Mathematics for New Technologies in Finance

Solution sheet 1

Exercise 1.1

(a) What is the formal definition of shallow or deep neural networks? (You might need the words affine function and activation function in the definition)

(b) Why we always like to consider non-linear activation functions?

(c) If the activation functions of neural network are bounded, e.g. $x \mapsto \tanh(x)$, is the neural network bounded for all possible inputs? If the activation functions are unbounded, e.g. ReLU, is the neural network bounded on compact input spaces?

(d) Is the sum of two neural networks still a neural network? Is the product of two neural networks still a neural network? Is the composition of two neural networks still a neural network?

(e) Are neural networks with ReLU activation functions differentiable?

Solution 1.1

(a) Shallow neural network is a composition of affine function, activation function, and affine function. Deep neural network is a iteratively composition of affine functions and activation functions with the first and final ones be affine functions (You can also write it mathematically).

(b) Otherwise, neural network is affine function

(c)
1. Yes if we are talking about a neural network with fixed weights and the last layer is the activation function, and NO is we are talking about neural network with all possible weights (This shed some light on the UAT of neural network with activation function bounded by 1).
2. Yes because continuous function on compact set is bounded.

(d)
1. Yes.
2. In general NO unless you choose the nonlinear function for this purpose.
3. Yes but be careful about how this coincides with the def of neural network on the "composition layer".

(e) In general NO, but YES for special case (identity function expressed by neural network).

Exercise 1.2

(a) Build a tent neural network $h\colon \mathbb{R} \to \mathbb{R}$ s.t.

$$h(x) = \begin{cases} x+1, & x \in [-1,0] \\ 1-x, & x \in [0,1] \\ 0, & \text{otherwise} \end{cases}, \tag{1}$$

from a ReLU neural network.

(b) Use (a) to prove the universal approximation theory on $[0, 1]$ i.e. every continuous function f on [0, 1] can be uniformly approximated by neural networks (see the Faber-Schauder expansion of a continuous function).

Solution 1.2

(a)

$$h(x) = \text{ReLU}\Big(1 - \text{ReLU}(x) - \text{ReLU}(-x)\Big) \tag{2}$$

(b) By shifting and scaling we prove that Faber-Schauder basis are neural networks. Therefore by the Faber-Schauder expansion of a continuous function $[0, 1]$ we prove the UAT (universal approximation theory) see [1].

Exercise 1.3 See the notebook

(a) Code a neural network to approximate a function on $[-5, 5]$.

(b) Code a tent neural network on $\mathbb{R}^2$.

Solution 1.3 See notebook

Exercise 1.4 Prove the real continuous function on compact set with supremum norm i.e. $(C_0(K), \|\cdot\|_\infty)$ is a Banach space.

Solution 1.4 The vector space is from the definition so we remain to prove the completeness. Let f_n be a Cauchy sequence under uniform norm. Then pointwisely $f_n(x)$ is also a Cauchy sequence for all $x \in [0, 1]$. So by the completeness of $\mathbb{R}$, $f_n(x)$ has a limit which we denote by $f(x)$. By the uniform Cauchy, for all $\epsilon > 0$ there exists N_ϵ s.t. for all $m, n > N_\epsilon$, $\|f_m - f_n\| \leq \epsilon$. Letting $m \to \infty$ we have $\|f - f_n\| \leq \epsilon$ which proves that $fn \to f$ uniformly. Therefore $f \in (C_0(K), \|\cdot\|_\infty)$ as well, which concludes the proof.

References

[1] Wikipedia contributors. Haar wavelet — Wikipedia, the free encyclopedia, 2021. [Online; accessed 1-March-2022].