Hypercomplex-valued Neural Networks

Part 1 – Introduction and Basic Concepts.

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

Marcos Eduardo Valle
Universidade Estadual de Campinas (Unicamp)
Campinas, Brazil.

UNICAMP

Aims and Goals:

Introduce a broad class of hypercomplex-valued neural networks.

Why a broad class of hypercomplex-valued neural networks?

Neural networks (NNs) have demonstrated outstanding results in several application areas, including image detection and classification.

- Hypercomplex-valued neural networks showed competitive or superior performance but with fewer parameters than their equivalent neural networks defined on real numbers.
- The advantages of hypercomplex-valued NNs include reducing the number of parameters and treating multiple values as a single entity.

A search on Scopus on February 2022 by “neural network” combined with an additional hypercomplex algebra search term:

Hypercomplex algebra additional search term	**# Documents**
“complex number” OR “complex valued”	1,855
“quaternion”	704
“Clifford"	121
“hypercomplex”	80
“octonion”	22
“hyperbolic number” OR “hyperbolic valued”	18
“Bicomplex”	8
“Cayley-Dickson”	5
“tessarines”	1
“coquaternion”	1
“Klein four-group”	1

Although some documents appeared in more than one search, it is clear that complex and quaternion-valued models dominate the research on HvNNs.

Aims and goals:

Most research focuses on networks based on complex numbers and quaternions.

Alternative algebras can result in efficient hypercomplex-valued NNs!

The following presents a general framework for hypercomplex algebras.

- On the one hand, we will be able to work with a broad class of hypercomplex-valued neural networks.
- On the other hand, we will not be able to explore specific properties (like the geometric properties of Clifford algebras) of the hypercomplex algebra.

Addition and Multiplication are Key Concepts

Addition and multiplication are core concepts for developing hypercomplex-valued models.

For example, dense and convolutional layers are described by the following equations, respectively:

$$y_i = \varphi\left(s_i + b_i\right), \quad \text{with} \quad s_i = \sum_{j=1}^{N} w_{ij} x_j, \tag{1}$$

and

$$\mathbf{J}(p,k) = \varphi\big((\mathbf{I} * \mathbf{F})(p,k) + b(k)\big), \tag{2}$$

with

$$(\mathbf{I} * \mathbf{F})(p,k) = \sum_{c=1}^{C} \sum_{q \in G} \mathbf{I}(p + S(q), c)\mathbf{F}(q,c,k). \tag{3}$$

Basic Concepts on Algebra

Definition 1 (Algebra – Schafer (1961))

An algebra $\mathbb{V}$ over a field $\mathbb{F}$ is a vector space over $\mathbb{F}$ with an additional bilinear operation called multiplication.

In the following, we focus on algebras over the field of real numbers; that is, we only consider $\mathbb{F} = \mathbb{R}$.

As a bilinear operation, the multiplication of $x, y \in \mathbb{V}$, denoted by the juxtaposition xy, satisfies

$$(x+y)z = xz + yz \quad \text{and} \quad z(x+y) = zx + zy, \quad \forall x, y, z \in \mathbb{V}. \tag{4}$$

and

$$\alpha(xy) = (\alpha x)y = x(\alpha y), \quad \forall \alpha \in \mathbb{R} \quad \text{and} \quad x, y \in \mathbb{V}. \tag{5}$$

Finite-Dimensional Vector Algebra

We will be only concerned with finite dimensional vector spaces. In other words, we assume that $\mathbb{V}$ is a vector space of dimension n, i.e., $dim(\mathbb{V}) = n$.

Let $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ be an ordered basis for $\mathbb{V}$. Given $x \in \mathbb{V}$, there is an unique n-tuple $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ such that

$$x = \sum_{i=1}^{n} x_i e_i. \tag{6}$$

The scalars $x_1, \ldots, x_n$ are the coordinates of x relative to the ordered basis $\mathcal{E}$.

In computational applications, $x \in \mathbb{V}$ is given by its coordinates relative to the ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$. In other words, x is usually written as a vector in R^n.

Given an ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$, the mapping $\varphi : \mathbb{V} \to \mathbb{R}^n$ given by

$$\varphi(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \quad \forall x \in \mathbb{V}, \tag{7}$$

yields an isomorphism between $\mathbb{V}$ and $\mathbb{R}^n$.

The absolute value $x \in \mathbb{V}$ with respect to the basis $\mathcal{E} = \{e_1, \ldots, e_n\}$ is the Euclidean norm of $\varphi(x)$:

$$|x| := \|\varphi(x)\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}. \tag{8}$$

Remark:

The absolute value of x given by (8) is not an invariant; it depends on the basis $\mathcal{E} = \{e_1, \ldots, e_n\}$.

