Filomat 26:3 (2012), 597–606
DOI 10.2298/FIL1203597Z

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

A new iterative method for computing the Drazin inverse

Jin Zhong[a], Xiaoji Liu[b,c], Guangping Zhou[b], Yaoming Yu[d]

[a]*Faculty of Science, Jiangxi University of Science and Technology, Ganzhou, 341000, P.R. China*
[b]*College of Mathematics and Computer Science, Guangxi University for Nationalities, Nanning 530006, P.R. China*
[c] *Guangxi Key Laborarory of Hybrid Computational and IC Design Analysis, Nanning 530006, P.R. China.*
[d]*School of Mathematical Sciences, Monash University, VIC 3800, Australia*

Abstract. In this paper, we construct a new iterative method for computing the Drazin inverse and deduce the necessary and sufficient condition for its convergence to A^d. Moreover, we present the error bounds of the iterative methods for approximating A^d.

1. Introduction

The Drazin inverse has been applied to various fields, for instance, finite Markov chains, singular differential and difference equations, multibody system dynamics and so on (see, [1, 7, 9, 14] and references therein). It is well known that iteration algorithms are undoubtedly adopted to solve large sparse linear systems. So the iterative methods for computing the Drazin inverse have been widely researched (see, for example, [2, 4–6, 11–13, 15, 16]).

The paper is organized as follows: In the remainder of this section, we will introduce some notions and lemmas. In Section 2, we will construct a new iterative method for computing the Drazin inverse, present two iterations, and give the necessary and sufficient conditions for their convergence to the Drazin inverse and the error bounds. In Section 3, we will compare our iteration with (2). In Section 4, we will give an example for computing the Drazin inverse by exploiting our iterative method.

Throughout this paper, the symbol $\mathbb{C}^{n\times n}$ denotes the set of all $n \times n$ complex matrices, the symbol $\mathbb{C}^n$ denotes the n-dimensinal complex vector space, and the symbol $L \subset \mathbb{C}^n$ denotes that L is a subspace of $\mathbb{C}^n$. Let $L, M \subset \mathbb{C}^n$ with $L \oplus M = \mathbb{C}^n$. Then the symbol $P_{L,M}$ stands for the projector on L along M, i.e., $P_{L,M}x = x, x \in L$ and $P_{L,M}y = 0, y \in M$.

For $A \in \mathbb{C}^{n\times n}$, the symbols $\mathcal{R}(A)$, $\mathcal{N}(A)$, rank(A), $\sigma(A)$, $\rho(A)$ and $\|A\|$ denote its range, null space, rank, spectrum, spectral radius and norm, respectively. And recall that the index of A, denoted by Ind(A), is the smallest nonnegative integer k such that rank(A^k) = rank(A^{k+1}), and that a matrix $X \in \mathbb{C}^{n\times n}$ is called the Drazin inverse of A, denoted by A^d, if the following hold:

$$A^{k+1}X = A^k,\ XAX = X,\ AX = XA,$$

2010 *Mathematics Subject Classification*. Primary 65F10; Secondary 15A09
Keywords. Drazin inverse, iterative method, error estimate, matrices, Frobenius norm
Received: 18 March 2011; Accepted: 06 October 2011
Communicated by Predrag Stanimirović
Research supported by the NSFC grant 11061005 and the Ministry of Education Science and Technology Key Project under grant 210164 and Grants(HCIC201103) of Guangxi Key Laborarory of Hybrid Computational and IC Design Analysis Open Fund
Email addresses: zhongjin1984@126.com (Jin Zhong), xiaojiliu72@yahoo.com.cn, Corresponding author: Tel. +86-0771- 3264782 (Xiaoji Liu), zhougpshuxue@126.com (Guangping Zhou), yaoming.yu@sci.monash.edu.au (Yaoming Yu)

where $k = \mathrm{Ind}(A)$ (see [1, 3, 10]).

In this section we denote any nonzero eigenvalue of A by $\lambda_i(A)$. Let $A \in \mathbb{C}^{n\times n}$, and $\mathrm{Ind}(A) = k \geq 1$. Write $T = \mathcal{R}(A^k)$, $S = \mathcal{N}(A^k)$.

Recently, in [2], in light of the Neumann-type expansions, Chen constructed the two iterations (1) and (2) for computing A^d:

$$X_k = X_{k-1}(I - \alpha AY) + X_0, \quad X_0 = \alpha Y, \tag{1}$$

$$\widetilde{X}_k = \widetilde{X}_{k-1}\sum_{i=0}^{p-1}(I - A\widetilde{X}_{k-1})^i, \quad X_0 = \alpha Y, \tag{2}$$

where Y satisfies $\mathcal{R}(Y) = T$ and $\mathcal{N}(Y) = S$ and $\lambda_i(AY) > 0$. Take $e > 0$ such that $q = \rho(P_T - \alpha YA) + e < 1$. It is known that there is a Q–norm $\|\ \|_Q$ satisfying $\|P_T - \alpha YA\|_Q \leq q < 1$. And the error bounds for X_k and $\widetilde{X}_k$, respectively, are

