JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 87 (1997) 21–38

A semi-iterative method for real spectrum singular linear systems with an arbitrary index

Joan-Josep Climent[a,1], Michael Neumann[b,*], Avram Sidi[c]

[a] *Departament de Tecnologia Informàtica i Computació, Universitat d'Alacant, E-03071 Alacant, Spain*
[b] *Department of Mathematics, University of Connecticut, Storrs, CT 06269-03009, USA*
[c] *Computer Science Department, Technion-Israel Institute of Technology, Haifa 32000, Israel*

Received 14 February 1997; received in revised form 16 July 1997

Abstract

In this paper we develop a semi-iterative method for computing the Drazin-inverse solution of a singular linear system $Ax = b$, where the spectrum of A is real, but its index (i.e., the size of its largest Jordan block corresponding to the eigenvalue zero) is arbitrary. The method employs a set of polynomials that satisfy certain normalization conditions and minimize some well-defined least-squares norm. We develop an efficient recursive algorithm for implementing this method that has a fixed length independent of the index of A. Following that, we give a complete theory of convergence, in which we provide rates of convergence as well. We conclude with a numerical application to determine eigenprojections onto generalized eigenspaces. Our treatment extends the work of Hanke and Hochbruck (1993) that considers the case in which the index of A is 1.

Keywords: Singular systems; Iterative methods; Polynomial acceleration

AMS classification: 65F10; 65F20; 15A09

1. Introduction

Consider the linear system

$$Ax = b, \tag{1.1}$$

where $A \in \mathbb{C}^{n,n}$ is singular and $\operatorname{ind}(A) = a$ is arbitrary. Here $\operatorname{ind}(\cdot)$ denotes the *index* of a matrix, namely, the size of the largest Jordan block corresponding to its zero eigenvalue. The purpose of

* Corresponding author. E-mail: neumann@math.uconn.edu.

[1] The work of this author was done while he was visiting the Department of Mathematics at the University of Connecticut, supported by Spanish DGICYT grant number PR94-205.

PII S0377-0427(97)00169-6

this paper is to develop a semi-iterative method for computing the Drazin-inverse solution of (1.1), namely, the vector $A^{\mathrm{D}}b$, where A^{D} is the *Drazin inverse* of A, in an *efficient manner*. For the Drazin inverse and its properties, see, e.g., [1] or [2].

We shall assume that

$$\sigma(A) \subseteq \{0\} \cup [c-d, c+d], \quad 0<d<c, \tag{1.2}$$

where $\sigma(\cdot)$ denotes the spectrum of a matrix.

Our work here extends that of Hanke and Hockbruck [8] which treats the case of $a=1$ and utilizes the general theory of Eiermann et al. [4] of semi-iterative methods for computing the Drazin-inverse solution to singular systems.

We begin with some essential background. Let x_0 be an arbitrary initial vector and let $r_0=b-Ax_0$ be the corresponding residual vector. Then, beginning with x_0, the mth iterate x_m is given by

$$x_m = x_0 + q_{m-1}(A)r_0 = p_m(A)x_0 + q_{m-1}(A)b, \tag{1.3}$$

where $q_{m-1}(\lambda)$ is a polynomial of degree at most $m-1$ and $p_m(\lambda)$ is a polynomial of degree at most m given by

$$p_m(\lambda) = 1 - \lambda q_{m-1}(\lambda). \tag{1.4}$$

We call $p_m(\lambda)$ the mth *residual polynomial.* Note that

$$p_m(0) = 1. \tag{1.5}$$

As is shown by Eiermann et al. [4, Lemma 2], necessary and sufficient conditions for the convergence of the sequence $\{x_m\}_{m=0}^{\infty}$ are that

$$\lim_{m\to\infty} p_m^{(i)}(0) = 0, \quad i=1,\ldots,a, \tag{1.6}$$

and

$$\lim_{m\to\infty} p_m^{(i)}(\lambda_j) = 0, \quad i=0,\ldots,k_j-1, \tag{1.7}$$

where λ_j are the nonzero eigenvalues of A and $k_j=\operatorname{ind}(A-\lambda_j I)$.

The conditions in (1.6) will, of course, be satisfied if

$$p_m^{(i)}(0) = 0, \quad i=1,\ldots,a, \quad \text{for all } m=0,1,\ldots\,. \tag{1.8}$$

Polynomials $p_m(\lambda)$ satisfying (1.8) and (1.7) were considered by Hanke and Hockbruck [8] for the case $a=1$. We mention in passing that the polynomials that arise in connection with the extrapolation methods for the Drazin-inverse solution studied by Sidi [10] satisfy (1.8) and (1.7) for arbitrary a.

The plan of this paper is as follows: In Section 2, using a weight function $w(\lambda)$, we provide an integral norm $|||\cdot|||$ and a set of polynomials $\{p_m(\lambda)\}_{m=0}^{\infty}$ satisfying (1.8) and (1.7) such that the norm of $p_m(\lambda)$ is minimal over the set of all polynomials $p(\lambda)$ of degree at most m which satisfy $p(0)=1$ and $p^{(i)}(0)=0$, $i=1,\ldots,a$. We use these polynomials to construct our semi-iterative method. Our work here extends directly the developments of Hanke and Hockbruck [8] to the case $a>1$.

In Section 3 we develop a recursive algorithm for implementing the semi-iterative method defined by (1.3) and (1.4), $p_m(\lambda)$ being the minimal polynomials of Section 2. This algorithm involves only four successive iterates x_m, independently of the index of A. Here we make use of the fact that the $p_m(\lambda)$ can be expressed in terms of a set of orthogonal polynomials that satisfy the usual 3-term recurrence relation.

In Section 4 we prove the convergence of the method and provide error bounds and the corresponding rates of convergence for the case in which

$$w(\lambda)=\frac{1}{\sqrt{(\lambda-c+d)(c+d-\lambda)}}. \tag{1.9}$$

In particular, we show that if A satisfies (1.2), then

$$\|x_m - A^{\mathrm{D}}b - \tilde{x}_0\| = \mathrm{O}(m^{a+s}\kappa^m) \quad \text{as } m \to \infty,$$

where $\tilde{x}_0$ is that part of x_0 that lies in the null space of A^a, s is a nonnegative integer, and

$$\kappa=\frac{c-\sqrt{c^2-d^2}}{d}<1. \tag{1.10}$$

The asymptotic estimates that we give for the bounds on our residual polynomials in equation (4.17) of Theorem 4.5 do not reach, except in the case of the index A being equal to 1, the near optimal rate achieved by the residuals of Berstein (see [6]) which is displayed here in (4.24). But we believe that our short recurrence relation for computing the residuals makes up for this deficiency.

