NUMERICAL BEHAVIOR OF GMRES

Miro Rozložník

Institute of Computer Science
Czech Academy of Sciences
CZ-182 07 Prague, Czech Republic
email: miro@cs.cas.cz

joint work with
A. Greenbaum, C. C. Paige and Z. Strakoš

Seminarium Pierścienie, macierze i algorytmy numeryczne,
Politechnika Warszawska, Warszawa, October 26, 2010

OUTLINE

1. ROUNDING ERROR EFFECTS: DELAY OF CONVERGENCE AND MAXIMUM ATTAINABLE ACCURACY
2. BACKWARD ERROR AND BACKWARD STABILITY
3. LOSS OF ORTHOGONALITY AND NUMERICAL BEHAVIOR OF GMRES: HOUSEHOLDER GMRES, MGS AND CGS GMRES
4. HOW TO MAKE SIMPLER GMRES AND GCR MORE STABLE: ADAPTIVE SIMPLER GMRES

ITERATIVE METHODS IN EXACT ARITHMETIC

generate approximate solutions to the solution of $Ax = b$

$$x_0, x_1, \ldots, x_n \longrightarrow x$$

with residual vectors $r_0 = b - Ax_0, \ldots, r_n = b - Ax_n \to 0$

METHODS IN FINITE PRECISION ARITHMETIC

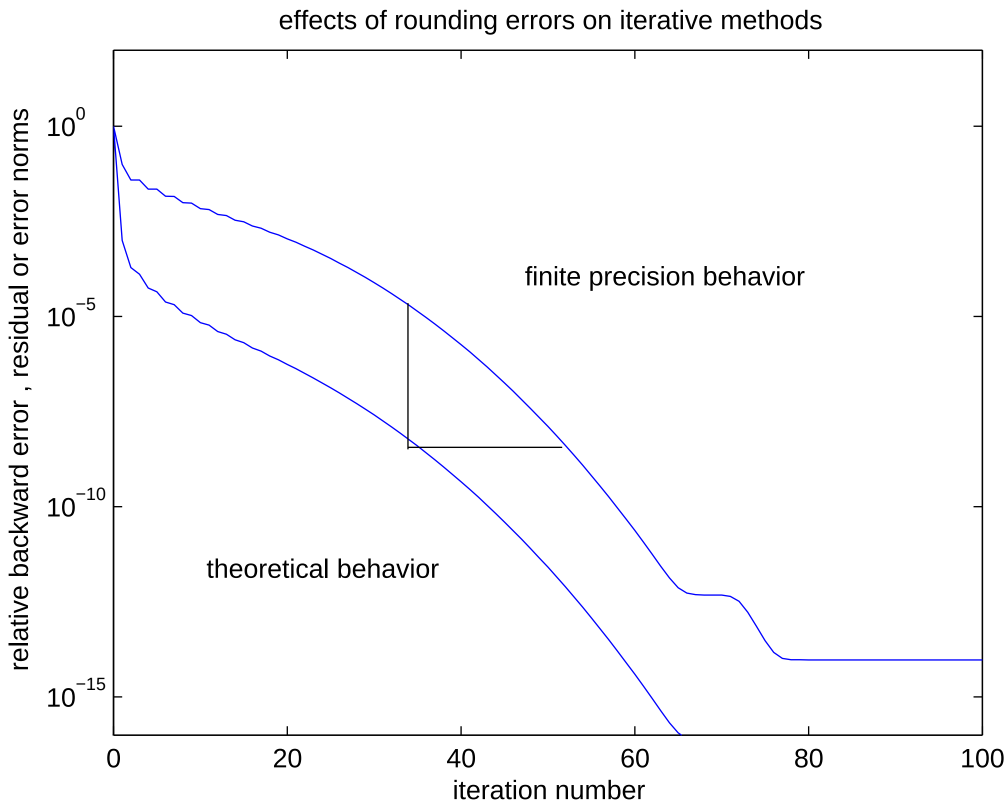
compute approximations $x_0, \bar{x}_1, \ldots, \bar{x}_n$ and updated residual vectors $\bar{r}_0, \bar{r}_1, \ldots, \bar{r}_n$ which are usually close to (but different from) the true residuals $b - A\bar{x}_n$

TWO MAIN QUESTIONS

- How good is the computed approximate solution $\bar{x}_n$? How many (extra) steps do we need to reach the same accuracy as one can get in the exact method?

- How well the computed vector $\bar{r}_n$ approximates the (true) residual $b - A\bar{x}_n$? Is there a limitation on the accuracy of the computed approximate solution?

TWO EFFECTS OF ROUNDING ERRORS: DELAY OF CONVERGENCE AND LIMITING (MAXIMUM ATTAINABLE) ACCURACY


effects of rounding errors on iterative methods
relative backward error , residual or error norms
10^0
10^−5
10^−10
10^−15
finite precision behavior
theoretical behavior
0
20
40
60
80
100
iteration number

THE CONCEPT OF BACKWARD STABILITY

A backward stable algorithm eventually computes the exact answer to a nearby problem, i.e. the vector $\bar{x}_n$ satisfying

$$(A + \Delta A_n)\bar{x}_n = b + \Delta b_n$$

$$\|\Delta A_n\|_{(F)}/\|A\|_{(F)} \leq O(\varepsilon), \ \|\Delta b_n\|/\|b\| \leq O(\varepsilon)$$

$\Longleftrightarrow$ The normwise backward error associated with the approximate solution $\bar{x}_n$ satisfies

$$\frac{\|b - A\bar{x}_n\|}{\|b\| + \|A\|_{(F)}\|\bar{x}_n\|} \leq O(\varepsilon)$$

