Training/Testing and Regularization

Il-Chul Moon
Dept. of Industrial and Systems Engineering
KAIST
icmoon@kaist.ac.kr

Copyright © 2010 by Il-Chul Moon, Dept. of Industrial and Systems Engineering, KAIST

Weekly Objectives

- Understand the concept of bias and variance
  - Know the concept of over-fitting and under-fitting
  - Able to segment two sources, bias and variance, of error
- Understand the bias and variance trade-off
  - Understand the concept of Occam's razor
  - Able to perform cross-validation
  - Know various performance metrics for supervised machine learning
- Understand the concept of regularization
  - Know how to apply regularization to
    - Linear regression
    - Logistic regression
    - Support vector machine

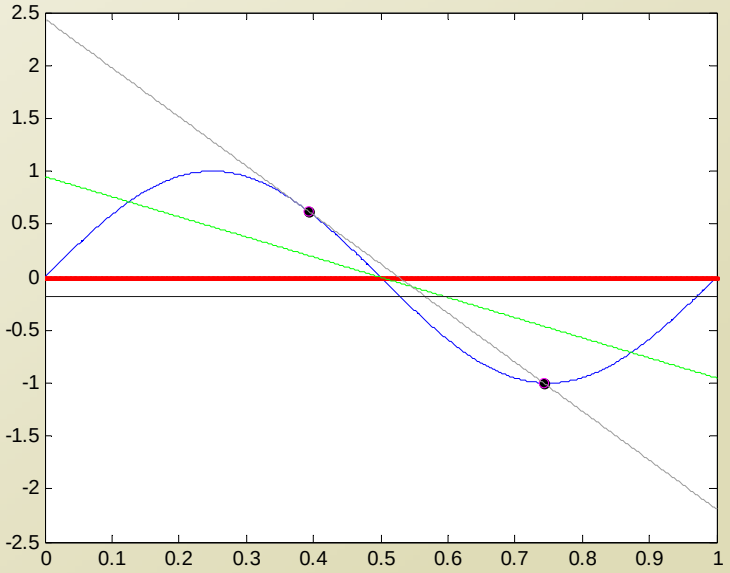
KAIST Copyright © 2010 by Il-Chul Moon, Dept. of Industrial and Systems Engineering, KAIST

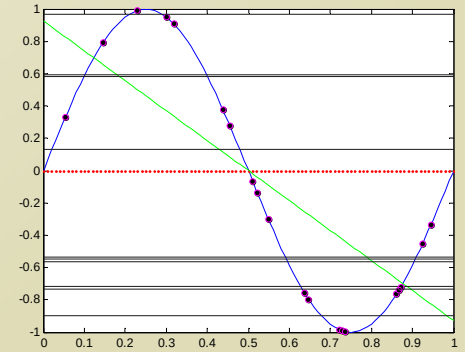
PERFORMANCE MEASUREMENT

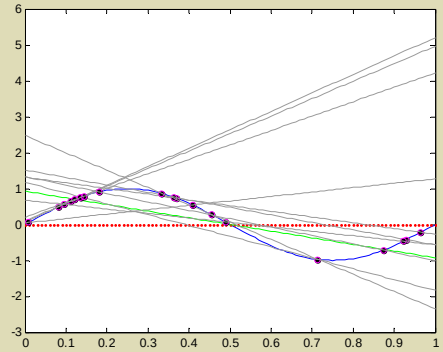
Copyright © 2010 by Il-Chul Moon, Dept. of Industrial and Systems Engineering, KAIST

Empirical Bias and Variance Trade-off

- Consider
  - f(x)=sin(2*pi*x)
  - D={two points|point=(x,sin(2*pi*x)), 0<=x<=1)
  - Two g(x)
    - Zero degree: dark grey line
    - One degree: light grey line
  - Two $\bar{g}(x)$
    - Zero degree: red line
    - One degree: green line
- Which has a greater bias and a greater variance between one degree and zero degree?

