Chapter five

Numerical behavior of GMRES

1. Orthogonality and numerical stability
2. Householder GMRES
3. Modified Gram-Schmidt GMRES
4. What should follow

5.1 Orthogonality and numerical stability

[Liesen, Rozložník, S - 02]: The choice of the subspace (basis) which is used in computations can have a fundamental impact.

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

$$\Leftrightarrow \;\; r_n \perp A\,K_n(A, r_0)\,.$$

The straightforward approach

Based on the orthonormal basis of $A\mathcal{K}_n(A, r_0)$. Define $w_1 \equiv Ar_0/\|Ar_0\|$, $v_1 = r_0/\|r_0\|$. Then the recursive columnwise QR-factorization yields

$$[Av_1,\, AW_{n-1}] \;=\; A\,[v_1,\, W_{n-1}] \;\equiv\; W_n\, R_n\,,$$

$$W_n \;\equiv\; [w_1, \ldots, w_n], \quad W_n^T W_n \;=\; I_n\,,$$

$$\text{span}\,\{w_1, \ldots, w_n\} \;=\; A\mathcal{K}_n(A, r_0)\,,$$

$$\kappa([v_1, W_{n-1}]) \;\le\; \kappa(R_n) \;\le\; \kappa(A)\,\kappa([v_1, W_{n-1}])\,.$$

Using the orthonormal basis $w_1, \ldots, w_n$ of $A\mathcal{K}_n(A, r_0)$:

$$x_n = x_0 + [v_1, W_{n-1}]\, t_n \in x_0 + \mathcal{K}_n(A, r_0);$$

$$\Rightarrow \quad r_n = r_0 - A\,[v_1, W_{n-1}]\, t_n = r_0 - W_n R_n\, t_n;$$

$$\Rightarrow \quad t_n = (W_n R_n)^{+}\, r_0 = R_n^{-1}\, W_n^T\, r_0 .$$

How does this affect the numerical stability?

Theorem

$$\frac{\|r_n\|}{\|r_0\|} = \sigma_{n+1}([v_1, W_n])\,\sigma_1([v_1, W_n]) = \frac{2\kappa([v_1, W_n])}{\kappa([v_1, W_n])^2 + 1},$$

$$\frac{\|r_0\|}{\|r_n\|} \leq \kappa([v_1, W_n]) \leq 2\,\frac{\|r_0\|}{\|r_n\|},$$

$$\frac{\|r_0\|}{\|r_n\|} \leq \kappa(R_n) \leq 2\kappa(A)\,\frac{\|r_0\|}{\|r_n\|}.$$

Consequently, $\kappa(R_n)$ must inevitably increase as $\|r_n\|$ decreases, even for small $\kappa(A)$ and with the most stable way of computing $w_1, \ldots, w_n$.

$\Rightarrow$ Computation of $t_n = R_n^{-1} W_n^T r_0$ is inherently unstable!

Surprise: Numerical behaviour gets worse when orthogonality of $w_1, \ldots, w_n$ is maintained better; Householder implementation performs worse than Modified-Gram-Schmidt implementation.

The straightforward approach is used in "Simpler GMRES" [Walker, Lu Zhou - 94], and it is related to other implementations, e.g. Orthodir.

Classical GMRES implementation

Based on the orthonormal basis of $\mathcal{K}_n(A, r_0)$. Let $v_1 \equiv r_0/\|r_0\|$. Then the Arnoldi process yields