Prager, Oettli, 1964; Rigal, Gaches, 1967

see also Higham, 2nd ed. 2002; Stewart, Sun, 1990; Meurant 1999

THE LEVEL OF MAXIMUM ATTAINABLE ACCURACY

We are looking for the difference between the updated $\bar{r}_n$ and true residual $b - A\bar{x}_n$ (divided by $\|A\|\|\bar{x}_n\| + \|b\|$ or $\|A\|_F\|\bar{x}_n\| + \|b\|$)

$$\frac{\|b - A\bar{x}_n - \bar{r}_n\|}{\|A\|\|\bar{x}_n\| + \|b\|} \leq ?$$

$$\|\bar{r}_n\| \longrightarrow 0 \Longrightarrow \lim_{n\to\infty} \frac{\|b - A\bar{x}_n\|}{\|A\|\|\bar{x}_n\| + \|b\|} \leq ?$$

In the optimal case the bound is of $O(\varepsilon)$; then we have a backward stable solution

Chris Paige, R, Strakoš, 2006

MOTIVATION: SYMMETRIC LANCZOS PROCESS AND CONJUGATE GRADIENT METHOD

THE CONCEPT OF CONVERGENCE DELAY

Greenbaum, 1989; Strakoš, Greenbaum, 1991

Paige, Strakoš, 1999

Meurant, Strakoš, Acta Numerica 2006

Strakoš, Liesen, ZAMM 2006

Delay in convergence of the conjugate gradient method (due to rounding errors) is given by the rank-deficiency of the computed Lanczos basis!

NONSYMMETRIC ARNOLDI PROCESS AND THE GMRES METHOD

Saad, Schultz 1986

THE CORRECTED PRINCIPLE OF CONVERGENCE DELAY:

Once the rank-deficiency occurs in the Arnoldi process the GMRES method stagnates on its final accuracy level

THEORETICAL JUSTIFICATION?

BASIC QUESTION:

How important is the orthogonality of computed basis vectors in the GMRES method?

ANSWER:

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!

$$Ax = b$$

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

THE GMRES METHOD:

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

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

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

IMPLEMENTATION OF GMRES

$$\bar{x}_n = fl(x_0 + \bar{V}_n \bar{y}_n)$$

Arnoldi (orthogonalization) process:

The loss of orthogonality (loss of rank) in $\bar{V}_n = [\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n]$
$\|I - \bar{V}_n^T \bar{V}_n\| \leq ?$, $\sigma_n(\bar{V}_n) \leq ?$

Upper-Hessenberg least squares problem:

The (in)accurate $\bar{y}_n = fl(\arg\min_y \| \|r_0\| e_1 - \bar{H}_{n+1,n} y\|)$
$\kappa(\bar{H}_{n+1,n}) \leq ?$

THE NONSYMMETRIC ARNOLDI PROCESS

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

Arnoldi process is a (recursive) column-oriented QR decomposition of $[r_0, AV_n]$**!**

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

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

WELL-PRESERVED ORTHOGONALITY $\Rightarrow$ BACKWARD STABILITY

HOUSEHOLDER GMRES:

