



A Motivational Example for the Numerical Solution of the Algebraic Eigenvalue Problem

Stephen M. Alessandrini

SIAM Review, Vol. 40, No. 4. (Dec., 1998), pp. 935-940.

Stable URL:

<http://links.jstor.org/sici?sici=0036-1445%28199812%2940%3A4%3C935%3AAMEFTN%3E2.0.CO%3B2-E>

SIAM Review is currently published by Society for Industrial and Applied Mathematics.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/siam.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

A MOTIVATIONAL EXAMPLE FOR THE NUMERICAL SOLUTION OF THE ALGEBRAIC EIGENVALUE PROBLEM*

STEPHEN M. ALESSANDRINI†

Abstract. This paper presents an example of an algebraic eigenvalue problem which can be used to motivate the study of numerical techniques for solving such problems. The problem consists of finding the axis and angle of rotation from a 3×3 rotation matrix and is referred to as the *axis-angle problem*. The problem is used to demonstrate the inverse power method for finding eigenvectors. The axis-angle problem is developed and numerical results are given.

Key words. inverse power method, inverse iteration, algebraic eigenvalue problem

AMS subject classifications. 65F15, 15A18

PII. S0036144597328353

1. Introduction. In a previous note [1], an example was given to motivate the study of the numerical solution of two-point boundary-value problems. This note describes another example which can be used to motivate the numerical solution of the algebraic eigenvalue problem and, in particular, is a good application of the inverse power method, also called inverse iteration or Wielandt iteration [9].

2. The example. Consider the following situation which arises in computer graphics. Usually objects in a scene are constructed in some reference coordinate system, the *local coordinates*, and then rotated and translated into *world coordinates*, the coordinates of the scene. This allows a particular object which may occur several times in a scene to be created once in the local coordinates and then copied to the various locations in the world coordinates of the scene. For example, if the scene is a classroom there may be 30 of the same chairs in the room. The chair is modeled once in local coordinates and then 30 occurrences are rotated and translated into their positions in the room. Sometimes several rotations are performed to achieve the correct orientation of the object. These rotations can be combined into one rotation about some arbitrary axis of rotation by some angle of rotation. The rotation matrix is determined by just multiplying the individual rotations together. However, it is sometimes necessary to determine the axis and angle of rotation of the final rotation. Let the axis of rotation be denoted by the unit vector \mathbf{a} and the angle of rotation be θ . Then the general 3×3 rotation matrix which rotates a point counterclockwise by θ looking down the axis of rotation \mathbf{a} towards the origin is given in [4] by

$$R(\mathbf{a}, \theta) = \begin{pmatrix} a_x^2 + C(1 - a_x^2) & a_x a_y(1 - C) - a_z S & a_x a_z(1 - C) + a_y S \\ a_x a_y(1 - C) + a_z S & a_y^2 + C(1 - a_y^2) & a_y a_z(1 - C) - a_x S \\ a_x a_z(1 - C) - a_y S & a_y a_z(1 - C) + a_x S & a_z^2 + C(1 - a_z^2) \end{pmatrix},$$

where $C = \cos \theta$ and $S = \sin \theta$. Note that rotating by $-\theta$ about $-\mathbf{a}$ is equivalent to rotating by θ about \mathbf{a} . It is not very easy to just equate this expression to the matrix representing the combined rotation and solve for \mathbf{a} and θ . However, after some thought, we realize that any point on the axis of rotation remains fixed when it is

*Received by the editors January 23, 1997; accepted for publication (in revised form) September 24, 1997.

