

A TECHNIQUE TO SHIFT AN EIGENVALUE OF A COMPLEX MATRIX TO ACCELERATE CONVERGENCE OF THE POWER AND INVERSE POWER METHOD

K. J. BUNCH and R. W. GROW

Microwave Device and Physical Electronics Laboratory, Department of Electrical Engineering
University of Utah, 3280 Merrill Engineering Building, Salt Lake City, UT 84112, USA

(Received August 1989; in revised form January 1990)

Abstract—A technique is presented to shift an eigenvalue of a complex matrix. It can be used in the power and inverse power method to accelerate convergence to a single eigenvalue or to eliminate an eigenvalue degeneracy. Error bounds are presented that indicate how the remaining eigenvalues are perturbed with each shift.

BACKGROUND

The power method and inverse power method are two convenient numerical techniques for calculating the maximum and minimum eigenvalues of a matrix [1-4]. These methods begin with a guess to an eigenvector and successively refine it until an eigenvalue/eigenvector pair is obtained. For example, in the power method, the largest eigenvalue and its corresponding eigenvector is obtained. The initial guess is assumed to be composed of a sum of the eigenvectors of a matrix \overline{M} :

$$\vec{v}^{(1)} = \sum_{i=1}^n \alpha_i \vec{v}_i \quad (1)$$

where \vec{v}_i is an eigenvector of the $n \times n$ matrix \overline{M} so that

$$\overline{M} \vec{v}_i = \lambda_i \vec{v}_i \quad (2)$$

Repeated multiplication of equation (1) by the matrix produces

$$(\overline{M})^k \vec{v}^{(1)} = \vec{v}^{(k)} = \lambda_1^{(k)} \left\{ \alpha_1 \vec{v}_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \right\} \vec{v}_i \quad (3)$$

where λ_1 is assumed to be the largest eigenvalue. If this eigenvalue is assumed to be unique, then the sum in equation (3) will approach zero as k approaches infinity and

$$\lambda_1 = \frac{\vec{v}^{(k+1)\dagger} \vec{v}^{(k)}}{\vec{v}^{(k)\dagger} \vec{v}^{(k)}} \quad \text{as } k \rightarrow \infty \quad (4)$$

where \dagger denotes the complex conjugate transpose. In a similar manner, the inverse power method produced the smallest eigenvalue. For k iterations, equation (3) indicates that the error term is proportional to $(\lambda_2/\lambda_1)^k$, where λ_2 is the eigenvalue closest to λ_1 . If λ_2 and λ_1 are close together, then convergence to the proper eigenvalue/eigenvector pair will be slow. If it were possible to shift the maximum eigenvalue in this case, then convergence would become proportional to $[\lambda_2/(\lambda_1 + \lambda_s)]^k$, where λ_s is a shift variable. This paper presents a method to create this shift.

Present shifting methods used in the inverse power method shift the entire spectrum of eigenvalues according to [1-4]:

$$(\overline{M} - \lambda_s \overline{I}) \vec{v}^{(k+1)} = \vec{v}^{(k)} \quad (5)$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ -TEX

where \bar{I} is the identity matrix. By decomposing the matrix into its eigenvector decomposition, it is easily seen that this formula shifts all the eigenvalues [5,6]

$$\begin{aligned}\bar{M} - \lambda_s \bar{I} &= \bar{v}^\dagger \bar{\lambda} \bar{v} - \lambda_s \bar{I} \bar{v}^\dagger \bar{v} \\ &= \bar{v}^\dagger (\bar{\lambda} - \lambda_s \bar{I}) \bar{v}\end{aligned}\quad (6)$$

where $\bar{\lambda}$ is a diagonal matrix of eigenvalues and \bar{v} is the matrix of normalized eigenvectors. Equation (6) shows that the shift variable is subtracted from every one of the eigenvalues.

