RAYLEIGH QUOTIENT ITERATION
FOR NONSYMMETRIC MATRICES

STEVE BATTERSON AND JOHN SMILLIE

ABSTRACT. Rayleigh quotient iteration is an iterative algorithm for the calculation of approximate eigenvectors of a matrix. Given a matrix, the algorithm supplies a function whose iteration of an initial vector, \( v_0 \), produces a sequence of vectors, \( v_n \). If the matrix is symmetric, then for almost any choice of \( v_0 \) the sequence will converge to an eigenvector at an eventually cubic rate. In this paper we show that there exist open sets of real matrices, each of which possesses an open set of initial vectors for which the algorithm will not converge to an eigenspace. The proof employs techniques from dynamical systems and bifurcation theory.

1. Introduction

One measure of the success of an iterative algorithm is the probability that a random choice of initial vector will produce a sequence that converges to a solution. We say that an algorithm succeeds provided the set of initial points for which the sequence of iterates converges to a solution is a set of full measure. An algorithm which does not succeed is said to fail. Rayleigh quotient iteration is an iterative algorithm for finding approximate eigenvectors of a matrix. If the matrix is symmetric, then Rayleigh quotient iteration succeeds (see [12, 3]). Parlett has asked whether the algorithm (or some modification) succeeds for nonsymmetric matrices. We will show that there exists an open set of nonsymmetric real matrices for which the algorithm fails. Since the characteristic polynomials of our counterexamples will have both real and complex roots, we begin with a discussion of invariant subspaces that are reasonable targets for an eigenvector algorithm.

Let \( A \) be an \( n \times n \) matrix with real entries. Associated with each real eigenvalue, \( \lambda \), of \( A \) is an eigenspace and a generalized eigenspace. The eigenspace is the largest invariant subspace on which the restriction of \( A \) can be represented as \( \lambda I \). The generalized eigenspace is the maximal invariant subspace for which \( \lambda \) is the only eigenvalue of the restriction. If \( a \pm bi \) is a conjugate pair of solutions to the characteristic equation, then there are analogous...
maximal subspaces. Corresponding to the eigenspace there is the maximal invariant subspace for which the restriction of $A$ can be represented by a direct sum of copies of the $2 \times 2$ matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. Similarly, there is a subspace analogous to the generalized eigenspace of the real case.

Now suppose that we are given an iterative eigenvector algorithm. In decreasing order of desirability we might hope that the algorithm would converge to:

1. an eigenvector,
2. the one-dimensional span of an eigenvector,
3. an eigenspace, or
4. a generalized eigenspace.

For a symmetric matrix, Rayleigh quotient iteration succeeds with the solution having the form 2R above. If the matrix is nonsymmetric and the characteristic polynomial has complex roots, then it is perhaps unreasonable to expect an eigenvalue algorithm to converge to one of the above types of solution. For each of the four types of solutions there is a complex analogue (1C–4C, where the space referred to in 2C is two-dimensional). An algorithm is said to weakly succeed for a matrix $A$ provided that the set of initial points for which the sequence converges to either 4R or 4C is a set of full measure. An algorithm which does not weakly succeed is said to strongly fail.

Now we shift the problem to projective space. Note that a nonzero vector is an eigenvector of $A$ if and only if each element in its one-dimensional span is also an eigenvector. Thus, one can view projective space, $\mathbb{R}P^{n-1}$, as providing a space of eigenvector candidates. An iterative algorithm which preserves one-dimensional subspaces induces a dynamical system on $\mathbb{R}P^{n-1}$. In the Rayleigh quotient iteration (RQI) algorithm the map $F_A$ is obtained from the map on $\mathbb{R}^n$ given by $F_A(x) = (A - \rho_A(x)I)^{-1}(x)$, where $\rho_A(x) = \langle Ax, x \rangle / \langle x, x \rangle$. It is shown in [14] that $F_A$ is well defined and that RQI succeeds in the sense of 2R for a $2 \times 2$ matrix with real eigenvalues. There is an incorrect argument in [14] for the case of a $2 \times 2$ matrix for which the roots of the characteristic polynomial are complex. In any event, the algorithm trivially succeeds in the sense of 2C for such a matrix. In §3 of this paper we prove the following theorem, which was announced in [4].

