IMA Journal of Numerical Analysis (2002) **22**, 521–528

When modified Gram–Schmidt generates a well-conditioned set of vectors

L. GIRAUD† AND J. LANGOU‡
Cerfacs, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex 1, France

[Received on 21 July 2001; revised on 17 December 2001]

In this paper, we show why the modified Gram–Schmidt algorithm generates a well-conditioned set of vectors. This result holds under the assumption that the initial matrix is not 'too ill-conditioned' in a way that is quantified. As a consequence we show that if two iterations of the algorithm are performed, the resulting algorithm produces a matrix whose columns are orthogonal up to machine precision. Finally, we illustrate through a numerical experiment the sharpness of our result.

Keywords: Gram–Schmidt algorithm; orthogonalization schemes; floating point arithmetic.

1. Introduction

In this paper we study the condition number of the set of vectors generated by the Modified Gram–Schmidt (MGS) algorithm in floating-point arithmetic. After a quick review, in Section 2, of the fundamental results that we use, we devote Section 3 to our main theorem. Through this central theorem we give an upper bound close to one for the condition number of the set of vectors produced by MGS. This theorem applies to matrices that are not 'too ill-conditioned'. In Section 4, we give another way to prove a similar result. This other point of view throws light on the key points of the proof. In Section 4.2 we combine our theorem with a well known result from Björck to obtain that two iterations of MGS are indeed enough to get a matrix whose columns are orthogonal up to machine precision. We conclude Section 4 by exhibiting a counter-example matrix. This matrix shows that if we relax the constraint on the condition number of the studied matrices, no pertinent information on the upper bound of the condition number of the set of vectors generated by MGS can be gained. For the sake of completeness, we give explicitly the constants that appear in our assumptions and formula: Appendix A details the calculus of those constants.

2. Previous results and notation

We consider the MGS algorithm applied to a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ with full rank $n \leqslant m$ and singular values: $\sigma_1 \geqslant \cdots \geqslant \sigma_n > 0$; we define the condition number of $\mathbf{A}$ as $\kappa(\mathbf{A}) = \sigma_1/\sigma_n$.

Using results from Björck (1967) and Björck & Paige (1992), we know that, in floating-point arithmetic, MGS computes $\bar{\mathbf{Q}} \in \mathbb{R}^{m\times n}$ and $\bar{\mathbf{R}} \in \mathbb{R}^{n\times n}$ so that there exists $\bar{\mathbf{E}} \in \mathbb{R}^{m\times n}$,

†Email: Luc.Giraud@cerfacs.fr
‡Email: Julien.Langou@cerfacs.fr

$\hat{\mathbf{E}} \in \mathbb{R}^{m\times n}$ and $\hat{\mathbf{Q}} \in \mathbb{R}^{m\times n}$, where

$$\mathbf{A} + \bar{\mathbf{E}} = \bar{\mathbf{Q}}\bar{\mathbf{R}} \quad \text{and} \quad \|\bar{\mathbf{E}}\|_2 \leqslant \bar{c}_1 u \|\mathbf{A}\|_2, \tag{2.1}$$

$$\|\mathbf{I} - \bar{\mathbf{Q}}^T\bar{\mathbf{Q}}\|_2 \leqslant \bar{c}_2 \kappa(\mathbf{A}) u, \tag{2.2}$$

$$\mathbf{A} + \hat{\mathbf{E}} = \hat{\mathbf{Q}}\bar{\mathbf{R}}, \quad \hat{\mathbf{Q}}^T\hat{\mathbf{Q}} = \mathbf{I} \quad \text{and} \quad \|\hat{\mathbf{E}}\|_2 \leqslant cu\|\mathbf{A}\|_2. \tag{2.3}$$

$\bar{c}_i$ and c are constants depending on m, n and the details of the arithmetic, and $u = 2^{-t}$ is the unit round-off.

Result (2.1) shows that $\bar{\mathbf{Q}}\bar{\mathbf{R}}$ is a backward-stable factorization of $\mathbf{A}$, that is the product $\bar{\mathbf{Q}}\bar{\mathbf{R}}$ represents accurately $\mathbf{A}$ up to machine precision.

