Journal of the
Nigerian Mathematical Society
Vol. 33, pp. 205-230, 2014

A COMPARISON OF THE IMPLICIT DETERMINANT METHOD AND INVERSE ITERATION

R. O. AKINOLA[1] AND A. SPENCE

ABSTRACT. It is well known that if the largest or smallest eigenvalue of a matrix has been computed by some numerical algorithms and one is interested in computing the corresponding eigenvector, one method that is known to give such good approximations to the eigenvector is inverse iteration with a shift. However, in a situation where the desired eigenvalue is defective, inverse iteration converges harmonically to the eigenvalue close to the shift. In this paper, we extend the implicit determinant method of Spence and Poulton [13] to compute a defective eigenvalue given a shift close to the eigenvalue of interest. For a defective eigenvalue, the proposed approach gives quadratic convergence and this is verified by some numerical experiments.

Keywords and phrases: Implicit determinant, Inverse iteration, Quadratic convergence, Eigenvalues, Eigenvector, Jordan blocks
AMS Subject Classification: 65F15, 15A18.

1. INTRODUCTION

Inverse iteration was first discovered by Wielandt in his paper [16] and later studied extensively by Wilkinson in ([17], [18]). Given that an eigenvalue of a matrix has been computed by QR factorization or any other algorithm, inverse iteration is used to calculate the corresponding eigenvector. Inverse iteration can also be used to find an eigenvalue and its corresponding eigenvector, infact, Wilkinson linked it with Newton's method.

Consider the nonlinear eigenvalue problem

$$\mathbf{A}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}, \quad \boldsymbol{\phi} \in \mathbb{R}^n, \quad \boldsymbol{\phi} \neq \mathbf{0} \quad \text{and} \quad \lambda \in \mathbb{R}. \tag{1}$$

equation Inverse iteration is an iterative method for computing the eigenvalue and corresponding eigenvector of $\mathbf{A}$ simultaneously [18]. Let α be a shift or an approximation to an eigenvalue of $\mathbf{A}$, inverse iteration is also used to compute an eigenvector associated with an eigenvalue closest to α. Inverse iteration can also be described as

Received by the editors February 11, 2013; Revised: June 20, 2013; Accepted: July 18, 2013

[1]Corresponding author

applying the Power method to $(\mathbf{A} - \alpha I)^{-1}$ in exact arithmetic [8]. However, in situations where the desired eigenvalue is defective, inverse iteration converges harmonically to the eigenvalue close to the shift. It is as a result of this that in this article, we extend the implicit determinant method of Spence and Poulton [13] to compute a defective eigenvalue given a shift close to the eigenvalue of interest. For a defective eigenvalue, the proposed approach gives quadratic convergence. In otherwords, we give the nonsymmetric version of inverse iteration and then extend the implicit determinant method of Spence and Poulton (in their case, they considered a parameter-dependent Hermittian matrix) to a nonsymmetric matrix. We conclude by comparing the nonsymmetric version of the implicit determinant method with inverse iteration. The discussion on inverse iteration in this paper is a special case of [5] for the standard eigenvalue problem.

The implicit determinant method is actually an application of Newtons method in finding the zeros of a nonlinear function $f(\lambda)$ (to be described shortly). The main crux of this article centres around the theory of Jordan blocks and precisely two-dimensioal Jordan blocks.

Two-dimensional Jordan blocks arise when one considers the problem of computing the nearest defective matrix from a simple one. Alam and Bora [2] developed a numerical algorithm for computing the distance of a simple matrix from the set of matrices having a Jordan block of at least dimension two. Jordan blocks appear in Freitag and Spence [6], who computed the distance of a stable matrix to the set of unstable matrices by computing a 2-dimensional Jordan block in a special class of parameter-dependent Hamiltonian matrices. Next, we define some Linear Algebra terms and summarize the theory of Jordan blocks.

We first define what a Jordan block is, algebraic and geometric multiplicities of the eigenvalue of a matrix, as well as what it means for a matrix to have a 2-dimensional Jordan block.

Let $\mathbf{A} \in \mathbb{R}$ be a real n by n matrix and $\lambda \in \mathbb{C}$ an eigenvalue of $\mathbf{A}$ corresponding to the nonzero eigenvector $\boldsymbol{\phi} \in \mathbb{C}^n$, such that

$$\mathbf{A}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}. \tag{2}$$

The vector $\boldsymbol{\phi}$ is often referred to as a right eigenvector [11]. A left eigenvector corresponding to the eigenvalue λ is defined as any nonzero vector $\boldsymbol{\psi}$ that satisfies $\boldsymbol{\psi}^T\mathbf{A} = \lambda\boldsymbol{\psi}^T$. The term geometric multiplicity of an eigenvalue λ of $\mathbf{A}$ is defined as the dimension of the nullspace of $(\mathbf{A}-\lambda\mathbf{I})$. The algebraic multiplicity of an eigenvalue

λ of $\mathbf{A}$ is its multiplicity as a root of the characteristic polynomial of $\mathbf{A}$ (see, for example, [15, p. 184]). We say λ is algebraically simple if it is a simple root of the characteristic polynomial. If λ is algebraically simple, then its corresponding left and right eigenvectors are not orthogonal, that is $\boldsymbol{\psi}^T\boldsymbol{\phi} \neq 0$ (see, for example, [5, p. 29, equation 2.2]).

Organization of the Paper

The plan of this paper is as follows. In Section 1.1, we describe two mathematical tools used: the 'ABCD' Lemma and the implicit determinant method. Furthermore, in Section 1.2, we give the nonsymmetric version of inverse iteration and then extend the implicit determinant method of Spence and Poulton to a nonsymmetric matrix. We also compare this nonsymmetric version of the implicit determinant method with inverse iteration. Numerical experiments are given which confirms and backs up our theoretical discussion and we used `Matlab`® Version 7.9.0.529 (R2009b). As mentioned earlier in the introduction, since the implicit determinant method is an application of Newton's method, we give a 'rigorous' proof on the quadratic convergence of Newton's method in Appendix A.

In the next section, we present a review on the implicit determinant method of Spence and Poulton for the solution of a nonlinear eigenvalue problem arising from photonic crystals [13] as well as Keller's [10] ABCD Lemma.

1. PRELIMINARY

1.1. Background: ABCD Lemma and the Implicit Determinant Method.

In this section, we present two key mathematical tools that will be of great use in this paper. In the first case, we present Keller's [10] ABCD Lemma. Secondly, we review the implicit determinant method of Spence and Poulton [13] which makes use of a special case of the ABCD Lemma and Cramer's rule. The key results in this section are Lemmas 1.1 and 1.2.

We begin by defining what it means for a matrix to have a 2-dimensional Jordan block. But before we do that, it is important to know what a Jordan block is first.

Definition 1.1. [12, p. 358] *A square upper-triangular matrix* $\mathbf{J}(\lambda)$ *that satisfies the following properties*

(a). all its main diagonal entries equal λ*,*

(b). all its entries on the first superdiagonal equal to one,
(c). all other entries are zero,
is called a Jordan block.

The following result explains the relationship between the Jordan decomposition of $\mathbf{A}$ and the Jordan block of $\mathbf{A}$.

Theorem 1.1. [7, p. 317] *If $\mathbf{A} \in \mathbb{C}^{n\times n}$, then there exists a nonsingular $\mathbf{Y} \in \mathbb{C}^{n\times n}$ such that*

$$\mathbf{J} = \mathbf{Y}^{-1}\mathbf{A}\mathbf{Y} = \mathrm{diag}\left(\mathbf{J}(\lambda_1), \mathbf{J}(\lambda_2), \ldots, \mathbf{J}(\lambda_t)\right),$$

where

$$\mathbf{J}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \cdots & 0 \\ 0 & \lambda_i & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 0 & \lambda_i \end{bmatrix},$$

is a $m(\lambda_i)$ by $m(\lambda_i)$ and $m(\lambda_1)+m(\lambda_2)+\cdots+m(\lambda_t) = n$, $m(\lambda_i)$ is the algebraic multiplicity of λ_i and t is the number of linearly independent eigenvectors of $\mathbf{A}$ corresponding to the number of blocks.

A way to recognise if the matrix $\mathbf{A}(\gamma)$ has a 2-dimensional Jordan block is given in the next definition.

