Bias-Variance Tradeoff

Shaobo Han

Duke University

Graphical Illustration

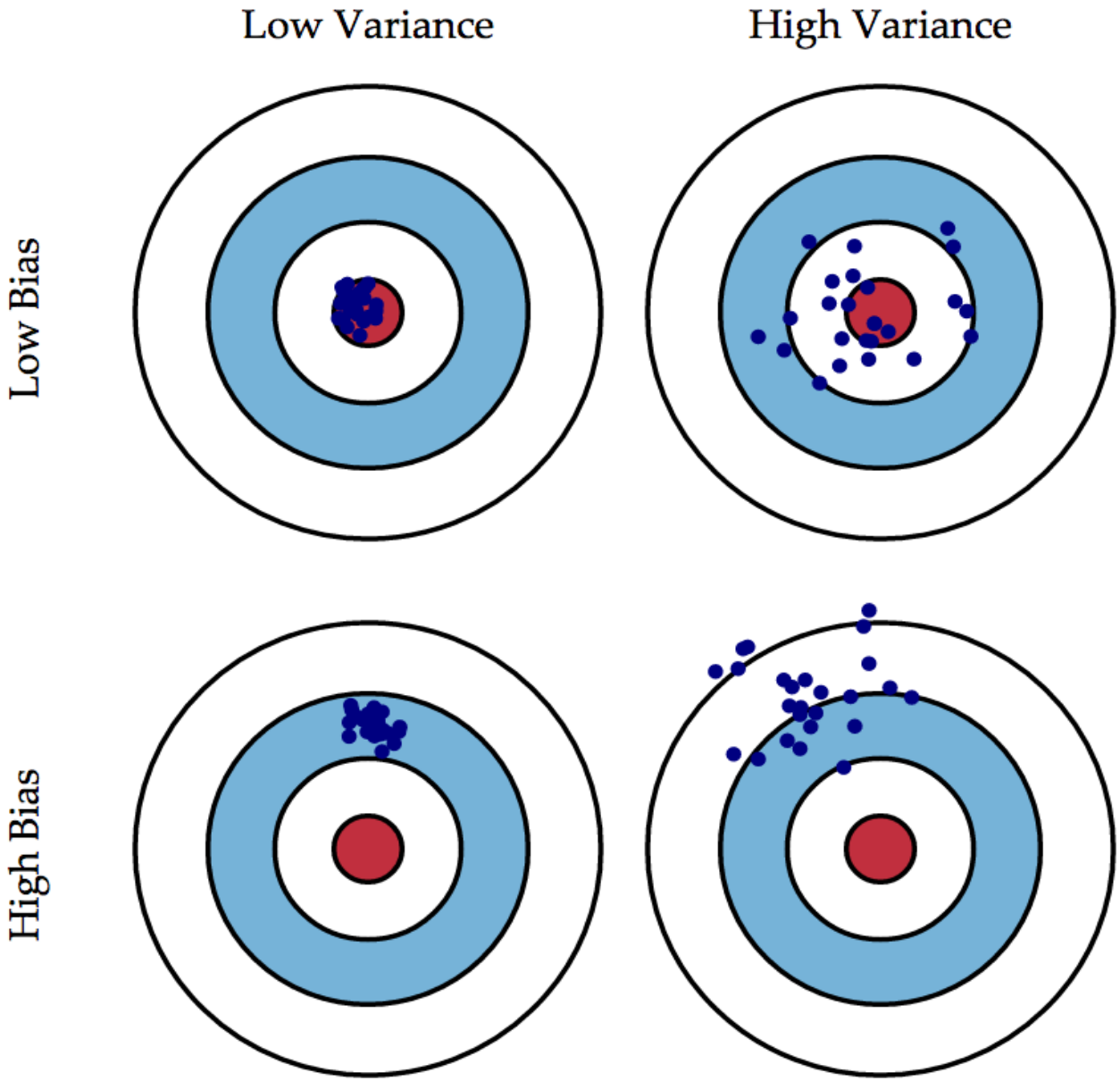

Figure 1 : Dart Game: 4 players taking multiple shots

Intuitions

In a world with imperfect models and finite data, the performance of players (models, algorithms, or estimators) can be measured by **prediction errors**, which decompose into bias and variance terms.

