GRAM-SCHMIDT PROCESS: ROUNDING ERROR ANALYSIS AND ITS APPLICATIONS

Miro Rozložník

Institute of Computer Science,
Czech Academy of Sciences,
Prague, Czech Republic and
Technical University of Liberec,
email: miro@cs.cas.cz

joint results on CGS with
Luc Giraud, Jasper van den Eshof and Julien Langou

SNA'05, Ostrava, February 7-11, 2005

GRAM-SCHMIDT ORTHOGONALIZATION

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

$$m \geq n = rank(A)$$

orthogonal basis Q of $span(A)$

$Q = (q_1, \ldots, q_n) \in \mathcal{R}^{m,n}$, $Q^T Q = I_n$
$A = QR$, R upper triangular

CLASSICAL & MODIFIED GRAM-SCHMIDT ALGORITHMS

classical Gram-Schmidt (CGS) algorithm

Schmidt, 1907,1908

modified Gram-Schmidt (MGS) algorithm

Laplace, 1816, Cauchy, 1837

CLASSICAL & 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 & MODIFIED GRAM-SCHMIDT ALGORITHMS

finite precision arithmetic:

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

classical and **modified** Gram-Schmidt are mathematically equivalent, but they have **"different"** numerical properties

but are they really so different?

ILLUSTRATION

$$A = \begin{pmatrix} 1 & 1 & 1 \\ \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma \end{pmatrix}$$

Läuchli, 1961, Björck, 1967

$\kappa(A) = \sigma^{-1}(3+\sigma^2)^{1/2} \approx \sigma^{-1}\sqrt{3}$, $\sigma \ll 1$

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

ILLUSTRATION

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

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

ILLUSTRATION

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

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

$\sigma^2 \leq \varepsilon$ (CGS), $\sigma \leq \varepsilon$ (MGS):

complete loss of orthogonality ($\Longleftrightarrow$ loss of lin. independence, loss of (numerical) rank)

MGS: numerical full rank of A: $\varepsilon\kappa(A) \ll 1$

CGS: numerical nonsingularity of $A^T A$: $\varepsilon\kappa^2(A) \ll 1$

MODIFIED GRAM-SCHMIDT ORTHOGONALIZATION

assuming $c_1 \varepsilon \kappa(A) < 1$

$$\|I - \bar{Q}^T \bar{Q}\| \leq \frac{c_2 \varepsilon \kappa(A)}{1 - c_1 \varepsilon \kappa(A)}$$

