International Journal of Innovation in Science and Mathematics
Volume 2, Issue 5, ISSN (Online): 2347–9051

Two-Norm Normalization for the Matrix Pencil: Inverse Iteration with a Complex Shift

R. O. Akinola
Department of Mathematics,
University of Jos, P. M. B. 2084, Nigeria
Email: roakinola@yahoo.com

A. Spence
Department of Mathematical Sciences,
University of Bath, BA2 7AY, UK.

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. For complex eigenpairs, instead of using Ruhe's normalization, we show that the natural two norm normalization for the matrix pencil, yields a quadratically convergent algorithm. Numerical experiment is given which confirms the theory.**

Keywords **– Eigenvalue, Defective, Quadratic Convergence.**

I. INTRODUCTION

Let A be a large sparse, real n by n nonsymmetric matrix and $\mathbf{B} \in \mathbb{R}^{n\times n}$ a symmetric positive definite matrix. In this paper, we consider the problem of computing the eigenpair ($\mathbf{z}$, λ) from the following generalised complex eigenvalue problem

$$\mathbf{Az} = \lambda\mathbf{Bz}, \quad \mathbf{z} \in \mathbb{C}^n, \ \mathbf{z} \neq \mathbf{0}, \qquad (1)$$

where $\lambda \in \mathbb{C}$ is the eigenvalue of the pencil ($\mathbf{A}$, $\mathbf{B}$) and $\mathbf{z}$ its corresponding complex eigenvector. We assume that the eigenpair of interest ($\mathbf{z}$, λ) is algebraically simple, so that ψ^H the corresponding left eigenvector is such that [9, p. 136]

$$\boldsymbol{\psi}^H\mathbf{Bz} \neq 0. \qquad (2)$$

By adding the normalization

$$\mathbf{z}^H\mathbf{Bz} = 1, \qquad (3)$$

to (1) and $v = \left[z^T, \lambda\right]$ with $\mathbf{v}$ = [$\mathbf{z}$T, l], the combined system of equations can be expressed in the form $\mathbf{F}(\mathbf{v}) = \mathbf{0}$ as

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B})\mathbf{z} \\ -\frac{1}{2}\mathbf{z}^H\mathbf{Bz} + \frac{1}{2} \end{bmatrix} = \mathbf{0}. \qquad (4)$$

Note that $\mathbf{z}^H\mathbf{Bz}$ is real since $\mathbf{B}$ is symmetric and positive definite. This results in solving a system of n complex and one real nonlinear equation for the (n +1) complex unknowns $v = \left[z, \lambda\right]^T$. Note that, if $\mathbf{z}$ from ($\mathbf{z}$, λ) solves (4), then so does $e^{i\theta}z$ for any $\theta \in [0, 2\pi)$. Hence, (4) does not have a unique solution. Another drawback of the normalisation (3) is that $\bar{z}$ in $\mathbf{z}^H\mathbf{Bz} = \bar{\mathbf{z}}^T\mathbf{Bz}$ is not differentiable[1]. Therefore, we cannot just differentiate (4) and apply the standard Newton's method. In this article, we shall show how these drawbacks can be overcome, at least for the B = I case.

Recall that for a real eigenpair ($\mathbf{z}$, λ), (4) gives (n +1) real equations for (n + 1) real unknowns and Newton's method for solving (4) involves the solution of the (n + 1) square linear systems

$$\begin{bmatrix} \mathbf{A} - \lambda^{(k)}\mathbf{B} & -\mathbf{Bz}^{(k)} \\ -(\mathbf{Bz}^{(k)})^T & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{z}^{(k)} \\ \Delta\lambda^{(k)} \end{bmatrix} = - \begin{bmatrix} (\mathbf{A} - \lambda^{(k)}\mathbf{B})\mathbf{z}^{(k)} \\ -\frac{1}{2}\mathbf{z}^{(k)^T}\mathbf{Bz}^{(k)} + \frac{1}{2} \end{bmatrix}$$

for the (n + 1) real unknowns $\Delta v^{(k)} = \left[\Delta z^{(k)^T}, D\lambda^{(k)}\right]$, and updating $v^{(k+1)} = v^{(k)} + \Delta v^{(k)}$ for k = 0, 1, 2, ...

