An optimal Q-OR Krylov method for solving linear systems

Gérard MEURANT

March 2016

Many Krylov methods have been proposed over the years for solving linear systems $Ax = b$

Most of them can be classified as quasi-orthogonal (Q-OR) or quasi-minimum residual (Q-MR)

Q-OR: FOM, BiCG, Hessenberg, ...

Q-MR: GMRES, truncated GMRES, QMR, CMRH, ...

Whatever their definition, these methods share many fundamental properties

See M. Eiermann and O.G. Ernst, *Geometric aspects in the theory of Krylov subspace methods*, Acta Numerica, v 10 n 10 (2001), pp. 251–312

They differ by the basis of the Krylov space that is constructed:

- orthogonal for FOM/GMRES,
- bi-orthogonal for BiCG/QMR,
- based on an LU factorization for Hessenberg/CMRH

What do we know about GMRES?

Let

$$K = \begin{pmatrix} b & Ab & A^2b & \cdots & A^{n-1}b \end{pmatrix}$$

be the Krylov matrix that we assume of full rank. Then

$$K = VU$$

with V orthogonal (or unitary) and U upper triangular with positive real diagonal entries

The matrix $H = V^*AV$ is upper Hessenberg

We have

$$H = UCU^{-1}$$

where C is the companion matrix for the eigenvalues of A

Let x_k^G (resp. x_k^F) be the iterates for GMRES (resp. FOM) and the residual vectors $r_k^G = b - Ax_k^G$, $r_k^F = b - Ax_k^F$

We assume $x_0 = 0$ and $\|b\| = 1$

We know that

- every residual norm convergence curve is possible for GMRES (and FOM)
- $|(U^{-1})_{1,k}| = 1/\|r_{k-1}^F\|$
- $\|r_k^G\|^2 = 1/(\mathcal{M}_{k+1}^{-1})_{1,1}$ with $\mathcal{M}_{k+1} = U_{k+1}^* U_{k+1} = K_{k+1}^* K_{k+1}$
- one can construct matrices A with a prescribed spectrum and right-hand sides b such that GMRES yields a prescribed decreasing residual norm convergence curve
- we have two parametrizations of this class of matrices

For all these properties see

A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent advances in iterative methods, G.H. Golub, A. Greenbaum and M. Luskin, eds., Springer, (1994), pp. 95–118

A. Greenbaum, V. Pták and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., v 17 (1996), pp. 465–469

M. Arioli, V. Pták and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT Numerical Mathematics, v 38 n 4 (1998), pp. 636–643

J. Duintjer Tebbens and G. Meurant, *Any Ritz value behavior is possible for Arnoldi and for GMRES*, SIAM J. Matrix Anal. Appl., v 33 n 3 (2012), pp. 958–978

G. Meurant and J. Duintjer Tebbens, *The role eigenvalues play in forming GMRES residual norms with non-normal matrices*, Numerical Algorithms, v 68, n 1 (2015), pp. 143–165

Q-OR and Q-MR methods

We assume that we have a basis V of the Krylov space (with columns of unit norm) such that $K = VU$ with V nonsingular and U upper triangular

We define $H = UCU^{-1}$. As a consequence $AV = VH$. The iterates are

$$x_k = V_k y^{(k)}$$

where V_k is the matrix of the k first columns of V. The residual r_k is

$$V_k e_1 - AV_k y^{(k)} = V_k(e_1 - H_k y^{(k)}) - h_{k+1,k} y_k^{(k)} v_{k+1} = V_{k+1}(e_1 - \underline{H}_k y^{(k)})$$

The Q-OR method is defined (provided that H_k is nonsingular) by

$$H_k y^{(k)} = e_1$$

This annihilates the first term in the residual

In the Q-MR method $y^{(k)}$ is computed as the solution of the least squares problem

$$\min_y \|e_1 - \underline{H}_k y\|$$

where $\underline{H}_k$ is $(k+1) \times k$. The vector $z_k^M = e_1 - \underline{H}_k y^{(k)}$ is referred as the quasi-residual

The residual vector is $r_k^M = V_{k+1} z_k^M$

Generally, the two problems are solved using Givens rotations with sines s_j. It is known that

$$\|z_k^M\| = |s_1 s_2 \cdots s_k|$$

Moreover we have a relation between the Q-OR residual norms and the Q-MR quasi-residual norms

$$\frac{1}{\|r_k^O\|^2} = \frac{1}{\|z_k^M\|^2} - \frac{1}{\|z_{k-1}^M\|^2}$$

Properties of Q-OR and Q-MR methods

From these results we can show by induction that

$$|(U^{-1})_{1,k}| = \frac{1}{\|r^O_{k-1}\|}$$

The inverses of the Q-OR residual norms can be read from the first row of the inverse of U

