Quadratic Convergence of a Special Quasi-cyclic Jacobi Method

Vjeran Hari*

Abstract

The paper considers the ultimate asymptotic convergence of a block-oriented, quasi-cyclic Jacobi method for symmetric matrices. The conclusion applies to the new one-sided Jacobi method for computing the singular value decomposition, recently proposed by Drmač and Veselić. Using the simple qualitative analysis, the discussion indicates that the quadratic off-norm reduction per quasi-sweep is to be expected in all perceivable cases.

1 Introduction

In [6, 7] Drmač and Veselić have proposed a greatly improved modification of the one-sided Jacobi method for computing the singular value decomposition (SVD) of rectangular matrices. Their modification is relatively accurate and almost as fast as Divide and Conquer method, which is considered the fastest, but not relatively accurate method. Their method prepares the initial matrix for the iteration by applying to it one or two QR factorizations (cf. [24, 4, 6]) and in the iterative part, uses many ideas from [3]. To better exploit the fast (cache) memory, it uses a new block-oriented, quasi-cyclic pivot strategy. In [16] the global convergence of the modified method has been proved. What is still lacking, is the asymptotic convergence consideration.

Note that the global and the quadratic convergence of a one-sided SVD Jacobi method for the matrix G actually means the global convergence of the corresponding two-sided Jacobi method for the non-negative definite Gram matrix $G^T G$. For this reason, we consider here the two-sided Jacobi method for general symmetric matrices, under the special quasi-cyclic pivot strategy which has been used in [7].

It is known that in the case of simple eigenvalues, the cyclic Jacobi methods are quadratically convergent [25]. Under certain constraints [13], the same is true for the case of multiple eigenvalues. Some quasi-cyclic Jacobi methods are even cubically convergent per quasi-sweep [18, 22]). However, in practical computation, the cubic convergence show itself very late, often in the moment when the process has to be is stopped. Therefore, for the speed of the Jacobi method, the quadratic reduction per quasi-sweep, which starts sooner, is more important.

The aim of this paper is to show that under usual assumptions, the method from [7] should be quadratically convergent per quasi-sweep. The difficulties appear from the fact that we actually work with a class of methods. Each method if defined by some matrix block-partition, which can be almost arbitrary. The second block-partition which influences the asymptotic convergence is the inherent, natural partition of almost diagonal matrices. This one depends on the multiplicities of the eigenvalues and on affiliation (i.e. on the ordering) of the diagonal elements. Because of this interplay, for each ordered pair of these two partitions, a different asymptotic convergence analysis should be made. Note, however that the proper proofs are very complicated [13].

A way how to avoid these obstacles is to apply the qualitative analysis which uses $\mathcal{O}(\epsilon)$ symbols. But, even with this simplification, it is almost impossible to encompass all the possibilities. Therefore, we have chosen

*Department of Mathematics, University of Zagreb, Croatia

several characteristic examples of block-partitions and we show by the qualitative analysis that the quadratic convergence is always present.

The paper is divided into four sections. In Section 2, we introduce notation, the quasi-cyclic strategy and define the method. In Section 3, some important results on almost diagonal symmetric matrices are presented. Finally, in Section 4, we discuss for each case, the asymptotic convergence of the considered method.

2 The Quasi-cyclic Jacobi Method $J_{\mathcal{M}}$

Although, the method and the analysis can be considered for complex Hermitian matrices, for simplicity we restrict our consideration to real symmetric matrices.

For any quadratic matrix $X = (x_{ij})$, the function $Off(X) = \|X - \text{diag}\,(X)\|_F$ is referred to as *departure from the diagonal form* of X. Here, $\|\cdot\|_F$ is the Frobenius norm and diag (X) is the diagonal part of X. Then, for any symmetric A, a useful measure of almost diagonallity and of advancing of Jacobi methods, is

$$S(A) = Off(A)/\sqrt{2} = \sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (a_{ij})^2}.$$

2.1 The Quasi-Cyclic Jacobi Method

Here we introduce the two-sided Jacobi method for symmetric matrices, which will be analyzed (cf. [16]). Let A be a symmetric matrix of order n. The method performs the sequence of similarity transformations

$$A^{(k+1)} = [R^{(k)}]^T A^{(k)} R^{(k)}, \quad V^{(k+1)} = V^{(k)} R^{(k)}, \quad k \geq 0\,, \tag{2.1}$$

where $A^{(0)} = A$, $V^{(0)} = I_n$ and $R^{(0)}$, $R^{(1)}, \ldots$ are plane rotations. For each k, the rotation $R^{(k)}$ is defined by a pair of indices $(p, q) = (p(k), q(k))$ called *pivot pair* and by four essential elements $R_{pp}^{(k)} = R_{qq}^{(k)} = \cos\phi_k$, $R_{pq}^{(k)} = -R_{qp}^{(k)} = \sin\phi_k$, which constitute the 2×2 *pivot submatrix* of $R^{(k)}$. All other elements of $R^{(k)}$ are as in the identity matrix I_n. The process (2.1) is defined by a rule for computing the elements of the pivot submatrix of $R^{(k)}$ and by a way of selecting pivot pairs (*pivot strategy*). We assume the usual angle choice

$$\tan 2\phi_k = \frac{2a_{pq}^{(k)}}{a_{qq}^{(k)} - a_{pp}^{(k)}}, \quad \phi_k \in \left[\frac{-\pi}{4}, \frac{\pi}{4}\right], \tag{2.2}$$

which makes the *pivot element* $a_{pq}^{(k)}$ zero, i.e. $a_{p(k)q(k)}^{(k+1)} = 0, \ \ k \geq 0$ holds, where $A^{(k)} = (a_{ij}^{(k)})$. Consequently, we have

$$S^2(A^{(k+1)}) = S^2(A^{(k)}) \ - \ (a_{p(k)q(k)}^{(k)})^2, \ \ k \geq 0$$

and $\lim_{k\to\infty} a_{p(k)q(k)}^{(k)} = 0$. If A is 2×2 the process is completed for $k = 1$, since $S(A^{(1)}) = 0$. So, we assume $n \geq 3$.