Secondly, for ($\mathbf{z}$, λ) complex, Ruhe [7] added the normalization $\mathbf{c}^H\mathbf{z}$ = 1, where $\mathbf{c}$ is a fixed complex vector instead of (3), so that (1) and $\mathbf{c}^H\mathbf{z}$ = 1 provide (n + 1) complex equations for (n + 1) complex unknowns, and the Jacobian of this system is

$$\begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B}) & -\mathbf{Bz} \\ \mathbf{c}^H & 0 \end{bmatrix}$$

The above Jacobian is square and can be easily shown to be nonsingular, using the ABCD Lemma if the eigenvalue of interest is algebraically simple and c^Hz 6= 0 at the root. One major distinction between our normalization and Ruhe's is that, ours is the natural normalization for an eigenvector and we do not worry about how to choose $\mathbf{c}$.

Parlett and Saad in [6], studied inverse iteration with a complex shift $\sigma = \alpha + i\beta$ where a and b are real. They showed that by replacing the shifted complex system $(A - \sigma B)\phi = B\varphi$, with a real one, the size of the problem is doubled, where $\varphi = \varphi_1 + i\varphi_2, \phi = \phi_1 + i\phi_2$ for $\varphi_1, \varphi_2, \phi_1, \phi_2 \in R^n$ and $i = \sqrt{-1}$ is the imaginary unit of a complex number. This is because solving a complex linear system of equations takes twice the storage and is roughly three times the cost of solving a real system [4]. When real arithmetic rather than complex arithmetic is used, we lose any band structure in $\mathbf{A}$ and $\mathbf{B}$ [6]. The numerical examples in [6], show linear convergence to the eigenvalue closest to the fixed shift.

Next, Tisseur in [8] considered the symmetric definite generalised eigenvalue problem $A\phi = \lambda B\phi, \lambda \in R$ as a special case of (1) where $\mathbf{A}$ is symmetric and $\mathbf{B}$ is symmetric positive definite but with the real normalization

$$\tau\mathbf{e}_s^T\boldsymbol{\phi} = \tau; \quad \text{for some fixed s,} \qquad (6)$$

where $\tau = \max(\|\mathbf{A}\|, \|\mathbf{B}\|)$, (see, for example, [8, p. 1049]) and $\mathbf{e}_j$ is the jth column of the identity matrix.

The real scalar t is introduced to scale $\mathbf{F}(\mathbf{w})$ and $\mathbf{F}_\mathbf{w}(\mathbf{w})$ when $\mathbf{A}$ and[1] $\mathbf{B}$ are multiplied by a scalar. In this case,

[1]For a single variable, if z = x + iy, $\bar{z}$ = x – iy, then the Cauchy-Riemann equations are not satisfied because, with u(x, y) = x, v(x, y) = –y, then u_x(x, y) = 1 and v_y(x, y) = –1, whereas the Cauchy-Riemann equations (see, for example [3]) require that u_x(x, y) = v_y(x, y). This shows that $\bar{z}$ is not differentiable at (x, y).

$$\mathbf{F}(\mathbf{w}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B})\boldsymbol{\phi} \\ \tau \mathbf{e}_s^T \boldsymbol{\phi} - \tau \end{bmatrix},$$

and

$$\mathbf{F}_{\mathbf{w}}(\mathbf{w}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B}) & -\mathbf{B}\boldsymbol{\phi} \\ \tau \mathbf{e}_s^T & 0 \end{bmatrix}$$

Tisseur [8], showed that the Jacobian $\mathbf{F}_{\mathbf{w}}(\mathbf{w})$ above is singular at the root if and only if $l\_$ is a finite multiple eigenvalue of the pencil (**A**,**B**). The main result in [8] is Theorem 2.4 [8, pp. 1044-1046]. It shows that if the linear system to be solved is not too ill conditioned, the solver is not completely unstable, the Jacobian is approximated accurately enough and we have a good initial guess very close to the solution, then the norm of the residual reduces after one step of Newton's method in floating point arithmetic. The main point is that both [7] and [8] used two different differentiable normalisations, while in this paper we analyse the natural extension of the distance norm, which is a non differentiable normalization and so leads to interesting theoretical questions.