In Section 5 we present several numerical examples in which we compute the *projections onto the generalized eigenspaces of matrices* whose spectrum is real and satisfies the condition of (1.2). The algorithm does well when the transforming matrix of A to its Jordan canonical form has a relatively low condition number.

2. Minimal polynomials

Let $w(\lambda)$ be a nonnegative weight function over the interval $[c-d, c+d]$ and let f and g be functions defined on $[c-d, c+d]$. Define the inner product $\langle\cdot,\cdot\rangle$ on $[c-d, c+d]$ by

$$\langle f, g\rangle = \int_{c-d}^{c+d} w(\lambda) f(\lambda) g(\lambda)\,\mathrm{d}\lambda.$$

Next define the norm $|||\cdot|||$ via

$$|||f|||^2 = \left\langle f, \frac{1}{\lambda^a} f\right\rangle.$$

Let Π_m denote the set of all real polynomials of degree at most m and define

$$\Pi_m^0 = \{p \in \Pi_m\colon\ p(0)=1,\ \ p^{(i)}(0)=0,\ \ i=1,\dots,a\}.$$

Note that $p_m(\lambda)=1$ is the only member of the set Π_m^0 for $m=0,1,\dots,a$.

Theorem 2.1. *Let a be a positive integer. Then, for $m \geqslant a+1$, the minimization problem*

$$\min_{p \in \Pi_m^0} |||p||| \tag{2.1}$$

admits a unique solution $p_m(\lambda)$ which is characterized by

$$\langle p_m, \lambda^j \rangle = 0, \quad j = 1, \ldots, m-a. \tag{2.2}$$

Proof. We start by noting that $p \in \Pi_m^0$ implies that $p(\lambda) = 1 - \lambda^{a+1} u(\lambda)$ with $u \in \Pi_{m-a-1}$. Thus, we have

$$|||p|||^2 = \int_{c-d}^{c+d} w(\lambda)\lambda^{a+2} \left[\lambda^{-a-1} - u(\lambda)\right]^2 \, \mathrm{d}\lambda. \tag{2.3}$$

As $w(\lambda)\lambda^{a+2}$ is nonnegative on $[c-d, c+d]$, there exists a unique polynomial $u^*(\lambda)$ in Π_{m-a-1} that minimizes the integral on the right-hand side of (2.3); $u^*(\lambda)$ is the best approximation from Π_{m-a-1} to λ^{-a-1} in the norm induced by the inner product $\langle \cdot, \lambda^{a+2} \cdot \rangle$, see, e.g., [3]. Consequently, $p_m(\lambda) = 1 - \lambda^{a+1} u^*(\lambda)$ is the unique solution to (2.1).

Consider the polynomial $p(\lambda) = p_m(\lambda) + \alpha\lambda^{a+j}$, where $\alpha \in \mathbb{R}$ and $j = 1, \ldots, m-a$. Obviously, $p \in \Pi_m^0$. Hence,

$$|||p_m|||^2 \leqslant |||p|||^2 = |||p_m|||^2 + 2\alpha\langle p_m, \lambda^j \rangle + \alpha^2 |||\lambda^{a+j}|||^2.$$

Since α has an arbitrary sign, this inequality holds if and only if (2.2) holds.

Conversely, assume that (2.2) holds for some $p_m \in \Pi_m^0$. Let $p \in \Pi_m^0$. Then $p(\lambda) - p_m(\lambda)$ has a zero of multiplicity at least $a+1$ at $\lambda = 0$. Thus,

$$v(\lambda) = \frac{p(\lambda) - p_m(\lambda)}{\lambda^a} \in \operatorname{span}\{\lambda, \lambda^2, \ldots, \lambda^{m-a}\}.$$

Consequently,

$$|||p|||^2 = |||p_m|||^2 + 2\langle p_m, v \rangle + |||\lambda^a v|||^2 = |||p_m|||^2 + |||\lambda^a v|||^2 \geqslant |||p_m|||^2,$$

since $\langle p_m, v \rangle = 0$ by (2.2). This means that $p_m(\lambda)$ is the unique solution to (2.1). □

Now, let us define

$$u_m(\lambda) = \frac{p_m(\lambda) - p_{m+1}(\lambda)}{\lambda}, \quad m \geqslant a. \tag{2.4}$$

Clearly, $u_m(\lambda)$ is a polynomial in $\operatorname{span}\{\lambda^a, \lambda^{a+1}, \ldots, \lambda^m\}$.

Theorem 2.2. *The polynomials $\lambda^{-a} u_m(\lambda)$, $m = a, a+1, \ldots,$ are orthogonal with respect to the inner product $\langle \cdot, \lambda^{a+2} \cdot \rangle$.*

Proof. First, $\lambda^{-a} u_m(\lambda)$ is of degree precisely $m-a$. For any polynomial $p(\lambda)$ in Π_{m-a-1}, with $m \geqslant a+1$, we then have that

$$\langle \lambda^{-a} u_m, \lambda^{a+2} p \rangle = \langle u_m, \lambda^2 p \rangle = \langle p_m - p_{m+1}, \lambda p \rangle = 0$$

by (2.2). □

From Theorem 2.2 we now have that the $u_m(\lambda)$ satisfy a 3-term recursion relation of the form

$$u_m(\lambda)=(\omega_m\lambda+\mu_m)u_{m-1}(\lambda)+v_m u_{m-2}(\lambda), \quad m\geqslant a+1 \tag{2.5}$$

for some constants ω_m, μ_m, and v_m, with $v_{a+1}=0$.

Let us denote by $t_m(\lambda)$ the orthogonal polynomial of degree m with respect to the inner product $\langle\cdot,\cdot\rangle$ and normalized such that $t_m(0)=1$. As a result of this normalization, the $t_m(\lambda)$ satisfy a 3-term recursion relation of the form

$$t_{m+1}(\lambda)=-\alpha_m\lambda t_m(\lambda)+(1+\beta_m)t_m(\lambda)-\beta_m t_{m-1}(\lambda), \quad m\geqslant 0 \tag{2.6}$$

with

$$t_{-1}(\lambda)=0 \quad \text{and} \quad t_0(\lambda)=1,$$

for some constants α_m and β_m.