Multiplication Table

Given an ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$ of $\mathbb{V}$, the multiplication is completely determined by the n^3 parameters $p_{ijk} \in \mathbb{R}$ which appear in the products

$$e_i e_j = \sum_{k=1}^{n} p_{ijk} e_k, \quad \forall i, j = 1, \ldots, n. \tag{9}$$

The products in (9) can be arranged in the multiplication table:

	e_1	e_j	e_n
		$\vdots$	
e_i	$\cdots$	$\sum_{k=1}^{n} p_{ijk} e_k$	$\cdots$
		$\vdots$	

Commutative Algebras

Definition 2 (Commutative Algebra)

An algebra $\mathbb{V}$ is commutative if

$$xy = yx, \quad \forall x, y \in \mathbb{V}. \tag{10}$$

The properties of an algebra can be derived from the same property from the basis elements.

Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be an ordered basis for an algebra $\mathbb{V}$. The algebra is commutative if and only if

$$e_i e_j = e_j e_i, \quad \forall i, j = 1, \ldots, n.$$

From the multiplication table, an algebra is commutative if

$$p_{ijk} = p_{jik}, \quad \forall i, j, k = 1, \ldots, n.$$

Associative Algebra

Definition 3 (Associative Algebra)

An algebra $\mathbb{V}$ is associative if

$$(xy)z = x(yz), \quad \forall x, y, z \in \mathbb{V}. \tag{11}$$

Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be an ordered basis for an algebra $\mathbb{V}$. The algebra is associative if and only if

$$(e_i e_j)e_k = e_i(e_j e_k), \quad \forall i, j = 1, \ldots, n.$$

Therefore, an algebra is associative if

$$\sum_{\mu=1}^{n} p_{ij\mu} p_{k\mu\nu} = \sum_{\mu=1}^{n} p_{jk\mu} p_{i\mu\nu}, \quad \forall i, j, k, \nu = 1, \ldots, n.$$

Hypercomplex Algebra

A hypercomplex algebra, denoted by $\mathbb{H}$, is a finite-dimensional algebra in which the product has a two-sided identity (Catoni et al., 2008; Kantor and Solodovnikov, 1989).

A hypercomplex algebra $\mathbb{H}$ is equipped with an (unique) element e_0 such that

$$xe_0 = e_0 x = x, \quad \forall x \in \mathbb{V}.$$

Identity is usually the first element of the ordered basis. Thus, $\mathcal{E} = \{e_0, e_1, \ldots, e_n\}$ is an ordered basis of an hypercomplex algebra and $dim(\mathbb{H}) = n + 1$.

We often consider the canonical basis $\tau = \{1, \boldsymbol{i}_1, \ldots, \boldsymbol{i}_n\}$. Thus, a hypercomplex number is given by

$$x = x_0 + x_1 \boldsymbol{i}_1 + \ldots + x_n \boldsymbol{i}_n. \tag{12}$$

The multiplication table of a hypercomplex algebra with respect to the canonical basis $\tau = \{1, \boldsymbol{i}_1, \ldots, \boldsymbol{i}_n\}$ is

	1	$\boldsymbol{i}_1$	$\boldsymbol{i}_j$	$\boldsymbol{i}_n$
1	1	$\boldsymbol{i}_1$	$\boldsymbol{i}_j$	$\boldsymbol{i}_n$
			$\vdots$	
$\boldsymbol{i}_i$	$\boldsymbol{i}_i$	$\cdots$	$p_{ij0} + \sum_{k=1}^{n} p_{ijk} \boldsymbol{i}_k$	$\cdots$
			$\vdots$	

Thus, we have

$$p_{0jk} = p_{i0k} = \begin{cases} 1, & i = k \text{ or } j = k, \\ 0, & \text{otherwise.} \end{cases}$$

Example – Quaternions

Quaternions, introduced by Hamilton in the late 19th, are hypercomplex numbers that generalize real and complex numbers.

Using the canonical basis $\{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$, a quaternion is given by

$$x = x_0 + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}.$$

Quaternions are effective mathematical tools for describing 3D rotations.

They have also been effectively used to develop neural networks (Arena et al., 1997; Parcollet et al., 2020).