To introduce the quasi-cyclic strategy from [7], let $N = n(n-1)/2$, $\mathbf{N}_0 = \{0, 1, 2, \ldots\}$ and $\mathbf{P}_n = \{(i, j) : 1 \leq i < j \leq n\}$. Each pivot strategy can be identified with a function $\mathcal{I} \ : \ \mathbf{N}_0 \mapsto \mathbf{P}_n$, defined by $\mathcal{I}(k) = (p(k), q(k)), \ \ k \geq 0$. If $\mathcal{I}$ is periodic, then $\mathcal{I}$ is called *periodic strategy*.

Let $\mathcal{I}$ be a periodic strategy with period M. If $\{\mathcal{I}(k) : 0 \leq k \leq M-1\} = \mathbf{P}_n$ and $M > N$ ($M = N$), then $\mathcal{I}$ is called *quasi-cyclic* (*cyclic*) *strategy*. For $\mathbf{S} \subseteq \mathbf{P}_n$, let $\nu(\mathbf{S})$ denote the number of elements of $\mathbf{S}$. With $\mathbf{O}(\mathbf{S})$ we denote the collection of all finite sequences made of the elements of $\mathbf{S}$. We assume that each $O \in \mathbf{O}(\mathbf{S})$ contains all elements from $\mathbf{S}$; some may appear more than once in O. So, each sequence from $\mathbf{O}(\mathbf{S})$ contains at least $\nu(\mathbf{S})$ terms. A cyclic or a quasi-cyclic strategy can be specified in the following way. For any $O = \{(i_r, j_r)\}_{r=0}^{M-1} \in \mathbf{O}(\mathbf{P}_n)$, the cyclic or the quasi-cyclic strategy $\mathcal{I}_O$, generated by O, is given by

$$\mathcal{I}_O(rM + k) \equiv (p(rM + k), q(rM + k)) = (i_k, j_k), \quad 0 \leq k \leq M - 1, \quad r \geq 0.$$

Let us denote the Jacobi method, which defines ϕ_k by (2.2) and uses the pivot strategy $\mathcal{I}$, by $J(\mathcal{I})$. $J(\mathcal{I})$ is called quasi-cyclic (cyclic) Jacobi method if $\mathcal{I}$ is the quasi-cyclic (cyclic) strategy.

In [18] Mascarenhas has proved that the diagonal elements of the iterated symmetric matrix $A^{(k)}$, obtained by the Jacobi method under any periodic pivot strategy, converge. Hence, the sequence of iterates $A^{(k)}$ is convergent iff $S(A^{(k)}) \to 0$ as $\to \infty$ holds for any initial symmetric matrix A. Quadratic convergence per quasi-sweep means that the sequence $S(A^{(rM)})$, $S(A^{((r+1)M)})$, $S(A^{((+2)rM)})$, ... converges with quadratic rate, provided that r is large enough (or more generally, provided $S(A^{(rM)})$ is small enough).

Let $\mathbf{S}$ be any subset of $\mathbf{P}_n$. By $O_R(\mathbf{S})$ we denote the "row-wise ordering" of $\mathbf{S}$, that is the sequence satisfying the following two conditions: (i) each element $(i, j) \in \mathbf{S}$ appears exactly once in $O_R(\mathbf{S})$; (ii) for any two terms (i_1, j_1) and (i_2, j_2) in $O_R(\mathbf{S})$, (i_1, j_1) precedes (i_2, j_2) if $i_1 < i_2$ or $i_1 = i_2$ and $j_1 < j_2$. In an obvious manner we can define the "column-wise" ordering of $\mathbf{S}$ denoted by $O_C(\mathbf{S})$.

Let $\mathbf{S}_i$, $1 \leq i \leq \sigma$, be the subsets of $\mathbf{P}_n$ and let $O_i \in \mathbf{O}(\mathbf{S}_i)$, $1 \leq i \leq \sigma$ be arbitrary sequences. By $[O_1, O_2, \ldots, O_\sigma]$, we mean the sequence which is obtained by the concatenation of the sequences O_1,O_2, ..., O_σ. Often, the brackets are omitted.

2.2 The Special Quasi-cyclic Method $J_\mathcal{M}$

Let $A = (a_{rs})$ be a real symmetric matrix of order n and let

$$A = \left[\begin{array}{ccc} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{array} \right] \begin{array}{l} \} n_1 \\ \vdots \\ \} n_m \end{array} ,$$

be any block-partition of A. Here A_{ij} is $n_i \times n_j$ block of A. Note that $\mathcal{M} = (n_1, \ldots, n_m)$ is the appropriate partition of n.

With each submatrix B contained in the upper–triangle of A (or with each principal submatrix B of A), one can associate the set $\mathbf{S}(B) \subseteq \mathbf{P}$ of those pairs which are subscripts of the elements of B (in the upper triangle of B). By abuse of notation, the row- and the column- wise orderings of $\mathbf{S}(B)$ are denoted by $O_R(B)$ ($= O_R(\mathbf{S}(B))$) and $O_C(B)$ ($= O_C(\mathbf{S}(B))$), respectively. We shall also use $\mathbf{S}_{ij} = \mathbf{S}(A_{ij})$ and $\mathcal{R}_{ij} = O_R(A_{ij})$ ($= O_R(\mathbf{S}_{ij})$), $\mathcal{C}_{ij} = O_C(A_{ij})$ ($= O_C(\mathbf{S}_{ij})$) for $i \leq j$.

