A Seventh-Order Scheme for Computing the Generalized Drazin Inverse

Dilan Ahmed 1, Mudhafar Hama 2, Karwan Hama Faraj Jwamer 2, and Stanford Shateyi 3, *

1 Department of Mathematics, College of Education, University of Sulaimani, Kurdistan Region, Sulaimani 46001, Iraq
2 Department of Mathematics, College of Science, University of Sulaimani, Kurdistan Region, Sulaimani 46001, Iraq
3 Department of Mathematics, University of Venda, Private Bag X5050, Thohoyandou 0950, South Africa
* Correspondence: stanford.shateyi@univen.ac.za

Received: 26 May 2019; Accepted: 22 June 2019; Published: 12 July 2019

Abstract: One of the most important generalized inverses is the Drazin inverse, which is defined for square matrices having an index. The objective of this work is to investigate and present a computational tool in the form of an iterative method for computing this task. This scheme reaches the seventh rate of convergence as long as a suitable initial matrix is chosen and by employing only five matrix products per cycle. After some analytical discussions, several tests are provided to show the efficiency of the presented formulation.

Keywords: drazin inverse; generalized inverse; iterative methods; higher order; efficiency index

JEL Classification: 15A09; 65F30.

1. Introduction

Drazin, in the pioneering work in [1], proposed and generalized a different type of outer inverse in associative rings and semigroups that does not possess the reflexivity feature but commutes with the element. The significance of this type of inverse and its calculation was then discussed wholly in [2]. Accordingly, several authors attempted to propose procedures for the calculation of generalized inverses. See, e.g., [3–5].

It is recalled that the smallest non-negative integer k such that [6]

\[ \text{rank}(A^k) = \text{rank}(A^{k+1}), \] (1)

is named as the matrix A's index and is shown by \( \text{ind}(A) \). Furthermore, assume that A is an \( N \times N \) matrix with complex entries. The Drazin inverse of A, shown by \( A^D \), is the unique matrix X reading the following identities [7]:

1. \( A^kXA = A^k \),
2. \( XAX = X \),
3. \( AX =XA \).

Throughout the paper, with \( A^* \), \( R(A) \), \( N(A) \), and \( \text{rank}(A) \), we show the conjugate transpose, the range, the null space, and the rank of \( A \in \mathbb{C}^{N \times N} \), respectively [8]. It is remarked that if \( \text{ind}(A) = 1 \), X is named as the g-inverse or group inverse of A. In addition, if A is nonsingular, then it is easily seen that

\[ \text{ind}A = 0, \text{ and } A^D = A^{-1}. \] (2)
For the square system \( Ax = b \), the general solution can be represented using the concept of the Drazin inverse as follows [7]:

\[
x = A^D b + (I - AA^D)z,
\]

(3)

wherein \( z \in \mathcal{R}(A^{k-1}) + \mathcal{N}(A) \).

Methods in the category of iteration schemes for computing generalized inverses like the Drazin inverse are quite sensitive to the choice of the initial approximation. In fact, the convergence of the Schulz-type method can only be observed if the initial value is chosen correctly [9,10]. This selection can be done under special care on some criteria, as already discussed, and given in the literature for different types of outer inverses. For some in-depth discussions about this, one may refer to [11].

Authors in [12] showed that iterative Schulz-type schemes can be used for finding the Drazin inverse of square matrices both possessing real or complex spectra. Actually, the authors investigated the initial matrix below:

\[
X_0 = \alpha A^l, \quad l \geq \text{ind}(A) = k,
\]

(4)

wherein the parameter \( \alpha \) should be selected such that the criterion \( \| I - AX_0 \| < 1 \) is read. Employing the starting value (4) yields an iterative scheme for computing the famous Drazin inverse with second-order convergence.

It is in fact necessary to apply an appropriate initial matrix when calculating the Drazin inverse. One way is as follows [12,13]:

\[
X_0 = \frac{2}{\text{Tr}(A^{k+1})} A^k,
\]

(5)

wherein \( \text{Tr}(\cdot) \) stands for the trace of an arbitrary square matrix. Another fruitful initial matrix which could lead in converging sequence of matrices for computing the generalized Drazin inverse can be written as

\[
X_0 = \frac{1}{2\|A\|_{k+1}} A^k.
\]

(6)

The Schulz method of the quadratic convergence rate for doing this task can be defined by [14]

\[
X_{n+1} = X_n(2I - AX_n), \quad n = 0, 1, 2, \cdots ,
\]

(7)

where \( I \) is the identity matrix and requires only two matrix products to achieve this rate per cycle.

