A norm-reducing Jacobi-like algorithm for the eigenvalues of non-normal matrices

I. Kiessling and A. Paulik *

ABSTRACT

A new norm decreasing Jacobi-like method for reducing a non-normal matrix to a normal one is described. The method is an improved version of the Huang-Gregory's procedure [6], in which certain norm reducing non-optimal steps are replaced by the optimal ones, which correspond no more to regular matrix transformations. The method renders itself particularly effective in dealing with defective matrices of special forms. Theory and experiments alike indicate that this algorithm is in all computational aspects – accuracy, convergence rate, computing time, complexity of the computer program – better or at least as good as the original one.
Both procedures, the original and the improved one, end in all of the authors computed examples with the diagonal matrix containing the eigenvalues of the non-normal initial matrix.

1. INTRODUCTION

The aim of this paper is to present an improved version of the norm-decreasing Jacobi-like algorithm for reducing a non-normal matrix to a normal one, which was introduced by Huang and Gregory [6]. The above-mentioned algorithm, like all norm-decreasing Jacobi-like algorithms, cf. e.g. [1, 2, 3, 10, 11, 14, 15] originates in the following well known facts :
Let $A = (a_{j,k} | j,k=1,...,n)$, $a_{j,k} \in \mathbb{C}$ be some $n \times n$-matrix, λ_j $(j=1,...,n)$ its – not necessary different – eigenvalues and

$$\|A\| = \Big(\sum_{j,k=1}^{n} |a_{j,k}|^2\Big)^{1/2} \tag{1.1}$$

its Euclidean norm. Then there holds not only the Schur-inequality, cf. [13],

$$\sum_{k=1}^{n} |\lambda_k|^2 \leqslant \|A\|^2, \tag{1.2}$$

with equality only iff A is normal, but – since the spectrum of A does not change by a similarity transformation $A \to ZAZ^{-1}$ – also its generalisation

$$\sum_{k=1}^{n} |\lambda_k|^2 \leqslant \|ZAZ^{-1}\|^2, \text{ for all } Z \in R_n, \tag{1.3}$$

where R_n denotes the set of all regular complex-valued $n \times n$-matrices. Mirski [9] has shown that there holds

$$\sum_{k=1}^{n} |\lambda_k|^2 = \inf_{Z \in R_n} \|ZAZ^{-1}\|^2. \tag{1.4}$$

Let now $(Z_p) \subset R_n$ be some sequence with

$$\sum_{k=1}^{n} |\lambda_k|^2 = \lim_{p \to \infty} \|A_p\|^2, \quad A_p := Z_p A Z_p^{-1}, \tag{1.5}$$

then it can be shown that this is the case then and only then, if

$$\lim_{p \to \infty} \|A_p^* A_p - A_p A_p^*\| = 0. \tag{1.6}$$

Hence, choosing p large enough, A_p can be made arbitrarily close to a normal matrix with the same eigenvalues as A. But for normal matrices there exists an extension of the Jacobi algorithm [7, 12] for symmetric matrices, cf. [4], so the following simple model algorithm for the computation of the eigenvalues of a general matrix may be proposed :

i) Find the minimizing sequence of regular matrices $(Z_p) \subset R_n$ so that (1.5) holds.

ii) Choose $\tilde{p}$ large enough (or take some normal accumulation point $A_\infty = \lim_{k \to \infty} A_{p_k}$) so that $A_{\tilde{p}}$ is nearly normal.

iii) If $A_{\tilde{p}}$ (or A_∞) is not diagonal, apply the sequence (U_p) of the unitary matrices of the extended Jacobi algorithm [4] on $A_{\tilde{p}}$ in order to obtain the eigenvalues λ_k :

$$\Lambda = \operatorname{diag}(\lambda_1,...,\lambda_n) = \lim_{p \to \infty} U_p^* A_{\tilde{p}} U_p \; (\lim_{p \to \infty} U_p^* A_\infty U_p)$$

