Title
INVERSE ITERATION ON DEFECTIVE MATRICES

Author
Chen, Nai-Fu.

Publication Date
1976

Submitted to Mathematics of Computation

LBL-4653
~~Preprint~~ c.1

INVERSE ITERATION ON DEFECTIVE MATRICES

Nai-fu Chen

January 1976

Prepared for the U. S. Energy Research and Development Administration under Contract W-7405-ENG-48

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA BERKELEY

DISCLAIMER

INVERSE ITERATION ON DEFECTIVE MATRICES*

By Nai-fu Chen

ABSTRACT

Very often, inverse iterations are used with shifts to accelerate convergence to an eigenvector. In this paper, it is shown that, if the eigenvalue belongs to a nonlinear elementary divisor, the vector sequences may diverge even when the shift sequences converge to the eigenvalue. The local behavior is further displayed through a 2 x 2 example.

INTRODUCTION

If an accurate approximation σ to an eigenvalue λ of a matrix B is available, then inverse iteration is an attractive technique for computing the associated eigenvector. We choose an arbitrary unit vector v_o and a fixed shift σ. Then for $j = 1,2,\ldots$ we solve

$$(1) \qquad (B - \sigma I)\, w_j = v_{j-1}, \qquad v_j = w_j/\|w_j\|$$

AMS (MOS) subject classifications (1970). Primary 65F15; Secondary 15A18.

Key Words and Phrases: Eigenvector, eigenvalue, nonlinear divisors, convergence, inverse iterations, shifts.

*This work was partially supported by the Office of Naval Research under Contract No. N00014-69-A-0200-1017.

This work was performed under the auspices of the Energy Research Development Administration

where $\|\cdot\|$ is the user's preferred vector norm.

If λ is a simple eigenvalue with unit eigenvector x and if v_o is not an unfortunate choice, then the vector sequence $\{v_j\}$ converges linearly to x and the convergence factor is very favorable. This well known result holds also for multiple eigenvalues λ provided that:

(*i*) the dimension of λ's eigenspace is equal to λ's algebraic multiplicity (i.e. linear elementary divisor),

(*ii*) the spectral projection of v_o onto λ's eigenspace is not zero (i.e. the starting vector is not deficient in x).

If σ is known to equal λ to within the working precision of the computer, then only one or two steps of the iteration are necessary.

Wilkinson and Varah [2&3] pointed out that the situation is not so nice if λ has generalized eigenvectors of grade higher than one, i.e., when λ belongs to a nonlinear elementary divisor. In exact arithmetic the iteration converges not linearly, but harmonically like 1/j as $j \to \infty$. Even worse is the fact that except for very special choices of v_o, the vectors v_2 and v_3 will be poorer approximations than v_1! This fact runs counter to our natural intuition and is hard to understand.

Inverse iteration can also be used with variable shifts. If the sequence of shifts $\{\sigma_j\}$ converges to λ as $j \to \infty$ then the vector sequence generated by

$$(B - \sigma_j I)\, w_j = -v_{j-1} \quad \text{(the - sign is for convenience)}$$

(2) $$v_j = w_j / \|W_j\|$$

converges quadratically or better to x whenever λ is a simple eigenvalue. However, when λ has eigenvectors of grade higher than one, the situation is again complicated and the shifts can make things worse. In fact we shall prove the following surprising result. THE SEQUENCE $\{v_j\}$ GENERATED BY (2) MAY FAIL TO CONVERGE TO x EVEN THOUGH THE SEQUENCE $\{|\sigma_j - \lambda|\}$ CONVERGES MONOTONICALLY TO 0 AS $j \to \infty$.

The 2 x 2 Case

There is no loss of generality in studying

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

(for analogous results on n x n matrices, see Chen [1]. We observe that $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is the eigenvector and any other vector is an eigenvector of grade 2. Also

$$(N - \sigma I)w = -e_1 \text{ yields } w = \begin{cases} \tau e_1 - e_2, \text{ for any } \tau, & \text{if } \sigma = 0 \\ \sigma^{-1} e_1 & , \text{if } \sigma \neq 0. \end{cases}$$

This shows that e_1 is a fixed point of the iteration (2) provided that $\sigma \neq 0$.