$$\|I - \bar{V}_N^T \bar{V}_N\| \leq O(\varepsilon)$$

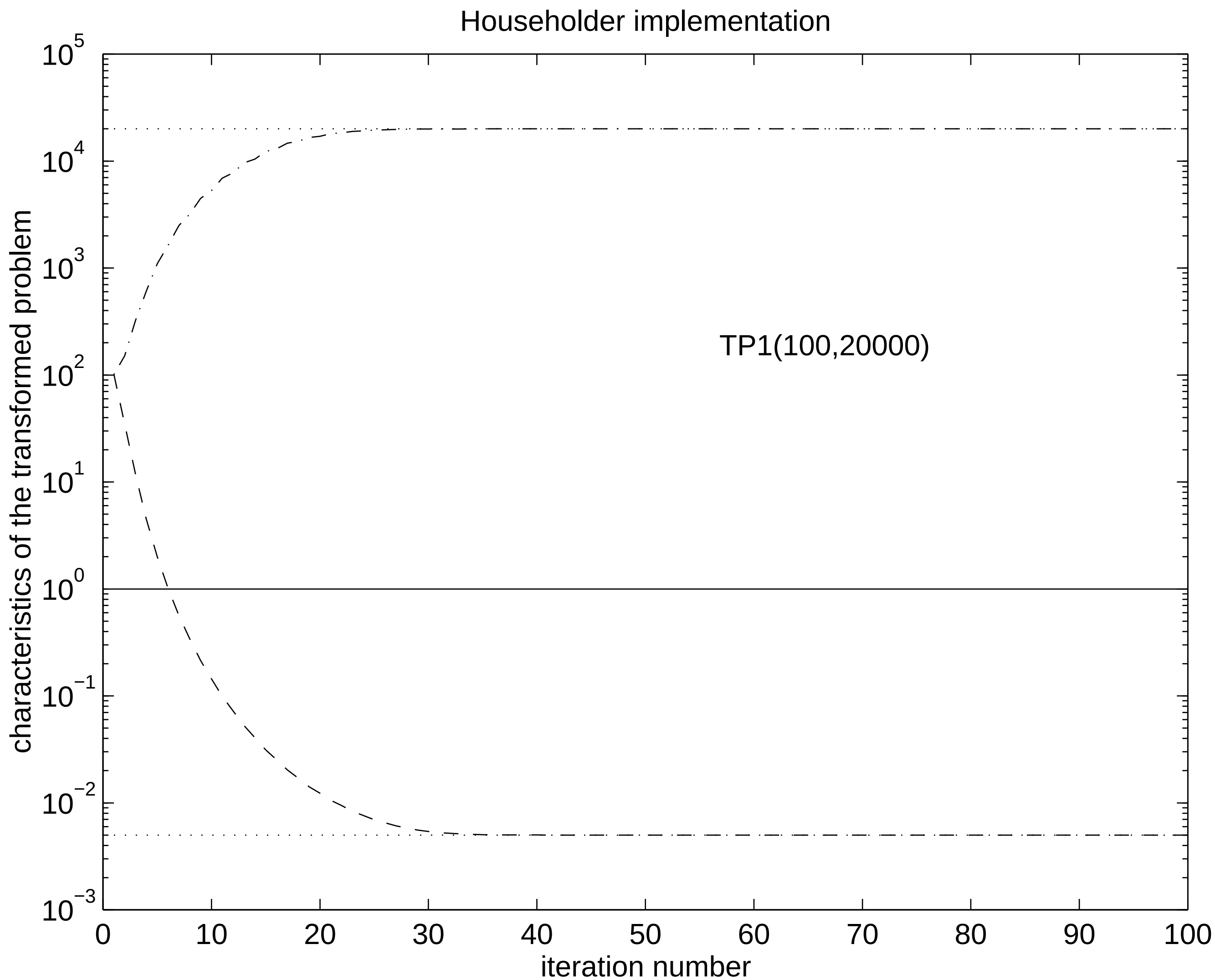
Wilkinson 1967, Walker 1988, 1989

$$\frac{\|b - A\bar{x}_N\|}{\|A\|\|\bar{x}_N\| + \|b\|} \leq O(\varepsilon)$$

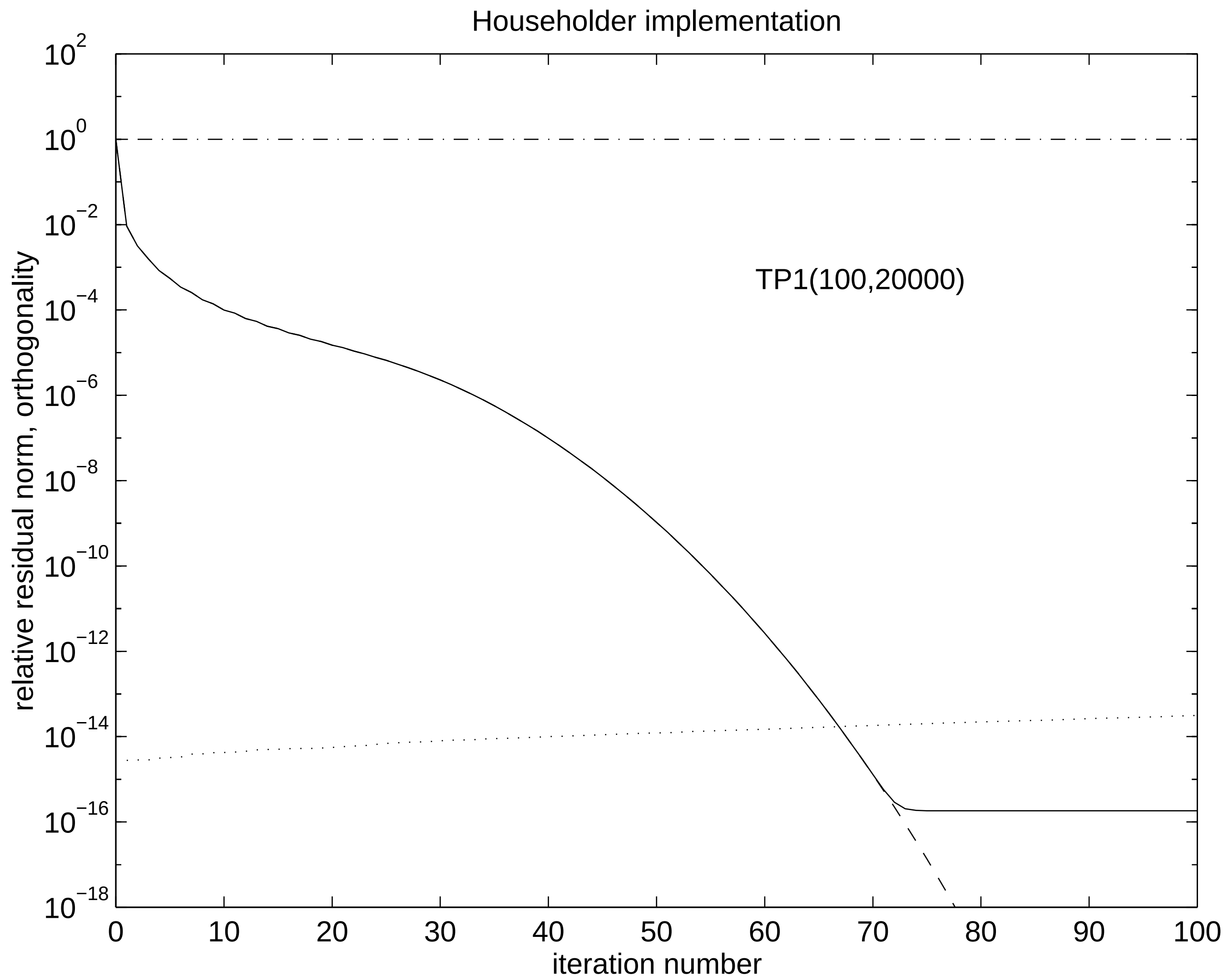
$\bar{x}_N$ represents an exact solution to the nearby problem

$$(A + \Delta A)\bar{x}_N = b + \Delta b$$

Greenbaum, Drkošová, R, Strakoš, 1995



Householder implementation
characteristics of the transformed problem
iteration number
TP1(100,20000)
10^5
10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
60
70
80
90
100


