Gram-Schmidt process with respect to inner products and bilinear forms: rounding error analysis

Miro Rozložník
joint work with several coauthors mentioned during the talk

Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic

Tenth Workshop on Mathematical Modelling of Environmental and Life Sciences Problems, October 16-19, 2014, Constantza, Romania

STANDARD INNER PRODUCT: GRAM-SCHMIDT PROCESS AS QR ORTHOGONALIZATION

$$A = (a_1, \ldots, a_n) \in \mathcal{R}^{m,n}, \quad m \geq rank(A) = n$$

orthogonal basis Q of $span(A)$

$Q = (q_1, \ldots, q_n) \in \mathcal{R}^{m,n}$, $Q^T Q = I$

$A = QR$, R upper triangular

$$A^T A = R^T R$$

HISTORICAL REMARKS I

origination of the QR factorization used for orthogonalization of functions:
J. P. Gram: Über die Entwicklung reeller Funktionen in Reihen mittelst der Methode der Kleinsten Quadrate. Journal f. r. a. Math., 94: 41-73, 1883.

algorithm of the QR decomposition but still in terms of functions:
E. Schmidt: Zur Theorie der linearen und nichlinearen Integralgleichungen. I Teil. Entwicklung willkürlichen Funktionen nach system vorgeschriebener. Mathematische Annalen, 63: 433-476, 1907.

name of the QR decomposition in the paper on nonsymmetric eigenvalue problem, rumor: the "Q" in QR was originally an "O" standing for orthogonal:
J.G.F. Francis: The QR transformation, parts I and II. Computer Journal 4:265-271, 332-345, 1961, 1962.

HISTORICAL REMARKS II

”modified” Gram-Schmidt (MGS) interpreted as an elimination method using weighted row sums not as an orthogonalization technique:
P.S. Laplace: Theorie Analytique des Probabilités. Courcier, Paris, third edition, 1820. Reprinted in P.S. Laplace. (Evres Compeétes. Gauthier-Vilars, Paris, 1878-1912).

”classical” Gram-Schmidt (CGS) algorithm to solve linear systems of infinitely many solutions:
E. Schmidt: Über die Auflösung linearen Gleichungen mit unendlich vielen Unbekanten, Rend. Circ. Mat. Palermo. Ser. 1, 25 (1908), pp. 53-77.

first application to finite-dimensional set of vectors:
G. Kowalewski: Einfuehrung in die Determinantentheorie. Verlag von Veit & Comp., Leipzig, 1909.

CLASSICAL AND MODIFIED GRAM-SCHMIDT ALGORITHMS

classical (CGS)

for $j = 1, \dots, n$
$u_j = a_j$
for $k = 1, \dots, j-1$

$$u_j = u_j - (a_j, q_k)q_k$$

end
$q_j = u_j / \|u_j\|$
end

modified (MGS)

for $j = 1, \dots, n$
$u_j = a_j$
for $k = 1, \dots, j-1$

$$u_j = u_j - (u_j, q_k)q_k$$

end
$q_j = u_j / \|u_j\|$
end

CLASSICAL AND MODIFIED GRAM-SCHMIDT ALGORITHMS

- **classical** and **modified** Gram-Schmidt are mathematically equivalent, but they have **"different"** numerical properties
- **classical** Gram-Schmidt can be **"quite unstable",** can **"quickly" lose** all semblance of **orthogonality**
- Gram-Schmidt with **reorthogonalization**: **"two-steps are enough"** to preserve the orthogonality to working accuracy
- **Cholesky** QR algorithm: the triangular factor computed as the Cholesky factor of the **cross-product** matrix and the orthogonal vectors recovered from the inverse of the triangular factor

GRAM-SCHMIDT ALGORITHMS WITH COMPLETE REORTHOGONALIZATION

classical (CGS2)

for $j = 1, \ldots, n$
$u_j = a_j$
for $i = 1, 2$
for $k = 1, \ldots, j-1$

$$u_j = u_j - (a_j, q_k)q_k$$

end
end
$q_j = u_j / \|u_j\|$
end

modified (MGS2)

for $j = 1, \ldots, n$
$u_j = a_j$
for $i = 1, 2$
for $k = 1, \ldots, j-1$

$$u_j = u_j - (u_j, q_k)q_k$$

end
end
$q_j = u_j / \|u_j\|$
end

ORTHOGONALIZATION WITH A NON-STANDARD INNER PRODUCT

$B \in \mathcal{R}^{m,m}$ symmetric positive definite, inner product $\langle \cdot, \cdot \rangle_B$

$A = [a_1, \ldots, a_n] \in \mathcal{R}^{m,n}$, $m \geq n = rank(A)$

B-orthogonal basis of the range of A:

