Fundamental theorem of matrix representations of hyper-dual numbers for computing higher-order derivatives

Yusuke Imoto¹, Naoya Yamanaka², Takeo Uramoto¹, Masato Tanaka³, Masaki Fujikawa⁴ and Naoto Mitsume⁵

¹ Institute for Advanced Study, Kyoto University, Yoshida Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan

² School of Information Science, Meisei University, 2-1-1 Hodokubo, Hino, Tokyo 191-8506, Japan

³ Toyota Central R&D Laboratories, INC., 41-1 Yokomichi, Nagakute, Aichi 480-1192, Japan

⁴ Faculty of Engineering Department of Mechanical Systems Engineering, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan

⁵ Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan

E-mail imoto.yusuke.4e@kyoto-u.ac.jp

Received March 6, 2020, Accepted April 23, 2020

Abstract

Hyper-dual numbers (HDN) are numbers defined by using nilpotent elements that differ from each other. The introduction of an operator to extend the domain of functions to HDN space based on Taylor expansion allows higher-order derivatives to be obtained from the coefficients. This study inductively defines matrix representations of HDN and proposes a numerical method for higher-order derivatives, called *HDN-M differentiation*, based on the matrix representations of HDN. The proposed method is characterized so that higher-order derivatives can be computed with matrix operation rules without implementations of the operation rules of HDN.

Keywords hyper-dual numbers, matrix representation, higher-order derivatives, HDN-M differentiation, automatic differentiation

Research Activity Group Mathematical Aspects of Continuum Mechanics

1. Introduction

In numerical analysis, it is often necessary to strictly calculate higher-order derivatives of given functions. For example, the numerical simulations of nonlinear materials based on incremental variational formulation [1, 2], usually applied to steel, organic, inorganic, and composite materials, requires higher-order derivatives of two potentials: a free energy and a dissipation function, which are given depending on the models. The automatic differentiation is a representative method of derivative computations and computes strict derivatives based on derivative rules such as the Leibniz rule and the Faà di Bruno's formula. However, as algebraic structures based on derivative rules are quite complex, this disturbs its acceleration.

Hyper-dual numbers (HDN) [3] are extended numbers of dual numbers (DN) [4, 5], and are defined using nilpotent elements that differ from each other. By defining an operator on the space spanned by HDN based on Taylor expansion, and referring to its coefficients of HDN, we can obtain the higher-order derivatives of functions. The computational method of higher-order derivatives utilizing the property of HDN is referred to as *HDN differentiation*. This method only utilizes implementations

of HDN operation rules and the operator, both being very simple, without any implementations of the derivative rules. The property can redress us from the typical complexity of implementations. Therefore, HDN differentiation has been used for numerical simulations for elastic and inelastic materials at large strains [2, 6].

However, because HDN differentiation requires additional implementations based on HDN operation rules, we need to develop exclusive methods to enhance performance. It follows that the prospects of high-speed or self-validating numerics for HDN differentiation are low. Therefore, we develop a new derivative method without the implementation of HDN operation rules.

In this study, as an extension of the matrix representation of DN [7], we construct matrix representations of HDN. Moreover, we introduce an operator that extends the domain of functions to the space spanned by the matrix representations of HDN. The operator is proven to obtain higher-order derivatives by referring appropriate elements of the matrix induced by the operator. Furthermore, we propose a derivative method utilizing the theorem, called *HDN-M differentiation*, and give two examples. As the operation rules of HDN are replaced with those of matrices, HDN-M differentiation is character-

Table 1. Comparison of methods for computing higher-order derivatives.

	accuracy	implementation of operations
Numerical differentiation (finite difference method)	low	unnecessary
Automatic differentiation	high	necessary
HDN differentiation	high	necessary
Proposed method (HDN-M differentiation)	high	unnecessary (replaced with matrix operations)

ized by not requiring the implementation of operation rules; see Table 1. Furthermore, we can easily accelerate HDN-M differentiation by utilizing existing libraries for linear algebra.