THE SHIFTING OF A SINGLE EIGENVALUE

Equation (6) suggests a simple way to shift a single eigenvalue. Create a new matrix defined by

$$\widetilde{\bar{M}} = \bar{v}^\dagger (\bar{\lambda} - \bar{\lambda}_s) \bar{v} \quad (7)$$

where the matrix $\bar{\lambda}_s$ has zeros everywhere except at the k th diagonal element at which the element is λ_s . This new matrix has eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{k-1}, (\lambda_k - \lambda_s), \lambda_{k+1}, \dots, \lambda_n$, i.e., the same eigenvalues as before, but with a single one shifted by λ_s . Multiplying equation (7) produces

$$\widetilde{\bar{M}} = \bar{v}^\dagger (\bar{\lambda} - \lambda_s \bar{e}_k \bar{e}_k^\dagger) \bar{v} = \bar{M} - \lambda_s \bar{v}_k \bar{v}_k^\dagger \quad (8)$$

where \bar{e}_k is the unit vector with one at its k th element and with zeros elsewhere; $\bar{e}_k \bar{e}_k^\dagger$ is an outer product, and it produces an $n \times n$ matrix with a nonzero element only at the k th diagonal location.

For the shift defined by equation (8) to work, the eigenvector \bar{v}_k must be known. During the process of the power and inverse power methods, each iteration creates a successively better approximation to this eigenvector. In the power method, this eigenvector corresponds to the largest eigenvalue, whereas in the inverse power, it corresponds to the smallest eigenvalue. This shift can then be applied to one of these two eigenvalues using equation (8) during the course of iteration.

PERTURBATION BOUNDS OF AN INEXACT EIGENVECTOR FOR SYMMETRIC MATRICES

The shift defined by equation (8) assumes that the eigenvector \bar{v}_k is exact. Unfortunately, the eigenvector used is likely to contain error, and the shift defined by equation (8) will likely perturb all the eigenvalues of the original matrix. It will be necessary to know bounds for these perturbations in order to determine when this shift can be applied.

Perturbation bounds for the eigenvalues are derived from theorems due to Wielandt and Hoffman [7] and discussed in Lawson and Hanson [8]. These theorems are derived for symmetric matrices; the authors know of no similar theorems for general matrices.

Given three $n \times n$ symmetric matrices such that

$$\widetilde{\bar{M}} = \bar{M} + \bar{P} \quad (9)$$

denote their eigenvalues by $\tilde{m}_i, m_i, p_i, i = 1, \dots, n$. Then

$$\sum_{i=1}^n (\tilde{m}_i - m_i)^2 \leq |\bar{P}|_F^2 \quad (10)$$

where F denotes the Frobenius norm [9].

Assume now that the eigenvector used in equation (8) is close to the actual eigenvector so that

$$\tilde{\bar{v}}_k = \bar{v}_k + \delta \bar{v} \quad (11)$$

$$\tilde{\bar{\lambda}}_k = \lambda_k + \delta \lambda \quad (12)$$

$$\overline{M}\tilde{v}_k = \lambda_k \tilde{v}_k \quad (13)$$

Creating the shifted matrix based on the vector \tilde{v}_k produces

$$\overline{S} = \overline{M} - \lambda_s \tilde{v}_k \tilde{v}_k^\dagger = \widetilde{\overline{M}} - \lambda_s (\tilde{v}_k \delta \tilde{v}^\dagger + \delta \tilde{v} \tilde{v}_k^\dagger + \delta \tilde{v} \delta \tilde{v}^\dagger) \quad (14)$$

where $\widetilde{\overline{M}}$ is defined by equation (8). The remaining terms in equation (14) are perturbational terms so that, from equation (10),

$$\begin{aligned} \sum_{i=1}^n (\tilde{m}_i - m_i)^2 &\leq |\lambda_s|^2 |\tilde{v}_k \delta \tilde{v}^\dagger + \delta \tilde{v} \tilde{v}_k^\dagger + \delta \tilde{v} \delta \tilde{v}^\dagger|_F^2 \\ &\leq |\lambda_s|^2 n (2|\delta \tilde{v}| + |\delta \tilde{v}|^2) \end{aligned} \quad (15)$$