Let $\mathcal{O}_\mathcal{M}$ denote the sequence of pairs which defines one quasi-sweep of the method from [7]. It is defined by (cf. [16])

$$\mathcal{O}_\mathcal{M} = [\mathcal{R}_{11}, \mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_{m-1}, \mathcal{R}_{mm}]. \tag{2.3}$$

where

$$\mathcal{R}_i = [\mathcal{R}_{i+1,i+1}, \mathcal{R}_{ii}, \mathcal{R}_{i,i+1}, \mathcal{R}_{i,i+2}, \ldots, \mathcal{R}_{im}]. \quad 1 \leq i \leq m - 1.$$

The quasi-cyclic Jacobi method from [7], defined by $\mathcal{I}_\mathcal{M} = \mathcal{I}_{\mathcal{O}_\mathcal{M}}$, is denoted $J_\mathcal{M}$. Its pivot strategy $\mathcal{I}_{\mathcal{O}_\mathcal{M}}$ uses the row-wise ordering within each block A_{ij}. $\mathcal{I}_{\mathcal{O}_\mathcal{M}}$ fetches the blocks in the row-wise fashion, and all the diagonal blocks A_{ii} are operated twice. Therefore, we call it (cf. [16]), the *special quasi-cyclic block-oriented method*. The quasi-sweep contains M ordinary Jacobi steps, where

$$M = N + \sum_{i=1}^{m} n_i(n_i-1)/2\,.$$

In [16] it has been shown that $\mathcal{O}_{\mathcal{M}}$ is equivalent to

$$\begin{aligned}\mathcal{O}'_{\mathcal{M}} \;=\; & [\mathcal{C}_{11},\mathcal{C}_{11},\ \mathcal{C}_{22},\mathcal{C}_{12},\mathcal{C}_{22},\ \mathcal{C}_{13},\mathcal{C}_{33},\mathcal{C}_{23},\mathcal{C}_{33},\ldots\mathcal{C}_{1j},\mathcal{C}_{2j},\ldots,\mathcal{C}_{j-2,j},\mathcal{C}_{jj},\mathcal{C}_{j-1,j},\mathcal{C}_{jj},\ldots \\ & \mathcal{C}_{1m},\mathcal{C}_{2m},\ldots,\mathcal{C}_{m-2,m},\mathcal{C}_{mm},\mathcal{C}_{m-1,m},\mathcal{C}_{mm}].\end{aligned} \tag{2.4}$$

This means that the both methods $J_{\mathcal{M}}$ and $J'_{\mathcal{M}} = J(\mathcal{I}_{\mathcal{O}'_{\mathcal{M}}})$, when applied to A, yield the same matrix $A^{(M)}$ after one quasi-sweep. Note that most of the intermediate matrices (iterates for $1 \le k \le M-1$) determined by $J_{\mathcal{M}}$ and $J'_{\mathcal{M}}$ are different. Since we consider the quadratic asymptotic convergence, which uses only the matrices $A^{(rM)}$, $r \ge r_0$ for some r_0, we can replace $J_{\mathcal{M}}$ by $J'_{\mathcal{M}}$.

3 Almost Diagonal Symmetric Matrices

Let $A = (a_{ij})$ be a symmetric matrix of order n and let its eigenvalues be ordered non-increasingly, so that

$$\lambda_1 = \cdots = \lambda_{\sigma_1} > \lambda_{\sigma_1+1} = \cdots = \lambda_{\sigma_2} > \cdots > \lambda_{\sigma_{p-1}+1} = \cdots = \lambda_{\sigma_p} \tag{3.1}$$

holds. Here λ_{σ_1}, λ_{σ_2},…,λ_{σ_p} are all mutually different eigenvalues. Then for each k, $\nu_k = \sigma_k - \sigma_{k-1}$ is the multiplicity of λ_{σ_k}. Here $\sigma_0 = 0$. The absolute gap (or separation) of λ_{σ_i} from the other eigenvalues δ_i, and the minimum absolute gap in the spectrum δ are defined as follows

$$\delta_i = \min_{\substack{1\le j\le p\\ j\ne i}} \mid \lambda_{\sigma_i} - \lambda_{\sigma_j} \mid\,, \qquad \delta = \min_{1\le i\le p} \delta_i.$$

We assume that A is almost diagonal i.e. that all of its off-diagonal elements are small compared to δ, i.e. $S(A) \ll \delta$. Then it is known (see [26, 12, 13]) that A has a special structure. To simplify notation, let us denote the value of $S(A)$ by ε. This structure reveals several properties of A. First, all diagonal elements are ε^2 good approximations of the corresponding eigenvalues. Because of (3.1), for each k there exist exactly ν_k diagonal elements which are ε^2 close to λ_{σ_k}. Second, the off-diagonal elements which link those diagonals which approximate the same eigenvalue are ε^2 small.

To make the exposition easier, we assume that the diagonal elements of A are ordered the same way as are the eigenvalues, i.e. non-increasingly:

$$a_{11} \ge a_{22} \ge \cdots \ge a_{nn}\,. \tag{3.2}$$

A bit less stringent requirement, but sufficient for all further estimates and conclusions would read: *the diagonals which are affiliated with the same eigenvalue take successive positions on the diagonal*.

We partition A according to $\mathcal{N} = (\nu_1,\nu_2,\ldots,\nu_p)$ and denote the corresponding matrix block-partition by $(\mathcal{A}_{ij})$. We call $\mathcal{N}$ and $(\mathcal{A}_{ij})$ *natural* or *inherent* partitions. The following result provides the best known estimates concerning the structure of almost diagonal matrices

Theorem 3.1 *[13] Let A be a symmetric matrix of order n satisfying (3.1), (3.2) and $S(A) \le (\sqrt{2}/6)\delta$. Then*

$$S(\mathcal{A}_{ii}) \le \frac{\sqrt{2}}{2}\|\mathcal{A}_{ii} - \lambda_{\sigma_i} I_{\nu_i}\|_F \le \frac{0.66\sqrt{2}}{\delta_i}\sum_{\substack{j=1\\ j\ne i}}^{p}\|\mathcal{A}_{ij}\|_F^2 \le \frac{S^2(A)}{\delta_i},\ 1 \le i \le p.$$