As one can see, the crucial point in designing any algorithm for the eigenvalue computation of non-normal matrices is to find the sequence (Z_p). The Jacobi-like algorithms [1,2,3,6,10,11,14,15] differ mainly in the way how one chooses the norm-minimizing sequence (Z_p). In each algorithm ii) and iii) are done simultaneously. In the following we confine ourself to the description

* I. Kiessling, A. Paulik, Institut für Numerische und Angewandte Mathematik der Georg August Universität, D-3400 Göttingen, BRD.

and analysis of the algorithm of Huang and Gregory [6].

2. HUANG-GREGORY'S ALGORITHM

The algorithm [6] uses three kinds of similarity matrix transformation. Given a matrix A_p and the pivot pair (k, m), it generates the matrix A_{p+1} by the rule :

$$A_{p+1} = \begin{cases} D^{-1} A_p D, \text{ if } k=m \\ L^{-1} FL \text{ and } F=G^* A_p G, \text{ if } k\neq m, \end{cases} \tag{2.1}$$

with

$$D = \text{diag}(1,\dots,1,d,1,\dots,1), \text{ d in the k-th row.} \tag{2.2}$$

$$G = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1_c & & -se^{i\alpha} & \\ & & & 1 & & \\ & & & & \ddots & \\ & & se^{i\alpha} & & 1_c & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow \text{k-th row} \\ \\ \\ \leftarrow \text{m-th row} \\ \\ \\ \end{matrix} \quad , c=\cos\theta, \; s=\sin\theta, \tag{2.3}$$

$$\begin{matrix} & & \uparrow & & \uparrow \\ & & \text{k-th} & & \text{m-th column} \end{matrix}$$

and

$$L = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & & l_{k,m} & & \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{k-th row} \\ \\ \\ \\ \end{matrix} \quad , l_{k,m}=\rho e^{i\beta}, \tag{2.4}$$

The similarity transformations using matrices D and L are norm reducing, the similarity transformation with the unitary matrix G is diagonalising. The transformation parameters d, θ, α, ρ, β are chosen in the following way :

Parameter d :

Let $A_p = (b_{j,k} \mid j,k=1,\dots,n)$ and

$$\mu_k^2 = \sum_{\substack{j=1 \\ j\neq k}}^{n} |b_{k,j}|^2, \quad \xi_k^2 = \sum_{\substack{j=1 \\ j\neq k}}^{n} |b_{j,k}|^2,$$

$$\hat{\mu}_k^2 = \mu_k^2 + |b_{k,k}|^2, \quad \hat{\xi}_k^2 = \xi_k^2 + |b_{k,k}|^2. \tag{2.5}$$

If $\mu_k \cdot \xi_k \neq 0$, set $d := (\mu_k/\xi_k)^{1/2}$, achieving the maximum possible norm reduction

$$\Delta = \|A_p\|^2 - \|A_{p+1}\|^2 = (\mu_k - \xi_k)^2. \tag{2.6a}$$

If $\mu_k \cdot \xi_k = 0$ and $\hat{\mu}_k \cdot \hat{\xi}_k \neq 0$, set $d := (\hat{\mu}_k/\hat{\xi}_k)^{1/2}$, achieving the norm reduction

$$\Delta = \|A_p\|^2 - \|A_{p+1}\|^2 = (\hat{\mu}_k - \hat{\xi}_k)^2 + |b_{k,k}|^2 (d - \tfrac{1}{d})^2. \tag{2.6b}$$

If $\hat{\mu}_k \cdot \hat{\xi}_k = 0$, i.e. $b_{j,k} = b_{k,j}$ $(j=1,\dots,n)$, deflate the matrix and continue (2.6c)

Parameters θ and α :

Choose G so as to annihilate the element $f_{m,k}$ in $F = (f_{i,k} \mid i,j=1,\dots,n)$, i.e. the parameters θ and α must satisfy the equation

$$\tan\theta = \frac{2b_{m,k}e^{i\alpha}}{t_{k,m} + (t_{k,m}^2 + 4b_{k,m}b_{m,k})^{1/2}},$$

$$\text{with } t_{k,m} = b_{k,k} - b_{m,m}. \tag{2.7}$$

Here α is chosen so that $\tan\theta$ is real and θ is chosen to be the smaller (in absolute value) of the rotation angles.

