On the Quadratic Convergence of Kogbetliantz's Algorithm for Computing the Singular Value Decomposition*

C. C. Paige
Computer Science
McGill University
Montreal, Canada

and

P. Van Dooren
Philips Research Laboratory Brussels
Brussels, Belgium

Submitted by G. W. Stewart

ABSTRACT

It is shown that the cyclic Kogbetliantz algorithm ultimately converges quadratically when no pathologically close singular values are present.

1. INTRODUCTION

Kogbetliantz's algorithm for computing the singular value decomposition (SVD) of an arbitrary real or complex matrix A [5,6] has received a great deal of attention recently because of its efficiency as a parallel algorithm [1] and also because of its possible extensions to various other decompositions [3,8]. This article is concerned with the speed of convergence of the Kogbetliantz method. As proved in [2], the method converges under the assumption that the pairs of rotation angles $\{\phi_k, \psi_k\}$ lie in a closed interval J, independent of k:

$$\phi_k, \psi_k \in J, \qquad k=1,2,\dots, \tag{1.1}$$

where J is interior to the interval $(-\pi/2, \pi/2)$. For the Jacobi method

*This work was partly supported by Natural Sciences and Engineering Research Council of Canada Grant A8652.

0024-3795/86/$3.50

applied to Hermitian matrices, this result has been extended to *quadratic* convergence under some mild conditions [4,9,11–13]. Here we extend the property of quadratic convergence to the Kogbetliantz method and discuss under what conditions and for what variants of the method this property can be proved.

The basic method of Kogbetliantz [5,6] for computing the SVD of an $m \times n$ matrix

$$A = U \Sigma V^* \tag{1.2}$$

(where * stands for conjugate transpose) consists of applying transformations U_k, V_k^* such that

$$U^{(0)} = I_m, \qquad V^{(0)} = I_n, \qquad A^{(0)} = A, \tag{1.3a}$$

$$U^{(k+1)} = U_k U^{(k)}, \qquad V^{(k+1)} = V_k V^{(k)}, \qquad A^{(k+1)} = U_k A^{(k)} V_k^*. \tag{1.3b}$$

The updating transformations U_k and V_k are chosen as orthogonal transformations acting only on the pair of rows and columns $\{i_k, j_k\}$, in order to yield zeros in the positions (i_k, j_k) and (j_k, i_k):

$$U_k A^{(k)} V_k^* = \begin{array}{c} \begin{bmatrix} x & \cdot & \cdot & \cdot & x \\ \cdot & x & \cdots & 0 & \cdot \\ \cdot & \vdots & & \vdots & \cdot \\ \cdot & 0 & \cdots & x & \cdot \\ x & \cdot & \cdot & \cdot & x \end{bmatrix} \begin{array}{l} \\ \leftarrow i_k \\ \\ \leftarrow j_k \\ \\ \end{array} \\ \begin{array}{ccccc} \quad & \uparrow & \quad & \uparrow & \quad \\ & i_k & & j_k & \end{array} \end{array} \tag{1.4}$$

Notice that if m and n are unequal and say $m < j_k \leqslant n$, then only one element is zeroed and only one transformation is applied (V_k in the above case). The relevant 2×2 submatrices $\hat{U}_k, \hat{V}_k$ of U_k, V_k, respectively, are denoted as [2]

$$\hat{U}_k = \begin{bmatrix} \alpha_k \cos\phi_k & \beta_k \sin\phi_k \\ -\gamma_k \sin\phi_k & \delta_k \cos\phi_k \end{bmatrix}, \qquad \hat{V}_k = \begin{bmatrix} \xi_k \cos\psi_k & \eta_k \sin\psi_k \\ -\zeta_k \sin\psi_k & \omega_k \cos\psi_k \end{bmatrix}, \tag{1.5}$$

where $\alpha_k, \beta_k, \gamma_k, \delta_k, \xi_k, \eta_k, \zeta_k, \omega_k$ are complex numbers of norm 1, satisfying

$$\alpha_k \delta_k = \beta_k \gamma_k, \qquad \xi_k \omega_k = \eta_k \zeta_k. \tag{1.6}$$