$Q = [q_1, \ldots, q_n] \in \mathcal{R}^{m,n}$, $Q^T B Q = I$

$A = QR$, $R \in \mathcal{R}^{n,n}$ upper triangular

$$A^T B A = R^T R$$

REFERENCE AND GRAM-SCHMIDT IMPLEMENTATIONS

backward stable eigendecomposition + backward stable QR:
$B = V\Lambda V^T,\ \Lambda^{1/2}V^T A = SR,\ Q = V\Lambda^{-1/2}S$

Cholesky QR factorization $A^T BA = R^T R$, $Q = AR^{-1}$

Gram-Schmidt process:

$$u_j^{(0)} = a_j, \quad u_j^{(k)} = u_j^{(k-1)} - r_{kj}q_k, \quad k = 1, \ldots, j-1$$
$$q_j = u_j^{(j-1)}/r_{jj}, \quad r_{jj} = \|u_j^{(j-1)}\|_B$$

modified Gram-Schmidt $\equiv$ SAINV: $r_{kj} = \langle u_j^{(k-1)}, q_k\rangle_B$
classical Gram-Schmidt: $r_{kj} = \langle a_j, q_k\rangle_B$
AINV algorithm: $r_{kj} = \langle u_j^{(k-1)}, a_k/r_{kk}\rangle_B$

CLASSICAL AND MODIFIED GRAM-SCHMIDT ALGORITHMS

classical (CGS)

for $j = 1, \ldots, n$
$u_j^{(0)} = a_j$
for $k = 1, \ldots, j-1$

$$u_j^{(k)} = u_j^{(k-1)} - \langle a_j, q_k \rangle_B q_k$$

end
$q_j = u_j^{(j-1)} / \|u_j^{(j-1)}\|_B$
end

modified (MGS)

for $j = 1, \ldots, n$
$u_j^{(0)} = a_j$
for $k = 1, \ldots, j-1$

$$u_j^{(k)} = u_j^{(k-1)} - \langle u_j^{(k-1)}, q_k \rangle_B q_k$$

end
$q_j = u_j^{(j-1)} / \|u_j^{(j-1)}\|_B$
end

AINV **algorithm**

for $j = 1, \ldots, n$
$u_j^{(0)} = a_j$
for $k = 1, \ldots, j-1$

$$u_j^{(k)} = u_j^{(k-1)} - \langle u_j^{(k-1)}, \frac{a_k}{\|z_k^{(k-1)}\|_B} \rangle_B q_k$$

end
$q_j = u_j^{(j-1)} / \|u_j^{(j-1)}\|_B$
end

GRAM-SCHMIDT ALGORITHMS WITH COMPLETE REORTHOGONALIZATION

classical (CGS2)

for $j = 1, \ldots, n$
  $u_j^{(0)} = a_j$
  for $i = 1, 2$
  for $k = 1, \ldots, j-1$

$$u_j^{(k)} = u_j^{(k-1)} - \langle a_j, q_k \rangle_B q_k$$

  end
  end
  $q_j = u_j^{(j-1)} / \|u_j^{(j-1)}\|_B$
end

modified (MGS2)

for $j = 1, \ldots, n$
  $u_j^{(0)} = a_j$
  for $i = 1, 2$
  for $k = 1, \ldots, j-1$

$$u_j^{(k)} = u_j^{(k-1)} - \langle u_j^{(k-1)}, q_k \rangle_B q_k$$

  end
  end
  $q_j = u_j^{(j-1)} / \|u_j^{(j-1)}\|_B$
end

ORTHOGONALIZATION WITH RESPECT TO A SYMMETRIC BILINEAR FORM

$B \in \mathcal{R}^{m,m}$ symmetric indefinite and nonsingular, bilinear form

$$A = (a_1, \ldots, a_n) \in \mathcal{R}^{m,n},\ m \geq n = rank(A)$$

B-orthonormal basis of $span(A)$:

$$Q = (q_1, \ldots, q_n) \in \mathcal{R}^{m,n},\ Q^T B Q = \Omega \in \mathrm{diag}(\pm 1)$$