In addition, it was shown numerically in [8, pp. 1053-1054] that if Newton's method is applied in floating point arithmetic with mixed precision iterative refinement, the linear solver is unstable and there are inaccuracies in computing the Jacobian, then this may affect the rate of convergence of Newton's method but not the accuracy and stability of the computed eigenvalues.

Our approach for analysing the solution of (4) for **v** begins by splitting the eigenpair (**z**, λ) into their real and imaginary parts: $\mathbf{z} = \mathbf{z}_1 + \mathrm{i}\mathbf{z}_2$, $\lambda = \alpha + \mathrm{i}\beta$ where $z_1, z_2 \in R^n$, and $\alpha, \beta \in R$. After expanding (4), we obtain a real system of $(2n + 1)$ under-determined nonlinear equations in $(2n + 2)$ real unknowns $\mathbf{v} = [\mathbf{z}_1, \mathbf{z}_2, \alpha, \beta]^{\mathrm{T}}$, and it is natural to use the Gauss-Newton method (see, for example, Deuflhard [**?**, pp. 222-223]) to obtain a solution. By linearising the system of under-determined nonlinear equations, we obtain a system of under-determined linear equations involving the corresponding Jacobian. The key result in this paper is Theorem 2.1.

This article is structured as follows. In Section II, we show that for an algebraically simple eigenvalue, the Jacobian is of full rank at the root with a known nullvector. A numerical example is given in Section III which confirms the validity of the theory. Throughout this paper, $\|.\| = \|.\|2$.

II. Computation of Complex Eigenpairs by Solving an Under-Determined System of Nonlinear Equations

In this section, we will expand the system of n complex and one real nonlinear equations in (n + 1) complex unknowns (4) by writing **z** and λ as $\mathbf{z} = \mathbf{z}_1 + \mathrm{i}\mathbf{z}_2$, and $\lambda = \alpha + \mathrm{i}\beta$, respectively. The reason for having an under-determined system of equations instead of a square system of equations is because, expanding $\mathbf{z}^H\mathbf{B}\mathbf{z} = 1$ gives only one real equation, since **B** is symmetric positive definite, while $(\mathbf{A}\text{-}\lambda\mathbf{B})\mathbf{z} = \mathbf{0}$ results in 2n real equations. This results in a real (2n+1) underdetermined system of nonlinear equations in (2n + 2) real unknowns. This will then be followed by presenting the real under-determined system of nonlinear equations and an explicit expression for its Jacobian. Furthermore, we will show in the main result of this paper-Theorem 2.1 that, if the eigenvalue of interest in (**A**,**B**) is algebraically simple, then the Jacobian has linearly independent rows at the root. We will find the right nullvector of the Jacobian at the root. We conclude the section by presenting an algorithm for computing the complex eigenpair of the matrix pencil (**A**, **B**).

If we let $\mathbf{z} = \mathbf{z}_1 + \mathrm{i}\mathbf{z}_2$, $\lambda = \alpha + \mathrm{i}\beta$, then the nonlinear system of equations (4) can be written as

$$\begin{aligned}(\mathbf{A} - \lambda\mathbf{B})\mathbf{z} &= [\mathbf{A} - (\alpha + i\beta)\mathbf{B}](\mathbf{z}_1 + i\mathbf{z}_2) \\ &= (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 + i[(\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1],\end{aligned} \tag{7}$$

and

$$\mathbf{z}^H\mathbf{B}\mathbf{z} = \mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2. \tag{8}$$

