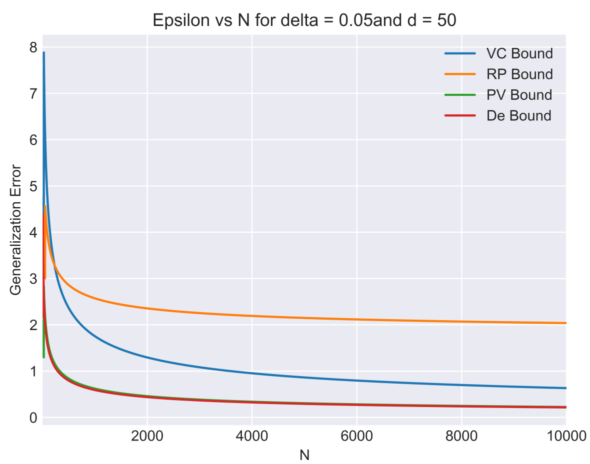
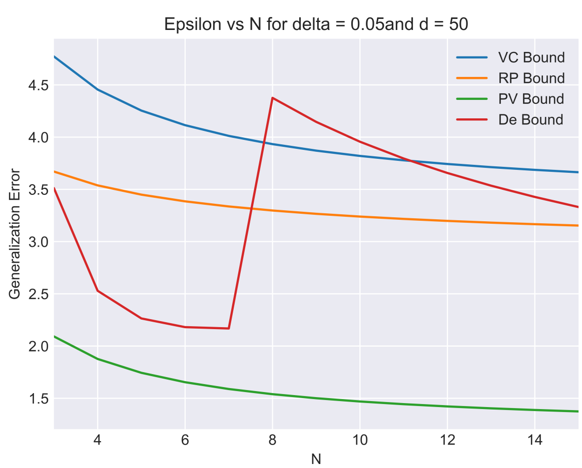
1. (d)
$$0.05 = \delta = 4m_{\mathcal{H}}(2N)\exp\left(-\frac{1}{8}\epsilon^2 N\right)$$
Making the large sample approximation
$$m_{\mathcal{H}}(N) \approx N^{d_{\text{VC}}}$$
So
$$0.05 = \delta \approx 4(2N)^{d_{\text{VC}}}\exp\left(-\frac{1}{8}\epsilon^2 N\right)$$
Solving for N must be done numerically, and this gives
$$N \approx 453,000$$

2. (d)



3. (c)

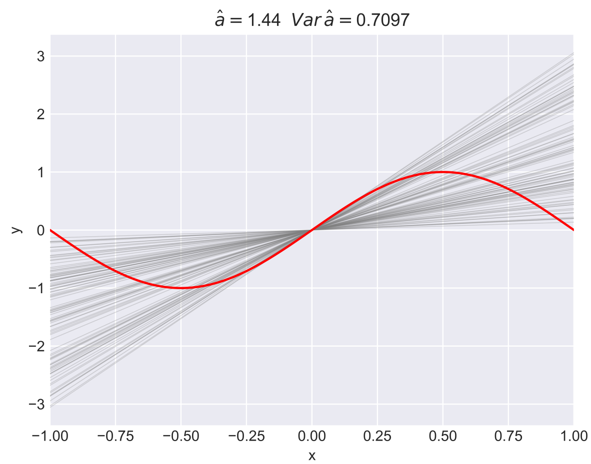


4. (e) We can get the expected value $\hat{a}$ by running a linear regression multiple times and averaging out, or we can calculate the value

$$\hat{a} = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} dx_1\, dx_2 \frac{x_1 \sin \pi x_1 + x_2 \sin \pi x_2}{x_1^2 + x_2^2} = \frac{1}{2}\int_{-1}^{1}\int_{0}^{1} dx_1\, dx_2 \frac{x_1 \sin \pi x_1 + x_2 \sin \pi x_2}{x_1^2 + x_2^2}$$

$$\hat{a} \approx 1.43$$

$$\operatorname{Var} \hat{a} = \frac{1}{2}\int_{-1}^{1}\int_{0}^{1} dx_1\, dx_2 \left(\frac{x_1 \sin \pi x_1 + x_2 \sin \pi x_2}{x_1^2 + x_2^2} - \hat{a}\right)^2 = 0.71$$



5. (b) The bias is given by

$$\text{bias} = \frac{1}{2}\int_{-1}^{1} dx\, (\sin \pi x - \hat{a}x)^2 = 0.27$$

6. (d) The variance of $\hat{a}$ is given in problem 4.

$$\operatorname{Var} = \frac{1}{3}\operatorname{Var} \hat{a} = 0.24$$

7. (b)

```
CALCULATING EOUT FOR h = b, ax, ax + b, ax2, ax2 + b
[  0.75027242   0.51089643   1.93203402   7.18532571 354.21041544]
```

8. (c) Assume that k is the VC dimension, we have

$$2^k = m_{\mathcal{H}}(k) = 2 \cdot 2^{k-1} - \binom{k-1}{q} = m_{\mathcal{H}}(k) - \binom{k-1}{q}$$

which implies that $q = k$. This implies that that the VC dimension is q.

9. (b) We can think of the size of

$$\bigcap_{k=1}^{K} \mathcal{H}_k$$

being at most the smallest set. So

$$0 \leq d_{\text{VC}} \left(\bigcap_{k=1}^{K} \mathcal{H}_k \right) \leq \min\{d_{\text{VC}}(\mathcal{H}_k)\}_{k=1}^{K}$$

10. Same as the previous problem, the lower bound is bound by the maximum size, or the VC dimension of the highest VC-dimensional hypothesis set. Now consider two hypothesis sets $\mathcal{H}_1$ and $\mathcal{H}_2$ with VC dimension d_1 and d_2 respectively. Then

$$m_{\mathcal{H}_1 \cup \mathcal{H}_2}(N) \leq \sum_{i=0}^{d_1} \binom{N}{i} + \sum_{i=0}^{d_2} \binom{N}{i},$$

and using the fact that

$$\binom{N}{i} = \binom{N}{N-i}$$

yields

$$m_{\mathcal{H}_1 \cup \mathcal{H}_2}(N) \leq \sum_{i=0}^{d_1} \binom{N}{i} + \sum_{i=0}^{d_2} \binom{N}{N-i} = \sum_{i=0}^{d_1} \binom{N}{i} + \sum_{i=N-d_2}^{N} \binom{N}{i}$$

If $N - d_2 > d_1 + 1$, that is $N \geq d_1 + d_2 + 2$, then

$$m_{\mathcal{H}_1 \cup \mathcal{H}_2}(N) \leq \sum_{i=0}^{N} \binom{N}{i} - \binom{N}{d_1+1} = 2^N - \binom{N}{d_1+1} < 2^N$$

so the VC dimension of the union set is at most $d_1 + d_2 + 1$.

Now we can prove inductively for many sets $\mathcal{H}_1, \cdots, \mathcal{H}_k$ that the VC dimension is

$$k - 1 + \sum_{i=0}^{k} d_i.$$