Journal of Computational Mathematics, Vol.15, No.2, 1997, 138–148.

RQI DYNAMICS FOR NON-NORMAL MATRICES WITH REAL EIGENVALUES*[1)]

D.L. Hu and D.Y. Cai
(*Department of Applied Mathematics of Tsignhua Universty, Beijing 100084*)

Abstract

RQI is an approach for eigenvectors of matrices. In 1974, B.N Parlett proved that it was a "succeessful algorithm" with cubic convergent speed for normal matrices. After then, several authors developed relevant theory and put this research into dynamical frame. [3] indicated that RQI failed for non-normal matrices with complex eigenvalues.

In this paper, RQI fornon-normal matrices with only real spectrum is analyzed. The authers proved that eigenvectors are super-attractive fixed points of RQI. The geometrical and topological behaviours of two periodic orbits are considered in detail.

The existness of three or higher periodic orbits and their geometry are considered in detail.

The existness of three or higher periodic orbits and their geometry are still open and of interest. It will be reported in our forthcomming paper.

1. Introduction

As well known, RQI (Rayleigh Quotient Iteration) is a practical algorithm for eigenvalue problems of symmetric matrices. In 1974, B.N. Parlett proved that the sequence generated by RQI always converges to an eigenvector for almost all of initial vectors if the matrix in question is a normal one. Namely the set of vectors in R^n, for which RQI diverges, has zero measure. Nevertheless, he also pointed out the convergent speed being cubic[1]. In 1989, S. Barttson and J. Smillie considered RQI for symmetric matrix again. They discovered that the dynamics of RQI is, in a sense, similar to that of Morse-Smale diffeomorphism except its discontinuity. In their paper[2] the geometry and topology of initial vectors for which RQI is covergent are characterized.

Based upon discrete dynamical system and bifurcation theory, S. Barttson and J. Smillie, in 1990, constructed an example to show the existness of a nonempty open set of matrices for which RQI strongly fails. This set is refered to be a bad set. Note that the example given by the authers has eigenvalues with nonzero imaginary parts. Comparing with the counterpart of Newton iteration for polynomail equations[4,5], the authers of [3] gave rise to an open question:

* Received Feburary 9, 1995.
1) This author is supported by the State Key Project "Large Scale Scientific Computing" and supported by Reseach Grants Council of the Hong Kong UPGC

Is successful RQI for matrices having only real spectrum?

In this paper, the authers solved this problem partially. In section 2, RQI is overviewed briefly. Then the relationship between it and discrete dynamical system is described. Section 3 is devoted to our new results. Finally, a conjecture is presented.

2. R.Q.I. Algorithm and Discrete Dynamical Systems

R.Q.I. Algorithm is a will-known method for symmetric eigenproblems. In fact, it is nothing but the inverse power iteration with shifts. We summarize RQI briefly as follows:

Let A be a $n \times n$ real matrix, $\rho(x)$ be Rayleigh Quotient defined on $R^n \setminus \{0\}$ as

$$\rho A(x) = \frac{(x, Ax)}{(x, x)}$$

where $(\cdot, \cdot)$ is Euclid inner product.

Algorithm 2.1. (*R.Q.I. Algorithm*)

Step 1. Choose an initial vector $x_0 in R^n \setminus \{0\}$.

Step 2. For $k = 0, 1, 2, \cdots$,

if $(A - \rho(x_k)I)$ is singular

then get an eigenvector and normalize it, Stop

else

$$Y_{k+1} = (A - \rho A(x_k)I)^{-1}(x_k) \equiv F_A(x_k)$$

Step 3. Normalize y_{k+1} and x_{k+1}

Step 4. Go to Step 2.

First of all, well define a discrete dynamical system for RQI. Note that a nonzero vector is an eigenvector of matrix A. if and only if each element in its one-dimensional span is also an eiginvector. The set $\{\alpha x | \alpha \in R\}$ forms a one-dimensional subspace in R^n. All such subspaces compose a manifold of $n-1$ dimension, RP^{N-1}, refered as a projective space. One can view the projective space as providing a space of eigenvector candidates.

