MA4199 Project – Bias Variance Tradeoff

Ng Wei Le

1 Approximation Theorem

The below theorem gives us some justification as to why the minimum norm interpolating function was chosen, though this only works under noiseless conditions:

Theorem 1. *Fix* $h^* \in \mathcal{H}_\infty$ *. Let* $(x_1, y_1), ..., (x_n, y_n)$ *be i.i.d. random variables where* x_i *drawn randomly from a compact cube* $\Omega \subset \mathbb{R}^d$*,* $y_i = h^*(x_i)\, \forall i$*. There exists* $A, B > 0$ *such that for any interpolating* $h \in \mathcal{H}_\infty$ *with high probability*

$$\sup_{x \in \Omega} |h(x) - h^*(x)| < Ae^{-B(n/log\, n)^{1/d}}(\|h^*\|_{\mathcal{H}_\infty} + \|h\|_{\mathcal{H}_\infty})$$