Householder implementation
TP1(100,20000)
relative residual norm, orthogonality
iteration number
10^2
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
10^−16
10^−18
0
10
20
30
40
50
60
70
80
90
100

GRAM-SCHMIDT GMRES

MODIFIED GRAM-SCHMIDT:

$$\|I - \bar{V}_n^T \bar{V}_n\| \leq O(\varepsilon)\kappa([r_0, A\bar{V}_n])$$

Björck, 1967

Björck, Paige, 19992

CLASSICAL GRAM-SCHMIDT:

$$\|I - \bar{V}_n^T \bar{V}_n\| \leq O(\varepsilon)\kappa^2([r_0, A\bar{V}_n])$$

van den Eshof, Giraud, Langou, R, 2005

Smoktunowicz, Barlow, Langou, 2006

The GRAM-SCHMIDT GMRES IMPLEMENTATION

The (modified) Gram-Schmidt version of GMRES (MGS-GMRES) is efficient, but looses orthogonality.

The rank-deficiency (total loss of orthogonality $\equiv$ loss of linear independence of computed basis vectors) in the Arnoldi process with (modified) Gram-Schmidt can occur **only after** GMRES reaches its final accuracy level!

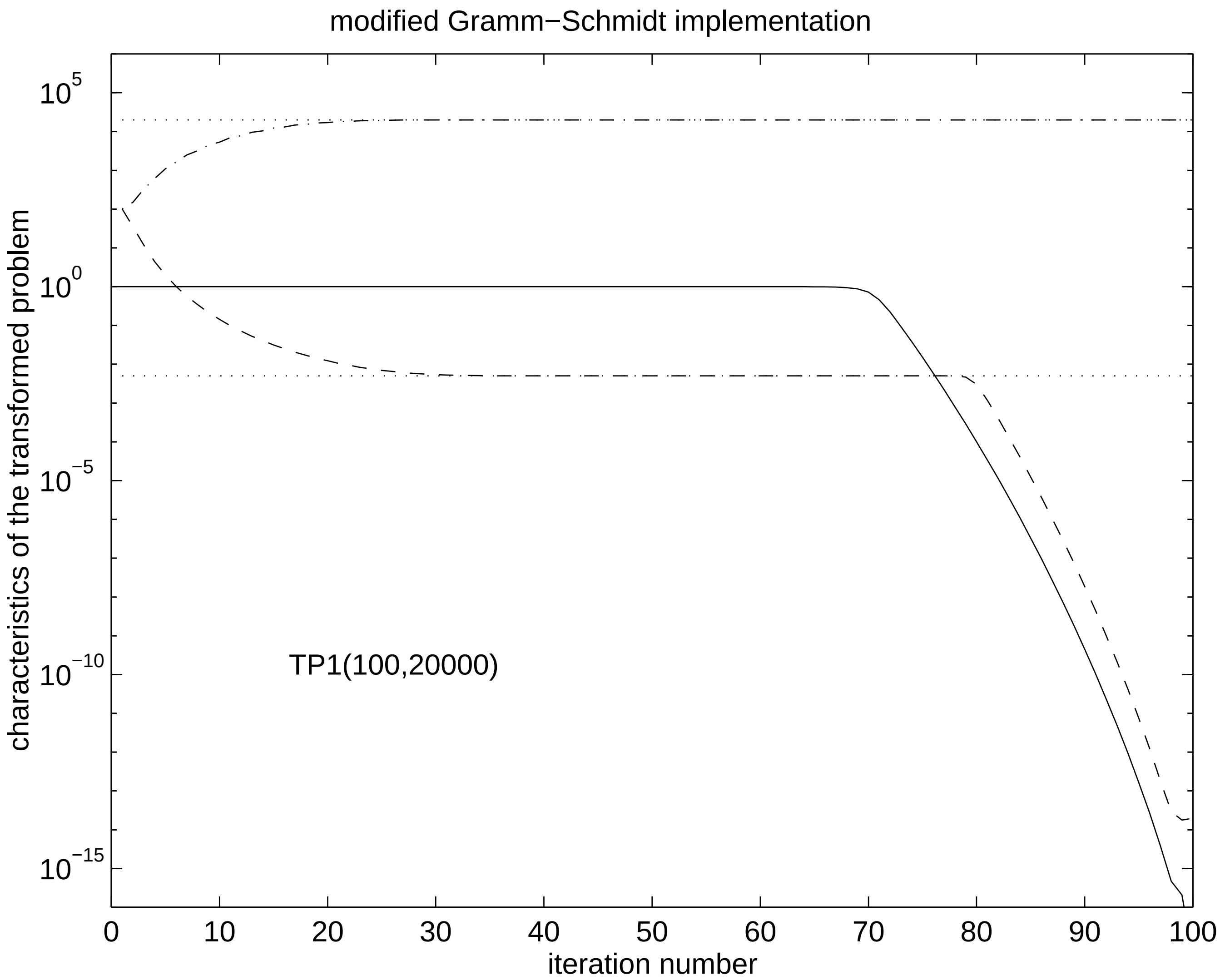