Theorem 2.3. *For $m\geqslant a$, the polynomials $p_m(\lambda)$ can be expressed in terms of the polynomials $t_j(\lambda)$ as*

$$\lambda p_m(\lambda)=\sum_{j=m-a}^{m+1}\pi_{m,j}t_j(\lambda) \tag{2.7}$$

for some constants $\pi_{m,j}$ which satisfy the linear system

$$\begin{aligned}
&\sum_{j=m-a}^{m+1}\pi_{m,j}=0,\\
&\sum_{j=m-a}^{m+1}\pi_{m,j}\tau_j^{(1)}=1,\\
&\sum_{j=m-a}^{m+1}\pi_{m,j}\tau_j^{(i)}=0, \quad i=2,3,\ldots,a+1,
\end{aligned} \tag{2.8}$$

where $\tau_j^{(i)}=t_j^{(i)}(0)$.

Proof. Since $\lambda p_m(\lambda)$ is in Π_{m+1}, we have that

$$\lambda p_m(\lambda)=\sum_{j=0}^{m+1}\pi_{m,j}t_j(\lambda).$$

But

$$\langle\lambda p_m,t_j\rangle=\langle p_m,\lambda t_j\rangle=0, \quad j=0,1,\ldots,m-a-1,$$

by (2.2). Therefore, $\pi_{m,j}=0$ for $j=0,1,\ldots,m-a-1$. This proves (2.7). The first of the equations in (2.8) follows by letting $\lambda=0$ in (2.7). The remaining equations in (2.8) can be obtained by taking the ith derivative of both sides of (2.7), $i = 1, 2, \ldots, a+1$, noting that

$$[\lambda p_m(\lambda)]^{(i)} = \lambda p_m^{(i)}(\lambda) + i p_m^{(i-1)}(\lambda), \quad i = 1, 2, \ldots,$$

and letting $\lambda = 0$ on both sides, and recalling that $p_m \in \Pi_m^0$. □

We now show how the coefficients ω_m, μ_m, and ν_m in (2.5) can be determined from the constants $\pi_{m,j}$ in (2.7) and the coefficients α_m and β_m in (2.6).

Theorem 2.4. *Let*

$$\gamma_m := \pi_{m,m+1}, \quad \delta_m := \pi_{m,m} \quad \text{and} \quad \varepsilon_m := \pi_{m,m-a}. \tag{2.9}$$

Then the coefficients ω_m, μ_m, and ν_m of (2.5) can be computed from

$$\begin{aligned} \omega_m &= -\frac{\gamma_{m+1}}{\gamma_m}\alpha_{m+1}, \\ \mu_m &= -\frac{1}{\gamma_m}\left[\gamma_m - \delta_{m+1} + \frac{\omega_m(\gamma_{m-1} - \delta_m)}{\alpha_m} - \gamma_{m+1}(1 + \beta_{m+1})\right], \\ \nu_m &= \frac{\omega_m \varepsilon_{m-1}\beta_{m-a-1}}{\alpha_{m-a-1}\varepsilon_{m-2}}. \end{aligned} \tag{2.10}$$

Proof. Using (2.7) in (2.4) we see that

$$u_m = \frac{1}{\lambda^2}\left[\pi_{m,m-a}t_{m-a} + \sum_{j=m-a+1}^{m+1}(\pi_{m,j} - \pi_{m+1,j})t_j - \pi_{m+1,m+2}t_{m+2}\right], \quad m \geqslant a, \tag{2.11}$$

where we have written t_j instead of $t_j(\lambda)$ for short.

Now, let $m \geqslant a+2$. Substituting (2.11) in (2.5), and multiplying throughout by λ^2, we obtain that

$$\begin{aligned} &\pi_{m,m-a}t_{m-a} + \sum_{j=m-a+1}^{m+1}(\pi_{m,j} - \pi_{m+1,j})t_j - \pi_{m+1,m+2}t_{m+2} \\ &\quad -\omega_m\left[\pi_{m-1,m-1-a}\lambda t_{m-1-a} + \sum_{j=m-a}^{m}(\pi_{m-1,j} - \pi_{m,j})\lambda t_j - \pi_{m,m+1}\lambda t_{m+1}\right] \\ &\quad -\mu_m\left[\pi_{m-1,m-1-a}t_{m-1-a} + \sum_{j=m-a}^{m}(\pi_{m-1,j} - \pi_{m,j})t_j - \pi_{m,m+1}t_{m+1}\right] \\ &\quad -\nu_m\left[\pi_{m-2,m-2-a}t_{m-2-a} + \sum_{j=m-1-a}^{m-1}(\pi_{m-2,j} - \pi_{m-1,j})t_j - \pi_{m-1,m}\lambda t_m\right] \\ &= 0. \end{aligned} \tag{2.12}$$

Next, from (2.6) we know that

$$\lambda t_j = -\frac{1}{\alpha_j} t_{j+1} + \frac{1+\beta_j}{\alpha_j} t_j - \frac{\beta_j}{\alpha_j} t_{j-1}, \quad j=0,1,\ldots. \tag{2.13}$$

Substituting (2.13) in (2.12) yields that

$$\begin{aligned}
&\pi_{m,m-a}t_{m-a} + \sum_{j=m-a+1}^{m+1} (\pi_{m,j} - \pi_{m+1,j})t_j - \pi_{m+1,m+2}t_{m+2} \\
&\quad -\omega_m \Bigg[\pi_{m-1,m-1-a} \left(-\frac{1}{\alpha_{m-1-a}} t_{m-a} + \frac{1+\beta_{m-1-a}}{\alpha_{m-1-a}} t_{m-1-a} - \frac{\beta_{m-1-a}}{\alpha_{m-1-a}} t_{m-2-a} \right) \\
&\quad + \sum_{j=m-a}^{m} (\pi_{m-1,j} - \pi_{m,j}) \left(-\frac{1}{\alpha_j} t_{j+1} + \frac{1+\beta_j}{\alpha_j} t_j - \frac{\beta_j}{\alpha_j} t_{j-1} \right) \\
&\quad -\pi_{m,m+1} \left(-\frac{1}{\alpha_{m+1}} t_{m+2} + \frac{1+\beta_{m+1}}{\alpha_{m+1}} t_{m+1} - \frac{\beta_{m+1}}{\alpha_{m+1}} t_m \right) \Bigg] \\
&\quad -\mu_m \left[\pi_{m-1,m-1-a}t_{m-1-a} + \sum_{j=m-a}^{m} (\pi_{m-1,j} - \pi_{m,j})t_j - \pi_{m,m+1}t_{m+1} \right] \\
&\quad -\nu_m \left[\pi_{m-2,m-2-a}t_{m-2-a} + \sum_{j=m-1-a}^{m-1} (\pi_{m-2,j} - \pi_{m-1,j})t_j - \pi_{m-1,m}\lambda t_m \right] \\
&= 0,
\end{aligned}$$

which is of the form $\sum_{j=m-a-2}^{m+2} \eta_{m,j} t_j = 0$ for some constants $\eta_{m,j}$. Thus, we must have $\eta_{m,j} = 0$ for all $j = m-a-2, m-a-1, \ldots, m+2$. Now, from $\eta_{m,m+2} = \eta_{m,m+1} = \eta_{m,m-a-2} = 0$, we obtain the expressions for ω_m, μ_m, and ν_m, respectively, as given in (2.10). $\square$

3. The algorithm

We now return to the general framework of semi-iterative methods for computing the Drazin-inverse solution of singular linear systems that was discussed in Section 1. We choose the polynomials $p_m(\lambda)$ that appear in (1.3) and (1.4) to be precisely those given in Theorem 2.1, the integer a in the latter being $\operatorname{ind}(A)$. As they are in Π_m^0, these $p_m(\lambda)$ already satisfy (1.5) and (1.8).