Parameters ρ and β :

Let

$$v_{k,m} := \sum_{\substack{j=1 \\ j\neq k,m}}^{n} (f_{k,j}\bar{f}_{m,j} - f_{j,m}\bar{f}_{j,k}) - f_{k,m}\bar{\tau}_{k,m},$$

$$\tau_{k,m} := f_{k,k} - f_{m,m},$$

$$u_{k,m} := |\tau_{k,m}|^2 + \sum_{\substack{j=1 \\ j\neq k,m}}^{n} (|f_{j,k}|^2 + |f_{m,j}|^2). \tag{2.8}$$

Since

$$\Delta = \|A_p\|^2 - \|A_{p+1}\|^2 = 2\rho \,\text{Re}(v_{k,m}e^{-i\beta}) - \rho^2 u_{k,m},$$

choose β so as to satisfy the equations

$$v_{k,m}e^{-i\beta} = |v_{k,m}| \tag{2.9}$$

and ρ according to

$$\rho = \begin{cases} \dfrac{|v_{k,m}|}{u_{k,m}}, \text{ for } 0 < |v_{k,m}| \leq u_{k,m}, & (2.10a) \\ 1, \text{ for } 0 < u_{k,m} < |v_{k,m}|, & (2.10b) \\ 0, \text{ for } 0 = v_{k,m}, & (2.10c) \end{cases}$$

achieving the norm reduction

$$\Delta = \begin{cases} \dfrac{|v_{k,m}|^2}{u_{k,m}}, \text{ for } 0 < |v_{k,m}| \leq u_{k,m}, & (2.11a) \\ 2|v_{k,m}| - u_{k,m}, \text{ for } 0 < u_{k,m} < |v_{k,m}|, & (2.11b) \\ 0, \text{ for } 0 = v_{k,m}. & (2.11c) \end{cases}$$

3. IMPROVED ALGORITHM

Let us now analyse the norm reducing transformations. Looking closely at the choice of the parameters in the non-optimal (not maximal) norm reduction transformations, one can see that letting drop the requirement that the norm reduction should happen by the similarity transformation there is a possibility to reach maximal norm reduction, cf. e.g. [8]. Not to get lost in details, we mention only the well known and in the following extensively used principle : in an upper or lower triangular block matrix one may set all the off-diagonal blocks

equal to zero, without changing eigenvalues and their algebraic multiplicities.
We improve now the choice of the transformation parameters d and ρ in (2.6b) and (2.10c) in the following way :

Proposition 1 (improves 2.6b).

If $\mu_k \cdot \xi_k = 0$ then form A_{p+1} from A_p by the rule : set all off diagonal elements in the k-th row and the k-th column of the matrix A_p equal to zero, achieving the maximum possible norm reduction

$$\Delta = \|A_p\|^2 - \|A_{p+1}\|^2 = (\mu_k - \xi_k)^2. \tag{3.1}$$

Proposition 2 (improves 2.10c)