2. Hyper-dual numbers

We consider n -elements $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$ satisfying

$$\begin{aligned} \varepsilon_i &\neq 0 \quad (i = 1, 2, \dots, n); \quad \varepsilon_i^2 = 0 \quad (i = 1, 2, \dots, n); \\ \varepsilon_i &\neq \varepsilon_j \quad (i \neq j); \quad \varepsilon_i \varepsilon_j = \varepsilon_j \varepsilon_i \neq 0 \quad (i \neq j). \end{aligned} \quad (1)$$

Then, n th order hyper-dual numbers are defined as the number

$$\sum_{\alpha \in \Lambda_n} a_\alpha \varepsilon^\alpha, \quad a_\alpha \in \mathbb{R}^m. \quad (2)$$

Here, Λ_n is the set of n -dimensional multi-indices whose components are 0 or 1 and $\varepsilon_i^0 := 1$ ($i = 1, 2, \dots, n$). Let Ω be a domain in $\mathbb{R}^m$ ($m \in \mathbb{N}$). Let $\mathbb{D}_n(\Omega)$ be the space of n th order hyper-dual numbers defined as

$$\mathbb{D}_n(\Omega) := \left\{ \sum_{\alpha \in \Lambda_n} a_\alpha \varepsilon^\alpha \mid a_0 \in \Omega, a_\alpha \in \mathbb{R}^m \ (\alpha \neq 0) \right\}.$$

We define operator T with respect to $f \in C^n(\Omega : \mathbb{R})$, which corresponds to the Taylor expansion, as

$$\begin{aligned} Tf : \mathbb{D}_n(\Omega) &\ni \sum_{\alpha \in \Lambda_n} a_\alpha \varepsilon^\alpha \\ &\mapsto \sum_{|\beta| \leq n} \frac{\partial^\beta f(a_0)}{\beta!} \left(\sum_{\alpha \in \Lambda_n \setminus \{0\}} a_\alpha \varepsilon^\alpha \right)^\beta \in \mathbb{D}_n(\mathbb{R}). \end{aligned}$$

Note that β is the m -dimensional multi-index. The operator has the following algebraic properties: for $f, g \in C^n(\Omega : \mathbb{R})$, $h \in C^n(\mathbb{R} : \mathbb{R})$,

$$\begin{aligned} T(f + g) &= Tf + Tg; \\ T(f \cdot g) &= (Tf) \cdot (Tg); \\ T(f \circ h) &= (Tf) \circ (Th). \end{aligned} \quad (3)$$

Moreover, for $\eta = (\eta_1, \eta_2, \dots, \eta_n) \in \{1, 2, \dots, m\}^n$, we have

$$\begin{aligned} Tf \left(a + \sum_{i=1}^n \delta_{\eta_i} \varepsilon_i \right) &= \sum_{|\beta| \leq n} \frac{\partial^\beta f(a)}{\beta!} \left(\sum_{i=1}^n \delta_{\eta_i} \varepsilon_i \right)^\beta \\ &= \sum_{\alpha \in \Lambda_n} \partial_\eta^\alpha f(a) \varepsilon^\alpha. \end{aligned}$$

Here, $\partial_\eta := (\partial_{\eta_1}, \partial_{\eta_2}, \dots, \partial_{\eta_n})^T$ and $\delta_i := (\delta_{i1}, \delta_{i2}, \dots, \delta_{im})^T \in \{0, 1\}^m$, where δ_{ij} is the Kronecker delta. Therefore, we obtain

$$\mathfrak{J}_{\varepsilon^\alpha} \left[Tf \left(a + \sum_{i=1}^n \delta_{\eta_i} \varepsilon_i \right) \right] = \partial_\eta^\alpha f(a), \quad \alpha \in \Lambda_n \quad (4)$$

with

$$\mathfrak{J}_{\varepsilon^\alpha} [\cdot] : \mathbb{D}_n(\mathbb{R}) \ni \sum_{\beta \in \Lambda_n} a_\beta \varepsilon^\beta \mapsto a_\alpha \in \mathbb{R}.$$

From above, for function $f \in C^n(\Omega : \mathbb{R})$ which is represented in the sum, product, and composition of some components, if the components and their derivatives are given, we can obtain up to n th order ordinary or partial derivatives of f on $a \in \Omega$ by computing (4).

