On Smooth LU Decompositions with Applications to Solutions of Nonlinear Eigenvalue Problems

Article *in* Journal of Computational Mathematics · November 2010

DOI: 10.4208/jcm.1004-m0009

CITATIONS
12

READS
121

2 authors, including:

Hua Dai
Nanjing University of Aeronautics & Astronautics
149 PUBLICATIONS **916** CITATIONS

SEE PROFILE

Journal of Computational Mathematics
Vol.28, No.?, 2010, 1–22.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1004-m0009

ON SMOOTH LU DECOMPOSITIONS WITH APPLICATIONS TO SOLUTIONS OF NONLINEAR EIGENVALUE PROBLEMS*

Hua Dai
Department of Mathematics, Nanjing University of Aeronautics and Astronautics
Nanjing 210016, China
Email: hdai@nuaa.edu.cn
Zhong-Zhi Bai
LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences
Beijing 100190, China
Email: bzz@lsec.cc.ac.cn

Abstract

We study the smooth LU decomposition of a given analytic functional λ-matrix $A(\lambda)$ and its block-analogue. Sufficient conditions for the existence of such matrix decompositions are given, some differentiability about certain elements arising from them are proved, and several explicit expressions for derivatives of the specified elements are provided. By using these smooth LU decompositions, we propose two numerical methods for computing multiple nonlinear eigenvalues of $A(\lambda)$, and establish their locally quadratic convergence properties. Several numerical examples are provided to show the feasibility and effectiveness of these new methods.

Mathematics subject classification: 15A18, 15A23, 65F15.
Key words: Matrix-valued function, Smooth LU decomposition, Pivoting, Nonlinear eigenvalue problem, Multiple eigenvalue, Newton method.

1. Introduction

The importance of constant matrix decompositions cannot be overstressed, as they are not only basic methods in matrix computations and analyses, but also applicable tools in other areas beyond mathematics. However, the smooth decompositions of matrices depending on some parameters are by no means less important.

In fact, fundamental theory on decompositions of a matrix-valued function $A(\lambda)$ is given in Kato's book [21], in which one of the strongest results is that $A(\lambda)$ has an analytic spectral decomposition in the case that $A(\lambda)$ is real, analytic and Hermitian. Here, $A(\lambda) = (a_{ij}(\lambda))$ denotes an n-by-n matrix with all elements $a_{ij}(\lambda)$, $i, j = 1, 2, \ldots, n$, being analytic functions with respect to a real or a complex parameter λ; it is called a functional λ-matrix in [20] or a matrix-valued function in [26]; in particular, if $a_{ij}(\lambda)$ are polynomials in λ, then it is called a λ-matrix or a matrix polynomial in [25]. Based on the work in [20,23,34], Li [27,28] developed QR decomposition and its block-analogue for a differentiable matrix-valued function, and gave sufficient conditions for guaranteeing the existence of a differentiable QR decomposition. In [6] the authors showed that a real analytic matrix-valued function admits an analytic singular value decomposition, and Wright further presented a numerical method in [43] for finding a

* Received February 3, 2009 / Revised version received xxx, 2009 / Accepted June 6, 2009 / Published online xxx /

smooth singular value decomposition of a matrix solely depending on a single parameter. It is noticed that Gingold and Hsieh [13] used a different approach to show that a real analytic matrix-valued function with only real eigenvalues admits an analytic Schur decomposition. Then, Dieci and Eirola [10] considered smooth QR, smooth Schur, and smooth singular value decompositions as well as their block-analogues, gave sufficient conditions for guaranteeing the existence of such matrix decompositions, and derived differential equations for the involved factors. Recently, Rebaza [33] applied smooth block Schur decompositions of matrix-valued functions to numerical computations of the separatrices in dynamical systems, and analyzed the actual implementations of the correspondingly induced numerical method.

One important application of the above-described smooth matrix decompositions is that they may lead to effective numerical methods for solving *nonlinear eigenvalue problems* (**NEPs**) of the form

$$A(\lambda)x = 0, \tag{1.1}$$

where λ and x, known as the eigenvalue and the eigenvector, respectively, are the variables to be determined. The NEP (1.1), including the typical linear and quadratic eigenvalue problems as special cases, is of great importance in a large number of disciplines of scientific computing and engineering applications such as density functional theory calculations [4], vibration of viscoelastic structures [9], dynamic finite element method [11], photonic band structure calculations [38], vibration of fluid-solid structures [40], and so on.

When $A(\lambda)$ is a matrix polynomial, the NEP (1.1) can be reformulated as a linear eigenvalue problem [24, 31] and, hence, it may be effectively solved by the Jacobi-Davidson method; see, e.g., [19,36]. In general, the NEP (1.1) may be first reformulated as a system of nonlinear equations through either adding a normalization equation $v^*x = 1$ or computing the characteristic polynomial $c(\lambda) := \det(A(\lambda))$, and then solve the so-obtained nonlinear system by utilizing the classical Newton method, see [2, 22, 25, 34], where and in the sequel, $\det(\cdot)$ is used to represent the determinant of the corresponding matrix. Here, the vector v^*, the conjugate transpose of the complex vector v, can be chosen flexibly so that either the nonlinear function may satisfy certain desired property or various iteration methods can be produced. For example, for $v = x$ we can obtain the Rayleigh quotient iteration, see [24, 25, 34].

It was shown in [34] that the Newton method resulted from the first approach is equivalent to the inverse iteration discussed in [39], which converges quadratically under the nondegeneracy condition [37, 38]. However, the inverse iteration requires to solve a linear system with the coefficient matrix $A(\lambda^{(k)})$ at each iteration step k. The Newton method resulted from the second approach has the succinct expression

$$\lambda^{(k+1)} = \lambda^{(k)} - \frac{1}{\mathrm{tr}\left(A(\lambda^{(k)})^{-1}A'(\lambda^{(k)})\right)}, \qquad k = 0, 1, 2, \ldots,$$

where $\mathrm{tr}(\cdot)$ denotes the trace of the corresponding matrix and $A'(\lambda^{(k)})$ the derivative with respect to the variable λ of the matrix-valued function $A(\lambda)$ at $\lambda = \lambda^{(k)}$. For details, we refer to [22, 25]. Once the eigenvalues are available, the eigenvectors can be approximated by the inverse iteration [32, 37]. However, we should mention that the latter approach is less efficient for larger matrices. We refer to [1] for another elegant derivation of a scalar function having the same zeros as $c(\lambda)$ and for several iteration methods.

To overcome the disadvantages of the afore-described Newton-type methods, many authors have presented and analyzed various iteration methods such as the modified inverse iteration [29], the Arnoldi method [41], the rational Krylov subspace method [35], and the Jacobi-

Davidson method [5]. However, none of these methods can exempt from the intrinsic difficulties caused from the nonlinearity.

Differently, based on the QR decomposition of the matrix-valued function $A(\lambda)$ with column pivoting, Kublanovskaya [23] developed a more implicit strategy for constructing iterative methods of the Newton-type for finding eigenvalues of the NEP (1.1). Let

$$A(\lambda)P(\lambda) = Q(\lambda)R(\lambda)$$

be the column-pivoting QR decomposition of $A(\lambda)$, where $P(\lambda)$ is a permutation matrix, $Q(\lambda)$ a unitary matrix, and $R(\lambda) = (r_{ij}(\lambda))$ an upper-triangular matrix with the diagonal elements being ordered as

$$r_{11}(\lambda) \geq r_{22}(\lambda) \geq \cdots \geq r_{nn}(\lambda).$$

Then the zero points of the function $e(\lambda) := r_{nn}(\lambda)$ are the eigenvalues of the NEP (1.1). The classical Newton method can be employed to compute the zero points of $e(\lambda)$, which is well defined by using the smooth QR decomposition developed in [27] for a differentiable $A(\lambda)$.

Of course, a Newton-type method based on the smooth singular value decomposition was also discussed for solving the NEP (1.1) in [15], but it could be computationally too complicated and costly to be used in actual applications.

Because an LU decomposition of a constant matrix is more efficient and easily realizable than a QR or a singular value decomposition, we can alternatively establish Newton-type methods for solving the NEP (1.1) through computing the zeros of $f(\lambda) := u_{nn}(\lambda)$ determined by the following LU decomposition of the matrix-valued function $A(\lambda)$ with partial pivoting:

$$P(\lambda)A(\lambda) = L(\lambda)U(\lambda),$$

where $P(\lambda)$ is a permutation matrix, $L(\lambda)$ a unit lower-triangular matrix, and $U(\lambda) = (u_{ij}(\lambda))$ an upper-triangular matrix; see [20,42]. Note that a theoretical guarantee for the well-definiteness of the so-derived Newton-type methods is the continuous differentiability of the decomposition factor $u_{nn}(\lambda)$. Here, we remark that in general it is not necessary for every eigenvalue of $A(\lambda)$ to be a zero point of $f(\lambda)$, but a complete pivoting strategy may guarantee the validity of this statement. For example, we consider the matrix-valued function

$$A(\lambda) = \begin{pmatrix} \lambda - 1 & 1 \\ 0 & \lambda - 3 \end{pmatrix}$$

of an upper-triangular form. Evidently, $f(\lambda) := u_{22}(\lambda) = \lambda - 3$ only has a single zero point, whereas $A(\lambda)$ has two eigenvalues 1 and 3. However, when a complete pivoting strategy is adopted, the correspondingly obtained function $f(\lambda) := u_{22}(\lambda)$ just have two zero points 1 and 3 with respect to different initial values, which gives exactly all eigenvalues of the $A(\lambda)$.

In this paper, we will develop the smooth LU decomposition and its block-analogue for the matrix-valued function $A(\lambda)$, so that a theoretical foundation for using the Newton-type methods to solve the NEP (1.1) can be established. Moreover, we will give sufficient conditions for guaranteeing the existence of such a matrix decomposition, derive several explicit expressions for derivatives of certain elements in the decomposition factors, and present efficient procedures for computing these derivatives. Based on the smooth LU decomposition and its theory, we will establish new Newton-type methods for computing the multiple eigenvalues of the NEP (1.1), and prove their locally quadratic convergence.

The organization of the paper is as follows. In Section 2, we develop the smooth LU decomposition and its block-analogue for the matrix-valued function, prove the existence theorems,

and give derivative expressions of certain involved elements. In Section 3, two Newton-type methods based on the smooth LU decompositions with complete pivoting are established for computing the multiple eigenvalues of the NEP (1.1). The locally quadratic convergence of these new methods are analyzed in Section 4, and several numerical examples are given in Section 5 to show the feasibility, effectiveness and accuracy of the new methods. Finally, in Section 6, we use some conclusions and remarks to end this paper.