Authors in [15] re-studied Chebyshev’s method for calculating \( A^\dagger_{MN} \) using a suitable initial value as follows:

\[
X_{n+1} = X_n(3I - AX_n(3I - AX_n)), \quad n = 0, 1, 2, \cdots ,
\]

(8)

with a third-order of convergence having three matrix products per cycle. Another scheme, having a cubic rate of convergence and a greater number of matrix products, was given by the same authors as follows:

\[
X_{n+1} = X_n \left[ I + \frac{1}{2}(I - AX_n)(I + (2I - AX_n)^2) \right], \quad n = 0, 1, 2, \cdots .
\]

(9)

A general procedure for having \( p \)-order methods with a \( p \) number of matrix–matrix products was given in [16] (Chapter 5). For instance, the authors presented the following fourth-order scheme:

\[
X_{n+1} = X_n(I + B_n(I + B_n(I + B_n))), \quad n = 0, 1, 2, \cdots ,
\]

(10)

in which \( B_n = I - AX_n \).

The main goal and motivation for investigating novel or useful matrix schemes for computing the Drazin inverse is not only because of the applications of such solvers in different kinds of mathematical problems [17,18] but also to improve the computational efficiency index. In fact, the hyperpower structure as given for a sample case in (10) requires a \( p \) number of matrix–matrix products to achieve a
rate of convergence. This leads to more and more inefficient methods from this class of methods, particularly when the order increases.

As such, motivated by extending efficient methods of higher orders for calculating generalized inverses, here we focus on a seventh-order scheme and discuss how we can reach this higher rate by employing only five matrix by matrix products. This will reveal a higher efficiency index for the discussed scheme in computing the Drazin inverse. For more studies and investigations in this field and related issues of generalized matrix inverses, interested readers are guided to [19–22].

The organization of this paper is as follows. In Section 1, we quickly review the definition, literature, and the need for development of higher order schemes. Section 2 is devoted to extending and proposing an efficient iterative expression for the Drazin inverse. It is derived that the scheme requires only five matrix–matrix products to achieve this rate.

Theoretical discussions with some concrete proofs are provided in Section 3, while Section 4 is oriented on the application of this scheme for computing the Drazin inverses. Results will reveal the effectiveness of this scheme for calculating the Drazin inverse. Lastly, some concluding comments are given in Section 5.

2. Derivation of an Efficient Formulation

Here, the aim is to present a competitive formulation for a member of the hyperpower family of iterations, so as to not only gain a high rate of convergence but also improve the computational efficiency index. In fact, we must factorize the formulation so as to gain the same high convergence rate but a lower number of matrix products.

Toward this objective, let us take into account the following seventh-order method from the family of hyperpower iteration schemes [23]:

$$X_{n+1} = X_n (I + B_n + B_n^2 + B_n^3 + B_n^4 + B_n^5 + B_n^6).$$ (11)

It is necessary to emphasize that we are looking for a seventh-order scheme and not a higher-order one, since we wish to hit several targets at the same time. First, the derived scheme for the Drazin inverse must be efficient, viz., it must improve the computational efficiency index of the existing solvers, as will be shown at the end of this section. Second, very high-order schemes might occasionally become hard for coding purposes, and this limits their application, so we aim to have this order be high but not so high. Besides, higher-order schemes mean that fewer stopping criteria (the computation of matrix norms) should be calculated per cycle, which is useful in terms of the elapsed time.

Now, to improve the performance of (11), we factorize (11) in what follows:

$$X_{n+1} = X_n \left( I + (B_n + B_n^2)(I - B_n + B_n^2)(I + B_n + B_n^2) \right).$$ (12)

However, the formulation (12) needs six matrix–matrix multiplications, which is still not that useful for improving the computational index of efficiency theoretically. As such, doing more factorization would yield the following scheme:

$$X_{n+1} = X_n \left( I + (B_n + B_n^2)(I - B_n + B_n^2)(I + B_n + B_n^2) \right).$$ (13)

The scheme (13) requires five matrix products per cycle to achieve the seventh order of convergence. Noting that one reason for the need to propose and have an efficient higher scheme in the category of matrix Schulz-type methods is also in the fact that Schulz-type schemes of lower orders are quite slow at the initial stage of iterates, and this could yield a greater computational burden for finding the Drazin inverse [24]. In fact, it sometimes takes several iterates for the scheme to arrive at its convergence phase, and due to imposing some stopping termination based on matrix norms, this might add some more elapsed time for the application of lower-order schemes.
It is recalled that the definition of index of efficiency is given by [25]:

\[ EI = \rho^\frac{1}{\kappa}, \]  

(14)

wherein \( \rho \) and \( \kappa \) are the convergence rate and the total cost per cycle, respectively.

