An accurate bidiagonal reduction for the computation of singular values

Rui M. S. Ralha *

Abstract

We present a new bidiagonalization technique which is competitive with the standard bidiagonalization method and analyse the numerical accuracy of such technique. We show that, for some matrices, our algorithm produces singular values with low relative errors.

1 Introduction

In [10] we proposed a new algorithm for reducing a rectangular matrix A to bidiagonal form from where the singular values of A can be computed iteratively using the classical method due to Golub and Kahan or any of the alternative algorithms given in [6], [9] and [12]. In [13] we presented results of numerical tests carried out on a CRAY Y-MP EL and compared the speed of our code (in Fortran77) to the speed of: the CRAY's library module for SVD, the NAG's routine where R-bidiagonalization as proposed in [4] is implemented, and the LAPACK's routine; these results have shown that our algorithm is very competitive in terms of performance and we claimed that it is a serious alternative for computing the singular values of large matrices. However, at the time we had not fully understood the behaviour of our method for matrices with very small singular values, more precisely we were not able to explain the errors in the smallest singular values of almost rank defficient matrices. We have done significant progress in this matter since we have identified the cause of such errors and know how to improve the original method in order to produce accurate small singular values. Furthermore, we found that for matrices $A = DX$ where D is diagonal and X is well conditioned, our bidiagonalization scheme, just like the one-sided Jacobi algorithm, produces singular values with low relative errors.

2 A new bidiagonalization method

Our bidiagonalization strategy consists of two stages. In the first one (dubbed *"triorthogonalization"*) we perform $n-2$ Householder transformations

$$A_r := A_{r-1}H_r, \quad r = 1, \cdots, n-2 \tag{1}$$

where A_0 represents the initial m-by-n matrix A. The basic idea is to select the appropriate Householder matrices H_r in (1) such that the resulting rectangular matrix A_{n-2} is *"triorthogonal"*, i.e. the columns a_i of A_{n-2} satisfy

$$a_i^T a_j = 0 \ \text{ for } \ |i-j| > 1 \tag{2}$$

Noting that (2) is equivalent to say that the n-by-n symmetric matrix $A_{n-2}^T A_{n-2}$ is tridiagonal, we conclude that to produce in (1) a rectangular "triorthogonal" matrix, we just need to use the Householder matrices that, if applied on both sides of the symmetric matrix $A^T A$, would produce a similar tridiagonal form T in the well-known reduction

$$T := H_{n-2} \cdots H_1 \left(A^T A\right) H_1 \cdots H_{n-2} = \left(AH_1 \cdots H_{n-2}\right)^T \left(AH_1 \cdots H_{n-2}\right) \tag{3}$$

*Centre of Mathematics, University of Minho, 4710-057 Braga, Portugal. email: r_ralha@math.uminho.pt

In the rth step, the computation of the Householder matrix H_r requires only the $n-r$ last elements of the rth column of $A_{r-1}^T A_{r-1}$; since we do not have this product explicitly formed, we just need to compute the $n-r$ *dot* products with the appropriate columns of A_{r-1}.

The second stage of our algorithm is a variant of the modified Gram-Schmidt orthogonalization method (*MGS*) and produces an upper bidiagonal matrix B in the decomposition $A_{n-2} = QB$, where Q is orthogonal. Such decomposition requires only $O(mn)$ *flops* since the property (2) holds. In the following we will refer to the procedure that implements this decomposition as *TRIMGS*.

3 Accuracy of the singular values of A_{n-2}

Since the reduction of the initial matrix to "triorthogonal" form involves $n-2$ Householder transformations, we can expect the process to be numerically stable. In fact, we have, with ϵ being the machine precision,

Theorem 1 *The "triorthogonalization" algorithm is backward stable, i.e. for the computed "triorthogonal" matrix* $\widetilde{A}_{n-2} := A\widetilde{H}_1\widetilde{H}_2\cdots\widetilde{H}_{n-2}$, *we have*

$$\widetilde{A}_{n-2} = (A+E)\,H_1H_2\cdots H_{n-2}, \quad \textit{with } \|E\|_2 = (n-2).O(\epsilon).\,\|A\|_2 \tag{4}$$

Proof. To prove this result, we simply need to adapt the proof given in [14], p. 124, for a sequence of orthogonal transformations applied to both sides of A ■