Hence, (3) implies that

$$-\frac{1}{2}\mathbf{z}^H\mathbf{B}\mathbf{z} + \frac{1}{2} = -\frac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \frac{1}{2} = \mathbf{0}.$$

Since (A – λB)z = 0, we equate the real and imaginary parts of (7) to zero and obtain the 2n real equations

$$(\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 = \mathbf{0},$$

and

$$(\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 = \mathbf{0}$$

This means, F(v) consists of the 2n real equations arising from (7) and one real equation

$$-\tfrac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \tfrac{1}{2} = 0;$$

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 \\ -\beta\mathbf{B}\mathbf{z}_1 + (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 \\ -\frac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \frac{1}{2} \end{bmatrix} = \mathbf{0}, \tag{9}$$

where $\mathbf{F} : \mathbb{R}^{(2n+2)} \to \mathbb{R}^{(2n+1)}$. The Jacobian, $\mathrm{F_v}(\mathrm{v})$ of F(v) with $v = [z_1, z_2 \alpha.\beta]^T$ has the following explicit expression

$$\mathbf{F}_{\mathrm{v}}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} & -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) & -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \\ -(\mathbf{B}\mathbf{z}_1)^T & -(\mathbf{B}\mathbf{z}_2)^T & 0 & 0 \end{bmatrix}, \tag{10}$$

and is a (2n + 1) by (2n + 2) real matrix. We define the real 2n by 2n matrix M as

$$\mathbf{M} = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) \end{bmatrix}. \tag{11}$$

Also, we form the 2n by 2 real matrix

$$\mathbf{N} = \begin{bmatrix} -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \end{bmatrix} = [-\mathbf{B}_2\mathbf{w} \quad \mathbf{B}_2\mathbf{w}_1], \tag{12}$$

consisting of the product of $\mathbf{B}_2 = \begin{bmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{bmatrix}$ and the matrix of right nullvectors (given in the next equation) of M at the root, where

$$\mathbf{w} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \qquad \mathbf{w}_1 = \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix}, \tag{13}$$

and O is the n by n zero matrix. The Jacobian (10) can be rewritten in the following partitioned form

International Journal of Innovation in Science and Mathematics
Volume 2, Issue 5, ISSN (Online): 2347–9051

$$\mathbf{F_v(v)} = \begin{bmatrix} \mathbf{M} & -\mathbf{B_2 w} & \mathbf{B_2 w_1} \\ -(\mathbf{B_2 w})^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{N} \\ -(\mathbf{B_2 w})^T & \mathbf{0}^T \end{bmatrix}, \quad (14)$$

with M, N defined in (11) and (12) respectively. Note that because at the root,

$$\begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A}-\alpha\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 \\ (\mathbf{A}-\alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 \end{bmatrix} = \mathbf{0},$$

this implies that $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ or its nonzero scalar multiple is a right nullvector of M. In the same vein, we find

$$\begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A}-\alpha\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix} = \begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 \\ -\{(\mathbf{A}-\alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2\} \end{bmatrix} = \mathbf{0},$$

and $\begin{bmatrix} z_2 \\ -z_1 \end{bmatrix}$ or its nonzero scalar multiple is also a right nullvector of **M** at the root.

Since the eigenvalue λ of (**A**,**B**) is algebraically simple by assumption, then by (2), we need to give explicit expressions for the left nullvector of (**A** – λ**B**) in order to prove that the Jacobian has full row rank at the root.

Observe that for all $\psi \in \mathcal{N}(\mathbf{A}-\lambda\mathbf{B})^H \backslash \{0\},$, we define $\psi = \psi_1 + i\psi_2,$, where $\psi_1, \psi_2 \in R^n,$, then this implies

$$\psi^H(\mathbf{A}-\lambda\mathbf{B}) = (\psi_1^T - i\psi_2^T)[(\mathbf{A}-\alpha\mathbf{B}) - i\beta\mathbf{B}]$$

$$= \psi_1^T(\mathbf{A}-\alpha\mathbf{B}) - \beta\psi_2^T\mathbf{B} - i[\beta\psi_1^T\mathbf{B} + \psi_2^T(\mathbf{A}-\alpha\mathbf{B})] = \mathbf{0}^T.$$

