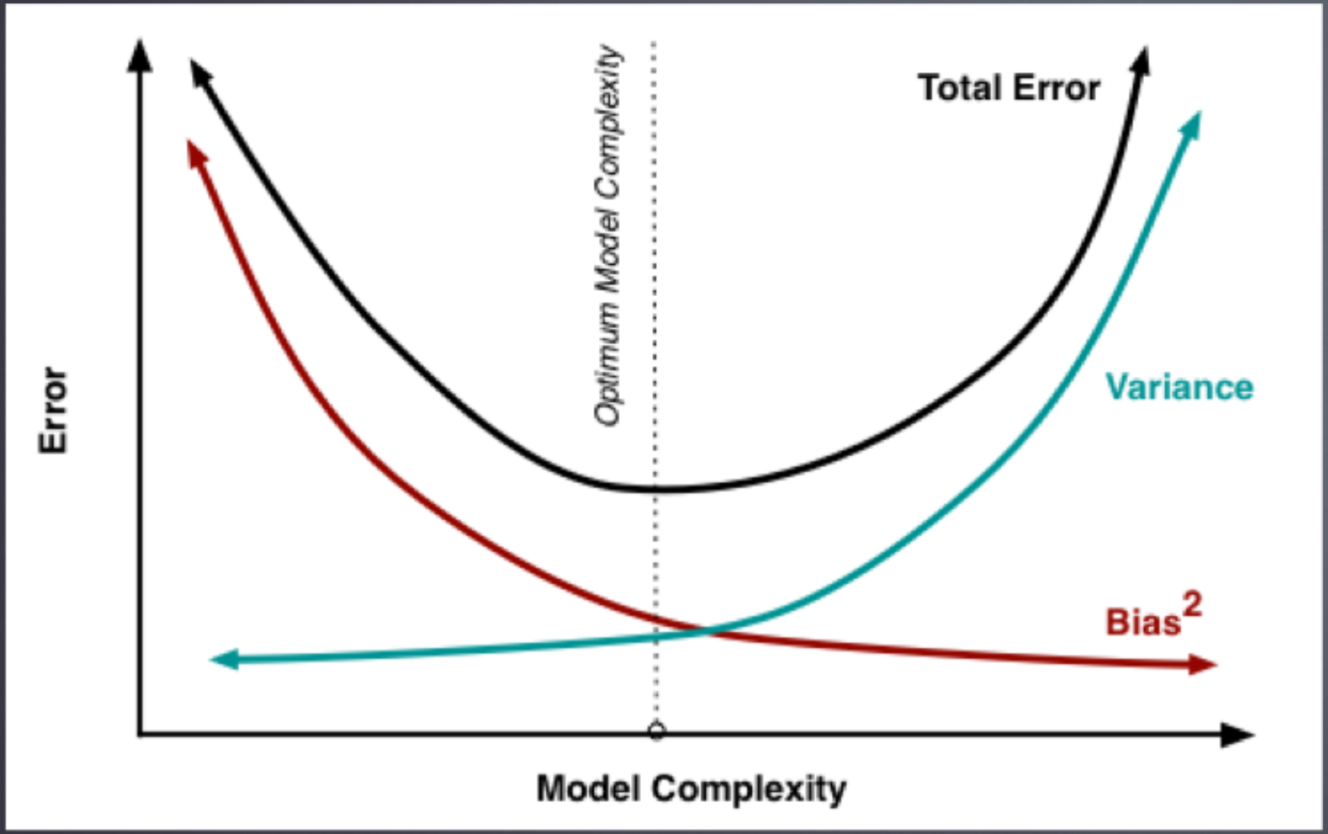
Bias & variance – at an intuitive level



- **Note:** this visual is for intuition building and shouldn't be interpreted literally!
- Roughly, bias is tendency to miss, while variance is tendency to be inconsistent
- Ideally we get the top left outcome: highly consistent predictions that are close to perfect on average