Hence, the efficiency indexes of different methods are reported by

\[ EI(7) = 2^{\frac{1}{2}} \approx 1.41421, \quad EI(8) = 3^{\frac{1}{3}} \approx 1.44225, \]  

(15)

\[ EI(9) = 4^{\frac{1}{4}} \approx 1.41421, \quad EI(12) = 7^{\frac{1}{2}} \approx 1.38309, \quad EI(13) = 7^{\frac{1}{2}} \approx 1.47577. \]  

(16)

This shows that we have achieved our motivation by improving the efficiency index for calculating the Drazin inverse via a competitive formulation.

3. Seventh Rate of Convergence

Let us now recall some of the well-known lemmas we need in the rest of this section.

**Proposition 1** ([26]). Assume that \( M \in \mathbb{C}^{n \times n} \) and \( \epsilon > 0 \) are given. There is at least one matrix norm \( \| \cdot \| \) such that

\[ \rho(M) \leq \| M \| \leq \rho(M) + \epsilon, \]  

(17)

wherein \( \rho(M) \) shows the collection of all of \( M \)'s eigenvalues (in the maximum of absolute value sense).

**Proposition 2** ([26]). If \( P_L, M \) shows the projector on a space \( L \) on space \( M \), then

(i) \( P_L, M Q = Q \) if and only if \( R(Q) \subseteq L \),

(ii) \( Q P_L, M = Q \) if and only if \( N(Q) \supseteq M \).

The proof of the main theorem concerning the convergence as well as its rate of (13) for calculating the generalized Drazin inverse is now addressed as follows.

**Theorem 1.** Consider that \( A \in \mathbb{C}^{N \times N} \) is a square singular matrix. In addition, let the initial value \( X_0 \) be selected via (4) or (5). Thence, the matrices \( \{ X_n \}_{n=0}^{\infty} \) generated via (13) satisfy the following error estimate for calculating the Drazin inverse:

\[ \| A^D - X_n \| \leq \| A^D \| \| I - AX_0 \|^{\frac{1}{7}}. \]  

(18)

In addition, the convergence order is seven.

**Proof.** To prove that the sequence is converging, we first take into consideration that

\[ R_{n+1} = I - AX_{n+1} \]

\[ = I - A(X_n \left( I + (B_n + B_n^2)(I - B_n + B_n^2)(I + B_n + B_n^2) \right)) \]

\[ = I - A(X_n \left( I + B_n(I + B_n)(I - B_n + B_n^2)(I + B_n + B_n^2) \right)) \]

\[ = I - A(X_n \left( I + B_n + B_n^2 + B_n^3 + B_n^4 + B_n^5 + B_n^6 \right)) \]

\[ = (I - AX_n)^7 \]

\[ = R_{n+1}^7, \]  

(19)

wherein \( R_n = I - AX_n \). Employing a matrix norm on (19), we obtain that

\[ \| R_{n+1} \| \leq \| R_n \|^7. \]  