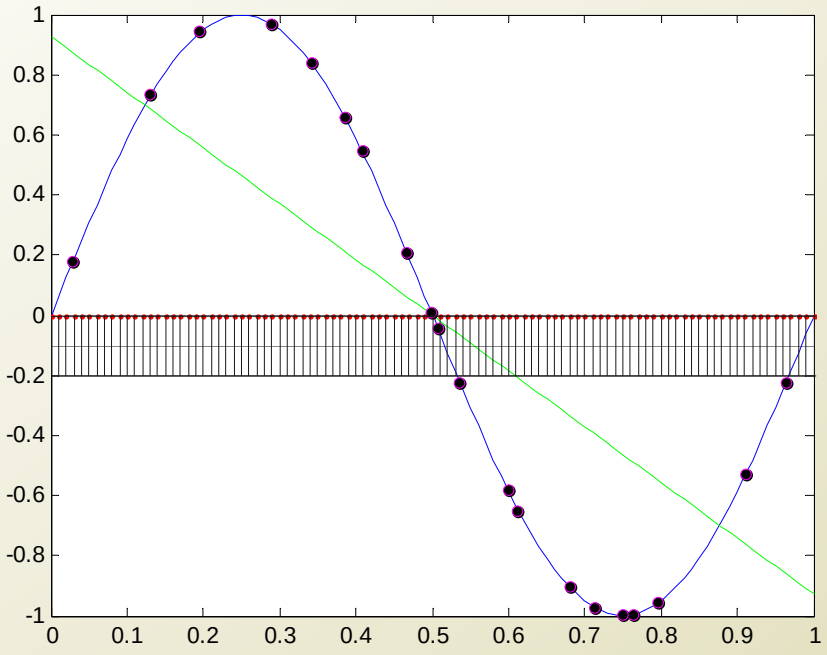




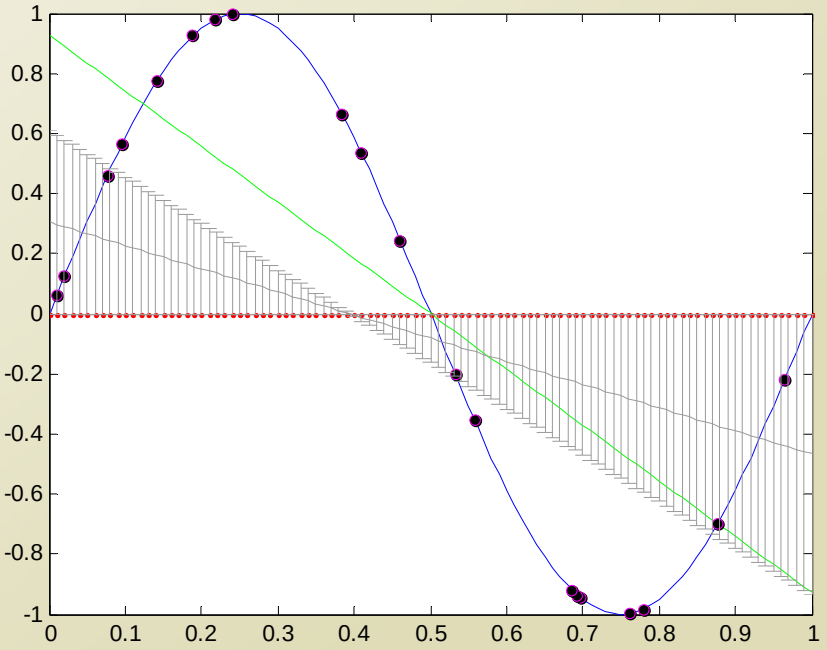


KAIST
Copyright © 2010 by Il-Chul Moon, Dept. of Industrial and Systems Engineering, KAIST

Bias and Variance of Two Hypotheses



Bias = 0.5051
Var. = 0.2410



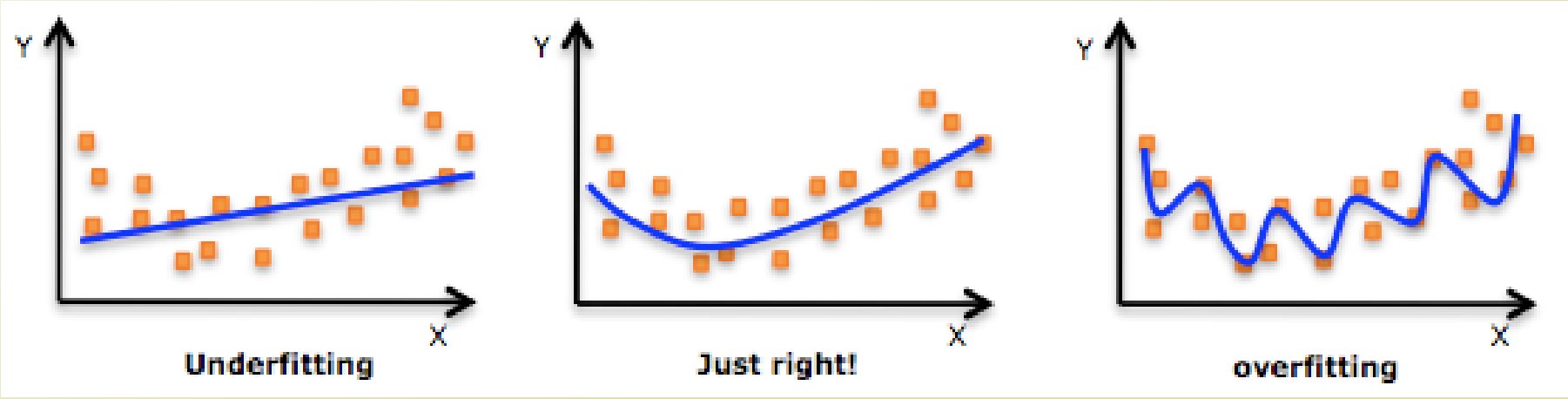
Bias = 0.3092
Var. = 2.0708

- A complex model has a higher variance and a lower bias.
- A simple model has a lower variance and a higher bias.
- Need a balance in the complexity of a ML algorithm

Copyright © 2010 by Il-Chul Moon, Dept. of Industrial and Systems Engineering, KAIST

Occam's Razor

- Occam's Razor
  - Among competing hypotheses, the one which makes the fewest assumption should be selected
- Competing?
  - Relevantly similar error in the prediction
- Fewest assumption
  - Less complex model
- Given the approximately same error, a simple model should be selected
- Reflection of Bias and Variance tradeoff!
  - By the way, is it possible to calculate the bias and the variance in the real world setting?



Copyright © 2010 by Il-Chul Moon, Dept. of Industrial and Systems Engineering, KAIST