The rotations (1.5) can be constructed to satisfy (1.1) and (1.4) simultaneously.[1] Note that in the real case all $\alpha_k, \beta_k, \gamma_k, \delta_k, \xi_k, \eta_k, \zeta_k, \omega_k$ can be chosen equal to ± 1. Defining $A^{(k)}$ as

$$A^{(k)} = D^{(k)} + E_u^{(k)} + E_l^{(k)}, \tag{1.7}$$

where $D^{(k)}$, $E_u^{(k)}$, and $E_l^{(k)}$ are the matrices containing the elements on, above, and below the diagonal, respectively, we define

$$S^{(k)} = \|E_u^{(k)} + E_l^{(k)}\|_F = \sqrt{\sum_{p \neq q} |a_{p,q}^{(k)}|^2}, \tag{1.8}$$

where $\|\cdot\|_F$ stands for the Frobenius norm. This quantity is used in the sequel as a measure of convergence of $A^{(k)}$ to a diagonal matrix.

2. THE CASE OF DISTINCT SINGULAR VALUES

Let us assume that the singular values of A satisfy

$$|\sigma_i - \sigma_j| \geqslant 2\delta, \tag{2.1}$$

and suppose we have reached the stage when

$$\|A^{(k)} - D^{(k)}\|_F = S^{(k)} < \delta/2. \tag{2.2}$$

Then we have, according to [7, Theorem 5.10],

$$\big| |a_{ii}^{(k)}| - \sigma_i \big| \leqslant \|D^{(k)} - A^{(k)}\|_F < \delta/2 \tag{2.3}$$

[1] In fact, a riguorous analysis [2] shows that this can only be ensured for a weaker version of either (1.1) or (1.4). For details we refer to [2], since this is not important for the sequel.

for some ordering of the σ_i. From this, it also follows that

$$\begin{aligned}\left||a_{ii}^{(k)}| - |a_{jj}^{(k)}|\right| &= \left|\left(|a_{ii}^{(k)}| - \sigma_i\right) - \left(|a_{jj}^{(k)}| - \sigma_j\right) + (\sigma_i - \sigma_j)\right| \\ &\geqslant |\sigma_i - \sigma_j| - \left||a_{ii}^{(k)}| - \sigma_i\right| - \left||a_{jj}^{(k)}| - \sigma_j\right| \\ &> 2\delta - \frac{\delta}{2} - \frac{\delta}{2} \\ &= \delta. \end{aligned} \tag{2.4}$$

Since the rotations (1.4) reduce $S^{(k)}$ at each step, (2.4) also holds for all subsequent k. We now show that under the above assumption (2.1), after one cycle of

$$N = \frac{\max\{m,n\}(\max\{m,n\} - 1)}{2} \tag{2.5}$$

rotations (1.4) including all possible (i_k, j_k) pairs, we have

$$S^{(k+N)} \leqslant \frac{[S^{(k)}]^2}{\delta}. \tag{2.6}$$

The result is proved in a similar fashion to [13] but first requires the following lemma, obtained from [10, Theorem 6.3].

LEMMA 1. *Let the* 2×2 *matrix*

$$M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \tag{2.7}$$

be such that

$$\gamma/\delta < \tfrac{1}{2} \tag{2.8}$$

for

$$\gamma = \sqrt{|m_{12}|^2 + |m_{21}|^2}, \qquad \delta = \left||m_{11}| - |m_{22}|\right|. \tag{2.9}$$

Then there exist angles ϕ, ψ and complex numbers $\alpha, \beta, \gamma, \delta, \xi, \eta, \zeta, \omega$ defining unitary transformations $\hat{U}, \hat{V}$ as in (1.5), *such that*

$$\hat{U}M\hat{V}^* = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \tag{2.10}$$

with σ_i real positive and

$$\sqrt{\sin^2\phi + \sin^2\psi} < \sqrt{\tan^2\phi + \tan^2\psi} < 2\gamma/\delta. \tag{2.11}$$

