The Projected Newton Method Has Order $1 + \sqrt{2}$ for the Symmetric Eigenvalue Problem

by

R.A. Tapia

and

David L. Whitley

Technical Report 87-14, May 1987
(Revised November 1987)

The Projected Newton Method Has Order $1 + \sqrt{2}$ for the Symmetric Eigenvalue Problem

R.A. Tapia* David L. Whitley*

Abstract

In their study of the classical inverse iteration algorithm, Peters and Wilkinson considered the closely related algorithm that consists of applying Newton's method, followed by a 2-norm normalization, to the nonlinear system of equations consisting of the eigenvalue-eigenvector equation and an equation requiring the eigenvector to have the square of its 2-norm equal to one. They argue that in practice the ∞-norm is easier to work with, and they therefore replace the 2-norm normalization equation with a linear equation requiring that a particular component of the eigenvector be equal to one (effectively an ∞-norm normalization). Next, they observe that, because of the linearity of the normalization equation, the normalization step is automatically satisfied; the algorithm thus reduces to Newton's method and quadratic convergence follows from standard theory. Peters and Wilkinson choose to dismiss the 2-norm formulation in favor of the ∞-norm formulation; one factor in their choice seems to be that quadratic convergence is not so immediate for the 2-norm formulation. In this work we establish the surprising result that the 2-norm formulation gives a convergence rate of $1 + \sqrt{2}$, significantly superior to that given by the Peters and Wilkinson formulation.

*Mathematical Sciences Department, Rice University, Houston, Texas 77251-1892. Research sponsored by DOE DE-FG05-86ER25017, SDIO/IST managed under ARO DAAG-03-86-K-0113, and AFOSR 85-0243.

1 Introduction

In their study of the classical inverse iteration algorithm, Peters and Wilkinson (1979) considered two closely related algorithms. They began by observing that in finding an eigenvector-eigenvalue pair $(x_*,\ \lambda_*)$ of a given real symmetric matrix A, if we require that $\|x\|_2 = 1$, then the pair must satisfy

$$\begin{aligned} (A-\lambda I)x &= 0, \\ \frac{1}{2}(1-x^Tx) &= 0. \end{aligned} \tag{1}$$

They then suggest the following scheme for solving this system: let $x_0^Tx_0 = 1$ for the initial iterate $(x_0,\ \lambda_0)$; now given a current iterate $(x,\ \lambda)$, let $(\Delta x,\ \Delta\lambda)$ solve

$$\begin{pmatrix} A-\lambda I & -x \\ -x^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} (A-\lambda I)x \\ \frac{1}{2}(1-x^Tx) \end{pmatrix} \tag{2}$$

and let

$$\begin{aligned} x_+ &= \frac{(x+\Delta x)}{\|x+\Delta x\|}, \\ \lambda_+ &= \lambda+\Delta\lambda. \end{aligned} \tag{3}$$

Without the normalization in (3) this would be Newton's method applied to the nonlinear system (1). With the normalization it is the Projected Newton Method.

Motivated by the fact that in practice it is usually simpler to scale successive x-iterates so that a particular component is equal to one, Peters and Wilkinson propose using a different normalization: instead of finding a solution to (1), find one to

$$\begin{aligned} (A-\lambda I)x &= 0, \\ 1-e_m^Tx &= 0; \end{aligned} \tag{4}$$

where it is assumed that the mth component of x_* is one of its larger components. The analogous iterative scheme for solving this problem requires that $e_m^Tx_0 = 1$, and that $(\Delta x, \Delta\lambda)$ solve

$$\begin{pmatrix} A-\lambda I & -x \\ -e_m^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} (A-\lambda I)x \\ 1-e_m^Tx \end{pmatrix}. \tag{5}$$

Here, we take

$$\begin{aligned} x_+ &= x + \Delta x, \\ \lambda_+ &= \lambda + \Delta\lambda. \end{aligned} \tag{6}$$

Because $e_m^T(\Delta x) = 0$, each iterate satisfies the new normalization; therefore, Projected Newton coincides with Newton's method on the system (4). In Section 4 of Peters and Wilkinson (1979), there is an argument that establishes that the matrix in Equation (5) is nonsingular at (x_*, λ_*) if λ_* is a simple eigenvalue. Thus, when λ_* is a simple eigenvalue, the convergence of this scheme is clearly q-quadratic in the pair (x, λ), as a consequence of the standard theory for Newton's Method.