Hence, $\psi_1^T(\mathbf{A}-\alpha\mathbf{B}) - \beta\psi_2^T\mathbf{B} = \mathbf{0}^T$ and $\beta\psi_1^T\mathbf{B} + \psi_2^T(\mathbf{A}-\alpha\mathbf{B}) = \mathbf{0}^T.$ The implication of this is that

$$[\psi_1^T \quad \psi_2^T]\mathbf{M} = [\psi_1^T \quad \psi_2^T] \begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A}-\alpha\mathbf{B}) \end{bmatrix}$$

$$= [\psi_1^T(\mathbf{A}-\alpha\mathbf{B}) - \beta\psi_2^T\mathbf{B} \qquad \beta\psi_1^T\mathbf{B} + \psi_2^T(\mathbf{A}-\alpha\mathbf{B})] = \mathbf{0}^T,$$

which means, $[\psi_1^T, \psi_2^T]$or its nonzero scalar multiple is a left null-vector of **M**. Similarly,

$$[\psi_2^T \quad -\psi_1^T]\mathbf{M} = [\psi_2^T \quad -\psi_1^T] \begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A}-\alpha\mathbf{B}) \end{bmatrix}$$

$$= [\beta\psi_1^T\mathbf{B} + \psi_2^T(\mathbf{A}-\alpha\mathbf{B}) \qquad -\{\psi_1^T(\mathbf{A}-\alpha\mathbf{B}) - \beta\psi_2^T\mathbf{B}\}] = \mathbf{0}^T,$$

and it shows that $[\psi_2^T, \ -\psi_1^T]$is also a left nullvector of **M**. So we form the matrix **C** consisting of the 2-dimensional left nullvectors of **M** at the root (in practice **C** is not computed), as

$$\mathbf{C} = \begin{bmatrix} \psi_1 & \psi_2 \\ \psi_2 & -\psi_1 \end{bmatrix}. \quad (15)$$

Now, observe that the condition (2), implies

$$\psi^H\mathbf{B}\mathbf{z} = [\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2] + i[\psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1] \neq 0.$$

Since $\psi^H\mathbf{B}\mathbf{z} \neq 0,$ this implies

$$[\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2]^2 + [\psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1]^2 \neq 0. \quad (16)$$

Before we continue with the rest of the analysis, wepresent the main result of this section which shows that the Jacobian (10) has a one dimensional nullvector at the root.

Theorem 2.1: Assume that the eigenpair (z, λ) of the pencil (A,B) is algebraically simple. If $\mathbf{z}_1$ and $\mathbf{z}_2$ are nonzero vectors, then $\phi = \{\tau[\mathbf{z}_2^T, -\mathbf{z}_1^T, 0, 0], \tau \in \mathbb{R}\}$ is the eigenspace corresponding to the zero eigenvalue of $\mathbf{F_v(v)}$ at the root.

Proof. Post-multiply $\mathbf{F_v(v)}$ by the unknown nonzero vector $\phi = [\mathbf{p}', \mathbf{q}']^T,$ equate to the zero vector and solve

$$\begin{bmatrix} \mathbf{M} & \mathbf{N} \\ -(\mathbf{B_2 w})^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{p}' \\ \mathbf{q}' \end{bmatrix} = \mathbf{0},$$

where **M** and **N** are as defined in (11) and (12) respectively. After expanding, we have the following set of equations

$$\mathbf{M}\mathbf{p}' + \mathbf{N}\mathbf{q}' = \mathbf{0} \quad (17)$$

$$\mathbf{w}^T\mathbf{B}_2\mathbf{p}' = 0. \quad (18)$$

Let $\mathbf{H} = \mathbf{C}^T\mathbf{N}$, for all $\mathbf{C} \in \mathcal{N}(\mathbf{M}^T)\backslash\{\mathbf{0}\}$ as in (15). This means,