Greenbaum, R, Strakoš, 1997

Paige, R, Strakoš, 2006

GMRES WITH CGS ARNOLDI PROCESS

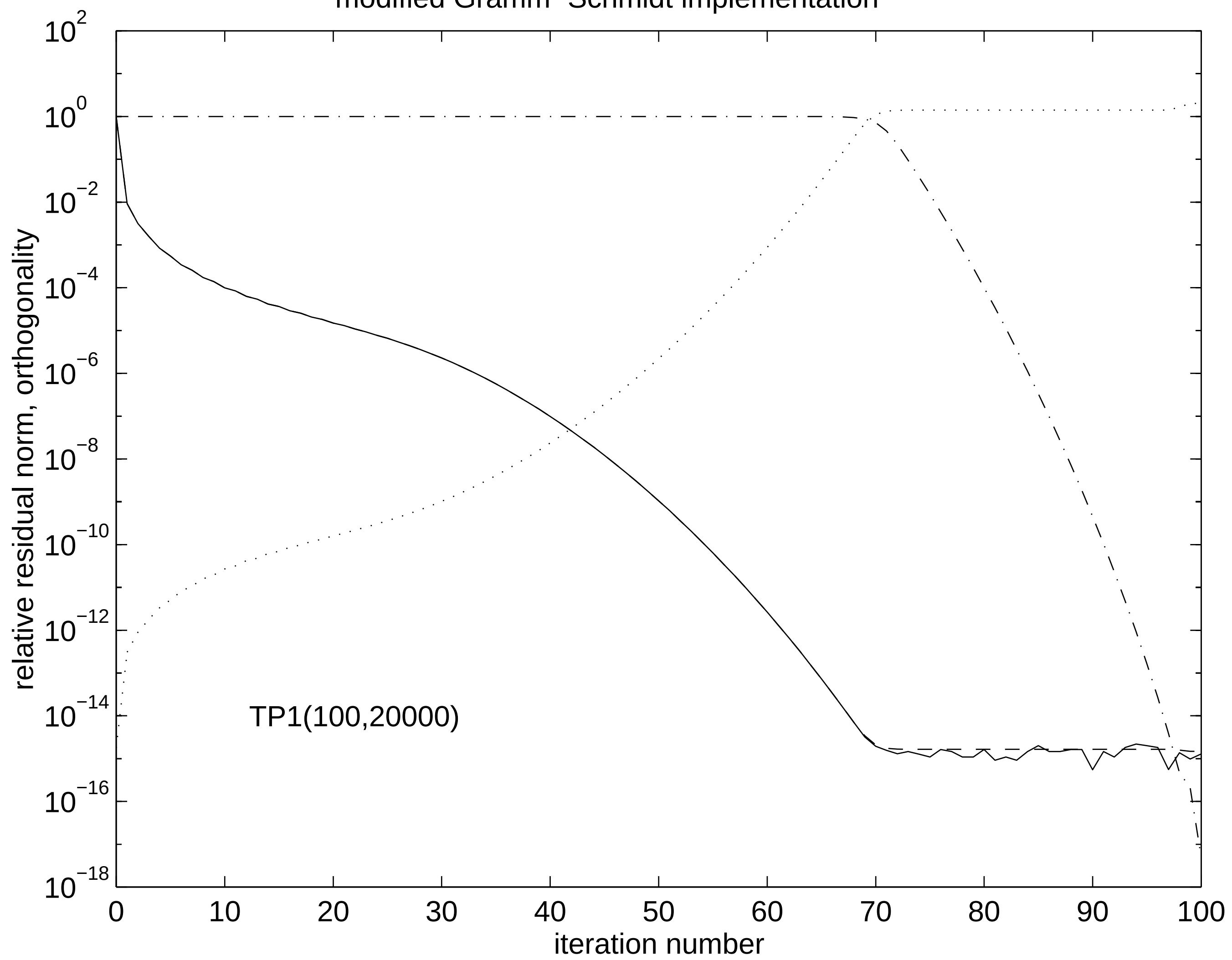
van den Eshof, Giraud, Langou, R, 2005

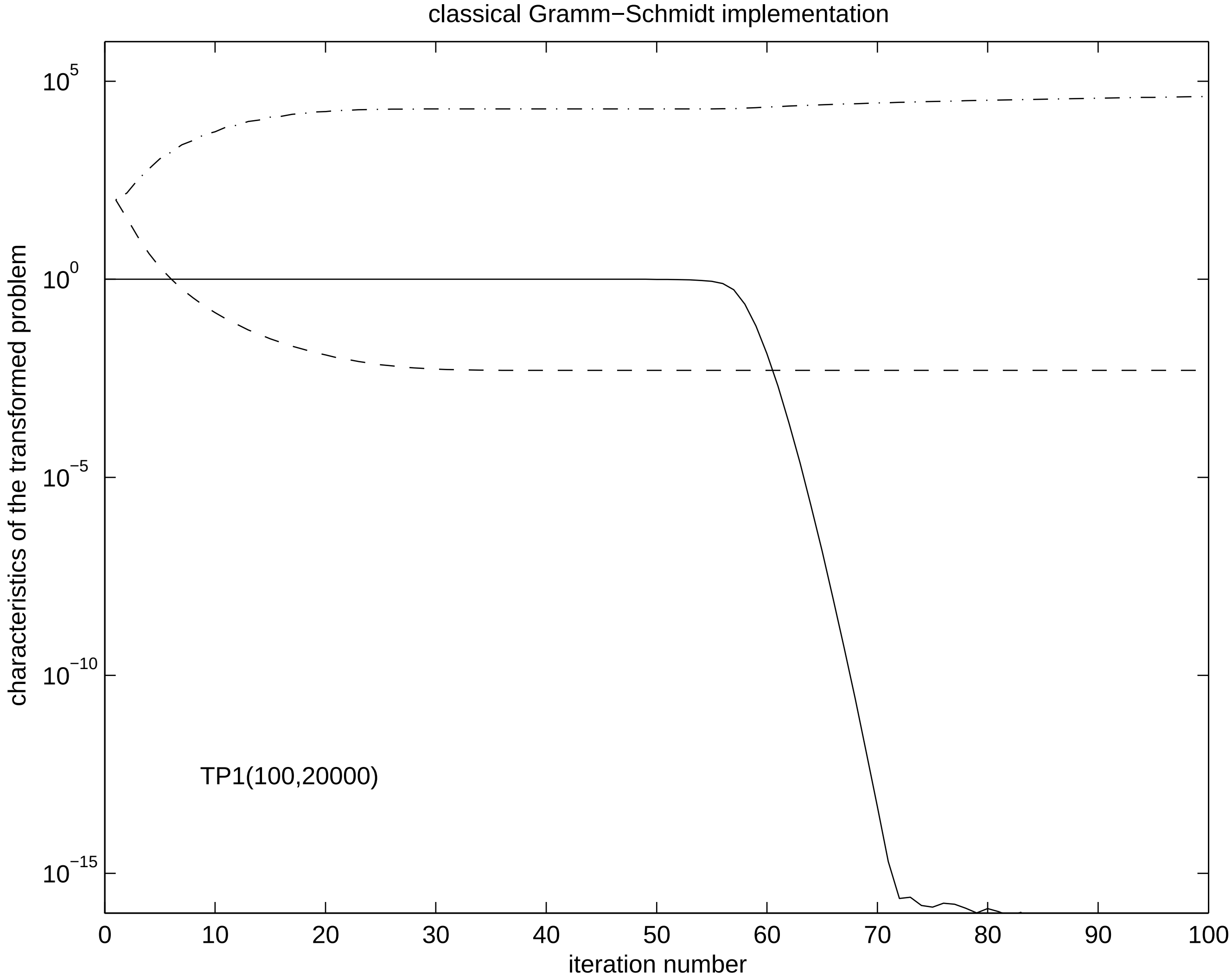
Smoktunowicz, Barlow, Langou, 2006



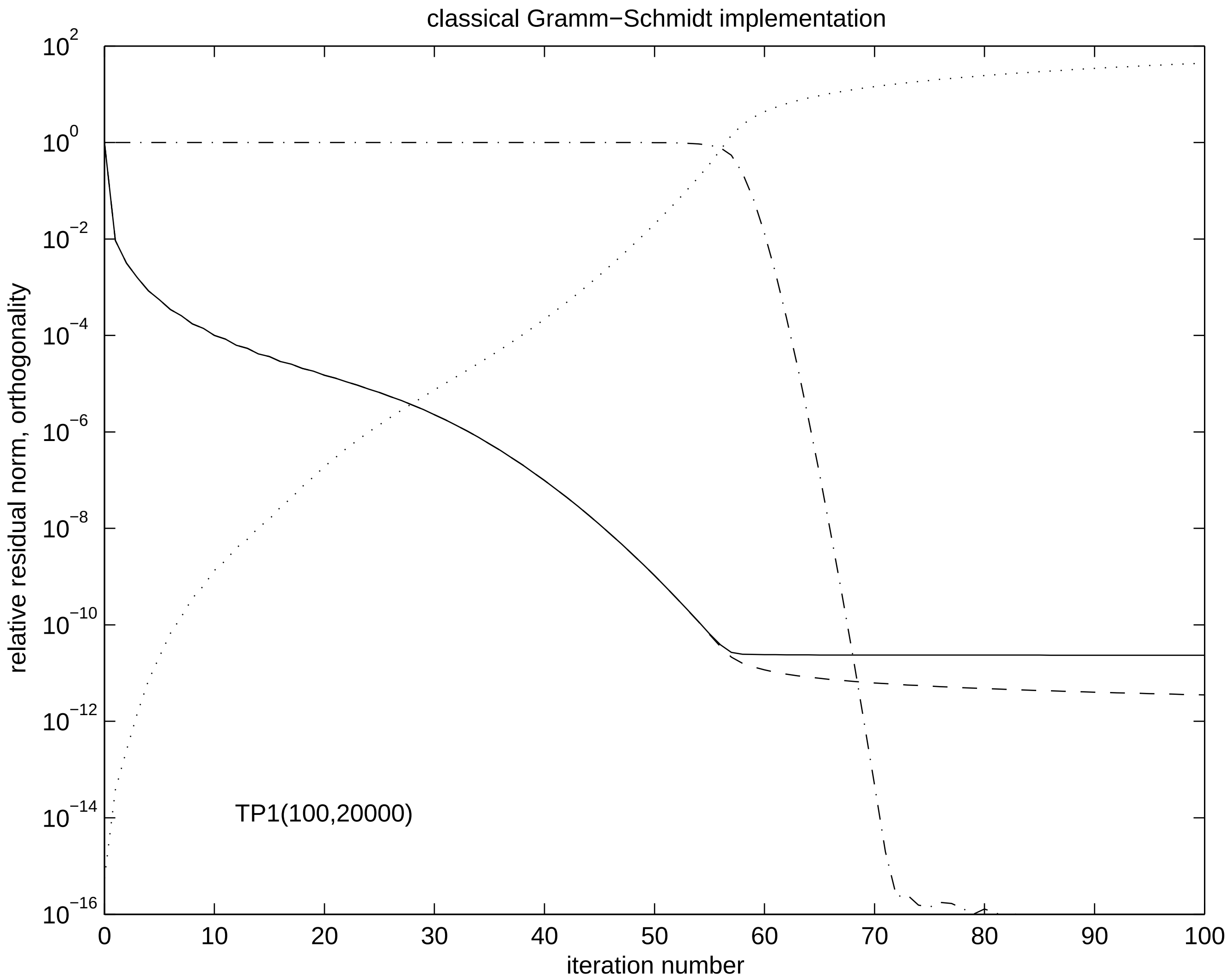
modified Gramm−Schmidt implementation
characteristics of the transformed problem
iteration number
TP1(100,20000)
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
0
10
20
30
40
50
60
70
80
90
100