If $|\delta \tilde{v}|$ is small, then the root mean square perturbation of the eigenvalues is approximately bounded by

$$\left\{ \sum_{i=1}^n (\tilde{m}_i - m_i)^2 \right\}^{1/2} \leq \sqrt{n} |\lambda_s| \sqrt{2|\delta \tilde{v}|} \quad (16)$$

Equation (16) indicates that for small systems (n small), small shifts (λ_s small), and small eigenvector errors ($\delta \tilde{v}$ small), the shifting process disturbs the eigenvalues little.

USING SHIFTS TO COMPUTE MULTIPLE EIGENVALUES

The shift defined by equation (8) can be applied after an accurate value for the eigenvector has been obtained, and the resulting eigenvalues will be perturbed little. This fact suggests an algorithm to compute several eigenvalues of a matrix. Assume that the largest eigenvalue λ_1 and its corresponding eigenvector \tilde{v}_1 has been found using the power method. Apply the shift

$$\widetilde{\overline{M}}^{(1)} = \overline{M} - \lambda_1 \tilde{v}_1 \tilde{v}_1^\dagger \quad (17)$$

This shift creates a new matrix with the next largest eigenvalue λ_2 as its largest eigenvalue. Continuing with the power method will produce the eigenvalue λ_2 and its corresponding eigenvector \tilde{v}_2 . The next shift can be applied:

$$\widetilde{\overline{M}}^{(2)} = \widetilde{\overline{M}}^{(1)} - \lambda_2 \tilde{v}_2 \tilde{v}_2^\dagger \quad (18)$$

This process can be continued, but errors will accumulate after each shift. This process is therefore not recommended for more than a limited number of eigenvalues.

Note that in the case of eigenvalue degeneracy, if the eigenvectors are not degenerate, this shift will remove the degeneracy.

NUMERICAL RESULTS

This section presents results testing the error bounds of an eigenvalue shift (equation 16). Matrices with known eigenvalues were created and a single eigenvalue was shifted. The error E is defined as the left-hand side of equation (16):

$$E = \left\{ \sum_{i=1}^n (\tilde{m}_i - m_i)^2 \right\}^{1/2} \quad (19)$$

Shifting used with the power method is considered first. Figure 1 shows the effect of an eigenvalue shift for different eigenvector errors ($|\delta \tilde{v}|$ in equation 16). The matrices tested had 4, 8, 16, and 32 rows, and convergence was to the largest eigenvalue whose value was 4. The staircase nature