Proof. In fact, [10, Theorem 6.3] proves the bound (2.11), where ϕ and ψ are the angles over which left and right singular vectors, respectively, are rotated in the transformation (2.10). From this, one then easily constructs $\hat{U}, \hat{V}$ as in (1.5). Notice that therefore, σ_1 and σ_2 are not necessarily ordered. Each σ_i is the singular value closest to $|m_{ii}|$, and this choice is crucial for obtaining (2.11). ■

In the general cyclic method, all off diagonal pairs of elements $(a_{i_s j_s}, a_{j_s i_s})$ are annihilated successively in some order. Let us denote the values of these pairs immediately before annihilation by (x_s, y_s) and the angles of the corresponding pair of rotations by (ϕ_s, ψ_s), for $s = k, \dots, k+N-1$. If (x_s, y_s) is annihilated in step s we have, according to Lemma 1,

$$\sin^2\phi_s + \sin^2\psi_s < 4\frac{|x_s|^2 + |y_s|^2}{\delta^2}. \tag{2.12}$$

Since $[S^{(s)}]^2 - [S^{(s+1)}]^2 = |x_s^2| + |y_s|^2$, we have

$$[S^{(k)}]^2 - \sum_{s=k}^{k+N-1} \left(|x_s|^2 + |y_s|^2\right) = [S^{(k+N)}]^2 \geqslant 0, \tag{2.13}$$

and thus, from (2.12),

$$\sum_{s=k}^{k+N-1} \left(\sin^2\phi_s + \sin^2\psi_s\right) \leqslant \frac{4}{\delta^2} \sum_{s=k}^{k+N-1} \left(|x_s|^2 + |y_s|^2\right) \leqslant \frac{4[S^{(k)}]^2}{\delta^2}. \tag{2.14}$$

In order to bound $S^{(k+N)}$, we now look at the history of the pair of elements in the x_s and y_s positions subsequent to their annihilation. These elements are only affected by a subset of the later rotations, namely those involving column or row p or q, where (p,q) was the rotation annihilating the pair (x_s, y_s). Let us therefore denote the subscripts of the rotations affecting (x_s, y_s) by $s_1, \dots, s_r$, where r is a function of s. We also denote the values of the pair following the transformation involving the rotations (ϕ_{s_j}, ψ_{s_j}) by x_{s,s_j} and y_{s,s_j}, respectively.

We then have one of the following two situations:

$$x_{s,s_r} = x_{s,s_{r-1}} \alpha \cos\phi_{s_r} + a_{s_r} \beta \sin\phi_{s_r},$$

$$y_{s,s_r} = y_{s,s_{r-1}} \xi \cos\psi_{s_r} + b_{s_r} \eta \sin\psi_{s_r} \tag{2.15}$$

or

$$x_{s,s_r} = x_{s,s_{r-1}} \gamma \cos\psi_{s_r} + a_{s_r} \delta \sin\psi_{s_r},$$

$$y_{s,s_r} = y_{s,s_{r-1}} \zeta \cos\phi_{s_r} + b_{s_r} \omega \sin\phi_{s_r}, \tag{2.16}$$

where $\alpha, \beta, \gamma, \delta, \xi, \eta, \zeta, \omega$ are complex numbers of norm 1, and a_{s_r}, b_{s_r} denote elements of the matrix $A^{(s_r)}$ whose positions depend on the rotation of step s_r and on x_s, y_s. From this we find (with $|\sin\theta_{s_r}| \triangleq \max\{|\sin\phi_{s_r}|, |\sin\psi_{s_r}|\}$)

$$|x_{s,s_r}| \leqslant |x_{s,s_{r-1}}| + |a_{s_r}|\,|\sin\theta_{s_r}|, \tag{2.17}$$

$$|y_{s,s_r}| \leqslant |y_{s,s_{r-1}}| + |b_{s_r}|\,|\sin\theta_{s_r}|. \tag{2.18}$$

Using this recursively and $x_s = x_{s,s_1} = y_s = y_{s,s_1} = 0$, we obtain

$$|x_{s,s_r}| \leqslant |a_{s_1}|\,|\sin\theta_{s_1}| + |a_{s_2}|\,|\sin\theta_{s_2}| + \cdots + |a_{s_r}|\,|\sin\theta_{s_r}|, \tag{2.19}$$

$$|y_{s,s_r}| \leqslant |b_{s_1}|\,|\sin\theta_{s_1}| + |b_{s_2}|\,|\sin\theta_{s_2}| + \cdots + |b_{s_r}|\,|\sin\theta_{s_r}|. \tag{2.20}$$

