AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 9(12): 35125–35150.
DOI: 10.3934/math.20241670
Received: 10 October 2024
Revised: 30 November 2024
Accepted: 09 December 2024
Published: 16 December 2024

Research article

Characterizations and properties of hyper-dual Moore-Penrose generalized inverse

Qi Xiao and Jin Zhong*

Faculty of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China

* **Correspondence:** Email: jzhong@jxust.edu.cn.

Abstract: In this paper, the definition of the hyper-dual Moore-Penrose generalized inverse of a hyper-dual matrix is introduced. Characterizations for the existence of the hyper-dual Moore-Penrose generalized inverse are given, and a formula for the hyper-dual Moore-Penrose generalized inverse is presented whenever it exists. Least-squares properties of the hyper-dual Moore-Penrose generalized inverse are discussed by introducing a total order of hyper-dual numbers. We also introduce the definition of a dual matrix of order n. A necessary and sufficient condition for the existence of the Moore-Penrose generalized inverse of a dual matrix of order n is given.

Keywords: dual matrix; hyper-dual matrix; hyper-dual Moore-Penrose generalized inverse; least-squares property; dual matrix of order n
Mathematics Subject Classification: 15A09, 15A24, 15B33

1. Introduction

Dual numbers were introduced by Clifford [1] in order to expand quaternions to bi-quaternions that represent both rotations and translations. Dual numbers since then have been important and convenient mathematical tools in dealing with some problems in various fields of science and engineering, such as kinematic synthesis [2, 3], robotics [4], scara kinematics [5] and displacement analysis [6, 7]. A matrix with dual number entries is called a dual matrix. Dual matrices are used today in a variety of fields like kinematic analysis and synthesis of spatial mechanisms, and also in robotics [8]. There are many investigations where the kinematic analysis and synthesis problems are addressed through the solution of overdetermined systems of linear dual equations, and dual generalized inverses of dual matrices have been shown to be very useful in studying the solutions of systems of linear dual equations [9]. For example, the dual Moore-Penrose generalized inverse (DMPGI, for short) provides minimum-norm least-squares solution for the system of linear dual equations [10]

$$\widehat{A}\widehat{x} = \widehat{b}.$$

However, many research results have shown that various dual generalized inverses of dual matrices may not exist. Based on this fact, in the past few years, numerous articles were dedicated to characterizing the existence of different kinds of dual generalized inverses, for example, DMPGI [11, 12], weak dual generalized inverse [13], dual core generalized inverse [14, 15]. Especially, Wang [16] gave some necessary and sufficient conditions for a dual matrix to have the DMPGI, and a compact formula for the computation of the DMPGI was also given. Zhong and Zhang [17, 18] presented some necessary and sufficient conditions for a square dual matrix to have the dual group inverse and the dual Drazin inverse.

Throughout this paper, we use $\widehat{\mathbb{R}}$ to denote the set of dual numbers over the real field. A dual number $\widehat{a} \in \widehat{\mathbb{R}}$ has the form

$$\widehat{a} = a + \epsilon a_0,$$

where a and a_0 are real numbers, and ϵ is the dual unity that satisfies the rules

$$\epsilon \neq 0 \quad \text{and} \quad \epsilon^2 = 0.$$

Hyper-dual numbers are an extension of dual numbers and were first introduced by Fike et al. [19–21] to derive the kinematics of a multi-body system. They introduced the hyper-dual numbers to perform second-order numerical differentiation that leads to smaller numerical (subtractive and cancellation) errors as well as to reduced computational time. A hyper-dual number $\widetilde{a}$ is a number consisting of four real numbers a_0–a_3 and two dual units ϵ_1, ϵ_2 with the following rules:

$$\epsilon_1^2 = \epsilon_2^2 = (\epsilon_1\epsilon_2)^2 = 0, \quad \epsilon_1, \epsilon_2, \epsilon_1\epsilon_2 \neq 0,$$

and $\widetilde{a}$ is of the form

$$\widetilde{a} = a_0 + \epsilon_1 a_1 + \epsilon_2 a_2 + \epsilon_1\epsilon_2 a_3. \tag{1.1}$$

Notice that we can rewrite the hyper-dual number $\widetilde{a}$ in (1.1) as

$$\widetilde{a} = (a_0 + \epsilon_1 a_1) + \epsilon_2(a_2 + \epsilon_1 a_3) \triangleq \widehat{a} + \epsilon_2\widehat{a}_0, \tag{1.2}$$

i.e., a hyper-dual number is a combination of two dual numbers, where $\widehat{a}$ is called the primal part and $\widehat{a}_0$ is called the hyper-dual part of $\widetilde{a}$, respectively. In other words, a hyper-dual number can be obtained by replacing the two real numbers in a dual number by two dual numbers. The physical meaning of these two dual numbers in the context of kinematics was discussed in [22, 23] by introducing the hyper-dual angle. We denote the set of all hyper-dual numbers over the real field by $\widetilde{\mathbb{R}}$. For the sake of convenience, we replace ϵ_1, ϵ_2 by ϵ, ϵ^* in (1.1) and (1.2).

For $\widetilde{a} \in \widetilde{\mathbb{R}}$, the Taylor series expansion of a dual function of order 2 is given by (see [21])

$$f(\widetilde{a}) = f(a_0) + \epsilon a_1 f'(a_0) + \epsilon^* a_2 f'(a_0) + \epsilon\epsilon^* \left[a_3 f'(a_0) + a_1 a_2 f''(a_0)\right].$$

For example, for a hyper-dual number

$$\widetilde{a} = a_0 + \epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3$$

with $a_0 > 0$, the square root of $\widetilde{a}$ is given by

$$\sqrt{\widetilde{a}} = \sqrt{a_0} + \epsilon\frac{a_1}{2\sqrt{a_0}} + \epsilon^*\left[\frac{a_2}{2\sqrt{a_0}} + \epsilon(\frac{a_3}{2\sqrt{a_0}} - \frac{a_1a_2}{4\sqrt{a_0^3}})\right]. \tag{1.3}$$

According to (1.3), for

$$\widetilde{a} = a_0 + \epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3 \in \widetilde{\mathbb{R}}$$

with $a_0 \neq 0$, the absolute value and the Euclidean norm of $\widetilde{a}$ can be respectively defined by

$$|\widetilde{a}| = |a_0| + \epsilon\ \mathrm{sgn}(a_0)a_1 + \epsilon^*\mathrm{sgn}(a_0)a_2 + \epsilon\epsilon^*\mathrm{sgn}(a_0)a_3$$

and

$$\|\widetilde{a}\| = \|a_0\| + \epsilon\ \frac{a_0^{\mathrm{T}}a_1}{\|a_0\|} + \epsilon^*\frac{a_0^{\mathrm{T}}a_2}{\|a_0\|} + \epsilon\epsilon^*(\frac{a_0^{\mathrm{T}}a_3 + a_1^{\mathrm{T}}a_2}{\|a_0\|} - \frac{a_0^{\mathrm{T}}a_1a_0^{\mathrm{T}}a_2}{\|a_0\|^3}).$$

A matrix with hyper-dual number entries is called a *hyper-dual matrix*. Analogous to the forms of hyper-dual numbers, an $m \times n$ hyper-dual matrix $\widetilde{A}$ is defined as

$$\widetilde{A} = A_0 + \epsilon A_1 + \epsilon^* A_2 + \epsilon\epsilon^* A_3 = (A_0 + \epsilon A_1) + \epsilon^*(A_2 + \epsilon A_3) \triangleq \widehat{A} + \epsilon^*\widehat{A}_0,$$

where A_0–A_3 are $m \times n$ real matrices, and ϵ and ϵ^* are dual units. The set of all $m \times n$ hyper-dual matrices over the real field is denoted by $\widetilde{\mathbb{R}}^{m\times n}$. Some studies on hyper-dual matrices can be found in [24, 25].

For a given hyper-dual matrix $\widetilde{A} \in \widetilde{\mathbb{R}}^{m\times n}$, if there exists a hyper-dual matrix $\widetilde{X} \in \widetilde{\mathbb{R}}^{n\times m}$ satisfying

$$\widetilde{A}\widetilde{X}\widetilde{A} = \widetilde{A},\ \ \widetilde{X}\widetilde{A}\widetilde{X} = \widetilde{X},\ \ (\widetilde{A}\widetilde{X})^{\mathrm{T}} = \widetilde{A}\widetilde{X},\ \ (\widetilde{X}\widetilde{A})^{\mathrm{T}} = \widetilde{X}\widetilde{A}, \tag{1.4}$$

then we call $\widetilde{X}$ the *hyper-dual Moore-Penrose generalized inverse* (HDMPGI) of $\widetilde{A}$, and denoted by $\widetilde{A}^{\dagger}$.

In this paper, we aim to give some theoretical findings of HDMPGI. The rest of this paper is organized as follows. In Section 2, we give some necessary and sufficient conditions for a hyper-dual matrix to have the HDMPGI, and present a compact formula for HDMPGI whenever it exists. In Section 3, analogous to the applications of the dual Moore-Penrose generalized inverse in linear dual equations, we discuss the least-squares properties of HDMPGI. In Section 4, based on the forms of dual matrices and hyper-dual matrices, we introduce the definition of dual matrix of order n. We also study the existence of the Moore-Penrose generalized inverse of such matrices. The theoretical results are illustrated by some numerical examples.

Throughout this paper, we use $\mathbb{R}^n$, $\widehat{\mathbb{R}}^n$, and $\widetilde{\mathbb{R}}^n$ to denote the set of all n-dimensional real column vectors, dual column vectors, and hyper-dual column vectors, respectively. $\mathbb{R}^{m\times n}$, $\widehat{\mathbb{R}}^{m\times n}$, and $\widetilde{\mathbb{R}}^{m\times n}$ are, respectively, the set of all $m\times n$ real matrices, dual matrices, and hyper-dual matrices. For a real matrix A, $r(A)$ is the rank of A, the superscript "T" is the transpose of a matrix, and I_n is the identity of order n. $\|\cdot\|$ is the Euclidean norm of a vector. We will use

$$G \triangleq \cdots$$

to mean that we define G to be something.

The following lemma is well-known as singular value decomposition, which will be a basic tool for proving Theorem 2.1.

Lemma 1.1. [26] Let $A \in \mathbb{R}^{m\times n}$ be such that

$$r(A) = r.$$

Then, there exist real orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such that

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^{\mathrm{T}},$$

where $\Sigma \in \mathbb{R}^{r\times r}$ is a diagonal positive definite matrix. Then,

$$A^{\dagger} = V \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}}.$$

The following lemma will also be used in the proof of Theorem 2.1, which is a rank equality that involves a special 2×2 block matrix and Moore-Penrose generalized inverse.

Lemma 1.2. [27] *Let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times k}$, and $C \in \mathbb{R}^{l\times n}$. Then,*

$$r\begin{bmatrix} A & B \\ C & 0 \end{bmatrix} = r(B) + r(C) + r\left[(I_m - BB^{\dagger})A(I_n - C^{\dagger}C)\right].$$

2. Characterizations of HDMPGI of hyper-dual matrices

In this section, we study the existence and computation of the HDMPGI. We first give a necessary and sufficient condition for a hyper-dual matrix to be the HDMPGI of a given hyper-dual matrix, which can be obtained directly from the definition of the HDMPGI in (1.4), and we omit the proof.

Lemma 2.1. *Let*

$$\widetilde{A} = \widehat{A} + \epsilon^{*}\widehat{A_0} \in \widetilde{\mathbb{R}}^{m\times n}.$$

Then, a hyper-dual matrix

$$\widetilde{X} = \widehat{X} + \epsilon^{*}\widehat{X_0} \in \widetilde{\mathbb{R}}^{n\times m}$$

is the HDMPGI of $\widetilde{A}$ if and only if

$$\widehat{X} = \widehat{A}^{\dagger}$$

and

$$\begin{cases} \widehat{A}\widehat{X}\widehat{A_0} + \widehat{A}\widehat{X_0}\widehat{A} + \widehat{A_0}\widehat{X}\widehat{A} = \widehat{A_0}, \\ \widehat{X}\widehat{A}\widehat{X_0} + \widehat{X}\widehat{A_0}\widehat{X} + \widehat{X_0}\widehat{A}\widehat{X} = \widehat{X_0}, \\ (\widehat{A}\widehat{X_0} + \widehat{A_0}\widehat{X})^{\mathrm{T}} = \widehat{A}\widehat{X_0} + \widehat{A_0}\widehat{X}, \\ (\widehat{X}\widehat{A_0} + \widehat{X_0}\widehat{A})^{\mathrm{T}} = \widehat{X}\widehat{A_0} + \widehat{X_0}\widehat{A}. \end{cases}$$

Analogous to the DMPGI of dual matrices, the HDMPGI of hyper-dual matrices may not exist. We present some necessary and sufficient conditions for the existence of the HDMPGI in the following theorem. A compact formula for the computation of the HDMPGI is also given whenever it exists.