A consequence of this result is the following

Let $M_{k+1} = U^*_{k+1} U_{k+1} \neq K^*_{k+1} K_{k+1}$. Then

$$\|z^M_k\|^2 = \frac{1}{(M^{-1}_{k+1})_{1,1}}$$

The difference with GMRES is that we only have the norm of the quasi-residual

For all these properties and more see,

G. Meurant and J. Duintjer Tebbens, *On the convergence of Q-OR and Q-MR Krylov methods for solving nonsymmetric linear systems*, BIT Numerical Mathematics, to appear (2016)

Construction of “good” bases

We would like to find bases which lead to a “good” convergence of the Q-OR method

The matrix V of the basis is related to the Krylov matrix K by $K = VU$ with U upper triangular

The entries of the first row of U^{-1} are the inverses of the Q-OR residual norms

Constructing a “good” basis may seem easy since one can think that we can just construct any upper triangular matrix U^{-1} with entries of large modulus on the first row

But, it is not so since the columns of V have to be of unit norm

The inverse of U_k is given recursively as

$$U_k^{-1} = \begin{pmatrix} U_{k-1}^{-1} & -\frac{1}{u_{k,k}} U_{k-1}^{-1} U_{1:k-1,k} \\ 0 & \frac{1}{u_{k,k}} \end{pmatrix}$$

Let $\nu_{i,j}$ be the entries of U^{-1}
At step k we already know the entries of U_{k-1}^{-1}. We would like to choose the entries $u_{j,k}$ of column k of U to maximize the quantity

$$\frac{1}{u_{k,k}} \left| \sum_{j=1}^{k-1} \nu_{1,j} u_{j,k} \right|$$

However, the columns of V have to be of unit norm and, using $K = VU$ we must have

$$u_{k,k} = \left\| \sum_{j=1}^{k-1} u_{j,k-1} A v_j - \sum_{j=1}^{k-1} u_{j,k} v_j \right\|$$

At step k the first sum is known

Something simpler

A first step and a way to approximately obtain large entries on the first row of of U^{-1} is to have small diagonal entries $u_{k,k}$

We choose $u_{2,2} = \|Av_1 - u_{1,2}v_1\|$ and we would like to pick $u_{1,2}$ to minimize $u_{2,2}$
This is a least squares problem whose solution is given by

$$u_{1,2} = \frac{v_1^T A v_1}{v_1^T v_1} = v_1^T A v_1$$

It yields $v_2 = \frac{1}{u_{2,2}}[Av_1 - (v_1^T A v_1)v_1]$
It follows that

$$v_1^T v_2 = \frac{1}{u_{2,2}} v_1^T [Av_1 - (v_1^T A v_1)v_1] = 0$$

Therefore we have v_2 orthogonal to v_1

At the next step we choose $u_{1,3}$ and $u_{2,3}$ to minimize
$\|u_{2,2}Av_2 + u_{1,2}Av_1 - u_{2,3}v_2 - u_{1,3}v_1\|$
This is a least squares problem $\min \|b - By\|$ with

$$b = u_{2,2}Av_2 + u_{1,2}Av_1, \quad B = \begin{pmatrix} v_1 & v_2 \end{pmatrix}, \quad y = \begin{pmatrix} u_{1,3} \\ u_{2,3} \end{pmatrix}$$

The normal equations are

$$B^T By = \begin{pmatrix} 1 & v_1^T v_2 \\ v_2^T v_1 & 1 \end{pmatrix} y = B^T b = \begin{pmatrix} v_1^T [u_{2,2}Av_2 + u_{1,2}Av_1] \\ v_2^T [u_{2,2}Av_2 + u_{1,2}Av_1] \end{pmatrix}$$

But $B^T B = I$ since $v_2 \perp v_1$ and we obtain $y = B^T b$

$$v_3 = \frac{1}{u_{3,3}}[b - By] = \frac{1}{u_{3,3}}[b - BB^T b]$$

Let us consider $v_1^T v_3$. We have $v_1^T B = [1\ 0]$, therefore

$$v_1^T BB^T b = v_1^T [u_{2,2} A v_2 + u_{1,2} A v_1] = v_1^T b$$

Hence, $v_1^T v_3 = 0$. Similarly $v_2^T v_3 = 0$

And we can go on...