2. The Smooth LU Decompositions

In this section, we will develop the smooth LU decomposition and its block-analogue for the matrix-valued function $A(\lambda)$.

For convenience, we introduce the following notations. $\mathbf{R}^{m\times n}$ and $\mathbf{C}^{m\times n}$ represent the sets of all real and complex $m\times n$ matrices, respectively. In particular, it stands that $\mathbf{R}^m = \mathbf{R}^{m\times 1}$, $\mathbf{C}^m = \mathbf{C}^{m\times 1}$, $\mathbf{R} = \mathbf{R}^1$ and $\mathbf{C} = \mathbf{C}^1$. $\mathcal{L}_n$ and $\mathcal{U}_n$ denote the sets of all $n\times n$ unit lower-triangular matrices and all $n\times n$ upper-triangular matrices, respectively. We use I to denote the identity matrix of suitable size, e_j the j-th column vector of I, $I^{(j)} = [e_1, e_2, \ldots, e_j]$, and $K^{(j)} = [e_{n-j+1}, e_{n-j+2}, \ldots, e_n]$. In addition, $\|\cdot\|_2$ and $\|\cdot\|_F$ represent the Euclidean and the Frobenius matrix norms, respectively, and $\|\cdot\|$ is used to simply denote the Euclidean vector or matrix norm without causing any confusion. An $m\times n$ matrix A can be partitioned into columns as $A = [a_1, a_2, \ldots, a_n]$, with a_j its j-th column vector. Then we can define the vector $\mathrm{col}(A) := [a_1^T, a_2^T, \ldots, a_n^T]^T$, where $(\cdot)^T$ denotes the transpose of either a vector or a matrix.

2.1. The Smooth LU Decompositions

We first develop the smooth LU decomposition in the pointwise form for the matrix-valued function $A(\lambda)$. To this end, we need the following preparatory lemmas.

Lemma 2.1. *Let $\mathbb{D}$ be a domain in $\mathbf{R}$ or $\mathbf{C}$, and $A(\lambda)$ be an analytic matrix-valued function defined on $\mathbb{D}$. Then there exist a unit lower-triangular matrix $L(\lambda)$ and an upper-triangular matrix $U(\lambda)$, both unique and analytic in $\mathbb{D}$, such that*

$$A(\lambda) = L(\lambda)U(\lambda),$$

provided all leading principal minors of $A(\lambda)$ are nonzero for any $\lambda\in\mathbb{D}$.

Lemma 2.1 can be easily proved by induction. By directly extending Lemma 2.1 in [10], we can get the following result.

Lemma 2.2. *Assume that $B(\lambda)\in\mathbf{C}^{m\times m}$, $C(\lambda)\in\mathbf{C}^{n\times n}$ and $F(\lambda)\in\mathbf{C}^{m\times n}$ are analytic matrix-valued functions defined on a domain $\mathbb{D}\subseteq\mathbf{C}$, and both $B(\lambda)$ and $C(\lambda)$ are invertible on $\mathbb{D}$. Then the solution of the matrix equation*

$$B(\lambda)X(\lambda)C(\lambda) = F(\lambda)$$

is analytic on $\mathbb{D}$.

Now, we establish the main theorem about the existence of the smooth LU decomposition in the pointwise form for the matrix-valued function $A(\lambda)$.

Theorem 2.1. *Let $A(\lambda) \in \mathbf{C}^{n\times n}$ be a differentiable matrix-valued function on the domain $\mathbb{D} \subseteq \mathbf{C}$ such that $\det(A(\lambda_0)) \neq 0$ holds at a given point $\lambda_0 \in \mathbb{D}$. Assume that there exists an $n \times n$ permutation matrix P such that $PA(\lambda_0)$ has an LU decomposition*

$$PA(\lambda_0) = L_0U_0, \tag{2.1}$$

with $L_0 \in \mathcal{L}_n$ and $U_0 \in \mathcal{U}_n$. Then there exists a neighborhood of λ_0, say, $\mathbb{N}(\lambda_0) \subseteq \mathbb{D}$, such that for all $\lambda \in \mathbb{N}(\lambda_0)$ the matrix-valued function $PA(\lambda)$ has an LU decomposition

$$PA(\lambda) = L(\lambda)U(\lambda), \quad \textit{for} \quad \lambda \in \mathbb{N}(\lambda_0),$$

which satisfies

$$L(\lambda_0) = L_0 \ \textit{ and } \ U(\lambda_0) = U_0.$$

In addition, both $L(\lambda)$ and $U(\lambda)$ are differentiable at the point λ_0. Moreover, it holds that

$$\begin{aligned} e_n^TU(\lambda)e_n =& e_n^TU_0e_n + e_n^TL_0^{-1}PA'(\lambda_0)\Big(e_n - I^{(n-1)}(I^{(n-1)^T}U_0I^{(n-1)})^{-1}I^{(n-1)^T}U_0e_n\Big) \\ & \cdot(\lambda-\lambda_0) + o(|\lambda-\lambda_0|). \end{aligned} \tag{2.2}$$

Proof. Clearly, $A(\lambda)$ can be expressed as

$$A(\lambda) = A(\lambda_0) + A'(\lambda_0)(\lambda-\lambda_0) + o(|\lambda-\lambda_0|). \tag{2.3}$$

Denote

$$E(\lambda) = A'(\lambda_0)(\lambda-\lambda_0) + o(|\lambda-\lambda_0|)$$

satisfying $E(\lambda_0) = 0$. Hence, it holds that

$$L_0^{-1}PA(\lambda) = L_0^{-1}PA(\lambda_0) + L_0^{-1}PE(\lambda) := U_0 + \widehat{E}(\lambda),$$

with

$$\widehat{E}(\lambda) = L_0^{-1}PA'(\lambda_0)(\lambda-\lambda_0) + o(|\lambda-\lambda_0|)$$

satisfying

$$\widehat{E}(\lambda_0) = \lim_{\lambda\to\lambda_0}\widehat{E}(\lambda) = 0.$$

Now, partition the matrices U_0 and $\widehat{E}(\lambda)$ into the blockwise forms as

$$U_0 = \begin{pmatrix} U_{11} & U_{12} \\ 0 & u_{nn} \end{pmatrix} \quad \text{and} \quad \widehat{E}(\lambda) = \begin{pmatrix} \widehat{E}_{11}(\lambda) & \widehat{E}_{12}(\lambda) \\ \widehat{E}_{21}(\lambda) & \widehat{e}_{nn}(\lambda) \end{pmatrix},$$

respectively, with $U_{11}, \widehat{E}_{11}(\lambda) \in \mathbf{C}^{(n-1)\times(n-1)}$. Then we know from the hypothesis $\det(A(\lambda_0)) \neq 0$ that U_{11} is invertible and u_{nn} is nonzero. Therefore, there exists a neighborhood of λ_0, say, $\mathbb{N}(\lambda_0) \subseteq \mathbb{D}$, such that

$$\widehat{U}_{11}(\lambda) := U_{11} + \widehat{E}_{11}(\lambda)$$

is invertible. Moreover, because $A(\lambda)$ is differentiable on the domain $\mathbb{D} \subseteq \mathbf{C}$, the matrix-valued functions $\widehat{E}_{11}(\lambda)$, $\widehat{E}_{12}(\lambda)$ and $\widehat{E}_{21}(\lambda)$ as well as the λ-function $\widehat{e}_{nn}(\lambda)$ are differentiable at λ_0. It follows from Lemma 2.2 that the matrix-valued function

$$L_1(\lambda) := \begin{pmatrix} I & 0 \\ -\widehat{E}_{21}(\lambda)(U_{11}+\widehat{E}_{11}(\lambda))^{-1} & 1 \end{pmatrix}$$

is differentiable at λ_0. Moreover, straightforward operations can yield

$$L_1(\lambda)^{-1}L_0^{-1}PA(\lambda) = \begin{pmatrix} \widehat{U}_{11}(\lambda) & \widehat{U}_{12}(\lambda) \\ 0 & u_{nn}(\lambda) \end{pmatrix},$$

where

$$\widehat{U}_{12}(\lambda) = U_{12} + \widehat{E}_{12}(\lambda)$$

and

$$\begin{aligned} u_{nn}(\lambda) &= u_{nn} + \widehat{e}_{nn}(\lambda) - \widehat{E}_{21}(\lambda)(U_{11} + \widehat{E}_{11}(\lambda))^{-1}(U_{12} + \widehat{E}_{12}(\lambda)) \\ &= u_{nn} + e_n^T L_0^{-1} PA'(\lambda_0)\Big(e_n - I^{(n-1)}U_{11}^{-1}U_{12}\Big)(\lambda - \lambda_0) + o(|\lambda - \lambda_0|). \end{aligned} \tag{2.4}$$

Therefore, $\widehat{U}_{11}(\lambda)$, $\widehat{U}_{12}(\lambda)$ and $u_{nn}(\lambda)$ are differentiable at λ_0, and satisfy

$$\widehat{U}_{11}(\lambda_0) = U_{11}, \quad \widehat{U}_{12}(\lambda_0) = U_{12} \quad \text{and} \quad u_{nn}(\lambda_0) = u_{nn}.$$

Here, we have used the fact that $\widehat{E}(\lambda_0) = 0$.

By Lemma 2.1, there exist a unit lower-triangular matrix $\widehat{L}_2(\lambda)$ and an upper-triangular matrix $\widehat{U}_2(\lambda)$, both unique and differentiable at λ_0, such that

$$\widehat{U}_{11}(\lambda) = \widehat{L}_2(\lambda)\widehat{U}_2(\lambda)$$

holds true, with

$$\widehat{L}_2(\lambda_0) = I \quad \text{and} \quad \widehat{U}_2(\lambda_0) = U_{11}.$$

Define the matrix-valued functions

$$L(\lambda) = L_0 L_1(\lambda) \begin{pmatrix} \widehat{L}_2(\lambda) & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad U(\lambda) = \begin{pmatrix} \widehat{U}_2(\lambda) & \widehat{L}_2(\lambda)^{-1}\widehat{U}_{12}(\lambda) \\ 0 & u_{nn}(\lambda) \end{pmatrix}.$$

Then it holds that

$$L(\lambda_0) = L_0 \quad \text{and} \quad U(\lambda_0) = U_0,$$

and $PA(\lambda)$ possesses the LU decomposition

$$PA(\lambda) = L(\lambda)U(\lambda), \quad \text{for} \quad \lambda \in \mathbb{N}(\lambda_0).$$

Moreover, both $L(\lambda)$ and $U(\lambda)$ are differentiable at λ_0. Finally, the expression (2.2) about the last diagonal element of $U(\lambda)$ follows directly from (2.1) and (2.4). □

The proof of Theorem 2.1 readily implies the following observation.