The multiplication table of quaternions is

$$\begin{array}{c|cccc} & 1 & \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \hline 1 & 1 & \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \boldsymbol{i} & \boldsymbol{i} & -1 & \boldsymbol{k} & -\boldsymbol{j} \\ \boldsymbol{j} & \boldsymbol{j} & -\boldsymbol{k} & -1 & \boldsymbol{i} \\ \boldsymbol{k} & \boldsymbol{k} & \boldsymbol{j} & -\boldsymbol{i} & -1 \end{array}$$

Because $\boldsymbol{ij} = \boldsymbol{k}$ and $\boldsymbol{ji} = -\boldsymbol{k}$, the multiplication of quaternions is non-commutative.

The multiplicative inverse of any quaternion $x \neq 0$ is

$$x^{-1} = \frac{\bar{x}}{x\bar{x}},$$

where $\bar{x} = x_0 - x_1\boldsymbol{i} - x_2\boldsymbol{j} - x_3\boldsymbol{k}$ is the conjugate of x.

Product and Bilinear Forms

Using the distributive law and the multiplication table, the product of $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{j=1}^n y_j e_j$ satisfies

$$xy = \left(\sum_{i=1}^n x_i e_i\right)\left(\sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j (e_i e_j)$$

$$= \sum_{i=1}^n \sum_{j=1}^n x_i y_j \left(\sum_{k=1}^n p_{ijk} e_k\right) = \sum_{k=1}^n \left(\sum_{i=1}^n \sum_{j=1}^n x_i y_j p_{ijk}\right) e_k$$

Because the product is bilinear, the function $\mathcal{B}_k : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ given by

$$\mathcal{B}_k(x, y) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j p_{ijk}, \quad \forall k = 1, \ldots, n, \tag{13}$$

is a bilinear form.

Proposition 1

Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be an ordered basis of an algebra $\mathbb{V}$. The multiplication of $x = \sum_{i=1}^{n} x_i e_i$ and $y = \sum_{j=1}^{n} y_j e_j$ satisfies

$$xy = \sum_{k=1}^{n} \mathcal{B}_k(x, y) e_k, \tag{14}$$

where $\mathcal{B}_k : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ is a bilinear form whose matrix representation in the ordered basis $\mathcal{E}$ is

$$B_k = \begin{bmatrix} p_{11k} & p_{12k} & \ldots & p_{1nk} \\ p_{21k} & p_{22k} & \ldots & p_{2nk} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1k} & p_{n2k} & \ldots & p_{nnk} \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad \forall k = 1, \ldots, n. \tag{15}$$

Thus, we have $\mathcal{B}_k(x, y) = \varphi(x)^T B_k \varphi(y)$.

Let $\mathbb{H}$ be a hypercomplex algebra. The matrix representation of the bilinear forms $\mathcal{B}_k : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ with respect to the canonical basis $\tau = \{1, \boldsymbol{i}_1, \ldots, \boldsymbol{i}_n\}$ are

$$B_0 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & p_{110} & \cdots & p_{1n0} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & p_{n10} & \cdots & p_{nn0} \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)},$$

and, for $k = 1, \ldots, n$,

$$B_k = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & p_{11k} & p_{12k} & \cdots & p_{1kk} & \cdots & p_{1nk} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & p_{k1k} & p_{j2k} & \cdots & p_{kkk} & \cdots & p_{knk} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & p_{n1k} & p_{n2k} & \cdots & p_{nkk} & \cdots & p_{nnk} \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

Non-Degenerate Algebra

A bilinear form $\mathcal{B}_k : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ is non-degenerate if its matrix representation B_k is non-singular with respect to any ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$.

Definition 4 (Non-degenerate algebra)

An algebra $\mathbb{V}$ is non-degenerate if all the bilinear forms in the coordinates of the multiplication are non-degenerate. Otherwise, we say that the algebra $\mathbb{V}$ is degenerate.

Remark

Non-degenerate algebras play an important role in the approximation capability of hypercomplex-valued (or vector-valued) multi-layer networks.

Matrix Representation of the Product

The multiplication to the left by $a = \sum_{i=1}^{n} a_i e_i$ yields a linear operator $\mathcal{A}_L : \mathbb{V} \to \mathbb{V}$ defined by

$$\mathcal{A}_L(x) = ax, \quad \forall x \in \mathbb{V}.$$

The matrix representation of $\mathcal{A}_L$ relative to an ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$ yields a mapping $\mathcal{M}_L : \mathbb{V} \to \mathbb{R}^{n \times n}$ given by

$$\mathcal{M}_L(a) = \begin{bmatrix} | & | & & | \\ \varphi(ae_1) & \varphi(ae_2) & \ldots & \varphi(ae_n) \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{n} a_i p_{i11} & \sum_{i=1}^{n} a_i p_{i21} & \ldots & \sum_{i=1}^{n} a_i p_{in1} \\ \sum_{i=1}^{n} a_i p_{i12} & \sum_{i=1}^{n} a_i p_{i22} & \ldots & \sum_{i=1}^{n} a_i p_{in2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_i p_{i1n} & \sum_{i=1}^{n} a_i p_{i2n} & \ldots & \sum_{i=1}^{n} a_i p_{inn} \end{bmatrix}.$$

