The Parametrized Newton-Secant Method for Finding an Eigenpair of the Symmetric Quadratic Eigenvalue Problem in an Interval

Karabi Datta [1], Yoopyo Hong [1] and Mohan Thapa [2]

[1]*Department of Mathematics, Northern Illinois University, USA*

[2]*Department of Mathematics, University of Wisconsin Milwaukee - Washington County, USA*

***Corresponding** Author: Mohan Thapa*

Email: *thapam@uwm.edu*

Abstract: *Solving a large sparse quadratic eigenvalue problem (QEP),*

$$Q(\lambda)u = (\lambda^2 M + \lambda C + K)u = 0$$

numerically is a difficult task. We are interested in obtaining an eigenpair (λ, u) of the QEP where the eigenvalue λ lies in a specified interval $[a, b]$ from an initial pair (α, x) in which $x \in \mathbb{R}^n$ is chosen arbitrarily. We propose a method, the Parametrized Newton-Secant (PNS) that first transforms the QEP to an approximated linear form which has the same dimension as QEP, by using a secant slope matrix. Then the modified Newton method is applied to the secant linear form to obtain the desired eigenpair in an interval. This method is especially useful to expeditiously obtain a good initial pair. This initial pair can then be utilized in other methods (such as Jacobi-Davidson) that require such a pair to guarantee convergence to a target eigenpair.

Keywords: Quadratic eigenvalue problem(QEP), Standard eigenvalue problem(SEP), Generalized eigenvalue problem(GEP), Jacobi-Davidson method, Second Order Arnoldi method(SOAR), Lanczos method, Secant slope linear approximation(SSLA), Parametrized Newton-Secant method(PNS).

1   Introduction

The Quadratic Eigenvalue Problem (QEP) finds the scalars λ, called the **eigenvalues**, and the nonzero vectors u, called the **eigenvectors**, of the quadratic matrix pencil $Q(\lambda)$:

$$Q(\lambda)u = (\lambda^2 M + \lambda C + K)u = 0,$$

where M is mass matrix, C is damping matrix, and K is stiffness matrix. The QEP arises in a wide variety of applications including vibration of structures, vibro-acoustic systems, fluid dynamics, signal processing, and micro-electric mechanical systems. The vibration of mechanical structures is determined by natural frequencies and mode shapes where the frequencies are related to the eigenvalues and the mode shapes are the associated eigenvectors of the quadratic pencil.

The quadratic eigenvalue problem is difficult to solve numerically because there does not exist any canonical form analogous to the real-Schur form for the standard eigenvalue problem or the generalized eigenvalue problem. In view of this fact, an obvious approach for solving the QEP is to transform the QEP either to a standard eigenvalue problem:

$$\begin{bmatrix} -M^{-1}C & -M^{-1}K \\ I & O \end{bmatrix} \begin{bmatrix} \lambda u \\ u \end{bmatrix} = \lambda \begin{bmatrix} \lambda u \\ u \end{bmatrix} \tag{1}$$

and solve the problem by the QR algorithm or to a generalized eigenvalue problem:

$$\begin{bmatrix} -C & -K \\ I & 0 \end{bmatrix} \begin{bmatrix} \lambda u \\ u \end{bmatrix} = \lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \lambda u \\ u \end{bmatrix} \tag{2}$$

and solve by the QZ algorithm [3, 10]. Some difficulties with this conventional approach are encountered during the process of linearization. For example the dimension of the problem becomes double in size and unfortunately, both the QR and the QZ algorithm are known to be ineffective for large sparse problems. Certain desirable aspects of the QEP such as the symmetric, positive or negative definite, sparse or tridiagonal etc. can be destroyed. Another difficulty is that the matrix M

needs to be inverted, and if M is close to a singular matrix, then this approach fails. For large sparse problems, the Lanczos method computes the largest and smallest eigenvalues by using the shift and invert method which requires a sizeable effort [1]. There are two methods, the Second Order Arnoldi (SOAR) method and the Jacobi-Davidson method, that work directly with the quadratic pencil without transforming the QEP to a linear form. The SOAR method generates an orthonormal basis for the generalized Krylov subspace, $G_n(A, B, v)$ where $A = M^{-1}K$ and $B = M^{-1}C$, from an initial vector v [2]. The Jacobi-Davidson method is a projection method. By using a few initial vectors, the method projects the QEP into a small dimension problem. The reduced QEP is then solved by a standard technique. It is noted that these direct methods efficiently compute only the largest eigenvalue and the associated eigenvector [2]. Also, these direct methods clearly depend on how wisely the initial vector(s) is chosen. Unfortunately, there are no definite criteria on how to choose an initial vector. One object of the paper is to expeditiously produce an effective initial vector.

In some applications such as in the control of resonance, what is needed is a few eigenvalues in a given interval. The primary objective of this paper is to describe a method that effectively computes an eigenpair whose eigenvalue lies in this intervals $[a, b]$. A Newton type method is preferred for this purpose because it can be immediately implemented without changing the QEP to a linear form (1) or (2). One chooses an initial point (α, x), where α is in the interval $[a, b]$ but $x \in \mathbb{R}^n$ is arbitrarily chosen. With this method, the structure of the QEP is preserved. Despite these desirable aspects of the Newton method, given the complexity of the QEP, one obstacle remains. We must first approximate the quadratic pencil to a linear form by using a secant slope matrix. The secant slope linear form is then iterated by a parametrized Newton method which is a modification of the Newton method for the symmetric standard eigenvalue problem [11, 5]. We call the combined method the parametrized Newton-Secant Method (PNS). The method is particularly suited for finding a targeted eigenvalue in a prescribed interval and the associated eigenvector. However, the PNS method computes one eigenpair at a time. Our numerical experiments on test matrices of dimensions up to 2000 show that the PNS method is much more accurate and faster than the Jacobi-Davidson, SOAR or Lanczos methods.