Remark 2.1. If $A(\lambda) \in \mathbf{C}^{n\times n}$ is further assumed to be twice continuously differentiable on the domain $\mathbb{D} \subseteq \mathbf{C}$, then the term $o(|\lambda - \lambda_0|)$ involved in (2.2), as well as (2.3) and (2.4), may be replaced by a higher-order quantity $\mathcal{O}(|\lambda - \lambda_0|^2)$.

Furthermore, through a slight and technical modification of the demonstration of Theorem 2.1, we can obtain the following result about the smooth LU decomposition for the rank-deficient case.

Corollary 2.1. *Let $A(\lambda) \in \mathbf{C}^{n\times n}$ be a differentiable matrix-valued function on the domain $\mathbb{D} \subseteq \mathbf{C}$ such that* $\operatorname{rank}(A(\lambda_0)) = n - 1$ *holds at a given point $\lambda_0 \in \mathbb{D}$. Assume that there exist two $n \times n$ permutation matrices P_l and P_r such that $P_l A(\lambda_0) P_r$ has an LU decomposition*

$$P_l A(\lambda_0) P_r = L_0 U_0,$$

with $L_0 \in \mathcal{L}_n$ and $U_0 \in \mathcal{U}_n$. Then there exists a neighborhood of λ_0, say, $\mathbb{N}(\lambda_0) \subseteq \mathbb{D}$, such that for all $\lambda \in \mathbb{N}(\lambda_0)$ the matrix-valued function $P_l A(\lambda) P_r$ has an LU decomposition

$$P_l A(\lambda) P_r = L(\lambda) U(\lambda), \quad \textit{for} \quad \lambda \in \mathbb{N}(\lambda_0),$$

which satisfies $L(\lambda_0) = L_0$ and $U(\lambda_0) = U_0$. In addition, both $L(\lambda)$ and $U(\lambda)$ are differentiable at the point λ_0.

2.2. The Smooth Block LU Decomposition

We now turn to develop the smooth LU decomposition in the blockwise form for the matrix-valued function $A(\lambda)$.

Given a positive integer m $(1 \leq m < n)$. For a constant matrix $A \in \mathbf{C}^{n\times n}$, we may define its block LU decomposition of index m, abbreviated as BLU(m) decomposition, as follows:

$$A = LU, \quad \text{with} \quad L = \begin{pmatrix} L_{11} & 0 \\ L_{21} & I \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix}, \tag{2.5}$$

where $L_{11} \in \mathcal{L}_{n-m}$, $U_{11} \in \mathcal{U}_{n-m}$, and $U_{22} \in \mathbf{C}^{m\times m}$. Note that U_{22} is a general matrix block. The set of all matrices $L \in \mathcal{L}_n$ of the special form shown in (2.5) is denoted by $\mathcal{L}_{n,m}$.

The definition of the BLU(m) decomposition is, in spirit, similar to the QR-like decomposition presented in [8, 28]. In particular, when $m = 1$, it naturally reduces to the standard LU decomposition. Alternatively, we may perform a BLU(m) decomposition by only constructing a so-called "partial" LU decomposition; see, e.g., [14].

It is obvious that a matrix with a singular leading principal sub-matrix may still admit a BLU(m) decomposition, but this decomposition may not be unique.

By using the uniqueness of the standard LU decomposition, we can easily prove the existence and uniqueness of the BLU(m) decomposition defined in the above.

Theorem 2.2. *Let $A \in \mathbf{C}^{n\times n}$ be a matrix with its first $n-m$ $(1 \leq m \leq n)$ leading principal minors being nonzero. Then it has a unique BLU(m) decomposition.*

By making use of Lemma 2.2, analogously to the proof of Theorem 2.1 we can demonstrate the existence of a locally smooth BLU(m) decomposition for the matrix-valued function $A(\lambda)$.

Theorem 2.3. *Given a positive integer m $(1 \leq m < n)$. Let $A(\lambda) \in \mathbf{C}^{n\times n}$ be a differentiable matrix-valued function on the domain $\mathbb{D} \subseteq \mathbf{C}$ such that $\operatorname{rank}(A(\lambda_0)) \geq n - m$ holds at a given point $\lambda_0 \in \mathbb{D}$. Let there exist two $n \times n$ permutation matrices P_l and P_r such that the first $n-m$ leading principal minors of the matrix $P_l A(\lambda_0) P_r$ are nonzero, and $P_l A(\lambda_0) P_r$ has a BLU(m) decomposition*

$$P_l A(\lambda_0) P_r = L_0 U_0,$$

with $L_0 \in \mathcal{L}_{n,m}$ and

$$U_0 = \begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix}, \quad U_{11} \in \mathcal{U}_{n-m}.$$

Then there exists a neighborhood of λ_0, say, $\mathbb{N}(\lambda_0) \subseteq \mathbb{D}$, such that for all $\lambda \in \mathbb{N}(\lambda_0)$ the matrix-valued function $P_l A(\lambda) P_r$ has a BLU(m) decomposition

$$P_l A(\lambda) P_r = L(\lambda) U(\lambda), \quad \textit{for} \quad \lambda \in \mathbb{N}(\lambda_0),$$

where $L(\lambda) \in \mathcal{L}_{n,m}$ *and*

$$U(\lambda) = \begin{pmatrix} U_{11}(\lambda) & U_{12}(\lambda) \\ 0 & U_{22}(\lambda) \end{pmatrix}, \quad \textit{with} \quad U_{11}(\lambda) \in \mathcal{U}_{n-m}.$$

Moreover, this BLU(m) decomposition has the following properties:

(*a*) $L(\lambda_0) = L_0$ *and* $U(\lambda_0) = U_0$*;*

(*b*) *both* $L(\lambda)$ *and* $U(\lambda)$ *are differentiable at the point* λ_0*; and*

(*c*) *if the following condition is satisfied:*

$$L_0^{-1} P_l A'(\lambda_0) P_r = \begin{pmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{pmatrix}, \quad \textit{with} \quad \widetilde{A}_{11} \in \mathbf{C}^{(n-m)\times(n-m)},$$

then it holds that

$$U_{22}(\lambda) = U_{22} + (\widetilde{A}_{22} - \widetilde{A}_{21} U_{11}^{-1} U_{12})(\lambda - \lambda_0) + o(|\lambda - \lambda_0|). \tag{2.6}$$

Similar to Remark 2.1 about the smooth LU decomposition, we also have the following observation about the smooth BLU(m) decomposition.

Remark 2.2. If $A(\lambda) \in \mathbf{C}^{n\times n}$ is further assumed to be twice continuously differentiable on the domain $\mathbb{D} \subseteq \mathbf{C}$, then the term $o(|\lambda - \lambda_0|)$ involved in (2.6) may be replaced by a higher-order quantity $\mathcal{O}(|\lambda - \lambda_0|^2)$.

3. The Gauss-Newton-Type Methods for NEPs

In this section, we construct numerical methods of the Gauss-Newton-type for finding a multiple eigenvalue $\lambda_\star$ of the NEP (1.1), for which the matrix $A(\lambda_\star)$ has rank $n-m$ ($1 \le m < n$), through repeatedly using the smooth BLU(m) decomposition with complete pivoting.

To this end, we will further assume in the sequel that the matrix-valued function $A(\lambda) \in \mathbf{C}^{n\times n}$, defined on the domain $\mathbb{D} \subseteq \mathbf{C}$, is twice continuously differentiable in a neighborhood $\mathbb{N}(\lambda_\star)$ of $\lambda_\star$.

For an eigenvalue $\lambda_\star$ of $A(\lambda)$, if there exists a positive integer μ and a function $q(\lambda)$ differentiable at $\lambda_\star$ such that

$$\det(A(\lambda)) = (\lambda - \lambda_\star)^\mu q(\lambda) \tag{3.1}$$

holds, then the largest positive integer μ satisfying (3.1) is called the algebraic multiplicity of $\lambda_\star$. We remark that the eigenvalue $\lambda_\star$ is said to be simple if $\mu = 1$, and multiple if $\mu \ge 2$. For example, the rational polynomial $(\lambda - \lambda_\star)^{5/2}$ has a zero of algebraic multiplicity 2 at $\lambda_\star$; our definition, however, does not apply to the rational polynomial $(\lambda - \lambda_\star)^{1/2}$ as it is not differentiable at $\lambda_\star$, though it also vanishes at $\lambda = \lambda_\star$. The geometric multiplicity of the eigenvalue $\lambda_\star$ is defined to be the dimension of the null space $\ker(A(\lambda_\star))$ of the matrix $A(\lambda_\star)$.

Because $\operatorname{rank}(A(\lambda_\star)) = n - m$, with $1 \le m < n$, from [16] we know that there exist two $n \times n$ permutation matrices $P_l^\star$ and $P_r^\star$ such that

$$P_l^\star A(\lambda_\star) P_r^\star = L(\lambda_\star) U(\lambda_\star), \tag{3.2}$$

where $L(\lambda_\star) \in \mathcal{L}_{n,m}$ and

$$U(\lambda_\star) = \begin{pmatrix} U_{11}(\lambda_\star) & U_{12}(\lambda_\star) \\ 0 & 0 \end{pmatrix}, \quad \text{with} \quad U_{11}(\lambda_\star) \in \mathcal{U}_{n-m} \quad \text{nonsingular.} \tag{3.3}$$

Let $\lambda^{(c)} \in \mathbb{N}(\lambda_\star)$ be the current approximation to $\lambda_\star$, and

$$P_l^{(c)} A(\lambda^{(c)}) P_r^{(c)} = L(\lambda^{(c)}) U(\lambda^{(c)})$$

be the BLU(m) decomposition of the matrix $A(\lambda^{(c)})$, where $P_l^{(c)}$ and $P_r^{(c)}$ are two $n \times n$ permutation matrices, $L(\lambda^{(c)}) \in \mathcal{L}_{n,m}$, and

$$U(\lambda^{(c)}) = \begin{pmatrix} U_{11}(\lambda^{(c)}) & U_{12}(\lambda^{(c)}) \\ 0 & U_{22}(\lambda^{(c)}) \end{pmatrix}, \quad \text{with} \quad U_{11}(\lambda^{(c)}) \in \mathcal{U}_{n-m}.$$