Alternatively, we can write

$$\mathcal{M}_L(a) = \sum_{i=1}^{n} a_i P_{i:}^T, \quad \text{with} \quad P_{i:}^T = \begin{bmatrix} p_{i11} & p_{i21} & \dots & p_{in1} \\ p_{i12} & p_{i22} & \dots & p_{in2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{i1n} & p_{i2n} & \dots & p_{inn} \end{bmatrix}. \tag{16}$$

Using the matrix representation, we have

$$\varphi(ax) = \mathcal{M}_L(a)\varphi(x) = \sum_{i=1}^{n} a_i P_{i:}^T \varphi(x), \tag{17}$$

for all $a = \sum_{i=1}^{n} a_i e_i \in \mathbb{V}$ and $x \in \mathbb{V}$.

Analogously, the multiplication to the right by $a = \sum_{i=1}^{n} a_i e_i$ yields a linear operator $\mathcal{A}_R : \mathbb{V} \to \mathbb{V}$ defined by

$$\mathcal{A}_R(x) = xa, \quad \forall x \in \mathbb{V}.$$

The matrix of $\mathcal{A}_R$ relative to an ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$ yields the mapping $\mathcal{M}_R : \mathbb{V} \to \mathbb{R}^{n\times n}$ given by

$$\mathcal{M}_R(a) = \sum_{i=1}^{n} a_i P_{:i}, \quad \text{with} \quad P_{:i} = \begin{bmatrix} p_{1i1} & p_{1i2} & \cdots & p_{1in} \\ p_{2i1} & p_{2i2} & \cdots & p_{2in} \\ \vdots & \vdots & \ddots & \vdots \\ p_{ni1} & p_{ni2} & \cdots & p_{nin} \end{bmatrix}. \tag{18}$$

Using the matrix representation, we have

$$\varphi(xa)^T = \varphi^T(x)\mathcal{M}_R(a) = \sum_{i=1}^{n} a_i \varphi^T(x) P_{:i}, \quad \forall x \in \mathbb{V}, \tag{19}$$

where $a = \sum_{i=1}^{n} a_i e_i \in \mathbb{V}$.

Example – Quaternions

Consider the quaternions with the canonical basis $\tau = \{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$. The product of $x = x_0 + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$ and $y = y_0 + y_1\boldsymbol{i} + y_2\boldsymbol{j} + y_3\boldsymbol{k}$ satisfies

$$\varphi(xy) = \begin{bmatrix} x_0 & -x_1 & -x_2 & -x_3 \\ x_1 & x_0 & -x_3 & x_2 \\ x_2 & x_3 & x_0 & -x_1 \\ x_3 & -x_2 & x_1 & x_0 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = \mathcal{M}_L(x)\varphi(y).$$

Note that

$$\mathcal{M}_L(x) = x_0 P_{0:} + x_1 P_{1:} + x_2 P_{2:} + x_n P_{n:},$$

where$P_{0:} = \mathbf{I}_{4\times 4}$ is the identity matrix and

$$P_{1:}^T = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, P_{2:}^T = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, P_{3:}^T = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Similarly, we have

$$\varphi(xy)^T = [x_0, x_1, x_2, x_3]^T \begin{bmatrix} y_0 & y_1 & y_2 & y_3 \\ -y_1 & y_0 & -y_3 & y_2 \\ -y_2 & y_3 & y_0 & -y_1 \\ -y_3 & y_2 & -y_1 & y_0 \end{bmatrix} = \varphi(x)^T \mathcal{M}_R(y).$$

Note that

$$\mathcal{M}_R(y) = y_0 P_{:0} + y_1 P_{:1} + y_2 P_{:2} + y_3 P_{:3},$$

where $P_{:0} = \mathbf{I}_{4\times 4}$ is the identity matrix and

$$P_{:1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, P_{:2} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, P_{:3} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

Parametrized “Hypercomplex” Algebra

Recently, Zhang et al. introduced the so-called *parametrized “hypercomplex” algebra* (Grassucci et al., 2022; Zhang et al., 2021).

However, a parametrized “hypercomplex” algebra does not necessarily has an identity.