$$\|A^d - X_k\|_Q \leq \alpha q^{k+1}(1-q)^{-1}\|Y\|_Q, \tag{3}$$

$$\|A^d - \widetilde{X}_k\|_Q \leq \alpha q^{p^k}(1-q)^{-1}\|Y\|_Q. \tag{4}$$

Obviously, the error bound of the iteration (2) is very small, but it requires to compute a large amount of matrix multiplications at each step. In practice, the iterative process (2) is very expensive. It motivates us to construct a new iteration whose quantity of matrix multiplications at each step is less than that of (2) and whose error bound is between (3) and (4). In addition, the restriction of the initial value is relaxed.

Lemma 1.1. ([10]) *Let $A \in \mathbb{C}^{n\times n}$. Then $k = \mathrm{Ind}(A)$ if and only if $\mathcal{R}(A^k)$ and $\mathcal{N}(A^k)$ are complementary subspaces, i.e. $\mathcal{R}(A^k) \oplus \mathcal{N}(A^k) = \mathbb{C}^n$.*

Lemma 1.2. ([10]) *Let $A \in \mathbb{C}^{n\times n}$ with $k = \mathrm{Ind}(A)$. Then, for any nonnegative integer $l \geq k$,*

$$\mathcal{R}(A^l) = \mathcal{R}(A^k) = \mathcal{R}(A^d), \quad \mathcal{N}(A^l) = \mathcal{N}(A^k) = \mathcal{N}(A^d),$$
$$AA^d = A^dA = P_{\mathcal{R}(A^l),\mathcal{N}(A^l)}, \quad I - AA^d = I - A^dA = P_{\mathcal{N}(A^l),\mathcal{R}(A^l)}.$$

By the lemmas above, we can easily obtain the following results.

Lemma 1.3. *Let $A \in \mathbb{C}^{n\times n}$ with $k = \mathrm{Ind}(A)$, and let nonnegative integer $l \geq k$. Then*
(i) *$P_{\mathcal{R}(A^l),\mathcal{N}(A^l)}X = X$ if and only if $\mathcal{R}(X) \subset \mathcal{R}(A^l)$.*
(ii) *$XP_{\mathcal{R}(A^l),\mathcal{N}(A^l)} = X$ if and only if $\mathcal{N}(X) \supset \mathcal{N}(A^l)$.*

2. Computational methods

In this section, we discuss the iterative methods for computing the Drazin inverse A^d and deduce the necessary and sufficient conditions for their convergence to A^d.

Theorem 2.1. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$, and let p, l be positive integers with $l \geq k$. Suppose $Y \in \mathbb{C}^{n\times n}$ with $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$ and $\mathcal{N}(Y) \supset \mathcal{N}(A^l)$. For any initial approximation $X_0 \in \mathbb{C}^{n\times n}$ satisfying $\mathcal{N}(X_0) \supset \mathcal{N}(A^l)$, define the sequence $\{X_k\}$ in the following way:*

$$X_k = \alpha Y\sum_{i=0}^{p-1}(I - \alpha AY)^i + X_{k-1}(I - \alpha AY)^p, k = 1, 2, \ldots, \tag{5}$$

where α is a nonzero real parameter. Then the iteration (5) converges to A^d if and only if $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY) < 1$.

In this case, when $q = \|P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY\| < 1$, we have

$$\|X_k - X_{k-1}\| \leq \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}(1+q^p)q^{(k-1)p}, \tag{6}$$

$$\|A^d - X_k\| \leq \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}q^{kp}. \tag{7}$$

Proof. Write $P = P_{\mathcal{R}(A^l),\mathcal{N}(A^l)}$. So, by Lemma 1.2, $P = AA^d$. First, we will prove the following equation by induction on p:

$$A^d(I-\alpha AY)^p = A^d - \alpha Y\sum_{i=0}^{p-1}(I-\alpha AY)^i, p \geq 1. \tag{8}$$

When $p = 1$, the equation obviously holds. Suppose that (8) is true for $p = k$, namely,

$$A^d(I-\alpha AY)^k = A^d - \alpha Y\sum_{i=0}^{k-1}(I-\alpha AY)^i. \tag{9}$$

Consider the case $p = k+1$. By (9) and $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$,

$$\begin{aligned} A^d(I-\alpha AY)^{k+1} &= [A^d - \alpha Y\sum_{i=0}^{k-1}(I-\alpha AY)^i](I-\alpha AY) \\ &= A^d(I-\alpha AY) - \alpha Y\sum_{i=1}^{k}(I-\alpha AY)^i \\ &= A^d - \alpha Y\sum_{i=0}^{k}(I-\alpha AY)^i. \end{aligned}$$

Thus (8) holds.

