DOI: 10.37943/20ERZJ2964

Bimurat Sagindykov
Ph.D. in Physics and Mathematics, associate professor, department of Higher Mathematics and Modeling
b.sagindykov@satbayev.university, orcid.org/0000-0002-5349-1961
Satbayev University, Kazakhstan

Zhanar Bimurat
Ph.D. in Information systems, researcher scientist, R&D
bimuratzhanar@gmail.com orcid.org/0000-0001-8283-898X
D.A. Kunayev Mining Institute, Kazakhstan

IMPLEMENTATION OF THE ALGEBRA OF HYPERDUAL NUMBERS IN NEURAL NETWORKS

Abstract: For the numerical solution of problems arising in various fields of mathematics and mechanics, it is often necessary to determine the values of derivatives included in the model. Currently, numerical values of derivatives can be obtained using automatic differentiation libraries in many programming languages. This paper discusses the use of the Python programming language, which is widely used in the scientific community. It should be noted that the principles of automatic differentiation are not related to numerical or symbolic differentiation methods. The work consists of three parts. The introduction reviews the historical development of the general theory of complex numbers and the use of simple complex, double and dual numbers, which are a subset of the set of general complex numbers, in various fields of mathematics. The second part is devoted to the algebra of dual and hyperdual numbers and their properties. This section presents tables of the basis element of elementary functions with dual and hyperdual arguments, based on multiplication rules. Two important formulas for finding the numerical values of a complex function's first and second derivatives by expanding functions with dual and hyperdual arguments in the Taylor series are also obtained. A simple test function was used to verify the correctness of these formulas, the results of which were checked analytically as well as through implementation in a programming language. The third part of the paper focuses on practical applications and the implementation of these methods in Python. It includes detailed examples of case studies demonstrating the effectiveness of using hyperdual numbers in automatic differentiation. The results highlight the accuracy and computational efficiency of these methods, making them valuable tools for researchers and engineers. This comprehensive approach not only validates the theoretical aspects but also showcases the practical utility of dual and hyperdual numbers in solving complex mathematical and mechanical problems.

Keywords: dual numbers, hyperdual numbers, automatic differentiation, Taylor series expansion.

Literature review

The algebra of dual numbers was first proposed by W. Clifford in 1873 as an extension of real numbers [1]. The Cayley-Klein models used in theoretical calculation to study the structure of spaces in problems of mathematical physics require the application of the general algebra of complex numbers. Initially, the system of complex numbers was introduced to de-

Received: 03.08.2024 Accepted: 17.10.2024 Published: 30.10.2024

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

scribe the Cayley-Klein models [2], [3]. However, it has found other applications as well.

The use of dual and hyperdual numbers in automatic differentiation has been extensively explored in recent years. These mathematical constructs have proven to be highly effective in obtaining numerically exact derivatives, which are crucial for various computational and engineering applications.

Fike introduced the concept of hyperdual numbers and demonstrated their application in calculating numerically exact derivatives. This foundational work laid the groundwork for further developments in the field by providing a robust mathematical framework for derivative calculations using hyperdual numbers [4].

Building on this initial work, Fike and Alonso developed hyperdual numbers specifically for exact second-derivative calculations. Their research, presented at the New Horizons Forum and Aerospace Exposition, delved into the theoretical underpinning and practical applications of hyperdual numbers in engineering problems. The authors highlighted the accuracy and efficiency of this method, making it a valuable tool for engineers and researchers [5].

Further advancements were made by Fike and Alonso in their publication on automatic differentiation through the use of hyperdual numbers for second derivatives. This study, published in Computational Science and Engineering focused on the implementation of hyperdual numbers in computational algorithms. The authors emphasized the advantages of hyperdual numbers over traditional differentiation methods, particularly in terms of accuracy and computational efficiency [6].

In addition to these theoretical developments, the practical applications of hyperdual numbers in optimization problems have also been explored. Fike, Jogsma, Alonso and van der Weida demonstrated how gradient and Hessian information can be efficiently computed using hyperdual numbers. Their research, presented at the 29th AIAA Applied Aerodynamics Conference, provided valuable insights into the benefits of hyperdual numbers in aerodynamic optimization [7].

Neuenhofen reviewed the theory and implementation of hyperdual numbers for first and second-order automatic differentiation. This review highlighted the correctness of hyperdual number implementation in various programming environments and their application in numerical optimization, sensitivity analysis, and risk analysis [8].

The algebra of dual and hyperdual numbers has also been discussed in the context of generalized complex numbers. These numbers, which include ordinary complex numbers, double numbers, and dual numbers, are used to carry first-derivative information in their non-real parts. However, for second-derivative information, higher-dimensional extensions like hyperdual numbers are necessary.

The role of information technology (IT) in enhancing these methodologies cannot be overstated. IT has revolutionized the way automatic differentiation is implemented and utilized in various fields. For instance, the integration of IT in research methodologies has enabled more efficient data processing, storage, and analysis, thereby accelerating the pace of innovation and discovery [9]. The systematic review by Mamonov and Peterson highlights how IT investments and capabilities are crucial for organizational innovation, emphasizing the need for granular insights into the role of IT in innovation processes [10].

Moreover, the use of IT in research has evolved significantly, with advancements in artificial intelligence (AI) and machine learning (ML) playing a pivotal role. AI and ML have been instrumental in automating complex calculations and simulations, which are essential for the practical application of dual and hyperdual numbers in automatic differentiation [11]. The empirical study by Chubb, Cowling and Reed explores the impact of AI on research practices, highlighting its potential to enhance data analysis and streamline research processes [12].