It is easy to verify, $\alpha \neq 0$,

$$\begin{aligned} \rho A(\alpha x) &= \rho A(x) \\ F_A(\alpha x) \equiv (A - \rho A(\alpha x)I)^{-1}(\alpha x) &= \alpha(A - \rho A(x)I)^{-1} x = \alpha F_A(x). \end{aligned}$$

RQI defines a smooth map F_A on the subset of RP^{n-1} for which ρA does not yield an eigenvalue of A. If $\rho A(x)$. In the event that $\rho A(x)$ is a repeated eigenvalue then the dimension of the eigenspace is greater than 1. To have a well-defined iteration we must specify a method for the selection of the particular eigenvector. For dynamical reasons we define $F_A(x)$ to be the one-dimensional subspace spanned by the orthogonal projection of x onto the eigenspace corresponding to $\rho A(x)$. If x is orthogonal to the eigenspace the choice of eiggenvector is dynamically unimportant and we can specify any algorithm for choosing the eigenvector. Of course, the fiscrete dynamical system may be possibly discontinuous with respect to x.

Secondly, from Schur theorem in matrix algebra, we know that there exists a orthogonal matrix Q such that $QAQ^* = T$ for nay $n \times n$ matrix A whose eigenvalues are real, where T is an upper triangular matrix. It is also straightforward to verify:

Proposition. *Let A be an $\times n$ matrix and $x \in R^n \setminus \{0\}$, then*

1. $\rho A(x) = \rho T(Qx)$,
2. $F_T \bar{Q} = \bar{Q} F_A$

holds, where $\rho_\alpha(\cdot)$ and $F_\beta(\cdot)$ are defined as before. $\bar{Q}$ is the induced operator of Q on RP^{n-1}.

From the proposition, F_A and F_T are topological conjugate mappings.The dynamicsof F_A and F_T are identical globally. Therefore, without loss of generality, we assume the matrix in question beign an upper triangular one thoughout this paper.

3. Dynamics of RQI for Non-normal Matrices

Let

$$A = \begin{pmatrix} \lambda & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda \end{pmatrix}$$

be an upper triangular matrix and $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. Obviously, the spans of eigenvectors are the fixed points of the induced operator, F_A, of RQI on RP^{n-1}. In general, to clearly observe the dynamics of F_A, we study the projective map f of F_A on the chart $x_j = 1$. In fact, if we denote $F_A(x)$ as $F(x) = (F_1(x), \cdots, F_n(x))^T$, then $F(x) = (A - \rho(x)I)^{-1}x$, $x \in R^n$. On the chart $x_j = 1$, we have

$$\begin{aligned} f(x) &= \Big(\frac{F_1(x1,d,x_2)}{F_j(x_1,1,x_2)}, \cdots, \frac{F_{j-1}(x_1,1,x_2)}{F_j(x_1,1,x_2)}, \frac{F_{j+1}(x_1,1,x_2)}{F_j(x,1,x_2)}, \frac{F_n(x_1,1,x_2)}{F_j(x_1,1,x_2)}\Big)^T \\ &= \Big(\frac{\mathcal{F}_1(x_1,1,x_2)^T}{F_j(x_1,1,x_2)}, \frac{\mathcal{F}_2(x_1,1,x_2)^T}{F_j(x_1,1,x_2)}\Big)^T \triangleq (f_1(x)^T, f_2(x)^T)^T \end{aligned}$$

where

$$\begin{aligned} &x_1 \in R^{j-1}, \qquad x_2 \in R^{n-j}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in R^{n-1} \\ &\mathcal{F}_1(x_1,1,x_2) = (F_1(x_1,1,x_2), \cdots, F_{j-1}(x_1,1,x_2))^T \\ &\mathcal{F}_2(x_1,1,x_2) = (f_{j+1}(x_1,1,x_2), \cdots, F_n(x_1,1,x_2))^T. \end{aligned}$$

We assume that

$$A = \begin{pmatrix} B & \alpha & E \\ 0 & \lambda_j & \beta^T \\ 0 & 0 & C \end{pmatrix}$$

where B, C, E, α, β have their proper dimensions. Thus for $\forall x = \binom{x_1}{x_2} \in R^{n-1}$,

$$\rho(x) = \frac{1}{1 + x^T x}(x_1^T B x_1 + x_1^T E x_2 + x_1^T \alpha + \lambda_j + \beta^T x_2 + x_2^T C x_2)$$

and

$$\begin{pmatrix} B-\rho I & \alpha & E \\ 0 & \lambda_j-\rho & \beta^T \\ 0 & 0 & C-\rho I \end{pmatrix} \begin{pmatrix} \mathcal{F}_1 \\ F_j \\ \mathcal{F}_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 1 \\ x_2 \end{pmatrix}.$$

