Extension of the Universal Approximation Theorem to Complex Functions

Adrián Pérez-Salinasssssssssssssssssssss[1]

[1] *Barcelona Supercomputing Center*

In this document we will extend the universal approximation theorem from Ref. [1] to complex functions. For doing so, we must beware that the UAT is an existence theorem, and not a constructive one. The proof of this theorem is based on topology and functional analysis, and does not provide any tool for knowing how fine must be the tuning of our approximations for getting close to a given function.

I. STATEMENT OF THE UNIVERSAL APPROXIMATION THEOREM

In its original form, the Universal Approximation Theorem (UAT from now) is written as follows

Theorem 1 *Let I_n denote the n-dimensional cube $[0,1]^n$. The space of continuous functions on I_n is denoted by $C(I_n)$, and we use $|\cdot|$ to denote the uniform norm of any function in $C(I_n)$. Let σ be a discriminatory and sigmoidal function. Given a function $f \in C(I_n)$ there exists a function*

$$G(\vec{x}) = \sum_j \alpha_j \sigma(\vec{w}_j \cdot \vec{x} + b_j) \tag{1}$$

such that

$$|G(\vec{x}) - f(\vec{x})| < \varepsilon \qquad \forall \vec{x} \in I_n \tag{2}$$

for $w_j \in \mathbf{R}^n$ and $b_j \in \mathbf{R}$

We also need to define the assumptions of this theorem

Definition 1 *A function σ is sigmoidal if*

$$\sigma(t) \rightarrow \begin{cases} 1 \text{ as } t \rightarrow +\infty \\ 0 \text{ as } t \rightarrow -\infty \end{cases} \tag{3}$$

Definition 2 *A function σ is discriminatory if for a measure $\mu \in M(I_n)$,*

$$\int_{I_n} \sigma(\vec{w}_j \cdot \vec{x} + b_j) d\mu(x) = 0 \leftrightarrow \mu = 0, \tag{4}$$

where $M(I_n)$ is the space of signed regular Borel measures on I_n.

This theorem is an existence theorem, and thus there is no contribution on how many terms from Eq. (1) are needed for reaching a level precision ε. This theorem can be proved by means of topology and some previous results from functional analysis. The assumptions of the UAT are given for matching the assumptions of the results supporting it.

The extension made in Ref. [2] extends the UAT to any bounded nonconstant and nonlinear function. This generalization is important from a mathematical and practical point of view, but most common functions can be constructed my means of sigmoidal functions. For instance, a sine (cosine) function can be understood as a sum of sigmoidal functions with the proper shape. Two sigmoids can cancel each other properly and the result would be one cycle of the sine (cosine) function. In principle, this can be repeated periodically. As sine (cosine) functions are a sum of sigmoids, then every function can be represented as a sum of sines (cosines).

II. EXTENSION TO COMPLEX VARIABLE

We will follow the steps of the demonstration of this UAT as in Ref. [1] for complex variables. Let us assume that this time the quantities α_j are allowed to be complex variables. The norm $|\cdot|$ implies now the modulus of a complex quantity. Thus, the set $C(I_n)$ is defined on the body of complex, although I_n is still real. The UAT still holds under these conditions.

Let $S \subset C(I_n)$ be the set of functions

$$S = \{G(x) : G(x) = \sum_j^N \alpha_j \sigma(\vec{w}_j \cdot \vec{x} + b_j)\}, \tag{5}$$

where $\alpha_j \in \mathbf{C}$, $\vec{w}_j \in \mathbf{R}^n$, $b_j \in \mathbf{R}$ and $\vec{x} \in I_n$. We claim that the closure of S, namely $\bar{S}$ is the whole space $C(I_n)$.

Let us proceed by contradiction, as in Ref. [1]. Let us suppose that $\bar{S} \subset C(I_n)$, but $\bar{S} \neq C(I_n)$. By the Hahn-Banach theorem there is a linear functional L acting on $C(I_n)$ such that

$$L(S) = L(\bar{S}) = 0, \qquad L \neq 0. \tag{6}$$

Theorem 2 : Hahn-Banach *[3, 4]*

Set $\mathbf{K} = \mathbf{R}$ or $\mathbf{C}$. *Let V be a $\mathbf{K}-$ vector space with a seminorm $p : V \rightarrow \mathbf{R}$. If $\varphi : U \rightarrow \mathbf{K}$ is a $\mathbf{K}-$linear functional on a $\mathbf{K}-$linear subspace $U \subset V$ such that*

$$|\varphi(x)| \leq p(x) \qquad \forall x \in U, \tag{7}$$

then there exist a linear extension $\psi : V \rightarrow \mathbf{K}$ of φ to the whole space V such that

$$\psi(x) = \varphi(x) \qquad \forall x \in U \tag{8}$$

$$|\psi(x)| \leq p(x) \qquad \forall x \in V \tag{9}$$

This theorem is also used in the original paper. As our assumptions add no additional constraints, then the theorem can be used here too.