**Theorem B.** For each $n \geq 3$ there is a nonempty open set of matrices for which Rayleigh quotient iteration strongly fails.

In §2 it is shown that there is a one-parameter family of $2 \times 2$ matrices for which $F_A$ has a period doubling bifurcation. Theorem B is proved in §3 using another bifurcation argument. Section 4 contains a discussion of further problems.

2. Period Doubling

In this section we will investigate the dynamics of $2 \times 2$ matrices whose characteristic polynomials have complex roots. Note that for these matrices, $\mathbb{R}^2$ is a
subspace of type $2C$, and thus the algorithm succeeds trivially regardless of the precise dynamics of $F_A$. Nevertheless, an understanding of these dynamics is interesting and will be useful in §3.

Define a one-parameter family of matrices by

$$A_c = \begin{bmatrix} 0 & -c \\ 1/c & 0 \end{bmatrix}.$$

Note that the map is defined on $\mathbb{RP}^1$, which is homeomorphic to the circle.

For the following theorem, let $a = \sqrt{3} - \sqrt{5}/\sqrt{2} \approx .618$. Note that $a < 1/\sqrt{2} \approx .707$ and $a$ is a root of $c^2 - 3c + 1$.

**Theorem A.** $F_{A_c}$ has a smooth inverse when $1/\sqrt{2} < c < 1$, has a continuous inverse which is not smooth when $c = 1/\sqrt{2}$, and is a (noninjective) endomorphism for $c < 1/\sqrt{2}$. The dynamics of $F_{A_c}$ depend on $c$ as follows:

- a. At $c = 1$ the map is a rotation of period two.
- b. If $1/\sqrt{2} < c < 1$ the map has a periodic set consisting of a sink of period two and a source of period two.
- c. At $c = \sigma$ a period-two sink bifurcates into a period-four sink and a period-two source.
- d. If $c$ is sufficiently small ($c < .271$), then the map exhibits “chaos” in the sense that it has periodic points of infinitely many different periods.

**Note.** In case (a), $A_c$ is normal.

The projective space $\mathbb{RP}^1$ is covered by two standard coordinate charts. One chart is given by mapping $y$ to the one-dimensional subspace spanned by $(1, y)$. We call this chart the $y$-chart. The image of this chart is the complement of the line spanned by $(0, 1)$. The $x$-chart is defined analogously. It is often convenient to describe functions on projective space in terms of these charts.

**Proof of Theorem A.** The following computations are straightforward:

$$\rho_A \begin{bmatrix} 1 \\ y \end{bmatrix} = \begin{bmatrix} y(1 - c^2) \\ c(1 + y^2) \end{bmatrix},$$

$$(A - \rho_A I)^{-1} \begin{bmatrix} 1 \\ y \end{bmatrix} = \frac{1}{\rho^2 + 1} \begin{bmatrix} -\rho + cy \\ -1/c - \rho y \end{bmatrix}.$$

In terms of the $y$-chart the map (suppressing the $A$ in the notation) is given by

$$F_c(y) = \frac{1 + (2 - c^2)y^2}{y(1 - 2c^2 - c^2y^2)}.$$

When $c = 1$, $F_c$ is the map $y \to -1/y$. This is a rotation of period two on $\mathbb{RP}^1$. This proves assertion (a).
The derivative of $F_c$ is given by

$$F_c'(y) = \frac{c^2(2 - c^2)(y^2 + c^{-2})(y^2 + (2c^2 - 1)/(2 - c^2))}{y^2(1 - 2c^2 - c^2y^2)^2}.$$