2 The Parametrized Newton-Secant Method for Symmetric QEP

To describe the parametrized Newton-Secant method for the symmetric QEP, first we approximate the quadratic pencil $Q(\lambda)$ by a linear form $Q_s(\lambda)$ in an interval $[a, b]$ which is obtained by using a secant slope matrix. Then the parametrized Newton method is applied to the approximated linear pencil $Q_S(\lambda)$. Throughout this paper, we assume the Euclidean norm (l_2) on a vector and the induced matrix norm, the spectral norm $||| \cdot |||$, on a matrix.

2.1 Secant Slope Linear Approximation(SSLA) $Q_s(\lambda)$:

Suppose $Q(\lambda) = \lambda^2 M + \lambda C + K$ is a symmetric quadratic pencil where M, C and $K \in \mathbb{R}^{n \times n}$ are symmetric. Let $f : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ be such that

$$f(\lambda, u) = \begin{pmatrix} Q(\lambda)u \\ u^T u - 1 \end{pmatrix} = \begin{pmatrix} (\lambda^2 M + \lambda C + K)u \\ u^T u - 1 \end{pmatrix}, \ u \in \mathbb{R}^n, \lambda \in \mathbb{R}.$$

The symmetric QEP is to solve $f(\lambda, u) = 0$. For any two real numbers α_1 and α_2, $\alpha_1 < \alpha_2$, we find the secant slope matrix R_s,

$$R_s \equiv \frac{(\alpha_1^2 M + \alpha_1 C + K) - (\alpha_2^2 M + \alpha_2 C + K)}{\alpha_1 - \alpha_2} = (\alpha_1 + \alpha_2)M + C.$$

The secant slope linear approximation (SSLA) $Q_s(\lambda)$ is defined by,

$$\begin{aligned} Q_s(\lambda) &\equiv R_s(\lambda - \alpha_1) + Q(\alpha_1) \\ &= (\lambda - \alpha_1)((\alpha_1 + \alpha_2)M + C) + \alpha_1^2 M + \alpha_1 C + K \\ &= \lambda((\alpha_1 + \alpha_2)M + C) - \alpha_1 \alpha_2 M + K. \end{aligned}$$

If we set $T = -\alpha_1 \alpha_2 M + K$, then $Q_s(\lambda) = \lambda R_s + T$. We define

$$f_s(\lambda, u) \equiv \begin{pmatrix} (\lambda R_s + T) u \\ u^T u - 1 \end{pmatrix}.$$

The following verifies that for $\lambda \in [\alpha_1, \alpha_2]$, if $|\alpha_2 - \alpha_1| \to 0$ then $Q_s(\lambda) \approx Q(\lambda)$, and hence $f_s(\lambda, u) \approx f(\lambda, u)$. Let $\delta_1 = (\lambda - \alpha_1)$ and $\delta_2 = (\alpha_2 - \lambda)$ for $\lambda \in [\alpha_1, \alpha_2]$. Then

$$Q_s(\lambda) \quad = (\lambda R_s + T) = \lambda((\alpha_1 + \alpha_2)M + C) - \alpha_1 \alpha_2 M + K = (\lambda(\alpha_1 + \alpha_2) - \alpha_1 \alpha_2) M + \lambda C + K$$

$$= (\lambda(\lambda - \delta_1 + \lambda + \delta_2) - (\lambda - \delta_1)(\lambda + \delta_2)) M + \lambda C + K$$
$$= (\lambda^2 + \delta_1\delta_2)M + \lambda C + K. \tag{3}$$

Thus,
$$|||Q(\lambda) - Q_s(\lambda)||| = |\delta_1\delta_2||||M|||.$$

Since $\delta_1 = t(\alpha_2 - \alpha_1)$ and $\delta_2 = (1 - t)(\alpha_2 - \alpha_1)$ for some t, $0 \leq t \leq 1$,

$$\delta_1\delta_2 = t(1 - t)(\alpha_2 - \alpha_1)^2 \leq \frac{1}{4}(\alpha_2 - \alpha_1)^2.$$

We conclude $|||Q(\lambda) - Q_s(\lambda)||| \leq \frac{1}{4}(\alpha_2 - \alpha_1)^2|||M|||$. Thus, if $|\alpha_2 - \alpha_1| \to 0$ then $Q(\lambda) \approx Q_s(\lambda)$.

The above analysis shows our method depends strongly on the successful reduction of the gap, $|\alpha_1 - \alpha_2|$. In the parametrized Newton-Secant method, a parameter s is chosen to achieve the goal.

2.2 The Parametrized Newton Method for the SSLA :

We write an eigenpair (λ, u) as a column vector $\begin{pmatrix} u \\ \lambda \end{pmatrix} \in \mathbb{R}^{n+1}$. Let $\begin{pmatrix} x^{(i)} \\ \alpha^{(i)} \end{pmatrix}$ be the ith approximation of an eigenpair $\begin{pmatrix} u \\ \lambda \end{pmatrix}$ where λ is assumed to be in the given interval $[a, b]$. The parametrized Newton method refines the approximation of $\begin{pmatrix} u \\ \lambda \end{pmatrix}$, iteratively, as follows:

$$\begin{pmatrix} x^{(i+1)} \\ \alpha^{(i+1)} \end{pmatrix} = \begin{pmatrix} x^{(i)} \\ \alpha^{(i)} \end{pmatrix} - \begin{bmatrix} Q_s(\alpha^{(i)}) & R_s^{(i)}x^{(i)} \\ 2x^{(i)T} & 0 \end{bmatrix}^{-1} \begin{pmatrix} Q_s(\alpha^{(i)})x^{(i)} \\ x^{(i)T}x^{(i)} - 1 \end{pmatrix}$$

where
$$\begin{bmatrix} Q_s(\lambda) & R_s u \\ 2u^T & 0 \end{bmatrix}$$ is the Jacobian matrix of $f_s(\lambda, u)$.