3. HDN-M differentiation

For $k \in \mathbb{N}$, let $M_k(\mathbb{R})$ be the set of all square k -by- k matrices over $\mathbb{R}$. Let $I_k \in M_k(\mathbb{R})$ and $O_k \in M_k(\mathbb{R})$ be the k -by- k identity matrix and k -by- k zero matrix, respectively. For $n \in \mathbb{N}$, matrices $\mathcal{E}_{n,i} \in M_{2^n}(\mathbb{R})$ ($n \geq 2, i = 1, 2, \dots, n$) are inductively defined as

$$\mathcal{E}_{n,i} = \begin{pmatrix} \mathcal{E}_{n-1,i} & O_{2^{n-1}} \\ O_{2^{n-1}} & \mathcal{E}_{n-1,i} \end{pmatrix}, \quad i = 1, 2, \dots, n-1, \quad (5)$$

$$\mathcal{E}_{n,n} = \begin{pmatrix} O_{2^{n-1}} & I_{2^{n-1}} \\ O_{2^{n-1}} & O_{2^{n-1}} \end{pmatrix}. \quad (6)$$

Here, $\mathcal{E}_{1,1} \in M_2(\mathbb{R})$ is

$$\mathcal{E}_{1,1} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then, matrices $\mathcal{E}_{n,i}$ ($i = 1, 2, \dots, n$) satisfy the following properties, which correspond to (1),

$$\begin{aligned} \mathcal{E}_{n,i} &\neq O_{2^n} \quad (i = 1, 2, \dots, n); \\ \mathcal{E}_{n,i}^2 &= O_{2^n} \quad (i = 1, 2, \dots, n); \\ \mathcal{E}_{n,i} &\neq \mathcal{E}_{n,j} \quad (i \neq j); \\ \mathcal{E}_{n,i} \mathcal{E}_{n,j} &= \mathcal{E}_{n,j} \mathcal{E}_{n,i} \neq O_{2^n} \quad (i \neq j). \end{aligned}$$

Therefore, matrix

$$\sum_{\alpha \in \Lambda_n} a_\alpha \otimes \mathcal{E}_n^\alpha, \quad a_\alpha \in \mathbb{R}^m \quad (7)$$

with $\mathcal{E}_n := (\mathcal{E}_{n,1}, \mathcal{E}_{n,2}, \dots, \mathcal{E}_{n,n})^T$ is a matrix representation of hyper-dual numbers (2). Here, $\mathcal{E}_n^\alpha = \mathcal{E}_{n,1}^{\alpha_1} \mathcal{E}_{n,2}^{\alpha_2} \dots \mathcal{E}_{n,n}^{\alpha_n}$ and $\mathcal{E}_{n,i}^0 = I_{2^n}$ ($i = 1, 2, \dots, n$). Let $\mathbb{M}_n(\Omega)$ be the space spanned by matrix (7) defined as

$$\mathbb{M}_n(\Omega) := \left\{ \sum_{\alpha \in \Lambda_n} a_\alpha \otimes \mathcal{E}_n^\alpha \mid a_0 \in \Omega, a_\alpha \in \mathbb{R}^m \ (\alpha \neq 0) \right\}.$$

We define operator $\mathcal{T}$ with respect to $f \in C^n(\Omega : \mathbb{R})$ as

$$\begin{aligned} \mathcal{T}f : \mathbb{M}_n(\Omega) &\ni \sum_{\alpha \in \Lambda_n} a_\alpha \otimes \mathcal{E}_n^\alpha \mapsto \\ &\sum_{|\beta| \leq n} \frac{\partial^\beta f(a_0)}{\beta!} \left(\sum_{\alpha \in \Lambda_n \setminus \{0\}} a_\alpha \otimes \mathcal{E}_n^\alpha \right)^\beta \in \mathbb{M}_n(\mathbb{R}). \end{aligned} \quad (8)$$

Here,

$$\left(\sum_{\alpha \in \Lambda_n \setminus \{0\}} a_\alpha \otimes \mathcal{E}_n^\alpha \right)^\beta = \prod_{i=1}^m \left(\sum_{\alpha \in \Lambda_n \setminus \{0\}} a_{\alpha,i} \mathcal{E}_n^\alpha \right)^{\beta_i}.$$