$$AV_n = V_{n+1} H_{n+1,n},$$

$$V_n \equiv [v_1, \dots, v_n], \quad V_n^T V_n = I_n,$$

$$\text{span}\,\{v_1, \dots, v_n\} = \mathcal{K}_n(A, r_0),$$

$$\kappa(H_{n+1,n}) \le \kappa(A).$$

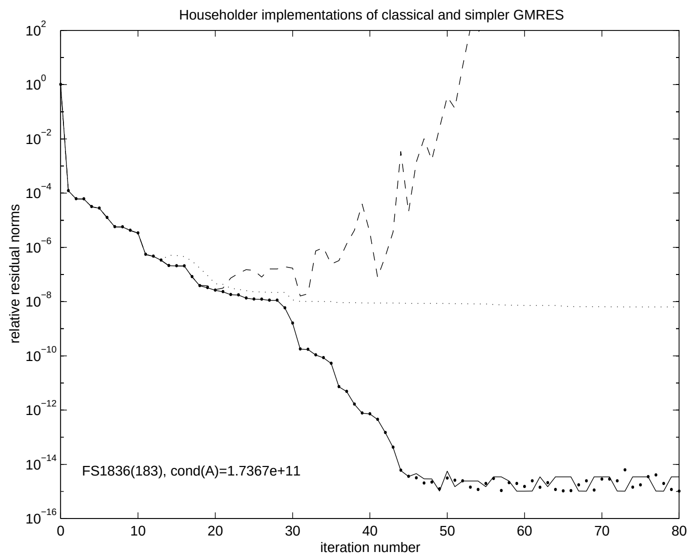
Using the orthonormal basis $v_1, \ldots, v_n$ of $\mathcal{K}_n(A, r_0)$:

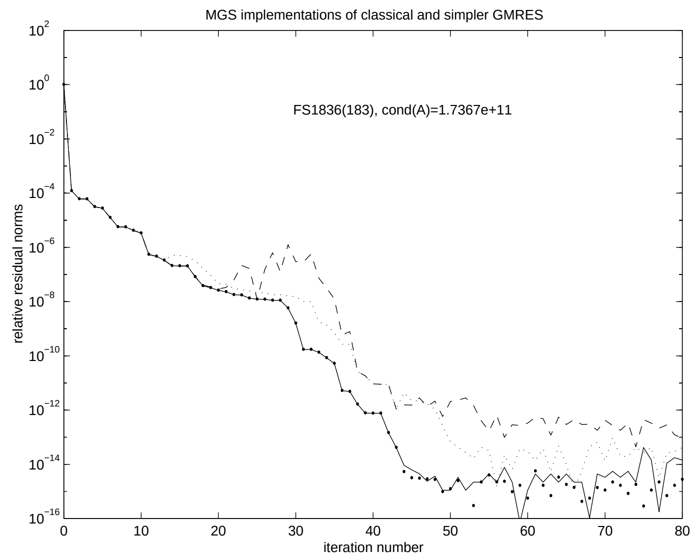
$$x_n = x_0 + V_n z_n \in \mathcal{K}_n(A, r_0),$$

$$\Rightarrow \quad r_n = r_0 - AV_n z_n = r_0 - V_{n+1} H_{n+1,n} z_n,$$

$$\Rightarrow \quad z_n = (V_{n+1} H_{n+1,n})^+ r_0 = (H_{n+1,n})^+ V_{n+1}^T r_0.$$

How does this affect the numerical stability?





Householder implementations (residual and backward error). Excessive ill-conditioning of computed R_n leads in the straightforward implementation to divergence.

MGS implementations (residual and backward error). Due to rounding errors the identity is violated, and the computed R_n is not so badly ill-conditioned as it ideally should be!

Moral

The choice of the right subspace (basis) is fundamental for numerical stability of the Krylov subspace methods.

Even the best orthogonalization technique in computing the basis (here Householder reflections) can not compensate for instabilities artificially created due to a bad choice of the subspace. Paradoxically - preserving orthogonality of the computed basis can even make things worse!

