

Ch 2.2: Assessing Model Accuracy

Lecture 3 - CMSE 381

Michigan State University

::

Dept of Computational Mathematics, Science & Engineering

Fri, Jan 16, 2026

Last time:

- Ch 2.1, Vocab day!

Announcements:

- First homework due Sunday, 1/25. Covers:
 - ▶ Mon 1/12 lecture
 - ▶ Weds 1/14 Lecture
 - ▶ Today 1/16 Lecture

Covered in this lecture

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

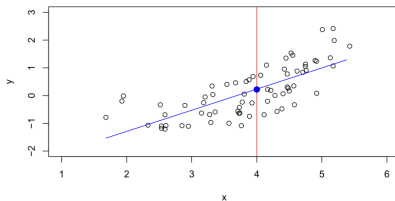
Quick review of notation

Section 1

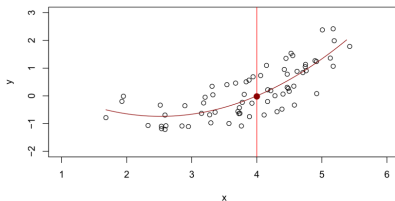
Mean Squared Error

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.

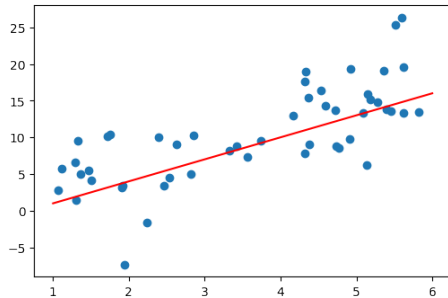


No free lunch

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

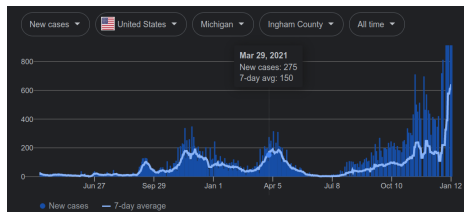
Given the following data, you decide to use the model

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

What is the MSE?

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

Training MSE



Train vs test

Training set:

The observations

$\{(x_1, y_1), \dots, (x_n, y_n)\}$ used to get
the estimate $\hat{f}$

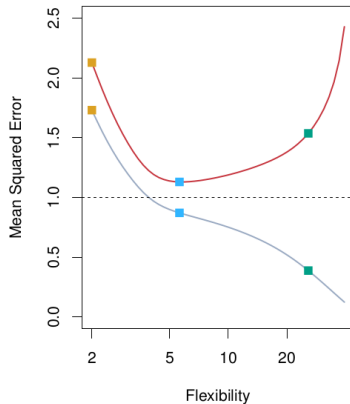
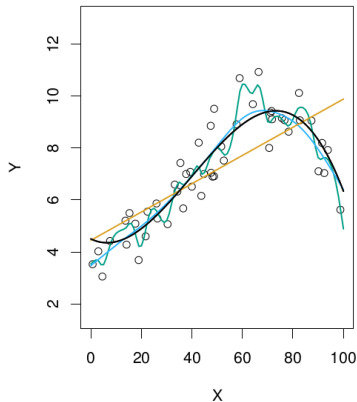
Test set:

The observations

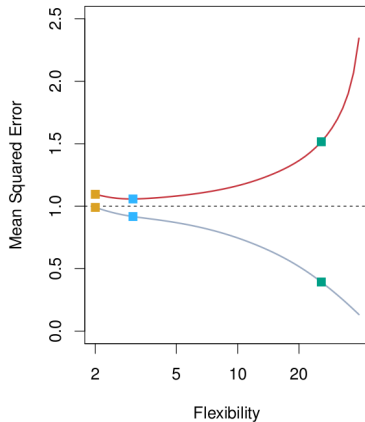
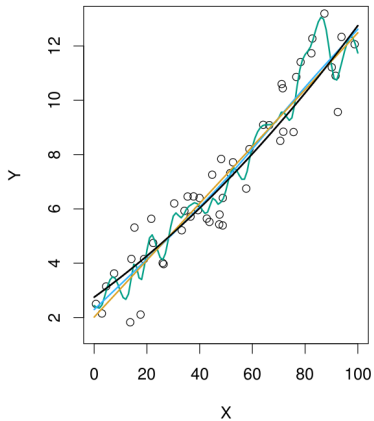
$\{(x'_1, y'_1), \dots, (x'_{n'}, y'_{n'})\}$ used to
compute the average squared
prediction error

$$\frac{1}{n'} \sum_i (y'_i - \hat{f}(x'_i))^2$$

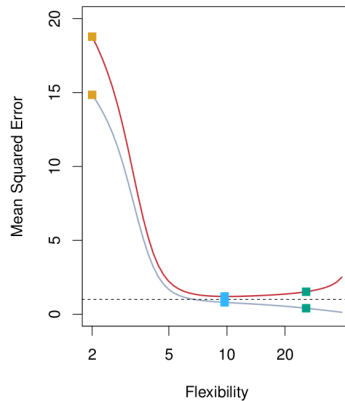
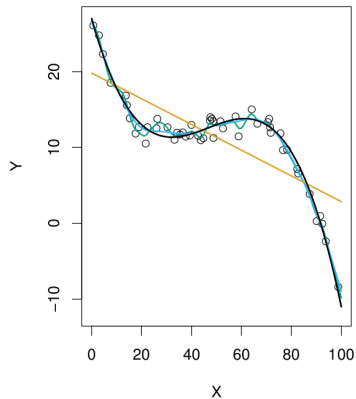
Why not just get the best model for the training data?



