

# Performance analysis of Krylov iterative solvers using Ritz values

IRINA CRACIUN  
 Advisor: Dr. Burak Aksoylu  
 Center for Computation and Technology  
 Louisiana State University - Baton Rouge

(irina@cct.lsu.edu)



## 1. Introduction

The problem of solving linear systems has become a crucial issue for scientific computing during the past 50 years. A majority of problems in physics, astrophysics, computational fluid dynamics, chemistry, engineering etc. reduce to solving a differential or integral equation, which in turn, after discretisation and linearisation, reduce to a linear system of the form  $Ax = b$ . Coming from real-life problems, the matrices are usually non-normal and very large - often reaching orders of millions. For such large matrices, any linear solver fails because of computational errors.

On the positive side, they are also often sparse matrices (contain very few non-zero entries), which makes the computation of matrix-vector products a very easy, non-costing task. This is why we use iterative methods to solve large sparse linear systems: iterative methods generate successive approximations of the solution, such that each approximation is computed from the previous one through a series of matrix-vector products.

Currently there is no "perfect" iterative method; each known iterative method has been proved efficient for some particular types of matrices and inefficient for others, each having many circumstances where it fails. In our research we focus on two of the most well-known such methods, FOM and GMRES, which have been found to fail or stall in only one case. The failure of these methods depends on Ritz and harmonic Ritz values, both relatively "newly-discovered" topics in linear algebra. New properties of Ritz vectors and Ritz values are still being discovered, but using them to "fix" FOM and GMRES would be of great importance, as it might yield the "Holy Grail" of solving linear systems, i.e. an iterative method which could be applied without failure for any matrix.

## 2. Eigenvalues and Eigenvectors

Let a matrix  $A \in \mathbb{C}^{n \times n}$ . A complex scalar  $\lambda$  is called an **eigenvalue** of  $A$  if there exists a nonzero  $n$ -vector  $u$  such that

$$Au = \lambda u \quad (1)$$

The vector  $u$  is called an **eigenvector** of  $A$  associated with  $\lambda$ .

### Example

Let  $A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ . We need to solve the following equation, with unknowns  $\lambda$  and  $u$ .

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \lambda \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

First we find the eigenvalues  $\lambda$  by solving the equation

$$\det(A - \lambda I) = 0 \quad (2)$$

known as the **characteristic equation** of  $A$ . In our case, this yields  $(1 - \lambda)(4 - \lambda) = 0$ , so the eigenvalues of  $A$  are  $\lambda_1 = 1$  and  $\lambda_2 = 4$ . For the case  $\lambda = 1$  we find  $u$  by solving the equation

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

which yields the eigenvector  $u = \begin{bmatrix} t \\ 0 \end{bmatrix}$ ,  $t \in \mathbb{R}$ . In the same way, we find the eigenvectors associated

with  $\lambda_2$ :  $u = \begin{bmatrix} 0 \\ t \end{bmatrix}$ ,  $t \in \mathbb{R}$

## 3. Iterative Methods

Many problems in computational physics, astrophysics, chemistry, engineering etc reduce to solving a linear system of the form  $Ax = b$ . A key issue in science today, the problem of solving linear systems for real-life simulations is often complicated, even for the most powerful supercomputers: coming from real applications, the matrix  $A$  is usually non-normal, and often very large (sometimes reaching orders in the range of millions). Therefore one of the most important on-going struggles in computational science is finding new algorithms to solve such systems, as well as improving the current algorithms.

Iterative methods for solving linear systems begin with an initial guess of the solution, and then proceed to find an approximate solution to the system by finding successive, increasingly accurate

approximations in an **invariant subspace** of  $A$ . An invariant subspace  $W$  of a matrix  $A$  is a subspace with the property that

$$\forall w \in W, Aw \in W \quad (3)$$

The most successful iterative methods make use of **Krylov subspaces**: for a non-singular matrix  $A \in \mathbb{C}^{n \times n}$  and a vector  $r_0 \in \mathbb{C}^n$ , we define the Krylov subspaces as:

$$\mathcal{K}_m(A, r_0) = \text{span}\{r_0, Ar_0, A^2r_0, \dots, A^{m-1}r_0\} \quad (4)$$

where  $m = 1, 2, \dots, n$ . We often need to form an orthonormal basis for the Krylov subspace  $\mathcal{K}_m$ , and we do this by orthogonalizing the Krylov power basis. We thus obtain the **Arnoldi basis**  $\{v_1, v_2, \dots, v_m\}$ .