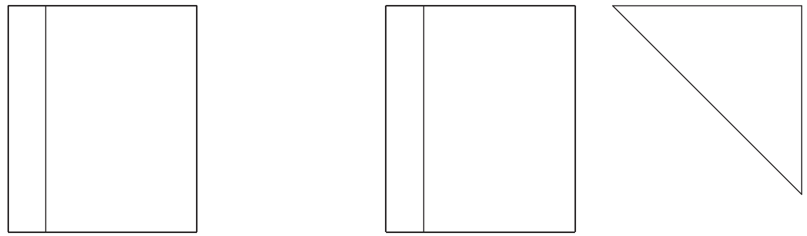
In the rest of the text we concentrate on classical GMRES.

Numerical stability of the GMRES implementation based on the (ideally) orthonormal basis of $K_n(A, r_0)$ – many related publications, e.g.

[Björck - 67], [Björck, Paige - 92], [Karlson - 91], [Arioli, Fassino - 96], [Drkošová, Greenbaum, Rozložník, S - 95], [Greenbaum, Rozložník, S - 96], [Rozložník 1997], [Paige, S - 02, 02, 02], [Giraud, Langou, Rozložník, van der Eshof - 05], [Rozložník, Paige, S - in progress]

Arnoldi process $\equiv$ recursive columnwise QR decomposition

$$[r_0, AV_n] \quad = \quad V_{n+1} \quad R_{n+1}$$



5.2 Householder GMRES

Implementation using Householder reflections was suggested in [Walker 1988, 89].

A tedious but straightforward proof in [DGRS -95]:

Householder - reflections based implementation of GMRES computes a backward stable approximate solution.

5.3 Modified Gram-Schmidt GMRES

$$
\begin{aligned}
v_{n+1} &= (I - v_n v_n^* - \ \ldots \ - v_1 v_1^*)\, A v_n \\
&= (I - v_n v_n^*)\ \ldots\ (I - v_1 v_1^*)\, A v_n
\end{aligned}
$$

A common general belief: MGS orthogonalization is a good compromise between propagation of errors (loss of orthogonality) and algorithm efficiency (computational cost). Price - the computation is recursive!

Comparison with classical Gram-Schmidt, Householder reflections, Givens rotations.

However:

Despite the loss of orthogonality, some algorithms with MGS provide results as good as algorithms using the most stable orthogonalization processes (with the loss of orthogonality among the basis vectors kept close to the machine precision level).

Theoretical justification?

- Linear least squares: [Björck, Paige - 92];

- Our case: MGS GMRES.

In MGS GMRES, loss of orthogonality is **controlled by convergence**.

Statement:

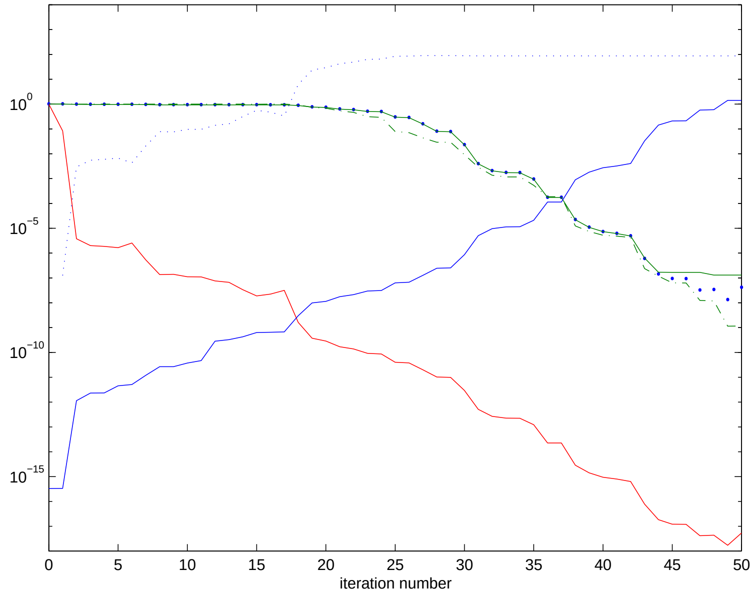
Loss of orthogonality $\|I - V_{n+1}^* V_{n+1}\|_F$ in the modified Gram–Schmidt Arnoldi process