Then by making use of Theorem 2.3 and Remark 2.2 at the point $\lambda_0 := \lambda^{(c)}$, we know that the matrix-valued function $P_l^{(c)} A(\lambda) P_r^{(c)}$ has the smooth BLU(m) decomposition

$$P_l^{(c)} A(\lambda) P_r^{(c)} = L(\lambda) U(\lambda), \quad \text{for} \quad \lambda \in \mathbb{N}(\lambda_\star),$$

provided the neighborhood $\mathbb{N}(\lambda_\star)$ is small enough, where $L(\lambda) \in \mathcal{L}_{n,m}$ and

$$U(\lambda) = \begin{pmatrix} U_{11}(\lambda) & U_{12}(\lambda) \\ 0 & U_{22}(\lambda) \end{pmatrix}, \quad \text{with} \quad U_{11}(\lambda) \in \mathcal{U}_{n-m}.$$

Moreover, if

$$L(\lambda^{(c)})^{-1} P_l^{(c)} A'(\lambda^{(c)}) P_r^{(c)} = \begin{pmatrix} \widetilde{A}_{11}(\lambda^{(c)}) & \widetilde{A}_{12}(\lambda^{(c)}) \\ \widetilde{A}_{21}(\lambda^{(c)}) & \widetilde{A}_{22}(\lambda^{(c)}) \end{pmatrix},$$

with $\widetilde{A}_{11}(\lambda^{(c)}) \in \mathbf{C}^{(n-m)\times(n-m)}$, then it holds that

$$U_{22}(\lambda) = U_{22}(\lambda^{(c)}) + U_{22}'(\lambda^{(c)})(\lambda - \lambda^{(c)}) + \mathcal{O}(|\lambda - \lambda^{(c)}|^2), \tag{3.4}$$

where

$$U_{22}'(\lambda^{(c)}) = \widetilde{A}_{22}(\lambda^{(c)}) - \widetilde{A}_{21}(\lambda^{(c)}) U_{11}(\lambda^{(c)})^{-1} U_{12}(\lambda^{(c)}). \tag{3.5}$$

Recalling from Theorem 2.3 that $U(\lambda)$ is differentiable at the point $\lambda_\star$, by continuity we may anticipate that $U(\lambda^{(c)}) \approx U(\lambda_\star)$. It follows immediately from the BLU(m) decomposition (3.2)-(3.3) of the matrix $A(\lambda_\star)$ that $U_{11}(\lambda^{(c)})$ is nonsingular and $U_{22}(\lambda^{(c)}) \approx 0$. Therefore, a feasible strategy for computing a multiple eigenvalue of the matrix-valued function $A(\lambda)$ may be to minimize the Frobenius norm of the linear part of $U_{22}(\lambda)$. This naturally results in the following Gauss-Newton-type iteration [12] for solving the NEP (1.1):

$$\lambda^{(+)} = \lambda^{(c)} - \|U_{22}'(\lambda^{(c)})\|_F^{-2} [\mathrm{col}(U_{22}'(\lambda^{(c)}))]^* \mathrm{col}(U_{22}(\lambda^{(c)})).$$

Here, we have applied the formulas (3.4)-(3.5), and used $(\cdot)^*$ to denote the conjugate transpose of a vector.

Method 3.1 (The Gauss-Newton-type Method). Assume that $\lambda_\star$ is an eigenvalue of the NEP (1.1) with geometric multiplicity m, and let $\lambda^{(0)}$ be an initial guess of $\lambda_\star$. Then, for $k = 0, 1, 2, \ldots$ until the sequence $\{\lambda^{(k)}\}$ is convergent,

Step 1. evaluate the matrix $A(\lambda)$ and compute its derivative $A'(\lambda)$ at $\lambda = \lambda^{(k)}$;
Step 2. compute a BLU(m) decomposition of the matrix $A(\lambda^{(k)})$ with complete pivoting:

$$P_l^{(k)} A(\lambda^{(k)}) P_r^{(k)} = L(\lambda^{(k)}) U(\lambda^{(k)}),$$

where $P_l^{(k)}$ and $P_r^{(k)}$ are two $n \times n$ permutation matrices, $L(\lambda^{(k)}) \in \mathcal{L}_{n,m}$, and

$$U(\lambda^{(k)}) = \begin{pmatrix} U_{11}(\lambda^{(k)}) & U_{12}(\lambda^{(k)}) \\ 0 & U_{22}(\lambda^{(k)}) \end{pmatrix}, \quad \text{with} \quad U_{11}(\lambda^{(k)}) \in \mathcal{U}_{n-m};$$

Step 3. compute the derivative of $U_{22}(\lambda)$ at $\lambda = \lambda^{(k)}$ by

$$U_{22}'(\lambda^{(k)}) = K^{(m)^T}(\lambda^{(k)})^{-1} P_l^{(k)} A'(\lambda^{(k)}) P_r^{(k)} (K^{(m)} - I^{(n-m)} U_{11}(\lambda^{(k)})^{-1} U_{12}(\lambda^{(k)}));$$

Step 4. compute

$$\lambda^{(k+1)} = \lambda^{(k)} - \|U_{22}'(\lambda^{(k)})\|_F^{-2} [\mathrm{col}(U_{22}'(\lambda^{(k)}))]^* \mathrm{col}(U_{22}(\lambda^{(k)}));$$

Step 5. if the prescribed accuracy is achieved, then stop. Otherwise, go to Step 1.

Evidently, a practically important problem in the actual implementation of Method 3.1 is how to determine the geometric multiplicity m of the target eigenvalue $\lambda_\star$. Theoretically, this problem is equivalent to the determination of $\mathrm{rank}(A(\lambda_\star))$. The rank-revealing LU decomposition based on the Gaussian elimination [30] can be used to effectively compute a numerical rank of the matrix $A(\lambda_\star)$; see also [7, 17, 18].

Alternatively, we can give a simpler strategy for numerically approximating $\mathrm{rank}(A(\lambda_\star))$. Assume that $\lambda^{(k)}$ is a good approximation to $\lambda_\star$ and

$$P_l^{(k)} A(\lambda^{(k)}) P_r^{(k)} = L(\lambda^{(k)}) U(\lambda^{(k)})$$

is a BLU(m) decomposition with complete pivoting; see Theorem 2.3. Then the entries of $U(\lambda^{(k)}) := (u_{ij}(\lambda^{(k)}))$ possess the property

$$\min_{1 \le i \le n-m} |u_{ii}(\lambda^{(k)})| \gg \max_{n-m+1 \le i,j \le n} |u_{ij}(\lambda^{(k)})|.$$

This property may be useful for approximately determining a numerical rank of the matrix $A(\lambda^{(k)})$, although we have realized that a small $U_{22}(\lambda^{(k)})$ (small in the sense of certain norm) does not generally guarantee that the numerical rank of $A(\lambda^{(k)})$ is $n - m$.

Based on this observation, we have the following strategy for estimating the geometric multiplicity m of $\lambda_\star$:

Choose a threshold $\varepsilon > 0$, and find the largest integer m such that

$$\max_{n-m+1 \le i,j \le n} |u_{ij}(\lambda^{(k)})| < \varepsilon \max_{1 \le i \le n-m} |u_{ii}(\lambda^{(k)})| \le \min_{1 \le i \le n-m} |u_{ii}(\lambda^{(k)})|.$$

By incorporating this strategy with Method 3.1, we can obtain the following modified Gauss-Newton-type method for solving the NEP (1.1).

Method 3.2 (The Modified Gauss-Newton-type Method). Assume that $\lambda_\star$ is an eigenvalue of the NEP (1.1) with geometric multiplicity m, and let $\lambda^{(0)}$ be an initial guess of $\lambda_\star$. Then, for $k = 0, 1, 2, \ldots$ until the sequence $\{\lambda^{(k)}\}$ is convergent,

Step 0. set a threshold $\varepsilon > 0$;
Step 1. evaluate the matrix $A(\lambda)$ and compute its derivative $A'(\lambda)$ at $\lambda = \lambda^{(k)}$;
Step 2. compute a BLU(m) decomposition of the matrix $A(\lambda^{(k)})$ with complete pivoting:

$$P_l^{(k)} A(\lambda^{(k)}) P_r^{(k)} = L(\lambda^{(k)}) U(\lambda^{(k)}),$$

where $P_l^{(k)}$ and $P_r^{(k)}$ are two $n \times n$ permutation matrices, $L(\lambda^{(k)}) \in \mathcal{L}_{n,m}$, and

$$U(\lambda^{(k)}) := (u_{ij}(\lambda^{(k)})) = \begin{pmatrix} U_{11}(\lambda^{(k)}) & U_{12}(\lambda^{(k)}) \\ 0 & U_{22}(\lambda^{(k)}) \end{pmatrix}, \quad \text{with} \quad U_{11}(\lambda^{(k)}) \in \mathcal{U}_{n-m};$$

Step 3. find the largest integer ℓ $(1 \le \ell < n)$ such that

$$\max_{n-\ell+1 \le i,j \le n} |u_{ij}(\lambda^{(k)})| < \varepsilon \max_{1 \le i \le n-\ell} |u_{ii}(\lambda^{(k)})| \le \min_{1 \le i \le n-\ell} |u_{ii}(\lambda^{(k)})|$$

and set $m := \ell$. If there does not exist such an ℓ, then set $m := 1$;
Step 4. partition $U(\lambda^{(k)})$ into

$$U(\lambda^{(k)}) = \begin{pmatrix} U_{11}(\lambda^{(k)}) & U_{12}(\lambda^{(k)}) \\ 0 & U_{22}(\lambda^{(k)}) \end{pmatrix}, \quad \text{with} \quad U_{11}(\lambda^{(k)}) \in \mathcal{U}_{n-m};$$

Step 5. compute the derivative of $U_{22}(\lambda)$ at $\lambda = \lambda^{(k)}$ by

$$U_{22}'(\lambda^{(k)}) = K^{(m)^T} L(\lambda^{(k)})^{-1} P_l^{(k)} A'(\lambda^{(k)}) P_r^{(k)} \Big(K^{(m)} - I^{(n-m)} U_{11}(\lambda^{(k)})^{-1} U_{12}(\lambda^{(k)}) \Big);$$

Step 6. compute

$$\lambda^{(k+1)} = \lambda^{(k)} - \|U_{22}'(\lambda^{(k)})\|_F^{-2} [\mathrm{col}(U_{22}'(\lambda^{(k)}))]^* \mathrm{col}(U_{22}(\lambda^{(k)}));$$

Step 7. if the prescribed accuracy is achieved, then stop. Otherwise, go to Step 1.

In actual implementations of Method 3.2, it is expected that the integer m increases as the iteration index k is growing. Therefore, we can set $m := 1$ as the initial choice and use it in the first few iterates. Then, m may be modified after each number of several iteration steps, and the threshold ε may be periodically reassessed with proceeding of the computation, too.