From (1.3), (1.4), and (2.4), the iterates x_m and x_{m+1} of the semi-iterative method satisfy

$$x_{m+1} - x_m = u_m(A) r_0. \tag{3.1}$$

But the $u_m(\lambda)$ satisfy the 3-term recursion relation given in (2.5). Consequently, the x_m satisfy the 4-term recursion relation

$$x_{m+1} = x_m + \omega_m A(x_m - x_{m-1}) + \mu_m(x_m - x_{m-1}) + \nu_m(x_{m-1} - x_{m-2}), \quad m \geqslant a+1, \tag{3.2}$$

which is exactly of the form given in [8] for the case $a=1$. Note that this recursion relation has the same length *independent* of a.

As the recursion relation above is valid for $m \geqslant a+1$ and as $v_{a+1}=0$, we see that in order to start the algorithm we need x_a and x_{a+1}. Now because $p_a(\lambda)=1$, we have that $q_{a-1}(\lambda)=0$ so that $x_a=x_0$. As for x_{a+1}, we proceed as follows: First, we know that

$$p_{a+1}(\lambda)=1-\rho\lambda^{a+1}$$

for some constant ρ that can be uniquely determined from the characterization property in (2.2). As $\langle p_{a+1},\lambda\rangle=0$, we evidently have that

$$\rho=\frac{\langle 1,\lambda\rangle}{\langle 1,\lambda^{a+2}\rangle}. \tag{3.3}$$

Next, on recalling (2.4), we see that

$$u_a(\lambda)=\rho\lambda^a. \tag{3.4}$$

Finally, from (3.1) and (3.4) we have that

$$x_{a+1}=x_a+\rho A^a r_0=x_a+\rho A^a(b-Ax_0). \tag{3.5}$$

Assuming that the polynomials $t_m(\lambda)$ and the constants α_m and β_m in (2.6) are known, our algorithm now reads as follows:

Step 0: Choose x_0 and set $x_a=x_0$.
Set $\pi_{a,0}=\frac{-1}{t_1'(0)}$, $\pi_{a,1}=-\pi_{a,0}$, and $\pi_{a,j}=0$ for $j\neq 0,1$.
Determine ρ from (3.3).
Compute x_{a+1} from (3.5).

Step 1: For $m=a+1,a+2,\ldots$, until convergence, do:
Solve (2.8) for the $\pi_{m,j}$.
Compute ω_m, μ_m, and v_m from (2.9) and (2.10).
Compute x_{m+1} from (3.2).

For the special case in which the weight function $w(\lambda)$ is that defined by (1.9), the polynomials $t_m(\lambda)$ and the corresponding constants α_m and β_m are given by

$$t_m(\lambda)=\frac{T_m(z(\lambda))}{T_m(z(0))},\quad \text{with } z(\lambda)=\frac{c-\lambda}{d}, \tag{3.6}$$

where $T_m(z)$ are the Chebyshev polynomials of the first kind normalized so that $T_m(1)=1$, and

$$\alpha_0=\frac{1}{c},\quad \beta_0=0,$$
$$\alpha_1=\frac{2c}{2c^2-d^2},\quad \beta_1=c\alpha_1-1,$$
$$\alpha_m=\frac{1}{c-(\frac{1}{2}d)^2\alpha_{m-1}},\quad \beta_m=c\alpha_m-1,\quad m\geqslant 2.$$

In addition, the constant ρ in (3.3) is now given by

$$\rho = \frac{1}{c^{a+1}\sum_{k=0}^{\lfloor a/2\rfloor+1}\binom{a+2}{2k}\binom{2k}{k}\left(\frac{d}{2c}\right)^{2k}}.$$

4. Error bounds and convergence analysis

4.1. General preliminaries and error bounds

Let us denote by $\hat{\mathscr{S}}$ the direct sum of the invariant subspaces of A corresponding to its nonzero eigenvalues λ_j, and by $\tilde{\mathscr{S}}$, its invariant subspace corresponding to its zero eigenvalue. Thus, $\hat{\mathscr{S}} = \mathscr{R}(A^a)$, the range of A^a, and $\tilde{\mathscr{S}} = \mathscr{N}(A^a)$, the nullspace of A^a. Every vector in $\mathbb{C}^n$ can be written as the sum of two unique vectors, one in $\hat{\mathscr{S}}$ and the other in $\tilde{\mathscr{S}}$.

Resolve $b = \hat{b} + \tilde{b}$, where $\hat{b} \in \hat{\mathscr{S}}$ and $\tilde{b} \in \tilde{\mathscr{S}}$. Then $A^{\mathrm{D}}b$, the Drazin-inverse solution of $Ax = b$, is the unique vector in $\hat{\mathscr{S}}$ that satisfies the consistent linear system $Ax = \hat{b}$. From (1.3) and (1.4) we see that

$$\begin{aligned} x_m - A^{\mathrm{D}}b &= p_m(A)x_0 + q_{m-1}(A)(\hat{b} + \tilde{b}) - A^{\mathrm{D}}b \\ &= p_m(A)x_0 + q_{m-1}(A)AA^{\mathrm{D}}b + q_{m-1}(A)\tilde{b} - A^{\mathrm{D}}b \\ &= p_m(A)(x_0 - A^{\mathrm{D}}b) + q_{m-1}(A)\tilde{b}. \end{aligned} \tag{4.1}$$

Decompose $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x}_0 \in \hat{\mathscr{S}}$ and $\tilde{x}_0 \in \tilde{\mathscr{S}}$. Then (4.1) becomes

$$x_m - A^{\mathrm{D}}b = p_m(A)(\hat{x}_0 - A^{\mathrm{D}}b) + p_m(A)\tilde{x}_0 + q_{m-1}(A)\tilde{b}. \tag{4.2}$$

Because

$$p_m(\lambda) = 1 - \lambda^{a+1}u(\lambda) \tag{4.3}$$

for some $u \in \Pi_{m-a-1}$, we have that

$$p_m(A)\tilde{x}_0 = \tilde{x}_0 - u(A)A^{a+1}\tilde{x}_0 = \tilde{x}_0 \tag{4.4}$$

as $\tilde{x}_0 \in \tilde{\mathscr{S}} = \mathscr{N}(A^a)$. Similarly, $q_{m-1}(\lambda) = \lambda^a u(\lambda)$ by (1.4) and (4.3), so that

$$q_{m-1}(A)\tilde{b} = u(A)A^a\tilde{b} = 0 \tag{4.5}$$

as $\tilde{b} \in \tilde{\mathscr{S}} = \mathscr{N}(A^a)$.