Accordingly, a parametrized “hypercomplex” algebra is defined as follows using the matrix representation of multiplication:

Given matrices $P_1, \ldots, P_n \in \mathbb{R}^{n \times n}$ and an ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$, the product of a parametrized “hypercomplex” algebra is defined by the equation

$$xy = \varphi^{-1}\Big(\sum_{i=1}^{n} x_i P_i \varphi(y)\Big), \tag{20}$$

Concluding Remarks:

We identify an algebra $\mathbb{V}$ with $\mathbb{R}^n$ equipped with a multiplication.

A hypercomplex algebra, denoted by $\mathbb{H}$, is a finite-dimensional algebra with a two-sided multiplication identity.

Given an ordered basis $\mathcal{E} = \{e_1, \ldots, e_n\}$, the multiplication is characterized by the multiplication table:

$$(e_i e_j) = \sum_{k=1}^{n} p_{ijk} e_k, \quad \forall i, j = 1, \ldots, n.$$

The multiplication satisfies

$$xy = \sum_{i=1}^{n} \mathcal{B}_k(x, y) e_k, \quad \forall x, y \in \mathbb{V},$$

where $\mathcal{B}_k : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ are bilinear forms.

The multiplication of a and x also satisfies

$$\varphi(ax) = \mathcal{M}_L(a)\varphi(x) \quad \text{and} \quad \varphi(xa)^T = \varphi(x)^T \mathcal{M}_R(a), \quad \forall x \in \mathbb{V}.$$

where

$$\mathcal{M}_L(a) = \sum_{i=1}^{n} a_i P_{i:}^T, \quad \text{with} \quad P_{i:}^T = \begin{bmatrix} p_{i11} & p_{i21} & \cdots & p_{in1} \\ p_{i12} & p_{i22} & \cdots & p_{in2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{i1n} & p_{i2n} & \cdots & p_{inn} \end{bmatrix},$$

and

$$\mathcal{M}_R(a) = \sum_{i=1}^{n} a_i P_{:i}, \quad \text{with} \quad P_{:i} = \begin{bmatrix} p_{1i1} & p_{1i2} & \cdots & p_{1in} \\ p_{2i1} & p_{2i2} & \cdots & p_{2in} \\ \vdots & \vdots & \ddots & \vdots \\ p_{ni1} & p_{ni2} & \cdots & p_{nin} \end{bmatrix}.$$

Thanks for your attention!

Acknowledge

These slides are part of a mini-course given during the workshop on **hypercomplex-valued neural networks**, which took place at the **Institute for Research and Applications of Fuzzy Modeling**, **University of Ostrava**, Ostrava, Checz Republic, *06-10 February 2023*, with the support of

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

My research on hypercomplex-valued neural networks has been partially supported by:

FAPESP and CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico

References (1)

P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia. Multilayer perceptrons to approximate quaternion valued functions. *Neural Networks*, 10(2):335–342, 3 1997. doi: 10.1016/S0893-6080(96)00048-2.

F. Catoni, D. Boccaletti, R. Cannata, V. Catoni, E. Nichelatti, and P. Zampetti. *The Mathematics of Minkowski Space-Time*. Birkhäuser Basel, 2008. doi: 10.1007/978-3-7643-8614-6.

E. Grassucci, A. Zhang, and D. Comminiello. PHNNs: Lightweight Neural Networks via Parameterized Hypercomplex Convolutions. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–13, 10 2022. ISSN 2162-237X. doi: 10.1109/TNNLS.2022.3226772. URL https://ieeexplore.ieee.org/document/9983846/.

References (2)

I. L. Kantor and A. S. Solodovnikov. *Hypercomplex Numbers: An Elementary Introduction to Algebras*. Springer New York, 1989. doi: 10.1007/978-1-4612-3650-4.

T. Parcollet, M. Morchid, and G. Linarès. A survey of quaternion neural networks. *Artificial Intelligence Review*, 53(4):2957–2982, 4 2020. doi: 10.1007/s10462-019-09752-1.

R. Schafer. *An Introduction to Nonassociative Algebras*. Project Gutenberg, 1961. URL `https://www.gutenberg.org/ebooks/25156`.

A. Zhang, Y. Tay, S. Zhang, A. Chan, A. T. Luu, S. C. Hui, and J. Fu. Beyond Fully-Connected Layers with Quaternions: Parameterization of Hypercomplex Multiplications with $1/n$ Parameters. 2 2021. doi: 10.48550/arxiv.2102.08597. URL `https://arxiv.org/abs/2102.08597v1`.