Remark 3.1. When $m \equiv 1$, both Methods 3.1 and 3.2 reduce to the approach given in [20,42]. In this case, Steps 3-4 in Method 3.2 is correspondingly skipped.

Remark 3.2. The standard LU and QR decompositions require about $\frac{1}{3}n^3$ and $\frac{2}{3}n^3$ multiplications, respectively; see [14]. Therefore, Methods 3.1 and 3.2 are generally cheaper than those given in [28].

4. Convergence Theory

We first establish the perturbation theories about the standard LU and the BLU(m) decompositions for a constant matrix, which are necessary for demonstrating the local convergence of the Gauss-Newton-type methods in Section 3.

Lemma 4.1. *Let $B \in \mathbf{C}^{n\times n}$ be a constant matrix, with its first $n-1$ leading principal minors being nonzero, and $\Delta B \in \mathbf{C}^{n\times n}$ be a given perturbation increment to the matrix B. For a given sufficiently small $\varepsilon > 0$, assume that $\|\Delta B\| < \varepsilon$.*

(a) If $B = L_B U_B$ is an LU decomposition of the matrix B, then the matrix $\overline{B} := B + \Delta B$ possesses an LU decomposition $\overline{B} = L_{\overline{B}} U_{\overline{B}}$ such that

$$\|L_{\overline{B}} - L_B\| \le \kappa_L \varepsilon \quad \text{and} \quad \|U_{\overline{B}} - U_B\| \le \kappa_U \varepsilon,$$

where κ_L and κ_U are constants independent of $\overline{B}$.

(b) If $B = L_B U_B$ is a BLU(m) decomposition of the matrix B, with $1 \le m < n$ being a prescribed integer, then the matrix $\overline{B} := B + \Delta B$ possesses a BLU(m) decomposition $\overline{B} = L_{\overline{B}} U_{\overline{B}}$ such that

$$\|L_{\overline{B}} - L_B\| \le \kappa_L \varepsilon \quad \text{and} \quad \|U_{\overline{B}} - U_B\| \le \kappa_U \varepsilon,$$

where κ_L and κ_U are constants independent of $\overline{B}$.

Proof. We only prove (a), as (b) can be demonstrated in an analogous fashion. Let $\overline{L}$ be any unit lower-triangular matrix, and define

$$L_{\overline{B}} = L_B \overline{L} \quad \text{and} \quad U_{\overline{B}} = \overline{L}^{-1}(U_B + L_B^{-1}\Delta B).$$

Then it holds that

$$L_{\overline{B}} U_{\overline{B}} = L_B U_B + \Delta B = \overline{B}.$$

Note that

$$L_B - L_{\overline{B}} = L_B(I - \overline{L}) \quad \text{and} \quad U_B - U_{\overline{B}} = (I - \overline{L}^{-1})U_B - \overline{L}^{-1} L_B^{-1}\Delta B.$$

We now partition the matrices U_B and $L_B^{-1}\Delta B$ into blocks as

$$U_B = \begin{pmatrix} U_{11}^{(B)} & U_{12}^{(B)} \\ 0 & u_{nn}^{(B)} \end{pmatrix} \quad \text{and} \quad L_B^{-1}\Delta B = \begin{pmatrix} \Delta E_{11} & \Delta E_{12} \\ \Delta E_{21} & \Delta e_{nn} \end{pmatrix},$$

where $U_{11}^{(B)}, \Delta E_{11} \in \mathbf{C}^{(n-1)\times(n-1)}$. It then follows that all leading principal minors of $U_{11}^{(B)} + \Delta E_{11}$ are nonzero, provided ε is small enough. Therefore, we have the LU decomposition $U_{11}^{(B)} + \Delta E_{11} = \widehat{L}_1 \widehat{U}_1$, with $\widehat{L}_1 \in \mathcal{L}_{n-1}$ and $\widehat{U}_1 \in \mathcal{U}_{n-1}$. According to [3, Theorem 3.1], we know that there exist positive constants ν_1 and ν_2, independent of ΔE_{11}, such that there hold

$$\|\widehat{L}_1 - I\| \le \nu_1 \varepsilon \quad \text{and} \quad \|\widehat{U}_1 - U_{11}^{(B)}\| \le \nu_2 \varepsilon.$$

Hence, we can further obtain the estimates

$$\|\widehat{L}_1^{-1} - I\| \le \kappa_1\varepsilon \quad \text{and} \quad \|\Delta E_{21}(U_{11}^{(B)} + \Delta E_{11})^{-1}\| \le \kappa_2\varepsilon,$$

where κ_1 and κ_2 are constants independent of ΔB.

By constructing the unit lower-triangular matrix

$$\overline{L} = \begin{pmatrix} \widehat{L}_1 & 0 \\ \Delta E_{21}(U_{11}^{(B)} + \Delta E_{11})^{-1}\widehat{L}_1 & 1 \end{pmatrix},$$

we can easily verify that $L_{\overline{B}} \in \mathcal{L}_n$ and $U_{\overline{B}} = \overline{L}^{-1}(U_B + L_B^{-1}\Delta B) \in \mathcal{U}_n$ hold true, and $L_{\overline{B}}$ and $U_{\overline{B}}$ possess the required properties. □

Let $A(\lambda) \in \mathbf{C}^{n\times n}$ be a twice continuously differentiable matrix-valued function on the domain $\mathbb{D} \subseteq \mathbf{C}$, and $\lambda_\star \in \mathbb{D}$ be an eigenvalue of $A(\lambda)$ with geometric multiplicity m $(1 \le m < n)$. Then we easily know from Theorem 2.2 that the BLU(m) decomposition of the matrix-valued function $P_l A(\lambda) P_r$, defined in Theorem 2.3 by setting $\lambda_0 = \lambda_\star$, is uniquely determined for all $\lambda \in \mathbb{N}(\lambda_\star)$, a neighborhood of $\lambda_\star$, provided the permutation matrices P_l and P_r have been finally computed by a complete pivoting. In general, the permutation matrices are not uniquely determined by a complete pivoting during the BLU(m)-decomposition process. This may cause some difficulty in the convergence analyses of Methods 3.1 and 3.2. However, the following lemma about the property of the function

$$f(\lambda, m, P_l, P_r) := \|U_{22}'(\lambda)\|_F^2$$

affords a tool for overcoming such a difficulty.

Lemma 4.2. *Let $A(\lambda) \in \mathbf{C}^{n\times n}$ be a twice continuously differentiable matrix-valued function on the open connected domain $\mathbb{D} \subseteq \mathbf{C}$, and $\lambda_\star \in \mathbb{D}$ be an eigenvalue of $A(\lambda)$ with geometric multiplicity m $(1 \le m < n)$.*

(*a*) *If the first $n-1$ leading principal minors of $A(\lambda)$ are nonzero, then $f(\lambda, m, P_l, P_r)$ is continuous with respect to $\lambda \in \mathbb{D}$.*

(*b*) *If $(P_l^{(1)}, P_r^{(1)})$ and $(P_l^{(2)}, P_r^{(2)})$ are two pairs of permutation matrices with respect to two BLU(m) decompositions of the matrix $A(\lambda_\star)$ with complete pivoting, respectively, then it holds that*

$$f(\lambda_\star, m, P_l^{(1)}, P_r^{(1)}) > 0 \Longleftrightarrow f(\lambda_\star, m, P_l^{(2)}, P_r^{(2)}) > 0. \tag{4.1}$$

Proof. (a) is a straightforward conclusion of Theorem 2.3 and Lemma 4.1, as well as (2.6). So, now we turn to prove (b).

Assume on the contrary that (4.1) does not hold true. Without loss of generality, we suppose that

$$f(\lambda_\star, m, P_l^{(1)}, P_r^{(1)}) > 0, \quad \text{and} \quad f(\lambda_\star, m, P_l^{(2)}, P_r^{(2)}) = 0.$$

According to (3.2) and (3.3), we have

$$P_l^{(\ell)} A(\lambda_\star) P_r^{(\ell)} = L^{(\ell)}(\lambda_\star) U^{(\ell)}(\lambda_\star), \quad \ell = 1, 2,$$

where $L^{(\ell)}(\lambda_\star) \in \mathcal{L}_{n,m}$ and

$$U^{(\ell)}(\lambda_\star) = \begin{pmatrix} U_{11}^{(\ell)}(\lambda_\star) & U_{12}^{(\ell)}(\lambda_\star) \\ 0 & 0 \end{pmatrix}, \quad \text{with} \quad U_{11}^{(\ell)}(\lambda_\star) \in \mathcal{U}_{n-m} \quad \text{nonsingular.}$$

It follows from Theorem 2.3 that there exists a neighborhood $\mathbb{N}(\lambda_\star)$ of $\lambda_\star$ such that for any $\lambda \in \mathbb{N}(\lambda_\star)$ the matrix-valued functions $P_l^{(\ell)} A(\lambda) P_r^{(\ell)}$, $\ell = 1, 2$, have the BLU(m) decompositions

$$P_l^{(\ell)} A(\lambda) P_r^{(\ell)} = L^{(\ell)}(\lambda) U^{(\ell)}(\lambda), \quad \ell = 1, 2,$$

with $L^{(\ell)}(\lambda) \in \mathcal{L}_{n,m}$ and

$$U^{(\ell)}(\lambda) = \begin{pmatrix} U_{11}^{(\ell)}(\lambda) & U_{12}^{(\ell)}(\lambda) \\ 0 & U_{22}^{(\ell)}(\lambda) \end{pmatrix}, \quad U_{11}^{(\ell)}(\lambda) \in \mathcal{U}_{n-m},$$

which possess the properties:

- (P_1) for $\ell = 1, 2$, $U_{11}^{(\ell)}(\lambda)$ are invertible;
- (P_2) $U_{22}^{(1)}(\lambda) = \frac{d\, U_{22}^{(1)}(\lambda_\star)}{d\, \lambda}(\lambda - \lambda_\star) + \mathcal{O}(|\lambda - \lambda_\star|^2)$, with $\frac{d\, U_{22}^{(1)}(\lambda_\star)}{d\, \lambda} \neq 0$; and
- (P_3) $U_{22}^{(2)}(\lambda) = \mathcal{O}(|\lambda - \lambda_\star|^2)$.