Another, important and similar result deals with *scaled almost diagonal* symmetric matrices. Let A be a nonsingular almost diagonal matrix with non-zero diagonal elements. Then

$$A_S = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}, \qquad D = \text{diag}\,(|a_{11}|, \dots, |a_{nn}|)\,,$$

is called *scaled* A. If $\|A_S\| \leq \alpha < 1$, then A is called α-*scaled diagonally dominant* with respect to the considered norm, or shorter α-s.d.d. In this context we presume the spectral matrix norm, $\|A\|_2 = \max\{|\lambda_{\sigma_1}|, |\lambda_{\sigma_p}|\}$. Note that scaled matrices are dimension-free. Therefore the corresponding estimates should use dimension-free gaps in the spectrum. Typically, such are the relative gaps

$$\gamma_i = \min_{\substack{1\leq j\leq p\\ j\neq i}} \frac{|\,\lambda_{\sigma_i} - \lambda_{\sigma_j}\,|}{|\lambda_{\sigma_i}| + |\lambda_{\sigma_j}|}\,, \quad 1 \leq i \leq p\,, \qquad \gamma = \min_{1\leq i\leq p} \gamma_i\,.$$

Note that $0 \leq \gamma_i \leq 1$ for $1 \leq i \leq p$ and $\gamma_i = 1$ if either $\lambda_{\sigma_i} = 0$ or λ_{σ_i} is a single spectral point in $(-\infty, 0)$ or in $(0, \infty)$. The natural partition of A_S is denoted by $(\mathcal{A}_{ij}^S)$.

Theorem 3.2 *[14] Let A be a nonsingular symmetric matrix of order n satisfying (3.1) and (3.2). If $\alpha < \gamma/(\gamma+3)$ and A is α-s.d.d., then*

$$\sum_{j=\sigma_{i-1}+1}^{\sigma_i} \left|1 - \frac{\lambda_{\sigma_i}}{|a_{jj}|}\right|^2 + 2S^2(\mathcal{A}_{ii}^S) \leq \left[\frac{4}{\gamma_i} \sum_{\substack{j=1\\ j\neq i}}^{p} \|\mathcal{A}_{ij}^S\|_F^2\right]^2 \leq \left[\frac{2}{\gamma_i} S^2(A_S)\right]^2 .$$

holds for $1 \leq i \leq p$. In addition, if A is positive definite, then the constants 4 and 2 in the numerators can be replaced by 2 and 1, respectively.

We see that $S(A)$ and $S(A_S)$ measure the absolute and the relative distance between the diagonal elements and the eigenvalues, respectively. So, they both can be used for stopping the process and for the definition of the quadratic asymptotic convergence.

4 Asymptotic behavior of $J_{\mathcal{M}}$

Here we briefly discuss the behavior of the method $J_{\mathcal{M}}$ on almost diagonal symmetric matrices. As we have mentioned earlier, for this purpose, we can replace the method $J_{\mathcal{M}}$ by the method $J'_{\mathcal{M}}$. This means replacing the strategy $\mathcal{I}_{\mathcal{O}_{\mathcal{M}}}$ by $\mathcal{I}_{\mathcal{O}'_{\mathcal{M}}}$.

In the case of simple eigenvalues every cyclic and especially every quasi-cyclic Jacobi method is quadratically (at least quadratically) convergent. Therefore our main interest is to investigate the case of non-trivial partitions $\mathcal{M}$ and $\mathcal{N}$.

The classic definition of the quadratic convergence of $J_{\mathcal{M}}$ means that $S(A^{(M)}) \leq c\,S^2(A)$ whenever $S(A)$ is sufficiently small. The constant c may depend on n or the eigenvalues of A, but not on $S(A)$. Usually (see [25, 13], c has the form $const/\delta$.

Instead of $S(A)$, one can consider the quadratic convergence of $S(A_S)$ (see [19, 20]), since this measure sometimes better determines when to stop the process. Since $S(A_S)$ is dimension-free, the constant c has the form $const/\gamma$.

Here we assume that the Jacobi method $J_{\mathcal{M}}$ has reached the stage when

$$S(A^{(k)}) \ll \delta \quad \text{or} \quad S(A_S^{(k)}) \ll \gamma\,, \tag{4.3}$$

where $A_S^{(k)}$ stands for the scaled $A^{(k)}$. To simplify notation, we can assume that the initial A already satisfies (4.3).

We shall consider what happens with $S(A^{(rM)})$ for several consecutive values of r. The estimates in [20] are pretty analogous to those in [13], but more complicated, so we shall assume that the first condition in (4.3) holds. Another reason for using the first condition in [13] lies in the fact that we use the qualitative analysis, so working with $A^{(rM)}$ and $A_S^{(rM)}$ would yield the same analysis and conclusion. The second condition in (4.3) is added just to include the case $\delta \ll \gamma$, which can happen when there is a cluster of eigenvalues around the origin. In that case the estimates from [20] have to be used.

Then we can write $\delta = \mathcal{O}(1)$, $a_{ij}^{(k)} = \mathcal{O}(\varepsilon)$. In [13], it has been shown that $a_{ii}^{(k)} - a_{ii}^{(k)} > (2.56/3)\delta$, $i < j$ holds whenever $a_{ii}^{(k)}$ and $a_{jj}^{(k)}$ are not affiliated with the same eigenvalue. In that case, we also have [13, relation (2.7)], $|\sin\varphi_k| \leq |\tan\varphi_k| \leq 1.2|a_{p(k)q(k)}^{(k)}|/\delta$. We need these results just to conclude that $\varphi_k = \mathcal{O}(\varepsilon) = \sin\varphi_k$ and $\cos\varphi_k = 1 - x_k$, $x_k = \mathcal{O}(\varepsilon^2)$.

We shall use the technique from ([25, 26]) which estimates how much an element can grow after its annihilation. Since our strategy is block-oriented, we shall consider how much can the spectral norm of each block (or the largest element in each block) grow after all of its elements have once been annihilated.