Theorem 2.1. *Let*

$$\widetilde{A} = \widehat{A} + \epsilon^* \widehat{A_0} = A_0 + \epsilon A_1 + \epsilon^* A_2 + \epsilon\epsilon^* A_3 \in \widetilde{\mathbb{R}}^{m\times n}.$$

Then, the following statements are equivalent:

(i) *The HDMPGI of* $\widetilde{A}$ *exists;*

(ii)

$$\widetilde{A} = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} R_1 & R_2 \\ R_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon^*\left(U\begin{bmatrix} Y_1 & Y_2 \\ Y_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix}V^{\mathrm{T}}\right),$$

where U *and* V *are real orthogonal matrices of orders* m *and* n*, respectively,* Σ *is a diagonal positive definite matrix, and* R_1*–*R_3*,* Y_1*–*Y_3*,* Z_1*–*Z_4 *are real matrices of appropriate sizes that satisfy*

$$Z_4 = R_3\Sigma^{-1}Y_2 + Y_3\Sigma^{-1}R_2;$$

(iii) $\widehat{A}^{\dagger}$ *exists, and*

$$(I_m - \widehat{A}\widehat{A}^{\dagger})\widehat{A_0}(I_n - \widehat{A}^{\dagger}\widehat{A}) = 0;$$

(iv)

$$\begin{aligned}(I_m - A_0A_0^{\dagger})A_1(I_n - A_0^{\dagger}A_0) &= (I_m - A_0A_0^{\dagger})A_2(I_n - A_0^{\dagger}A_0)\\ &= (I_m - A_0A_0^{\dagger})(A_3 - A_2A_0^{\dagger}A_1 - A_1A_0^{\dagger}A_2)(I_n - A_0^{\dagger}A_0)\\ &= 0;\end{aligned}$$

(v)

$$r\begin{bmatrix} A_1 & A_0 \\ A_0 & 0 \end{bmatrix} = r\begin{bmatrix} A_2 & A_0 \\ A_0 & 0 \end{bmatrix} = r\begin{bmatrix} A_3 - A_2A_0^{\dagger}A_1 - A_1A_0^{\dagger}A_2 & A_0 \\ A_0 & 0 \end{bmatrix} = 2r(A_0).$$

Furthermore, if the HDMPGI of $\widetilde{A}$ *exists, then*

$$\widetilde{A}^{\dagger} = \widehat{A}^{\dagger} + \epsilon^*\left[-\widehat{A}^{\dagger}\widehat{A_0}\widehat{A}^{\dagger} + (\widehat{A}^{\mathrm{T}}\widehat{A})^{\dagger}\widehat{A}_0^{\mathrm{T}}(I_m - \widehat{A}\widehat{A}^{\dagger}) + (I_n - \widehat{A}^{\dagger}\widehat{A})\widehat{A}_0^{\mathrm{T}}(\widehat{A}\widehat{A}^{\mathrm{T}})^{\dagger}\right]. \tag{2.1}$$

Proof. In order to show the equivalence of the five items, we will prove that (i)⇔(ii), (ii)⇔(iii), (iii)⇔(iv), and (iv)⇔(v).

(i)⇔(ii): If the HDMPGI of

$$\widetilde{A} = \widehat{A} + \epsilon^*\widehat{A_0}$$

exists, then we may assume that

$$\widetilde{A}^{\dagger} = \widehat{X} + \epsilon^*\widehat{X_0}.$$

It follows from Lemma 2.1 that the DMPGI of $\widehat{A}$ exists and

$$\widehat{X} = \widehat{A}^{\dagger}.$$

Then, by [16], using the singular value decomposition of real matrices in Lemma 1.1, $\widehat{A}$ and $\widehat{X}$ have the forms

$$\widehat{A} = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} R_1 & R_2 \\ R_3 & 0 \end{bmatrix}V^{\mathrm{T}} \tag{2.2}$$

and

$$\widehat{X} = V\begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} -\Sigma^{-1}R_1\Sigma^{-1} & \Sigma^{-2}R_3^{\mathrm{T}} \\ R_2^{\mathrm{T}}\Sigma^{-2} & 0 \end{bmatrix}U^{\mathrm{T}}, \tag{2.3}$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are real orthogonal matrices, $\Sigma \in \mathbb{R}^{r\times r}$ is a diagonal positive definite matrix,

$$r = r(A_0),$$

and R_1–R_3 are real matrices of appropriate sizes.

Let

$$\begin{aligned}\widehat{A}_0 &= U\begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix}V^{\mathrm{T}},\\ \widehat{X}_0 &= V\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} W_1 & W_2 \\ W_3 & W_4 \end{bmatrix}U^{\mathrm{T}}.\end{aligned}$$

Then, a direct calculation shows that

$$\begin{aligned}\widehat{A}\widehat{X}\widehat{A}_0 &= U\begin{bmatrix} Y_1 & Y_2 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 + \Sigma^{-1}R_3^{\mathrm{T}}Y_3 & Z_2 + \Sigma^{-1}R_3^{\mathrm{T}}Y_4 \\ R_3\Sigma^{-1}Y_1 & R_3\Sigma^{-1}Y_2 \end{bmatrix}V^{\mathrm{T}},\\ \widehat{A}\widehat{X}_0\widehat{A} &= U\begin{bmatrix} \Sigma X_1\Sigma & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} \Theta & \Sigma X_1 R_2 \\ R_3X_1\Sigma & 0 \end{bmatrix}V^{\mathrm{T}},\end{aligned}$$

where

$$\Theta = \Sigma X_1R_1 + \Sigma X_2R_3 + \Sigma W_1\Sigma + R_1X_1\Sigma + R_2X_3\Sigma.$$

$$\widehat{A}_0\widehat{X}\widehat{A} = U\begin{bmatrix} Y_1 & 0 \\ Y_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 + Y_2R_2^{\mathrm{T}}\Sigma^{-1} & Y_1\Sigma^{-1}R_2 \\ Z_3 + Y_4R_2^{\mathrm{T}}\Sigma^{-1} & Y_3\Sigma^{-1}R_2 \end{bmatrix}V^{\mathrm{T}}.$$

Hence,

$$\widehat{A}\widehat{X}\widehat{A}_0 + \widehat{A}\widehat{X}_0\widehat{A} + \widehat{A}_0\widehat{X}\widehat{A} = U\begin{bmatrix} 2Y_1 + \Sigma X_1\Sigma & Y_2 \\ Y_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} \Gamma_1 & \Gamma_2 \\ \Gamma_3 & \Gamma_4 \end{bmatrix}V^{\mathrm{T}},$$

where

$$\begin{cases} \Gamma_1 = 2Z_1 + \Sigma^{-1}R_3^{\mathrm{T}}Y_3 + Y_2R_2^{\mathrm{T}}\Sigma^{-1} + \Theta,\\ \Gamma_2 = Z_2 + \Sigma^{-1}R_3^{\mathrm{T}}Y_4 + \Sigma X_1R_2 + Y_1\Sigma^{-1}R_2,\\ \Gamma_3 = Z_3 + Y_4R_2^{\mathrm{T}}\Sigma^{-1} + R_3\Sigma^{-1}Y_1 + R_3X_1\Sigma,\\ \Gamma_4 = R_3\Sigma^{-1}Y_2 + Y_3\Sigma^{-1}R_2. \end{cases}$$

According to Lemma 2.1,

$$\widehat{A}\widehat{X}\widehat{A}_0 + \widehat{A}\widehat{X}_0\widehat{A} + \widehat{A}_0\widehat{X}\widehat{A} = \widehat{A}_0,$$

i.e.,

$$U\begin{bmatrix} 2Y_1 + \Sigma X_1\Sigma & Y_2 \\ Y_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} \Gamma_1 & \Gamma_2 \\ \Gamma_3 & \Gamma_4 \end{bmatrix}V^{\mathrm{T}} = U\begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix}V^{\mathrm{T}}.$$

Equating the real part and the dual part of both sides of the above equality yields

$$Y_4 = 0$$

and

$$\Gamma_4 = R_3\Sigma^{-1}Y_2 + Y_3\Sigma^{-1}R_2 = Z_4.$$

Therefore, $\widetilde{A}$ has the form

$$\widetilde{A} = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} R_1 & R_2 \\ R_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon^*\left(U\begin{bmatrix} Y_1 & Y_2 \\ Y_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix}V^{\mathrm{T}}\right).$$

Conversely, if

$$\widetilde{A} = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} R_1 & R_2 \\ R_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon^*\left(U\begin{bmatrix} Y_1 & Y_2 \\ Y_3 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix}V^{\mathrm{T}}\right),$$

where U and V are real orthogonal matrices of orders m and n, respectively, Σ is a diagonal positive definite matrix, and

$$Z_4 = R_3\Sigma^{-1}Y_2 + Y_3\Sigma^{-1}R_2.$$

Let

$$\begin{aligned}\widetilde{G} =& V\begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} -\Sigma^{-1}R_1\Sigma^{-1} & \Sigma^{-2}R_3^{\mathrm{T}} \\ R_2^{\mathrm{T}}\Sigma^{-2} & 0 \end{bmatrix}U^{\mathrm{T}} \\ &+ \epsilon^*\left(V\begin{bmatrix} -\Sigma^{-1}Y_1\Sigma^{-1} & \Sigma^{-2}Y_3^{\mathrm{T}} \\ Y_2^{\mathrm{T}}\Sigma^{-2} & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix}U^{\mathrm{T}}\right),\end{aligned} \tag{2.4}$$

where

$$\begin{cases} M_1 = -\Sigma^{-2}R_3^{\mathrm{T}}Y_3\Sigma^{-1} - \Sigma^{-1}Y_2R_2^{\mathrm{T}}\Sigma^{-2} - \Sigma^{-2}Y_3^{\mathrm{T}}R_3\Sigma^{-1} - \Sigma^{-1}R_2Y_2^{\mathrm{T}}\Sigma^{-2}, \\ M_2 = \Sigma^{-2}Z_3^{\mathrm{T}} - \Sigma^{-2}R_1^{\mathrm{T}}\Sigma^{-1}Y_3^{\mathrm{T}} - \Sigma^{-1}R_1\Sigma^{-2}Y_3^{\mathrm{T}} - \Sigma^{-2}Y_1^{\mathrm{T}}\Sigma^{-1}R_3^{\mathrm{T}} - \Sigma^{-1}Y_1\Sigma^{-2}R_3^{\mathrm{T}}, \\ M_3 = Z_2^{\mathrm{T}}\Sigma^{-2} - R_2^{\mathrm{T}}\Sigma^{-1}Y_1^{\mathrm{T}}\Sigma^{-2} - Y_2^{\mathrm{T}}\Sigma^{-2}R_1\Sigma^{-1} - Y_2^{\mathrm{T}}\Sigma^{-1}R_1^{\mathrm{T}}\Sigma^{-2} - R_2^{\mathrm{T}}\Sigma^{-2}Y_1\Sigma^{-1}, \\ M_4 = R_2^{\mathrm{T}}\Sigma^{-3}Y_3^{\mathrm{T}} + Y_2^{\mathrm{T}}\Sigma^{-3}R_3^{\mathrm{T}}. \end{cases}$$

