Ch 2.2: Assessing Model Accuracy
Lecture 3 - CMSE 381

Prof. Guanqun Cao

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Friday, Aug 29, 2025

Announcements

Last time:
- Ch 2.1, Vocab day!

Announcements:
- Get on slack!
  - +1 point on the first homework if you post a gif in the thread
- First homework due Sunday, 9/7.
- Office hours: see website

Covered in this lecture

- Mean Squared Error (regression)
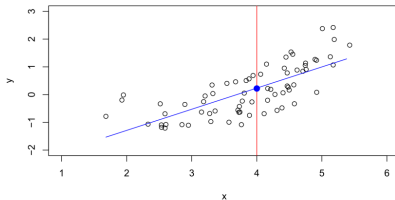- Train vs Test
- Bias Variance Trade off

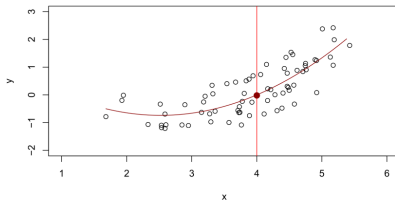# Quick review of notation

Section 1

Mean Squared Error

Which is better?

A linear model $\hat{f}_L(X) = \hat{\beta}_0 + \hat{\beta}_1 X$ gives a reasonable fit here

A quadratic model $\hat{f}_Q(X) = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{\beta}_2 X^2$ fits slightly better.
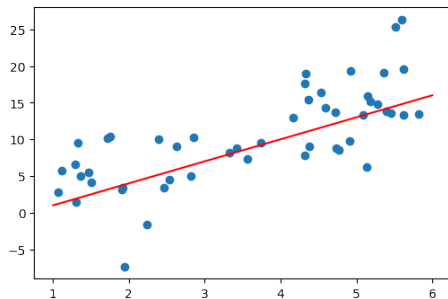
No free lunch

Mean Squared Error

Error in the regression setting

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

Group Work

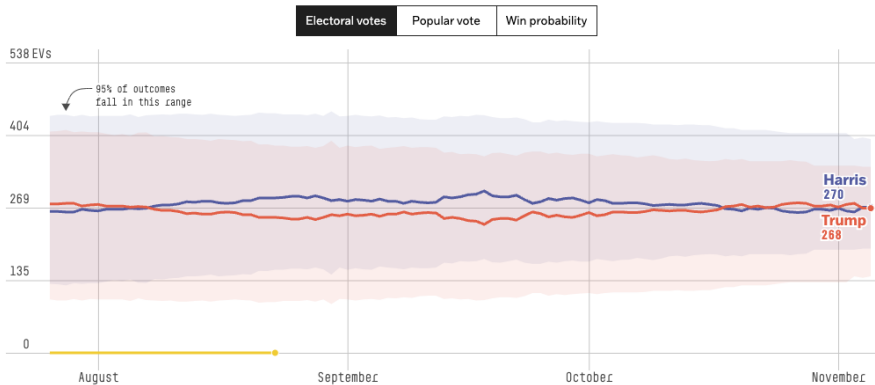Given the following data, you decide to use the model

$$\hat{f}(X_1, X_2) = 1 - 3X_1 + 2X_2.$$

What is the MSE?

X_1	X_2	Y
0	7	14
1	-3	-6
5	2	-10
-1	1	7

Training MSE

How has the forecast changed over time?

The forecast updates at least once a day and whenever we get new data. Uncertainty will decrease as we get closer to Election Day.

| Electoral votes | Popular vote | Win probability |

Train vs test

Training set:
The observations
$\{(x_1, y_1), \cdots, (x_n, y_n)\}$ used to get the estimate $\hat{f}$