<http://www.siam.org/journals/sirev/40-4/32835.html>

†Lockheed Martin Corporation, Moorestown, NJ 08057 (stephen.m.alessandrini@lmco.com).

multiplied by the rotation. Thus,

$$R\mathbf{a} = \mathbf{a}.$$

This says that $\lambda = 1$ is an eigenvalue of the rotation matrix R and \mathbf{a} is a corresponding eigenvector. The other eigenvalues are $\cos \theta \pm i \sin \theta$. The direct solution to this problem is given in [8]. It involves what are called *quaternions*, which are 4-vectors with certain algebraic properties discovered by W. R. Hamilton in 1844 while he was trying to generalize complex numbers [5]. It can be shown that a unit quaternion represents a rotation matrix and a direct method can be constructed which determines the axis and angle of rotation. We will consider an iterative solution using the inverse power method and compare it to the direct method.

Once the axis of rotation has been determined, we need to determine θ . To do this we note that the trace of R , which is the sum of the diagonal elements, is

$$\begin{aligned} \text{tr } R &= r_{11} + r_{22} + r_{33} \\ &= a_x^2 + C(1 - a_x^2) + a_y^2 + C(1 - a_y^2) + a_z^2 + C(1 - a_z^2) \\ &= a_x^2 + a_y^2 + a_z^2 + C(3 - a_x^2 - a_y^2 - a_z^2) \\ &= 1 + 2 \cos \theta, \end{aligned}$$

since \mathbf{a} is a unit vector. Thus,

$$(2.1) \quad \cos \theta = \frac{\text{tr } R - 1}{2}.$$

To determine θ in the range 0 to 2π , we need to determine $\sin \theta$. Let $\|\mathbf{a}\| = \max\{|a_x|, |a_y|, |a_z|\}$. Then, we can determine $\sin \theta$ using

$$(2.2) \quad \sin \theta = \begin{cases} \frac{r_{32} - a_y a_z (1 - \cos \theta)}{a_x} & \text{if } \|\mathbf{a}\|_\infty = |a_x|, \\ \frac{r_{13} - a_x a_z (1 - \cos \theta)}{a_y} & \text{if } \|\mathbf{a}\|_\infty = |a_y|, \\ \frac{r_{21} - a_x a_y (1 - \cos \theta)}{a_z} & \text{if } \|\mathbf{a}\|_\infty = |a_z|. \end{cases}$$

Finally, θ is determined from

$$(2.3) \quad \theta = \text{atan2}(\sin \theta, \cos \theta). \quad \square$$

3. The direct solution. In this section, we summarize the direct solution given by Shoemake [8]. We first describe how to convert a rotation matrix into a unit quaternion. Let R be a rotation matrix, $\mathbf{q} = (w, x, y, z)$ be the corresponding unit quaternion, and $\epsilon > 0$ be a zero tolerance. Let $w^2 = (1 + \text{tr } R)/4$. If $w^2 > \epsilon$, we set

$$w = \sqrt{w^2}, \quad x = \frac{r_{32} - r_{23}}{4w}, \quad y = \frac{r_{13} - r_{31}}{4w}, \quad z = \frac{r_{21} - r_{12}}{4w}.$$

Otherwise, set $w = 0$ and $x^2 = -(r_{22} + r_{33})/2$. If $x^2 > \epsilon$, we set

$$x = \sqrt{x^2}, \quad y = \frac{r_{12}}{2x}, \quad z = \frac{r_{13}}{2x}.$$

Otherwise, set $x = 0$ and $y^2 = (1 - r_{33})/2$. If $y^2 > \epsilon$, we set

$$y = \sqrt{y^2}, \quad z = \frac{r_{23}}{2y}.$$

Otherwise, set $y = 0$ and $z = 1$. The axis and angle of rotation are then determined by the fact that

$$\cos \frac{\theta}{2} = w, \quad a_x \sin \frac{\theta}{2} = x, \quad a_y \sin \frac{\theta}{2} = y, \quad a_z \sin \frac{\theta}{2} = z.$$

Note that this differs from [8] in that we assume the matrix multiplies a vector on the left; that is, $\mathbf{y} = R\mathbf{x}$.