At every step (in exact arithmetic) the matrices $B^T B$ in the least squares problems are identity matrices

Minimizing the diagonal entries of U yields an orthonormal basis

Then Q-OR $\rightarrow$ FOM, Q-MR $\rightarrow$ GMRES

This is optimal for Q-MR

What about Q-OR?

We can try directly computing U^{-1} from $V = KU^{-1}$

In this way we obtain the vectors v_j straightforwardly, but, again, the columns of V have to be of unit norm

Let $\nu_{i,j}$ be the entries of U^{-1} and

$$v_k = \nu_{1,k} v + \nu_{2,k} A^2 v + \cdots + \nu_{k,k} A^{k-1} v$$

We would like to have $\|v_k\| = 1$ and $|\nu_{1,k}|$ as large as possible

Can we solve this problem?

Let $\tilde{\nu}$ be the vector of the components $\nu_{i,k}, i = 1, \ldots, k$

We want $\|K_k\tilde{\nu}\| = 1$. This corresponds to

$$\tilde{\nu}^T K_k^T K_k \tilde{\nu} = \tilde{\nu}^T \mathcal{M}_k \tilde{\nu} = 1$$

This is the equation of an (hyper) ellipsoid in $\mathbb{R}^k$ centered at the origin

We have to find a point on the surface of this ellipsoid with a maximum of the absolute value of the first coordinate

Let us consider $k = 2$. The matrix is

$$\mathcal{M}_2 = \begin{pmatrix} 1 & v^T A v \\ v^T A^T v & v^T A^T A v \end{pmatrix}$$

and the quadratic form is
$x^2 + (v^T A v + v^T A^T v)xy + (v^T A^T A v)y^2 - 1 = 0$ with $x = \nu_{1,2}$
and $y = \nu_{2,2}$
We need to maximize $|x|$ on the ellipse

Let us write the equation as $x^2 + 2\alpha xy + \beta y^2 - 1 = 0$
We find that the solution is

$$x = \pm \left(\frac{\beta}{\beta - \alpha^2} \right)^{\frac{1}{2}}$$

The value of y is

$$y = \pm \frac{\alpha}{\sqrt{\beta(\beta - \alpha^2)}}$$

We remark that $x = \pm\sqrt{(\mathcal{M}_2^{-1})_{1,1}}$

More generally the solution is obtained by writing the equation of a tangent hyperplane and asking that it is orthogonal to the x-axis

One can show that a solution is $x = \sqrt{(\mathcal{M}_k^{-1})_{1,1}}$ and the other components are obtained by solving a linear system of order $k-1$ whose matrix and right-hand side are $\mathcal{M}_{2:k,2:k}$ and $-x\mathcal{M}_{2:k,1}$

This yields U^{-1}. If we apply Q-OR with the basis $V = KU^{-1}$ we obtain residual vectors whose norms are

$$\|r_k^O\|^2 = \frac{1}{(\mathcal{M}_{k+1}^{-1})_{1,1}}$$

These values are those that are obtained from GMRES
Therefore, they are the best ones that we can get with the given Krylov subspace. In a sense it defines an optimal Q-OR method

Avoiding the use of U

The previous construction is not practical because

1) we do not want to compute $\mathcal{M}_k^{-1}$
2) in many cases the matrix U is almost singular and must be (numerically) avoided

Instead we would like to construct H column by column. We have

$$H_j = U_j E_j U_j^{-1} + \begin{pmatrix} 0 & \cdots & 0 & \frac{1}{u_{j,j}} U_{1:j,j+1} \end{pmatrix}$$

It yields

$$\sum_{j=1}^{k+1} \nu_{1,j} h_{j,k} = 0 \;\Rightarrow\; \nu_{1,k+1} = -\frac{1}{h_{k+1,k}} \sum_{j=1}^{k} \nu_{1,j} h_{j,k}$$

At the first step, we have

$$\nu_{1,2} = -\frac{1}{h_{2,1}}\nu_{1,1}h_{1,1}$$

But, $\nu_{1,1} = 1$ and $\nu_{1,2} = -h_{1,1}/h_{2,1}$ with $h_{2,1} = \|Av_1 - h_{1,1}v_1\|$

