On the residual norm in FOM and GMRES

Gérard MEURANT

June 2010

Introduction

We solve

$$Ax = b$$

with non symmetric Krylov solvers : FOM or GMRES

We would like to obtain expressions for the (l_2) residual norms

We will show that the decrease of the residual norm depends on a submatrix of the Hessenberg matrix H_k

On suppose que H_k est non singulière, $\sigma_{min}(H_k) > 0$

FOM and GMRES : notations

Let A be of order n

$$x^k = x^0 + V_k z^k$$

where V_k is the Arnoldi orthogonal basis given by

$$AV_k = V_k H_k + h_{k+1,k} v^{k+1} (e^k)^T$$

H_k is an upper Hessenberg matrix

$$\text{FOM}: H_k z^k = V_k^T r^0 = \|r^0\| V_k^T v^1 = \|r^0\| e^1$$

$$\text{GMRES}: \min_z \| \|r^0\| e^1 - H_k^{(e)} z \|$$

where

$$H_k^{(e)} = \begin{pmatrix} H_k \\ h_{k+1,k} (e^k)^T \end{pmatrix}$$

The FOM residual norm I

Theorem

Assuming that H_k is nonsingular, the norm of the residual in FOM is given by

$$\|r^k\|^2 = \|r^0\|^2 h_{k+1,k}^2 (H_k^{-1} e^1, e^k)^2$$

Proof.

$$\begin{aligned} r^k &= b - Ax^k \\ &= b - A(x^0 + V_k z^k) \\ &= r^0 - (V_k H_k z^k + h_{k+1,k} v^{k+1} (e^k)^T z^k) \\ &= -h_{k+1,k} z_k^k v^{k+1} \end{aligned}$$

but $z^k = \|r^0\| H_k^{-1} e^1$

□

The FOM residual norm II

Theorem

Assuming that H_k is nonsingular, let

$$H_k = \begin{pmatrix} (h^{k-1})^* & h_{1,k} \\ \tilde{H}_{k-1} & w^{k-1} \end{pmatrix}$$

Then

$$(H_k^{-1} e^1, e^k) = \frac{1}{h_{1,k} - (h^{k-1}, \tilde{H}_{k-1}^{-1} w^{k-1})}$$

$$h_{k+1,k}(H_k^{-1} e^1, e^k) = \frac{1}{(e^k, \tilde{H}_k^{-*} h^k)}$$

Proof. Compute an LU factorization of a row permutation of H_k □

Note that $\tilde{H}_k$ is upper triangular

Theorem

Assuming that H_k is nonsingular, the FOM the residual norm is given by

$$\|r^k\|^2 = \frac{\|r^0\|^2}{(e^k, \tilde{H}_k^{-*} h^k)^2}$$

Hence the FOM residual norm is small if and only if $(e^k, \tilde{H}_k^{-*} h^k)^2$ is large

Bounds for the FOM residual norm

Theorem

$$\|r^0\|^2 \frac{[\sigma_{min}(\tilde{H}_k)]^2}{\|h^k\|^2} \leq \|r^0\|^2 \frac{1}{\|\tilde{H}_k^{-T} h^k\|^2} \leq \|r^k\|^2$$

$$\leq \|r^0\|^2 \, h_{k+1,k}^2 \frac{[\sigma_{min}(\tilde{H}_{k-1})]^2}{[\sigma_{min}(H_k)]^2 \, [\sigma_{k-1}(H_k)]^2}.$$

The GMRES residual norm

Theorem

$$\|r_M^k\|^2 = \frac{\|r^0\|^2}{1 + \|\tilde{H}_k^{-*} h^k\|^2}$$

This result was first proven by E. Ayachour, J. Comp. and Appl. Math., v 159, n 2, (2003)

We have a simpler proof

Skip the proof

Proof

$$(H_k^{(e)})^* H_k^{(e)} = \tilde{H}_k^* \tilde{H}_k + h^k (h^k)^*$$

Moreover, we have $(H_k^{(e)})^* e^1 = h^k$. Hence

$$z^k = [\tilde{H}_k^* \tilde{H}_k + h^k (h^k)^*]^{-1} h^k$$

Let $\tilde{z}_k = z^k / \|r^0\|$ and $\xi_k = 1/(1 + \|\tilde{H}_k^{-*} h^k\|^2)$, using the Sherman-Morrison formula we obtain

$$\tilde{z}_k = \xi_k \tilde{H}_k^{-*} h^k$$

$$\begin{aligned} \|r^k\| &= \|r^0\| \left\| e^1 - \begin{pmatrix} (h^k)^* \\ \tilde{H}_k \end{pmatrix} \tilde{z}^k \right\| \\ &= \|r^0\| \left\| \begin{pmatrix} 1 - (h^k)^* \tilde{z}^k \\ -\tilde{H}_k \tilde{z}^k \end{pmatrix} \right\| \end{aligned}$$

