CERFACS **Technical Report No. TR/PA/02/52**

Luc Giraud and Julien Langou

ROBUST SELECTIVE GRAM-SCHMIDT REORTHOGONALIZATION

Abstract. A new criterion for selective reorthogonalization in the Gram-Schmidt procedure is given. We establish its comportment in presence of rounding errors when the criterion is used with modified Gram-Schmidt algorithm and show counter-example matrices which prove that standard criteria are not always valid. Experimentally, our criterion is fine also for the classical Gram-Schmidt algorithm with reorthogonalization.
AMS Subject Classification : 65F25, 65G50, 15A23.

1. Introduction

Let $A=(a_1,\dots,a_n)$ be a real $m\times n$ matrix whose columns are linearly independent and $\kappa(A)$ be its condition number. In many applications, it is required to have an orthonormal basis for the space spanned by the columns of A. This amounts to knowing a matrix $Q\in\mathbb{R}^{m\times n}$ with orthonormal columns such that $A=QR$, $R\in\mathbb{R}^{n\times n}$. Moreover it is possible to impose R being triangular, when this is the case one speak of QR-factorization. For all j, the first j columns of Q are an orthonormal basis for the space spanned by the first j columns of A.

Starting from A, there is many algorithms to get such a factorization. We focus in this paper on the Gram-Schmidt algorithm [1] which consists in projecting successively the columns of A on the space orthogonal to the space spanned by the already constructed columns of Q. Depending how the projections are made, there are two versions of this algorithm [2]: the classical Gram-Schmidt algorithm (CGS) and the modified Gram-Schmidt algorithm (MGS). In exact arithmetic, both algorithms produce exactly the same results and the resulting matrix Q has its columns orthonormal. In presence of round-off errors, the computed Q by CGS differs a lot from the one computed by MGS but in both cases the columns of Q may be poorly orthogonal. To remedy this problem, a solution is to project, instead of one time, a few times each column of A on the space orthogonal to the space spanned by the constructed columns of Q. Giraud, Langou and Rozložník [15] have shown that either for CGS or MGS, two projections were enough when the initial matrix A is numerically nonsingular. This confirms what was already experimentally well-known and generalized the result with $n=2$ vectors of Kahan and Parlett [10]. In this paper, we focus only on Gram-Schmidt algorithms where the number of projections for each column of A is either one or two. We call the resulting algorithm classical (resp. modified) Gram-Schmidt algorithm with reorthogonalization (CGS2, resp. MGS2) or Gram-Schmidt algorithm with reorthogonalization (GS2) when we speak about both versions.

In one hand, GS2 gives a high quality result, in the other hand the computational cost of GS2 is twice as much as the one of GS. In many applications, we observe that GS is enough, the additionnal reorthogonalizations done by GS2 are useless. A good compromise in term of quality and time between GS and GS2 is to use a selective criterion that checks for each columns of A whether reorthogonalization is needed or not. To be interesting, this selective criterion has to be cheap to compute. Historically, the first criterion

---

Luc Giraud: CERFACS, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex 1, France. e-mail: `Luc.Giraud@cerfacs.fr`
Julien Langou: CERFACS, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex 1, France. e-mail: `Julien.Langou@cerfacs.fr`

