КРАТКАЯ ИСТОРИЯ ДУАЛЬНЫХ И ГИПЕР-ДУАЛЬНЫХ ЧИСЕЛ

В элементарной алгебре наряду с действительными числами рассматривается и более широкая система комплексных чисел. История комплексных чисел начинается с XVI века. Итальянские математики Джироламо Кардано и Рафаэль Бомбелли, решая квадратные уравнения, ввели в рассмотрение символ -1 - формальное решение уравнения $x^2 = -1$. Впоследствии эти числа стали называться «мнимыми», а затем «комплексными» числами и записываться A = $a + bi$.

В последующем комплексные числа нашли широкое применение не только в самой математике, но и в физике, механике и многих других областях естествознания. Именно это обстоятельство послужило причиной поиска новых систем чисел, которые, являясь обобщением действительных и комплексных чисел, обладают если не всеми, то хотя бы частью основных свойств последних. Так возникли системы двойных и дуальных чисел. Существует три типа обобщенных комплексных чисел [1, 2, 3, 4]:

- комплексные числа (обычные) с $i^2 = -1$
- двойные числа (гиперболические) $i^2 = 1$, но $i \neq 1$
- дуальные числа (параболические) $i^2 = 0$, но $i \neq 0$

Дуальные числа были введены в 1871 году Уильям Клиффорд [1] , и были использованы в начале двадцатого века немецким математиком Эдуардом Штуди [5], который применял их для представления дуального угла, измеряющего относительное положение двух косых линий в пространстве. Дуальным числом или (гипер) комплексным числом параболического типа «по Клиффорду» называется комплексное число вида $A = a + b\omega,\ \omega^2 = 0$.

По-видимому, еще Л. Эйлер использовал нули различных порядков, например, $\omega = \sqrt{0}$, $\omega^2 = 0,\ 0 < \omega < \varepsilon$ (ε – любое положительное число) в исчислении бесконечно-малых величин. В современном нестандартном анализе [6] такие величины называются нестандартными или актуальными бесконечно малыми, они могут иметь разный порядок малости.

В конце XIX века А. П. Котельников [7] фактически использовал дуальные числа для развития теории винтов, которая нашла широкое применение в различных разделах физики, механики, геометрии [8]. Им был предложен принцип перенесения, который в упрощенном виде может быть сформулирован так: «Решение, полученное на множестве вещественных чисел, может быть перенесено в область дуальных чисел путем замены вещественных чисел на дуальные числа». Справедливости ради, похожие утверждения были использованы в работах У. Клиффорд, Э. Штуди и их последователей.

Уже в нашем веке с развитием вычислительной техники возникло новое направление применения дуальных чисел (точнее функций дуального аргумента) – автоматическое дифференцирование [9, 10], связанное с точным (точностью представления чисел в компьютерной системе) вычислением первых производных сложных математических функций. Автоматическое дифференцирование (АД) позволяет избежать дублирование функциональности программного кода (изменение кода функции не требует изменения кода ее производной). Для компьютерной реализации АД необходимо создать новый тип данных, перезагрузить базовые математические функции и операции над ними. Если новый тип данных строится на основе дуальных чисел, то за одно обращение к перезагруженной функции точно вычисляются значения самой функции и ее производной.

Учитывая, что многие градиентные методы вычислительной математики используют значения производных более высоких порядков, было предложено расширение дуальных чисел, получивших название гипер-дуальные числа [11, 12]. За одно обращение к функции гипер-дуального аргумента можно получить точные значения функции и ее первой и второй производных. Отправной точкой для построения гипер-дуальных чисел являются дуальные числа.

Дуальных числа

Классические дуальные числа [13] это – гиперкомплексные параболические числа вида $X = x + x_1\boldsymbol{\varepsilon}$, где x и x_1 – вещественные числа, а $\boldsymbol{\varepsilon}$ – абстрактный элемент, квадрат которого равен нулю [8]. Число x называется главной (Re - действительной) частью дуального числа, а x_1 – мнимой (Im - инфинитезимальной) его частью. Абстрактный элемент $\boldsymbol{\varepsilon}$ является *основным базисом* мнимой части дуального числа. Причем $\boldsymbol{\varepsilon} \neq 0$ и $\boldsymbol{\varepsilon}^n = 0$ при $n > 1$.