From (4) we conclude, using a well-kown perturbation theorem, that the singular values $\widetilde{\sigma}_i$ of the "triorthogonal" matrix $\widetilde{A}_{n-2}$ are close to the singular values σ_i of A, in the sense that

$$|\sigma_i - \widetilde{\sigma}_i| \leq (n-2).O(\epsilon).\,\|A\|_2 \tag{5}$$

holds for $i = 1,\cdots,n$. This is a satisfactory error bound for large singular values (those near $\|A\|_2$) since most of the computed digits of $\widetilde{\sigma}_i$ will coincide with the correct ones. However, such bound is not so good for singular values much smaller than $\|A\|_2$ which may exhibit very large relative errors. It is known that the one-sided Jacobi method can compute all singular values to high relative accuracy for matrices $A = DX$, where D is diagonal and X is well-conditioned (see [14], pp.250-251). The relative errors in the singular values $\widehat{\sigma}_i$ of the matrix $\widehat{G}$ obtained from A with post-multiplication of m successive Givens rotations, satisfy the following bound

$$\frac{|\sigma_i - \widehat{\sigma}_i|}{\sigma_i} \leq O(m\varepsilon)\kappa_2(X) \tag{6}$$

Interestingly, the proof of this result can be adapted to our "triorthogonalization" method, since it can be shown that such proof, given in [14], can be extended to any one-sided orthogonal transformation. Therefore, the bound (6) holds, with $m = n-2$, for the singular values of the "triorthogonal" matrix $\widetilde{A}_{n-2}$.

4 Loss of orthogonality and errors

In practice, we have experienced that very small singular values of the bidiagonal form produced with our method, let us call it *ORTHOGSVD*, may exhibit large absolute errors. Is this because we are using *MGS* that is known to produce a matrix which may be far from orthogonal in the case of ill-conditioned matrices? Since Björck has shown [8] that *MGS* produces a triangular matrix which is numerically as good as that from the Householder QR factorization,we conclude that the singular values of the triangular matrix are close to those of the initial matrix. However, our procedure *TRIMGS* builds an upper bidiagonal form, assuming that a complete QR decomposition would produce a triangular matrix R with negligible elements r_{ij}, for $j > i+1$. Unfortunately, we can not guarantee this, in general, since for ill-conditioned matrices there may appear large $|r_{ij}|$ above the bidiagonal form; for instance, for the "triorthogonal" matrix obtained from the Hilbert matrix of order $n = 11$, the last four columns of R are given in Table 1.

$\cdots$	$-2.7493e-17$	$4.4252e-18$	$-5.9938e-18$	$1.1687e-17$
$\cdots$	$-7.3356e-17$	$1.3068e-17$	$-1.8988e-17$	$3.7100e-17$
$\cdots$	$2.9805e-16$	$-5.3282e-17$	$7.8085e-17$	$-1.5086e-16$
$\cdots$	$1.3420e-15$	$-2.3995e-16$	$3.5153e-16$	$-6.7919e-16$
$\cdots$	$-6.7391e-15$	$1.2050e-15$	$-1.7654e-15$	$3.4108e-15$
$\cdots$	$3.8096e-14$	$-6.8119e-15$	$9.9794e-15$	$-1.9281e-14$
$\cdots$	$-4.0374e-08$	$4.3893e-14$	$-6.4303e-14$	$1.2424e-13$
	$5.3952e-09$	$7.4512e-10$	$-4.8120e-13$	$9.2972e-13$
		$8.3168e-11$	$-2.8831e-12$	$-8.3296e-12$
			$3.0252e-13$	$7.2449e-13$
				$8.8244e-15$

Table 1: Last four columns of R produced by the MGS method applied to the "triorthogonal" matrix obtained from the Hilbert matrix of order $n = 11$.

We notice that there are elements above the upper bidiagonal form of absolute value as large as 10^{-12} and 10^{-13}, therefore errors of this size may affect the singular values of the bidiagonal form produced with *TRIMGS*. This actually happens since the two smallest singular values computed with *ORTHOGSVD* are $5.7089e-013$ and $8.4136e-012$ whereas the correct values, with five significant digits, are $3.3932e-015$ and $7.8071e-013$.