We note that if $\alpha^{(i)}$ is not an eigenvalue of $Q_s(\lambda)$ then $Q_s(\alpha^{(i)})$ is invertible. Assuming $\alpha^{(i)} \neq 0$, we can write

$$\begin{bmatrix} Q_s(\alpha^{(i)}) & R_s^{(i)}x^{(i)} \\ 2x^{(i)T} & 0 \end{bmatrix} \begin{pmatrix} x^{(i+1)} \\ \alpha^{(i+1)} \end{pmatrix} = \begin{bmatrix} 0 & R_s^{(i)}x^{(i)} \\ x^{(i)T} & \frac{1}{\alpha^{(i)}} \end{bmatrix} \begin{pmatrix} x^{(i)} \\ \alpha^{(i)} \end{pmatrix}.$$

Choose a parameter $\tau > 0$ so that the method takes the form [5]:

$$\begin{bmatrix} Q_s(\alpha^{(i)}) & R_s^{(i)}x^{(i)} \\ 2x^{(i)T} & 0 \end{bmatrix} \begin{pmatrix} x^{(i+1)} \\ \alpha^{(i+1)} \end{pmatrix} = \begin{bmatrix} I & 0 \\ 0 & \tau \end{bmatrix} \begin{bmatrix} 0 & R_s^{(i)}x^{(i)} \\ x^{(i)T} & \frac{1}{\alpha^{(i)}} \end{bmatrix} \begin{pmatrix} x^{(i)} \\ \alpha^{(i)} \end{pmatrix}.$$

Then we have:
$$Q_s(\alpha^{(i)})x^{(i+1)} + \alpha^{(i+1)}R_s^{(i)}x^{(i)} = \alpha^{(i)}R_s^{(i)}x^{(i)}$$

or
$$x^{(i+1)} = \left(\alpha^{(i)} - \alpha^{(i+1)}\right)Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)} \tag{4}$$

and
$$2x^{(i)T}x^{(i+1)} = \tau\left(x^{(i)T}x^{(i)} + 1\right). \tag{5}$$

From (4) and (5), we obtain

$$x^{(i)T}x^{(i+1)} = \left(\alpha^{(i)} - \alpha^{(i+1)}\right)x^{(i)T}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}.$$

Set

$$\beta^{(i)} = x^{(i)T}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}.$$

Then we have

$$\frac{\tau}{2}\left(x^{(i)T}x^{(i)} + 1\right) = \beta^{(i)}\left(\alpha^{(i)} - \alpha^{(i+1)}\right)$$

or

$$\left(\alpha^{(i)} - \alpha^{(i+1)}\right) = \frac{\tau}{2\beta^{(i)}}\left(x^{(i)T}x^{(i)} + 1\right).$$

Since $x^{(i)}$ is normalized vector,

$$\alpha^{(i+1)} = \alpha^{(i)} - \frac{\tau}{\beta^{(i)}}. \tag{6}$$

Now from (4) and (6) we obtain,

$$x^{(i+1)} = \frac{\tau}{\beta^{(i)}}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}.$$

Set

$$\hat{\beta}^{(i)} = \left\|Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}\right\| = \left(x^{(i)T}R_s^{(i)}Q_s^{-2}(\alpha^{(i)})R_s^{(i)}x^{(i)}\right)^{\frac{1}{2}}.$$

After normalizing the vector $x^{(i+1)}$, we have

$$x^{(i+1)} = \frac{\tau Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}}{\left\|\tau Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}\right\|} = \frac{1}{\hat{\beta}^{(i)}}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}. \tag{7}$$

Thus, from (6) and (7) the parametrized Newton iteration for $Q_s(\lambda)$ takes the form:

$$x^{(i+1)} = \frac{1}{\hat{\beta}^{(i)}}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)} \tag{8}$$

$$\alpha^{(i+1)} = \alpha^{(i)} - \frac{\tau}{\beta^{(i)}} \tag{9}$$

The parametrized Newton method with the iteration (8) and (9) is called the parametrized Newton-Secant (PNS) method if the parameter τ is chosen to minimize the residual and also to ensure $|\alpha^{(i+1)} - \alpha^{(i)}| \le |\alpha^{(i)} - \alpha^{(i-1)}|$. For the purpose, we need the following result.

Lemma 2.1: *Suppose $\alpha^{(i)}$ is not an eigenvalue of $Q_s(\lambda)$. Then $\left|\frac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right| \le 1$.*

Proof: Since $|\beta^{(i)}| = |x^{(i)T}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}|$ is the usual inner product of the vectors $x^{(i)}$ and $Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}$, by the Cauchy-Schwarz inequality, we have $|\beta^{(i)}| \le \|x^{(i)}\|\,\|Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)}\|$. Thus, $\left|\frac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right| \le \|x^{(i)}\| = 1$. $\square$

Since τ is chosen to decrease the residual, we compute the following.

$$
\begin{aligned}
\|Res(\alpha^{(i+1)})\|^2 &= \left\|R_s^{-1}\left(\alpha^{(i+1)}R_s + T\right)x^{(i+1)}\right\|^2 \\
&= \left\|R_s^{-1}\left(\left(\alpha^{(i)} - \frac{\tau}{\beta^{(i)}}\right)R_s + T\right)\frac{Q_s^{-1}(\alpha^{(i)})R_sx^{(i)}}{\hat{\beta}^{(i)}}\right\|^2 \\
&= \frac{1}{\hat{\beta}^{(i)2}}\left\|R_s^{-1}\left(Q_s(\alpha^{(i)}) - \frac{\tau}{\beta^{(i)}}R_s\right)Q_s^{-1}(\alpha^{(i)})R_sx^{(i)}\right\|^2 \\
&= \frac{1}{\hat{\beta}^{(i)2}}\left\|\left(I - \frac{\tau}{\beta^{(i)}}Q_s^{-1}(\alpha^{(i)})R_s\right)x^{(i)}\right\|^2 \\
&= \frac{1}{\hat{\beta}^{(i)2}}\left(1 - \frac{\tau}{\beta^{(i)}}\left(x^{(i)T}Q_s^{-1}(\alpha^{(i)})R_sx^{(i)} + x^{(i)T}R_sQ_s^{-1}(\alpha^{(i)})x^{(i)}\right) + \right.\\
&\qquad\qquad \left. \frac{\tau^2}{\beta^{(i)2}}x^{(i)T}\left(Q_s^{-1}(\alpha^{(i)})R_s\right)^T\left(Q_s^{-1}(\alpha^{(i)})R_s\right)x^{(i)}\right) \\
&= \frac{1}{\hat{\beta}^{(i)2}}\left(1 - 2\tau + \frac{\tau^2}{\beta^{(i)2}}\hat{\beta}^{(i)2}\right).
\end{aligned}
$$