In summary, dual and hyperdual numbers have been effectively utilized in automatic differ-

DOI: 10.37943/20ERZJ2964

entiation problems. The works of Fike and colleagues collectively highlight the development, implementation, and application of hyperdual numbers in various computational and engineering contexts. These studies underscore the importance of hyperdual numbers in achieving numerically exact derivatives and optimizing complex engineering systems. Additionally, the integration of IT has played a crucial role in advancing these methodologies, enabling more efficient and accurate computations.

In many areas of physics and engineering, it is important to find the numerical value of derivatives. An even more significant task is to calculate the exact values of the derivatives included in the models. To implement automatic differentiation using dual numbers in a programming language, it is necessary to define the type of dual numbers, as well as the corresponding arithmetic operations and elementary functions. This is especially relevant for object-oriented languages that support function overloading.

Methods and materials

Algebra of Dual and Hyperdual Numbers

This section presents a brief overview of the concepts of dual and hyperdual numbers.

1 Dual Numbers

The set of all dual numbers is defined as follows:

$$\mathbb{D} = \{A = a + \varepsilon a^* : a, a^* \in \mathbb{R}\}.$$

Here, the symbol ε represents the dual unit that satisfies conditions $\varepsilon \neq 0$, $\varepsilon^2 = 0$ and the equality $r\varepsilon = \varepsilon r$ for all values of $r \in \mathbb{R}$ [13].

The addition and multiplication of dual numbers $A = a + \varepsilon a^*$ and $B = b + \varepsilon b^*$ are defined by the equalities:

$$A + B = (a + b) + \varepsilon(a^* + b^*),$$
$$AB = ab + \varepsilon(ab^* + a^*b).$$

The multiplicative inverse of a dual number $A = a + \varepsilon a^*$ is defined by the equality:

$$A^{-1} = \frac{1}{a} - \varepsilon\frac{a^*}{a^2},\ a \neq 0.$$

At the same time, this formula implies that the dual number $A = a + \varepsilon a^*$ does not have a multiplicative inverse if $a = 0$.

For $a > 0$, the square root of a dual number $A = a + \varepsilon a^*$ is defined as follows:

$$\sqrt{A} = \sqrt{a} + \varepsilon\frac{a^*}{2\sqrt{a}}.$$

The absolute value of a dual number $A = a + \varepsilon a^*$ is defined by the equality $|A| = a$.

The conjugate of a dual number $A = a + \varepsilon a^*$ is defined by the equality: $\overline{A} = a - \varepsilon a^*$. Then, the modulus of a dual number can be negative. It can be defined by the following formula:

$$|A| = \frac{1}{2}\left(A + \overline{A}\right) = \frac{1}{2}(a + \varepsilon a^* + a - \varepsilon a^*) = a.$$

The division of dual numbers A and B is defined by the formula:

$$\frac{A}{B} = \frac{a + \varepsilon a^*}{b + \varepsilon b^*} = \frac{(a + \varepsilon a^*)(b - \varepsilon b^*)}{(b + \varepsilon b^*)(b - \varepsilon b^*)} = \frac{a}{b} + \varepsilon\frac{a^*b - ab^*}{b^2}$$

provided that $|B| \neq 0$.

From the properties of dual numbers considered above, it is clear that their imaginary parts do not affect the real parts of these numbers.

Consider a dual number of the form $C = 0 + \varepsilon c^*$, where $c^* \in \mathbb{R}$. This dual number has a zero

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

absolute value. For such numbers, the following equality holds:

$$(\varepsilon c^*)\cdot(\varepsilon d^*) = c^* d^* \varepsilon^2 = 0.$$

Such numbers are also called zero divisors [14].

The expansion of a hyperdual function $f(x+\varepsilon c^*)$ in a Taylor series around the point $x+\varepsilon c^* = a+\varepsilon a^*$can be expressed as follows [15]:

$$f(a+\varepsilon a^*) = f(a) + \varepsilon a^* f'(a). \tag{1}$$

If in this formula $a^* = 0$, then the derivative of the function $f(x)$ is determined by the equality:

$$f'(x) = f'(x+\varepsilon 0) = \frac{d}{dx} f(x).$$

The expansion of elementary functions with dual arguments in a Taylor series is as follows:

Exponent:

$exp(x+\varepsilon x^*) = e^x(1+\varepsilon x^*).$

Trigonometric functions:

$sin(x+\varepsilon x^*) = sinx + \varepsilon x^* cosx;$
$cos(x+\varepsilon x^*) = cosx - \varepsilon x^* sinx;$
$tg(x+\varepsilon x^*) = tgx + \varepsilon \frac{x^*}{cos^2 x};$
$ctg(x+\varepsilon x^*) = ctgx - \varepsilon \frac{x^*}{sin^2 x}.$

Inverse trigonometric functions:

$arcsin(x+\varepsilon x^*) = arcsinx + \varepsilon \frac{x^*}{\sqrt{1-x^2}};$
$arccos(x+\varepsilon x^*) = arccosx - \varepsilon \frac{x^*}{\sqrt{1-x^2}};$
$arctg(x+\varepsilon x^*) = arctgx + \varepsilon \frac{x^*}{1+x^2};$
$arcctg(x+\varepsilon x^*) = arcctgx - \varepsilon \frac{x^*}{1+x^2}.$

Logarithmic functions:

$log_a(x+\varepsilon x^*) = log_a x + \varepsilon \frac{x^*}{xlna};$
$ln(x+\varepsilon x^*) = lnx + \varepsilon \frac{x^*}{x}.$

2 Hyperdual Numbers

It is considered that the hyperdual numbers as extensions of dual numbers. The set of all hyperdual numbers is defined by the following formula:

$$\widehat{D} = \left\{\hat{A} = a_0 + \varepsilon_1 a_1 + \varepsilon_2 a_2 + \varepsilon_1\varepsilon_2 a_3 : \ a_0, a_1, a_2, a_3 \in \mathbb{R}\right\}.$$

Additionally, the dual units ε_1 and ε_2 satisfy the following conditions:

$$\varepsilon_1^2 = \varepsilon_2^2 = (\varepsilon_1\varepsilon_2)^2 = 0 \ \text{ and } \ \varepsilon_1, \varepsilon_2, \varepsilon_1\varepsilon_2 \neq 0.$$

The commutative law between the imaginary dual units and real numbers holds: $r\varepsilon = \varepsilon r$, where $r \in \mathbb{R}$.