For $s = \sin\theta \neq 0$,

$$(N - \sigma I)w = -\begin{pmatrix} c \\ s \end{pmatrix} \text{ yields } w = \begin{cases} \infty & , \text{if } \sigma = 0 \\ c\sigma^{-1}[(1 + t/\sigma)e_1 + te_2], & \sigma \neq 0 \end{cases}$$

where $t = \tan\theta$, $c = \cos\theta$ (thus θ is the angle between $\begin{pmatrix} c \\ s \end{pmatrix}$ and e_1).

Let θ' be the angle between w and e_1, and $t' = \tan\theta'$. Then the iteration function for a typical step is given by

(3) $$t' \equiv \Phi_\sigma(t) = t/(1 + t/\sigma).$$

If we study inverse iteration with shifts $\{\sigma_j\}$ yielding vectors $\{v_j\}$, then in applying (3), we have the following correspondence:

$\{\sigma_j\}$ converges to $\lambda \sim \sigma$ small

$\{v_j\}$ converges to $e_1 \sim t$ small

The fact that $\Phi'_\sigma(0) = 1$, $\sigma \neq 0$, corresponds to the harmonic convergence of the fixed shift sequence. We are now in a position to make our perverse construction, provided $v_o \neq e_1$.

<u>Construction</u>

Set $\sigma = 1$ ($=\|N\|$). While $|t_j| \geq \frac{1}{3}|\sigma|$ continue iterating with current σ. When $|t_j| < \frac{1}{3}|\sigma|$, then set $\sigma \leftarrow t/(t - 1)$, i.e.,

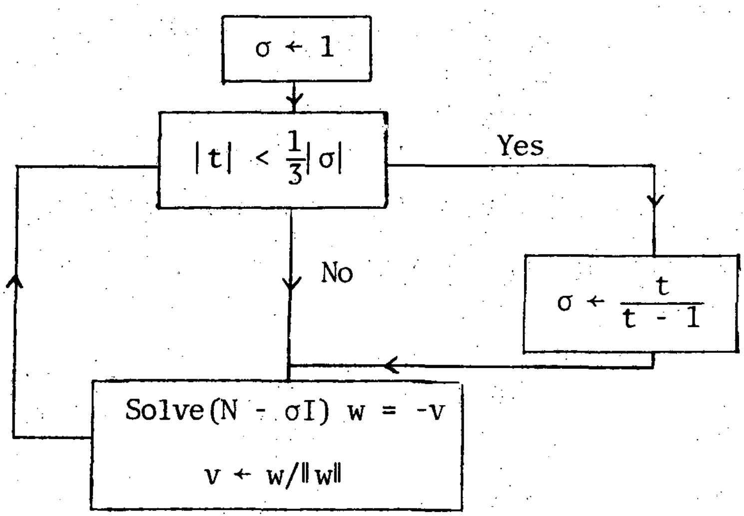

FIGURE 1 THE ITERATION

Since we have convergence for constant σ, the condition must be satisfied with "Yes" eventually. The new σ yields $t' = 1$, i.e., the new v will be $\begin{pmatrix}1\\1\end{pmatrix}/\sqrt{2}$. Moreover since $|t| < 1/3$, the $|\sigma'| < |\frac{3}{2}t| < |\sigma|/2$. Hence the $|\sigma|$'s are monotonic. Do they converge to zero? Yes, because they are at least halved each time. The vector sequence thus generated does not converge since $t = 1$ infinitely often. The trick was to make the factor t/σ in

$$t' = \frac{t}{1 + t/\sigma}$$

go negative periodically when t/σ is small.

It is clear that <u>without this sign reversal</u> (or something like it in the complex case) the variable shifts <u>will</u> accelerate convergence.

<u>Proposition</u>: In the real case if $\{\sigma_j\}$ converges to 0 <u>from one side</u>, then $\{v_j\} \to e_1$ as $j \to \infty$.

<u>Proof</u>: The formula $t' = -\frac{t}{1 + t/\sigma}$ shows that

$$\text{sign}\ (t') = \begin{cases}\text{sign}\ (t), \text{ if } 1 + t/\sigma > 0,\\ \text{sign}\ (\sigma), \text{ if } |t/\sigma| > 1\end{cases}$$

and $|t'| \geq |t|$ if sign $(t) \neq$ sign $(\sigma) \neq$ sign (t'), therefore eventually for some J sign (t_J) = sign (σ_J) and $|t_j|$, $j \geq J$, converges monotonically to 0.