We want to find the minimum of the inverse

$$\frac{\|Av_1 - y\, v_1\|}{|\nu_{1,1}y|}$$

with $\nu_{1,1} = 1$. We consider the square of this ratio and we look for a lower bound γ

$$\frac{v_1^T A^T A v_1 - 2y\, v_1^T A v_1 + y^2}{y^2} \geq \gamma$$

We would like to find the largest possible value of γ such that

$$v_1^T A^T A v_1 - 2y\, v_1^T A v_1 + y^2(1-\gamma) \geq 0$$

In matrix form this can be written as

$$\begin{pmatrix} y & 1 \end{pmatrix} \begin{pmatrix} 1-\gamma & -v_1^T A v_1 \\ -v_1^T A v_1 & v_1^T A^T A v_1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} \geq 0$$

The value

$$\gamma_{opt} = 1 - \frac{(v_1^T A v_1)^2}{v_1^T A^T A v_1}$$

yields

$$\begin{pmatrix} 1-\gamma & -v_1^T A v_1 \\ -v_1^T A v_1 & v_1^T A^T A v_1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} = 0$$

The solution is

$$h_{1,1} = y = \frac{v_1^T A^T A v_1}{v_1^T A v_1}$$

and $h_{2,1} = \|Av_1 - h_{1,1} v_1\|$

More generally, at step k we wish to find the largest γ satisfying

$$\frac{1}{|\nu_{1,k+1}|^2} = \frac{\|b - By\|^2}{(\nu^T y)^2} \geq \gamma$$

with

$$b = Av_k, \quad B = \begin{pmatrix} v_1 & \cdots & v_k \end{pmatrix} = V_k, \quad y = \begin{pmatrix} h_{1,k} & \cdots & h_{k,k} \end{pmatrix}^T,$$

$$\nu = \begin{pmatrix} \nu_{1,1} & \cdots & \nu_{1,k} \end{pmatrix}$$

Since $(\nu^T y)^2 = y^T \nu \nu^T y$, in matrix form we have

$$\begin{pmatrix} y^T & 1 \end{pmatrix} \begin{pmatrix} B^T B - \gamma \nu \nu^T & -B^T b \\ -b^T B & b^T b \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} \geq 0$$

The value

$$\gamma_{opt} = \frac{\alpha}{\alpha \nu^T (B^T B)^{-1} \nu + \omega^2} = \frac{1}{\nu^T (B^T B)^{-1} \nu + \frac{\omega^2}{\alpha}}.$$

with $\alpha = b^T b - b^T B (B^T B)^{-1} B^T b \geq 0$ and $\omega = b^T B (B^T B)^{-1} \nu$ yields the solution

$$(B^T B - \gamma_{opt} \nu \nu^T) y = B^T b$$

for which we have

$$\begin{pmatrix} B^T B - \gamma_{opt} \nu \nu^T & -B^T b \\ -b^T B & b^T b \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} = 0$$

Using the Sherman-Morrison formula, we can write

$$y = (B^TB)^{-1}B^Tb + \frac{\gamma}{1-\gamma\nu^T(B^TB)^{-1}\nu}[\nu^T(B^TB)^{-1}B^Tb](B^TB)^{-1}\nu$$

Hence, to compute y we have to solve two $k \times k$ linear systems for $s = (B^TB)^{-1}B^Tb$ and $p = (B^TB)^{-1}\nu$

Finally $h_{k+1,k} = \|b - By\|$, $\nu_{1,k+1} = -\nu^Ty/h_{k+1,k}$ and $v_{k+1} = (b - By)/h_{k+1,k}$

We can do this by computing incrementally the Cholesky factorization of $B^TB = V_k^TV_k = L_kL_k^T$

The Q-OR optimal algorithm (first version)

Let A be real and $v_k^A = Av_k$

1- $L_{k-1}\ell_k = V_{k-1}^T v_k$ and $\ell_{k,k} = \sqrt{1 - \ell_k^T \ell_k}$

$$L_k = \begin{pmatrix} L_{k-1} & 0 \\ \ell_k & \ell_{k,k} \end{pmatrix}$$

2 - solve for p

$$q_k = \frac{1}{\ell_{k,k}} \left[\nu_{1,k} - \ell_k^T \begin{pmatrix} q_1 \\ \vdots \\ q_{k-1} \end{pmatrix} \right]$$

$$p_k = \frac{1}{\ell_{k,k}^2} \left[\nu_{1,k} - \ell_k^T \begin{pmatrix} q_1 \\ \vdots \\ q_{k-1} \end{pmatrix} \right]$$

$$L_{k-1}^T \begin{pmatrix} p_1 \\ \vdots \\ p_{k-1} \end{pmatrix} = \begin{pmatrix} q_1 \\ \vdots \\ q_{k-1} \end{pmatrix} - p_k \ell_k$$