Sometimes, dual units are also called the bases of the imaginary parts of a hyperdual number. The basis of the real number a_0 is 1. The dual units ε_1, ε_2 and $\varepsilon_1\varepsilon_2$ of the same index form the primary basis, whereas the dual unit $\varepsilon_1\varepsilon_2$, which is an extension of these dual units, forms the secondary basis.

The properties of the dual units ε_1, ε_2 and $\varepsilon_1\varepsilon_2$ are determined by their multiplication table.

Table 1: Multiplication Rules for Dual Units

x	1	ε_1	ε_2	$\varepsilon_1\varepsilon_2$
1	1	ε_1	ε_2	$\varepsilon_1\varepsilon_2$
ε_1	ε_1	0	$\varepsilon_1\varepsilon_2$	0
ε_2	ε_2	$\varepsilon_2\varepsilon_1$	0	0
$\varepsilon_1\varepsilon_2$	$\varepsilon_1\varepsilon_2$	0	0	0

The addition and multiplication of hyperdual numbers $\mathbb{A} = a_0 + \varepsilon_1 a_1 + \varepsilon_2 a_2 + \varepsilon_1\varepsilon_2 a_3$ and $\mathbb{B} = b_0 + \varepsilon_1 b_1 + \varepsilon_2 b a_2 + \varepsilon_1\varepsilon_2 b_3$ are defined by the formulas:

$$\mathbb{A} + \mathbb{B} = (a_0 + b_0) + \varepsilon_1(a_1 + b_1) + \varepsilon_2(a_2 + b_2) + \varepsilon_1\varepsilon_2(a_3 + b_3),$$
$$\mathbb{A} \cdot \mathbb{B} = (a_0 \cdot b_0) + \varepsilon_1(a_0 b_1 + a_1 b_0) + \varepsilon_2(a_0 b_2 + a_2 b_0) +$$
$$+\varepsilon_1\varepsilon_2(a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0).$$

To determine the inverse of a hyperdual number $\mathbb{A} = a_0 + \varepsilon_1 a_1 + \varepsilon_2 a_2 + \varepsilon_1\varepsilon_2 a_3$ in the form $\mathbb{A}^{-1} = y_0 + \varepsilon_1 y_1 + \varepsilon_2 y_2 + \varepsilon_1\varepsilon_2 y_3$, consider the equality:

$$\mathbb{A} \cdot \mathbb{A}^{-1} = \mathbb{E},$$

where $\mathbb{E} = 1 + \varepsilon_1 0 + \varepsilon_2 0 + \varepsilon_1\varepsilon_2 0$.

From the properties of hyperdual numbers and considering the equality $\mathbb{A} \cdot \mathbb{A}^{-1} = \mathbb{E}$, the inverse of a hyperdual number $\mathbb{A} = a_0 + \varepsilon_1 a_1 + \varepsilon_2 a_2 + \varepsilon_1\varepsilon_2 a_3$ is determined as follows:

$$\mathbb{A}^{-1} = \frac{1}{\mathbb{A}} = \frac{1}{a_0} - \varepsilon_1\frac{a_1}{a_0^2} - \varepsilon_2\frac{a_2}{a_0^2} + \varepsilon_1\varepsilon_2\left(-\frac{a_3}{a_0^2} + \frac{2a_1a_2}{a_0^3}\right), \quad (2)$$

where $a_0 \neq 0$.

Additionally, this formula implies that a hyperdual number $\mathbb{A} = 0 + \varepsilon_1 a_1 + \varepsilon_2 a_2 + \varepsilon_1\varepsilon_2 a_3$ cannot have a multiplicative inverse. In other words, any hyperdual number with a non-zero real part is invertible.

Using formula (2), the division operation can be derived for hyperdual numbes. For hyperdual numbers

$$\mathbb{B} = b_0 + \varepsilon_1 b_1 + \varepsilon_2 b_2 + \varepsilon_1\varepsilon_2 b_3 \text{ and } \mathbb{A} = a_0 + \varepsilon_1 a_1 + \varepsilon_2 a_2 + \varepsilon_1\varepsilon_2 a_3$$

the division operation is expressed as:

$$\frac{\mathbb{B}}{\mathbb{A}} = \mathbb{B} \cdot \mathbb{A}^{-1} = (b_0 + \varepsilon_1 b_1 + \varepsilon_2 b_2 + \varepsilon_1\varepsilon_2 b_3) \cdot$$
$$\cdot\left(\frac{1}{a_0} - \varepsilon_1\frac{a_1}{a_0^2} - \varepsilon_2\frac{a_2}{a_0^2} + \varepsilon_1\varepsilon_2\left(-\frac{a_3}{a_0^2} + \frac{2a_1a_2}{a_0^3}\right)\right) =$$
$$= \frac{b_0}{a_0} + \left(\frac{b_1}{a_0} - \frac{b_0a_1}{a_0^2}\right)\varepsilon_1 + \left(\frac{b_2}{a_0} - \frac{b_0a_2}{a_0^2}\right)\varepsilon_2 + \quad (3)$$
$$+\left(b_0\left(\frac{2a_1a_2}{a_0^3} - \frac{a_3}{a_0^2}\right) - \frac{b_1a_2}{a_0^2} - \frac{b_2a_1}{a_0^2} + \frac{b_3}{a_0}\right)\varepsilon_1\varepsilon_2,$$
$$a_0 \neq 0.$$