- **Bias** is due to **erroneous model assumptions**
- **Variance** is from **the variability of data gathered and model**

Example: 1-Nearest Neighbor Estimator (1/2)

Task: Predict y_0 at the test-point $x_0 = 0$

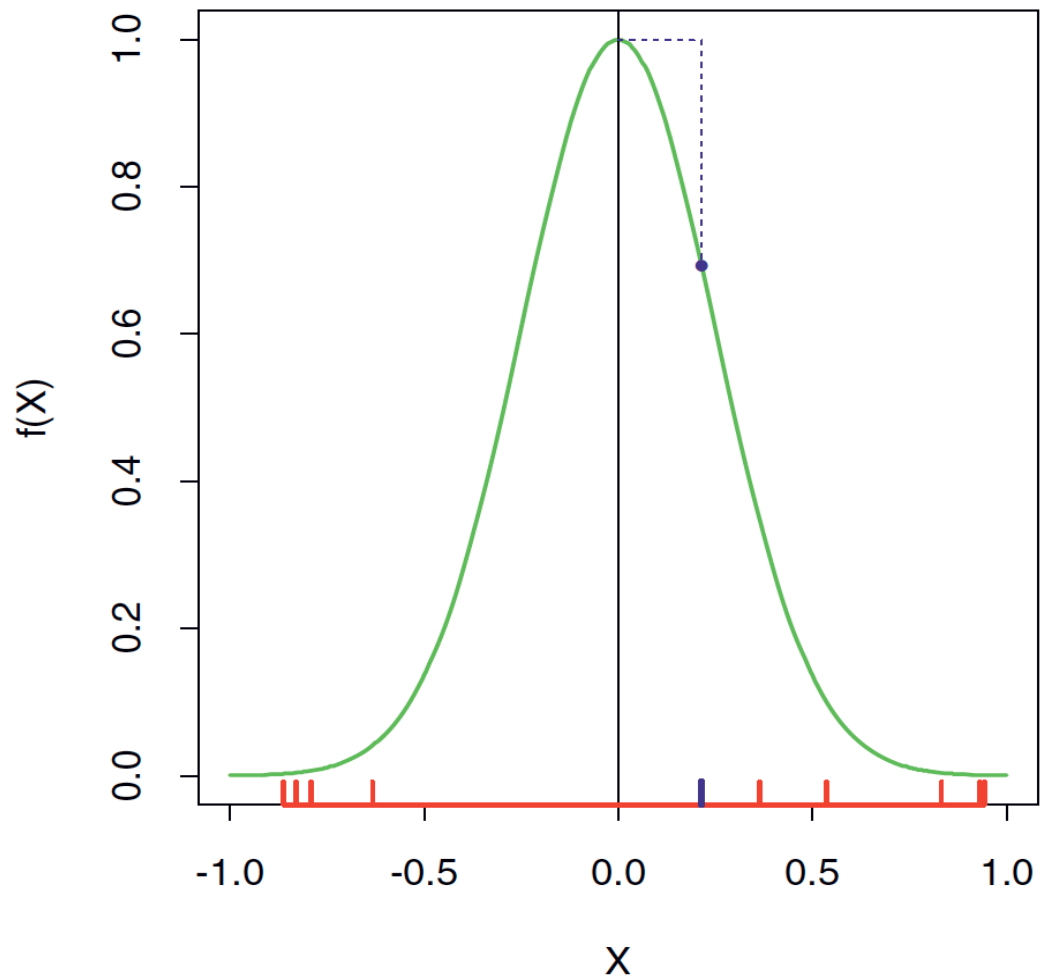

Figure 2 : Simulation example, $x \in [-1, 1]^n$, target function (no noise): $Y = f(X) \equiv \exp(-8||x||^2)$

Example: Mathematical Derivations (2/2)

Mean squared error (MSE) for estimating $f(0)$:

$$\begin{aligned}\mathrm{MSE}(x_0) &= \mathbb{E}[(f(x_0) - \hat{y}_0)^2] \\ &= \mathbb{E}[(\hat{y}_0 - \mathbb{E}(\hat{y}_0))^2] + [\mathbb{E}(\hat{y}_0 - f(x_0))]^2 \\ &= \mathrm{Var}(\hat{y}_0) + \mathrm{Bias}^2(\hat{y}_0) \end{aligned} \tag{1}$$

This is called **Bias-Variance Decomposition**.

Bias-Variance Tradeoff (1/2): How to balance?

Thinkings:

1. Bias indicates something basically wrong

Bias-Variance Tradeoff (1/2): How to balance?

Thinkings:

1. Bias indicates something basically wrong
2. Variance is also bad

Bias-Variance Tradeoff (1/2): How to balance?

