

Ch 5.1.4-5: More Cross-Validation

Lecture 14 - CMSE 381

Prof. Elizabeth Munch

Michigan State University

::

Dept of Computational Mathematics, Science & Engineering

Fri, Oct 4, 2024

Announcements

Last time:

- k-fold CV

This lecture:

- More k -fold CV
- Bias-Variance Tradeoff
- CV for classification

Announcements:

- Exam 1 grades
- HW #4 Posted.
 - ▶ Changed Deadline! Due Wednesday Oct 9.

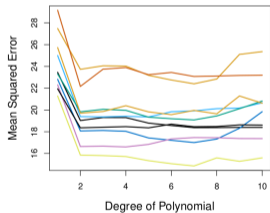
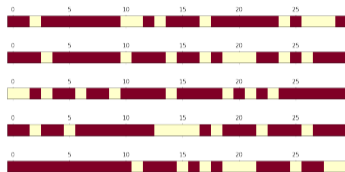
Lec #	Date			Reading	HW
12	Mon	9/30	Leave one out CV	5.1.1, 5.1.2	
13	Wed	10/2	k-fold CV	5.1.3	
14	Fri	10/4	More k-fold CV,	5.1.4-5	
15	Mon	10/7	k-fold CV for classification	5.1.5	
16	Wed	10/9	Resampling methods: Bootstrap	5.2	HW #4 Due Weds 10/9
17	Fri	10/11	Subset selection	6.1	
18	Mon	10/14	Shrinkage: Ridge	6.2.1	
19	Wed	10/16	Shrinkage: Lasso	6.2.2	
20	Fri	10/18	Dimension Reduction	6.3	HW #5 Due Fri 10/18
	Mon	10/21	No class - Fall break		
	Wed	10/23	Review		
	Fri	10/25	Midterm #2		

Section 1

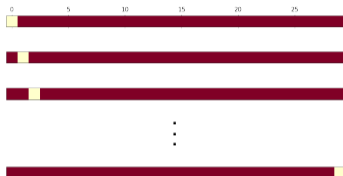
k-fold CV

Approximations of Test Error

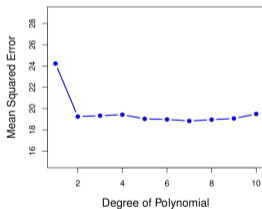
Validation Set



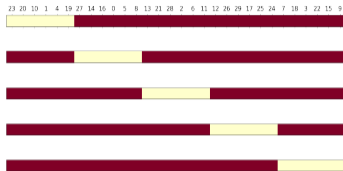
LOOCV



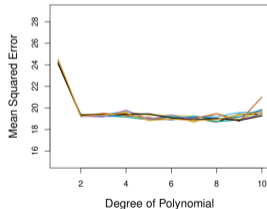
LOOCV



K-fold CV

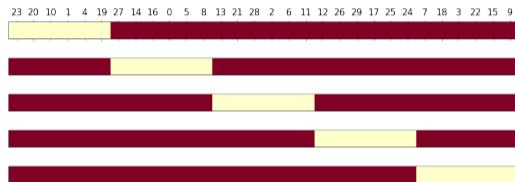


10-fold CV



Definition of k -fold CV

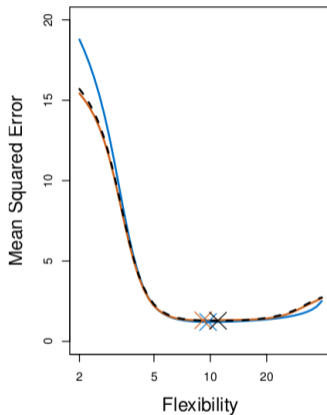
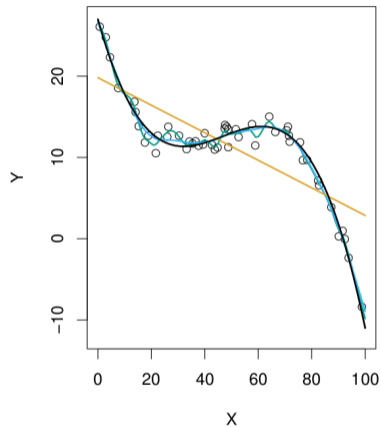
- Randomly split data into k -groups (folds)
- Approximately equal sized. For the sake of notation, say each set has ℓ points
- Remove i th fold U_i and reserve for testing.
- Train the model on remaining points
- Calculate
$$\text{MSE}_i = \frac{1}{\ell} \sum_{(x_j, y_j) \in U_i} (y_j - \hat{y}_j)^2$$
- Rinse and repeat