(20)
Since \( X_0 \) is selected as in (4) or (5), we have
\[
\mathcal{R}(X_0) \subseteq \mathcal{R}(A^k) . \tag{21}
\]
This could now be stated as
\[
\mathcal{R}(X_n) \subseteq \mathcal{R}(X_{n-1}) . \tag{22}
\]
Thus, we can conclude that
\[
\mathcal{R}(X_n) \subseteq \mathcal{R}(A^k), \, n \geq 0 . \tag{23}
\]
In a similar way, by defining the scheme by the left-multiplying of \( X_n \), we can state that
\[
\mathcal{N}(X_n) \supseteq \mathcal{N}(A^k), \, n \geq 0 . \tag{24}
\]
Now, an application of the definition of the Drazin inverse yields
\[
AA^D = A^D A = P_{\mathcal{R}(A^k) \mathcal{N}(A^k)}. \tag{25}
\]
Proposition 2, along with (23), (24), and (25) could lead to
\[
X_n A A^D = X_n = A^DAX_n, \, n \geq 0 . \tag{26}
\]
To complete the proof, we proceed in what follows. The error matrix \( \delta_n = A^D - X_n \) satisfies
\[
\delta_n = A^D - X_n = A^D - A^DAX_n = A^D (I - AX_n) = A^D R_n . \tag{27}
\]
Using (20), we obtain the following inequality
\[
||\delta_n|| = \|A^D\|\|R_n\| \leq \|A^D\|\|R_0\|^m , \tag{28}
\]
which is an affirmation of (18). Employing (28) and Proposition 2, one gets that
\[
A \delta_{n+1} = A A^D - A X_{n+1} = A A^D - I + I - A X_{n+1} = A A^D - I + R_{n+1} . \tag{29}
\]
Note that the idempotent matrix \( AA^D \) is the projector on \( \mathcal{R}(A^k) \) along \( \mathcal{N}(A^k) \), where \( \mathcal{R}(A^k) \) denotes the range of \( A^k \), and \( \mathcal{N}(A^k) \) is the null space of \( A^k \). Considering (19) and applying several simplifications, one obtains
\[
A \delta_{n+1} = A A^D - I + R_n^\pi . \tag{30}
\]
Now, by taking into account the following feature,
\[
(I - A A^D)^t = (I - A A^D), \, t \geq 1, \tag{31}
\]
we can get that
\[
(I - A A^D) A \delta_n = (I - A A^D) A (A^D - X_n) = X_n - A A^D V_n = 0 . \tag{32}
\]
We obtain, for each $t \geq 1$, that (here we use (32) in simplifications)

\[
(R_n)^t + AA^D - I = (I - AV_n)^t + AA^D - I
\]
\[
= (I - AA^D + AA^D - AV_n)^t + AA^D - I
\]
\[
= (I - AA^D + A\delta_n)^t + AA^D - I
\]
\[
= I - AA^D + (A\delta_n)^t + AA^D - I
\]
\[
= (A\delta_n)^t.
\]

From (33) and (30), we have

\[
A\delta_{n+1} = (A\delta_n)^7.
\]

Taking matrix norms from both sides yields

\[
\|A\delta_{n+1}\| \leq \|A\delta_n\|^7.
\]

Considering (35) and the second criterion of (1), we obtain that

\[
\|\delta_{n+1}\| = \|X_{n+1} - A^D\|
\]
\[
= \|A^DAV_{n+1} - A^DAA^D\|
\]
\[
= \|A^D(AV_{n+1} - AA^D)\|
\]
\[
\leq \|A^D\|\|A\delta_{n+1}\|
\]
\[
\leq \|A^D\|\|\delta_n\|^7.
\]

The relations in (36) yield the point that $X_n \to A^D$ as $n \to +\infty$ with the seventh order of convergence. The proof is complete. \(\Box\)

4. Computational Tests

The purpose of this section is to investigate the efficiency of our competitive formulation for computing the Drazin inverse, both theoretically and numerically. For such a task, we employ the challenging schemes (7), (8), and (13).

Here, we have simulated the tests in Mathematica 11.0, \cite{27,28} and the time shown is in seconds. Noting that the compared methods are programmed in the same environment using the hardware CPU Intel Core i5 2430–M, 16 GB RAM in Windows 7 Ultimate with an SSD hard disk.

**Test Problem 1.** The aim of this test is to testify the computation of the Drazin inverse for the following input \cite{12}:

\[
A = \begin{bmatrix}
2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2 & 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\
-1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 \\
0 & 0 & 0 & -1 & -2 & 0.4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0.4 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4 & -2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4 \\
\end{bmatrix},
\]
at which \( k = \text{ind}(A) = 3 \). Here, the Drazin inverse is expressed by

\[
A^D = \begin{bmatrix}
0.25 & -0.25 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
1.25 & 1.25 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
-1.66406 & -0.992187 & 0.25 & -0.25 & 0. & 0. & 0. & 0. & 0. \\
-1.19531 & -0.679687 & -0.25 & 0.25 & 0. & 0. & 0. & 0. & 0. \\
-2.76367 & -1.04492 & -1.875 & -1.25 & -1.25 & -1.25 & 1.25 & 1.25 & 1.25 \\
-2.76367 & -1.04492 & -1.875 & -1.25 & -1.25 & -1.25 & 1.25 & 1.25 & 1.25 \\
-0.625 & -0.3125 & 0. & 0. & 0. & 0. & 0. & 0. & 0.25 \\
-1.25 & -0.9375 & 0. & 0. & 0. & 0. & 0. & 0. & -0.25 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 1.25 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.25 \\
\end{bmatrix}
\]