Equation (2.3) says that $\bar{\mathbf{R}}$ solves the QR-factorization problem in a backward-stable sense; that is, there exists an exact orthonormal matrix $\hat{\mathbf{Q}}$ so that $\hat{\mathbf{Q}}\bar{\mathbf{R}}$ is a QR factorization of a slight perturbation of $\mathbf{A}$.

We notice that results (2.1) from Björck (1967) and (2.3) from Björck & Paige (1992) are proved under assumptions

$$2.12 \cdot (m+1)u < 0.01, \tag{2.4}$$

$$cu\kappa(\mathbf{A}) < 1. \tag{2.5}$$

For clarity, it is important to explicitly define the constants that are involved in the upper bounds of the inequalities. Complying with assumptions (2.4) and (2.5) we can set the constants c and $\bar{c}_1$ to

$$c = 18.53 \cdot n^{\frac{3}{2}} \quad \text{and} \quad \bar{c}_1 = 1.853 \cdot n^{\frac{3}{2}} = 0.1 \cdot c. \tag{2.6}$$

The value of $\bar{c}_1$ is given explicitly by Björck (1967). The details on the calculus of the constant c is given in Appendix A. It is worth noticing that the value of c depends only on n, the number of vectors to be orthogonalized, and not on m, the size of the vectors, since (2.4) holds.

Assumption (2.5) prevents $\bar{\mathbf{R}}$ being singular. Under this assumption and defining

$$\eta = \frac{1}{1 - cu\kappa(\mathbf{A})}, \tag{2.7}$$

Björck & Paige (1992) obtain an upper bound for $\|\bar{\mathbf{R}}^{-1}\|_2$ as

$$\|\mathbf{A}\|_2\|\bar{\mathbf{R}}^{-1}\|_2 \leqslant \eta\kappa(\mathbf{A}). \tag{2.8}$$

Assuming (2.5), we note that (2.1) and (2.3) are independent of $\kappa(\mathbf{A})$. This is not the case for inequality (2.2): the level of orthogonality in $\bar{\mathbf{Q}}$ is dependent on $\kappa(\mathbf{A})$. If $\mathbf{A}$ is well-conditioned then $\bar{\mathbf{Q}}$ is orthogonal to machine precision. But for an ill-conditioned matrix $\mathbf{A}$, the set of vectors $\bar{\mathbf{Q}}$ may lose orthogonality. An important question that arises then is whether MGS manages to preserve the full rank of $\bar{\mathbf{Q}}$ or not. In order to investigate this, we study in the next section the condition number of $\bar{\mathbf{Q}}$. For this purpose, we define the singular values of $\bar{\mathbf{Q}}$, $\sigma_1(\bar{\mathbf{Q}}) \geqslant \cdots \geqslant \sigma_n(\bar{\mathbf{Q}})$. When $\bar{\mathbf{Q}}$ is nonsingular, $\sigma_n(\bar{\mathbf{Q}}) > 0$, we also define the condition number $\kappa(\bar{\mathbf{Q}}) = \sigma_1(\bar{\mathbf{Q}})/\sigma_n(\bar{\mathbf{Q}})$.

3. Conditioning of the set of vectors $\bar{\mathbf{Q}}$

This section is fully devoted to the key theorem of this paper and to its proof. For the sake of completeness, we establish a similar result using different arguments in the next section. The central theorem is the following.

THEOREM 3.1 Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a matrix with full rank $n \leqslant m$ and condition number $\kappa(\mathbf{A})$ such that

$$2.12 \cdot (m+1)u < 0{\cdot}01 \quad \text{and} \quad cu\kappa(\mathbf{A}) \leqslant 0.1, \tag{3.1}$$

where $c = 18.53 \cdot n^{\frac{3}{2}}$ and u is the unit round-off. Then MGS in floating-point arithmetic computes $\bar{\mathbf{Q}} \in \mathbb{R}^{m\times n}$ as

$$\kappa(\bar{\mathbf{Q}}) \leqslant 1.3. \tag{3.2}$$

Note that assumption (3.1) is just slightly stronger than assumption (2.5) made by Björck & Paige (1992).

