Ch 5.1.4-5: More Cross-Validation

Lecture 14 - CMSE 381

Prof. Lianzhang Bao

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

Weds, Feb 19, 2025

Announcements

Last time:

- k-fold CV

This lecture:

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

Announcements:

- Exam 1 grades.... hopefully soon
- HW #4 will be posted soon.
    - Due Sunday 3/2.

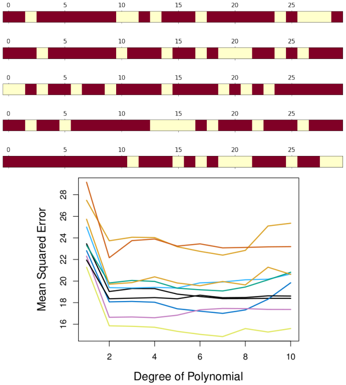
12	F	2/14	Leave one out CV	5.1.1, 5.1.2		
13	M	2/17	k-fold CV	5.1.3		
14	W	2/19	More k-fold CV	5.1.4-5		Q5
15	F	2/21	k-fold CV for classification	5.1.5		
16	M	2/24	Subset selection	6.1		
17	W	2/26	Shrinkage: Ridge	6.2.1		
18	F	2/28	Shrinkage: Lasso	6.2.2	HW #4 Due Sun 3/2	
	M	3/3	Spring Break			
	W	3/5	Spring Break			
	F	3/7	Spring Break			
19	M	3/10	PCA	6.3		
20	W	3/12	PCR	6.3		Q6
	F	3/14	***Review***		HW #5 Due Sun 3/16	
	M	3/17	**Midterm #2**			

Section 1

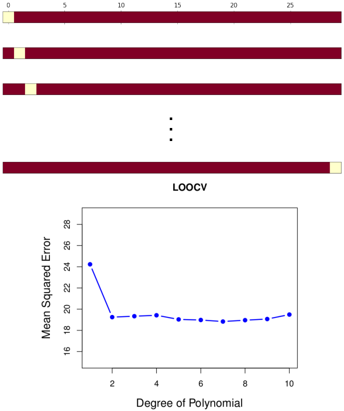
k-fold CV

Approximations of Test Error

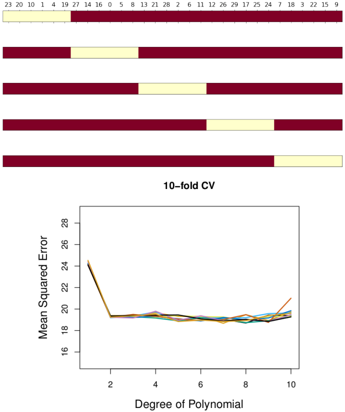
Validation Set



LOOCV

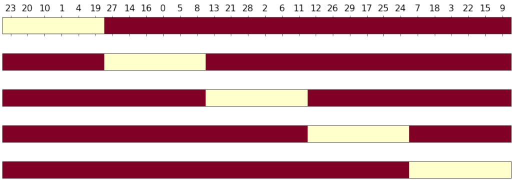


K-fold CV



Definition of k-fold CV

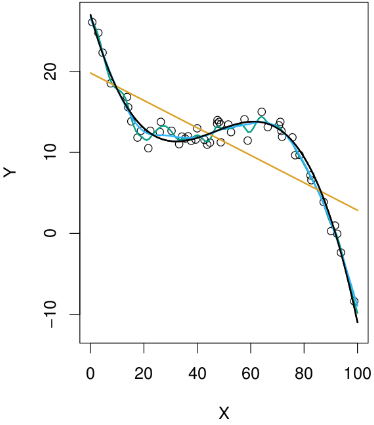
- Randomly split data into k-groups (folds)
- Approximately equal sized. For the sake of notation, say each set has ℓ points
- Remove ith fold U_i and reserve for testing.
- Train the model on remaining points
- Calculate $\mathrm{MSE}_i = \frac{1}{\ell} \sum_{(x_j, y_j) \in U_i} (y_j - \hat{y}_j)^2$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

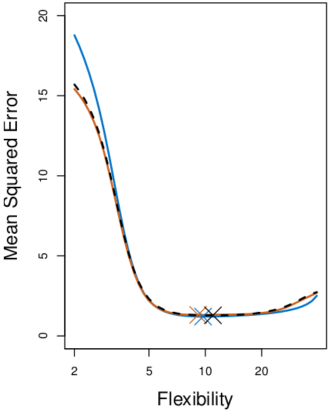


Return

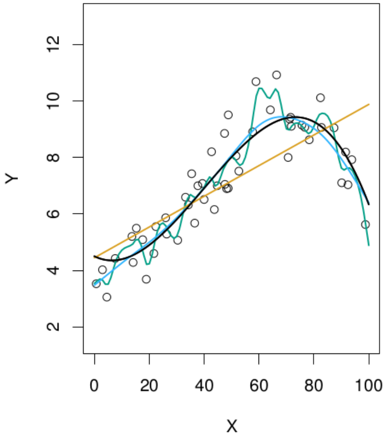
$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^{k} \mathrm{MSE}_i$$

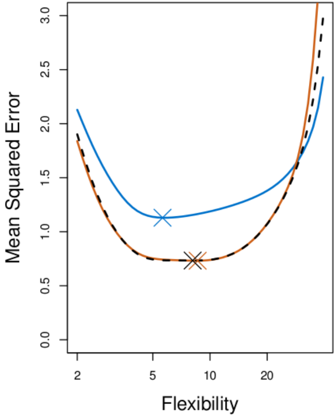
Comparison with simulated data: Ex 3



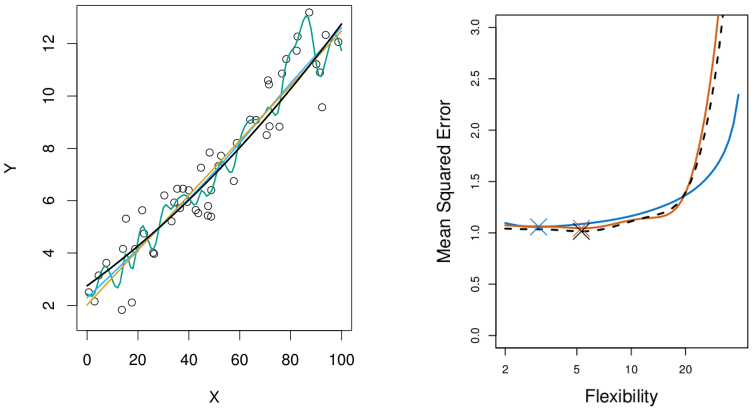


Comparison with simulated data: Ex 1





Comparison with simulated data: Ex 2



Takeaways from the examples

Bias-Variance Tradeoff: Bias

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

Bias-Variance Tradeoff: Variance

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

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

12	F	2/14	Leave one out CV	5.1.1, 5.1.2		
13	M	2/17	k-fold CV	5.1.3		Q5
14	W	2/19	More k-fold CV	5.1.4-5		
15	F	2/21	k-fold CV for classification	5.1.5		
16	M	2/24	Subset selection	6.1		
17	W	2/26	Shrinkage: Ridge	6.2.1		
18	F	2/28	Shrinkage: Lasso	6.2.2	HW #4 Due Sun 3/2	
	M	3/3	Spring Break			
	W	3/5	Spring Break			
	F	3/7	Spring Break			
19	M	3/10	PCA	6.3		Q6
20	W	3/12	PCR	6.3		
	F	3/14	***Review***		HW #5 Due Sun 3/16	
	M	3/17	**Midterm #2**			