Definition 1.2. [6] *Let λ^* be an eigenvalue of $\mathbf{A}$. $\mathbf{A}$ has a 2-dimensional Jordan block corresponding to the eigenvalue λ^* if λ^* has algebraic multiplicity 2 and geometric multiplicity 1.*

An immediate consequence of λ^* being algebraically double and geometrically simple in the above definition is explained as follows. If $\boldsymbol{\phi}^* \in \mathcal{N}\left(\mathbf{A} - \lambda^*\mathbf{I}\right)\backslash\{\mathbf{0}\}$ and $\boldsymbol{\psi}^* \in \mathcal{N}\left(\mathbf{A} - \lambda^*\mathbf{I}\right)^T\backslash\{\mathbf{0}\}$, then

$$\boldsymbol{\psi}^{*T}\boldsymbol{\phi}^* = 0, \tag{3}$$

and there exists a generalised eigenvector $\hat{\boldsymbol{\phi}}^*$ corresponding to λ^* which satisfies

$$\left(\mathbf{A} - \lambda^*\mathbf{I}\right)\hat{\boldsymbol{\phi}}^* = \boldsymbol{\phi}^*, \quad \textit{and} \qquad \boldsymbol{\psi}^{*T}\hat{\boldsymbol{\phi}}^* \neq 0. \tag{4}$$

We have used the Jordan chain equations (see, for example [12, pp. 359]) to arrive at the last equation and the condition $\boldsymbol{\psi}^{*T}\hat{\boldsymbol{\phi}}^* \neq 0$ ensures that the dimension of the Jordan block is exactly 2. After premultiplying both sides of (4) by $\left(\mathbf{A} - \lambda^*\mathbf{I}\right)$, we obtain

$$\left(\mathbf{A} - \lambda^*\mathbf{I}\right)^2\hat{\boldsymbol{\phi}}^* = \left(\mathbf{A} - \lambda^*\mathbf{I}\right)\boldsymbol{\phi}^* = \mathbf{0}.$$

This shows that the algebraic multiplicity of λ^* is at least two and that $\hat{\boldsymbol{\phi}}^*$ is indeed a generalised eigenvector. Before we continue, we give some further definitions in use.

Definition 1.3. [2] *An $n \times n$ matrix is simple if it has n distinct eigenvalues.*

Definition 1.4. [15, p. 185] *An eigenvalue is said to be defective if its algebraic multiplicity is greater than its geometric multiplicity. A matrix is said to be defective if it has one or more defective eigenvalues.*

First, we present the one-dimensional version of Keller's [10] ABCD Lemma.

Lemma 1.1. *The "ABCD" Lemma*
Let $\mathbf{A}$ be an n by n matrix, $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ and $d \in \mathbb{R}$. Let

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^T & d \end{bmatrix}, \tag{5}$$

be an $(n+1)$ by $(n+1)$ real matrix.

(a). Suppose that $\mathbf{A}$ is nonsingular, then there exists the following decomposition of $\mathbf{M}$,

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^T & d \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{c}^T\mathbf{A}^{-1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0}^T & d - \mathbf{c}^T\mathbf{A}^{-1}\mathbf{b} \end{bmatrix}. \tag{6}$$

The matrix $\mathbf{M}$ is nonsingular if and only if $d - \mathbf{c}^T\mathbf{A}^{-1}\mathbf{b} \neq 0$.

(b). If $\mathbf{A}$ is singular of $\operatorname{rank}(\mathbf{A}) = n-1$, then $\mathbf{M}$ is nonsingular if and only if $\boldsymbol{\psi}^T\mathbf{b} \neq 0$, for all $\boldsymbol{\psi} \in \mathcal{N}(\mathbf{A}^T)\backslash\{\mathbf{0}\}$ and $\mathbf{c}^T\boldsymbol{\phi} \neq 0$, for all $\boldsymbol{\phi} \in \mathcal{N}(\mathbf{A})\backslash\{\mathbf{0}\}$. Where $\mathcal{N}(\mathbf{A})$ is the nullspace of $\mathbf{A}$.

Proof: See [10]. ■

Next, we describe Spence and Poulton's implicit determinant method as formulated in [13]. The aim of presenting the implicit determinant method is because we want to extend it to the parameter-dependent nonsymmetric matrix case to find a 2-dimensional Jordan block.

The implicit determinant method of Spence and Poulton [13] is a method of converting a nonlinear problem for square matrices into an equivalent scalar problem. We can solve the scalar problem in a number of ways, for example, using the bisection method. The fact that it is efficient to implement Newton's method is an added advantage. In Spence and Poulton (2005), the theory of the implicit determinant method was given for the case in which $\mathbf{A}(\gamma)$ is Hermitian, and comparisons were made on the convergence of the implicit determinant method and nonlinear inverse iteration applied to a nonlinear eigenvalue problem arising in a photonic crystal problem.

Given a parameter-dependent Hermittian matrix $\mathbf{A}(\gamma)$. In addition, let $\mathbf{A}(\gamma)$ be a smooth function of γ. Let [13, p. 69]

$$\mathbf{A}(\gamma)\mathbf{x} = \mathbf{0}, \quad \text{where} \quad \mathbf{x} \neq \mathbf{0}, \tag{7}$$

be a parameter-dependent eigenvalue problem.

Consider the following $(n+1)$ by $(n+1)$ bordered linear system of equations [13, p. 70],

$$\begin{bmatrix} \mathbf{A}(\gamma) & \mathbf{b} \\ \mathbf{b}^H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ f \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \tag{8}$$

which shows that the eigenvector $\mathbf{x}$ is normalised using $\mathbf{b}^H\mathbf{x} = 1$. The following result is the main mathematical tool of Spence and Poulton's implicit determinant method.

Lemma 1.2. [13, pp. 70] *Let $(\mathbf{x}^*, \gamma^*)$ solve (7) with $\mathbf{A}(\gamma)$ Hermitian. Assume that zero is a simple eigenvalue of $\mathbf{A}(\gamma^*)$, such that*

(a). $\dim \mathcal{N}\big[\mathbf{A}(\gamma^*)\big] = 1$.
(b). For some $\mathbf{b} \in \mathbb{C}^n \backslash \{\mathbf{0}\}$, assume

$$\mathbf{b}^H\mathbf{x}^* \neq 0. \tag{9}$$

Then the $(n+1)$ by $(n+1)$ matrix $\mathbf{M}(\gamma)$ defined by

$$\mathbf{M}(\gamma) = \begin{bmatrix} \mathbf{A}(\gamma) & \mathbf{b} \\ \mathbf{b}^H & 0 \end{bmatrix},$$

is nonsingular at $\gamma = \gamma^*$.

Proof: See [10]. ■

From the result of Lemma 1.2, $\mathbf{M}(\gamma)$ is nonsingular at the root. Following [13], this means that by an application of the implicit function theorem (see, for example, [14, p. 186]) $\mathbf{M}(\gamma)$ is nonsingular for γ near γ^* because $\mathbf{A}(\gamma)$ is a smooth function of γ. Therefore, from equation (8), $\mathbf{x}$ and f are smooth functions of γ and we can write $\mathbf{x} = \mathbf{x}(\gamma)$ and $f = f(\gamma)$. So that (8) becomes

$$\begin{bmatrix} \mathbf{A}(\gamma) & \mathbf{b} \\ \mathbf{b}^H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}(\gamma) \\ f(\gamma) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{10}$$

Now, by applying Cramer's rule (see [9, p. 414]) to (10), we obtain

$$f(\gamma) = \frac{\det \mathbf{A}(\gamma)}{\det \mathbf{M}(\gamma)}. \tag{11}$$

Observe that because $\mathbf{A}(\gamma)$ and $\mathbf{M}(\gamma)$ are both Hermitian, this means that $f(\gamma)$ is real. We conclude by saying that the main idea behind the implicit determinant method is that if $\mathbf{M}(\gamma)$ is nonsingular, then $f(\gamma) = 0$ if and only if $\mathbf{A}(\gamma)$ is singular. So we seek zeros of $f(\gamma)$ as a way of finding the zeros of the determinant of $\mathbf{A}(\gamma)$. Spence and Poulton continue by finding the solution of $f(\gamma) = 0$ using Newton's method, which requires the calculation of $f'(\gamma)$, where $f'(\gamma) = \frac{\mathrm{d}}{\mathrm{d}\gamma} f(\gamma)$. This is accomplished by solving

$$\begin{bmatrix} \mathbf{A}(\gamma) & \mathbf{b} \\ \mathbf{b}^H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}'(\gamma) \\ f'(\gamma) \end{bmatrix} = - \begin{bmatrix} \mathbf{A}'(\gamma)\mathbf{x}(\gamma) \\ 0 \end{bmatrix},$$

obtained by differentiating both sides of (10) with respect to γ. After which the sequence of γ iterates is computed by $\gamma^{(k+1)} = \gamma^{(k)} - f(\gamma^{(k)})/f'(\gamma^{(k)})$, for $k = 0, 1, 2, \ldots$. Using the above matrix equation, $f'(\gamma^*)$ was shown to be equal to $-\mathbf{x}^{*H}\mathbf{A}'(\gamma^*)\mathbf{x}(\gamma^*)$. Hence, $f'(\gamma^*)$ is nonzero provided $\mathbf{x}^{*H}\mathbf{A}'(\gamma^*)\mathbf{x}(\gamma^*)$ is nonzero.

In the next section, we present the implicit determinant method for a nonsymmetric matrix and compare it with inverse iteration.

3. THE HEART OF THE MATTER

1.2. **A Comparison of the Implicit Determinant Method and Inverse Iteration.** Let $\mathbf{A}$ be a real n by n nonsymmetric matrix. In this section, we give the nonsymmetric version of inverse iteration and then extend the implicit determinant method of Spence and Poulton to a nonsymmetric $\mathbf{A}$. We conclude by comparing this version of the implicit determinant method with inverse iteration. The discussion on inverse iteration in this section is a special case of [5] for the standard eigenvalue problem.