So we have

$$F_j = \frac{1-\beta^T\mathcal{F}_2}{\lambda_j-\rho} \Longrightarrow \begin{cases} (Bg\rho I)f_1 + \alpha + Ef_1 = \dfrac{\lambda_j-\rho}{1-\beta^T\mathcal{F}_2}x_1 \\ (C-\rho I)f_2 = \dfrac{\lambda_j-\rho}{1-\beta^T\mathcal{F}_2}x_2 \end{cases}.$$

Let

$$\tilde{A} = \begin{pmatrix} B & E \\ 0 & C \end{pmatrix}, \quad P\tilde{a} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, \tilde{\beta} = \begin{pmatrix} & \alpha \\ & \beta \end{pmatrix}$$

Then

$$\rho(x) = \frac{1}{\Delta}(x^T\tilde{A}x + \lambda_j + \tilde{\beta}^Tx), \quad \text{where} \quad \Delta = 1 + x^Tx.$$

and

$$(\tilde{A}-\rho I)f + \tilde{\alpha} = \frac{\lambda_j\rho}{1-\beta^T\mathcal{F}_2}. \tag{3.1}$$

The equality (3.1) is very important and will be used in the proofs of the following theorem 3.1 and theorem 3.3. Assume that ξ_j is an eigenvector of A corresponding to λ_j, and x_j^* is the projective coordinate of ξ on the chart $x_j = 1$. It is clear that f is smooth at x_j^*, then we have

Theorem 3.1. *$Df(x_j^*) \equiv 0$, where $Df(\cdot)$ is the Jacobian matrix of $f(x)$.*

Remark*:* This theorem asserts that all the eigenvectors of A must be the super-attractive fixed points of RQI.

Proof. Since $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, we assume that $\xi_j = \begin{pmatrix} \zeta \\ 1 \\ 0 \end{pmatrix} \in R^n$. So $s_j^* = \begin{pmatrix} \zeta \\ 0 \end{pmatrix} \in R^{n-1}$ on the chart $x_j = 1$ is a fixed point of the map f, then we have

$$f(x_j^*) = x_j^*, \qquad \rho(x_j^*) = \lambda_j \tag{3.2}$$

Differentiating both sides of (3.1) with respect to x gives

$$(\tilde{A}-\rho I)\cdot Df - f\cdot\Delta\rho = \frac{\lambda_j-\rho}{1-\beta^T\mathcal{F}_2}I + \frac{x}{(1-\beta^T\mathcal{F}_2)^2}[(\lambda_j-\rho)\beta^T\cdot D\mathcal{F}_2 - \Delta\rho\cdot(1-\beta^T\mathcal{F}_2)] \tag{3.3}$$

where

$$\Delta\rho(x) = \frac{1}{\Delta}[(\tilde{A}+\tilde{A}^T-2\rho I)x + \tilde{\beta}]^T \tag{3.4}$$

$$F\mathcal{F}_2(x) = (C-\rho i)^{-1}[(0I) + \mathcal{F}_2\cdot\Delta\rho(x)] \tag{3.5}$$

Note that $\mathcal{F}_2(x_j^*) = 0$ and (3.2), hence we have

$$(\tilde{A}-\lambda_j I) \equiv 0.$$

Finally we get

$$Df(x_j^*) \equiv 0.$$

In order to characterize the orbits of period 2 of RQI, we prove two lemmas.

Lemma A. *Let $Ax = \lambda_i x$, $Ay = \lambda x$, $Ay = \lambda_j y$ and $\|x\|_2 = \|y\|_2 = 1$ then*

$$\rho A(x \pm y) = \frac{1}{2}(\lambda_i + \lambda_j), \qquad \textit{if} \;\; x \neq y.$$

Proof. The proof is straighforward and omitted.

For the simplification of our proofs, a hypothesis is needed. We call it as a standard hypothesis.

Standard hypothesis: Assume that A a $n \times n$ matrix with distinct eigenvalues $\lambda_1, \cdots, \lambda_n$ and that $\frac{1}{2}(\lambda_i + \lambda_j) \neq \lambda_k$, for different i, j, k.