Denoting $E_u^{(k)} + E_l^{(k)}$ as the matrix $E^{(k)}$ of off diagonal elements of $A^{(k)}$, we thus have that $E^{(k+N)}$ can be bounded elementwise by

$$|E^{(k+N)}| \leqslant |\sin\theta_{k+1}|\,|P_1| + |\sin\theta_{k+2}|\,|P_2| + \cdots + |\sin\theta_{k+N}|\,|P_N|, \tag{2.21}$$

where P_i are the matrices containing the appropriate elements a_{s_i} and b_{s_i} of (2.15)–(2.20). Notice that the elements of each P_i are all zero except for two columns and rows, which are those involved in the rotations through the angles ϕ_{k+i} and ψ_{k+i}. These two columns and rows contain the elements $|a_{s_r}|$ and $|b_{s_r}|$ of (2.15)–(2.20) when $s_r = k+i$. Since these are also off diagonal elements of $A^{(k+i)}$, their sum of squares is less than $[S^{(k+i)}]^2$ and thus we have

$$\| \, |P_i| \, \|_F \leqslant S^{(k+i)} \leqslant S^{(k)}. \tag{2.22}$$

From this we then obtain [because of (2.14)]

$$\begin{aligned} S^{(k+N)} &= \| E^{(k+N)} \|_F = \| \, |E^{(k+N)}| \, \|_F \\ &\leqslant S^{(k)} \left[|\sin\theta_{k+1}| + \cdots + |\sin\theta_{k+N}| \right] \\ &\leqslant S^{(k)} \left[N \sum_{i=1}^{N} \sin^2\theta_{k+i} \right]^{1/2} \\ &\leqslant S^{(k)} \left[N \sum_{k=1}^{N} \left(\sin^2\phi_{k+i} + \sin^2\psi_{k+i} \right) \right]^{1/2} \\ &\leqslant S^{(k)} \left[\frac{4N}{\delta^2} \left(S^{(k)} \right)^2 \right]^{1/2} = \frac{2\sqrt{N}}{\delta} \left[S^{(k)} \right]^2, \end{aligned} \tag{2.23}$$

which shows the quadratic convergence of the general cyclic scheme.

For the special row by row or column by column scheme, better bounds can be obtained, as was also the case for the Jacobi method [13]. The proof is here again strongly inspired by [13]. We make the same choice as there also, and illustrate the proof for a matrix of moderate size in order to simplify the notation. The considered matrix is 5×4, and only the row cyclic method is discussed, since the proof for the column cyclic method is completely analogous. Below we show the effect of annihilating the elements in the first row and column. The off diagonal elements are those we are interested in, and they are denoted with an index which is updated only when these elements are affected by the current rotations. On each matrix $A^{(k+i)}$ we also mark with arrows the columns and rows affected by the transformations U_{k+i}

and V_{k+i} yielding the next matrix $A^{(k+i+1)}$:

$$
\begin{array}{cc}
 & A^{(k)} \\
 & \begin{matrix} \downarrow & \downarrow & & \end{matrix} \\
\begin{matrix} \to \\ \to \\ \\ \\ \\ \end{matrix} &
\begin{bmatrix}
x & a_0 & b_0 & c_0 \\
d_0 & x & e_0 & f_0 \\
g_0 & h_0 & x & i_0 \\
j_0 & k_0 & l_0 & x \\
m_0 & n_0 & o_0 & p_0
\end{bmatrix}
\end{array},
\qquad
\begin{array}{cc}
 & A^{(k+1)} \\
 & \begin{matrix} \downarrow & & \downarrow & \end{matrix} \\
\begin{matrix} \to \\ \\ \to \\ \\ \\ \end{matrix} &
\begin{bmatrix}
x & 0 & b_1 & c_1 \\
0 & x & e_1 & f_1 \\
g_1 & h_1 & x & i_0 \\
j_1 & k_1 & l_0 & x \\
m_1 & n_1 & o_0 & p_0
\end{bmatrix}
\end{array},
$$