4. The iterative solution. We now describe the iterative solution to the axis-angle problem. The solution uses the inverse power method which Wilkinson [10] claims is the most powerful method for determining eigenvectors once the eigenvalues have been determined. For completeness, we will first describe the power method [2], [3], [10].

The power method is an iterative method for finding the dominant eigenvalue (if one exists) and corresponding eigenvector of a matrix A . The dominant eigenvalue is the one with the largest absolute value. Let A be an $n \times n$ matrix which is similar to a diagonal matrix $A = PDP^{-1}$ with D diagonal and P nonsingular. Let $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of A and $\mathbf{x}_1, \dots, \mathbf{x}_n$ denote the corresponding eigenvectors which form a basis for C^n so that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$. Also, assume that

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|;$$

that is, A has a single dominant eigenvalue λ_1 . Let $\mathbf{z}^{(0)} \in C^n$ be an arbitrarily chosen initial column vector having n components which can be expressed as a linear combination of the eigenvectors \mathbf{x}_i as

$$\mathbf{z}^{(0)} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \dots + \alpha_n\mathbf{x}_n$$

with $\alpha_1 \neq 0$. We then form the sequence $\mathbf{z}^{(0)}, \mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \dots$, where

$$\mathbf{z}^{(1)} = A\mathbf{z}^{(0)}, \quad \mathbf{z}^{(m)} = A\hat{\mathbf{z}}^{(m-1)}, \quad \text{and} \quad \hat{\mathbf{z}}^{(m-1)} = \frac{\mathbf{z}^{(m-1)}}{\|\mathbf{z}^{(m-1)}\|}$$

for $m = 2, 3, \dots$, which will converge to an eigenvector of λ_1 . Also, the sequence

$$r_m = \frac{\hat{\mathbf{z}}^{(m)T} \cdot A\hat{\mathbf{z}}^{(m)}}{\hat{\mathbf{z}}^{(m)T} \cdot \hat{\mathbf{z}}^{(m)}}$$

for $m = 1, 2, \dots$ will approach the dominant eigenvalue λ_1 .

The power method is *not* very general. It assumes the existence of a dominant eigenvalue, which is not always satisfied! Consider an *orthogonal* matrix; that is, a matrix such that $OO^T = O^T O = I$ so that $O^{-1} = O^T$. Now let λ be an eigenvalue of O and \mathbf{x} be a corresponding eigenvector. Then

$$\|O\mathbf{x}\|_2^2 = (O\mathbf{x}, O\mathbf{x}) = (O\mathbf{x})^T \cdot O\mathbf{x} = \mathbf{x}^T O^T O\mathbf{x} = \mathbf{x}^T \cdot \mathbf{x} = \|\mathbf{x}\|_2^2.$$

So $\|O\mathbf{x}\|_2 = \|\mathbf{x}\|_2$. Also, since $O\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$, we have

$$\|\mathbf{x}\|_2 = \|O\mathbf{x}\|_2 = \|\lambda\mathbf{x}\|_2 = |\lambda| \|\mathbf{x}\|_2$$

so that $|\lambda| = 1$. Since λ was arbitrary, this says that all of the eigenvalues of an orthogonal matrix have absolute value 1. Thus, orthogonal matrices do not have dominant eigenvalues. This is the case of the rotation matrix in our example. It is an example of an orthogonal matrix. Therefore, the power method fails for our application.

We now describe the inverse power method. Recall that if A is invertible and has eigenvalues $\lambda_1, \dots, \lambda_n$, then A^{-1} has eigenvalues $\lambda_1^{-1}, \dots, \lambda_n^{-1}$. Let s be some scalar and construct $B = A - sI$. Then, if B is invertible, the eigenvalues of B^{-1} are $(\lambda_1 - s)^{-1}, \dots, (\lambda_n - s)^{-1}$. Now suppose we apply the power method to B^{-1} . Starting with the initial vector $\mathbf{z}^{(0)}$, we get

$$\mathbf{z}^{(m)} = B^{-m} \mathbf{z}^{(0)} = (A - sI)^{-m} \mathbf{z}^{(0)} = \sum_{i=1}^n \alpha_i (\lambda_i - s)^{-m} \mathbf{x}_i.$$