Для дуальных чисел определены операции сложения, умножения и деления:

$$X + Y = (x + x_1\boldsymbol{\varepsilon}) + (y + y_1\boldsymbol{\varepsilon}) = (x + y) + (x_1 + y_1)\boldsymbol{\varepsilon},$$
$$X \cdot Y = (x + x_1\boldsymbol{\varepsilon}) \cdot (y + y_1\boldsymbol{\varepsilon}) = (x \cdot y) + (x \cdot y_1 + y \cdot x_1)\boldsymbol{\varepsilon},$$
$$\frac{X}{Y} = \frac{(x + x_1\boldsymbol{\varepsilon})}{(y + y_1\boldsymbol{\varepsilon})} = \left(\frac{x}{y}\right) + \frac{(y \cdot x_1 - x \cdot y_1)}{y^2}\boldsymbol{\varepsilon} \quad \text{при} \quad y \neq 0$$

Параметром дуального числа $X = x + x_1\boldsymbol{\varepsilon}$ называется отношение $p(X) = x_1/x$, тогда дуальное число представимо в виде $X = x[1 + p(X)\boldsymbol{\varepsilon}]$.

Вводятся также следующие определения: сопряженное дуальное число $\bar{X} = x - x_1\boldsymbol{\varepsilon}$; обратное дуальное число $X^{-1} = x^{-1} + x_1 x^{-2}\boldsymbol{\varepsilon}$ (не определено при $x = 0$); дуальный ноль $0 = 0 + 0\boldsymbol{\varepsilon}$; дуальная единица $E = 1 + 0\boldsymbol{\varepsilon}$; противоположное дуальное число $-X = -x - x_1\boldsymbol{\varepsilon}$; модуль дуального числа $|X| = x$; выделения вещественной части $Re(X) = x$; выделения мнимой части $Im(X) = x_1$; дуальный полюс $P(r) = 1 + r\boldsymbol{\varepsilon}$, где r – вещественно. Нетрудно видеть, что $E = P(0) = 1 + 0\boldsymbol{\varepsilon}$, т.е. дуальная единица – суть нулевой дуальный полюс. С дуальным полюсом связана операция приведения дуального числа к полюсу (reduction to the pole) $P(r)X = x + (x_1 + rx)\boldsymbol{\varepsilon}$. Все арифметические операции над дульными числами предполагают, что компоненты приведены к одному и тому же полюсу. Считается, что изначально все дуальные числа приведен к нулевому полюсу.

Отношение порядка для дуальных чисел определяется следующим образом:

$$X = Y,\ \text{если}\ x = y\ \text{и}\ x_1 = y_1;\quad X > Y,\ \text{если}\quad x > y \quad \text{или}\quad x_1 > y_1\ \text{и}\ x = y$$

Аффинное преобразование дуального объекта (числа) осуществляется при помощи аффинатора (A) вида 4-х блочной матрицы:

$$(A) = \begin{Vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Vmatrix}$$

Умножение аффинатора (A) на дуальный объект имеет вид

$$(A) \cdot (x + x_1\boldsymbol{\varepsilon}) = A_{11} \cdot x_1 + A_{12} \cdot x + (A_{21} \cdot x_1 + A_{22} \cdot x)\boldsymbol{\varepsilon},$$

в котором прослеживается диагональная природа произведения дуальный объектов.
С помощью единично-диагонального аффинатора $(E)= diag\{1, 1\}$ можно осуществить инверсию (перестановку) компонент главной и мнимой частей дуального числа $(E) \cdot (x + x_1\boldsymbol{\varepsilon}) = x_1 + x\boldsymbol{\varepsilon}$.
Разложение аналитической функции дуального аргумента (с учетом свойств $\boldsymbol{\varepsilon}$) дает

$$F(x + x_1\boldsymbol{\varepsilon}) = f(x) + x_1 f'(x),$$

где: $f(x) = F(x + 0\boldsymbol{\varepsilon})$, $f'(x) = \partial f(x)/\partial x$.
Нетрудно видеть, что значение функции дуального аргумента $F(x + x_1\boldsymbol{\varepsilon})$ определяется через значения этой функции $f(x) = F(x + 0\boldsymbol{\varepsilon})$ и ее производной $f'(x)$ от главной части дуального числа x и величины мнимой части x_1. При $x_1 = 1$ получим $F(x + \boldsymbol{\varepsilon}) = f(x) + f'(x)$ – представление функции дуального аргумента с мнимой частью, содержащей только значение производной.