The results are obtained by applying the stop termination

\[ ||X_{n+1} - X_n||_1 \leq 10^{-6}, \]  

and now by employing the definition in Section 1, we have for (13),

\[
\begin{align*}
||A^{k+1}X_{n+1} - A^k||_\infty & \simeq 3.69638 \times 10^{-12}, \\
||X_{n+1}AX_{n+1} - X_{n+1}||_\infty & \simeq 8.43992 \times 10^{-10}, \\
||AX_{n+1} - X_{n+1}A||_\infty & \simeq 3.75205 \times 10^{-10}.
\end{align*}
\]  

(38)

It is also necessary to mention that the domain of validity for the proposed formulation (13) is not only limited to the Drazin inverse, and if a suitable initial approximation is used, under some assumptions we can construct a converging sequence of matrix iterates for other types of generalized inverses.

**Test Problem 2.** In this test, we compare the results of various schemes for computing the regular inverse using the initial matrix

\[ X_0 = \frac{1}{||A||_F} A^*, \]  

the stopping condition (37), and the following complex matrices constructed in Mathematica:

\[
\text{N} = 5000; \ \text{no} = 25; \\
\text{ParallelTable[} \\
\quad \text{A[j] = SparseArray[} \\
\quad \quad \{\text{Band}\{-100, 1100\} \to \text{RandomReal}[20], \text{Band}\{1, 1\} \to 2., \}
\quad \quad \text{Band}\{1000, -50\}, \{\text{N} - 20, \text{N} - 25\} \to \{2.8, \text{RandomReal[]} + \text{I}, \}
\quad \quad \text{Band}\{600, 150\}, \{\text{N} - 100, \text{N} - 400\} \to \{-\text{RandomReal}[3], 3. + 3 \text{I}\}\}, \\
\quad \{\text{N}, \text{N}\}, 0.\}] \\
\text{];}
\]

The plot of the large sparse matrices in Test Problem 2 is plotted in Figure 1, showing the sparsity pattern of these matrices, while the pattern of sparsity for the inverse matrix is provided in Figure 2. Figure 3 shows the clear superiority of the proposed formulation in computing the inverse of large sparse matrices.

Here, we give a simple Mathematica implementation of (13) in solving Test Problem 2:
For[j = 1, j <= number, j++,
{
X = A[j]/(Norm[A[j], "Frobenius"]^2);
k = 1;
X1 = 20 X;
Time[j] = Part[
    While[k <= 75 && N[Norm[X - X1, 1]] >= 10^-6,
        X1 = SparseArray[X];
        XX = Id - A[j].X1;
        X2 = XX.XX;
        X = Chop@
        SparseArray[
            X1.(Id + (XX + X2).(Id - XX + X2).(Id + XX + X2))];
        k++]; // AbsoluteTiming,
1];
}];

To apply our scheme in modern applications of numerical linear algebra getting involved with sparse large matrices, one may use some commands such as SparseArray[] for tackling matrices in sparse forms and subsequently reduce the computational effort and time for preserving the sparsity pattern and finding an approximate inverse. Such applications may occur in various types of problems like the ones in [29,30].

Figure 1. The sparsity pattern of matrices in Test Problem 2.

Figure 2. The sparsity pattern of the inverse matrix $X = A_{25}^{-1}$ in Test Problem 2.
5. Conclusions

Following the motivation of proposing and extending efficient iteration schemes for computing generalized inverses, and particularly for Drazin inverses, in this work we have extended and discussed theoretically how we could achieve a seventh rate in a hyperpower structure for an iterative method. The scheme is a matrix product method and employs only five products to reach this rate. Clearly, the efficiency index will hit the bound $7^{1/5} \approx 1.47577$.

Several computational tests for calculating the Drazin inverses of several randomly generated matrices were provided to show the superiority and stability of the scheme in doing this task. Other computational problems of different sizes for different matrices were also done and showed similar behavior and the superiority of (13) for the Drazin inverse.

Author Contributions: All authors contributed equally in preparing and writing this work.

Funding: This manuscript receives no funding.

Acknowledgments: We are grateful to four anonymous referees for several comments which improved the readability of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References