For the orbits of period 2 of RQI, we have

Lemma B. *Let A be a $n \times n$ upper triangular matrix with only real eigenvalues and satsfy the standard hypothesis, then $x_i \pm x_j$ for $1 \leq i \neq j \leq n$, are the orbits of period 2 of RQI, where x_i for $1 \leq i \leq n$ are the unit eigenvectors of A, i.e., $Ax_i = \lambda_i x_i$ and $\|x_i\|_2 = 1$, $\forall i$.*

Proof. By the standard hypothesis, the matrix $(A - \rho(x_i \pm x_j)I)$ is invertible. Since $(A - \rho(X_i \pm x_j)I)^{-1}$. $(A - \rho(x_i \pm x_j)I) = I$, so it is easy to obtain that

$$\begin{cases} (A - \rho(x_i \pm x_j)I)^{-1} x_i = \dfrac{x_i}{\lambda_i - \rho(x_i \pm x_j)} \\ (A - \rho(x_i \pm x_j)I)^{-1} x_j = \dfrac{x_j}{\lambda_j - \rho(x, \pm x_j)} \end{cases}.$$

Adding the above equalities gives

$$f(x_i \pm x_j) = \frac{x_i}{\lambda - \rho(x_i \pm x_j)} \pm \frac{x_j}{\lambda_j - \rho(x_i \pm x_j)}.$$

From Lemma A we know $\rho(x_i \pm x_j) = \frac{1}{2}\; \lambda_i + \lambda_j)$. Consequently, we obtain

$$F(x_i \pm x_j) = \frac{2}{\lambda_i - \lambda_j}(x_i \mp x_j) \tag{3.6}$$

Recall that the scaler $\dfrac{2}{\lambda_i - \lambda_j}$ can be omitted, because that a one-dimensional span is viewed as a point in the projective space. (3.6) indicates that the bisectors of every pair of eigenvectors are the orbiys of period 2 of RQI.

Theorem 3.2. *Let A be a $n \times n$ upper triangular matrix with real distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, and satisfy the standard hypothesis, then any 2-periodic orbit of RQI must be the bisector of certain pair of eigenvectors of A.*

Proof. Since $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, so $\{\xi_i | \xi_i$ is the unit eigenvector of A corresponding to λ_i, $i = 1, 2, \cdots, n\}$ are linearly independent. Assume $\{c, y\}$ is a 2-periodic orbit of RQI and $\|x\|_2 = \|y\|_2 = 1$, there exists $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ such that

$$x = \sum_{i=1}^n a_i \xi_i, \quad y = \sum_{i=1}^n b_i \xi_i \tag{3.7}$$

Also, there are constants $\mu \neq 0, \theta \neq 0$ such that

$$\begin{cases} (A-\rho(x)I)y = \mu x & (3.8) \\ (A-\rho(y)I)x = \theta y & (3.9) \end{cases}$$

(Recall that x and y denote the entries of one dimensional span in the projective space.) Substituting (3.7) into (3.8—3.9) gives that

$$\begin{cases} (A-\rho(x)I)\cdot \sum_{i=1}^{n} b_i\xi_i = \mu \sum_{i=1}^{n} a_i\xi_i \\ (A-\rho(y)I)\cdot \sum_{i=1}^{n} a_i\xi_i = \theta \sum_{i=1}^{n} b_i\xi_i \end{cases}.$$

By the linear independendence of $\{\xi_i\}_{i=1}^n$, we have

$$\begin{cases} b_i(\lambda_i-\rho(x)) = \mu a_i \\ a_i(\lambda_i-\rho(y)) = \theta b_i \end{cases} \quad \text{for} \quad i = 1, 2, \cdots, n. \tag{3.10}$$

If there exists $i < j < k$ s.t. $a_m \neq 0$, $M = i, j, k$, then we have

$$(\lambda_m-\rho(x))(\lambda_m-\rho(y)) = \mu\theta, \qquad m = i, j, k.$$

Thus

$$\begin{cases} \lambda_i^2-(\rho(x)+\rho(y))\lambda_i+\rho(x)\rho(y) = \lambda_j^2-(\rho(x)+\rho(y))\lambda_j+\rho(x)\rho(y) \\ \lambda_i^2-\lambda_i^2-(\rho(x)+\rho(y))\lambda_i+\rho(x)\rho(y) = \lambda_k^2-(\rho(x)+\rho(y))\lambda_k+\rho(x)\rho(y) \end{cases}.$$

As a result,

$$\begin{cases} \rho(x)+\rho(y) = \lambda_i+\lambda_j \\ \rho(x)+\rho(y) = \lambda_i+\lambda_k \end{cases}$$

So we can get $\lambda_k = \lambda_j$. Thes contradicts the assumption.