3 - solve of $L_k L_k^T s = w = V_k^T v_k^A$
4 - $\omega = w^T p$ and $\alpha = (v_k^A)^T v_k^A - w^T s$
5 -

$$H_{1:k,k} = \begin{pmatrix} h_{1,k} \\ \vdots \\ h_{k,k} \end{pmatrix} = s + \frac{\alpha}{\omega} p$$

6 -

$$h_{k+1,k} = \left\| v_k^A - V_k H_{1:k,k} \right\|, \; \nu_{1,k+1} = -\frac{1}{h_{k+1,k}} \nu^T H_{1:k,k}$$

7 - $v_{k+1} = \frac{1}{h_{k+1,k}} [v_k^A - V_k H_{1:k,k}]$ and $v_{k+1}^A = A v_{k+1}$

Using Q-OR with this basis must yield the same residual norms as GMRES

However, computing this basis is more costly than Arnoldi

The basis being non-orthogonal, we have more inner products to compute, but they are all independent

The triangular solves (which are recursive) can also slow down the computation

But we can modify the algorithm to directly compute the inverses of the triangular Cholesky factors

The Q-OR optimal algorithm (second version)

Let $v_k^A = Av_k$

1- $v_k^V = V_{k-1}^T v_k, \;\; v_k^{tA} = V_k^T v_k^A$

2- $\ell_k = \tilde{L}_{k-1} v_k^V$

3- $y_k^T = \ell_k^T \tilde{L}_{k-1}, \; (p_k^v)^T = y_k^T V_{k-1}^T$

4- $\ell_{k,k} = \| v_k - p_k^v \|$

5-
$$\tilde{L}_k = \begin{pmatrix} \tilde{L}_{k-1} & 0 \\ -\frac{1}{\ell_{k,k}} y_k^T & \frac{1}{\ell_{k,k}} \end{pmatrix}$$

6- $\ell_\nu = \tilde{L}_k \nu, \;\; \ell_A = \tilde{L}_k v_k^{tA}$

7- $\omega = \ell_A^T \ell_\nu, \;\; \alpha = (v_k^A)^T v_k^A - \ell_A^T \ell_A$

8 -
$$H_{1:k,k} = \begin{pmatrix} h_{1,k} \\ \vdots \\ h_{k,k} \end{pmatrix} = \tilde{L}_k^T \left[\ell_A + \frac{\alpha}{\omega} \ell_\nu \right]$$

9 -

$$\tilde{v} = v_k^A - V_k\, h_{1:k,k},\ \ h_{k+1,k} = \|\tilde{v}\|,\ \ \nu_{1,k+1} = -\frac{1}{h_{k+1,k}}\nu^T h_{1:k,k}$$

$$\nu = \begin{pmatrix} \nu_{1,1} & \cdots & \nu_{1,k+1} \end{pmatrix}$$

10 - $v_{k+1} = \frac{1}{h_{k+1,k}}\tilde{v}$

When we have H_k we solve $H_k y^{(k)} = e_1$ and $x_k = V_k y^{(k)}$

In this algorithm almost everything is expressed in terms of matrix-matrix or matrix-vector products

A first numerical experiment

From Liesen and Strakoš, SIAM J. Sci. Comput., v 26 n 6 (2005)

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

with $w = [0,1]^T$ in $\Omega = (0,1)^2$ with Dirichlet boundary conditions

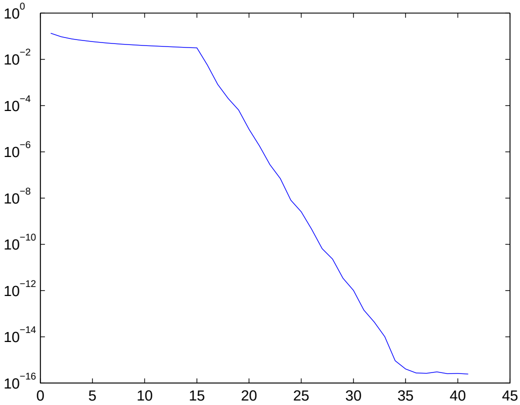
Stabilized Petrov–Galerkin SUPG method with bilinear finite elements on a regular Cartesian mesh