Now, denote by

$$V^{(\ell)}(\lambda) = \begin{pmatrix} U_{11}^{(\ell)}(\lambda) & U_{12}^{(\ell)}(\lambda) \\ 0 & I \end{pmatrix} \quad \text{and} \quad D^{(\ell)}(\lambda) = \begin{pmatrix} I & 0 \\ 0 & U_{22}^{(\ell)}(\lambda) \end{pmatrix}, \qquad \ell = 1, 2.$$

Then it holds that

$$A(\lambda) = L_l^{(\ell)}(\lambda) D^{(\ell)}(\lambda) M^{(\ell)}(\lambda), \qquad \ell = 1, 2,$$

where

$$L_l^{(\ell)}(\lambda) = {P_l^{(\ell)}}^T L^{(\ell)}(\lambda) \quad \text{and} \quad M^{(\ell)}(\lambda) = V^{(\ell)}(\lambda) {P_r^{(\ell)}}^T,$$

which are invertible when $\mathbb{N}(\lambda_\star)$ is sufficiently small. Hence,

$$D^{(1)}(\lambda) N^{(1)}(\lambda) = N^{(2)}(\lambda) D^{(2)}(\lambda),$$

where

$$N^{(1)}(\lambda) = M^{(1)}(\lambda) M^{(2)}(\lambda)^{-1} \quad \text{and} \quad N^{(2)}(\lambda) = L_l^{(1)}(\lambda)^{-1} L_l^{(2)}(\lambda).$$

Based on the above investigation, by following a similar way to that in [28] we can deduce the inequalities

$$0 \neq \frac{d\, U_{22}^{(1)}(\lambda_\star)}{d\, \lambda} = \lim_{\lambda \to \lambda_\star} \frac{U_{22}^{(1)}(\lambda)}{\lambda - \lambda_\star} = 0,$$

which is obviously a contradiction. Thus, the statement (4.1) holds true. □

Just before stating the convergence theorem, we introduce the following useful index, which is similar to Definition 4.1 in [28].

To this end, we again let $A(\lambda) \in \mathbf{C}^{n \times n}$ be a continuously differentiable matrix-valued function on the domain $\mathbb{D} \subseteq \mathbf{C}$, and $\lambda_\star \in \mathbb{D}$ an eigenvalue of $A(\lambda)$ with geometric multiplicity m ($1 \leq m < n$). Let P_l and P_r be permutation matrices from a BLU(m) decomposition of the matrix $A(\lambda_\star)$ with complete pivoting. If $m = n$, then we define

$$\mathrm{Ind}(\lambda_\star) = \begin{cases} 1, & \text{for} \quad \|A'(\lambda_\star)\| > 0, \\ 0, & \text{otherwise;} \end{cases}$$

and if $1 \leq m < n$, then we define

$$\mathrm{Ind}(\lambda_\star) = \begin{cases} 1, & \text{for} \quad f(\lambda_\star, m, P_l, P_r) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

At present, Theorem 2.3 and Lemma 4.2 have readily led to the locally quadratic convergence for Method 3.1 and, hence, for Method 3.2, too.

Theorem 4.1. (*The Local Convergence Theorem*). *Let $A(\lambda) \in \mathbf{C}^{n\times n}$ be a twice continuously differentiable matrix-valued function on the open connected domain $\mathbb{D} \subseteq \mathbf{C}$, and $\lambda_\star \in \mathbb{D}$ be an eigenvalue of $A(\lambda)$ with geometric multiplicity m $(1 \leq m < n)$. Let in Method* 3.1 *the initial guess $\lambda^{(0)}$ satisfy $\lambda^{(0)} \in \mathbb{N}(\lambda_\star)$, the permutation matrices $P_l^{(k)}$ and $P_r^{(k)}$ be independent of k, and the index* $\mathrm{Ind}(\lambda_\star)$ *be constantly equal to* 1 *for $\lambda^{(k)} \in \mathbb{N}(\lambda_\star)$, where $\mathbb{N}(\lambda_\star) \subseteq \mathbb{D}$ is a neighborhood of $\lambda_\star$. Then the iteration sequence $\{\lambda^{(k)}\}$, generated by Method* 3.1*, is locally quadratically convergent to $\lambda_\star$, with $\{\lambda^{(k)}\} \subset \mathbb{N}(\lambda_\star)$.*

Proof. The proof is similar to that of Theorem 4.1 in [28]. Hence, it is omitted. □

Remark 4.1. By appropriately choosing the threshold ε such that $\mathrm{Ind}(\lambda_\star) \equiv 1$, we can analogously obtain the locally quadratic convergence of Method 3.2, too.

5. Numerical Results

In this section, we use numerical experiments to show the feasibility and effectiveness of Methods 3.1 and 3.2 for computing the eigenvalues of the NEP (1.1), and also their advantages over the methods given in [28], in the sense of iteration step, the absolute solution error $|\lambda_\star - \lambda^{(k)}|$ and the absolute residual error $\|U_{22}(\lambda^{(k)})\|_F$ (for Methods 3.1 and 3.2) or $\|R_{22}(\lambda^{(k)})\|_F$ (for Methods 3.1 and 3.2 in [28]).

In our implementations, all iterations are started from initial guesses close to the exact solution $\lambda_\star$ so that a few iterations are required for the iteration sequences to achieve the convergence, and terminated once the current iterate satisfies $\|U_{22}(\lambda^{(k)})\|_F \leq 10^{-8}$ (for Methods 3.1 and 3.2) or $\|R_{22}(\lambda^{(k)})\|_F \leq 10^{-8}$ (for Methods 3.1 and 3.2 in [28]). In addition, all codes are run in MATLAB (version 6.5) with machine precision 10^{-16} on a Pentium IV personal computer.

With the numerical results, we are interested in practically verifying the locally quadratic convergence for the case that $\lambda_\star$ is either a distinct or a multiple eigenvalue of the matrix-valued function $A(\lambda)$. Besides, in the tables, we use $k(m)$ to denote the value of m determined at the k-th iteration step.

Example 5.1. ([25]) *Consider the quadratic eigenvalue problem $A(\lambda) = \lambda^2 G + \lambda B + C$, where*

$$G = \begin{pmatrix} -4 & 4 & 0 & 8 \\ -8 & 8 & 0 & 16 \\ 4 & -4 & 3 & -8 \\ -12 & 12 & -3 & 25 \end{pmatrix}, \qquad B = \begin{pmatrix} 12 & -12 & 0 & -24 \\ 24 & -26 & -4 & -50 \\ -12 & 14 & -5 & 26 \\ 36 & -40 & 1 & -78 \end{pmatrix},$$

Table 5.1: Numerical results for Example 5.1 using Method 3.1

$\lambda_\star$	$\lambda^{(0)}$	k	Method 3.1		Method 3.1 in [28]	
			$\lvert\lambda_\star-\lambda^{(k)}\rvert$	$\lVert U_{22}(\lambda^{(k)})\rVert_F$	$\lvert\lambda_\star-\lambda^{(k)}\rvert$	$\lVert R_{22}(\lambda^{(k)})\rVert_F$
1	1.25	0	2.50e-01	1.72e-01	2.50e-01	1.50e-01
		1	2.44e-02	1.71e-02	9.20e-03	6.04e-03
		2	1.28e-04	9.04e-05	3.20e-05	2.09e-05
		3	4.43e-09	3.13e-09	3.84e-10	2.51e-10
$\frac{3+\sqrt{7}\imath}{2}$	$1.46+1.3\imath$	0	4.61e-02	6.42e-01	4.61e-02	5.54e-01
		1	1.62e-02	2.90e-01	4.66e-02	6.00e-01
		2	2.14e-03	4.15e-02	3.59e-02	4.84e-01
		3	3.65e-05	7.12e-04	2.26e-02	3.60e-01
		4	1.08e-08	2.10e-07	4.50e-03	7.76e-02
		5	1.11e-15	2.49e-14	5.34e-05	9.27e-04
		6			1.75e-09	3.05e-08

and

$$C=\begin{pmatrix} -16 & 16 & 0 & 32\\ -32 & 34 & 4 & 66\\ 16 & -18 & 8 & -34\\ -48 & 52 & -4 & 101 \end{pmatrix}.$$

We can easily check that the matrix-valued function $A(\lambda)$ has a real and multiple eigenvalue $\lambda_\star = 1$ of algebraic multiplicity 3 and geometric multiplicity $m = 2$, and two complex and multiple eigenvalues $\lambda_\star = \frac{1}{2}(3 \pm \sqrt{7}\imath)$ of algebraic multiplicity 2 and geometric multiplicity $m = 2$, where $\imath$ is used to represent the imaginary unit.

In Tables 5.1 and 5.2, we list the numerical results computed by Methods 3.1 and 3.2, respectively, for $m = 2$. In the implementations of Method 3.2 we take the threshold $\varepsilon = 10^{-2}$ and set $m := 1$ as the initial value. Then m is adaptively determined by the procedure. From Tables 5.1 and 5.2 we easily see that both Methods 3.1 and 3.2 show locally quadratic convergence rates. Moreover, Table 5.2 shows that the geometric multiplicity of the eigenvalues adaptively determined by Method 3.2 can rapidly achieve the exact ones.

Example 5.2. ([25]) *Consider the quadratic eigenvalue problem* $A(\lambda) = \lambda^2 G + \lambda B + C$, *where* $G = I$ *is the* 4×4 *identity matrix, and*

$$B=\begin{pmatrix} 3\alpha & -(1+\alpha^2+2\beta^2) & \alpha(1+2\beta^2) & -\beta^2(\alpha^2+\beta^2)\\ 2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 2 & 0 \end{pmatrix}$$

and

$$C=\begin{pmatrix} -1+2\alpha^2 & \alpha(1-\alpha^2-2\beta^2) & 2\alpha^2\beta^2 & -\alpha\beta^2(\alpha^2+\beta^2)\\ 2\alpha & -(\alpha^2+2\beta^2) & 2\alpha\beta^2 & -\beta^2(\alpha^2+\beta^2)\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 \end{pmatrix},$$

Table 5.2: Numerical results for Example 5.1 using Method 3.2

$\lambda_\star$	$\lambda^{(0)}$	Method 3.2			Method 3.2 in [28]		
		$k(m)$	$\lvert\lambda_\star-\lambda^{(k)}\rvert$	$\lVert U_{22}(\lambda^{(k)})\rVert_F$	$k(m)$	$\lvert\lambda_\star-\lambda^{(k)}\rvert$	$\lVert R_{22}(\lambda^{(k)})\rVert_F$
1	1.25	0(1)	2.50e-01	1.75e-01	0(1)	2.50e-01	1.50e-01
		1(2)	1.68e-02	1.18e-02	1(2)	9.20e-03	6.04e-03
		2(2)	6.60e-05	4.66e-05	2(2)	3.20e-05	2.09e-05
		3(2)	1.18e-09	8.34e-10	3(2)	3.84e-10	2.51e-10
$\frac{3+\sqrt{7}\imath}{2}$	$1.46+1.3\imath$	0(1)	4.61e-02	1.66e-01	0(1)	4.61e-02	1.09e-01
		1(1)	5.34e-03	8.02e-02	1(1)	3.31e-03	5.74e-02
		2(2)	2.30e-04	4.48e-03	2(2)	1.82e-05	3.16e-04
		3(2)	4.26e-07	8.30e-06	3(2)	2.12e-10	3.69e-09
		4(2)	1.46e-13	2.85e-11			