The problem is that two columns, say $a_i^{(1)}$ and $a_j^{(1)}$, of A_{n-2} may be far from orthogonal, even if ${a_i^{(1)}}^T a_j^{(1)} \leq \epsilon$, when at least one of the norms $||a_i^{(1)}||$ or $||a_j^{(1)}||$ is very small; in other words, A_{n-2} is truly "triorthogonal", to working precision, only if the quantities

$$c_{ij}^{(1)} := \frac{{a_i^{(1)}}^T a_j^{(1)}}{||a_i^{(1)}||\,||a_j^{(1)}||} \tag{7}$$

are close to the machine precision ϵ, for $j > i + 1$.

We have investigated how large values $|c_{ij}^{(1)}|$ introduce important errors in the bidiagonal form produced with *TRIMGS*. For the elements r_{1j} of the first row of R we have

$$r_{1j} := ||a_j^{(1)}||.c_{1j}^{(1)}, \qquad j = 2, ..., n \tag{8}$$

Therefore, $|r_{1j}|$ can be much smaller than $|c_{1j}^{(1)}|$ provided $||a_j^{(1)}||$ is small, i.e. loss of orthogonality of the columns $a_j^{(1)}$, $j > 2$, relatively to the first column $a_1^{(1)}$, is harmless to the accuracy of the computed R since such loss of accuracy occurs gradually with decreasing $||a_j^{(1)}||$. For this reason, we expect to have in all cases

$$|r_{1j}| \approx ||A||\,\epsilon, \quad j = 3, ..., n \tag{9}$$

Representing by $a_i^{(j)}$ the ith column of the matrix under transformation, after making it orthogonal to q_{j-1}, we have proved that, for each $j = 3, \cdots, n$

$$|r_{ij}| \approx |r_{i-1,j}| \frac{||a_i^{(i-1)}||}{||a_i^{(i)}||}, \quad i = 2, \cdots, j-2 \tag{10}$$

This gives an estimate of the growth of the size of the elements $r_{2,j}, \cdots\ r_{j-2,j}$ in the jth column of R , starting with r_{1j}. In practical tests, we found this estimate to be quite accurate.

5 Reorthogonalization of columns

From the analysis carried out in the previous section, it becomes clear that for the bidiagonal form produced with *TRIMGS* to have accurate singular values, the elements r_{ij} have to be negligible, for $j = 3, \cdots, n$ and $i = 1, \cdots, j-2$. This will happen if A_{n-2} is "triorthogonal" to working precision and in some cases this may require reorthogonalization of the columns of A_{n-2}, i.e. we apply the same procedure twice, the first time to produce A_{n-2} and the second time to improve the "triorthogonality" of this matrix.

Having completed this procedure with the Hilbert matrix of order $n = 11$, the MGS method applied to the resulting "triorthogonal" matrix produces an upper triangular matrix whose last four columns are given in Table 2.

$\cdots$	$-5.9847e-025$	$8.3919e-027$	$4.3511e-030$	$-6.2775e-029$
$\cdots$	$-1.2327e-024$	$1.5856e-026$	$1.6832e-028$	$-1.6234e-028$
$\cdots$	$-2.8172e-024$	$7.6665e-026$	$-8.9748e-028$	$2.4253e-028$
$\cdots$	$-7.6227e-024$	$-6.5261e-026$	$-4.6777e-027$	$1.1409e-027$
$\cdots$	$4.5097e-023$	$-2.8395e-025$	$2.4386e-026$	$-5.8864e-027$
$\cdots$	$-2.4042e-022$	$1.3565e-024$	$-1.3398e-025$	$3.4580e-026$
$\cdots$	$4.0374e-008$	$-8.7048e-024$	$8.7033e-025$	$-2.1923e-025$
	$5.3952e-009$	$7.4513e-010$	$6.5099e-024$	$-1.6409e-024$
		$8.3159e-011$	$-8.8968e-012$	$1.4707e-023$
			$7.8320e-013$	$5.6476e-014$
				$3.4089e-015$

Table 2: Last four columns of R produced by the MGS method applied to the triorthogonal matrix (with reorthogonalization) obtained from the Hilbert matrix of order $n = 11$.

These values are to be compared with those given in Table 1 and show that the elements r_{ij} above the bidiagonal form are negligible. Therefore, we expect the bidiagonal form computed with the procedure *TRIMGS* to be accurate (in fact, its elements coincide with those of the corresponding diagonals of R up to the machine precisison).

