

Newton bases for GMRES and Q-MR

Gérard MEURANT

Joint work with Lothar Reichel

Algorithmy, March 2016

- 1 Introduction
- 2 Newton bases
- 3 Choices of shifts
- 4 Example
- 5 Q-MR
- 6 Q-MR with loose orthogonalization

Introduction

GMRES is one of the best Krylov method for solving non symmetric linear systems

$$Ax = b$$

where A is a real non symmetric matrix

The Krylov subspace is

$$\mathcal{K}_m(A, r) = \{r, Ar, A^2r, \dots, A^{m-1}r\}$$

where $r = b - Ax_0$ (usually $x_0 = 0$)

Generally GMRES is restarted every m iterations \rightarrow GMRES(m)

GMRES uses an “orthonormal” basis constructed with the Arnoldi process

It is modified Gram-Schmidt adapted to the Krylov subspace

$$v_1 = r / \|r\|$$

For $k = 1, \dots, m$

$$w = Av_k$$

For $j = 1, \dots, k$

$$h_{j,k} = v_j^T w$$

$$w = w - h_{j,k} v_j$$

end (for j)

$$h_{k+1,k} = \|w\|, \quad v_{k+1} = w / h_{k+1,k}$$

end (for k)

There are data dependencies which are not good on a parallel computer

How to introduce more parallelism in GMRES?

Many people have been working on this problem; see, for instance:

[Z. Bai, D. Hu, and L. Reichel](#), *A Newton basis GMRES implementation*, IMA J. Numer. Anal., 14 (1994)

[J. Erhel](#), *A parallel GMRES version for general sparse matrices*, Electron. Trans. Numer. Anal., 3 (1995)

[R.B. Sidje](#), *Alternatives to parallel Krylov subspace basis computation*, Numer. Linear Algebra Appl., 4 (1997)

[B. Philippe and L. Reichel](#), *On the generation of Krylov subspace bases*, Applied Numerical Mathematics, 62 (2012)

Newton basis

One possibility is to use a **Newton** basis constructed as

$$v_1 = r/\|r\|, \quad v_{i+1} = \frac{1}{\eta_i}(A - \xi_i I)v_i, \quad i = 1, \dots, m-1$$

where the η_i 's are normalizing factors to have vectors of unit norm and the ξ_i 's $\in \mathbb{C}$ are known as *shifts*

If $V_m = [v_1, v_2, \dots, v_m]$ is of rank m , we have a basis of the **Krylov** subspace

If we use a preconditioner M , A has to be replaced either by $M^{-1}A$ or AM^{-1}

Note that $v_{m+1} = p_m(A)v_1$ and the ξ_i 's are the roots of p_m

If the ξ_i 's appear in complex conjugate pairs, everything can be implemented in real arithmetic

Algorithm GMRESN(m)

For $k = 1, \dots, m$

 Add a vector to the known Newton basis V_k
end (for k)

Orthogonalize V_{m+1} with (parallel) QR

Compute $H_{m+1,m}$

Compute the iterates x_k by reducing $H_{m+1,m}$ to upper triangular structure

Note that V_m is tall and skinny ($m \ll n$)

There exist good (recent) parallel QR algorithms for such matrices (for instance, TSQR)

Problem: How to compute “good” shifts?

- ▶ $\|V_m\|$ small
- ▶ $\sigma_{\min}(V_m)$ large

We consider two possibilities:

- ▶ Spoke sets (Reichel and Philippe)
- ▶ “Good” interpolation points in the complex plane

We first do p initial steps of GMRES and we compute the Ritz values (eigenvalues of H_p)

Typically $p = 10$

H_p is real, the Ritz values are real or complex conjugate pairs

Spoke sets

We compute the center of gravity G (a point on the x -axis) of the Ritz values R_i

For each segment GR_i we compute fast Leja points (good interpolation points on a segment)

see [J. Baglama, D. Calvetti and L. Reichel, *Fast Leja points*, Electron. Trans. Numer. Anal., 7 \(1998\)](#)

“Good” interpolation points

A minimal volume ellipse containing the **Ritz** values is computed

Then a set X of $N \geq m$ points is computed in the upper half-ellipse to obtain a small **Lebesgue** constant

The N **Lagrange** polynomials l_j related to these points are such that the polynomial l_j is equal to 1 at point ξ_j and zero at the other points