Таким образом, производя вычисления не над вещественными, а над дуальными числами, можно автоматически получать значение производной функции в точке. Особенно удобно рассматривать таким образом сложные композиции функций.
Для получения точных значений вторых производных необходимо расширение дуальных чисел.

Гипер-дуальные числа

Гипер-дуальное число [11, 12] является расширением дуального числа [13]. Если дуальное число определяется как сумма $x + x_1\boldsymbol{\varepsilon}$, то гипер-дуальное число имеет следующее представление:

$$X = x + x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + x_3\boldsymbol{\varepsilon}_{12} \quad (1)$$

где: x, x_1, x_2 и x_3 – действительные числа, а $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$ и $\boldsymbol{\varepsilon}_{12}$ – мнимые символы (объекты), причем ε_1 является символом Клиффорда $\boldsymbol{\varepsilon}_1 \neq 0$, $\boldsymbol{\varepsilon}_1^k = 0$ при $k > 1$ и мнимые символы $\boldsymbol{\varepsilon}_2$, $\boldsymbol{\varepsilon}_{12}$ аналогичной природы. Мнимые символы коммутируют с действительными числами при умножении ($a\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}a$) и их иногда называют мнимым базисом или базисом мнимой части, при этом базис числа a равен 1. Мнимые символы с одним индексом образуют основной базис ($\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$), а с двумя его расширение ($\boldsymbol{\varepsilon}_{12} = \boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2$).

Свойства мнимых символов $\boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_1$ и $\boldsymbol{\varepsilon}_{12}$ определяются следующей таблицей их произведений:

$\times$	**1**	$\boldsymbol{\varepsilon}_1$	$\boldsymbol{\varepsilon}_2$	$\boldsymbol{\varepsilon}_{12}$
1	1	$\boldsymbol{\varepsilon}_1$	$\boldsymbol{\varepsilon}_2$	$\boldsymbol{\varepsilon}_{12}$
$\boldsymbol{\varepsilon}_1$	$\boldsymbol{\varepsilon}_1$	0	$\boldsymbol{\varepsilon}_{12}$	0
$\boldsymbol{\varepsilon}_2$	$\boldsymbol{\varepsilon}_2$	$\boldsymbol{\varepsilon}_{12}$	0	0
$\boldsymbol{\varepsilon}_{12}$	$\boldsymbol{\varepsilon}_{12}$	0	0	0

Таблица 1.Правила умножения мнимых символов

Следует заметить, что структура гипер-дуального числа (1) с точностью до определения $\boldsymbol{\varepsilon}_1,\ \boldsymbol{\varepsilon}_2$ и $\boldsymbol{\varepsilon}_{12}$ соответствует структуре кватерниона [13].
Число x называется главной (Re - действительной) частью, а выражение $x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + x_3\boldsymbol{\varepsilon}_{12}$ - мнимой (Im - инфинитезимальной) частью гипер-дуального числа (1).
Нетрудно видеть, что мнимая часть гипер-дуального числа A состоит из трех слагаемых, т.е. $Im(X) = x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + x_3\boldsymbol{\varepsilon}_{12} = Im_1(X) + Im_2(X) + Im_{12}(X)$.
Алгебраические операции сложения и умножения (с учетом таблицы 1) определены по правилам:

$$
\begin{aligned}
&X = x + x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + x_3\boldsymbol{\varepsilon}_{12}, \quad Y = y + y_1\boldsymbol{\varepsilon}_1 + y_2\boldsymbol{\varepsilon}_2 + y_3\boldsymbol{\varepsilon}_{12},\\
&X + Y = x + y + (x_1 + y_1)\boldsymbol{\varepsilon}_1 + (x_2 + y_2)\boldsymbol{\varepsilon}_2 + (x_3 + y_3)\boldsymbol{\varepsilon}_{12},\\
&X{\cdot}Y = x{\cdot}y + (x{\cdot}y_1 + x_1{\cdot}y)\boldsymbol{\varepsilon}_1 + (x{\cdot}y_2 + x_2{\cdot}y)\boldsymbol{\varepsilon}_2 + (x{\cdot}y_3 + x_1{\cdot}y_2 + x_2{\cdot}y_1 + x_3{\cdot}y)\boldsymbol{\varepsilon}_{12}
\end{aligned}
\tag{2}
$$