introduced in a Gram-Schmidt algorithm is by Rutishauser [4]. We refer to it as the K-criterion because it is dependent of a single parameter $K \geq 1$ and refer to the resulting algorithm as GS2(K). Other criterion like super-orthogonalization [8] have also been tested but the K-criterion is the most commonly used reorthogonalization criterion. We give below the modified Gram-Schmidt algorithm with reorthogonalization and K-criterion (MGS2(K)). Kahan and Parlett [10] have shown that for two vectors the orthogonality obtained ($|q_1^T q_2|$) is bounded by a term proportionnal to the parameter K and the machine precision denoted by ε. Intuitively this is natural: if K is high then we have few reorthogonalizations, if K is small then more reorthogonalization are performed and the orthogonality of the obtained set of vectors Q is better. For n vectors, the choice of the parameter K is not so clear. Giraud and al. [15] show that if K is greater than the condition number of A, $\kappa(A)$, then GS2($K = \kappa(A)$) does not perform any reorthogonalization. Interesting values for K are therefore ranging from 1 (this corresponds to GS2) to $\kappa(A)$ (this corresponds to GS). Initially, in 1967, Rutishauser [4] chose the value $K = 10$. We find an explanation of this value in Gander [9, p. 12]: "*in particular one may state the rule of thumb that at least one decimal digit is lost by cancellation if* $10\|a_j^{(1)}\|_2 \leq \|a_j\|_2$. *This equation is the criterion used by Rutishauser to decide whether reorthogonalization is necessary.*" The value $K = \sqrt{2}$ is also often used since the publication of the paper of Daniel, Gragg, Kaufman and Stewart [6] (e.g. by Ruhe [11] or by Reichel and Gragg [13]). More exotic values like $K = 100.05$ [8] or $K = \sqrt{5}$ [14] have also been implemented. In 1989, Hoffmann [12] tested a wide range of value $K = 2, 10, \ldots, 10^{10}$. The conclusion of his experiments is that, for what he observed, the resulting orthogonality is proportional to the parameter K and the machine precision, exactly as for 2 vectors.

Algorithm 1 MGS2(K)

```
for $j = 1$ to $n$ do
  $a_j^{(1)} = a_j$
  for $k = 1$ to $j-1$ do
    $r_{kj}^{(1)} = q_k^T a_j^{(1)}$
    $a_j^{(1)} = a_j^{(1)} - q_k r_{kj}^{(1)}$
  end for
  if $K_1^{(j)} = \frac{\|a_j\|_2}{\|a_j^{(1)}\|_2} \leq K$ then
    $r_{jj} = \|a_j^{(1)}\|_2$
    $q_j = a_j^{(1)} / r_{jj}$
    $r_{kj} = r_{kj}^{(1)}, 1 \leq k \leq j-1$
  else
    $a_j^{(2)} = a_j^{(1)}$
    for $k = 1$ to $j-1$ do
      $r_{kj}^{(2)} = q_k^T a_j^{(2)}$
      $a_j^{(2)} = a_j^{(2)} - q_k r_{kj}^{(2)}$
    end for
    $r_{jj} = \|a_j^{(2)}\|_2$
    $q_j = a_j^{(2)} / r_{jj}$
    $r_{kj} = r_{kj}^{(1)} + r_{kj}^{(2)}, 1 \leq k \leq j-1$
  end if
end for
```

The goal of this paper is to give new ideas on the subject of selective reorthogonalization. In Section 2, we derive a new criterion, depending on a single parameter L, that we call the L-criterion. The Gram-Schmidt algorithm with reorthogonalization and L-criterion is denoted by GS2(L). We prove that under numerical nonsingularity assumption of A if the paramater L is lower than 1 then the algorithm MGS2(L) gives a matrix Q orthogonal to machine precision. To justify the fact that the statement $L < 1$ is really needed, in Section 3, we develop counter-example matrices for which when $L > 1$ the algorithm GS(L) gives a set of vectors Q far from being orthogonal . Finally, in Section 4, we give other counter-example matrices. These latter matrices have the property that the Gram-Schmidt algorithm with reorthogonalization and K-criterion fails to give an acccurately orthogonal basis for any $K > 1$.

2. A new selective orthogonalization criterion.

2.1. A second proof for MGS2

In [15], defining $\zeta(m,n)$, $\psi(m,n)$ and $\phi(m,n)$ low polynomial in m and n depending on the details of the arithmetic, we have shown that if we assume

$$\zeta(m,n)\varepsilon\kappa(A) < 1, \tag{2.1}$$

then the set of vectors Q generated by MGS2 is such that

$$\|I - Q^T Q\| \leq 2\sqrt{2}n\phi(m,n)\varepsilon. \tag{2.2}$$

This means that if A is numerically nonsingular (2.1) then the matrix Q given by MGS2 is orthogonal to machine precision (2.2).