## 4. FOM and GMRES

The most used Krylov subspace methods are the full orthogonalization method (FOM), also known as the Lanczos method, and the generalized minimal residual method (GMRES). The final goal of iterative methods is to find the best approximation by minimizing the **residual vector**:

$$r = Ax_{approx} - b \quad (5)$$

Say an iterative method selects its approximate solutions from an invariant subspace  $W$ . Obviously the residual,  $r$ , is minimal when it is orthogonal to  $W$  (see Figure 1).

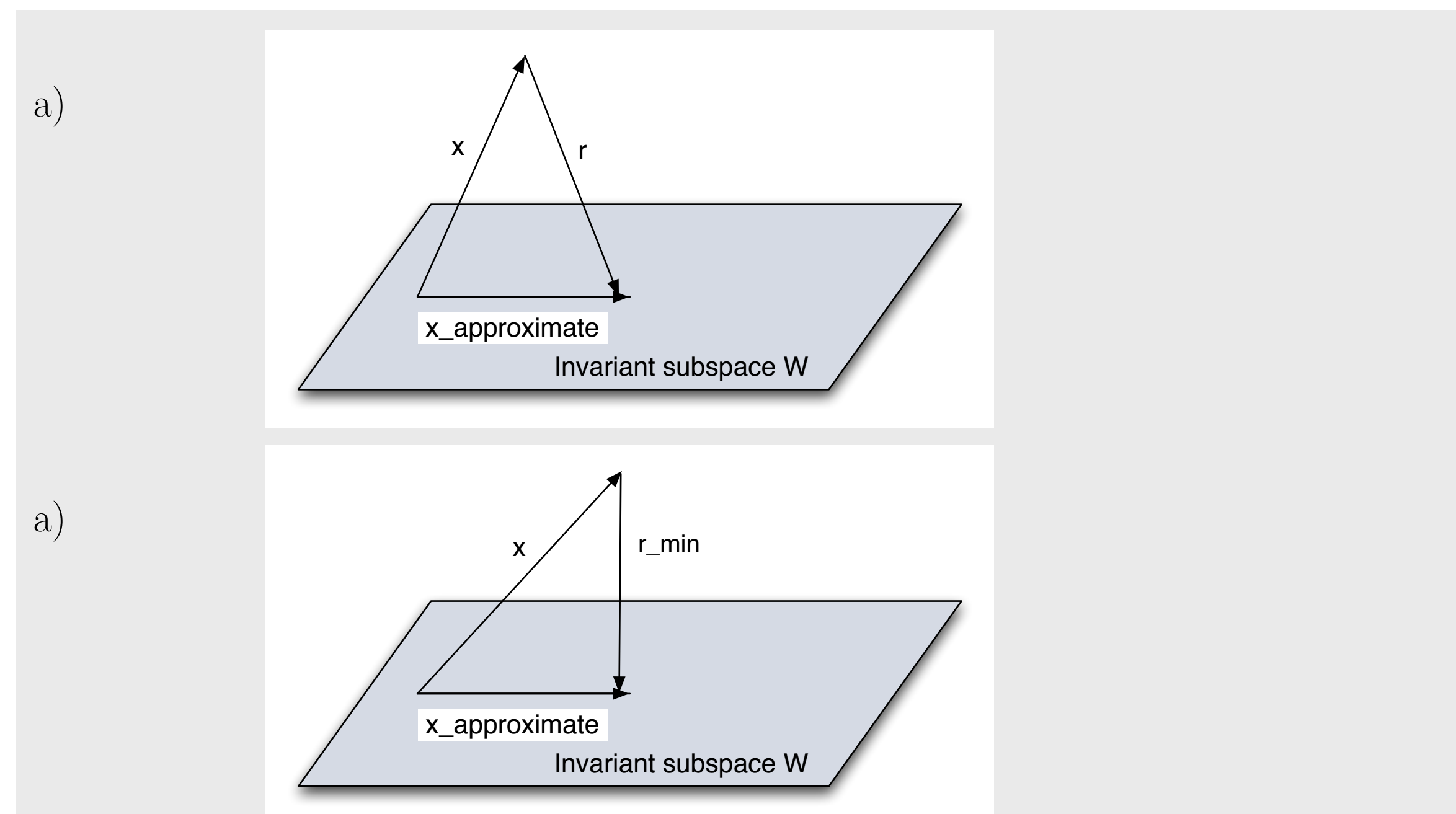


FIGURE 1: a). Residual vector for an approximate solution  
 b). Minimal residual obtained through orthogonalisation

