STAT406- Methods of Statistical Learning Lecture 3

Matias Salibian-Barrera

UBC - Sep / Dec 2019

Cross validation

- We can show that

$$E_{\text{data},Y|\mathbf{X}_0}\left[\left(Y-\hat{f}(\mathbf{X}_0)\right)^2\right]$$
$$= V(\hat{f}(\mathbf{X}_0)) + B^2\left(\hat{f}(\mathbf{X}_0)\right) + V(\epsilon)\,,$$

where V denotes variance and

$$B^2\left(\hat{f}(\mathbf{X}_0)\right) = \left[E_{\text{data}}\left(\hat{f}(\mathbf{X}_0) - f(\mathbf{X}_0)\right)\right]^2 \,.$$

is the squared bias:

Activity!

Cross validation

- Conservative lower bound for MSPE
- Discussion
- Proper way of using CV