Choose $\tau = \left(\frac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right)^2 s$ for some positive number s. Then,

$$\|Res(\alpha^{(i+1)})\|^2 = \frac{1}{\hat{\beta}^{(i)2}}\left[1 - 2\left(\frac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right)^2 s + \left(\frac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right)^2 s^2\right]. \tag{10}$$

It is easy to verify that the expression in (10) is decreasing for $0 < s \leq 1$ and is minimized when $s = 1$ with its minimum value $\frac{1}{\hat{\beta}^{(i)^2}}\left(1 - \left(\frac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right)^2\right)$. For this reason, the Parametrized Newton-Secant (PNS) method takes the form;

$$x^{(i+1)} = \frac{1}{\hat{\beta}^{(i)}}Q_s^{-1}(\alpha^{(i)})R_s^{(i)}x^{(i)} \qquad (11)$$

$$\alpha^{(i+1)} = \alpha^{(i)} - \frac{(r^{(i)})}{\hat{\beta}^{(i)}}s, \ 0 < s \leq 1. \qquad (12)$$

In the algorithm for the Parametrized Newton-Secant method, the parameter $s \approx 1$ is chosen to accomplish the reduction of both the residual and the gap, $|\alpha^{(i+1)} - \alpha^{(i)}|$. In the next section, we analyse the convergence of the method.

3 Convergence of the Parametrized Newton-Secant Method

Suppose $[a, b]$ is the given interval where a target eigenvalue λ exists. There are obvious initial choices for the starting secant slope matrix R_s, those are $\alpha^{(0)} = a$ and $\alpha^{(1)} = b$. However, there is no definite criteria for choosing an initial approximation of the eigenvector. We verify that if the parameter s, $0 < s \leq 1$ is chosen suitably so that the gap, $|\alpha^{(i+1)} - \alpha^{(i)}| \to 0$ then the parameterized Newton-Secant method converges to an eigenpair of the QEP with an arbitrarily initial vector $x^{(0)}$ inside the given interval.

Lemma 3.1: *Suppose $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $x \in \mathbb{R}^n$, such that $\|x\| = 1$. Then*

$$\|Ax\|^2 = \sum_{j=1}^{n} \sigma_j^2 y_j^2,$$

where $\sigma_j > 0$ are the singular values of A and $\sum_{j=1}^{n} y_j^2 = 1$.

Proof: Since $A^T A$ is a symmetric positive definite matrix, there exists a real orthogonal matrix U such that
$\|Ax\|^2 = x^T A^T A x = x^T U^T \Sigma U x = (Ux)^T \Sigma (Ux)$,

where $\Sigma = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}$, $\sigma_j > 0$ are the singular values of A.

If we set $Ux = y = (y_1, y_2, ..., y_n)^T$, then $\|y\|^2 = \sum_{j=1}^{n} y_j^2 = 1$ and $y^T \Sigma y = \sum_{j=1}^{n} \sigma_j^2 y_j^2$. $\qquad \square$

Lemma 3.2: *Suppose $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $\sigma_j > 0$, $\sigma_1 > \sigma_2 > \cdots > \sigma_n$ are the singular values of A. Then there exists a constant $k > 0$ such that $\|Ax\| = k\sigma_1$.*

Proof: By Lemma (3.1), $\|Ax\| = \left(\sum_{j=1}^{n} \sigma_j^2 y_j^2\right)^{\frac{1}{2}} = \left(\sum_{j=1}^{n} \left(\frac{\sigma_j}{\sigma_1}\right)^2 y_j^2\right)^{\frac{1}{2}} \sigma_1 = k\sigma_1,$

where $k = \left(\sum_{j=1}^{n} \left(\frac{\sigma_j}{\sigma_1}\right)^2 y_j^2\right)^{\frac{1}{2}} > 0$ for $\sum_{j=1}^{n} y_j^2 = 1$. $\qquad \square$

The following is the well known **Jensen's Inequality** [6]. *Suppose*
$f : I \to \mathbb{R}$ *is a convex function on* $I \subset \mathbb{R}$, *and let* $S = \{\alpha_i : \alpha_i \in I, \ i = 1, \ldots, n\}$.

Then, $f\left(\sum_{j=1}^{n} \alpha_j y_j^2\right) \le \sum_{j=1}^{n} f(\alpha_j) y_j^2$ *for* $\sum_{j=1}^{n} y_j^2 = 1$.