Recall from (2) that $(\mathbf{A} - \lambda \mathbf{I})\boldsymbol{\phi} = \mathbf{0}$. So, if we add to (2) the eigenvector normalization $\mathbf{c}^T\boldsymbol{\phi} = 1$, then the extended system of nonlinear equations becomes: (see, also [5, p. 29])

$$\mathbf{F}(\mathbf{w}) = \begin{bmatrix} (\mathbf{A} - \lambda \mathbf{I})\boldsymbol{\phi} \\ \mathbf{c}^T\boldsymbol{\phi} - 1 \end{bmatrix} = \mathbf{0}, \tag{12}$$

where $\mathbf{w} = [\boldsymbol{\phi}^T, \lambda]$. Using the ABCD Lemma [10], it can be shown that the Jacobian $\mathbf{F}_{\mathbf{w}}(\mathbf{w})$ is nonsingular,

$$\mathbf{F}_{\mathbf{w}}(\mathbf{w}) = \begin{bmatrix} \mathbf{A} - \lambda\mathbf{I} & -\boldsymbol{\phi} \\ \mathbf{c}^T & 0 \end{bmatrix}. \tag{13}$$

at the root. Hence, its inverse exists at an algebraically simple eigenvalue (*i.e.*, $\boldsymbol{\psi}^T\boldsymbol{\phi} \neq 0$, for all $\boldsymbol{\psi} \in \mathcal{N}(\mathbf{A}^T - \lambda\mathbf{I})\backslash\{\mathbf{0}\}$ and $\boldsymbol{\phi} \in \mathcal{N}(\mathbf{A} - \lambda\mathbf{I})\backslash\{\mathbf{0}\}$) and if $\mathbf{c}$ is chosen such that $\mathbf{c}^T\boldsymbol{\phi} \neq 0$. Newton's method

$$\begin{aligned} \mathbf{F}_{\mathbf{w}}(\mathbf{w}^{(k)})\Delta\mathbf{w}^{(k)} &= -\mathbf{F}(\mathbf{w}^{(k)}) \\ \mathbf{w}^{(k+1)} &= \mathbf{w}^{(k)} + \Delta\mathbf{w}^{(k)}, \end{aligned} \tag{14}$$

with $\mathbf{c}^T\boldsymbol{\phi}^{(k)} = 1$, now becomes

$$\begin{bmatrix} \mathbf{A} - \lambda^{(k)}\mathbf{I} & -\boldsymbol{\phi}^{(k)} \\ \mathbf{c}^T & 0 \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{\phi}^{(k)} \\ \Delta\lambda^{(k)} \end{bmatrix} = -\begin{bmatrix} (\mathbf{A} - \lambda^{(k)}\mathbf{I})\boldsymbol{\phi}^{(k)} \\ \mathbf{c}^T\boldsymbol{\phi}^{(k)} - 1 \end{bmatrix}. \tag{15}$$

By expanding the above, we have the following system of equations

$$(\mathbf{A} - \lambda^{(k)}\mathbf{I})\Delta\boldsymbol{\phi}^{(k)} - \Delta\lambda^{(k)}\boldsymbol{\phi}^{(k)} = -(\mathbf{A} - \lambda^{(k)}\mathbf{I})\boldsymbol{\phi}^{(k)} \tag{16}$$

$$\mathbf{c}^T\Delta\boldsymbol{\phi}^{(k)} = 0. \tag{17}$$

After collecting like terms in the first equation above and using the relation
$\boldsymbol{\phi}^{(k+1)} = \boldsymbol{\phi}^{(k)} + \Delta\boldsymbol{\phi}^{(k)}$, we obtain

$$(\mathbf{A} - \lambda^{(k)}\mathbf{I})\boldsymbol{\phi}^{(k+1)} = \Delta\lambda^{(k)}\boldsymbol{\phi}^{(k)} \tag{18}$$

Upon division of both sides by $\Delta\lambda^{(k)}$ and letting $\mathbf{w}^{(k)} = \frac{\boldsymbol{\phi}^{(k+1)}}{\Delta\lambda^{(k)}}$ we have

$$(\mathbf{A} - \lambda^{(k)}\mathbf{I})\mathbf{w}^{(k)} = \boldsymbol{\phi}^{(k)}, \tag{19}$$

using the fact that $\mathbf{c}^T\Delta\boldsymbol{\phi}^{(k)} = 0$, we have $\mathbf{c}^T\boldsymbol{\phi}^{(k+1)} = \mathbf{c}^T(\boldsymbol{\phi}^{(k)} + \Delta\boldsymbol{\phi}^{(k)}) = 1$. Hence, $\mathbf{c}^T\mathbf{w}^{(k)} = \frac{1}{\Delta\lambda^{(k)}}$, from which $\Delta\lambda^{(k)} = \frac{1}{\mathbf{c}^T\mathbf{w}^{(k)}}$. Therefore,

$$\begin{aligned} \boldsymbol{\phi}^{(k+1)} &= \Delta\lambda^{(k)}\mathbf{w}^{(k)} \\ &= \frac{\mathbf{w}^{(k)}}{\mathbf{c}^T\mathbf{w}^{(k)}}. \end{aligned} \tag{20}$$

By making use of (14) we have

$$\begin{aligned} \lambda^{(k+1)} &= \lambda^{(k)} + \Delta\lambda^{(k)} \\ &= \lambda^{(k)} + \frac{1}{\mathbf{c}^T\mathbf{w}^{(k)}}. \end{aligned} \tag{21}$$

From the above analysis, Algorithm 1.1 is immediate.

Algorithm 1.1. *Inverse Iteration and Newton's Method*

(1) *Choose* $\boldsymbol{\phi}^{(0)}$, $\lambda^{(0)}$, $\mathbf{c}$ *such that* $\mathbf{c}^T\boldsymbol{\phi}^{(0)} = 1$, *and such that the conditions of Lemma 1.1 is satisfied and tol.*
(2) *For* $k = 1, 2, \cdots$, *until convergence*

(3) *Solve* $(\mathbf{A} - \lambda^{(k)}\mathbf{I})\mathbf{w}^{(k)} = \boldsymbol{\phi}^{(k-1)}$.
(4) *Compute* $\Delta\lambda^{(k)} = \frac{1}{\mathbf{c}^T\mathbf{w}^{(k)}}$.
(5) *Compute* $\lambda^{(k+1)} = \lambda^{(k)} + \Delta\lambda^{(k)}$.
(6) *Update* $\boldsymbol{\phi}^{(k+1)} = \Delta\lambda^{(k)}\mathbf{w}^{(k)}$.
(7) *Test for convergence.*
(8) *Output* $\boldsymbol{\phi}^*$ *and* λ^*.

Next, we describe the implicit determinant method for a nonsymmetric $\mathbf{A}$. Consider the following $(n+1)$ by $(n+1)$ bordered linear system of equations [13, p. 70],

$$\begin{bmatrix} (\mathbf{A} - \lambda\mathbf{I}) & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ f \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{22}$$

Lemma 1.3. *Let* $(\mathbf{x}^*, \lambda^*)$ *solve* (22). *Assume that zero is a simple eigenvalue of* $(\mathbf{A} - \lambda^*\mathbf{I})$, *such that*

(a). $\dim \mathcal{N}\big[(\mathbf{A} - \lambda^*\mathbf{I})\big] = 1$.
(b). For some $\mathbf{b}$, $\mathbf{c} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$, *assume*

$$\boldsymbol{\psi}^{*T}\mathbf{b} \neq 0, \quad \textit{and} \quad \mathbf{c}^T\mathbf{x}^* \neq 0, \tag{23}$$

for all $\boldsymbol{\psi}^* \in N\big[(\mathbf{A}^T - \lambda^*\mathbf{I})\big]$.

Then the $(n+1)$ *by* $(n+1)$ *matrix* $\mathbf{M}(\lambda^*)$ *defined by*

$$\mathbf{M}(\lambda^*) = \begin{bmatrix} (\mathbf{A} - \lambda^*\mathbf{I}) & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix},$$

is nonsingular.

Proof: See [10]. ∎

Since the result of Lemma 1.3 shows that $\mathbf{M}(\lambda^*)$ is nonsingular, then following [13], this means that by an application of the implicit function theorem (see, for example, [14, p. 186]), $\mathbf{M}(\lambda)$ is nonsingular for λ near λ^* because $(\mathbf{A} - \lambda\mathbf{I})$ is a smooth function of λ. Therefore, from (22) $\mathbf{x}$ and f are smooth functions of λ and we can write $\mathbf{x} = \mathbf{x}(\lambda)$ and $f = f(\lambda)$. So that (22) becomes

$$\begin{bmatrix} (\mathbf{A} - \lambda\mathbf{I}) & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}(\lambda) \\ f(\lambda) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{24}$$

Now, by applying Cramer's rule (see [9, p. 414]) to (24), we obtain

$$f(\lambda) = \frac{\det(\mathbf{A} - \lambda \mathbf{I})}{\det \mathbf{M}(\lambda)}. \tag{25}$$

By the implicit determinant method, if $\mathbf{M}(\lambda)$ is nonsingular, then $f(\lambda) = 0$ if and only if $(\mathbf{A} - \lambda \mathbf{I})$ is singular, which is attainable at the root. So we seek zeros of $f(\lambda)$ as a way of finding the zeros of the determinant of $(\mathbf{A} - \lambda \mathbf{I})$. To find the solution of $f(\lambda) = 0$ using Newton's method, we need $f'(\lambda)$, where $f'(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda} f(\lambda)$. This means we have to differentiate (24) with respect to λ and solve

$$\begin{bmatrix} (\mathbf{A} - \lambda \mathbf{I}) & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}'(\lambda) \\ f'(\lambda) \end{bmatrix} = \begin{bmatrix} \mathbf{x}(\lambda) \\ 0 \end{bmatrix}. \tag{26}$$