Hyperdual numbers have the following properties concerning addition and multiplication:

1. Commutativity of addition:

$$\mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A}.$$

2. Associativity of addition:

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

$$\mathbb{A} + (\mathbb{B} + \mathbb{C}) = (\mathbb{A} + \mathbb{B}) + \mathbb{C}.$$

3. Existence of a zero hyperdual number:

There exists a zero hyperdual number $\mathbb{O} = 0 + \varepsilon_1 0 + \varepsilon_2 0 + \varepsilon_1\varepsilon_2 0$ such that $\mathbb{A} + \mathbb{O} = \mathbb{A}$.

4. Additive inverse property:

For every hyperdual number $\mathbb{A}$, there exists an additive inverse $-\mathbb{A}$ such that $\mathbb{A} + (-\mathbb{A}) = (-\mathbb{A}) + \mathbb{A} = \mathbb{O}$.

5. Commutativity of multiplication:

$$\mathbb{A} \cdot \mathbb{B} = \mathbb{B} \cdot \mathbb{A}.$$

6. Associativity of multiplication:

$$(\mathbb{A} \cdot \mathbb{B}) \cdot \mathbb{C} = \mathbb{A} \cdot (\mathbb{B} \cdot \mathbb{C}).$$

7. Existence of a multiplicative identity:

There exists a multiplicative identity $\mathbb{E} = 1 + \varepsilon_1 0 + \varepsilon_2 0 + \varepsilon_1\varepsilon_2 0$, such that:

$$\mathbb{A} \cdot \mathbb{E} = \mathbb{E} \cdot \mathbb{A}.$$

Since the singular hyperdual number $\mathbb{E}$ corresponds to the real number 1, multiplying any hyperdual number $\mathbb{A}$ by $\mathbb{E}$ does not change $\mathbb{A}$, and therefore this equality is always true.

8. Distributive properties:

a) Left distributivity: $\mathbb{A} \cdot (\mathbb{B} + \mathbb{C}) = \mathbb{A} \cdot \mathbb{B} + \mathbb{A} \cdot \mathbb{C}$

b) Right distributivity: $(\mathbb{A} + \mathbb{B}) \cdot \mathbb{C} = \mathbb{A} \cdot \mathbb{C} + \mathbb{B} \cdot \mathbb{C}$.

These distributivity properties hold for hyperdual numbers in the same way as for ordinary numbers.

To verify correctness of formula (2) and (3), we introduce the conjugate hyperdual number:

$$\overline{\mathbb{A}} = a_0 - \varepsilon_1 a_1 - \varepsilon_2 a_2 - \varepsilon_1\varepsilon_2 a_3.$$

Then $\mathbb{A} \cdot \overline{\mathbb{A}} = a_0^2 - 2\varepsilon_1\varepsilon_2 a_1 a_2$ gives a simple dual number. Therefore:

$$\mathbb{A}^{-1} = \frac{1}{\mathbb{A}} = \frac{\overline{\mathbb{A}}}{\mathbb{A} \cdot \overline{\mathbb{A}}} = \frac{a_0 - \varepsilon_1 a_1 - \varepsilon_2 a_2 - \varepsilon_1\varepsilon_2 a_3}{a_0^2 - 2\varepsilon_1\varepsilon_2 a_1 a_2} =$$

$$= \frac{(a_0 - \varepsilon_1 a_1 - \varepsilon_2 a_2 - \varepsilon_1\varepsilon_2 a_3)(a_0^2 + 2\varepsilon_1\varepsilon_2 a_1 a_2)}{(a_0^2 - 2\varepsilon_1\varepsilon_2 a_1 a_2)(a_0^2 + 2\varepsilon_1\varepsilon_2 a_1 a_2)} =$$

$$= \frac{1}{a_0} - \varepsilon_1 \frac{a_1}{a_0^2} - \varepsilon_2 \frac{a_2}{a_0^2} + \varepsilon_1\varepsilon_2 \left(-\frac{a_3}{a_0^2} + \frac{2a_1 a_2}{a_0^3} \right),$$

Here, $a_0 \neq 0$. Thus, we have proven the validity of formula (2).

The modulus of a hyperdual number is determined by the formula:

$$|\mathbb{A}| = \sqrt{(a_0^2 - 2\varepsilon_1\varepsilon_2 a_1 a_2)(a_0^2 + 2\varepsilon_1\varepsilon_2 a_1 a_2)} = \sqrt{a_0^4} = a_0^2.$$

From this, it is concleded that the comparison of two hyperdual numbers should be based only on the real parts of these hyperdual numbers.

Therefore, to determine if two hyperdual numbers are equal, it is sufficient to compare their real parts and the coefficients of the dual units.

3 Expansion of Elementary Hyperdual Functions in Taylor Series

Consider a hyperdual differentiable function:

$$f(x) = p(x) + \varepsilon_1 q(x) + \varepsilon_2 r(x) + \varepsilon_1\varepsilon_2 s(x), \tag{4}$$

where $x = x_0 + \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_1\varepsilon_2 x_3$.

For complex-valued differentiable functions, the Cauchy-Riemann conditions provide the necessary and sufficient conditions for complex differentiability. Similarly, for hyperdual functions, analogous conditions to the Cauchy-Riemann conditions can by defined.