Regarding the positive constant k in Lemma (3.2), we make the following observation for the case $A = Q_s^{-1}(\alpha^{(i)})R_s$ when $\alpha^{(i)} \approx \alpha^{(i+1)}$. Suppose σ_j, $\sigma_1 > \sigma_2 > \cdots > \sigma_n > 0$ are the singular values of $Q_s^{-1}(\alpha^{(i)})R_s$, the SSLA at the ith step of the PNS iteration. Then

$$\hat{\beta}^{(i)^2} = \|Q_s^{-1}(\alpha^{(i)})R_s x^{(i)}\|^2 = \sigma_1^2 k^{(i)^2} \text{ where } k^{(i)^2} = \sum_{j=1}^{n} \left(\frac{\sigma_j}{\sigma_1}\right)^2 y_j{}^2.$$

At the $(i+1)^{th}$ step of the iteration, $\hat{\beta}^{(i+1)^2} = \|Q_s^{-1}(\alpha^{(i+1)})R_s x^{(i+1)}\|^2$ where
$x^{(i+1)} = \frac{1}{\hat{\beta}^{(i)}} Q_s^{-1}(\alpha^{(i)})R_s x^{(i)}$. Thus, if $\alpha^{(i)} \approx \alpha^{(i+1)}$, then $\||Q_s^{-1}(\alpha^{(i+1)})R_s\|| \approx \||Q_s^{-1}(\alpha^{(i)})R_s\||$: by continuity of nonsingular quadratic form of $Q_s^{-1}(\alpha^{(i)})R_s$. In this case,

$$\hat{\beta}^{(i+1)^2} \approx \frac{1}{\hat{\beta}^{(i)^2}} \|(Q_s^{-1}(\alpha^{(i+1)})R_s)^2 x^{(i)}\|^2 = \frac{1}{\hat{\beta}^{(i)^2}}\left(\sum_{j=1}^{n} \sigma_j{}^4 y_j{}^2\right) = \frac{\sigma_1^4}{\hat{\beta}^{(i)^2}}\left(\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j{}^2\right)$$

$$= \frac{\sigma_1^4}{\sigma_1^2 k^{(i)^2}}\left(\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j{}^2\right) = \frac{\sigma_1^2}{k^{(i)^2}}\left(\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j{}^2\right) = \sigma_1^2\left(\frac{1}{k^{(i)^2}}\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j{}^2\right). \tag{13}$$

Since $\hat{\beta}^{(i+1)^2} \approx \sigma_1^2 k^{(i+1)^2}$, the equations in (13) shows that $k^{(i+1)^2} = \frac{1}{k^{(i)^2}}\left(\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j{}^2\right)$.

Consider $f(x) = x^2$. Then f is a convex function on $\mathbb{R}$, and the Jensen's inequality asserts

$$f\left(k^{(i)^2}\right) = f\left(\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^2 y_j^2\right) \le \sum_{j=1}^{n} f\left(\left(\frac{\sigma_j}{\sigma_1}\right)^2\right) y_j^2 = \sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j^2. \text{ Therefore,}$$

$$k^{(i+1)^2} = \frac{1}{k^{(i)^2}}\left(\sum_{j=1}^{n}\left(\frac{\sigma_j}{\sigma_1}\right)^4 y_j^2\right) \ge \frac{k^{(i)^4}}{k^{(i)^2}} = k^{(i)^2}.$$

We conclude that when the gap, $|\alpha^{(i+1)} - \alpha^{(i)}|$ is sufficiently small, the positive constants $k^{(i)}$ increases as i increases. We write the observation as a lemma.

Lemma 3.3: *Suppose* $|\alpha^{(i+1)} - \alpha^{(i)}| = s\frac{|\beta^{(i)}|}{\hat{\beta}^{(i)^2}} \le \epsilon$ *for some small* $\epsilon \ge 0$. *Then* $\hat{\beta}^{(i)} = \sigma_1 k^{(i)}$ *for* $k^{(i)} > 0$ *is an increasing function of* i .

The set of all eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$ is called the spectrum of A and is denoted by $\sigma(A)$. The spectral radius of A is denoted and defined as $\rho(A) = \max_j\{|\lambda_j| : \lambda_j \in \sigma(A)\}$. It is well known that $\rho(A)$ is a lower bound of every matrix norms on A, i.e., $\rho(A) \le \||A\||$ for $\|| \cdot \||$ is a matrix norm. Since σ_1, the largest singular value of A is the spectral matrix norm of A, we obtain the following immediate consequence.

Corollary 3.1: *Suppose* $Q_s(\alpha^{(i)}) \in \mathbb{R}^{n \times n}$ *is nonsingular. Then* $\hat{\beta}^{(i)} = \sigma_1 k^{(i)} \ge \rho(Q_s^{-1}(\alpha^{(i)})R_s) k^{(i)}$ *for* $k^{(i)} > 0$ *is an increasing function of* i.

Before we present the main convergence result, we note some other facts of the eigenvalues of $Q_s(\lambda) = \lambda R_s + T$. Assume $R_s \in \mathbb{R}^{n \times n}$ is invertible. Since $Q_s(\lambda) = \lambda R_s + T = R_s(\lambda I + R_s^{-1}T)$, the eigenvalues of $Q_s(\lambda)$ are the eigenvalues of the matrix $-R_s^{-1}T$, i.e., if the set $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ is the eigenvalues of $Q_s(\lambda)$ then $\sigma(R_s^{-1}T) = \{-\lambda_1, -\lambda_2, \cdots, -\lambda_n\}$. For $\alpha^{(i)} \in \mathbb{R}$, we have $Q_s^{-1}(\alpha^{(i)})R_s = \left(R_s^{-1}(\alpha^{(i)}R_s + T)\right)^{-1} = (\alpha^{(i)}I + R_s^{-1}T)^{-1}$. Thus,

$$\sigma(Q_s^{-1}(\alpha^{(i)})R_s) = \left\{\frac{1}{\alpha^{(i)} - \lambda_1}, \frac{1}{\alpha^{(i)} - \lambda_2}, \cdots, \frac{1}{\alpha^{(i)} - \lambda_n}\right\}. \text{ If } \lambda_p \text{ is the eigenvalue of } Q_s(\lambda) \text{ that is closest to } \alpha^{(i)},$$

then we see that $\rho(Q_s^{-1}(\alpha^{(i)})R_s) = \frac{1}{|\alpha^{(i)} - \lambda_p|}$. Combining with Lemma (3.3), we have the following result.