Proof 2

$$\frac{\|r^k\|^2}{\|r^0\|^2} = (1-(h^k)^*\tilde{z}^k)^2 + \|\tilde{H}_k\tilde{z}^k\|^2$$

But we have

$$\begin{aligned} 1-(h^k)^*\tilde{z}^k &= 1-(h^k)^*\tilde{H}_k^{-1}[1-\xi_k\|\tilde{H}_k^{-*}h^k\|^2]\tilde{H}_k^{-*}h^k \\ &= 1-[1-\xi_k\|\tilde{H}_k^{-*}h^k\|^2]\,\|\tilde{H}_k^{-*}h^k\|^2 \\ &= 1-\xi_k\|\tilde{H}_k^{-*}h^k\|^2 \\ &= \xi_k \end{aligned}$$

It implies that

$$\frac{\|r^k\|^2}{\|r^0\|^2} = \xi_k^2 + \xi_k^2\|\tilde{H}_k^{-*}h^k\|^2 = \xi_k$$

Summary

FOM :

$$\|r_O^k\|^2 = \frac{\|r^0\|^2}{(e^k, \tilde{H}_k^{-*} h^k)^2}$$

GMRES :

$$\|r_M^k\|^2 = \frac{\|r^0\|^2}{1 + \|\tilde{H}_k^{-*} h^k\|^2}$$

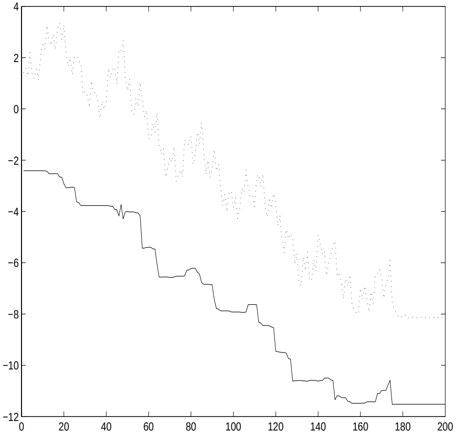
Bounds for the GMRES residual norm

Theorem

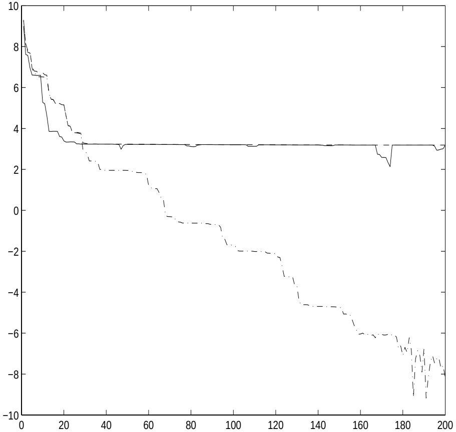
$$\|r^0\|^2 \frac{[\sigma_{min}(\tilde{H}_k)]^2}{[\sigma_{min}(\tilde{H}_k)]^2 + \|h^k\|^2} \leq \|r_M^k\|^2 \leq \|r_O^k\|^2$$

First example : Steam2

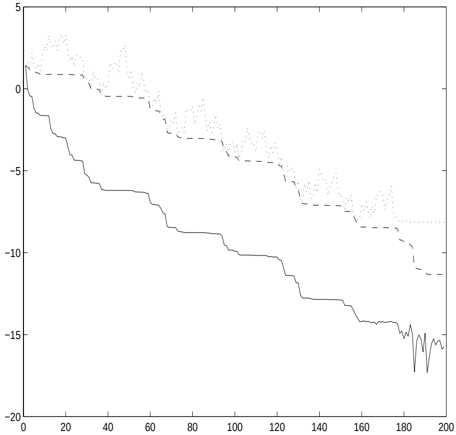
$n = 600,\ \kappa(A) = 3.78\ 10^6$



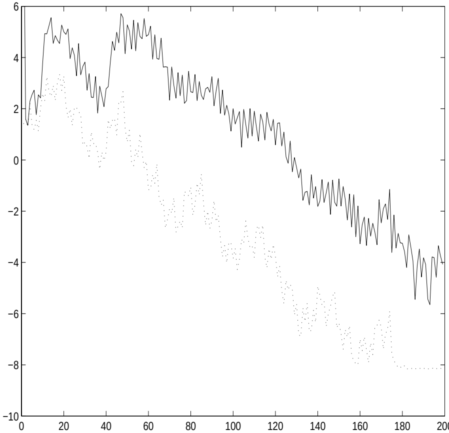
E1 : FOM, $\log_{10}$ of the norm of the error (plain), residual (dotted)