Unfortunately the proof used in [15] for GS2 can not be extended to an algorithm using a selective orthogonalization criterion like GS2(K). In this Section, we explain how to derive another proof for MGS2. This is a tedious work that is fully developped in a working note [16]. We combine the approaches of Björck [3] with the result of Giraud and al. [15].

Defining U the strictly upper part of $Q^T Q$, Björck shows that for MGS, we have the equality

$$UR = S, \tag{2.3}$$

where R is the computed R-factor in the QR-factorization and E is such that

$$\|s_j\|_2 \leq \xi(m,n)\varepsilon\|a_j\|_2. \tag{2.4}$$

$\xi(m,n)$ being a low polynomial in m and n depending on the details of the arithmetic. From this Björck derives the well-known formula

$$\|I - Q^T Q\|_2 \leq 2\xi(m,n)\varepsilon\|A\|_E\|R^{-1}\|_2,$$

which means that the loss of orthogonality is proportionnal to the condition number of A times the machine precision ε.

We remark that the Equation (2.3) fits for the second loop of MGS2. Normalizing the Equation (2.3), we get

$$UM = T \tag{2.5}$$

where M is the matrix with a unit upper triangular matrix such as the (k,j) component m_{kj} is

$$\begin{cases} m_{kj} = \frac{r_{kj}^{(2)}}{r_{jj}^{(2)}}, \text{ if } k < j, \\ m_{jj} = 1, \\ m_{kj} = 0, \quad \text{if } k < j, \end{cases}$$

and T is such that

$$\|t_j\|_2 \leq \xi(m,n)\varepsilon\|a_j^{(1)}\|_2/r_{jj}^{(2)}. \tag{2.6}$$

In [15], it is shown that if we assume

$$\sqrt{n}\psi(m,n)\varepsilon\kappa(A) < 1 \quad \text{and} \quad \frac{3n^{3/2}\omega(m,n)\varepsilon\kappa(A)}{\sqrt{1-n\psi^2(m,n)\varepsilon^2\kappa^2(A)}} \leq c < 1, \tag{2.7}$$

for the second loop we have

$$\|a_j^{(1)}\|_2/r_{jj}^{(2)} \leq \left[1 - \frac{3n^{3/2}\omega(m,n)\varepsilon\kappa(A)}{\sqrt{1-n\psi^2(m,n)\varepsilon^2\kappa^2(A)}}\right]^{-1}. \tag{2.8}$$

Therefore

$$\|T\|_2 \leq c_1\xi(m,n)\sqrt{n}\varepsilon.$$

Meaning that the 2-norm of T is close to machine precision.
From (2.5), as M has full rank, we have

$$U = TM^{-1}.$$

It remains to show that M is well-conditioned. For this we use the decomposition $M = I + B$ with B nilpotent, we have the well-known formula (e.g. [7])

$$\|M^{-1}\|_1 \leq (1 - \|B\|_1)^{-1}. \tag{2.9}$$

From arguments of [15], at the second loop, if A is numerically nonsingular, we have

$$\|B\|_1 = \max_j \sum_{k=1}^{j-1} \frac{|r_{kj}^{(2)}|}{r_{jj}^{(2)}} \leq L < 1 \tag{2.10}$$

where L is a constant independent of the problem and therefore

$$\|I - Q^TQ\|_2 \leq \frac{2}{1-L}c_1\xi(m,n)\sqrt{n}\varepsilon. \tag{2.11}$$

We recover a result developped in [15]: the set of vectors computed by MGS2 is orthonormal up to machine precision.