Table 5.3: Numerical results for Example 5.2 using Method 3.1, $\alpha = 0$

$\lambda_\star$	$\lambda^{(0)}$	k	Method 3.1		Method 3.1 in [28]	
			$\lvert\lambda_\star-\lambda^{(k)}\rvert$	$\lVert U_{22}(\lambda^{(k)})\rVert_F$	$\lvert\lambda_\star-\lambda^{(k)}\rvert$	$\lVert R_{22}(\lambda^{(k)})\rVert_F$
0	-0.2	0	2.00e-01	4.71e-02	2.00e-01	3.33e-02
		1	1.14e-01	1.38e-02	1.14e-01	9.70e-03
		2	6.01e-02	3.66e-03	5.99e-02	2.57e-03
		3	3.05e-02	9.31e-04	3.04e-02	6.55e-04
		4	1.53e-02	2.34e-04	1.52e-02	1.65e-04
		5	7.65e-03	5.85e-05	7.63e-03	4.12e-05
		6	3.83e-03	1.46e-05	3.82e-03	1.03e-05
		7	1.91e-03	3.66e-06	1.91e-03	2.57e-06
		8	9.56e-04	9.15e-07	9.54e-04	6.44e-07
		9	4.78e-04	2.29e-07	4.77e-04	1.61e-07
		10	2.39e-04	5.72e-08	2.39e-04	4.02e-08
		11	1.20e-04	1.43e-08	1.19e-04	1.01e-08
		12	5.98e-05	3.57e-09	5.96e-05	2.51e-09
$\imath$	$1.2\imath$	0	2.00e-01	3.80e-01	2.00e-01	3.35e-01
		1	3.02e-02	5.20e-02	2.13e-02	3.33e-02
		2	4.43e-04	7.54e-04	1.41e-04	2.19e-04
		3	7.59e-08	9.29e-09	5.34e-09	8.31e-09

with $\beta = 1+\alpha$ and α a problem parameter.

It is clear that the corresponding NEP (1.1) becomes more ill-conditioned and difficult when $|\alpha|$ is approaching zero; see [1].

We can easily check that the matrix-valued function $A(\lambda)$ has eight eigenvalues as follows:

$$-\alpha, \quad -\alpha \pm (1+\alpha)\imath, \quad \pm(1+\alpha)\imath, \quad \pm\imath, \quad \text{and} \quad 0.$$

In particular, when $\alpha = 0$, it has three eigenvalues: $\lambda_\star = \pm\imath$ of algebraic multiplicity 3 and

Table 5.4: Numerical results for Example 5.2 using Method 3.1, $\alpha = -1$

$\lambda_\star$	$\lambda^{(0)}$	k	Method 3.1		Method 3.1 in [28]	
			$\lvert\lambda_\star - \lambda^{(k)}\rvert$	$\lVert U_{22}(\lambda^{(k)})\rVert_F$	$\lvert\lambda_\star - \lambda^{(k)}\rvert$	$\lVert R_{22}(\lambda^{(k)})\rVert_F$
0	-0.2	0	2.00e-01	2.98e-01	2.00e-01	2.55e-01
		1	6.30e-02	7.15e-02	4.88e-02	4.73e-02
		2	7.70e-03	7.82e-03	4.96e-03	4.35e-03
		3	1.19e-04	1.19e-04	5.87e-05	5.08e-05
		4	2.83e-08	8.83e-09	8.33e-09	7.21e-09
1	1.2	0	2.00e-01	2.97e-01	2.00e-01	2.74e-01
		1	3.03e-02	4.08e-02	3.30e-02	3.92e-02
		2	3.75e-04	5.01e-04	9.29e-04	1.07e-03
		3	4.13e-08	5.51e-08	7.19e-07	8.31e-07
		4	4.44e-16	5.92e-16	4.31e-13	4.98e-13
$\imath$	$1.2\imath$	0	2.00e-01	1.95e-01	2.00e-01	1.38e-001
		1	5.85e-02	4.21e-02	5.68e-02	2.89e-02
		2	6.83e-03	4.38e-03	6.46e-03	2.93e-03
		3	1.07e-04	6.78e-05	9.69e-05	4.34e-05
		4	2.68e-08	9.69e-09	2.23e-08	9.95e-09

geometric multiplicity $m = 2$, and $\lambda_\star = 0$ of algebraic multiplicity 2 and geometric multiplicity 1; and when $\alpha = -1$, it has four eigenvalues: $\lambda_\star = 1$ of algebraic multiplicity 3 and geometric multiplicity 2, $\lambda_\star = \pm\imath$ of both algebraic and geometric multiplicities 1, and $\lambda_\star = 0$ of algebraic multiplicity 3 and geometric multiplicity 2.

In the computations, we choose the initial guesses $\lambda^{(0)}$ as follows:

$$\lambda^{(0)} = \begin{cases} -0.2 \quad \text{or} \quad 1.2\imath, & \text{for} \quad \alpha = 0, \\ -0.2 \quad \text{or} \quad 1.2 \quad \text{or} \quad 1.2\imath, & \text{for} \quad \alpha = -1. \end{cases}$$

In Tables 5.3 and 5.4 we list the numerical results computed by Method 3.1 for the exact geometric multiplicity, and in Tables 5.5 and 5.6 those computed by Method 3.2 for the adaptively determined geometric multiplicity. In the implementations of Method 3.2 we again take the threshold $\varepsilon = 10^{-2}$ and set $m := 1$ as the initial value. Then m is adaptively determined by the procedure. From these tables we can see again that both Methods 3.1 and 3.2 show locally quadratic convergence rates, except for the case $\lambda_\star = 0$ when $\alpha = 0$, for which only linear convergence rate is observed. Moreover, Tables 5.5-5.6 show that the geometric multiplicity of the eigenvalues adaptively determined by Method 3.2 can rapidly achieve the exact ones.

In fact, when $\alpha = 0$, the eigenvalue $\lambda_\star = 0$ has algebraic multiplicity 2 and geometric multiplicity 1. Through straightforward calculations we have

$$\begin{aligned} &P(3,4)P(1,2)A(0)P(1,2)P(3,4) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -0.5 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

where $P(i, j)$ is the elementary matrix that exchanges the i-th and the j-th rows or columns of

Table 5.5: Numerical results for Example 5.2 using Method 3.2, $\alpha = 0$

$\lambda_\star$	$\lambda^{(0)}$	Method 3.2			Method 3.2 in [28]		
		$k(m)$	$\lvert\lambda_\star - \lambda^{(k)}\rvert$	$\lVert U_{22}(\lambda^{(k)})\rVert_F$	$k(m)$	$\lvert\lambda_\star - \lambda^{(k)}\rvert$	$\lVert R_{22}(\lambda^{(k)})\rVert_F$
0	-0.2	0(1)	2.00e-01	4.71e-02	0(1)	2.00e-01	3.33e-02
		1(1)	1.14e-01	1.38e-02	1(1)	1.14e-01	9.70e-03
		2(1)	6.01e-02	3.66e-03	2(1)	5.99e-02	2.57e-03
		3(1)	3.05e-02	9.31e-04	3(1)	3.04e-02	6.55e-04
		4(1)	1.53e-02	2.34e-04	4(1)	1.52e-02	1.65e-04
		5(1)	7.65e-03	5.85e-05	5(1)	7.63e-03	4.12e-05
		6(1)	3.83e-03	1.46e-05	6(1)	3.82e-03	1.03e-05
		7(1)	1.91e-03	3.66e-06	7(1)	1.91e-03	2.57e-06
		8(1)	9.56e-04	9.15e-07	8(1)	9.54e-04	6.44e-07
		9(1)	4.78e-04	2.29e-07	9(1)	4.77e-04	1.61e-07
		10(1)	2.39e-04	5.72e-08	10(1)	2.39e-04	4.02e-08
		11(1)	1.20e-04	1.43e-08	11(1)	1.19e-04	1.01e-08
		12(1)	5.98e-05	3.57e-09	12(1)	5.96e-05	2.51e-09
$\imath$	$1.2\imath$	0(1)	2.00e-01	5.38e-02	0(1)	2.00e-01	3.14e-02
		1(1)	1.01e-01	1.36e-02	1(1)	1.03e-01	8.06e-03
		2(1)	5.05e-02	3.40e-03	2(1)	5.24e-02	2.05e-03
		3(1)	2.53e-02	3.66e-02	3(1)	2.65e-02	4.15e-02
		4(2)	6.71e-04	1.14e-03	4(2)	2.25e-04	3.51e-04
		5(2)	1.74e-07	2.96e-07	5(2)	1.37e-08	2.13e-08
		6(2)	1.18e-14	2.00e-14	6(2)	4.35e-15	1.14e-15

the identity matrix, $i, j = 1, 2, 3, 4$. It then follows that

$$
\begin{aligned}
&P(3,4)P(1,2)A(\lambda)P(1,2)P(3,4)\\
&= \begin{pmatrix} 1 & 0 & 0 & 0\\ -\frac{3\lambda}{\lambda^2-2} & 1 & 0 & 0\\ \frac{1}{\lambda^2-2} & -\frac{2\lambda}{\lambda^4+3\lambda^2+2} & 1 & 0\\ \frac{2\lambda}{\lambda^2-2} & -\frac{3\lambda^2+2}{\lambda^4+3\lambda^2+2} & \frac{2\lambda-\lambda^3}{(\lambda^2-1)^2(\lambda^2+2)-2\lambda^2} & 1 \end{pmatrix}\\
&\quad \cdot \begin{pmatrix} \lambda^2-2 & 2\lambda & -1 & 0\\ 0 & \frac{\lambda^4+3\lambda^2+2}{\lambda^2-2} & -\frac{\lambda^3+\lambda}{\lambda^2-2} & 0\\ 0 & 0 & \frac{(\lambda^2-1)^2(\lambda^2+2)-2\lambda^2}{(\lambda^2-2)(\lambda^2+2)} & 2\lambda\\ 0 & 0 & 0 & \frac{\lambda^2[(\lambda^2-1)^2(\lambda^2+2)-4]}{(\lambda^2-1)^2(\lambda^2+2)-2\lambda^2} \end{pmatrix}.
\end{aligned}
$$

Clearly, it holds that $\mathrm{Ind}(\lambda_\star) = 0$. Therefore, the convergence rates of both Methods 3.1 and 3.2 are only linear for the eigenvalue $\lambda_\star = 0$.