Now we will investigate $A^d - X_k$. By (5) and (8),

$$A^d - X_k = (A^d - X_{k-1})(I-\alpha AY)^p = (A^d - X_{k-2})(I-\alpha AY)^{2p} = \cdots = (A^d - X_0)(I-\alpha AY)^{kp}. \tag{10}$$

From $\mathcal{N}(Y) \supset \mathcal{N}(A^l)$ and $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$, it follows that $AA^dY = Y = YAA^d$. Thus

$$P(P-\alpha AY) = P(I-\alpha AY) = P - \alpha AY = (I-\alpha AY)P = (P-\alpha AY)P. \tag{11}$$

Since $\mathcal{N}(A^l) \subset \mathcal{N}(X_0)$, $X_0 = PX_0$ and then, by Lemma 1.2, (10) and (11), we have

$$A^d - X_k = (A^d - X_0)P(I-\alpha AY)^{kp} = (A^d - X_0)P^{kp}(I-\alpha AY)^{kp} = (A^d - X_0)(P-\alpha AY)^{kp}. \tag{12}$$

If $\rho(P-\alpha AY) < 1$, then from (12) we have obviously $X_k \to A^d$ as $k \to \infty$ for any X_0 satisfying $\mathcal{N}(X_0) \supset \mathcal{N}(A^l)$.

Conversely, suppose that $X_k \to A^d$ for any X_0 satisfying $\mathcal{N}(X_0) \supset \mathcal{N}(A^l)$. Take $X_0 = (A-I)A^d$. Obviously, $\mathcal{N}(X_0) \supset \mathcal{N}(A^d) = \mathcal{N}(A^l)$. From (11) and (12),

$$(P-\alpha AY)^{kp} = AA^d(P-\alpha AY)^{kp} = A^d - X_k \to 0,$$

and then $\rho(P-\alpha AY) < 1$.

Finally, we will show (6) and (7). Since $X_0P = X_0$, by (5) and (11),

$$X_kP = X_k,$$

by induction on k.

Since $\rho(P-\alpha AY) \leq q < 1$, $I-(P-\alpha AY)$ and $I-(P-\alpha AY)^p$ are invertible. So

$$\begin{aligned} X_1 - X_0 &= \alpha Y\sum_{i=0}^{p-1}(I-\alpha AY)^i + X_0(I-\alpha AY)^p - X_0 \\ &= \alpha Y\sum_{i=0}^{p-1}(P-\alpha AY)^i + X_0(P-\alpha AY)^p - X_0 \\ &= \Big[\alpha Y[I-(P-\alpha AY)]^{-1} - X_0\Big][I-(P-\alpha AY)^p]. \end{aligned}$$

Then, by (5),

$$\begin{aligned} X_k - X_{k-1} &= (X_{k-1} - X_{k-2})(P - \alpha AY)^p \\ &= (X_1 - X_0)(P - \alpha AY)^{(k-1)p} \\ &= \Big[\alpha Y[I - (P - \alpha AY)]^{-1} - X_0\Big][I - (P - \alpha AY)^p](P - \alpha AY)^{(k-1)p}. \end{aligned}$$

Hence,

$$\begin{aligned} \|X_k - X_{k-1}\| &\le \|\alpha Y(I - (P - \alpha AY))^{-1} - X_0\| \, \|I - (P - \alpha AY)^p\| \, \|(P - \alpha AY)^{(k-1)p}\| \\ &\le \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}(1 + q^p)q^{(k-1)p}. \end{aligned}$$

Since $I - (P - \alpha AY)^p$ is invertible, by (8),

$$A^d[I - (P - \alpha AY)^p] = \alpha Y[I - (P - \alpha AY)]^{-1}[I - (P - \alpha AY)^p],$$

and then $A^d = \alpha Y[I - (P - \alpha AY)]^{-1}$.

By (12) and the equation above,

$$A^d - X_k = \Big[\alpha Y[I - (P - \alpha AY)]^{-1} - X_0\Big](P - \alpha AY)^{kp}.$$

Hence

$$\begin{aligned} \|A^d - X_k\| &\le \|\alpha Y(I - (P - \alpha AY))^{-1} - X_0\| \, \|(P - \alpha AY)^{kp}\| \\ &= \|\alpha Y(I - (P - \alpha AY))^{-1} - X_0\| q^{kp} \\ &\le \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q} q^{kp}. \end{aligned}$$

□

From the proof of Theorem 2.1, we have the following result.

Corollary 2.2. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$. Suppose that $Y \in \mathbb{C}^{n\times n}$ with $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$ and $\mathcal{N}(Y) \supset \mathcal{N}(A^l)$, where positive integer $l \ge k$, and that α is a nonzero real number. If $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY) < 1$, then*

$$A^d = \alpha Y[I - (P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY)]^{-1} = \alpha Y(P_{\mathcal{N}(A^l),\mathcal{R}(A^l)} + \alpha AY)^{-1}.$$

In Theorem 2.1, taking $p = 1$, we have the following result.