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

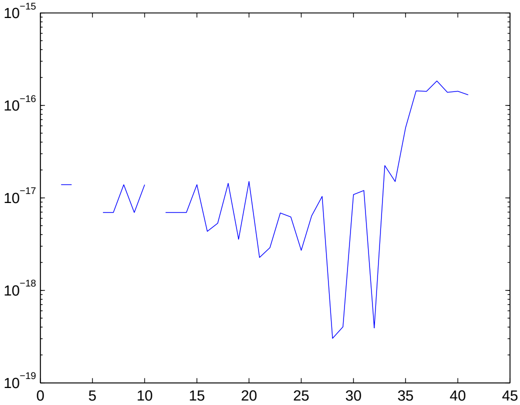
where δ is the stabilization parameter and

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

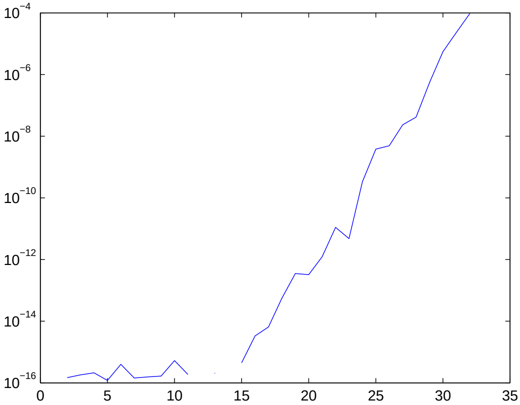
$$C = \frac{1}{2}\mathrm{tridiag}(-1,0,-1)$$



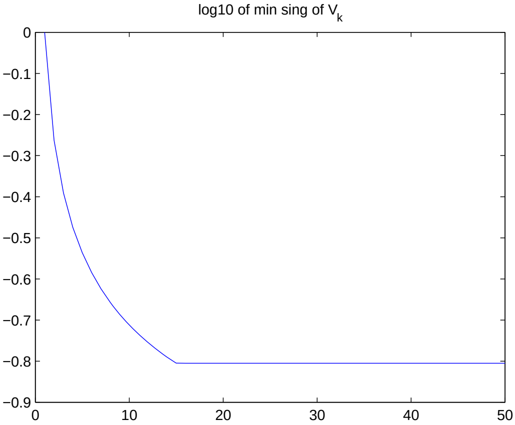
Residual norms for GMRES-MGS, SUPG, $n = 225$



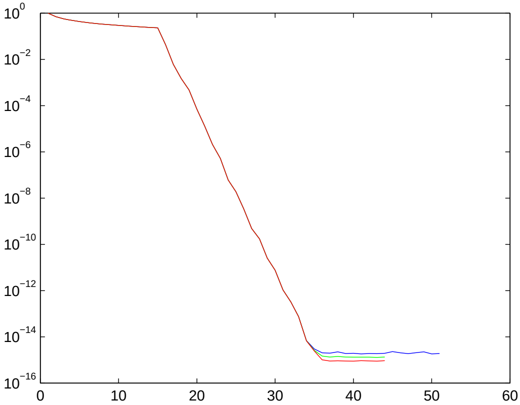
Difference of the residual norms of GMRES-MGS and Q-OR optimal, SUPG, $n = 225$



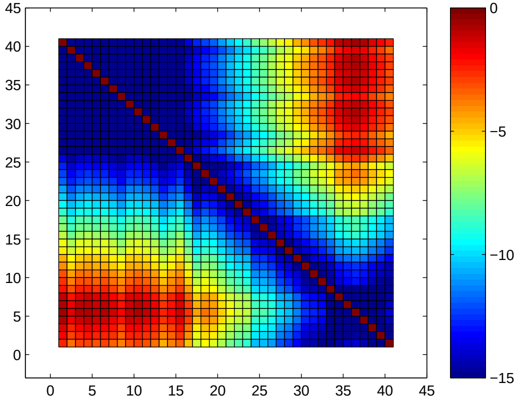
Relative difference of the residual norms of GMRES-MGS and Q-OR optimal, SUPG, $n = 225$



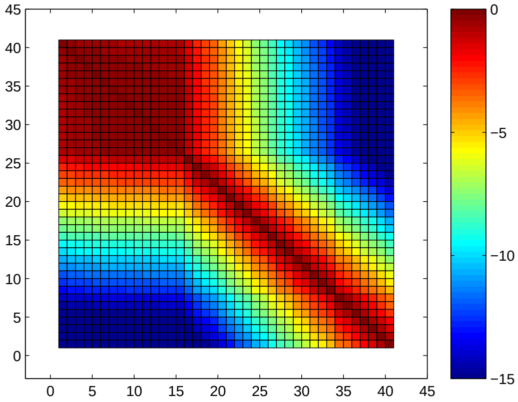
Minimum singular value of V_k, Q-OR optimal, SUPG, $n = 225$



Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), SUPG, $n = 225$

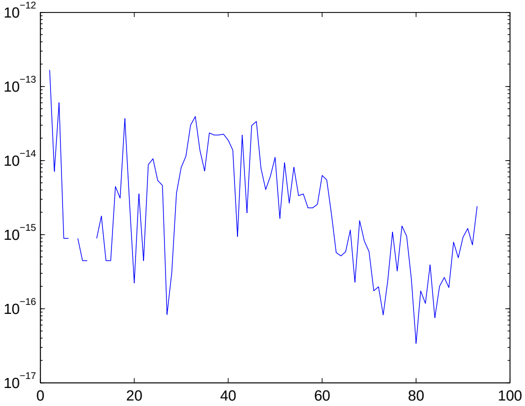


$\log_{10}$ of $|V^T V|$, GMRES-MGS (Arnoldi) without reorthogonalization

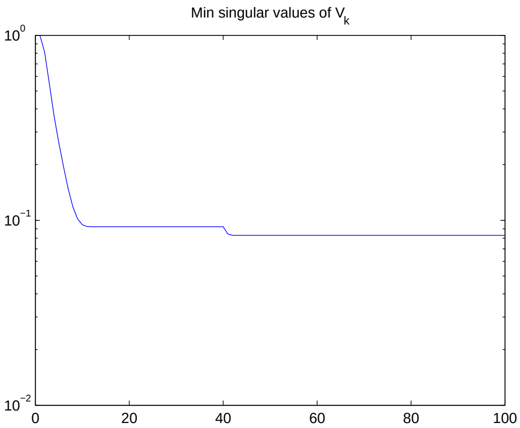


$\log_{10}$ of $|V^T V|$, Q-OR optimal

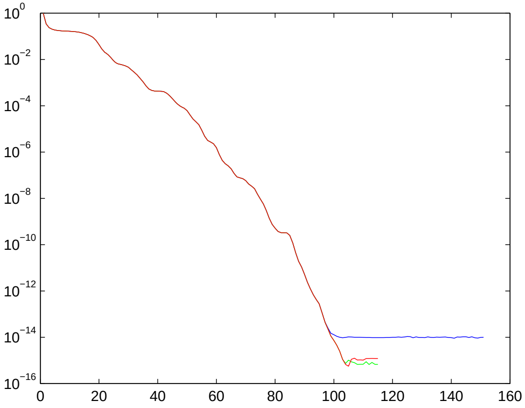
More experiments



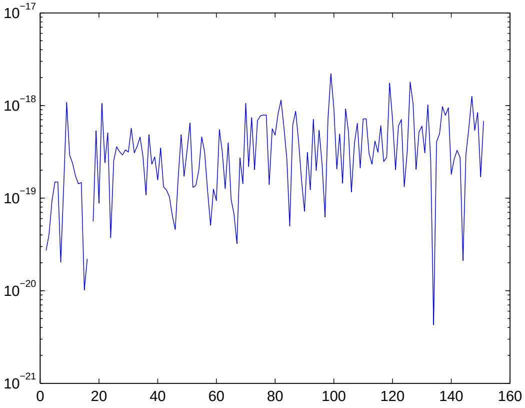
Difference of the residual norms of GMRES-MGS and Q-OR optimal, fs 680 1c, $n = 680$



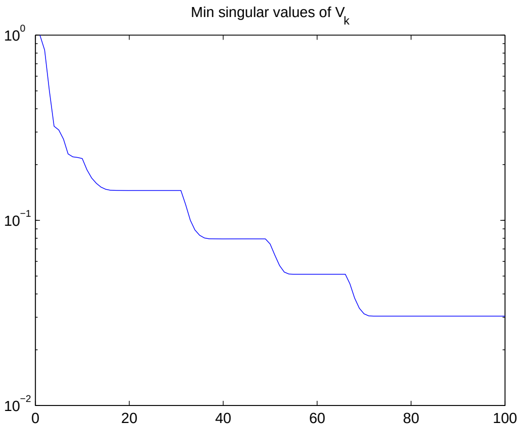
Minimum singular value of V_k, Q-OR optimal, fs 680 1c, $n = 680$