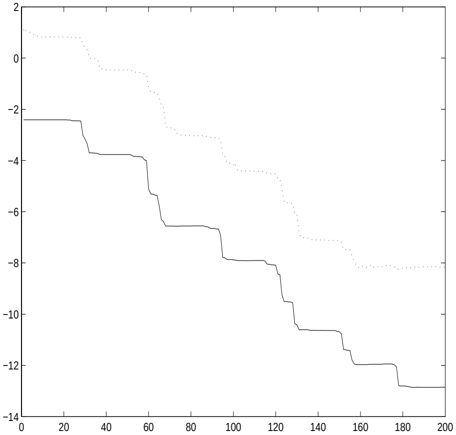
E1 : FOM, $\log_{10}$ of $\sigma_{min}(H_k)$ (plain), $\sigma_{min}(H_k^{(e)})$ (dashed) and $\sigma_{min}(\tilde{H}_k)$ (dot–dashed)



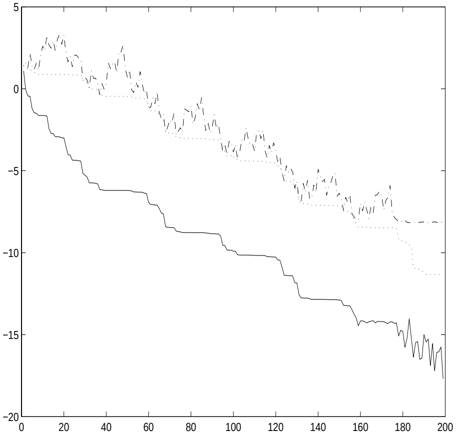
E1 : FOM, $\log_{10}$ of $\|r^k\|$ (dotted), flbound1 (plain) and flbound2 (dashed)



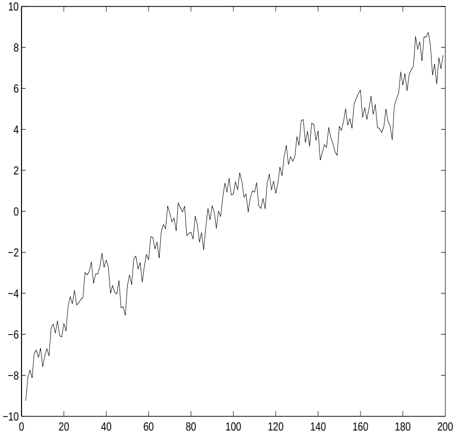
E1 : FOM, $\log_{10}$ of $\|r^k\|$ (dotted) and fubound (plain)



E1 : GMRES, $\log_{10}$ of the norm of the error (plain), residual (dotted)



E1 : GMRES, $\log_{10}$ of $\|r^k\|$ (dotted), glbound (plain) and $\|r_O^k\|$ (dot–dashed)



E1 : $\log_{10}$ of the norm of last column of the inverse of $\tilde{H}_k$

Second example

From Liesen and Strakoš, SUPG bilinear finite element discretization of

$$-\nu\Delta u + w \cdot \nabla u = 0$$

with $w = [0,1]^T$ in $\Omega = (0,1)^2$ with Dirichlet b.c.

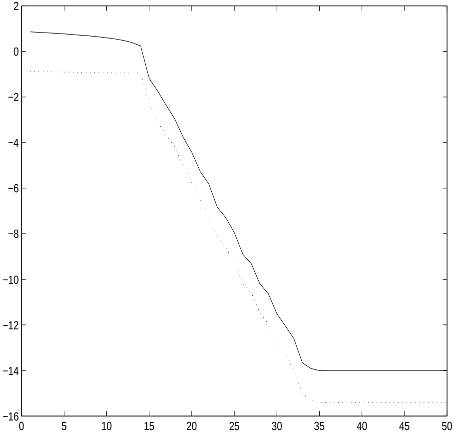
$$A = \nu K \otimes M + M \otimes ((\nu + \delta h)K + C)$$

$$M = \frac{h}{6}\mathrm{tridiag}(1,4,1), \;\; K = \frac{1}{h}\mathrm{tridiag}(-1,2,-1)$$

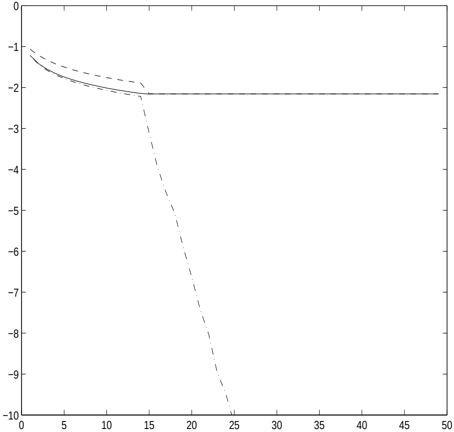
$$C = \frac{1}{2}\mathrm{tridiag}(-1,0,-1)$$