modified Gramm−Schmidt implementation





classical Gramm−Schmidt implementation
characteristics of the transformed problem
iteration number
TP1(100,20000)
10^5
10^0
10^−5
10^−10
10^−15
0
10
20
30
40
50
60
70
80
90
100


classical Gramm−Schmidt implementation
relative residual norm, orthogonality
iteration number
TP1(100,20000)
10^2
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
10^−16
0
10
20
30
40
50
60
70
80
90
100

GMRES WITH MGS ARNOLDI PROCESS

The MGS-GMRES implementation is a **backward stable** iterative method.

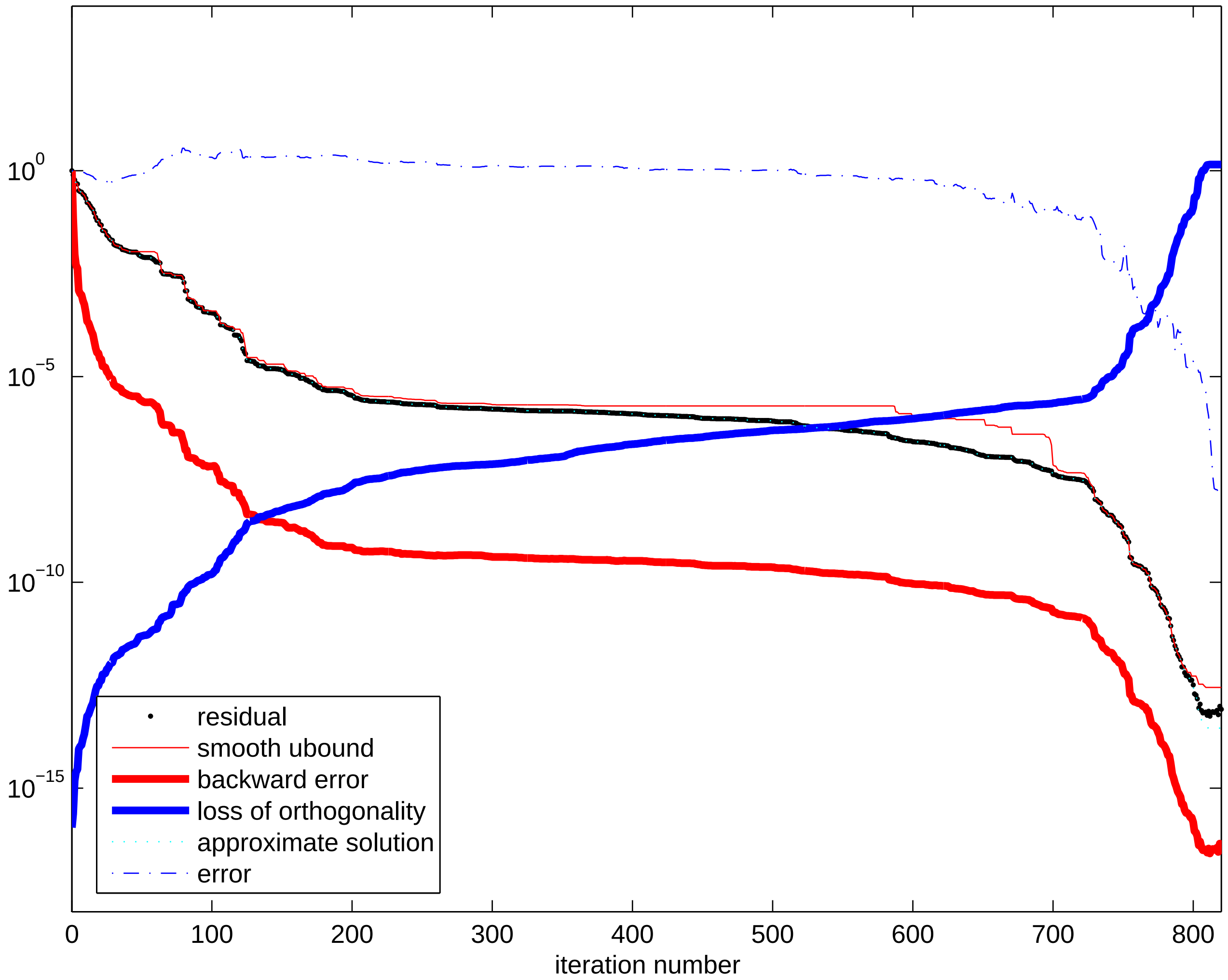
STATEMENT:

For some iteration step $n \leq N$ the computed approximate solution $\bar{x}_n$ satisfies

$$(A + \Delta A_n)\bar{x}_n = b + \Delta b_n$$

$$\|\Delta A_n\|/\|A\| \leq O(\varepsilon),\ \|\Delta b_n\|/\|b\| \leq O(\varepsilon)$$

Paige, R, Strakoš, 2006



10^{0}
10^{-5}
10^{-10}
10^{-15}
0
100
200
300
400
500
600
700
800
iteration number
residual
smooth ubound
backward error
loss of orthogonality
approximate solution
error