Set ρ according to

$$\rho = \begin{cases} \dfrac{|v_{k,m}|}{u_{k,m}}, & \text{for } 0 < |v_{k,m}| \leqslant u_{k,m}, \quad (3.2a) \\ 1, & \text{for } 0 < u_{k,m} < |v_{k,m}|, \quad (3.2b) \\ 0, & \text{for } 0 = v_{k,m} \text{ and } u_{k,m} \neq |\tau_{k,m}|^2, \quad (3.2c) \end{cases}$$

achieving, except for (3.2b), the maximum possible norm reduction

$$\Delta = \begin{cases} \dfrac{|v_{k,m}|^2}{u_{k,m}}, & \text{for } 0 < |v_{k,m}| \leqslant u_{k,m}, \quad (3.3a) \\ 2|v_{k,m}| - u_{k,m}, & \text{for } 0 < u_{k,m} < |v_{k,m}|, \quad (3.3b) \\ 0, & \text{for } 0 = v_{k,m} \text{ and } u_{k,m} \neq |\tau_{k,m}|^2. \quad (3.3c) \end{cases}$$

If $v_{k,m} = 0$ and $u_{k,m} = |\tau_{k,m}|^2$, then form A_{p+1} from F by setting all off-diagonal elements of the k-th and m-th rows and columns equal to zero, achieving the maximum possible norm reduction

$$\Delta = \|A_p\|^2 - \|A_{p+1}\|^2$$

$$= |f_{k,m}|^2 + \sum_{\substack{j=1 \\ j \neq k,m}}^{n} |f_{k,j}|^2 + |f_{j,m}|^2. \tag{3.4}$$

Proof of the proposition 1

Let $\mu_k \cdot \xi_k = 0$. Since this fact implies that A_p has one of the forms

$$\text{k-th row} \begin{cases} \left[\begin{array}{c} x \\ \vdots \\ x \\ 0 \ldots 0 a_{k,k} 0 \ldots 0 \\ x \\ \vdots \\ x \end{array}\right] \text{ or } \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ x \ldots x a_{k,k} x \ldots x \\ 0 \\ \vdots \\ 0 \end{array}\right], \\ \qquad \text{k-th column} \end{cases} \tag{3.5}$$

we may set all the remaining off-diagonal elements of the k-th column or k-th row, respectively, equal to zero, without changing the eigenvalues of A_p. Hence (3.1) follows.

Proof of the proposition 2

Since we have chosen β according to (2.9) we obtain for Δ the expression

$$\Delta = \|A_p\|^2 - \|A_{p+1}\|^2 = 2\rho |v_{k,m}| - \rho^2 u_{k,m}, \tag{3.6}$$

which is for $u_{k,m} \neq 0$ a quadratic function of ρ. If $u_{k,m} \neq 0$, it is positive for

$$0 < \rho < \frac{2|v_{k,m}|}{u_{k,m}} =: \hat{\rho} \tag{3.7}$$

and achieves a maximum (3.3a) for ρ chosen according to (3.2a). If $\hat{\rho} > 1$, then we choose for computational purposes $\rho = 1$ according to (3.2b) and obtain the (not maximum possible) norm reduction (3.3b). If $v_{k,m} = 0$, then (3.6) reduces to

$$\Delta = -\rho^2 u_{k,m}, \tag{3.8}$$

implying that the only possible regular transformation, which does not increase the norm of A_{p+1} is to choose $\rho = 0$. However, there is an important exception for $u_{k,m} = |\tau_{k,m}|^2$. In this case we obtain from (2.8)

$f_{j,k} = f_{m,j} = 0 \quad (j = 1, \ldots, n; j \neq k)$

so that F has the form

$$\left[\begin{array}{ccc} 0 & & x \\ \vdots & & \vdots \\ 0 & & x \\ x \ldots x & f_{k,k} \; x \ldots x & f_{k,m} \; x \ldots x \\ 0 & & x \\ \vdots & & \vdots \\ 0 & & x \\ 0 \ldots 0 \; 0 & 0 \ldots 0 & f_{m,m} \; 0 \ldots 0 \\ 0 & & x \\ \vdots & & \vdots \\ 0 & & x \end{array}\right] \tag{3.9}$$

Hence, we may set all the remaining off-diagonal elements of the k-th and m-th rows and columns equal to zero, without changing the eigenvalues of A_p, obtaining the maximum possible norm reduction (3.4).