For both FOM and GMRES, the "initial guess"  $x_0$  is 0, and therefore the initial residual  $r_0$  is equal to  $b$ . Both algorithms select their approximations from the Krylov subspace  $\mathcal{K}_m(A, r_0)$ . Therefore, the approximate solutions found by FOM and GMRES respectively, can be written as:

$$x^{\text{FOM/GMRES}} = (a_{m-1}A^{m-1} + \dots + a_2A^2 + a_1A + a_0)r_0 = p_{m-1}(A)r_0 \in \mathcal{K}_m(A, r_0) \quad (6)$$

where  $p_{m-1}$  is a polynomial of degree  $m - 1$ . The residual is therefore

$$r = b - Ax = (I - Ap_{m-1}(A))r_0 = q_m(A)r_0 \in \mathcal{K}_{m+1}(A, r_0) \quad (7)$$

FOM selects the approximate solution  $x^{\text{FOM}}$  such that the residual  $r^{\text{FOM}}$  is orthogonal to the Krylov subspace  $\mathcal{K}_m(A, r_0)$ , and GMRES selects the approximate solution  $x^{\text{GMRES}}$  which minimizes the norm of the residual vector,  $r^{\text{GMRES}} = b - Ax^{\text{GMRES}}$ . Since the approximate solution vector  $x^{\text{GMRES}}$  is selected from the Krylov subspace  $\mathcal{K}_m(A, r_0)$ , this condition is equivalent to finding a vector  $t$  which minimizes

$$\|A\mathcal{K}_m(A, r_0)t - b\|$$

or, in other words, the residual must be orthogonal to the subspace  $A\mathcal{K}_m(A, r_0)$  (Figure 2).

## 5. Ritz and Harmonic Ritz Values

It has been discussed previously that we need an orthonormal basis for the Krylov subspaces. The process of orthogonalising the vectors  $r_0, Ar_0, \dots, A^{m-1}r_0$  as to obtain an orthogonal matrix

$V_m = [v_1 \ v_2 \ \dots \ v_m]$ , is known as **Arnoldi iteration**. In this process, we need to first compute scalars  $h_{i,j}$ , which lead to the orthogonal vectors  $v_i$ . These scalars form an upper Hessenberg matrix  $H_m$ , which proves to be crucial to the algorithm:

$$H_m = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,m-1} & h_{1,m} \\ h_{2,1} & h_{2,2} & \dots & h_{2,m-1} & h_{2,m} \\ & h_{3,2} & \dots & h_{3,m-1} & h_{3,m} \\ & & \dots & \vdots & \vdots \\ & & & h_{m,m-1} & h_{m,m} \end{bmatrix} \quad (8)$$

The eigenvalues of  $H_m$  are known as the **Ritz values** of  $A$ : they are approximations for the eigenvalues of  $A$ , and they have been shown to be the roots of the FOM residual polynomial - in other words, the residual reaches the desired zero-value when evaluated at the Ritz values. **Harmonic Ritz** values are the reciprocals of the Ritz values of  $A^{-1}$ : if  $\lambda$  is a Ritz value of  $A^{-1}$ , then  $1/\lambda$  is a harmonic Ritz value of  $A$ . Harmonic Ritz values have been shown to be the roots of the residual polynomial of GMRES.

Numerical studies have shown that FOM breaks down when a Ritz value is 0, and that GMRES stalls when a harmonic Ritz value goes to infinity. Of course, these "exact" values need not be achieved in computation, it suffices that a Ritz value is very small, or that a harmonic Ritz value is very large, for these algorithms to fail.

In our research we are studying the properties of the Ritz and harmonic Ritz values, with the final goal of fixing the problems in FOM and GMRES, by either modifying the algorithms in the cases where it is likely to fail, or by coming up with completely new algorithms to be used especially in these problem cases.