6. Concluding Remarks

We have established the existence theories for the smooth LU decomposition and its block-analogue about a matrix-valued function. Moreover, based on these matrix decompositions we have presented and analyzed two numerical methods for computing multiple eigenvalues of the

Table 5.6: Numerical results for Example 5.2 using Method 3.2, $\alpha = -1$

$\lambda_\star$	$\lambda^{(0)}$	Method 3.2			Method 3.2 in [28]		
		$k(m)$	$\lvert\lambda_\star - \lambda^{(k)}\rvert$	$\lVert U_{22}(\lambda^{(k)})\rVert_F$	$k(m)$	$\lvert\lambda_\star - \lambda^{(k)}\rvert$	$\lVert R_{22}(\lambda^{(k)})\rVert_F$
0	-0.2	0(1)	2.00e-01	2.14e-02	0(1)	2.00e-01	1.48e-02
		1(1)	1.06e-01	5.72e-03	1(1)	9.72e-02	3.63e-03
		2(1)	5.39e-02	1.46e-03	2(1)	4.73e-02	8.87e-04
		3(1)	2.71e-02	3.68e-04	3(1)	2.33e-02	2.13e-02
		4(1)	1.36e-02	1.39e-02	4(2)	1.23e-03	1.07e-03
		5(2)	3.61e-04	3.61e-04	5(2)	3.65e-06	3.16e-06
		6(2)	8.60e-08	9.60e-09	6(2)	3.22e-11	2.79e-11
1	1.2	0(1)	2.00e-01	4.07e-02	0(1)	2.00e-01	2.33e-02
		1(1)	1.01e-01	1.03e-02	1(1)	1.01e-01	5.87e-03
		2(1)	5.10e-02	2.60e-03	2(1)	5.05e-02	1.47e-03
		3(1)	2.55e-02	2.42e-02	3(1)	2.53e-02	2.98e-02
		4(2)	2.21e-04	2.94e-04	4(2)	5.42e-04	6.26e-04
		5(2)	1.43e-08	9.90e-09	5(2)	2.44e-07	2.82e-07
					6(2)	5.00e-14	5.78e-14
$\imath$	$1.2\imath$	0(1)	2.00e-01	1.95e-01	0(1)	2.00e-01	1.38e-01
		1(1)	5.85e-02	4.21e-02	1(1)	5.68e-02	2.89e-02
		2(1)	6.83e-03	4.38e-03	2(1)	6.46e-03	2.93e-03
		3(1)	1.07e-04	6.78e-05	3(1)	9.69e-05	4.34e-05
		4(1)	2.68e-08	9.69e-09	4(1)	2.23e-08	9.95e-09

nonlinear eigenvalue problems. The new methods possess locally quadratic convergence rates, in the viewpoints of both theory and applications. Hence, they are effective solvers for nonlinear eigenvalue problems.

We remark that the smooth LU decompositions for the matrix-valued functions analytically depending on several parameters and their applications to solving the corresponding multiparameter eigenvalue problems, in particular, the smooth LDL^T decompositions of symmetric matrix-valued functions and their applications, may be developed in an analogous fashion. Moreover, the obtained results can be easily specified to the real field.

Acknowledgments. The research is supported by the National Basic Research Program (No. 2005CB321702), the China Outstanding Young Scientist Foundation (No. 10525102) and the National Natural Science Foundation (No. 10471146), P.R. China. This work was initiated when the first author was visiting State Key Laboratory of Scientific/Engineering Computing of Chinese Academy of Sciences during March-May in 2007. The support and hospitality from the Laboratory are very much appreciated.

References

[1] A.L. Andrew, K.E. Chu and P. Lancaster, On the numerical solution of nonlinear eigenvalue problems, *Computing*, **55** (1995), 91-111.

[2] P.M. Anselone and L.B. Rall, The solution of characteristic value-vector problems by Newton's method, *Numer. Math.*, **11** (1968), 38-45.

[3] A. Barrlund, Perturbation bounds for the LDL^H and LU decompositions, *BIT Numer. Math.*, **31** (1991), 358-363.

[4] C. Bendtsen, O.H. Nielsen and L.B. Hansen, Solving large nonlinear generalized eigenvalue problems from density functional theory calculations in parallel, *Appl. Numer. Math.*, **37** (2001), 189-199.

[5] T. Betcke and H. Voss, A Jacobi-Davidson-type projection method for nonlinear eigenvalue problems, *Future Gener. Comput. Sys.*, **20** (2004), 363-372.

[6] A. Bunse-Gerstner, R. Byers, V. Mehrmann and N.K. Nichols, Numerical computation of an analytic singular value decomposition of a matrix valued function, *Numer. Math.*, **60**(1991), 1-39.

[7] T.F. Chan, On the existence and computation of LU-factorizations with small pivots, *Math. Comput.*, **42** (1984), 535-547.

[8] H. Dai and P. Lancaster, Numerical methods for finding multiple eigenvalues of matrices depending on parameters, *Numer. Math.*, **76** (1997), 189-208.

[9] E.M. Daya and M. Potier-Ferry, A numerical method for nonlinear eigenvalue problems – application to vibrations of viscoelastic structures, *Comput. Struct.*, **79** (2001), 533-541.

[10] L. Dieci and T. Eirola, On smooth decompositions of matrices, *SIAM J. Matrix Anal. Appl.*, **20** (1999), 800-819.

[11] N.A. Dumont, On the solution of generalized non-linear complex-symmetric eigenvalue problems, *Int. J. Numer. Meth. Engrg.*, **71** (2007), 1534-1568.

[12] R. Fletcher, Practical Methods of Optimization, Vol. 1: Unconstrained Optimization, John Wiley & Sons, Chichester, 1980.

[13] H. Gingold and P.-F. Hsieh, Globally analytic triangularization of a matrix function, *Linear Algebra Appl.*, **169** (1992), 75-101.

[14] G.H. Golub and C.F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins Univ. Press, Baltimore, 1996.

[15] J.-S. Guo, W.-W. Lin and C.-S. Wang, Nonequivalence deflation for the solution of matrix latent value problems, *Linear Algebra Appl.*, **231** (1995), 15-45.

[16] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge Univ. Press, Cambridge, 1985.

[17] T.-M. Hwang, W.-W. Lin and E.-K. Yang, Rank revealing LU factorizations, *Linear Algebra Appl.*, **175** (1992), 115-141.

[18] T.-M. Hwang, W.-W. Lin and D. Pierce, Improved bound for rank revealing LU factorizations, *Linear Algebra Appl.*, **261** (1997), 173-186.

[19] T.-M. Hwang, W.-W. Lin, J.-L. Liu and W.-C. Wang, Jacobi-Davidson methods for cubic eigenvalue problems, *Numer. Linear Algebra Appl.*, **12** (2005), 605-624.

[20] N.K. Jain and K. Singhal, On Kublanovskaya's approach to the solution of the generalized latent value problem for functional λ-matrices, *SIAM J. Numer. Anal.*, **20** (1983), 1062-1070.

[21] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1966.

[22] V.B. Khazanov and V.N. Kublanovskaya, Spectral problems for matrix pencils: methods and algorithms II, *Sov. J. Numer. Anal. Math. Modelling*, **3** (1988), 467-485.

[23] V.N. Kublanovskaya, On an approach to the solution of the generalized latent value problem for λ-matrices, *SIAM J. Numer. Anal.*, **7** (1970), 532-537.

[24] P. Lancaster, A generalized Rayleigh quotient iteration for lambda-matrices, *Arch. Ration. Mech. Anal.*, **8** (1961), 309-322.

[25] P. Lancaster, Lambda-Matrices and Vibrating Systems, Pergamon Press, Oxford, 1966.

[26] P. Lancaster and M. Tismenetsky, The Theory of Matrices, 2nd ed., Academic Press, Orlando, 1985.

[27] R.-C. Li, QR decomposition and nonlinear eigenvalue problems, *Math. Numer. Sinica*, **11** (1989), 374-385. (In Chinese)

[28] R.-C. Li, Compute multiply nonlinear eigenvalues, *J. Comput. Math.*, **10** (1992), 1-20.

[29] A. Neumaier, Residual inverse iteration for the nonlinear eigenvalue problem, *SIAM J. Numer. Anal.*, **22** (1985), 914-923.

[30] C.-T. Pan, On the existence and computation of rank-revealing LU-factorizations, *Linear Algebra Appl.*, **316** (2000), 199-222.

[31] G. Peters and J.H. Wilkinson, $Ax = \lambda Bx$ and the generalized eigenproblem, *SIAM J. Numer. Anal.*, **7** (1970), 479-492.

[32] G. Peters and J.H. Wilkinson, Inverse iteration, ill-conditioned equations and Newton's method, *SIAM Rev.*, **21** (1979), 339-360.

[33] J. Rebaza, Smooth Schur factorizations in the continuation of separatrices, *Linear Algebra Appl.*, **421** (2007), 138-156.

[34] A. Ruhe, Algorithms for the nonlinear eigenvalue problem, *SIAM J. Numer. Anal.*, **10** (1973), 674-689.

[35] A. Ruhe, A rational Krylov algorithm for nonlinear matrix eigenvalue problems, *J. Math. Sci.*, **114** (2003), 1854-1856.

[36] G.L.G. Sleijpen, A.G.L. Booten, D.R. Fokkema and H.A. Van Der Vorst, Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems, *BIT Numer. Math.*, **36** (1996), 595-633.

[37] A. Spence and C. Poulton, Inverse iteration for nonlinear eigenvalue problems, in Electromagnetic Scattering in IUTAM Symposium on Asymptotics, Singularities and Homogenisation in Problems of Mechanics, A.B. Movchan ed., 585-594, Kluwer Academic Publishers, 2003.

[38] A. Spence and C. Poulton, Photonic band structure calculations using nonlinear eigenvalue techniques, *J. Comput. Phys.*, **204** (2005), 65-81.

[39] H. Unger, Nichtlineare Behandlung von Eigenwertaufgaben, *Z. Angew. Math. Mech.*, **30** (1950), 281-282.

[40] H. Voss, A maxmin principle for nonlinear eigenvalue problems with application to a rational spectral problem in fluid-solid vibration, *Appl. Math.*, **48** (2003), 607-622.

[41] H. Voss, An Arnoldi method for nonlinear eigenvalue problems, *BIT Numer. Math.*, **44** (2004), 387-401.

[42] R. Wobst, The generalized eigenvalue problem and acoustic surface wave computations, *Computing*, **39** (1987), 57-69.

[43] K. Wright, Differential equations for the analytic singular value decomposition of a matrix, *Numer. Math.*, **63** (1992), 283-295.