Remark 1

The norm reduction in proposition 2 occurs also in the important case of the defective eigenvalues : the difference between the algebraic and the geometric multiplicity of the eigenvalues $f_{k,k}$, $f_{m,m}$ in the case $f_{k,k} = f_{m,m}$ is at least 1, if $\Delta > 0$.

4. NUMERICAL RESULTS

All calculations were performed with a UNIVAC Series 1100 computer using 60 bits for mantissa. The authors of [6] performed their calculations with an IBM 360/65 using 56 bits for mantissa. Our implementation of the Huang-Gregory's algorithm reproduced therefore the results for the matrices given in [6] with little higher accuracy.
Let us now consider the conditions which must be satis-

TABLE 1

ex. 4 of [6]	stopping criterion	A	B		C	
			$\epsilon = 0$	$\epsilon = 10^{-12}$	$\epsilon = 0$	$\epsilon = 10^{-12}$
$\|A_p - \mathrm{diag}(A_p)\|$	i)	$.90.10^{-21}$	$.20.10^{-35}$		$.20.10^{-35}$	0
	ii)		$.66.10^{-144}$	$.17.10^{-3}$	0	0
$\|A_p A_p^* - A_p^* A_p\|$	i)		$.23.10^{-38}$		$.23.10^{-88}$	0
	ii)		$.4.10^{-144}$	$.45.10^{-7}$	0	0
$\frac{\|A - A_p\|}{\|A\|}$	i)	$.40.10^{-7}$	$.41.10^{-8}$		$.41.10^{-8}$	$.41.10^{-8}$
	ii)		$.41.10^{-8}$	$.23.10^{-4}$	$.41.10^{-8}$	$.41.10^{-8}$
Number of sweeps p	i)	9	9		9	8
	ii)		30	30	11	8
CPU [s]	i)		.650		.697	.542
	ii)		1.899	2.139	.799	.542

fied applying the improved algorithm. Consider $\xi_k \cdot \mu_k = 0$ and $\hat{\xi}_k \cdot \hat{\mu}_k \neq 0$ (cf. (2.6b)).This case occurs for instance, if $\xi_k = 0$; μ_k, $\hat{\xi}_k$, $\hat{\mu}_k \neq 0$. Due to the rounding errors the evaluation of ξ_k may give a small value > 0 instead of zero. Then proposition 1 giving maximal norm reduction would not be applied. To suppress this influence of rounding errors we define a parameter $\epsilon > 0$ and evaluate the relational expression $\xi_k > \epsilon$ instead of $\xi = 0$. If a matrix occurs having the form (3.9) we proceed in an analogous way.

The numerous examples computed by us can be divided into three classes.

1. General class, containing examples 2 to 7 from [6] for instance. The improved algorithm is at least as good as the original one.
2. Special matrices having essentially triangular or Jordan's normal form. These structures are recognized by the improved algorithm giving the eigenvalues after the first sweep (1). Curiously, the Huang-Gregory's algorithm needs a multiple number of sweeps or it fails.
3. In exceptional cases, see ex. 1 [6], both algorithms give a normal, but not a diagonal matrix. Performing the algorithms in such a way that small perturbations of the matrices A_p due to the rounding errors are admitted does not change the eigenvalues. However, the computation algorithms are instable in that sense that after slightly disturbing the matrices A_p converging against some normal matrix, convergence occurs against a diagonal matrix. So we obtain in ex. 1 [6] the eigenvalues with high accuracy after 6 sweeps.