Corollary 2.3. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$, and let $Y \in \mathbb{C}^{n\times n}$ with $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$ and $\mathcal{N}(Y) \supset \mathcal{N}(A^l)$, where positive integer $l \ge k$. For any initial approximation $Z_0 \in \mathbb{C}^{n\times n}$ satisfying $\mathcal{N}(Z_0) \supset \mathcal{N}(A^l)$, define the sequence $\{Z_k\}$ in the following way:*

$$Z_k = \alpha Y + Z_{k-1}(I - \alpha AY), \quad k = 1, 2, \ldots, \tag{13}$$

where α is a nonzero real parameter. Then the iteration (13) *converges to A^d if and only if $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY) < 1$.*

In this case, when $q = \|P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY\| < 1$, we have

$$\begin{aligned} \|Z_k - Z_{k-1}\| &\le \frac{|\alpha|\|Y\| + \|Z_0\|(1-q)}{1-q}(1+q)q^{k-1}, \\ \|A^d - Z_k\| &\le \frac{|\alpha|\|Y\| + \|Z_0\|(1-q)}{1-q} q^k. \end{aligned} \tag{14}$$

Proof. (Remark 2.1.) (i) By (12), we can get $A^d - Z_k = [A^d - Z_0](I - \alpha AY)^k$. If taking $X_0 = Z_0$ in Theorem 2.1 and Corollary 2.3, we easily obtain that $X_k = Z_{kp}$. So $\{X_k\}$ in Theorem 2.1 is regarded as a subsequence of $\{Z_k\}$ in Corollary 2.3. Therefore, the iteration (5) converges faster than the iteration (13) when $p > 1$.

(ii) Obviously, the iteration (3.2.7a2) in [2] is a special case of (13). □

When $Y = A^k$ in Theorem 2.1, we have the following result.

Corollary 2.4. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$, and let p, l be positive integers with $l \geq k$. For any initial approximation $X_0 \in \mathbb{C}^{n\times n}$ satisfying $\mathcal{N}(X_0) \supset \mathcal{N}(A^l)$, define the sequence $\{X_k\}$ in the following way:*

$$X_k = \alpha A^l \sum_{i=0}^{p-1} (I - \alpha A^{l+1})^i + X_{k-1}(I - \alpha A^{l+1})^p, k = 1, 2, \ldots, \tag{15}$$

where α is a nonzero real parameter. Then the iteration (15) converges to A^d if and only if $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha A^{l+1}) < 1$.

Dually, we can give the following iterative method for computing the Drazin inverse A^d, whose proof is analogous to that of Theorem 2.1, and then omit it.

Theorem 2.5. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$, and let p, l be positive integers with $l \geq k$. Suppose $Y \in \mathbb{C}^{n\times n}$ with $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$ and $\mathcal{N}(Y) \supset \mathcal{N}(A^l)$. For any initial approximation $X_0 \in \mathbb{C}^{n\times n}$ satisfying $\mathcal{R}(X_0) \subset \mathcal{R}(A^l)$, define the sequence $\{X_k\}$ in the following way:*

$$X_k = \alpha \sum_{i=0}^{p-1} (I - \alpha YA)^i Y + (I - \alpha YA)^p X_{k-1}, \quad k = 1, 2, \ldots, \tag{16}$$

where α is a nonzero real parameter. Then the iteration (16) converges to A^d if and only if $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha YA) < 1$.
In this case, when $q = \|P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha YA\| < 1$, we have

$$\begin{aligned} \|X_k - X_{k-1}\| &\leq \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}(1+q^p)q^{(k-1)p}, \\ \|A^d - X_k\| &\leq \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}q^{kp}. \end{aligned}$$

Similarly, we can obtain the corollaries below.

Corollary 2.6. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$, and let $Y \in \mathbb{C}^{n\times n}$ with $\mathcal{R}(Y) \subset \mathcal{R}(A^l)$ and $\mathcal{N}(Y) \supset \mathcal{N}(A^l)$, where positive integer $l \geq k$. For any initial approximation $X_0 \in \mathbb{C}^{n\times n}$ satisfying $\mathcal{R}(X_0) \subset \mathcal{R}(A^l)$, define the sequence $\{X_k\}$ in the following way:*

$$X_k = \alpha Y + (I - \alpha YA)X_{k-1}, \quad k = 1, 2, \ldots, \tag{17}$$

where α is a nonzero real parameter. Then the iteration (17) converges to A^d if and only if $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha YA) < 1$.
In this case, when $q = \|P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha YA\| < 1$, we have

$$\begin{aligned} \|X_k - X_{k-1}\| &\leq \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}(1+q)q^{k-1}, \\ \|A^d - X_k\| &\leq \frac{|\alpha|\|Y\| + \|X_0\|(1-q)}{1-q}q^{k}. \end{aligned}$$

Proof. (Remark 2.2) When $Y = A^k$, from the iteration (17), we can obtain the iteration (11) in [12]. □

Corollary 2.7. *Let $A \in \mathbb{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$, and let p, l be positive integers with $l \geq k$. For any initial approximation $X_0 \in \mathbb{C}^{n\times n}$ satisfying $\mathcal{R}(X_0) \subset \mathcal{R}(A^l)$, define the sequence $\{X_k\}$ in the following way:*

$$X_k = \alpha \sum_{i=0}^{p-1} (I - \alpha A^{l+1})^i A^l + (I - \alpha A^{l+1})^p X_{k-1}, k = 1, 2, \ldots, \tag{18}$$

where α is a nonzero real parameter. Then the iteration (18) converges to A^d if and only if $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha A^{l+1}) < 1$.