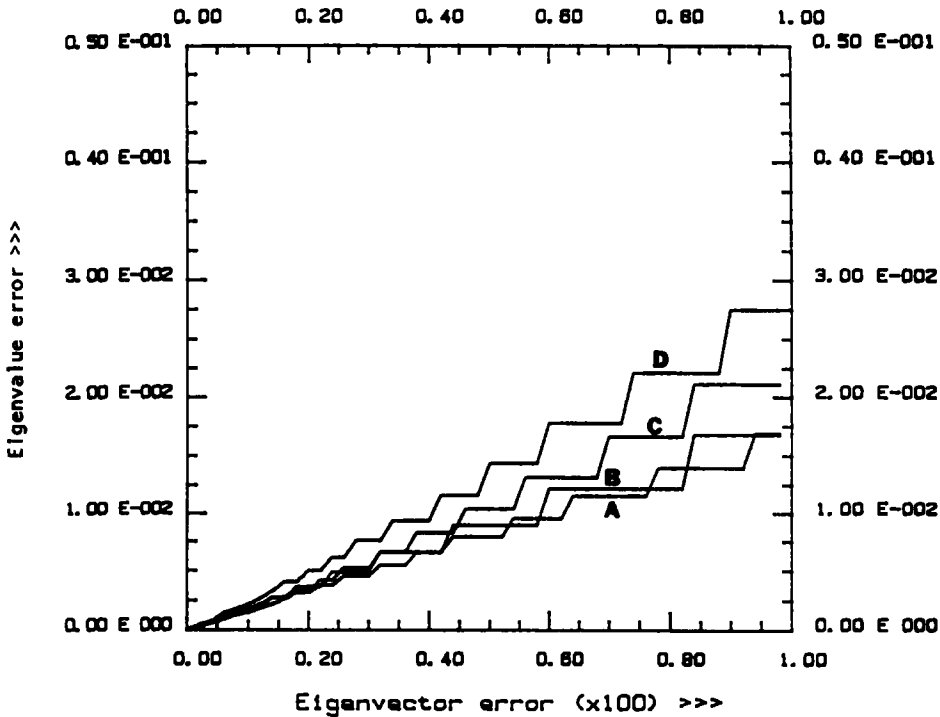


Figure 1. The error of the perturbed eigenvalues of the matrix \bar{M} . The largest eigenvalue (4) is shifted to zero at the eigenvector errors shown for (a) 4×4 , (b) 8×8 , (c) 16×16 , and (d) 32×32 matrices.

of the plot is from the discrete nature of the shifting process. The shift is applied only after a certain number of iterations in which the error is below that specified.

Figure 1 shows two aspects of the eigenvalue shift. First, increasing the matrix size from 4 to 32 only approximately doubles the error instead of increasing it by 2.8, as equation (16) suggests. Second, the approximately linear increase in the eigenvalue error with an increase in $|\delta\bar{v}|$ is somewhat worse than expected.

Figure 2 shows how variations in the amount of eigenvalue shift ($|\lambda_s|$ in equation 16) affect the error of the perturbed eigenvalues for 16×16 matrices. The largest eigenvalue is shifted below an eigenvector error $|\delta\bar{v}|$ of 1×10^{-4} . Although the eigenvalue errors are widely scattered, they seem to show a linear increase with an increase in the amount of shift $|\lambda_s|$, as expected with equation (16).

Since equation (16) is an upper bound on the eigenvalue error, it is useful to take a specific case and show how restrictive this bound is. Consider the maximum error in Figure 2. Equation (16) gives

$$E \leq \sqrt{16}(12.75)\sqrt{2 \times 10^{-4}} = 0.7212$$

The actual error is about 1×10^{-3} , almost three orders of magnitude less than that given by equation (16).

Consider now the inverse power method. Figure 3 shows the perturbation error on the eigenvalues as the smallest eigenvalue varies between 0 and 1. This eigenvalue is shifted to (nearly) zero and iteration proceeds. The final eigenvalue is the sum of this shift and the new eigenvalue. The next eigenvalue is 1, so that this case tests the range of shifting an isolated eigenvalue to a degenerate one. Shifting is at an eigenvector error $|\delta\bar{v}|$ of 1×10^{-4} , as in the previous case.

Little difference exists between the 8×8 and the 32×32 matrices tested. The error increases with larger shift, as with the power method, but the error is considerably smaller ($< 1 \times 10^{-8}$). The reason for this difference is that iterations continue to the smallest eigenvalue after the shift. The final eigenvalue is a sum of the first shift and a "residual" eigenvalue left from an inexact