Proof. On the one hand, MGS computes $\bar{\mathbf{Q}}$, on the other hand, the matrix $\hat{\mathbf{Q}}$ has exactly orthonormal columns. It seems natural to study the distance between $\bar{\mathbf{Q}}$ and $\hat{\mathbf{Q}}$. For that we define $\mathbf{F}$ as

$$\mathbf{F} = \bar{\mathbf{Q}} - \hat{\mathbf{Q}}, \tag{3.3}$$

and look at its 2-norm. For this purpose, we subtract (2.3) from (2.1) to get

$$\begin{aligned}(\bar{\mathbf{Q}} - \hat{\mathbf{Q}})\bar{\mathbf{R}} &= \mathbf{A} + \bar{\mathbf{E}} - \mathbf{A} - \hat{\mathbf{E}},\\ \mathbf{F}\bar{\mathbf{R}} &= \bar{\mathbf{E}} - \hat{\mathbf{E}}.\end{aligned}$$

Assuming $cu\kappa(\mathbf{A}) < 1$, $\bar{\mathbf{R}}$ is nonsingular and we can write

$$\mathbf{F} = (\bar{\mathbf{E}} - \hat{\mathbf{E}})\bar{\mathbf{R}}^{-1}.$$

We bound, in terms of norms, this equality

$$\|\mathbf{F}\|_2 \leqslant (\|\bar{\mathbf{E}}\|_2 + \|\hat{\mathbf{E}}\|_2)\|\bar{\mathbf{R}}^{-1}\|_2.$$

Using inequality (2.1) on $\|\bar{\mathbf{E}}\|_2$ and inequality (2.3) on $\|\hat{\mathbf{E}}\|_2$, we obtain

$$\|\mathbf{F}\|_2 \leqslant (c + \bar{c}_1)u\|\mathbf{A}\|_2\|\bar{\mathbf{R}}^{-1}\|_2.$$

Using inequality (2.8) on $\|\mathbf{A}\|_2\|\bar{\mathbf{R}}^{-1}\|_2$ and (2.6), we have

$$\|\mathbf{F}\|_2 \leqslant 1.1 \cdot cu\eta\kappa(\mathbf{A}). \tag{3.4}$$

This is the desired bound on $\|\mathbf{F}\|_2$.

Since we are interested in an upper bound on $\kappa(\bar{\mathbf{Q}})$, the condition number of $\bar{\mathbf{Q}}$, we then look for an upper bound for the largest singular value of $\bar{\mathbf{Q}}$ and a lower bound for its smallest singular value.

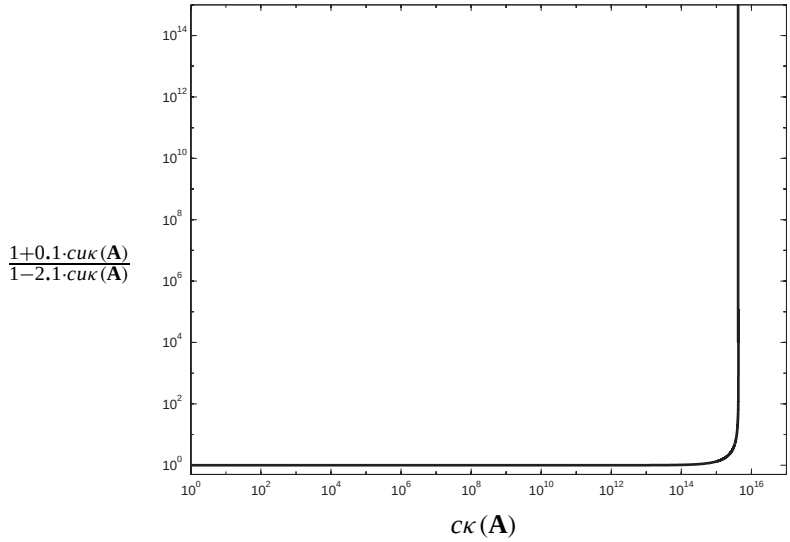

FIG. 1. Behaviour of the upper bound on $\kappa(\bar{\mathbf{Q}})$ as a function of $c\kappa(\mathbf{A})$.