The derivative of the function f at the point $x = x_0 + \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_1\varepsilon_2 x_3$ defined by the formula

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

To determine the derivative of the hyperdual function $f(x)$ in different incremental directions, as outlined in [16], four cases of argument increments are considered. The values of the function $f(x)$ derivative in these directions is determined as follows:

1) increment $\Delta x = \Delta x_0$:

$$f'(x) = \frac{\partial p(x)}{\partial x_0} + \varepsilon_1 \frac{\partial q(x)}{\partial x_0} + \varepsilon_2 \frac{\partial r(x)}{\partial x_0} + \varepsilon_1\varepsilon_2 \frac{\partial s(x)}{\partial x_0}.$$

2) increment $\Delta x = \varepsilon_1 x_1$:

$$\varepsilon_1 f'(x) = \frac{\partial p(x)}{\partial x_1} + \varepsilon_1 \frac{\partial q(x)}{\partial x_1} + \varepsilon_2 \frac{\partial r(x)}{\partial x_1} + \varepsilon_1\varepsilon_2 \frac{\partial s(x)}{\partial x_1}.$$

3) increment $\Delta x = \varepsilon_2 x_2$:

$$\varepsilon_2 f'(x) = \frac{\partial p(x)}{\partial x_2} + \varepsilon_1 \frac{\partial q(x)}{\partial x_2} + \varepsilon_2 \frac{\partial r(x)}{\partial x_2} + \varepsilon_1\varepsilon_2 \frac{\partial s(x)}{\partial x_2}.$$

4) increment $\Delta x = \varepsilon_1\varepsilon_2 \Delta x_3$:

$$\varepsilon_1\varepsilon_2 f'(x) = \frac{\partial p(x)}{\partial x_3} + \varepsilon_1 \frac{\partial q(x)}{\partial x_3} + \varepsilon_2 \frac{\partial r(x)}{\partial x_3} + \varepsilon_1\varepsilon_2 \frac{\partial s(x)}{\partial x_3}.$$

In [17] various combinations were created from the equations obtained for the values of the derivatives of hyperdual functions in different directions of increments. These combinations help in constructing a more general understanding of the derivatives and allow for consideration of all components of the hyperdual number. As a result, for a differentiable hyperdual function $f(x)$ with a hyperdual argument, the expansion is obtained in the form:

$$\begin{aligned} f(x_0 + \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_1\varepsilon_2 x_3) = p(x_0) + \varepsilon_1 \left(x_1 \frac{dp(x_0)}{dx_0} + \varphi_1(x_0) \right) + \\ + \varepsilon_2 \left(x_2 \frac{dp(x_0)}{dx_0} + \varphi_2(x_0) \right) + \varepsilon_1\varepsilon_2 \left(x_1 x_2 \frac{d^2 p(x_0)}{dx_0^2} + x_3 \frac{dp(x_0)}{dx_0} + \right. \\ \left. + x_1 \frac{d\varphi_2(x_0)}{dx_0} + x_2 \frac{d\varphi_1(x_0)}{dx_0} + \varphi_3(x_0) \right). \end{aligned} \quad (5)$$

4 Taylor Series for Elementary Functions

Based on the structure of equation (5), the hyperdual function $f(x)$ can be expanded into a Taylor series using the equation:

$$f(x) = f(x_0) + x_1 f'(x_0)\varepsilon_1 + x_2 f'(x_0)\varepsilon_2 + \big(x_3 f'(x_0) + x_1 x_2 f''(x_0)\big)\varepsilon_1\varepsilon_2, \quad (6)$$

where $x = x_0 + \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_1\varepsilon_2 x_3$.

In this expansion, the real part $f(x_0)$ remains unchanged, meaning the real part behaves as it does in the real code. These actions do not affect the real part, regardless of whether the imaginary parts are combined or not.

The validity of equation (6) can be inferred from the Taylor series expansion of the function $f(x)$ around the point x_0:

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

$$f(x_0+\Delta v)=f(x_0)+\Delta v f'(x_0)+\frac{1}{2!}\Delta v^2 f''(x_0)+\frac{1}{3!}\Delta v^3 f'''(x_0)+\cdots,$$

where $\Delta v = x_1\varepsilon_1 + x_2\varepsilon_2 + x_3\varepsilon_3$ is the increment of the point x_0 composed of imaginary dual units.

Using equation (6), it is considered the Taylor series expansion of elementary functions with hyperdual arguments:

Exponential function:

$$e^x = exp(x_0+x_1\varepsilon_1+x_2\varepsilon_2+x_3\varepsilon_1\varepsilon_2)=e^{x_0}\big(1+x_1\varepsilon_1+x_2\varepsilon_2+(x_3+x_1x_2)\big).$$

Hyperdual numbers can be used for accurate computation of the first and second derivatives of a function. To verify the validity and accuracy of equation (6) using hyperdual numbers, it can be used a simple analytical test function whose derivatives are easily computed manually.

For this purpose, it is considered a special case of equation (6) where $x_1=x_2=1$ and $x_3=0$. Then, equation (6) can be written as:

$$f(x)=f(x_0)+f'(x_0)\varepsilon_1+f'(x_0)\varepsilon_2+f(x_0)\varepsilon_1\varepsilon_2{}''. \quad (7)$$

Now, let's take the function $f(x)=\frac{sinx}{x}$ as a test function. The first and second derivatives of this function at the point x_0 are given by:

$$f'(x_0)=\frac{cosx_0}{x_0}-\frac{sinx_0}{x_0^2} \text{ and } f''(x_0)=-\frac{sinx_0}{x_0}-\frac{2cosx_0}{x_0^2}+\frac{2sinx_0}{x_0^3}.$$

Next, it is determined by the values of $f'(x_0)$ and $f''(x_0)$ using hyperdual numbers. For this, it is used the Taylor series expansion of the functions $sin\,x$ and $\frac{1}{x}$. Then

$$\frac{sinx}{x}=\frac{sinx_0}{x_0}+\left(\frac{x_1}{x_0}cosx_0-\frac{x_1}{x_0^2}sinx_0\right)\varepsilon_1+\left(\frac{x_2}{x_0}cosx_0-\frac{x_2}{x_0^2}sinx_0\right)\varepsilon_2+$$
$$+\left(\frac{x_3}{x_0}cosx_0-\frac{x_1x_2}{x_0}sinx_0-\frac{2x_1x_2}{x_0^2}cosx_0-\frac{x_3}{x_0^2}sinx_0+\frac{2x_1x_2}{x_0^3}sinx_0\right)\varepsilon_1\varepsilon_2,$$

where $x=x_0+\varepsilon_1x_1+\varepsilon_2x_2+\varepsilon_1\varepsilon_2x_3$.