The form of the graph of $F_c$ depends on whether $c$ is greater or less than $1/\sqrt{2}$. Since $F_c' \geq 0$ and $F_c$ has topological degree one, it follows that $F_c$ is a homeomorphism for $1/\sqrt{2} \leq c \leq 1$. For $c < 1/\sqrt{2}$, $F_c$ is not injective because $F_c$ has a change in sign. To determine whether $F_c$ is a diffeomorphism in the former case, it remains to check the derivatives at 0 and $\infty$. Note that these points are in a periodic orbit of period two. We have

$$(F_c^2)'(0) = 5 - 2(c^2 + c^{-2}).$$

$(F_c^2)'(0)$ is a strictly increasing function of $c$ for $0 < c \leq 1$. It assumes the value $-1$ at $c = \sigma$, the value 0 at $c = 1/\sqrt{2}$, and the value 1 at $c = 1$. This shows that $F_c$ is a diffeomorphism precisely when $c > 1/\sqrt{2}$. Furthermore, 0 changes from a period-two sink to a period-two source as $c$ decreases through $\sigma$. Figure 1 shows the behavior of $F_c$ for $c$ in the range $1/\sqrt{2} < c < 1$.

\[\text{Figure 1. } F_c \text{ for } 1/\sqrt{2} < c < 1.\]

\[\text{Figure 2. } F_c \text{ for } 0 < c < 1/\sqrt{2}.\]
Figure 2 shows the behavior of $F_c$ for $c$ in the range $0 < c < 1/\sqrt{2}$. The figures on the left are the graphs of $F_c$ with respect to the coordinates given by the $y$-chart. The figures on the right show the topological behavior of $F_c$ as a map from the inner circle to itself. The map is obtained by identifying the inner circle homeomorphically with the outer curve and then radially projecting back onto the inner circle.

If $1 > c \geq 1/\sqrt{2}$, then we know that $F_c$ is a homeomorphism with a sink of period two. It can be shown that an orientation-preserving homeomorphism of the circle cannot have periodic orbits of different periods. An analysis of the equation $F_c^2(y) = y$ shows that there is precisely one additional orbit of period two. The theory of homeomorphisms of the circle implies that this orbit must be a source. The proof of assertion (b) is now complete.

By a standard result in one-dimensional bifurcation theory, in order to prove assertion (c) of the theorem it suffices to show that

$$(F^2_c)'(0) > 0 \quad \text{and} \quad \frac{\partial^2 (F^4_c)}{\partial c \partial y}(y = 0, c = \sigma) < 0$$

(see [6, p. 89]). We have, indeed,

$$(F^2_c)'(0) = 6\sigma^{-4}[16\sigma^{10} - 58\sigma^8 + 84\sigma^6 - 60\sigma^4 + 20\sigma^2 - 2]
= 180 - 72\sqrt{5} \approx 19.003,$$

$$\frac{\partial^2 (F^4_c)}{\partial c \partial y}(y = 0, c = \sigma) = 8\sigma^{-5}[2\sigma^8 - 5\sigma^6 + 5\sigma^2 - 2] = -8\sqrt{5}\sigma^{-1} \approx -28.944.$$

Before proving assertion (d), we introduce some notation for the positive critical value and singularity of $F_c$. Let $m_c = [(2 - c^2)/(1 - 2c)]^{3/2}$ and $\xi_c = c^{-1}(1 - 2c)^{1/2}$. For $c$ small, $m_c < \xi_c$. Assume that $c$ is sufficiently small that this inequality holds ($c < .271$). On $\mathbb{RP}^1$ consider the closed arc, $I_c$, whose endpoints are represented by 0 and $\xi_c$. We have $F_c^3(I_c) \supset I_c$. A careful analysis shows that $F_c$ has a point of period three. Assertion (d) now follows from the existence of points of period two and three (see [5]). □

### 3. Failure of Rayleigh quotient iteration

In this section we will prove Theorem B. Let

$$B_{cd} = \begin{bmatrix} 0 & -c & 0 \\ 1/c & 0 & 0 \\ 0 & 0 & d \end{bmatrix},$$

and let $F_{cd}$ be the associated map. For each matrix $B_{cd}$ the $z$-axis is an eigenspace. The $xy$-plane is a subspace of type $2C$. We wish to find a sink orbit which does not lie on either subspace.

