

Outline

- Bias and Variance
- Learning Curves

Expected E_{out} and E_{in}

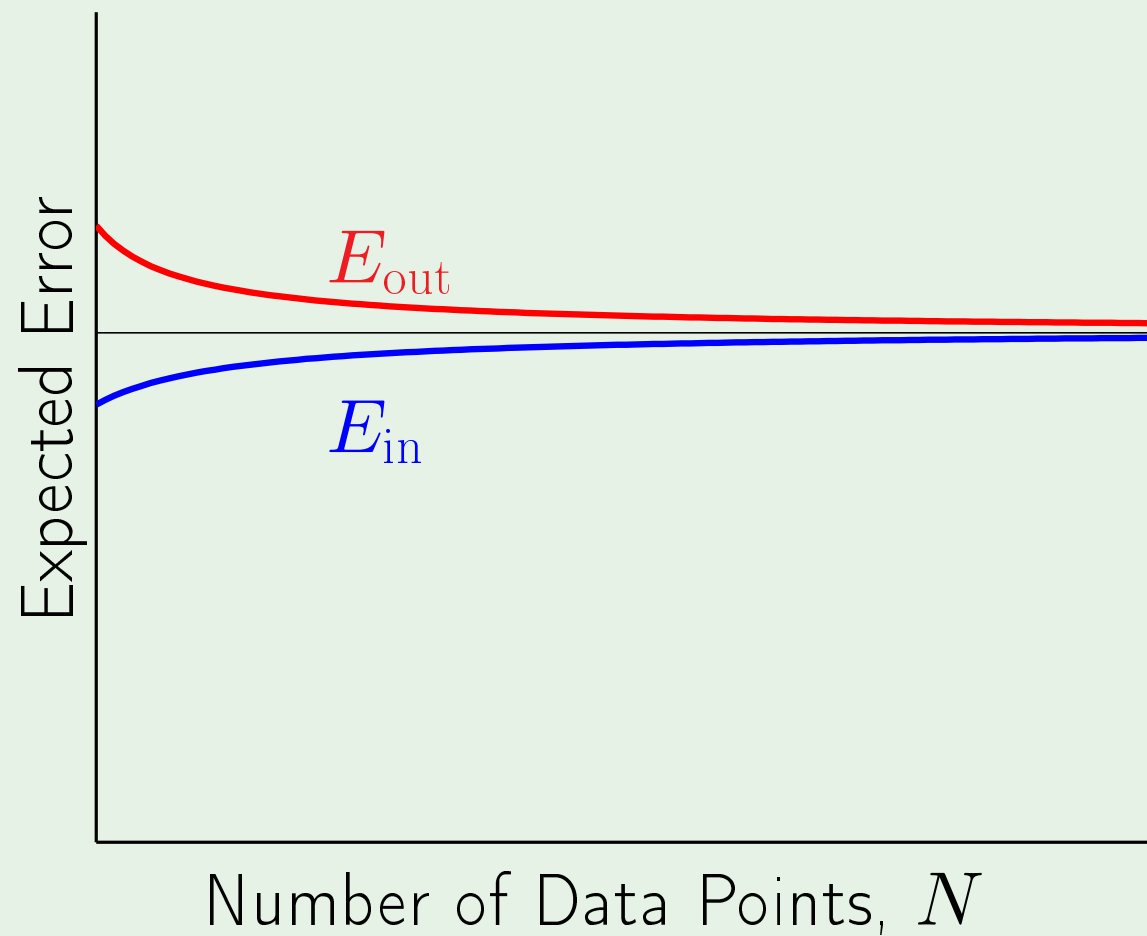
Data set \mathcal{D} of size N

Expected out-of-sample error $\mathbb{E}_{\mathcal{D}}[E_{\text{out}}(g^{(\mathcal{D})})]$