In the remainder of this section, we will consider how to choose the scalar α in the iteration (5).

As we know, the Drazin inverse has a classical representation [1, Theorem 7.2.1]: If $A \in \mathbb{C}^{n\times n}$ and $\mathrm{Ind}(A) = k$, then there exists a nonsingular matrix W such that $A = W\begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}W^{-1}$, where C is an $r \times r$ nonsingular matrix and N is nilpotent of index k, and so $A^d = W\begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}W^{-1}$, $AA^d = W\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}W^{-1}$.

Now partition $W^{-1}YW$ in conformity with $W^{-1}AA^DW$ as follows $W^{-1}YW = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}$. From $AA^dY = Y = YAA^d$, it follows that $Y_{12} = 0, Y_{21} = 0$ and $Y_{22} = 0$. Then

$$P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY = W\begin{pmatrix} I_r - \alpha CY_{11} & 0 \\ 0 & 0 \end{pmatrix}W^{-1},$$

where α is a nonzero real parameter. Thus $\sigma(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY) = \sigma(I_r - \alpha CY_{11}) \cup \{0\}$. Since

$$AY = W\begin{pmatrix} CY_{11} & 0 \\ 0 & 0 \end{pmatrix}W^{-1},$$

$\rho(AY) = \rho(CY_{11})$ and then $\rho(P_{\mathcal{R}(A^l),\mathcal{N}(A^l)} - \alpha AY) = \rho(I_r - \alpha CY_{11}) = |1 - \alpha\lambda_0|$, where $\lambda_0 \in \sigma(AY)$. Write $\lambda_0 = |\lambda_0|(\cos\varphi + i\sin\varphi)$, where $\varphi = \arg(\lambda_0)$. Then $\rho(P - \alpha AY) = (|\alpha\lambda_0|^2 + 1 - 2\alpha|\lambda_0|\cos\varphi)^{1/2}$. Thus $\rho(P - \alpha AY) < 1$ if and only if $|\alpha\lambda_0|^2 < 2\alpha|\lambda_0|\cos\varphi$, namely, the sign of α is the same as that of $\mathrm{Re}(\lambda_0)$. Therefore, when

$$0 < \alpha < \frac{2\cos\varphi}{\rho(AY)} \quad \text{or} \quad 0 > \alpha > \frac{2\cos\varphi}{\rho(AY)}, \tag{19}$$

$\rho(P - \alpha AY) < 1$.

If $\sigma(AY)$ is a subset of $\mathbb{R}$, then it follows from $\rho(P - \alpha AY) < 1$ that each elements in $\sigma(AY) \setminus \{0\}$ has the same sign. Indeed, let $\lambda_{min} = \min\{\lambda : \lambda \in \sigma(AY)\}\setminus\{0\}$ and $\lambda_{max} = \max\{\lambda : \lambda \in \sigma(AY)\}\setminus\{0\}$. If $\lambda_{min} < 0$ and $\lambda_{max} > 0$, then $\max\{1 - \alpha\lambda_{min}, 1 - \alpha\lambda_{max}\} > 1$, which contradicts the condition $\rho(P - \alpha AY) < 1$.

So, by [8, Example 4.1], the best value α_{opt} for the parameter α is

$$\alpha_{opt} = \frac{2}{\lambda_{min} + \lambda_{max}}. \tag{20}$$

In the iterations (16), since $\rho(AY) = \rho(YA)$, we likewise take scalar α satisfying (19) (or (20)).

3. Example

Here is an example for computing A^d by exploiting the iteration (5), where the symbol $\|\cdot\|$ denotes the Frobenius norm.

Example 3.1. Consider the matrix

$$A = \begin{pmatrix} 2.0000 & -1.6000 & 5.6000 & -5.6000 & 0 & 5.6000 \\ 0 & 1.0000 & 6.0000 & -6.0000 & 0 & 6.0000 \\ 0 & 0 & 4.0000 & -4.0000 & 0.1000 & 3.9000 \\ 0 & 0 & 0 & 0 & 0.1000 & -0.1000 \\ 0 & 0 & 0 & 0 & 0 & 0.1000 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

with $\mathrm{Ind}(A) = 3$. Let

$$Y = \begin{pmatrix} 0.975 & 2.0000 & 2.0000 & -2.0000 & 0 & 2.0000 \\ 0 & 2.0500 & 2.0000 & -2.0000 & 0 & 2.0000 \\ 0 & 0 & 0.5000 & -0.5000 & 0 & 0.5000 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then