δ is the regularization parameter, h is the mesh size

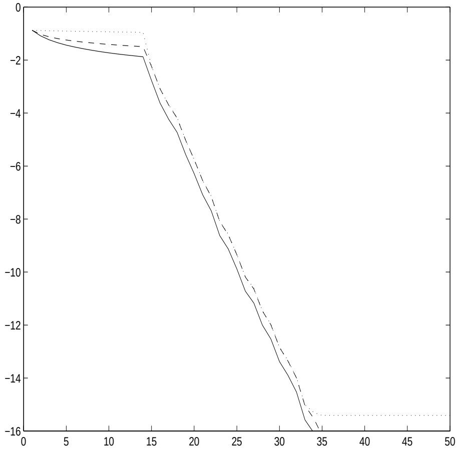
$h = 1/16 \Rightarrow n = 225, \nu = 0.01$ and $\delta = 0.34$



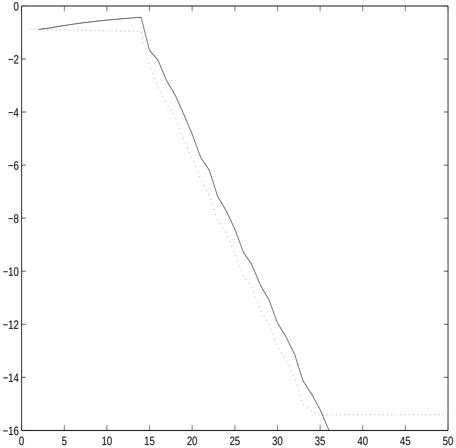
E2 : FOM, $\log_{10}$ of the norm of the error (plain), residual (dotted)



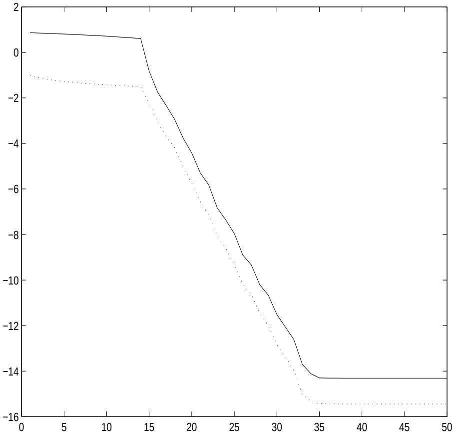
E2 : FOM, $\log_{10}$ of $\sigma_{min}(H_k)$ (plain), $\sigma_{min}(H_k^{(e)})$ (dashed) and $\sigma_{min}(\tilde{H}_k)$ (dot–dashed)



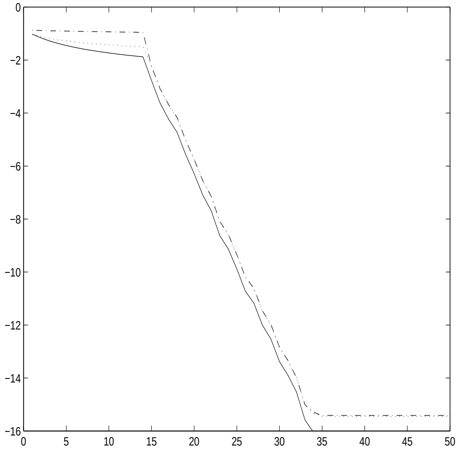
E2 : FOM, $\log_{10}$ of $\|r^k\|$ (dotted), flbound1 (plain) and flbound2 (dashed)



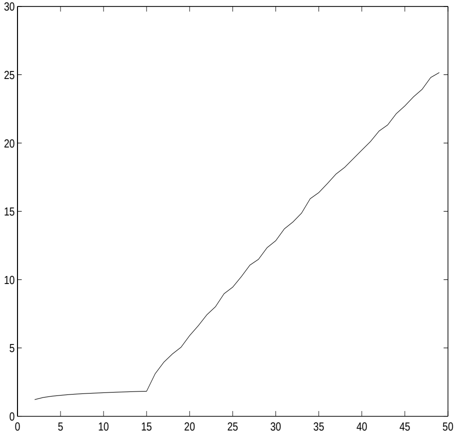
E2 : FOM, $\log_{10}$ of $\|r_M^k\|$ (dotted) and ubound (plain)



E2 : GMRES, $\log_{10}$ of the norm of the error (plain), residual (dotted)



E2 : GMRES, $\log_{10}$ of $\|r_M^k\|$ (dotted), glbound (plain) and $\|r_O^k\|$ (dot–dashed)



E2 : $\log_{10}$ of the norm of last column of the inverse of $\tilde{H}_k$

The main question

When is $\sigma_{min}(\tilde{H}_k) \to 0$?