Of course, the use of reorthogonalization doubles the arithmetic complexity of the process of producing a "triorthogonal" matrix. We have not devised so far an efective sheme to implement selective reorthogonalization in this procedure. It must be stressed out that not all ill-conditioned matrices require the use of reorthogonalization. In the next section we give some examples of such matrices.

6 Numerical results

In this section we discuss the numerical results obtained with our procedure in the case of selected matrices.

- Lauchli matrices

A Lauchli matrix is a $(n+1) \times n$ matrix of the type

$$L(n,\mu) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \mu & & & & \\ & \mu & & & \\ & & \mu & & \\ & & & \ddots & \\ & & & & \mu \end{bmatrix}$$

The singular values of $L(n,\mu)$ are $\sigma_1 = \mu$, of multiplicity $n-1$, and $\sigma_2 = \sqrt{n+\mu^2}$. For small values of μ, $L(n,\mu)$ is ill-conditioned, therefore we are interested in testing *ORTHOGSVD* for matrices of this type. We found that, in most cases, the method produces singular values with very low relative errors, even for very small values μ. For instance, with $\mu = \sqrt{\epsilon}$ almost all singular values exhibit a relative error of the order of magnitude of ϵ, although for certain sizes n, one of the singular values is affected by a larger relative error. In Table 3, the maximum relative error produced with ORTHOGSVD is shown for some values n. With $\mu = \epsilon$, again all but one of the singular values have relative errors close to ϵ but one singular value exhibits a very large relative error, even for small sizes n. For instance, for $n = 7$, we get one singular value equal to $5.4390e - 016$ which has not a single significant digit correct, since $\epsilon \approx 2.2204e - 016$. The gain in accuracy due to the use of reorthogonalization is dramatic in the case of $\mu = \epsilon$, as it can be appreciated in Table 4. From Table 3, we can also conclude that *ORTHOGSVD* with reorthogonalization is more accurate than the procedure *svd* of *Matlab* (for $\mu = \epsilon$ the relative errors of the singular values produced with *svd* are essentially the same as for $\mu = \sqrt{\epsilon}$).

n		50	100	200	300	400	500
$\max \frac{\lvert\sigma_i - \tilde{\sigma}_i\rvert}{\sigma_i}$	(a)	5.5e-16	2.4e-15	4.9e-14	8.2e-15	3.9e-13	2.2e-13
	(b)	8.8e-16	1.5e-15	1.8e-15	1.8e-15	2.8e-15	2.7e-15
	(c)	4.2e-15	7.4e-15	2.0e-14	5.9e-14	6.3e-14	1.4e-13

Table 3: Maximum relative error in the singular values of $L(n, \sqrt{\epsilon})$, computed with *ORTHOGSVD*, with (a) and without reorthogonalization (b) and with the function svd of Matlab (c).

n		50	100	200	300	400	500
$\max \frac{\lvert\sigma_i - \tilde{\sigma}_i\rvert}{\sigma_i}$	(a)	5.5e-16	4.0e+00	2.0e+01	7.6e+00	5.8e+01	4.3e+01
	(b)	8.8e-16	1.2e-15	1.7e-15	2.2e-15	2.1e-15	2.7e-15

Table 4: Maximum relative error in the singular values of $L(n, \epsilon)$, computed with *ORTHOGSVD*, without (a) and with reorthogonalization (b).

6.1 Random matrices produced with RANDSVD

The Matlab's function *randsvd* [7] generates a random matrix with pre-assigned singular values. Used in the form $A = randsvd(n, k, 1)$, it produces a square matrix of order n, with a single singular value equal to one and $n - 1$ singular values equal to $1/k$. In Table 5 the maximum absolute error in the smallest $n - 1$ singular values of matrices of this type, produced with $k = 10^7$, is given (with $k = 10^{15}$, we obtained similar absolute errors).

n		50	100	200	300	400	500
$\max_{2 \le i \le n} \lvert\sigma_i - \tilde{\sigma}_i\rvert$	(a)	5.5e-17	4.4e-17	4.3e-17	4.1e-17	4.2e-17	4.2e-17
	(b)	5.5e-17	5.9e-17	6.8e-17	5.3e-17	5.5e-17	8.9e-17

Table 5: Maximum absolute error in the smallest $n-1$ singular values of $A = randsvd(n, 10^7, 1)$, computed with *ORTHOGSVD*, without reorthogonalization (a), and with the function svd of Matlab (b).