$A = QR$, $R \in \mathcal{R}^{n,n}$ upper triangular with positive diagonal

if no principal minor of C vanishes (if C is strongly nonsingular)

$$C = A^T B A = R^T \Omega R$$

Bunch 1971, Bunch-Parlett 1971

Della Dora 1975, Elsner 1979, Bunse-Gerstner 1981

Slapnicar 1999, Singer and Singer 2000, Singer 2006

CLASSICAL GRAM-SCHMIDT AND CLASSICAL GRAM-SCHMIDT WITH REORTHOGONALIZATION

classical (CGS)

for $j = 1, \ldots, n$
  $u_j = a_j$

  for $k = 1, \ldots, j-1$

  $$u_j = u_j - \omega_k^{-1}(a_j, Bq_k)q_k$$

  end

  $\omega_j = \text{sign}[(u_j, Bu_j)]$
  $q_j = u_j / \sqrt{|(u_j, Bu_j)|}$
end

classical with reorthogonalization (CGS2)

for $j = 1, \ldots, n$
  $u_j = a_j$
  for $i = 1, 2$
  for $k = 1, \ldots, j-1$

  $$u_j = u_j - \omega_k^{-1}(u_j, Bq_k)q_k$$

  end
  end
  $\omega_j = \text{sign}[(u_j, Bu_j)]$
  $q_j = u_j / \sqrt{|(u_j, Bu_j)|}$
end

ORTHOGONALIZATION WITH RESPECT TO A SKEW-SYMMETRIC BILINEAR FORM

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \in \mathcal{R}^{2m,2m} \text{ skew-symmetric and orthogonal,}$$

$$A = (a_1, \dots, a_n) \in \mathcal{R}^{m,n},\ m \geq n = rank(A)$$

J-orthonormal basis of $span(A)$: $Q = (q_1, \dots, q_n) \in \mathcal{R}^{m,n}$

$$Q^T J Q = \mathrm{diag}(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}) \in \mathcal{R}^{2n,2n}$$

$A = QR$, $R \in \mathcal{R}^{n,n}$ upper triangular with positive diagonal

if no minor of C with even dimension vanishes

$$C = A^T J A = R^T \mathrm{diag}(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}) R$$

Della Dora 1975, Elsner 1979, Bunse-Gerstner 1981
Mehrmann 1979, Bunse-Gerstner and Mehrmann 1986
Benner, Byers, Fassbender, Mehrmann, Watkins 2000

CLASSICAL GRAM-SCHMIDT

classical Gram-Schmidt (CGS)

for $j = 1, \ldots, n$

$[u_{2j-1}, u_{2j}] = [a_{2j-1}, a_{2j}]$

for $k = 1, \ldots, j-1$

$$[u_{2j-1}, u_{2j}] = [u_{2j-1}, u_{2j}] - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} [q_{2k-1}, q_{2k}]^T J [a_{2j-1}, a_{2j}]$$

end

$$\begin{pmatrix} r_{11} & 0 \\ r_{12} & r_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix} = [u_{2j-1}, u_{2j}]^T J [u_{2j-1}, u_{2j}]$$

$$[q_{2j-1}, q_{2j}] = [u_{2j-1}, u_{2j}] \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}^{-1}$$

end

$$\bar{Q} = (\bar{q}_1, \ldots, \bar{q}_n),\ \bar{Q}^T\bar{Q} \neq I_n,\ \|I - \bar{Q}^T\bar{Q}\| \leq ?$$

$$A \neq \bar{Q}\bar{R},\ \|A - \bar{Q}\bar{R}\| \leq ?$$

$$\bar{R}?\ ,\ \mathrm{cond}(\bar{R}) \leq ?$$

ILLUSTRATION, EXAMPLE

Läuchli, 1961, Björck, 1967: $A = \begin{pmatrix} 1 & 1 & 1 \\ \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma \end{pmatrix}$

$$\kappa(A) = \sigma^{-1}(n + \sigma^2)^{1/2} \approx \sigma^{-1}\sqrt{n},\ \sigma \ll 1$$
$$\sigma_{min}(A) = \sigma,\ \|A\| = \sqrt{n + \sigma^2}$$

assume first that $\sigma^2 \leq u$, so $\mathrm{fl}(1 + \sigma^2) = 1$

ILLUSTRATION, EXAMPLE

if no other rounding errors are made, the matrices computed in CGS and MGS have the following form:

$$\begin{pmatrix} 1 & 0 & 0 \\ \sigma & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ \sigma & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}$$

CGS: $(\bar{q}_3, \bar{q}_1) = -\sigma/\sqrt{2}$, $(\bar{q}_3, \bar{q}_2) = 1/2$,

MGS: $(\bar{q}_3, \bar{q}_1) = -\sigma/\sqrt{6}$, $(\bar{q}_3, \bar{q}_2) = 0$

complete loss of orthogonality ($\iff$ loss of lin. independence, loss of (numerical) rank): $\sigma^2 \leq u$ (CGS), $\sigma \leq u$ (MGS)

STANDARD INNER PRODUCT: ROUNDING ERRORS

- **modified** Gram-Schmidt:
  assuming $\mathcal{O}(u)\kappa(A) < 1$
  $\|I - \bar{Q}^T\bar{Q}\| \leq \frac{\mathcal{O}(u)\kappa(A)}{1-\mathcal{O}(u)\kappa(A)}$