The last equation, when $x_1=x_2=1$ and $x_3=0$ gives the expression:

$$\frac{sinx}{x}=\frac{sinx_0}{x_0}+\left(\frac{cosx_0}{x_0}-\frac{sinx_0}{x_0^2}\right)\varepsilon_1+\left(\frac{cosx_0}{x_0}-\frac{sinx_0}{x_0^2}\right)\varepsilon_2+$$
$$+\left(-\frac{sinx_0}{x_0}-\frac{2cosx_0}{x_0^2}+\frac{2sinx_0}{x_0^3}\right)\varepsilon_1\varepsilon_2. \quad (8)$$

Comparing equation (8) with equation (7), we see that the first and second derivatives of the function $f(x)=\frac{sinx}{x}$ are exactly unchanged.

Next, we consider the expression of the function $f(x)=\frac{sinx}{x}$ written in computer code. To find the first and second derivatives of this function, it is used a method called automatic differentiation.

Results

By transforming the program for computing derivatives of complex-valued functions presented in [17], it will be demonstrated the use of hyperdual numbers for automatic differentiation of the function $\frac{\sin(x)}{x}$.

1. Library Imports

```
import numpy as np
import matplotlib.pyplot as plt
import cmath
```

These libraries are necessary for performing mathematical operations, plotting graphs, and working with complex numbers.

2. Class HyperDualCplx

This class describes hyperdual numbers with components f_0, f_1, f_2 and f_{12}, analogous to the components of the hyperdual number $x = x_0 + \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_1 \varepsilon_2 x_3$ used for automatic differentiation.

```
class HyperDualCplx:
    def __init__(self, f0=0+0j, f1=0+0j, f2=0+0j, f12=0+0j):
        self.f0 = f0
        self.f1 = f1
        self.f2 = f2
        self.f12 = f12
```

3. Defining Addition and Multiplication Methods

Methods for the addition and multiplication of hyperdual numbers are defined, allowing them to be easily combined in computations.

```
def __add__(self, rhs):
    temp = HyperDualCplx()
    temp.f0 = self.f0 + rhs.f0
    temp.f1 = self.f1 + rhs.f1
    temp.f2 = self.f2 + rhs.f2
    temp.f12 = self.f12 + rhs.f12
    return temp

def __mul__(self, rhs):
    temp = HyperDualCplx()
    temp.f0 = self.f0 * rhs.f0
    temp.f1 = self.f0 * rhs.f1 + self.f1 * rhs.f0
    temp.f2 = self.f0 * rhs.f2 + self.f2 * rhs.f0
    temp.f12 = self.f0 * rhs.f12 + self.f1 * rhs.f2 + self.f2 * rhs.f1 + self.f12 * rhs.f0
    return temp
```

4. Defining Functions for $sin(x)$ and $^1/_x$ of Hyperdual Numbers

```
def sin(x):
    temp = HyperDualCplx()
    funval = cmath.sin(x.f0)
    deriv = cmath.cos(x.f0)
    temp.f0 = funval
    temp.f1 = deriv * x.f1
    temp.f2 = deriv * x.f2
    temp.f12 = deriv * x.f12 - funval * x.f1 * x.f2
    return temp

def reciprocal(x):
    temp = HyperDualCplx()
    inv_f0 = 1.0 / x.f0
    temp.f0 = inv_f0
    temp.f1 = -inv_f0**2 * x.f1
    temp.f2 = -inv_f0**2 * x.f2
    temp.f12 = 2 * inv_f0**3 * x.f1 * x.f2 - inv_f0**2 * x.f12
    return temp
```

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

5. Function sin_over_x(x)

This function computes $\frac{\sin(x)}{x}$ for hyperdual numbers, automatically obtaining the function values and their derivatives for real numbers.

```
def sin_over_x(x):
    return sin(x) * reciprocal(x)
```

6. Computing Hyperdual Components of the Function $\frac{\sin(x)}{x}$

```
hyperdual_f0 = []
hyperdual_f1 = []
hyperdual_f12 = []

for x in x_values:
    h = HyperDualCplx(x+0j, 1+0j, 1+0j, 0+0j)
    result = sin_over_x(h)
    hyperdual_f0.append(result.f0.real)
    hyperdual_f1.append(result.f1.real)
    hyperdual_f12.append(result.f12.real)
```

Next, the derivatives of the function $\frac{\sin(x)}{x}$ are plotted, comparing analytical results with hyperdual computations, demonstrating the accuracy and utility of automatic differentiation using hyperdual numbers.

