ϵ误差估计

> [Shiyu Liang and R Srikant. Why deep neural networks? arXiv preprint arXiv:1610.04161 , 2016.]

神经网络$\tilde{f}(x)$拟合函数$f(x)$，满足$|f(x)-\tilde{f}(x)|\le\varepsilon$

结论

以ε的误差拟合函数：

- 分段光滑函数，$\Theta(\log\frac{1}{\varepsilon})$层，$\mathcal{O}(polylog(\frac{1}{\varepsilon})$个神经元
- 分段光滑函数，$o(\log\frac{1}{\varepsilon})$层，$\Omega(poly(\frac{1}{\varepsilon})$个神经元
- 可微凸函数， $\Omega(\log\frac{1}{\varepsilon})$个神经元

单变量函数上限

拟合函数	层数	神经元个数
$f(x)=x^2\ (x\in[0,1])$	$O(\log\frac{1}{\varepsilon})$	$O(\log\frac{1}{\varepsilon})$
$f(x)=\sum_{i=0}^{p}a_ix^i\ (x\in[0,1],\sum_{i=0}^{p}\lvert a_i\rvert\le 1)$	$O(p+\log\frac{p}{\varepsilon})$	$O(p\log\frac{p}{\varepsilon})$
$f(x)\ (x\in[0,1])$ (对$\forall n\in[\lceil\log\frac{2}{\varepsilon}\rceil+1]$，$\lvert\lvert f^{(n)}\rvert\rvert_\infty\le n!$)	$O(\log\frac{1}{\varepsilon})$	$O((\log\frac{1}{\varepsilon})^2)$
$f=\prod_{i=1}^{k}h_i\ (x\in[0,1])$ (对 $\forall n\in[\lceil 4k\log 4k+4k+2\log\frac{2}{\varepsilon}\rceil+1]$，$\lvert\lvert h_i^{(n)}\rvert\rvert_\infty\le n!$)	$O(k\log k+\log\frac{1}{\varepsilon})$	$O((k\log k)^2+(\log\frac{1}{\varepsilon})^2)$
$f(x)=h_1(h_2(\cdots(h_k(x))))\ (x\in[0,1])$ (对$\forall n\in[\lceil\log\frac{2}{\varepsilon}\rceil+1]$ ，$\lvert\lvert h_i^{(n)}\rvert\rvert_\infty\le n!$，$h_i:[0,1]\to[0,1]$)}	$O(k\log k\log\frac{1}{\varepsilon}+\log k(\log\frac{1}{\varepsilon})^2)$	$O(k\log k\log\frac{1}{\varepsilon}+k^2(\log\frac{1}{\varepsilon})^2+(\log\frac{1}{\varepsilon})^4)$

下限

对强凸函数（$<\nabla f(x)-\nabla f(y),x-y>\ge\lambda||x-y||_2^2?$），$(x\in[0,1])$，要求$N\ge L(\frac{\mu}{16\varepsilon})^{\frac{1}{2L}}$