Now we can say that there are two nonzero pairs $\{a_i, a_j\}$, $\{b_i, b_j\}$ such that

$$x = a_i a_i\xi_i + a_j\xi_j, \qquad y = b_i\xi_i + b_j\xi_j.$$

According to (3.10), we have

$$\begin{cases} b_i((\lambda_i\rho(x)) = \mu a_i, & b_j(\lambda_j-\rho(x)) = \mu a_j \\ a_i((\lambda_i\rho(y)) = \theta b_i, & a_j(\lambda_j-\rho(y)) = \theta b_j \end{cases} \tag{3.11}$$

and

$$\rho(x)+\rho(y) = \lambda_i+\lambda_j. \tag{3.12}$$

We make the inner product on both sides of (3.8) with y, then we can obtain

$$\rho(y)+\rho(x) = \mu(x, y) \equiv \mu e \qquad \text{where} \quad e = (x, y).$$

Siminarly,

$$\rho(x)\rho(y) = \theta e.$$

Sowe get

$$(\mu+\theta)e = 0.$$

We claim that e must be zero. If not so, then

$$\mu+\theta = 0 \Longrightarrow \theta = -\mu.$$

Since

$$\rho(x) = (x, Ax) = (a_i\xi + a_j\xi_j, A(a_i\xi_i + a_j\xi_j)) = \lambda_i a_i^2 + (\lambda_i+\lambda_j)a_ia_j(\xi_i,\xi_j) + \lambda_j a_j^2$$

and

$$\|x\|_2 = 1 \Longrightarrow a_ia_j(\xi_i,\xi_j) = \frac{1}{2}(1 - a_i^2 - a_j^2).$$

So we can get

$$\rho(x) = \frac{1}{2}(\lambda_i+\lambda_j) + \frac{1}{2}(\lambda_i-\lambda_j)(a_i^2 - a_j^2). \tag{3.13}$$

In the sininar way, we can get

$$\rho(y) = \frac{1}{2}(\lambda_i+\lambda_j) + \frac{1}{2}(\lambda_i-\lambda_j)(b_i^2 - b_j^2).$$

Here

$$\rho(x) + \rho(y) = \lambda_i + \lambda_j + \frac{1}{2}(\lambda_i+\lambda_j)(a_i^2 + b_i^2 - a_j^2 - b_j^2).$$

Comparing with (3.12) gives that

$$a_i^2 + b_i^2 = a_j^2 + b_j^2. \tag{3.14}$$

According to (3.11) we have

$$\begin{cases} \mu a_i^2 = a_ib_i(\lambda_i - \rho(x)), & \mu a_j^2 = a_jb_j(\lambda_j - \rho(x)) \\ \theta b_i^2 = a_ib_i(\lambda_i - \rho(y)), & \theta b_j^2 = a_jb_j(\lambda_j - \rho(y)) \end{cases}.$$

Since $\theta = -\mu$ and (3.14), so we can get

$$a_ib_i(\rho(y) - \rho(x)) = a_jb_j(\rho(y) - \rho(x)) \Longrightarrow a_ib_i = a_jb_j. \tag{3.15}$$

Since $\rho(y) \neq \rho(x)$, ogherwise $\rho = 0$ from (3.8) or (3.9).

From (3.11) we can obtain

$$a_ib_j(\lambda_j - \rho(x)) = a_jb_i(\lambda_i - \rho(x)).$$

Substituting (3.15) into the above equality gives that

$$(a_i^2 - a_j^2)\rho(x) = \lambda_j a_i^2 - \lambda_i a_j^2.$$

Again, we substitute (3.13) into the above equality. It shows that

$$\begin{aligned} \lambda_j a_i^2 - \lambda_i a_j^2 &= \frac{1}{2}(\lambda_i+\lambda_j)(a_i^2 - a_j^2) + \frac{1}{2}(\lambda_i-\lambda_j)(a_i^2 - a_j^2)^2 \\ &\Longrightarrow (a_i^2 - a_j^2)^2 + a_i^2 + a_j^2 = 0 \Longrightarrow a_i = a_j + 0 \Longrightarrow x = 0. \end{aligned}$$

This contradicts the assumption of x.

Hence $e = 0$, thus we can get

$$\rho(x) = \rho(y) = \frac{1}{2}(\lambda_i + \lambda_j).$$

According to (3.13), we have $a_i = \pm a_j$. Recall that the one-dimensional span is one point of RP^{n-1}, we can obtain $b_i = \mp b_j$.

So the result is followed.