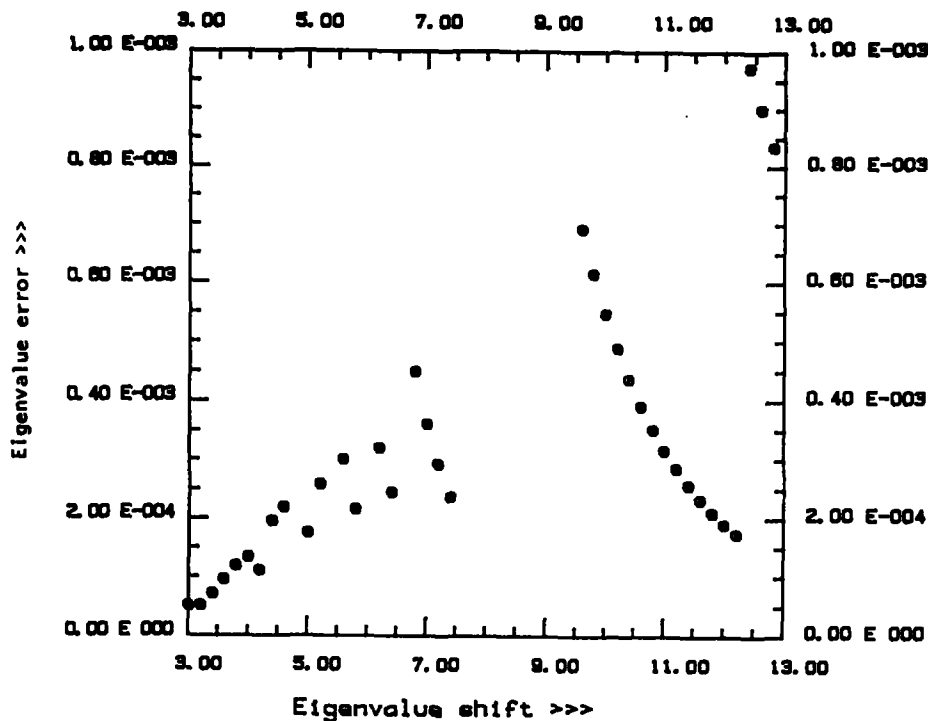


Figure 2. The error of the perturbed eigenvalues of the 16×16 matrix \bar{M} . The largest eigenvalue (shifted to zero) varies between 3 and 13.

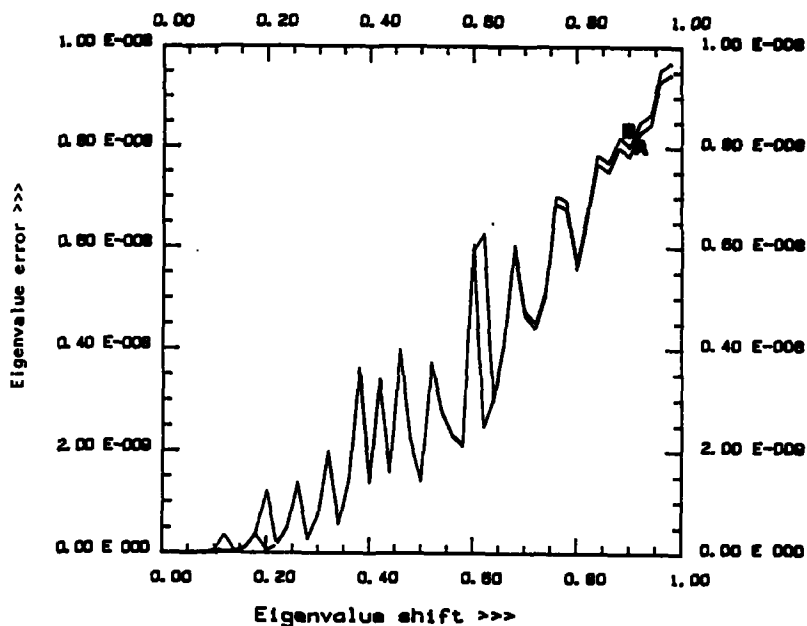


Figure 3. The eigenvalue error in the inverse power method with varying shift for (a) 8×8 and (b) 32×32 matrices.

shift to zero. This residual corrects the final eigenvalue to be considerably more accurate than that suggested by equation (16).