Then,

$$\widetilde{A}\widetilde{G} = U\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon U\begin{bmatrix} 0 & \Sigma^{-1}R_3^{\mathrm{T}} \\ R_3\Sigma^{-1} & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon^* U\begin{bmatrix} 0 & \Sigma^{-1}Y_3^{\mathrm{T}} \\ Y_3\Sigma^{-1} & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon\epsilon^* U\begin{bmatrix} N_1 & N_2 \\ N_3 & N_4 \end{bmatrix}U^{\mathrm{T}},$$

where

$$\begin{cases} N_1 = -\Sigma^{-1}R_3^{\mathrm{T}}Y_3\Sigma^{-1} - \Sigma^{-1}Y_3^{\mathrm{T}}R_3\Sigma^{-1}, \\ N_2 = \Sigma^{-1}Z_3^{\mathrm{T}} - \Sigma^{-1}R_1^{\mathrm{T}}\Sigma^{-1}Y_3^{\mathrm{T}} - \Sigma^{-1}Y_1^{\mathrm{T}}\Sigma^{-1}R_3^{\mathrm{T}}, \\ N_3 = Z_3\Sigma^{-1} - Y_3\Sigma^{-1}R_1\Sigma^{-1} - R_3\Sigma^{-1}Y_1\Sigma^{-1}, \\ N_4 = Y_3\Sigma^{-2}R_3^{\mathrm{T}} + R_3\Sigma^{-2}Y_3^{\mathrm{T}}. \end{cases}$$

$$\widetilde{G}\widetilde{A} = V\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon V\begin{bmatrix} 0 & \Sigma^{-1}R_2 \\ R_2^{\mathrm{T}}\Sigma^{-1} & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon^* V\begin{bmatrix} 0 & \Sigma^{-1}Y_2 \\ Y_2^{\mathrm{T}}\Sigma^{-1} & 0 \end{bmatrix}V^{\mathrm{T}} + \epsilon\epsilon^* V\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix}V^{\mathrm{T}},$$

where

$$\begin{cases} P_1 = -\Sigma^{-1}Y_2R_2^{\mathrm{T}}\Sigma^{-1} - \Sigma^{-1}R_2Y_2^{\mathrm{T}}\Sigma^{-1}, \\ P_2 = \Sigma^{-1}Z_2 - \Sigma^{-1}Y_1\Sigma^{-1}R_2 - \Sigma^{-1}R_1\Sigma^{-1}Y_2, \\ P_3 = Z_2^{\mathrm{T}}\Sigma^{-1} - R_2^{\mathrm{T}}\Sigma^{-1}Y_1^{\mathrm{T}}\Sigma^{-1} - Y_2^{\mathrm{T}}\Sigma^{-1}R_1^{\mathrm{T}}\Sigma^{-1}, \\ P_4 = Y_2^{\mathrm{T}}\Sigma^{-2}R_2 + R_2^{\mathrm{T}}\Sigma^{-2}Y_2. \end{cases}$$

Then, $\widetilde{A}\widetilde{G}$ and $\widetilde{G}\widetilde{A}$ are symmetric.

Furthermore,

$$\begin{aligned}
\widetilde{A}\widetilde{G}\widetilde{A} =& U\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix}0 & 0\\ R_3 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon^* U\begin{bmatrix}0 & 0\\ Y_3 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon\epsilon^* U\begin{bmatrix}-\Sigma^{-1}R_3^{\mathrm{T}}Y_3 - \Sigma^{-1}Y_3^{\mathrm{T}}R_3 & 0\\ Z_3 - Y_3\Sigma^{-1}R_1 - R_3\Sigma^{-1}Y_1 & 0\end{bmatrix}V^{\mathrm{T}} \\
&+ \epsilon U\begin{bmatrix}R_1 & R_2\\ 0 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon\epsilon^* U\begin{bmatrix}\Sigma^{-1}Y_3^{\mathrm{T}}R_3 & 0\\ Y_3\Sigma^{-1}R_1 & Y_3\Sigma^{-1}R_2\end{bmatrix}V^{\mathrm{T}} + \epsilon\epsilon^* U\begin{bmatrix}\Sigma^{-1}R_3^{\mathrm{T}}Y_3 & 0\\ R_3\Sigma^{-1}Y_1 & R_3\Sigma^{-1}Y_2\end{bmatrix}V^{\mathrm{T}} \\
&+ \epsilon^* U\begin{bmatrix}Y_1 & Y_2\\ 0 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon\epsilon^* U\begin{bmatrix}Z_1 & Z_2\\ 0 & 0\end{bmatrix}V^{\mathrm{T}} \\
=& U\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix}R_1 & R_2\\ R_3 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon^* U\begin{bmatrix}Y_1 & Y_2\\ Y_3 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon\epsilon^* U\begin{bmatrix}Z_1 & Z_2\\ Z_3 & Z_4\end{bmatrix}V^{\mathrm{T}} \\
=& \widetilde{A}, \\
\widetilde{G}\widetilde{A}\widetilde{G} =& V\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix}0 & 0\\ R_2^{\mathrm{T}}\Sigma^{-2} & 0\end{bmatrix}U^{\mathrm{T}} + \epsilon^* V\begin{bmatrix}0 & 0\\ Y_2^{\mathrm{T}}\Sigma^{-2} & 0\end{bmatrix}U^{\mathrm{T}} \\
&+ \epsilon\epsilon^* V\begin{bmatrix}-\Sigma^{-1}Y_2R_2^{\mathrm{T}}\Sigma^{-2} - \Sigma^{-1}R_2Y_2^{\mathrm{T}}\Sigma^{-2} & 0\\ Z_2^{\mathrm{T}}\Sigma^{-2} - R_2^{\mathrm{T}}\Sigma^{-1}Y_1^{\mathrm{T}}\Sigma^{-2} - Y_2^{\mathrm{T}}\Sigma^{-1}R_1^{\mathrm{T}}\Sigma^{-2} & 0\end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix}-\Sigma^{-1}R_1\Sigma^{-1} & \Sigma^{-2}R_3^{\mathrm{T}}\\ 0 & 0\end{bmatrix}U^{\mathrm{T}} \\
&+ \epsilon\epsilon^* V\begin{bmatrix}\Sigma^{-1}Y_2R_2^{\mathrm{T}}\Sigma^{-2} & 0\\ -Y_2^{\mathrm{T}}\Sigma^{-1}R_1^{\mathrm{T}}\Sigma^{-1} & Y_2^{\mathrm{T}}\Sigma^{-3}R_3^{\mathrm{T}}\end{bmatrix}U^{\mathrm{T}} + \epsilon^* V\begin{bmatrix}-\Sigma^{-1}Y_1\Sigma^{-1} & \Sigma^{-2}Y_3^{\mathrm{T}}\\ 0 & 0\end{bmatrix}U^{\mathrm{T}} \\
&+ \epsilon\epsilon^* V\begin{bmatrix}\Sigma^{-1}R_2Y_2^{\mathrm{T}}\Sigma^{-2} & 0\\ -R_2^{\mathrm{T}}\Sigma^{-1}Y_1^{\mathrm{T}}\Sigma^{-1} & R_2^{\mathrm{T}}\Sigma^{-3}Y_3^{\mathrm{T}}\end{bmatrix}U^{\mathrm{T}} + \epsilon\epsilon^* V\begin{bmatrix}M_1 & M_2\\ 0 & 0\end{bmatrix}U^{\mathrm{T}} \\
=& V\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix}-\Sigma^{-1}R_1\Sigma^{-1} & \Sigma^{-2}R_3^{\mathrm{T}}\\ R_2^{\mathrm{T}}\Sigma^{-2} & 0\end{bmatrix}U^{\mathrm{T}} \\
&+ \epsilon^* V\begin{bmatrix}-\Sigma^{-1}Y_1\Sigma^{-1} & \Sigma^{-2}Y_3^{\mathrm{T}}\\ Y_2^{\mathrm{T}}\Sigma^{-2} & 0\end{bmatrix}U^{\mathrm{T}} + \epsilon\epsilon^* V\begin{bmatrix}M_1 & M_2\\ M_3 & M_4\end{bmatrix}U^{\mathrm{T}} \\
=& \widetilde{G}.
\end{aligned}$$

Hence, $\widetilde{A}$ and $\widetilde{G}$ satisfy the four Penrose conditions in (1.4), i.e., $\widetilde{G}$ is the HDMPGI of $\widetilde{A}$.

(ii)⇔(iii): If

$$\widetilde{A} = \widehat{A} + \epsilon^*\widehat{A}_0,$$

where

$$\widehat{A} = U\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix}R_1 & R_2\\ R_3 & 0\end{bmatrix}V^{\mathrm{T}}, \quad \widehat{A}_0 = U\begin{bmatrix}Y_1 & Y_2\\ Y_3 & 0\end{bmatrix}V^{\mathrm{T}} + \epsilon U\begin{bmatrix}Z_1 & Z_2\\ Z_3 & Z_4\end{bmatrix}V^{\mathrm{T}},$$

then by [16], the DMPGI of $\widehat{A}$ exists and $\widehat{A}^\dagger$ has the matrix form in (2.3). Substituting the matrix forms of $\widehat{A}$, $\widehat{A}^\dagger$, and $\widehat{A}_0$ into $(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A})$, we obtain

$$
\begin{aligned}
(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) =& \left(U \begin{bmatrix} 0 & 0 \\ 0 & I_{m-r} \end{bmatrix} U^{\mathrm{T}} - \epsilon U \begin{bmatrix} 0 & \Sigma^{-1}R_3^{\mathrm{T}} \\ R_3\Sigma^{-1} & 0 \end{bmatrix} U^{\mathrm{T}}\right) \\
&\times \left(U \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & 0 \end{bmatrix} V^{\mathrm{T}} + \epsilon U \begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix} V^{\mathrm{T}}\right) \\
&\times \left(V \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix} V^{\mathrm{T}} - \epsilon V \begin{bmatrix} 0 & \Sigma^{-1}R_2 \\ R_2^{\mathrm{T}}\Sigma^{-1} & 0 \end{bmatrix} V^{\mathrm{T}}\right) \\
=&\epsilon U \begin{bmatrix} 0 & 0 \\ 0 & Z_4 - R_3\Sigma^{-1}Y_2 - Y_3\Sigma^{-1}R_2 \end{bmatrix} V^{\mathrm{T}}.
\end{aligned}
$$

Therefore, if

$$Z_4 = R_3\Sigma^{-1}Y_2 + Y_3\Sigma^{-1}R_2,$$

then

$$(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) = 0.$$

On the other hand, if $\widehat{A}^\dagger$ exists, then $\widehat{A}$ and $\widehat{A}^\dagger$ have the matrix forms in (2.2) and (2.3), respectively. By a direct calculation, we have

$$(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) = U \begin{bmatrix} 0 & 0 \\ 0 & Y_4 \end{bmatrix} V^{\mathrm{T}} + \epsilon U \begin{bmatrix} 0 & -\Sigma^{-1}R_3^{\mathrm{T}}Y_4 \\ -Y_4R_2^{\mathrm{T}}\Sigma^{-1} & Z_4 - R_3\Sigma^{-1}Y_2 - Y_3\Sigma^{-1}R_2 \end{bmatrix} V^{\mathrm{T}}.$$

Hence, if

$$(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) = 0,$$

then $Y_4 = 0$ and

$$Z_4 = R_3\Sigma^{-1}Y_2 + Y_3\Sigma^{-1}R_2.$$

(iii)⇔(iv): By [16], if $\widehat{A}^\dagger$ exists, then

$$(I_m - A_0A_0^\dagger)A_1(I_n - A_0^\dagger A_0) = 0.$$

Moreover, if

$$(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) = 0,$$

then substituting

$$\widehat{A} = A_0 + \epsilon A_1, \quad \widehat{A}_0 = A_2 + \epsilon A_3$$

and

$$\widehat{A}^\dagger = A_0^\dagger + \epsilon\left[-A_0^\dagger A_1 A_0^\dagger + (A_0^{\mathrm{T}}A_0)^\dagger A_1^{\mathrm{T}}(I_m - A_0A_0^\dagger) + (I_n - A_0^\dagger A_0)A_1^{\mathrm{T}}(A_0A_0^{\mathrm{T}})^\dagger\right]$$

into

$$(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) = 0$$

gives

$$(I_m - A_0A_0^\dagger)A_2(I_n - A_0^\dagger A_0) + \epsilon[(I_m - A_0A_0^\dagger)(A_3 - A_2A_0^\dagger A_1 - A_1A_0^\dagger A_2)(I_n - A_0^\dagger A_0)$$

$$
\begin{aligned}
&- (I_m - A_0A_0^\dagger)A_2(I_n - A_0^\dagger A_0)A_1^{\mathrm{T}}(A_0A_0^{\mathrm{T}})^\dagger A_0 - A_0(A_0^{\mathrm{T}}A_0)^\dagger A_1^{\mathrm{T}}(I_m - A_0A_0^\dagger)A_2(I_n - A_0^\dagger A_0)] \\
&= 0,
\end{aligned}
$$

which implies

$$(I_m - A_0A_0^\dagger)A_2(I_n - A_0^\dagger A_0) = 0$$

and

$$(I_m - A_0A_0^\dagger)(A_3 - A_2A_0^\dagger A_1 - A_1A_0^\dagger A_2)(I_n - A_0^\dagger A_0) = 0.$$

Conversely, if

$$(I_m - A_0A_0^\dagger)A_1(I_n - A_0^\dagger A_0) = 0,$$

then by [16], $\widehat{A}^\dagger$ exists. Moreover, if

$$(I_m - A_0A_0^\dagger)A_2(I_n - A_0^\dagger A_0) = 0$$

and

$$(I_m - A_0A_0^\dagger)(A_3 - A_2A_0^\dagger A_1 - A_1A_0^\dagger A_2)(I_n - A_0^\dagger A_0) = 0,$$

then it is not difficult to see that

$$(I_m - \widehat{A}\widehat{A}^\dagger)\widehat{A}_0(I_n - \widehat{A}^\dagger\widehat{A}) = 0.$$

(iv)⇔(v): It follows directly from Lemma 1.2.