Björck, 1967

Björck, Paige, 1992

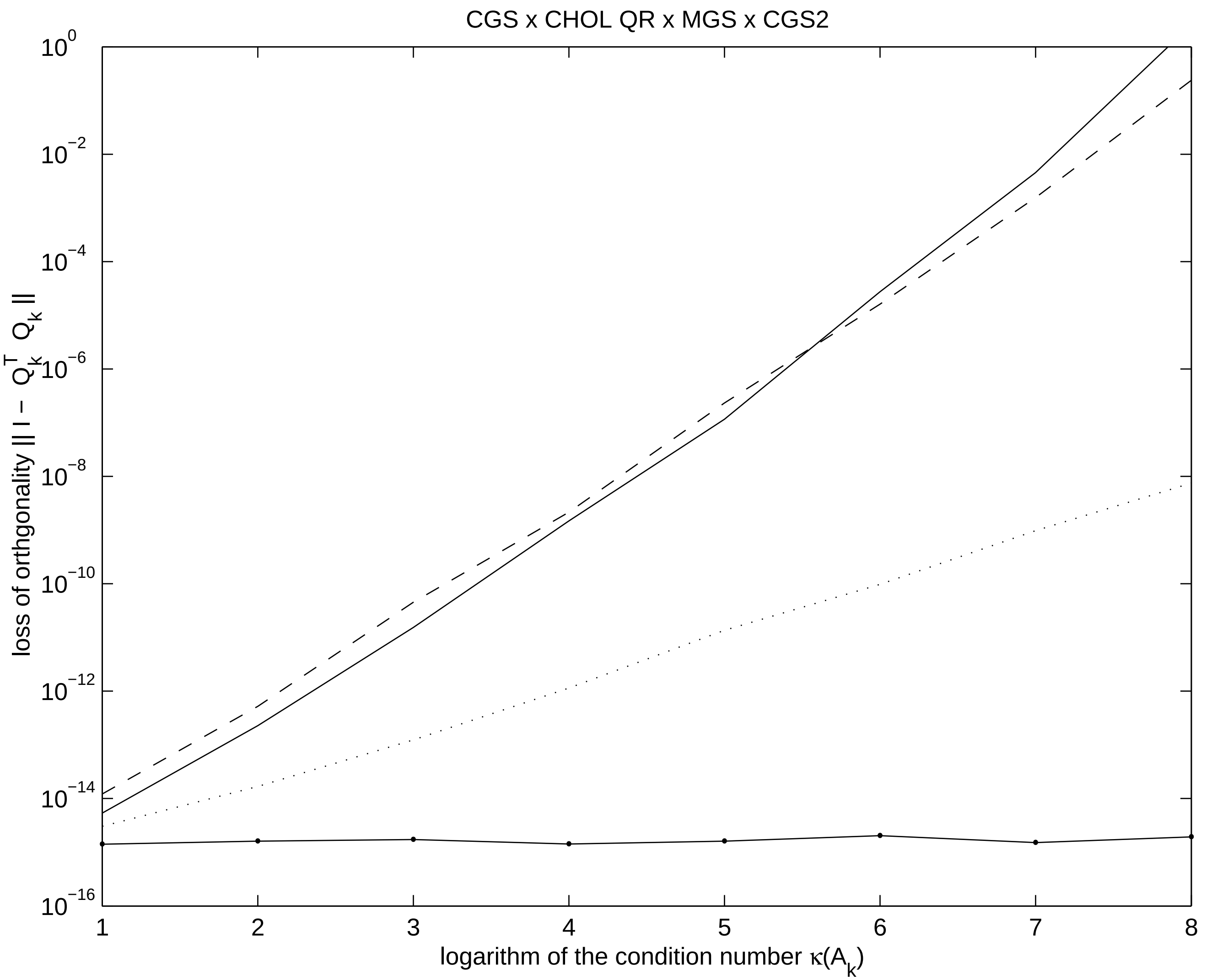
CLASSICAL GRAM-SCHMIDT ORTHOGONALIZATION

assuming $c_3 \varepsilon \kappa^2(A) < 1$

$$\|I - \bar{Q}^T \bar{Q}\| \leq \frac{c_4 \varepsilon \kappa^2(A)}{1 - c_3 \varepsilon \kappa^2(A)}$$

Giraud, Van den Eshof, Langou, R, 2004

Kielbasinski, Schwettlik, 1994



CGS x CHOL QR x MGS x CGS2
loss of orthgonality || I − Q_k^T Q_k ||
logarithm of the condition number κ(A_k)
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
10^−16
1
2
3
4
5
6
7
8

GRAM-SCHMIDT PROCESS VERSUS ROUNDING ERRORS

- modified Gram-Schmidt (MGS): assuming $c_1 \varepsilon \kappa(A) < 1$

  $$\|I - \bar{Q}^T \bar{Q}\| \leq \frac{c_2 \varepsilon \kappa(A)}{1 - c_1 \varepsilon \kappa(A)}$$

  Björck, 1967, Björck, Paige, 1992

- **classical Gram-Schmidt (CGS):** assuming $c_3 \varepsilon \kappa^2(A) < 1$

  $$\|I - \bar{Q}^T \bar{Q}\| \leq \frac{c_4 \varepsilon \kappa^2(A)}{1 - c_3 \varepsilon \kappa^2(A)}!$$

  Giraud, Van den Eshof, Langou, R, 2004

THE GMRES METHOD

$$Ax = b$$

$A \in R^{N,N}$, A nonsingular, $b \in R^N$

$$x_0, r_0 = b - Ax_0,$$
$$K_n(A, r_0) = span\ \{r_0, Ar_0, \ldots, A^{n-1}r_0\}$$

$$x_n \in x_0 + K_n(A, r_0)$$

$$\|b - Ax_n\| = \min_{u \in x_0 + K_n(A, r_0)} \|b - Au\|$$

IMPLEMENTATION OF GMRES

$$x_n = x_0 + V_n y_n$$

Arnoldi (orthogonalization) process: V_n
The loss of orthogonality (loss of rank)
in the computed $\bar{V}_n = [\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n]$

$$\| I - \bar{V}_n^T \bar{V}_n \| \leq ? \ , \ \sigma_n(\bar{V}_n) \leq ?$$

Upper-Hessenberg least squares problem: y_n
The (in)accurate computation of
$\bar{y}_n = fl(\arg\min_y \| \, \|r_0\| e_1 - \bar{H}_{n+1,n} y \|)$