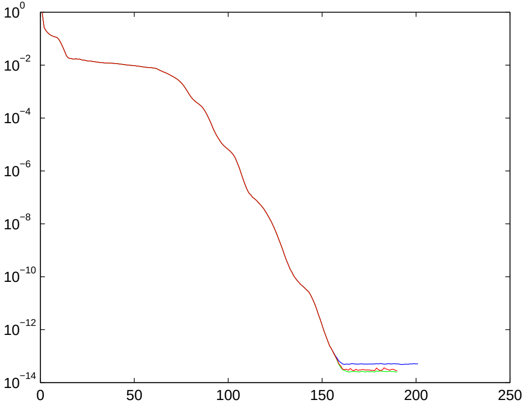
Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), fs 680 1c, $n = 680$



Difference of the residual norms of GMRES-MGS and Q-OR optimal, raefsky1 3242, $n = 3242$

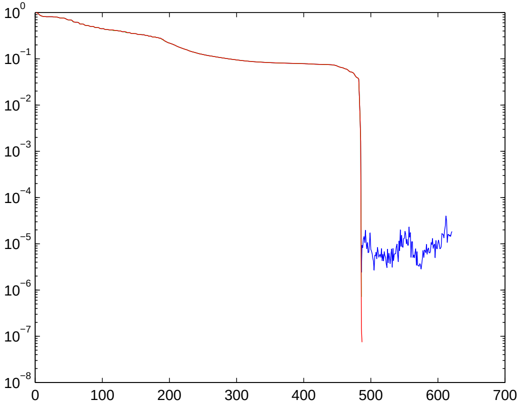


Minimum singular value of V_k, Q-OR optimal, raefsky1 3242, $n = 3242$

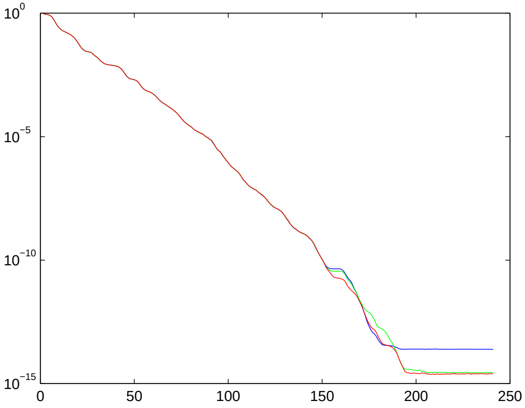


Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), raefsky1 3242, $n = 3242$

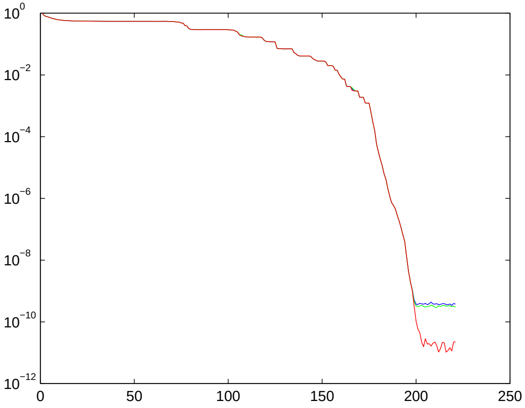
In many cases the maximum attainable accuracy is (slightly) better than with GMRES-MGS



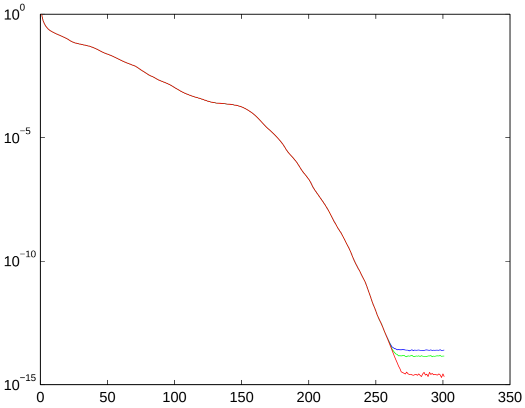
Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), bcsstk20 485, $n = 485$



Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), rdb2048, $n = 2048$



Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), steam1 240, $n = 240$



Residual norms of GMRES-MGS (blue), GMRES-MGS with double re-orthogonalization (green) and Q-OR optimal (red), Trefethen 500, $n = 500$

Conclusion

Using the properties of the Q-OR methods we were able to construct a non-orthogonal basis for which Q-OR gives the same residual norms as GMRES

However, the algorithm is more expensive than GMRES

But, it is more parallel than GMRES-MGS and most of the operations are matrix-matrix or matrix-vector products

In many cases the maximum attainable accuracy is better than with GMRES-MGS

However, (at least theoretically), the algorithm is not breakdown-free

It remains to study its stability in finite precision arithmetic