Wilkinson (1981) considered extensions, refinements or applications of the scheme (6), as did Dongarra, Moler and Wilkinson (1983). The latter paper also contains a proof of r-quadratic convergence of the x-iterates for this scheme, which follows immediately from the q-quadratic convergence of the pair (x, λ) (The definitions of r- and q-order of convergence may be found in Dennis and Schnabel (1983), pp. 19-21. For further detail, see Chapter 9 of Ortega and Rheinboldt (1970)). In numerical experiments to assess the behavior of the two schemes described, we discovered that the second did indeed seem to be q-quadratically convergent, but no better; the first, however, was undeniably faster than q-quadratic, yet not q-cubic. In §2 we prove our main result, that the convergence of *each* of the sequences x_k and λ_k generated by the first scheme is actually of q-order $1 + \sqrt{2}$. In §3, we add concluding remarks.

Before proceeding to the main result, we note that the two algorithms presented above are closely related to two other well-known methods for finding an eigenvalue-eigenvector pair for a symmetric matrix A: inverse iteration and Rayleigh Quotient Iteration. To see this, note that the first equation in either (2) or (5) implies that

$$(A - \lambda I)\hat{x}_+ = (A - \lambda I)(x + \Delta x) = (\Delta\lambda)x. \tag{7}$$

Given a current eigenvector estimate x and eigenvalue estimate λ (for inverse iteration the eigenvalue estimate is fixed, for the other three it changes from iteration to iteration), equation (7) shows that each of the latter three methods produces a new eigenvector estimate that is a scalar multiple of the one given by inverse iteration. The new eigenvalue estimates for these three methods are also related to each other in an interesting way. By premultiplying both sides of (2) by $(x^T, -\lambda)$ and rearranging terms, we obtain the expression

$$\lambda_+ = \frac{x^T A x_+}{x^T x_+} \tag{8}$$

for the Projected Newton formulation. Similarly, we can obtain from (5) the expression

$$\lambda_+ = \frac{e_m^T A x_+}{e_m^T x_+}$$

for the Peters-Wilkinson formulation. For the Rayleigh Quotient Iteration, the new eigenvalue estimate is given by

$$\lambda_+ = \frac{x_+^T A x_+}{x_+^T x_+}.$$

Thus each method has the effect of solving $(A - \lambda I)\hat{x}_+ = x$, scaling appropriately, and updating λ using the appropriate formula. This shows that our result is of a theoretical nature, since it offers few if any real computational advantages over the Rayleigh Quotient Iteration, the convergence of which is q-cubic.

2 Main result

The following Lemma gives a sufficient condition for a q-convergence rate of $1 + \sqrt{2}$.

Lemma *Let $\{x_k\}$ be a sequence that converges to x_*. If there exist positive constants m, M, and $\hat{k}$ such that*

$$m \leq \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2 \|x_{k-1} - x_*\|} \leq M \tag{9}$$

for all $k \geq \hat{k}$, then x_k converges to x_* with q-order $1 + \sqrt{2}$.

Proof. Clearly, we can assume $\hat{k} = 1$ without loss of generality. If we define α_k by

$$\alpha_k = \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^{1+\sqrt{2}}},$$

then we can rewrite (9) as

$$m \leq \alpha_{k+1}\alpha_k^{\sqrt{2}-1} \leq M.$$

This pair of inequalities can be applied repeatedly to yield

$$m\alpha_1^{-(\sqrt{2}-1)} \leq \alpha_2 \leq M\alpha_1^{-(\sqrt{2}-1)},$$

$$mM^{-(\sqrt{2}-1)}\alpha_1^{(\sqrt{2}-1)^2} \leq \alpha_3 \leq Mm^{-(\sqrt{2}-1)}\alpha_1^{(\sqrt{2}-1)^2},$$

and so on (as can be shown by induction); in general,

$$\alpha_k \leq Mm^{-(\sqrt{2}-1)}M^{(\sqrt{2}-1)^2}m^{-(\sqrt{2}-1)^3}\cdots\alpha_1^{(-(\sqrt{2}-1))^{k-1}} \equiv \beta_k.$$

Therefore, $\limsup \alpha_k \leq \lim \beta_k = M^{\left(\frac{1}{2(\sqrt{2}-1)}\right)}m^{-\frac{1}{2}}$. Hence for all sufficiently large k,