Simple Eigenvalues

In this case $\mathcal{N} = (1, 1, \ldots, 1)$, so that $\lambda_{\sigma_i} = \lambda_i$ for $1 \leq i \leq n$. In Table 1 are displayed the data for a matrix of order 16 partitioned as indicated in Figure 1. We have assumed simple eigenvalues, and have applied the Jacobi method $J'_{\mathcal{M}}$ to a matrix whose all off-diagonal elements are of order ε.

n_1			n_2			n_3		n_4	n_5				n_6		
■	4	5	10	13	16	22	25	36	44	47	50	53	92	95	98
1	■	6	11	14	17	23	26	37	45	48	51	54	93	96	99
2	3	■	12	15	18	24	27	38	46	49	52	55	94	97	100
			■	19	20	29	32	39	56	59	62	65	101	104	107
			7	■	21	30	33	40	57	60	63	66	102	105	108
			8	9	■	31	34	41	58	61	64	67	103	106	109
						■	35	42	68	70	72	74	110	112	114
						28	■	43	69	71	73	75	111	113	115
								■	82	83	84	85	116	117	118
									■	86	87	89	122	126	130
									76	■	88	90	123	127	131
									77	78	■	91	124	128	132
									79	80	81	■	125	129	133
													■	134	135
													119	■	136
													120	121	■

Figure 1: The annihilation ordering of $J'_{\mathcal{M}}$

In Figure 1 we have depicted the matrix A for the case $n = 16$, together with its basic partition (A_{ij}), which defines the quasi-cyclic pivot strategy. We have chosen

$$\mathcal{M} = (3, 3, 1, 4, 3).$$

But instead of displaying the matrix A or $A^{(k)}$, we have displayed at the (i, j) position the value of $k + 1$, where k is such that $(i, j) = (p(k), q(k))$. Remember that $a_{p(k)q(k)}^{(k+1)} = 0$. So Figure 1 describes the pivot strategy. Since some matrix elements are annihilated twice during the quasi-sweep, we have also used the lower-triangle of diagonal blocks.

The annihilation ordering depicted in Figure 1, $\mathcal{O}'_{\mathcal{M}}$ is defined in (2.4). It is equivalent to $\mathcal{O}_{\mathcal{M}}$ from the relation (2.3.

To see what happens with the magnitudes of matrix elements during the first quasi-cycle we display a table. Each column of Table 1 corresponds to one block of the matrix. The subscripts of that block are written on the top of that column. Each row of the table corresponds to a stage of the process.

	A_{ij}																			
	11	12	13	14	15	16	22	23	24	25	26	33	34	35	36	45	46	55	56	66
A_{11}^*	8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A_{22}	8	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A_{12}	2	3	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A_{22}	2	3	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1
A_{13}	2	2	2	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1
A_{33}	2	2	2	1	1	1	4	1	1	1	1	∞	1	1	1	1	1	1	1	1
A_{23}	2	2	2	1	1	1	2	3	1	1	1	2	1	1	1	1	1	1	1	1
A_{33}	2	2	2	1	1	1	2	3	1	1	1	∞	1	1	1	1	1	1	1	1
A_{14}	2	2	2	3	1	1	2	3	1	1	1	∞	1	1	1	1	1	1	1	1
A_{24}	2	2	2	3	1	1	2	2	3	1	1	∞	1	1	1	1	1	1	1	1
A_{34}	2	2	2	3	1	1	2	2	3	1	1	2	3	1	1	1	1	1	1	1
A_{15}	2	2	2	2	2	1	2	2	3	1	1	2	3	1	1	1	1	1	1	1
A_{25}	2	2	2	2	2	1	2	2	2	2	1	2	3	1	1	1	1	1	1	1
A_{35}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	1	1	1	1	1
A_{55}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	1	1	2	1	1
A_{45}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	3	1	2	1	1
A_{55}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	3	1	4	1	1
A_{16}	2	2	2	2	2	2	2	2	2	2	1	2	2	2	1	3	1	4	1	1
A_{26}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	3	1	4	1	1
A_{36}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	1	4	1	1
A_{46}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	4	1	1
A_{66}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	4	1	2
A_{56}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2
A_{66}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	4

Table 1: The block norms: entries denote exponents of ε

The first row corresponds to the stage, just after all off-diagonal elements of A_{11} have *twice* been annihilated (rotated). The second row corresponds to the stage, just after all off-diagonal elements of A_{22} have been rotated for the *first time*. The third row corresponds to the stage, just after the elements of A_{12} have been rotated. The forth row corresponds to the stage, just after all off-diagonal elements of A_{22} have been rotated for the *second time*. And so on. The blocks that have just been operated upon are displayed at the left edge of that row. The entries of the table are the powers of ε. For example, if ε^2 is the largest possible order of magnitude of the norm of the (i_2, j_2)-block of the iterated matrix at stage, just after the annihilations in the (i_1, j_1)-block have been completed, then the number 2 is the table entry on the intersection of the row starting with A_{i_1,j_1} and the column whose top is designated by $i_2 j_2$. If the whole block becomes zero we write ∞ as the table entry.

The Table 1 is obtained from the output to the MATLAB code, whose essential part (the transformation which rotates `a(i,j)`) has the form

```
t=a(i,j); x=a(:,j);
a(:,j)=min(x,a(:,i)+t*z);
a(:,i)=min(a(:,i),x+t*z);
a(i,j)=aij;
a(j,:)=a(:,j)';
a(i,:)=a(:,i)';
a(i,i)=0;a(j,j)=0;
```

Here `z=ones(n,1)` is the vector of ones (defined earlier) and `aij` is a variable whose value `9*10^(ic)` is set at the beginning of each quasi-sweep (`ic` counts the quasi-sweeps). It serves as large enough power of ε to represent zero (annihilated element). Diagonals are represenred by 0. Here `t` has the role of $\sin \varphi_k$. The rest of the program just selects the pivot indices (which are in MATLAB programs denoted by `i` and `j`) according to the quasi-cyclic strategy and calls to the M-file containing the described code.