$$\mathbf{H} = \begin{bmatrix} \psi_1^T & \psi_2^T \\ \psi_2^T & -\psi_1^T \end{bmatrix} \begin{bmatrix} \mathbf{B} & \\ & \mathbf{B} \end{bmatrix} \begin{bmatrix} -\mathbf{z}_1 & \mathbf{z}_2 \\ -\mathbf{z}_2 & -\mathbf{z}_1 \end{bmatrix}$$

$$= \begin{bmatrix} \psi_1^T & \psi_2^T \\ \psi_2^T & -\psi_1^T \end{bmatrix} \begin{bmatrix} -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \end{bmatrix}$$

$$= \begin{bmatrix} -(\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) & \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 \\ \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 & (\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) \end{bmatrix}.$$

By premultiplying both sides of (17) by $\mathbf{C}^T$, we obtain

$$\mathbf{C}^T\mathbf{M}\mathbf{p}' + \mathbf{C}^T\mathbf{N}\mathbf{q}' = \mathbf{0}. \quad (19)$$

But, $\mathbf{C}^T\mathbf{M} = \mathbf{0}^T$. Consequently, we are left with $\mathbf{C}^T\mathbf{N}\mathbf{q}' = \mathbf{0}$, or

$$\mathbf{H}\mathbf{q}' = \mathbf{C}^T\mathbf{N}\mathbf{q}'$$

$$= \begin{bmatrix} -(\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) & \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 \\ \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 & (\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) \end{bmatrix} \mathbf{q}' = \mathbf{0}.$$

Now,

$$\det\mathbf{H} = -\{(\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2)^2 + (\psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1)^2\} \neq 0,$$

using (16), which implies **H** is nonsingular. Thus, $\mathbf{q}' = \mathbf{0}$. Equation (17) now becomes $\mathbf{M}\mathbf{p}' = \mathbf{0}$, meaning that $\mathbf{p}' \in \mathrm{N}(\mathbf{M})$, $\mathbf{p}' = \mu\mathbf{w} + \tau\mathbf{w}_1$. From (18),

$$0 = \mathbf{w}^T\mathbf{B}_2\mathbf{p}' = \mu\mathbf{w}^T\mathbf{B}_2\mathbf{w} + \tau\mathbf{w}^T\mathbf{B}_2\mathbf{w}_1.$$

Now, because $\mathbf{w}^T\mathbf{B}_2\mathbf{w}_1 = 0$ and $\mathbf{w}^T\mathbf{B}_2\mathbf{w} \neq 0$, we have $\mu = 0$ and so

$$\mathbf{p}' = \tau\mathbf{w}_1$$

Hence, for all $\tau \in \mathbb{R}\backslash\{0\}$, $\mathbf{p}' = [\tau\mathbf{z}_2, -\tau\mathbf{z}_1]^T \in \mathcal{N}(\mathbf{M})$ also satisfies equation (18). Therefore, we obtain $\phi = \tau[z_2, -z_1, 0, 0]^T$ as the only nonzero nullvector of $\mathbf{F_v(v)}$. The next result is a corollary to Theorem 2.1 and it shows that the Jacobian (10) has linearly independent rows at the root.

Corollary 2.1: If the eigenpair (**z**, λ) of (**A**, **B**) is algebraically simple, then the Jacobian $\mathbf{F_v(v)}$ in (14) is of full rank at the root.

Proof. Since Theorem 2.1 guarantees the existence of a single nonzero nullvector of $\mathbf{F_v(v)}$ at the root, then *rank* ($\mathbf{F_v(v)}$) = 2n+1 (using the dimension theorem, see, for example, [5]). Therefore, the Jacobian (10) is of full rank at the root.