After which the sequence of λ iterates is computed by

$$\lambda^{(k+1)} = \lambda^{(k)} - \frac{f(\lambda^{(k)})}{f'(\lambda^{(k)})}, \tag{27}$$

for $k = 0, 1, 2, \ldots$, until convergence. At the root, observe that by expanding the first row of (26), one obtains

$$(\mathbf{A} - \lambda^* \mathbf{I})\mathbf{x}'(\lambda^*) + f'(\lambda^*)\mathbf{b} = \mathbf{x}(\lambda^*). \tag{28}$$

Hence, after premultiplying both sides by $\boldsymbol{\psi}^{*T}$, then

$$f'(\lambda^*) = \frac{\boldsymbol{\psi}^{*T}\mathbf{x}(\lambda^*)}{\boldsymbol{\psi}^{*T}\mathbf{b}}, \quad \textit{since} \quad \boldsymbol{\psi}^{*T}\mathbf{b} \neq 0. \tag{29}$$

But for an algebraically simple eigenvalue, the left and right eigenvector are not orthogonal *i.e.*, $\boldsymbol{\psi}^{*T}\mathbf{x}(\lambda^*) \neq 0$. Therefore,

$$f'(\lambda^*) \neq 0. \tag{30}$$

Algorithm 1.2 is now immediate.

Algorithm 1.2. *Implicit Determinant Method Algorithm for a Simple Matrix*

(1) Choose $\mathbf{b}, \mathbf{c}$, $\lambda^{(0)}$, *such that* $\mathbf{M}(\lambda^{(0)})$ *is nonsingular, tol.*
(2) For $k = 1, 2, \cdots$, *until convergence*
(3) Solve (24) *for* $\mathbf{x}(\lambda)$ *and* $f(\lambda)$.
(4) Solve (26) *for* $\mathbf{x}'(\lambda)$ *and* $f'(\lambda)$.
(5) Update

$$\lambda^{(k+1)} = \lambda^{(k)} - \frac{f(\lambda^{(k)})}{f'(\lambda^{(k)})}.$$

(6) Test for convergence.
(7) Output $\mathbf{x}(\lambda^*)$ *and* λ^*.

Stop Algorithm 1.2 as soon as

$$\|f(\lambda^{(k)})\| \leq tol.$$

Now, we present the theory to explain the link between the implicit determinant method and inverse iteration. For ease of notation, we shall drop the superscripts k and write $\lambda^{(k+1)} = \lambda^+$ and $\lambda^{(k)} = \lambda$.

We start by assuming that $\lambda \neq \lambda^*$, which implies $(\mathbf{A} - \lambda\mathbf{I})$ is nonsingular. Observe by expanding along the first row of (24), that

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x}(\lambda) + \mathbf{b}f(\lambda) = \mathbf{0}, \quad \text{and} \quad \mathbf{x}(\lambda) + (\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{b}f(\lambda) = \mathbf{0}.$$

Premultiply both sides by $\mathbf{c}^T$ and solve for $f(\lambda)$ using the second row of (24) to obtain

$$f(\lambda) = -\frac{1}{\mathbf{c}^T(\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{b}}. \tag{31}$$

Similarly, it can be shown by using the first row of (26) and $\mathbf{c}^T\mathbf{x}'(\lambda)$ from the second row that

$$f'(\lambda) = \frac{\mathbf{c}^T(\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{x}(\lambda)}{\mathbf{c}^T(\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{b}}. \tag{32}$$

Note from (19), that if we replace $\mathbf{w}$ with $\mathbf{y}$ and $\boldsymbol{\phi}$ with $\mathbf{x}$, that is,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{y} = \mathbf{x}, \quad \text{then} \quad \mathbf{y} = (\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{x}(\lambda),$$

and we can rewrite (32) as

$$f'(\lambda) = \frac{\mathbf{c}^T\mathbf{y}}{\mathbf{c}^T(\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{b}}.$$

Now, it is easy to see that (27) reduces to

$$\lambda^+ = \lambda + \frac{1}{\mathbf{c}^T\mathbf{y}}.$$

The above equation, which is the same update for λ as that obtained using inverse iteration, see (21). What remains now is to give the implicit determinant method's analogue for the eigenvector update which can be explained as follows.

Set $(\mathbf{A} - \lambda\mathbf{I})\mathbf{z} = \mathbf{b}$ in (24) and expand the first row to obtain

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x}(\lambda) + f(\lambda)(\mathbf{A} - \lambda\mathbf{I})\mathbf{z} = \mathbf{0}, \quad \text{then} \quad \mathbf{x}(\lambda) = -f(\lambda)\mathbf{z}.$$

By premultiplying both sides of $\mathbf{x}(\lambda) = -f(\lambda)\mathbf{z}$, by $\mathbf{c}^T$, we have $f(\lambda) = -\frac{1}{\mathbf{c}^T\mathbf{z}}$. Hence, $\mathbf{x}(\lambda) = -f(\lambda)\mathbf{z}$ can be rewritten as $\mathbf{x}(\lambda) = \frac{\mathbf{z}}{\mathbf{c}^T\mathbf{z}}$.

Example 1.1. *Consider*

$$\mathbf{D} = \mathrm{diag}\{-10, 9, -8, 7, -6, 5, -4, 3, -2, 1\} \tag{33}$$

and the matrix $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1}$, *where* $\mathbf{X}$ *is a nonsymmetric random* 10 *by 10 matrix. So* $\mathbf{A}$ *is nonsingular. We compare the convergence to a simple eigenvalue of inverse iteration and the implicit determinant method. We make the fixed choice for* $\mathbf{c}$ *and* $\mathbf{b}$ *as* $\mathbf{c} = \mathbf{1}/\|\mathbf{1}\|$ *and* $\mathbf{b} = [(\mathbf{A} - \lambda^{(0)}\mathbf{I})^T]^{-1}\mathbf{c}$, *where* $\lambda^{(0)} = 0.1$, *and* $\mathbf{c} = \mathbf{1}$ *is the vector of all ones. See results in Tables 1 and 2.*

We see from Table 1 that the implicit determinant Algorithm 1.2 converged to the eigenvalue $\lambda^* \approx 1$ *and quadratic convergence is seen in the third column.*

TABLE 1. Values of $\lambda^{(k)}$ and $f(\lambda^{(k)})$ of Example 1.1 using the implicit determinant method (Algorithm 1.2). Column 3 show that the results converged quadratically for $k = 5$ and 6. The last column shows the angle between $\mathbf{x}(\lambda^{(k)})$ and $\mathbf{x}(\lambda^*)$ in radians.

k	$\lambda^{(k)}$	$f(\lambda^{(k)})$	$\theta^{(k)}$
0	1.0000000000000001e-01	Inf	0.0e+00
1	1.0000000000000001e-01	1.5e-01	4.7e-01
2	6.2059513954948309e-01	4.5e-02	1.6e-01
3	9.4639992744280654e-01	5.6e-03	2.0e-02
4	9.9905160715023478e-01	9.8e-05	3.6e-04
5	9.9999970803503402e-01	3.0e-08	1.1e-07
6	9.9999999999987121e-01	4.0e-15	0.0e+00
7	9.9999999999990963e-01	8.3e-16	0.0e+00

TABLE 2. Values of $\lambda^{(k)}$ and $|\Delta\lambda^{(k)}|$ of Example 1.1. Column 3 show that the inverse iteration Algorithm 1.1 converged quadratically for $k = 3, 4$ and 5. However, Matlab complained of the singularity of $(\mathbf{A} - \lambda^*\mathbf{I})$ for $k = 7$. The last column shows the angle between $\boldsymbol{\phi}^{(k)}$ and $\mathbf{x}(\lambda^*)$ in radians.

k	$\lambda^{(k)}$	$|\Delta\lambda^{(k)}|$	$\theta^{(k)}$
0	1.0000000000000001e-01	Inf	0.0e+00
1	1.0000000000000001e-01	1.4e+00	4.7e-01
2	1.4835721862680826e+00	3.5e-01	4.7e-01
3	1.1307842488543134e+00	1.2e-01	5.3e-02
4	1.0108595562087610e+00	1.1e-02	2.4e-03
5	9.9994813631139773e-01	5.2e-06	1.4e-05
6	9.9999999987802568e-01	1.2e-10	0.0e+00
7	1.0000000000000322e+00	-	0.0e+00

We conclude from the third columns of Tables 1 and 2 that both the implicit determinant method and inverse iteration converge quadratically with a good enough starting guess.

We conclude this section by saying that the implicit determinant method is an inefficient way of carrying out inverse iteration. This is because it involves solving two linear systems at each iteration. However, one advantage of the implicit determinant method over inverse iteration is that it converges quadratically when the dimension of the nullspace of $(\mathbf{A} - \lambda^*\mathbf{I})$ is one, which includes the case when λ^* is a defective eigenvalue. Inverse iteration converges with $\mathcal{O}(1/k)$ for $k \to \infty$, $k \in \mathbb{N}$ when λ^* is a defective eigenvalue as illustrated in [18]. Other advantages of the implicit determinant method can be seen in [1].