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^{1+\sqrt{2}}} = \alpha_k < M^{\left(\frac{1}{2(\sqrt{2}-1)}\right)}m^{-\frac{1}{2}} + 1.$$

Therefore x_k converges to x_* with q-order $1 + \sqrt{2}$.□

From now on, let us write (x_-, λ_-), (x, λ), and (x_+, λ_+), where convenient, in place of (x_{k-1}, λ_{k-1}), (x_k, λ_k), and (x_{k+1}, λ_{k+1}). We now state our main result:

Theorem *Let A be a symmetric matrix, x_0 a vector with $\|x_0\| = 1$, and λ_0 a scalar that is not an eigenvalue of A. Taking $(x, \lambda) = (x_0, \lambda_0)$ initially, apply the following steps iteratively to generate a sequence (x_k, λ_k):*

$$\begin{pmatrix} \hat{x}_+ \\ \lambda_+ \end{pmatrix} = \begin{pmatrix} x \\ \lambda \end{pmatrix} - \begin{pmatrix} A - \lambda I & -x \\ -x^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} (A - \lambda I)x \\ 0 \end{pmatrix}$$

$$x_+ = \frac{\hat{x}_+}{\|\hat{x}_+\|}.$$

If λ_k converges to an eigenvalue λ_* and $\lambda_k \neq \lambda_*$ for all k, then λ_k converges with q-order $1+\sqrt{2}$, and x_k converges with the same q-order to a corresponding eigenvector x_*.

Remark. We note that neither the hypothesis requiring convergence of λ_k nor the hypothesis that $\lambda_k \neq \lambda_*$ should be considered restrictive. When λ_* is a simple eigenvalue, the Jacobian matrix is nonsingular at (x_*, λ_*), so the standard theory for Newton's method ensures that the sequence is locally quadratically convergent (at least) in this case. When λ_* is not simple, convergence does not follow from the standard theory, but in view of this algorithm's resemblance to inverse iteration, it is reasonable to assume that x_k will converge to an eigenvector, and Equation (8) then implies that λ_k converges to an eigenvalue. The second hypothesis excludes the possibility that a finite number of iterations might lead to an exact solution. If x_k were an eigenvector, then λ_{k+1} would be the corresponding eigenvalue, as can be readily seen from (8). If λ_k were equal to an eigenvalue, then the iteration as described above might not be defined, since the matrix might be singular; in any case, however, $(\Delta x, \Delta\lambda) = (\hat{x}_* - x, 0)$ would solve the associated linear system (2) for some eigenvector $\hat{x}_*$ corresponding to the eigenvalue λ_k. Thus an exact solution would be found in one more iteration, provided that the iteration is defined.

Proof. There are two main parts to the proof of the Theorem: we show first that

$$l_1 \leq \frac{\|x_{k+1} - x_*\|}{|\lambda_k - \lambda_*| \|x_k - x_*\|} \leq u_1 \tag{10}$$

and then that

$$l_2 \leq \frac{|\lambda_k - \lambda_*|}{\|x_k - x_*\| \|x_{k-1} - x_*\|} \leq u_2, \tag{11}$$

for k sufficiently large, where l_1, u_1, l_2, and u_2 are positive constants. From (10) and (11), it is straightforward to show that

$$l_1 l_2 \leq \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2 \|x_{k-1} - x_*\|} \leq u_1 u_2,$$

and

$$\frac{l_1^2 l_2}{u_2} \leq \frac{|\lambda_{k+1} - \lambda_*|}{|\lambda_k - \lambda_*|^2 |\lambda_{k-1} - \lambda_*|} \leq \frac{u_1^2 u_2}{l_2}.$$

We can then apply the Lemma to conclude that both x_k and λ_k converge with q-order $1+\sqrt{2}$ whenever (10) and (11) hold.