From Golub & Van Loan (1983, p. 449), we know that (3.3) implies

$$\sigma_1(\bar{\mathbf{Q}}) \leqslant \sigma_1(\hat{\mathbf{Q}}) + \|\mathbf{F}\|_2 \quad \text{and} \quad \sigma_n(\bar{\mathbf{Q}}) \geqslant \sigma_n(\hat{\mathbf{Q}}) - \|\mathbf{F}\|_2.$$

Since $\hat{\mathbf{Q}}$ has exactly orthonormal columns, we have $\sigma_1(\hat{\mathbf{Q}}) = \sigma_n(\hat{\mathbf{Q}}) = 1$. Using the bound (3.4) on $\|\mathbf{F}\|_2$, we get

$$\sigma_1(\bar{\mathbf{Q}}) \leqslant 1 + 1.1 \cdot cu\eta\kappa(\mathbf{A}) \quad \text{and} \quad \sigma_n(\bar{\mathbf{Q}}) \geqslant 1 - 1.1 \cdot cu\eta\kappa(\mathbf{A}).$$

With (2.7), these inequalities can be written as

$$\sigma_1(\bar{\mathbf{Q}}) \leqslant \eta(1 - cu\kappa(\mathbf{A}) + 1.1 \cdot cu\kappa(\mathbf{A})) = \eta(1 + 0.1 \cdot cu\kappa(\mathbf{A}))$$

and

$$\sigma_n(\bar{\mathbf{Q}}) \geqslant \eta(1 - cu\kappa(\mathbf{A}) - 1.1 \cdot cu\kappa(\mathbf{A})) = \eta(1 - 2.1 \cdot cu\kappa(\mathbf{A})).$$

If we assume

$$2.1 \cdot cu\kappa(\mathbf{A}) < 1, \tag{3.5}$$

$\sigma_n(\bar{\mathbf{Q}}) > 0$ so $\bar{\mathbf{Q}}$ is nonsingular.

Under this assumption, we have

$$\kappa(\bar{\mathbf{Q}}) \leqslant \frac{1 + 0.1 \cdot cu\kappa(\mathbf{A})}{1 - 2.1 \cdot cu\kappa(\mathbf{A})}. \tag{3.6}$$

To illustrate the behaviour of the upper bound of $\kappa(\bar{\mathbf{Q}})$, we plot in Fig. 1 the upper bound as a function of $c\kappa(\mathbf{A})$. We fix $u = 1.12 \cdot 10^{-16}$.

It can be seen that this upper bound explodes when $2.1 \cdot cu\kappa(\mathbf{A}) \lesssim 1$ but in the main part of the domain where $2.1 \cdot cu\kappa(\mathbf{A}) < 1$ it is small and very close to one. For instance, if we slightly increase the constraint (2.5) used by Björck & Paige (1992) and assume that $cu\kappa(\mathbf{A}) < 0.1$ then $\kappa(\bar{\mathbf{Q}}) < 1.3$.

4. Some remarks

4.1 *Another way to establish a similar result as Theorem 3.1*

It is also possible to get a bound on $\kappa(\bar{\mathbf{Q}})$ by using inequality (2.2). In this aim, we need explicitly the constant $\bar{c}_2$ given by Björck & Paige (1992). Using assumptions (2.4) and (3.1), $\bar{c}_2$ can be set to

$$\bar{c}_2 = 31.6863 \cdot n^{\frac{3}{2}} = 1.71 \cdot c. \tag{4.1}$$

The details on the calculus of the constant $\bar{c}_2$ are given in Appendix A. Let $\bar{\mathbf{Q}}$ have the polar decomposition $\bar{\mathbf{Q}} = \mathbf{UH}$. The matrix $\mathbf{U}$ is the closest orthonormal matrix to $\bar{\mathbf{Q}}$ in any unitarily invariant norm. We define

$$\mathbf{G} = \bar{\mathbf{Q}} - \mathbf{U}.$$

From Higham (1994), we know that in 2-norm the distance from $\bar{\mathbf{Q}}$ to $\mathbf{U}$ is bounded by $\|\mathbf{I} - \bar{\mathbf{Q}}^T\bar{\mathbf{Q}}\|_2$. This means

$$\|\mathbf{G}\|_2 = \|\bar{\mathbf{Q}} - \mathbf{U}\|_2 \leqslant \|\mathbf{I} - \bar{\mathbf{Q}}^T\bar{\mathbf{Q}}\|_2$$

and using (2.2) we get

$$\|\mathbf{G}\|_2 \leqslant \bar{c}_2 u\kappa(\mathbf{A}) = 1.71 \cdot cu\kappa(\mathbf{A}). \tag{4.2}$$

Using the same arguments as in Section 3 for the proof of Theorem 3.1, but replacing (3.4) with (4.2), we get a similar result: that is

$$\text{assuming (2.4) and (3.1)}, \quad \kappa(\bar{\mathbf{Q}}) < 1.42.$$

This result should be compared with that of Theorem 3.1. With the same assumptions, we obtain a slightly weaker result.