In the next section, we present a description of the implicit determinant method and inverse iteration for a defective eigenvalue.

1.3. **Implicit Determinant Method and Inverse Iteration for a Defective Eigenvalue.** At the tail end of the last section, we mentioned that the implicit determinant method converges quadratically for a defective eigenvalue and inverse iteration converges harmonically. In this section, we present the theory of the implicit determinant method for a defective matrix and present numerical results to compare its convergence with inverse iteration.

We begin by making the following remark backed up with a numerical example. For a defective eigenvalue λ^*, Algorithm 1.2 converges linearly.

Example 1.2. *We consider the* $\mathbf{A} = \mathbf{X}\mathbf{J}\mathbf{X}^{-1}$*, where* $\mathbf{X}$ *is a nonsymmetric but random matrix and*

$$\mathbf{J} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 \end{bmatrix}. \qquad (34)$$

The structure of $\mathbf{J}$ *shows that* $\mathbf{A}$ *has a 2-dimensional Jordan block corresponding to the double eigenvalue -1. The aim is to compare* *the convergence of the implicit determinant method and inverse iteration. We use the same values for* $\mathbf{c}$ *and* $\mathbf{b}$ *as in the preceding example with* $\lambda^{(0)} = -0.1$*. Results of a numerical experiment which shows that Algorithm 1.2 converges linearly is presented in Table 3.*

We observe linear convergence to the double eigenvalue $\lambda^* = -1$ *as shown in Table 3.*

TABLE 3. Values of $\lambda^{(k)}$ and $f(\lambda^{(k)})$ of Example 1.1 using the implicit determinant method (Algorithm 1.2). Column 3 show that the results converged linearly to the eigenvalue $\lambda^* \approx -1$. The last column shows the angle between $\mathbf{x}(\lambda^{(k)})$ and $\mathbf{x}(\lambda^*)$ in radians.

k	$\lambda^{(k)}$	$f(\lambda^{(k)})$	$\theta^{(k)}$
0	*-1.0000000000000001e-01*	*Inf*	*0.0e+00*
1	*-1.0000000000000001e-01*	*2.9e-02*	*2.9e-01*
2	*-6.9283591626360030e-01*	*5.1e-03*	*1.3e-01*
3	*-8.7380099988202919e-01*	*1.1e-03*	*7.1e-02*
4	*-9.4317106696159647e-01*	*2.4e-04*	*3.7e-02*
5	*-9.7304975133933103e-01*	*5.7e-05*	*1.9e-02*
6	*-9.8687638851037474e-01*	*1.4e-05*	*9.4e-03*
7	*-9.9352418513524987e-01*	*3.4e-06*	*4.7e-03*
8	*-9.9678335110488692e-01*	*8.5e-07*	*2.4e-03*
9	*-9.9839696011811008e-01*	*2.1e-07*	*1.2e-03*
10	*-9.9919979743351017e-01*	*5.3e-08*	*5.9e-04*
11	*-9.9960022759069178e-01*	*1.3e-08*	*3.0e-04*
12	*-9.9980019596300884e-01*	*3.3e-09*	*1.5e-04*
13	*-9.9990011851389116e-01*	*8.3e-10*	*7.4e-05*
14	*-9.9995006439987510e-01*	*2.1e-10*	*3.7e-05*
15	*-9.9997503347743444e-01*	*5.2e-11*	*1.9e-05*
16	*-9.9998751727209412e-01*	*1.3e-11*	*9.3e-06*
17	*-9.9999375865895801e-01*	*3.2e-12*	*4.6e-06*
18	*-9.9999687988314290e-01*	*8.1e-13*	*2.3e-06*
19	*-9.9999844051901110e-01*	*2.0e-13*	*1.2e-06*
20	*-9.9999922146580489e-01*	*5.0e-14*	*5.8e-07*
21	*-9.9999961158516693e-01*	*1.3e-14*	*2.9e-07*
22	*-9.9999981779099434e-01*	*3.1e-15*	*1.3e-07*
23	*-9.9999992159144380e-01*	*6.0e-16*	*5.8e-08*

For a defective eigenvalue λ^*, with algebraic multiplicity two and geometric multiplicity one, then $\boldsymbol{\psi}^{*T}\mathbf{x}(\lambda^*) = 0$, and so from (29)

$$f'(\lambda^*) = 0. \tag{35}$$

Thus, (28) gives

$$(\mathbf{A}-\lambda^*\mathbf{I})\mathbf{x}'(\lambda^*)=\mathbf{x}(\lambda^*), \tag{36}$$

so that $\mathbf{x}'(\lambda^*)$ is a generalised eigenvector of λ^*. Since λ^* has algebraic multiplicity two,

$$\boldsymbol{\psi}^{*T}\mathbf{x}'(\lambda^*)\neq 0. \tag{37}$$

Now, differentiate (26) with respect to λ again to obtain

$$\begin{bmatrix}(\mathbf{A}-\lambda\mathbf{I}) & \mathbf{b}\\ \mathbf{c}^T & 0\end{bmatrix}\begin{bmatrix}\mathbf{x}''(\lambda)\\ f''(\lambda)\end{bmatrix}=\begin{bmatrix}2\mathbf{x}'(\lambda)\\ 0\end{bmatrix}. \tag{38}$$

After expanding along the first row and premultiplying both sides by $\boldsymbol{\psi}^{*T}$ at the root, then we obtain

$$\boldsymbol{\psi}^{*T}(\mathbf{A}-\lambda^*\mathbf{I})\mathbf{x}''(\lambda^*)+f''(\lambda^*)\boldsymbol{\psi}^{*T}\mathbf{b}=2\boldsymbol{\psi}^{*T}\mathbf{x}'(\lambda^*).$$

Since $\boldsymbol{\psi}^*\in\mathcal{N}(\mathbf{A}^T-\lambda^*\mathbf{I})$, this implies

$$f''(\lambda^*)=\frac{2\boldsymbol{\psi}^{*T}\mathbf{x}'(\lambda^*)}{\boldsymbol{\psi}^{*T}\mathbf{b}}.$$

Consequently, from (37), we conclude that

$$f''(\lambda^*)\neq 0. \tag{39}$$

It is well known that for a double root, the standard Newton's method of a scalar nonlinear problem

$$\lambda^{(k+1)}=\lambda^{(k)}-\frac{f(\lambda^{(k)})}{f'(\lambda^{(k)})},$$

can be altered to

$$\lambda^{(k+1)}=\lambda^{(k)}-2\frac{f(\lambda^{(k)})}{f'(\lambda^{(k)})},$$

for $k=0,1,2,3,\ldots$. This now leads to the following implicit determinant method algorithm for an algebraically double and geometrically simple eigenvalue.

Algorithm 1.3. *Implicit Determinant Method Algorithm for a Defective Matrix*

(1) *Choose* $\mathbf{b},\mathbf{c}$, $\lambda^{(0)}$, *such that* $\mathbf{M}(\lambda^{(0)})$ *is nonsingular, tol.*
(2) *For* $k=1,2,\cdots$, *until convergence*
(3) *Solve* (24) *for* $\mathbf{x}(\lambda)$ *and* $f(\lambda)$.
(4) *Solve* (26) *for* $\mathbf{x}'(\lambda)$ *and* $f'(\lambda)$.
(5) *Update*

$$\lambda^{(k+1)}=\lambda^{(k)}-2\frac{f(\lambda^{(k)})}{f'(\lambda^{(k)})}.$$

(6) Test for convergence.
(7) $\mathbf{x}(\lambda^*)$ and λ^*.

Algorithm 1.3 gives quadratic convergence and should be stopped as soon as

$$\|f(\lambda^{(k)})\| \leq tol.$$

Example 1.3. *We now apply Algorithm 1.3 to the same matrix in Example 1.2. The results are presented in Table 4. Quadratic convergence is observed in the third column of Table 4 for $k = 5$ and 6.*

TABLE 4. Values of $\lambda^{(k)}$ and $f(\lambda^{(k)})$ of Example 1.2 using the implicit determinant method (Algorithm 1.3). Column 3 show that the results converged quadratically for $k = 5$ and 6. The last column shows the angle between $\mathbf{x}(\lambda^{(k)})$ and $\mathbf{x}(\lambda^*)$ in radians.

k	$\lambda^{(k)}$	$f(\lambda^{(k)})$	$\theta^{(k)}$
0	-1.0000000000000001e-01	Inf	0.0e+00
1	-1.0000000000000001e-01	2.9e-02	2.9e-01
2	-1.2856718325272007e+00	2.6e-02	6.1e-01
3	-1.1834901395004076e+00	4.9e-03	2.5e-01
4	-1.0559116144280840e+00	2.9e-04	4.8e-02
5	-1.0036805192714482e+00	1.1e-06	2.8e-03
6	-1.0000140602287790e+00	1.6e-11	1.0e-05
7	-1.0000000000051936e+00	2.4e-16	0.0e+00

Observe that for a defective eigenvalue, because $f(\lambda^*) = 0$ and $f'(\lambda^*) = 0$, we can use the Gauss-Newton method to solve the combined two nonlinear equations in one unknown. This forms the basis for our next discussion below.