$$
\begin{array}{cc}
 & A^{(k+2)} \\
 & \begin{matrix} \downarrow & & & \downarrow \end{matrix} \\
\begin{matrix} \to \\ \\ \\ \to \\ \\ \end{matrix} &
\begin{bmatrix}
x & a_2 & 0 & c_2 \\
d_2 & x & e_2 & f_1 \\
0 & h_2 & x & i_1 \\
j_2 & k_1 & l_1 & x \\
m_2 & n_1 & o_1 & p_0
\end{bmatrix}
\end{array},
$$

$$
\begin{array}{cc}
 & A^{(k+3)} \\
\begin{matrix} \to \\ \\ \\ \\ \to \end{matrix} &
\begin{bmatrix}
x & a_3 & b_3 & 0 \\
d_3 & x & e_2 & f_2 \\
g_3 & h_2 & x & i_2 \\
0 & k_2 & l_2 & x \\
m_3 & n_1 & o_1 & p_1
\end{bmatrix}
\end{array},
\qquad
\begin{array}{c}
A^{(k+4)} \\
\begin{bmatrix}
x & a_4 & b_4 & c_4 \\
d_4 & x & e_2 & f_2 \\
g_4 & h_2 & x & i_2 \\
j_4 & k_2 & l_2 & x \\
0 & n_2 & o_2 & p_2
\end{bmatrix}
\end{array}
\tag{2.24}
$$

As was noted before, the last transformation is only a one-sided transformation because the numbers of columns and rows are unequal. Using (1.5) and reasoning similar to that of (2.17)–(2.20), we find the following inequalities for the elements of the first row and column of $A^{(k+4)}$:

$$
\begin{aligned}
|a_4| &\leqslant |h_1|\,|\sin\phi_2| + |k_1|\,|\sin\phi_3| + |n_1|\,|\sin\phi_4|, \\
|b_4| &\leqslant \qquad\qquad\qquad\;\; |l_1|\,|\sin\phi_3| + |o_1|\,|\sin\phi_4|, \\
|c_4| &\leqslant \qquad\qquad\qquad\qquad\qquad\qquad\quad |p_1|\,|\sin\phi_4|
\end{aligned}
\tag{2.25}
$$

and

$$\begin{aligned}
|d_4| &= |d_3| \leqslant |e_1|\,|\sin\psi_2| + |f_1|\,|\sin\psi_3|,\\
|g_4| &= |g_3| \leqslant \qquad\qquad\qquad\quad |i_1|\,|\sin\psi_3|,\\
|j_4| &= |j_3| = 0,\\
|m_4| &= 0.
\end{aligned} \tag{2.26}$$

Using the Cauchy-Schwartz inequality on each of the above lines and adding yields

$$|a_4|^2 + |b_4|^2 + |c_4|^2$$
$$\leqslant \left[|h_1|^2 + |k_1|^2 + |n_1|^2 + |l_1|^2 + |o_1|^2 + |p_1|^2\right] \sum_{i=2}^{4} \sin^2\phi_i \tag{2.27}$$

and

$$|d_4|^2 + |g_4|^2 + |j_4|^2 + |m_4|^2 \leqslant \left[|e_1|^2 + |f_1|^2 + |i_1|^2\right] \sum_{i=2}^{3} \sin^2\psi_i. \tag{2.28}$$

The following groups of equalities:

$$\begin{aligned}
|m_3|^2 + |n_1|^2 + |o_1|^2 + |p_1|^2 &= |m_0|^2 + |n_0|^2 + |o_0|^2 + |p_0|^2,\\
|j_2|^2 + |k_1|^2 + |l_1|^2 &= |j_0|^2 + |k_0|^2 + |l_0|^2,\\
|g_1|^2 + |h_1|^2 &= |g_0|^2 + |h_0|^2
\end{aligned} \tag{2.29}$$

and

$$\begin{aligned}
|c_2|^2 + |f_1|^2 + |i_1|^2 &= |c_0|^2 + |f_0|^2 + |i_0|^2,\\
|b_1|^2 + |e_1|^2 &= |b_0|^2 + |e_0|^2,
\end{aligned} \tag{2.30}$$

are easily obtained from (2.24) and the fact that rotations on elements of a column or row do not affect the sum of the squares of their elements. Using (2.29)–(2.30) on (2.27)–(2.28), one then finally obtains (on putting $|\sin\theta_i| =$