Lemma B and Theorem 3.2 depict the geometry of 2-periodic orbits of RQI. For the dynamics of 2-periodic orbits we give

Theorem 3.3. *Let A be a $n \times n$ upper triangular matrix with only real eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, and satisfy the standard hypothesis. Assume $P_{ij}^{\pm} (i < j)$ be a 2-periodic orbit of RQI, which is given by the bisector of ξ_i and ξ_j, where ξ_i and ξ_j are two eigenvectors of A corresponding to λ_j respectively, then*

$$\begin{aligned} \dim(W_l^u(P_{ij}^{\pm})) &= j - i \\ \dim(W_l^s(P_{ij}^{\pm})) &= n - 1j + i \end{aligned}$$

where dim $(\cdot)$ denotes the dimension of a manifold, $W_l^u(P_{ij}^{\pm}$ and $W_l^s(P_{ij}^{\pm})$ are the local unstable and stable manifolds of $P_{ij}^{\pm}$ respectively.

Proof. Assume that

$$\begin{aligned} \xi_i =&(x_1, \cdots, x_{i-1}, x_i, 0, \cdots, 0)^T, \qquad \|\xi\|_2 = 1 \\ \xi_j =&(y_1, \cdots, y_{j-1}, y_j, 0, \cdots, 0)^T, \qquad \|\xi\|_2 = 1 \end{aligned}$$

Obviously, we can say $x \neq 0$, $y_j \neq 0$, and $P_{ij}^{\mp} = \xi_j \pm \xi_i$.

Since $y_j \neq 0$, so $P_{ij}^{\pm}/y_j$ must be on the chart $x_j = 1$. Thus we consider the problem on the chart $x_j = 1$:

Without loss of generality, we assume that $P_{\pm}$ (corresponding to $P_{ij}^{\pm}$) is a 2-periodic point of f.

Let $P_{\pm} = (p\pm, 0)^T$, where $p\pm = (y_1 \pm x_1, \cdots, y_i \pm x_i, y_{i+1}, \cdots, y_{j-1})^T / y_j$. Then we have

$$\rho(P_{\pm}) = \frac{1}{2}(\lambda_i + \lambda_j) \overset{\triangle}{=} \rho, \quad f(P_{pm}) = P_{\mp}.$$

Note that $\mathcal{F}_2(P_{\pm}) = 0$, hence according to (3.3—3.5) we can obtain

$$(\tilde{A} - \rho I) \cdot Df(P_{\pm}) = (\lambda_j - \rho)I + P_{\mp} \cdot \nabla\rho(P_{\pm}) + P_{\pm} \cdot [(\lambda - j\rho)\beta^T \cdot D\mathcal{F}_2(P_{\pm}) - \nabla\rho(P_{\pm})]$$

where

$$\begin{aligned} D\mathcal{F}_2(P_{\pm}) &= (C - \rho I)^{-1} \cdot (0 \quad I) \\ \nabla\rho(P_{\pm}) &= \frac{1}{\triangle_{\pm}}[(\tilde{A} + \tilde{A}^T - 2\rho I)P_{\pm} + \tilde{\beta}]^T \quad (\triangle_{\pm} \text{ denotes } \triangle(P_{\pm}) = 1 + P_{\pm}^T p_{\pm}) \end{aligned}$$

Sowe have

$$(C - \rho I) \cdot df_2(P_{\pm}) = (\lambda_j - \rho)(0 \quad I). \tag{3.16}$$

and

$$\begin{aligned}&(B-\rho I)\cdot Df_1(P_\pm)+E\cdot Df_2(P_\pm)\\=&(\lambda_j-\rho)(I\quad 0)+p_\mp\cdot\nabla_\rho(P_\pm)+p_\pm\cdot[(\lambda_j-\rho)\beta^T\cdot(C-\rho I)^{-1}(0\quad I)-\nabla_\rho(P_\pm)].\end{aligned}\tag{3.17}$$

From (3.16) we can get

$$Df_2(P_\pm)=(\lambda_j-\rho)\cdot(0\quad (C-\rho I)^{-1}).\tag{3.18}$$

So we know that the latter $(n-j)$ eigenvalues of $Df(P_\pm)$ are

$$(\lambda_j-\rho)\cdot(\lambda_{k+1}-\rho)^{-1},\quad k=j,j+1,\cdots,n-1.$$

Hence we only need to inspect the matrix $G(P_\pm)\triangleq\dfrac{\partial f_1}{\partial x_1}(P_\pm)$.