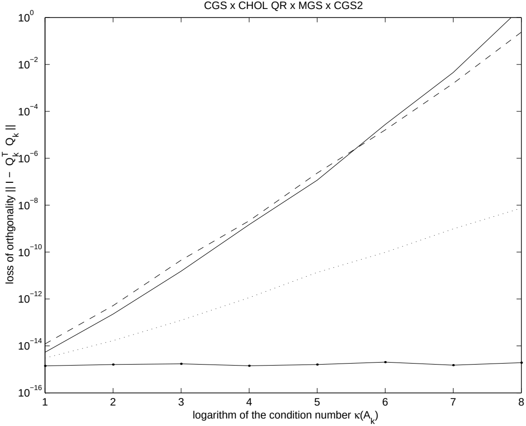
  Björck, 1967, Björck, Paige, 1992

- **classical** Gram-Schmidt:
  assuming $\mathcal{O}(u)\kappa(A) < 1$
  $\|I - \bar{Q}^T\bar{Q}\| \leq \frac{\mathcal{O}(u)\kappa^2(A)}{1-\mathcal{O}(u)\kappa(A)}$

  Giraud, van den Eshof, Langou, R, 2005
  Barlow, Smoktunowicz, Langou, 2006

- classical or modified Gram-Schmidt with **reorthogonalization**:
  assuming $\mathcal{O}(u)\kappa(A) < 1$
  $\|I - \bar{Q}^T\bar{Q}\| \leq \mathcal{O}(u)$

  Giraud, van den Eshof, Langou, R, 2005
  Barlow, Smoktunowicz, 2011



Stewart, "Matrix algorithms" book, p. 284, 1998

TRIANGULAR FACTOR FROM CLASSICAL GRAM-SCHMIDT VS. CHOLESKY FACTOR OF THE CROSS-PRODUCT MATRIX

exact arithmetic:

$$
\begin{aligned}
r_{i,j} = (a_j, q_i) &= \left(a_j, \frac{a_i - \sum_{k=1}^{i-1} r_{k,i} q_k}{r_{i,i}} \right) \\
&= \frac{(a_j, a_i) - \sum_{k=1}^{i-1} r_{k,i} r_{k,j}}{r_{i,i}}
\end{aligned}
$$

The computation of R in the classical Gram-Schmidt is closely related to the left-looking Cholesky factorization of the cross-product matrix $A^T A = R^T R$

CLASSICAL GRAM-SCHMIDT PROCESS: COMPUTED TRIANGULAR FACTOR

$$\sum_{k=1}^{i} \bar{r}_{k,i}\bar{r}_{k,j} = (a_i, a_j) + \Delta e_{i,j}, \; i < j$$

$$[A^T A + \Delta E_1]_{i,j} = [\bar{R}^T \bar{R}]_{i,j}!$$

$$\|\Delta E_1\| \leq c_1 u \|A\|^2$$

The CGS process is another way how to compute **a backward stable Cholesky factor of the cross-product matrix** $A^T A$!

Giraud, van den Eshof, Langou, R, 2005

CLASSICAL GRAM-SCHMIDT PROCESS: DIAGONAL ELEMENTS!

$$u_j = (I - Q_{j-1}Q_{j-1}^T)a_j$$

$$\|u_j\| = \|a_j - Q_{j-1}(Q_{j-1}^T a_j)\| = (\|a_j\|^2 - \|Q_{j-1}^T a_j\|^2)^{1/2}$$

$$q_j = \frac{u_j}{(\|a_j\| - \|Q_{j-1}^T a_j\|)^{1/2}(\|a_j\| + \|Q_{j-1}^T a_j\|)^{1/2}},$$

$$\|a_j\|^2 = \sum_{k=1}^{j} \bar{r}_{k,j}^2 + \Delta e_{j,j},\ \|\Delta e_{j,j}\| \leq c_1 u \|a_j\|^2$$

J. Barlow, A. Smoktunowicz, Langou, 2006

CLASSICAL GRAM-SCHMIDT PROCESS: THE LOSS OF ORTHOGONALITY

$$A^T A + \Delta E_1 = \bar{R}^T \bar{R}, \; A + \Delta E_2 = \bar{Q}\bar{R}$$

$$\bar{R}^T (I - \bar{Q}^T \bar{Q}) \bar{R} = -(\Delta E_2)^T A - A^T \Delta E_2 - (\Delta E_2)^T \Delta E_2 + \Delta E_1$$