Combining (4.4) and (4.5) in (4.2), we deduce the following result.

Theorem 4.1. *Let $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x}_0 \in \hat{\mathscr{S}}$ and $\tilde{x}_0 \in \tilde{\mathscr{S}}$. Then*

$$x_m - A^{\mathrm{D}}b = p_m(A)(\hat{x}_0 - A^{\mathrm{D}}b) + \tilde{x}_0. \tag{4.6}$$

Now, as the vector $\hat{x}_0 - A^{\mathrm{D}}b$ is in $\hat{\mathscr{S}}$, we observe that the behavior of $x_m - A^{\mathrm{D}}b$ is determined by the action of $p_m(A)$ on $\hat{\mathscr{S}}$.

Recall that, by $\mathrm{ind}(A-\lambda_j I)=k_j$, $\lambda_j\in\sigma(A)\backslash\{0\}$, the fact that $\hat{x}_0-A^{\mathrm{D}}b\in\hat{\mathscr{S}}$ implies that

$$p_m(A)(\hat{x}_0-A^{\mathrm{D}}b)=\sum_{\lambda_j\in\sigma(A)\backslash\{0\}}\sum_{i=0}^{k_j-1} r_{ji}\,p_m^{(i)}(\lambda_j) \tag{4.7}$$

for some vectors r_{ji} that lie in the invariant subspace of A corresponding to λ_j. Thus, from (4.6) and (4.7),

$$\|x_m-A^{\mathrm{D}}b-\tilde{x}_0\|=\|p_m(A)(\hat{x}_0-A^{\mathrm{D}}b)\|\leqslant C\left(\max_{\lambda_j\in\sigma(A)\backslash\{0\}}\max_{0\leqslant i\leqslant k_j-1}|p_m^{(i)}(\lambda_j)|\right) \tag{4.8}$$

for some positive constant C. Replacing the maximum over the $\lambda_j\in\sigma(A)\backslash\{0\}$ by the maximum over the interval $[c-d,c+d]$, and using

$$\max_{\lambda\in[c-d,c+d]}|p_m^{(i)}(\lambda)|\leqslant D_i m^{2i}\left(\max_{\lambda\in[c-d,c+d]}|p_m(\lambda)|\right)\quad\text{for some}\quad D_i>0,$$

which follows from one of Markoff's inequalities, see. e.g., Meinardus [9, p. 67], (4.8) becomes

$$\|x_m-A^{\mathrm{D}}b-\tilde{x}_0\|=\|p_m(A)(\hat{x}_0-A^{\mathrm{D}}b)\|\leqslant Mm^{2\hat{k}-2}\left(\max_{\lambda\in[c-d,c+d]}|p_m(\lambda)|\right), \tag{4.9}$$

where M is a positive constant and

$$\hat{k}=\max\{k_j\colon \lambda_j\in\sigma(A)\backslash\{0\}\}. \tag{4.10}$$

Hence, all we have to analyze is $\max_{\lambda\in[c-d,c+d]}|p_m(\lambda)|$.

Before we go on, we observe from (4.6) and (4.7) that the conditions in (1.7) ensure the convergence of $\{x_m\}_{m=0}^{\infty}$ to $A^{\mathrm{D}}b+\tilde{x}_0$, as guaranteed also by Eiermann et al. [4, Lemma 2]. Also, if $\tilde{x}_0=0$, which can be enforced by picking $x_0=0$, then $\lim_{m\to\infty}x_m=A^{\mathrm{D}}b$ under (1.7).

4.2. Convergence analysis

In the sequel, we analyze the case in which the weight function $w(\lambda)$ is that defined by (1.9). Obviously, we first need to know the behavior of the $\pi_{m,j}$ in Theorem 2.3 for $m\to\infty$. For this we have to start with the behavior of the $t_m^{(i)}(0)$ for $m\to\infty$, as is obvious from (2.8). Recall that in this case $t_m(\lambda)$ are as in (3.6).

Lemma 4.2. *Suppose that $t_m(\lambda)$ are the polynomials given in* (3.6). *If $\lambda\in[c-d,c+d]$, then, for $i=0,1,2,\ldots,$*

$$\begin{aligned} t_m^{(i)}(\lambda)&=P_i(\lambda,m)\frac{\mathrm{e}^{m\cosh^{-1}z(\lambda)}+\mathrm{e}^{-m\cosh^{-1}z(\lambda)}}{\mathrm{e}^{m\cosh^{-1}z(0)}+\mathrm{e}^{-m\cosh^{-1}z(0)}}\\ &\quad+N_i(\lambda,m)\frac{\mathrm{e}^{m\cosh^{-1}z(\lambda)}-\mathrm{e}^{-m\cosh^{-1}z(\lambda)}}{\mathrm{e}^{m\cosh^{-1}z(0)}+\mathrm{e}^{-m\cosh^{-1}z(0)}}, \end{aligned} \tag{4.11}$$

where $P_i(\lambda, m)$ and $N_i(\lambda, m)$ are polynomials in m, whose coefficients are functions of λ and whose degree is dependent on the parity of i, given by

$$\begin{aligned} P_{2r}(\lambda, m) &= \left[-\frac{1}{\sqrt{(c-\lambda)^2 - d^2}}\right]^{2r} m^{2r} + \mathrm{O}(m^{2r-2}), \\ N_{2r}(\lambda, m) &= \frac{(2r-1)2r}{2}(c-\lambda)\left[-\frac{1}{\sqrt{(c-\lambda)^2 - d^2}}\right]^{2r+1} m^{2r-1} + \mathrm{O}(m^{2r-3}), \end{aligned} \tag{4.12}$$

with the terms $\mathrm{O}(m^{2r-2})$ and $\mathrm{O}(m^{2r-3})$ missing for $r = 0, 1$, and

$$\begin{aligned} P_{2r+1}(\lambda, m) &= \frac{2r(2r+1)}{2}(c-\lambda)\left[-\frac{1}{\sqrt{(c-\lambda)^2 - d^2}}\right]^{2r+2} m^{2r} + \mathrm{O}(m^{2r-2}), \\ N_{2r+1}(\lambda, m) &= \left[-\frac{1}{\sqrt{(c-\lambda)^2 - d^2}}\right]^{2r+1} m^{2r+1} + \mathrm{O}(m^{2r-1}), \end{aligned} \tag{4.13}$$

with the terms $\mathrm{O}(m^{2r-2})$ and $\mathrm{O}(m^{2r-1})$ missing for $r = 0, 1$, and $r = 0$, respectively.