is inversely proportional

to the value of the **GMRES backward error**

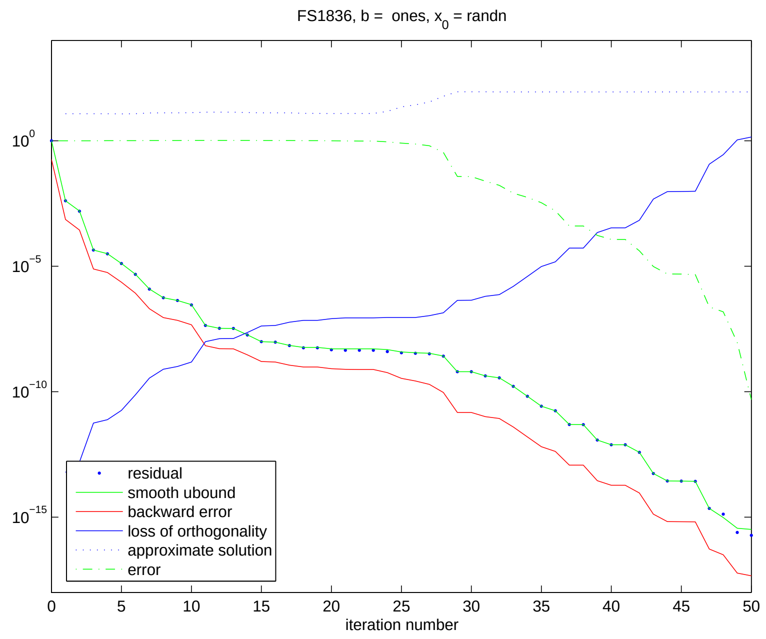
$$\frac{\|b - Ax_n\|}{\|b\| + \|A\| \|x_n\|}\,.$$