Moreover, $a_\alpha = (a_{\alpha,1}, a_{\alpha,2}, \dots, a_{\alpha,m})^\top$ and multi-index $\beta = (\beta_1, \beta_2, \dots, \beta_m)$. The operator has the following algebraic properties, which correspond to (3): for $f, g \in C^n(\Omega : \mathbb{R})$, $h \in C^n(\mathbb{R} : \mathbb{R})$,

$$\begin{aligned} \mathcal{T}(f + g) &= \mathcal{T}f + \mathcal{T}g; \\ \mathcal{T}(f \cdot g) &= (\mathcal{T}f) * (\mathcal{T}g); \\ \mathcal{T}(f \circ h) &= (\mathcal{T}f) \circ (\mathcal{T}h). \end{aligned}$$

Here, $*$ is the matrix product. Now, for implementations, we emphasize that the definition (8) is equivalent to the following matrix-based definition:

$$\mathcal{T}f : \mathbb{M}_n(\Omega) \ni M \mapsto \sum_{|\beta| \leq n} \frac{\partial^\beta f(x)}{\beta!} Q(M)^\beta \Big|_{x=P(M)} \in \mathbb{M}_n(\mathbb{R}). \quad (9)$$

Here, $P(M) := ((M_1)_{11}, (M_2)_{11}, \dots, (M_m)_{11})^\top \in \Omega$ with $M := (M_1, M_2, \dots, M_m)^\top$ and $Q(M) := M - P(M) \otimes I_{2^n}$.

For $k \in \mathbb{N}$, let $\mathfrak{M}_{i,j}[\cdot] : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ ($1 \leq i, j \leq n$) be a function picking (i, j) element up from a matrix defined in $M_k(\mathbb{R})$, i.e., $\mathfrak{M}_{i,j}[B] = B_{ij}$. Then, we obtain the following theorem.

Theorem 1 Let $n, m \in \mathbb{N}$, $f \in C^n(\Omega : \mathbb{R})$, and $\eta = (\eta_1, \eta_2, \dots, \eta_n) \in \{1, 2, \dots, m\}^n$. For all $a \in \Omega$ and $k = 1, 2, \dots, n$, we have

$$\begin{aligned} \mathfrak{M}_{1,2^k} \left[\mathcal{T}f \left(a \otimes I_{2^n} + \sum_{i=1}^n \delta_{\eta_i} \otimes \mathcal{E}_{n,i} \right) \right] \\ = \partial_{\eta_1} \partial_{\eta_2} \dots \partial_{\eta_k} f(a). \end{aligned} \quad (10)$$

Proof Let $A_{n,\eta}$ with respect to $f \in C^n(\Omega : \mathbb{R})$ be

$$A_{n,\eta}(a; f) := \mathcal{T}f \left(a \otimes I_{2^n} + \sum_{i=1}^n \delta_{\eta_i} \otimes \mathcal{E}_{n,i} \right).$$

We prove (10), i.e.,

$$\mathfrak{M}_{1,2^k} [A_{n,\eta}(a; f)] = \partial_{\eta_1} \partial_{\eta_2} \dots \partial_{\eta_k} f(a)$$

with mathematical induction for $n \in \mathbb{N}$. Since

$$A_{1,\eta}(a; f) = \begin{pmatrix} f(a) & \partial_{\eta_1} f(a) \\ 0 & f(a) \end{pmatrix},$$

(10) with $n = 1$ holds. Next, we show that if (10) with $n = \ell$ ($\ell \in \mathbb{N}$) holds, then also (10) with $n = \ell + 1$ holds. From (5), (6), and (8), we have

$$\begin{aligned} A_{\ell+1,\eta}(a; f) \\ &= \sum_{\alpha \in \Lambda_{\ell+1}} \partial_\eta^\alpha f(a) \mathcal{E}_n^\alpha \\ &= \sum_{\alpha \in \Lambda_{\ell+1}, \alpha_{\ell+1}=0} \partial_\eta^\alpha f(a) \mathcal{E}_n^\alpha + \sum_{\alpha \in \Lambda_{\ell+1}, \alpha_{\ell+1}=1} \partial_\eta^\alpha f(a) \mathcal{E}_n^\alpha \end{aligned}$$

$$\begin{aligned} &= \sum_{\alpha \in \Lambda_\ell} \partial_\eta^\alpha f(a) \begin{pmatrix} \mathcal{E}_\ell^\alpha & O_{2^\ell} \\ O_{2^\ell} & \mathcal{E}_\ell^\alpha \end{pmatrix} \\ &\quad + \sum_{\alpha \in \Lambda_\ell} \partial_\eta^\alpha \partial_{\eta_{\ell+1}} f(a) \begin{pmatrix} O_{2^\ell} & \mathcal{E}_\ell^\alpha \\ O_{2^\ell} & O_{2^\ell} \end{pmatrix} \\ &= \begin{pmatrix} A_{\ell,\eta}(a; f) & \sum_{\alpha \in \Lambda_\ell} \partial_\eta^\alpha \partial_{\eta_{\ell+1}} f(a) \mathcal{E}_\ell^\alpha \\ O_{2^\ell} & A_{\ell,\eta}(a; f) \end{pmatrix}. \end{aligned}$$