<u>Example</u>: the Rayleigh Quotient Iteration on the matrix N given above. Here σ is constrained to be $sc = \frac{1}{2}\sin 2\theta$, thus

$$t' = \frac{t}{1 + t/\sigma} = \frac{t}{1 + t/sc} = \frac{t}{1 + 1/c^2} < t$$

Thus establishes the linear convergence of RQI on N.

In the following diagram, we demonstrate how v' compares with v as an approximation to the eigenvector. Each point on the diagram represents one step of inverse iteration with a shift σ and the vector v whose components ratio is t:

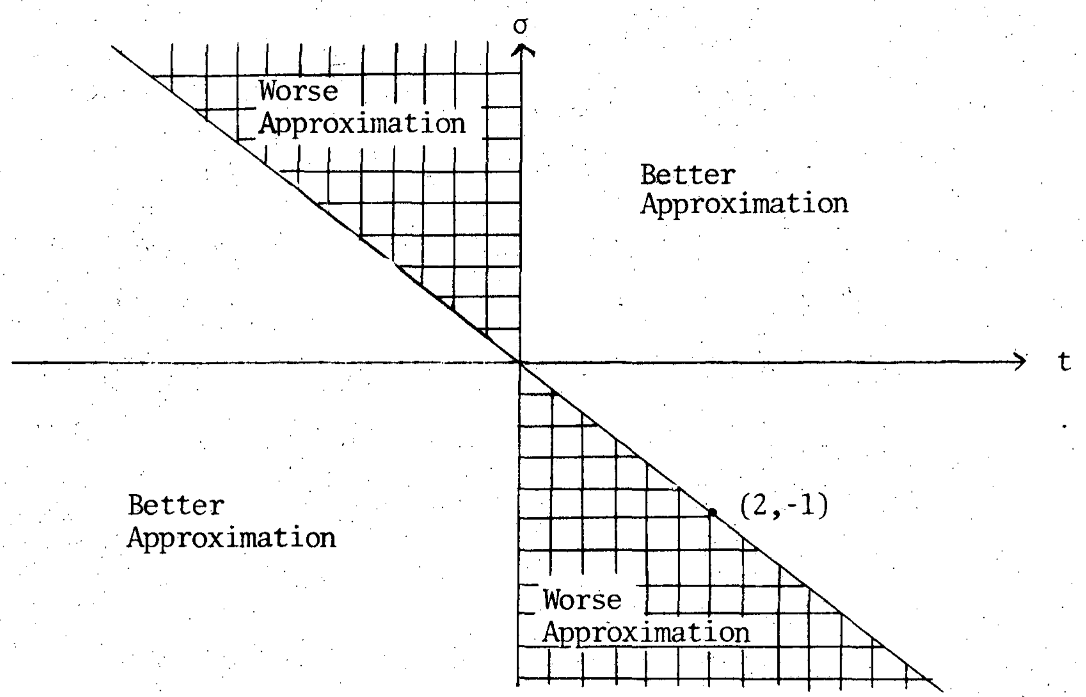

FIGURE 2 BEHAVIOR OF INVERSE ITERATION

Remember that when σ is small σ is a good approximation to the eigenvalue and when t is small v is a good approximation to the eigenvector. The diagram shows that no matter how good our approximation to the eigenvector to eigenvalue or both, inverse iteration can still give a much worse approximation in exact arithmetics.

We can summarize the situation for Inverse Iteration as follows:

ACKNOWLEDGMENT

The author is very grateful to Professor Bereford Parlett for his many suggestions and comments.

REFERENCES

1. N. Chen, The Rayleigh Quotient Iteration for non-normal matrices, Thesis, Electronic Research Laboratory memorandum No. ERL-M548, University of California, Berkeley. 1975.
2. J.M. Varah, The computation of bounds for the invariant subspaces of a general matrix operator, Thesis, Stanford 1967.
3. J.H. Wilkinson, Inverse iteration in theory and in practice, Symposia Mathematica X(1972), pp. 361-379.

TABLE I

vector \ shift	good $\sigma << 1$	bad $\sigma = 1$
good $t << 1$	results variable	good results no improvement
bad $t \geqslant 1$	good result great improvement	results variable

Department of Mathematics and Computing
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720