According to (3.17) and (3.18), we can get

$$\begin{aligned}&(B-\rho I)\cdot G)P_\pm)=(\lambda_j-\rho)I+\frac{1}{\triangle_{pm}}(p_{mp}-p_\pm)[(B+B^T-1\rho I)p_{pm}+\alpha]^T\\ \Longrightarrow\ &G(P_\pm)=(\lambda_j-\rho)(B-\rho I)^{-1}+\frac{1}{\triangle_\pm}(B-\rho I)^{-1}(p_\pm-p_\pm)[(B+B^T-2\rho I)p_\pm+\alpha]^T\end{aligned}$$

From (3.1) we know that

$$(B-\rho I)p_{mp}+\alpha=(\lambda_j-\rho)p_\pm\Longrightarrow(B-\rho I)^{-1}(p_\mp-p_\pm)=\frac{p_\pm-p_{mp}}{\lambda_j-\rho}.$$

Thus we can obtain

$$G(P_\pm)=(\lambda_j-\rho)(B-\rho I)^{-1}+\frac{1}{\triangle_\mp}\frac{p_\pm-p_\mp}{\lambda_j-\rho}[(B^T-\rho I)p_\pm+(\lambda_j-\rho)p_\mp]^T.$$

Let

$$B=\begin{pmatrix}U & S\\ 0 & V,\end{pmatrix}\quad G=\begin{pmatrix}G_{11} & G_{12}\\ G_{21} & G_{22}\end{pmatrix}$$

where $U,G_{11}\in R^{i\times i},S,G_{12}\in R^{i\times(j-1-i)}$, $G_{21}\in R^{(j-1-i)\times i},G_{22}\in R^{(j-1-i)\times(j-1-i)}$. Since

$$p_\pm-p_\mp=\frac{\pm 2}{y_j}(x_1,\cdots,x_i,0,\cdots,0)^T\triangleq(q_\pm^T,0)^T\in R^{j-1}.\tag{3.19}$$

Then we have

$$(G_{21}\quad G_{22})=(\lambda_j-\rho)(0\quad (V-\rho I)^{-1}).$$

So we kmow that the latter $(j-1-i)$ eigenvalues of $G(P_\pm)$ are

$$(\lambda_j-\rho)\cdot(\lambda_k-\rho)^{-1},\quad k=i+1,\cdots,j-1.$$

Hence only the matrix G_{11} needs too be considered. We denote $\tilde p_\pm$ as the former i entries of $p_\pm$, then

$$G_{11}=(\lambda_j-\rho)(U-\rho I)^{-1}+\frac{1}{\triangleq}\frac{q_\pm}{\lambda_j-\rho}[(U^T-\rho I)\tilde p_\pm+(\lambda_j\rho)\tilde p_\mp]^T.$$

Let

$$(\lambda_j-\rho)(U-\rho I)^{-1}=\begin{pmatrix} H & r \\ 0 & -1 \end{pmatrix}$$
$$(U^T-\rho I)\tilde{p}_{\mp}+(\lambda_j-\rho)\tilde{p}_{\mp}\triangleq t_{\pm}=\begin{pmatrix} \tilde{t}_{\pm} \\ t_i \end{pmatrix}$$
$$\frac{1}{\triangle_{\pm}}\frac{1}{\lambda_j-\rho}\triangleq d_{\pm},\quad q_{\pm}\triangleq\begin{pmatrix} \tilde{q}_{\pm} \\ q_i \end{pmatrix}.$$

Thus we have

$$G_{11}=\begin{pmatrix} H & r \\ 0 & -1 \end{pmatrix}+d_{\pm}\begin{pmatrix} \tilde{q}_{\pm} \\ q_i \end{pmatrix}(\tilde{t}_{\pm}^T \quad t_i)=\begin{pmatrix} H+d_{\pm}\tilde{q}_{\pm}\tilde{t}_{\pm}^T & r+d_{\pm}\tilde{q}_{\pm}t_i \\ d_{\pm}q_i\tilde{t}_{\pm}^T & -1+d_{\pm}q_it_i \end{pmatrix}.$$

Since $\xi_i=(x_1,\cdots,x_i,0,\cdots,0)^T$ is a unit eigenvector of A corresponding to λ_i, from (3.19) we can have

$$\begin{aligned} Uq_{\pm}=\lambda_iq_{\pm}\Longrightarrow&\frac{q_{\pm}}{\lambda_i-\rho}=(U-\rho I)^{-1}q_{\pm}=\frac{1}{\lambda_j-\rho}\begin{pmatrix} H & r \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \tilde{q}_{\pm} \\ q_i \end{pmatrix} \\ \Longrightarrow& H\tilde{q}_{\pm}+rq_i=-\tilde{q}_{\pm}. \end{aligned}\tag{3.20}$$

