Topics in Applied Mathematics: Infinitely Large Neural Networks, 3341.751
E. Ryu
Spring 2022

Homework 1
Due 5pm, Friday, March 18, 2022

Problem 1: Let $d = 1$ and $\Omega = [-10, 10]$. Show that the activation function $\sigma(x) = x^2$ is not discriminatory.

Remark. Despite not having a universality result for wide neural networks, quadratic activation functions are often considered in deep learning theory research.

Problem 2: *Uni. apx. with ReLU activation functions.* Let $\Omega \subset \mathbb{R}^d$ be compact. Let $\sigma(r) = \max\{0, r\}$ be the ReLU activation function. Show that $\mathrm{span}\{\sigma(a^\intercal x + b)\}_{a\in\mathbb{R}^d,\, b\in\mathbb{R}}$ is dense in $(\mathcal{C}(\Omega), \|\cdot\|_\infty)$.

Hint. Consider $\sigma(r) - \sigma(r-1)$.

Problem 3: *Uni. apx. with exponential activation functions.* Let $\Omega \subset \mathbb{R}^d$ be compact. Let $\sigma(r) = \exp(r)$. Show that $\mathrm{span}\{\sigma(a^\intercal x + b)\}_{a\in\mathbb{R}^d,\, b\in\mathbb{R}}$ is dense in $(\mathcal{C}(\Omega), \|\cdot\|_\infty)$.

Hint. Follow the reasoning used with sinusoidal activation functions.

Problem 4: Let $\Omega \subset \mathbb{R}^d$ be compact. Let $(\mathcal{C}(\Omega; \mathbb{R}^k), \|\cdot\|_\infty)$ be the Banach space of continuous functions $f\colon \Omega \to \mathbb{R}^k$ with norm

$$\|f\|_\infty = \sup_{x\in\Omega} \|f(x)\|_\infty = \sup_{\substack{x\in\Omega\\ i=1\dots,k}} |f_i(x)|.$$

Assume $\sigma\colon \mathbb{R} \to \mathbb{R}$ is an activation function with a universal approximation theorem for wide 2-layer neural networks. In other words, $f_\theta\colon \mathbb{R}^d \to \mathbb{R}$ of the form

$$f_\theta(x) = \sum_{i=1}^N u_i \sigma(a_i^\intercal x + b_i),$$

where $a_1, \dots, a_N \in \mathbb{R}^d$, $b_1, \dots, b_N \in \mathbb{R}$, and $u_1, \dots, u_N \in \mathbb{R}$, is dense in $(\mathcal{C}(\Omega), \|\cdot\|_\infty)$. Show that $f_\theta\colon \mathbb{R}^d \to \mathbb{R}^k$

$$f_\theta(x) = \sum_{i=1}^N u_i \sigma(A_i x + b_i),$$

where σ applies element-wise, $A_1, \dots, A_N \in \mathbb{R}^{k\times d}$, $b_1, \dots, b_N \in \mathbb{R}^k$, and $u_1, \dots, u_N \in \mathbb{R}$, is dense in $(\mathcal{C}(\Omega; \mathbb{R}^k), \|\cdot\|_\infty)$.

Problem 5: Let $\Omega \subseteq \mathbb{R}^d$ be compact. Show that if $\mu \in \mathcal{M}(\Omega)$ such that

$$\hat{\mu}(a) = \int_\Omega e^{ia^\intercal x}\, d\mu(x) = 0$$

for all $a \in \mathcal{R}^d$, then $\mu = 0$.

Hint. Define $L_\mu[h] = \int_\Omega h[x]\, d\mu(x)$. Show that $L_\mu\colon \mathcal{C}(\Omega) \to \mathbb{R}$ is linear and bounded. Using the Stone–Weierstrass theorem, show that $L_\mu[h] = 0$ for all $h \in \mathcal{C}(\Omega)$. Finally, appeal to the Riesz–Markov–Kakutani representation theorem to conclude that $L_\mu = 0$ implies $\mu = 0$.