Test set:
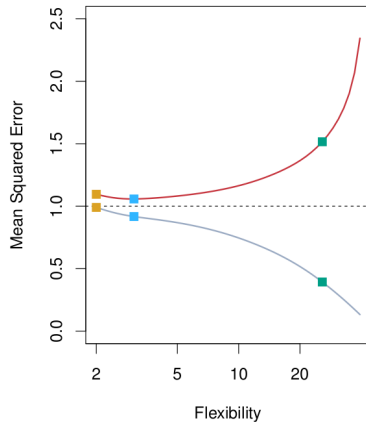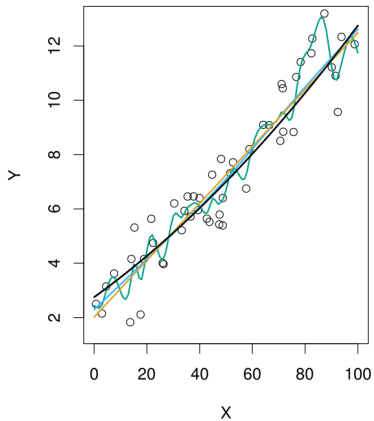The observations
$\{(x_1', y_1'), \cdots, (x_{n'}', y_{n'}')\}$ used to compute the average squared prediction error

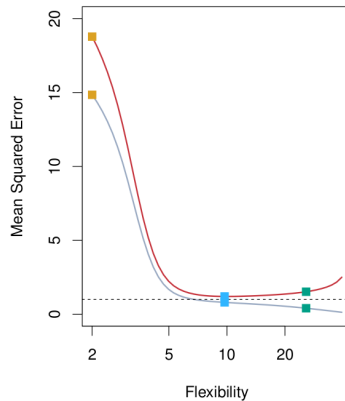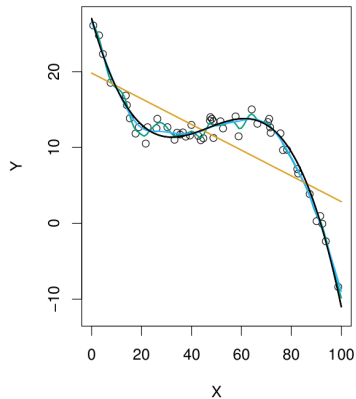$$\frac{1}{n'} \sum_i (y_i' - \hat{f}(x_i'))^2$$

Why not just get the best model for the training data?
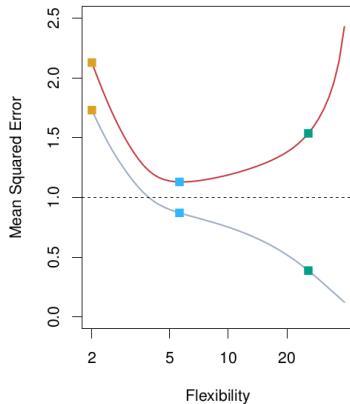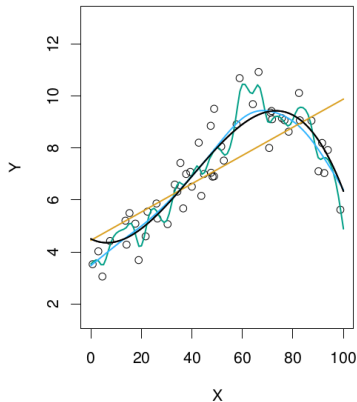
A more linear example

A more non-linear example
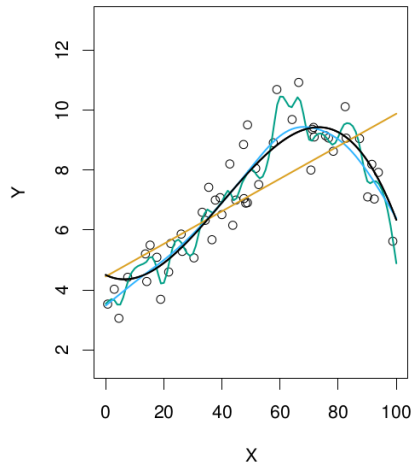
A simple solution: Train/test split
More on this in Ch 5

Section 2

Bias-Variance Trade-Off

Bias-variance

$$E(y_0 - \hat{f}(x_0))^2 = \mathrm{Var}(\hat{f}(x_0)) + \left[\mathrm{Bias}(\hat{f}(x_0))\right]^2 + \mathrm{Var}(\varepsilon)$$
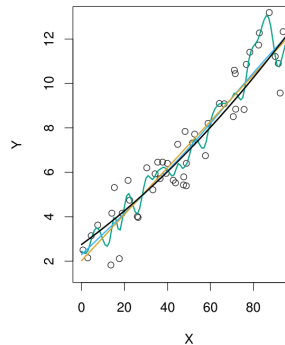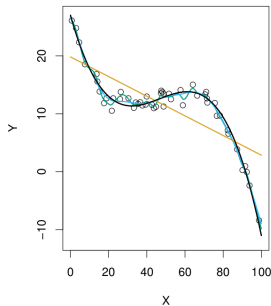
Variance