Thus, if s is close to one of the λ_i (which we will assume has multiplicity one even though it is possible to handle the case of an eigenvalue of multiplicity greater than one) and if $|\lambda_i - s| < |\lambda_j - s|$ for $j \neq i$ (s is closer to λ_i than to any other eigenvalue), then, as long as $\alpha_i \neq 0$, the term $\alpha_i (\lambda_i - s)^{-m} \mathbf{x}_i$ will become very large relative to the other terms in the sum for sufficiently large m . The vector $\mathbf{z}^{(m)}$ is then a very good approximation to the eigenvector corresponding to the eigenvalue λ_i . To keep things bounded, we normalize the vector at each iteration as in the power method. Also, we do not compute B^{-1} but rather solve the system $B\mathbf{z}^{(i+1)} = \hat{\mathbf{z}}^{(i)}$.

The inverse power method can be summarized as follows: If $\mathbf{z}^{(0)}$ has a nonzero component in the direction of an eigenvector for λ_i , then the sequence $\mathbf{z}^{(0)}, \mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \dots$, where

$$B\mathbf{z}^{(1)} = \mathbf{z}^{(0)}, \quad B\mathbf{z}^{(m)} = \hat{\mathbf{z}}^{(m-1)}, \quad \text{and} \quad \hat{\mathbf{z}}^{(m-1)} = \frac{\mathbf{z}^{(m-1)}}{\|\mathbf{z}^{(m-1)}\|}$$

for $m = 2, 3, \dots$, will converge to an eigenvector of λ_i . In addition, the estimate for λ can be improved using

$$\lambda_m = \begin{cases} s + \frac{1}{\mathbf{z}^{(0)T} \mathbf{z}^{(1)}} & \text{for } m = 1, \\ s + \frac{1}{\hat{\mathbf{z}}^{(m-1)T} \mathbf{z}^{(m)}} & \text{for } m = 2, 3, \dots \end{cases}$$

Looking at this procedure carefully, we see that we are repeatedly solving a system of linear equations with the same coefficient matrix but a different right-hand side. Therefore, we should use the triangular factorization of B to perform this algorithm; that is, we factorize B once and repeatedly use forward and back substitution to produce each approximation. This is an example of the advantage of having the triangular factorization. Thus, if $B = LU$ where L is a unit lower-triangular matrix and U is an upper-triangular matrix, then we have

$$\begin{aligned} L\mathbf{y}^{(m+1)} &= \mathbf{z}^{(m)} \quad (\text{solve by forward substitution}), \\ U\mathbf{w}^{(m+1)} &= \mathbf{y}^{(m+1)} \quad (\text{solve by back substitution}), \\ \mathbf{z}^{(m+1)} &= \frac{\mathbf{w}^{(m+1)}}{\|\mathbf{w}^{(m+1)}\|}. \end{aligned}$$

The norm used to normalize the approximations can be any norm which may be appropriate for the application.

In addition, if s is very close to the eigenvalue λ_i , the matrix B will be close to noninvertible. However, as long as we use some form of partial pivoting in computing the factorization, the algorithm should not lose accuracy [7]. See Peters and Wilkinson [6] for a more detailed discussion of this situation.

In the above algorithm, we need to choose an initial vector $\mathbf{z}^{(0)}$ for the iteration process. One approach taken by Wilkinson [9] which has been very successful is to set $\mathbf{z}^{(0)} = L\mathbf{e}$ where $\mathbf{e} = (1, 1, \dots, 1)$. Therefore, $\mathbf{z}^{(0)}$ is never directly computed unless we need it to update λ in the case $m = 1$. If $\mathbf{z}^{(0)}$ is not needed, the first iteration just solves $U\mathbf{w}^{(1)} = \mathbf{e}$ by back substitution.