Lemma 3.4: *Suppose* λ_p *is the eigenvalue of* $Q_s(\lambda)$ *such that* $|\alpha^{(i)} - \lambda_p| = \min\limits_{j} |\alpha^{(i)} - \lambda_j|, \lambda_j \in \sigma(Q_s(\lambda))$. *Then*

$\hat{\beta}^{(i)} \geq \dfrac{k^{(i)}}{|\alpha^{(i)} - \lambda_p|}$ *for* $k^{(i)} > 0$ *is an increasing function of* i.

Proof: If $|\alpha^{(i)} - \lambda_p| = \min\limits_{j} |\alpha^{(i)} - \lambda_j|$, then $\rho(Q_s^{-1}(\alpha^{(i)})R_s) = \rho\left((\alpha^{(i)}I + R_s^{-1}T)^{-1}\right) = \dfrac{1}{|\alpha^{(i)} - \lambda_k|}$. The result follows from Corollary 3.1 . $\square$

Theorem 3.1: *Suppose* $\alpha^{(i)}$ *and* $\alpha^{(i+1)}$ *are two successive points obtained during the Parametrized Newton-Secant(PNS) iteration such that the interval* $[\alpha^{(i)}, \alpha^{(i+1)}]$ *contains an eigenvalue* λ_p *of* $Q_s(\lambda)$. *If the gap* $|\alpha^{(i+1)} - \alpha^{(i)}| \to 0$ *as* i *increases,* $\|Res^{(i+1)}\| \to 0$.

Proof: The equation (10) $\|Res^{(i+1)}\|^2 = \dfrac{1}{\hat{\beta}^{(i)^2}}\left[1 - 2\left(\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right)^2 s + \left(\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}\right)^2 s^2\right] \leq \dfrac{1}{\hat{\beta}^{(i)^2}}$ for $s \approx 1$. By Corollary(3.1)

and Lemma(3.4), there is a positive constant $k^{(i)}$ such that $\dfrac{1}{\hat{\beta}^{(i)^2}} \leq \dfrac{1}{k^{(i)^2}}|\alpha^{(i)} - \lambda_p|^2$. Since, λ_p is in the interval

$[\alpha^{(i)}, \alpha^{(i+1)}]$, $|\alpha^{(i)} - \lambda_p| \leq |\alpha^{(i+1)} - \alpha^{(i)}|$, and therefore, $\dfrac{1}{\hat{\beta}^{(i)^2}} \leq \dfrac{1}{k^{(i)^2}}|\alpha^{(i+1)} - \alpha^{(i)}|^2$, where $\dfrac{1}{k^{(i)^2}}$ is decreas-

ing as i increases. We conclude $\|Res^{(i+1)}\| \to 0$ when $|\alpha^{(i+1)} - \alpha^{(i)}| \to 0$, as desired. $\square$

Theorem (3.1) assures that the Parametrized Newton-Secant method converges inside the given interval provided the gap, $|\alpha^{(i+1)} - \alpha^{(i)}| \to 0$.

Algorithm: (Parametrized Newton-Secant Method for a symmetric QEP)
Input:

- The matrix M, C, and $K \in R^{n \times n}$ are symmetric positive definite matrices

- Interval $[a, b]$

- Chose initial points, $\alpha^{(0)}$ and $\alpha^{(1)} \in [a, b]$

- Initial eigenvector $x^{(0)}$

- Tolerance ϵ, δ where $0 < \epsilon < \delta$

Output:

- An approximate eigenpair of the QEP whose eigenvalue is in the interval $[a, b]$

For $i = 1, 2, \ldots$
Step 1.

  (a) Compute $R^{(i)} = (\alpha^{(i-1)} + \alpha^{(i)})M + C$ and $T = -(\alpha^{(i-1)} * \alpha^{(i)})M + K$

  (b) Compute $Q^{(i)} = \alpha^{(i)}R + T$

  (c) Compute $y^{(i)} = R^{(i)}x^{(i)}$

  (d) Solve $Q^{(i)}y^{(i+1)} = y^{(i)}$ for $y^{(i+1)}$

  (e) Compute $\hat{\beta}^{(i)} = \|y^{(i+1)}\|$

  (f) Compute $\beta^{(i)} = x^{(i)^T}y^{(i+1)}$

  (g) Compute $r^{(i)} = \dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$

Step 2. update $\alpha^{(i+1)}$

  (a) if $\left(r^{(i)}\right)^2 > \epsilon$ then $\alpha^{(i+1)} = \alpha^{(i)} - \dfrac{r^{(i)}s}{\hat{\beta}^{(i)}}$, for $0 < s \leq 1$

(b) otherwise, $\alpha^{(i+1)} = \alpha^{(i)} - \dfrac{s}{\hat{\beta}^{(i)}}$, for $0 < s \leq 1$

(c) Check updated $\alpha^{(i+1)}$
If $\alpha^{(i+1)}$ is outside the interval $[a, \ b]$, then choose appropriate $0 < s \leq 1$)
go to Step 2, so that $\alpha^{(i+1)}$ falls inside the interval.

Step 3. Update $x^{(i+1)} = \dfrac{1}{\hat{\beta}^{(i)}} y^{(i+1)}$

Step 4. Check the following:

(a) If $\|Q^{(i+1)}x^{(i+1)}\| \leq \epsilon$ **STOP**.

(b) If $\|Q^{(i+1)}x^{(i+1)}\| \leq \delta$ then set $s = 1$ and go to Step 5.

(c) If $\|Q^{(i+1)}x^{(i+1)}\| > \delta$, then go to Step 5.