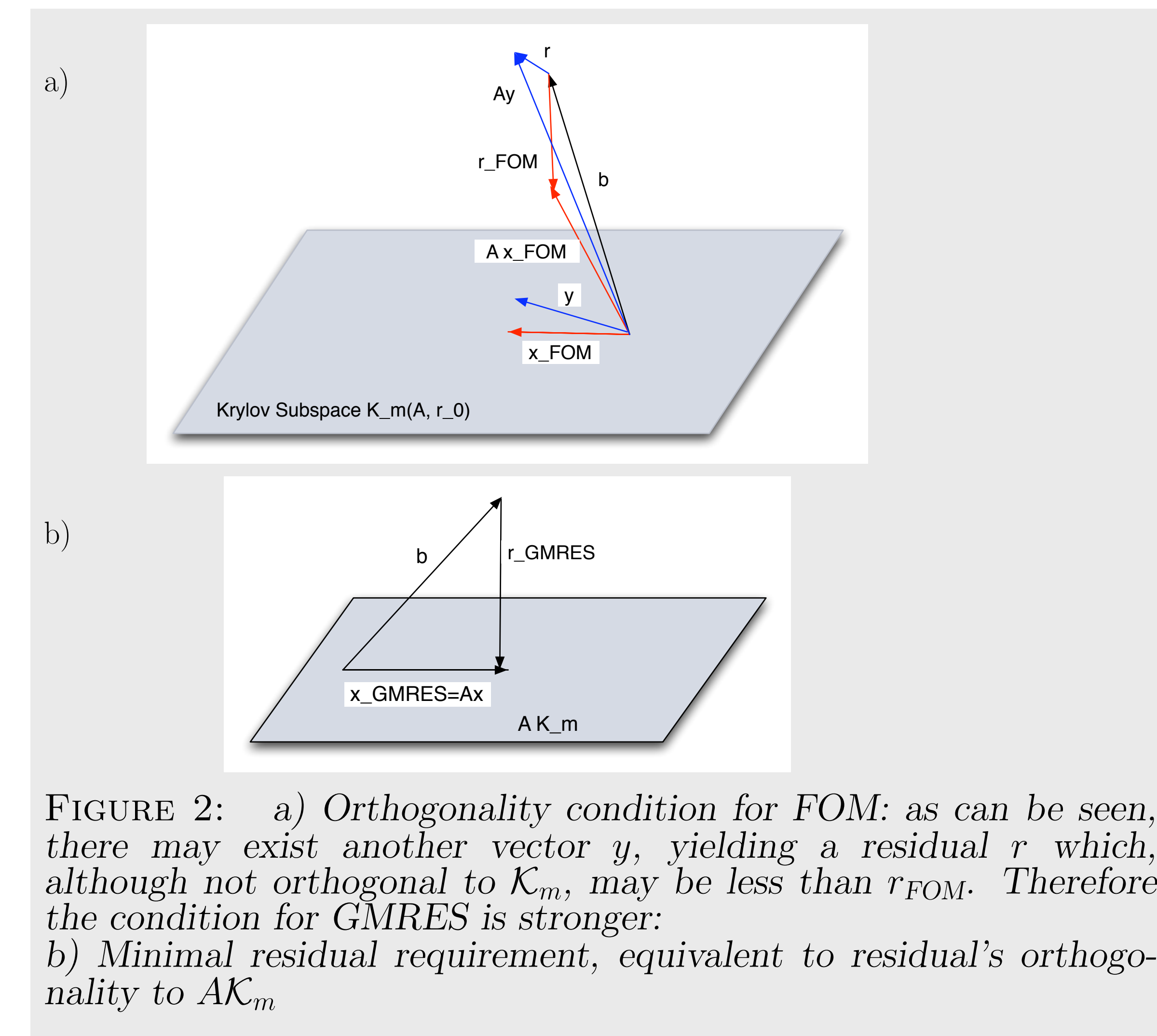


FIGURE 2: a) Orthogonality condition for FOM: as can be seen, there may exist another vector  $y$ , yielding a residual  $r$  which, although not orthogonal to  $\mathcal{K}_m$ , may be less than  $r_{\text{FOM}}$ . Therefore the condition for GMRES is stronger:  
 b) Minimal residual requirement, equivalent to residual's orthogonality to  $A\mathcal{K}_m$

## 6. Acknowledgments

I would like to thank Dr. Burak Aksoylu, Dr. Ed Seidel, Dr. Gabrielle Allen, Elena Caraba and all the Center for Computation & Technology for support, advice and help in this research.

## References

- [1] Serge Goossens and Dirk Roose, *Ritz and harmonic Ritz values and the convergence of FOM and GMRES*, Numerical linear algebra with applications 6 (1999) 281-293
- [2] Stewart, Gilbert W., *Matrix algorithms*, 1998, SIAM, Philadelphia
- [3] Lloyd N. Trefethen and David Bau, III, *Numerical linear algebra*, 1997, SIAM, Philadelphia