It remains to show that

$$\widetilde{G} = \widehat{A}^\dagger + \epsilon^* \left[-\widehat{A}^\dagger\widehat{A}_0\widehat{A}^\dagger + (\widehat{A}^{\mathrm{T}}\widehat{A})^\dagger\widehat{A}_0^{\mathrm{T}}(I_m - \widehat{A}\widehat{A}^\dagger) + (I_n - \widehat{A}^\dagger\widehat{A})\widehat{A}_0^{\mathrm{T}}(\widehat{A}\widehat{A}^{\mathrm{T}})^\dagger\right].$$

By a direct calculation, we have

$$\widehat{A}^\dagger = V\begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} -\Sigma^{-1}R_1\Sigma^{-1} & \Sigma^{-2}R_3^{\mathrm{T}} \\ R_2^{\mathrm{T}}\Sigma^{-2} & 0 \end{bmatrix}U^{\mathrm{T}}, \tag{2.5}$$

$$\widehat{A}^\dagger\widehat{A}_0\widehat{A}^\dagger = V\begin{bmatrix} \Sigma^{-1}Y_1\Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} Q_1 & \Sigma^{-1}Y_1\Sigma^{-2}R_3^{\mathrm{T}} \\ R_2^{\mathrm{T}}\Sigma^{-2}Y_1\Sigma^{-1} & 0 \end{bmatrix}U^{\mathrm{T}}, \tag{2.6}$$

where

$$Q_1 = \Sigma^{-1}(-R_1\Sigma^{-1}Y_1\Sigma^{-1} - Y_1\Sigma^{-1}R_1\Sigma^{-1} + Z_1\Sigma^{-1} + \Sigma^{-1}R_3^{\mathrm{T}}Y_3\Sigma^{-1} + Y_2R_2^{\mathrm{T}}\Sigma^{-2}).$$

$$(\widehat{A}^{\mathrm{T}}\widehat{A})^\dagger\widehat{A}_0^{\mathrm{T}}(I_m - \widehat{A}\widehat{A}^\dagger) = V\begin{bmatrix} 0 & \Sigma^{-2}Y_3^{\mathrm{T}} \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} -\Sigma^{-2}Y_3^{\mathrm{T}}R_3\Sigma^{-1} & Q_2 \\ 0 & R_2^{\mathrm{T}}\Sigma^{-3}Y_3^{\mathrm{T}} \end{bmatrix}U^{\mathrm{T}},$$

where

$$Q_2 = \Sigma^{-2}Z_3^{\mathrm{T}} - \Sigma^{-2}R_1^{\mathrm{T}}\Sigma^{-1}Y_3^{\mathrm{T}} - \Sigma^{-1}R_1\Sigma^{-2}Y_3^{\mathrm{T}} - \Sigma^{-2}Y_1^{\mathrm{T}}\Sigma^{-1}R_3^{\mathrm{T}}.$$

$$(I_n - \widehat{A}^\dagger\widehat{A})\widehat{A}_0^{\mathrm{T}}(\widehat{A}\widehat{A}^{\mathrm{T}})^\dagger = V\begin{bmatrix} 0 & 0 \\ Y_2^{\mathrm{T}}\Sigma^{-2} & 0 \end{bmatrix}U^{\mathrm{T}} + \epsilon V\begin{bmatrix} -\Sigma^{-1}R_2Y_2^{\mathrm{T}}\Sigma^{-2} & 0 \\ Q_3 & Y_2^{\mathrm{T}}\Sigma^{-3}R_3^{\mathrm{T}} \end{bmatrix}U^{\mathrm{T}}, \tag{2.7}$$

where

$$Q_3 = Z_2^{\mathrm{T}}\Sigma^{-2} - R_2^{\mathrm{T}}\Sigma^{-1}Y_1^{\mathrm{T}}\Sigma^{-2} - Y_2^{\mathrm{T}}\Sigma^{-2}R_1\Sigma^{-1} - Y_2^{\mathrm{T}}\Sigma^{-1}R_1^{\mathrm{T}}\Sigma^{-2}.$$

Now, it can be seen from (2.4)–(2.7) that

$$\widetilde{G} = \widehat{A}^{\dagger} + \epsilon^* \left[-\widehat{A}^{\dagger}\widehat{A}_0\widehat{A}^{\dagger} + (\widehat{A}^{\mathrm{T}}\widehat{A})^{\dagger}\widehat{A}_0^{\mathrm{T}}(I_m - \widehat{A}\widehat{A}^{\dagger}) + (I_n - \widehat{A}^{\dagger}\widehat{A})\widehat{A}_0^{\mathrm{T}}(\widehat{A}\widehat{A}^{\mathrm{T}})^{\dagger}\right],$$

and thus $\widetilde{A}^{\dagger}$ has the expression in (2.1). □

Remark that we can know whether the HDMPGI of a hyper-dual matrix exists by checking one of the four conditions in Theorem 2.1, especially by condition (v). Once the HDMPGI exists, we can obtain it by the formula given in (2.1). We illustrate this by the following example:

Example 2.1. Let

$$\widetilde{A} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \epsilon \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + \epsilon^* \begin{bmatrix} 0 & 0 \\ 2 & 2 \end{bmatrix} + \epsilon\epsilon^* \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} \triangleq A_0 + \epsilon A_1 + \epsilon^* A_2 + \epsilon\epsilon^* A_3.$$

Since

$$r\begin{bmatrix} A_1 & A_0 \\ A_0 & 0 \end{bmatrix} = r\begin{bmatrix} A_2 & A_0 \\ A_0 & 0 \end{bmatrix} = r\begin{bmatrix} A_3 - A_2A_0^{\dagger}A_1 - A_1A_0^{\dagger}A_2 & A_0 \\ A_0 & 0 \end{bmatrix} = 2 = 2r(A_0),$$

then by Theorem 2.1(v), the HDMPGI of $\widetilde{A}$ exists.

A direct computation shows that

$$\widehat{A}^{\dagger} = (A_0 + \epsilon A_1)^{\dagger} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix} + \epsilon \begin{bmatrix} -1 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

and

$$\begin{aligned} \widetilde{A}^{\dagger} &= \widehat{A}^{\dagger} + \epsilon^* \left[-\widehat{A}^{\dagger}\widehat{A}_0\widehat{A}^{\dagger} + (\widehat{A}^{\mathrm{T}}\widehat{A})^{\dagger}\widehat{A}_0^{\mathrm{T}}(I_m - \widehat{A}\widehat{A}^{\dagger}) + (I_n - \widehat{A}^{\dagger}\widehat{A})\widehat{A}_0^{\mathrm{T}}(\widehat{A}\widehat{A}^{\mathrm{T}})^{\dagger}\right] \\ &= \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix} + \epsilon \begin{bmatrix} -1 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + \epsilon^* \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + \epsilon\epsilon^* \begin{bmatrix} -\frac{5}{2} & -\frac{5}{2} \\ -\frac{3}{2} & -\frac{3}{2} \end{bmatrix}. \end{aligned}$$

3. Least-squares properties of HDMPGI

Qi et al. [28] introduced a total order $\leq$ over $\widehat{\mathbb{R}}$. Suppose

$$\widehat{p} = p + \epsilon p_0, \quad \widehat{q} = q + \epsilon q_0 \in \widehat{\mathbb{R}}.$$

We have $\widehat{p} < \widehat{q}$ if $p < q$, or $p = q$ and $p_0 < q_0$; $\widehat{p} = \widehat{q}$ if $p = q$ and $p_0 = q_0$. The total order provides an efficient way to compare the magnitude of two dual numbers. Based on the total order $\leq$ over $\widehat{\mathbb{R}}$, Wang et al. [29] extended it to dual vectors and introduced a QLY total order $\overset{Q}{\leq}$ over $\widehat{\mathbb{R}}^m$. We introduce a total order over $\widetilde{\mathbb{R}}$ as follows. For two hyper-dual numbers

$$\widetilde{p} = \widehat{p} + \epsilon^*\widehat{p}_0, \quad \widetilde{q} = \widehat{q} + \epsilon^*\widehat{q}_0 \in \widetilde{\mathbb{R}}.$$

We have $\widetilde{p} < \widetilde{q}$ if $\widehat{p} < \widehat{q}$, or $\widehat{p} = \widehat{q}$ and $\widehat{p_0} < \widehat{q_0}$; $\widetilde{p} = \widetilde{q}$ if $\widehat{p} = \widehat{q}$ and $\widehat{p_0} = \widehat{q_0}$. If $\widetilde{a} > 0$, then we say that $\widetilde{a}$ is a positive hyper-dual number. If $\widetilde{a} \geq 0$, then we call $\widetilde{a}$ a nonnegative hyper-dual number.

Recall that for a dual vector

$$\widehat{x} = x + \epsilon x_0 \in \widehat{\mathbb{R}}^n,$$

the Euclidean norm of $\widehat{x}$ is defined as [28]

$$\|\widehat{x}\| = \begin{cases} \|x\| + 2\epsilon \frac{x^{\mathrm{T}} x_0}{\|x\|}, & \text{if } x \neq 0, \\ \|x_0\| \epsilon, & \text{if } x = 0. \end{cases}$$

For a hyper-dual number $\widetilde{a}$, $\|\widetilde{a}\|^2$ is also a hyper-dual number. We may study least-squares properties of HDMPGI by the total order. However, $\|\widetilde{a}\|^2$ is not always nonnegative, for example,

$$\|\epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3\|^2 = (\epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3)^{\mathrm{T}} (\epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3) = 2\epsilon\epsilon^* a_1^{\mathrm{T}} a_2.$$

For this reason, we introduce the following set:

$$\widetilde{\mathbb{R}}_0^m = \{a_0 + \epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3 \mid a_0, a_1, a_2, a_3 \in \mathbb{R}^m, a_0 \neq 0 \text{ or } a_0 = 0 \text{ and } a_1^{\mathrm{T}} a_2 \geq 0\}.$$

For a hyper-dual vector

$$\widetilde{a} = a_0 + \epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3 \in \widetilde{\mathbb{R}}^m,$$
$$\|\widetilde{a}\|^2 = \widetilde{a}^{\mathrm{T}} \widetilde{a} = \|a_0\|^2 + 2\epsilon a_0^{\mathrm{T}} a_1 + 2\epsilon^* a_0^{\mathrm{T}} a_2 + 2\epsilon\epsilon^* (a_0^{\mathrm{T}} a_3 + a_1^{\mathrm{T}} a_2).$$

Hence, if $\widetilde{a} \in \widetilde{\mathbb{R}}_0^m$, then $\|\widetilde{a}\|^2 \geq 0$.

