## €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 $\varepsilon$ 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
$f(x)=x^2(x\in[0,1])$	$O(\log rac{1}{arepsilon})$	$O(\log \frac{1}{\varepsilon})$
$f(x) = \sum_{i=0}^p a_i x^i \; (x \in [0,1], \sum_{i=0}^p  a_i  \leq 1)$	$O(p + \log rac{p}{arepsilon})$	$O(p\log rac{p}{arepsilon})$
$f(x)$ $(x\in[0,1])$ (স্থే $orall n\in[\lceil\lograc{2}{arepsilon} ceil+1]$ , $  f^{(n)}  _{\infty}\leq n!$ )	$O(\log rac{1}{arepsilon})$	$O((\log \frac{1}{arepsilon})^2)$
$f=\prod_{i=1}^k h_i\ (x\in[0,1])$ (저 $orall n\in[\lceil 4k\log 4k+4k+2\lograc{2}{arepsilon} ceil+1]$ , $  h_i^{(n)}  _\infty\leq n!$ )	$O(k\log k + \log rac{1}{arepsilon})$	$O(((k\log k)^2 + (\log rac{1}{arepsilon})^2)$
$f(x)=h_1(h_2(\cdots(h_k(x))))\;(x\in[0,1])\;($ X $orall orall n\in[\lceil\lograc{2}{arepsilon} ceil+1]$ , $  h_i^{(n)}  _\infty\leq n!,\;h_i:[0,1] o[0,1])$ }	$O(k\log k\log rac{1}{arepsilon} + \log k(\log rac{1}{arepsilon})^2)$	$O(k \log k \log \frac{1}{arepsilon} + k^2 (\log \frac{1}{arepsilon})^2 + (\log \frac{1}{arepsilon})^4)$

## 下限

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