Figure 4 shows how shifting can accelerate the convergence of the inverse power method. 8×8 matrices are tested, and the smallest eigenvalue varies between 0 and 1, with a degeneracy at 1. The number of iterations needed for convergence increases the closer the minimum eigenvalue

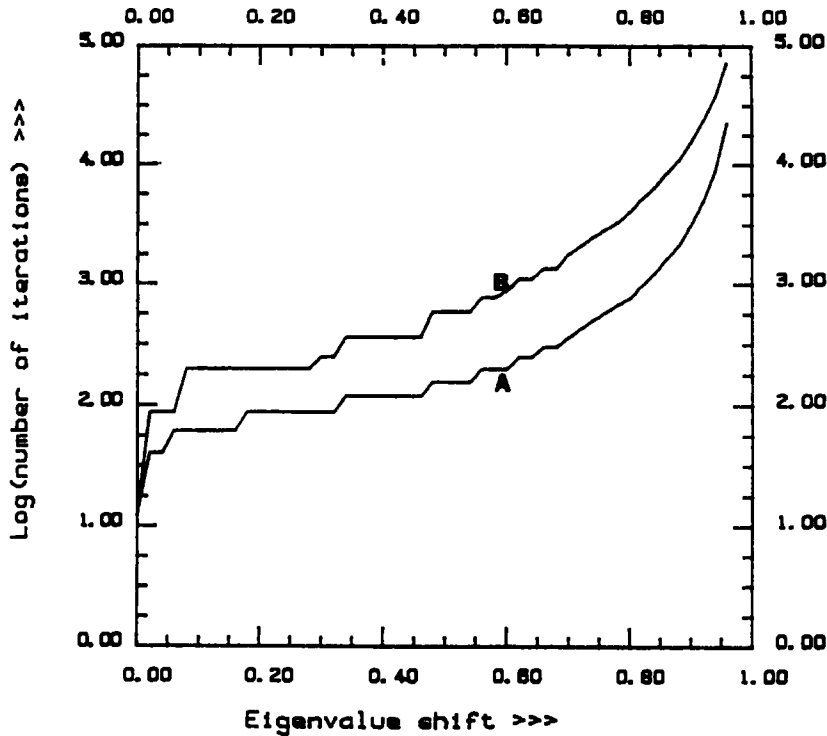


Figure 4. The natural logarithm of the number of iterations needed (a) with and (b) without shifting with variations in the size of this shift (8×8 matrices).

gets to being degenerate, as expected. Shifting approximately decreases the number of iterations needed by one half. Of course, this case has not been optimized to better this factor, so greater acceleration is probably possible, especially considering the possibility of multiple shifts.

CONCLUSIONS

A technique has been presented that is able to shift a single eigenvalue of a matrix. This shift can be used in the power or inverse power methods to accelerate convergence to an eigenvalue. It can also be used to compute successive eigenvalues by shifting the largest eigenvalue to zero.

REFERENCES

1. L.W. Johnson and R.D. Riess, *Numerical Analysis*, Addison-Wesley, Reading, MA, pp. 74-89, (1977).
2. A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, McGraw-Hill Book Company, New York, pp. 492-501, (1978).
3. G. Dahlquist and A. Björck, *Numerical Methods*, Prentice-Hall, Inc., Englewood Cliffs, NJ, pp. 209-211, (1974).
4. A.W. Al-Khafaji and J.R. Tooley, *Numerical Methods in Engineering Practice*, Holt, Rinehart, and Winston, Inc., New York, pp. 234-237, (1986).
5. L.W. Johnson and R.D. Riess, *Numerical Analysis*, Addison-Wesley, Reading, MA, pp. 105-110, (1977).
6. A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, McGraw-Hill Book Company, New York, pp. 488-492, (1978).
7. A.J. Hoffman and H.W. Wielandt, The variation of the spectrum of a normal matrix, *Duke Mathematics Journal* 20, 37-39 (1953).
8. C.L. Lawson and R.J. Hanson, *Solving Least-Squares Problems*, Prentice-Hall, Inc., Englewood Cliffs, NJ, pp. 23-27.
9. A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, McGraw-Hill Book Company, New York, pp. 6-9, (1978).