Next, in order to solve the under-determined system of nonlinear equations (9), we need to linearize $\mathbf{F}(\mathbf{v}) = \mathbf{0}$. After linearizing $\mathbf{F}(\mathbf{v}) = \mathbf{0}$, we have to solve the following under-determined linear system of equations

$$\mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)})\Delta\mathbf{v}^{(k)} = -\mathbf{F}(\mathbf{v}^{(k)}). \qquad (20)$$

Hence, solving for $\Delta\mathbf{v}^{(k)}$ in $\mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)})\Delta\mathbf{v}^{(k)} = -\mathbf{F}(\mathbf{v}^{(k)})$, involves solving a 2n + 1 real under-determined linear system of equations for the 2n + 2 real unknowns $\Delta\mathbf{v}^{(k)} = [\Delta\mathbf{z}_1^{(k)}, \Delta\mathbf{z}_2^{(k)}, \Delta\alpha^{(k)}, \Delta\beta^{(k)}]^T$. We find the reduced QR factorization $\mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)})^T = \mathbf{QR}$, where in this case $\mathbf{Q}$ and $\mathbf{R}$ are (2n + 2) by (2n + 1) and (2n + 1) by (2n + 1) real matrices respectively. Hence, we solve $\mathbf{R}^T\mathbf{g}^{(k)} = -\mathbf{F}(\mathbf{v}^{(k)})$ for $\mathbf{g}^{(k)}$ and then obtain the solution to (20) as

$$\Delta\mathbf{v}^{(k)} = \mathbf{Q}\mathbf{g}^{(k)},$$

and update $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$. Since we have shown that the Jacobian has linearly independent rows in Corrolary 2.1, the whole analysis now gives rise to the following algorithm, namely, the **Gauss-Newton method** applied to $\mathbf{F}(\mathbf{v}) = \mathbf{0}$.

Eigenpair Computation using Gauss-Newton's method

- $\mathbf{A}, \mathbf{B}, \mathbf{v}^{(0)} = [\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}, \alpha^{(0)}, \beta^{(0)}]^T, k_{\max}$ and tol.
- for $k = 0, 1, 2, \ldots,$ until convergence
  - Find the reduced QR factorisation of $\mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)})^T = \mathbf{QR}$.
  - Solve $\mathbf{R}^T\mathbf{g}^{(k)} = -\mathbf{F}(\mathbf{v}^{(k)})$ for $\mathbf{g}^{(k)}$ in (10).
  - Compute $\Delta\mathbf{v}^{(k)} = \mathbf{Q}\mathbf{g}^{(k)}$ for $\Delta\mathbf{v}^{(k)}$ using (9).
  - Update $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$.
  - $\mathbf{v}^{(k_{\max})}$.

The stopping condition for the algorithm above is

$$\|\Delta\mathbf{v}^{(k)}\| \leq tol.$$

Next, we give the following numerical example to illustrate the above theory.

III. Numerical Experiment

Consider the 200 by 200 matrix **A** bwm200.mtx from the matrix market library [1]. It is the discretised Jacobian of the Brusselator wave model for a chemical reaction. The resulting eigenvalue problem with $\mathbf{B} = \mathbf{I}$ was also studied in [6] and we are interested in finding the rightmost eigenvalue of **A** which is closest to the imaginary axis and its corresponding eigenvector.

In this example, we take $\alpha^{(0)} = 0.0, \beta^{(0)} = 2.5$ in line with [6] and took $\mathbf{z}_1^{(0)} = 1/2\|\mathbf{1}\|$ and $\mathbf{z}_2^{(0)} = \frac{\sqrt{3}}{2}\mathbf{1}/\|\mathbf{1}\|$, where **1** is the vector of all ones. The algorithm is stopped as soon as $\|\Delta v^{(k)}\|$ is less than or equal to 5.6×10^{-14}. The computed eigenpairs are shown in Table I. Observe that we obtained quadratic convergence from the second to the last and the last columns of Table I for k = 3, 4, 5, 6 and 7. At the root, the condition number of $\mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)})$ is approximately 3×10^3. $\mathbf{w}^{(k)}$ in the above table represents $[\mathbf{z}_1^{(k)T}, \mathbf{z}_2^{(k)T}]$ and $\lambda^{(k)} = [\alpha^{(k)}, \beta^{(k)}]$.