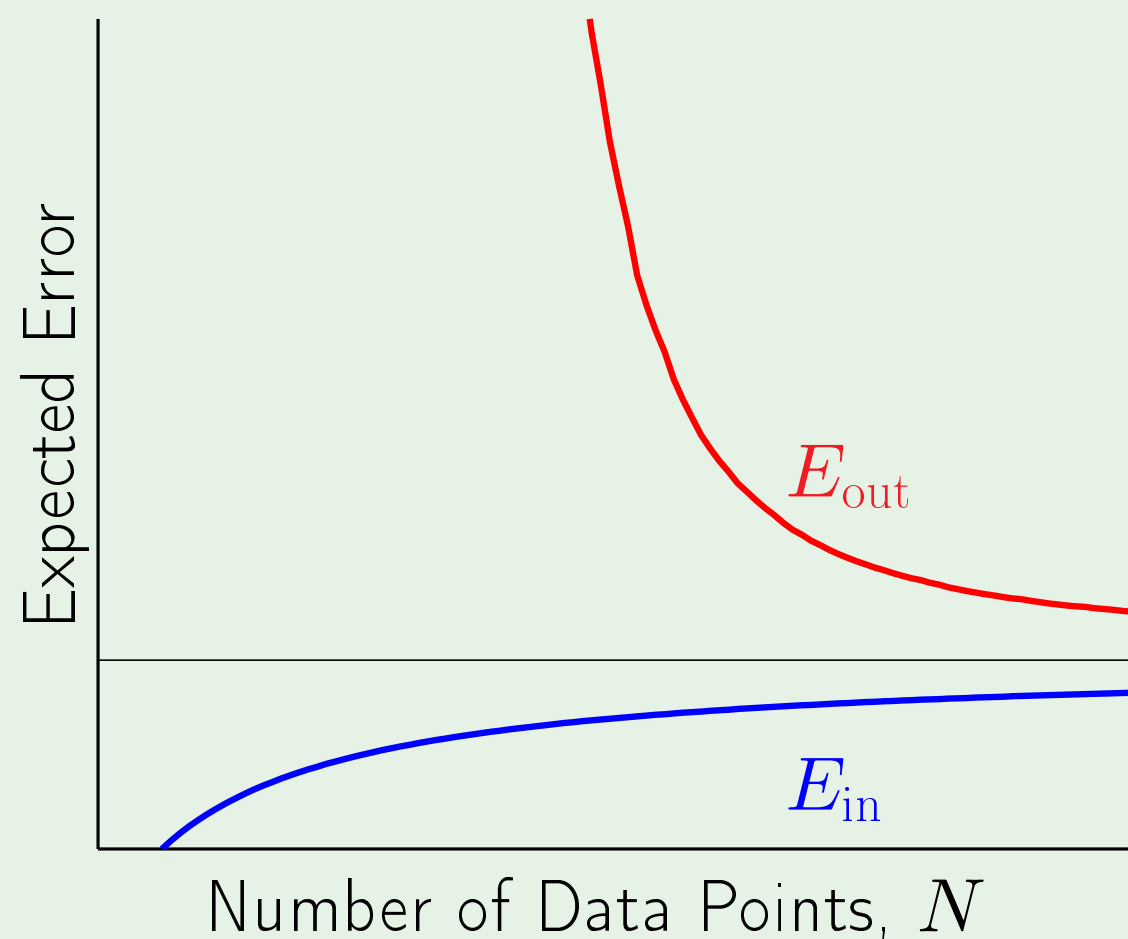
Expected in-sample error $\mathbb{E}_{\mathcal{D}}[E_{\text{in}}(g^{(\mathcal{D})})]$

How do they vary with N ?