6.2 A=DX

Finally, we consider the example used in [14], pp. 252-253, to illustrate the hability of the one-sided Jacobi algorithm to compute the singular values of A with small relative errors, according to (6), when $A = DX$, where D is diagonal and X is well-conditioned. The matrix

$$G \equiv \begin{bmatrix} \eta & 1 & 1 & 1 \\ \eta & \eta & 0 & 0 \\ \eta & 0 & \eta & 0 \\ \eta & 0 & 0 & \eta \end{bmatrix}$$

with $\eta = 10^{-20}$, has the singular values (to at least 16 digits) $\sqrt{3}$, $\sqrt{3}\eta$, η, η. As shown by Demmel, the classical reduction to bidiagonal form produces a matrix whose three smallest singular values are very inaccurate approximations of the true values. In constrast, the one-sided Jacobi method produces accurate singular values. We now show, that in the case of this matrix our method also produces very accurate singular values. A straightforward Matlab's implementation of our method applied to matrix G produces a matrix B which has the following singular values (computed with MAPLE using 30 decimal digits arithmetic)

$$\begin{aligned}
\sigma_1 &= 1.732\,050\,807\,568\,877\,270\,145\,585 \\
\sigma_2 &= 1.732\,050\,807\,568\,877\,316\,909\,308 \times 10^{-20} \\
\sigma_3 &= 9.999\,999\,999\,999\,996\,710\,898\,576 \times 10^{-21} \\
\sigma_4 &= 9.999\,999\,999\,999\,995\,289\,101\,424 \times 10^{-21}
\end{aligned}$$

which coincide with $\sqrt{3}$, $\sqrt{3}\eta$, η, η, respectively, to at least 16 decimal digits.

References

[1] J.K.H.Wilkinson, The Algebraic Eigenvalue Problem, The Oxford University Press, 1965

[2] A. Bjorck, Solving Linear Least Squares Problems by Gram-Schmidt Orthogonalization, BIT 7:1-21 (1967).

[3] B.N.Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Inc., 1980.

[4] T. F. Chan, An Improved Algorithm for Computing the Singular Value Decomposition, ACM Trans. on Math. Software, vol.8: 72-83 (1982)

[5] G. H. Golub, C. F. Van Loan, Matrix Computations, 2nd ed., The Johns Hopkins University Press, 1989.

[6] J. Demmel, W. Kahan, Accurate Singular Values of Bidiagonal Matrices, SIAM J. Sci. Sta Comput. 11:873-912 (1990).

[7] N. Higham, A Collection of Test Matrices in Matlab, ACM Trans. on Math. Software, vol.17, N.3: 289-305 (1991).

[8] A. Bjorck, Numerics of Gram-Schmidt Orthogonalization, Linear Algebra and its Applications 197, 198:297-316(1994).

[9] K. Vince Fernando, Beresford N. Parlett, Accurate Singular Values and Differential QD Algorithms, Numerische Mathematik, vol.67: 191-229 (1994).

[10] R. M. S. Ralha, A New Algorithm for Singular Value Decompositions, Proc. of the II Euromicro Work. on Parallel and Distributed Processing, 240-244, IEEE Computer Society Press (1994).

[11] C. L. Lawson, R. J. Hanson, Solving Least Squares Problems, Classics in Applied Mathematics, SIAM, 1995.

[12] T. Y. Li, Noah H. Rhee, Zhonggang Zeng, An Efficient and Accurate Parallel Algorithm for the Singular Value Problem of Bidiagonal Matrices, Numerische Mathematik, vol.69: 283-301 (1995).

[13] R. M. S. Ralha, A. Mackiewicz, An Efficient Algorithm for the Computation of Singular Values, Proc. of the III International Congress of Numerical Methods in Engineering, M. Doblaré, J.M. Correas, E. Alarcón, L. Gavete and M. Pastor (Eds.), 1371-1380, SEMNI (1996).

[14] J. Demmel, Applied Numerical Linear Algebra, SIAM (1997).