FS1836, $b =$ ones, $x_0 = 0$

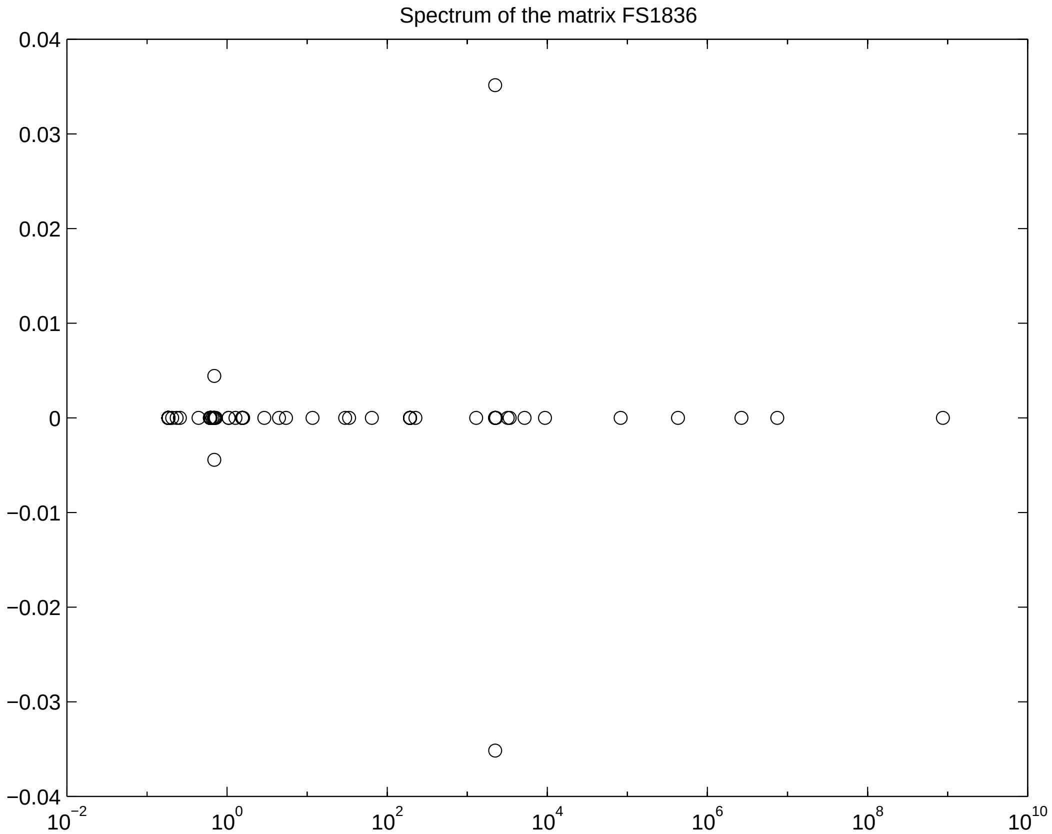


See the difference between the relative residual and backward error (it is interesting to see the same for $x_0 \neq 0$).

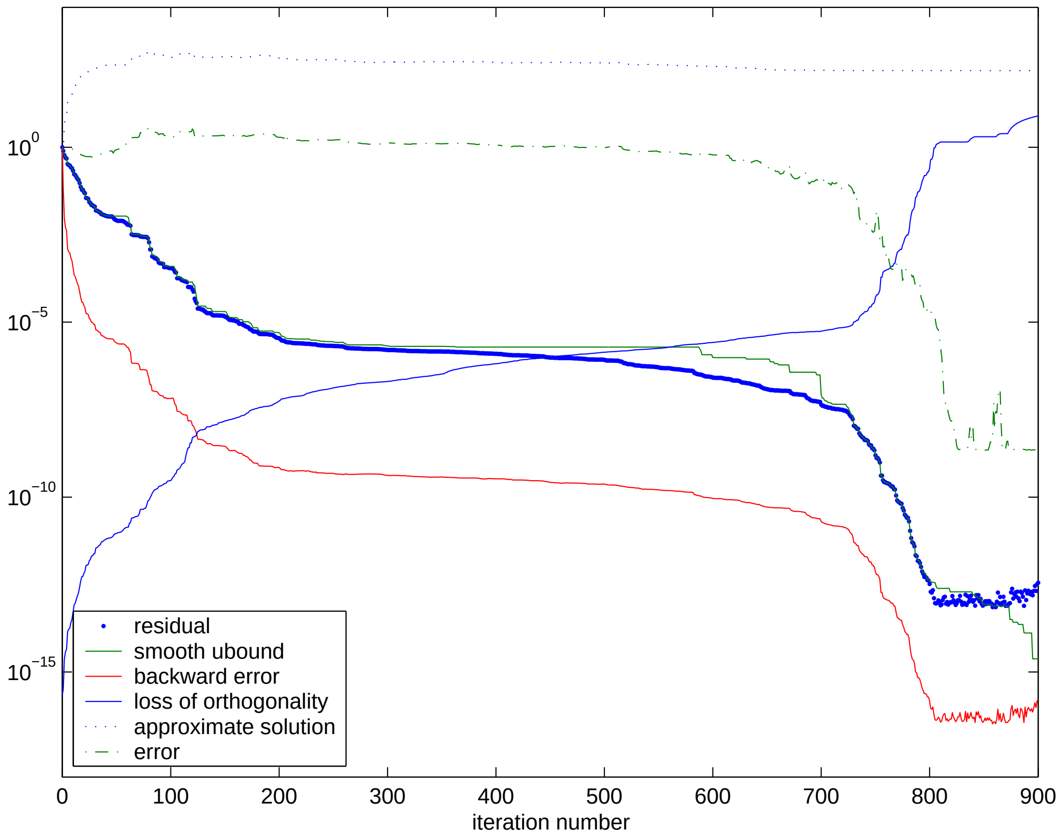
FS1836, $b =$ ones, $x_0 =$ randn



link the spectrum with convergence ?