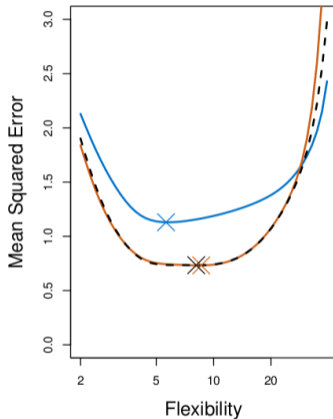
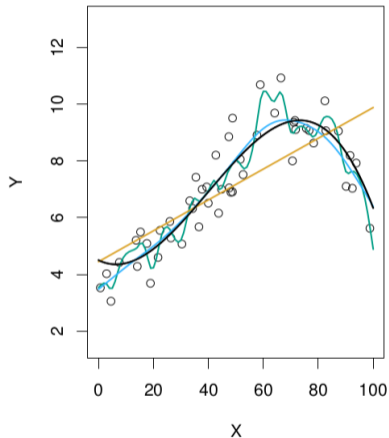
Return

$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^k \text{MSE}_i$$

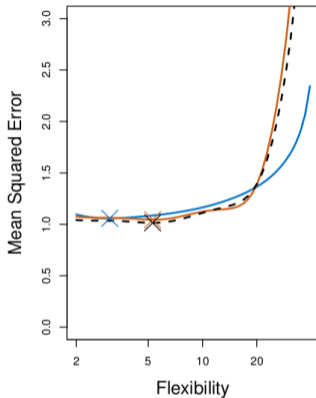
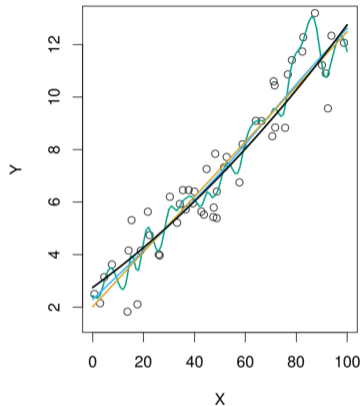
Comparison with simulated data: Ex 3



Comparison with simulated data: Ex 1



Comparison with simulated data: Ex 2



Takeaways from the examples

Bias-Variance Tradeoff: Bias

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

Bias-Variance Tradeoff: Variance

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

Section 2

Using K-Fold CV on Polynomial Linear Regression

Polynomial regression

Replace linear model

$$y_i = \beta_0 + \beta_1 x_1 + \varepsilon_i$$

with

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \cdots + \beta_d x_1^d + \varepsilon_i$$

Faking linear regression into doing our work for us

Coding - Build a plot for train/test scores vs flexibility

Next time

Lec #	Date		Reading	HW	
12	Mon	9/30	Leave one out CV	5.1.1, 5.1.2	
13	Wed	10/2	k-fold CV	5.1.3	
14	Fri	10/4	More k-fold CV,	5.1.4-5	
15	Mon	10/7	k-fold CV for classification	5.1.5	
16	Wed	10/9	Resampling methods: Bootstrap	5.2	HW #4 Due Weds 10/9
17	Fri	10/11	Subset selection	6.1	
18	Mon	10/14	Shrinkage: Ridge	6.2.1	
19	Wed	10/16	Shrinkage: Lasso	6.2.2	
20	Fri	10/18	Dimension Reduction	6.3	HW #5 Due Fri 10/18
	Mon	10/21	No class - Fall break		
	Wed	10/23	Review		
	Fri	10/25	Midterm #2		