Гипер-дуальное число $\bar{X} = Re(X) - Im(X)$ называется сопряженным гипер-дуальным числом, причем $X{\cdot}\bar{X} = Re(X)$.
Если ввести единичное гипер-дуального числа $E = 1 + 0\boldsymbol{\varepsilon}_1 + 0\boldsymbol{\varepsilon}_2 + 0\boldsymbol{\varepsilon}_{12}$, то из условия $X{\cdot}X^{-1} = E$ можно определить обратное гипер-дуальное число

$$
X^{-1} = \frac{1}{x} - \frac{x_1}{x^2}\boldsymbol{\varepsilon}_1 - \frac{x_2}{x^2}\boldsymbol{\varepsilon}_2 + \left(\frac{2x_1{\cdot}x_2}{x^3} - \frac{x_3}{x^2}\right)\boldsymbol{\varepsilon}_{12}, \tag{3}
$$

которое не определено при $x = 0$.
Используя X^{-1}, нетрудно получить формулу операции деления гипер-дуальных чисел

$$
\begin{aligned}
\frac{X}{Y} = X{\cdot}Y^{-1} = \frac{x}{y} + \left(\frac{x_1}{y} - \frac{x{\cdot}y_1}{y^2}\right)\boldsymbol{\varepsilon}_1 + \left(\frac{x_2}{y} - \frac{x{\cdot}y_2}{y^2}\right)\boldsymbol{\varepsilon}_2 \ +\\
+\left[x{\cdot}\left(\frac{2y_1{\cdot}y_2}{y^3} - \frac{y_3}{y^2}\right) - \frac{x_1{\cdot}y_2}{y^2} - \frac{x_2{\cdot}y_1}{y^2} + \frac{x_3}{y}\right]\boldsymbol{\varepsilon}_{12},
\end{aligned}
\tag{4}
$$

Операция деления при $y = 0$ не определена (деление на ноль).
Как видно из (3) и (4), деление ограничено делителем с ненулевой действительной частью. Одно из требований, чтобы система счисления была алгеброй с делением, состоит в том, что для каждого числа, которое не равно нулю, существует обратное число. Другими словами, гипер-дуальное число имеет обратное значение для каждого гипер-дуального числа, кроме $0 +$

$0\boldsymbol{\varepsilon}_1 + 0\boldsymbol{\varepsilon}_2 + 0\boldsymbol{\varepsilon}_{12}$. Однако на практике это не должно быть проблемой. Деление на ноль никогда не должно происходить в вещественной функции, поэтому деление на гипер-дуальные числа с действительными частями, $Re(X) = 0$, также никогда не должно происходить.

Для гипер-дуальных чисел, предполагается, что определение нормы гипер-дуального числа должно основываться только на действительной части этого числа, $norm(X) = \sqrt{Re^2(X)}$. Это определение нормы обладает тем свойством, что $norm(X{\cdot}Y) = norm(X){\cdot}norm(Y$. Определение нормы гипер-двойного числа, включающей только действительную часть числа, предполагает, что сравнение между двумя гипер-дуальными числами должно основываться только на действительной части этих чисел.

В дуальных числах существует операция приведения к полюсу, которую можно распространить и на гипер-дуальные числа: $P(X, r) = Re(X) + (x_1 + r{\cdot}x)\boldsymbol{\varepsilon}_1 + (x_2 + r{\cdot}x)\boldsymbol{\varepsilon}_2 + (x_3 + r{\cdot}x)\boldsymbol{\varepsilon}_{12}$, где r – действительное число, называемое полюсом приведения. По умолчанию все гипер-дуальные числа приведен к нулевому полюсу.

Разложение аналитической функции от гипер-дуального аргумента $X = x + x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + x_3\boldsymbol{\varepsilon}_{12}$ (с учетом таблицы 1) дает

$$f(X) = f(x) + x_1{\cdot}f'(x)\boldsymbol{\varepsilon}_1 + x_2{\cdot}f'(x)\boldsymbol{\varepsilon}_2 + (x_3{\cdot}f'(x) + x_1{\cdot}x_2{\cdot}f''(x))\boldsymbol{\varepsilon}_{12}, \qquad (5)$$

где $f'(x) = \frac{\partial f(x)}{\partial x}$ (градиент), $f''(x) = \frac{\partial^2 f(x)}{\partial x^2}$ (гессиан).