For a defective eigenvalue λ^*, since $f(\lambda^*) = 0$ and $f'(\lambda^*) = 0$ but $f''(\lambda^*) \neq 0$. This shows that we can write the resulting nonlinear system of equations as an over-determined nonlinear system of two equations

$$\mathbf{F}(\lambda) = \begin{bmatrix} f(\lambda) \\ f'(\lambda) \end{bmatrix} = \mathbf{0}, \tag{40}$$

in one real unknown λ. The last equation can be re-written as the following nonlinear least squares problem;

$$\min_{\lambda \in \mathbb{R}} \ \|\mathbf{F}(\lambda)\|,$$

which is the same as minimizing $\frac{1}{2}\mathbf{F}(\lambda)^T\mathbf{F}(\lambda)$. Now, observe that the Jacobian of (40) is

$$\mathbf{F}_\lambda(\lambda) = \begin{bmatrix} f'(\lambda) \\ f''(\lambda) \end{bmatrix}.$$

Using standard Gauss-Newton techniques (see, for example [3]), then we have for $k = 0, 1, 2, \ldots$ that

$$\mathbf{F}_\lambda(\lambda^{(k)})^T\mathbf{F}_\lambda(\lambda^{(k)})\Delta\lambda^{(k)} = -\mathbf{F}_\lambda(\lambda^{(k)})^T\mathbf{F}(\lambda).$$

Therefore,

$$\big(f'^2(\lambda^{(k)}) + f''^2(\lambda^{(k)})\big)\Delta\lambda^{(k)} = -[f'(\lambda^{(k)})f(\lambda^{(k)}) + f''(\lambda^{(k)})f'(\lambda^{(k)})],$$

and

$$\Delta\lambda^{(k)} = -\frac{[f'(\lambda^{(k)})f(\lambda^{(k)}) + f''(\lambda^{(k)})f'(\lambda^{(k)})]}{\big(f'(\lambda^{(k)})^2 + f''(\lambda^{(k)})^2\big)}.$$

Since at the root, $f(\lambda) = f'(\lambda) = 0$ but $f''(\lambda)$ is nonzero, this means that the above expression is valid at the root. Hence, the following algorithm, Algorithm 1.4 can be used to compute an algebraically double and geometrically simple eigenvalue of a defective matrix.

Algorithm 1.4. *Implicit Determinant Method Algorithm for a Defective Matrix*

(1) *Choose* $\mathbf{b}, \mathbf{c}$, $\lambda^{(0)}$, *such that* $\mathbf{M}(\lambda^{(0)})$ *is nonsingular, tol.*
(2) *For* $k = 1, 2, \cdots$, *until convergence*
(3) *Solve* (24) *for* $\mathbf{x}(\lambda^{(k)})$ *and* $f(\lambda^{(k)})$.
(4) *Solve* (26) *for* $\mathbf{x}'(\lambda^{(k)})$ *and* $f'(\lambda^{(k)})$.
(5) *Solve*

$$\begin{bmatrix} (\mathbf{A} - \lambda^{(k)}\mathbf{I}) & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}''(\lambda^{(k)}) \\ f''(\lambda^{(k)}) \end{bmatrix} = \begin{bmatrix} 2\mathbf{x}'(\lambda^{(k)}) \\ 0 \end{bmatrix},$$

for $\mathbf{x}''(\lambda^{(k)})$ *and* $f''(\lambda^{(k)})$.
(6) *Compute*

$$\Delta\lambda^{(k)} = -\frac{[f'(\lambda^{(k)})f(\lambda^{(k)}) + f''(\lambda^{(k)})f'(\lambda^{(k)})]}{\big(f'(\lambda^{(k)})^2 + f''(\lambda^{(k)})^2\big)}.$$

(7) *Update* $\lambda^{(k+1)} = \lambda^{(k)} + \Delta\lambda^{(k)}$.
(8) *Test for convergence.*
(9) *Output* $\mathbf{x}(\lambda^*)$ *and* λ^*.

Stop Algorithm 1.4 as soon as

$$\|\Delta\lambda^{(k)}\| \le tol.$$

Table 5. Values of $\lambda^{(k)}$, $|f(\lambda^{(k)})|$ and $|\Delta\lambda^{(k)}|$ of Example 1.2 using the implicit determinant method and the Gauss-Newton method (Algorithm 1.4). Column 4 show that the results converged quadratically for $k = 4, 5$ and 6.

k	$\lambda^{(k)}$	$\lvert f(\lambda^{(k)})\rvert$	$\lvert\Delta\lambda^{(k)}\rvert$
0	-1.0000000000000001e-01	Inf	Inf
1	-1.0000000000000001e-01	4.2e-02	8.2e-01
2	-9.2060596788524907e-01	5.7e-04	8.9e-02
3	-1.0100862296075745e+00	1.0e-05	9.9e-03
4	-1.0001831090567705e+00	3.3e-09	1.8e-04
5	-1.0000000596542082e+00	1.2e-15	6.0e-08
6	-1.0000000000000111e+00	1.5e-15	6.9e-15
7	-1.0000000000000042e+00	6.2e-16	1.4e-15

Example 1.4. *Table 6 shows the result obtained after applying the inverse iteration algorithm, Algorithm 1.1 with the same starting guesses as those in the implicit determinant method.*

Note that inverse iteration for a matrix with a defective eigenvalue gives $\mathcal{O}(\frac{1}{k})$ convergence as against the quadratic convergence of the implicit determinant method. Moreover, after the $k = 21$ iterate, Matlab gave a warning about the singularity of $(\mathbf{A} - \lambda^*\mathbf{I})$. For a matrix $\mathbf{A}$ in which $\dim\big(\mathcal{N}(\mathbf{A} - \lambda^*\mathbf{I})\big) = n - 1$, the implicit determinant method converge quadratically to a defective eigenvalue while inverse iteration converges much slowly as shown in Tables 3 and 6 respectively.

Remark 1.1. *We remark that if we perturb the Jordan block matrix $\mathbf{J}$ in (34) by adding ε, as in*

$$\mathbf{J} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1+\varepsilon & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 \end{bmatrix}, \tag{41}$$

TABLE 6. Values of $\lambda^{(k)}$ and $|\Delta\lambda^{(k)}|$ of Example 1.1. Column 3 show that the inverse iteration Algorithm 1.1 gave $\mathcal{O}(\frac{1}{k})$ convergence, which is slow compared to the quadratic convergence obtained in column 3 of Table 3. However, Matlab complained of the singularity of $(\mathbf{A} - \lambda^*\mathbf{I})$ for $k = 22$.

k	$\lambda^{(k)}$	$\|\Delta\lambda^{(k)}\|$	$\theta^{(k)}$
0	-9.8999999999999999e-01	Inf	0.0e+00
1	-9.8999999999999999e-01	9.6e-02	5.997142553134e-01
2	-1.0857076510151373e+00	9.7e-02	6.166383316980e-01
3	-9.8858594680621992e-01	5.9e-03	6.177058985192e-01
4	-9.9444440554346691e-01	2.8e-03	6.194551424680e-01
5	-9.9722293199093404e-01	1.4e-03	6.204420440981e-01
6	-9.9861146791050637e-01	6.9e-04	6.209364177328e-01
7	-9.9930573395791322e-01	3.5e-04	6.211836853986e-01
8	-9.9965286698119782e-01	1.7e-04	6.213073389238e-01
9	-9.9982643348181011e-01	8.7e-05	6.213691705903e-01
10	-9.9991321674642697e-01	4.3e-05	6.214000876538e-01
11	-9.9995660836771694e-01	2.2e-05	6.214155464877e-01
12	-9.9997830417135514e-01	1.1e-05	6.214232759786e-01
13	-9.9998915212523509e-01	5.4e-06	6.214271407617e-01
14	-9.9999457603133857e-01	2.7e-06	6.214290731328e-01
15	-9.9999728777802699e-01	1.4e-06	6.214300392461e-01
16	-9.9999864473463207e-01	6.8e-07	6.214305226889e-01
17	-9.9999932267868641e-01	3.4e-07	6.214307642201e-01
18	-9.9999966195557344e-01	1.6e-07	6.214308850943e-01
19	-9.9999981720530673e-01	8.6e-08	6.214309404051e-01
20	-9.9999990344742584e-01	6.5e-08	6.214309711306e-01
21	-9.9999996833868421e-01	8.9e-08	6.214309942494e-01
22	-1.0000000577077797e+00	3.9e-08	6.214310260890e-01

and with $\mathbf{A} = \mathbf{XJX}^{-1}$ as in the last example, then we obtained similar results. We conclude this section that we made use of the same random matrix $\mathbf{X}$ in all the numerical examples of this section.

4. CONCLUDING REMARKS

We compared numerically the rate of convergence of inverse iteration with the implicit determinant method for an algebraically simple eigenvalue and a defective one. Numerical experiments show that for an algebraically simple eigenvalue, both inverse iteration and the implicit determinant method give quadratic convergence. However, for a defective eigenvalue, results show that while two versions of the implicit determinant method show quadratic convergence at a double root, inverse iteration shows $\mathcal{O}(1/k)$ convergence as predicted by Wilkinson [19].