The projective space $\mathbb{RP}^2$ is covered by three standard coordinate charts. One of these charts is given by mapping $(y, z)$ to the one-dimensional subspace spanned by $(1, y, z)$. We call this chart the $yz$-chart. With respect to this
chart, the $xy$-plane corresponds to the $y$-axis, and the $z$-axis corresponds to a point at infinity. In the coordinate system given by this chart we want to find a sink orbit which does not lie on the $y$-axis.

The map $F_{cd}$ commutes with the reflection through the $z$-axis and with reflection through the $y$-axis. Both these axes are invariant. The restriction of the map $F_{cd}$ to the $y$-axis is the map $F_c$ considered in the previous section. The origin is a point of period two. The derivative of $F^2_{cd}$ in the $y$-direction is controlled by the $c$-parameter. It is $-1$ when $c = \sigma$, and decreases as $c$ decreases. The derivative at the origin in the $z$-direction is given by $d^{-2}$.

An obvious approach to creating a sink orbit is to choose the $c$-parameter so that the $z$-axis is normally attracting and then to choose the $d$-parameter to create a “pitchfork bifurcation”, making the origin into a saddle and creating two nearby sinks. Unfortunately, this does not succeed. The bifurcation on the $z$-axis is degenerate and does not create new sinks. When the origin becomes repelling, all points on the $z$-axis are attracted to infinity.

The approach we use is slightly more complicated. We adjust the $c$-parameter to create a period doubling bifurcation on the $y$-axis. Let $(y_c, 0)$ and $(-y_c, 0)$ be the period-four points created. We adjust the $d$-parameter to make the points $(\pm y_c, 0)$ neither attracting nor repelling in the normal direction (i.e., the $z$-direction). (Actually, we make these two adjustments in a single step.) We appeal to the center stable manifold theorem to create an invariant curve through $(\pm y_c, 0)$ transverse to the $y$-axis. We then adjust the $d$-parameter to create pitchfork bifurcations along these invariant curves.

Let $\Omega_{cd}$ denote the map obtained by composing $F^4_{cd}$ with the projection onto the $z$-coordinate. We will produce a path in parameter space passing through the point $(\sigma, 1)$ such that for each $(c, d)$ on the path, $\left(\partial \Omega_{cd}/\partial z\right)(y_c, 0) = 1$. Applying the implicit function theorem to the function

$$\frac{\partial \Omega_{cd}}{\partial z} \bigg|_{z=0} - 1$$

at $(y = 0, c = \sigma, d = 1)$, we obtain a function $d(y, c)$ and a curve in parameter space given by $\gamma(c) = (c, d(y_c, c))$, which will be defined for $c$ slightly less than $\sigma$. Using the chain rule, it follows that $\gamma'(\sigma)$ is parallel to the $c$-axis.

At any point $\gamma(c)$ the derivative of $F^4_{\gamma(c)}$ at $(y_c, 0)$ is $[\tau^0]$. The off-diagonal elements are zero because of symmetry and the invariance of the $y$-axis. Now $\tau = (F^4_{\gamma})'(y_c)$. Since $c < \sigma$, it follows that $\tau < 1$, where $\tau$ depends on $c$. Consider a path in parameter space that begins at a point $\gamma(c)$ and decreases the $d$-parameter while fixing the $c$-parameter. Since $(\partial^2 \Omega_{cd}/\partial d\partial z)(0, 0)$ is negative when $d = 1$, the effect of a decrease in $d$ on the derivative of $F^4_{\gamma(c)}$ at $(y_c, 0)$ is to increase the lower diagonal entry in the above matrix and create a saddle (see Figure 3).