Proof. The proof is straightforward and proceeds by induction on i. □

Taking $\lambda = 0$, (4.11) becomes

$$t_m^{(i)}(0) = P_i(0, m) + N_i(0, m)\left(1 - \frac{2}{\kappa^{-2m} + 1}\right), \tag{4.14}$$

where κ, defined by (1.10), satisfies also

$$\kappa = \mathrm{e}^{-\cosh^{-1} z(0)}.$$

Upon substituting (4.12) and (4.13) in (4.14), we now have the following result.

Theorem 4.3. *Suppose that $t_m(\lambda)$ are the polynomials defined in* (3.6). *Then, for $i = 0, 1, 2, \ldots,$*

$$t_m^{(i)}(0) = \begin{cases} 1, & i = 0, \\ -\dfrac{1}{(c^2 - d^2)^{1/2}} m + \mathrm{O}(m\kappa^{2m}), & i = 1, \\ (-1)^i \dfrac{1}{(c^2 - d^2)^{i/2}} m^i + \displaystyle\sum_{k=0}^{i-1} \eta_{i,k} m^k + \mathrm{O}(m^l \kappa^{2m}), & i \geqslant 2 \end{cases}$$

as $m \to \infty$, where $\eta_{i,k}$ are some constants and $l = i - 1$ if i is even, and $l = i$ if i is odd.

Theorem 4.3 has the following implication. In order to solve the system in (2.8) for the $\pi_{m,j}$, we first introduce the matrices B and E in $\mathbb{C}^{a+2,a+2}$ and the vector h in $\mathbb{C}^{a+2}$ as follow:

- For $i,j=0,1,2,\ldots,a+1$, set

$$b_{i+1,j+1}=\begin{cases}1, & i=0,\\ -\dfrac{1}{(c^2-d^2)^{1/2}}(m-a+j), & i=1,\\ (-1)^i\dfrac{1}{(c^2-d^2)^{i/2}}(m-a+j)^i+\displaystyle\sum_{k=0}^{i-1}\eta_{i,k}(m-a+j)^k, & 2\leqslant i\leqslant a+1,\end{cases}$$

  with the $\eta_{i,k}$ as in Theorem 4.3.
- For $j=0,1,2,\ldots,a+1$, set

$$e_{i+1,j+1}=\begin{cases}0, & i=0,\\ \mathrm{O}(m^l\kappa^{2m}), & 1\leqslant i\leqslant a+1,\end{cases}$$

  where l is defined in Theorem 4.3, and observe that $E\to\mathrm{O}$ as $m\to\infty$.
- Introduce the vector h via

$$h_{i+1}=\begin{cases}0, & i=0,\\ 1, & i=1,\\ 0, & 2\leqslant i\leqslant a+1.\end{cases}$$

Then the linear system in (2.8) can be written as

$$(B+E)\pi=h, \tag{4.15}$$

where $\pi\in\mathbb{C}^{a+2}$ is the unknown vector whose $(j+1)$th entry, $0\leqslant j\leqslant a+1$, is $\pi_{m,m-a+j}$.

To solve (4.15) for π we apply elementary row operations to obtain the equivalent system

$$(B^{(2)}+E^{(2)})\pi=h^{(2)},$$

where

$$b^{(2)}_{i+1,j+1}=\begin{cases}1, & i=0,\\ (m-a+j)^i, & 1\leqslant i\leqslant a+1,\end{cases}$$

i.e., $B^{(2)}$ is a Vandermonde matrix, and

$$h^{(2)}_{i+1}=\begin{cases}0, & i=0,\\ -(c^2-d^2)^{1/2}, & i=1,\\ K_i(c^2-d^2)^{1/2}, & 2\leqslant i\leqslant a+1,\end{cases}$$

where K_i is a constant that depends only on the coefficients $\eta_{i,k}$.

Now, using the algorithm to solve Vandermonde systems (see [7, p. 122]), we obtain the equivalent system

$$(I+E^{(3)})\pi=h^{(3)},$$

where

$$h_{i+1}^{(3)} = (-1)^i \binom{a+1}{i} \frac{(c^2-d^2)^{1/2}}{a!} m^a + \mathrm{O}(m^{a-1}) \quad \text{for } 0 \leqslant i \leqslant a+1 \tag{4.16}$$

and $E^{(3)} \to O$, as $m \to \infty$. Therefore, we have the following result.

Theorem 4.4. *For $i = 0, 1, 2, \ldots, a, a+1$, the $\pi_{m,m-a+i}$ in* (2.8) *are given by $\pi_{m,m-a+i} = h_{i+1}^{(3)}$, where $h_{i+1}^{(3)}$ are defined in* (4.16).

We now combine Theorem 4.4 with the expansion of $p_m(\lambda)$ in Theorem 2.3 to derive an asymptotically optimal upper bound on $|p_m(\lambda)|$ for $\lambda \in [c-d, c+d]$.

Theorem 4.5. *Consider the polynomials $p_m(\lambda)$ of Theorem* 2.1 *with the weight function $w(\lambda)$ given by* (1.9). *Then*

$$\max_{\lambda \in [c-d,c+d]} |p_m(\lambda)| = \frac{2(\kappa^{-1}-\kappa)(1+\kappa)^{a-1}}{a!} m^a \kappa^{m-a+1} + \mathrm{O}(m^{a-1}\kappa^m) \quad \text{as } m \to \infty, \tag{4.17}$$

where κ is given by (1.10).

Proof. For the weight function $w(\lambda)$ given by (1.9) we have that the polynomials $t_m(\lambda)$ are defined by (3.6).