Bias-Variance Tradeoff, Visualized

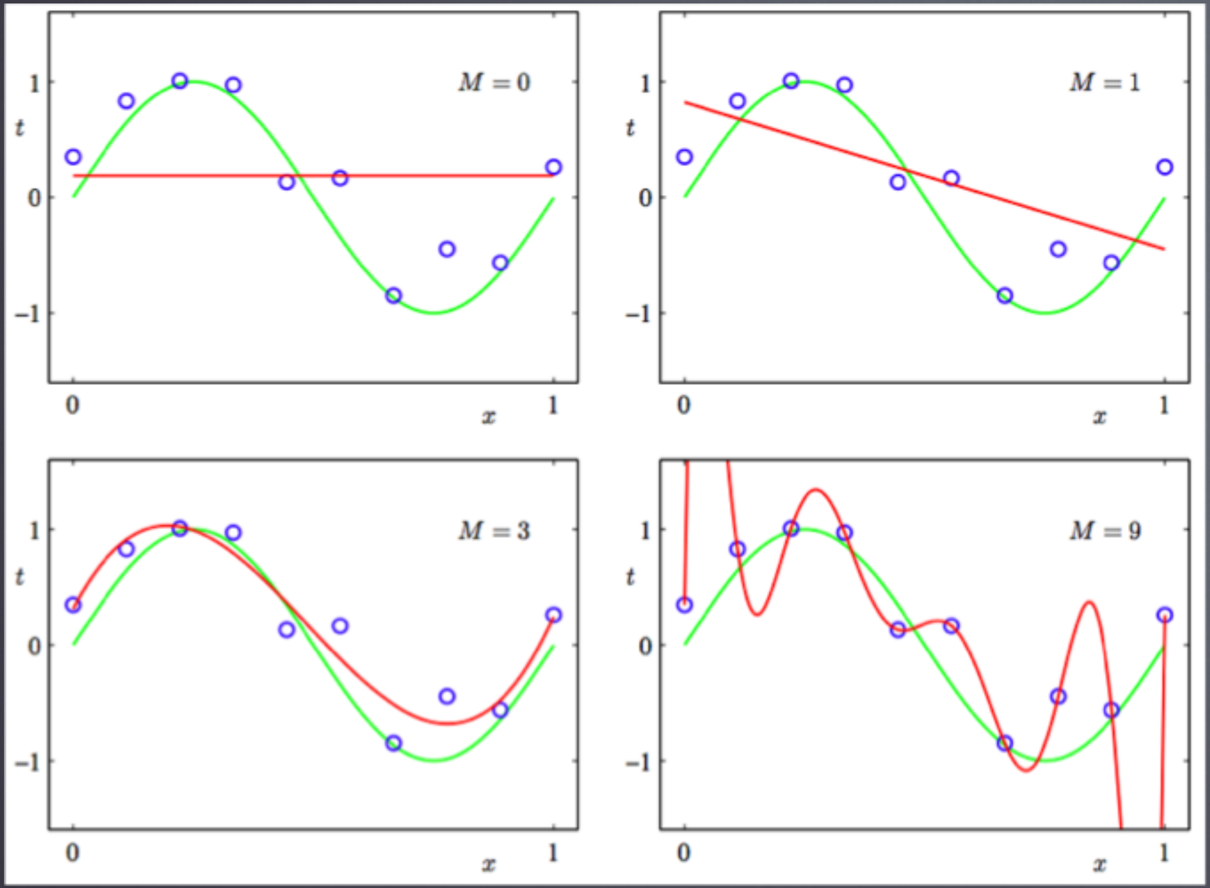
Visualizing the complexity tradeoff



- Usually we analogize the bias-variance tradeoff to a *complexity tradeoff*
- Model adjustments that decrease bias often increase variance, and vice versa
- Finding an optimally predictive model is essentially an exercise in finding the right balance of complexity
- We search for a model that is elaborate enough to describe the feature-target relationship (not underfit), but not so elaborate that it fits to spurious patterns in the training data (not overfit)

Bias-Variance Tradeoff: Example

Complexity tradeoff: polynomial regression



- The higher the degree of a polynomial regression, the more complex the model (lower bias, higher variance)
- At degrees 0 and 1, we can see *visual signs of bias*: the predictions are too rigid to capture the curve pattern in the data
- At degree 9, we can see *visual signs of variance*: the predictions fluctuate wildly because of the model's sensitivity
- Degree 3 is *just right*: the model has sufficient complexity to describe the data without overfitting to noise

Lesson Recap

- Optimizing predictive models is all about finding the right bias/variance i.e. complexity tradeoff
- We need models that are sufficiently complex to capture patterns in data, but not so complex that they overfit to noise

Thank You!

Image Citations

- Slides 6-7: Scott Fortmann-Roe
- Slide 8: Justin Domke