Остаточный член в (5) отсутствует из-за свойств $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$ и $\boldsymbol{\varepsilon}_{12}$, т. е. (5) является точным представлением $f(X)$. Таким образом, для описания $f(X)$ необходимо иметь три компоненты $f(x)$, $f'(x)$ и $f''(x)$. Заметим, что при $x_1 = x_2 = 1$ и $x_3 = 0$ соотношение (5) упрощается и принимает вид:

$$f(x + \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + 0\boldsymbol{\varepsilon}_{12}) = f(x) + f'(x)\boldsymbol{\varepsilon}_1 + f'(x)\boldsymbol{\varepsilon}_2 + +f''(x)\boldsymbol{\varepsilon}_{12} \qquad (6)$$

Соотношения (5) и (6) позволяют точно определять значения самой функции, ее первой и второй производных. В работах [11, 12] используются гипер-дуальные числа и функций от них для решения некоторых практических задач, требующих многократного точного вычисления значений градиента и гессиана.

В заключении отметим, что в некоторых прикладных дисциплинах используются производные функций до 6-го порядка включительно. Они получают специальные названия: speed (скорость), acceleration (ускорение), jerk (рывок), snap (щелчок), crackle (треск) и pop (хлопок) – производные с 1-го по 6-й порядок соответственно. Так, например, в теории балок поперечную перерезывающую силу можно ассоциировать с рывком, ускорение с изгибающим моментом, скорость с углом поворота поперечного сечения и т. д. Для каждого конкретного случая следует использовать соответствующее расширение гипер-дуальных чисел, обеспечивающее нужный порядок производных.

Литература

1. Clifford W. K. Preliminary Sketch of Biquaternions. Proc. London Math. Soc., s1-4(1):381–395, 1871.
2. Deakin M. A. B Functions of a dual or duo variable. Mathematics Magazine, 39(4):215–219, 1966.
3. R. W. H. T. Hudson R. W. H. T. Review: Geometrie der dynamen. von E. Study. The Mathematical Gazette, 3(44):15–16, 1904.
4. Kantor I. L., Solodovnikov A. S. Hypercomplex Numbers: An Elementary Introduction to Algebras. Springer-Verlag, New York, 1989.
5. Study E. Geometrie der Dynamen. Verlag Teubner, Leipzig ,1903.
6. Гордон Е. И., Кусраев А. Г., С. С. Кутателадзе С. С. Инфинитезимальный анализ. – Изд-во ин-та математики, Новосибирск, 2006. 526 с.
7. Котельников А. П. Винты и комплексные числа. Казань, 1896.
8. Диментберг Ф. М. Теория винтов и ее приложения. М.: Наука, 1978. 328 с.
9. Naumann U. The art of differentiating computer programs. Society for industrial and applied mathematics, Philadelphia, USA, 2012.
10. Corliss G., Faure C., Griewank A., Hascoлт L., Naumann U. Automatic Differentiation Bibliography // Automatic Differentiation of Algorithms: From Simulation to Optimization. Springer, 2002. p. 383—425
11. Fike J.A., Alonso J.J. The development of hyper-dual numbers for exact second derivative calculations. AIAA paper 2011-886, 49th AIAA Aerospace Sciences Meeting, January 4-7, 2011.
12. Fike J.A. Multi-objective optimization using hyper-dual numbers. Ph.D. Dissertation, Stanford University, 2013.
13. Гордеев В. Н. Кватернионы и бикватернионы с приложениями в геометрии и механике. Киев: Сталь, 2016. - 315 с.

Абстракт

Кратко описана история возникновения дуальных чисел, их определение и использование. Приводится понятие гипер-дуальных чисел как расширение дуальных чисел. Рассматриваются основные операции над гипер-дуальными числами и понятие функции гипер-дуального аргумента. Указывается на применение гипер-дуальных чисел и функций от них для решения некоторых практических задач, требующих многократного точного вычисления значений градиента и гессиана в вычислительных процессах автоматического дифференцирования.

Ключевые слова: *дуальные числа, гипер-дуальные числа, автоматическое дифференцирование, dual numbers, hyper-dual numbers, automatic differentiation*

---

12 апрель 2019