For each $c$, consider the product of a neighborhood of $d(y_c, c)$ with $\mathbb{R}^2$. Applying the center manifold theorem to the map $\Xi_c(t, w) = (t, F^4_{cd}(w))$ at
the point \((d(yc, c), (yc, 0))\) yields a one-parameter family of \(F^4_{ct}\)-invariant one-dimensional submanifolds (see [11]). For each of these one-dimensional submanifolds the tangent vector to the submanifold at \((yc, 0)\) is parallel to the \(z\)-axis. Thus, for any fixed \(c\) we can coordinatize the one-parameter family by \((ht(z), z)\), where \(t\) is defined on some neighborhood of \(d(yc, c)\). Restricting the map \(F^4_{ct}\) to the family of submanifolds yields a one-parameter family of maps, \(f_t(z)\), for any fixed \(c\). To obtain the bifurcation in the family \(f_t(z)\), we will employ the following result, which is discussed in [7, p. 150].

**Pitchfork bifurcation.** If \(f(t, z) = f_t(z)\) is a one-parameter family of mappings with the following properties:

- \(a.\) \(f(t, 0) = 0\),  
- \(b.\) \(\frac{\partial f}{\partial z}(t_0, 0) = 1\),  
- \(c.\) \(\frac{\partial^2 f}{\partial z^2}(t_0, 0) = 0\),  
- \(d.\) \(\frac{\partial^2 f}{\partial t \partial z}(t_0, 0) < 0\),
- \(e.\) \(\frac{\partial^3 f}{\partial z^3}(t_0, 0) < 0\),

then there is a bifurcation at \(t_0\). If \(t > t_0\), \(f_t\) has a sink at \(z = 0\). If \(t < t_0\), there is a source at \(z = 0\) and \(f_t\) has sinks on each side of \(0\).

Letting \(t_0 = d(yc, c)\), conditions (a), (b), and (d) follow from the previous discussion. Condition (c) follows from the symmetry \(f_t(-z) = -f_t(z)\). To establish condition (e), it suffices to show that the third Maclaurin coefficient of \(f_t\) is negative. We begin with the first-degree Taylor approximation of our iteration map at \((yc, 0)\). We assume that \(c\) has been fixed and \(d = d(yc, c)\).

Then

\[
F^4_{d(yc)}(y, z) = (\tau(y - yc) + R(y, z), z + S(y, z)).
\]
The functions have the following symmetries in $z$:
\[ p(x, y, -z) = p(x, y, z), \quad R(y, -z) = R(y, z), \quad S(y, -z) = -S(y, z). \]

From these symmetries it follows that the powers of $z$ in Taylor polynomials will be odd for both $S$ and $f_d$, and even for both $R$ and $h_d$. Denote the coefficients as follows:
\[ h_d(z) = az^2 + \cdots, \quad R(y, z) = bz^2 + \cdots, \quad S(y, z) = uyz + ez^3 + \cdots. \]

From the invariance of the submanifold $(h_d(z), z)$ we obtain
\[ h_d(z + S(h_d(z), z)) = \tau(h_d(z) - y_c) + R(h_d(z), z). \]

Computing the $z^2$ coefficient yields
\[ a[z + S(h_d(z), z)]^2 = \tau az^2 + bz^2, \quad a = b/(1 - \tau). \]

Since
\[ f_d(z) = z + S(h_d(z), z), \]
the third-degree Maclaurin polynomial of $f_d$ is given by
\[ z + (au + e)z^3 = z + (bu/(1 - \tau) + e)z^3. \]