$\max\{|\sin\phi_i|, |\sin\psi_i|\})$

$$|a_4|^2+|b_4|^2+|c_4|^2+|d_4|^2+|g_4|^2+|j_4|^2+|m_4|^2$$

$$\leqslant \|E_l\|_F^2 \sum_{i=2}^{4} \sin^2\phi_i + \|E_u\|_F^2 \sum_{i=2}^{4} \sin^2\psi_i$$

$$\leqslant \|E\|_F^2 \sum_{i=2}^{4} \sin^2\theta_i = [S^{(k)}]^2 \sum_{i=2}^{4} \sin^2\theta_i, \tag{2.31}$$

where we have implicitly taken $\sin\psi_4 = 0$, since no column transformation is performed at that stage. The sum of squares of the elements bounded in (2.31) remains unchanged in subsequent steps. One then bounds in a similar fashion the sum of squares of the following elements in the second row and second column after step $k+7$, which completes their successive annihilation:

$$|e_5|^2+|f_5|^2+|h_5|^2+|k_5|^2+|n_5|^2$$

$$\leqslant [S^{(k+4)}]^2 \sum_{i=6}^{7} \sin^2\theta_i \leqslant [S^{(k)}]^2 \sum_{i=6}^{7} \sin^2\theta_i. \tag{2.32}$$

Here again the sum of squares on the left hand side remains constant in subsequent steps. Similarly for the third row and column we have

$$|i_6|^2+|l_6|^2+|o_6|^2 \leqslant [S^{(k+7)}]^2 \sin^2\theta_9 \leqslant [S^{(k)}]^2 \sin^2\theta_9, \tag{2.33}$$

and finally the last transformation $(k+10)$ yields $p_7 = 0$. Adding (2.31)–(2.32) and (2.33), we then have

$$[S^{(k+10)}]^2 \leqslant [S^{(k)}]^2 \sum_{i=1}^{10} \sin^2\theta_i \leqslant [S^{(k)}]^2 \left[\frac{2S^{(k)}}{\delta}\right]^2. \tag{2.34}$$

From this one derives the general inequality

$$S^{(k+N)} \leqslant \frac{2[S^{(k)}]^2}{\delta}. \tag{2.35}$$

3. THE CASE OF REPEATED OR VERY CLOSE SINGULAR VALUES

For the analysis of the quadratic convergence of the Jacobi method in the presence of repeated or clustered eigenvalues, one relies on a lemma of the following type [14]:

LEMMA 2. *Let the symmetric matrix A be decomposed as*

$$A = D + E \tag{3.1}$$

with D diagonal and $\|E\|_F = \varepsilon < \delta/2$, and let all eigenvalues of A be separated at least by 2δ except for a cluster $\lambda_1, \dots, \lambda_k$ of pathologically close (or repeated) eigenvalues:

$$|\lambda_i - \lambda_j| < \eta \ll \delta \qquad \text{for} \quad i, j \leqslant k. \tag{3.2}$$

Then after a suitable symmetric permutation of rows and columns one has

$$|\lambda_i - a_{ii}| < \delta, \tag{3.3}$$

and the off diagonal elements E_k of the leading $k \times k$ principal submatrix A_k of A are bounded by

$$\|E_k\|_F < \frac{\varepsilon^2}{\delta} + \eta. \qquad \blacksquare$$

An analogous theorem for the singular values of a nearly diagonal matrix A does *not* hold, as is shown by the following example:

$$A = \begin{bmatrix} -1 & \frac{1}{2}\varepsilon & 0 \\ \frac{1}{2}\varepsilon & 1+\frac{1}{2}\varepsilon^2 & -\frac{1}{2}\varepsilon \\ 0 & -\frac{1}{2}\varepsilon & 2 \end{bmatrix} \tag{3.4}$$