4.2 *Iterative modified Gram–Schmidt*

If the assumption (3.1) on the condition number of $\mathbf{A}$ holds, then we obtain, after a first sweep of MGS, $\bar{\mathbf{Q}}_1$ satisfying (3.6). If we run MGS a second time on $\bar{\mathbf{Q}}_1$ to obtain $\bar{\mathbf{Q}}_2$, we deduce using (2.2) that $\bar{\mathbf{Q}}_2$ is such that

$$\|\mathbf{I} - \bar{\mathbf{Q}}_2^T\bar{\mathbf{Q}}_2\|_2 \leqslant 1.71 \cdot c\kappa(\bar{\mathbf{Q}}_1)u,$$

so we get

$$\|\mathbf{I} - \bar{\mathbf{Q}}_2^T\bar{\mathbf{Q}}_2\|_2 < 40.52 \cdot un^{\frac{3}{2}}, \tag{4.3}$$

meaning that $\bar{\mathbf{Q}}_2$ has columns orthonormal to machine precision. Two MGS sweeps are indeed enough to have an orthonormal set of vectors $\mathbf{Q}$.

We recover, in a slightly different framework, the famous sentence of Kahan: *'Twice is enough.'* Based on unpublished notes of Kahan, Parlett (1980) shows that an iterative Gram–Schmidt process on two vectors with a selective criterion (optional) produces two vectors orthonormal up to machine precision. In this paper, inequality (4.3) show that *twice is enough* for n vectors under assumptions (2.4) and (3.1) with MGS and a complete *a posteriori* re-orthogonalization (i.e. no selective criterion).

4.3 *What can be said on $\kappa(\bar{\mathbf{Q}})$ when $cu\kappa(\mathbf{A}) > 0.1$*

For $2.1 \cdot cu\kappa(\mathbf{A}) < 1$, the bound (3.6) on $\kappa(\bar{\mathbf{Q}})$ is well defined but when $cu\kappa(\mathbf{A}) > 0{\cdot}1$, this bound explodes and very quickly nothing interesting can be said about the condition number of $\bar{\mathbf{Q}}$. For $2.1 \cdot cu\kappa(\mathbf{A}) > 1$, we even do not have any bound.

Here, we ask whether or not there can exist an interesting upper bound on $\bar{\mathbf{Q}}$ when $cu\kappa(\mathbf{A}) > 0.1$. In order to answer this problem, we consider the matrix *CERFACS* $\in \mathbb{R}^{3\times 3}$. (See Appendix B.)

When we run MGS with Matlab on *CERFACS*, we obtain with $u = 1.12 \cdot 10^{-16}$

$$\kappa(\mathbf{A}) = 3 \cdot 10^{15}, \quad cu\kappa(\mathbf{A}) = 37 \quad \text{and} \quad \kappa(\bar{\mathbf{Q}}) = 2 \cdot 10^{14}.$$

Matrix *CERFACS* generates a very ill-conditioned set of vectors $\bar{\mathbf{Q}}$ with $cu\kappa(\mathbf{A})$ not too far from 0.1.

If we are looking for an upper bound of $\kappa(\bar{\mathbf{Q}})$, we can take the value 1.3 up to $cu\kappa(\mathbf{A}) = 0.1$ and then this upper bound has to be greater than $2 \cdot 10^{14}$ for $cu\kappa(\mathbf{A}) = 37$.

Matrix *CERFACS* proves that it is not possible to increase by much the domain of validity (i.e. $cu\kappa(\mathbf{A}) < 0{\cdot}1$) of Theorem (3.1) in order to get a more interesting result.