For $\widetilde{a} \in \widetilde{\mathbb{R}}_0^m$, we define the Euclidean norm of $\widetilde{a}$ as follows:

$$\|\widetilde{a}\| = \begin{cases} \|a_0\| + \epsilon \frac{a_0^{\mathrm{T}} a_1}{\|a_0\|} + \epsilon^* \frac{a_0^{\mathrm{T}} a_2}{\|a_0\|} + \epsilon\epsilon^* \left(\frac{a_0^{\mathrm{T}} a_3 + a_1^{\mathrm{T}} a_2}{\|a_0\|} - \frac{a_0^{\mathrm{T}} a_1 a_0^{\mathrm{T}} a_2}{\|a_0\|^3} \right), & \text{if } a_0 \neq 0, \\ \epsilon \sqrt{a_1^{\mathrm{T}} a_2} + \epsilon^* \sqrt{a_1^{\mathrm{T}} a_2} + \epsilon\epsilon^* \|a_3\|, & \text{if } a_0 = 0, a_1 \neq 0, a_2 \neq 0 \text{ and } a_1^{\mathrm{T}} a_2 \geq 0, \\ \epsilon \|a_1\| + 2\epsilon\epsilon^* \frac{a_1^{\mathrm{T}} a_3}{\|a_1\|}, & \text{if } a_0 = a_2 = 0, a_1 \neq 0, \\ \epsilon^* \|a_2\| + 2\epsilon\epsilon^* \frac{a_2^{\mathrm{T}} a_3}{\|a_2\|}, & \text{if } a_0 = a_1 = 0, a_2 \neq 0, \\ \epsilon\epsilon^* \|a_3\|, & \text{if } a_0 = a_1 = a_2 = 0, \\ 0, & \text{if } a_0 = a_1 = a_2 = a_3 = 0. \end{cases} \tag{3.1}$$

Upon expansion into its primal and hyper-dual parts, the system of linear hyper-dual equations

$$\widetilde{A}\widetilde{x} = \widetilde{b}$$

reveals four systems of real linear equations,

$$\begin{cases} A_0 x_0 = b_0, \\ A_0 x_1 = b_1 - A_1 x_0, \\ A_0 x_2 = b_2 - A_2 x_0, \\ A_0 x_3 = b_3 - A_3 x_0 - A_2 x_1 - A_1 x_2. \end{cases} \tag{3.2}$$

We will consider the least-squares solutions of the system of linear hyper-dual equations

$$\widetilde{A}\widetilde{x} = \widetilde{b}$$

under some constraints. We suppose that the real linear equation

$$A_0 x_0 = b_0$$

in (3.2) is inconsistent, and thus

$$\widetilde{A}\widetilde{x} = \widetilde{b}$$

is also inconsistent. Remark that the symbol $\widetilde{A}^{(1,3)}$ is the set of hyper-dual matrices $\widetilde{X}$ that satisfies the two equations

$$\widetilde{A}\widetilde{X}\widetilde{A} = \widetilde{A}$$

and

$$(\widetilde{A}\widetilde{X})^{\mathrm{T}} = \widetilde{A}\widetilde{X}$$

in (1.4), which is important for studying least-squares solutions of systems of linear hyper-dual equations.

Theorem 3.1. *Let $\widetilde{A} \in \widetilde{\mathbb{R}}^{m\times n}$ be such that $\widetilde{A}^{\dagger}$ exists, $\widetilde{b} \in \widetilde{\mathbb{R}}^{m}$, and*

$$(\widetilde{A}\widetilde{A}^{(1,3)} - I_m)\widetilde{b} \in \widetilde{\mathbb{R}}_0^{m}.$$

Denote

$$\widetilde{x}_0 = \widetilde{A}^{(1,3)}\widetilde{b} - (I_n - \widetilde{A}^{(1,3)}\widetilde{A})\widetilde{w} \in \widetilde{\mathbb{R}}^{n},$$

where $\widetilde{w} \in \widetilde{\mathbb{R}}^{n}$ is an arbitrary hyper-dual vector. Then,

$$\|\widetilde{A}\widetilde{x}_0 - \widetilde{b}\| \leq \|\widetilde{A}\widetilde{x} - \widetilde{b}\|$$

for any hyper-dual vector $\widetilde{x}$ that satisfies

$$\widetilde{A}(\widetilde{x} - \widetilde{A}^{(1,3)}\widetilde{b}) \in \widetilde{\mathbb{R}}_0^{m}.$$

Proof. Adding and subtracting $\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b}$, we get

$$\widetilde{e} = \widetilde{A}\widetilde{x} - \widetilde{b} = \widetilde{A}(\widetilde{x} - \widetilde{A}^{(1,3)}\widetilde{b}) + (\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b} - \widetilde{b}) \triangleq \widetilde{u} + \widetilde{v}. \tag{3.3}$$

Since

$$\widetilde{v}^{\mathrm{T}}\widetilde{u} = \widetilde{b}^{\mathrm{T}}(\widetilde{A}\widetilde{A}^{(1,3)} - I_m)\widetilde{A}(\widetilde{x} - \widetilde{A}^{(1,3)}\widetilde{b}) = 0$$

in (3.3), then $\widetilde{u}^{\mathrm{T}}\widetilde{v}$ is also zero and

$$\|\widetilde{e}\|^2 = \|\widetilde{u} + \widetilde{v}\|^2 = (\widetilde{u} + \widetilde{v})^{\mathrm{T}}(\widetilde{u} + \widetilde{v}) = \|\widetilde{u}\|^2 + \|\widetilde{v}\|^2 + 2\widetilde{u}^{\mathrm{T}}\widetilde{v} = \|\widetilde{u}\|^2 + \|\widetilde{v}\|^2. \tag{3.4}$$

Let

$$\widetilde{u} = u_0 + \epsilon u_1 + \epsilon^* u_2 + \epsilon\epsilon^* u_3.$$

Then,

$$\|\widetilde{u}\|^2 = \widetilde{u}^{\mathrm{T}}\widetilde{u} = \|u_0\|^2 + 2\epsilon u_0^{\mathrm{T}} u_1 + 2\epsilon^* u_0^{\mathrm{T}} u_2 + 2\epsilon\epsilon^* (u_0^{\mathrm{T}} u_3 + u_1^{\mathrm{T}} u_2). \tag{3.5}$$

If $\widetilde{u} \in \widetilde{\mathbb{R}}_0^m$, then it can be observed from (3.5) that

$$\|\widetilde{u}\|^2 \geq 0,$$

and thus

$$\|\widetilde{e}\|^2 \geq \|\widetilde{v}\|^2$$

by (3.4), and equality holds if and only if

$$\|\widetilde{u}\|^2 = 0.$$

Let

$$\widetilde{e} = e_0 + \epsilon e_1 + \epsilon^* e_2 + \epsilon\epsilon^* e_3, \quad \widetilde{v} = v_0 + \epsilon v_1 + \epsilon^* v_2 + \epsilon\epsilon^* v_3.$$

Then,

$$\|\widetilde{e}\|^2 = \|e_0\|^2 + 2\epsilon e_0^{\mathrm{T}} e_1 + 2\epsilon^* e_0^{\mathrm{T}} e_2 + 2\epsilon\epsilon^* (e_0^{\mathrm{T}} e_3 + e_1^{\mathrm{T}} e_2), \tag{3.6}$$

$$\|\widetilde{v}\|^2 = \|v_0\|^2 + 2\epsilon v_0^{\mathrm{T}} v_1 + 2\epsilon^* v_0^{\mathrm{T}} v_2 + 2\epsilon\epsilon^* (v_0^{\mathrm{T}} v_3 + v_1^{\mathrm{T}} v_2). \tag{3.7}$$

Since the system of real linear equations

$$A_0 x_0 = b_0$$

is inconsistent, then $e_0 \neq 0$, and thus

$$\|\widetilde{e}\|^2 > 0.$$

In this case, it follows from (3.1) that

$$\|\widetilde{e}\| = \|e_0\| + \epsilon \frac{e_0^{\mathrm{T}} e_1}{\|e_0\|} + \epsilon^* \frac{e_0^{\mathrm{T}} e_2}{\|e_0\|} + \epsilon\epsilon^* \left(\frac{e_0^{\mathrm{T}} e_3 + e_1^{\mathrm{T}} e_2}{\|e_0\|} - \frac{e_0^{\mathrm{T}} e_1 e_0^{\mathrm{T}} e_2}{\|e_0\|^3}\right). \tag{3.8}$$

By the assumption, $\widetilde{v} \in \widetilde{\mathbb{R}}_0^m$, and then

$$\|\widetilde{v}\|^2 \geq 0.$$

We consider the following two cases:

Case 1. $\|\widetilde{v}\|^2 > 0$. In this case, either $v_0 \neq 0$ or $v_0 = 0$ and $v_1^{\mathrm{T}} v_2 > 0$. If $v_0 = 0$ and $v_1^{\mathrm{T}} v_2 > 0$, then by (3.1),

$$\|\widetilde{v}\| = \epsilon \sqrt{v_1^{\mathrm{T}} v_2} + \epsilon^* \sqrt{v_1^{\mathrm{T}} v_2} + \epsilon\epsilon^* \|v_3\|.$$

Hence, by (3.8), $\|\widetilde{e}\| > \|\widetilde{v}\|$.

If $v_0 \neq 0$, then

$$\|\widetilde{v}\| = \|v_0\| + \epsilon \frac{v_0^{\mathrm{T}} v_1}{\|v_0\|} + \epsilon^* \frac{v_0^{\mathrm{T}} v_2}{\|v_0\|} + \epsilon\epsilon^* \left(\frac{v_0^{\mathrm{T}} v_3 + v_1^{\mathrm{T}} v_2}{\|v_0\|} - \frac{v_0^{\mathrm{T}} v_1 v_0^{\mathrm{T}} v_2}{\|v_0\|^3}\right). \tag{3.9}$$

Subcase 1. $\|\widetilde{e}\|^2 > \|\widetilde{v}\|^2$.

In this case, by (3.6) and (3.7),

$$\|e_0\| > \|v_0\| \quad \text{or} \quad \|e_0\| = \|v_0\|,$$

$$e_0^{\mathrm{T}} e_1 > v_0^{\mathrm{T}} v_1 \quad \text{or} \quad \|e_0\| = \|v_0\|,$$

$$e_0^{\mathrm{T}} e_1 = v_0^{\mathrm{T}} v_1, e_0^{\mathrm{T}} e_2 > v_0^{\mathrm{T}} v_2 \quad \text{or} \quad \|e_0\| = \|v_0\|,$$

$$e_0^{\mathrm{T}} e_1 = v_0^{\mathrm{T}} v_1, \quad e_0^{\mathrm{T}} e_2 = v_0^{\mathrm{T}} v_2, \quad e_0^{\mathrm{T}} e_3 + e_1^{\mathrm{T}} e_2 > v_0^{\mathrm{T}} v_3 + v_1^{\mathrm{T}} v_2.$$

Then, it can be observed from (3.8) and (3.9) that $\|\widetilde{e}\| > \|\widetilde{v}\|$.

Subcase 2. $\|\widetilde{e}\|^2 = \|\widetilde{v}\|^2$.

In this case,

$$\|e_0\| = \|v_0\|, \quad e_0^{\mathrm{T}} e_1 = v_0^{\mathrm{T}} v_1, \quad e_0^{\mathrm{T}} e_2 = v_0^{\mathrm{T}} v_2$$

and

$$e_0^{\mathrm{T}} e_3 + e_1^{\mathrm{T}} e_2 = v_0^{\mathrm{T}} v_3 + v_1^{\mathrm{T}} v_2.$$

Hence, it can be easily seen from (3.8) and (3.9) that $\|\widetilde{e}\| = \|\widetilde{v}\|$.

Case 2. $\|\widetilde{e}\|^2 > \|\widetilde{v}\|^2 = 0$.