assuming $c_2 u \kappa(A) < 1$

$$\|I - \bar{Q}^T \bar{Q}\| \leq \frac{c_3 u \kappa^2(A)}{1 - c_2 u \kappa(A)}$$

Giraud, van den Eshof, Langou, R, 2005

GRAM-SCHMIDT WITH THE REORTHOGONALIZATION

$$u_j = (I - Q_{j-1}Q_{j-1}^T)a_j,\ v_j = (I - Q_{j-1}Q_{j-1}^T)^2 a_j$$
$$\|u_j\| = |r_{j,j}| \geq \sigma_{min}(R_j) = \sigma_{min}(A_j) \geq \sigma_{min}(A)$$
$$\frac{\|a_j\|}{\|u_j\|} \leq \kappa(A),\ \frac{\|u_j\|}{\|v_j\|} = 1,\ Q_{j-1}^T\left(\frac{v_j}{\|v_j\|}\right) = 0$$
$$A + \Delta E_2 = \bar{Q}\bar{R},\ \|\Delta E_2\| \leq c_2 u \|A\|$$

$$\frac{\|a_j\|}{\|\bar{u}_j\|} \leq \frac{\kappa(A)}{1-\tilde{c}_1 u\kappa(A)},\ \frac{\|\bar{u}_j\|}{\|\bar{v}_j\|} \leq [1 - \tilde{c}_2 u\kappa(A)]^{-1},\ \frac{\|\bar{Q}_{j-1}^T\bar{v}_j\|}{\|\bar{v}_j\|} \leq ?$$

assuming $\tilde{c}_3 u\kappa(A) < 1$

$$\|I - \bar{Q}^T\bar{Q}\| \leq \frac{\tilde{c}_4 u}{1-\tilde{c}_3 u\kappa(A)}$$

Giraud, van den Eshof, Langou, R, 2005

GRAM-SCHMIDT FLOPS VERSUS PARALLELISM

- classical Gram-Schmidt (CGS): mn^2 saxpys
- classical Gram-Schmidt with reorthogonalization (CGS2): $2mn^2$ saxpys
- Householder orthogonalization: $2(mn^2 - n^3/3)$ saxpys

in parallel environment and using BLAS2, CGS2 may be faster than (plain) MGS!

Frank, Vuik, 1999, Lehoucq, Salinger, 2001

ON THE WAY FROM THE STANDARD TO THE NONSTANDARD INNER PRODUCT

- Axel Ruhe. Numerical aspects of Gram-Schmidt orthogonalization of vectors, Lin. Alg. and its Appl., 52/53 (591-601), 1983.
- T. Ericsson, An analysis of orthogonalization in elliptic norms, to appear.
- M. Gulliksson: Backward error analysis for the constrained and weighted linear least squares problem when using the weighted QR factorization. SIAM J. Matrix Anal. Appl. 16(2), 675-687 (1995)
  M. Gulliksson: On the modified GramSchmidt algorithm for weighted and constrained linear least squares problems. BIT Numer. Math. 35(4), 453-468 (1995)
- S.J. Thomas, R.V.M. Zahar: Efficient orthogonalization in the M-norm. Congr. Numer. 80, 23-32 (1991) 36.
  S.J. Thomas, R.V.M. Zahar, : An analysis of orthogonalization in elliptic norms. Congr. Numer. 86, 193-222 (1992)

NON-STANDARD INNER PRODUCT AND STANDARD QR

$$B^{1/2}A = (B^{1/2}Q)R, \;\; Q^T B Q = (B^{1/2}Q)^T (B^{1/2}Q) = I$$

$$\kappa(Q) \ll \kappa^{1/2}(B)$$

$$\kappa(R) = \kappa(B^{1/2}A) \leq \kappa^{1/2}(B)\kappa(A)$$

$$A = I\text{: } Q = \left[\begin{array}{c} R^{-1} \\ 0 \end{array} \right] \in \mathcal{R}^{m,n} \text{ upper triangular}$$

$$\kappa(Q) = \kappa(R) = \kappa^{1/2}(A)$$

INVERSE FACTORIZATION AND APPROXIMATE INVERSE PRECONDITIONING

$$QQ^T = AR^{-1}R^{-T}A^T = A[A^TBA]^{-1}A^T$$

$A = I$ square and nonsingular: inverse factorization $QQ^T = B^{-1}$

$Bx = b$, approximate inverse $\bar{Q}\bar{Q}^T \approx B^{-1}$

$$\downarrow$$

$$\bar{Q}^TB\bar{Q}y = \bar{Q}^Tb\ ,\ \ x = \bar{Q}y,\ \ \|\bar{Q}^TB\bar{Q} - I\| \leq ?$$

finite precision arithmetic:

$$\bar{Q} = (\bar{q}_1, \ldots, \bar{q}_n),\ \bar{Q}^TB\bar{Q} \neq I,\ \|I - \bar{Q}^TB\bar{Q}\| \leq ?$$