The **Lebesgue** function is

$$\Lambda_X(\xi) = \sum_{j=1}^m |l_j(\xi)|$$

The **Lebesgue** constant l_X is the maximum of the **Lebesgue** function over $\xi \in \Omega$ where Ω is our given half-ellipse

We solve the problem

$$\min_{x, \xi_j \in \bar{\Omega}} I_{\Omega, \partial\Omega}, \quad I_{\Omega, \partial\Omega} = \left\{ I_{\Omega}^2 + \mu \int_{\partial\Omega} [l_1(\xi)^2 + \dots + l_N(\xi)^2] ds \right\}^{\frac{1}{2}}$$

where μ is a positive real number and

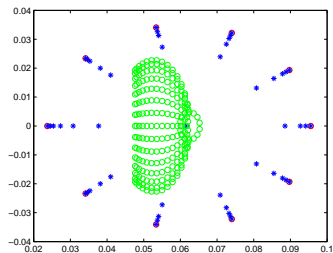
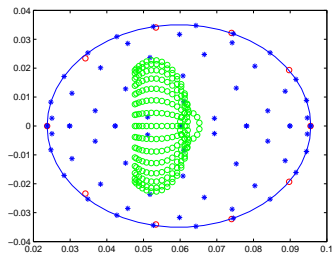
$$I_{\Omega} = \left\{ \int_{\Omega} [l_1(x, y)^2 + \dots + l_N(x, y)^2] dx dy \right\}^{\frac{1}{2}}$$

This does not always give the smallest possible **Lebesgue** constants therefore we added heuristic refinement algorithms which, in many cases, improve the results

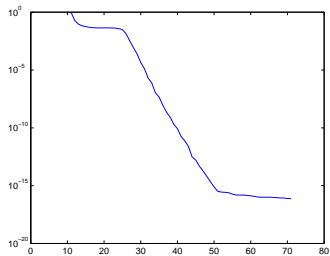
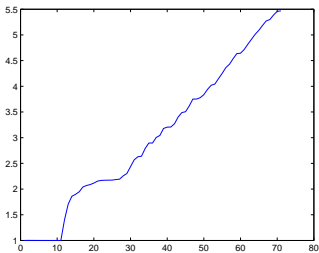
We pre-computed points for the half-ellipse Ω_0 of center $(0, 0)$ with semi-axes 2 (in x) and 1 (in y) for different total degrees of the bivariate polynomial

When we have the minimal volume ellipse Ω we compute the needed degree for having at least m points and we map the (symmetrized) points in Ω_0 to Ω

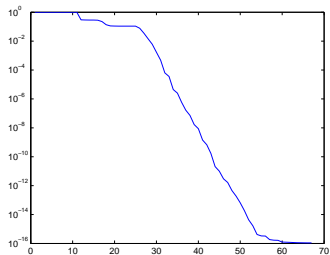
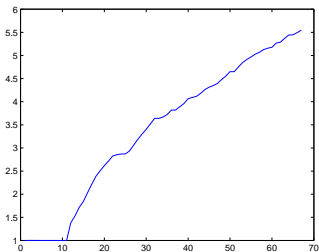
Example



supg001 225, eigenvalues, half-ellipse and spoke sets



supg001 225, half-ellipse, norm of V_k and $\sigma_{\min}(V_k)$



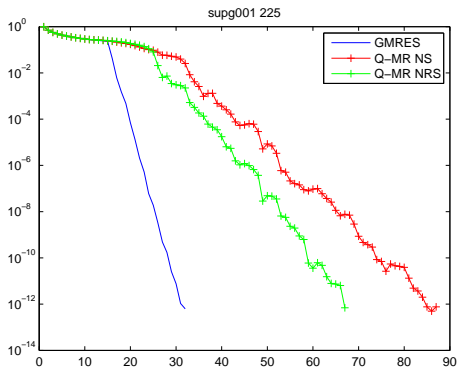
supg001 225, spoke sets, norm of V_k and $\sigma_{\min}(V_k)$

As long as the basis is full rank (for m small) we obtain the same results as with **GMRES(m)**

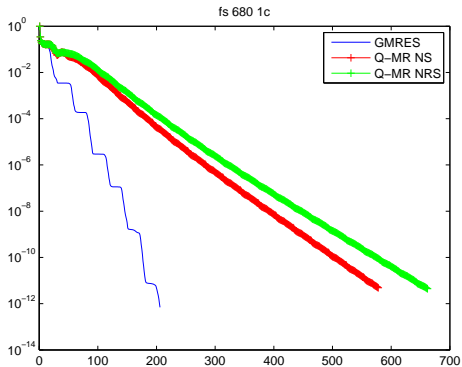
Could we use the **Newton** bases with **Q-MR**?