$$AY = \begin{pmatrix} 1.9500 & 0.7200 & 3.6000 & -3.6000 & 0 & 3.6000 \\ 0 & 2.0500 & 5.0000 & -5.0000 & 0 & 5.0000 \\ 0 & 0 & 2.0000 & -2.0000 & 0 & 2.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A^d = \begin{pmatrix} 0.5000 & 0.8000 & -1.9000 & 1.9000 & 0 & -1.9000 \\ 0 & 1.0000 & -1.5000 & 1.5000 & 0 & -1.5000 \\ 0 & 0 & 0.2500 & -0.2500 & 0 & 0.2500 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Apparently, $\mathcal{R}(Y) \subset \mathcal{R}(A^l), \mathcal{N}(Y) \supset \mathcal{N}(A^l)$.

If

$$X_0 = \begin{pmatrix} 1.0000 & 2.0000 & 2.0000 & -2.0000 & 0 & 2.0000 \\ 0 & 1.0000 & 2.0000 & -2.0000 & 0 & 2.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

then $\mathcal{N}(A^l) \subset \mathcal{N}(X_0)$ clearly hold. In order to satisfy $\rho(P - \alpha AY) < 1$, take α satisfying $0 < \alpha < 2/\rho(AY)$. Hence, those conditions in Theorem 2.1 are satisfied. Using iteration format (5) to Example 3.1, we have the following four tables.

Table 3.1: Results for Example 3.1 using the iteration (5) for $p = 1, 2$

		$p = 1$		$p = 2$	
α	Step	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$
0.4	$k = 13$	6.2504e-007	1.8749e-006	3.7521e-015	6.6921e-014
	$k = 14$	1.5329e-007	4.7175e-007	5.6814e-016	3.4613e-015
	$k = 15$	3.7017e-008	1.1627e-007	5.2296e-016	3.8459e-016
	$k = 16$	8.8276e-009	2.8190e-008	5.2296e-016	0
	$k = 27$	1.0053e-015	3.0767e-015	5.2296e-016	0
	$k = 28$	4.7103e-016	7.6919e-016	5.2296e-016	0
	$k = 29$	4.7103e-016	0	5.2296e-016	0
0.5	$k = 6$	3.3998e-007	2.9798e-006	3.8459e-016	1.3268e-013
	$k = 7$	1.6500e-009	3.4163e-007	3.8459e-016	0
	$k = 12$	3.8459e-016	7.6919e-016	3.8459e-016	0
	$k = 13$	3.8459e-016	0	3.8459e-016	0
0.6	$k = 14$	7.4623e-007	3.8571e-006	1.7631e-015	2.0000e-014
	$k = 15$	1.7788e-007	9.2411e-007	1.0660e-015	1.5384e-015
	$k = 16$	4.2178e-008	2.2005e-007	1.0660e-015	0
	$k = 28$	1.7200e-015	4.6151e-015	1.0660e-015	0
	$k = 29$	1.4729e-015	7.6919e-016	1.0660e-015	0
	$k = 30$	1.4729e-015	0	1.0660e-015	0

Table 3.2: Results for Example 3.1 using the iteration (5) for $p = 3, 4$

		$p = 3$		$p = 4$	
α	Step	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$
0.4	$k = 6$	4.8738e-010	3.6530e-008	7.0643e-014	2.6047e-011
	$k = 7$	5.9943e-012	4.8139e-010	6.0809e-016	7.0384e-014
	$k = 8$	7.0899e-014	5.9234e-012	6.0809e-016	0
	$k = 9$	1.0175e-015	6.9998e-014	6.0809e-016	0
	$k = 10$	4.9651e-016	7.6919e-016	6.0809e-016	0
	$k = 11$	4.9651e-016	0	6.0809e-016	0
0.5	$k = 4$	3.8459e-016	1.0312e-012	3.8459e-016	3.8459e-016
	$k = 5$	3.8459e-016	0	3.8459e-016	0
0.6	$k = 7$	2.9949e-011	2.3717e-009	1.8578e-015	3.7498e-013
	$k = 8$	3.7627e-013	3.0325e-011	1.2363e-015	1.1538e-015
	$k = 9$	4.9101e-015	3.8101e-013	1.2363e-015	0
	$k = 10$	1.2755e-015	4.6151e-015	1.2363e-015	0
	$k = 11$	1.2755e-015	0	1.2363e-015	0