We have the following theorem about the convergence of the inverse power method [2].

THEOREM 1. *Let A be diagonalizable and λ be an eigenvalue which may or may not be simple. Suppose that $\tilde{\lambda}$ satisfies*

$$\tilde{\lambda} \neq \lambda \text{ and } |\tilde{\lambda} - \lambda| < |\tilde{\lambda} - \mu| \text{ for all } \mu \in \sigma(A) - \lambda$$

and that $\mathbf{z}^{(0)}$ is not in the subspace spanned by the eigenvectors corresponding to the eigenvalues which are distinct from λ . Then, if $\|\mathbf{x}\|$ is any vector norm,

$$\lim_{m \rightarrow \infty} \left(\frac{(\lambda - \tilde{\lambda})^m}{|\lambda - \tilde{\lambda}|^m} \frac{\mathbf{z}^{(m)}}{\|\mathbf{z}^{(m)}\|} \right) = \mathbf{x},$$

where \mathbf{x} is an eigenvector corresponding to the eigenvalue λ .

Note that if λ and $\tilde{\lambda}$ are real, then either

$$\lim_{m \rightarrow \infty} \frac{\mathbf{z}^{(m)}}{\|\mathbf{z}^{(m)}\|} = \mathbf{x} \text{ if } \tilde{\lambda} < \lambda \text{ or } \lim_{m \rightarrow \infty} (-1)^m \frac{\mathbf{z}^{(m)}}{\|\mathbf{z}^{(m)}\|} = \mathbf{x} \text{ if } \tilde{\lambda} > \lambda.$$

Thus, by just looking at the convergence of the method, we can tell if $\tilde{\lambda}$ is an overapproximation to λ or an underapproximation to λ .

5. Numerical results. In this section we present numerical results based on a C implementation of the inverse power method as described above. All floating point computations were performed using the float type. The triangular factorization section of the code is based on the scaled partial pivoting algorithm described in Conte and de Boor [3]. All results are shown to nine decimal places so that the full single precision accuracy is displayed.

Example. Recall the problem of finding the axis and angle of rotation of a general three-dimensional rotation matrix. Let $\mathbf{x} = (1.0, 2.0, 3.0)$ and $\mathbf{a} = \mathbf{x}/\|\mathbf{x}\|_2$. Let $\theta = \pi/4$. Then, the rotation matrix is given by

$$R = \begin{pmatrix} 0.728027761 & -0.525104821 & 0.440727293 \\ 0.608788550 & 0.790790558 & -0.063456595 \\ -0.315201640 & 0.314507872 & 0.895395279 \end{pmatrix}.$$

Using (2.1), we have $\cos \theta = 0.707106829$. Since $\lambda = 1$ is the eigenvalue corresponding to the axis of rotation, we apply inverse iteration with $\tilde{\lambda} = 0.9999$. This gives the results in Table 5.1. Finally, (2.2) gives $\sin \theta = 0.707106769$, and (2.3) gives $\theta = 0.785398126$. Notice that since $\tilde{\lambda} < \lambda$, there is no sign change in the approximations to \mathbf{a} .

Now applying inverse iteration with $\tilde{\lambda} = 1.0001$, we have the results given in Table 5.2. Finally, (2.2) gives $\sin \theta = -0.707106709$, and (2.3) gives $\theta =$

TABLE 5.1
Axis-angle problem numerical results: $\tilde{\lambda} = 0.9999$.

| Iteration | a_x | a_y | a_z | λ |
|-----------|-------------|-------------|-------------|-------------|
| 0 | 0.553420901 | 1.000000000 | 0.148788154 | |
| 1 | 0.267348856 | 0.534426033 | 0.801818788 | 1.000055552 |
| 2 | 0.267261237 | 0.534522474 | 0.801783681 | 0.999999940 |
| 3 | 0.267261267 | 0.534522474 | 0.801783681 | 0.999999940 |