As a brief comment, we notice the following important facts. The process advances by block-columns, within each block-column by blocks and within each block the elements are rotated column-wise. After the element at position (i, j) has been rotated, it first becomes zero, then it becomes ε^3 small and so small it remains as long as the elements in the same column, within the same block, are rotated. This must be so because the contributions to its magnitude have the form: $\sin\varphi_k a_{it}^{(k)} = \mathcal{O}(\varepsilon)\cdot\mathcal{O}(\varepsilon^2)$. Here $a_{it}^{(k)} = \mathcal{O}(\varepsilon^2)$ because this one has already been annihilated. After that, when the elements in the next column, within the same block are rotated, the contribution to its magnitude becomes $\sin\varphi_k a_{jt}^{(k)} = \mathcal{O}(\varepsilon)\cdot\mathcal{O}(\varepsilon)$, since the element $a_{jt}^{(k)}$ lies in the diagonal block and has not yet been annihilated. This conclusion does not hold only for the elements of the block next to the diagonal block, due to the pivot strategy. But the norms of this one and the diagonal block will eventually become ε^3 and will remain so until the transformations in the next block-column change them to become $\mathcal{O}(\varepsilon^2)$.

In Table 2 the intermediate steps are omitted and the data are displayed after every quasi-sweep.

	A_{ij}																			
	11	12	13	14	15	16	22	23	24	25	26	33	34	35	36	45	46	55	56	66
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	4
2	4	4	4	4	4	6	4	4	4	4	6	4	4	4	6	5	6	6	9	12
3	8	8	8	9	10	14	8	8	9	10	14	8	9	10	14	15	20	18	27	36
4	16	16	16	25	26	30	16	18	25	26	30	18	25	28	33	44	49	48	72	98
5	32	34	36	57	58	62	36	50	61	62	66	50	72	76	95	105	113	112	168	226
6	68	86	100	125	126	130	100	133	161	164	168	144	181	190	220	233	241	240	360	482

Table 2: Block norms after each of the first 6 quasi-sweeps

We can notice, the larger the sum of the block's subscripts, the faster is its convergence to zero. This effect can be noticed in the real process.

Multiple Eigenvalues

Here we use the same analysis as above, but additionally assume $\mathcal{N} = (1,1,3,6,4,1)$.

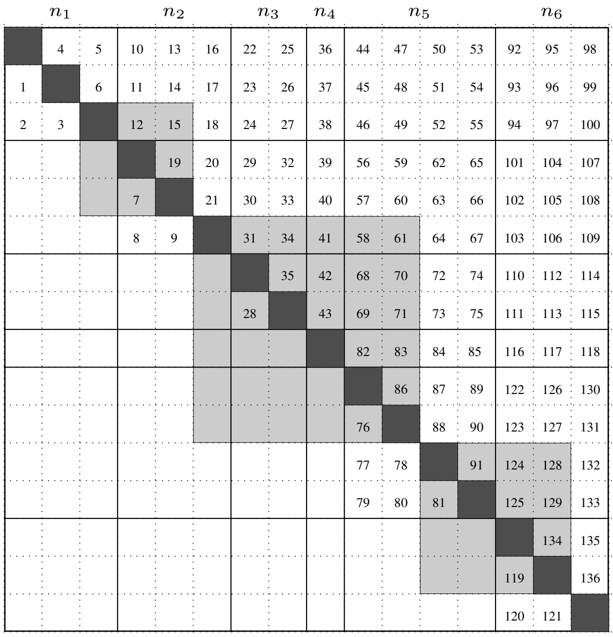

Figure 2: $\mathcal{N} = (1,1,3,6,4,1)$, $\mathcal{M} = (3,3,2,1,4,3)$.

In Figure 2 we have depicted the matrix A together with its basic partition (A_{ij}), which defines the quasi-cyclic pivot strategy (same as above) and with its natural partition $(\mathcal{A}_{pq})$ defined by the multiplicities of its eigenvalues. The partition $(\mathcal{A}_{pq})$ is indicated by the diagonal blocks colored grey.

The method $J'_{\mathcal{M}}$ is applied to the almost diagonal symmetric matrix A which satisfies the conditions (4.3) and whose off-diagonal elements are of order ε. Note that A and all the subsequent matrix iterates satisfy the estimates from Theorem 3.1 and Theorem 3.2. Namely, the affiliation of the diagonal elements cannot be changed during the process (see [13, 20]).

Therefore the MATLAB codes have to be changed to take these relations into account.

Below are the essential parts of the MATLAB program.

```
u=a(i,j);
t=sinij(i,j,u,nu);
x=a(:,j);
a(:,j)=min(x,a(:,i)+t*z);
a(:,i)=min(a(:,i),x+t*z);
a(i,j)=aij;
a(j,:)=a(:,j)';
a(i,:)=a(:,i)';
a(i,i)=0;
a(j,j)=0;
```

In addition, we have added the following statements after `a(j,j)=0;`.

```
l=0; for p=1:nn
  ll=l+nu(p);
  if(nu(p) > 1)
    if (l == 0)
      amin = 2*min(min(a(l+1:ll,ll+1:n)));
    elseif (ll == n)
      amin = 2*min(min(a(1:l,l+1:ll)));
    else
      amin1=min(min(a(1:l,l+1:ll)));
      amin2=min(min(a(l+1:ll,ll+1:n)));
      amin=2*min([amin1 amin2]);
    end
    for jj=l+2:ll
      for ii=l+1:jj-1
        a(ii,jj)=max([a(ii,jj) amin]);
        a(jj,ii)=a(ii,jj);
      end
    end
  end
  l=ll;
end
```

```
function t=sinij(i,j,x,nu);
nn=length(nu);
l=0;
for r=1:nn
 l=l+nu(r);
 if (l >= i)
     if (j-i < nu(r) )
         t=0;
     else
         t=x;
     end
     return
 end
end;
```

The above code computes `t` by taking into account whether the pivot element `a(i,j)` is contained in some diagonal block $\mathcal{A}_{ll}$ or not. If it is, then $|\varphi_k|$ can be large and `t` is set zero.

The statements to the left, ensure that the off-diagonal elements within the diagonal blocks of the natural partition $(\mathcal{A}_{pq})$ remain quadratically small in comparison with the largest elements in the corresponding block-rows and block-columns.