To give a numerical example of the 1-st class we select ex. 4 of [6]. This is a defective 5*5-matrix. In the following table 1 column A contains values taken from [6], column B values of the Huang-Gregory's algorithm implemented by us, column C values of the improved algorithm, both for $\epsilon = 0$ and $\epsilon = 10^{-12}$, respectively. All rows marked by i) contain the data obtained by that sweep after which condition i) $\|A_p - \mathrm{diag}(A_p)\| < 10^{-8}$ was satisfied for the first time. An empty field in a row marked by i) means that condition i) was not satisfied after the 30-th sweep. All rows marked by ii) contain data obtained by that sweep after which $\|A_p - \mathrm{diag}(A_p)\|$ or $\|A_p A_p^* - A_p^* A_p\|$ became zero for the first time or those computed by the 30-th sweep. An empty field in a row marked by ii) means that condition i) was satisfied for the first time by $\|A_p - \mathrm{diag}(A_p)\| = 0$. One can see that for $\epsilon = 0$ the two algorithms do not differ remarkably. This behaviour is not surprising since the non-regular transformations do not occur in the first 9 sweeps due to the fact that under the influence of the rounding errors generally μ_k, ξ_k, $u_{k,m} \neq 0$.

For $\epsilon = 10^{-12}$ the number of sweeps p of the improved algorithm needed to satisfy stopping criterion i) decreases to p = 8 whereas the Huang-Gregory's algorithm cannot satisfy this condition.

We set in the table $\Lambda := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with exact eigenvalues λ_k of the matrix A.

(1) A sweep is by definition a sequence of $\frac{n}{2}(n+1)$ transformations (2.1), in which no pivot pair appears more than once.

REFERENCES

1. EBERLEIN P. J. : 'A Jacobi-like method for the automatic computation of eigenvalues and eigenvectors of an arbitrary matrix', J. SIAM, 10 (1962) 74-88.
2. EBERLEIN P. J., BOOTHROYD J. : 'Solution to the eigenproblem by a norm-reducing Jacobi-type method', Numer. Math., 11 (1968) 1-12.
3. EBERLEIN P. J. : 'Solution to the complex eigenproblem by a norm-reducing Jacobi-type method', Numer. Math., 14 (1970) 232-245.

4. GOLDSTINE H. H., HORWITZ L. P. : 'A procedure for the diagonalisation of normal matrices', J. ACM, 6 (1959), 176-195.

5. GREGORY R. T., KARNEY D. L. : *A collection of matrices for testing computational algorithms,* Wiley-Interscience, 1969.

6. HUANG C. P., GREGORY R. T. : 'A norm-reducing Jacobi-like algorithm for the eigenvalues of non-normal matrices, Colloquia Mathematica Societatis Janos Bolyai, Keszthely (Hungary), (1977) 365-393.

7. JACOBI C. G. J. : 'Über ein leichtes Verfahren, die in der Theorie der Säkularstörungen vorkommenden Gleichungen numerisch aufzulösen', J. Reine Angew. Math., 30 (1864) 51-94.

8. KRESS R., DE VRIES H. L., WEGMANN R. : 'On nonnormal matrices', Linear Algebra and Appl., 8 (1974) 108-120.

9. MIRSKY L. : 'On the minimization of matrix norms', Amer. Math. Monthly, 65 (1958) 106-107.

10. RUHE A. : 'On the quadratic convergence of a generalization of the Jacobi method to arbitrary matrices', BIT 8 (1968) 210-231.

11. RUTISHAUSER H. 'Une méthode pour le calcul des valeurs propres des matrices non symétriques', Comptes Rendus, 259 (1964) 2758.

12. RUTISHAUSER H. : 'The Jacobi method for real symmetric matrices', in J. H. Wilkinson; C. Reinsch, *Handbook for automatic computation, Vol. II, Linear algebra,* Springer Verlag, 1971

13. SCHUR I. : 'Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen', Math. Ann., 66 (1909) 488-510.

14. VESELIČ K. 'On a class of Jacobi-like procedures for diagonalizing arbitrary real matrices', Numer. Math. 33, (1979) 157-172.

15. VESELIČ K., WENZEL H. J. : 'A quadratically convergent Jacobi-like method for real matrices with complex eigenvalues', Numer. Math. 33 (1979) 425-435.