TABLE 5.2
Axis-angle problem numerical results: $\tilde{\lambda} = 1.0001$.

| Iteration | a_x | a_y | a_z | λ |
|-----------|--------------|--------------|--------------|-------------|
| 0 | 0.553092360 | 1.000000000 | 0.149040759 | |
| 1 | -0.267173707 | -0.534618914 | -0.801748633 | 0.999944448 |
| 2 | 0.267261237 | 0.534522474 | 0.801783741 | 1.000000000 |
| 3 | -0.267261267 | -0.534522474 | -0.801783741 | 1.000000000 |

-0.785398066. Notice that since $\tilde{\lambda} > \lambda$, there are sign changes in the approximations to \mathbf{a} . The direct solution gives

$$\mathbf{q} = (0.923879564, 0.102276444, 0.204552889, 0.306829333)$$

and

$$\theta = 0.785398006, \quad \mathbf{a} = (0.267261267, 0.534522533, 0.801783860),$$

which agree with the above iterative results to single precision accuracy. \square

6. Conclusion. The above example has been used successfully in the classroom to motivate the study on numerical techniques for solving the algebraic eigenvalue problem.

REFERENCES

- [1] S. M. ALESSANDRINI, *A motivational example for the numerical solution of two-point boundary-value problems*, SIAM Rev., 37 (1995), pp. 423–427.
- [2] P. G. CIARLET, *Introduction to Numerical Linear Algebra and Optimisation*, Cambridge University Press, Cambridge, 1989.
- [3] S. D. CONTE AND C. DE BOOR, *Elementary Numerical Analysis: An Algorithmic Approach*, 3rd ed., McGraw-Hill, New York, 1980.
- [4] I. D. FAUX AND M. J. PRATT, *Computational Geometry for Design and Manufacture*, John Wiley and Sons, New York, 1985.
- [5] W. R. HAMILTON, *On quaternions; or on a new system of imaginaries in algebra*, Philosophical Magazine, XXV (1844), pp. 10–13.
- [6] G. PETERS AND J. H. WILKINSON, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Rev., 21 (1979), pp. 339–360.
- [7] A. RALSTON AND P. RABINOWITZ, *A First Course in Numerical Analysis*, 2nd ed., McGraw-Hill, New York, 1978.
- [8] K. SHOEMAKE, *Animating rotations with quaternion curves*, SIGGRAPH '85, 19 (1985), pp. 245–254.
- [9] J. H. WILKINSON, *Rounding Errors in Algebraic Processes*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- [10] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

LINKED CITATIONS

- Page 1 of 1 -



You have printed the following article:

A Motivational Example for the Numerical Solution of the Algebraic Eigenvalue Problem

Stephen M. Alessandrini

SIAM Review, Vol. 40, No. 4. (Dec., 1998), pp. 935-940.

Stable URL:

<http://links.jstor.org/sici?sici=0036-1445%28199812%2940%3A4%3C935%3AAMEFTN%3E2.0.CO%3B2-E>

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

References

¹ **A Motivational Example for the Numerical Solution of Two-Point Boundary-Value Problems**

Stephen M. Alessandrini

SIAM Review, Vol. 37, No. 3. (Sep., 1995), pp. 423-427.

Stable URL:

<http://links.jstor.org/sici?sici=0036-1445%28199509%2937%3A3%3C423%3AAMEFTN%3E2.0.CO%3B2-A>

⁶ **Inverse Iteration, Ill-Conditioned Equations and Newton's Method**

G. Peters; J. H. Wilkinson

SIAM Review, Vol. 21, No. 3. (Jul., 1979), pp. 339-360.

Stable URL:

<http://links.jstor.org/sici?sici=0036-1445%28197907%2921%3A3%3C339%3AIIIEAN%3E2.0.CO%3B2-T>

NOTE: *The reference numbering from the original has been maintained in this citation list.*