$$\bar{R}^T\bar{R} \approx A^TBA,\ \ \|A^TBA - \bar{R}^T\bar{R}\| \leq ?$$

$$\bar{Q}\bar{R} \approx A,\ \ \|A - \bar{Q}\bar{R}\| \leq ?$$

CLASSICAL GRAM-SCHMIDT PROCESS

$$A^T BA = R^T R$$

$$\|A - \bar{Q}\bar{R}\| \leq \mathcal{O}(u)\|\bar{Q}\|\|\bar{R}\|$$
$$\|\bar{R} - \bar{Q}^T BA\| \leq \mathcal{O}(u)\|A\|\|B\|\|\bar{Q}\|$$

$$A^T BA = \bar{R}^T \bar{R} - (\bar{R} - \bar{Q}^T BA)^T \bar{R} + A^T B(A - \bar{Q}\bar{R})$$

assuming $\mathcal{O}(u)\kappa(B)\kappa(B^{1/2}A)\kappa(A) < 1$

$$\|I - \bar{Q}^T B\bar{Q}\| \leq \frac{\mathcal{O}(u)\|B\|^{1/2}\|\bar{Q}\|\kappa(B^{1/2}A)\kappa^{1/2}(B)\kappa(A)}{1-\mathcal{O}(u)\|B\|^{1/2}\|\bar{Q}\|\kappa(B^{1/2}A)\kappa^{1/2}(B)\kappa(A)}$$

GRAM-SCHMIDT PROCESS WITH REORTHOGONALIZATION

$$u_j^{(1)} == (I - Q_{j-1}Q_{j-1}^T B)a_j,\ u_j^{(2)} = (I - Q_{j-1}Q_{j-1}^T B)u_j^{(1)} = (I - Q_{j-1}Q_{j-1}^T B)^2 a_j = u_j^{(1)}$$

$$\|\bar{Q}_{j-1}^T B \left(\frac{\bar{u}_j^{(2)}}{\bar{r}_{j,j}} \right) \| \lessapprox \|I - \bar{Q}_{j-1}^T B \bar{Q}_{j-1}\|^2 \frac{\|\bar{Q}_{j-1}^T B a_j\|}{\bar{r}_{j,j}}$$

$$1/r_{j,j} \le \sigma_{min}^{-1}(R) = \sigma_{min}^{-1}(B^{1/2}A),\ \|\|\|\bar{Q}_{j-1}^T B a_j\|\|\|/r_{j,j} \le \kappa(B^{1/2}A),$$

$$\mathcal{O}(u)\kappa^{1/2}(B)\kappa(B^{1/2}A) < 1$$

$$\|I - \bar{Q}^T B \bar{Q}\| \le \mathcal{O}(u)\|B\|\|\bar{Q}\|\|\bar{Q}^{(1)}\|$$

NON-STANDARD INNER PRODUCT: THE LOSS OF B-ORTHOGONALITY

modified Gram-Schmidt:

$$\mathcal{O}(u)\kappa(B)\kappa(B^{1/2}A) < 1$$

$$\|I - \bar{Q}^T B\bar{Q}\| \leq \frac{\mathcal{O}(u)\|B\|\|\bar{Q}\|^2\kappa(B^{1/2}A)}{1 - \mathcal{O}(u)\|B\|\|\bar{Q}\|^2\kappa(B^{1/2}A)}$$

classical Gram-Schmidt and AINV algorithm:

$$\mathcal{O}(u)\kappa(B)\kappa(B^{1/2}A)\kappa(A) < 1$$

$$\|I - \bar{Q}^T B\bar{Q}\| \leq \frac{\mathcal{O}(u)\|B\|^{1/2}\|\bar{Q}\|\kappa(B^{1/2}A)\kappa^{1/2}(B)\kappa(A)}{1 - \mathcal{O}(u)\|B\|^{1/2}\|\bar{Q}\|\kappa(B^{1/2}A)\kappa^{1/2}(B)\kappa(A)}$$

classical Gram-Schmidt with reorthogonalization:

$$\mathcal{O}(u)\kappa^{1/2}(B)\kappa(B^{1/2}A) < 1$$

$$\|I - \bar{Q}^T B\bar{Q}\| \leq \mathcal{O}(u)\|B\|\|\bar{Q}\|\|\bar{Q}^{(1)}\|$$