A more linear example



A more non-linear example



A simple solution: Train/test split

More on this in Ch 5

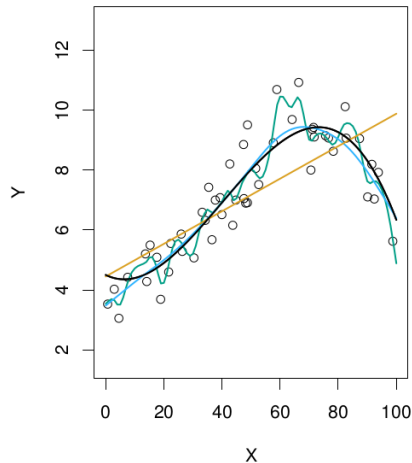
Section 2

Bias-Variance Trade-Off

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

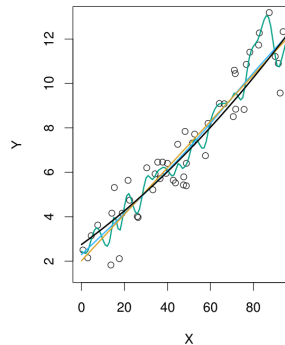
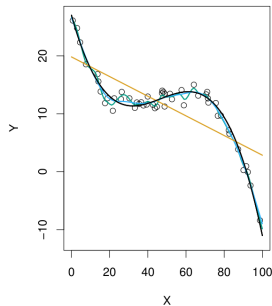
Variance

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

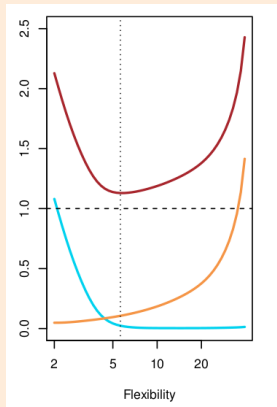
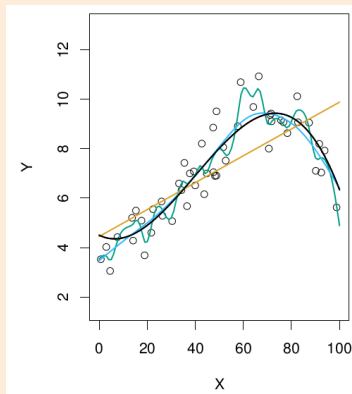


Bias

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



Group work

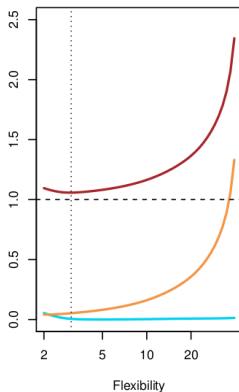
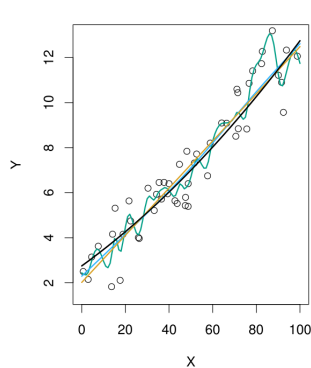


Label the line corresponding to each of the following:

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

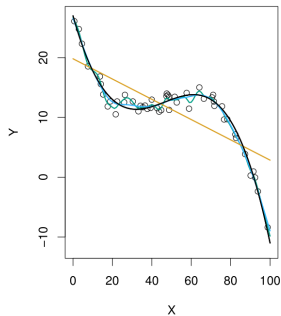
$$E(y_0 - \hat{f}(x_0))^2 = \text{Var}(\hat{f}(x_0)) + [\text{Bias}(\hat{f}(x_0))]^2 + \text{Var}(\varepsilon)$$

Another example

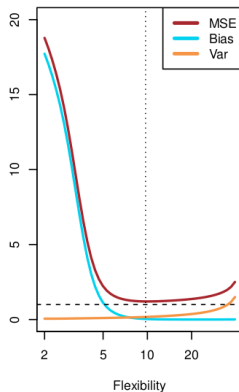


$$E(y_0 - \hat{f}(x_0))^2 = \text{Var}(\hat{f}(x_0)) + [\text{Bias}(\hat{f}(x_0))]^2 + \text{Var}(\varepsilon)$$

Yet another example

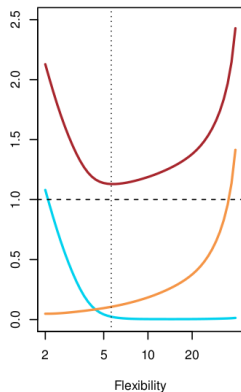


Mean Squared Error



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

Bias-variance trade off



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

Group work: coding

See jupyter notebook

Next time

- Wednesday:
 - ▶ 3.1 Linear Regression
- Sunday
 - ▶ Homework due midnight on www.msu-stt-learning.com

Lec #		Date	Topic	Reading	HW	Pop Quizzes	Notes
1	M	1/12	Intro / Python Review	1		Q1	
2	W	1/14	What is statistical learning	2.1			
3	F	1/16	Assessing Model Accuracy	2.2.1, 2.2.2			
	M	1/19	MLK - No Class			Q2	
4	W	1/21	Linear Regression	3.1			
5	F	1/23	More Linear Regression	3.1	HW #1 Due Sun 1/25		
6	M	1/26	Multi-linear Regression	3.2		Q3	
7	W	1/28	Probably More Linear Regression	3.3			
8	F	1/30	Last of the Linear Regression		HW #2 Due Sun 2/1		
9	M	2/2	Intro to classification, Bayes classifier, KNN classifier	2.2.3			