One can also remark that with *CERFACS* two MGS sweeps are no longer enough since

$$\|\mathbf{I} - \bar{\mathbf{Q}}_2^T \bar{\mathbf{Q}}_2\|_2 = 2 \cdot 10^{-3}.$$

Acknowledgements

We would like to thank Miroslav Rozložník for fruitful discussions on the Modified Gram–Schmidt algorithm and in particular for having highlighted that the sentence *twice is enough* required the assumption of a not 'too ill-conditioned' matrix $\mathbf{A}$. We also thank the anonymous referees for their comments that helped to improve the paper. The work of J.L. was supported by EADS, Corporate Research Centre, Toulouse.

REFERENCES

BJÖRCK, Å. (1967) Solving linear least squares problems using Gram–Schmidt orthogonalization. *BIT*, **7**, 1–21.

BJÖRCK, Å. & PAIGE, C. C. (1992) Loss and recapture of orthogonality in the modified Gram–Schmidt Algorithm. *SIAM J. Matrix Analysis and Applications*, **13**, 176–190.

GOLUB, G. H. & VAN LOAN, C. F. *Matrix Computations*. Baltimore, MA: Johns Hopkins University Press.

HIGHAM, N. J. (1994) The matrix sign decomposition and its relation to the polar decomposition. *Linear Algebra and its Applications*, **212/213**, 3–20.
PARLETT, B. N. (1980) *The Symmetric Eigenvalue Problem*. Englewood Cliffs, NJ: Prentice-Hall.
WILKINSON, J. H. (1965) *The Algebraic Eigenvalue Problem*. Oxford: Oxford University Press.

Appendix A. Details on the calculus of the constants

In this Appendix, we justify the values of the constants $\bar{c}_1$, $\bar{c}_2$ and c as fixed in the paper. We state that
$\bar{c}_1 = 1.853 \cdot n^{\frac{3}{2}}$ verifies (2.1) under assumption (2.4),
$\bar{c}_2 = 31.6863 \cdot n^{\frac{3}{2}}$ verifies (2.2) under assumptions (2.4) and (3.1),
$c = 18.53 \cdot n^{\frac{3}{2}}$ verifies (2.3) under assumptions (2.4) and (2.5).

A value for $\bar{c}_1$. Under assumption (2.4) Björck (1967) has shown that

$$\mathbf{A} + \bar{\mathbf{E}} = \bar{\mathbf{Q}}\bar{\mathbf{R}} \quad \text{with} \quad \|\bar{\mathbf{E}}\|_E \leqslant 1.5 \cdot (n-1)u\|\mathbf{A}\|_E.$$

where $\|\cdot\|_E$ denotes the Frobenius norm.

$\bar{c}_1 = 1.853 \cdot n^{\frac{3}{2}}$ verifies $\|\bar{\mathbf{E}}\|_E \leqslant \bar{c}_1 u\|\mathbf{A}\|_2$.

A value for c. Björck & Paige (1992) explained that the sequence of operations to obtain the R-factor with the MGS algorithm applied on $\mathbf{A}$ is exactly the same as the sequence of operations to obtain the R-factor with the Householder process applied on the augmented matrix $\begin{pmatrix}\mathbf{0}_n \\ \mathbf{A}\end{pmatrix} \in \mathbb{R}^{(m+n)\times n}$. They deduce that the R-factor from the Householder process applied on the augmented matrix is equal to $\bar{\mathbf{R}}$. We first present the results from Wilkinson (1965) related to the Householder process on the matrix $\begin{pmatrix}\mathbf{0}_n \\ \mathbf{A}\end{pmatrix} \in \mathbb{R}^{(m+n)\times n}$. Wilkinson (1965) works with a square matrix but in the case of a rectangular matrix, proofs and results remain the same. All the results of Wilkinson holds under the assumption $(m+n)\cdot u < 0.1$ which is true because of (2.4).

Defining $x = 12.36 \cdot u$, Wilkinson proves that there exists $\mathbf{P} \in \mathbb{R}^{(m+n)\times n}$ with orthonormal columns such that

$$\|\mathbf{P}\bar{\mathbf{R}} - \mathbf{A}\|_E \leqslant (n-1)(1+x)^{n-2}x\|\mathbf{A}\|_E. \tag{A1}$$

With assumption (2.4), we get $(1+x)^{n-2} \leqslant 1.060053$.