For the first part of the proof, we parallel the proof of the linear convergence of the power method given by Parlett (1980). We begin with some notation: let S be the subspace of eigenvectors of A with eigenvalue λ_*, and denote the orthogonal complement of S by $\mathrm{S}^{\perp}$. Define x_* to be the vector obtained by normalizing the projection of x_0 onto the subspace S. Denote the angle between x and x_* by θ, and the angle between x_+ and x_* by θ_+. Note that we can write x as

$$x = x_* \cos\theta + u \sin\theta, \tag{12}$$

where u is a unit vector orthogonal to x_*. If we pre-multiply both sides of (12) by $(A-\lambda I)^{-1}$, we obtain

$$(A-\lambda I)^{-1}x = x_* \left(\frac{\cos\theta}{\lambda_* - \lambda}\right) + \left(\frac{(A-\lambda I)^{-1}u}{\|(A-\lambda I)^{-1}u\|}\right)(\|(A-\lambda I)^{-1}u\| \sin\theta).$$

By (7), $(A-\lambda I)^{-1}x$ is a scalar multiple of x_+; since x_* is orthogonal to $(A-\lambda I)^{-1}u$, the above expression shows that an orthogonal decomposition of x_+ is

$$x_+ = x_* \cos\theta_+ + u_+ \sin\theta_+,$$

where

$$u_+ = \frac{(A-\lambda I)^{-1}u}{\|(A-\lambda I)^{-1}u\|} \tag{13}$$

and

$$\begin{aligned} \tan\theta_+ &= \frac{\|(A-\lambda I)^{-1}u\| \sin\theta}{\cos\theta/(\lambda_* - \lambda)} \\ &= (\lambda_* - \lambda)\|(A-\lambda I)^{-1}u\| \tan\theta. \end{aligned}$$

By the choice of x_*, $u_0 \in \mathrm{S}^{\perp}$; hence, by Equation (13) and induction, we can conclude that $u_k \in \mathrm{S}^{\perp}$ for all k. Thus we can obtain upper and lower bounds on $\|(A-\lambda I)^{-1}u\|$ by considering the restriction of $(A-\lambda I)^{-1}$ to $\mathrm{S}^{\perp}$ (i.e. the reduced resolvent), which we denote by $(A-\lambda I)^{-\perp}$:

$$\begin{aligned} \|(A-\lambda I)^{-1}u\| &= \|(A-\lambda I)^{-\perp}u\| \\ &\le \|(A-\lambda I)^{-\perp}\| \\ &= \frac{1}{|\nu_\lambda - \lambda|}, \end{aligned} \tag{14}$$

where ν_λ is the nearest eigenvalue of A to λ other than λ_*. Similarly,

$$\|(A-\lambda I)^{-1}u\| \geq \frac{1}{|\phi_\lambda - \lambda|} \tag{15}$$

where ϕ_λ is the eigenvalue of A farthest from λ. So we now have that

$$\left|\frac{\lambda_* - \lambda}{\phi_\lambda - \lambda}\right| \leq \left|\frac{\tan\theta_+}{\tan\theta}\right| \leq \left|\frac{\lambda_* - \lambda}{\nu_\lambda - \lambda}\right|. \tag{16}$$

It is clear from (16) that if λ_k converges to λ_*, then the vector we have defined to be x_* is indeed the eigenvector to which x_k converges.

In terms of the error angle θ, the error in x is given by

$$\begin{aligned} \|x - x_*\| &= 2\sin\frac{\theta}{2} && (17) \\ &= \sqrt{2(1-\cos\theta)}. && (18) \end{aligned}$$

From (18), the following relationship can be derived in a straightforward manner:

$$\frac{\|x_+ - x_*\|}{\|x - x_*\|} = \left|\frac{\tan\theta_+}{\tan\theta}\right| \left|\frac{\cos\theta_+}{\cos\theta}\right| \left(\frac{1+\cos\theta}{1+\cos\theta_+}\right)^{\frac{1}{2}}. \tag{19}$$

Combining (16) with (19) and taking limits, we may conclude that

$$\begin{aligned} \frac{1}{|\phi_{\lambda_*} - \lambda_*|} &\leq \liminf \frac{\|x_{k+1} - x_*\|}{|\lambda_k - \lambda_*|\|x_k - x_*\|} \\ &\leq \limsup \frac{\|x_{k+1} - x_*\|}{|\lambda_k - \lambda_*|\|x_k - x_*\|} \\ &\leq \frac{1}{|\nu_{\lambda_*} - \lambda_*|}. \end{aligned}$$

Hence (10) holds with, for instance, $l_1 = \frac{1}{2|\phi_{\lambda_*} - \lambda_*|}$ and $u_1 = \frac{2}{|\nu_{\lambda_*} - \lambda_*|}$.