Here, $\partial_\eta := (\partial_{\eta_1}, \partial_{\eta_2}, \dots, \partial_{\eta_\ell})^\top$. Therefore, for $k = 1, 2, \dots, \ell$, we have

$$\begin{aligned} \mathfrak{M}_{1,2^k} [A_{\ell+1,\eta}(a; f)] &= \mathfrak{M}_{1,2^k} [A_{\ell,\eta}(a; f)] \\ &= \partial_{\eta_1} \partial_{\eta_2} \dots \partial_{\eta_k} f(a). \end{aligned} \quad (11)$$

Specifically, from (11) with $k = \ell$ and arbitrariness of f , for all $a_\alpha \in \Omega$ ($\alpha \in \Lambda_\ell$), we have

$$\mathfrak{M}_{1,2^\ell} \left[\sum_{\alpha \in \Lambda_\ell} a_\alpha \mathcal{E}_\ell^\alpha \right] = a_{(1,1,\dots,1)}.$$

Therefore, we conclude

$$\begin{aligned} \mathfrak{M}_{1,2^{\ell+1}} [A_{\ell+1,\eta}(a; f)] &= \mathfrak{M}_{1,2^\ell} \left[\sum_{\alpha \in \Lambda_\ell} \partial_\eta^\alpha \partial_{\eta_{\ell+1}} f(a) \mathcal{E}_\ell^\alpha \right] \\ &= \partial_{\eta_1} \partial_{\eta_2} \dots \partial_{\eta_{\ell+1}} f(a). \end{aligned}$$

Consequently, by the mathematical induction, (10) holds for all $n \in \mathbb{N}$.

(QED)

From Theorem 1, for function $f \in C^n(\Omega : \mathbb{R})$, we can obtain up to n th order ordinary or partial derivatives of f on $a \in \Omega$. We call the method computing higher-order derivatives from (10) *HDN-M differentiation*. Note that because division $\{f(x)\}^{-1}$ is represented as $h \circ f(x)$ with $h(x) = x^{-1}$, HDN-M differentiation does not require the division (the inverse matrix operation) by implementing $\mathcal{T}h$. We show two examples of the computational procedure of HDN-M differentiation in the next section. Note that the implementation of the proposed method follows the matrix-based definition (9) and main results (10).

4. Examples

In this section two examples of HDN-M differentiation are shown. First, we show the case of solving the 2nd ordinary derivative $f^{(2)}$ on $a \in \mathbb{R}$ with $f(x) = e^{\sin(x)}$, $n = 2$ and $m = 1$. We represent f as $f = f_2 \circ f_1$, where $f_1(x) := \sin(x)$ and $f_2(x) := e^x$. Let $M_0 := aI_4 + \mathcal{E}_{2,1} + \mathcal{E}_{2,2}$. From $\mathcal{T}f(\cdot) = \mathcal{T}f_2(\mathcal{T}f_1(\cdot))$, we solve the derivative by the following steps: First, from $P(M_0) = a$, we compute $M_1 := \mathcal{T}f_1(M_0)$ as

$$\begin{aligned} M_1 &= \sum_{k=0}^2 \frac{f_1^{(k)}(x)}{k!} Q(M_0)^k \Big|_{x=a} \\ &= \sin(a)I_4 + \cos(a)Q(M_0) - \frac{1}{2} \sin(a)Q(M_0)^2. \end{aligned}$$