That is, without the QR step

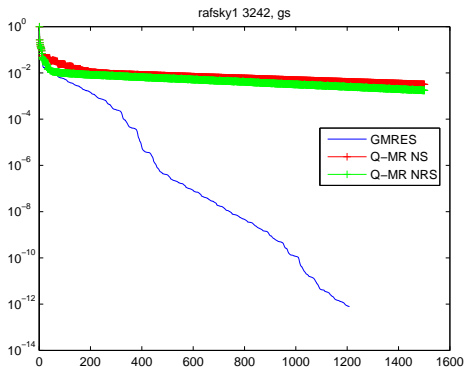
⇒ We just minimize the norm of the quasi-residual



supg001 225, GMRES(30), Q-MR(30) with half-ellipse and spoke sets



fs 680 1c, GMRES(30), Q-MR(30) with half-ellipse and spoke sets



rafsky1 3242, GMRES(30), Q-MR(30) with half-ellipse and spoke sets

Q-MR with loose orthogonalization

Q-MR is not always working well with the **Newton** basis

We just need the “R” factor of a QR factorization of V_{m+1}

$$AV_m = V_{m+1} T_{m+1,m}$$

where $T_{m+1,m}$ is with at most an upper bandwidth of 2

$$V_{m+1} = Q_{m+1} R_{m+1,m+1}$$

It yields

$$AV_m = Q_{m+1} R_{m+1,m+1} T_{m+1,m}$$

and if $x_m = x_0 + V_m z$, we have

$$\|b - Ax_m\| = \|r_0 - AV_m z\| = \|Q_{m+1}(\|r_0\|e_1 - R_{m+1,m+1} T_{m+1,m} z)\|$$

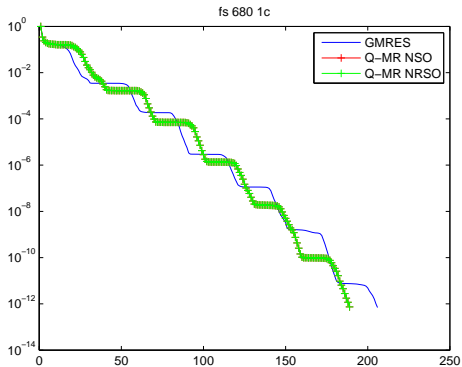
We can use **CholQR** that is, the **Cholesky** factorization

$$V_{m+1}^T V_{m+1} = R_{m+1,m+1}^T R_{m+1,m+1}$$

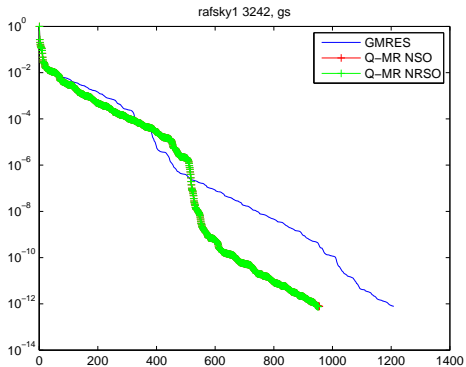
However, when $\sigma_{\min}(V_{m+1})$ is small $V_{m+1}^T V_{m+1}$ might not be numerically positive definite

If this happens, we use CGS (classical **Gram-Schmidt**)

Note that this is very close to what we did for **GMRES** except that the orthogonalization is done differently (loosely)



fs 680 1c, GMRES(30), CholQR Q-MR(30) with half-ellipse and spoke sets



raefsky1 3242, GMRES(30), ChoQR Q-MR(30) with half-ellipse
and spoke sets

Conclusion

- ▶ **Newton** bases + **QR** work well for **GMRES(m)**
- ▶ Interpolation points and spoke sets yield similar results
- ▶ **Newton** bases do not always work well with **Q-MR**
- ▶ But **Newton** bases + loose orthogonalization could give nice results