Table 3.3: Results for Example 3.1 using the iteration (5) for $p = 5, 6$

		$p = 5$		$p = 6$	
α	Step	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$
0.4	$k = 3$	3.7017e-008	3.6804e-005	4.8738e-010	2.4995e-006
	$k = 4$	2.6117e-011	3.6991e-008	7.0901e-014	4.8731e-010
	$k = 5$	1.6157e-014	2.6101e-011	5.0877e-016	7.0768e-014
	$k = 6$	3.3307e-016	1.6153e-014	5.0877e-016	0
	$k = 7$	3.3307e-016	0	5.0877e-016	0
0.5	$k = 2$	1.3230e-013	2.6400e-006	4.0030e-016	3.3998e-007
	$k = 3$	3.8459e-016	1.3268e-013	4.0030e-016	0
	$k = 4$	3.8459e-016	0	4.0030e-016	0
0.6	$k = 4$	1.2838e-010	1.7801e-007	3.7627e-013	2.3414e-009
	$k = 5$	8.7188e-014	1.2847e-010	1.4771e-015	3.7614e-013
	$k = 6$	1.4771e-015	8.7304e-014	1.4771e-015	0
	$k = 7$	1.4771e-015	0	1.4771e-015	0

Table 3.4: Results for Example 3.1 using the iteration (5) for $p = 7, 10$

		$p = 7$		$p = 10$	
α	Step	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$	$\|A^d - X_k\|$	$\|X_k - X_{k-1}\|$
0.4	$k = 2$	1.5329e-007	0.0014	2.6118e-011	3.6841e-005
	$k = 3$	5.9940e-012	1.5328e-007	5.6610e-016	2.6117e-011
	$k = 4$	5.6610e-016	5.9938e-012	5.6610e-016	0
	$k = 5$	5.6610e-016	0	5.6610e-016	0
0.5	$k = 2$	3.8459e-016	1.6500e-009	3.8459e-016	1.3268e-013
	$k = 3$	3.8459e-016	0	3.8459e-016	0
0.6	$k = 2$	7.4623e-007	0.0129	1.2838e-010	2.1383e-004
	$k = 3$	2.9949e-011	7.4626e-007	1.3911e-015	1.2838e-010
	$k = 4$	1.8584e-015	2.9950e-011	1.3911e-015	0
	$k = 5$	1.2372e-015	1.1538e-015	1.3911e-015	0
	$k = 6$	1.2372e-015	0	1.3911e-015	0

From the above four tables, we can see that the larger is p, the better is results of iteration, since $X_k = S_{kp}$ is an increasing function of p.

The above tables illustrate that $\alpha = 0.5$ is the best value for which the iteration (5) fastest converges to A^d. The reason is that α is calculated by using (20). Thus, for an appropriate α, the iteration is better.

In practice, we generally consider the quantity $\|X_k - X_{k-1}\|$ in order to abort iteration since there exist such cases as $\alpha = 0.5$. For example, for $\|X_k - X_{k-1}\| < \mu X_k$, where μ is the machine precision, the iteration for $\alpha = 0.5$ and $p = 5$ only needs 4 steps, and the iteration for $\alpha = 0.5$ and $p = 6$ only needs 3 steps (see Table 3.3).

Comparing (7) with (14), it is not difficult to see that in order to obtain the same error, if the iteration (13) in Corollary 2.3 requires N steps, then the iteration (5) in Theorem 2.1 requires $\lceil N/p \rceil$ steps, where $p > 1$ and the symbol $\lceil \beta \rceil$ stands for a largest integer less than β+1. To this end, the iteration (5) requires to operate $\lceil N/p \rceil + p + 1$ matrix multiplications in all.

Hence, when $p = \lceil \sqrt{N} \rceil$ or $p = \lceil \sqrt{N} \rceil - 1$, the value of $\lceil N/p \rceil + p + 1$ reaches the minimum. The following tables illustrate that an appropriate p lessens the number of iterative steps and matrix multiplications in iteration processes.

Table 3.5: $\alpha = 0.4$, $\varepsilon_k \leq$ 6.09e-016

p	1	2	3	4	**5**	**6**	7	10
k	N=28	14	10	7	**6**	**5**	4	3
η	29	17	14	12	**12**	**12**	12	14

Table 3.6: $\alpha = 0.6$, $\varepsilon_k \leq$ 1.48e-015

p	1	2	3	4	**5**	**6**	7	10
k	N=29	15	10	8	**6**	**5**	5	3
η	30	18	14	13	**12**	**12**	13	14

where η is the number of required matrix multiplication and $\varepsilon_k = \|A^d - X_k\|$.

4. Comparison

In this section, first, we will compare the iteration (5) with the iteration (2).

The two iterations (5) (if $X_0 = \alpha Y$) and (2) stem from the series

$$\sum_{j=0}^{\infty} \alpha Y (I - \alpha AY)^j. \tag{21}$$

Let S_n denote the partial sum to n of (21). Then $\widetilde{X}_k = S_{p^k-1}$ in (2) and $X_k = S_{kp}$ in (5). Thus, it is obvious that there exist positive integers k_i, i=1,2, which would make the two iterations obtain the same results after k_1 and k_2 steps, respectively.

For example, in Example 3.1, let the initial approximation $X_0 = \alpha Y (\alpha = 0.4)$. Then, obviously, $\mathcal{N}(A^l) \subset \mathcal{N}(X_0)$ and $0 < \alpha < 2/\rho(AY)$ hold. Hence, those conditions in Theorem 2.1 are satisfied. Using the iteration (5) for Example 3.1, we have the following table,where the symbol $\|\cdot\|$ denotes the Frobenius norm.