2.2. Link with selective reorthogonalization

The key property of the matrix M is (2.10). In term of column, for each $j = 1, \ldots, n$ we have after the reorthogonalization loop

$$L_j^{(2)} = \sum_{k=1}^{j-1} \frac{|r_{kj}^{(2)}|}{r_{jj}^{(2)}} \leq L < 1.$$

However this property may occur after the first orthogonalization that is to say

$$L_j^{(1)} = \sum_{k=1}^{j-1} \frac{|r_{kj}^{(1)}|}{\|a_j^{(1)}\|_2} \leq L < 1 \tag{2.12}$$

In this case, it is clear that if we do not reorthogonalize $a_j^{(1)}$ the proof is not changed and the result (2.11) holds.

Naturally we propose the L-criterion define by (2.12). The resulting algorithm is called GS2(L) and is exactly the same as Algorithm 1 except that the line 7 is changed by:

$$\textbf{if} \quad \frac{\sum_{k=1}^{j-1} |r_{kj}^{(1)}|}{\|a_j^{(1)}\|_2} \leq L \quad \textbf{then}$$

The parameter L may be any real value greater than 0. The lower L is the more reorthogonalization are performed. We have seen that the algorithm MGS2(L) gives good results for $L < 1$. If one want an orthogonal Q and perform less reothogonalization than MGS2 an optimal value for L seems to be 0.99.

Assuming that $\sum_{k=1}^{j-1}(r_{kj}^{(1)})^2 + \|a_j^{(1)}\|_2^2 = \|a_j\|_2^2$ (which corresponds to the theorem of Pythagorus if Q_{j-1} has orthogonal columns), we can rewrite the K-criterion as

$$\frac{\sqrt{\sum_{k=1}^{j-1}(r_{kj}^{(1)})^2}}{\|a_j^{(1)}\|_2} \leq \sqrt{K^2 - 1}. \tag{2.13}$$

The K-criterion consists in compare the 2-norm of the non-diagonal entries $r_{kj}^{(1)}$, $k \neq j$, to the diagonal entry $\|a_j^{(1)}\|_2$. The L criterion consists in compare the 1-norm of the non-diagonal entries $r_{kj}^{(1)}$, $k \neq j$, to the diagonal entry $\|a_j^{(1)}\|_2$. The value $L = 1$ for the L criterion is to relate to the value $K = \sqrt{L^2+1} = \sqrt{2}$ for the K-criterion. Clearly, GS2($L = 1$) performs systematically more reorthogonalization than GS2($K = \sqrt{2}$).

3. The case $L > 1$

We want to find counter-example matrices such that for each fixed $L > 1$ and fixed η the algorithm MGS2(K) applied on A gives $\bar{Q}$ so as

$$\|I - \bar{Q}^T\bar{Q}\|_2 > \eta. \tag{3.1}$$

This means that for $L > 1$, the orthogonality obtained can be as bad as desired. We restrict our study to $\eta \leq 1$.

Let us define the matrix $A(n,\alpha) \in \mathbb{R}^{n\times n}$ such that

$$A(n,\alpha) = Q \begin{pmatrix} \alpha & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha \end{pmatrix}$$

where $Q \in \mathbb{R}^{n\times n}$ is such that $Q^TQ = I$.

For all α, $0 < \alpha < 1$,the condition number of this matrix $\kappa(A)$ can be set arbitrarily large by choosing an appropriate n, in fact for these matrices, we have

$$\kappa(A(n,\alpha)) \geq \alpha^{-n}(1-\alpha^2) \tag{3.2}$$

It is also clear that if we apply GS2(L) in exact arithmetic we get

$$L_1^{(j)} = \sum_{k=1}^{j-1} \frac{|r_{kj}^{(1)}|}{\|\bar{a}_j^{(1)}\|_2} = \frac{1}{\alpha}.$$

The goal in this section is to have the algorithm GS2(L) going as worse as possible, therefore if we choose $\alpha > 1/L$ for all column j of $A(n,\alpha)$, the reorthogonalizations are never performed because we always have $L_1^{(j)} = 1/\alpha > L$. As the parameter α is set such that $\alpha < 1$, using Equation (3.2), we increase n, the size of the matrix $A(n,\alpha)$, in order to have a sufficiently ill-conditioned matrix for which GS is unable to give an orthogonal basis.