For $\lambda \in [c-d, c+d]$, it is easy to see that

$$t_m(\lambda) = 2\kappa^m \Re(s(\lambda)^m) + \mathrm{O}(\kappa^{3m}) \quad \text{as } m \to \infty, \tag{4.18}$$

where

$$s(\lambda) = \mathrm{e}^{\mathrm{i} \arccos z(\lambda)} \quad \text{with } |s(\lambda)| = 1$$

and, therefore,

$$\lambda = c - \mathrm{d}z(\lambda) = c - \mathrm{d}\Re(s(\lambda)). \tag{4.19}$$

Now, from (2.7), (3.6), Theorem 4.4, and (4.18), we obtain that for $m \to \infty$,

$$\begin{aligned} p_m(\lambda) &= \frac{1}{\lambda} \sum_{j=0}^{a+1} (-1)^j \binom{a+1}{j} \frac{\sqrt{c^2-d^2}}{a!} m^a 2\kappa^{m-a+j} \Re(s(\lambda)^{m-a+j}) + \mathrm{O}(m^{a-1}\kappa^m) \\ &= \frac{2\sqrt{c^2-d^2}}{a!\lambda} m^a \kappa^{m-a+1} \Re\left(\frac{s(\lambda)^{m-a}}{\kappa} [1-\kappa s(\lambda)]^{a+1} \right) + \mathrm{O}(m^{a-1}\kappa^m). \end{aligned} \tag{4.20}$$

On the other hand, from (1.10) we have that

$$\kappa^{-1} + \kappa = 2\frac{c}{d} \quad \text{and} \quad \kappa^{-1} - \kappa = \frac{2\sqrt{c^2-d^2}}{d}. \tag{4.21}$$

Now, using (4.19) and the fact that $|s(\lambda)| = 1$, we conclude that

$$\lambda = \tfrac{1}{2}d\kappa^{-1}[1 - \kappa s(\lambda)][1 - \kappa\overline{s(\lambda)}].$$

Inserting this expression into (4.20) and using (4.21), we obtain that

$$p_m(\lambda) = \frac{2}{a!}(\kappa^{-1} - \kappa)m^a\kappa^{m-a+1}\Re\left(\frac{s(\lambda)^{m-a}[1 - \kappa s(\lambda)]^a}{1 - \kappa\overline{s(\lambda)}}\right) + \mathrm{O}(m^{a-1}\kappa^m) \quad \text{as } m \to \infty.$$

Then

$$\frac{1}{m^a\kappa^{m-a+1}}p_m(\lambda) = \frac{2}{a!}(\kappa^{-1} - \kappa)\Re\left(\frac{s(\lambda)^{m-a}[1 - \kappa s(\lambda)]^a}{1 - \kappa\overline{s(\lambda)}}\right) + \mathrm{O}(m^{-1}) \quad \text{as } m \to \infty. \tag{4.22}$$

Next, using the facts that $|s(\lambda)| = 1$ and $|1 - \kappa s(\lambda)| = |1 - \kappa\overline{s(\lambda)}|$, we conclude that

$$\max_{\lambda\in[c-d,c+d]}\left|\frac{s(\lambda)^{m-a}[1 - \kappa s(\lambda)]^a}{1 - \kappa\overline{s(\lambda)}}\right| = \max_{\lambda\in[c-d,c+d]}|1 - \kappa s(\lambda)|^{a-1} = (1 + \kappa)^{a-1} \tag{4.23}$$

and this maximum is attained at $\lambda = c+d$ for which $s(\lambda) = -1$. Finally, combining (4.23) and (4.22), we obtain that

$$\frac{1}{m^a\kappa^{m-a+1}}\max_{\lambda\in[c-d,c+d]}|p_m(\lambda)| = \frac{2(\kappa^{-1} - \kappa)(1 + \kappa)^{a-1}}{a!} + \mathrm{O}(m^{-1}) \quad \text{as } m \to \infty,$$

from which (4.17) follows. □

Now, using (4.9) and (4.17), we have the following convergence result that is the main result of this section.

Corollary 4.6. *With the same notation as in Theorem* 4.5, *we have*

$$\|x_m - A^{\mathrm{D}}b - \tilde{x}_0\| = \mathrm{O}(m^{a+2\hat{k}-2}\kappa^m) \quad \textit{as } m \to \infty,$$

where $\hat{k}$ is as defined in (4.10).

Theorem 4.5 implies that

$$\max_{\lambda\in[c-d,c+d]}|p_m(\lambda)| \approx \frac{2(\kappa^{-1} - \kappa)(1 + \kappa)^{a-1}}{a!}m^a\kappa^{m-a+1}.$$

On the other hand, the Berstein result as applied by Eiermann and Starke to the polynomials $\{p_m\}$ developed in their paper, see [6, p. 314], gives that their residual polynomials satisfy that

$$\max_{\lambda\in[c-d,c+d]}|p_m(\lambda)| \approx \frac{2(\kappa^{-1} - \kappa)^a}{a!}m^a\kappa^m. \tag{4.24}$$

Now,

$$\frac{2(\kappa^{-1} - \kappa)^a}{a!}m^a\kappa^m < \frac{2(\kappa^{-1} - \kappa)(1 + \kappa)^{a-1}}{a!}m^a\kappa^{m-a+1},$$

because $1 > -\kappa$. Therefore, our polynomials are not "near-optimal". However, the residual polynomials $\{p_m\}$ constructed by Eiermann and Starke in [6] cannot be computed by means of short recurrences as we have developed for the present residuals in Section 3. Such short recurrences make for the efficient implimentation of semi-iterative methods. In this regard please see also the comments on Hanke and Hochbruck [8, pp. 90, 93].

5. Numerical examples

In this section we use the algorithm developed in Section 3 to compute the eigenprojection $Z_A := I - AA^{\mathrm{D}}$ onto the eigenspace of A corresponding to the eigenvalue 0 of three singular matrices whose index exceeds 1.

If we take $b = 0$ in (1.1) then Corollary 4.6 implies

$$\lim_{m \to \infty} x_m = \tilde{x}_0 = (I - AA^{\mathrm{D}})x_0.$$

Now, if we choose x_0 as the ith column of I, the above expression represents the ith column of the eigenprojection Z_A.