In the defective case, because there is a possibility of stagnation in the implicit determinant method since $f(\lambda^*) = 0$ and $f'(\lambda^*) = 0$, we showed that this short-coming can be overcomed by using the Gauss-Newton method to solve an over-determined nonlinear system of two real equations in one real unknown.

ACKNOWLEDGEMENTS

R. O. Akinola acknowledges financial support from the University of Bath, United Kingdom, in the form of a University Research Studentship from April 2007 to June 2010. The authors would like to thank three anonymous referees whose comments improved the original version of this manuscript.

Appendix A. Quadratic Convergence of Newton's Method

Since the implicit determinant method is an application of Newton's method, the aim of this appendix is to prove the quadratic convergence of Newton's method for a system of nonlinear equations. The materials in this section can be found in [3]. To prove that Newton's method converged quadratically, we need some preliminary results. We start with the following definition.

Definition A.1. : *The Jacobian matrix* $\mathbf{F_x} \in Lip_\gamma(B(\mathbf{x}^{(0)}, r))$ *if there exists a* $\gamma > 0$ *such that for all* $\mathbf{x}, \mathbf{y} \in B(\mathbf{x}^{(0)}, r)$,

$$\|\mathbf{F_x}(\mathbf{y}) - \mathbf{F_x}(\mathbf{x})\| \le \gamma \|\mathbf{y} - \mathbf{x}\|. \tag{42}$$

Consider the problem of solving the following nonlinear system of equations $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, where $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$. Let $\mathbf{\Phi}(t) : \mathbb{R} \to \mathbb{R}^n$ be defined as

$$\mathbf{\Phi}(t) = \mathbf{F}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})),$$

and

$$\mathbf{F}(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) = [f_1(\mathbf{x}+t(\mathbf{y}-\mathbf{x})), f_2(\mathbf{x}+t(\mathbf{y}-\mathbf{x})), \ldots, f_n(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))]^T.$$

Let $\mathbf{\Phi}_i(t)$ denote the ith component of $\mathbf{\Phi}(t)$ for $i = 1, 2, \ldots, n$, then by the Chain rule

$$\begin{aligned}\mathbf{\Phi}_i'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\big\{f_i(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))\big\} \\ &= \frac{\partial}{\partial x_1} f_i(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))(y_1-x_1)+\ldots \\ &+ \frac{\partial}{\partial x_n} f_i(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))(y_n-x_n) \\ &= \sum_{j=1}^{n} \frac{\partial}{\partial x_j} f_i(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))(y_j-x_j) \\ &= \begin{bmatrix} \frac{\partial}{\partial x_1} f_i(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) & \ldots & \frac{\partial}{\partial x_n} f_i(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) \end{bmatrix} \times \\ &\quad [(y_1-x_1),(y_2-x_2),\ldots,(y_n-x_n)]^T.\end{aligned}$$

By the definition of $\mathbf{F}$, we have

$$\mathbf{\Phi}'(t) = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) & \ldots & \frac{\partial}{\partial x_n} f_1(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) \\ \frac{\partial}{\partial x_1} f_2(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) & \ldots & \frac{\partial}{\partial x_n} f_2(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) \\ \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_1} f_n(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) & \ldots & \frac{\partial}{\partial x_n} f_n(\mathbf{x}+t(\mathbf{y}-\mathbf{x})) \end{bmatrix} \times$$
$$[(y_1-x_1),(y_2-x_2),\ldots,(y_n-x_n)]^T. \tag{43}$$

Now,

$$\begin{aligned}\mathbf{F}(\mathbf{y})-\mathbf{F}(\mathbf{x}) &= \mathbf{\Phi}(1)-\mathbf{\Phi}(0) \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\big\{\mathbf{\Phi}(t)\big\}\mathrm{d}t \qquad (44) \\ &= \left\{\int_0^1 \mathbf{F}_{\mathbf{x}}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))\mathrm{d}t\right\}(\mathbf{y}-\mathbf{x}),\end{aligned}$$

hence, $\mathbf{F}(\mathbf{y}) = \mathbf{F}(\mathbf{x}) + \int_0^1 \mathbf{F}_{\mathbf{x}}(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))\mathrm{d}t(\mathbf{y}-\mathbf{x})$. Assuming we have a good guess $\mathbf{x}^{(0)}$ to a solution $\mathbf{x}^*$ of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, then by replacing $\mathbf{y}$ by $\mathbf{x}^*$ and $\mathbf{x}$ by $\mathbf{x}^{(0)}$ in (44), we obtain

$$\mathbf{0} = \mathbf{F}(\mathbf{x}^*) = \mathbf{F}(\mathbf{x}^{(0)}) + \int_0^1 \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)}+t(\mathbf{x}^*-\mathbf{x}^{(0)}))\mathrm{d}t(\mathbf{x}^*-\mathbf{x}^{(0)}).$$

If the integral in the above expression is approximated by

$$\int_0^1 \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)}+t(\mathbf{x}^*-\mathbf{x}^{(0)}))\mathrm{d}t \approx \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)}),$$

then we have

$$\mathbf{0} \approx \mathbf{F}(\mathbf{x}^{(0)}) + \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})(\mathbf{x}^* - \mathbf{x}^{(0)}).$$

So that if we define $\Delta\mathbf{x}^{(0)} = \mathbf{x}^{(1)} - \mathbf{x}^{(0)}$, then,

$$\mathbf{F}(\mathbf{x}^{(0)}) + \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = \mathbf{0} \quad \text{or} \quad \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = -\mathbf{F}(\mathbf{x}^{(0)}).$$

This means that by iterating the above for $k = 1, 2, \ldots$, yields

$$\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} = -\mathbf{F}(\mathbf{x}^{(k)}), \tag{45}$$

which is a linear system of equations. Thus, we have linearized the nonlinear system of equations $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ into (45). Therefore, if the Jacobian matrix $\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})$ is nonsingular, then

$$\Delta\mathbf{x}^{(k)} = -\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})^{-1}\mathbf{F}(\mathbf{x}^{(k)}). \tag{46}$$

The above expression for $\Delta\mathbf{x}^{(k)}$ and the solution to the nonlinear system of equations $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ for $k = 0, 1, 2, \ldots$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}. \tag{47}$$

is called Newton's method (see, for example, [4]).

The following n-dimensional version of the integral Mean Value Theorem is given below.

Lemma A.1. : *Let $\mathbf{x}^{^{(0)}} \in \mathbb{R}^n$, $r > 0$ and $\mathbf{F}$ is continuously differentiable on $B(\mathbf{x}^{^{(0)}}, r)$. Then for all $\mathbf{x}, \mathbf{y} \in B(\mathbf{x}^{^{(0)}}, r)$,*

$$\mathbf{F}(\mathbf{y}) = \mathbf{F}(\mathbf{x}) + \left\{ \int_0^1 \mathbf{F}_{\mathbf{x}}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\mathrm{d}t \right\} (\mathbf{y} - \mathbf{x}).$$

Proof: Let $\mathbf{\Phi} : \mathbb{R} \mapsto \mathbb{R}^n$ be defined as $\mathbf{\Phi}(t) := \mathbf{F}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$. If $\mathbf{\Phi}_i(t)$ is as defined above and using the result of (43), then $\mathbf{\Phi}_i'(t) = \mathbf{F}_{\mathbf{x}}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})$. Using equation (44), the result holds. ■

The following corrolary is an immediate consequence of the last result.

Corollary A.1. :

$$\|\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x})\| \leq \max_{\mathbf{z} = \mathbf{x} + t(\mathbf{y} - \mathbf{x})} \|\mathbf{F}_{\mathbf{z}}(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\|.$$

We state the following well known result without a proof.

Theorem A.1. *If $\mathbf{A}$ and $\mathbf{C}$ are square nonsingular matrices and*

$$\|\mathbf{C} - \mathbf{A}\| < \frac{1}{\|\mathbf{A}^{-1}\|},$$

then

$$\|\mathbf{C}^{-1}\| \le \frac{\|\mathbf{A}^{-1}\|}{1 - \|\mathbf{C} - \mathbf{A}\|\|\mathbf{A}^{-1}\|}. \tag{48}$$

The following lemma, contains two important results that will be used in the final proof for the convergence of Newton's method.