THE LOCAL ERRORS IN A NON-STANDARD INNER PRODUCT

general positive definite B :

$$|\text{fl}[\langle \bar{u}_j^{(k-1)}, \bar{q}_k \rangle_B] - \langle \bar{u}_j^{(k-1)}, \bar{q}_k \rangle_B| \leq \mathcal{O}(u)\|B\|\|\bar{u}_j^{(k-1)}\|\|\bar{q}_k\|$$

$$|1 - \|\bar{q}_j\|_B^2| \leq \mathcal{O}(u)\|B\|\|\bar{q}_j\|^2$$

diagonal positive (weight matrix) B :

$$|\text{fl}[\langle \bar{u}_j^{(k-1)}, \bar{q}_k \rangle_B] - \langle \bar{u}_j^{(k-1)}, \bar{q}_k \rangle_B| \leq \mathcal{O}(u)\|\bar{u}_j^{(k-1)}\|_B\|\bar{q}_k\|_B$$

$$|1 - \|\bar{q}_j\|_B^2| \leq \mathcal{O}(u)$$

DIAGONAL CASE IS SIMILAR TO STANDARD CASE

modified Gram-Schmidt:

$$\mathcal{O}(u)\kappa(B^{1/2}A) < 1$$

$$\|I - \bar{Q}^T B \bar{Q}\| \leq \frac{\mathcal{O}(u)\kappa(B^{1/2}A)}{1 - \mathcal{O}(u)\kappa(B^{1/2}A)}$$

classical Gram-Schmidt and AINV algorithm

$$\mathcal{O}(u)\kappa^2(B^{1/2}A) < 1$$

$$\|I - \bar{Q}^T B \bar{Q}\| \leq \frac{\mathcal{O}(u)\kappa^2(B^{1/2}A)}{1 - \mathcal{O}(u)\kappa^2(B^{1/2}A)}$$

classical Gram-Schmidt with reorthogonalization:

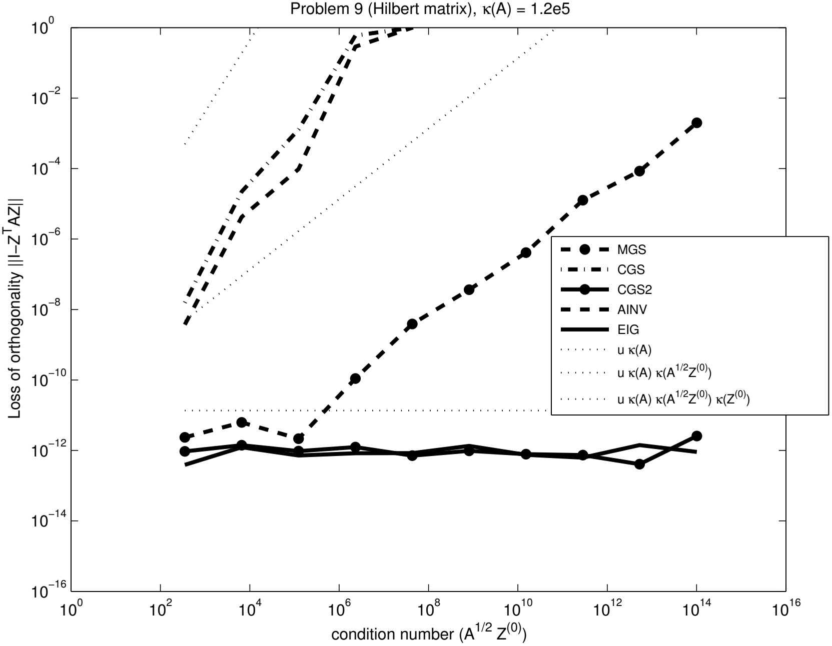
$$\mathcal{O}(u)\kappa(B^{1/2}A) < 1$$

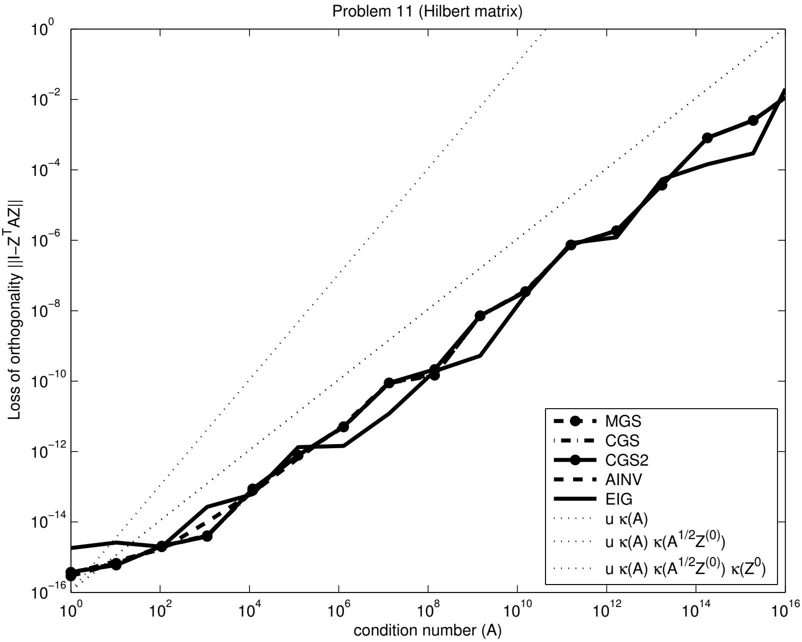
$$\|I - \bar{Q}^T B \bar{Q}\| \leq \mathcal{O}(u)$$