We now show that there exist positive l_2 and u_2 satisfying (11). We first use (8) to write

$$\lambda - \lambda_* = \frac{x_-^T(A - \lambda_* I)x}{x_-^T x}. \tag{20}$$

If we substitute the orthogonal decompositions of x_- and x given by (12) into (20), we obtain

$$\begin{aligned}\lambda - \lambda_* &= \frac{(x_* \cos\theta_- + u_- \sin\theta_-)^T(A - \lambda_* I)(x_* \cos\theta + u \sin\theta)}{x_-^T x} \\ &= \frac{(u_- \sin\theta_-)^T(A - \lambda_* I)(u \sin\theta)}{x_-^T x},\end{aligned}$$

since (x_*, λ_*) is an eigenvector-eigenvalue pair for A. Now we use (17) to relate the sines of the error angles to the sizes of the errors in x_- and x, and we use (13) to write u in terms of u_- and λ_-:

$$\begin{aligned}\frac{\lambda - \lambda_*}{\|x_- - x_*\|\|x - x_*\|} &= \frac{\cos\frac{\theta_-}{2}\cos\frac{\theta}{2}}{x_-^T x}(u_-^T(A - \lambda_- I)u + (\lambda_- - \lambda_*)u_-^T u) \\ &= \frac{\cos\frac{\theta_-}{2}\cos\frac{\theta}{2}}{x_-^T x}\left(\frac{1}{\|(A - \lambda_- I)^{-1}u_-\|} + (\lambda_- - \lambda_*)u_-^T u\right).\end{aligned}$$

Using the bounds for $\|(A - \lambda_- I)^{-1}u_-\|$ from (14) and (15), then taking limits, it follows that

$$\begin{aligned}\frac{1}{|\phi_{\lambda_*} - \lambda_*|} &\leq \liminf \frac{|\lambda_k - \lambda_*|}{\|x_k - x_*\|\|x_{k-1} - x_*\|} \\ &\leq \limsup \frac{|\lambda_k - \lambda_*|}{\|x_k - x_*\|\|x_{k-1} - x_*\|} \\ &\leq \frac{1}{|\nu_{\lambda_*} - \lambda_*|}.\end{aligned}$$

Hence (11) holds with, for example, $l_2 = \frac{1}{2|\phi_{\lambda_*} - \lambda_*|}$ and $u_2 = \frac{2}{|\nu_{\lambda_*} - \lambda_*|}$. This completes the proof. □

3 Conclusion

The motivation for this work was the intriguing and, to us, surprising non-integral superquadratic convergence rate indicated in our numerical experiments. We were also quite pleased that we were able to establish a q-rate of convergence of $1 + \sqrt{2}$ both for the x-iterate alone and for the λ-iterate alone. The classical Newton's method theory gives q-quadratic convergence in the pair (x, λ) for the Peters-Wilkinson algorithm, which implies an r-quadratic rate in each of x and λ. The direct proof given by Dongarra, Moler and Wilkinson (1983) also only establishes r-quadratic convergence in x for the same algorithm.

We experimented with several problem formulations that used p-norms other than the 2-norm and ∞-norm; for no values of p except $p = 2$ did our estimated convergence rates consistently exceed 2. On the basis of these experiments, we consider it extremely unlikely that any p-norm formulation other than the one presented here produces superquadratic convergence.

REFERENCES

Dennis, J.E., Jr., and Robert B. Schnabel (1983). *Numerical Methods for Nonlinear Equations and Unconstrained Optimization*, Prentice-Hall, Englewood Cliffs, New Jersey.

Dongarra, J.J., C.B. Moler and J.H. Wilkinson (1981). Improving the accuracy of computed eigenvalues and eigenvectors, SIAM J. Numer. Anal., 20, pp. 23–45.

Ortega, J.M., and W.C. Rheinboldt (1970), *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York.

Parlett, Beresford N. (1980). *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, New Jersey.

Peters, G. and J.H. Wilkinson (1979). Inverse iteration, ill-conditioned equations and Newton's method, SIAM Review, 21, pp. 339–360.

Wilkinson, J.H. (1962). Inverse iteration in theory and practice, in: *Symposia Mathematica Volume X* (Istituto Nazionale di Alta Mathematica Monograf, Bologna, 1972) pp. 361-379.

Wilkinson, J.H. (1981). Error bounds for computed invariant subspaces, Proc. of Rutishauser Symposium on Numerical Analysis, Research Report 81-02, Eidgenössische Technische Hochschule, Zürich, Switzerland, February, 1981.