Table 4.1: Results for Example 3.1 with $\alpha = 0.4$ and $X_0 = \alpha Y$, using (5)

step	$k = 3$	$k = 4$	k**=5**	$k = 6$
$\|A^d - X_k\|$	9.1800e-009	6.1105e-012	**3.7370e-015**	2.0260e-015

Taking $p = 5$ and $k_1 = 2$, we have $\|A^d - \widetilde{X}_{k_1}\|$=1.633e-014 by the iteration (2). From Table 4.1, $\|A^d - X_{k_2}\|$= 3.7370e-015 < 1.633e-014 where $k_2 = 5$.

If $k > 1$, note that the stopping criterion of iterations generally is based on the quantity of $\|X_k - X_{k-1}\|$. So, in contrast with (5), (2) often lavishes more operations from X_{k-1} to X_k when X_k is close to A^d. Indeed, computing $\widetilde{X}_k$ requires p times of the multiplications and additions of matrix. So, it would take kp times operations from X_0 to X_k in all. Whereas computing X_k requires only one matrix multiplication and two matrix additions. So it would take $k + p + 1$ times operations from X_0 to X_k in all. Next, we consider elapsed time in calculating process. In Example 3.1, if $p = 5$ and $X_0 = \alpha Y$, then the two iterations (5) and (2) elapsed

0.009231 seconds and 0.009993 seconds, respectively, where we exploited the Mathematic functions "tic" and "toc" in Matlab 7.8.0(R2009a), which return the CPU time consumed seconds, and the two tests run on a Lenovo QiTian M550E desktop with AMD Athlon(tm) 64 X2 Dual Core CPU 5000+@2.6GHz and 768 MB of memory, using Windows XP Professional(SP3).

Now, we consider the choice of p in (5). Evidently, the p's size just affects the operation quantity of computing X_1. Because matrices $\alpha Y \sum_{i=0}^{p-1}(I - \alpha AY)^i$ and $(I - \alpha AY)^p$ have been defined before computing $X_k, k > 1$. In spite of that, oversize p is expensive. If $X_k = S_{kp}$ is an ideal approximate solution, then k and p are reciprocal. In practice, k is often less than theoretical iterative number to reach the ideal approximate solution when p becomes small. For example, in Table 3.5, when $p = 5$ or $p = 6$, it would take the least work of calculations.

References

[1] S. L. Campbell, C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman (Advanced Publishing Program), Boston, MA, 1979, (reprinted by Dover, 1991).

[2] Y. Chen, Representation and approximation for the Drazin inverse A^d, Appl. Math. Comput. 119 (2001) 147-160.

[3] M.P. Drazin, Pseudoinverse in associative rings and semigroups, Amer. Math. Monthly. 65 (1958) 506-514.

[4] R.E. Hartwig, Schur's theorem and Drazin inverse, Pacific J. Math. 78 (1978) 133-238.

[5] R.E. Hartwig, A method for calculating A^d, Math. Japon. 26 (1981) 37-43.

[6] X. Li, Y. Wei, Iterative methods for the Drazin inverse of a matrix with a complex spectrum, Appl. Math. Comput. 147(3) (2004) 855-862.

[7] B. Načevska, Iterative methods for computing generalized inverses and splittings of operators, Appl. Math. Comput. 228(1) (2009) 186-188.

[8] Y. Saad, Iterative Methods for Sparse Linear Systems, (2nd ed.), Philadelphia, PA: Society for Industrial and Applied Mathematics, 2003.

[9] B. Simeon, C. Führer, P. Rentrop, The Drazin inverse in multibody system dynamics, Numer. Math. 64 (1993) 521-539.

[10] G. Wang, Y. Wei, S. Qiao, Generalized Inverses: Theory and Computations, Science Press, Beijing, 2004.

[11] Y. Wei, A characterization and representation of the Drazin inverse, SIAM J. Matrix Anal. Appl. 17 (1996) 744-747.

[12] Y. Wei, Successive matrix squaring algorithm for computing the Drazin inverse, Appl. Math. Comput. 108 (2-3) (2000) 67-75.

[13] Y. Wei, H. Wu, The representation and approximation for Drazin inverse, J. Comput. Appl. Math. 126 (1-2) (2000) 417-432.

[14] N. Castro González, On the convergence of semiiterative methods to the Drazin inverse solution of linear equations in Banach spaces, Collect. Math. 46 (3) (1995) 303-314.

[15] X. Liu, C. Hu, Y. Yu, Further results on iterative methods for computing generalized inverses, J. Comput. Appl. Math. 234 (3) (2010) 684-694.

[16] X. Sheng, G. Chen, An oblique projection iterative method to compute Drazin inverse and group inverse, Appl. Math. Comput. 211 (2) (2009) 417-421.