THE ARNOLDI PROCESS

$$V_n = [v_1, v_2, \ldots, v_n]$$

Arnoldi process is a (recursive) column-oriented QR decomposition!

$$[r_0, AV_n] = V_{n+1}[\|r_0\|e_1, H_{n+1,n}]$$

$H_{n+1,n}$ is an upper Hessenberg matrix

QR ORTHOGONALIZATION IN THE ARNOLDI CONTEXT

- modified Gram-Schmidt (MGS):

  $\|I - \bar{V}_{n+1}^T \bar{V}_{n+1}\| \leq c_2 \varepsilon \kappa([\bar{v}_1, A\bar{V}_n])$

Björck, Paige 1967, 1992

- classical Gram-Schmidt (CGS):
  $\|I - \bar{V}_{n+1}^T \bar{V}_{n+1}\| \leq c_4 \varepsilon \kappa^2([\bar{v}_1, A\bar{V}_n])$

Giraud, Langou, R, Van den Eshof 2004

CONDITION NUMBER IN ARNOLDI VERSUS RESIDUAL NORM IN GMRES

$$\frac{\|\widehat{r}_n\|}{\|\bar{r}_0\|[1+\frac{\|\widehat{y}_n\|^2}{1-\delta_n^2}]^{1/2}} \leq \sigma_{n+1}([\bar{v}_1, A\bar{V}_n])$$

Paige, Strakoš, 2002

$$\frac{\|\widehat{r}_n\|}{\|\bar{r}_0\|} = \|\bar{v}_1 - A\bar{V}_n\widehat{y}_n\| = min_y\|\bar{v}_1 - A\bar{V}_n y\|$$

$$\delta_n = \frac{\sigma_{n+1}([\bar{v}_1, A\bar{V}_n])}{\sigma_n(A\bar{V}_n)} < 1$$

Paige, Strakoš, 2002

Greenbaum, R, Strakoš, 1997

ARNOLDI WITH GRAM-SCHMIDT PROCESS

The loss of orthogonality in the Arnoldi process is controlled by the convergence of the residual norm:

$$\|I - \bar{V}_{n+1}^T \bar{V}_{n+1}\| \leq c_0 \varepsilon \kappa^{\alpha}([\bar{v}_1, A\bar{V}_n]), \qquad \alpha = 1,\ \alpha = 2$$

Björck, Paige 1967, 1992

Giraud, Langou, R, Van den Eshof 2003

$$\kappa([\bar{v}_1, A\bar{V}_n]) \leq \frac{\|[\bar{v}_1, A\bar{V}_n]\|}{\frac{\|\hat{r}_n\|}{\|\bar{r}_0\|}[1 + \frac{\|\hat{y}_n\|^2}{1-\delta_n^2}]^{1/2}}$$

Paige, Strakoš, 2000-2002

THE ROLE GRAM-SCHMIDT PROCESS IN THE GMRES METHOD:

The total loss of orthogonality (rank-deficiency) in the Arnoldi process with Gram-Schmidt can occur **only after** GMRES reaches its final accuracy level!

FINAL (MAXIMUM ATTAINABLE) ACCURACY LEVEL OF GMRES

MGS GMRES ($\alpha = 1$):

$$\frac{\|\widehat{r}_n\|||}{\|\bar{r}_0\|[1+\frac{\|\widehat{y}_n\|^2}{1-\delta_n^2}]^{1/2}} \approx c_2[\bar{v}_1, A\bar{V}_n]\|\varepsilon$$

CGS GMRES ($\alpha = 2$):

$$\frac{\|\widehat{r}_n\|}{\|\bar{r}_0\|[1+\frac{\|\widehat{y}_n\|^2}{1-\delta_n^2}]^{1/2}} \approx \left[c_4\|[\bar{v}_1, A\bar{V}_n]\|\varepsilon\right]^{1/2}$$

CONCLUSION:

How important is the orthogonality in GMRES?

For solving the system accurately we **do not** need fully orthogonal vectors - we need their **linear independence**! The crucial thing is a **complete loss** of their orthogonality!