By the assumption, $\widetilde{v} \in \widetilde{\mathbb{R}}_0^m$. If $\|\widetilde{v}\|^2 = 0$, then by (3.7), $v_0 = 0$ and

$$v_1^{\mathrm{T}} v_2 = 0.$$

We need only to consider the following five subcases:

(i) $v_0 = 0$, $v_1 \neq 0$, $v_2 \neq 0$, $v_1^{\mathrm{T}} v_2 = 0$. In this subcase, by (3.1),

$$\|\widetilde{v}\| = \epsilon\epsilon^* \|v_3\|.$$

(ii) $v_0 = v_1 = 0$, $v_2 \neq 0$. In this subcase, by (3.1),

$$\|\widetilde{v}\| = \epsilon^* \|v_2\| + 2\epsilon\epsilon^* \frac{v_2^{\mathrm{T}} v_3}{\|v_2\|}.$$

(iii) $v_0 = v_2 = 0$, $v_1 \neq 0$. In this subcase, by (3.1),

$$\|\widetilde{v}\| = \epsilon \|v_1\| + 2\epsilon\epsilon^* \frac{v_1^{\mathrm{T}} v_3}{\|v_1\|}.$$

(iv) $v_0 = v_1 = v_2 = 0$. In this subcase, by (3.1),

$$\|\widetilde{v}\| = \epsilon\epsilon^* \|v_3\|.$$

(v) $v_0 = v_1 = v_2 = v_3 = 0$. In this subcase, by (3.1), $\|\widetilde{v}\| = 0$.

For all these five subcases, by the total order defined above,

$$\|\widetilde{e}\| > \|\widetilde{v}\|.$$

Therefore, if

$$\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b} - \widetilde{b} \in \widetilde{\mathbb{R}}_0^m,$$

then

$$\|\widetilde{A}\widetilde{x}_0-\widetilde{b}\|=\|\widetilde{A}[\widetilde{A}^{(1,3)}\widetilde{b}-(I_n-\widetilde{A}^{(1,3)}\widetilde{A})\widetilde{w}]-\widetilde{b}\|=\|\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b}-\widetilde{b}\|\leq\|\widetilde{A}\widetilde{x}-\widetilde{b}\|$$

for any $\widetilde{x}$ that satisfies

$$\widetilde{A}(\widetilde{x}-\widetilde{A}^{(1,3)}\widetilde{b})\in\widetilde{\mathbb{R}}_0^m.$$

This completes the proof. □

Theorem 3.1 gives an analogous result to those of the least-squares problem of linear real equations and linear dual equations. It should be noted that the condition

$$\|\widetilde{u}\|^2\geq 0$$

is necessary for studying least-squares problem of linear hyper-dual equations, and this is the reason why we introduce the vector set $\widetilde{\mathbb{R}}_0^m$ and the total order over $\widetilde{R}$.

Example 3.1. Consider the inconsistent hyper-dual equation

$$\widetilde{A}\widetilde{x}\approx\widetilde{b},$$

where $\widetilde{A}$ is the hyper-dual matrix in Example 2.1, and

$$\widetilde{b}=\begin{bmatrix}2.8\\7.3\end{bmatrix}+\epsilon\begin{bmatrix}1.6\\5.3\end{bmatrix}+\epsilon^*\begin{bmatrix}21.6\\18.5\end{bmatrix}+\epsilon\epsilon^*\begin{bmatrix}31.2\\35.2\end{bmatrix}.$$

Then, a direct calculation shows that

$$(\widetilde{A}\widetilde{A}^{(1,3)}-I_n)\widetilde{b}=(\widetilde{A}\widetilde{A}^{\dagger}-I_n)\widetilde{b}=\begin{bmatrix}0\\-7.3\end{bmatrix}+\epsilon\begin{bmatrix}7.3\\-2.5\end{bmatrix}+\epsilon^*\begin{bmatrix}14.6\\-12.9\end{bmatrix}+\epsilon\epsilon^*\begin{bmatrix}10.6\\16\end{bmatrix}\in\widetilde{\mathbb{R}}_0^2.$$

Let

$$\widetilde{x}_1=\begin{bmatrix}1.6\\4.3\end{bmatrix}+\epsilon\begin{bmatrix}16.3\\2.8\end{bmatrix}+\epsilon^*\begin{bmatrix}8.3\\7.6\end{bmatrix}+\epsilon\epsilon^*\begin{bmatrix}6.2\\22.6\end{bmatrix}.$$

Then,

$$\begin{aligned}\widetilde{A}(\widetilde{x}_1-\widetilde{A}^{(1,3)}\widetilde{b})&=\widetilde{A}\widetilde{x}_1-\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b}=\widetilde{A}\widetilde{x}_1-\widetilde{A}\widetilde{A}^{\dagger}\widetilde{b}\\&=\begin{bmatrix}3.1\\0\end{bmatrix}+\epsilon\begin{bmatrix}20.4\\3.1\end{bmatrix}+\epsilon^*\begin{bmatrix}-20.3\\6.2\end{bmatrix}+\epsilon\epsilon^*\begin{bmatrix}13.2\\17.4\end{bmatrix}\\&\in\widetilde{\mathbb{R}}_0^2\end{aligned}$$

and

$$\widetilde{A}\widetilde{x}_1-\widetilde{b}=\begin{bmatrix}3.1\\-7.3\end{bmatrix}+\epsilon\begin{bmatrix}27.7\\0.6\end{bmatrix}+\epsilon^*\begin{bmatrix}-5.7\\-6.7\end{bmatrix}+\epsilon\epsilon^*\begin{bmatrix}23.8\\33.4\end{bmatrix}.$$

Therefore, by (3.1),

$$\|\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b}-\widetilde{b}\|=\|\widetilde{A}\widetilde{A}^{\dagger}\widetilde{b}-\widetilde{b}\|=7.3+\epsilon 2.5+\epsilon^*12.9-\epsilon\epsilon^*1.4$$

and

$$\|\widetilde{A}\widetilde{x}_1 - \widetilde{b}\| = 7.93 + \epsilon 10.3 + \epsilon^* 4 - \epsilon\epsilon^* 47.$$

Now, by the total order,

$$\|\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b} - \widetilde{b}\| < \|\widetilde{A}\widetilde{x}_1 - \widetilde{b}\|.$$

We choose another hyper-dual vector $\widetilde{x}_2$ as follows:

$$\widetilde{x}_2 = \begin{bmatrix} 1.6 \\ 1.2 \end{bmatrix} + \epsilon \begin{bmatrix} -2.5 \\ -2.8 \end{bmatrix} + \epsilon^* \begin{bmatrix} 11.6 \\ -6.8 \end{bmatrix} + \epsilon\epsilon^* \begin{bmatrix} 24.6 \\ -32.2 \end{bmatrix}.$$

Then,

$$\begin{aligned} \widetilde{A}(\widetilde{x}_2 - \widetilde{A}^{(1,3)}\widetilde{b}) &= \widetilde{A}\widetilde{x}_2 - \widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b} \\ &= \widetilde{A}\widetilde{x}_2 - \widetilde{A}\widetilde{A}^{\dagger}\widetilde{b} \\ &= \epsilon \begin{bmatrix} -10.2 \\ 0 \end{bmatrix} + \epsilon^* \begin{bmatrix} -31.4 \\ 0 \end{bmatrix} + \epsilon\epsilon^* \begin{bmatrix} -51.8 \\ -51.8 \end{bmatrix} \\ &\in \widetilde{\mathbb{R}}_0^2 \end{aligned}$$

and

$$\widetilde{A}\widetilde{x}_2 - \widetilde{b} = \begin{bmatrix} 0 \\ -7.3 \end{bmatrix} + \epsilon \begin{bmatrix} -2.9 \\ -2.5 \end{bmatrix} + \epsilon^* \begin{bmatrix} -16.8 \\ -12.9 \end{bmatrix} + \epsilon\epsilon^* \begin{bmatrix} -41.2 \\ -35.8 \end{bmatrix}.$$

It follows from (3.1) that

$$\|\widetilde{A}\widetilde{x}_2 - \widetilde{b}\| = 7.3 + \epsilon 2.5 + \epsilon^* 12.9 + \epsilon\epsilon^* 42.5.$$

Hence,

$$\|\widetilde{A}\widetilde{A}^{(1,3)}\widetilde{b} - \widetilde{b}\| < \|\widetilde{A}\widetilde{x}_2 - \widetilde{b}\|.$$

Corollary 3.1. *Let $\widehat{A} \in \widehat{\mathbb{R}}^{m\times n}$ be such that $\widehat{A}^{\dagger}$ exists, $\widehat{b} \in \widehat{\mathbb{R}}^{m}$. Denote*

$$\widehat{x}_0 = \widehat{A}^{(1,3)}\widehat{b} - (I_n - \widehat{A}^{(1,3)}\widehat{A})\widehat{w},$$

where $\widehat{w} \in \widehat{\mathbb{R}}^{m}$ is an arbitrary dual vector. Then,

$$\|\widehat{A}\widehat{x}_0 - \widehat{b}\| \le \|\widehat{A}\widehat{x} - \widehat{b}\|$$

for all $\widehat{x} \in \widehat{\mathbb{R}}^{n}$.

For a hyper-dual number

$$\widetilde{a} = a_0 + \epsilon a_1 + \epsilon^* a_2 + \epsilon\epsilon^* a_3,$$

if $a_0 \neq 0$, then we say that $\widetilde{a}$ is *appreciable*. Appreciable hyper-dual vectors and appreciable hyper-dual matrices can be defined similarly. We now consider minimum-norm least-squares solution of

$$\widetilde{A}\widetilde{x} = \widetilde{b}$$

under some certain restrictions.

Theorem 3.2. *Let* $\widetilde{A} \in \widetilde{\mathbb{R}}^{m\times n}$ *be such that* $\widetilde{A}^{\dagger}$ *exists,* $\widetilde{b} \in \widetilde{\mathbb{R}}^{m}$*, and* $\widetilde{A}^{\dagger}\widetilde{b} \in \widetilde{\mathbb{R}}^{n}_{0}$*. If* $\widetilde{A}\widetilde{A}^{\dagger}\widetilde{b}$ *is appreciable, then*

$$\|\widetilde{A}^{\dagger}\widetilde{b}\| \leq \|\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|$$

for any hyper-dual vector $\widetilde{h}$ *that satisfies*

$$(I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h} \in \widetilde{\mathbb{R}}^{n}_{0}.$$

Proof. Since

$$[(I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}]^{\mathrm{T}}\widetilde{A}^{\dagger}\widetilde{b} = \widetilde{h}^{\mathrm{T}}(I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{A}^{\dagger}\widetilde{b} = 0,$$

then

$$\|\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|^2 = \|\widetilde{A}^{\dagger}\widetilde{b}\|^2 + \|(I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|^2. \tag{3.10}$$

If a hyper-dual vector $\widetilde{h}$ satisfies

$$(I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h} \in \widetilde{\mathbb{R}}^{n}_{0},$$

then

$$\|(I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|^2 \geq 0.$$

Hence, it can be observed from (3.10) that

$$\|\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|^2 \geq \|\widetilde{A}^{\dagger}\widetilde{b}\|^2.$$

On the other hand, let

$$\widetilde{A} = A_0 + \epsilon A_1 + \epsilon^* A_2 + \epsilon\epsilon^* A_3, \quad \widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h} = x_0 + \epsilon x_1 + \epsilon^* x_2 + \epsilon\epsilon^* x_3.$$

Then,

$$\begin{aligned} \widetilde{A}\widetilde{A}^{\dagger}\widetilde{b} &= \widetilde{A}\left[\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\right] \\ &= A_0x_0 + \epsilon(A_0x_1 + A_1x_0)\epsilon^*(A_0x_2 + A_2x_0) + \epsilon\epsilon^*(A_0x_3 + A_3x_0 + A_1x_2 + A_2x_1). \end{aligned} \tag{3.11}$$

If $\widetilde{A}\widetilde{A}^{\dagger}\widetilde{b}$ is appreciable, it follows from (3.11) that $A_0x_0 \neq 0$. Hence, $x_0 \neq 0$ and $\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}$ is appreciable. In this case,

$$\|\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|^2 > 0.$$

Moreover, $\widetilde{A}^{\dagger}\widetilde{b} \in \widetilde{\mathbb{R}}^{n}_{0}$ implies

$$\|\widetilde{A}^{\dagger}\widetilde{b}\|^2 \geq 0.$$

Therefore, by an analogous discussion as the proof of Theorem 3.1, we conclude that

$$\|\widetilde{A}^{\dagger}\widetilde{b}\| \leq \|\widetilde{A}^{\dagger}\widetilde{b} + (I_n - \widetilde{A}^{\dagger}\widetilde{A})\widetilde{h}\|.$$

This completes the proof. □

Corollary 3.2. *Let* $\widehat{A} \in \widehat{\mathbb{R}}^{m\times n}$ *be such that* $\widehat{A}^{\dagger}$ *exists,* $\widehat{b} \in \widehat{\mathbb{R}}^{m}$*. If* $\widehat{A}\widehat{A}^{\dagger}\widehat{b}$ *is appreciable, then*

$$\|\widehat{A}^{\dagger}\widehat{b}\| \leq \|\widehat{A}^{\dagger}\widehat{b} + (I_n - \widehat{A}^{\dagger}\widehat{A})\widehat{h}\|$$

for all $\widehat{h} \in \widehat{\mathbb{R}}^{n}$.