Let as define $\mathbf{E}_1 \in \mathbb{R}^{n\times n}$ and $\mathbf{E}_2 \in \mathbb{R}^{m\times n}$ by

$$\begin{pmatrix}\mathbf{E}_1 \\ \mathbf{E}_2\end{pmatrix} = \mathbf{P}\bar{\mathbf{R}} - \begin{pmatrix}\mathbf{0}_n \\ \mathbf{A}\end{pmatrix},$$

We deduce with (A1) that

$$\left\|\begin{pmatrix}\mathbf{E}_1 \\ \mathbf{E}_2\end{pmatrix}\right\|_E \leqslant 13.1023 \cdot n^{\frac{3}{2}}u\|\mathbf{A}\|_2. \tag{A2}$$

If we set

$$c_1 = c_2 = 13.1023 \cdot n^{\frac{3}{2}}, \tag{A3}$$

then we get $\|E_1\|_2 \leqslant c_1 u \|\mathbf{A}\|_2$ and $\|E_2\|_2 \leqslant c_2 u \|\mathbf{A}\|_2$.

Note that we also have

$$\|\mathbf{E}_1\|_2 + \|\mathbf{E}_2\|_2 \leqslant \sqrt{2} \left\| \begin{pmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{pmatrix} \right\|_E \leqslant \sqrt{2} c_1 u \|\mathbf{A}\|_2. \tag{A4}$$

With respect to MGS, Björck & Paige (1992) have proved that there exists $\hat{\mathbf{E}} \in \mathbb{R}^{m \times n}$ and $\hat{\mathbf{Q}} \in \mathbb{R}^{m \times n}$ such that

$$\mathbf{A} + \hat{\mathbf{E}} = \hat{\mathbf{Q}} \bar{\mathbf{R}}, \quad \hat{\mathbf{Q}}^T \hat{\mathbf{Q}} = \mathbf{I} \quad \text{and} \quad \|\hat{\mathbf{E}}\|_2 \leqslant \|\mathbf{E}_1\|_2 + \|\mathbf{E}_2\|_2.$$

With (A4) we get

$$\|\hat{\mathbf{E}}\|_2 \leqslant \|\mathbf{E}_1\|_2 + \|\mathbf{E}_2\|_2 \leqslant \sqrt{2} c_1 u \|\mathbf{A}\|_2 \leqslant 18.53 \cdot n^{\frac{3}{2}},$$

and $c = 18.53 \cdot n^{\frac{3}{2}}$ verifies $\|\hat{\mathbf{E}}\|_2 \leqslant cu\|\mathbf{A}\|_2$.

A value for $\bar{c}_2$. Björck (1967) defines a value for $\bar{c}_2$. In this paper, we do not consider this value because the assumptions on n and $\kappa(\mathbf{A})$ that we obtain are too restricted. The value of $\bar{c}_2$ from Björck & Paige (1992) requires weaker assumptions that fit the context of this paper. From (3.1), we have $(c + c_1)u\kappa < 1$. Under this assumption, Björck & Paige (1992) have proved that

$$\|\mathbf{I} - \bar{\mathbf{Q}}^T \bar{\mathbf{Q}}\|_2 \leqslant \frac{2c_1}{1 - (c + c_1)u\kappa} \kappa u. \tag{A5}$$

With $\bar{c}_2 = 31.6863 \cdot n^{\frac{3}{2}}$ and using assumption (3.1), we have $\|\mathbf{I} - \bar{\mathbf{Q}}^T \bar{\mathbf{Q}}\|_2 \leqslant \bar{c}_2 \kappa u$.

Appendix B. Matrix *CERFACS*

We have developed a Matlab code that generates as many as desired matrices with relatively small $cu\kappa(\mathbf{A})$ and large $\kappa(\bar{\mathbf{Q}})$. *CERFACS* is one of these:

$$\textit{CERFACS} = \begin{pmatrix} 0.12100300219993308 & 2.09408775152625060 & 1.26139640819301024 \\ -0.10439395064078592 & -1.80665016070527140 & -1.08825526624380808 \\ 0.21661355806776747 & 0.49451660567698374 & -0.84174336538575500 \end{pmatrix}. \tag{B1}$$