At this point the derivation is analogous to the one in Ref. [1].

Using the Riesz Representation theorem we may write L as

$$L(h) = \int_{I_n} h(x) d\mu(x) \tag{10}$$

for $\mu \in M(I_n)$ and $\forall h \in C(I_n)$.

Theorem 3 : Riesz Representation *[5]*

Let X be a locally compact Hausdorff space. For any positive linear functional ψ on $C(X)$, there exists a uniruq regular Borel measure μ such that

$$\forall f \in C_c(X): \qquad \psi(f) = \int_X f(x) d\mu(x) \tag{11}$$

In particular, σ is valid as the only imaginary quantities are given by the coefficients α_k, then

$$\int_{I_n} \sigma(\vec{w}_j \cdot \vec{x} + b_j) d\mu(x) = 0 \tag{12}$$

always.

It is also assumed that σ is discriminatory, and thus $\mu = 0$, which contradicts the assumptions. Thus $\bar{S} = C(I_n)$, the set S is dense within $C(I_n)$ and the theorem is demonstrated. We still have to show that sigmoid functions are discriminatory.

Lemma 1 *Continuous sigmoidal functions are discriminatory.*

Let us take $\sigma\left(\lambda(\vec{w}_j \cdot \vec{x} + b_j) + c\right)$. As $\lambda \to \infty$ it is equivalent to the function

$$\gamma(x) = \begin{cases} = 1 & \text{for } \vec{w} \cdot \vec{x} + b > 0 \\ = 0 & \text{for } \vec{w} \cdot \vec{x} + b < 0 \\ = \sigma(c) & \text{for } \vec{w} \cdot \vec{x} + b = 0. \end{cases} \tag{13}$$

If we call the Lebesgue Bounded Convergence Theorem we have

$$0 = \int_{I_n} \sigma_\lambda(x) d\mu(x) = \int_{I_n} \gamma(x) d\mu(x) = \sigma(c)\mu(\Pi_{\vec{w},b}) + \mu(H_{\vec{w},b}) \tag{14}$$

where $\Pi_{\vec{w},b} = \{x : \vec{w} \cdot \vec{x} + b = 0\}$, and $H_{\vec{w},b} = \{x : \vec{w} \cdot \vec{x} + b > 0\}$.

Theorem 4 : Lebesgue Bounded Convergence *[6]* *Let $\{f_n\}$ be a sequence of complex-valued measurable functions on a measure space (S, Σ, μ). Suppose that $\{f_n\}$ converges pointwise to a function f and is dominated by some integrable function $g(x)$ in the sense*

$$|f_n(x)| \leq g(x), \qquad \int_S |g| d\mu < \infty \tag{15}$$

then

$$\lim_{n \to \infty} \int_S f_n d\mu = \int_S f d\mu \tag{16}$$

The measure of all half-planes being 0 implies that $\mu = 0$. Let us fix $\vec{w}$, and for a bounded measurabe function h we define the linear functional

$$F(h) = \int_{I_n} h(\vec{w} \cdot \vec{x}) d\mu(x), \tag{17}$$

which is bounded on $L^\infty(\mathbf{R})$ since μ is a finite signed measure. Let h be an indicator of the half planes $h(u) = 1$ if $u \geq -b$ and $h(u) = 0$ otherwise, then

$$F(h) = \int_{I_n} h(\vec{w} \cdot \vec{x}) d\mu(x) = \mu(\Pi_{\vec{w},b}) + \mu(H_{\vec{w},b}) = 0. \tag{18}$$

By linearity, $F(h) = 0$ for any simple function, such as sum of indicator functions of intervales [7].

In particular, for the bounded measurable functions $s(u) = \sin(\vec{w} \cdot \vec{x}), c(u) = \cos(\vec{w} \cdot \vec{x})$ we can write

$$F(c + is) = \int_{I_n} \exp i\vec{w} \cdot x d\mu(x) = 0. \tag{19}$$

The Fourier Transform of this F is null, thus $\mu = 0$.

III. LINK TO QUANTUM COMPUTING

$$R_z(\phi) R_y(\vec{w} \cdot \vec{x} + a)$$

[1] George Cybenko. Approximations by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2:183–192, 1989.

[2] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4:251–257, 1991.

[3] Hans Hahn. Über lineare gleichungssysteme in linearen räumen. *Journal für die reine und angewandte Mathematik*, 157:214–229, 1927.

[4] Stefan Banach. Sur les fonctionnelles linéaires ii. *Studia Mathematica*, 1:223–239, 1929.

[5] Frédéric Riesz. Démonstration nouvelle d'un théorème concernant les opérations fonctionnelles linéaires. In

Annales scientifiques de l'École Normale Supérieure, volume 31, pages 9–14, 1914.

[6] Alan J Weir, Alan J Weir, et al. *General integration and measure*, volume 2. CUP Archive, 1974.

[7] RB Ash. Real analysis and probability academic press. *New York*, 1972.