Step 5. $\alpha^{(i-1)} \equiv \alpha^{(i)}$ and $\alpha^{(i)} \equiv \alpha^{(i+1)}$, go to Step 1.
end For

4 Numerical Experiments

4.1 Numerical Experiments on PNS Method

The matrices we used in the examples are: (1) symmetric positive definite random sparse matrix which is generated from Matlab routine SPRANDSYM. SPRANDSYM is a Matlab routine which generate random sparse symmetric matrix. This matrix is generated by random Jacobi rotations applied to a diagonal matrix with the given eigenvalues or condition number. (2) symmetric positive definite Toeplitz matrices (PDTOEP). This is also a Matlab routine. This routine, PDTOEP, generates dense Toeplitz matrix not sparse matrix.
We illustrate the parametrized Newton-Secant method on matrices of size 500×500, 1500×1500 and 2000×2000. Let M and C be symmetric positive definite random sparse matrices (SPRANDSYM) and K be a symmetric negative definite random sparse matrix. We did experiments with different intervals in the positive spectrum. We use the following strategy. In every iteration of the parametrized Newton-Secant method, s is chosen such that a new approximated eigenvalue stays within the interval. Initially, we set $s = 1$. After a few iterations of the parametrized Newton-Secant method, if the norm of the residual is less than δ which is set to be greater than the given tolerance ϵ, then we continue with $s = 1$.

In the following experiment, we let e_i be the ith column vector of the identity matrix and n be the dimension of matrices. We set $\delta = 10^{-1}$ and $\epsilon = 10^{-5}$.

Example (1) $n = 500$ with three different intervals.

(1a) Interval $[2, \ 3.14]$, initial shift 2.40 and initial vector e_1.

TABLE 4.1: $n = 500$

Iteration	s	$\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$	eigenvalue(λ)	Residual
1	0.75	-0.888591839866	2.788015137830	396.062550110683
2	0.75	-0.824506002734	2.911622728034	62.201443869010
3	0.75	-0.976583257381	2.965903989807	5.649735521849
4	1	-0.999584169663	2.983444288313	0.169568441295
5	1	0.999999572762	02.983429603720	4.558×10^{-6}

(1b) Interval $[0.7, \ 0.71]$, initial shift 0.701 and initial vector e_1.

TABLE 4.2: $n = 500$

Iteration	s	$\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$	eigenvalue(λ)	Residual
1	0.75000	-0.034777	0.70187	4.7214
2	0.75000	-0.71714	0.70212	0.19221
3	1.0000	-0.96777	0.70225	0.021382
4	1.0000	-0.99963	0.70226	3.2157×10^{-5}

(1c) Interval $[0, \ 0.02]$, initial shift 0.01 and initial vector e_1.

TABLE 4.3: $n = 500$

Iteration	s	$\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$	eigenvalue(λ)	Residual
1	0.75000	-0.53511	0.014223	1.1561
2	0.75000	-0.98086	0.014828	0.071046
3	1.0000	-0.99813	0.015032	0.0042717
4	1.0000	0.99999	0.015031	2.5001×10^{-7}

Example (2). $n = 1500$ with one interval and three different shifts.

(2a) Interval $[0.1, \ 0.11]$, initial shift 0.105 and initial vector e_1.

TABLE 4.4: $n = 1500$

Iteration	s	$\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$	eigenvalue(λ)	Residual
1	0.75000	0.022828	0.10458	35.845
2	0.75000	-0.88948	0.10466	0.061900
3	1.0000	-0.99961	0.10468	0.00083675
4	1.0000	1.0000	0.10468	5.9674×10^{-8}

(2b) Interval $[0.1, \ 0.11]$, initial shift 0.108 and initial vector e_1.

TABLE 4.5: $n = 1500$

Iteration	s	$\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$	eigenvalue(λ)	Residual
1	0.75000	-0.0045477	0.10753	4.7912
2	0.75000	0.99754	0.10753	0.00090665
3	1.0000	1.0000	0.10753	1.9113×10^{-7}

(2c) Interval $[0.1, \ 0.11]$, initial shift 0.109 and initial vector e_1.

TABLE 4.6: $n = 1500$

Iteration	s	$\dfrac{\beta^{(i)}}{\hat{\beta}^{(i)}}$	eigenvalue(λ)	Residual
1	0.10000	-0.17368	0.10959	40.254
2	0.10000	0.55878	0.10957	0.42158
3	1.0000	0.99208	0.10939	0.026613
4	1.0000	0.99983	0.10939	1.2154×10^{-6}

Example (3). $n = 2000$ with one interval $[2, \ 4]$ and two different shifts.

(3a) Interval $[2, \ 4]$, initial shift 2.5 and initial vector e_1.

TABLE 4.7: $n = 2000$

Iteration	s	eigenvalue(λ)	Residual
1	0.10000	2.61565	4534.217
2	0.10000	2.37239	1297.06
3	1.0000	2.50857	229.548
4	1.0000	2.53172	7.93789
5	1.0000	2.531496	0.005416535
6	1.0000	2.531496	2.469×10^{-8}

(3b) Interval $[2, \ 4]$, initial shift 3.0 and initial vector e_1.

TABLE 4.8: $n = 2000$

Iteration	s	eigenvalue(α)	Residual
1	0.10000	3.16275271286	8666.12071499020
2	0.10000	3.31508127996	8618.57912832938
3	0.10000	3.46692496474	8441.01289897658
4	0.10000	3.62031326568	8268.54640262043
5	1.0000	2.05358209480	2013.13374182046
6	1.0000	2.06315842496	96.96510414400
7	1.0000	2.06852741334	1.02638528028
8	1.0000	2.06850881549	2.0437152×10^{-4}
9	1.0000	2.06850883279	1.2×10^{-10}

4.2 Numerical Results: The PNS Method Compares with Other Methods

The matrices we used in these examples are same set as above.

We present the numerical experiments that compare the Parametrized Newton Secant Method, the shift-invert Jacobi-Davidson method, the shift-invert SOAR Method, and the shift-invert Lanczos Method. The experiments which compare the Jacobi-Davidson method and PNS method, have the maximum number of iterations, 20 and the tolerance, 10^{-5}.