Next, from $P(M_1) = \sin(a)$, we compute $M_2 := \mathcal{T}f_2(M_1)$ as

$$\begin{aligned} M_2 &= \sum_{k=0}^2 \frac{f_2^{(k)}(x)}{k!} Q(M_1)^k \Big|_{x=\sin(a)} \\ &= e^{\sin(a)} I_4 + e^{\sin(a)} Q(M_1) + \frac{1}{2} e^{\sin(a)} Q(M_1)^2 \\ &= \begin{pmatrix} s & t & t & u \\ 0 & s & 0 & t \\ 0 & 0 & s & t \\ 0 & 0 & 0 & s \end{pmatrix}, \end{aligned}$$

where $s := e^{\sin(a)}$, $t := \cos(a)e^{\sin(a)}$, and $u := (\cos^2(a) - \sin(a))e^{\sin(a)}$. Finally, from Theorem 1, we obtain

$$f^{(2)}(a) = \mathfrak{M}_{1,2^2}[M_2] = u = (\cos^2(a) - \sin(a))e^{\sin(a)}.$$

Next, we show the case of solving the 3rd partial derivative $\partial_1^2 \partial_2 g$ on $a = (a_1, a_2)^T \in \mathbb{R}^2$ with $g(x) = e^{\sin(x_1+2x_2)} (x = (x_1, x_2)^T)$, $n = 3$ and $m = 2$. We represent g as $g = g_4 \circ g_3 \circ (g_1 + g_2)$, where $g_1(x_1, x_2) := x_1$, $g_2(x_1, x_2) := 2x_2$, $g_3(x) := \sin(x)$, and $g_4(x) := e^x$. Let $\widehat{M}_0 := a \otimes I_8 + \delta_1 \otimes \mathcal{E}_{3,1} + \delta_1 \otimes \mathcal{E}_{3,2} + \delta_2 \otimes \mathcal{E}_{3,3}$. From $\mathcal{T}g(\cdot) = \mathcal{T}g_4(\mathcal{T}g_3(\mathcal{T}g_1(\cdot) + \mathcal{T}g_2(\cdot)))$, we solve the derivative by the following steps: First, based on the matrix-based definition (9), we obtain $\widehat{M}_1 := \mathcal{T}g_1(\widehat{M}_0)$ and $\widehat{M}_2 := \mathcal{T}g_2(\widehat{M}_0)$ as $\widehat{M}_1 = a_1 I_8 + \mathcal{E}_{3,1} + \mathcal{E}_{3,2}$ and $\widehat{M}_2 = 2a_2 I_8 + 2\mathcal{E}_{3,3}$, respectively. Next, from $P(\widehat{M}_1 + \widehat{M}_2) = a_1 + 2a_2$, we compute $\widehat{M}_3 = \mathcal{T}g_3(\widehat{M}_1 + \widehat{M}_2)$ as

$$\begin{aligned} \widehat{M}_3 &= \sum_{k=0}^3 \frac{g_3^{(k)}(x)}{k!} Q(\widehat{M}_1 + \widehat{M}_2)^k \Big|_{x=a_1+2a_2} \\ &= \sin(a_1 + 2a_2) I_8 + \cos(a_1 + 2a_2) Q(\widehat{M}_1 + \widehat{M}_2) \\ &\quad - \frac{1}{2} \sin(a_1 + 2a_2) Q(\widehat{M}_1 + \widehat{M}_2)^2 \\ &\quad - \frac{1}{6} \cos(a_1 + 2a_2) Q(\widehat{M}_1 + \widehat{M}_2)^3 \end{aligned}$$

Next, from $P(\widehat{M}_3) = \sin(a_1 + 2a_2)$, we compute $\widehat{M}_4 = \mathcal{T}g_4(\widehat{M}_3)$ as

$$\begin{aligned} \widehat{M}_4 &= \sum_{k=0}^3 \frac{g_4^{(k)}(x)}{k!} Q(\widehat{M}_3)^k \Big|_{x=\sin(a_1+2a_2)} \\ &= e^{\sin(a_1+2a_2)} I_8 + e^{\sin(a_1+2a_2)} Q(\widehat{M}_3) \\ &\quad + \frac{1}{2} e^{\sin(a_1+2a_2)} Q(\widehat{M}_3)^2 + \frac{1}{6} e^{\sin(a_1+2a_2)} Q(\widehat{M}_3)^3 \\ &= \begin{pmatrix} \widehat{s} & \widehat{t} & \widehat{t} & \widehat{u} & 2\widehat{t} & 2\widehat{u} & 2\widehat{u} & \widehat{v} \\ 0 & \widehat{s} & 0 & \widehat{t} & 0 & 2\widehat{t} & 0 & 2\widehat{u} \\ 0 & 0 & \widehat{s} & \widehat{t} & 0 & 0 & 2\widehat{t} & 2\widehat{u} \\ 0 & 0 & 0 & \widehat{s} & 0 & 0 & 0 & 2\widehat{t} \\ 0 & 0 & 0 & 0 & \widehat{s} & \widehat{t} & \widehat{t} & \widehat{u} \\ 0 & 0 & 0 & 0 & 0 & \widehat{s} & 0 & \widehat{t} \\ 0 & 0 & 0 & 0 & 0 & 0 & \widehat{s} & \widehat{t} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \widehat{s} \end{pmatrix}, \end{aligned}$$