First, consider the following singular M-matrix:

$$A_1 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & -1 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 & 0 \\ -1 & -1 & -1 & 0 & 2 & -1 \\ -1 & -1 & 0 & -1 & -1 & 2 \end{bmatrix}.$$

Observe that $\sigma(A_1) = \{0, 0, 1, 2, 2, 3\}$ and $a = 2$. So we can choose $c = 2$ and $d = 1$ in (1.2). Using the algorithm of Section 3, with the polynomials $t_m(\lambda)$ defined by (3.6), and stopping when

$$\frac{\|x_{m+1} - x_m\|_\infty}{\|x_m\|_\infty} \leqslant 10^{-15},$$

we obtain, after 35 iterations, that

$$Z_{A_1} = \begin{bmatrix} 0.50000000000001 & 0.49999999999999 & 0 & 0 & 0 & 0 \\ 0.49999999999999 & 0.50000000000001 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.50000000000001 & 0.49999999999999 & 0 & 0 \\ 0 & 0 & 0.49999999999999 & 0.50000000000001 & 0 & 0 \\ 0 & 0 & 0.50000000000005 & 0.49999999999997 & -0.00000000000004 & 0.00000000000002 \\ 0 & 0 & 0.49999999999997 & 0.50000000000005 & 0.00000000000002 & -0.00000000000004 \end{bmatrix}.$$

The exact eigenprojection is given by

$$
\begin{bmatrix}
0.5 & 0.5 & 0 & 0 & 0 & 0 \\
0.5 & 0.5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.5 & 0.5 & 0 & 0 \\
0 & 0 & 0.5 & 0.5 & 0 & 0 \\
0 & 0 & 0.5 & 0.5 & 0 & 0 \\
0 & 0 & 0.5 & 0.5 & 0 & 0
\end{bmatrix}.
$$

As a second example, we again consider a singular M-matrix, this time of index $a=4$:

$$
A_2 = \begin{bmatrix}
1.0 & -1.0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1.0 & 1.0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1.0 & -1.0 & 1.0 & -1.0 & 0 & 0 & 0 & 0 \\
-1.0 & -1.0 & -1.0 & 1.0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1.0 & -1.0 & -1.0 & -1.0 \\
0 & 0 & 0 & 0 & -1.0 & 1.0 & -1.0 & -1.0 \\
0 & 0 & 0 & -1.0 & 0 & 0 & 1.0 & -1.0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1.0 & 1.0
\end{bmatrix}.
$$

Here $\sigma(A_2)=\{0,0,0,0,2,2,2,2\}$. With $c=2$ and $d=1$ in (1.2) we get using the algorithm in Section 3 that after 25 iterations for columns $1,2,5,6$, and 7 and 45 iterations for columns 3 and 4 that

$$
Z_{A_2} = \begin{bmatrix}
5.0000\times10^{-1} & 5.0000\times10^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\
5.0000\times10^{-1} & 5.0000\times10^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 5.0000\times10^{-1} & 5.0000\times10^{-1} & 0 & 0 & 0 & 0 \\
0 & 0 & 5.0000\times10^{-1} & 5.0000\times10^{-1} & 0 & 0 & 0 & 0 \\
0 & 0 & 1.2500\times10^{-1} & -1.2500\times10^{-1} & 5.0000\times10^{-1} & 5.0000\times10^{-1} & 0 & 0 \\
0 & 0 & 1.2500\times10^{-1} & -1.2500\times10^{-1} & 5.0000\times10^{-1} & 5.0000\times10^{-1} & 0 & 0 \\
-1.2500\times10^{-1} & -1.2500\times10^{-1} & -5.7858\times10^{-12} & 2.5000\times10^{-1} & 0 & 0 & 5.0000\times10^{-1} & 5.0000\times10^{-1} \\
1.2500\times10^{-1} & 1.2500\times10^{-1} & -2.5000\times10^{-1} & 5.3423\times10^{-11} & 0 & 0 & 5.0000\times10^{-1} & 5.0000\times10^{-1}
\end{bmatrix}.
$$

The exact eigenprojection is given here by

$$
\begin{bmatrix}
0.5000 & 0.5000 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5000 & 0.5000 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.5000 & 0.5000 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.5000 & 0.5000 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.1250 & -0.1250 & 0.5000 & 0.5000 & 0 & 0 \\
0 & 0 & 0.1250 & -0.1250 & 0.5000 & 0.5000 & 0 & 0 \\
-0.1250 & -0.1250 & 0 & 0.2500 & 0 & 0 & 0.5000 & 0.5000 \\
0.1250 & 0.1250 & -0.2500 & 0 & 0 & 0 & 0.5000 & 0.5000
\end{bmatrix}.
$$

Finally, we consider a singular matrix A with $a=3$.

$$A_3 = \begin{bmatrix} 5 & -1 & -1 & -1 & -1 & 0 & -1 \\ 1 & 3 & -1 & -1 & -1 & 0 & -1 \\ 0 & 0 & 3 & -1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

Here $\sigma(A_3) = \{0,0,0,2,2,4,4\}$, so we can take $c=3$ and $d=1$. Then, using the algorithm in Section 3 we get, after 51 iterations for columns 1, 2, 3 and 4, 29 iterations for column 5, and 6 iterations for columns 6 and 7 that

$$Z_{A_3} = \begin{bmatrix} -0.3908\times10^{-12} & 0.1799\times10^{-12} & 0.0267\times10^{-12} & 0.1909\times10^{-12} & 1.0000 & 0 & 0 \\ -0.1846\times10^{-12} & -0.0252\times10^{-12} & 0.0270\times10^{-12} & 0.1909\times10^{-12} & 1.0000 & 0 & 0 \\ 0 & 0 & -0.1840\times10^{-12} & 0.1909\times10^{-12} & 1.0000 & 0 & 0 \\ 0 & 0 & -0.1890\times10^{-12} & 0.1970\times10^{-12} & 1.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.0000 \end{bmatrix}.$$

The exact value is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

References

[1] A. Ben-Israel, T.N.E. Greville, Generalized Inverses, Theory and Applications, Wiley, New York, 1974.

[2] S.L. Campbell, C.D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

[3] P.J. Davis, Interpolation and Approximation, Blaisdell, New York, 1963.

[4] M. Eiermann, I. Marek, W. Niethammer, On the solution of singular linear systems of algebraic equations by semiiterative methods, Numer. Math. 53 (1988) 265–283.

[5] M. Eiermann, L. Reichel, On the application of orthogonal polynomials to the iterative solution of singular linear systems of equations, in: J. Dongarra, I. Duff, P. Gaffney, S. McKee (Eds.), Vector and Parallel Computing, Wiley, New York, 1989, pp. 285–297.

[6] M. Eiermann, G. Starke, The near-best solution of a polynomial minimization problem by the Carathéodory–Fejér method, Constr. Approx. 6 (1990) 303–319.

[7] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, MD, 1983.
[8] M. Hanke, M. Hochbruck, A Chebyshev-like semiiteration for inconsistent linear systems, Electronic Transactions on Numerical Analysis 1 (1993) 89–103.
[9] G. Meinardus, Approximation of Functions: Theory and Numerical Methods, Springer, New York, 1967.
[10] A. Sidi, Development of iterative techniques and extrapolation methods for Drazin inverse solution of consistent or inconsistent singular linear systems, Linear Algebra Appl. 167 (1992) 171–203.