Lemma A.2. : *Let* $\mathbf{F}$ *be a continuously differentiable vector-valued function on* $B(\mathbf{x}^{(0)}, r)$ *and* $\mathbf{F}_{\mathbf{x}} \in Lip_{\gamma}(B(\mathbf{x}^{(0)}, r))$ *and* $\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})$ *is nonsingular such that*

$$\|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\| = \beta.$$

Hence,

(a). for all $\mathbf{x}, \mathbf{y} \in B(\mathbf{x}^{(0)}, r)$.

$$\|\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x}) - \mathbf{F}_{\mathbf{x}}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| \le \frac{\gamma}{2}\|\mathbf{y} - \mathbf{x}\|^2.$$

(b). For all $\mathbf{x} \in B(\mathbf{x}^{(0)}, \varepsilon)$ *and with* $\varepsilon := \min\left\{r, \frac{1}{2\beta\gamma}\right\}$

$$\|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\| \le 2\beta. \tag{49}$$

Proof:

(a). Let $\mathbf{x}, \mathbf{y} \in B(\mathbf{x}^{(0)}, r)$ and $\boldsymbol{\omega} = \mathbf{F}_{\mathbf{x}}\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big)$. Then,

$$\begin{aligned}
\|\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x}) - \mathbf{F}_{\mathbf{x}}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| &= \| \int_0^1 \big\{\boldsymbol{\omega} - \mathbf{F}_{\mathbf{x}}(\mathbf{x})\big\}(\mathbf{y} - \mathbf{x})\mathrm{d}t\| \\
&\le \int_0^1 \|\big\{\boldsymbol{\omega} - \mathbf{F}_{\mathbf{x}}(\mathbf{x})\big\}(\mathbf{y} - \mathbf{x})\|\mathrm{d}t \\
&\le \int_0^1 \|\big\{\boldsymbol{\omega} - \mathbf{F}_{\mathbf{x}}(\mathbf{x})\big\}\|\|\mathbf{y} - \mathbf{x}\|\mathrm{d}t \\
&\le \gamma\|\mathbf{y} - \mathbf{x}\|^2 \int_0^1 t\mathrm{d}t \\
&= \frac{\gamma}{2}\|\mathbf{y} - \mathbf{x}\|^2.
\end{aligned} \tag{50}$$

(b). Let $\mathbf{x} \in B(\mathbf{x}^{(0)}, \varepsilon) \subseteq B(\mathbf{x}^{(0)}, r)$, then

$$\begin{aligned}
\|\mathbf{F}_{\mathbf{x}}(\mathbf{x}) - \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})\| &\leq \gamma\|\mathbf{x} - \mathbf{x}^{(0)}\| \\
&< \gamma\varepsilon \\
&\leq \frac{\gamma}{2\beta\gamma} \\
&= \frac{1}{2\beta} \\
&\leq \frac{1}{2\|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\|}.
\end{aligned}$$

Using the result of Theorem A.1 and the nonsingularity of $\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})$,

$$\begin{aligned}
\|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\| &\leq \frac{\|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\|}{1 - \|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\|\|\mathbf{F}_{\mathbf{x}}(\mathbf{x}) - \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})\|} \\
&< \frac{\beta}{1 - \frac{1}{2}} \\
&= 2\beta.
\end{aligned}$$

Hence the proof.

■

The above results now paves the way for the proof of the quadratic convergence of Newton's method.

Theorem A.2. *Assuming there exists a solution* $\mathbf{x}^* \in \mathbb{R}^n$ *with* $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$ *and the conditions of Lemma A.2 are satisfied at* $\mathbf{x}^{(0)} = \mathbf{x}^*$*. By setting* $\varepsilon = \min\left\{r, \frac{1}{2\beta\gamma}\right\}$ *and provided* $\mathbf{x}^{(0)} \in B(\mathbf{x}^*, \varepsilon)$*, then the following results holds:*

(a). $[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})]^{-1}$ *exists, that is, Newton's method is well defined.*
(b). For all $k \geq 0$*,* $\mathbf{x}^{(k)} \in B(\mathbf{x}^*, \varepsilon)$
(c). For all $k \geq 0$*,* $\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| \leq \beta\gamma\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2$
(d). $\mathbf{x}^{(k)}$ *converges to* $\mathbf{x}^*$ *quadratically.*

Proof:

(a).&(b). The proofs of (a) and (b) will be by induction. For $k = 0$, $\|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(0)})]^{-1}\| < 2\beta$ by Lemma A.2. Hence, (a) holds. Similarly, because by choice $\mathbf{x}^{(0)} \in B(\mathbf{x}^*, \varepsilon)$ and so (b) also holds. Now, suppose that both (a) and (b) holds for some $k \in \mathbb{N}$, then using the second part of Lemma A.2, this means $[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})]^{-1}$ exists and $\mathbf{x}^{(k+1)}$ is well defined. Thus, (a) holds for $n = k+1$. Next, we want to show that $\mathbf{x}^{(k+1)} \in B(\mathbf{x}^*, \varepsilon)$. Now,

$$\begin{aligned}
\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| &= \|\mathbf{x}^{(k)} - \mathbf{x}^* - [\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})]^{-1}\mathbf{F}(\mathbf{x}^{(k)})\| \\
&= \|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})]^{-1}\big\{\mathbf{F}(\mathbf{x}^*) - \mathbf{F}(\mathbf{x}^{(k)}) - \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})(\mathbf{x}^{(k)} - \mathbf{x}^*)\big\}\| \\
&\leq \|[\mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})]^{-1}\|\|\mathbf{F}(\mathbf{x}^*) - \mathbf{F}(\mathbf{x}^{(k)}) - \mathbf{F}_{\mathbf{x}}(\mathbf{x}^{(k)})(\mathbf{x}^{(k)} - \mathbf{x}^*)\| \\
&\leq 2\beta \times \frac{\gamma}{2}\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2 \\
&= \beta\gamma\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2 \qquad (51) \\
&\leq (\beta\gamma\|\mathbf{x}^{(k)} - \mathbf{x}^*\|)\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \\
&\leq \frac{1}{2}\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \ (\text{since}\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \min\left\{r, \frac{1}{2\beta\gamma}\right\}).
\end{aligned}$$

Hence, $\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| \leq \frac{1}{2}\|\mathbf{x}^{(k)} - \mathbf{x}^*\|$. Therefore, $\mathbf{x}^{(k+1)} \in B(\mathbf{x}^*, \varepsilon)$ and (b) holds for all $n \in \mathbb{N}$

(c). This is proved from (51).

(d). Using $\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| \leq \frac{1}{2}\|\mathbf{x}^{(k)} - \mathbf{x}^*\|$, we have

$$\begin{aligned}
\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| &\leq \frac{1}{2}\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \\
&\leq \left(\frac{1}{2}\right)^{(k+1)}\|\mathbf{x}^{(0)} - \mathbf{x}^*\|.
\end{aligned}$$

Therefore, $\mathbf{x}^{(k)}$ converges to $\mathbf{x}^*$ as k tends to infinity. Quadratic convergence is obvious from (51) ■

References

[1] R. O. Akinola, *Numerical Solution of Linear and Nonlinear Eigenvalue Problems*, Ph.D. thesis, University of Bath, 2010.

[2] R. Alam, and S. Bora, *On Sensitivity of Eigenvalues and Eigendecompositions of Matrices*, Linear Algebra and its Applications **396** (2005), 273–301.

[3] J. E. Dennis, Jr, and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Classics In Applied Mathematics, no. 16, SIAM Philadelphia, 1996.

[4] P. Deuflhard, *Newton Methods for Nonlinear Problems*, ch. 4, pp. 174–175, Springer, 2004.

[5] M. A. Freitag, and A. Spence, *Convergence of Inexact Inverse Iteration with Application to Preconditioned Iterative Solves*, BIT Numerical Mathematics **47** (2006), 27–44.

[6] M. A. Freitag, and A. Spence, *A Newton-based method for the calculation of the distance to instability*, Linear Alg. Appl., no. 435(12) , 15 December 2011, pp. 3189-3205.

[7] G. H. Golub, and C. F. van Loan, *Matrix Computations*, 3rd ed., The John Hopkins University Press, London, 1996.
[8] I. C. F. Ipsen, *Computing an Eigenvector with Inverse Iteration*, SIAM Review **39** (1997), no. 2, 254–291.
[9] A. Jeffrey, *Mathematics for Engineers and Scientists*, Nelson, 1969.
[10] H. B. Keller, *Numerical Solution of Bifurcation and Nonlinear Eigenvalue Problems*, Applications of Bifurcation Theory (in : P. Rabinowitz, ed.), Academic Press, New York, 1977, pp. 359–384.
[11] B. Noble, *Applied Linear Algebra*, ch. 5, pp. 143–145, Prentice-Hall, Inc., 1969.
[12] B. Noble, and J. W. Daniel, *Applied Linear Algebra*, third ed., ch. 9, pp. 355–397, Prentice-Hall, 1988.
[13] A. Spence, and C. Poulton, *Photonic Band Structure Calculations using Nonlinear Eigenvalue Techniques*, Journal of Computational Physics **204** (2005), 65 – 81.
[14] A. Spence, and I. G. Graham, *The Graduate Student's Guide to Numerical Analysis '98*, Lecture Notes from the VIII EPSRC Summer School in Numerical Analysis 3, pp. 176–216, Springer, 1998.
[15] L. N. Trefethen, and D. Bau III, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
[16] H. Wielandt, *Das Iterationsverfahren bei nicht Selbstadjungierten Linearen Eigenwertaufgaben*, Mathematische Zeitschrift (1944), 93–143.
[17] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Ely House, London W., 1965.
[18] J. H. Wilkinson, *Inverse Iteration in Theory and in Practice*, Istituto Nazionale di Alta Mathematica, Symposia Mathematica *X* (1972), 361–379.
[19] J. H. Wilkinson, *Note on Matrices with a very Ill-Conditioned Eigenproblem*, Numerische Mathematik **19** (1972), 176–178.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF JOS, NIGERIA

E-mail addresses: roakinola@yahoo.com, akinolar@unijos.edu.ng

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF BATH, UNITED KINGDOM

E-mail address: a.spence@bath.ac.uk