Note that $q_i=\pm\dfrac{2x}{y_j}\neq 0$, using (3.20) we can get

$$\begin{pmatrix} I & -\dfrac{\tilde{q}_{\pm}}{q_i} \\ 0 & 1 \end{pmatrix}\cdot G_{11}\cdot\begin{pmatrix} I & \dfrac{\tilde{q}_{\pm}}{q_i} \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} H & 0 \\ d_{\pm}q_i\tilde{t}_{\pm}^T & -1+d_{\pm}q_{\pm}^Tt_{\pm} \end{pmatrix}.$$

Cleearly the eigenvalues of H are

$$(\lambda_j-\rho)\cdot(\lambda_k-\rho)^{-1},\quad k=1,\cdots,i-1.$$

Note that $q_+=-q_-$, then we can obtain that the eigenvalues of $Df(P_+)\cdot Df(P_-)$ are

$$\begin{cases} \mu_k=\Big(\dfrac{\lambda_j-\lambda_i}{2\lambda_k-\lambda_i-\lambda_j}\Big)^2, & k=1,\cdots,i-1,i+1,\cdots,j-1. \\ \mu_k=\Big(\dfrac{\lambda_j-\lambda_i}{2\lambda_{k+1}-\lambda_i-\lambda_j}\Big)^2, & k=j,\cdots,n-1. \\ \mu_i=(-1+d_+q_+^Tt_+)\cdot(-1+d_-q_-^Tt_-). \end{cases}$$

According to (3.19), we have

$$\begin{aligned} \mu_i=&\Big[-1+\frac{1}{\triangle_+}\frac{(p_+-p_-)^T}{\lambda_j-\rho}((B^T-\rho I)p_++(\lambda_j-\rho)p_-)\Big] \\ &\times\Big[-1+\frac{1}{\triangle_-}\frac{(p_--p_+)^T}{\lambda_j-\rho}((B^T-\rho I)p_-+(\lambda_j-\rho)p_+)\Big]. \end{aligned}$$

Let

$$\tilde{x}=(x_1,\cdots,x_i,0,\cdots,0)^T\in R^{j-1},\tilde{y}+(y_1,\cdots,y_{j-1})^T\in R^{j-1}.$$

Then

$$p_{\pm} = (\tilde{y} \pm \tilde{x})/y_j.$$

Note that $B\tilde{x} = \lambda_i \tilde{x}$ and $\|\tilde{x}\|_2 = \|\xi_i\|_2 = 1$, $\|\tilde{y}\|_2 < \|\xi_J\|_2 = 1$, after a vast and glgebraic calculation, we obtain

$$\mu_i = 1 + \frac{8}{1 - (\tilde{x}^T \tilde{y})^2}.$$

It is easy to verify that

$$\begin{cases} \mu_k > 1, & k = i, i+1, \cdots, j-1. \\ \mu_k < 1, & k = 1, \cdots, i-1, j, \cdots n-1. \end{cases}$$

So the conclusion is followed

The conclusion is the same as that of the case of diagonal matrix, but its peoof is more complex than that one. This indicates the complexity of RQI dynamics for nnnormal matrices with real eigenvalues. (The proof of the case of diagonal matrix see [2]).

Also, we designed algorithms for exploring the orbits of three or higher period or attractive spurious eigenvectors. In any they have not been found out as yet. This fact motivates us to propose a conjecture as below.

Conjecture: RQI for non-normal matrices with real eigenvalues is successful.

Acknowledgement: D.Y. Cai is grateful to professor H.C. Huang for several helpful discussions and comments while he visited the Mathematical Department of HKBC.

References

[1] B. Parlett, The rayleigh quotient iteration and some generalization for nonnormal matrices, *Math. Comp.,* 28(1974), 679–693.

[2] S. Batterson and J. Smillie, The dynamics of rayleigh quotient iteration, *SIAM J. Numer. Anal.,* 26(1989), 624–636.

[3] S. Batterson and J. Smillie, Rayleigh quotient iteration for nonsymmetric matrices, *Math. Comp.,* 55(1990), 169–178.

[4] M. Hurley, Multiple attractors in Newton's Method, *Ergodic Theory and Dynamical Systems,* 6(1986), 561–569.

[5] M. Hurley and C. Martin, Newton's algorithm and chaotic dynamical systems, *SIAM J. Math. Anal.,* 15(1984), 238–252.

[6] R. Devaney, An introduction to chaotic dynamical systems, Benjamin-Cummings, *Menlo Park,* CA, 1986.