1. Show the Bias-Variance Decomposition, and introduce Bias-Variance Tradeoff

Bias-Variance Tradeoff (1/2)

- **Bias-Variance Decomposition:** The expected test MSE can always be decomposed as

$$E\left[(Y - \hat{f}(X;\mathcal{D}))^2\right] =$$

2.

4. **(10%+10%)** Suppose that the input and output variables of the n training data points can be expressed as $\mathbf{X} = (x_1, \ldots, x_n)^T$ and $\mathbf{y} = (y_1, \ldots, y_n)^T$, respectively. In ridge regression, the coefficient vector $\beta = (\beta_0, \ldots, \beta_p)^T$ is chosen such that $\|\mathbf{y} - \mathbf{X}\beta\|^2 + \lambda\|\beta\|^2$ is minimized for some $\lambda \geq 0$.

(a) Show that the resulting coefficient estimate is given by

$$\hat{\beta}_\lambda = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}$$

and that $\|\beta_\lambda\|_{\lambda>0} \leq \|\beta_\lambda\|_{\lambda=0}$

(b) Show that the training error is

$$\overline{\text{err}} = \frac{1}{n}\mathbf{y}\left[\mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\right]^2\mathbf{y}$$

and that it is an increasing function of λ.

and find the standard errors of the coefficient SE($\hat{\beta}$).

3. Suppose that the available data set is D = {(x1, y1),(x2, y2),(x3, y3)} = {(1, 2),(3, 7),(5, 8)}. Find the overall MSE of 3-fold CV

4. Find the classification rule using QDA

5.

(a) Compare the Validation Set Approach and Cross-Validation

(b) 為什麼要有 logistic regression

(c) explain the meaning of t-statistic.