where $\widehat{s} := e^{\sin(a_1+2a_2)}$, $\widehat{t} := e^{\sin(a_1+2a_2)} \cos(a_1 + 2a_2)$, $\widehat{u} := e^{\sin(a_1+2a_2)} (\cos^2(a_1+2a_2) - \sin(a_1+2a_2))$, $\widehat{v} := -2e^{\sin(a_1+2a_2)} \cos(a_1 + 2a_2) \sin(a_1 + 2a_2)$. Finally, from

Theorem 1, we obtain

$$\begin{aligned} &\partial_1^2 \partial_2 g(a_1, a_2) \\ &= \mathfrak{M}_{1,2^3}[\widehat{M}_4] = \widehat{v} \\ &= -2e^{\sin(a_1+2a_2)} \cos(a_1 + 2a_2) \sin(a_1 + 2a_2). \end{aligned}$$

Note that, from above, we find that HDN-M differentiation only requires the implementation of the matrix-based definition (9) with one-dimensional space ($m = 1$) for fundamental functions.

5. Concluding remarks

We have proposed a numerical method “*HDN-M differentiation*” for higher-order derivatives based on matrix representations of hyper-dual numbers (HDN), and proved a fundamental theorem of HDN-M differentiation to obtain higher-order derivatives. The matrix representations of HDN have been inductively defined by utilizing a matrix representation of the dual number. HDN-M differentiation has been defined by utilizing the operator that extends the domain of function to space spanned matrix representations of HDN, based on Taylor expansion. Moreover, the computability of higher-order derivatives in HDN-M differentiation is ensured by the fundamental theorem we proved.

These are simple implementations even for multi-variable functions. This property poses future issues and opportunities for HDN-M differentiation, such as acceleration by applying existing libraries of linear algebra, or developing acceleration methods specialized for operations for upper triangular matrices. Likewise, the self-validating numerics to develop a numerical method for higher-order derivatives with high reliability, high accuracy, high versatility, and high speed should be explored further.

Acknowledgments

This work was supported by JST CREST Grant Number JPMJCR14D4, Japan. Moreover, this work was supported by the priority research project of Graduate School of Information Sciences, Tohoku University.

References

- [1] M. Ortiz and L. Stainier, The variational formulation of viscoplastic constitutive updates, *Comput. Methods Appl. Mech. Engrg.*, **171** (1999), 419–444.
- [2] M. Tanaka, D. Balzani and J. Schröder, Implementation of incremental variational formulations based on the numerical calculation of derivatives using hyper dual numbers, *Comput. Methods Appl. Mech. Engrg.*, **301** (2016), 216–241.
- [3] J. A. Fike, Multi-objective optimization using hyper-dual numbers, Stanford University, Ph.D. Thesis, 2013.
- [4] W. K. Clifford, Preliminary sketch of biquaternions, *Proc. Lond. Math. Soc.*, **4** (1873) 381–395.
- [5] E. Pennestrì and R. Stefaneli, Linear algebra and numerical algorithms using dual numbers, *Multibody Syst. Dyn.*, **18** (2007), 323–344.
- [6] M. Tanaka, T. Sasagawa, et al., A highly accurate 1st- and 2nd-order differentiation scheme for hyperelastic material models based on hyper-dual numbers, *Comput. Methods Appl. Mech. Engrg.*, **283** (2015), 22–45.
- [7] S. Demir, Matrix realization of dual quaternionic electromagnetism, *Centr. Eur. J. Phys.*, **5** (2007), 487–506.