4. Moore-Penrose generalized inverses of dual matrices of order n

Dual matrices and hyper-dual matrices may be referred to as dual matrices of orders 1 and 2, respectively. Specifically, real matrices are of order 0. Then, a dual matrix in $\widehat{\mathbb{R}}^{m\times n}$ is constituted of two dual matrices of order 0, and a hyper-dual matrix in $\widetilde{\mathbb{R}}^{m\times n}$ is constituted of two dual matrices of order 1. From this perspective, we define a dual matrix of order n as follows:

$$\widehat{A}^{(n)} = \widehat{B}^{(n-1)} + \epsilon_n \widehat{C}^{(n-1)},$$

where $\widehat{B}^{(n-1)}$ and $\widehat{C}^{(n-1)}$ are two dual matrices of order $n-1$, and ϵ_n is a dual unit. Hence, a dual matrix of order n can be obtained by two dual matrices of order $n-1$. For example, a dual matrix of order 3 is of the form

$$\widehat{A}^{(3)} = \widehat{B}^{(2)} + \epsilon_3 \widehat{C}^{(2)} = A_0 + \epsilon_1 A_1 + \epsilon_2 A_2 + \epsilon_1\epsilon_2 A_3 + \epsilon_3(A_4 + \epsilon_1 A_5 + \epsilon_2 A_6 + \epsilon_1\epsilon_2 A_7).$$

In this section, we study the conditions for the existence of the Moore-Penrose generalized inverse of dual matrices of order n. Denote the set of all $m \times n$ dual matrices of order n by $\widehat{\mathbb{R}}^{m\times n}_{(n)}$.

Theorem 4.1. *Let*

$$\widehat{A}^{(n)} = \widehat{B}^{(n-1)} + \epsilon_n \widehat{C}^{(n-1)} \in \widehat{\mathbb{R}}^{m\times n}_{(n)}.$$

Then, $\widehat{A}^{(n)}$ has a Moore-Penrose generalized inverse if and only if $(\widehat{B}^{(n-1)})^\dagger$ exists and

$$\left[I_m - \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^\dagger\right]\widehat{C}^{(n-1)}\left[I_n - (\widehat{B}^{(n-1)})^\dagger \widehat{B}^{(n-1)}\right] = 0.$$

Moreover, if the Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$ exists, then

$$(\widehat{A}^{(n)})^\dagger = (\widehat{B}^{(n-1)})^\dagger + \epsilon_n \widehat{Z}^{(n-1)},$$

where

$$\begin{aligned}\widehat{Z}^{(n-1)} = &-(\widehat{B}^{(n-1)})^\dagger \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^\dagger + \left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}\right]^\dagger (\widehat{C}^{(n-1)})^{\mathrm{T}} \\ &\times\left[I_m - \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^\dagger\right] + \left[I_n - (\widehat{B}^{(n-1)})^\dagger \widehat{B}^{(n-1)}\right](\widehat{C}^{(n-1)})^{\mathrm{T}}\left[\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^\dagger.\end{aligned}$$

Proof. If $\widehat{A}^{(n)}$ has a Moore-Penrose generalized inverse, we may suppose that

$$\widehat{X}^{(n)} = \widehat{Y}^{(n-1)} + \epsilon_n \widehat{Z}^{(n-1)}$$

is a Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$. Then, $\widehat{A}^{(n)}$ and $\widehat{X}^{(n)}$ satisfy the four Penrose equations, i.e.,

$$\widehat{A}^{(n)}\widehat{X}^{(n)}\widehat{A}^{(n)} = \widehat{A}^{(n)},\ \widehat{X}^{(n)}\widehat{A}^{(n)}\widehat{X}^{(n)} = \widehat{X}^{(n)},\quad (\widehat{A}^{(n)}\widehat{X}^{(n)})^{\mathrm{T}} = \widehat{A}^{(n)}\widehat{X}^{(n)},\ (\widehat{X}^{(n)}\widehat{A}^{(n)})^{\mathrm{T}} = \widehat{X}^{(n)}\widehat{A}^{(n)}.$$

Substituting

$$\widehat{A}^{(n)} = \widehat{B}^{(n-1)} + \epsilon_n \widehat{C}^{(n-1)}$$

and

$$\widehat{X}^{(n)} = \widehat{Y}^{(n-1)} + \epsilon_n \widehat{Z}^{(n-1)}$$

into the above four equations yields

$$\widehat{B}^{(n-1)}\widehat{Y}^{(n-1)}\widehat{B}^{(n-1)} = \widehat{B}^{(n-1)}, \qquad \widehat{Y}^{(n-1)}\widehat{B}^{(n-1)}\widehat{Y}^{(n-1)} = \widehat{Y}^{(n-1)},$$
$$(\widehat{B}^{(n-1)}\widehat{Y}^{(n-1)})^{\mathrm{T}} = \widehat{B}^{(n-1)}\widehat{Y}^{(n-1)}, \quad (\widehat{Y}^{(n-1)}\widehat{B}^{(n-1)})^{\mathrm{T}} = \widehat{Y}^{(n-1)}\widehat{B}^{(n-1)}.$$

Hence, the Moore-Penrose generalized inverse of $\widehat{B}^{(n-1)}$ exists and

$$\widehat{Y}^{(n-1)} = (\widehat{B}^{(n-1)})^{\dagger}.$$

On the other hand, equating the dual parts of both sides of the equation

$$\widehat{A}^{(n)}\widehat{X}^{(n)}\widehat{A}^{(n)} = \widehat{A}^{(n)}$$

gives

$$\widehat{C}^{(n-1)} = \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)} + \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)} + \widehat{B}^{(n-1)}\widehat{Z}^{(n-1)}\widehat{B}^{(n-1)},$$

which is equivalent to

$$\widehat{B}^{(n-1)}\widehat{Z}^{(n-1)}\widehat{B}^{(n-1)} = \widehat{C}^{(n-1)} - \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)} - \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)} \triangleq \widehat{D}^{(n-1)}.$$

Then,

$$\begin{aligned}\widehat{D}^{(n-1)} &= \widehat{B}^{(n-1)}\widehat{Z}^{(n-1)}\widehat{B}^{(n-1)}\\ &= \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\widehat{Z}^{(n-1)}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\\ &= \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{D}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\\ &= -\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}.\end{aligned}$$

Now we have

$$-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)} = \widehat{C}^{(n-1)} - \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)} - \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)},$$

that is,

$$\left[I_m - \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]\widehat{C}^{(n-1)}\left[I_n - (\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\right] = 0.$$

Conversely, if $(\widehat{B}^{(n-1)})^{\dagger}$ exists and

$$\left[I_m - \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]\widehat{C}^{(n-1)}\left[I_n - (\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\right] = 0,$$

then we will show that the Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$ exists, and the matrix

$$\widehat{X}^{(n)} = (\widehat{B}^{(n-1)})^{\dagger} + \epsilon_n\widehat{Z}^{(n-1)}$$

is a Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$, where

$$\widehat{Z}^{(n-1)} = -(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger} + \left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}\right]^{\dagger}(\widehat{C}^{(n-1)})^{\mathrm{T}}$$

$$\times\left[I_m-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]+\left[I_n-(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\right](\widehat{C}^{(n-1)})^{\mathrm{T}}\left[\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}.$$

Indeed, by checking the four Penrose equations, we have

$$\begin{aligned}\widehat{A}^{(n)}\widehat{X}^{(n)}\widehat{A}^{(n)}=&(\widehat{B}^{(n-1)}+\epsilon_n\widehat{C}^{(n-1)})\left[(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\widehat{Z}^{(n-1)}\right](\widehat{B}^{(n-1)}+\epsilon_n\widehat{C}^{(n-1)})\\=&\widehat{B}^{(n-1)}+\epsilon_n[\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}+\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\\&-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}].\end{aligned}$$

Note that the condition

$$\left[I_m-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]\widehat{C}^{(n-1)}\left[I_n-(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\right]=0$$

is equivalent to

$$\widehat{C}^{(n-1)}=\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}+\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)},$$

which means that

$$\widehat{A}^{(n)}\widehat{X}^{(n)}\widehat{A}^{(n)}=\widehat{A}^{(n)}.$$

Moreover,

$$\begin{aligned}\widehat{X}^{(n)}\widehat{A}^{(n)}\widehat{X}^{(n)}=&(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\Big\{-(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\\&+(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}\right]^{\dagger}(\widehat{C}^{(n-1)})^{\mathrm{T}}\left[I_m-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]\\&+\left[I_n-(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\right](\widehat{C}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\Big\}.\end{aligned}$$

Notice that

$$\begin{aligned}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}\right]^{\dagger}&=(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}\\&=(\widehat{B}^{(n-1)})^{\dagger}\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}\\&=\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}\right]^{\dagger}\end{aligned}$$

and

$$\begin{aligned}\left[\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}&=\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\\&=\left[(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}(\widehat{B}^{(n-1)})^{\dagger}\\&=\left[\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\mathrm{T}}\right]^{\dagger}.\end{aligned}$$

Therefore,

$$\widehat{X}^{(n)}\widehat{A}^{(n)}\widehat{X}^{(n)}=(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\widehat{Z}^{(n-1)}=\widehat{X}^{(n)}.$$

Furthermore,

$$\widehat{A}^{(n)}\widehat{X}^{(n)}=\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\left[\widehat{B}^{(n-1)}\widehat{Z}^{(n-1)}+\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]$$

$$=\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\Big\{\Big[I_m-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\Big]\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}$$
$$+\Big[\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\Big]^{\mathrm{T}}\Big[I_m-\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\Big]\Big\}$$

and

$$\begin{aligned}\widehat{X}^{(n)}\widehat{A}^{(n)}=&(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}+\epsilon_n\Big[(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}+\widehat{Z}^{(n-1)}\widehat{B}^{(n-1)}\Big]\\=&(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}+\epsilon_n\Big\{(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}\Big[I_m-(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\Big]\\&+\Big[I_m-(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\Big]\Big[(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}\Big]^{\mathrm{T}}\Big\}\end{aligned}$$

are symmetric, which completes the proof. □

We remark that the necessary and sufficient condition in Theorem 4.1 is a generalization of condition (iii) in Theorem 2.1. However, so far we can not give any other necessary and sufficient conditions due to the complex structure of dual matrices of order n.

Next, we show the uniqueness of the Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$ whenever it exists.

Theorem 4.2. Let $\widehat{A}^{(n)}\in\widehat{\mathbb{R}}^{m\times n}_{(n)}$. If the Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$ exists, then it is unique.

Proof. According to the proof of Theorem 4.1, if the Moore-Penrose generalized inverse of

$$\widehat{A}^{(n)}=\widehat{B}^{(n-1)}+\epsilon_n\widehat{C}^{(n-1)}$$

exists, then the Moore-Penrose generalized inverse of $\widehat{B}^{(n-1)}$ exists, and the Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$ is of the form $(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\widehat{Z}^{(n-1)}$.