Note that if $L < 1$ then $\alpha > 1/L > 1$ and then no reorthogonalization is performed. However in this case ($\alpha > 1$) the matrix $A(n,\alpha)$ is well-conditioned for any n.

One can also notice that if we apply GS2(K) in exact arithmetic we get

$$K_1^{(j)} = \frac{\|a_j\|_2}{\|\bar{a}_j^{(1)}\|_2} = \sqrt{1+(\frac{1}{\alpha})^2}$$

and therefore for $K > \sqrt{2}$, there exists α such that no reorthogonalization are performed with GS2(K). E.g. if we take $\alpha = 1.02(K^2-1)^{-1/2}$.

This has been implemented in a MATLAB program where $\varepsilon = 1.12 \cdot 10^{-16}$.

We take $\alpha = 0.98$ and $n = 1500$ with a random unitary matrix Q to obtain : $A(n,\alpha)$. The condition number of the matrix is: $\kappa(A(n,\alpha)) = 7.31 \cdot 10^{14}$. The results are given in Table 1. When $L = 1.05$ or $K = 1.43$, no reorthogonalization is performed, these algorithms are in fact GS apply on $A(n,\alpha)$, $\|I - Q^TQ\|_2$ is far from machine precision. Note that the values in this table for CGS2($L = 1.05$) and CGS2($K = 1.43$) match exactly and in fact corresponds to CGS. The same remark holds for MGS as well. When $L = 0.99$, the criterion permits all the reorthogonalizations, the algorithm is in fact exactly CGS2. As explained in [15], in this case, we observe effectively that Q is orthogonal up to machine precision.

CGS2($L = 1.08$)	$2.33 \cdot 10^0$
MGS2($L = 1.08$)	$2.29 \cdot 10^0$
CGS2($K = 1.43$)	$2.33 \cdot 10^0$
MGS2($K = 1.43$)	$2.29 \cdot 10^0$
CGS2($L = 0.99$)	$3.79 \cdot 10^{-14}$
MGS2($L = 0.99$)	$4.87 \cdot 10^{-14}$

Table 1 : $\|I - Q^TQ\|_2$ for Q obtained with different GS algorithms applied on $A(n = 1500, \alpha = 0.98)$.

4. To conclude with the K-criterion.

We call a diagonal dominant matrix by column a matrix A such that for all j,

$$|a_{jj}| > \sum_{i \neq j} |a_{ij}|. \tag{4.1}$$

In the proof of Section 2, dealing with MGS2(L), the key point is that a diagonal dominant matrix by column is well-conditioned, this result has been shown by [5]. The right hand side of (4.1) is in fact the 1-norm of the column j of A without the diagonal entry a_{jj}.

To apply a similar proof to the K-criterion, by analogy, we are interested into matrix that are *diagonal dominant matrix by column in 2-norm* that is to say

$$|a_{jj}| > \sqrt{\sum_{i \neq j} a_{ij}^2}. \tag{4.2}$$

The problem is that those matrices may be ill-conditioned.
Let us define the matrix $B(n,\alpha) \in \mathbb{R}^{n\times n}$ such as :

$$B(n,\alpha) = QT(n,\alpha) = Q\begin{pmatrix} 1 & -\alpha & -\alpha/\sqrt{2} & -\alpha/\sqrt{3} & & -\alpha/\sqrt{n-1} \\ & 1 & -\alpha/\sqrt{2} & -\alpha/\sqrt{3} & & -\alpha/\sqrt{n-1} \\ & & 1 & -\alpha/\sqrt{3} & & -\alpha/\sqrt{n-1} \\ & & & 1 & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & -\alpha/\sqrt{n-1} \\ & & & & & 1 \end{pmatrix}$$

where $Q \in \mathbb{R}^{n\times n}$ such that $Q^TQ = I$.