Table 3 and Table 4 display the data obtained by the MATLAB program.

	A_{ij}																			
	11	12	13	14	15	16	22	23	24	25	26	33	34	35	36	45	46	55	56	66
1	2	2	2	2	2	2	2	2	2	2	2	4	4	2	2	2	2	2	2	3
2	4	4	4	4	4	6	4	4	4	4	6	8	8	6	6	6	6	6	6	9
3	8	8	12	12	12	14	12	12	12	12	16	24	24	16	16	16	16	16	16	27
4	16	24	28	28	28	30	32	32	32	32	44	56	56	44	44	46	46	46	46	60
5	48	60	74	74	74	88	88	88	88	88	104	148	148	104	104	112	112	112	112	176
6	122	162	192	192	192	222	208	208	208	208	280	384	384	280	280	312	312	312	312	444

Table 3: Block norms after each of the first 6 quasi-sweeps

We can see that the quadratic reduction per quasi-sweep is likely to be somewhat better than in the case of simple eigenvalues.

Our next experiment deals with the case when the basic and the natural partitions coincide (Table 5). After that we have considered the case $\mathcal{N} = (1, 2, 3, 2, 1, 2, 2, 3)$ in which the natural partition is contained within the basic partition (Table 6).

	A_{ij}																			
	11	12	13	14	15	16	22	23	24	25	26	33	34	35	36	45	46	55	56	66
A_{11}^*	8	1	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1
A_{22}	8	1	1	1	1	1	2	1	1	1	1	2	2	1	1	1	1	1	1	1
A_{12}	2	2	1	1	1	1	2	1	1	1	1	2	2	1	1	1	1	1	1	1
A_{22}	2	2	1	1	1	1	2	1	1	1	1	2	2	1	1	1	1	1	1	1
A_{13}	2	2	3	1	1	1	2	1	1	1	1	2	2	1	1	1	1	1	1	1
A_{33}	2	2	3	1	1	1	2	1	1	1	1	∞	2	1	1	1	1	1	1	1
A_{23}	2	2	3	1	1	1	2	2	1	1	1	2	2	1	1	1	1	1	1	1
A_{33}	2	2	3	1	1	1	2	2	1	1	1	∞	2	1	1	1	1	1	1	1
A_{14}	2	2	3	3	1	1	2	2	1	1	1	∞	2	1	1	1	1	1	1	1
A_{24}	2	2	3	3	1	1	2	2	3	1	1	∞	2	1	1	1	1	1	1	1
A_{34}	2	2	3	3	1	1	2	2	2	1	1	2	2	1	1	1	1	1	1	1
A_{15}	2	2	2	2	2	1	2	2	2	1	1	2	2	1	1	1	1	1	1	1
A_{25}	2	2	2	2	2	1	2	2	2	2	1	2	2	1	1	1	1	1	1	1
A_{35}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	1	1	1	1	1
A_{55}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	1	1	2	1	1
A_{45}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	2	1	1	1	1
A_{55}	2	2	2	2	2	1	2	2	2	2	1	2	2	2	1	2	1	2	1	1
A_{16}	2	2	2	2	2	2	2	2	2	2	1	2	2	2	1	2	1	2	1	1
A_{26}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	1	2	1	1
A_{36}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	1	1
A_{46}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1
A_{66}	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2
A_{56}	2	2	2	2	2	2	2	2	2	2	2	4	4	2	2	2	2	2	2	1
A_{66}	2	2	2	2	2	2	2	2	2	2	2	4	4	2	2	2	2	2	2	3

Table 4: Block norms during the first quasi-sweep.

In fact, whatever natural partition we tried for this matrix of order 16, we had not found a case which would lead to the loss of the quadratic convergence. The only "dangerous" situation that we have met appeared for the case $\mathcal{N} = (1,1,3,6,4,1)$ during the annihilations in the block A_{56}. At that stage the elements of A_{66} which have already once been rotated, have increased in magnitude from $\mathcal{O}(\varepsilon^3)$ and from zero to $\mathcal{O}(\varepsilon)$. We shall briefly explain it.

$$\left[\begin{array}{c|c} A_{55} & A_{56} \\ \hline A_{56}^T & A_{66} \end{array}\right] = \left[\begin{array}{cccc|ccc} d & \varepsilon^2 & \varepsilon^3 & \varepsilon^3 & \varepsilon^1 & \varepsilon^1 & \varepsilon^1 \\ \varepsilon^2 & d & \varepsilon^3 & \varepsilon^3 & \varepsilon^1 & \varepsilon^1 & \varepsilon^1 \\ \varepsilon^3 & \varepsilon^3 & d & 0 & \varepsilon^2 & \varepsilon^2 & \varepsilon^1 \\ \varepsilon^3 & \varepsilon^3 & 0 & d & \varepsilon^2 & \varepsilon^2 & \varepsilon^1 \\ \hline \varepsilon^1 & \varepsilon^1 & \varepsilon^2 & \varepsilon^2 & d & \varepsilon^2 & \varepsilon^3 \\ \varepsilon^1 & \varepsilon^1 & \varepsilon^2 & \varepsilon^2 & \varepsilon^2 & d & 0 \\ \varepsilon^1 & \varepsilon^1 & \varepsilon^1 & \varepsilon^1 & \varepsilon^3 & 0 & d \end{array}\right]$$

Since ε^2 elements in A_{56} lie in $\mathcal{A}_{55}$ they can produce large Jacobi angle which can increase the 0 and ε^3 elements of A_{66}, to become ε. This is indicated in Table 3 at A_{56}-row. Nevertheless, the pivot strategy forces the elements of A_{66} to be rotated once again and the overall quadratic reduction will be ensured. In the off-diagonal blocks A_{ij} we cannot have such situation that an element, after being annihilated, grows to as high as ε. It can grow, say, from ε^3 or ε^4 to ε^2 but not to ε.

In fact, whatever n and whatever the natural and the basic partition chosen, we could not imagine the case that would lead to the loss of the quadratic convergence (provided the assumption on the affiliation of the diagonal elements holds).