Thinkings:

1. Bias indicates something basically wrong
2. Variance is also bad
3. But, a model with low bias and high variance could predict well on average, at least, it is not fundamentally wrong

Bias-Variance Tradeoff (1/2): How to balance?

Thinkings:

1. Bias indicates something basically wrong
2. Variance is also bad
3. But, a model with low bias and high variance could predict well on average, at least, it is not fundamentally wrong

Question: Can we minimize bias even at the expense of variance?

Bias-Variance Tradeoff (1/2): How to balance?

Thinkings:

1. Bias indicates something basically wrong
2. Variance is also bad
3. But, a model with low bias and high variance could predict well on average, at least, it is not fundamentally wrong

Question: Can we minimize bias even at the expense of variance?

Warning: We only get one life (one shot)!

- Long run averages are not available in practice

Bias-Variance Tradeoff (2/2): Model Complexity

Picking the right model complexity

- Simple model does not fit the data (underfitting)
- Complex model are flexible but sensitive (overfitting)

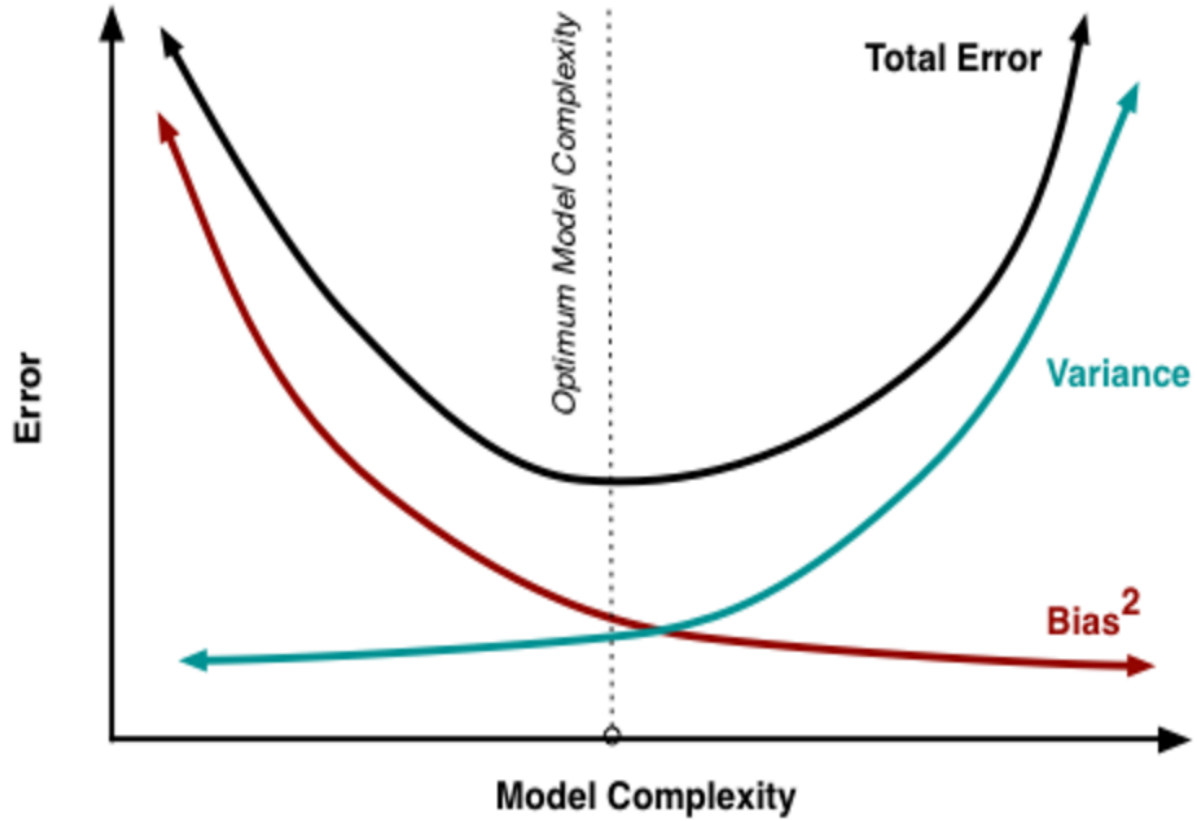

Figure 3 : Bias, variance and total error as a function of model complexity

More Readings

References:

1. Understanding the Bias-Variance Tradeoff, `http://scott.fortmann-roe.com/docs/BiasVariance.html`
2. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. (Second Edition)

Image credits: Figure 1 & 3 (reference 1). Figure 2 (reference 2)