HOW TO MAKE SIMPLER GMRES AND GCR MORE STABLE

Miro Rozložník

Institute of Computer Science
Czech Academy of Sciences
CZ-182 07 Prague, Czech Republic
email: miro@cs.cas.cz

joint work with
P. Jiránek and M.H. Gutknecht

Seminarium Pierścienie, macierze i algorytmy numeryczne,
Politechnika Warszawska, Warszawa, November 13, 2009

MINIMUM RESIDUAL METHODS

$$x_n \in x_0 + K_n(A, r_0), \qquad r_n \equiv b - Ax_n$$

$$K_n(A, r_0) \equiv \mathrm{span}\{r_0, Ar_0, \ldots, A^{n-1}r_0\}$$

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

$$\Leftrightarrow$$

$$r_n \perp AK_n(A, r_0)$$

THE SIMPLER GMRES APPROACH

$[q_1, Z_{n-1}]$: a basis of $K_n(A, r_0)$, $q_1 \equiv r_0/\|r_0\|$, $\|Z_{n-1}e_k\| = 1$

V_n: an orthonormal basis of $AK_n(A, r_0)$, $V_n^T V_n = I$

$$A[q_1, Z_{n-1}] = V_n U_n, \quad V_n \equiv [v_1, \ldots, v_n]$$

$$r_n \perp AK_n(A, r_0) = \mathcal{R}(V_n)$$

$$\Updownarrow$$

$$r_n = (I - V_n V_n^T) r_0 = r_{n-1} - \alpha_n v_n,\ \alpha_n = \langle r_{n-1}, v_n \rangle.$$

$$x_n = x_0 + [q_1, Z_{n-1}] t_n, \quad U_n t_n = V_n^T r_0 = (\alpha_1, \ldots, \alpha_n)^T.$$

Walker, Zhou, 1994

ROUNDING ERROR ANALYSIS

- The QR decomposition:
  $A[q_1, Z_{n-1}] = V_n U_n + F_n,\ \|F_n\| \leq O(\varepsilon)\|A\|\|[q_1, Z_{n-1}]\|$

  Wilkinson, 1963, Björck, 1967

- Solution of the triangular system:
  $(U_n + \Delta U_n)\widehat{t}_n = (\alpha_1, \ldots, \alpha_n)^T,\ |\Delta U_n| \leq O(\varepsilon)|U_n|$

  Wilkinson, 1963

MAXIMUM ATTAINABLE ACCURACY: THE BACKWARD AND FORWARD ERROR

$$\frac{\|b - A\widehat{x}_n - r_n\|}{\|A\|\|\widehat{x}_n\|} \leq O(\varepsilon)\kappa([q_1, Z_{n-1}])\left(1 + \frac{\|x_0\|}{\|\widehat{x}_n\|}\right).$$

$$\frac{\|x_n - \widehat{x}_n\|}{\|x\|} \leq O(\varepsilon)\kappa(A)\kappa([q_1, Z_{n-1}])\frac{\|\widehat{x}_n\| + \|x_0\|}{\|x\|}.$$

THE CHOICE $[q_1, Z_{n-1}] = [q_1, V_{n-1}]$

The **simpler** approach $\equiv$ Simpler GMRES of Walker and Zhou,1994.

$[q_1, V_{n-1}]$ is of full column rank if $r_0 \notin A\mathcal{K}_{n-1}(A, r_0)$.

Conditioning of $[q_1, V_{n-1}]$ related to the convergence of residuals, Walker and Zhou, 1994, Liesen,R, Strakoš 2002

$$\frac{\|r_0\|}{\|r_{n-1}\|} \leq \kappa([q_1, V_{n-1}]) \leq 2\frac{\|r_0\|}{\|r_{n-1}\|}.$$

THE CHOICE $[q_1, Z_{n-1}] = \tilde{R}_n \equiv [\frac{r_0}{\|r_0\|}, \ldots, \frac{r_{n-1}}{\|r_{n-1}\|}]$

The **simpler** approach $\equiv$ SGMRES (new implementation)

$\tilde{R}_n$ is of full-column rank if
$r_0 \notin AK_{n-1}(A, r_0)$ and $\|r_0\| > \cdots > \|r_{n-1}\|$.

Conditioning of $\tilde{R}_n$ related to the stagnation of residuals:

$$\kappa(\tilde{R}_n) \leq n^{1/2}\gamma_n, \quad \gamma_n \equiv \sqrt{1 + \sum_{i=1}^{n-1} \frac{\|r_{i-1}\|^2 + \|r_i\|^2}{\|r_{i-1}\|^2 - \|r_i\|^2}}$$

THANK YOU FOR YOUR ATTENTION!

THE UPDATE APPROACH: ORTHODIR, ORTHOMIN, GCR, GMRESR

V_n: an orthonormal basis of $AK_n(A, r_0)$, $V_n^T V_n = I$

$$A[q_1, Z_{n-1}] = V_n U_n, \quad V_n \equiv [v_1, \ldots, v_n]$$

$P_n \equiv A^{-1} V_n$: $A^T A$-orthonormal basis of $K_n(A, r_0)$

$$[q_1, Z_{n-1}] = P_n U_n, \quad P_n \equiv [p_1, \ldots, p_n]$$

$$\Downarrow$$

$$\left.\begin{array}{ll} x_n & = x_{n-1} + \alpha_n p_n, \\ r_n & = r_{n-1} - \alpha_n v_n, \end{array}\right\} \quad \alpha_n = \langle r_{n-1}, v_n \rangle$$

Vinsome,1976, Young, Jea, 1980, Elman et al. 1983, van der Vorst, Vuik, 1990