The curves

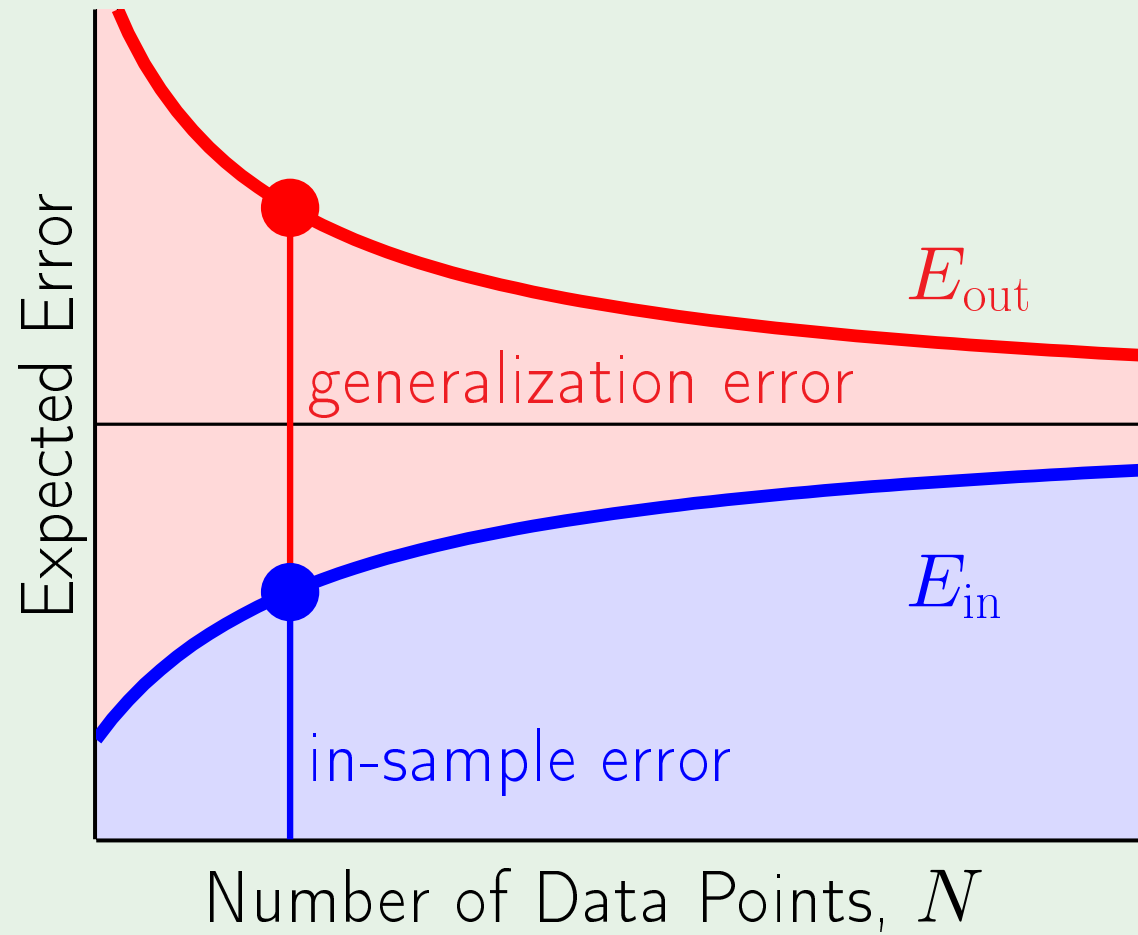


Simple Model

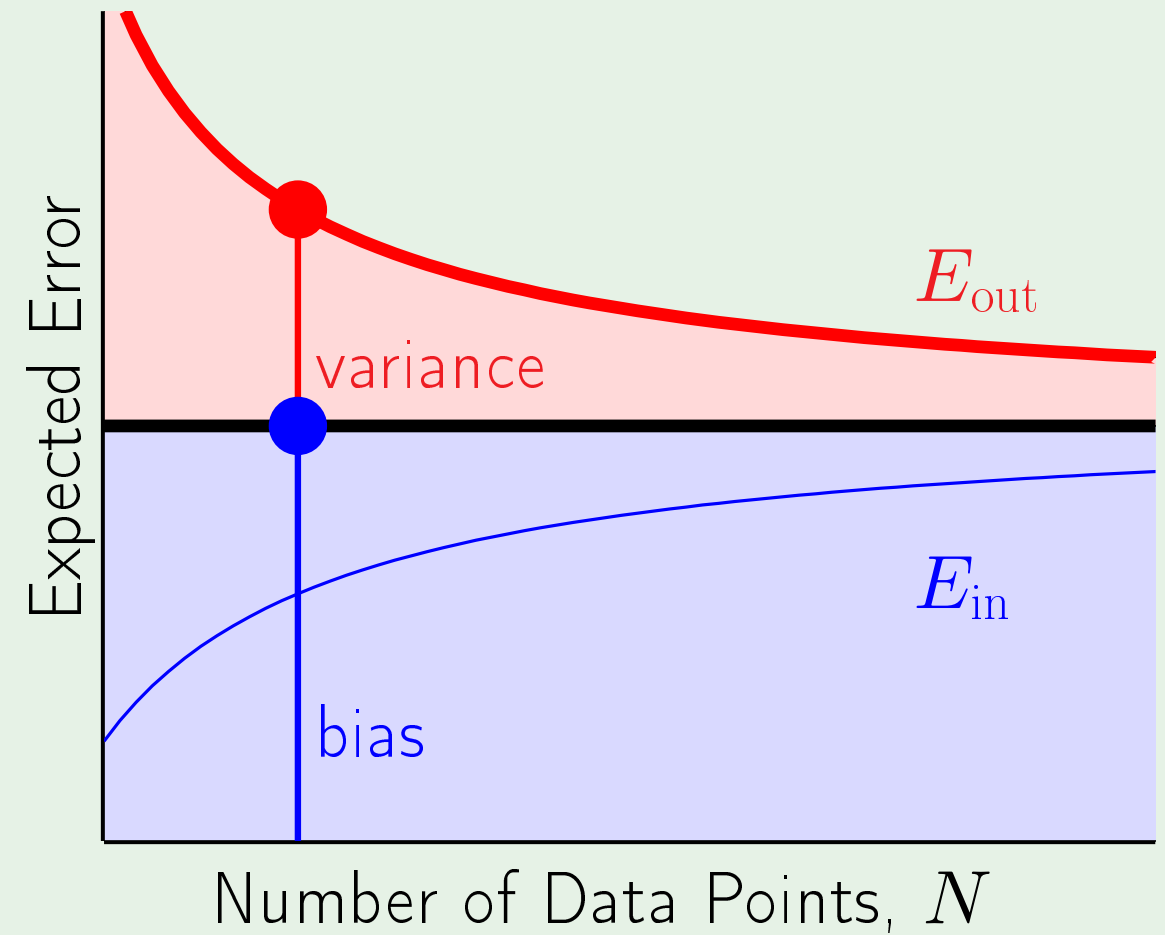


Complex Model

VC versus bias-variance



VC analysis



bias-variance