Variance: the amount by which $\hat{f}$ would change if we estimated it using a different training data set.
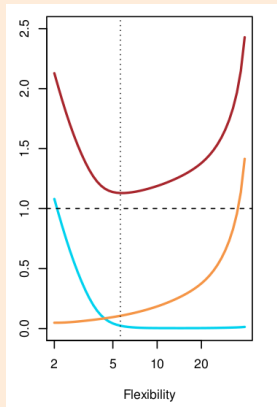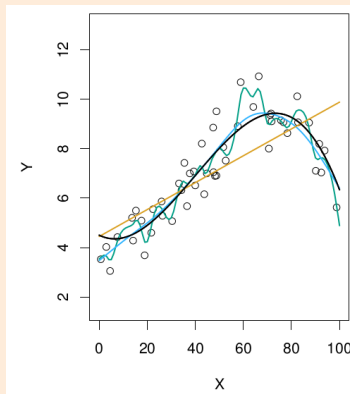
Bias

Bias: the error that is introduced by approximating a (complicated) real-life problem by a much simpler model.
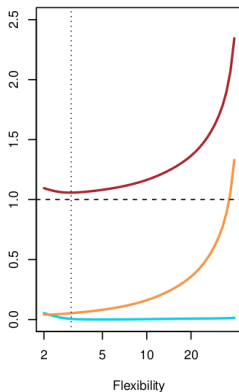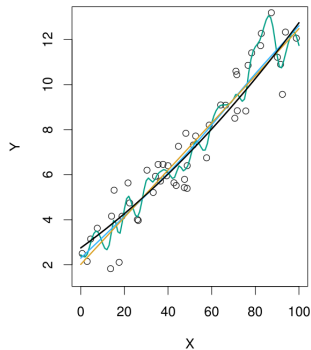
Group work

Label the line corresponding to each of the following:

- MSE
- Bias
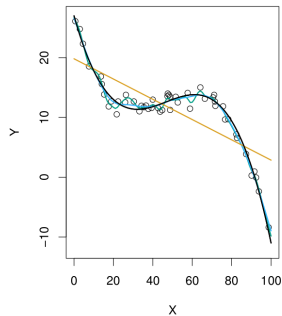- Variance of $\hat{f}(x_0)$
- Variance of ε

$$E(y_0 - \hat{f}(x_0))^2 = \mathrm{Var}(\hat{f}(x_0)) + \left[\mathrm{Bias}(\hat{f}(x_0))\right]^2 + \mathrm{Var}(\varepsilon)$$
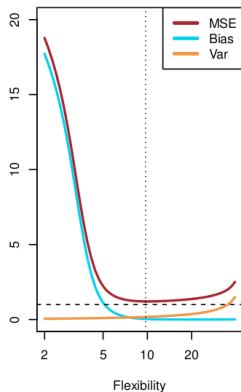
Another example

$$E(y_0 - \hat{f}(x_0))^2 = \text{Var}(\hat{f}(x_0)) + \left[\text{Bias}(\hat{f}(x_0))\right]^2 + \text{Var}(\varepsilon)$$
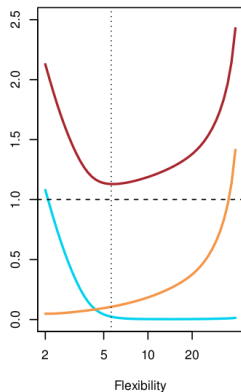
Yet another example

$$E(y_0 - \hat{f}(x_0))^2 = \text{Var}(\hat{f}(x_0)) + \left[\text{Bias}(\hat{f}(x_0))\right]^2 + \text{Var}(\varepsilon)$$

Bias-variance trade off

Flexibility

$$E(y_0 - \hat{f}(x_0))^2 = \mathrm{Var}(\hat{f}(x_0)) + \left[\mathrm{Bias}(\hat{f}(x_0))\right]^2 + \mathrm{Var}(\varepsilon)$$

Group work: coding

See jupyter notebook

Next time

- Next week:
  - Monday no class
  - 3.1 Linear Regression
- Sunday (9/7)
  - Homework due midnight on crowdmark