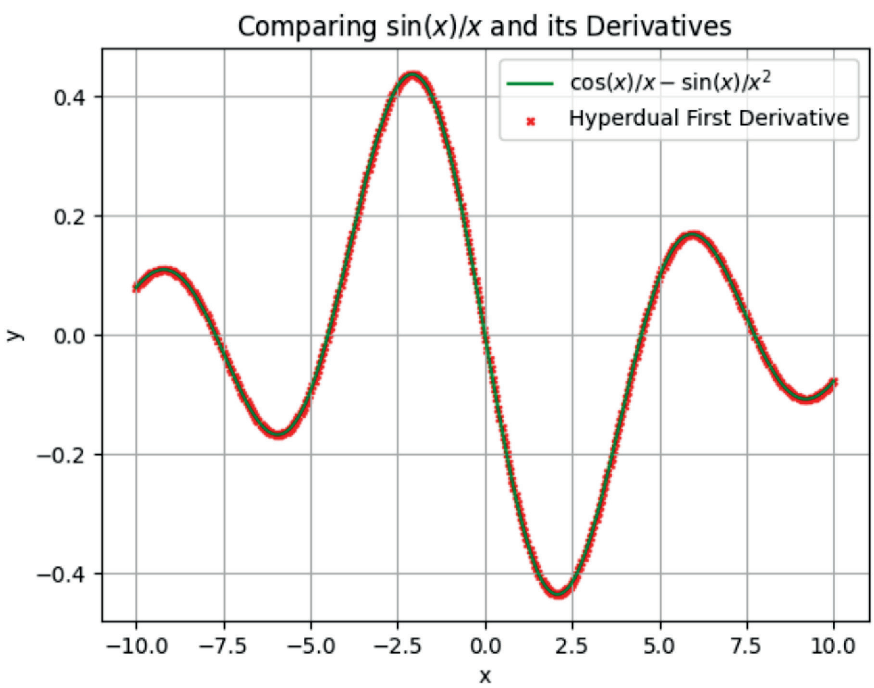

Figure 1. First derivative of the function $\frac{\sin(x)}{x}$

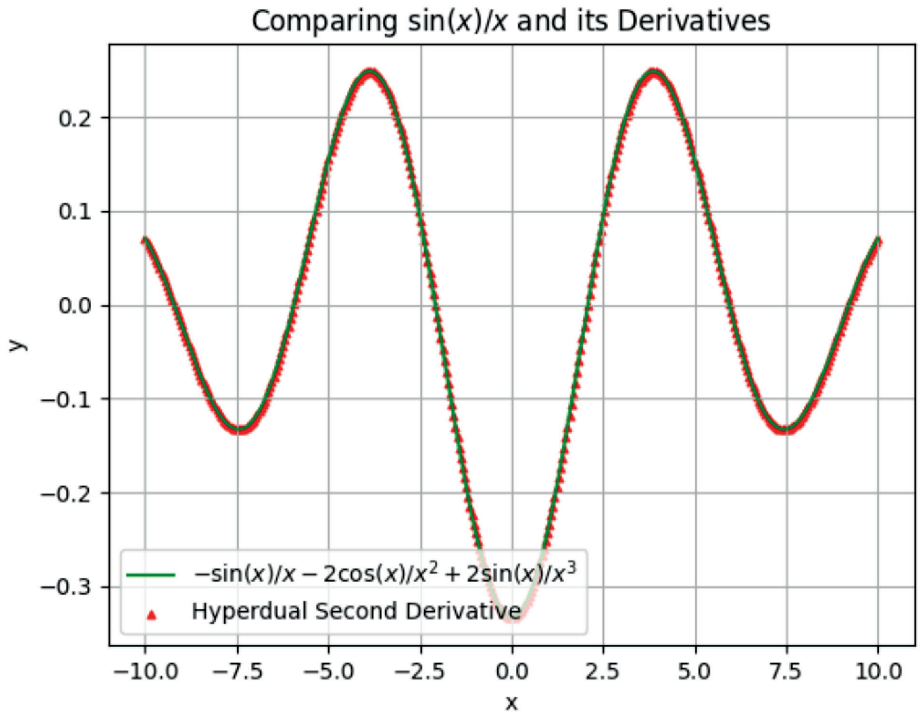

Figure 2. Second derivative of the function $\frac{\sin(x)}{x}$

Discussion

Hyperdual Numbers for Efficient Second-Order Optimization in Neural Networks

This paper proposes the application of hyperdual numbers to facilitate second-order optimization methods within neural networks. By extending neural network activation functions to support hyperdual numbers, both first and second derivatives can be efficiently computed within the forward pass. Unlike traditional automatic differentiation methods, which typically focus on first-order derivatives through forward or reverse mode differentiation, hyperdual numbers enable the calculation of second-order derivatives directly. This allows for the application of more advanced optimization techniques, such as Newton's method, which leverage curvature information for potentially faster and more accurate convergence.

1 Hyperdual Number Extension of Activation Functions

Let $h = a + \epsilon b + \delta c + \epsilon\delta d$ represent a hyperdual number, where a is the real component and b, c, and d are coefficients of the infinitesimal quantities ϵ and δ. For example, a hyperdual extension of the ReLU activation function could be defined as follows:

$$ReLU(h) = \begin{cases} a + \epsilon b + \delta c + \epsilon\delta d, & if\ a > 0 \\ 0, & if\ a \leq 0 \end{cases}$$

This extension enables each layer of the network to propagate not only the function value but also its first and second derivatives, essential for implementing second-order optimization. In this setup, the ReLU function can be programmed in Python to handle hyperdual numbers as shown below:

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 20, OCTOBER 2024

```
class Hyperdual:
    def __init__(self, a, b=0, c=0, d=0):
        self.a = a  # Real part
        self.b = b  # Coefficient of ε
        self.c = c  # Coefficient of δ
        self.d = d  # Coefficient of εδ

    def relu(self):
        if self.a > 0:
            return Hyperdual(self.a, self.b, self.c, self.d)
        else:
            return Hyperdual(0, 0, 0, 0)
```