Table I: Values of $\alpha^{(k)}$ and $\beta^{(k)}$. Columns 6 and 7 show that the results converged quadratically for $k = 3, 4, 5, 6$ and 7.

k	$\alpha^{(k)}$	$\beta^{(k)}$	$\|\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\|$	$\|\lambda^{(k+1)} - \lambda^{(k)}\|$	$\|\Delta\mathbf{v}^{(k)}\|$	$\|\mathbf{F}(\mathbf{v}^{(k)})\|$
0	0.00000e+00	2.50000	3.8e+00	7.8e-01	3.9e+00	3.6e+01
1	2.34253e-01	1.75371	1.8e+00	2.2e-01	1.8e+00	7.8e+00
2	1.18745e-01	1.94460	8.1e-01	1.4e-01	8.2e-01	1.7e+00
3	4.47044e-02	2.06484	2.5e-01	7.0e-02	2.6e-01	3.4e-01
4	8.82702e-03	2.12479	3.1e-02	1.7e-02	3.5e-02	3.7e-02
5	2.48114e-04	2.13905	4.8e-04	5.2e-04	7.1e-04	7.1e-04
6	1.80714e-05	2.13950	1.2e-07	2.5e-07	2.8e-07	2.8e-07
7	1.81999e-05	2.13950	2.1e-14	2.9e-14	3.6e-14	6.0e-14

IV. Conclusion

While Ruhe's normalization requires the choice of a starting vectore **c**, our normailization is the natural two-norm normalization for an eigenvector and we do not need to worry about how to choose **c**. The new approach gives quadratic convergence to the complex eigenpair of interest.

Acknowledgements

The first author acknowledges financial support from the University of Bath, United Kingdom, in the form of a University Research Studentship, the National Research Foundation and the Computational Biology Group at the University of Cape Town.

References

[1] B. Biosvert, R. Pozo, K. Remington, B. Miller, R. Lipman, Matrix Market, http://math.nist.gov/MatrixMarket/.
[2] P. Deuflhard, Newton Methods for Nonlinear Problems, ch. 4, pp. 174-175, Springer, 2004.
[3] E. Kreyszig, Advanced Engineering Mathematics, eighth edition, John Wiley & Sons Inc., New York, 1999.
[4] K. Meerbergen and D. Roose, Matrix Transformations for Computing Rightmost Eigenvalues of Large Sparse Non-Symmetric Eigenvalue Problems, IMA Journal of Numerical Analysis 16 (1996), pp, 297-346.
[5] E. D. Nering, Linear Algebra and Matrix Theory, second ed., ch. 1, pp, 31-32, John Wiley & Sons, Inc., 1970.
[6] B. N. Parlett and Y. Saad, Complex Shift and Invert Strategies for Real Matrices, Lin. Alg. Appl. (1987), 575-595.
[7] A. Ruhe,Algorithms for the Nonlinear Eigenvalue Problem, SIAM J. Matrix Anal. Appl. 10 (1973), no. 4, 674-689.

[8] F. Tisseur., Newton's Method in Floating Point Arithmetic and Iterative Refinement of Generalized Eigenvalue Problems, SIAM J. Matrix Anal. Appl. 22 (2001), 1038-1057.
[9] G. W. Stewart, Matrix Algorithms, vol. II: Eigensystems, SIAM 2001.

AUTHOR'S PROFILE

Dr. Richard Olatokunbo Akinola
is a lecturer in the Department of Mathematics, University of Jos. He attended Government Secondary School Laranto, Jos-Plateau State (Can anything good come out of Nazareth?) and a graduate of Mathematics from the University of Jos. Holds a masters degree cum laude from Stellenbosch University, South Africa and a Ph.D. in Applied Mathematics from the University of Bath, UK. He is currently doing postdoctoral research in Computational Biology at the Faculty of Health Sciences, University of Cape Town, South Africa. My research interests is in Numerical Analysis and Scientific Computing. To God be the glory for bringing me thus far because a man can receive nothing except it is given to him from above.