with singular values approximately $2 + \frac{1}{4}\varepsilon^2$, $1 + \frac{3}{8}\varepsilon^2$, $1 + \frac{1}{8}\varepsilon^2$, $\delta = \frac{1}{2}$, and $\eta \approx \frac{1}{4}\varepsilon^2$. Although there is a cluster of two close singular values around 1, (approximated by the diagonal elements $|a_{11}|$ and $|a_{22}|$), the off diagonal elements a_{12} and a_{21} are clearly of the order of ε and not ε^2/δ. However, when triangularizing this matrix by a left unitary transformation Q^*:

$$Q^*A=\begin{bmatrix} 1+\frac{1}{8}\varepsilon^2 & \frac{1}{4}\varepsilon^3 & -\frac{1}{4}\varepsilon^2 \\ 0 & 1+\frac{3}{4}\varepsilon^2 & -\frac{3}{2}\varepsilon+\frac{7}{8}\varepsilon^3 \\ 0 & 0 & 2-\frac{1}{2}\varepsilon^2 \end{bmatrix}+O(\varepsilon^4), \tag{3.5}$$

suddenly the element a_{12} becomes very small. This first step of the *triangular* version of the Kogbetliantz algorithm [8] was in fact always observed to yield indeed ε^2/δ off diagonal elements in the right places for several random matrices with pathologically close singular values. It is for this reason that we conjecture such a property to hold for *triangular* matrices. Notice that this does *not* imply that the triangular version of the algorithm has better convergence properties in the presence of pathologically close or multiple singular values: the conjectured lemma is indeed only a tool to derive appropriate error bounds for proving quadratic convergence. The obtained bounds are usually serious overestimates and do not always reflect the true behavior of the algorithm. This is particularly the case for the above example (3.4), since it is symmetric and Kogbetliantz's algorithm then becomes Jacobi's algorithm, which is known to converge quadratically (the eigenvalues are even well separated).

REFERENCES

1 P. Brent, F. Luk, and C. Van Loan, Computation of the singular value decomposition using mesh-connected processors, Tech. Rep. CS-528, Dept. of Computer Science, Cornell Univ., Ithaca, N.Y., 1983.
2 G. Forsythe and P. Henrici, The cyclic Jacobi method for computing the principal values of a complex matrix, Trans. Amer. Math. Soc. 94:1–23 (1960).
3 M. Heath, A. Laub, C. C. Paige, and R. Ward, Computing the SVD of a product of matrices, in preparation.
4 P. Henrici, On the speed of convergence of cyclic and quasi cyclic Jacobi methods for computing eigenvalues of Hermitian matrices, *J. SIAM* 6:144–162 (1958).
5 E. Kogbetliantz, Diagonalization of general complex matrices as a new method for solution of linear equations, in *Proceedings of the International Congress on Mathematics*, Amsterdam, Vol. 2, 1954, pp. 356–357.
6 E. Kogbetliantz, Solution of linear equations by diagonalization of coefficient matrices, *Quart. Appl. Math.* 13:123–132 (1955).
7 C. Lawson and R. Hanson, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, N.J., 1974.

8 C. C. Paige, Computing the generalized singular value decomposition, *SIAM J. Sci. Statist. Comput.* submitted for publication.

9 A. Schönhage, Zur Konvergenz des Jacobi-Verfahrens, *Numer. Math.* 3:374–380 (1961).

10 G. W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.* 15:727–764 (1973).

11 H. van Kempen, On the convergence of the classical Jacobi method for real symmetric matrices with non-distinct eigenvalues, *Numer. Math.* 9:11–18 (1966).

12 H. van Kempen, On the quadratic convergence of the special Jacobi method, *Numer. Math.* 9:19–22 (1966).

13 J. H. Wilkinson, Note on the quadratic convergence of the cyclic Jacobi process, *Numer. Math.* 4:296–300 (1962).

14 J. H. Wilkinson, Almost diagonal matrices with multiple or close eigenvalues, *Linear Algebra Appl.* 1:1–12 (1968).

Received 30 November 1984; revised 22 July 1985