2 Practical Implementation in PyTorch

To implement this approach within a deep learning framework like PyTorch, a hyperdual number layer could be created to support tensors that operate with hyperdual values. This layer would be specifically designed to propagate both first and second derivatives through hyperdual arithmetic. Below is a simplified example illustrating how a custom PyTorch layer might be constructed to incorporate hyperdual numbers for second-order derivative tracking:

```
import torch
from torch.nn import Module

class HyperdualLayer(Module):
    def __init__(self, in_features, out_features):
        super(HyperdualLayer, self).__init__()
        self.weight = Hyperdual(torch.randn(out_features, in_features))
        self.bias = Hyperdual(torch.randn(out_features))

    def forward(self, x):
        # Assume x is a hyperdual tensor as well
        return x @ self.weight + self.bias
```

Conclusion

In conclusion, the algebra of dual and hyperdual numbers represents a powerful tool for solving automatic differentiation problems, which are in demand in various fields of mathematics, mechanics, and applied sciences. The application of dual and hyperdual numbers allows for the efficient computation of derivatives of complex functions, significantly simplifying the solution of optimization and modeling problems.

The advantages of using dual numbers include the ability to accurately and quickly find the first derivative of a function, while hyperdual numbers also enable the computation of second derivatives with high precision. The principles of automatic differentiation, implemented using dual and hyperdual numbers, are independent of traditional numerical or symbolic differentiation methods, making them versatile and flexible for application in various computational tasks.

The effectiveness and accuracy of the proposed methods have been confirmed both analytically and in practice using a test function and its implementation in the Python programming language. This demonstrates the potential of dual and hyperdual numbers for broad application in scientific and engineering computations, opening new opportunities for researchers and developers.

Thus, the study of the algebra of dual and hyperdual numbers and their application in automatic differentiation not only deepens our understanding of this mathematical structure but also provides practical tools for solving complex problems in various fields of science and engineering. Furthermore, it has been demonstrated how to expand inverse trigonometric functions, natural logarithm, and functions such as $\frac{\sin x}{x}$ with dual arguments in Taylor series, which, to our knowledge, have not been explicitly explored in other works.

By effectively utilizing hyperdual numbers, it can be enhanced the optimization process in neural networks, particularly for models sensitive to curvature information. This approach offers the potential for faster convergence and improved performance.

References

[1] Clifford, W.K. (1873). Preliminary sketch of biquaternions. *Proceedings of the London Mathematical Society, 4(64)*, 381-395.
[2] Cayley, A. (1859). A sixth memoir upon quantis. *Philosophical Transactions of the Royal Society of London, 149*, 61-90.
[3] Klein, F. (1985). Ueber die sogenannte Nicht-Euklidische. *Springer-Verlag Wien, Teubner-Archiv zur Mathematik, 4*, 224-238.
[4] Fike, J.A. (2009). Numerically exact derivative calculations using hyper-dual numbers. *3rd Annual Student Joint Engineering and Design*.
[5] Fike, J.A., & Alonso, J.J. (2011). The development of hyper-dual numbers for exact second-derivative calculations. *Meeting Including the New Horizons Forum and Aerospace Exposition*, 4-7.
[6] Fike, J.A., & Alonso, J.J. (2011). Automatic differentiation through the use of hyper-dual numbers for second derivatives. *Computational Science and Engineering, 87(201)*, 163-173.
[7] Fike, J.A., Jongsma, S., Alonso, J.J., & van der Weida, E. (2011). Optimization with gradient and Hessian information using hyper-dual numbers. *29th AIAA Applied Aerodynamics Conference*.
[8] Neuenhofen, M.P. (2018). Review of theory and implementation of hyper-dual numbers for first and second order automatic differentiation. *arXiv*. https://doi.org/10.48550/arXiv.1801.03614
[9] Mamonov, S. & Peterson, R. (2021). The role of IT in organizational innovation – A systematic literature review. *Journal of Strategic Information Systems, 30(4)*, 101696.
[10] Mamonov, S. & Peterson, R. (2019). The role of IT in innovation at the individual and group level – a literature review. *Journal of Small Business and Enterprise Development, 26(6-7)*, 797-810.
[11] Boutet, A., Haile, S.S., Yang, A. Z., Son, H.J., Malik, M., Pai, V., Nasralla, M., Germann, J., Vetkas, A., & Ertl-Wagner, B.B. (2024). Assessing the Emergence and Evolution of Artificial Intelligence and Machine Learning Research in Neuroradiology. *American Journal of Neuroradiology, 45(9)*, 1269-1275.
[12] Chubb, J., Crowling, P., & Reed, D. (2021). Speeding up to keep up: exploring the use of AI in the research process. *AI & SOCIETY, 37*, 1439-1457.
[13] Kantor, I., & Solodovnikov, A. (1989). Hypercomplex numbers. *Springer-Verlag, New York*.
[14] Сағындықов, Б.Ж., & Бимұрат, Ж. (2024). Жалпы комплекс сандардың екі өлшемді алгебрасы. *Торайғыров университетінің хабаршысы, №1*, 89-101. Павлодар. ISSN 2959-068X.
[15] Бимұрат, Ж., & Сағындықов, Б.Ж. (2024). Дуал сандар негізінде автоматты дифференциалдау: әдістеме, мысалдар және іске асыру. *Радиоэлектроника және байланыс әскери-инженерлік институтының ғылыми еңбектері. Әскери ғылыми-техникалық журнал, №2(56)*, 155-164.
[16] Das, I., & Dennis, J.E. (1997). A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural Optimization, 14(1)*, 63-69.
[17] Fike, J.A. (2013). Multi-objective optimization using hyper-dual numbers (Doctoral dissertation, Stanford University). *Stanford Digital Repository*. https://purl.stanford.edu/jw107zn5044