Each table shows the results for the eigenvalue, norm of the residual and computation time in seconds, compared between different methods.

1. Symmetric positive definite Toeplitz matrices M and C, a symmetric negative definite matrix K (PDTOEP) with an initial vector e_1.

   TABLE 4.9: $n = 500$, interval $[1.1, \ 1.2]$, shift = 1.15

Methods	Eigenvalue	$\|Residual\|$	Time(seconds)
PNS	1.12490771776	2734892×10^{-8}	1.5288098
SOAR	1.124907717874	3.8595576×10^{-7}	4.88283
Lanczos	1.1249077177	2.1449555×10^{-7}	17.30051
Jacobi-Davison	-1.236321656	1.2208×10^{-11}	12.90128

2. Sparse symmetric positive definite matrices M and C, a symmetric negative definite matrix K (SPRANDSYM) with an initial vector e_1.

   TABLE 4.10: $n = 500$, interval $[2.5, \ 2.7]$, shift = 2.6

Methods	Eigenvalue	$\|Residual\|$	Time(seconds)
PNS	2.64285216201386	1.405562×10^{-10}	2.1684139
SOAR	26428521620	7.91300864993	1.4820095
Lanczos	2.55231673263	4.90611859441	7.3008468
Jacobi-Davison	*0.259886504671*	2.2008×10^{-5}	18.6577

3. In the experiments (3a) and (3b) we use the same matrices with different intervals.

   (3a) Sparse symmetric positive definite matrices M and C, a symmetric negative definite matrix K (SPRANDSYM) with an initial vector e_1.

   TABLE 4.11: $n = 1000$, interval $[1.5, \ 1.7]$, shift = 1.6

Methods	Eigenvalue	$\|Residual\|$	Time(seconds)
PNS	1.5474509664035	8.1415393×10^{-7}	7.86245
SOAR	1.547450966	2.2077×10^{-5}	23.6653
Lanczos	1.54745096636	91905×10^{-5}	110.1367
Jacobi-Davison	-0.91333	1.1973×10^{-9}	58.391

   (3b) Sparse symmetric positive definite matrices M and C, a symmetric negative definite matrix K (SPRANDSYM) with an initial vector e_1.

   TABLE 4.12: $n = 1000$, interval $[2.3, \ 2.5]$, shift = 2.4

Methods	Eigenvalue	$\|Residual\|$	Time(seconds)
PNS	2.33829838914641	4.034473×10^{-8}	11.60647
SOAR	2.338298389	1.4542×10^{-12}	24.5545
Lanczos	2.338298389	5.341×10^{-12}	129.8708
Jacobi-Davison	-0.91333	2.0529×10^{-10}	59.46758

4. Sparse symmetric positive definite matrices M and C, a symmetric negative definite matrix K (SPRANDSYM) with an initial vector e_1.

   TABLE 4.13: $n = 1500$, interval $[2.3, \ 2.5]$, shift = 2.4

Methods	Eigenvalue	$\|Residual\|$	Time(seconds)
PNS	2.35100491486958	4.386931×10^{-8}	27.28
SOAR	*2.35206825045*	0.0650630775	69.20
Lanczos	*2.3591806610*	44.319901271	388.91
Jacobi-Davison	*-0.69409*	8.4576×10^{-10}	6.29

5. Sparse symmetric positive definite matrices M and C, a symmetric negative definite matrix K (SPRANDSYM) with an initial vector e_1.

   TABLE 4.14: $n = 2000$, interval $[2.5, \ 2.7]$, shift = 2.6

Methods	Eigenvalue	$\|Residual\|$	Time(seconds)
PNS	2.551922233	2.355233×10^{-8}	83.320134
SOAR	2.554343874	144.901021	151.5549715
Lanczos	2.550342496	143.262457	1007.033255
Jacobi-Davison	*-0.75258*	4.6789×10^{-9}	35.66

In every table it shows that all methods except the Jacobi-Davidson method compute eigenpairs accurately in the given interval. However, the PNS method is the most efficient one by achieving the target eigenpair with the least computation time. We also observe that as the size of the matrices increases, the computing time of the SOAR or the Lanczos method increases at a much faster rate than the PNS method. The Jacobi-Davidson method converges to an eigenpair outside the interval.

References

[1] Bai, Z., Demmel J., Dongarra J., Ruhe A. and Van Der Vorst H. A.**(2000)**. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. SIAM, Philadelphia.

[2] Bai, Z. and Y. Su, SOAR **(2005)**. A Second-Order Arnoldi Method for the Solution of the Quadratic Eigenvalue Problem. *SIAM J. Matrix Anal. Appl.* , **26**(3)**:** 640-659.

[3] Datta, B.N. **(2010)**. *Numerical Linear Algebra and Application*. SIAM Publication.

[4] Datta, B. N., **(2004)**. *Numerical Methods for Linear Control Systems Design and Analysis*, Academic Press, New York.

[5] Datta, K., Hong Y.P. and Lee R. **(2003)**. A Parametrized Newton's Iteration for computing an eigenpair of a real symmetric matrix in an interval. *Computational Methods in Applied Mathematics*, **39**(4)**:** 517 - 535.

[6] Royden, H.L. **(1968)**. *Real Analysis*. Macmillan Publishing Company, New York.

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

[8] Sleijpen, G.L.G, and Van Der Vorst H.A.**(1996)**. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.* ,**17**(2): 401 - 425.

[9] Sleijpen, G.L.G, and Van Der Vorst H.A.**(2000)**. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM Review* ,**42**(2): 267 - 293.

[10] Tisseur, F. and Meerbergen K.**(2001)**. The quadratic eigenvalue problem. *SIAM REVIEW*, **43**(2): 235 -286.

[11] Thapa Mohan **(2010)**. A New Hybrid Method for Finding Eigenpairs of a Symmetric Quadratic Eigenvalue Problem in an Interval. [PhD Thesis], Northern Illinois University.

❑❑