Bias - Variance Tradeoff

<u>Recall</u> losses $l_i(\theta)$ are random functions of parameters θ (typically assume i.i.d.)

$$R_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} l_i(\theta) \qquad \text{empirical risk}$$

$$R(\theta) = \mathbb{E}\, l_1(\theta) \qquad \text{true risk}$$

<u>classification</u> 0-1 loss: $l_i(\theta) = \mathbb{1}\{g_\theta(x_i) \neq y_i\}$ $\left(\overset{\text{or}}{=} l_i(g_\theta)\right)$

$$R^* = \inf_g R(g) \quad \text{is Bayes Risk}$$

<u>def</u> 1-NN classifier for training set $\{x_i, y_i\}_{i=1}^n$, $x_i \in \mathbb{R}^p$, $y_i \in \{-1,1\}$, let $NN(x)$ be the

$$NN(x) = \operatorname*{argmin}_{i=1,\dots,n} \left\{ \|x_i - x\|_2^2 : \cancel{x_i \text{ in training data}} \right\}$$

then $\hat{g}_{NN}(x) = y_{NN(x)}$.

<u>thm</u> $\lim\limits_{n \to \infty} R(\hat{g}_{NN}) \leq 2R^*$ (Cover & Hart, 1967)

<u>regression</u> $l_i(f) = (y_i - f(x_i))^2$ for $f: \mathbb{R}^p \to \mathbb{R}$

$\eta(x) = \mathbb{E}[Y \mid X = x]$ minimizes $R(f)$,

$$R(\eta) = \mathbb{E}(Y - \eta(X))^2 =: R^*$$