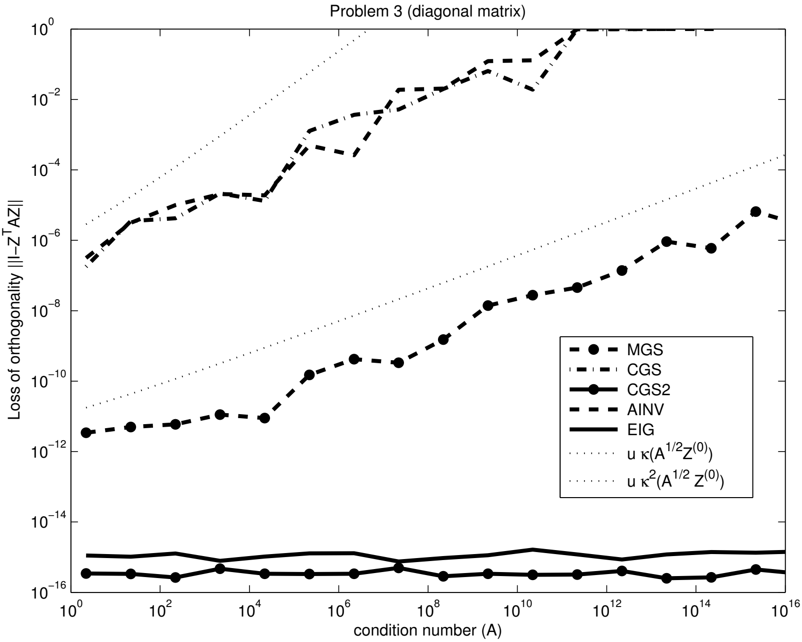
Gulliksson, Wedin 1992, Gulliksson 1995

NUMERICAL EXPERIMENTS - EXTREMAL CASES

1. $\kappa^{1/2}(B) \ll \kappa(B^{1/2}A)$
2. $\kappa(B^{1/2}A) \leq \kappa^{1/2}(B)$
3. B positive diagonal


Problem 9 (Hilbert matrix), κ(A) = 1.2e5
Loss of orthogonality ||I−Z^T AZ||
condition number (A^{1/2} Z^{(0)})
MGS
CGS
CGS2
AINV
EIG
u κ(A)
u κ(A) κ(A^{1/2}Z^{(0)})
u κ(A) κ(A^{1/2}Z^{(0)}) κ(Z^{(0)})


Problem 11 (Hilbert matrix)
Loss of orthogonality ||I−Z^T AZ||
condition number (A)
MGS
CGS
CGS2
AINV
EIG
u κ(A)
u κ(A) κ(A^{1/2}Z^{(0)})
u κ(A) κ(A^{1/2}Z^{(0)}) κ(Z^0)


Problem 3 (diagonal matrix)
Loss of orthogonality $||I-Z^TAZ||$
condition number (A)
MGS
CGS
CGS2
AINV
EIG
$u\,\kappa(A^{1/2}Z^{(0)})$
$u\,\kappa^2(A^{1/2}\,Z^{(0)})$

ON THE WAY FROM THE INNER PRODUCT TO THE BILINEAR FORM

- Symmetric indefinite eigenvalue problems. The bilinear form $\langle x, y\rangle_B = y^T Bx$ can have $\langle x, x\rangle_B < 0$ and $\langle x, x\rangle_B = 0$ for some $x \neq 0$.
- For B symmetric and nonsingular the eigenvectors Q can be chosen such that $Q^T BQ = \Omega$ where $\Omega = \mathrm{diag}(\pm 1)$ is a signature matrix. Isotropic vectors $x^T Bx = 0$.
- Structured eigenvalue problems. The SR factorization. The skew-symmetric bilinear form $\langle x, y\rangle_B = y^T Bx$, where $B^T = -B$. Each vector satisfies $x^T Bx = 0$.
- If B is skew-symmetric and nonsingular, then its dimension must be even.

Thank you for your attention!!!

References: R, J. Kopal, M. Tuma, A. Smoktunowicz: Numerical stability of orthogonalization methods with a non-standard inner product, BIT Numerical Mathematics (2012) 52:1035-1058.
R, F. Okulicka-Dluzewska, A. Smoktunowicz: Indefinite orthogonalization with rounding errors, submitted 2013.