	A_{ij}																			
	11	12	13	14	15	16	22	23	24	25	26	33	34	35	36	45	46	55	56	66
1	4	2	2	2	2	3	4	2	2	2	3	4	2	2	3	2	3	4	3	6
2	8	4	4	4	6	7	8	4	4	6	7	8	4	6	7	6	9	12	9	12
3	16	8	8	12	14	15	16	8	12	14	15	16	12	16	21	18	27	28	27	30
4	32	16	24	28	30	31	32	24	32	40	45	48	34	48	61	54	81	60	81	62
5	96	48	62	80	88	93	96	66	94	106	121	124	102	116	161	120	177	176	183	186
6	256	128	182	208	234	249	256	196	226	302	351	364	236	336	409	354	519	468	543	498

Table 5: The case $\mathcal{N} = (3, 3, 2, 1, 4, 3)$

	A_{ij}																			
	11	12	13	14	15	16	22	23	24	25	26	33	34	35	36	45	46	55	56	66
1	2	2	2	2	2	3	4	2	2	2	3	4	2	2	3	2	3	2	3	6
2	4	4	4	4	4	7	8	4	4	4	7	8	4	4	7	6	9	6	9	14
3	8	8	8	10	12	15	16	8	10	12	15	16	10	12	19	18	27	18	27	30
4	16	16	20	26	28	31	32	20	30	32	39	40	30	34	55	52	75	54	81	62
5	32	40	52	58	60	63	80	60	84	90	111	104	86	100	143	116	171	120	177	126
6	80	112	136	150	152	159	224	170	206	220	285	272	216	232	337	244	363	248	369	318

Table 6: The case $\mathcal{N} = (1, 2, 3, 2, 1, 2, 2, 3)$

References

[1] Demmel J., Veselić K.: Jacobi's method is more accurate than QR. SIAM J. Matrix Anal. Appl. 13 (1992) 1204–1245.

[2] Dongarra J, Duff I., Sorensen D., van der Vorst H.: Solving Linear Systems on Vector and Shared Memory Computers. SIAM 1991.

[3] Drmač Z.: Implementation of Jacobi rotations for accurate singular value computation in floating-point arithmetic. SIAM J. Sci. Comp. 18 (1997) 1200–1222.

[4] Drmač Z.: A posteriori computation of the singular vectors in a preconditioned Jacobi SVD algorithm. IMA J. Numer. Anal. 19 (1999) 191-213.

[5] Drmač Z., Hari V., Slapničar I.: Advances in Jacobi methods. Proccedings of the Conference on Applied and computational mathematics, Dubrovnik 2002. Kluwer 2003.

[6] Drmač Z., Veselić K.: New fast and accurate Jacobi SVD algorithm I. Part one. Preprint (2004).

[7] Drmač Z., Veselić K.: New fast and accurate Jacobi SVD algorithm I. Part one. Preprint (2004).

[8] Fernando K. V.: Linear Convergence of the Row Cyclic Jacobi and Kogbetliantz Methods. Numer. Math. 56 (1989) 73–94.

[9] Hansen E. R.: On Jacobi Methods and Block Jacobi Methods for Computing Matrix Eigenvalues, Ph. D. thesis, Stanford University 1960.

[10] Hansen E. R.: On cyclic Jacobi methods. SIAM J. Appl. Math. 11 (1963) 449–459.

[11] Hari V.: On the Convergence of Cyclic Jacobi–like Processes. Linear Algebra and Its Appl. 81 (1986) 105–127.

[12] V. Hari: On Pairs of Almost Diagonal Matrices. Linear Algebra and Its Appl. 148 (1991) 193–223.

[13] Hari V.: On sharp quadratic convergence bounds for the serial Jacobi methods. Numer. Math. 60: 375–406, 1991.

[14] Hari V., Drmač Z.:
On Scaled Almost Diagonal Hermitian Matrix Pairs. SIAM J. Matrix Anal. Appl. Vol. 18. No. 4 (1997) 1000-1012.

[15] Hari V., Veselić K.:
On Jacobi Methods for Singular Value Decompositions. SIAM J. Sci. Stat. Comput. Vol. 8, No.5 (1987) 741–754.

[16] Hari V.:
Convergence of a Block-oriented Quasi-cyclic Jacobi Method. Preprint, University of Zagreb 2005.

[17] P. Henrici and K. Zimmermann: An Estimate for the Norms of Certain Cyclic Jacobi Operators, Linear Algebra and Its Applications 1, 489–501, 1968.

[18] Mascarenhas W.: On the convergence of the Jacobi method for arbitrary orderings. SIAM Journal on Matrix Analysis and Applications, Vol. 16, No.4, (1995), pp. 1197-1206.

[19] Matejaš J.: Quadratic convergence of scaled matrices in Jacobi method. Numer. Math. 87 (1999) 171–199.

[20] Matejaš J.: Convergence of scaled iterates by Jacobi method. Linear Algebra and its Applications, Vol. 349, Number 1-3 (2002) 17-53.

[21] Nazareth L.: On the Convergence of the Cyclic Jacobi Method, Linear Algebra and Its Applications 12, 151–164, 1975.

[22] Rhee N. and Hari V.: On the global and cubic convergence of a quasy-cyclic Jacobi method. Numer. Math. 66, (1993) 97–122.

[23] G. Shroff and R. Schreiber: On the Convergence of the Cyclic Jacobi Method for Parallel Block Orderings, SIAM Journal on Matrix Analysis and Applications 10, 326–346, 1989.

[24] Veselić K. and Hari V.: A note on a one-sided Jacobi algorithm. Numer. Math. 56 (1989) 627-633.

[25] Wilkinson J.: Note on the quadratic convergence of the cyclic Jacobi process. Numer. Math. 4 (1962) 296–300.

[26] J. H. Wilkinson: Almost Diagonal Matrices with Multiple or Close Eigenvalues, Linear Algebra and its Applications 1, 1–12, 1968.