Let

$$\widehat{X}_1^{(n)}=(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\widehat{Z}_1^{(n-1)}$$

and

$$\widehat{X}_2^{(n)}=(\widehat{B}^{(n-1)})^{\dagger}+\epsilon_n\widehat{Z}_2^{(n-1)}$$

be two Moore-Penrose generalized inverses of $\widehat{A}^{(n)}$. In order to show the uniqueness of the Moore-Penrose generalized inverse of $\widehat{A}^{(n)}$, it suffices to shows that

$$\widehat{Z}_1^{(n-1)}=\widehat{Z}_2^{(n-1)}.$$

Equating the dual part of both sides of the equality

$$\widehat{A}^{(n)}\widehat{X}_1^{(n)}\widehat{A}^{(n)}=\widehat{A}^{(n)},$$

we get

$$\widehat{C}^{(n-1)}=\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}+\widehat{B}^{(n-1)}\widehat{Z}_1^{(n-1)}\widehat{B}^{(n-1)}+\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}. \tag{4.1}$$

Similarly, equating the dual part of both sides of the equality

$$\widehat{A}^{(n)}\widehat{X}_2^{(n)}\widehat{A}^{(n)}=\widehat{A}^{(n)}$$

gives

$$\widehat{C}^{(n-1)} = \widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)} + \widehat{B}^{(n-1)}\widehat{Z}_2^{(n-1)}\widehat{B}^{(n-1)} + \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}. \tag{4.2}$$

Subtracting (4.1) from (4.2) gives

$$\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\widehat{B}^{(n-1)} = 0. \tag{4.3}$$

On the other hand, equating the dual part of both sides of the equality

$$\widehat{X}_1^{(n)}\widehat{A}^{(n)}\widehat{X}_1^{(n)} = \widehat{X}_1^{(n)}$$

and the equality

$$\widehat{X}_2^{(n)}\widehat{A}^{(n)}\widehat{X}_2^{(n)} = \widehat{X}_2^{(n)}$$

respectively yields

$$\widehat{Z}_1^{(n-1)} = (\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\widehat{Z}_1^{(n-1)} + (\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger} + \widehat{Z}_1^{(n-1)}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger} \tag{4.4}$$

and

$$\widehat{Z}_2^{(n-1)} = (\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}\widehat{Z}_2^{(n-1)} + (\widehat{B}^{(n-1)})^{\dagger}\widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger} + \widehat{Z}_2^{(n-1)}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}. \tag{4.5}$$

Then, by subtracting (4.4) from (4.5), we have

$$\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)} = (\widehat{B}^{(n-1)})^{\dagger}\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)}) + (\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}. \tag{4.6}$$

Furthermore, equating the dual part of the equality

$$(\widehat{A}^{(n)}\widehat{X}_1^{(n)})^{\mathrm{T}} = \widehat{A}^{(n)}\widehat{X}_1^{(n)}$$

and the equality

$$(\widehat{A}^{(n)}\widehat{X}_2^{(n)})^{\mathrm{T}} = \widehat{A}^{(n)}\widehat{X}_2^{(n)},$$

we have

$$\left[\widehat{B}^{(n-1)}\widehat{Z}_1^{(n-1)} + \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]^{\mathrm{T}} = \widehat{B}^{(n-1)}\widehat{Z}_1^{(n-1)} + \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}$$

and

$$\left[\widehat{B}^{(n-1)}\widehat{Z}_2^{(n-1)} + \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]^{\mathrm{T}} = \widehat{B}^{(n-1)}\widehat{Z}_2^{(n-1)} + \widehat{C}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}.$$

It follows that

$$\begin{aligned}
\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)}) &= \left[\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\right]^{\mathrm{T}} = (\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})^{\mathrm{T}}(\widehat{B}^{(n-1)})^{\mathrm{T}} \\
&= (\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})^{\mathrm{T}}(\widehat{B}^{(n-1)})^{\mathrm{T}}\left[\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}\right]^{\mathrm{T}} \\
&= (\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})^{\mathrm{T}}(\widehat{B}^{(n-1)})^{\mathrm{T}}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger} \\
&= \left[\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\right]^{\mathrm{T}}\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger} \\
&= \widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\widehat{B}^{(n-1)}(\widehat{B}^{(n-1)})^{\dagger}.
\end{aligned}$$

Now, it can be seen from (4.3) that

$$\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)}) = 0.$$

We can also obtain

$$(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\widehat{B}^{(n-1)} = 0$$

in a similar way. Substituting

$$\widehat{B}^{(n-1)}(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)}) = 0$$

and

$$(\widehat{Z}_1^{(n-1)} - \widehat{Z}_2^{(n-1)})\widehat{B}^{(n-1)} = 0$$

into (4.6), we have

$$\widehat{Z}_1^{(n-1)} = \widehat{Z}_2^{(n-1)},$$

which completes the proof. □

5. Conclusions

In this paper, we studied the existence and properties of hyper-dual Moore-Penrose generalized inverse of hyper-dual matrices. We gave several sufficient and necessary conditions for the existence of the HDMPGI of a given hyper-dual matrix. A compact formula for the computation of the HDMPGI was presented whenever it exists. After introducing a total order of hyper-dual numbers and Euclidean norm of a hyper-dual vector in a special set, we studied least-squares solutions and minimum-norm least-squares solutions of systems of linear hyper-dual equations under some certain restrictions. Furthermore, we considered an extension of dual matrices and hyper-dual matrices, i.e., dual matrices of order n. We also gave a sufficient and necessary condition for the existence of the Moore-Penrose generalized inverse of such matrices. The availability of the conditions and formulas obtained in this paper allow the simultaneous solutions of overdetermined systems of linear hyper-dual equations that originate from many kinematic problems. We expect these results will be useful in the future applications. It is also worth considering constructing fast algorithms to find HDMPGI whenever it exists. For example, fast algorithms for finding generalized inverses of complex matrices can be found in [30].

Author contributions

Qi Xiao: conceptualization, methodology, writing-review and editing, software, validation; Jin Zhong: conceptualization, methodology, writing-original draft, writing-review and editing, validation. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant No. 12261043), and the Program of Qingjiang Excellent Young Talents, Jiangxi University of Science and Technology (JXUSTQJYX2017007).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. M. A. Clifford, Preliminary sketch of biquaternions, *Proc. London Math. Soc.*, **4** (1871), 381–395. https://doi.org/10.1112/plms/s1-4.1.381
2. J. Angeles, The dual generalized inverses and their applications in kinematic synthesis, In: J. Lenarcic, M. Husty, *Latest advances in robot kinematics*, Springer, 2012. https://doi.org/10.1007/978-94-007-4620-6_1
3. J. Angeles, The application of dual algebra to kinematic analysis, In: J. Angeles, E. Zakhariev, *Computational methods in mechanical systems*, Springer, 1998, 3–31. https://doi.org/10.1007/978-3-662-03729-4_1
4. Y. Gu, J. Luh, Dual-number transformation and its applications to robotics, *IEEE J. Robot. Autom.*, **3** (1987), 615–623. https://doi.org/10.1109/JRA.1987.1087138
5. Y. Jin, X. Wang, The application of the dual number methods to Scara kinematics, *International Conference on Mechanic Automation and Control Engineering*, 2010, 3871–3874. https://doi.org/10.1109/MACE.2010.5535409
6. E. Pennestrì, R. Stefanelli, Linear algebra and numerical algorithms using dual numbers, *Multibody Sys. Dyn.*, **18** (2007), 323–344. https://doi.org/10.1007/s11044-007-9088-9
7. E. Pennestrì, P. Valentini, Linear dual algebra algorithms and their application to kinematics, *Multibody Dyn.*, 2009, 207–229. https://doi.org/10.1007/978-1-4020-8829-2_11
8. H. Heiß, Homogeneous and dual matrices for treating the kinematic problem of robots, *IFAC Proc. Volumes*, **19** (1986), 51–55. https://doi.org/10.1016/S1474-6670(17)59452-5
9. E. Pennestrì, P. Valentini, D. de Falco, The Moore-Penrose dual generalized inverse matrix with application to kinematic synthesis of spatial linkages, *J. Mech. Des.*, **140** (2018), 102303. https://doi.org/10.1115/1.4040882
10. F. Udwadia, Dual generalized inverses and their use in solving systems of linear dual euqations, *Mech. Mach. Theory*, **156** (2021), 104158. https://doi.org/10.1016/j.mechmachtheory.2020.104158
11. D. de Falco, E. Pennestrì, F. Udwadia, On generalized inverses of dual matrices, *Mech. Mach. Theory*, **123** (2018), 89–106. https://doi.org/10.1016/j.mechmachtheory.2017.11.020
12. F. Udwadia, E. Pennestrì, D. de Falco, Do all dual matrices have dual Moore-Penrose generalized inverses? *Mech. Mach. Theory*, **151** (2020), 103878. https://doi.org/10.1016/j.mechmachtheory.2020.103878
13. H. Li, H, Wang, Weak dual generalized inverse of a dual matrix and its applications, *Heliyon*, **9** (2023), e16624. https://doi.org/10.1016/j.heliyon.2023.e16624
14. H. Wang, T. Jiang, Q. Ling, Y. Wei, Dual core-nilpotent decomposition and dual binary relation, *Linear Algebra Appl.*, **684** (2024), 127–157. https://doi.org/10.1016/j.laa.2023.12.014

15. H. Wang, J. Gao, The dual index and dual core generalized inverse, *Open Math.*, **21** (2023), 20220592. https://doi.org/10.1515/math-2022-0592
16. H. Wang, Characterizations and properties of the MPDGI and DMPGI, *Mech. Mach. Theory*, **158** (2021), 104212. https://doi.org/10.1016/j.mechmachtheory.2020.104212
17. J. Zhong, Y. Zhang, Dual group inverses of dual matrices and their applications in solving systems of linear dual equations, *AIMS Math.*, **7** (2022), 7606–7624. https://doi.org/10.3934/math.2022427
18. J. Zhong, Y. Zhang, Dual Drazin inverses of dual matrices and dual Drazin-inverse solutions of systems of linear dual equations, *Filomat*, **37** (2023), 3075–3089. https://doi.org/10.2298/FIL2310075Z
19. J. Fike, Numerically exact derivative calculations using hyper-dual numbers, *3rd Annural Student Joint Workshop in Simulation-Based Engineering and Design*, 2009.
20. J. Fike, J. Alonso, The development of hyper-dual numbers for exact second-derivative calculations, *49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2011.
21. J. Fike, S. Jongsma, J. Alonso, E. van der Weida, Optimization with gradient and Hessian information calculated using hyper-dual numbers, *29th AIAA Applied Aerodynamics Conference*, 2011.
22. A. Cohen, M. Shoham, Application of hyper-dual numbers to multibody kinematics, *J. Mech. Robot.*, **8** (2016), 011015. https://doi.org/10.1115/1.4030588
23. A. Cohen, M. Shoham, Application of hyper-dual numbers to rigid bodies equations of motion, *Mech. Mach. Theory*, **111** (2017), 76–84. https://doi.org/10.1016/j.mechmachtheory.2017.01.013
24. Ç. Ramis, Y. Yaylı, İ. Zengin, The application of Euler-Rodrigues formula over hyper-dual matrices, *Int. Electron. J. Geom.*, **15** (2022), 266–276. https://doi.org/10.36890/iejg.1127216
25. G. Yüca, Y. Yaylı, Hyper-dual matrices and dual transformations, *J. Geom. Phys.*, **175** (2022), 104473. https://doi.org/10.1016/j.geomphys.2022.104473
26. G. Wang, Y. Wei, S. Qiao, *Generalized inverses: theory and computations*, Springer, 2018. http://doi.org/10.1007/978-981-13-0146-9
27. G. Marsaglia, G. P. H. Styan, Equalities and inequalities for ranks of matrices, *Linear Multilinear Algebra*, **2** (1974), 269–292. https://doi.org/10.1080/03081087408817070
28. L. Qi, C. Ling, H. Yan, Dual quaternions and dual quaternion vectors, *Commun. Appl. Math. Comput.*, **4** (2022), 1494–1508. https://doi.org/10.1007/s42967-022-00189-y
29. H. Wang, C. Cui, Y. Wei, The QLY least-squares and the QLY least-squares minimal-norm of linear dual least squares problems, *Linear Multilinear Algebra*, **72** (2024), 1985–2002. https://doi.org/10.1080/03081087.2023.2223348
30. O. H. Ibarra, S. Moran, R. Hui, A generalization of the fast LUP matrix decomposition algorithm and applications, *J. Algorithms*, **3** (1982), 45–56. https://doi.org/10.1016/0196-6774(82)90007-4