For $\alpha < 1$, the unit triangular matrix $T(n,\alpha)$ is *diagonal dominant matrix by column in 2-norm*. A consequence is that if we apply GS2($K \geq \sqrt{1+\alpha^2}$) to $B(n,\alpha)$, no reorthogonalization is performed in exact arithmetic. When n increases, $T(n,\alpha)$ becomes ill-conditioned and therefore $B(n,\alpha)$ also (this means that a result like the one of Varah [5] does not exist if we just assume (4.2)).

The experimental results are in Table 2, we run different versions of Gram-Schmidt with reorthogonalization on a set of matrices $B(n,\alpha)$. The experiments are carry out with MATLAB. With $B(n = 2500, \alpha = 0.30)$, the algorithm GS2($K = 1.05$), either the classical version or the modified one, gives a matrix Q that is far from orthogonal. This means that to guarantee good accuracy K as to be taken lower than 1.05. By diminushing α and increasing n, we guess that it is possible to get lower value for K and to reach nearly 1. As expected we notice that the algorithm GS2($L = 0.99$) behaves well.

Note that if we want to construct such a matrix T that enforces no reorthogonalization with a L-criterion, $L < 1$, then the entry (i, j) of T for $i > j$, are $t_{i,j} = \alpha/(j-1)$ with $\alpha \leq L$. T is such that no reorthogonalization is performed with GS2(L). However, in this case, T is well-conditioned (using (2.9)), MGS is enough to have a well orthogonal basis.

Matrices $B(n,\alpha)$ are also good counter-example for $L > 1$ (see Section 3).

(L,K)	$L = 0.99 \quad K = 1.40$	$L = 0.99 \quad K = 1.30$	$L = 0.99 \quad K = 1.17$	$L = 0.99 \quad K = 1.05$
matrix B	$B(n = 400, \alpha = 0.97)$	$B(n = 500, \alpha = 0.82)$	$B(n = 1000, \alpha = 0.50)$	$B(n = 2500, \alpha = 0.30)$
$\kappa(B)$	$3.4\cdot10^{15}$	$8.6\cdot10^{14}$	$1.8\cdot10^{13}$	$5.9\cdot10^{12}$
CGS2(K)	$1.6\cdot10^{0}$	$1.6\cdot10^{0}$	$1.6\cdot10^{0}$	$1.6\cdot10^{0}$
MGS2(K)	$7.2\cdot10^{-1}$	$1.1\cdot10^{0}$	$1.0\cdot10^{-2}$	$7.6\cdot10^{-3}$
CGS2(L)	$1.2\cdot10^{-14}$	$1.5\cdot10^{-14}$	$2.8\cdot10^{-14}$	$6.0\cdot10^{-14}$
MGS2(L)	$1.5\cdot10^{-14}$	$1.9\cdot10^{-14}$	$3.5\cdot10^{-14}$	$8.0\cdot10^{-14}$

Table 2: $\|I - Q^TQ\|_2$ for Q obtained by different GS algorithms applied to four matrices $B(n,\alpha)$.

5. Conclusion

In this paper, we give a new reorthogonalization criterion for Gram-Schmidt algorithm with reorthogonalization, the L-criterion. This criterion depends on a single parameter L and costs less than n operations to compute. When L is taken smaller than 1 (e.g. $L = 0.99$), this criterion is able to realize the compromise between saving useless reorthogonalization and giving a set of vectors Q orthogonal up to machine precision level when used with the modified Gram-Schmidt algorithm with selective reorthogonalization. In another hand if we set $L > 1$, we have exhibited matrices for which the Gram-Schmidt algorithm with selective reorthogonalization based on the L-criterion (GS2(L)) can perform very badly. In order to justify the need of a new criterion, we also show counter-example matrices for which a standard criterion, the K-criterion, gives a matrix Q far from being orthogonal for any value of the parameter K. On all these counter-example matrices, we have verified the theory and observe that MGS2($L < 1$) behaves well.