It remains to show that if $c$ is sufficiently close to $\sigma$, and $d$ is sufficiently close to $d(y_c, c)$, then $bu/(1 - \tau) + e < 0$. To prove this, we will consider the corresponding coefficients in the Taylor expansion of $F_{cd}$ at the origin. All three of these coefficients are 0 when the parameters are given by $\gamma(\sigma)$. We now examine how these coefficients change as we move along a path in parameter space in the direction of the negative $c$-axis from $\gamma(\sigma)$. To do this, we evaluate at $\gamma(\sigma)$ the partial derivative with respect to $c$ of the coefficients which correspond to $b$, $u$, and $e$:
\[ b \sim -c^{-3}d^{-5}[(2c^6 - 9c^4 + 12c^2 - 4)d^6 + (-2c^6 + 5c^4 - 2c^2)d^4 + (2c^2 - c^4)d^2 + c^4], \]
\[ u \sim -c^{-3}d^{-5}[4c^8 - 20c^6 + 32c^4 - 20c^2 + 4], \]
\[ e \sim c^{-6}d^{-10}[(2c^{10} - 9c^8 + 12c^6 - 4c^4)d^6 + (-2c^{10} + 5c^8 - 2c^6)d^4 + (2c^6 - c^8)d^2 + c^8]. \]

Calculations reveal that the signs of the $c$-partials at $\gamma(\sigma)$ are $-, +$, and $+$, respectively. Thus, the derivatives in the direction of the negative $c$-axis are $+, -, \text{ and } -$, respectively. From continuity we conclude that if $c$ is slightly smaller than $\sigma$, then $b > 0$, $u < 0$, and $e < 0$; hence $bu/(1 - \tau) + e < 0$, and there is a pitchfork bifurcation (see Figure 4).

The pitchfork bifurcation produces a point of period four which is contracting in the $z$-direction and has positive $y$- and $z$-coordinates. Using continuity again, we can obtain parameter values for which the point is also contracting.
in the \( y \)-direction. Thus, we have a hyperbolic sink of period four which does not lie in a subspace of type 4R or 4C. Since hyperbolic sinks persist under perturbation of the map, the theorem is proved for \( n = 3 \).

To prove the theorem for \( n > 3 \), take a direct sum of \( B_{cd} \) with \( \kappa I \), where \( \kappa \) is chosen sufficiently large.

Note. While the above proof does not provide specific values of \( c \) and \( d \), failure of the algorithm can be observed numerically with parameter values \( c = .56 \) and \( d = 1.1 \) and the point \( x = -0.104, y = -0.995, \) and \( z = 0.107 \).

4. Further problems

We conclude with some questions concerning the Rayleigh quotient iteration algorithm. It may be relevant to compare Rayleigh quotient iteration with Newton's method. Newton's method may succeed or fail on a polynomial with real coefficients (see [15, 9, 10]). If all of the roots of the polynomial are real, then it succeeds [1, 2].

Problem 1. Does Rayleigh quotient iteration always succeed when all of the zeros of the characteristic polynomial are real?

Problem 2. Do there exist parameter values of \( c \) and \( d \) for which \( F_{cd} \) displays "chaotic" behavior on a set of positive measure? (Preliminary numerical investigations suggest that this might be the case for \( c = .5 \) and \( d = 1.1 \).)

Finally, one can define an analogous algorithm over \( \mathbb{C} \). For an \( n \times n \) complex-valued matrix \( B \), the Rayleigh quotient iteration formula induces a map of \( \mathbb{C}P^{n-1} \). This map has been studied in [13], and a related map is studied in [8]. Note that if \( B \) has real entries, then \( \mathbb{R}P^{n-1} \) is an invariant submanifold. Thus, the example of the previous section produces a map which has a periodic point which is attracting in this submanifold. However, this point would appear
to be a saddle point in the higher-dimensional setting of $\mathbb{CP}^{n-1}$, and thus our counterexample construction does not seem to work in the complex case.

**Problem 3.** Does Rayleigh quotient iteration (over $\mathbb{C}$) succeed for all matrices?

**Acknowledgment**

The authors would like to thank Michael Shub and Stephen Smale for useful discussions. Some of the calculations in this paper were done with MACSYMA.

**Bibliography**


**Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia 30322. E-mail: sb@mathcs.emory.edu**

**Department of Mathematics, Cornell University, Ithaca, New York 14850. E-mail: smillie@mssun7.msi.cornell.edu**