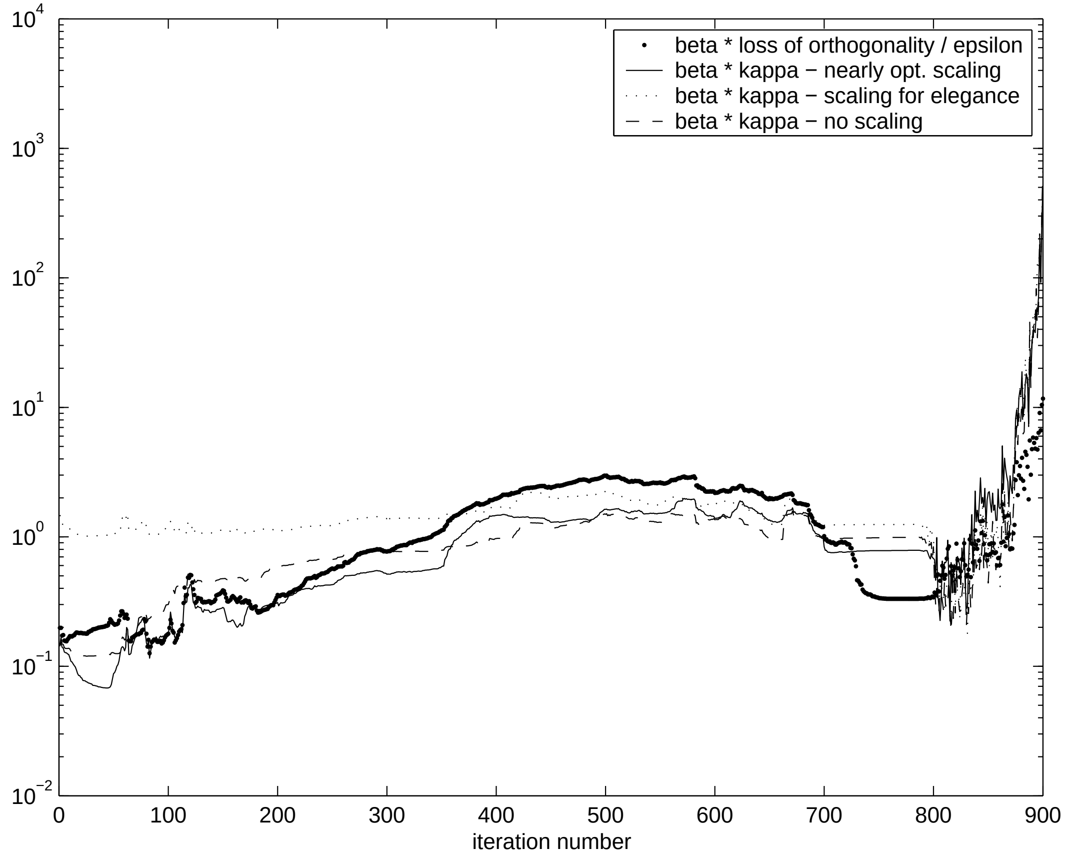


Sherman 2, b MM, $x_0 = 0$



Proof? A long and yet unfinished story, links with **Scaled Total Least Squares** and other strange problems.

Sherman 2, b MM, $x_0 = 0$



5.4 What should follow

Loss of orthogonality in MGS Arnoldi

$$\|I - V_{k+1}^T V_{k+1}\|_F \; \approx \; \kappa([r_0\gamma, AV_k D_k])\, \mathcal{O}(\varepsilon)\,.$$

Loss of orthogonality in CGS Arnoldi (could be based on recent result which will appear in [GLRvdE -05])

$$\|I - V_{k+1}^T V_{k+1}\|_F \; \approx \; \kappa([r_0\gamma, AV_k D_k])^2\, \mathcal{O}(\varepsilon)\,.$$

Numerical stability of restarted CGS GMRES.