We remark that this paper deals with real matrices however the same conclusions can be drawn for complex matrices.

An important point in the proof of Section 2, is that we have assumed the initial matrix A numerically non-singular. In the case of numerically singular matrices, it is clear that one reorthogonalization step may not be enough. The algorithm to be considered is then CGSI(L) that allows as many reorthogonalizations as necessary to satisfy the L-criterion.

An open question is still how the algorithm CGS2($L < 1$) theoretically behaves, the proof of Section 2 dealing only with MGS2($L < 1$). In all the experiments, CGS2($L < 1$) gives good results but nothing has been proved yet. From a practical point of view, CGS is interesting because of the parallel facilities it permits.

Finally we wanted to conclude by a remark on GS2(K). The matrix that are exhibited as counter-example have to be taken for what they are: isolated matrices within the wide set of matrices. It is clear that the K-criterion fails on these pathological matrices (as they were designed for that purpose) but, from having been widely tested, the K-criterion with the parameter K set to $\sqrt{2}$, 2, 10 or something else, is often fine and use for sure less reorthogonalization than GS2($L = 0.99$).

6. Acknowledgements

The work of J. L. was supported by EADS, Corporate Research Centre, Toulouse.

References

1. SCHMIDT E. Über die Auflösung linearer Gleichungen mit unendlich vielen Ubbekannten. *Rend. Circ. Mat. Palermo. Ser. 1, 25 (1908), 53-77.*
2. RICE J. R. Experiments on Gram-Schmidt orthogonalization. *Math. Comp. 20 (1966), 325-328.*
3. BJÖRCK Å. Solving linear least squares problems by Gram-Schmidt orthogonalization. *BIT 7 (1967), 1-21.*
4. RUTISHAUSER H. Description of Algol 60. handbook for automatic computation, Vol. 1a. *Springer Verlag, Berlin, 1967.*
5. VARAH J. M. A lower bound for the smallest singular value of a matrix. *Linear Algebra and its Applications 11 (1975), 3-5.*
6. DANIEL J. W. & GRAGG W. B. & KAUFMAN L. & STEWART G. W. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Math. Comp. 30 (1976), 772-795.*
7. VARGA R. S. On diagonal dominance arguments for bounding $\|A^{-1}\|_\infty$. *Linear Algebra and its Applications 14 (1976), 211-217.*
8. GANDER W. & MOLINARI L. & ŠVECOVÁ H. Numerische Prozeduren aus Nachlass und Lehre von Prof. Heinz Rutishauser. *Birkläuser Verlag, ISNM 33, (1977).*
9. GANDER W. Algorithms for the QR decomposition. *Research report No. 80-02, Eidgenoessische Technische Hochschule, CH-8092 Zuerich, 1980.*
10. PARLETT B. N. The Symmetric Eigenvalue Problem. *Englewood Cliffs, N.J., Prentice-Hall, 1980.*
11. RUHE A. Numerical aspects of Gram-Schmidt orthogonalization of vectors. *Linear Algebra and its Applications 52/53 (1983), 591-601.*
12. HOFFMANN W. Iterative algorithms for Gram-Schmidt orthogonalization. *Computing 41 (1989), 335-348.*
13. REICHEL L. & GRAGG W. B. FORTRAN subroutines for updating the QR decomposition, *ACM Trans. Math. Software, 16 (1990), 369-377.*
14. FRAYSSÉ V. & GIRAUD L. & GRATTON S. A set of GMRES routines for real and complex arithmetics. *Technical Report TR/PA/97/49, CERFACS,Toulouse, France, 1997.*

15. GIRAUD L. & LANGOU J. & ROZLOŽNÍK M. On the round-off error analysis of the Gram-Schmidt algorithm with reorthogonalization. *CERFACS Technical Report, TR/PA/02/33, (2002)*
16. GIRAUD L. & LANGOU J. Another proof for modified Gram-Schmidt with reorthogonalization. *CERFACS Working Note, WN/PA/02/53, (2002)*